Detecting Graves in GPR Data:

Assessing the viability of machine learning for the interpretation of graves in B-scan data using medieval Irish case studies

Bournemouth University

Ashely Green

Thesis submitted in partial fulfilment of the requirements for the degree 'Doctor of Philosophy,' awarded by Bournemouth University

June 2020

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with its author and due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

All images and illustrations in this document were created by the author unless otherwise indicated.

ABSTRACT

As commercial archaeogeophysical survey progressively shifts towards large landscape-scale surveys, small features like graves become more difficult to identify and interpret. In order to increase the rate and confidence of grave identification before excavation using geophysical methods, the accuracy and speed of survey outputs and reporting must be improved. The approach taken in this research was first to consider the survey parameters that govern the effectiveness of the four conventional techniques used in commercial archaeogeophysical evaluations (magnetometry, earth resistance, electromagnetic induction and ground-penetrating radar). Subsequently, in respect of ground-penetrating radar (GPR), this research developed machine learning applications to improve the speed and confidence of detecting inhumation graves. The survey parameters research combined established survey guidelines for the UK, Ireland, and Europe to account for local geology, soils and land cover to provide survey guidance for individual sites via a decision-based application linked to GIS database. To develop two machine learning tools for localising and probability scoring grave-like responses in GPR data, convolutional neural networks and transfer learning were used to analyse radargrams of medieval graves and timeslices of modern proxy clandestine graves. Models were *c.* 93% accurate at labelling images as containing a grave or no grave and *c.* 96% accurate in labelling and locating potential graves in radargram images. For timeslices, machine learning models achieved 94% classification accuracy. The >90% accuracy of the machine learning models demonstrates the viability of machine-assisted detection of inhumation graves within GPR data. While the expansion of the training dataset would further improve the accuracy of the proposed methods, the current machine-led interpretation methods provide valuable assistance for human-led interpretation until more data becomes available. The survey guidance tool and the two machine learning applications have been packaged into the *Reilig* web application toolset, which is freely available.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF DIGITAL APPENDICES

File Name	Description
TrainingData.json	JSON file containing training data
Reilig_SurveyParameters.kmz	Online version of survey parameters tool in .kmz format for import into Google Earth
Reilig_SurveyParameters Shapefile	Desktop version of survey parameters tool in shapefile format for import into ArcGIS
ArcGIS App Navigation	Video of survey parameters app in use
Model Training	Video of model being trained
Reilig for Radargrams	Video of example predictions using Reilig for Radargrams web app
Reilig for Timeslices	Video of example predictions using Reilig for Timeslices web app
Case Study Interpretations	Folder containing the interpretation outputs of the Reilig for Radargrams web app for detections on the case study survey data
Survey Reports	Folder containing the technical reports for the case study surveys
Jupyter Notebooks	Notebooks containing the code for training the machine learning models
Web Apps	Folder containing the CSS, HTML, and JavaScript files used to create the web app interfaces

LIST OF ABBREVIATIONS

1D	One-dimensional
2D	Two-dimensional
3D	Three-dimensional
API	Application Programming Interface
ATAGS	*Automated Tool for Archaeo-Geophysical Survey*
CAD	Computer-aided Design
CNN	Convolutional neural network
CRISP-DM	Cross-Industry Standard Process for Data Mining
CSS	Cascading Style Sheets
D	Depth
DART	*Detection of Archaeological residues using Remote sensing Techniques*
DC	Direct current
DL	Deep learning
EMI	Electromagnetic Induction
FFT	Fast Fourier Transform
FN	False negative
FP	False positive
FT	Fourier Transform
GAP	Global Average Pooling
GIS	Geographic Information System
GPR	Ground-penetrating Radar
GPS	Global Positioning System
GUI	Graphical user interface
HCP	Horizontal coplanar dipole
HTML	Hypertext Markup Language
IAFS	Irish Archaeology Field School
IED	Improvised Explosive Device
INHP	Irish National Heritage Park
ISSS	International Soil Science Society
L	Length

LFEM	Low frequency electromagnetic
LiDAR	Light Detection and Ranging
MCC	Matthew's Correlation Coefficient
ML	Machine learning
MNI	Minimum Number of Individuals
MNIST	Modified National Institute of Standards and Technology
MPN	Multilayer Perceptron Network
NASA	National Aeronautics and Space Administration
NCC	Normalised cross-correlation
NLCD	National Land Cover Database
NN	Neural network
PRP	Perpendicular dipole
R-CNN	Regions-based convolutional neural network
RDP	Relative dielectric permittivity
ReLU	Rectified linear unit
ResNet	Residual network
RGB	Red Green Blue
RTK	Real Time Kinematic
SMR	Scheduled Monument Record
SSD	Single shot multibox detector
TII	Transport and Infrastructure Ireland
TN	True negative
TP	True positive
USDA	United States Department of Agriculture
UK	United Kingdom
VCP	Vertical coplanar
VGG	Visual Geometry Group
W	Width

LIST OF PUBLICATIONS

PUBLICATIONS

McLoughlin, C. and **Green, A.**, 2019. Carrick town and manor: archaeological, geophysical and landscape assessment. In: *Carrick, County Wexford: Ireland's First Anglo-Norman Stronghold,* Shine, D., Potterton, M., Mandal, S., and McLoughlin, C., eds., Dublin: Four Courts Press.

Shine, D., Reid, R., Mandal, S., **Green, A.**, and Clutterbuck, R., 2019. New investigations at the Carrick ringwork site. In: *Carrick, County Wexford: Ireland's First Anglo-Norman Stronghold,* Shine, D., Potterton, M., Mandal, S., and McLoughlin, C., eds., Dublin: Four Courts Press.

Shine, D., Callaghan, S., **Green, A.**, and Arrigan, B., 2019. Rehearsing for the Theatre of War. *Archaeology Ireland,* 33(2).

Green, A., 2018. Visualization of ground-penetrating radar data. In: *Encyclopaedia of Archaeological Science*, Lopez-Verala, S., ed., London: Wiley.

In preparation: **Green, A.**, Milward, J., Cheetham, P., and Darvill, T., n.d. A broad characterisation of burial practices in Irish archaeological contexts. Intended Journal: European Journal of Archaeology.

In preparation: **Green, A.**, n.d., Archaeogeophysical Prospection of Potential Gravesites: Recommendations and a tool for determining survey parameters. Intended Journal: Archaeological Prospection.

In preparation: **Green, A.** and Budka, M., n.d., Reilig: Methods for the detection of graves in GPR data. Intended Journal: Journal of Computer Applications in Archaeology.

PRESENTATIONS

Green, A. and Cheetham, P., 2019. Rise of the Machines: Improving the identification of possible graves in GPR data with interactive survey guidance and machine learning. *AP2019 - 13th International Conference of Archaeological Prospection* 28-31 August 2019 Sligo, Ireland. **In Press.** [Conference Paper]

Green, A., 2018. Over the Top, Under the Ground: Geophysical survey at Birr Barracks, Crinkill, County Offaly. *In: Near Surface Geophysics Group Meeting on Recent Work in Archaeological Geophysics* 4 December 2018 The Geological Society, London. [Conference Poster]

Green, A., Cheetham, P. and Darvill, T., 2018. Grave New World: An Irish perspective on automating the interpretation of grave-like features in GPR data. *In: Institute for Archaeologists Ireland Conference* 23-24 March 2018 Sligo, Ireland. [Conference Poster]

Green, A., Cheetham, P. and Darvill, T., 2018. 'Mark the Graves': Adapting automation for improving the delineation of graves in GPR data. *In: CAA International Conference* 19-23 March 2018 University of Tubingen, Tubingen, Germany. [Conference Paper]

Green, A., Cheetham, P. and Darvill, T., 2017. Approaches to Improving the Pre-Excavation Detection of Inhumations. *In:* Jennings, B., Gaffney, C., Sparrow, T. and Gaffney, S., eds. *AP2017 - 12th International Conference of Archaeological Prospection* 12-16 September 2017 Bradford, United Kingdom. Oxford: Archaeopress Archaeology, 90-91. [Conference Poster]

Green, A., Cheetham, P. and Darvill, T., 2017. Information systems and ad hoc workflow models as viable applications to improve the pre-excavation detection of characteristic human interments. *In: CAA International Conference* 14-16 March 2017 Georgia State University, Atlanta, USA. [Conference Paper]

Green, A. and Holmes, R., 2017. Accounting for Environmental and Anthropogenic Factors: Approaches to enhancing horizontal resolution and interpretability in geophysical surveys. *In: CAA International Conference, Atlanta, USA.* [Conference Poster]

Green, A. and Cheetham, P., 2016. Reimaging the Black Friary: Recent Approaches to Seeing Beyond Modern Activity at the Dominican Friary, Trim, County Meath, Republic of Ireland. *In: Near Surface Geophysics Group (NSGG) Meeting on Recent Work in Archaeological Geophysics.* [Conference Poster]

ACKNOWLEDGEMENTS

"Archaeological anomalies are part of a wide range of causative bodies that are often difficult to differentiate…a geophysicist is not a magician; the interpretation should be logical and based on the theory of the techniques."

-Gaffney and Gater, 2010

1. INTRODUCTION

1.1. Project Rationale

Each year several thousand human skeletons are disturbed during development projects in England alone, with approximately 75% of the skeletons excavated from archaeological contexts post-date the 6th century AD (Advisory Panel on the Archaeology of Burials in England 2017, p. 3). One method for determining the potential for human remains on a site, aside from desk-based assessments, is geophysical survey. Since their development in the 1950s geophysical surveys, have been part of archaeological investigations, but have only regularly integrated into criminal investigative strategies in the last decade (Cheetham 2005, Conyers 2013). Conventional archaeogeophysical applications include delineating structural remains, ditches and enclosures, cultivation practices, former field boundaries, and subsurface graves, as well as differentiating modern services from archaeological remains (Davenport 2001). One aspect where geophysical survey has struggled since its inception is in detecting graves. The difficulty in detecting graves is often a result of the preservation state of graves as in most soil conditions the human remains and burial container, especially wooden coffins, decompose over time. Once the remains have fully decomposed the differences in the geochemical and physical properties of the remains, backfill, and surrounding matrix become challenging to detect, or even indiscernible. The minimal disturbance of the soil once returned to its original position, and lack of additives with enhanced geophysical characteristics such as charcoal or other material with enhanced magnetic properties result in marginal differences between the grave fill and surrounding material. A grave is comprised of the grave cut, grave fill, grave shaft, remains, and sometimes a burial container. Figure 1 depicts an illustrative example of interred skeletonised remains to demonstrate the relationship between the remains, grave fill, surrounding material, and ground surface..
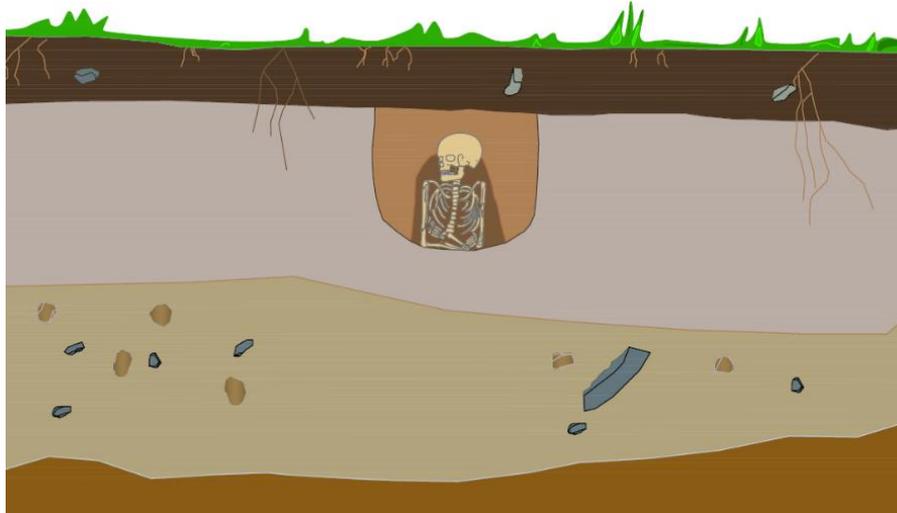
Figure 1: Illustrative example of interred skeletonised human remains that demonstrates the relationship between the grave fill, surrounding material and ground surface.

A compounding problem is introduced when the geophysical methods are successful in detecting a grave, but the surveyor or interpreter is inexperienced in interpreting the data. These situations can often lead to instances where graves are present in the dataset but are not delineated during the data interpretation stage, or the graves are not detected through geophysics and remain undiscovered until they are disturbed during excavation. An example of this phenomenon is the 793 individuals and associated interments not detected in a hillfort near Ranelagh, County Roscommon, Ireland (Delaney 2017, Murray 2017). The misinterpretation or lack of detection of graves has a significant monetary impact on clients in commercial archaeology as well as ethical implications when it comes to recovering the human remains.

Geophysical survey has been widely used in attempts to locate modern and archaeological unmarked graves and mass graves in a range of environments (Kenyon and Bevan 1977, Bevan 1991, Mellett 1992, Unterberger 1992, Miller 1996, Eaton 1999, Davis et al. 2000, Koppenjan et al. 2004, Cheetham 2005). Exploratory and controlled research surveys have proven geophysical methods can detect both modern and archaeological interments, dependent on environmental and anthropogenic factors, such as target morphology, physical and chemical properties of the fill and the surrounding matrix, and orientation (Gaffney and Gater 2003, Conyers 2013, Richardson and Cheetham 2013). Further discussion on the detection of modern graves is provided in Chapter 11.2.2.

Most geophysical techniques are used in the search for modern and archaeological unmarked graves. While magnetic techniques may aid in locating forensic evidence, GPR and resistivity techniques are suitable for detecting the moisture and soil compaction variations that result from a grave. It is important to note that because detecting the skeletal remains themselves is unlikely in a

modern clandestine grave, and nearly impossible in an archaeological example, it is the grave cut and grave fill that are detected (Bevan 1991; Bladon et al. 2011). Grave cuts are easily identifiable in highly stratified soil matrices (Bevan 1991; Conyers 2013). The grave cut is often detected because as the grave is dug, the cut truncates the stratigraphy and the redeposited material (grave fill) is mixed. The redeposited material disturbs the stratigraphy and can change the dielectric properties of the soil (Vaughan 1986; Bevan 1991; Bladon et al. 2011). Doolittle and Bellantoni (2010), however, demonstrated that the disturbed soil redeposited in the grave shaft becomes more challenging to detect in homogeneous soils because the dielectric properties do not change significantly from the surrounding soil matrix. Even in instances where an individual was interred in an untreated wooden coffin, the coffin is unlikely to cause a reflection as coffins are likely to deteriorate within ten years of interment in most environments (Doolittle and Bellantoni 2010). The deterioration of the coffin leaves only the skeletal remains, coffin nails, and any grave goods within the grave cut. There have been some instances where empty crania and thoracic cavities have produced individual hyperbolic reflections under optimal conditions (Damiata et al. 2013, p.273). However, responses from body cavities are rarely interpreted as a grave unless they form part of a targeted high-resolution survey following the identification of the burial ground's extent.

Graves are not only difficult to detect due to the ephemeral physical and chemical changes that minimise over time, but this difficulty is often compounded by post-interment disturbance and the spacing between a group of graves. Animals, plants, humans, and environmental factors can change or disturb the interface between the grave fill and the undisturbed soil matrix surrounding it, or produce a scattering effect (Conyers 2006, 2013; Nobes 2000; Vaughan 1986). Nobes (1999) also identified that a tight clustering or overlapping of graves could create an overlapping diffraction pattern in which individual graves cannot be isolated. However, Rial et al. (2009) later established that the overlapping diffractions could be minimised by using a high-frequency antenna. As is customary with responses from modern services, the data can also be migrated to reduce the effects of tightly clustered features/anomalies.

Ground-penetrating radar and earth resistance methods have proven to be the most successful in detecting graves (Gaffney and Gater 2003; Cheetham 2005; Conyers 2013; Ruffell and McAllister 2015). One of the earliest GPR surveys to attempt to locate graves was carried out by C.J. Vaughan (1986) from 1982 - 1983 in Red Bay, Labrador, Canada. The survey was challenged by waterlogged soils, complicated stratigraphy, and an overburden of cobbles (Vaughan 1986; Conyers 2013). Despite this, the survey still proved successful in identifying disturbed soil related to the grave fills.

However, the grave goods and skeletal remains were still undetectable due to their low contrast with the surrounding matrices (Vaughan 1986).

In their studies, NecroSearch (France et al. 1992, 1997) found that GPR was the most suitable technique for identifying clandestine graves in a forensic context. The depth of their pig cadavers (0.51 – 0.79m) was comparable to the average depth of medieval graves below the original topsoil.

Moffat (2015) has widely applied GPR to graves and potential cemetery sites in Australia with varying success due to environmental factors. More recently, Gaffney et al. (2014) achieved promising results from a multi-channel GPR survey of the Cistercian Fountains Abbey, Ripon, North Yorkshire. The survey demonstrated a clear demarcation of the high amplitude responses from individual graves.

As a result of technological developments and changing expectations much work has been carried out on data processing and visualisation, from the 3D visualisation of a dataset to animations that allow the interpreter to 'scan' through large datasets. Following the implementation of software to view radargrams and timeslices, the interpretation of graves in GPR data became much more straightforward, making the technique more applicable to locating such targets (Conyers 2013). However, the ability to identify such features still lies with the interpreter. Often this will bias the dataset as most geophysicists tend to focus on high amplitude responses; whereas, the responses of interest can often be the 'negative' areas or those with little to no reflection of the radar waves (Conyers 2013).

Geological factors, an uncontrollable aspect of the survey environment, play a significant role in determining the potential detectability of graves and the appropriate survey parameters for maximising the potential to detect them through geophysical survey. The principal disadvantage of using electromagnetic methods (including GPR) for archaeogeophysical survey is the variability of survey environments, as some subsurface materials cause high attenuation of the signal (Davenport 2001). It was initially thought that GPR surveys would not be suitable for clay soils, waterlogged soils, saltwater, or even brackish water (Conyers 2013). However, surveys over the past 20 years have shown that, while there is high signal attenuation, GPR can be successfully used on clayey soils, waterlogged soils, and freshwater (Conyers 2013). Conductive materials, such as volcanic bedrocks or saltwater, are still problematic for surveys. Electromagnetic geophysical survey techniques offer high-quality data across subsurface materials like sand or gravel because they are low loss materials and easily penetrable by the electromagnetic waves (Simms 1995; Ludwig et al. 2011).

Conversely, clays and loamy soils or highly conductive soils reduce the depth of penetration and the overall signal to noise ratio of data due to high attenuation (Simms 1995; Ludwig et al. 2011). Figure 2 shows the suitability of magnetic, electromagnetic (EMI and GPR), and resistance methods on a range of geologies present in Ireland, Northern Ireland, and Great Britain and shows that electromagnetic and resistance methods produce high-quality results on the broadest range of geologies. As Ireland's bedrock geology is primarily comprised of limestone formations with clayey silt, clayey sand, clay, sandy clay, and silty clay soils, when not accounting for topography, GPR is one of the most suitable techniques for archaeogeophysical survey in the country.

The Detection of Archaeological residues using Remote sensing Techniques (DART) project investigated the ability of a range of remote sensing techniques to detect archaeological features in challenging conditions. The project aimed to develop "analytical methods for identifying and quantifying gradual changes and dynamics in sensor responses associated with surface and near-surface archaeological features under different environmental and land management conditions" (Cohn 2019). DART focused on the detection of archaeological features in geologies and soils, which are known to hinder the detection and interpretation of archaeological features, such as clays. While several techniques were evaluated as part of the project, including quantum technology gravity sensors (Boddice et al. 2017) and LiDAR data (Stott et al. 2013, 2015), the evaluation most relevant to this research is the work undertaken by Fry (2014). Fry (2014) investigated the effects of weather conditions and "difficult" geologies (clay soils) on the quality of resistivity and earth resistance data. The study expanded the problem of seasonality effects to specifically focus on how and why multiple variables (e.g. moisture content, soils) influence the detection and interpretation of electrical anomalies caused by anthropogenic interference. Through a series of earth resistance, electrical resistivity imaging (ERI), and EMI surveys of four known archaeological features over 16 months, and subsequent statistical analysis of the survey data and cumulative moisture balance data, Fry determined that seasonality alone is not a suitable predictor for the detectability of archaeological features (Fry 2014, p. 189, 264). Instead, weather trends (as cumulative moisture balance) and the porosity of local soils are more reliable as predictors of the detectability of archaeological features even with the noted unpredictability in the effects of changes in weather conditions (Fry 2014, p. 264-266).

While the majority of southern and Northern Ireland is comprised of carboniferous limestone, sandstone, and shale, there are several regions of igneous bedrocks including basalts and granite which are less suitable for electromagnetic surveys. The regions with volcanic bedrock or rough terrain provide the opportunity to determine if the model will still perform well in unfavourable

conditions. The overall suitability of GPR for detecting a range of medieval monastic features in conjunction with the underuse of GPR in Ireland led to choosing Ireland as the first test region for the research project.
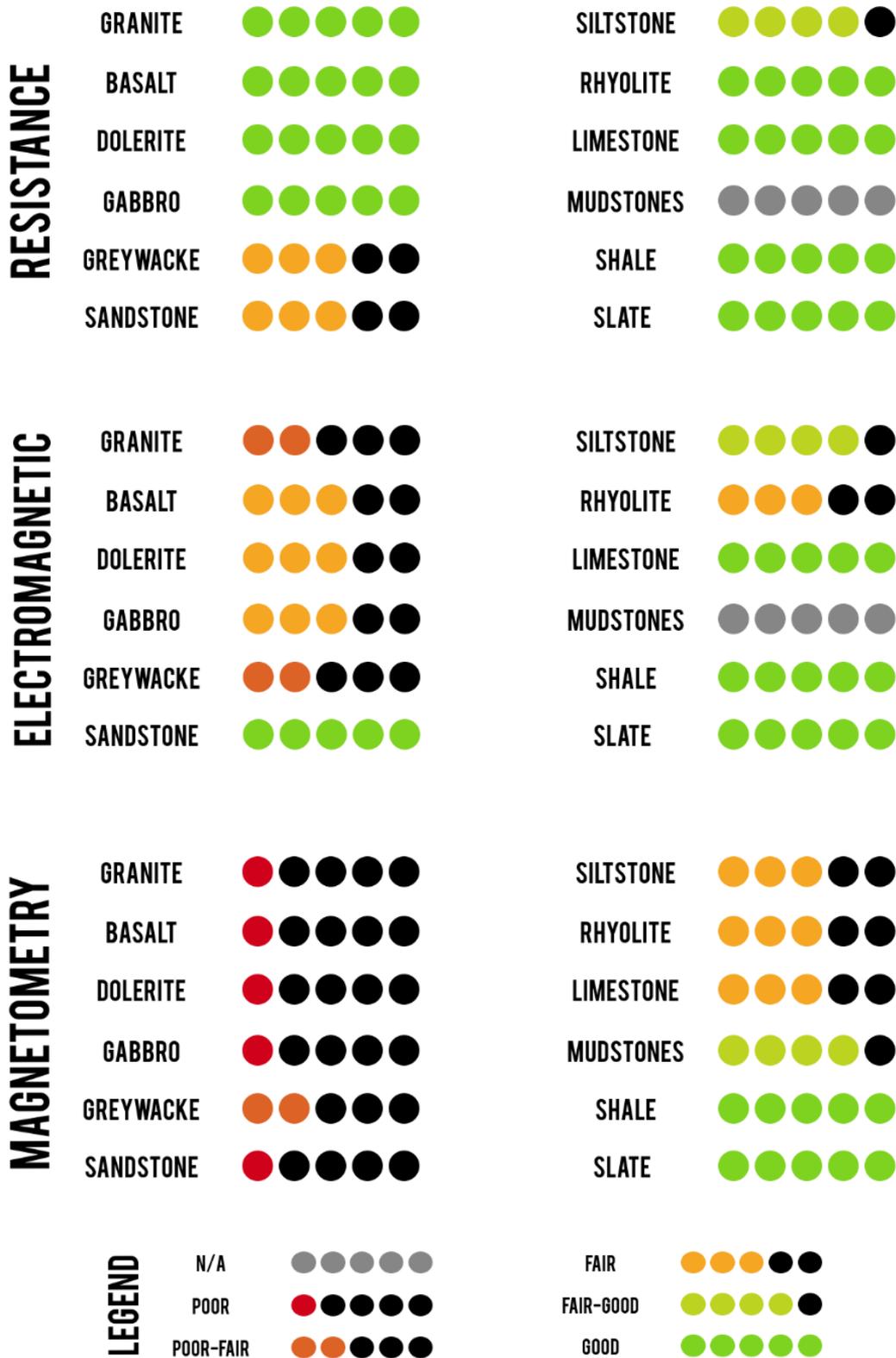
Figure 2: Suitability of resistance, electromagnetic (GPR and EMI) and magnetic techniques for a range of local geologies. N/A indicates there is no data available (adapted from Bonsall et al. 2014, p. 36)

Along with the geological factors, the survey parameters will have the most substantial impact on whether or not a surveyor can detect the expected target using a suitable technique. National and international guidelines are available to provide guidance on choosing the appropriate geophysical method, survey resolution, and data processing methods. However, while these guidelines provide accurate and suitable advice for surveyors and novices, the text format of the guidelines does not encourage their everyday use. As more local, amateur archaeologists try to incorporate geophysical survey into their projects, it is crucial to provide them with clear, accessible guidance on how to survey appropriately in a given set of conditions to minimise the risk of obtaining poor quality data. As previously established by Conyers (2013), GPR surveys are one of the most difficult to carry out, and the acquired datasets require more training and experience to interpret than most techniques. Therefore, it is necessary to improve the data collection process in order to improve the interpretation of the data. No one, geophysicist or amateur archaeologist, is immune to making survey or interpretation errors, especially with a complicated technique such as GPR. However, these errors can be minimised by implementing an interactive, input-output format for existing survey guidelines. An interactive approach to determining the appropriate survey parameters will hopefully encourage more practitioners to seek guidance before surveys and improve data quality.

The critical survey parameters (the traverse and sampling intervals) also affect survey results; however, these are easily adaptable for most techniques. Intrinsically, to increase the likelihood of detecting a target, the traverse interval should be no more than half the minimum dimension of the target(s), preferably without sample bias in any one direction, with minimal interpolation between data as well as an adequate characterisation of the background material to differentiate between the target and background responses (Green and Cheetham 2016; Green and Holmes 2017). From this, an interactive guidelines model for determining appropriate survey parameters while accounting for factors including local geology, presence of modern anthropogenic activity, weather conditions, land use, has the potential to enhance data collection methods which in turn will improve data interpretability (David et al. 2008a; Bonsall et al. 2014; Schmidt et al. 2015).

Further to this, introducing a computational approach into the interpretation process has the potential to reduce human error and improve the rate of identifying true positive results. In adapting existing computational geometry algorithms and machine learning libraries to classify responses in GPR data automatically, processing and interpretation speed are, in most cases, increased, and interpretation becomes more accurate as the intra- and inter-observer error is reduced.

Aside from not being detectable, graves are often overlooked in a dataset for several reasons, e.g. magnitude of the response, size of the dataset, and experience of the archaeologist. Overlooking or

misinterpreting grave-related responses can have severe impacts on commercial projects, particularly on the cost and timescale. By identifying potential graves present on a site prior to the start of the excavation stage of an investigation, archaeologists and the client(s) can prepare a strategy for dealing with the remains or mitigating the impact on them. While the magnitude and morphology of responses in a dataset may be indicative of inhumation(s), they may not be interpreted by many geophysicists as the responses of many archetypical interments are similar to those of other small-scale ground disturbances. Incorporating geophysical, spatial, and remote sensing data allows for the creation of an information system that can be used by the commercial sector and forensic archaeologists to detect interments during the pre-excavation stage of an investigation, thus maximising the potential for the recovery of human remains. By eliminating the subjective human factor in data analysis, the effectiveness of the program can be determined and present a more objective, thorough interpretation.

1.2. Assessment of the 'Expert' – Rationale for a Machine Learning Approach

Studies of machine learning and professionals in other fields, e.g. medicine, have shown that machine learning can assist and out-perform experts in data interpretation. Tschandl et al. (2019) compared the accuracy of human experts and machine learning methods in classifying pigmented skin lesions in dermatoscopic images. In comparing the responses from 511 medical professionals against those from the top three highest performing machine learning algorithms, Tschandl et al. (2019, p. 938) found that the machines out-performed the human interpreters. When compared with all respondents, the machines predicted an average of 2.01 more correct diagnoses than the humans. When compared with the responses from experts with more than 10 years of experience, the machines predicted an average of 6.65 more correct diagnoses than the experts (Tschandl et al. 2019, p. 938). Other studies, such as that carried out by Olczak et al. (2017, p. 584) on the classification of radiographs, which used outdated machine learning networks found that experts and the machine learning algorithm achieved similar accuracies on classifying fractures – 83% for the machine and 82% for the experts.

The state of human-led interpretation of graves in GPR was briefly assessed in this research project through a questionnaire. The questionnaire was distributed to individuals who were experienced in GPR survey in order to create a baseline evaluation of the agreement between multiple human interpreters without the potential for inexperienced surveyors and non-geophysicists to skew the results. The questionnaire was limited by the rage of data available that could be widely shared and the limits of the survey hosting platform. The questionnaire aimed to determine if human interpreters

would provide the same interpretations of the data, and subsequently if they were able to correctly identify responses from graves in GPR data as well assess their confidence in their interpretations. The interpretation and its confidence are also measured by the Reilig for Radargrams app during its interpretation of images, making the human- and machine-led interpretation comparable. The results of the machine learning interpretation are in Chapter 8.2.

1.2.1.   *Participant Characteristics*

The target population for the questionnaire was surveyors currently working primarily with GPR data and experts in GPR data collection and interpretation. The restriction of the population reduced the possibility of extraneous responses which might have skewed the outcome of the questionnaire. Responses were anonymised and cannot be traced back to a specific participant, only the 'years' experience' class. These restrictions limited the sample size and the analysis that could be performed on the responses.

10 responses were received from individuals with a range of experience levels. The majority of participants (n=6, 60%) had more than 10 years' experience in GPR survey. All participants had at least one year of experience in GPR survey.
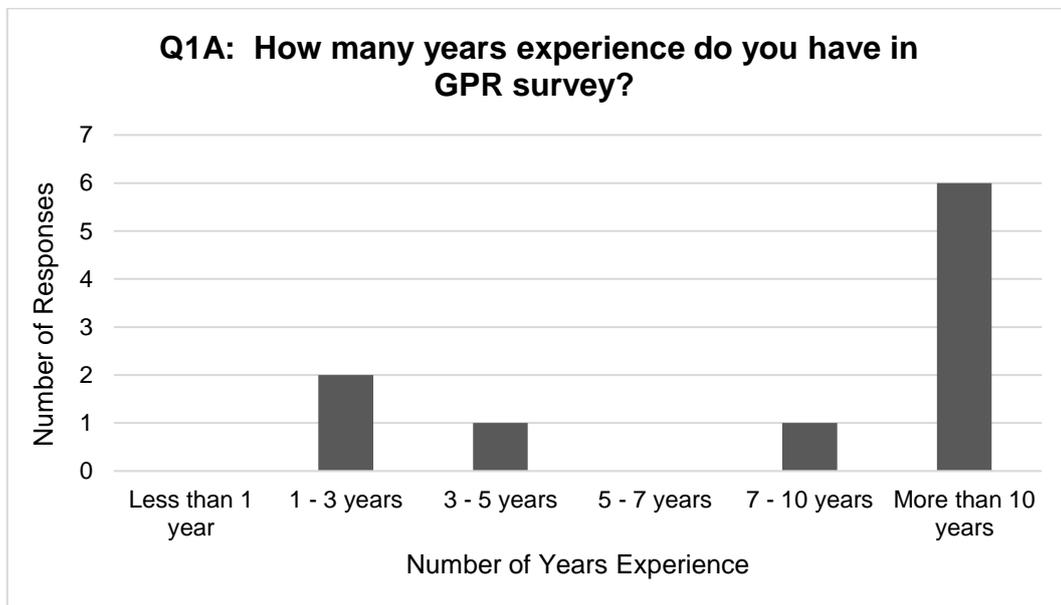
Figure 3: Distribution of participants' years of experience in GPR survey with most participants having more than 10 years of experience.

Participant experience was almost evenly distributed between commercial and research geophysics. A majority of the participants (90%) had experience in both commercial and research geophysics.

Q1B: Has your experience with GPR survey primarily been in research or commercial settings?
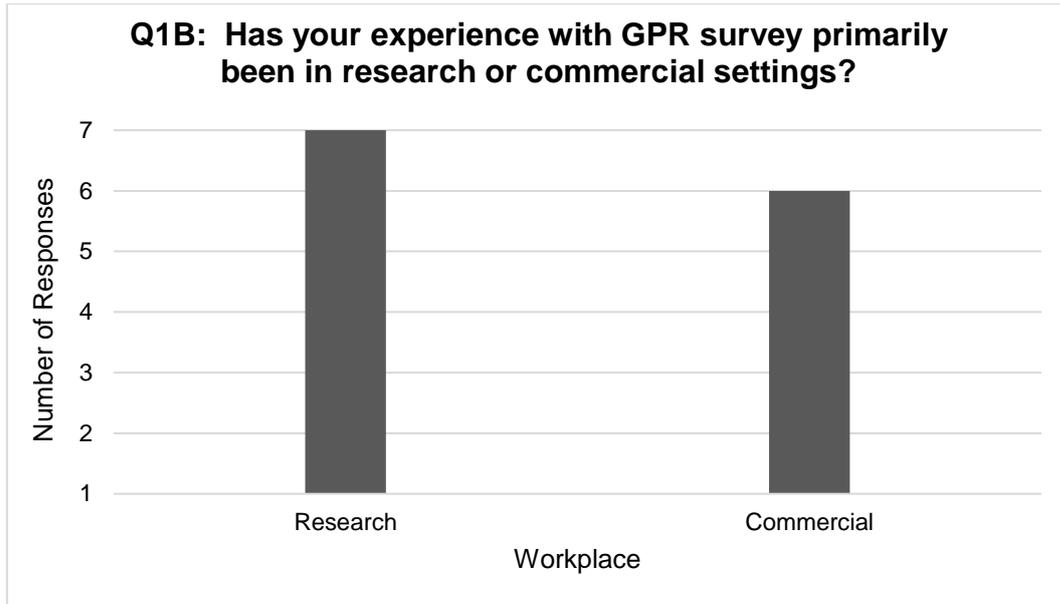
Figure 4: Distribution of participants' experience with GPR in commercial and research settings, showing that both feilds are represented.

1.2.2. *Data Visualisation Methods*

Participants were asked which methods they use to visualise GPR data for interpretation. The majority of participants used both timeslices and radargrams when interpreting data, with 60% of participants using animations (either of radargrams or timeslices) as well. No participants were using more sophisticated methods of data visualisation to interpret GPR data. One participant uses an "envelope of timeslices, and usually no more sophisticated processing..." to visualise GPR data.

Q2: When interpreting GPR data, you use:
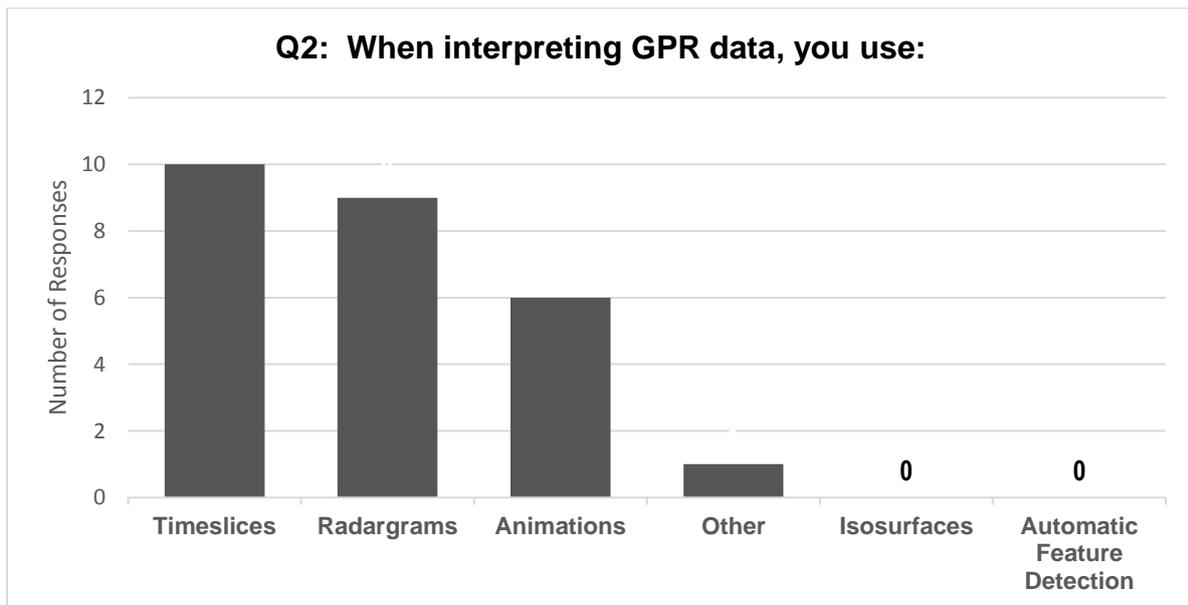
Figure 5: Distribution of participants' preferred visualisation methods for GPR data

1.2.3. *Interpretation*

Participants were presented with two radargrams and two timeslices to interpret. Each image was divided into four quadrants (labelled A – D). Participants were asked to identify whether there was no grave, a grave, or multiple graves present in each quadrant. The responses for each image are presented in Figure 6 - Figure 13. Grave responses are present in Radargram 1 Quadrant B, and Radargram 2 Quadrants A and C.

Q3A: In the image which quadrant(s), if any, contains grave-like response(s)?
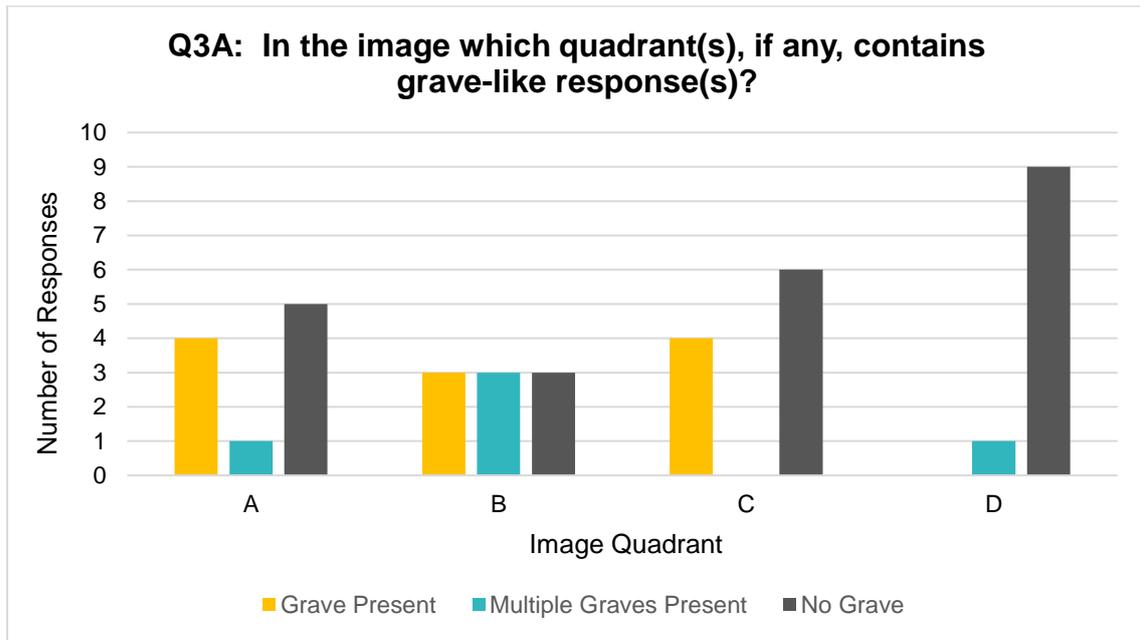
Figure 6: Distribution of all participants' interpretation of graves in Radargram 1

Q4A: In the image which quadrant(s), if any, contains grave-like response(s)?
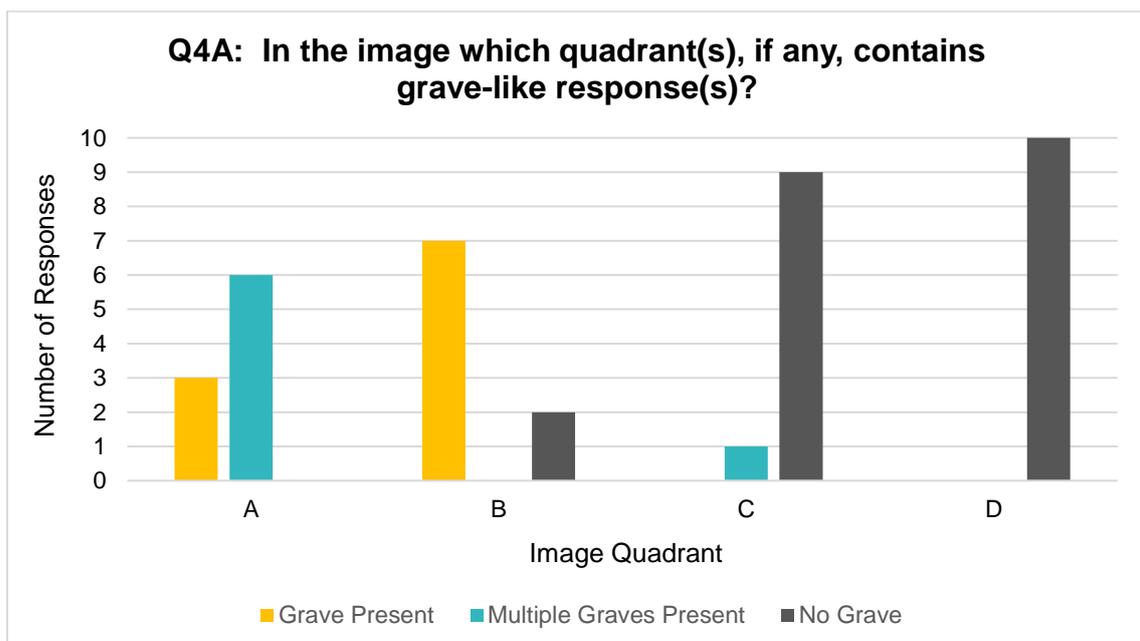
Figure 7: Distribution of all participants' interpretation of graves in Radargram 2

Participants were in total agreement, defined here as all participants providing the same response, on the interpretation one radargram quadrant (Radargram 2, Quadrant D). As demonstrated by the responses in Figure 6 and Figure 7, aside from Quadrant D in Radargram 2, while there were varying levels of agreement between participants, there was no total agreement between participants on the presence of graves in each radargram. When the responses from those participants with more than 10 years of experience in GPR survey are isolated, as shown in Figure 8 and Figure 9, total agreement amongst participants increases to two quadrants in Radargram 2 (25% of radargram quadrants).
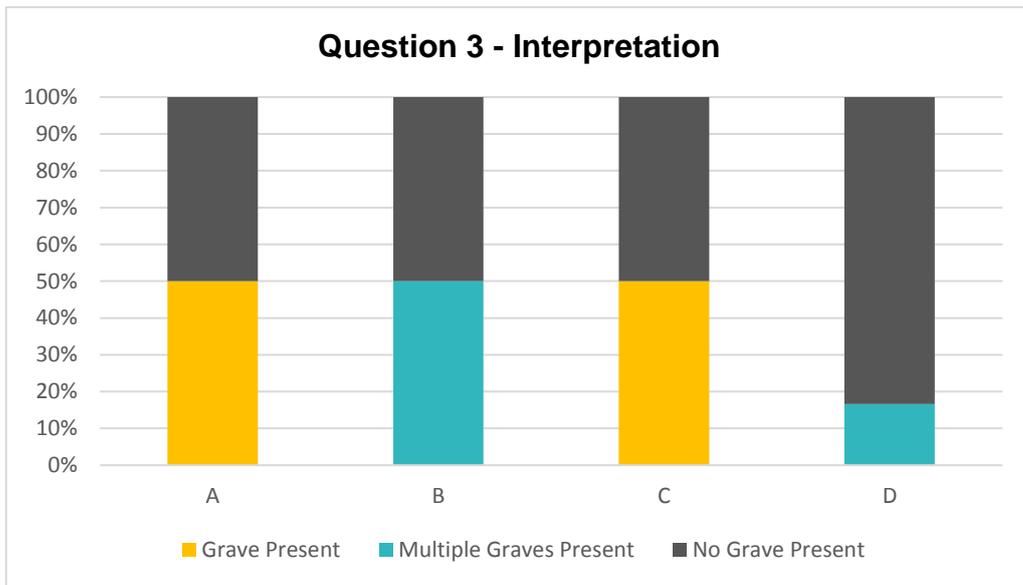
Figure 8: Distribution of the interpretation of graves in Radargram 1 by participants with more than 10 years of experience in GPR survey
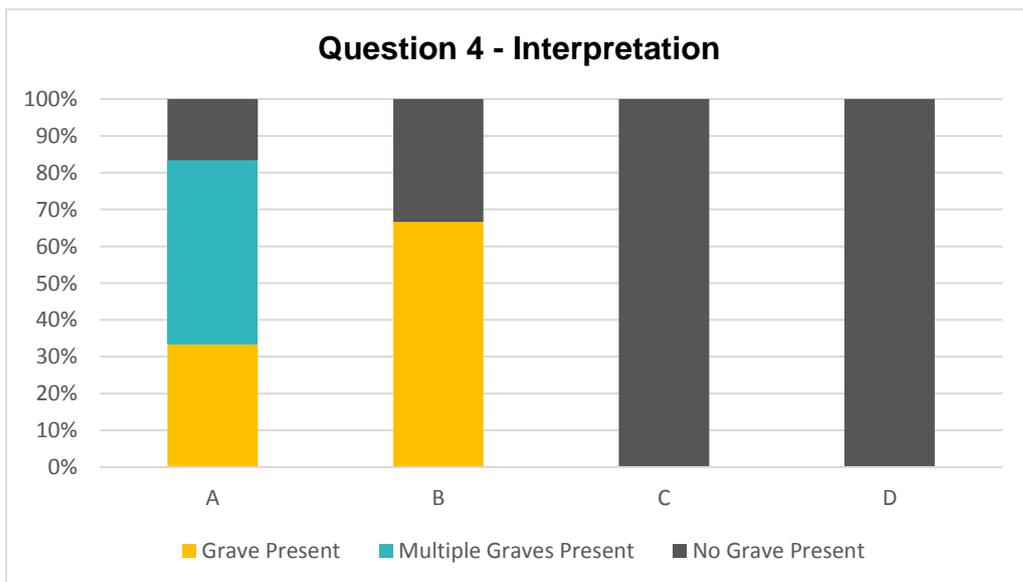
Figure 9: Distribution of the interpretation of graves in Radargram 2 by participants with more than 10 years of experience in GPR survey

Participants agreed more on the interpretation of the timeslice examples. There were no quadrants in which all participants were in total agreement on the interpretation of the presence of graves; however, there was less variation for each quadrant, as demonstrated in Figure 10 and Figure 11. Participants gave the most correct responses for Radargram 2 Quadrant D.
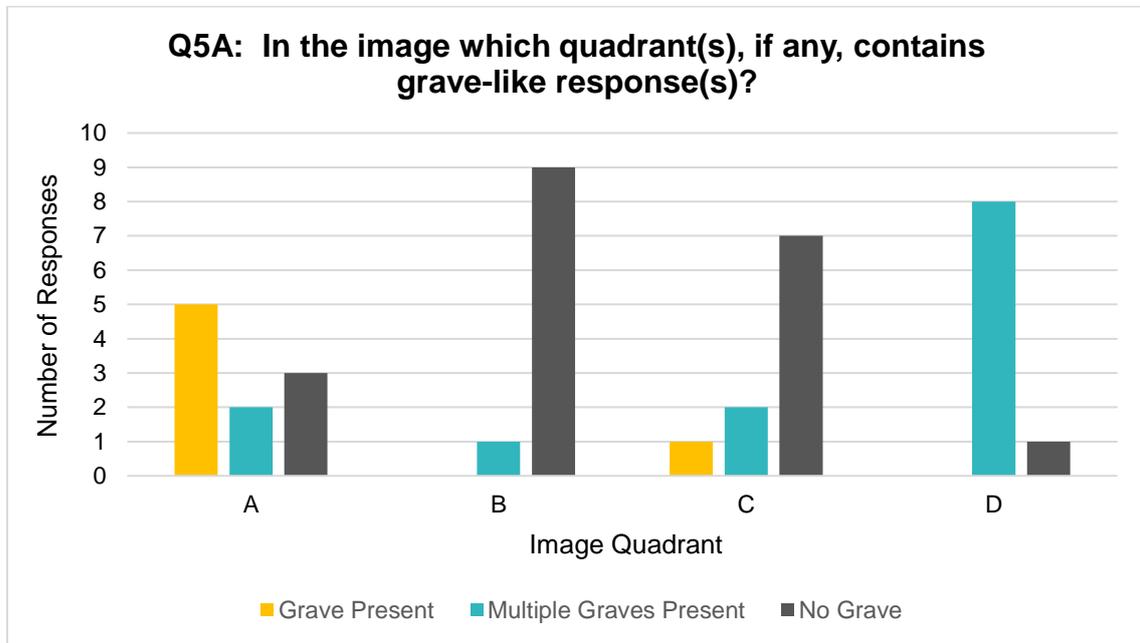
Figure 10: Distribution of participants' interpretation of graves in Timeslice 1
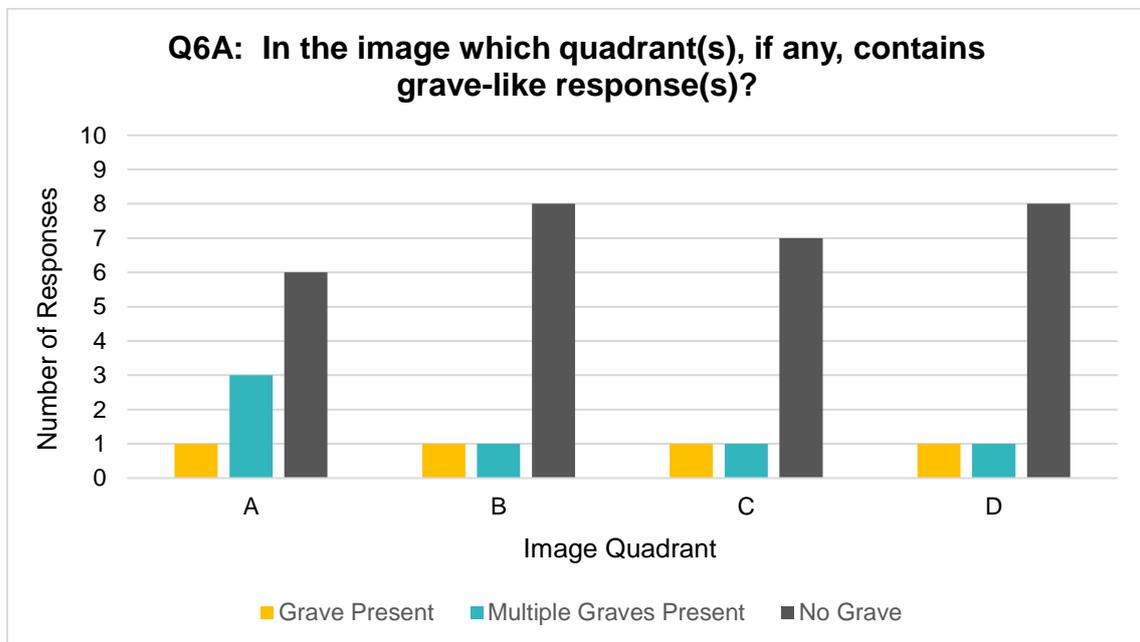
Figure 11: Distribution of participants' interpretation of graves in Timeslice 2

As demonstrated in Figure 12 and Figure 13, when participants with less than 10 years of experience are removed the number of quadrants in which the participants are in total agreement on the interpretation increases to two.

Question 5 - Interpretation
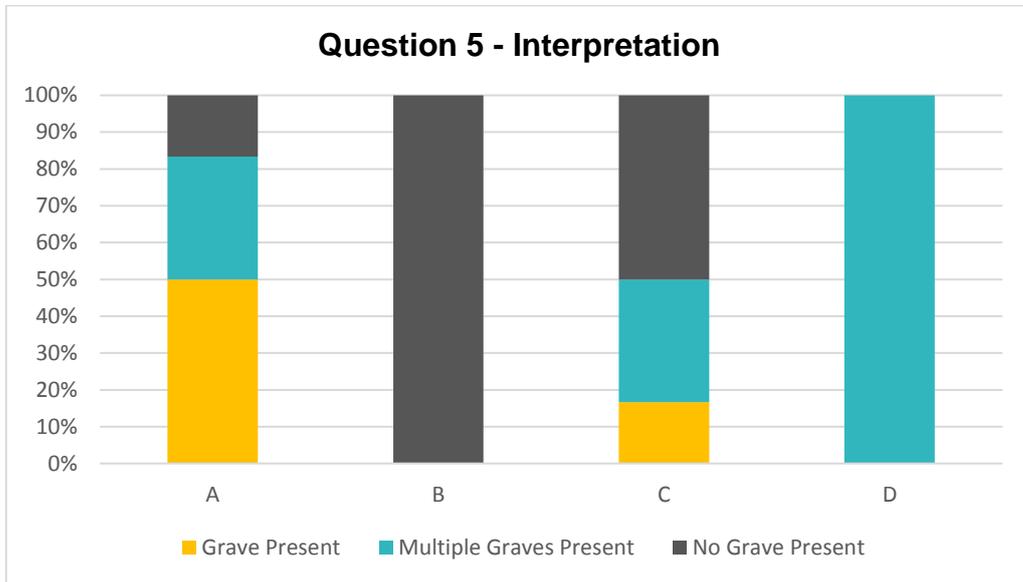
Figure 12: Distribution of the interpretation of graves in Timeslice 1 by participants with more than 10 years' experience in GPR survey

Question 6 - Interpretation
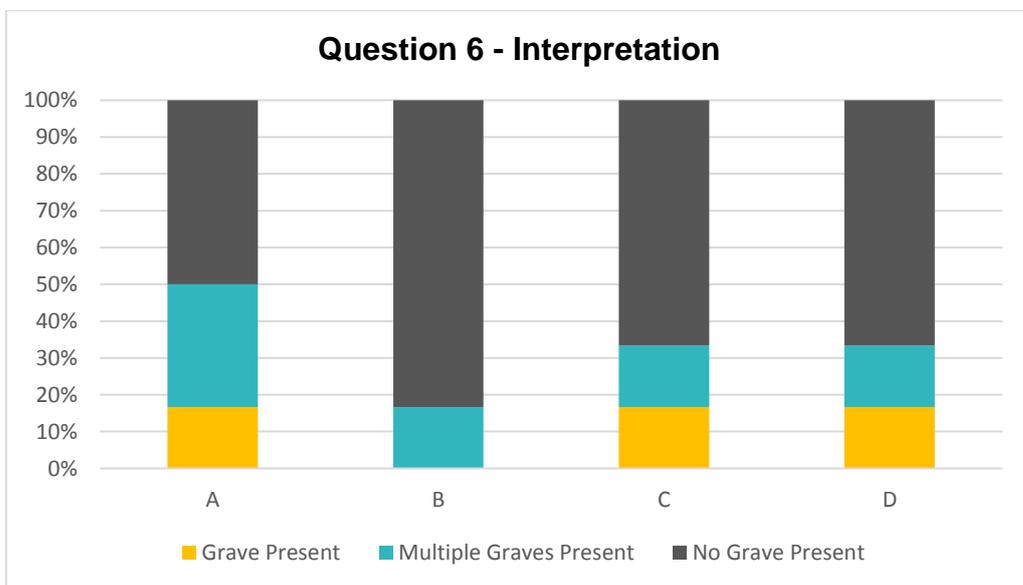
Figure 13: Distribution of the interpretation of graves in Timeslice 2 by participants with more than 10 years' experience in GPR survey

Participants were more successful in correctly interpreting the timeslice examples compared to the radargram examples. They were most successful on Timeslice 1 Quadrant B.

The lack of consensus among participants in interpretation of the radargrams and marginal improvement in the interpretation of the timeslices is a further indication of the need for a method for cohesive and quantifiable interpretation of GPR data.

1.2.4. *Confidence in Interpretation*

Alongside their interpretations, participants were also asked to rank their confidence in their interpretation for each quadrant in an image. Participants were able to indicate a low, medium, or high confidence in their interpretation. As demonstrated by Figure 14 - Figure 17, participants indicated they had low confidence in the majority of the interpretations. It should be noted, however, that this low confidence could be a result of the amount of data they were provided with and if it was processed or plotted different from their usual methods.
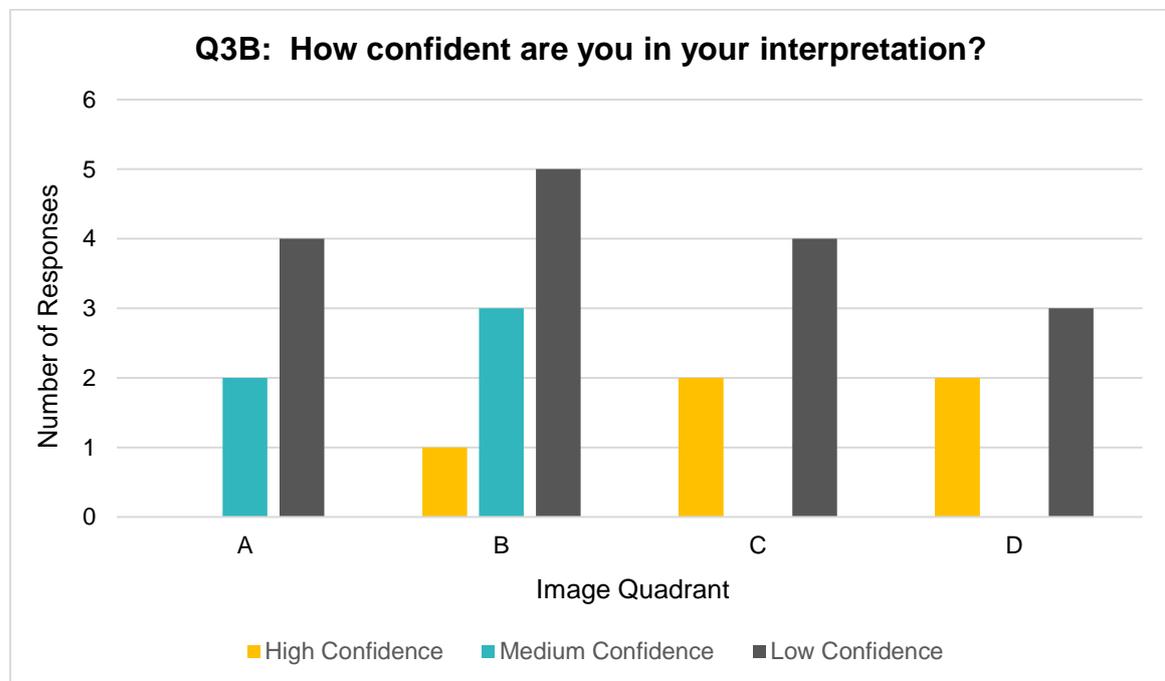
Figure 14: Distribution of participants' confidence in their interpretation of quadrants in Radargram 1
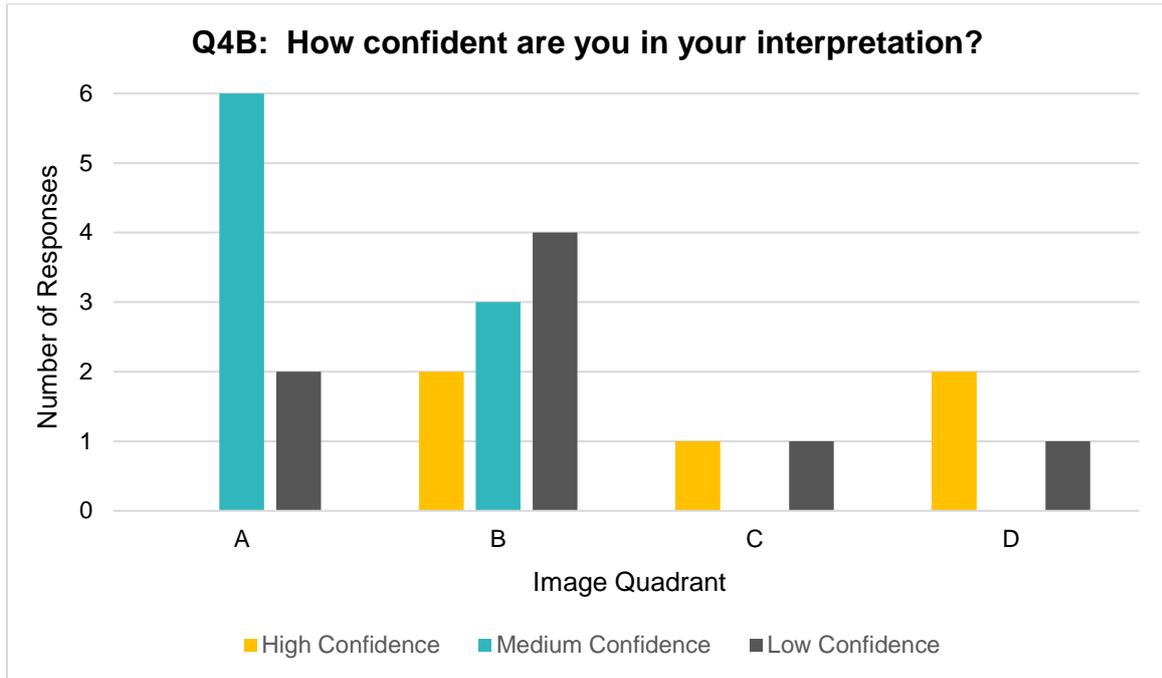
Figure 15: Distribution of participants' confidence in their interpretation of quadrants in Radargram 2
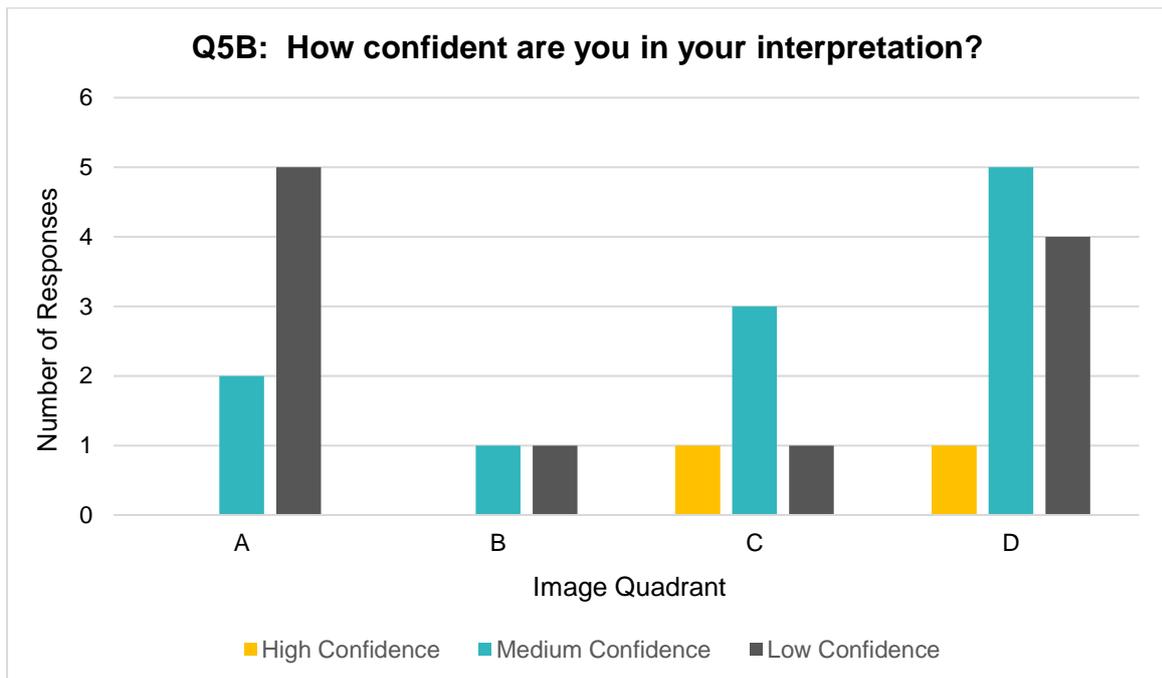
Figure 16: Distribution of participants' confidence in their interpretation of quadrants in Timeslice 1
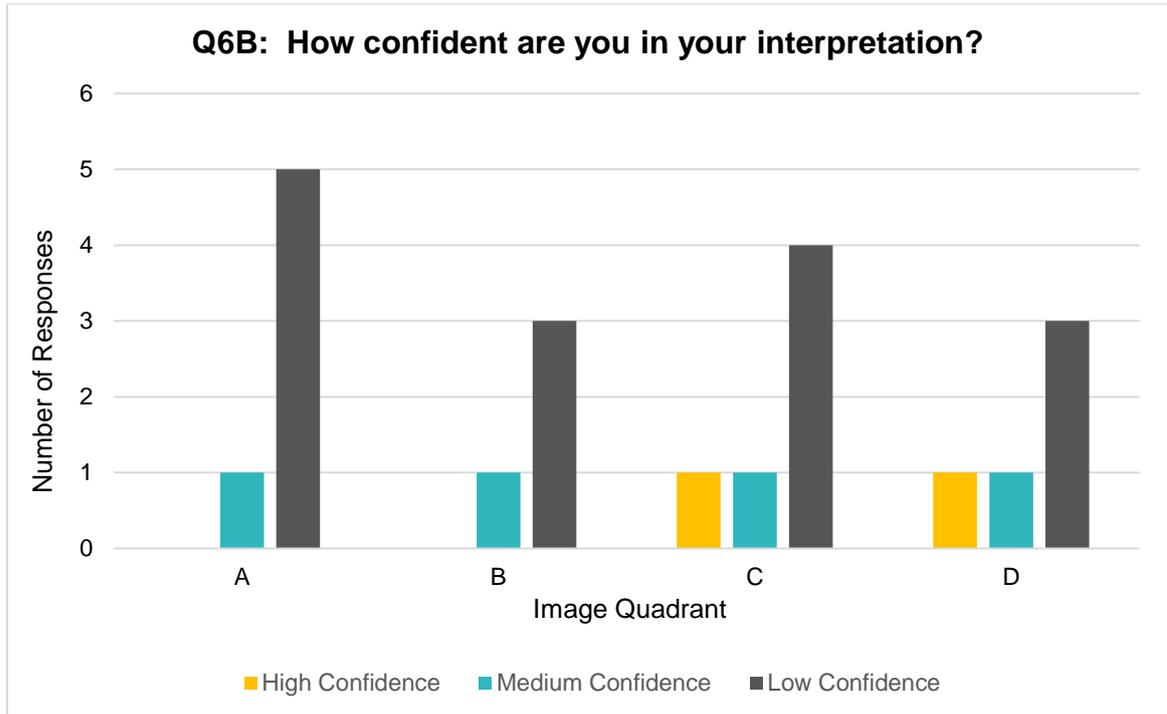
Figure 17: Distribution of participants' confidence in their interpretation of quadrants in Timeslice 2

When the responses from participants with less than 10 years' experience are excluded from the dataset, participants' confidence in their interpretations is still primarily low. While there was more agreement between participants in the interpretation of the timeslices compared to the radargrams, participants' confidence was lower in their interpretation of the timeslices. Figure 18 and Figure 19 show the ranked confidence of participants with more than 10 years of experience in their interpretations of the radargrams and timeslices.

Confidence in Interpreting Radargrams
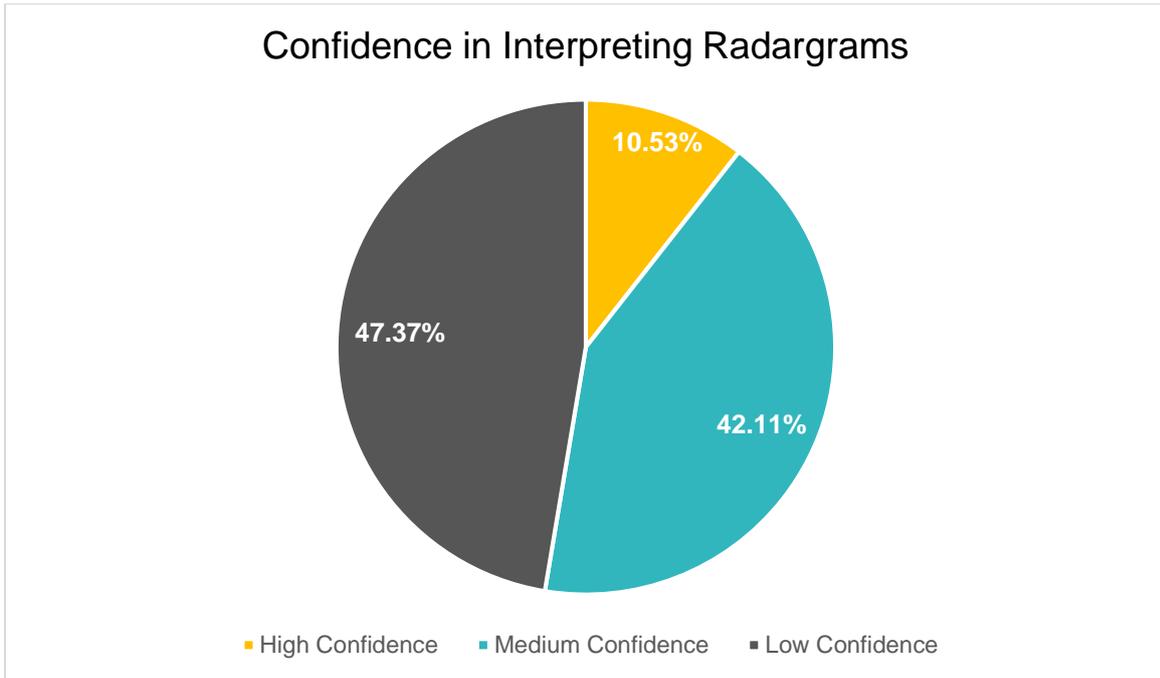
Figure 18: Confidence of participants with more than 10 years of experience in interpreting Radargram 1 and Radargram 2

Confidence in Interpreting Timeslices
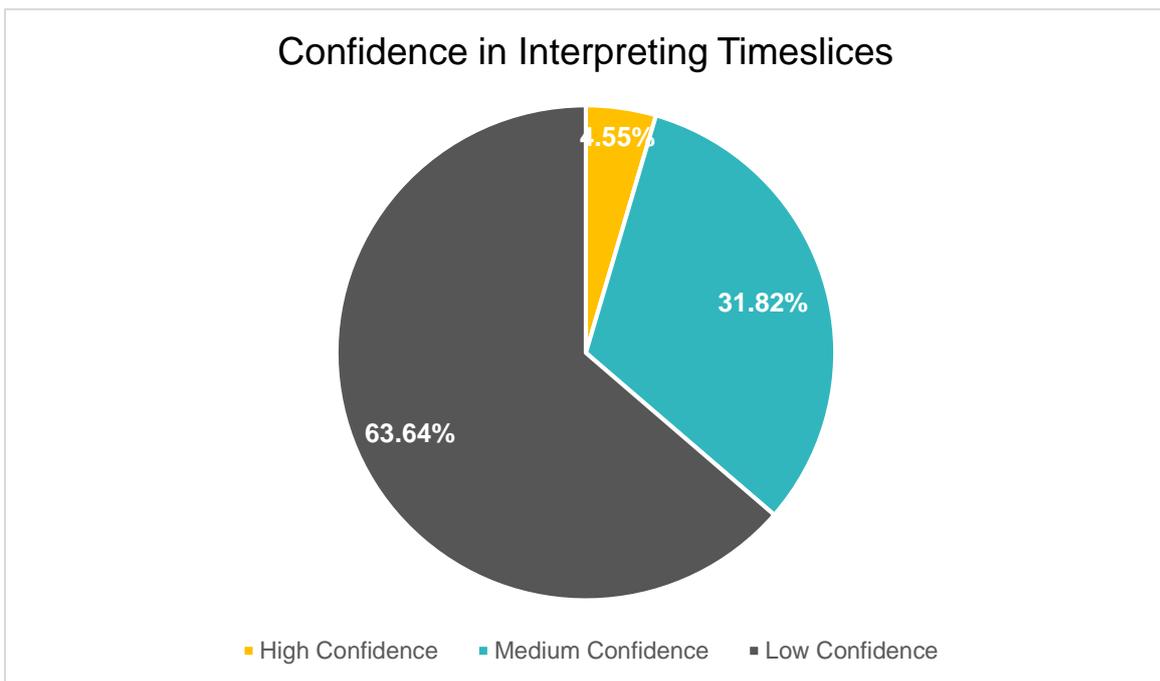
Figure 19: Confidence of participants with more than 10 years of experience in interpreting Timeslice 1 and Timeslice 2

Although the ranked confidence system is simplistic, the high rate of participants' low confidence in their interpretations is concerning if it is not a result of the amount of data provided. This low confidence further supports the need for a method of probability scoring interpretations to appropriately qualify their likelihood of being correct interpretations and assist human interpreters.

1.3. Aims and Objectives

1.3.1. *Aims*

Responses from graves in GPR data are often misinterpreted or overlooked by human interpreters. While this can negatively impact commercial archaeology companies and their clients as skeletal remains exponentially increase the cost of a project, there is also an ethical issue surrounding the disturbance of human remains. The overall aim of this project is to produce digital tools to improve data acquisition and interpretation of graves in archaeological GPR surveys to improve the rate of the pre-excavation detection of graves. This work is framed with relevance to the use of digital tools in commercial archaeology and community groups in Ireland and the UK,

1.3.2. *Objectives*

- Assess the appropriateness of current practices for the identification and interpretation of graves in GPR data;
- Develop an interactive decision-making tool for determining appropriate survey parameters for surveys in Ireland and the UK from existing guidelines;
- Develop input-output user-friendly automatic classification software for detecting and probability-scoring grave-like responses in GPR data;
- Apply machine learning to a range of sites in Ireland (five study areas);
- Assess the potential impact of the developed tools on commercial archaeogeophysics;
- Test the ability to develop a training dataset for standard modern clandestine burial practices;
- Assess the potential impact of the computational tools on geophysics applied in a forensic context.

1.4. Contribution to Knowledge

It is the aim that the research project and its outputs will contribute towards the improvement of the standard of archaeogeophysical survey and data interpretation as well as towards understanding the potential and challenges of detecting burials. As such, this research demonstrates that choosing the most appropriate survey parameters play a vital role in determining the detectability of burials, and therefore data should be collected systematically at the highest density possible within the means of each project. By integrating computational tools in archaeogeophysical surveys where GPR is a significant component, there is the potential to reduce the modern impact on archaeological skeletal remains as well as the monetary impact on commercial clients. Most importantly, and at the most

basic premise, the research aims to improve the confidence, consistency, and speed of interpreting GPR data.

1.5. Thesis Format

This thesis is comprised of an introductory chapter which details the research agenda, its components, the background to the research project, and the real-world implications of the research outputs.

Following the introduction are three chapters detailing relevant literature and background to the geophysical survey component of the project. Chapter 2 provides analysis and collation of the historic and excavation records for medieval burial practices in Ireland and the UK. From this discussion, Chapter 3 includes a discussion on the current survey guidance for the UK and Ireland and how their use can be encouraged to improve data quality and the detection of graves. Chapter 4 provides a detailed overview on the principles of the techniques often used in archaeogeophysics (as discussed in Chapter 3), with a particular focus on GPR, as well as their applications, limitations, data formats, and how these affect the detectability graves and grave goods using each technique.

From this, Chapter 5 outlines the methodology for the interactive survey parameters tools created and their preliminary use.

Chapters 6 - 8 are dedicated to machine learning.

Chapter 6 outlines current machine learning methods and their established applications to archaeological and geophysical data.

Chapter 7 outlines the methods used to create the neural network and machine learning algorithms. This chapter also compares the suitability of the eight pre-trained models tested using the GPR dataset.

Chapter 8 outlines the methods used to collect and interpret the data from case study surveys and compares the human-led interpretation methods with the machine-led interpretations. This chapter also includes analysis of the results from the automatic classification software, Reilig, and the performance of the system compared to the performance of archaeogeophysics professionals, as well as the results from the case study surveys

Chapter 9 includes additional preliminary tests on proxies for modern clandestine graves in GPR data.

Chapter 10 continues the discussion on the feasibility of applying the tools created in the project to commercial and research geophysics, as well as to forensic investigations, with particular focus on how real-world scenarios may affect the performance of the machine learning models.

Chapter 11 presents the overall conclusions, including the outputs and outcomes from the research and proposed future work.

2. ARCHAEOLOGICAL BACKGROUND

2.1. Introduction

The focus of this project is on the improvement and interpretation of geophysical data, primarily how machine learning approaches can be applied in "real-world scenarios." To explore these applications, the targeted archaeological features must first be discussed in context. Initial applications of machine learning in the project were limited to medieval graves as these are the most commonly encountered in commercial archaeology (Advisory Panel on the Archaeology of Burials in England 2017). However, discussion on the later application of machine learning to modern graves is provided in Chapter 9 and Chapter 11.2.2.

Burial traditions in Western Europe can often be explored together; however, the factors that influence burial styles, such as the arrival of new cultural groups or political mandates, vary across the region. There is a vast record of archaeological knowledge in Ireland due to the excavation and recovery strategies used in commercial archaeology over the last 30 years, as well as the presence of *in situ* remains across the country. While it could be argued that the surviving archaeology in Britain is less substantial than in Ireland due to development practices, the excavation records and presence of upstanding archaeology in Cornwall and South Wales demonstrate similarities in lifestyle, burial practice, and, at one point, language. As geophysical survey methods have proven capable of detecting medieval archaeological features and there is a wealth of information on these features, this research project was limited to burial styles of the medieval period, primarily in Ireland but with several similar examples from Britain.

2.1.1. *Chronological Divisions*

This study, which only aims to determine the viability of machine learning, focuses on common medieval grave types in Ireland and the UK. The timescales for the medieval period in Ireland and England are similar, with some minor transition differences due to variances in the political climate (see Figure 20). The timescale for the medieval period in Ireland and the UK is similar, with some minor transition differences due to variances in the political climate. While there is evidence for some overlap in burial traditions between phases, categorical burial styles can be associated with specific phases during the medieval period. It is also important to note that the minor discrepancies in the structure of the general population and the transition phases between periods in Ireland and Northern Ireland, England, and Scotland also led to minor variations in the burial practices of the countries.

While there is evidence for some overlap in burial traditions between phases, categorical burial styles can be associated with specific phases during the medieval period. It is also important to note that the minor discrepancies in the structure of the general population and the transitions between historical periods in Ireland and England also led to minor differences in the burial practices of the two countries.   The typological burial practices for each country are discussed hereafter.   The morphological and conductive properties for each typology used to create the training dataset and as classifiers are discussed in this chapter as well as Chapter 4.

IRELAND

AD 367
Offensive on Britain by the Irish, Picts and Saxons

AD 431
Pope sends Palladius to Ireland as Bishop

AD 432
St Patrick arrives in Ireland

c. AD 550-650
Growth of monasticism in Ireland

AD 563
Foundation of Iona by Columba

c. AD 670-700
Tírechán and Muirchú produce hagiographical works on St Patrick

AD 697
Proclamation of the 'Law of Innocents'

AD 795
First Viking raid on Iona, Rathlin, Inishmurray, and Inishbofin

AD 836
Increase in Viking raids

AD 914
Second wave of Viking raids

AD 1142
Foundation of the first Cistercian house in Ireland

AD 1169 (Anglo-Norman Invasion)
English military leaders arrive in Ireland

AD 1175
Treaty of Windsor between Henry II and Rory O'Connor (King of Ireland)

AD 1315
Edward Bruce invades Ireland

AD 1366
Statutes of Kilkenny – Prevent English settlers using the Irish language

AD 1509
Accession of Henry VIII

AD 1536-37
Meeting of the Irish Reformation Parliament

AD 1541
Parliamentary meeting declares Henry VII King of Ireland
Establishment of the 'Surrender and Re-grant' programme

ENGLAND

c. AD 600-1066
Anglo-Saxon period

AD 1066-1272
Norman Conquest
Reign of Edward I

AD 1169 (Anglo-Norman Invasion)
English military leaders arrive in Ireland

AD 1175
Treaty of Windsor between Henry II and Rory O'Connor (King of Ireland)

AD 1272-1540
Late Middle Ages
Tudor Era

Figure 20: Timeline of events and periods relevant to changes and variation in burial style (adapted from Barry 2003, Duffy 2005a, 2005b, Horning et al. 2007, Waddell 2010, Historic England 2019)

2.2. Archaeology of Ireland

Activity in medieval Ireland cannot be analysed irrespective of earlier activity in the Iron Age and Early Christian period as the political climate, landscape changes, and external influences impact later societies, politics, and economies. Even as a relatively remote island, the inhabitants of Ireland

had contact with Central Europe (Hallstatt culture) and the Mediterranean (La Tené) before the early medieval period. They yet remained politically divided amongst themselves (Stout 2018). Early written records by the Irish stress that the island was divided into northern and southern halves along glacial ridges formed between the two *Áth Cliath* (modern Dublin and Galway). The northern half, *Leth Cuinn*, was the territory of Conn, and the southern half, *Leth Moga*, was Mug/Mog's territory (Stout 2018). These dynasties are later recorded in 6th century records. The northern and southern halves were further divided into five *cóiceda* (modern: *cúigí*), or provinces – Ulaid, Uí Néill, Connacht, Mide, Laigin, and Mumu (Stout 2018). The approximate boundaries of these *cóiceda* are presented in Figure 21.

Approximate Extent of Early Christian Provinces
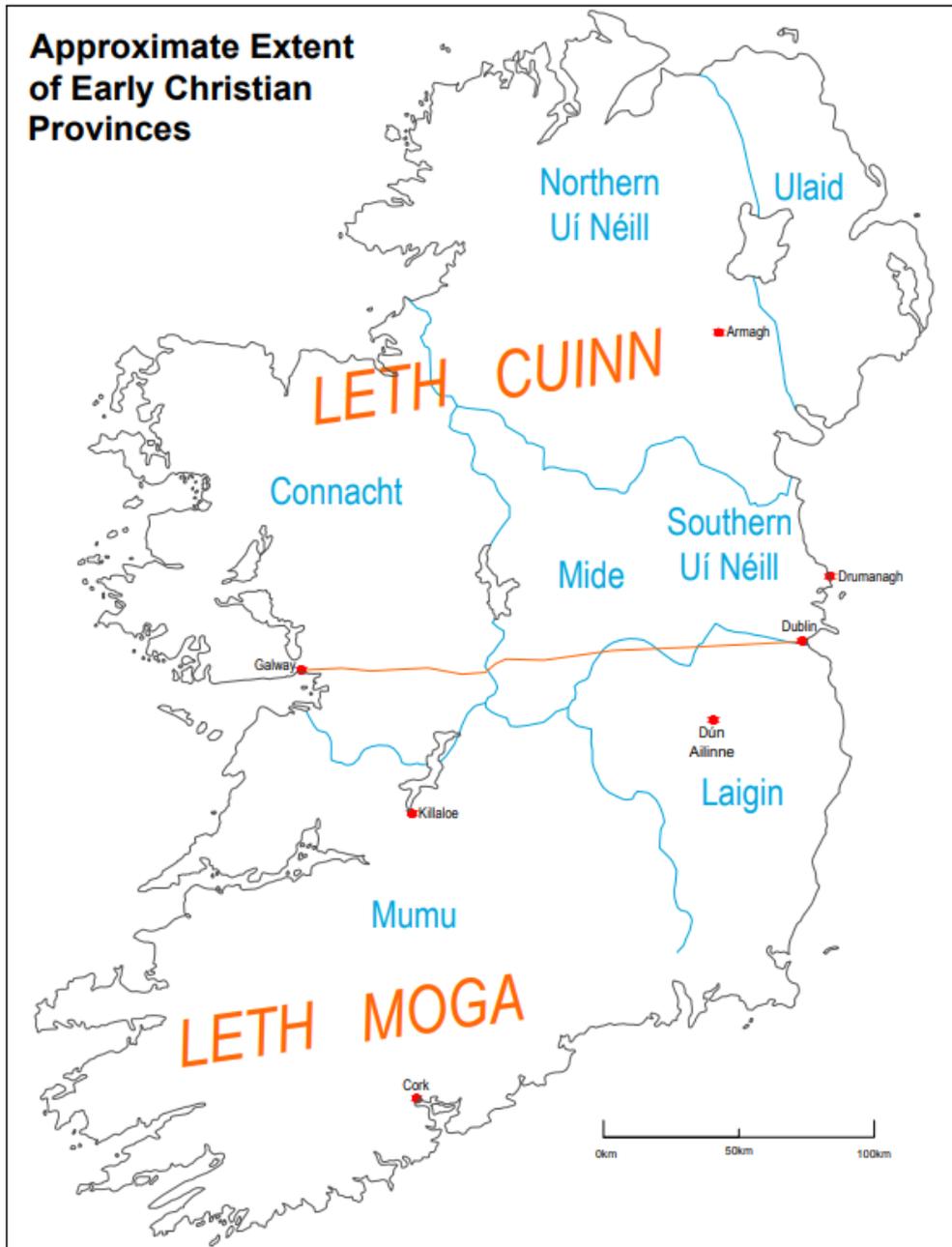
Figure 21: Map depicting *cóiceda* boundaries from *c*. 250 BC to 1167 AD (adapted from Stout 2017)

Around 300 AD landscapes throughout Ireland began to shift from woodlands to arable land and pasture, with a marked increase in agricultural activity. This shift was likely a direct result of the Roman impact on Ireland, mainly through contact with Roman Britain, which introduced new farming technologies, and subsequently foods, to the island. Irish raids on Britain, however, proved that some classes in Pre-Christian Ireland were literate. *Ogam* (also *ogham*), a twenty-letter alphabet created using intersecting lines (depicted in Figure 22), was initially used in Irish settlements in Britain to inscribe Irish names on stones, similar to that of Romano-Britain Latin inscriptions (Stout 2018). Later ogam spread into Ireland through Wales, where it was used to inscribe boundary-marking standing

stones and, eventually, ecclesiastical sites (Bhreathnach 2010). A number of these ogam stones are associated with burial grounds, indicating burial areas into the early medieval period.
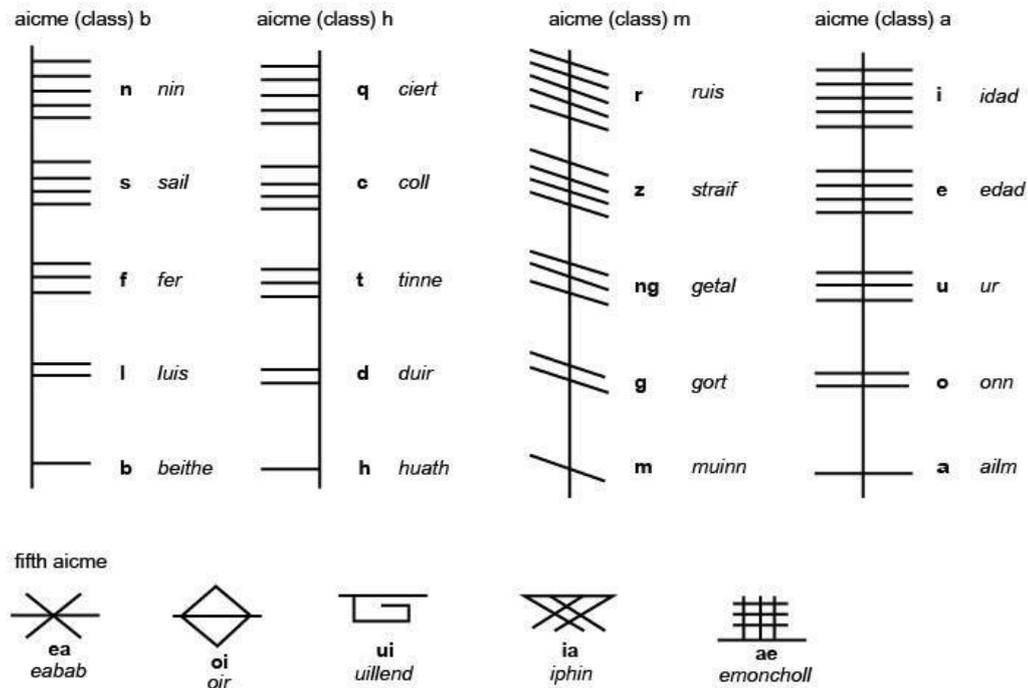
Figure 22: Ogam alphabet used to inscribe boundary-marking standing stones and stones to indicate ecclesiastical sites

2.2.1. *Early Medieval Ireland*

The beginning of the early medieval period in Ireland is marked by the pope's appointment of the first bishop in Ireland, Palladius, in AD 431. However, the spread of Christianity throughout Ireland is more often attributed to Patrick, a slave later becoming the patron saint of Ireland, as there is no evidence for Palladius in Irish records (Stout 2018). The temporal spread of Christianity throughout Ireland is evidence by place names and the annals. Before AD 500, the concentration of Christianity is demonstrated by the number of churches with the *domnach* place name; *domnach* being the Old Irish word for 'the Lord's Place' (from the Latin *dominicum*) (Stout 2018). Once the *domnach* place name went out of use, the annals demonstrate a rapid expansion of the Church, with the entire island becoming Christian by AD 700, as represented by the foundation dates of the churches shown in Figure 23 (Stout 2018). Until the ninth century, churches were timber-built structures; however, the 'burning' churches noted in the annals during the Viking raids contributed to the transition to stone churches (O'Sullivan et al. 2014, Stout 2018). The earliest documented stone church is Duleek (*Damliac* – 'house of stones') in County Meath.

Pre-Romanesque churches, such as Clonmacnoise in County Offaly, built around the tenth century were simple rectangular buildings with proportions between 1.3:1 and 2:1 and 'antae', a projection of

the side walls beyond the end walls (Stout 2018). Later in the 12th century churches, such as St Peter's Church in Waterford, were built with a rounded extension at their eastern end, similar to churches southern England at the time. Churches again changed style in the 12th century to the Romanesque style, a reflection of the Church reform at the time, which was typified by rounded arches and decorated mouldings (Stout 2018). This reform is reflected in the increased number of monasteries built in the 1140s. There was a shift in monastery styles in the 12th century as well with the arrival of the Cistercians, where monasteries became more isolated and estate-like to reflect the lifestyle demanded by the Rule of St Benedict (Stout 2018). The Rule of St Benedict required that in addition to the monks' daily prayers, they should practice farming or other forms of manual labour. The early medieval period ended in 1169 with the Anglo-Norman Invasion; however, the 1188 Annals of Ulster demonstrate many early medieval practices continued (Stout 2018).
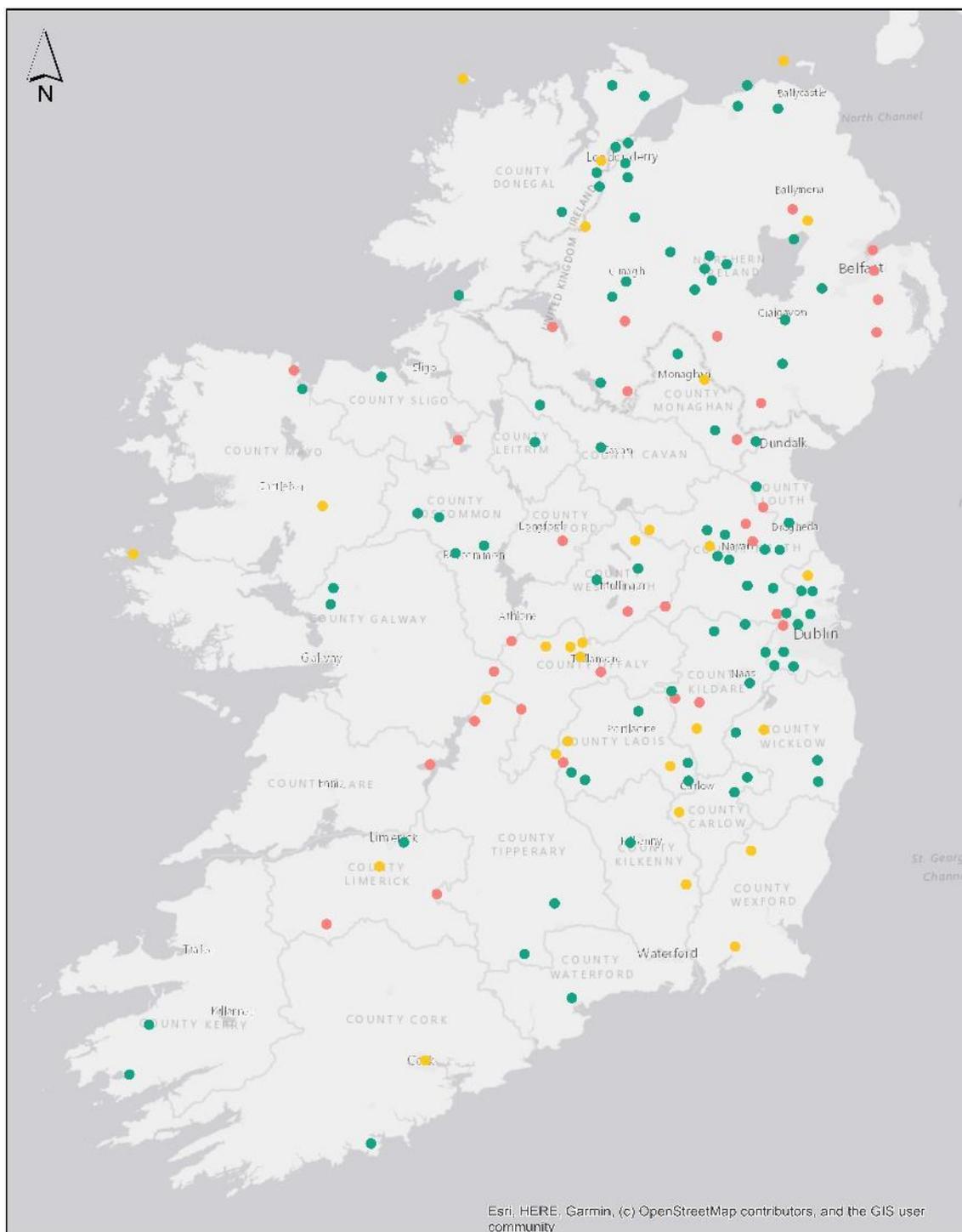
Figure 23: Approximate location of churches founded by AD 700 (adapted from Stout 2018)

2.2.2. *Medieval Ireland*

From the 12th century, churches and monastic building complexes followed a claustral plan – a layout which placed the cloister centrally, enclosed by the buildings. This layout places the church on the north or south side of the cloister with the refectory on the opposite side to the church and the chapter house and dormitories within the east range of buildings (Duffy 2005b). The first implementation of this plan in Ireland was at Mellifont Abbey in 1142 and continued until the Dissolution in 1536 (De Paor et al. 1969, Duffy 2005b). Monastic houses were primarily friaries (communities of brothers), with many the remaining houses being priories (communities presided over by priors/prioresses), and an exceedingly small minority being abbeys (communities presided over by abbots/abbesses) (Duffy 2005b). Abbeys and priories during the 12th – 13th centuries were primarily associated with Augustinian canons and Cistercian monks. Augustinian priories frequented urban areas as they required minimal space, were in Ireland at the time of colonisation and undertook pastoralism (e.g. animal husbandry and gardening) (Duffy 2005b). While the Augustinians and Cistercians comprised many abbeys and priories, there were other orders including one house of Cluniac monks founded by Tairrdelbach Ua Conchobair before his death in 1156, several houses of Premonstratensian canons, and one house of Carthusian monks (Bartlett and Jeffery 1997, Duffy 2005b). There were fewer Benedictine houses in Ireland than there were in England at the time; however, two notable Benedictine houses are the 12th century de Lacy founded a house at Fore, County Westmeath, and the Priory of St Peter in Athlone (Duffy 2005b).

Friaries were introduced to Ireland during the early part of the 13th century. Like the abbeys and priories, friaries were laid out in a claustral plan, the slight difference being that the church was to the south of the cloister and the refectory to the north. These clusters of buildings were then often surrounded by granges and precinct boundary (Duffy 2005b).

The changes in architectural styles with the introduction of monastic orders in Ireland is reflected in the change in burial style. Each order was responsible for burying their dead, commonly within their precinct boundaries. Some orders, such as the Dominicans, would often settle outside the town boundaries near the town cemetery. These settlement patterns help identify the potential for graves in a geophysical survey as the easily detected features related to the monastic settlement will indicate that a burial ground is likely to be in the vicinity of the site.

2.2.3. *Burial Practices*

Early medieval graves were often oblong or subrectangular simple earthen dug graves with the individual in an extended supine position orientated west-east (head to the west). The trend seen

across Ireland is shallow inhumations approximately 0.1-0.2m below the original ground surface (Bhreathnach 2010, Corlett and Potterton 2010, Cahill and Sikora 2011, O'Sullivan et al. 2014). In some instances, stones were added to the grave as a stone slab, stone lining, lintels, pillow stones, or foot stones (Corlett and Potterton 2010, Cahill and Sikora 2011, O'Sullivan et al. 2014) Ogham stones were used as grave markers on many occasions (Bhreathnach 2010). Ancestral burial grounds such as those in use during the Iron Age remained in continuous use into the medieval period (see Cloghermore Cave, County Kerry, Knoxspark, County Sligo, and Millockston, County Louth) (Corlett and Potterton 2010, O'Sullivan et al. 2014).

There is, however, a distinct transition from the *geinti* (non-Christian) burial practices to the Christian burial practices that continued into the later medieval period (Bhreathnach 2010). As cemeteries became focussed around the church and Christianity took hold as the prevailing religion, the preference for west-east aligned burials became common across Ireland and the rest of Europe (Duffy 2005b, Bhreathnach 2010, O'Sullivan et al. 2014). It is suggested that the transition from Paganism to Christianity was a slow, complicated process in which there was regional variation in the timescale of the adoption of Christianity (O'Sullivan et al. 2014). Non-Christian Viking burials still occurred in Ireland during the early medieval period up until the 10th century. These burials were distinguished by the grave goods that accompanied them. In the later medieval period, following the Anglo-Norman Invasion, graves were often more ornate. Organised cemeteries start to arise, and there is an increase in the use of stone as both grave markers and burial containers. Stone-lined graves are widely represented; less often, stone cists or stone-lined graves with cover slabs were used.

Based on archaeological evidence, multi-individual inhumations were rare, but when they did occur, they most often contained an adult and a juvenile (Corlett and Potterton 2010, Cahill and Sikora 2011). Individual juveniles were buried outside the consecrated ground in *cillíní*, which were often located in pre-existing monuments such as ringfort.

Following the Anglo-Norman invasion, there was a trend to mark graves with effigies or coffin-shaped floor slabs for tombs. An individual's burial location within the cemetery, and the method of deposition often varied by status or sex (Bhreathnach 2010). Those of a higher status received ornate burials, often in stone cists or tombs near or within the church, while poorer individuals were buried in wooden or stone coffins in the main cemetery. These primary burial types can be characterised by their size, shape, and depth, making it easier to interpret them in geophysical data.

Morphological characteristics of medieval Irish graves have been derived from the National Museum of Ireland's compilation of burial excavations and osteological reports (Cahill and Sikora 2011). The author compiled the excavation records for relevant grave types and used them to determine the average grave size characteristics. These characteristics are detailed in Table 1 as well as illustrated in Figure 24. Some grave types were not discussed in Cahill and Sikora (2011), yet still may be encountered in geophysical survey data. As such, stone-lined graves, truncated unlined graves, shroud burials, and kinship/multiple inhumations are illustrated in Figure 24. The morphological characteristics of these grave types are based on the author's excavation experience.

Table 1: Morphological characteristics of medieval Irish graves (adapted from Cahill and Sikora 2011)

Type	Sample Size	Average Length	Average Width	Average Depth
Lintel	23	1.895m	0.42m	0.295m
Cist	5	1.69m	0.40m	0.27m
Unlined (Simple)	42	1.8m	0.67m	0.27m
Slab-lined	3	1.58m	0.88m	0.40m
Pit	6	1.94m	0.87m	0.33m
Later Medieval Simple	1	-	-	0.3m
Later Medieval Pit	1	0.7m	0.7m	0.3m
Later Medieval Coffin	2	1.775m	0.525m	0.45m
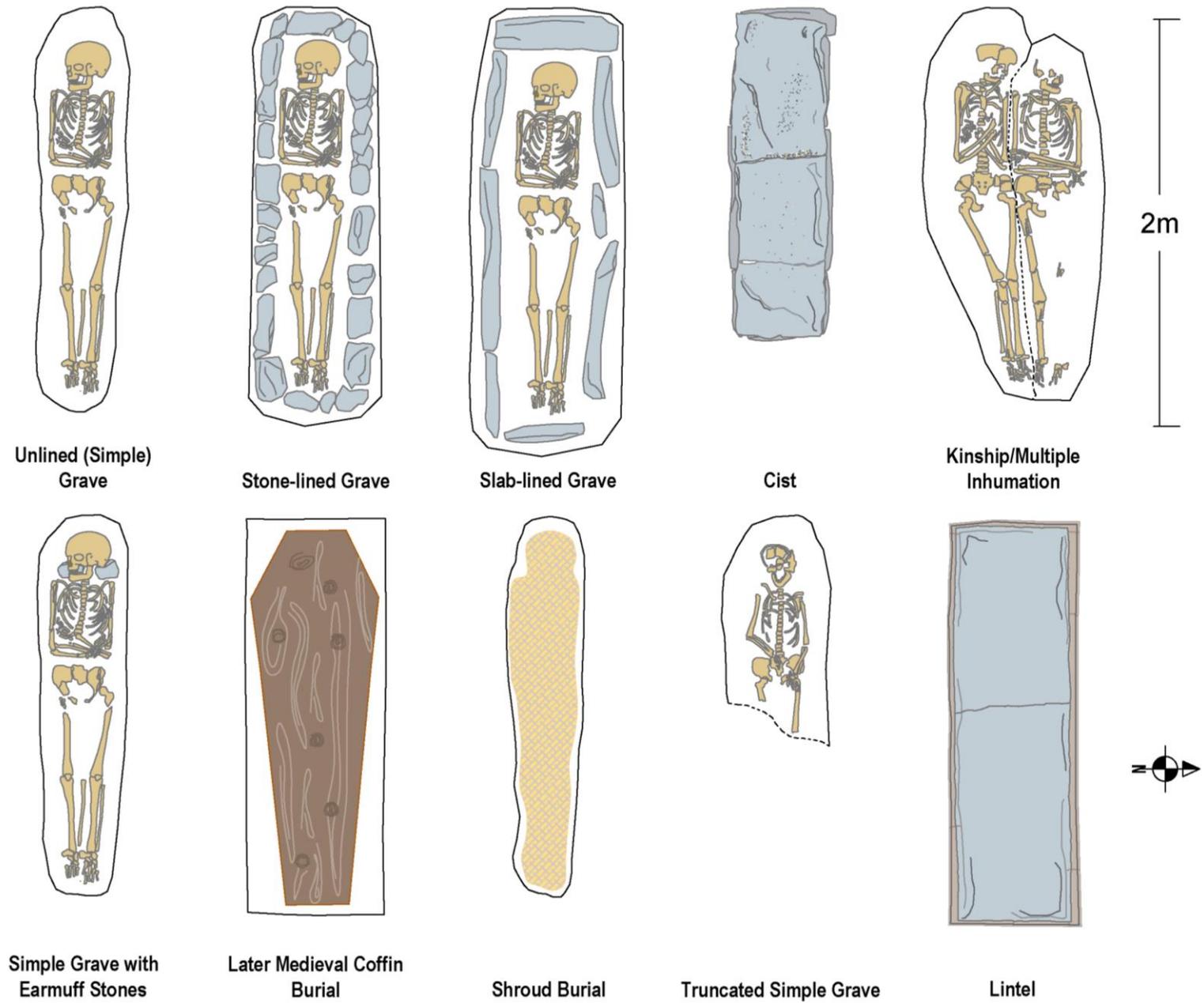
Figure 24: Illustrations of common Irish medieval grave types

Further to this, there is an extensive record of burials on the National Monuments Service database and Historic Environment Record. The prehistoric – post-medieval burial records for the National Monuments Service are shown with geographic distributions in Figure 25. A number of these burial records have corresponding excavation reports; as such, these burial records were compared with the associated excavation report to achieve a full characterisation of the burial. The full characterisation was used to inform subsequent simulations of GPR data during the creation of the machine learning training dataset. A review of these monument and excavation records by the author identified that upstanding remains in the vicinity of the study area are reliable indicators of high potential for graves within the study area. Figure 25 demonstrates the prevalence of burials throughout Ireland. In particular, the higher prevalence of simple graves ('Burial') and burial grounds. Pit burials were excluded from the machine learning training dataset due to their low prevalence.
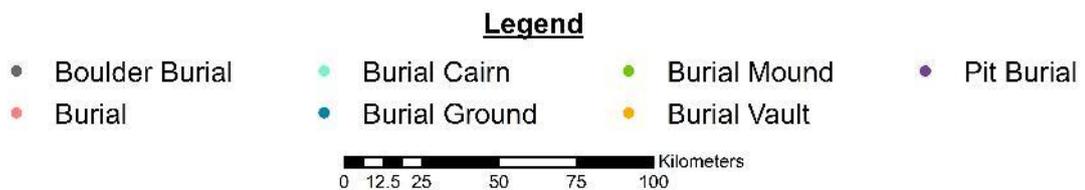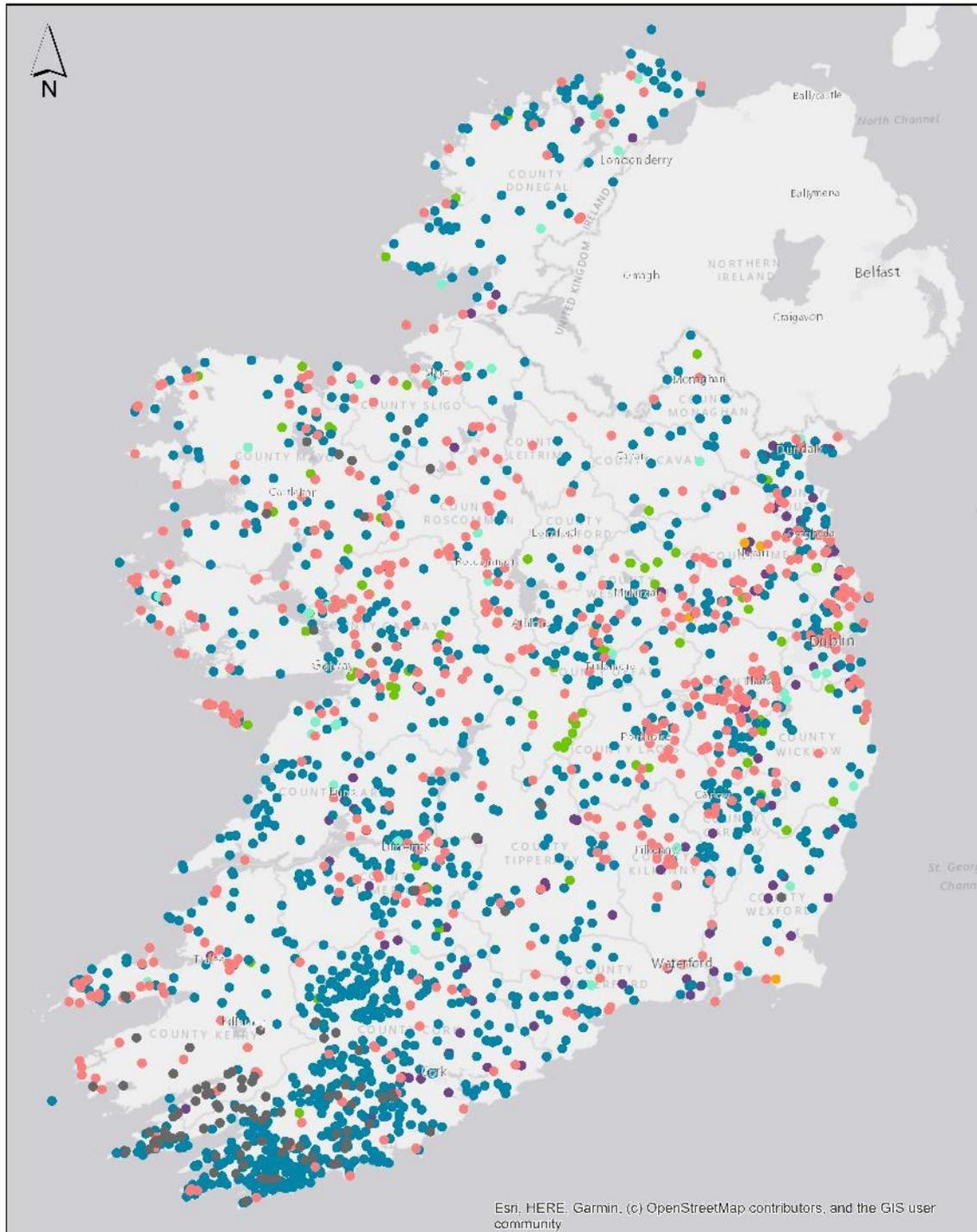
Figure 25: Distribution of the burial records on the National Monuments Service Database classified by typology (these include prehistoric – post-medieval burial types)

Burials may sometimes be marked with headstones, ogam stones, or grave slabs. In this case, the extent of the grave still requires delineation. However, development projects often avoid areas of marked burials; instead, they are more likely to encounter unmarked burial grounds or deserted medieval villages. The examples of churches included hereafter are representative of the ideal survey area and indicative of the structures or foundations that are likely to be detected through large-scale geophysical survey and suggest the further need for targeted high-resolution survey necessary for identifying graves. Examples of modern upstanding remains of churches and monastic buildings are provided in Figure 26 - Figure 29 to demonstrate the range of architectural styles and burial types. Figure 28 and Figure 29 also demonstrate the later use of medieval burial grounds into the post-medieval and modern periods.

Figure 26: Remains of the bell tower of Augustinian St Mary's Abbey in Trim, Co Meath (locally referred to as the Yellow Steeple). (Image copyright retained by author)

Figure 27: Remains of the Augustinian monastery Clonmacnoise near Shannonbridge, Co Offaly. Grave markers and grave slabs are still visible on the ground surface. (Image copyright retained by author)

Figure 28: Medieval and Post-medieval remains of St Brendan's Church on the site of St Brendan's monastery on William's Street, Birr, Co Offaly. Grave markers and lintels are still *in situ*. (Image copyright retained by author)

Figure 29: Romanesque-style church and nearby Holy Well dedicated to St Molua in Roscomroe, The Leap, Co Offaly. Possible medieval grave markers and modern headstones are still *in situ*. (Image copyright retained by author)

3. GEOPHYSICAL SURVEY GUIDANCE FOR TERRESTRIAL SURVEYS

3.1. Introduction

GPR is one of the least common techniques used in commercial and research archaeogeophysical surveys, and the datasets produced amongst the most difficult to interpret. As such, there is a pressing need to create software to assist geophysicists in interpreting this type of data. Figure 30 shows the effectiveness of conventionally used geophysical survey techniques in detecting a range of common archaeological features. The chart documents GPR's superior potential to detect graves successfully. Bonsall et al. (2014a) identified that, on average, magnetometry, earth resistance, and EMI have a poor-fair detection rate for inhumations; while, GPR has a fair-good detection rate for inhumations. For stone coffins or cists, GPR has a reasonable detection rate, followed by EMI and earth resistance which have a fair-good detection rate. There were no data available on GPR's ability to detect cremations; however, it is likely that even though magnetometry's detection rate is only fair, it would be the most suitable detection method due to the burning of the remains.

Figure 30:   Detectability of modern objects and archaeological features with GPR, EMI, earth resistance, and magnetometry. N/A indicates there is no data available (adapted from Bonsall et al. 2014)

3.2.   Established Survey Guidance

Guidelines for conducting archaeological geophysical surveys have been established at international and national levels. Guidance documents provide information regarding the most suitable survey techniques, instrument settings, and survey resolution for common archaeological features, geologies, and soils. The guidance provided is based on numerous studies on the detectability of archaeological features and effects of geologies and moisture content on survey data since the inception of archaeogeophysics. While the three primary sources for survey guidance (David et al. 2008a, Bonsall et al. 2014, Schmidt et al. 2015) recommend similar values, national guidelines offer methods and techniques guidance for mitigating for uncontrollable factors such as geology, nature of the potential archaeology, and modern activity on a site. This guidance is particularly important as some impeding factors may be region specific and used by agencies at several scales.

3.2.1.   *European Archaeological Council (Europae Archaeologiae Consilium)*

The European Archaeological Council established a set of guidelines for geophysical prospection in archaeology in 2015 (Schmidt et al.). These guidelines outline the most suitable techniques for the most common targets, controllable factors, and uncontrollable factors encountered during surveys in Europe, with a slight focus on the UK as they are adapted from Historic England's guidance (David et al. 2008a).   The recommended survey parameters for common techniques used in archaeogeophysical survey are in Table 2. As these recommendations by Schmidt et al. (2015) are reasonably text-based and inaccessible for some non-professional geophysicists, the text has been summarised in Table 2. These guidelines are quickly becoming the most commonly used. They are a suitable standard for all large-scale European surveys. However, where high loss geologies are present or ephemeral targets are the focus of the survey, local or regional guidance should also be consulted.

Table 2: A summary of the recommended survey parameters for European surveys (adapted from David et al. 2008a; Schmidt et al. 2015)

TECHNIQUE	EVALUATION (Reading x Traverse)	CHARACTERISATION (Reading x Traverse)
Earth Resistance	1m x 1m	0.5m x 1m **Graves: 0.5m x 0.5m
Electromagnetic (EMI)	Slingram: 1/3 of the coil separation Time-domain: 1/3 of the coil size	Slingram: 1/3 of the coil separation Time-domain: 1/3 of the coil size
EMI for Geomorphology	5m x 5m	-
GPR	0.05m x 0.5m	0.05m x 0.25m **Graves: 0.05m x 0.25m
Magnetometry	0.25m x 1m	0.25m x 0.5m
Topsoil Magnetic Susceptibility	10m x 10m	-

3.2.2. *Historic England*

Historic England (formerly English Heritage) established guidelines for the minimum resolution and guidance for choosing suitable techniques for archaeogeophysical surveys for England (David et al. 2008a). Table 2 also summarises the recommended survey strategies for a range of geophysical survey techniques as detailed in the decision table provided by David et al. (2008a). The suggested survey parameters are suitable for detecting a wide range of anomalies and even suggest suitable techniques based on soil types and the intended target. The parameters for evaluation surveys are often used in commercial geophysical surveys where timescales are short and have proven successful in detecting large- and small-scale features. However, as of June 2018, these guidelines have been superseded by the European Archaeological Council guidelines (Schmidt et al. 2015).

3.2.3. *Ireland*

In 2014, Transport Infrastructure Ireland (TII, formerly the National Roads Authority) produced guidelines for geophysical surveys of road corridors in Ireland (Bonsall et al. 2014). Bonsall et al. (2014a) make a case for the adoption of regional guidelines, rather than following the UK guidelines (David et al. 2008a). Due to the difference in bedrock geology, for example, the lack of chalk geology and presence of large areas of Carboniferous limestone in Ireland, Bonsall et al. (2014a, p. 34) stress that "the strategic response of surveys on these challenging Irish soils should not necessarily follow the English (or UK) model." It was also necessary to create guidance separate from those provided in the UK due to the nature of land cover Ireland. As in the UK, the most widely applied technique for large-scale survey is magnetometry. However, supplemental EMI or earth resistance surveys are also recommended.

The island (including Ireland and Northern Ireland) is comprised primarily of pasture, bog, and cultivated land. Uneven terrain in the west of Ireland can also impede survey quality and completion time as walking speed is hindered across the terrain. Also unlike the UK, the bedrock geology is largely Carboniferous limestones with some regions of concentrated igneous geologies as demonstrated by Figure 31, overlain by tills and, in some regions, peat. Soils, local geology, and land cover can all impact the effectiveness of magnetic techniques and, to an extent, electromagnetic techniques and therefore should be examined on a case by case basis.

Fluxgate gradiometers respond poorly to alluvium, colluvium, silts, sands, and gravels. The drift geologies overshadow archaeological features, and it becomes difficult to discern the archaeological features from the surrounding geology. In these instances, using a narrower traverse interval has the potential to improve the clarity of any archaeological features against the surrounding geology. Bonsall et al. (2014, p.39) identified that igneous geologies (e.g. basalt, diorite, granite, dolerite) create strong magnetically enhanced thermoremanent anomalies. Where the geology is near the ground surface without sufficient overburden, it creates widespread strong dipolar responses that obliterate any underlying features. In other instances, anomalies caused by the igneous geology may be misinterpreted as anomalies caused by anthropogenic interference. For example, Bonsall et al. (2014, p.39) provide an example of glacial erratics derived from igneous materials which could easily be misinterpreted as positive or dipolar anomalies associated with archaeological features such as hearths or in-filled pits.

Conversely, gradiometers respond very well in till (boulder clays) where other techniques such as GPR do not. The largest impact on electromagnetic techniques, including both GPR and EMI, are "noisy" volcanic geologies. GPR, especially, responds poorly on basalt and other igneous geologies where there is not suitably thick topsoil or overburden. Where the bedrock is near the ground surface, archaeological features cannot be discerned from the noise produced by the bedrock. However, because these geologies comprise only a small portion of Ireland (as demonstrated in Figure 31) and are constrained to small regions, electromagnetic techniques are widely applicable across Ireland. The ability to employ these techniques, which are most suited for detecting graves, can improve the detection and interpretation of graves before excavations if implemented correctly.

Bonsall et al. (2014a) clearly outline the effects of bedrock and superficial geologies on survey data and the suitable survey methodologies for reducing the impact of geologies and soils on the interpretability and quality of the survey data. As geology is one of the most limiting factors in geophysical survey, Bonsall et al. (2014a) provide guidance for the appropriate survey techniques based on the thickness of soils overlying the natural geology. This detail on the effects of geologies

on geophysical data is especially important due to the negative impacts some geologies can have on the detectability and interpretability of archaeological features as noted by both Bonsall (2014) and Fry (2014). Table 3 details the effects of igneous, metamorphic, and sedimentary bedrock geologies on geophysical techniques and how to overcome or counterbalance these effects. Additionally, Table 4 details the effects of the common superficial bedrocks in Ireland on geophysical data.   The advice presented in Table 3 and Table 4 further demonstrates that electromagnetic techniques are least likely to be negatively impacted by the local geologies. When the target features for a survey are small and likely to create a low magnitude response, as graves are, any background noise must be minimised either through choice of survey technique or processing in order to improve the potential for isolating small archaeological features. Thus, indicating that GPR and other electromagnetic techniques should be more widely applied across Ireland.

Table 3: Advice for surveying on typical bedrock geologies in Ireland (adapted from Bonsall et al. 2014a)

Geology		Formation Process	Advice for Areas with Significant Overburden or Surface Geology	Advice for areas with Near-surface Geology	Advice for Areas with a Variable Thickness of Overburden	Techniques Geology Does Not Negatively Impact
Igneous	• Basalt • Diorite • Dolerite • Felsite • Gabbro • Granite • Volcanic	Derived from molten rock and volcanic activity, creating magnetically strong thermoremanent anomalies	Magnetometer survey is often suitable as the strong magnetic response of igneous bedrock can be reduced by an overburden or surface geology	Magnetometer survey is not suitable as areas of insufficiently thick overburden will respond poorly	The area should be scanned using a magnetometer to identify areas unaffected by geology before commencing survey	Earth resistance EMI quadrature GPR
Metamorphic	• Graphite • Marble • Phyllite • Quartzite • Schist • Slate	Formed by a significant alteration of existing rocks through high temperature and pressure; there is potential for strong thermoremanent anomalies resulting from the geology	Use 0.5m traverse spacing to increase the resolution of archaeological features and to help discriminate against geological trends	Use 0.5m traverse spacing to increase the resolution of archaeological features and to help discriminate against geological trends	Use 0.5m traverse spacing to increase the resolution of archaeological features and to help discriminate against geological trends	Earth resistance EMI quadrature GPR
Sedimentary	• Carboniferous Limestone • Greywacke • Grit • Mudstone • Sandstone • Shale • Siltstone • Slate	Derived from fragments of pre-existing rocks and formed through the accumulation and lithification of sediment or by the precipitation from solution at normal surface temperatures (Crook et al. 2018)	In Counties Dublin, Meath, Kildare, Laois, Carlow, and Louth: Conduct detailed magnetometer surveys at a spatial resolution of 1m x 0.25m In the West and Southwest: Pilot studies should be carried out to determine suitable technique(s) and spatial resolution. A traverse spacing of 0.5m is recommended for magnetometer surveys. Cut-earth and thermoremanent features may be weakened or suppressed by waterlogging and organic matter and may not be suitable for magnetometer survey	In Counties Dublin, Meath, Kildare, Laois, Carlow, and Louth: Conduct detailed magnetometer surveys at a spatial resolution of 1m x 0.25m In the West and Southwest: Pilot studies should be carried out to determine suitable technique(s) and spatial resolution. A traverse spacing of 0.5m is recommended for magnetometer surveys. Cut-earth and thermoremanent features may be weakened or suppressed by waterlogging and organic matter and may not be suitable for magnetometer survey	In Counties Dublin, Meath, Kildare, Laois, Carlow, and Louth: Conduct detailed magnetometer surveys at a spatial resolution of 1m x 0.25m In the West and Southwest: Pilot studies should be carried out to determine suitable technique(s) and spatial resolution. A traverse spacing of 0.5m is recommended for magnetometer surveys. Cut-earth and thermoremanent features may be weakened or suppressed by waterlogging and organic matter and may not be suitable for magnetometer survey	Earth resistance EMI quadrature GPR

Table 4: Advice for surveying on common superficial geologies in Ireland (adapted from Bonsall et al. 2014a)

Geology	Advice for areas where the thickness of the geological layer is known	Advice for areas where the thickness of the geological layer is not known	Techniques Geology Does Not Negatively Impact
Alluvium	Gradiometers can respond very poorly to alluviated soils; instead, a total field magnetometer should be used. If carrying out a magnetometer survey, a 0.5m traverse spacing is recommended. Magnetometers should only be included in a multi-method survey approach. Increase the probe spacing in earth resistance surveys to assess layers below the surface geology.	Auger surveys should be carried out to determine the thickness of the alluvium. A multi-method survey approach which accounts for the penetration of the geological layer should be taken.	Total field caesium magnetometer GPR Earth resistance
Colluvium	If the survey area is small, the plough soil should be mechanically removed, if possible, to reduce the masking effects of the geology.	Auger surveys should be carried out to determine the thickness of the colluvium, and if possible, the plough soil should be mechanically removed.	Total field caesium magnetometer GPR Earth resistance
Fluvio-glacial Sand and Gravel	Magnetometer survey should not be relied upon as the primary survey technique; instead, it should only form part of a multi-technique survey methodology. However, if included as part of the methodology, magnetometer surveys should employ a 0.5m traverse separation to increase the probability of successfully identifying archaeological features.	Magnetometer survey should not be relied upon as the primary survey technique; instead, it should only form part of a multi-technique survey methodology. However, if included as part of the methodology, magnetometer surveys should employ a 0.5m traverse separation to increase the probability of successfully identifying archaeological features.	EMI (in-phase values) GPR
Peat	Peat deposit is 0 – 0.5m in thickness: Survey methodology should be determined by the target archaeological features and bedrock geology (see Table 3) Peat deposit is 0.5-1.0m in thickness: Survey methodology should be determined by the target archaeological features and bedrock geology (see Table 3). Prioritise earth resistance, detailed EMI, and detailed GPR surveys. Peat deposit is >1m in thickness: Low-resolution EMI survey combined with systematic augering or induced polarisation (IP) for wooden trackways	Auger surveys should be carried out to determine the thickness of the peat. Survey methodology should be determined by the target archaeological feature(s) and bedrock geology (see Table 3). Prioritise earth resistance, detailed EMI, and detailed GPR surveys.	Earth resistance EMI (quadrature values) GPR
Rock (Outcropping)	Use aerial photography to map areas of outcropping before commencing geophysical survey. Exclude areas of near-surface Carboniferous limestone and igneous geologies from the survey	Use aerial photography to map areas of outcropping before commencing geophysical survey. Exclude areas of near-surface Carboniferous limestone and igneous geologies from the survey	Dependent on the type of geology, see Table 3

Sand and Gravel	Magnetometer survey should not be relied upon as the primary survey technique; instead, it should only form part of a multi-technique survey methodology. However, if included as part of the methodology, magnetometer surveys should employ a 0.5m traverse separation to increase the probability of successfully identifying archaeological features.	Magnetometer survey should not be relied upon as the primary survey technique; instead, it should only form part of a multi-technique survey methodology. However, if included as part of the methodology, magnetometer surveys should employ a 0.5m traverse separation to increase the probability of successfully identifying archaeological features.	EMI (in-phase values) GPR
Silts and Sands	Magnetometer survey should not be relied upon as the primary survey technique; instead, it should only form part of a multi-technique survey methodology. However, if included as part of the methodology, magnetometer surveys should employ a 0.5m traverse separation to increase the probability of successfully identifying archaeological features.	Magnetometer survey should not be relied upon as the primary survey technique; instead, it should only form part of a multi-technique survey methodology. However, if included as part of the methodology, magnetometer surveys should employ a 0.5m traverse separation to increase the probability of successfully identifying archaeological features.	Earth resistance GPR
Till (Boulder Clays)	Detailed magnetometer surveys should always be a preferable option. The use of 0.5m traverse separations will increase the chances of successfully identifying archaeological features for detailed magnetometer surveys. Unrecorded magnetometer scanning surveys should not occur on tills	Detailed magnetometer surveys should always be a preferable option. The use of 0.5m traverse separations will increase the chances of successfully identifying archaeological features for detailed magnetometer surveys. Unrecorded magnetometer scanning surveys should not occur on tills	Magnetometer Earth resistance EMI (quadrature and in-phase values)

This guidance can also be widely applied where surveys are likely to encounter rough terrain or human interments. To reduce the impact of uncontrollable factors (e.g. land cover, geology, soils), the spatial resolution of magnetometry surveys should be increased, or electromagnetic techniques should be used where the local geology is not volcanic. While Bonsall (2014) provides detailed guidance for surveys in Ireland, a push for the use of electromagnetic techniques over magnetometry is necessary. Now that mechanised systems for GPR and EMI systems have become more accessible and are often a more manageable size compared to mechanised magnetometry systems, higher resolution electromagnetic surveys should be a requirement. It is understandable, in the west of Ireland especially, that some sites may not be accessible with mechanised arrays due to the land cover or layout of the field boundaries. In these instances, it should be recommended that high-resolution manual GPS-tracked surveys are carried out. In current practice, some commercial companies are heavily reliant on mechanised or cart-based surveys, but their use is still well below that in the UK.

Bonsall (2014) provides both tabular and flow chart guidance for determining the appropriate parameters for archaeogeophysical surveys in Ireland. This guidance, reproduced in flow chart format in Figure 33, accounts for potential controllable and uncontrollable factors which may be encountered in the planning and fieldwork stages of a project. The uncontrollable and controllable factors with the most significant impact on survey data quality and interpretability, as determined by Bonsall (2014), are provided in Table 5. Figure 33 is included here as an example of the most accessible form of survey guidance currently available. The flowchart format, while at some points can be confusing for the user, is more accessible than block text guidance and does not dilute the information necessary to make an informed decision on the appropriate survey parameters. However, presenting the data in an interactive map format where the user is not required to input the necessary geological and archaeological factor, as it is provided for them, would further increase the accessibility of the survey guidance.

Table 5: List of primary controllable and uncontrollable factors affects survey success (adapted from Bonsall et al. 2014a)

Controllable Factors	Uncontrollable Factors
Technique selection	Soil type
Data acquisition methods	Monument type
Spatial resolution	Land use
Competence of surveyors	Weather conditions
Seasonality	Bedrock geology
Data visualisation	Surface geology
	Sources of interference
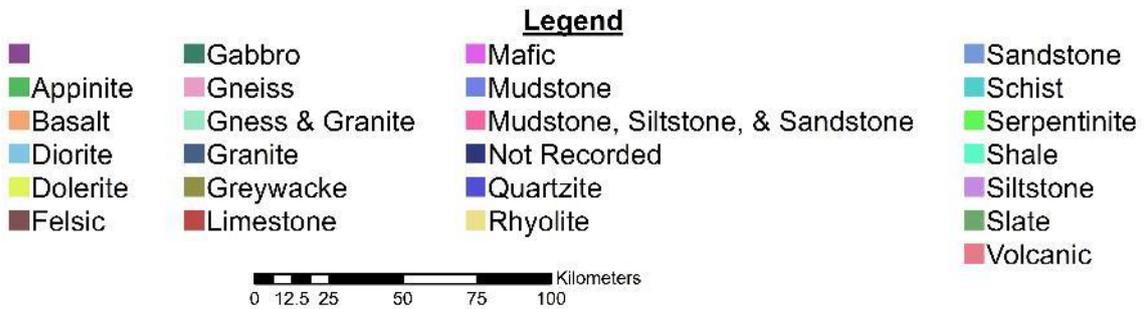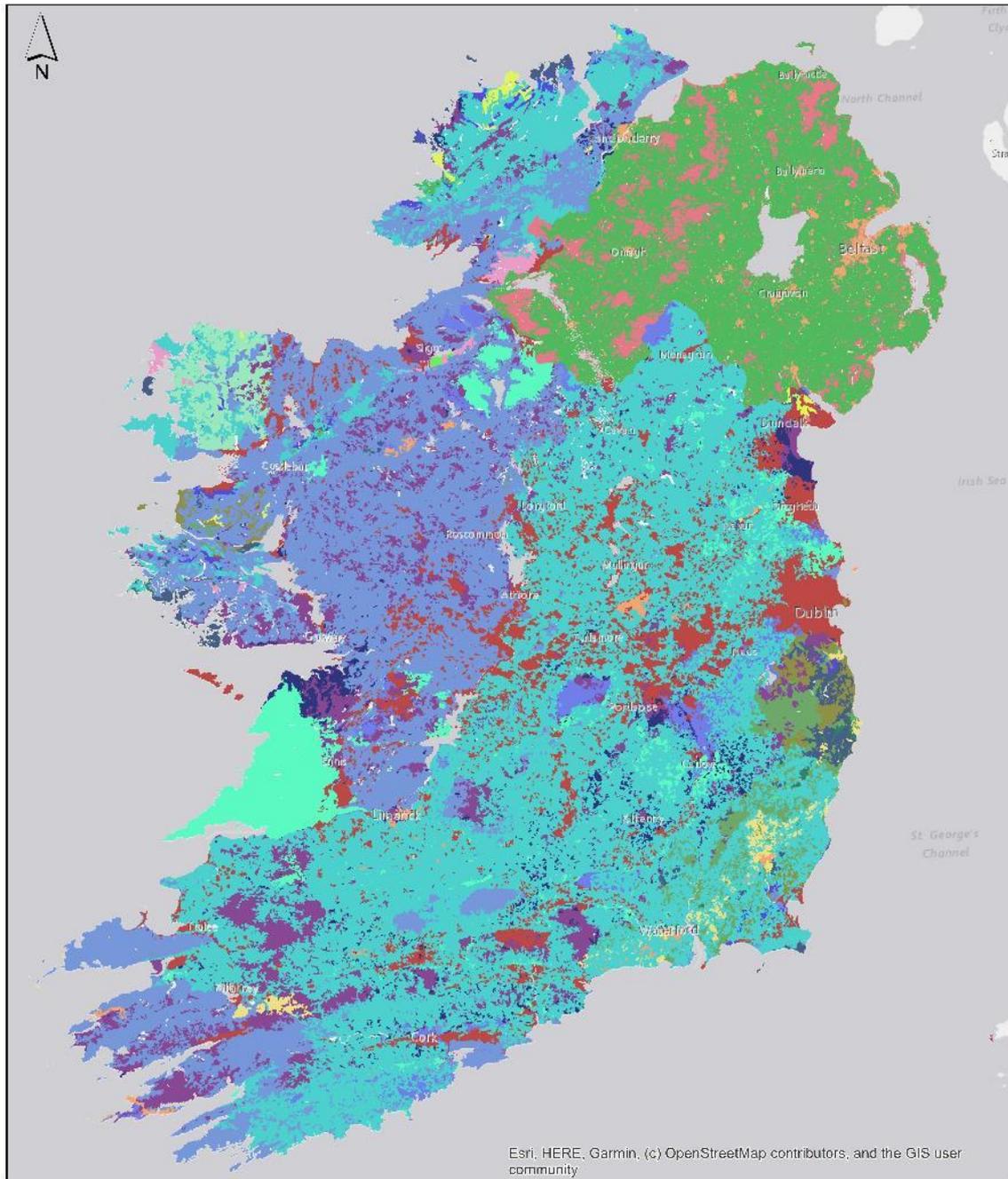
Figure 31: Large-scale map of the bedrock geology of Ireland, demonstrating a high proportion of sedimentary geologies in comparison to igneous geologies (adapted from Geological Survey Ireland 2018)

Figure 32: Large-scale soil texture map of Ireland depicting the broad coverage of loamy soils and peat (adapted from Creamer et al. 2016)

Choice of geophysical survey methods for use on Irish road schemes.

Other variables, such as seasonality, landscape, land use, sources of interference *etc.*, should also be considered, some of which cannot be controlled.
After David et al. (2008), with adaptations.

Flowchart Key

Question

Recommended Technique(s)

Alternative Recommended Technique(s)

Clearly state in the Briefing Document if the use of ATVs is permitted on the contract

What sort of location is the Survey Area in?

Urban (built up)

Rural or semi-urban

Try GPR. Electromagnetic, earth resistance and/or magnetometer survey may also be appropriate, if conditions permit

Or

Try a detailed magnetometer area survey at a 0.5m x 0.25m spatial resolution. Other techniques can and should be used to investigate features of interest and/or as a control

What is the depth of the overlying peat deposits?

Peat is greater than 1.0m deep

Peat is 0.5-1.0m deep

There is no Peat

Peat is less than 0.5m deep

Is the purpose of the survey over peat...

for general prospection of **unknown** deposits?

to trace the extent of **known** wooden trackways?

Are you are assessing a wooden trackway?

Yes

No

Try induced polarization (IP) profiles across the feature to trace its extent

Try EM transects across the peat at 5m transect intervals, to identify areas of raised ground. Consider combining the results with targeted hand-augering or on a regular 10m x 10m grid if mechanical augers are available. Geological investigations may also provide useful information. Augering could determine the depth below ground and palaeo-height-differences to identify archaeologically prospective areas. Further information could be gained by budgeting for laboratory magnetic susceptibility analysis of the cored soils, in advance of intrusive testing

deep (e.g. a **very large** ditch, greater than 1.0m depth)?

Are the Archaeological Features....

very magnetic (e.g. a kiln, furnace or other industrial deposits)?

shallow (less than 1.0m depth) or at unknown depth?

Weakly magnetic?

Try a detailed (0.5m x 1m) earth resistance area survey (Twin-probe, Wenner or Square arrays)

Or

a detailed (0.25m x 0.05m) GPR area survey

Or

a detailed (0.5m x 1m) EM area survey (apparent conductivity & apparent magnetic susceptibility)

Try a detailed magnetometer area survey at a 0.5m x 0.25m spatial resolution. Other techniques can and should be used to investigate features of interest and/or as a control

Try a detailed magnetometer area survey (1m x 0.25m)

Or, if on igneous geology

Use a pilot geological scan to determine areas suitable for detailed magnetometer area survey

Or

Try a detailed (0.5m x 0.25m) EM area survey (apparent conductivity & apparent magnetic susceptibility)

Or

Try a detailed (0.5m x 1m) earth resistance area survey (Twin-probe, Wenner or Square arrays)

Try a detailed (0.25m x 0.05m) GPR area survey

The Geology is

Metamorphic or Sedimentary

Igneous

Surface geology with magnetic pebble components

Is it the possible site of a stone building?

Yes

No

Is it the possible site of features greater than 1.0m depth, including souterrains?

Yes

No

Try electrical resistivity tomography (ERT) profiles across the feature or a detailed (0.5m x 0.05m) GPR area survey.

Try a detailed (0.5m x 1m) earth resistance area survey (Twin-probe, Wenner or Square arrays)

Or

a detailed (0.25m x 0.05m) GPR area survey

a detailed (0.5m x 1m) EM area survey (apparent conductivity & apparent magnetic susceptibility)

Try a detailed magnetometer area survey at a 0.5m x 0.25m spatial resolution. Other techniques can and should be used to investigate features of interest and/or as a control

The expected features include

masonry / stonework

cavities

large earth-cut features (e.g. channels / moats)

industrial features (e.g. hearths, furnaces, kilns, brick built structure)

diffuse / small scale (e.g. post-holes / palisades)

unknown feature types

other burnt feature (burnt mound of stones, building)

earth-cut features (ringfort / enclosures, ditches, pits)

Try electrical resistivity tomography (ERT) profiles across the feature or a detailed (0.5m x 0.05m) GPR area survey

a detailed (0.25m x 0.05m) GPR area survey

Or

a detailed (0.5m x 1m) EM area survey (apparent conductivity & apparent magnetic susceptibility)

Or

Try a detailed (0.5m x 1m) earth resistance area survey (Twin-probe, Wenner or Square arrays)

Try a detailed (1m x 0.25m) magnetometer area survey (strongly consider a 0.5m x 0.25m spatial resolution in areas of sedimentary geology to identify earth-cut features). Other techniques can and should be used to investigate features of interest and/or as a control - this is particularly important if stonework/masonry is also expected, which are unlikely to appear in magnetometer data. Pilot studies using a magnetometer, earth resistance or electromagnetic survey may help determine the most suitable method

Figure 33: Bonsall et al.'s (2014, p. 119) flowchart for determining appropriate survey parameters for archaeogeophysical surveys in Ireland which is an improvement on user accessibility from the block text and decision tables provided in other guidelines.

51

Guidance on Grave Detection

Bonsall et al. (2014) provide specific guidance for surveying small scale funerary monuments such as individual adult burials, cremations, and *Cillíní* (children's burial grounds). However, it is noted that even using the high-resolution survey methods suggested, it will still be difficult to detect small scale funerary monuments. Table 6 presents the suggested methods in which small-scale funerary monuments are likely, which required summarisation by the author. These methodologies are based on real-world geophysical surveys on road schemes as well as the grey literature (Bonsall et al. 2014).

Table 6: A summary of the suggested methods for the survey of potential small scale funerary monuments (adapted from Bonsall et al. 2014)

MONUMENT TYPE	TECHNIQUE	SPATIAL RESOLUTION
Inhumation	Combined earth resistance and magnetometry	0.25m x 0.25m
	GPR	0.05m x 0.5m Select antenna frequency based on suspected depth of the grave
	EMI	High sampling density
Cremation	Magnetometry	0.25m x 0.5m
	EMI (in-phase and conductivity)	0.25m x 0.5m
Cillín	Combined earth resistance and magnetometry	0.125m x 0.125m
	GPR	0.05m x 0.2m Select antenna frequency based on suspected depth of the grave

Recommended Improvements to the Guidance Format

Presenting a combination of Bonsall's guidance as well as English (David et al. 2008a) and European (Schmidt et al. 2015) guidelines in an interactive format will likely encourage the use of such guidelines, particularly by novices and new professionals. Overall, the primary guidance documents provide suitable methodologies for maximising the potential detection rate of expected targets. In accounting for controllable as well as uncontrollable impacting factors, Bonsall et al. (2014) can perceive and predict the majority of potential scenarios a surveyor could experience in Irish road corridor archaeology. However, the format of these guidelines is often inaccessible. Increasing the implementation of such guidance in survey planning will improve the data collected in the field, and therefore, the surveyor's ability to interpret the data appropriately.

Following from the discussion of grave morphology and composition in Chapter 2, the detectability and geophysical characteristics of each type included as a classifier in the automation tool are discussed in Chapter 6. To further understand the suitability of geophysical techniques for individual sites and the detection of graves, a discussion of the principles of widely applied geophysical techniques, data visualisation, and data processing is required.

3.3. Notes on the Format of Survey Guidance

Two main approaches to survey guidance have been used in archaeogeophysics – prescriptive and discursive (Schmidt 2019). Prescriptive guidance can be viewed as 'black and white' scenario; for example, if the local geology is volcanic only resistivity techniques should be used. Whereas, discursive guidance allows for more analysis of all contributing factors to the survey and the expert's knowledge. Both approaches have their merits depending on the intended audience.

From a North American perspective, Somers et al. (2003) developed *Automated Tool for Archaeo-Geophysical Survey* (ATAGS). ATAGS is a form of prescriptive guidance which uses an input-output method to determine appropriate parameters for a survey. The software outputs a 'survey design', which the authors define as "a set of decisions about the appropriate instrument, instrument configuration, data density, and data processing" (Somers et al. 2003, p. 3), as well as guidance for project management. This survey design is determined by the information provided by the user (e.g. rationale for the survey, soil characteristics, and the nature of the archaeological record). While the functionality and output of ATAGS are useful for both expert and non-expert surveyors, the tool is limited geographically and by technique. Somers et al. (2003, p. 3) note that ATAGS is only designed for use on sites in the Midwest and Plains regions, and some sites in the Mid-south and interior South regions, of the United States. The recommendations are also provided based on Geoscan Research instrument settings (Somers et al. 2003, p. 3).

Conversely, Schmidt (2019, p. 26) and Schmidt et al. (2015) advocate for a discursive approach to guidance over a prescriptive approach with the expectation that experts are best equipped to determine the appropriate survey strategies for sites with unknown or challenging characteristics. This discursive approach is suitable for surveyors with a wealth of knowledge and experience on a range of sites. However, it can be argued that less experienced surveyors, local/community groups, and companies commissioning surveyors would benefit from a more prescriptive approach. A prescriptive approach provides these groups with more detailed guidance they often require to achieve the same results as an expert. Combining discursive and prescriptive guidance into one tool allows all groups sufficient guidance for any survey.

The survey parameters tool created as part of the *Reilig* toolset, discussed in detail in Chapter 5, offers a middle ground between prescriptive and discursive approaches. It limits the amount of prior knowledge required by the user by providing the necessary background information (e.g. local geology, land cover, soils). The user is only required to input the site coordinates. The output for each coordinate displays:

- the most appropriate technique and survey density that should be used in surveys for site evaluation or characterisation, and potential graves;
- any additional techniques suitable for the site;
- reference to the relevant detailed guidance document where the user can find additional information on instrument configurations and detectability of archaeological features.

The most crucial aspect of any guidance, however, is that it provides a method for the user to determine the most suitable technique(s) for a site as the interpretability of a site is firstly determined by the technique(s) used and secondarily by the survey density.

4. PRINCIPLES OF GEOPHYSICAL SURVEY TECHNIQUES AFFECTING GRAVE DETECTION

4.1. Introduction

The measurements taken by geophysical instruments and the data visualisation methods are significant factors in the successful detection and interpretation of graves. The degree of processing on datasets can affect a human interpreter's ability to identify small features (as discussed in Chapter 6 and Chapter 1.2). In contrast, learning machines are reliant on the training data they are provided with. While GPR is often the most suitable technique for the detection of graves, it is important to note under which circumstances each technique is suitable to detect graves. In some instances, regardless of the processing on a dataset, graves are just not detectable by the instrument.

4.2. Ground-Penetrating Radar

Ground-penetrating radar (GPR) is an active electromagnetic technique which identifies subtle changes in the electromagnetic properties of subsurface materials including dielectric permittivity and conductivity. GPR systems are comprised of a transmitting antenna and a receiving antenna or, in multichannel systems, multiple sets of transmitting and receiving antennas. The transmitting antenna emits electromagnetic energy at a frequency determined by the length of the antenna element. Upon encountering variations in the physical and electromagnetic properties of the subsurface material, the signal is reflected to the receiving antenna(s), as shown in Figure 34. The amplitudes of the wavelets, corresponding to the dielectric properties of the reflective materials, are translated to numerical and colour-coded for almost immediate viewing on the computer system during survey. When material velocities are calculated, the two-way travel time is converted from time to depth.
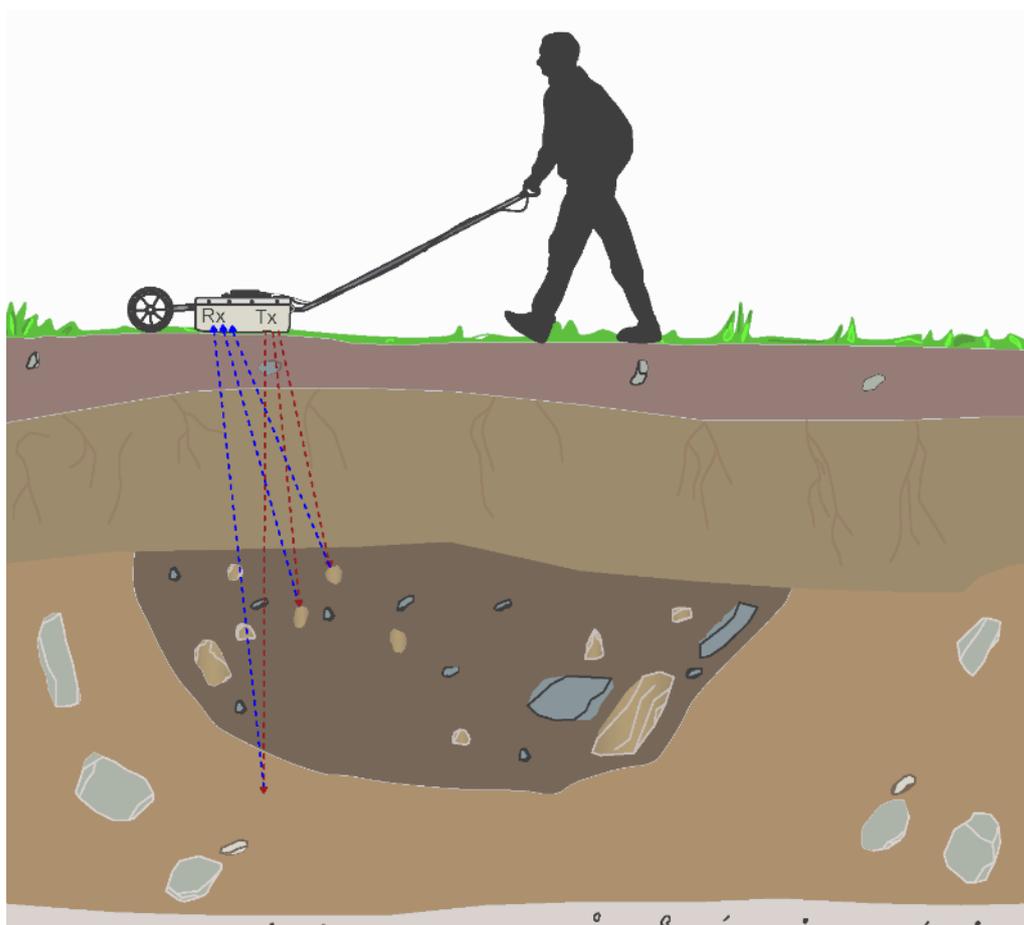
Figure 34: Electromagnetic principle of GPR

The first GPR system, albeit a crude one, was used in Austria in 1929 to measure ice thickness on glaciers (Conyers 2013). Following this, the applications of GPR expanded to geology and utility surveys in the 1970s, including work by the National Aeronautics and Space Administration (NASA) who built a GPR system to be sent to the moon (Conyers 2013). From the mid-1970s through to the 1990s the use of GPR systems, and geophysical survey in general, in archaeological science grew immensely (Bevan and Kenyon 1975, Vickers et al. 1976, Imai et al. 1987, Stove and Addyman 1989, Conyers and Goodman 1997, Conyers 2013). In the 1990s, significant improvements to survey and data processing came about with technological advances in computer hardware and processing software. After 1993 the field also saw the first use of amplitude slice maps, computer-simulated two-dimensional (2D) models (Goodman et al. 1994, 1995), and 3D reconstructions (Conyers 2013).

GPR is shown to be effective in locating buried objects depending on their size and orientation (Conyers 2012, 2016, Dupras 2012, Richardson and Cheetham 2013, Ruffell and McAllister 2015). While it is often used in engineering, utilities, infrastructure investigations, France et al. (1992, 1997) and Ruffell and McKinley (2005) have established GPR as an effective method for locating forensic inhumations. Experimental surveys have shown GPR to work in peat (Armstrong and Cheetham

2008), concrete (Toms et al. 2008, Harrison and Donnelly 2009, Ruffell et al. 2014), freshwater, chalk, and rock (Dupras 2012, Ruffell and McAllister 2015). Detectable subsurface features include geomorphological changes, large voids, structural remains, modern services, and burials (dependent upon the skeletal condition and soil matrix). Recent surveys have investigated the viability of GPR in detecting geomorphological features (Calder and Kennedy 2013, Lanzarone et al. 2016, Zaremba et al. 2016a), buried landforms (Gosar and Čeru 2016), and environmental features (Moore and Ryder 2015). Recent advances in the detection of landmines and improvised explosive devices (IEDs) have approached designing models to improve the rate and accuracy of detecting explosives which employ neural network and curve-fitting approaches (Singh and Nene 2013a).

During the post-processing stage, depth profiles are merged, and the space between traverses interpolated to present a three-dimensional (3D) rendering of the subsurface material. The 3D rendering can be 'sliced' to specific depths which are estimated by calculating the relative velocity of the subsurface material and the time it takes for the signal to return to the antenna(s). Once the data are processed, they can be exported as two-dimensional (2D) raster data or 3D animations.

A detailed discussion of the machine learning methods and composition of the training dataset is provided in Chapter 7.2; however, it is important to discuss the effect of data processing and visualisation on the accuracy of machine learning models. GPR data must be processed in order to correct for signal travel time, interference, and to view reflections of varying strengths. This processing, or filtering, of the data can impact on the learning machine's ability to detect reflections. Different degrees of processed data introduced in the training dataset will allow the machine to learn features in as many variations of processed data as possible. As surveyors do not process their data in a systematic manner across the discipline and there are numerous software packages available for data processing, no recommendations have been made to web app users on how to process their data; instead, users are notified that the level to which they survey and process their data will affect the output of the object detection model. It is expected that experienced surveyors will process their data cleanly and effectively. In contrast, less experienced surveyors (the primary target of the object detection tools) may not process the data as thoroughly. Therefore, the training dataset included in this pilot study has representative examples (see Chapter 7.2) of raw, improved, minimally processed, and over-processed data, as defined in Table 7, created in multiple mainstream software packages. It is understood that by not providing the machine with data processed under the same conditions the intersection over union (IoU) and classification accuracy may be impacted, while the accessibility of the tools by less experienced surveyors and surveyors without access to GPR-Slice software is improved. Sections 4.2.1. – 4.2.4. provide a discussion of the data processing and

visualisation factors which may impact on the interpretability of data and are established by the surveyors.

Table 7: Description of the data processing types used in this pilot study

Data Processing Type	Description
Raw	No corrections, first stage of data taken from the instrument
Improved	Gain corrected, gridded to correct offsets
Minimally Processed	Background removal, migration, regained
Over Processed	Filters have removed responses from archaeological features

4.2.1. *Signal Processing*

B-scans, or radargrams, are comprised of the specific 'signature' of pulses transmitted by the GPR antenna and the response from the ground/subsurface material. The data (signals) must be processed to analyse only the responses from below the surface. Most GPR systems record and present raw data in the field, meaning there are no filters added. Even with systems that display gained data, it may be necessary to filter the data in order to view later arriving reflections, determine and correct for the relative soil velocity, and remove any unwanted responses. Processed signals (radargrams) are then merged to be rendered as 2D images, 3D images, or animations. Radargrams are comprised of the impulse responses from the ground and the signal from the system itself (see Equation 4.1) whereby processing aims to isolate the impulse responses from the ground (Goodman and Piro 2013). The frequency, amplitude, velocity, and phasing of reflection waves are crucial in processing and interpreting GPR data. It is imperative not to over-process or arbitrarily apply filters to datasets as this could produce distorted, uninterpretable data (Conyers 2012, 2013).

4.1: Composition of radargram pulses

$$r = g \times i$$

$r =$ radargram pulses

$g =$ reflection response of the ground

$i =$ impulse response function of the antenna

Table 8: Description and effects of signal processes on GPR data (adapted from Green 2018)

Process	Purpose	Effect
Resampling	Correct the number of scans per the user-set navigation marker to create accurate spatial density information	Sets the number of scans per user-set distance unit to a constant value set by the user
Gain	Visualize later arriving, weakly visible reflections	Intensity, or amplitude, of radar reflections increases
Migration	Correct for signal spread and subsurface velocity distortion. Estimate the relative soil velocity based on the principle that broad hyperbolic reflections indicate a fast velocity and narrow hyperbolic reflections indicate a slow velocity	Removes extraneous hyperbolae by determining the "point source" of the reflections by averaging the reflection energy of each hyperbola in the dataset.
Hilbert Transform	Visualize weak reflections by considering the absolute value of the data and connecting the peak responses	A Fourier Transform (FT) is run on the radar pulses, the negative frequencies are shifted 90°, and then an inverse FT is run, to create a signal in the positive domain
Background Removal	Filter out horizontal banding caused by signal noise	Algorithms determine the average waveform across a radargram and subtract this value from all in the dataset
Bandpass Filter	Remove noise resulting from post-processing gain and signal noise	1D FFTs convert the radar pulses/signals into their spectral components (amplitude and phase at each detected frequency)
Spectral Whitening	Remove noise resulting from post-processing gain and "ringing" or signal noise from the antenna	Normalizes the real and imaginary spectral amplitudes by converting the data utilizing FFTs and setting the magnitude of the spectral frequencies to 1
Deconvolution	Minimize the negative effects of previous processes	Spectral division deconvolution: FFTs convert data to the frequency domain, remove the antenna's impulse frequency by dividing the frequency response of the radargram by the frequency response of the impulse, and revert the data to the time domain. Cepstrum deconvolution: Normalizes the data using a logarithmic scale. Predictive deconvolution: The preceding signal predicts the succeeding value in the radar scan

Resampling

Resampling is applied to correct for the variation in the number of scans per marker; where the scans per marker are the constant number of scans set by the GPR in a given unit of time per the user-determined navigation markers (which can vary between surveys due to ground conditions). This process must be completed before any others are applied to the dataset. In resampling, the number of scans per user set distance unit is set to a constant by either reducing the number of scans per

marker or interpolating the data between scans to create a new, resampled scan. By resampling the scans, accurate spatial density information can be used in subsequent processes.

Time Gain Adjustments

Time Gain adjustments are necessary in a majority of cases. Most GPR systems record non-gained data in the field; although, due to the severity of signal attenuation and antenna frequencies, it is also recommended that gained data collected in the field are regained during the processing and, sometimes, the post-processing stages depending on other processes performed on the data. Radargrams, especially those acquired using low-frequency antennas, will require post-processing gain to visualise the later arriving, weakly visible reflections. Automatic gain curves, either exponential or linear, or an alternative suitable custom gain curve can be applied. The post-processing gain will often amplify any signal noise in the dataset.

Bandpass Filtering

Bandpass Filters are applied after the data are gain corrected to remove noise. Noise is caused by post-processing gain and "ringing" or signal noise from the antenna. Processing applications utilise one-dimensional (1D) Fast Fourier Transforms (FFTs) to remove noise from radargrams rather than the 2D FFT used to remove noise from timeslices (see Equation 4.4). Equation 4.2 defines the FFT in a uniform dimension of the continuous domain, where $\omega = 2\pi f$ and f represents the temporal frequency given in hertz (Rao et al. 2011). Equation 4.3 defines the inverse FFT in a uniform dimension of the continuous domain, where $\omega = 2\pi f$ and f represents the temporal frequency given in hertz (Rao et al. 2011). The 1D FFT converts the radar pulses/signals into their spectral components (amplitude and phase at each detected frequency). By altering the amplitude of differing frequency, any superfluous noise can be removed from the dataset or corrected. To return the corrected data to radargram form, an inverse FFT is performed.

4.2: FFT in a uniform dimension of the continuous domain

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$$

$$\omega = 2\pi f$$

f = temporal frequency in hertz

t = time

4.3: Inverse FFT in a uniform dimension of the continuous domain

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)e^{j\omega t}d\omega$$

$$\omega = 2\pi f$$

f = temporal frequency in hertz

t = time

4.4: 2D FFT used to remove noise from time slices

$$X(z1k, z2k) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x(n1, n2)z_{1k}^{-n1} z_{2k}^{-n2}, \;\; k = 0, 1, \dots, N1N2 - 1$$

Background Removal

Background Removal serves to remove linear, horizontal banding caused by "ringing" from the antenna (Neal 2004, Goodman and Piro 2013). This process uses algorithms to determine the average pulse frequency across the given radargram and subtracts this value from all values in the dataset (Leckebusch 2003). Alternatively, where subsurface features are orientated parallel to the horizontal banding, it is recommended the average pulse is determined using data from the entire site, rather than a single radargram (Goodman and Piro 2013). As background removal processes are likely to remove any linear features of potential interest to the survey and create phantom responses and deriving values from the entire site is more likely to mitigate for this. It is still recommended for background removal processes to be applied to a dataset as it allows the interpreter to identify slight reflections near the surface.

Migration

Migration serves to correct the signal spread converging with electromagnetic changes multiple times at various angles, especially in surveys experiencing signal diffraction resulting from severe topographical variations (Jol and Bristow 2003, Neal 2004, Orlando and Slob 2009). This phenomenon will skew the two-way travel time, causing sections of an object to appear distorted and possibly larger than it is in reality. The manual migration process removes extraneous hyperbolae by determining the "point source" of the reflections, averaging the reflection energy of each hyperbola and placing the result at the apex. However, it is important to note that this averaging must be done

for each hyperbola in the dataset to collapse the migrated reflections into their source hyperbola (Conyers 2013, Goodman and Piro 2013). Kirchoff migration is a more sophisticated method used for archaeological survey data in which the angle of incidence and the depth of the reflective feature(s) are calculated. These undergo a correction based on an assumed velocity profile resulting in corrected positional errors, and hyperbolae collapsed back to their source (Neal 2004, Conyers 2013). Less common migration algorithms include FK, Stolt, phase shift, and finite difference, though they typically produce similar results (Leckebusch 2003, Goodman and Piro 2013).

Migration can also be used to estimate the relative velocity and dielectric constant of the subsurface material across the entire dataset based on the relative velocity of the soil and the two-way travel time (see Equation 4.5). Manual migrations fit computer-generated hyperbola of known velocity to hyperbolas in the dataset to calculate an overall or area-specific velocity. Collecting velocity data in the field is the most accurate. However, computer-generating values during processing is a quick and easy method, which is accurate depending on the number of reference hyperbolae available in the dataset (Conyers 2013). Velocity calculation is based on the principle that broad hyperbolic reflections indicate a fast velocity and narrow hyperbolic reflections indicate a slow velocity. Migrating data or using incorrect parameters for migration will cause phantom reflections, though these will appear markedly different from true reflections (Goodman and Piro 2013).

4.5: *Two-way travel time*

$$T = \frac{2\sqrt{x^2 + z^2}}{v}$$

T = two-way travel time

x = horizontal distance to the subsurface object

z = depth to the subsurface object

v = microwave velocity in the ground

Spectral Whitening

Spectral whitening is similar to bandpass filtering in that it removes noise, but the real and imaginary spectral amplitudes are normalized by converting the data utilizing FFTs and setting the magnitude of the spectral frequencies to 1 (Goodman and Piro 2013). Spectral whitening is particularly useful when processing data collected from multichannel systems as it serves as a secondary method of signal balancing.

Hilbert Transform

The Hilbert Transform is expressed mathematically as a function of the frequency of the radar pulses, where a Fourier Transform (FT) is run on the radar pulses (see Equation 4.6), the negative frequencies are shifted 90°, and then an inverse FT is run (see Equation 4.7) (Johansson 1999). This serves to create a signal in the positive domain and represents the envelope of the recorded pulse (Goodman and Piro 2013). Simply, the Hilbert Transform (see Equation 4.8) serves to rectify the signal such that only the absolute value of the data is considered and the peak responses are connected, where it is necessary to visualize strong or weak reflections.

4.6: Fourier Transform

$$\widehat{G}(f) = -j \, \text{sgn}(f) G(f)$$

4.7: Inverse Fourier Transform

$$f(x) = \int_{\mathbb{R}} \int_{\mathbb{R}} e^{2\pi i (x-y)\cdot\xi} f(y) dy d\xi$$

ξ = frequency

4.8: Hilbert Transform

$$\hat{g}(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{g(\tau)}{t-\tau} d\tau$$

τ = time-lag

Deconvolution

Deconvolution serves to minimize multiple reflections or echoes, and the effects of the transmitted pulses from the antenna. Spectral division deconvolution utilizes an FFT to convert the data to the frequency domain and removes the impulse frequency of the antenna by dividing the frequency response of the radargram by the frequency response of the impulse (see Equation 4.9) and reverting the data to the time domain using an inverse FFT (Goodman and Piro 2013). The downfall of using this method is the uncertainty of the antenna impulse signal. Alternate deconvolution methods include cepstrum deconvolution and predictive deconvolution. The cepstrum deconvolution method is similar to spectral whitening in that it normalizes the data but uses a logarithmic scale rather than setting all values to unity to smooth the spectral amplitudes and balance the spectral frequencies (Goodman and Piro 2013). It does, however, require gain renormalization as the spectral amplitudes

will have shifted. Predictive deconvolution, as the term suggests, utilizes the preceding signal to predict the succeeding value in the radar scan (Goodman and Piro 2013).

4.9: Spectral division deconvolution

$$g = G(w) = \frac{R(w)}{I(w)}$$

$g =$ value in the frequency domain

$G(w) =$ ground response

$R(w) =$ frequency response of radargram

$I(w) =$ frequency response of transmitted impulse

4.2.2. *Rendering*

Processed data can be presented as 2D or 3D plots, animations, and overlain on other geophysical and remote sensing data to construct a full characterisation of a site. In 3D rendering the space between radargrams is interpolated; as such the first step to producing high-quality raster outputs is to collect high-resolution data in the field (Leckebusch 2003, Verdonck et al. 2015, Green and Holmes 2017).

Kriging Interpolation

The primary methods of interpolating between data are kriging and inverse distance. Kriging solves an inverse covariance matrix to abate the error between data and interpolated points based on weighting, often producing superior results (Goodman and Piro 2013). Weighting is determined by the inverse covariance matrix of all searched point in the estimate and the vector containing points in the search radius (see Equation 4.10).

4.10: Kriging

$$w = c_{ij}^{-1}d$$

$$c_{ij} = c_0 + c_i \text{ if } h = 0$$
$$c_{ij} = c_1^{(-3h/a)} \text{ if } h > 0$$

$w = $ weighting

$a = $ range at which covariance value remains constant

$c_0 = $ nugget effect which provides a discontinuity at the origin

$h = $ distance between points

$c_0 + c_i = $ sill which describes the value of the estimate at large distances

$c_{ij}^{-1} = $ inverse covariance matrix of all searched point in the estimate

$d = $ vector containing points in the search radius

Inverse Distance Interpolation

Inverse distance interpolation is faster and more mathematically simplistic (see Equation 4.11), where nearby points are averaged to estimate an interpolated point, based on the distance to the interpolated point. Weighting is proportional to the inverse square of the distance between a known point and the estimated point (Goodman and Piro 2013).

4.11: Inverse distance interpolation

$$d = \sum_{i=1}^{N} \frac{w_i d_i}{\sum_{j=1}^{N} w_j}$$

$d = $ distance

$$w_j = \frac{1}{h^s}$$

$h = $ distance between nearby point and the point on the grid to be estimated

$s = $ smoothing factor (usually set to 2)

As demonstrated hereafter, particularly in Figure 35 and Figure 36, the processing applied to data affects the data quality and visualization, thereby potentially changing the interpretation of features in the data. The processing of GPR data is especially important in this project. The machine initially

learned on raw B-scans (radargrams) as a control to minimise the effects of data processing. Minimally processed (migrated and regained) data were then introduced to the training dataset with negligible change in the model accuracy. As the machine is learning to detect hyperbolae of within a size and shape range, the accuracy of the model is unlikely to be affected by processes that will not affect the overall shape of hyperbolae in the data. Any changes in accuracy were likely a result of the random split of the dataset into training and test datasets rather than the introduction of processed data. However, any processing that affects the size and shape of hyperbolae will reduce the accuracy of the feature detection and classification. Data processing and presentation is more likely to affect the interpretation of C-scans (timeslices). Over interpolated or smoothed data will obscure small features, like graves, decreasing the accuracy and confidence of both human- and machine-led interpretation.

Additional Corrections

Once girds are interpolated, grids may require further destaggering if data was collected in zig-zag or Global Positioning System-tracked (GPS-tracked) mode, and data collected over multiple days may require mosaic correction to reduce the effects of variable moisture content, weather conditions, equipment settings, and geological contrasts across a site (Conyers 2013).

4.2.3. *Visualization*

B-scans

B-scans (also called radargrams or depth profiles) are 2D representations of an individual traverse; thus, all additional renderings are based on the radargrams. The radargrams provide valuable information about the amplitude and depth of responses. Profiles are particularly useful for viewing hyperbolic reflections during the interpretation stage of a survey. The profiles should still be viewed in conjunction with slice maps as X-Y trace plots are viewed alongside greyscale plots when interpreting other geophysical data to obtain a view of the entire site.

C-scans

Timeslices (also called C-scans or amplitude maps) are blocks of data at a specified depth below the ground surface created from small sections of horizontal data, called timeslices. These demonstrate the scale and spatial distribution of responses in a given area or across an entire site. Rendering data this manner allows for a plane view of any responses of interest and provides the end-user with

3D data. Maps can be presented as a single slice or a series of stacked slices that demonstrate the change in an anomaly with a change in depth (Conyers 2016). Slice maps are the most common method of visualizing GPR data as they provide a 3D representation (top, side, and front views) of a site at depths set by the user. The slice thickness and overlap are set by the user, which can be altered to target specific depths. Once timeslices have been created, they can be interpolated and transformed into an animation and isosurface matrix. Animations allow the end-user to visualize the changes in the subsurface matrix in real-time, which aids in identifying subtle changes and patterns that change with depth (e.g. sloping features, stratigraphic layers).

Examples of stationary presentations of GPR data are shown in Figure 35; animations of the transition between slices from the ground surface to the end of the dataset can also be created. It is important to consider the range of methods used to display GPR data as an individual response can present differently in 2D and 3D forms.
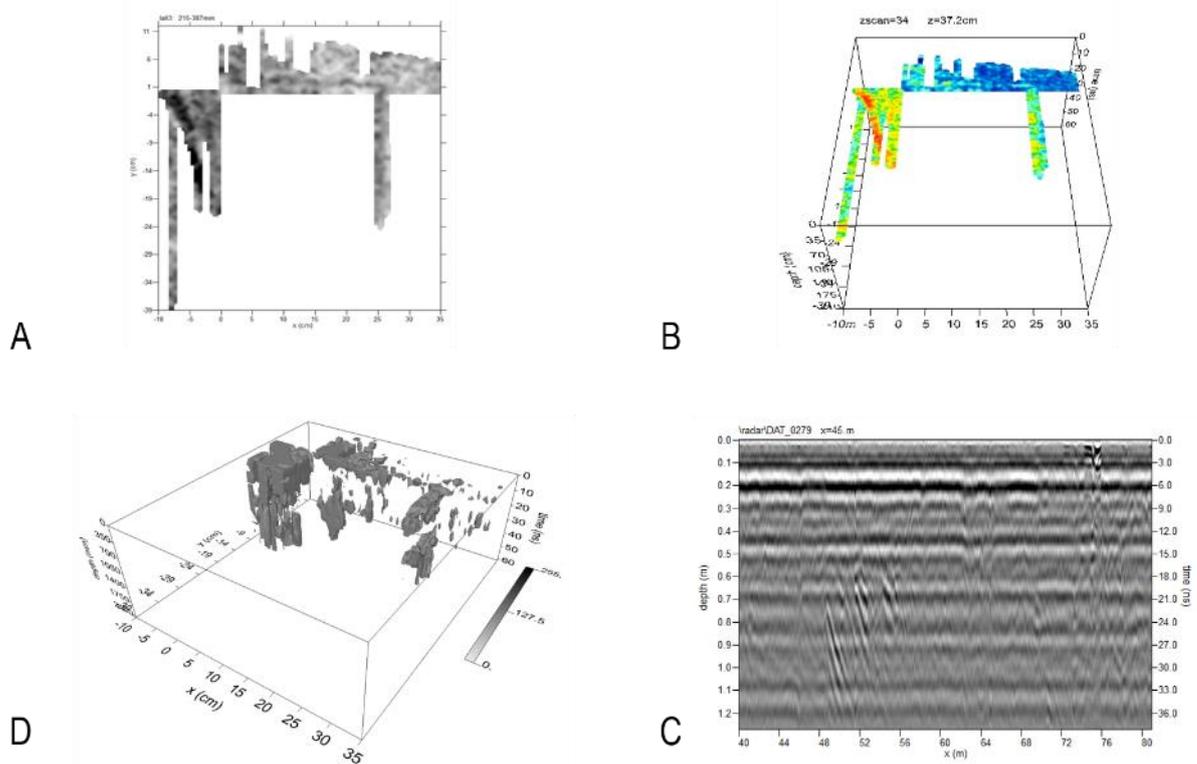
Figure 35: Examples of GPR data formats (A: Greyscale timeslice; B: Colour timeslice in a 3D matrix; C: 3D isosurface; D: Greyscale unfiltered radargram)

Modelling and Simulation

Modelling and simulation have developed since the 1990s when Dean Goodman (Goodman et al. 1994, Goodman et al. 1995, Goodman and Piro 2013) created software to simulate the GPR signatures of archaeological features based on their known or expected electromagnetic and

morphological properties (Conyers 2016). By establishing parameters such as relative dielectric permittivity, conductivity, and morphology, 2D reflection profiles are produced from artificially simulated wave paths (Conyers 2016, Warren et al. 2016). Simulated models are useful in demonstrating how a buried object may present differently within a range of geological background materials and weather conditions (Conyers 2012). Of importance to this research project are 3D models as they can be used to calculate the volume and represent the morphology of anomalies of interest. It is important to note that the accuracy of simulated models is entirely dependent on the quality of the known values used to create the model and present an idealised representation of a feature.

Isosurfaces

Isosurface rendering generates 3D vector data which represent points in space of a measured constant value. In GPR data, this is often the amplitude of reflected waves. Visualizing data in this orientation is particularly useful for isolating responses of interest. As isosurfaces can be rotated and viewed from any angle to show all aspects of an anomaly, they are a crucial tool in representing and understanding the full extent of buried objects. Isosurfaces can be imported into computer-aided design (CAD) programs to be georeferenced and combined with Geographic Information System (GIS) data and other forms of geophysical data.

"Data Fusion"

Data Fusion and multi-component integration allow the end-user to merge data from a range of geophysical techniques, remotes sensing, and spatial datasets into one GIS. As most geophysical techniques measure only one property, combining data to create a multi-component dataset can increase accuracy and confidence in interpretations. This is key when visualizing and interpreting GPR data which often contain an overprovision of anomalous responses by highlighting anomalies of archaeological potential hidden amongst the non-target responses. Kvamme (2007) has discussed the merits various integration approaches, such as mathematical and statistical operatives (e.g., Boolean union, Boolean intersection, thresholded binary sum, principal components analysis, k-means cluster analysis, binary logistic regression), computer graphics solutions (e.g., 2D overlays, RGB colour composites), and vectorised interpretative integration. The integrative approach taken is dependent on the types of data to be used. A simple 2D overlay or 3D composite is suitable for viewing most geophysical and spatial data. Statistical and mathematical approaches may be necessary when integrating data which measured non-comparable properties.

Data Presentation

GPR data is presented in red, green, blue (RGB) value colour scales or greyscale. The colour scale is particularly important in presenting GPR data, in part because of the range of responses and slight variations in measured amplitudes often encountered in datasets (Conyers 2012). Applying a colour scale based on the standard deviation from the mean is a more effective presentation style than assigning the colour scale based on the range of values in the dataset (Goodman and Piro 2013). Plotting in colour helps to visualize weaker responses as there are more colour assignments available. However, colour scales tend to produce contrasts in the dataset that might not exist. In greyscale, data is plotted with equal weighting on all reflections but often causes weaker reflections to become obscured in datasets with a broad range (Goodman and Piro 2013).

The colour scales relate to the amplitude of a reflected pulse. Sliced data can be plotted in 'relative normalization' or 'absolute' forms, as shown in Figure 36. In relative normalization, colours are assigned relative to the highest value in each time slice. Weak responses are easier to view in relative normalization, but responses are plotted differently across timeslices. Plotting in absolute is similar to display clipping magnetic or resistance data, whereby responses are colourised with respect to the maximum and minimum values of the dataset. This method produces a more consistent scale, but it becomes difficult to visualise weak reflections. The best practice is always to view the data in both aspects to interpret the origin of anomalous responses accurately.
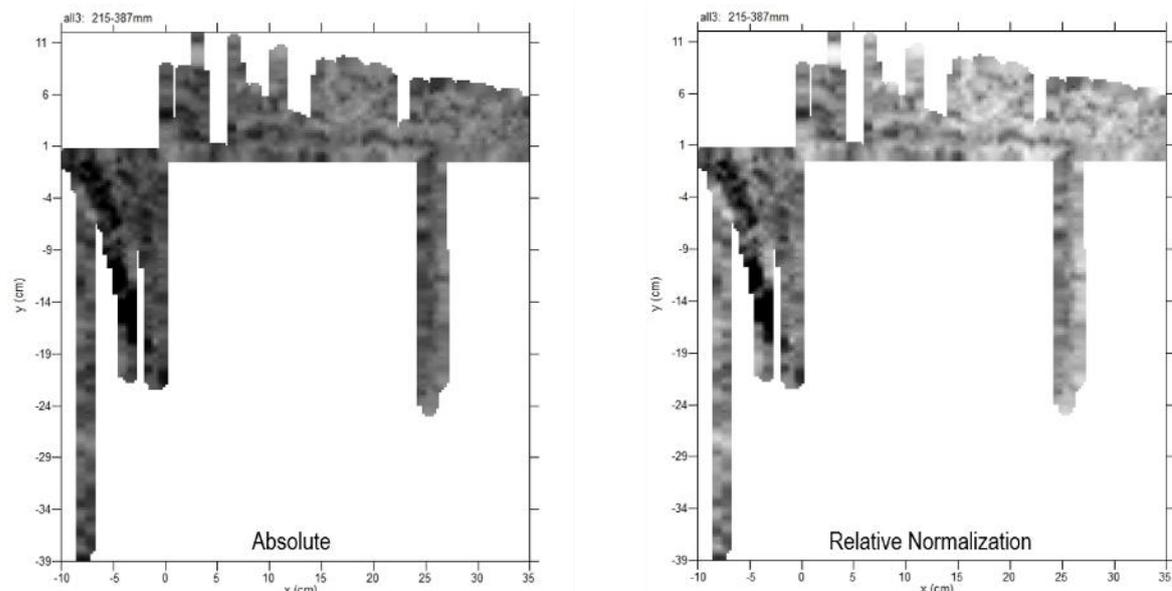
Figure 36: A comparison of absolute plotting parameters (left) and relative normalization plotting parameters (right) for GPR data. In this data, the lowest amplitude responses present noticeably different between the two timeslices, while the highest amplitude responses remain the same visually.

4.2.4. *Interpretation*

Interpreting GPR data is often considered more complex than other techniques (Conyers 2012, Goodman and Piro 2013), partially due to the nuances of processing the data but also the sheer number of objects GPR can detect. It is important in interpreting the data to not only look for anomalies but to view the entire "picture" as areas of no or low reflection can be as telling as stark high amplitude responses. The interpretation of GPR data is entirely reliant on characterising the amplitude, depth, and morphology of a response as there is no linear representation of entire datasets as there is in other geophysical techniques. While many advances in GPR data collection and processing have been made in the last 20 years, further enquiry is necessary to grasp its applications to archaeological investigations fully. Notably, the effects of traverse and sampling interval size on data quality and the ease and confidence of data interpretation require exhaustive analysis. Analysing these factors (Bonsall, Gaffney, et al. 2013, Bonsall et al. 2014, Verdonck et al. 2015) will aid in creating standard workflows necessary to improve data interpretability and overall survey quality. Standard guidelines (David et al. 2008a, Bonsall et al. 2014, Schmidt et al. 2015) require caveats for GPR survey as there are nuances which can reduce the quality of survey data.

Processing of the data is likely to affect humans' ability to identify ephemeral responses as background and general noise in the data can detract from the interpretability. In machine learning, however, the machine will learn from all data it is provided with. This allows the machine to learn on a range of clean to noisy data, encompassing the varying data it may later be asked to infer on. In this project, the training dataset includes a larger proportion of data that were only bandpass filtered, less data with bandpass filtering and background removal, and even fewer data with bandpass filtering, background removal, and migration. It is expected that there will be a trade-off between the accessibility of the tool for a range of data types/surveyor skill levels, and the accuracy of the trained model. To account for any potential adverse effects on the accuracy of the trained model, any noise and horizontal striations in the data were included as examples in the 'non-grave' class to inform the machine that these image features should be discarded and not learned as graves. Overall, this will allow for broader accessibility and applicability of the machine learning tool, both in the archaeogeophysics profession and among amateurs. Further discussion of the training data and processing steps are provided in Chapter 7.2.

4.2.5. *Detecting Graves*

GPR has proven one of the most successful techniques for detecting both modern and archaeological graves. Most recently, the survey of Fountains Abbey, Ripon, North Yorkshire, achieved clear,

successful results (Gaffney et al. 2014, p. 10). Individual graves were delineated as well as the *in situ* soil matrix between burials in the GPR survey (Gaffney et al. 2014, p. 10). Additional successful applications of GPR for detecting graves are discussed in Chapter 1 and Chapter 11.

4.3.   Supplemental and Secondary Techniques

In the scope of this research project, secondary techniques are those whose data are not the primary focus of the analysis software. This, however, is not to state that in commercial and research survey, these techniques are used as supplementary survey methods. Instead, these techniques, particularly magnetometry, are most often those implemented first in 'real-world' geophysical survey.

Supplementary survey techniques are those techniques which can be effectively implemented alongside GPR to improve the detection of archaeological features. These techniques are introduced in Sections 4.4 - 4.6. Employing these techniques in a multi-method survey approach often improve the confidence in data interpretation and are especially useful in identifying areas to target with high-resolution GPR survey.

4.4.   Electromagnetic Induction Methods

Electromagnetic induction (EMI) methods, also called low frequency electromagnetic (LFEM) methods, measure the apparent magnetic susceptibility (in-phase electromagnetic signals) and apparent electrical conductivity (quadrature electromagnetic signals) of the soil (Davenport 2001, Schmidt et al. 2015). Conductivity, being the real component, is the measure of the amplitude of the received waves in quadrature phase. Conversely, magnetic susceptibility, the imaginary component, is the measure of the amplitude of the received waves in the in-phase. EMI systems operate with separate transmitter and receiver coil(s). The transmitter coil acts to propagate an alternating magnetic field (the primary magnetic field) which interacts with the soil to create electrical currents which create a secondary magnetic field. The primary magnetic field is nulled electronically or by the orientation of the receiver coil(s), such that any changes in the primary field or the ground (eddy currents) are measured by the receiving coil(s).

4.4.1. *Arrays*

EMI systems come in Slingram and time-domain arrays, both of which are active techniques. In Slingram systems, the coil separation and orientation determine the maximum depth of investigation (Saey et al. 2013). Conversely, in time-domain systems, the depth of investigation, or measurement envelope, is dependent on coil size. Recently produced systems can simultaneously record data

from different coil separations and orientations (De Smedt et al. 2013). EMI systems are similar to other electromagnetic techniques but operate in the kHz range. Self-nulling multi-receiver systems offer a minimum of two coil pairings – one in horizontal coplanar dipole (HCP) mode and one in perpendicular dipole (PRP) mode (Saey et al. 2013). Rotating the instrument 90° will convert the instrument from HCP mode to vertical coplanar (VCP) mode. It is important to note how the how a change in the orientation of the coils will affect an EMI system's potential to detect archaeological features.
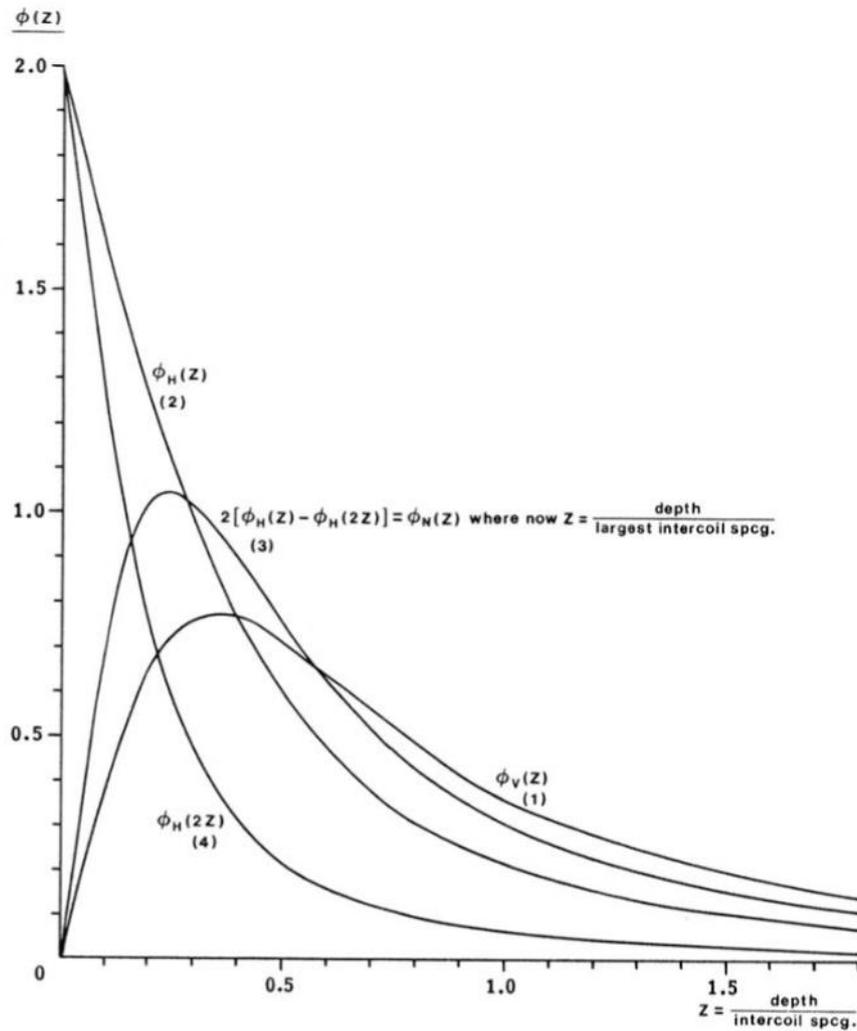
Figure 37: Relative response from a material at different depths for a multi-coil EMI system in vertical or horizontal dipole mode (McNeill 1985, p. 3)

Only quadrature values are discussed herein as in-phase measurements are not reliable or sensitive enough to be suitable for detecting graves. When operated in the vertical magnetic dipole mode quadrature sensitivity initially increases relative to depth, or signal penetration, thus overlooking materials near the ground surface. In the horizontal magnetic dipole mode, relative sensitivity is highest at the ground surface and decreases as depth increases (McNeill 1980, Geonics Ltd. 2003).

4.4.2. *Archaeological and Forensic Applications*

EMI surveys can delineate large earth features, e.g., remnants of mounds and backfilled ditches (Bevan 1983). In a forensic respect, EMI surveys can locate metallic objects relating to weapons and the disturbed ground of recent burials (Davenport 2001, Richardson and Cheetham 2013). EMI system performance is affected by coil orientation, coil separation, coil size, and soil properties. For archaeological targets, quadrature phase data is most useful as in-phase (magnetic susceptibility) data is often unreliable and unable to detect the slight variations in susceptibility necessary for interpreting archaeological features.

4.4.3. *Detecting Graves*

Bonsall et al. (2013a, p. 225) and Bigman (2012, p. 35) have both successfully used EMI to detect modern and earlier graves. Bonsall et al. (2013a, p. 224) conducted an EMI survey using the CMD Mini Explorer multireciever EM at the Asylum Cemetery at High Royds, West Yorkshire. The EM survey delineated modern graves (interred between 1890-1969) in the quadrature and in-phase data, with the grave more clearly defined in the quadrature data (Bonsall et al. 2013a, p. 224-225). In comparison with the earth resistance and magnetic surveys previously undertaken by Gaffney and Gaffney (2011), the EMI survey was able to define the extent of the graves clearly (Bonsall, Fry, et al. 2013, p. 225). However, it did not delineate the related structural features which are identifiable in the earth resistance data (Bonsall, Fry, et al. 2013, p. 225).

From a North American perspective, Bigman (2012, p. 31) carried out an EMI survey to locate graves within a Native American funeral mound in Georgia, USA. Burials at the site are known to date to AD 900 – AD 1100 (Mississippian occupation) and AD 1680 – AD 1720 (Creek occupation) (Bigman 2012, p. 33). The survey identified several low conductivity responses that are likely representative of graves within the funeral mound (Bigman 2012, p. 35).

4.5. Earth Resistance

Earth resistance techniques measure the electrical resistance (the inverse of conductance) of a subsurface matrix (soil). As such, a grave or other target object must create a significantly lower or higher resistance value compared to the surrounding survey environment (Hunter and Martin 2002, Cheetham 2005). Resistance and resistivity (the inverse of conductivity) are useful in detecting large stony structures, and graves backfilled with stones or containing remains wrapped in polythene (Hunter and Martin 2002, Gaffney and Gater 2003, Schmidt 2013).
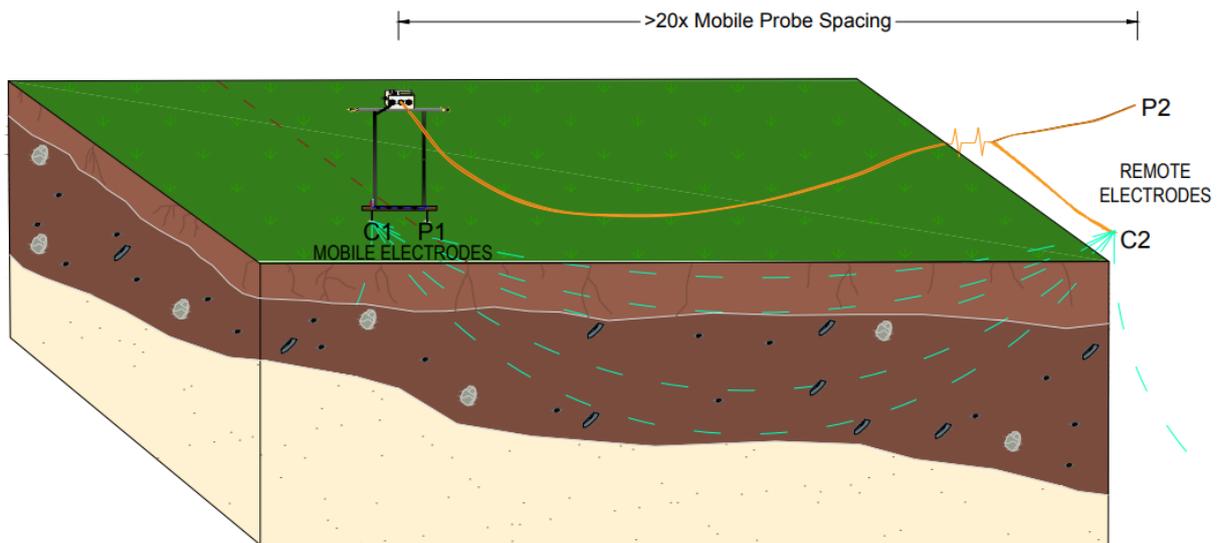
Figure 38: Principle of the twin probe earth resistance array

Resistance techniques measure the ability for soil or other material to allow an electric current to pass through it, as shown in Figure 38. In practice, resistance meters are detecting the presence or absence of interstitial water and salts present in the soil. The current flowing through the material is in proportion to the potential difference, or voltage, that is used (see Equation 4.12) (Gaffney and Gater 2003, Schmidt 2013). This flow of current is described by Ohm's law, which states that resistance (R) is equal to the change in voltage across a material (V) divided by the current flowing through said material (I, measured in amperes).

4.12: Measure of current flowing through a matrix

$$\Delta V = V_2 - V_1$$

V = voltage across a material

It is important to note that these resistance values are bulk measurements, and therefore rely on the type and volume of material the current is propagated through (Schmidt 2013). Resistivity, on the other hand, will measure changes in material itself irrespective of its volume.

4.5.1. *Detecting Graves*

Schmidt (2013, pp. 62–63, 88) discusses the application of earth resistance in the survey of a suspected cemetery in North Yorkshire. While the discussion is more specifically about the effect of electrode separation on the size and distribution of resistance anomalies, Schmidt (2013, pp. 62) present an example of a grave cut which was successfully delineated. In a twin-probe array, the

response from a grave is commonly a peak-and-trough, or triplet, response with two negative and one positive components (Schmidt 2013, p. 88).

4.6. Magnetic Techniques

Gradiometry systems detect changes in the Earth's magnetic field within a localized area by creating a gradient between two sensors situated 0.5m – 1m apart vertically and operating in the range of 2kHz (Hunter and Martin 2002, Aspinall et al. 2008, Schmidt et al. 2015).   Breiner (1981) demonstrates the proficiency of gradiometer surveys within small- and medium-scale forensic investigations.   While gradiometry is one of the most widely used techniques in commercial geophysics, the implementation of this method to detect burials is heavily dependent on the survey environment. It is adversely affected by ferrous objects in the surrounding area and has a maximum potential depth of investigation of only *c.* 1m.

Magnetometry systems operate on the principle that objects and materials can become magnetised through natural or anthropogenic processes. In archaeology materials primarily become magnetised through remnant magnetisation or induced magnetisation (Aspinall et al. 2008).   Remanent magnetisation, often thermoremanence, occurs when randomly oriented neighbouring magnetic domains become heated above their Curie temperature, causing them to become paramagnetic. After the materials cool, the domains reform around the newly aligned magnetic moments; this causes a significant difference between the heated, thermoremanent materials and the background material (Aspinall et al. 2008).   Stronger induced magnetisation occurs because of increased magnetic susceptibility as iron oxides in the soil undergo reduction and re-oxidation.   In an archaeological context, these processes often occur through heating and burning or the deposition of heated debris (Aspinall et al. 2008).

David et al. (2008a) and (Bonsall 2014) identified the effects of geological responses on magnetometry surveys and how these may affect the quality of the data. A review of the responses is presented in Table 9 to demonstrate the usability of magnetic techniques in the UK and Ireland.

Table 9: Response of geology in magnetometry survey (adapted from David et al. 2008a, Bonsall 2014)

Geology	Response in Magnetometry Survey
Igneous	Thermoremanent effects can preclude survey over some igneous rock types; however, others are relatively unaffected.
Metamorphic	Experience so far suggests that thermoremanence is not usually a significant problem and magnetometer survey can be effective but beware of adjacent intrusions.
Sedimentary:	Magnetometer survey can be recommended over any sedimentary geology. There are few significant distorting factors although a wide range of magnetic susceptibility in the parent rock results in a very variable background response to survey.
Conglomerates/Grits/Pebble beds	Response is average to poor, but good in places.
Sandstones	Average response is poor.
Limestones	Response is good.
Mudstones/Clays	Average response is poor.
Drift:	Quaternary deposits overlying the solid geology are a primary consideration. They often show a high degree of local variation, and the magnetic response usually depends on the magnetic mineralogy of the parent solid geology.
Sands/Gravels	Response is very variable – good on materials derived from Jurassic limestones, moderate – good in south-central England and the West Midlands.
Coversands	Response is uncertain - ?poor.
Boulder clay	Response is generally poor.
Clay-with-flints	Response is good.
Brickearth	Response is ?poor – average.
Alluvium/Colluvium	Response is poor-average, depending on the depth of features below this material.

4.6.1. *Detecting Graves*

While magnetic surveys have proven successful in detecting cremation burials due to their increased magnetic contrast post-burning, few surveys have been successful in delineating inhumations. Notably, however, (Linford 2004, p. 178) identified that, while potentially difficult to interpret, there was localised magnetic enhancement related to individual inhumations likely as a result of microbial colonisation within the grave in archaeological contexts.

Gaffney and Gaffney (2011) were also successful in delineating graves in magnetic survey (Bonsall, Fry, et al. 2013, p. 225). A fluxgate gradiometer survey detected individual graves within a modern cemetery (interment between 1890-1969) with more success than the twin-probe earth resistance survey, but with less clarity than the EMI survey (Bonsall, Fry, et al. 2013, p. 225).

4.7.   Data Processing

Survey can be traditional gridded surveys or GPS tracked surveys, which either continuously or intermittently record data. In traditional gridded survey, once grids have been assembled into a composite, they can be processed.

Depending on the survey methods, data will require improvement or processing. Improvement is limited to correcting defects resulting from the effects of the chosen surveys methods during data acquisition, such as destaggering, drift correction, edge matching, Zero Mean Traverse correction, and despiking (Schmidt et al. 2015). Improved data can then be processed, or filtered, to enhance anomalies of archaeological or possible archaeological origin without introducing additional noise to the dataset. It is often appropriate for magnetic and electromagnetic data to be interpolated and low pass filtered, and for resistance data to be interpolated and high pass filtered (Schmidt 2001). However, it is important to note that processing can change the size and shape of responses. Processed data can be exported and plotted as 2D and 3D georeferenced composites to aid in rendering and visualisation of interpretation diagrams at a preferred maximum scale of 1:1000 (Schmidt 2001, David et al. 2008a).

Table 10: Uses and effects of conventional data processes (adapted from Aspinall et al. 2008, Schmidt et al. 2015)

Process	Use	Effect
Interpolation	Smooth the appearance of greyscale plots where coarse data collection parameters were applied in the field.	Increases spatial density but not the number of *real* data points.
High-pass Filtering	Remove the effects of large-scale geological trends and substantial variations in the topsoil magnetic susceptibility allowing archaeological anomalies to become clearer.	A weighted local average is calculated around a data point (background value) and is deducted from the central reading to allow narrow, short-wavelength anomalies to become more prominent.
Low-pass Filtering	Remove spikes, smooth the image, and improve the visualisation of weak anomalies.	Changes value of a data point by calculating the average value within a block of data. Smooths the image, weak anomalies become more visible.
Despiking	Remove random spikes within the dataset caused by ferrous contamination within the topsoil and instrument instability.	Value of selected data points are replaced with a 'normalised' value, sometimes the average of surrounding data points (background value).
Destriping (Zero Mean Traverse)	Correct the banding in zig-zag surveys resulting from the directional sensitivity of magnetometers.	A constant value is subtracted from the affected traverses, so their mean is set to zero – this can obscure linear anomalies or remove those parallel to the affected traverses.
Destaggering	Correct offset readings caused by differences in the operator's speed and topographic variations.	Traverses are shifted.
Edge Matching	Correct variations in the background value between grids caused by the effects of changes in temperature, soil moisture content over long survey periods.	Mean or median of each grid is set to zero (or another common value).

4.8. Rendering, Visualization, and Interpretation

Data plots can be rendered in 2D or 3D greyscale and RGB colour plots. Greyscale 2D plots are the most common presentation for archaeogeophysical survey data. The 2D plots are especially useful for interpretation when accompanied by X-Y trace plots. 3D relief diagrams incorporate the 2D data and a 3D relief of the trace plots into a singular diagram. If data were collected using a multi-receiver EMI system or multiplexed resistance system additional 3D plots, similar to 3D GPR data plots, can be created to render isosurfaces and for use in multi-component integration rendering methods (Saey et al. 2013). In multi-receiver EMI data, particularly, data can be 'sliced' to specific depth boundaries (determined by the coil separations) to present 'non-apparent' electrical conductivity and magnetic susceptibility values, given Equations 4.13-4.16 which are solved using Levenberg-Marquardt non-linear least-squares algorithm (Marquardt 1963, Saey et al. 2013). This allows the interpreter to isolate features at approximate depths.

4.13: Apparent conductivity (ECa) in horizontal coplanar mode

$$\text{ECa}_{\text{HCP}} = \left[R_{\text{HCP}_S}(z_1) - R_{\text{HCP}_S}(z_s)\right]\text{EC}_1^* + \left[R_{\text{HCP}_S}(z_2) - R_{\text{HCP}_S}(z_1)\right]\text{EC}_2^* + \left[1 - R_{\text{HCP}_S}(z_2)\right]\text{EC}_3^*$$

4.14: Apparent conductivity (ECa) in perpendicular dipole mode

$$\text{ECa}_{\text{PRP}} = \left[R_{\text{PRP}_S}(z_1) - R_{\text{PRP}_S}(z_s)\right]\text{EC}_1^* + \left[R_{\text{PRP}_S}(z_2) - R_{\text{PRP}_S}(z_1)\right]\text{EC}_2^* + \left[1 - R_{\text{PRP}_S}(z_2)\right]\text{EC}_3^*$$

4.15: Apparent magnetic susceptibility (MSa) in horizontal coplanar mode

$$\text{MSa}_{\text{HCP}} = \left[R_{\text{HCP}_S}(z_1) - R_{\text{HCP}_S}(z_s)\right]\text{MS}_1^* + \left[R_{\text{HCP}_S}(z_2) - R_{\text{HCP}_S}(z_1)\right]\text{MS}_2^* + \left[1 - R_{\text{HCP}_S}(z_2)\right]\text{MS}_3^*$$

4.16: Apparent magnetic susceptibility (MSa) in perpendicular dipole mode

$$\text{MSa}_{\text{VCP}} = \left[R_{\text{VCP}_S}(z_1) - R_{\text{VCP}_S}(z_s)\right]\text{MS}_1^* + \left[R_{\text{VCP}_S}(z_2) - R_{\text{VCP}_S}(z_1)\right]\text{MS}_2^* + \left[1 - R_{\text{VCP}_S}(z_2)\right]\text{MS}_3^*$$

As with GPR data, multi-receiver EMI and multiplexed resistance datasets can be formatted as animations. Magnetic, EMI, and resistance data are commonly integrated in data fusion or multi-component data integration plots. Overlaying these data with spatial and topographic data or aerial imagery can improve the interpretability of data.

4.9. Applications of Archaeogeophysics for General Site Evaluation

Since its inception geophysical prospection has been applied to a range of archaeological sites, from prehistoric to modern, all of which have limitations on data interpretation. It is important to note that a geophysical survey is only as good as the survey parameters employed, the local geology, and, most importantly, the surviving archaeology. Prehistoric archaeology often proves difficult to geophysically detect, even though, as stated by Gaffney and Gater (2003), prehistoric anthropogenic activity is relatively simple compared to that in later history. As the majority of prehistoric archaeology is considered 'negative archaeology,' where features are cut into the subsoil, bedrock, or earlier deposits, geophysical prospection is useful for strata definition and identifying palaeochannels, field systems, settlements, settlement enclosures, and burial sites. It is agreed that the most suitable approach for locating these features is a magnetic survey (e.g., fluxgate or caesium gradiometry) due to the silting up of cut features and the increased potential for areas of repeated burning (Gaffney and Gater 2003). Although, additional techniques should not be discounted where features are likely to be more than 1m below the ground surface or require increased sampling in one or both directions.

Detecting early historic archaeology (Roman archaeology in British contexts, and pre-Christian archaeology in Irish contexts) is markedly easier and more likely to be detectable compared to

prehistoric and later historic archaeology (Gaffney and Gater 2003). The considerable magnetic enhancement of features allows for probable detections using magnetic systems, while the increased use of stone in buildings in this historic period is better detected with GPR and resistivity systems. The archaeology of this period often follows an established pattern which allows the surveyor to interpret archaeological features and their significance in geophysical datasets more quickly and confidently. It is posited by Gaffney and Gater (2003) that the majority of geophysical surveys, especially those in the UK, have been carried out on sites of this period, proving the great success of the majority of geophysical techniques in detecting features from this period as well as improving the confidence in interpreting such features in future surveys.

Geophysical prospection has had mixed success in detecting medieval and post-Roman to early modern archaeology. As is often the nature of ecclesiastical sites, they leave behind little to no enhanced magnetism, stone foundations are not guaranteed to survive, and they require an intensive sampling strategy during survey to maximise their detection potential (Gaffney and Gater 2003). Other archaeological features, such as ridge and furrow ploughing, land drains, and some military sites, are much easier to detect due to more magnetic enhancement. Surveying archaeology of these periods relies much more heavily on documentary evidence and a high resolution, multi-technique approach to surveying (Gaffney and Gater 2003).

4.10. Overall Detectability of Graves

While improvements in equipment sensitivity, data logging capabilities, and survey methods have led to a slight improvement in the potential to detect unmarked graves, the methodologies employed today still fall short where remains are fully skeletonised and the burial container has deteriorated (Killam 1990, Moffat 2015a). Inhumations are unlikely to be detected using large-scale reconnaissance methods often used in commercial archaeogeophysical surveys, and therefore require precise high-resolution, high-density data acquisition.

Where the orientation of the grave is known or assumed, it is essential traverses are aligned perpendicular to the grave's longest axis, with dense traverse and sampling intervals. The maximum traverse spacing as suggested by Moffat (2015), Bonsall et al. (2014), and Green (2015) is 0.5m, with a maximum sampling interval of 0.25m for magnetic and resistivity techniques and 0.05m for electromagnetic techniques.

The grave cut for an adult grave is typically *c.* 2m x 0.5m with some variation depending on if a coffin or other burial container was used. Pre-modern graves are often simple earthen dug graves or wood coffin burials. As graves are often dug and backfilled within hours, there is no measurable magnetic

enhancement in the grave fill. However, a grave cut may be identified by a contrast in the electrical conductivity of the grave fill and surrounding material caused by increased moisture retention. In more elaborate graves, there is potential for grave goods, coffin nails, and other remains to produce a magnetic response. However, these responses are unlikely to be interpreted to associated with a grave without evidence of grave markers or a grave cut. Juvenile graves are much smaller and nearly impossible to detect where there are only cuts for individual graves.

Cillíní are children's burial grounds, which often reuse prehistoric and medieval enclosures, for unbaptised children. It is unlikely any geophysical technique will detect individual graves; rather, if present, the enclosure ditch or wall is likely to be detected. As with adult inhumations, there is no measurable magnetic enhancement of the grave fill, but grave goods may produce a magnetic response.

4.10.1. *Associated Archaeological Features*

Stonework associated with graves will often contrast with surrounding materials when using electromagnetic and resistivity techniques. This is due to the compaction of the material and difference or variation in the conductive properties between the local geology, soil matrix, and stone. Stone can also cause moisture variations which would be easily detected by resistivity and some electromagnetic techniques depending upon the salt content of the retained water.

4.10.2. *Effects of Soil Characteristics on Detectability*

The primary factors in determining whether graves can be detected or easily identified are the physical and chemical properties of the surrounding soil matrix. Low conductivity and highly permeable soils are often the most suitable for GPR survey because they optimise the maximum penetration depth of the radar signal and reduce signal loss and strong reflections from standing water and waterlogged soils. While most archaeological graves are shallow (less than 1m below the original ground surface), there are instances where rubble or modern made ground can create a thick overburden over the original topsoil. In both scenarios, the soils and moisture content have a significant impact on the success of GPR detecting graves.

Soils are comprised of sand, silt, and clay particles. Sands are particles of quartz, feldspar, and mica that range in size from 2 – 0.02mm. Silts are particles of quartz, feldspar, and mica that range in size from 0.02 – 0.002mm (Osman 2013). Clays are particles of minerals such as kaolinite, smectite, vermiculite, illite, chlorite, and hydrated aluminium and iron oxides which are less than 0.002mm in size (Osman 2013). Soils are further defined by their texture and composition. For example, loams

are soils which have an equal representation of the properties of sand, silt, and clay (approximately 40% silt, 40% sand, and 20% clay) (Osman 2013).

Table 11: Size and properties of soil particles (adapted from Osman 2013)

Soil Particles*	Diameter (in mm)	Soil Particles**	Diameter (in mm)	Properties
Coarse sand	2.00-0.20	Very coarse sand	2.00-1.00	Little or no capacity to hold water and nutrients
		Coarse sand	1.00-0.50	Bind to other particles
		Medium sand	0.50-0.25	Loose when wet
Fine sand	0.20-0.02	Fine sand	0.25-0.10	Very loose when dry
		Very fine sand	0.10-0.05	
Silt	0.02-0.002	Silt	0.05-0.002	Low to medium capacity to hold water, nutrients, and other particles
Clay	<0.002	Clay	<0.002	High capacity to hold water, nutrients, and other particles

*ISSS System   **USDA System

12 soil classes are used to define soils as clayey, sandy, silty, or loamy by their texture. The 12 classes are based on the soil's mechanical composition – the proportion of sand, silt, and clay particles in the soil, as demonstrated in Table 12. The texture affects the soil's ability to retain water and nutrients, as well as its infiltration rate and leaching (see Table 13). As indicated by the particle size, sandy soils rapidly absorb and drain water due to their high proportion of medium-large coarse particles. Whereas, the smooth, compact nature of clay particles allow clayey soils to retain water and nutrients.

Table 12: Particle composition of the soil texture classes (adapted from Osman 2013)

Texture	Textural Class	Percentage Composition		
		Sand	Silt	Clay
Coarse	Sand	80-100	0-20	0-20
	Loamy Sand	70-80	0-30	10-15
	Sandy Loam	50-80	0-50	0-20
Medium	Loam	30-50	30-50	0-20
	Silt Loam	0-50	50-100	0-20
	Silt	0-20	90-100	0-10
Fine	Sandy Clay Loam	50-80	0-30	20-30
	Clay Loam	20-50	20-50	20-30
	Silty Clay Loam	0-30	50-80	20-30
	Sandy Clay	50-70	0-20	30-50
	Silty Clay	0-20	50-70	30-50
	Clay	0-50	0-50	30-100

Table 13: Physical characteristics of the soil texture classes (adapted from Osman 2013)

Soil Texture Class	Infiltration	Water-holding Capacity	Nutrient-holding Capacity	Aeration	Leaching
Sand	Good	Poor	Poor	Good	High
Silt	Medium	Medium	Medium	Medium	Medium
Clay	Poor	Good	Good	Poor	Low
Loam	Medium	Medium	Medium	Medium	Medium

4.10.3. *Effects of Land Cover on Detectability*

Survey data is also affected by the present and historic land coverage of the study area. As is demonstrated in Table 14, ground cover can have an impact on the data quality and subsequently, the interpretability of the data.

Table 14: Effects of land use/land cover on geophysical survey data (adapted from Bonsall et al. 2014a)

Land Use	Description	Effect on Geophysical Survey Data
Grassland Pasture	Open area of grasses where livestock graze	The groundcover is suitable for all geophysical survey methods, but livestock can interfere with magnetic and electromagnetic techniques.
Arable	Land that has been (or can be) ploughed for crops	Positional accuracy and data quality are reduced as crop height increases and crop brushes against the instruments.
		Ground contact often cannot be achieved for earth resistance and topsoil magnetic susceptibility surveys.
		Data quality can be reduced if the surveyor is walking on cut silage as it is slippery and produces an uneven surface.
		Surveyor may have to dummy around silage bales and other small obstacles.
Undifferentiated or Unmanaged Grassland	Open area of grasses that are not managed/grazed by livestock where scrub and coarse grasses grow	Hummocks and depressions formed by the unmanaged grasses can reduce positional accuracy and introduced surveyor walking effects.
Bog/Reclaimed Bog/Wetland/Marsh	Area of poorly drained ground that is often waterlogged or 'spongy' and surrounded by a body of water	The soft ground can reduce positional accuracy and introduced surveyor walking effects. In reclaimed bogs and unprotected wetlands, imported metals and other materials may reduce the quality of magnetic and electromagnetic surveys.
Woodland	Area of land covered with trees	Tree cover does not allow for a comprehensive, high-resolution survey and reduces the interpretability of the site.
Ploughed	Land where the soil is loosened or turned and seeds will be planted	Positional accuracy and ground contact are impeded by the uneven soil (except in recently seeded soils). Recently ploughed soils can also create "noise," sometimes making any geophysical survey ineffective.

If a grave is in woodland or unmanaged grasslands, its detectability will be reduced due to the availability of surveyable areas and noise introduced by the surveyor as they walk over the slight topographical variations caused by unmaintained grasses and roots. Moisture differences around

large trees will be detected by GPR, EMI, and resistance techniques. Large, widespread moisture differences can obscure graves or be misinterpreted as graves.

These factors impact the detectability of graves with any geophysical techniques. In some conditions, certain techniques are not suitable for detecting any archaeological features, let alone small features like graves. The first step in improving the detection and interpretation of graves is improving the quality of data collected. There are several guidance documents available for archaeogeophysical survey; however, this project primarily focused on those with European applications in order to create an interactive tool for determining the most appropriate survey parameters for a site in Western Europe.

5. SURVEY PARAMETERS TOOL METHODS

5.1. Introduction

As an initial step in improving the interpretability of data, the data acquired during a survey should be improved through the use of appropriate survey parameters. Currently, the guidance for surveys in Ireland is provided in flowchart and decision table format; while, the guidance for surveys in the rest of Europe (primarily England) is provided in text and decision table format. To make this guidance more accessible, this project aimed to translate the existing guidance to an interactive format which accounts for regional controllable and uncontrollable factors that may impact on the survey and data quality but encourages commercial clients and surveyors to consult the appropriate guidance documents before a survey. Further discussion on the structure of survey guidance is provided in Chapter 3.3. The appropriate survey parameters were assigned to polygons in shapefiles using openly available data and decision trees and tables implemented through SQL queries in ArcGIS. Additional details on the approach taken to create the interactive, user-friendly output for survey guidance is discussed hereafter.

5.2. Format

The survey parameters tool is implemented in shapefiles for desktop GIS software and ArcGIS web apps, and the outputs provided are based on results from a decision tree which was converted to a decision table. The rank table (divided into polygons with attributes) is available in the desktop version of the tool and can be viewed in Table 16. The compiled vector maps are provided in both web app and desktop versions. Due to file size limits, the dataset was split to create two web apps, one for the UK and one for Ireland. The online versions are limited to displaying only the suggested survey parameters due to file size limits and licencing concerns. However, the desktop version provides the user with all data available for a point (including the bedrock geology, superficial geology, soils, and land cover) and will be made available on a case-by-case basis according to licencing agreements.

5.3. Data

The primary decision-making variables were derived from controllable and uncontrollable factors which determine the quality of acquired data. These factors are provided by Historic England (David et al. 2008), Bonsall et al. (2014a), and Schmidt et al. (2015). The output (appropriate survey parameters) is heavily dependent on environmental factors such as bedrock and superficial geology,

and land cover. Table 15 details the factors accounted for in the decision table and their associated variables. In the output, users are referred to the relevant guidance document for further information on those factors, such as weather conditions and surveyor competency, which are not accounted for in the initial decision. The following data were ranked on a scale of 0 (do not attempt) to 5 (highly suitable) for the four most commonly used geophysical techniques in archaeological prospection, according to the available survey guidance document discussed in Chapter 3.

Table 15: Factors accounted for in deciding the best output data in the survey parameters tool

Controllable Factors	Uncontrollable Factors
Technique selection	Soil type
Data acquisition methods	Land cover
Spatial resolution	Bedrock geology
	Surface geology
	Sources of interference

Input data were derived from:

- British Geological Survey (2019) geology maps based on BGS data under an Open Government Licence,
- Esri UK satellite imagery and in-built automatic classification of land cover using the national land cover database in ArcGIS Pro v2.2,
- Corine land cover data for Ireland (Environmental Protection Agency Ireland 2019) licenced under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence,
- Irish soils data derived from Creamer et al. (2016),
- Geological Survey Ireland (2019) data licenced under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

Output data were derived from guidance supplied by Schmidt et al. (2015) and Bonsall et al. (2014a) which was initially translated to a decision table.

Land cover, bedrock geology, superficial geology, and, for the Republic of Ireland, soils data were imported into ArcGIS as shapefiles. Land cover data for Great Britain were acquired by classifying aerial imagery using the built-in unsupervised image segmentation and classification toolboxes based on the NLCD land cover classes in ArcGIS Pro. The shapefiles were simplified and data reclassified into classes conventionally used in the archaeogeophysics literature to reduce the file size and improve usability, as demonstrated in the classes shown in Table 16. The simplified

shapefiles were joined, and new polygons created from overlapping fields. The new polygons contained the land cover, geological, and soil characteristics available. These characteristics were used to derive the appropriate survey parameters for each polygon. Due to the assembly of multiple datasets from multiple countries and coordinate systems, the merged polygons were converted to the WGS 1984 coordinate system to allow the end-user to use their site coordinates in the search function.

There was limited data available for Northern Ireland. Freely available data were restricted to the bedrock geology classifications and raster datasets which were classified for land cover types. Due to this lack of data, guidance for Northern Ireland included in the first release of the tool should be used with caution. As more data becomes available, full guidance for Northern Ireland and additional European countries will be included in subsequent releases.

Table 16: List of simplified categories and classes used in the decision table for determining the most appropriate technique for each unique polygon in the survey parameters tool

Land Cover

Arable	Grassland	Outcrop	Transportation
Bog	Heather	Pasture	Urban
Burnt Areas	Marsh	Rock	Wetland
Coastal	Mineral Extraction Site	Sediment	Woodland
Cultivated	Moors & Heathland		

Bedrock Geology

Anorthosite	Gneissose Semipelite & Gneissose Psammite	Mudstone, Sandstone & Limestone	Sandstone & Mudstone
Appinite	Gneiss & Granite	Mudstone, Siltstone & Sandstone	Sandstone & Siltstone
Basalt	Granite	Mudstone, Siltstone, Sandstone & Limestone	Sandstone, Breccia & Conglomerate
Breccia & Metabreccia	Gravel, Silt, Sand & Clay	Mudstone, Siltstone, Sandstone, Coal, Ironstone & Ferricrete	Sandstone, Limestone & Argillaceous Rocks
Breccia, Conglomerate & Sandstone	Greywacke	Mylonitic Rock	Sandstone, Mudstone, Siltstone & Conglomerate
Chalk	Lava	Pelite	
Clay, Silt & Sand	Limestone	Psammite	Sandstone, Siltstone & Mudstone
Clay, Silt, Sand & Gravel	Limestone & Calcareous Sandstone	Psammite & Pelite	Schist
Conglomerate	Limestone & Mudstone	Psammite & Semipelite	Sedimentary Rock
Conglomerate, Sandstone, Siltstone, & Mudstone	Limestone, Sandstone, Siltstone & Mudstone	Psammite, Pelite, Semipelite & Calcsilicate Rock	Semipelite
Diamictite	Mafic		Semipelite & Pelite
Diorite	Mafite	Psammite, Semipelite & Pelite	Serpentinite
Dolerite	Metalimestone	Pyroclastic Rock	Shale
Dolomitised Limestone	Metasedimentary Rock	Quartz Arenite	Siltstone
Dolostone	Migmatitic Rock	Quartzite	Siltstone & Sandstone
Felsic	Mudstone	Rhyolite	Slate
Gabbro	Mudstone, Chert, & Smectite-Claystone	Sand, Silt & Clay	Syenitic Rock
Gneiss	Mudstone, Sandstone & Conglomerate	Sandstone	Ultramafitite
Gneissose Psammite & Gneissose Semipelite		Sandstone & Conglomerate	Volcanic
			Wacke

Drift/Superficial Geology

Airfield/Airport	Gravels Derived from Lower Palaeozoic & Devonian Sandstones	Lacustrine Clays	Till Derived from Devonian & Carboniferous Sandstones
Alluvium	Gravels Derived from Lower Palaeozoic Sandstones	None	Till Derived from Devonian & Carboniferous Sandstones & Shales
Alluvium (Clayey)		Peat	
Alluvium (Gravelly)	Gravels Derived from Lower Palaeozoic Sandstones & Shales	Pier	Till Derived from Acidic Volcanic Rocks
Alluvium (Sandy)		Lacustrine Sediments	Till Derived from Granites
Bedrock Outcrop or Subcrop	Gravels Derived from Lower Palaeozoic Shales	Lake Marl	Till Derived from Devonian Sandstones
Blanket Peat	Gravels Derived from Namurian Sandstones & Shales	Landslip	Till Derived from Lower Palaeozoic & Devonian Sandstones
Brickearth		Marine Beach Sands	
Clay with Flints	Gravels Derived from Granite	Marine Gravel & Sands (often Raised)	Till Derived from Lower Palaeozoic Sandstones & Shales
Crag Group	Gravels Derived from Quartzite	Raised Marine Deposits	Till Derived from Lower Palaeozoic Shales
Cut Over Raised Peat	Industrial	River Terrace Deposits	Till Derived from Metamorphic Rocks
Embankment	Irish Sea Till Derived from Cambrian Sandstones & Shales	Sand	Till Derived from Namurian Sandstones & Shales
Eskers Comprised of Gravels of Basic Reaction		Sand & Gravel	
Estuarine Silts & Clays	Irish Sea Till Derived from Limestones	Tidal Marsh	Till Derived from Silicified Limestone & Cherts
Glaciomarine Sediments	Irish Sea Till Derived from Lower Palaeozoic Sandstones & Shales	Till	Till Derived from Limestones
Gravels Derived from Cambrian Sandstones & Shales		Till Derived from Cambrian Sandstones & Shales	Till Derived from Quartzites
Gravels Derived from Devonian Sandstones	Kartsified Bedrock Outcrop or Subcrop	Till Derived from Carboniferous Sandstones & Cherts	Urban
	Lacustrine Deposits		Windblown Sands
Gravels Derived from Limestones			

Soils

Alluvium	Loamy	Rock	Tidal Marsh
Clay	Peat	Sandy	Urban
Island			

5.4.   Suitability Mapping Methods

Input data for the bedrock geology, superficial geology, soils and land cover were manually assigned a value on a scale of 0 (do not attempt) to 5 (highly suitable) for EMI, GPR, magnetometry and resistance techniques according to the advice given in by Bonsall et al. (2014), Schmidt et al. (2015) and David et al. (2008). Table 19 shows an excerpt of these raw suitability data. The complete data are shown in Appendix H. These initial rank data were used with analytic hierarchy process (AHP) as a tool for multi-criteria decision analysis (MCDA) in GIS.

AHP is a commonly used MCDA technique in GIS applications for suitability analysis, particularly for land use and resource management (Mendoza 2000, Duc 2006, Vaidya and Kumar 2006, Chandio et al. 2013, Malczewski and Rinner 2015, Al-shabeeb 2016, Chaudhary et al. 2016), that was initially developed by Saaty (1980). AHP uses multiple pairwise comparison matrices to determine the highest-ranking value by comparing all possible pairs of criteria, or input factors, against each other (Marinoni 2004, Duc 2006). As described in Figure 39, AHP can be used to create a suitability map based on the criteria weight (C_w), criterion score for each input (C_s), and consistency ratio.
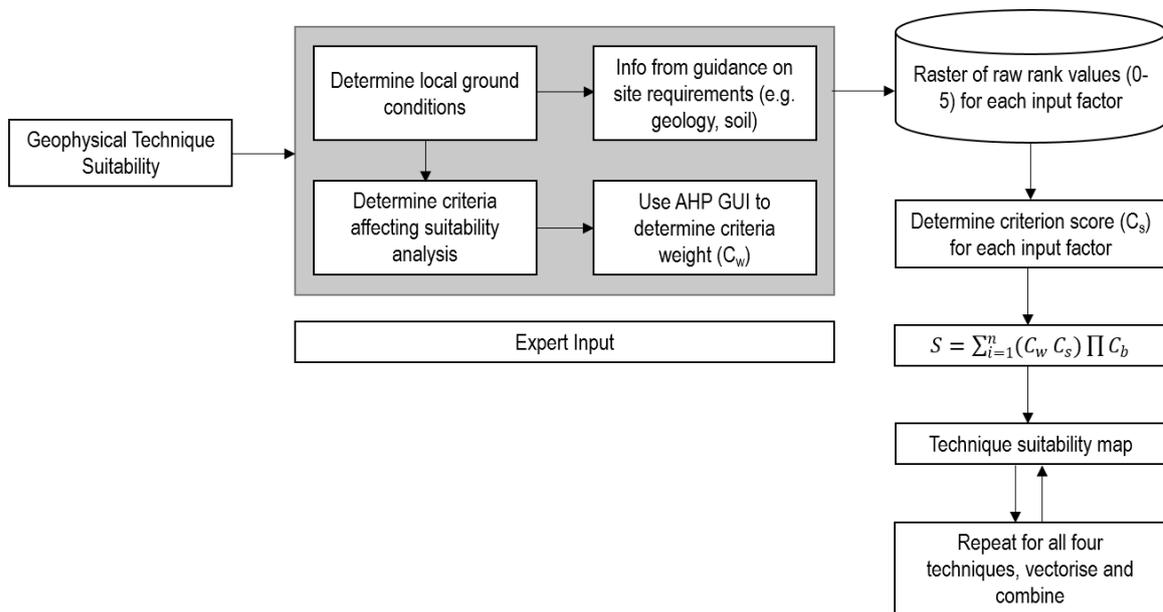
Figure 39: Flowchart demonstrating the AHP process for creating a suitability map for this project Adapted from Duc (2006)

The weighted linear combination using QR decomposition on the eigenvalues in the comparison matrix to calculate the suitability value from the C_w, C_s, and Boolean value of criterion on a scale set by the user (C_b), as shown in Equation 5.1: AHP Suitability Analysis Equation (Marinoni 2004, 2017, Duc 2006). In the model for this project, C_b was measured on a scale of 0-5.

5.1: AHP Suitability Analysis Equation

$$S = \sum_{i=1}^{n}(C_w\,C_s) \prod C_b$$

$$S = \text{suitability}$$

$$C_w = \text{criteria weight}$$

$$C_s = \text{criteria score for each input}$$

$$C_b = \text{Boolean value of criterion}$$

The consistency ratio (CR) is used to measure the variation allowed for the pairwise comparisons. Saaty (1980) recommends the CR is less than or equal to 0.1, to indicate suitable weights have been used in the comparison matrix and the pairwise comparisons are consistent.

For this project, the AHP was carried out using the AHP tool in the GIS Decisions add-in for ArcGIS Desktop (Marinoni 2017). The workflow for generating suitability maps is shown in Figure 40. The rank values stored in the attribute tables are used when converting the polygons to rasters to be used as inputs for the AHP tool. The outputs from the AHP tool runs are then converted into vectors and merged, so the suitability scores for all techniques are stored in the same attribute table. The values used in the Preference Matrix for the Ireland and Northern Ireland dataset are shown in Table 17, and the values for the GB data are shown in Table 18. A Preference Matrix uses a scale of 1 (equal preference for input criteria) to 9 (strong preference of one or more criteria over others) to compare and weight the input criteria, in this case, the soil, superficial geology, land cover and bedrock geology suitability (Marinoni 2017).

Table 17: AHP Preference Matrix for the Ireland and Northern Ireland dataset

	Soil	Superficial Geology	Land Cover	Bedrock Geology
Soil	1	1	1	1
Superficial Geology	1	1	1	0.5
Land Cover	1	1	1	0.5
Bedrock Geology	1	1	2	1
Superficial Geology:	20.36	Land Cover:	20.36	CR: 0.022
Soil:	24.627	Bedrock Geology:	34.654	

Table 18: AHP Preference Matrix for the Great Britain dataset

	Land Cover	Superficial Geology	Bedrock Geology	
Superficial Geology	1	1	1	
Land Cover	1	1	0.5	
Bedrock Geology	2	1	1	
Superficial Geology:	32.748	Land Cover:	25.992	CR: 0.062
		Bedrock Geology:	41.26	
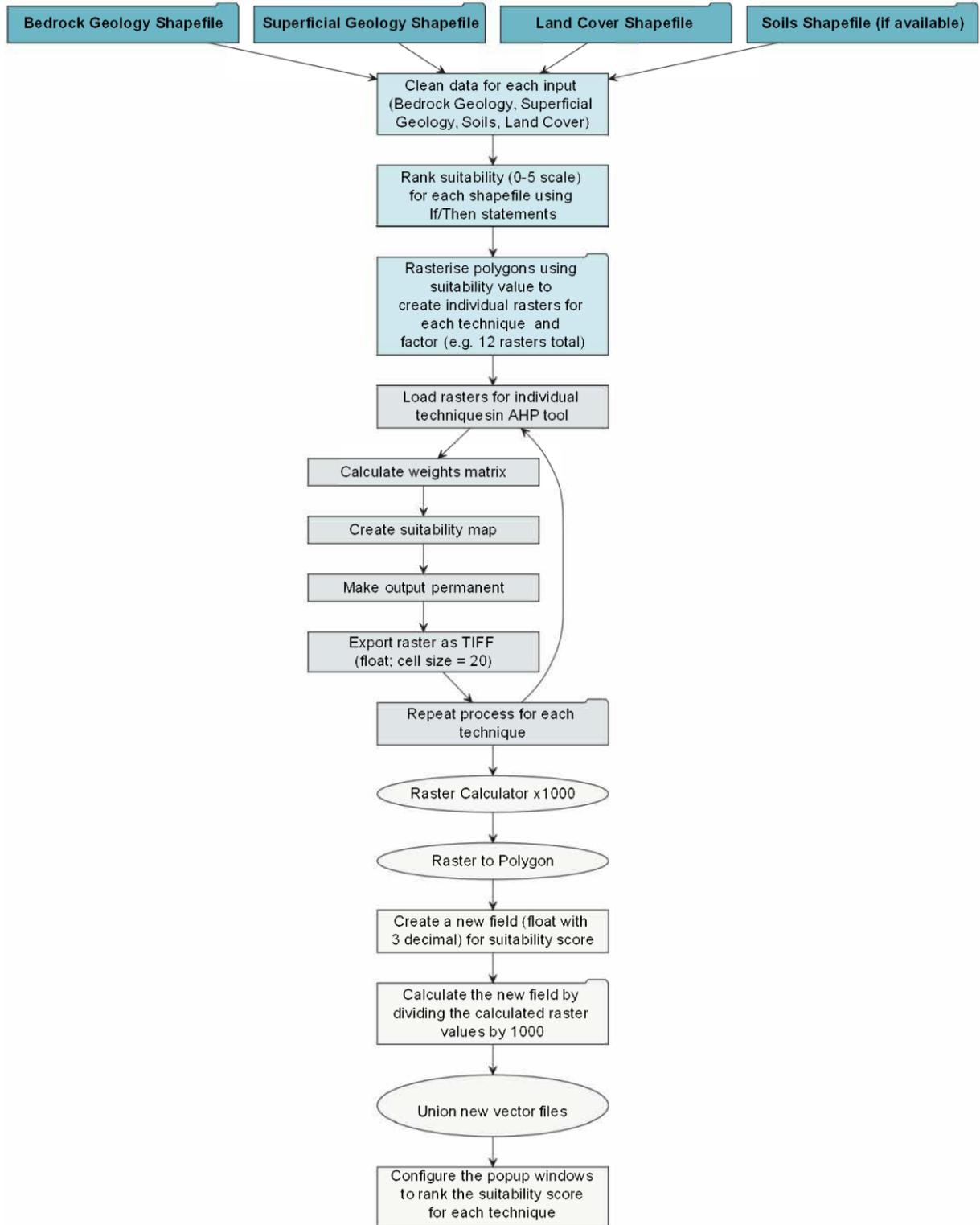
Figure 40: Workflow for creating the final combined suitability map for this project

Table 19: Excerpt of rank values used for suitability analysis

Bedrock Geology				
Class	**Magnetometry Suitability**	**EMI Suitability**	**Earth Resistance Suitability**	**GPR Suitability**
Anorthosite	1	2	5	2
Appinite	4	3	4	3
Basalt	1	2	5	2
Breccia and Metabreccia	3	4	3	4
Breccia, Conglomerate and Sandstone	3	4	3	4

Superficial Geology				
Class	**Magnetometry Suitability**	**EMI Suitability**	**Earth Resistance Suitability**	**GPR Suitability**
Airfield/Airport	1	4	3	4
Alluvium	3	2	4	3
Alluvium (Clayey)	3	2	4	3
Alluvium (Gravelly)	3	2	4	3
Alluvium (Sandy)	3	2	4	3

Soils				
Class	**Magnetometry Suitability**	**EMI Suitability**	**Earth Resistance Suitability**	**GPR Suitability**
Tidal Marsh	0	2	0	1
Rock	1	3	1	1
Island	0	1	0	0
Loamy	3	3	3	3
Alluvium	2	3	4	2

Land Cover				
Class	**Magnetometry Suitability**	**EMI Suitability**	**Earth Resistance Suitability**	**GPR Suitability**
Arable	5	5	5	4
Bog	1	4	1	3
Burnt Areas	1	2	3	3
Coastal	1	3	1	2
Cultivated	5	5	3	3

5.5. User Interface

The GUI was developed using the Esri web app builder with one app for the UK and one for Ireland due to file size limits. The design was heavily modified to support the extensive database as the primary aim of the map is to allow users to search for their site location easily. A limited dataset, consisting of only the relevant survey parameter fields, can be viewed in the web browser. The full dataset, including supplementary data fields, is available for use in desktop ArcGIS Pro or ArcMap. In both versions the user can search for a central coordinate of their survey area; once directed to their location they can select the polygon by clicking on it. Once a polygon is selected, a popup window containing the relevant data is displayed. The pop-up windows were designed using Arcade to show the techniques ranked from most suitable to least suitable for the selected coordinates, as well as placename information, the suggested minimum survey resolution for each technique and survey aim discussed and links to additional guidance documents should the user need more advanced information. The example output in the web browser is shown in Figure 41, and an example output of the full dataset provided in the desktop version is shown in Figure 42.
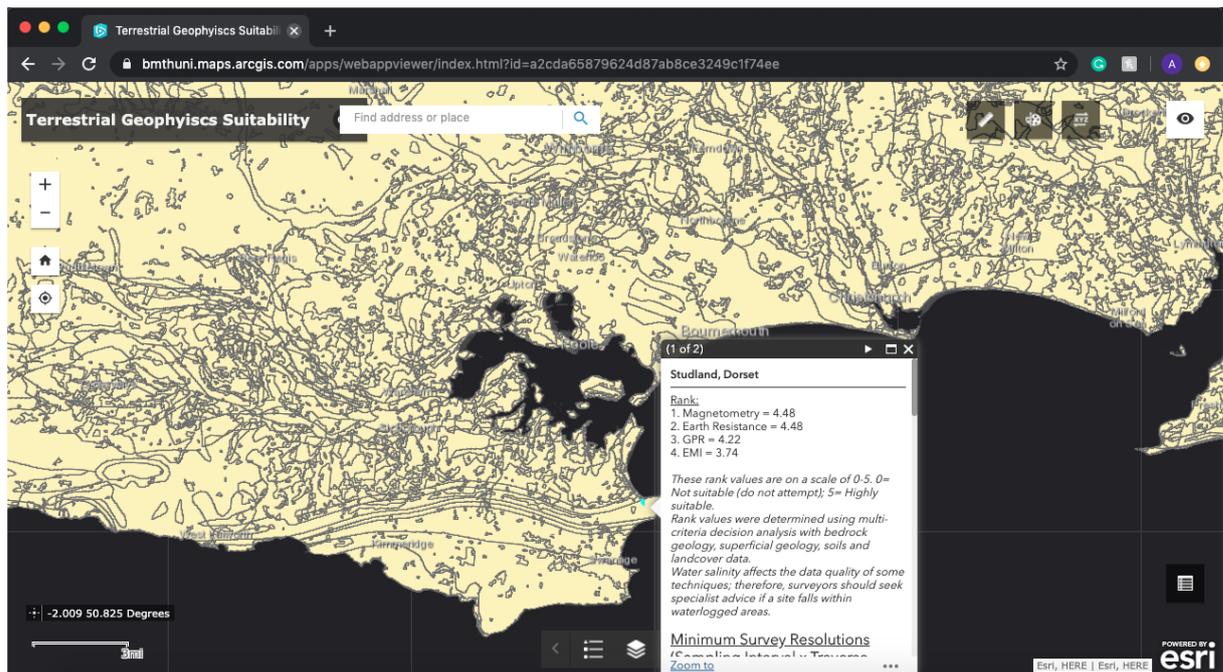
Figure 41: Example of an output in the web browser version of the survey parameters tool. The image shows the information the end-user is given once they have selected a polygon and the location search capabilities of the app.

The survey resolution guidance provided in this tool is derived from the relevant guidance document(s) for the region. The ArcGIS Online web app can be accessed through the browser using the default link provided by Esri. The parameters layers are also available in kmz file format for import into Google Earth for individuals who do not have access to ArcGIS.
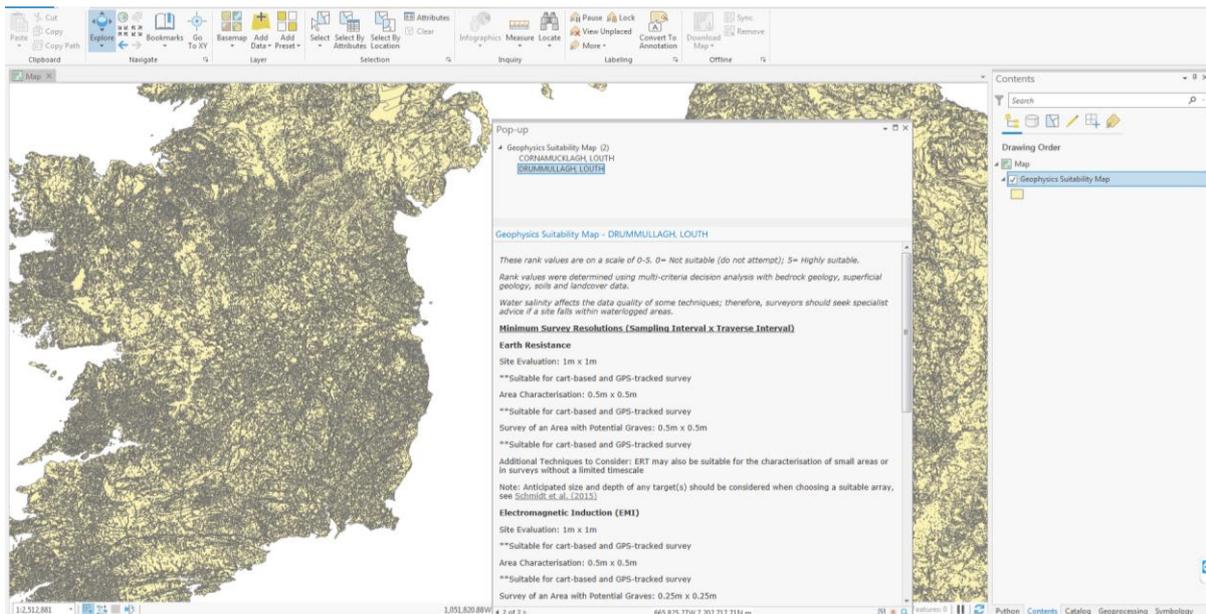
Figure 42: Example of an output in the desktop version of the survey parameters tool. This image demonstrates the end-users view of the shapefile and the information they are provided with in the configured pop-up window.

5.6. Guidance App Implementation

Using the survey data from Buckland Rings, Lymington, UK, in Figure 44 as an example, the survey parameters tool provides viable feedback to both commercial clients and surveyors. Three series of geophysical surveys have been carried out on the site since 2017 – gradiometry (Hagan et al. 2017), GPR (Howard 2018), and EMI (Monteith and Green 2019). The guidance app indicates that earth resistance would be the most suitable technique, with magnetometry and EMI as suitable supplementary techniques, for the geology and land cover at Buckland Rings (National Grid Reference: SZ315968). Results from the three surveys indicate that gradiometry was the most successful in delineating responses of archaeological potential. EMI was able to delineate larger features but was not able to clearly delineate any small features; however, this was likely due to variations in survey conditions and directional bias in the survey parameters. GPR was able to survey to a greater depth below ground level than the other techniques, but the survey area was limited by the topography and vegetation. The multiple surveys conducted at Buckland Rings demonstrate the validity of the survey guidance app and can be used to encourage the widespread use the guidance within commercial, volunteer, and research geophysics where necessary.

6. MACHINE LEARNING

6.1. Introduction

In recent years with the growth of mechanised, GPS-tracked survey equipment, the average size of surveys has also increased due to the ease and speed of surveying, as demonstrated in Table 20. However, with these large datasets comes the difficulty of identifying small features visually. It is well known that all humans see and focus on different aspects of an image, which can lead to a subjective, biased interpretation of a dataset (Verdonck et al. 2019). For example, a medium-sized survey dataset like those in Figure 43 and Figure 44 would have many different interpretations depending on the interpreter.

Table 20: Examples of large-scale surveys conducted 1996 - 2018 (adapted from Bonsall 2014)

Location	Hectares	Survey Year(s)
Land off Bury Road, Woolpit, Suffolk	19	2018
Marden Barrows and Wilsford Henge, Wiltshire	20	2012-2013
Stonehenge Southern WHS, Diamonds Field, Wiltshire	26.9	2015
Six Hills Road Solar Development, Leicestershire	27	2014
Land at Stonemead Farm, Worminster Down, Somerset	43	2014
Land at Moreton Hall, Bury St. Edmunds, Suffolk	46	2016
Land at Monks Cross, York	50	2014
Westacott, Barnstaple, Devon	55	2017
Selinus, Sicily	c. 65	1999-2003
Wroughton Airfield Solar Park, Swindon, Wiltshire	65	2013
Wroxeter, Shropshire, UK	70	Pre-1999
M3 Clonee to North of Kells Motorway, Ireland	105	2000
Carnuntum, Austria	>110	Pre-2003
Luton Hoo, Luton, Bedfordshire, UK	120	2003
Moorside Project, Sellafield, Cumbria	142	2015
N6 Galway to East Ballinasloe Motorway, Ireland	160	2004
M20 Cork to Limerick Motorway, Ireland	182	2010
BREBEMI Motorway, Italy	217	2009
Portus, Italy	c. 220	1998-2005
Commercial Project, UK	250	2008
Commercial Project, UK	400	2010
Stonehenge Hidden Landscapes Project, UK	820	2011-2013
Heslerton Parish Project, North Yorkshire, UK	>1200	1996-2006
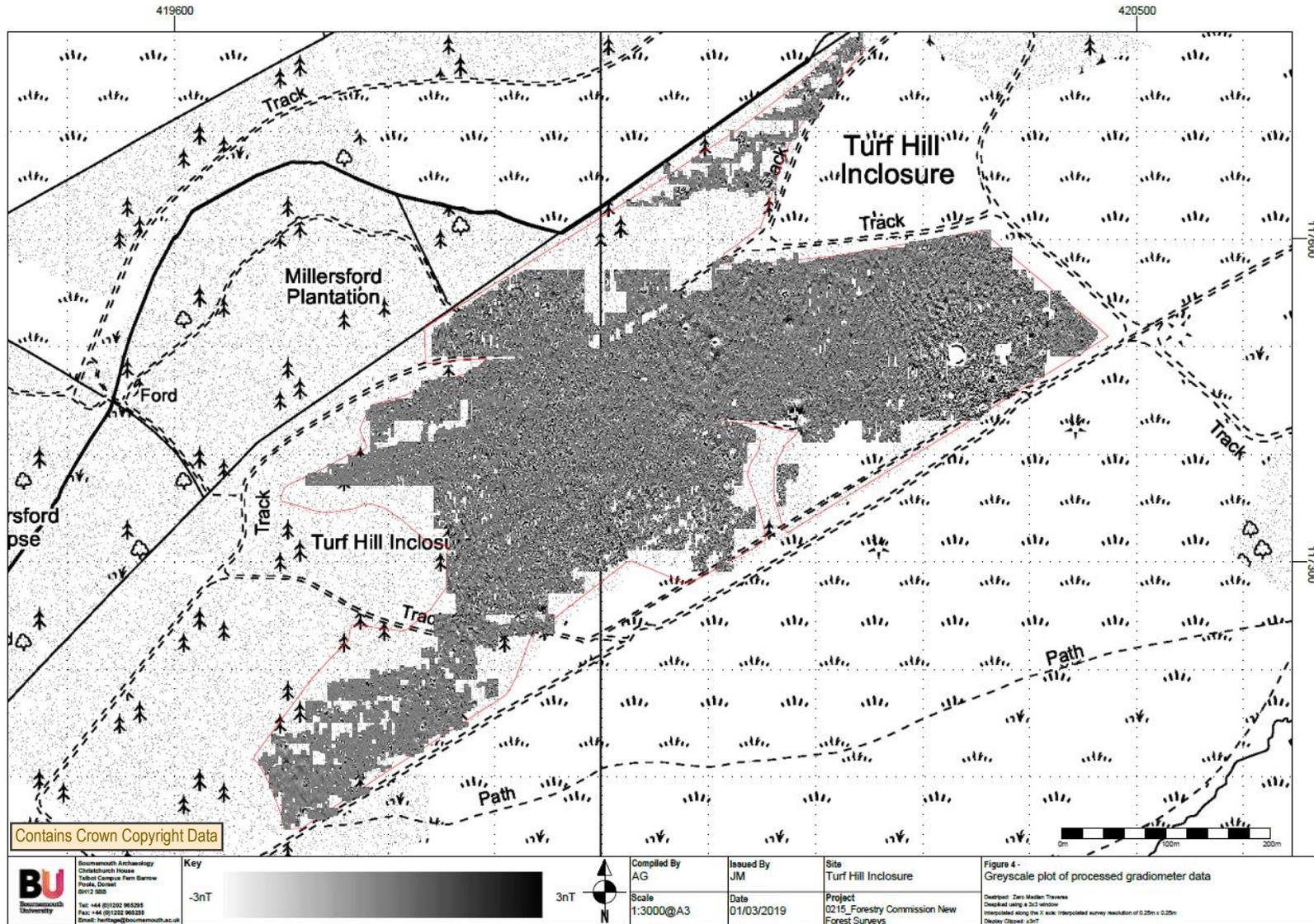
Figure 43: Example of a large magnetometer survey dataset (Bournemouth University 2019)

Figure 44: Example of a large EMI apparent conductivity dataset (Monteith and Green 2019)

Furthermore, it would not be unexpected for the human eye to miss small details like pits or postholes when there are much larger features to focus on. This natural bias can lead to misinterpretation or missed features, even in smaller datasets. In other fields which have complex small datasets, such as medical imaging, this bias also occurs as well as the discrepancy between human-led interpretations (Nattkemper et al. 2003, p. 35). In their study on the detection of lymphocytes in fluorescence micrographs, Nattkemper et al. (2003, p. 40-41) identified a correlation between the experience level of human interpreters and the accuracy of their detections, demonstrating that the most experienced interpreters achieved the highest accuracy. Overall, Nattkemper et al. (2003, p. 40) achieved a similar accuracy to a medium-trained human expert with artificial neural network classifiers and decreased the time required to analyse each micrograph. Similar trends are expected to occur in the comparison of human- and machine-led interpretation of geophysical data, especially where small features are present. Chapter 1.2 discusses the accuracy of human-led detection of potential graves in GPR data.

Overlooking other small features, such as a potential pit or kiln, in geophysical data has little consequence to archaeological evaluations because they are often associated with wider spread settlement activity and their presence is assumed. Whereas, not all cemeteries and burial grounds are associated with detectable settlement activity, making them more challenging to detect using surrounding features to localise areas of interest for high-resolution surveys. Not detecting graves also has a much more significant impact on evaluations as they require ethical considerations and additional expert staff and can cause delays while the appropriate licences are acquired if not already held for a site. Of utmost importance are the ethical implications for disturbing and recovering human remains. Where possible human remains should not be disturbed (Advisory Panel on the Archaeology of Burials in England 2017, p.1); however, if unavoidable, human remains should only be excavated by qualified excavators once the appropriate licenses are obtained (BABAO Working-group for ethics and practice 2010, p. 9; Institute of Archaeologists of Ireland 2016, p. 3-4). Skeletal remains can be costly for a client and require more time for the archaeologists to recover and record them to a suitable standard, which adds further timescale pressures for the archaeologists and client (OSSAFreelance 2012, p. 6-8). This impact could be mitigated by adapting existing machine learning methods to interpret complex geophysical datasets and an archaeologically complex feature. The knowledgebase of the viability of machine learning in fields like archaeology already exists, which begs the question, why has it not been widely applied to commercial geophysics? Is it because the data available is not a real-world representation of grave responses? Or is it merely because some

archaeological features are too difficult for the learning machine to identify them accurately and confidently?

Because these machine learning methods have been widely applied to many other complex problems in the last seven years since the development of AlexNet (Krizhevsky et al. 2012), this project aims to build on the existing knowledgebase to determine if machine learning is a viable approach for widespread automatic interpretation of features in commercial and research geophysics.

Automatic feature detection software exists for infrastructure monitoring, IED detection, and simplistic materials identification in geophysical datasets (Pasolli et al. 2008, 2009, Singh and Nene 2013b, Qiao et al. 2015, Zaremba et al. 2016b). The algorithms and learning processes can be readily applied to geophysical datasets to isolate grave-like responses through supervised and potentially semi-supervised machine learning. By analysing the presence of typical medieval burial practices while accounting for a range of uncontrollable factors (i.e. geology, weather conditions), classifiers for the expected geophysical response of these inhumations, particularly the response in GPR data, can be estimated to create a dataset of "knowns." Through machine learning, acquired GPR data ("unknowns") can be compared to the "known" or labelled dataset to probabilistically identify grave-like responses which may not be identified during manual interpretation. This research project aims to create computational tools that will improve survey data and subsequent interpretations in archaeological GPR surveys, mainly focusing on the implications of their implementation in commercial archaeology. Through retraining existing convolutional neural networks on raster outputs of 2D and 3D GPR data, grave-like responses will be highlighted in any GPR B- and C-scan datasets and given an associated confidence score.

6.2. Deep Learning

There are three primary types of machine learning: unsupervised, semi-supervised, and supervised as defined in Figure 46 (Singh et al. 2016, Maulik and Chakraborty 2017). A brief overview of how labelled and unlabelled datasets are used in the different learning machines is provided in Figure 45 and Figure 46.
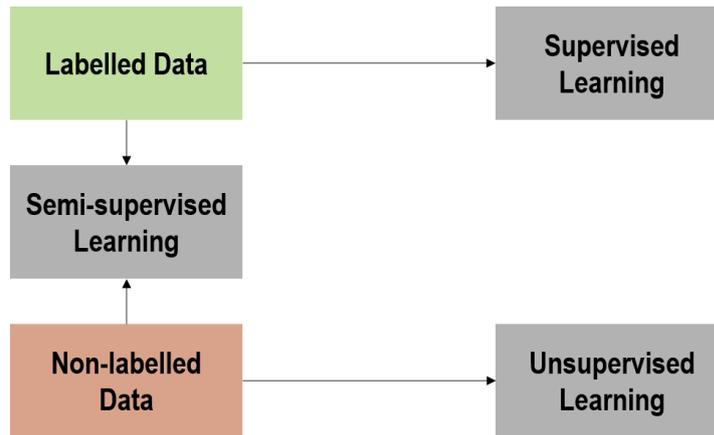
Figure 45: Brief overview of datasets and how they relate to the three main types of machine learning

In supervised learning, the learning machine is provided with labelled data (input data with the associated target value(s), such as an image broken down into labelled contents) and is expected to learn the data for target mapping, in order to predict (or estimate) target values for previously unobserved input data. The dataset (described as $D=\{x,y\}_{n=1}^{N}$) includes input features (x) and labels (y), which are often images and class labels or bounding boxes with associated class labels for classification or object detection tasks (Litjens et al. 2017, p. 62).

Contrary to supervised learning, in unsupervised learning, the learning machine is only provided with the unlabelled input data component, not its associated values. This unsupervised approach is often used to identify natural patterns (structure within the data) that cannot be identified or linked by the human eye. Unlabelled datasets are also frequently chosen over labelled datasets because they tend to be much larger, more easily obtainable, and require less annotation effort.

Semi-supervised learning is a middle ground between the supervised and unsupervised learning approaches in which the learning machine is supplied with a limited amount of labelled data. To improve the machine's predictive capabilities and accuracy, unlabelled data are also imputed to increase the size of the limited training dataset and deal with incomplete labels.

Supervised Learning

Learns patterns from labelled data

Uses learned patterns to predict future outcomes

Semisupervised Learning

Trains on a small labelled dataset in order to label large training datasets

Uses learned patterns to predict future outcomes

Unsupervised Learning

Identifies hidden patterns in unlabelled data

Clusters data based on the degree of similarity or dissimilarity

Figure 46: Comparison of the three primary types of machine learning approaches

A simplified machine learning workflow, the Cross-Industry Standard Process for Data Mining (CRISP-DM) model, demonstrating the major steps in the machine learning process and the importance of the feedback between the data preparation and modelling stages in order to improve the accuracy of the model is provided in Figure 47; a similar approach was taken in producing a suitable model in this research.
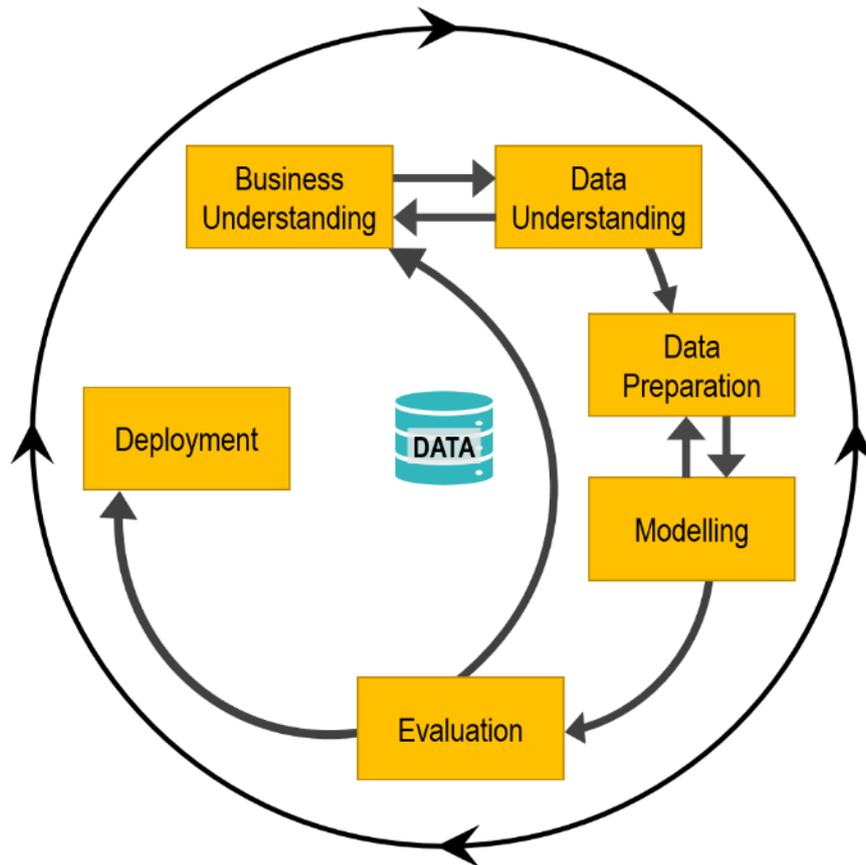
Figure 47: The Cross-Industry Standard Process for Data Mining (CRISP-DM) model indicating the six phases of data mining projects (adapted from Chapman et al. 2000, p. 10). The six phases are flexible, and their interactions will depend on the type of project and its objective.

Deep learning utilises large, deep neural networks with multiple processing layers in order to learn representations of data and patterns (LeCun et al. 2015, p. 436). Deep learning architectures are preferred over shallow architectures (e.g. decision trees and support vector machines) for computer vision tasks as they are more suitable when dealing with complex relationships in the input data (Najafabadi et al. 2015, p. 4). In deep learning, there are multiple levels of representations (extracted features/feature vectors) which are obtained by transforming the representation in layers of simple, non-linear modules (LeCun et al. 2015, p. 436). Supervised learning is more common than other machine learning methods in deep learning (LeCun et al. 2015, p. 436).

Several deep learning tools exist for training machine learning models. Table 21 summarises the most commonly used tools/libraries and applications. Fastai, PyTorch, Keras, and TensorFlow were primarily used in this project.

Table 21: Summary of commonly used deep learning tools (adapted from Ball et al. 2017, p. 11)

Tool	Summary	Website	Reference
AlexNet	A large-scale convolutional neural network with nonsaturating, neurons and a very efficient GPU parallel implementation of the convolution operation to make training faster	http://code.google.com/p/cuda-convnet/	(Krizhevsky et al. 2012)
fastai	Deep learning library for PyTorch	https://github.com/fastai	-
Keras	A high-level Python neural network library capable of running on top of either TensorFlow or Theano, which • allows for easy and fast prototyping • supports both convolutional networks and recurrent networks • supports arbitrary connectivity schemes	https://keras.io/ https://github.com/fchollet/keras	(Chollet 2015)
MatConvNet	A MATLAB® toolbox implementing convolutional neural networks with many pre-trained convolutional neural networks for image classification, segmentation	http://www.vlfeat.org/matconvnet/	(Vedaldi and Lenc 2014)
MXNet	Deep learning library for declarative symbolic expression with imperative tensor computation and differentiation to derive gradients	https://github.com/dmlc/mxnet/	(Chen et al. 2015)
TensorFlow	An open-source software library with a flexible architecture for tensor data flow graph computation.	https://www.tensorflow.org/	(Abadi et al. 2016)
PyTorch	Machine learning library for tensor computing and deep neural networks based on the Torch library	https://pytorch.org/	(Paszke et al. 2017)

A supervised machine learning approach is suitable for the grave detection problem because the variability in the morphological and geophysical characteristics between grave types will require labelling of datasets based on corresponding excavation data, whereby feature extraction algorithms, such as those using morphological parameters, are used to isolate and identify graves in images as the standard 'target.' An unsupervised approach will not be taken as it is predicted that the system may indiscriminately identify responses in the data that are not necessarily related to graves, thereby reducing the confidence of the interpretive classifications.

6.3. Neural Network Architectures

Artificial neural networks are comprised of "neurons" which are connected to each unit of the network receive input and formulate an output (Schmidhuber 2014, Maulik and Chakraborty 2017). One of the simplest neural networks is a multilayer perceptron network (MPN) with a minimum of one input layer, one hidden layer, and one output layer, as shown in Figure 48. The input layer receives the input data, where the number of nodes in the input layer is equal to the number of features in the input data, and after weights and bias are added the information is fed forward to the hidden layer(s). Weights, also called parameters, are the numeric array of a filter (e.g. edge detector). Hidden layers are any layers between the input layer and output where weights are applied, and the data can be transformed. When followed by an activation function in the output layer, the data from hidden layer(s) are fed as output and predictions are made based on the number of classes used (O'Shea and Nash 2015).
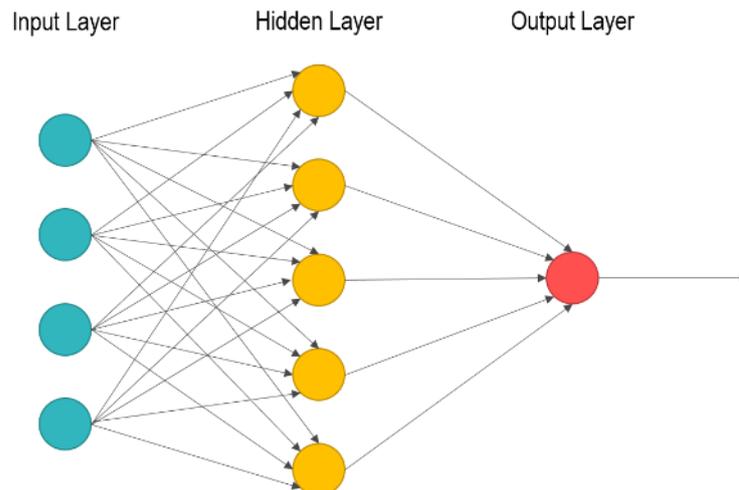
Figure 48: Architecture of the simplest version of a multilayer perceptron network, a basic feed-forward neural network

These simplistic artificial neural networks can still prove useful for less complex inputs, such as the handwritten digits in the Modified National Institute of Standards and Technology (MNIST) database. However, MPNs are now outperformed by convolutional neural networks (CNNs) and other deep learning architectures when solving complex computer vision tasks. MPNs are inefficient for complex tasks because each perceptron is connected to every perceptron, creating a large number of redundant parameters (as defined in Equation 6.1). More recent architectures can reduce the number of parameters while still maintaining high training accuracy and less cumbersome computing.

6.1: Definition of parameters in a multilayer perceptron network

$$\text{parameters} = perceptron_1 \times perceptron_2 \times perceptron_3 \cdots \times perceptron_n$$

For image classification and object detection tasks, deeper networks are necessary, which requires using many hidden layers with different transformation functions. Convolutional neural networks and residual neural networks are state-of-the-art for image classification.

6.3.1. *Convolutional Neural Networks*

Convolutional neural networks (CNNs) learn higher-level image features and objects by filtering the raw pixel data of images (TensorFlow 2019). Unlike a simple artificial neural network, a CNN uses numerous self-optimising neurons (O'Shea and Nash 2015, p. 2). CNNs are preferred over fully connected neural networks for image classification and detection tasks. Fully connected neural networks would require too many parameters (or weights) to be useful for image data as the considerable number of parameters puts the model at risk for overfitting. Overfitting often occurs when the model learns the features in the training data too well that it is then unable to generalise the features for prediction on new data. The filtering in each layer creates an equivariant representation of the input for spatially aware edge, shape, and colour detection, from simple to complex shapes, usable anywhere in an image.

There are three primary components to CNNs – convolutional layers, pooling layers (sometimes replaced by strided convolutional layers), and fully connected layers (sometimes replaced by global average pooling layers) (TensorFlow 2019). In a convolutional neural network, the structure of a simple neural network is multiplied to form *n*+1 convolutional and pooling layers which operate on the assumption that the input is an image with learnable features. A simplified representation of the layers in a convolutional neural network is provided in

Figure 49; this demonstrates the filters on the original input in order to produce a feature map which will ultimately produce a classification output in the fully connected layer.
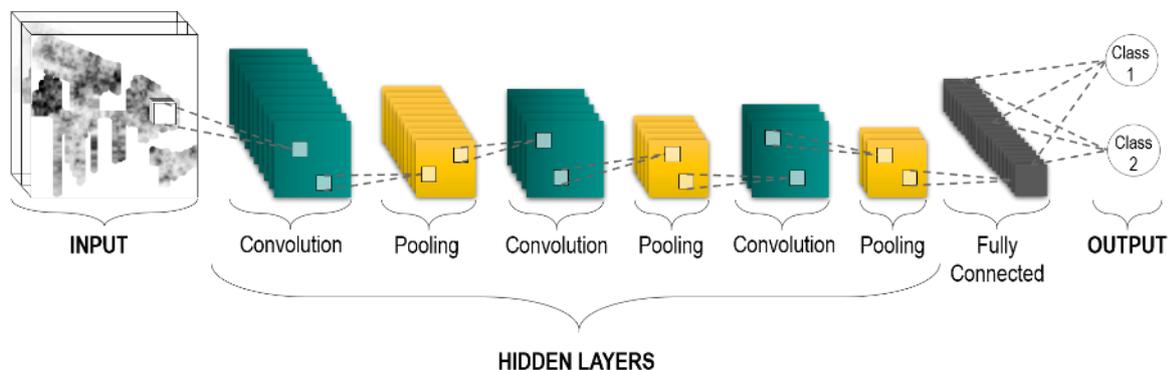
Figure 49: Simplified representation of a convolutional neural network

The features learned by the network are used to make class predictions. In deeper layers where high-level, class-wide features are learned, the learned features become more complex and activations are localised. Higher-level features are objects and shapes in an image. These are usually extracted following the detection of lower-level features in the first few convolutional layers. In contrast to higher-level features, lower-level features are minor parts of an image, such as edges, lines, and dots, that are detectable. These minor features usually comprise the higher-level features later learned by the network. The features learned by the network can be visualised using an activation, or feature, map.

Activation maps are useful for visualising the activations of a filter – which features or objects in an image each layer learns as they are visual representations of the edges, patterns, and objects learned from each filter (or kernel). Visualising the activation maps for a network can also help identify where some activations are zero for several inputs, indicating a dead filter and, potentially, a high learning rate (Karpathy 2019a). An example of activation maps for a small (17-layer) CNN that includes convolutional, rectified linear unit (ReLU) activation functions, pooling layers, and a fully connected layer is provided in Figure 50. In this example, the ReLU activation function facilitates the output of a node in the network, and the activation maps show higher-level features being learned in the deeper layers of the network to make class predictions. For deeper networks, gradient maps are often used to represent feature learning through histograms of the weights and biases. Examples of these gradient histograms from training for this project are shown in Figure 51.

Figure 50: Example of activations for each layer in a pre-trained ConvNetJS CNN with 17 layers and 7000 parameters (adapted from Karpathy 2019). No activation is shown in black. The activation maps prove the learned features become more localised in the deeper layers of the CNN.
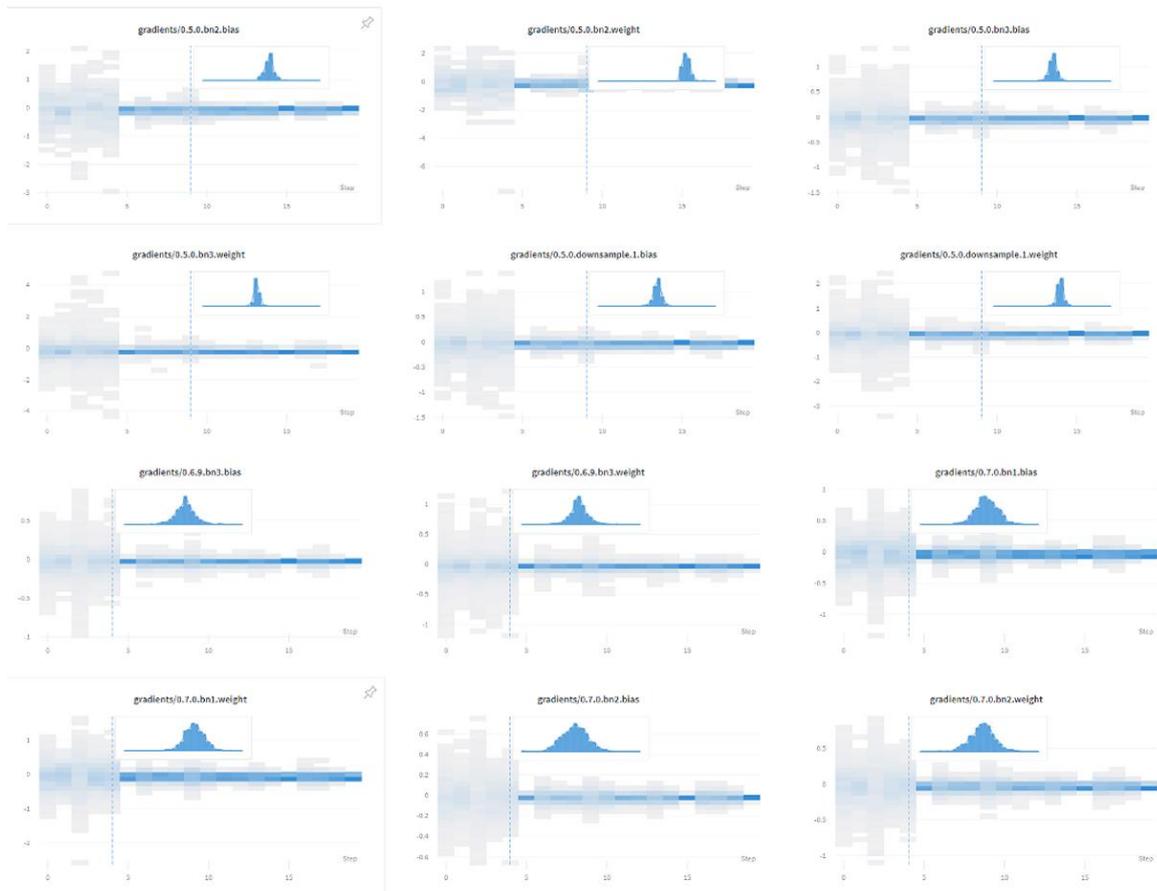
Figure 51: Gradient histograms of the weights and biases

Convolutional Layers

Convolutional layers use filters (or kernels) on images to extract subregions, where each layer creates a feature map comprised of individual values mathematically representative of each subregion. The network's architecture defines the size and number of filters. In the context of CNNs, filters (also called kernels) are matrices of a defined size which slide across an input (e.g. an image) and extract features. A 2D activation map of the response of each filter at any given spatial position is created by convolving each selected filter along the width and height of the input. This convolving acts as a sliding window across the input at a specified interval, called stride, as shown in Figure 52. The 2D activation maps for each filter in a layer are stacked in the depth dimension that compiles the output volume and aids in feature extraction.
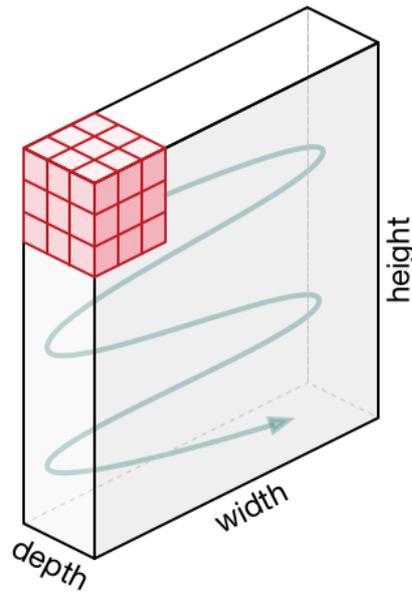
Figure 52: An example of the movement of the filter (shown in red) across the input using a sliding window path (highlighted by the curved line) in the depth, width and height aspects *(Saha 2018)*

Filters in CNNs usually extract edges, curves, shapes, and colours from images, especially in 2D convolutional (conv2D) layers, as number matrices that represent a known or target pattern. An example of an edge detection feature is shown in the Sobel edge detection filter in Figure 54 and Figure 55. Commonly, a 3x3x3 (see Figure 73) or 7x7x3 (see Figure 60) size filter is used in the first convolutional layer – where the filter is either 3 pixels wide and high, or 7 pixels wide and high, and three deep for each of the colour channels in an RGB image. Following this example, if using a converted greyscale image (an image with only one colour channel) as input on the first layer, the filter size would be 3x3x1 or 7x7x1. For subsequent layers, the third dimension in the filter size is equal to the number of slices along the depth dimension inherited from the output of the previous convolutional layer. If an image is comprised of greys, blacks, and whites but is not normalised to greyscale, it is an RGB image with pixel values on a scale of 0 to 255 but of the same intensity in the three channels. The training data included in this project were plotted using a white-black colour scale but were not converted into a single channel, greyscale image. Therefore, it is reasonable for the data to be treated as RGB images with three channels of equal colour intensity, particularly in transfer learning, where the original network was trained on RGB images. For clarity, a selection of pixel values from a sample training image are provided in Figure 53.
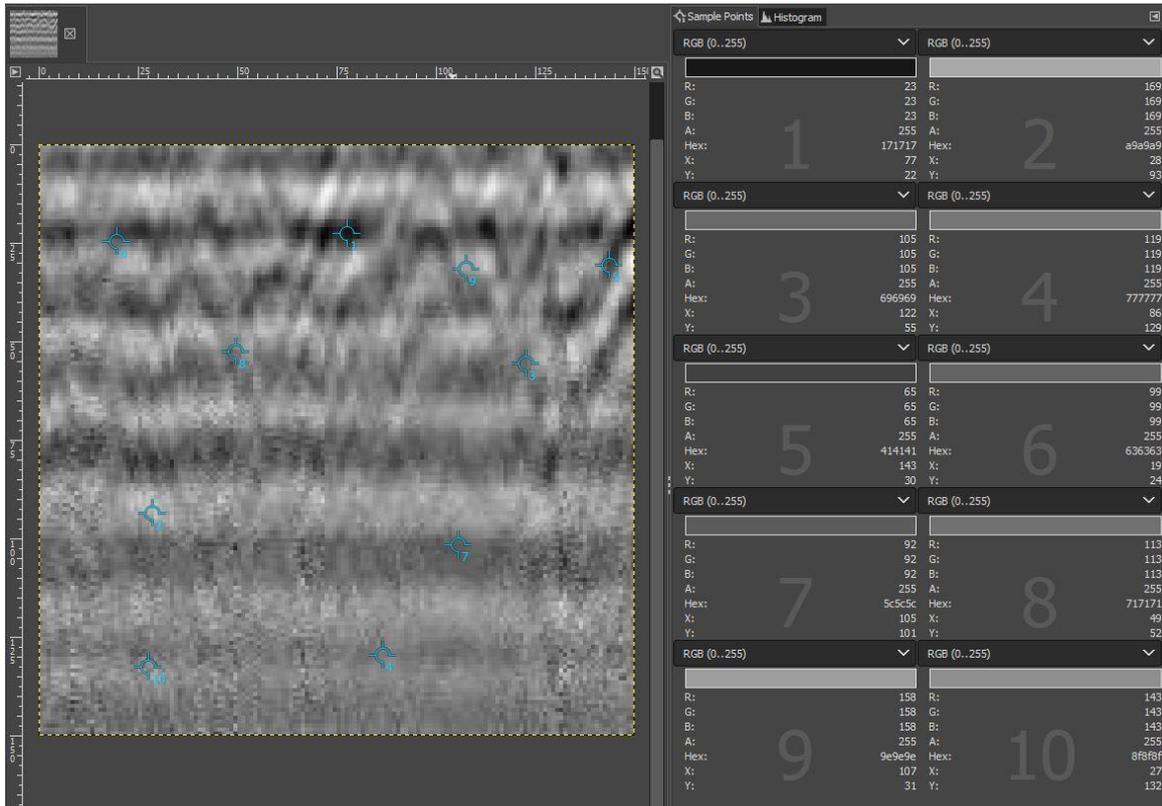
Figure 53: RGB values for select pixels in a sample training image
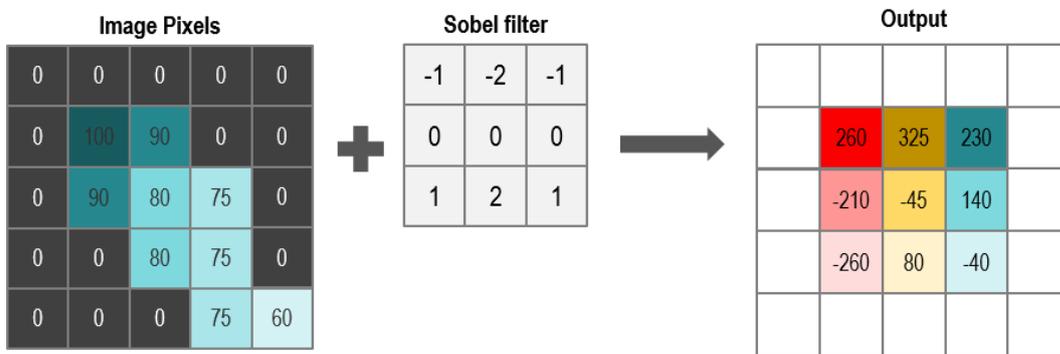
Figure 54: An example of the numerical representation of an edge detection filter and output in a convolutional neural network

These filters are applied to input images to identify a series of reoccurring shapes and patterns within a target or class. The network learns the patterns that comprise a class and apply them to new data to determine the likelihood that the new data contains patterns similar to that of known classes. Figure 55 shows a typical edge detection filter, the Sobel filter, and the pattern it produces from a randomly generated example of pixel values in an image.
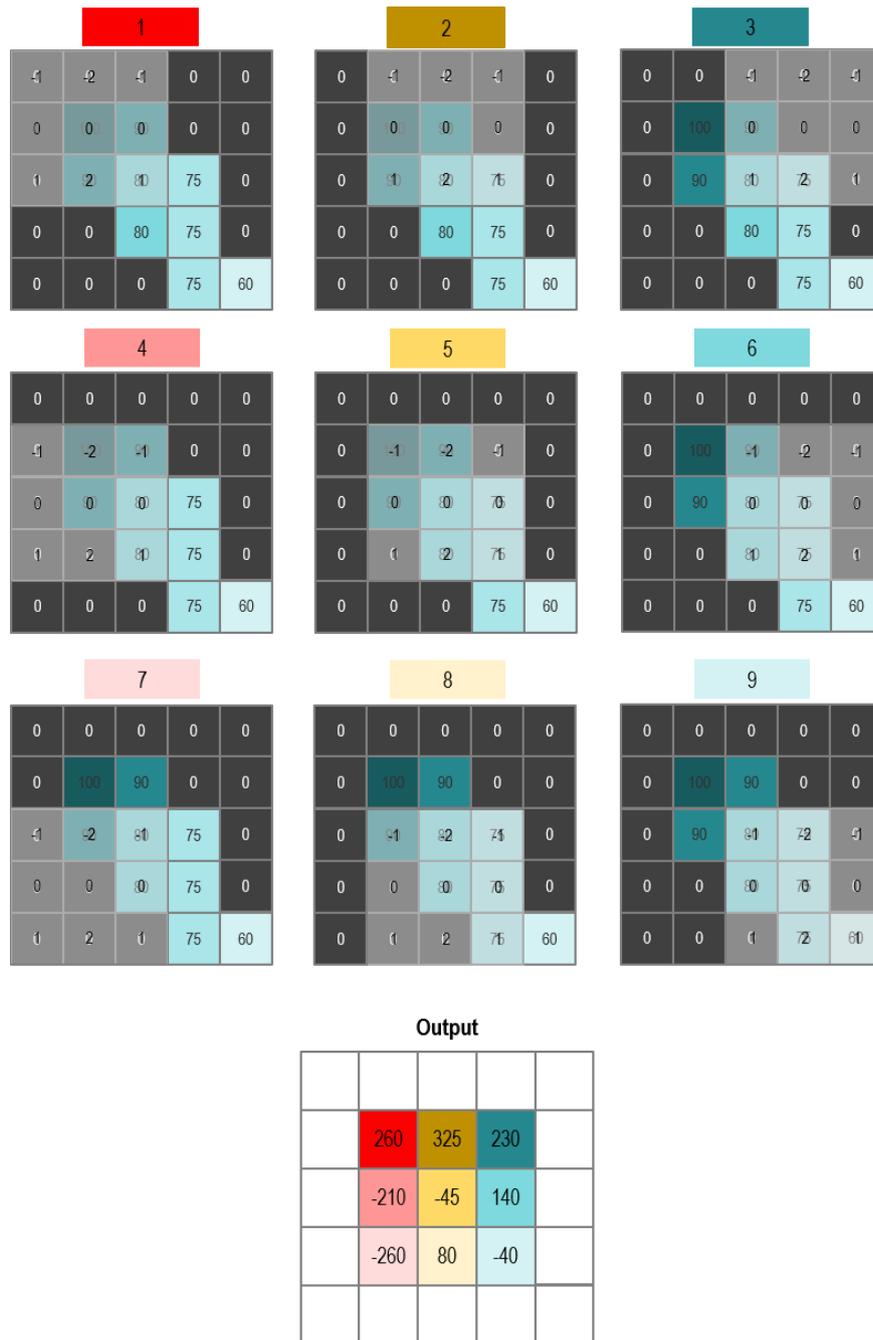
Figure 55: An example of the creation of the output from an edge detection filter. The 3x3 size filter, shown in grey, slides across the image. At each point, the value of the filter is multiplied by the value of the image. For each application of the 3x3 filter, the values are summed to produce a cell of the output. In this example, the colours of each stage (indicated above the matrix) correspond to the colour values in the output matrix.

An example of the activation maps for the first convolutional layer (with 16 5x5x3 filters) of a small CNN is shown in Figure 56, in which the general features (e.g. edges) of the focus object (deer) in the input image are outlined and discernible in the activation maps. In contrast, on the final convolutional layer (8x8x20 filter size and 20 filters) shown in Figure 57, the activation maps show higher-level localised features. The size and number of filters in each layer is defined in the network architecture. These can be altered if a network is built from scratch or an off-the-shelf network is not

suitable for a specific detection problem. The filters are often able to detect slight changes within an image that humans are not able to see. The output of the first convolutional layer (feature map) becomes the input in the second convolutional layer, and so on. Carrying the feature data through to the final, fully connected layer allows the network to learn more complex features in each subsequent convolutional layer as it builds upon what is already learned (Zeiler and Fergus 2014).
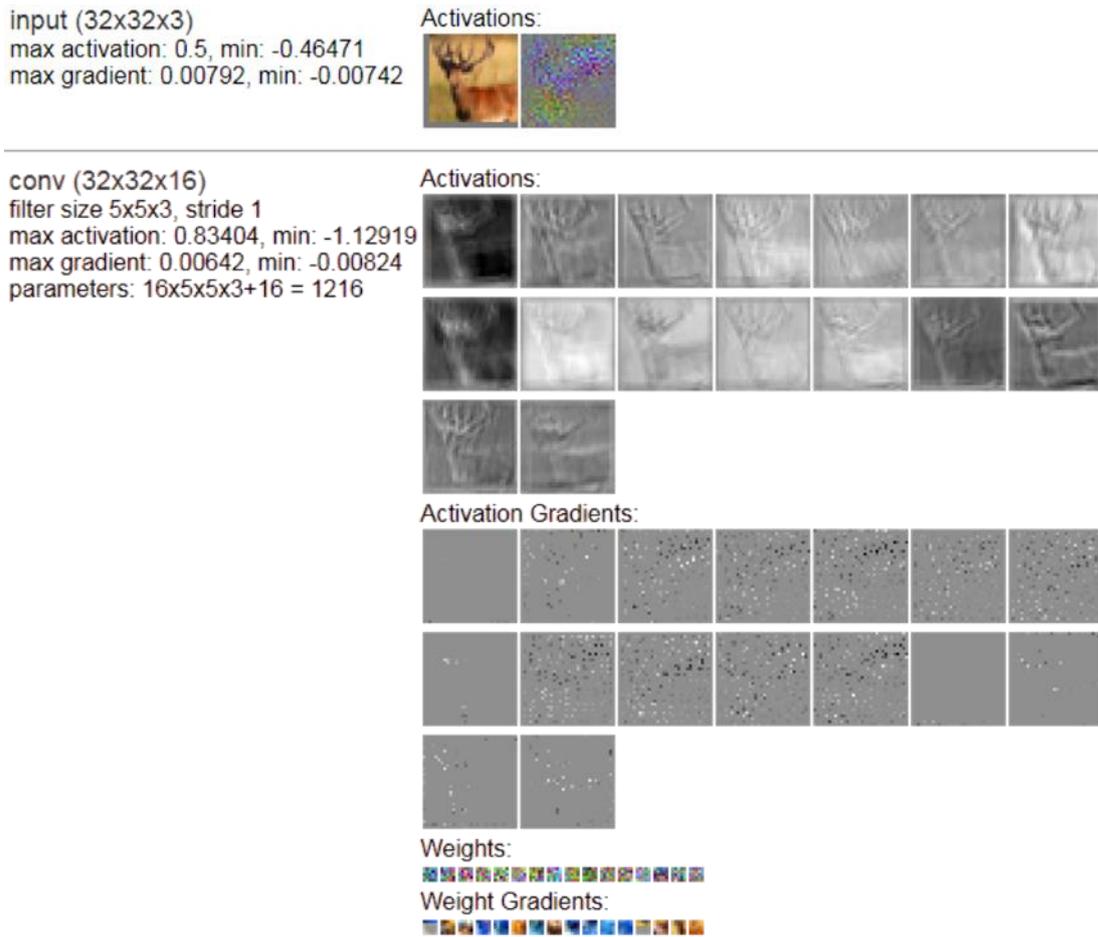
Figure 56: Activation maps for the input and first convolutional layer in a 17-layer ConvNetJS CNN. The general shape of the object (deer) in the input image is outlined in the activation maps. (Image is output from ConvNetJS CIFAR-10 demo)

input (32x32x3)
max activation: 0.5, min: -0.46471
max gradient: 0.00792, min: -0.00742

conv (8x8x20)
filter size 5x5x20, stride 1
max activation: 0.63027, min: -4.10268
max gradient: 0.15808, min: -0.11322
parameters: 20x5x5x20+20 = 10020

Figure 57: Activation maps for the input and final convolutional layer in a 17-layer ConvNetJS CNN. The general shape of the object (deer) in the input image is no longer discernible in the activation maps; instead, higher-level localised features are visible. (Image is output from ConvNetJS CIFAR-10 demo)

Pooling Layers

Pooling layers are often run between convolutional layers to reduce the spatial dimensions of the representation and computation in the network by downsampling. The depth slices generated during the preceding convolutional layer(s) are downsampled using filters with a stride of 2 (Karpathy 2019b). For example, a 2x2 filter with a stride of 2 will reduce a 32x32x16 input to 16x16x16, as shown in Figure 58.

relu (32x32x16)
max activation: 0.83404, min: 0
max gradient: 0.00879, min: -0.00824

Activations:

pool (16x16x16)
pooling size 2x2, stride 2
max activation: 0.83404, min: 0
max gradient: 0.00879, min: -0.00824

Activations:
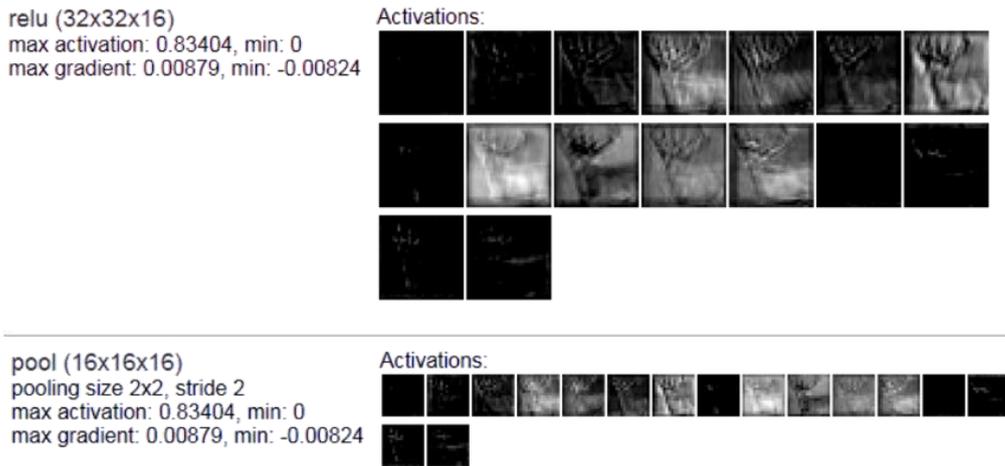
Figure 58: Example of the downsampling effect of a 2x2, stride = two pooling filter on 32x32x16 inputs. (Image is output from ConvNetJS CIFAR-10 demo)

Max pooling has recently been the favourable pooling function, but historically average pooling was favoured and is still used in some networks. With the input image as a numpy array, max pooling downsamples features by determining the maximum value for each subregion, determined by the size of the filter, of a depth slice. Whereas, average pooling downsamples features by determining the average value for each subregion of an input. In complex inputs, there is a noticeable difference in the outputs of max and average pooling filters, as shown in Figure 59. In some more recent architectures, pooling layers have been replaced with strided convolutional layers to produce the same downsampling effect.
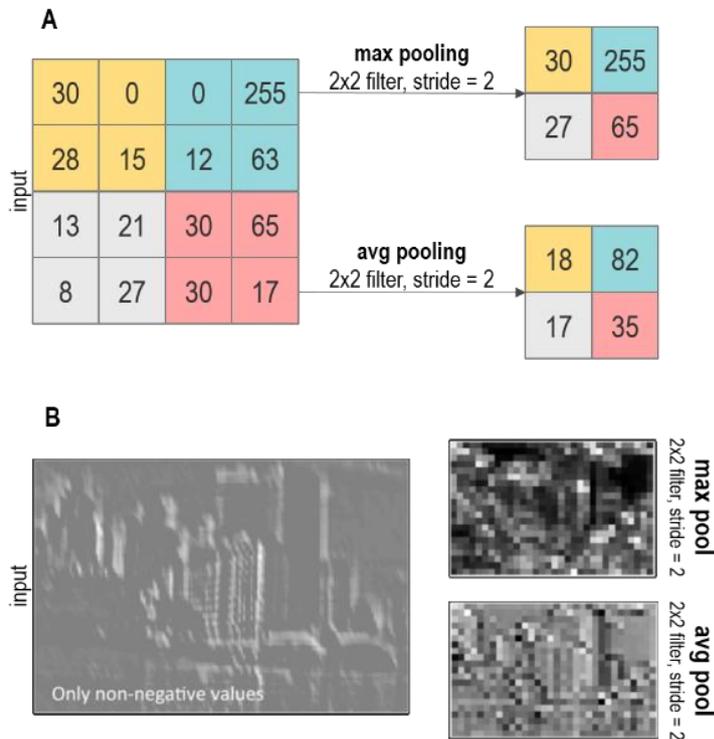
Figure 59: (A) Theoretical example of pooling filters with size 2x2 and stride 2; (B) Real example of pooling filters with size 2x2 and stride 2 (adapted from Fergus 2015, p. 39)

Fully Connected Layers

Once the feature maps have been produced, the final fully connected layer has an output of a vector with dimensions equal to the number of training classes, followed by a softmax layer which feeds to the prediction output to map the correlation of features in the image to features common to each class (O'Shea and Nash 2015). Lin et al. (2013, p. 4) note that using fully connected layers in this way uses the convolutional layers as feature extractors. However, as fully connected layers are at risk for overfitting, in recent convolutional neural networks fully connected layers are replaced with Global Average Pooling (GAP) layers (Lin et al. 2013, p. 4). GAP layers reduce the spatial dimension of feature maps as with other pooling layers discussed previously; however, rather than downsampling a map by 75%, GAP reduces a feature map of size $h \times w \times d$ to $1 \times 1 \times d$. In the ResNet models used here, a GAP layer is followed by a densely connected layer with a softmax activation function for predictions to make them better suitable for localisation tasks.

6.3.2. *Residual Neural Networks*

He et al. (2016, p. 1) identified that deeper CNNs (56-layer) had higher training and test errors than shallower CNNs (20-layer). This phenomenon is counterintuitive as deeper networks should inherently be more accurate but are more likely to overfit. Residual neural networks address the problem of training accuracy degradation in deep neural networks. He et al. (2016, p. 3) define the building block for residual networks as a function of the feed-forward connections and input vectors, as shown in Equation 6.2. Nonlinear 'identity shortcut connections' which skip layer(s) perform identity mapping and add outputs to the outputs of stacked layers, shown in Figure 60 (He et al. 2016, p. 2). He et al. (2016, p. 2) defined stacked nonlinear layers and original mapping, realised with the nonlinear shortcut connections, as part of the desired mapping, as shown in Equation 6.3. The connections do not add parameters or computational complexity, maintaining suitability for classification and localisation tasks.

6.2: Building block of residual neural network (He et al. 2016, p. 3)

$$y = \mathcal{F}(x, \{W_i\}) + x$$

$y = $ output vectors

$\mathcal{F}(x, \{W_i\}) = $ residual mapping to be learned

$x = $ input vectors

6.3: ResNet definition of a deep residual learning framework (He et al. 2016, p. 2)

$$\mathcal{F}(x) := \mathcal{H}(x) - x$$

$\mathcal{H}(x) =$ desired underlying mapping

$\mathcal{F}(x) + x =$ recast original mapping

ResNet (He et al. 2016) allowed deeper networks to achieve as high accuracy, if not better, as smaller networks. The ResNet networks were ground-breaking, increasing the size of useable CNNs to 152 layers and reducing susceptibility to overfitting; whereas, before ResNet, it was assumed that larger architectures would be too susceptible to overfitting.
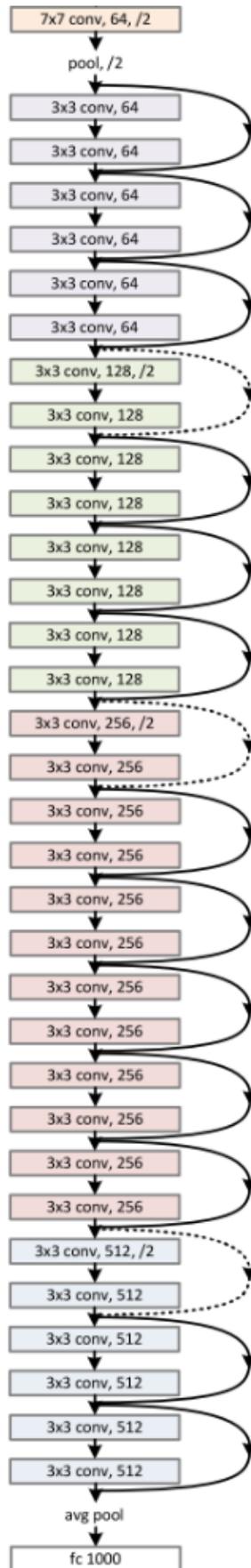
Figure 60: Example of residual CNN architecture where dashed lines indicate shortcut connections between blocks
(He et al. 2016, p. 4)

6.4.   Deep Transfer Learning

Deep transfer learning approaches were necessary for this project due to the size of the training dataset, difficulty in obtaining additional training data, and complexity of features. Deep learning algorithms can use supervised or unsupervised machine learning for multiple layers of successive output to input processing, such as a convolutional neural network (Schmidhuber 2014). Pan and Yang (2010, p. 1347) define transfer learning as:

> "Given a source domain $\mathcal{D}_S$ and learning task T_S, a target domain $\mathcal{D}_T$ and learning task T_T, transfer learning aims to help improve the learning of the target predictive function $f_T(\bullet)$ in $\mathcal{D}_T$ using the knowledge in $\mathcal{D}_S$ and T_S, where $\mathcal{D}_S \neq \mathcal{D}_T$, or $T_S \neq T_T$."

Simply, transfer learning is similar to inductive reasoning, in that it uses knowledge learned from processing one problem to solve another (Raina et al. 2006). Generally, the more training data available, the more re-training can be performed, from the final layer to fine-tuning the network. In most instances, where the dataset is limited, transfer learning is limited to retraining the last layers of an existing model on new data from classes that were not included in the original model. However, the whole network can be fine-tuned to improve the model's accuracy if a large enough dataset is available. Transfer learning reduces training, learning, and development time, leading it to be a widely applied method (Rawat and Wang 2017). Yosinski et al. (2014) have investigated how transferable features are for similar and different transfer learning tasks in deep neural networks. In this work, Yosinski et al. (2014, pp. 1–2) note that the first layers in a network are not specific to a dataset, but the last layers are highly specific to the dataset it was trained on. While transfer learning is still a widely accepted approach, Yosinski et al. (2014, p. 1) defined three issues impacting on the transferability of features:

> "1. the specialization of higher layer neurons to their original task at the expense of performance on the target task
>
> 2. optimization difficulties related to splitting networks between co-adapted neurons
>
> 3. the distance between the base task and target task."

As the final layers of a pre-trained network were retrained in this project, the model has learned those high-level features specific to the new task (detection of graves).

Many models are pre-trained on extensive collections (upwards of 60,000) of images. The standard image datasets are ImageNet, CIFAR-10, and CIFAR-100. ImageNet contains 14,197,122 images (as of April 2010) labelled according to the WordNet database of cognitive synonyms. The CIFAR-10 dataset contains 60,000 images divided evenly into ten classes (airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck). The CIFAR-100 dataset contains 60,000 images divided evenly into 100 classes of aquatic mammals, fish, flowers, food containers, fruit and vegetables, household

electrical devices, household furniture, insects, large carnivores, large man-made outdoor things, large natural outdoor scenes, large omnivores and herbivores, medium-sized mammals, non-insect invertebrates, people, reptiles, small mammals, trees, and vehicles.

If using a pre-trained network., this model is called in the training script as the base network architecture, as shown in Figure 61. Information, such as weights, is stored in a model file defined in the training script. These files will differ between the deep learning libraries used to define the training script. For example, in this project which primarily uses the Fastai library, the model parameters but not the architecture are saved to .pth files when learn.save is called. The model saved from learn.save is loaded using learn.load and layers are unfrozen for training using learn.unfreeze to retrain layers before the final layer in a pre-trained model. This process can be repeated or unfreeze redefined to unfreeze the entire model to retrain all the layers in the network if the training dataset allows for this. Once optimum training is achieved, the model (or learner) is exported using learn.export and loaded back using learn.load for inferences. Figure 61 and Figure 62 show excerpts of the working object detection script for this project which demonstrates the training process and how model information is saved and loaded. The complete training script is hosted on GitHub.

```
In [11]: learn = cnn_learner(data, models.resnet152, custom_head=MyHead(), metrics=[l1_loss, cross_entropy, accuracy])
         learn.loss_func = MyLoss()

In [12]: learn.lr_find(end_lr=10)

         LR Finder is complete, type {learner_name}.recorder.plot() to see the graph.

In [13]: learn.recorder.plot(skip_end=10)
```



```
In [14]: learn.fit_one_cycle(10, max_lr=4e-4)
```

epoch	train_loss	valid_loss	l1_loss	cross_entropy	accuracy	time
0	0.998631	0.754584	0.185334	0.383916	0.853896	03:27
1	0.894142	0.810546	0.201980	0.406586	0.857143	03:20
2	0.843934	0.665906	0.214866	0.236174	0.905844	03:22
3	0.769735	0.547540	0.156065	0.235410	0.905844	03:21
4	0.675532	0.468125	0.126036	0.216052	0.928571	03:20
5	0.603766	0.446525	0.117942	0.210640	0.915584	03:21
6	0.560311	0.423995	0.112504	0.198988	0.925325	03:20
7	0.509234	0.401160	0.109920	0.181320	0.935065	03:20
8	0.483372	0.384849	0.105122	0.174604	0.938312	03:21
9	0.459534	0.403714	0.108039	0.187637	0.931818	03:21

Figure 61: Excerpt of training script demonstrating how a pre-trained network is loaded for retraining and retrained

```
In [23]:  learn.save('stage-1152')

In [24]:  learn.load('stage-1152')

          Train: LabelList (1236 items)
          x: ObjectItemList
          Image (3, 150, 150),Image (3, 150, 150),Image (3, 150, 150),Image (3, 150, 150),Image (3, 150, 150)
          y: ObjectCategoryList
          ImageBBox (150, 150),ImageBBox (150, 150),ImageBBox (150, 150),ImageBBox (150, 150),ImageBBox (150, 150)
          Path: C:\Users\Student\Desktop\moderngraves_july\Downloaded;

          Valid: LabelList (308 items)
          x: ObjectItemList
          Image (3, 150, 150),Image (3, 150, 150),Image (3, 150, 150),Image (3, 150, 150),Image (3, 150, 150)
          y: ObjectCategoryList
          ImageBBox (150, 150),ImageBBox (150, 150),ImageBBox (150, 150),ImageBBox (150, 150),ImageBBox (150, 150)
          Path: C:\Users\Student\Desktop\moderngraves_july\Downloaded;

          Test: None, model=Sequential(
            (0): Sequential(
              (0): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2), padding=(3, 3), bias=False)
              (1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
              (2): ReLU(inplace)

In [25]:  learn.unfreeze()
```

Figure 62: Excerpt of training script demonstrating how models are saved and layers are unfrozen for retraining

6.5. Training Datasets

As is common practice within machine learning, training data comprise an 80% subset of the entire dataset; the remaining 20% comprises the testing dataset. Training datasets are labelled data. In this instance, a collection of images containing known features used to train classification algorithms to identify objects or features, improve the accuracy of the model, and create classifiers. Accuracy of the model refers to the number of correct predictions made divided by the total number of predictions made. Images containing any object imaginable can be used to train a model to detect specified objects or features within an image. Caches of labelled images are available in repositories online, or there is the option for the user to create and label their dataset. The accuracy of a model is mostly dependent on the size and quality of the training dataset. Generally, the larger a training dataset is the higher potential accuracy the model can achieve. The required number of instances (e.g. images) in a training dataset is dependent on the training algorithm and learning method used, it is most important, however, to ensure the model size and representation of features in the dataset are suitable to decrease the risk of overfitting the model (He et al. 2016). In this project, the training dataset was labelled according to two classes – grave and non-grave/background. The total dataset contained 537 "positive" examples of graves and 680 "negative" examples of responses that are similar to graves (false positives), noise, and general background responses. These images were fed into an artificial neural network to train it to identify similar patterns within the labelled data.

6.6. Use of Machine Learning on Non-Archaeological GPR Data

Aside from their applications to archaeogeophysical datasets, machine learning and neural networks have been widely investigated for rebar and IED detection in GPR data, and less frequently for hyperbola detection and fitting in GPR data and deposit detection in seismic and GPR data.

120

Asadi et al. (2019), Kaur et al. (2016), and Dinh et al. (2018) present examples of the automatic detection of rebar reflections in GPR data. Asadi et al. (2019, pp. 2619–2620) used histogram of oriented gradients (HOG) feature descriptors and AdaBoost, a supervised binary classifier, to detect hyperbolic reflections from rebar in bridge decks using radargram data. Kaur et al. (2016) also identified an automated method for the detection of rebar in bridge decks using radargram data. This work, however, used Support Vector Machines and RANSAC Hyperbola Fitting for pattern recognition (Kaur et al. 2016, p. 2269). Dinh et al. (2018) achieved a classification accuracy of 99.6%±0.85% using a 14-layer convolutional neural network for the detection of hyperbolic reflections from rebar in radargrams. Ishitsuka et al. (2018, p. 1) used deep convolutional neural networks for the detection of shallow geological features and other embedded objects with a maximum classification accuracy of 97.9%.

Lameri et al. (2017) and Núñez-Nieto et al. (2014) are recent examples of work carried out on the automatic detection of IEDs and unexploded ordnance in GPR data. Both papers present an evaluation of small (less than ten layers) neural networks, where Núñez-Nieto et al. (2014, p. 9737) have designed their 3-layer network and Lameri et al. (2017, p. 510) based their network on the 5-layer LeNet (LeCun et al. 1998). Lameri et al. (2017, p. 512) achieved a maximum classification accuracy of 95%; whereas, Núñez-Nieto et al. (2014, p. 9745) achieved a classification accuracy of 89% for a 1GHz GPR antenna and 92% for a 2.3GHz antenna.

Similar work has also been carried out on seismic data, which is comparable to GPR radargrams. Milosavljevic (2020) has used a more sophisticated approach with a self-designed CNN based on U-Net, ResNet, and DenseNet for semantic segmentation. Using this network through a Kaggle competition, classification on the seismic dataset achieved a maximum accuracy of 85% (Milosavljevic 2020, p. 10). Other supervised methods, such as Extremely Random Trees Ensemble, have been applied to the detection of deposits in seismic data (Guillen et al. 2015).

The above work carried out on the automatic detection of reflections in radargrams influenced the use of radargrams for grave detection in this project.

6.7. Use of Computational Approaches and Machine Learning on Archaeological Data

Within the last ten years, the integration of automatic classification algorithms has become increasingly popular for archaeological applications, particularly for analysing and interpreting remote sensing data (De Boer 2007, De Laet et al. 2007, Trier et al. 2009, Freeland et al. 2016, Traviglia et al. 2016, Cerrillo-Cuenca 2017). There have been several focus groups on the issues surrounding applying such machine learning techniques to complex datasets and problems seen in archaeology

(Traviglia et al. 2016). These focus groups culminated in a workshop on integrating citizen science in the automation workflow for archaeological targets in order to improve the validation of classifications and accuracy of learning machines (Lambers et al. 2019).

While still relatively unexplored, since 2006 (van der Maaten et al.) the implementation of machine learning in archaeology has expanded. There has been a marked increase in the number and variety of projects in the last five years. Machine learning-based approaches have been used in both practical field applications (e.g. remote sensing and geophysics data interpretation) and post-excavation analyses (e.g. finds identification and typology development).

In remote sensing, there have been several machine learning approaches to detecting archaeological objects. Most recently, Verschoof-van der Vaart and Lambers (2019) used regions-based convolutional neural networks (R-CNNs) to detect prehistoric barrows and Celtic fields in LiDAR data from the Netherlands. To detect these archaeological features, the workflow (WODAN) implements pre-processing, object detection, and post-processing steps to input and output data (Verschoof-van der Vaart and Lambers 2019). The Faster R-CNN, using ResNet50 (He et al. 2016), had an average precision value, *(True Positives/(True Positives + False Positives)),* of 0.64 for barrows and 0.46 for Celtic fields. The precision values and F_1-scores demonstrate that the Faster R-CNN model is better at detecting barrows than Celtic fields (Verschoof-van der Vaart and Lambers 2019). Verschoof-van der Vaart and Lambers' research demonstrates that an R-CNN is amongst the top-performing approaches to multi-class object detection. However, they do note that there is still a need to improve the accessibility of such tools to implement them in broader archaeological uses.

Similar to the R-CNN approach used by Verschoof-van der Vaart and Lambers, Kramer is using deep learning to detect round barrows in LiDAR data from the UK (Kramer 2015, Kramer et al. 2017). However, further complexity is added to this work because there are many levelled barrows in the dataset that could not be detected by LiDAR and exponentially increase the number of false negatives in the dataset. Other approaches, like that taken by Trier et al. (2016, 2018) and Schneider et al. (2015) have experienced performance issues in their models due to the complexity of their targets (charcoal kilns) and the surrounding terrain, as well as the small training datasets which are limited by the number of available examples.

Post-excavation applications of machine learning have ranged from text mining in archaeological reports (Richards et al. 2015) to developing typologies and identifying finds (Boon et al. 2009, Gansell et al. 2014, Hörr et al. 2014).

While the majority of machine learning applications in archaeology have been on remote sensing data, there have also been a few examples of machine learning on geophysical datasets (Moysey et al. 2006, Leckebusch et al. 2008, Zhao et al. 2015, Verdonck 2016, Verdonck et al. 2017). These examples of feature extraction and machine learning on geophysical datasets are fairly undemanding. There have been few other attempts to automatically classify GPR data, none of which have taken a deep learning approach until recently with the work of Küçükdemirci and Sarris (2020).

Ernenwein (2009) implemented a pixel-based segmentation approach to multidimensional data using GPR, magnetometry, and magnetic susceptibility data. Supervised and unsupervised classification were both tested. The unsupervised classification using K-means cluster analysis was able to define positive anomalies which were already easily identified in human-led interpretation (Ernenwein 2009, p. 147). However, the supervised approach using Mahalanobis classification for four classes (Burned, Enhanced, Non-magnetic, Background) achieved the best results with 95% accuracy (Ernenwein 2009, p. 153). In Mahalanobis classification, each pixel of an image is assigned to the class based on its typicality. Often, typicalities are "derived from Mahalanobis distance measurements with reference to an F distribution or a chi-squared approximation" (Ernenwein 2009, p. 153). Based on Ernenwein's (2009, p. 157) promising results using supervised classification, they advocate for further use of the method when interpreting multi-dimensional, multi-technique datasets, with a final human-led interpretation for confirmation of the automatic classification.

Linford et al. (2018) approach the problem of interpreting large, high-density GPR datasets using semi-automated vector identification. Specifically, they assessed the viability of the vector identification approach to locate pit-like anomalies by creating polygons from binary objects in images with size and orientation attributes that form a template for the target features. Linford et al. (2018) note that the interpretation and location of pit-like anomalies can be improved by characterising the morphology of the anomaly.

Verdonck (2016) uses a template matching approach for the extraction and detection of Roman villa walls in 2D and 3D GPR data. Verdonck has based his methodology on that which is established for remote sensing datasets. A 2D normalised cross-correlation (NCC) was used as a measure of the similarity between the template and original image (Verdonck 2016, p. 263). The original image was rotated around the template, rather than the template sliding across the original image, to identify the highest NCC (Verdonck 2016, p. 263). For the 3D data, the 2D rectangles from the template matching stage were extruded to the depth of the data cube to create a 3D volume. The shapes were filled with an isosurface created from the associated GPR data and plotted using the

coordinates of each subvolume with the complete 3D volume to plot the 3D features (Verdonck 2016, p. 265). While this approach was successful (*c.* 77% detection and correctness rate for 2D data and *c.* 71% detection and correctness rate for 3D data), the linearity and uniformity of the target features are likely the reason for the high success rate (Verdonck 2016, p. 257,269). With other archaeological features it is unlikely a template matching approach would achieve a similar success rate, due to their variability in morphology and preservation state. Although Verdonck's examples of feature extraction on Roman villa walls uses a simple, well-defined target, it is important to note that it is also an example of successful application of machine learning methods on a GPR dataset and hints to promising applications in the future. Verdonck, very importantly, notes that the interpretation of geophysical data is still heavily reliant on human verification and intervention due to the complexities of the data and target features, and further assessment of the success of automatic detection on more complex GPR datasets is needed (Verdonck 2016, p. 269).

While these applications of machine learning to the interpretation of archaeogeophysical data all achieved a degree of success, none of them achieved 100% accuracy. Verdonck et al. (2019) warn that these tools are still not a stand-alone method of data interpretation; instead, they should be used as tools to aid human-led data interpretation. Particularly, caution should be exercised when using the interpretation from only one data type or geophysical technique.

Küçükdemirci and Sarris (2020) used U-Net convolutional neural networks trained from scratch for semantic segmentation of archaeological features in RGB images of GPR timeslices. As in the examples from Verdonck (2016) and Verdonck et al. (2019), the dataset used by Küçükdemirci and Sarris (2020, p. 4) contains only clearly defined linear and circular features. Küçükdemirci and Sarris (2020, p. 6) used the Sørensen–Dice coefficient as a measure of the segmentation accuracy, where 1 indicates the prediction and label are perfectly matched and 0 indicates there is no match between the label and prediction. The U-Net approach achieved a maximum Sørensen–Dice coefficient of 0.92 (Küçükdemirci and Sarris 2020, p. 6). While the segmentation accuracy achieved in this study is promising, the approach taken does not account for the depth of features or reoccurrence of features in multiple timeslices. On 2D geophysical data, such as magnetometry, there would be no issue using segmentation approaches on plan-view images without a depth aspect; however, the interpretation of linear and circular features in single slices is problematic. Without accounting to the depth of features while training a network, there is no distinction made between modern features (e.g. services, excavation trenches, septic systems) and similar archaeological features by the machine. For this reason, radargrams were primarily used for the work discussed herein, except in the case of clearly delineated rectangular modern graves.

7. MACHINE LEARNING METHODOLOGY

7.1. Introduction

The primary tasks in creating a package which will have the ability to classify and detect grave-like objects within a GPR dataset are to adapt existing models to the data format, create a training dataset for the new model, and conduct geophysical surveys of sites to be used to test the model in "real-world" scenarios when complete. While models exist for classification, feature extraction, and object detection, they are not particularly suited to images of GPR data, and there are no labelled datasets of geophysical data available. Therefore, a suitable training dataset was collated and labelled for retraining existing models. Once the model achieved the maximum accuracy with the available data for the training dataset, it was tested on "known" and "unknown" datasets. The trained classification and detection models were compiled, packaged, and made available for public use as a web app.

To compare the accuracy and consistency of human-led interpretation and machine-led interpretation, a survey which aimed to assess archaeologists' ability to interpret graves in GPR data was circulated to trained geophysicists. The brief survey was hosted on SurveyMonkey and was therefore restricted in the number and format of questions that could be asked. The survey asked the participant to identify whether a grave, multiple graves, or no graves were present in two radargram samples and two timeslice examples. Participants were also asked to indicate their confidence in their response from low-high confidence. The survey responses were anonymised, but information on the participants' experience with GPR data was collected. The survey questions are included in Appendix D and the results from responses to these questions are discussed in further detail in Chapter 1.2.

7.2. Training Dataset

As there was no pre-existing dataset of suitable GPR data, a training dataset was created and labelled. An appropriately sized dataset was required to achieve high accuracy using a retrained model and reduce overfitting in the model. To obtain enough images without augmenting the data, ground-truthed GPR data retrieved from Bournemouth University surveys and grey literature reports, as well as simulated data form the training dataset. For the data retrieved from grey literature reports, the geophysical survey data were compared with excavation results to isolate true positive and true negative grave responses. By definition, the model is not able to identify false positive or false negative results. However, the outputs will feed into the geophysical investigations and any

subsequent excavations to identify true positive, false positive, true negative, and false negative responses.

Approximately 1000 images were pre-processed to be the same size and colour scale. Data produced directly by the author were bandpass filtered and gain adjusted using GPR-Slice. Some data were also migrated and background filtered depending on the quality of the survey data. The wide spectrum of processing techniques were included in the training dataset to mimic the processing methods of end-users of various skill levels. The data produced from GPR-Slice and GPRSIM (see Section 7.2.1) were plotted to a custom paper size to improve image quality while maintaining the proportions of each data grid. Future users of these models are provided with the plot size and scaling information in the user guide in Appendix F. Where possible, data not produced in GPR-Slice or GPRSIM were corrected to match the pixel size of the GPR-Slice and GPRSIM data. It is accepted that discrepancies in pixel size may account for a portion of the classification errors discussed in Sections 0 and 7.4. Images fed to the trained model for inference are rescaled proportionally on input and export; therefore, the size of the input image is irrelevant if it is larger than 150x150 pixels (the size of the training images). Further information on inferring with the trained model is provided in Section 7.5.

Additional training data was derived from simulated GPR data using gprMax (Warren et al. 2016) and GPRSIM in order to reduce overfitting of the model by increasing the number of images in the training dataset. These radargrams were simulated based on the morphology, known conductivities, and relative dielectric permittivity (RDP) of grave-related materials derived the software and a literature search. The pre-processed images were then cropped into random 150x150 pixel tiles and labelled as having a response from a grave in the image or not. Within the model, the input data were augmented by horizontally flipping the images to double the size of the dataset and simulate data collected in multiple traverse directions. Examples of the images included in the training and testing datasets are presented in Figure 63. The entire training dataset is included in Appendix G.
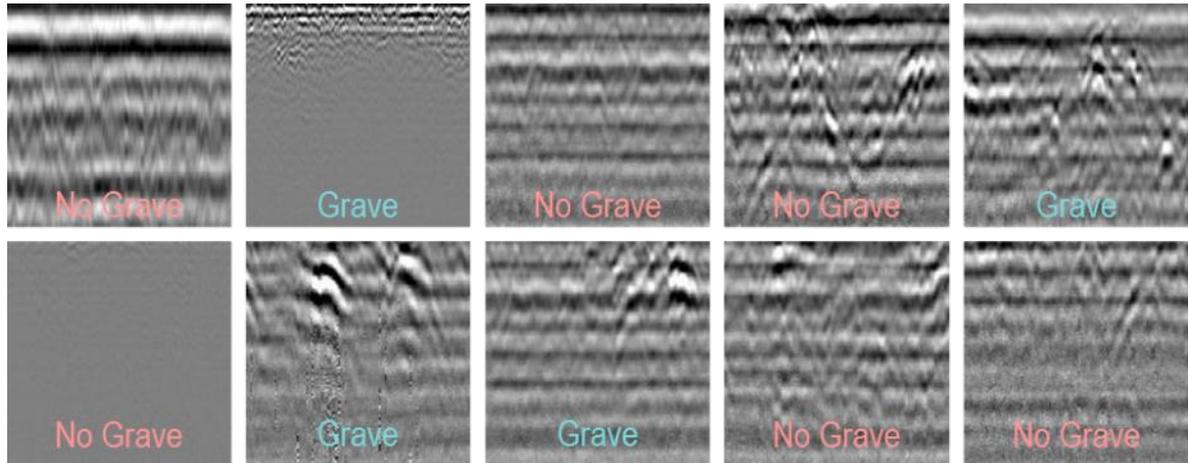
Figure 63: Examples of the 150x150 pixel tiles included in the training and testing datasets, with four images containing true positives (labelled 'Grave') and the remaining images containing false positives or true negatives (labelled 'No Grave')

7.2.1. *Simulated Data*

Complex simulations of true positive and true negative examples were created using GPRSIM. The following series of figures (Figure 64 - Figure 66) highlight the simulation process and focus on the parameters which can be altered within the software and the format of the output.

Models are drawn using polylines and polygons. A background grid is provided within the user interface to ensure that models are drawn to scale. Complex stratigraphic layers and archaeological features are drawn quickly; however, it should be noted that any model drawn will always be a simpler representation of the real-world scenario.

Polygons are assigned a parent material, the characteristics of which are set by the user. These materials can either be experimental or derived from real values. The material characteristics used to create the simulated data in this project were derived from known values. The relative dielectric permittivity, conductivity values, and attenuation constant for materials commonly associated with grave fills and local geologies are detailed in Table 22. These values were used to create new materials in the GPRSIM library to be included in the simulation models.
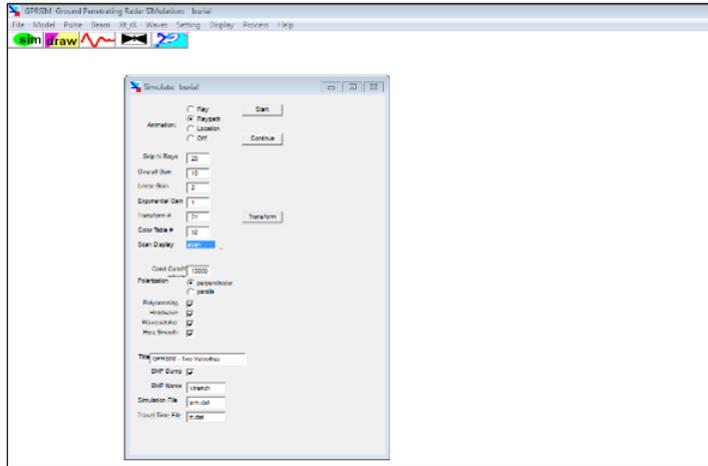
A. GPRSIM simulation visualisation menu which is used to set the parameters and output that will be shown in the final model visualisation.
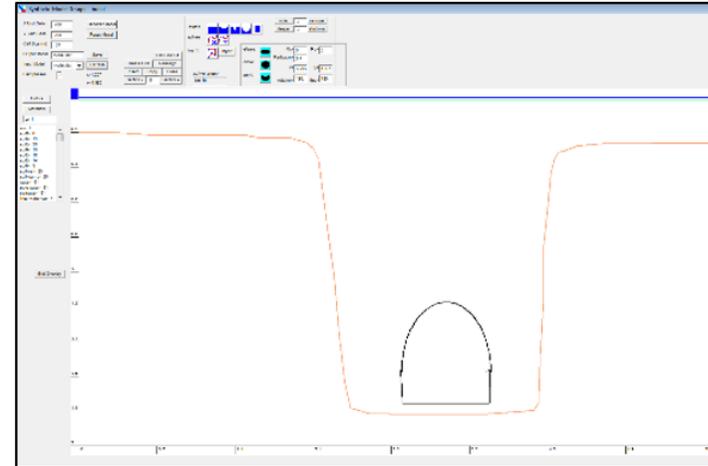
B. GPRSIM simulation drawing tool used to draw to-scale models.
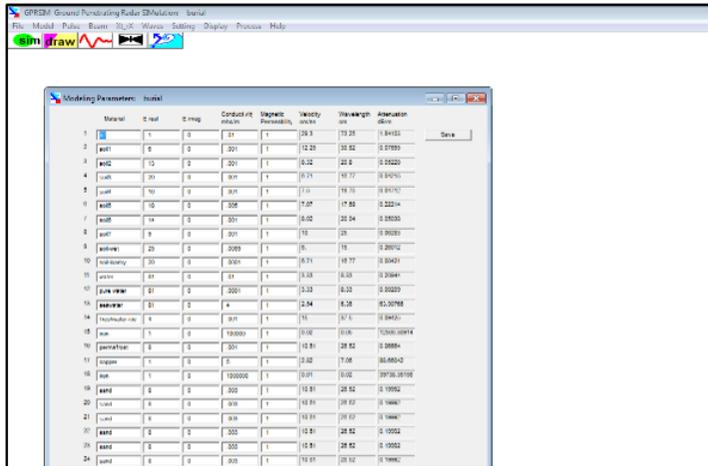
C. GPRSIM materials characterisation menu where existing materials can be chosen for the drawn model or new materials can be characterised and added to the library.
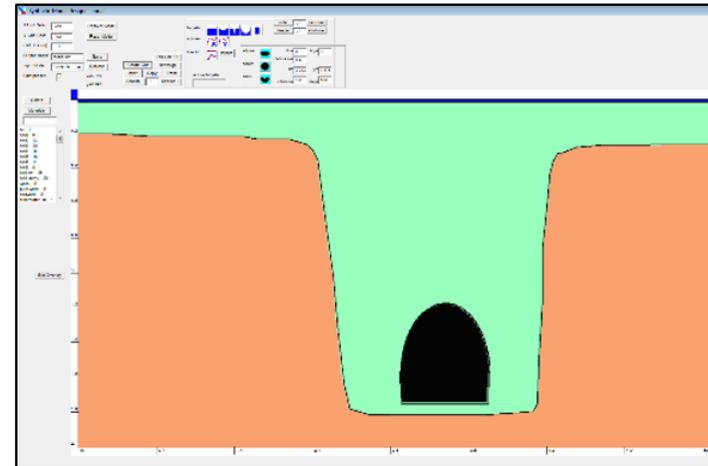
D. GPRSIM materials assignment and model gridding tool where each stratigraphic layer and features/objects are assigned a material.

Figure 64: GPRSIM simulation visualisation menus which are used to set the parameters and model that will be shown in the final model visualisation

Table 22: Known relative dielectric permittivity, conductivity values, and attenuation constant for common materials (adapted from Davis and Annan 1989; Daniels 1996; Olhoeft 1989; Keller 1989; Daniels 2004; Cassidy 2009; Martinez and Byrnes 2001)

Material	ε_r	Conductivity (S/m)	Attenuation Constant (dB/m)
Air	1	0	0
Distilled water	80		
Freshwater	80	10^{-6}-10^{-2}	0.01
Saltwater	80		
Freshwater ice	3-4		
Saltwater ice	4-8		
Snow	8-12		
Permafrost	4-8		
Dry sand	3-6		
Wet sand	10-30		
Dry sandstone	2-3		
Wet sandstone	5-10		
Limestone	4-8		
Dry limestone	7		
Wet limestone	8		
Shale	5-15		
Wet shale	6-9		
Silts	5-30		
Clays	5-40		
Dry clay	2-6	10^{-3}-10^{-1}	10-50
Wet clay	15-40	10^{-1}-10^{-0}	20-100
Dry sandy soil	4-6	10^{-7}-10^{-3}	0.01-1
Wet sandy soil	15-30	10^{-3}-10^{-2}	0.5-5
Dry loamy soil	4-6		
Wet loamy soil	10-20		
Dry clayey soil	4-6		
Wet clayey soil	10-15		
Dry coal	3.5		
Wet coal	8		
Granite	4-6		
Dry granite	5		
Wet granite	7		
Dry salt	4-7		
Calcite	6.4-8.5		
Gypsum	6.5		
Halite	5.9		
Kaolinite	11.8		
Mica	6.4		
Olivine	7.2		
Quartz	4.5		
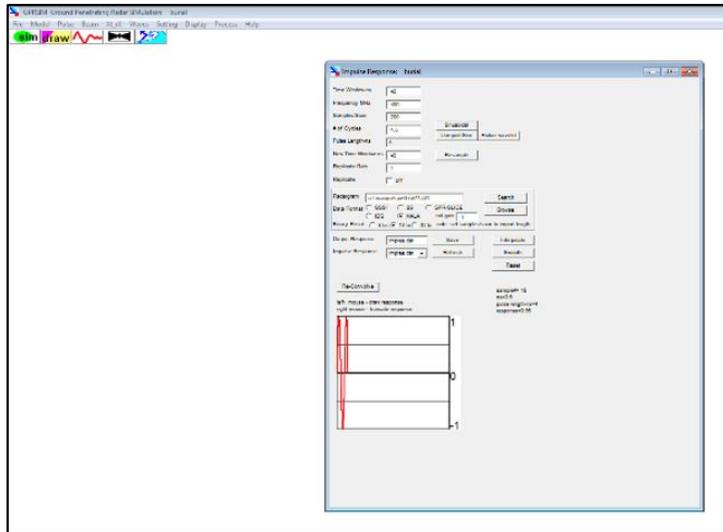
A. GPRSIM directional response menu
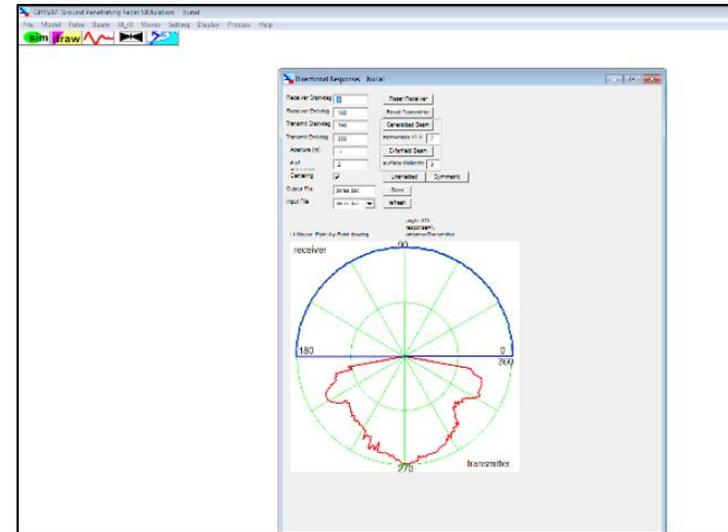
B. GPRSIM directional response menu
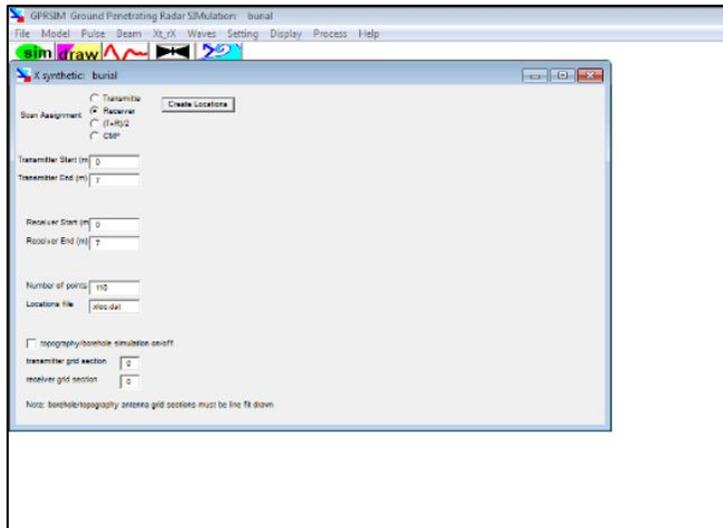
C. GPRSIM transmitter and receiver antenna positions menu
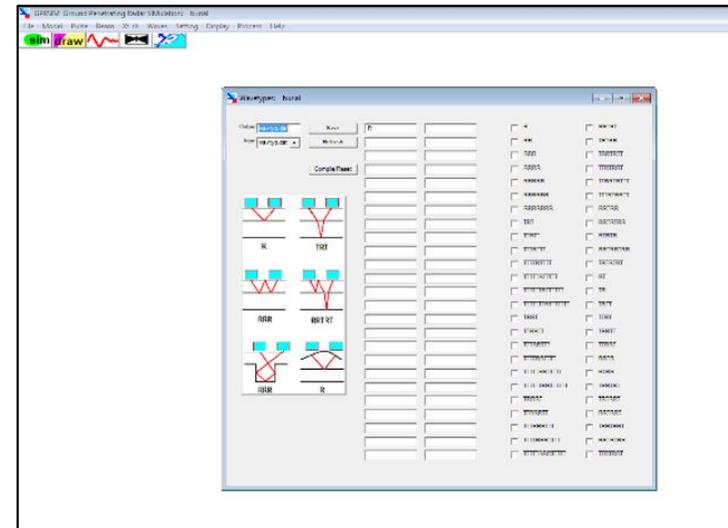
D. GPRSIM wavetype settings menu

Figure 65: GPRSIM GPR settings menus to set the radar parameters
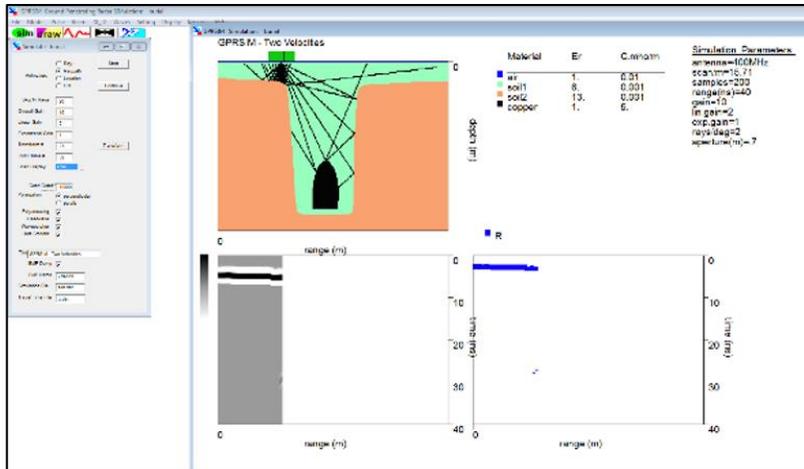
A. GPRSIM simulation showing the raypath and radargram as the model progresses.
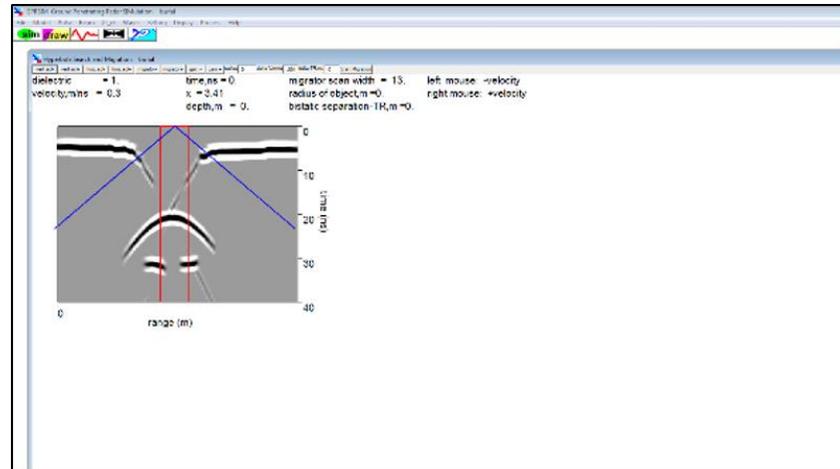
B. GPRSIM processing menu for radargrams where data is migrated, regained, and scaled.
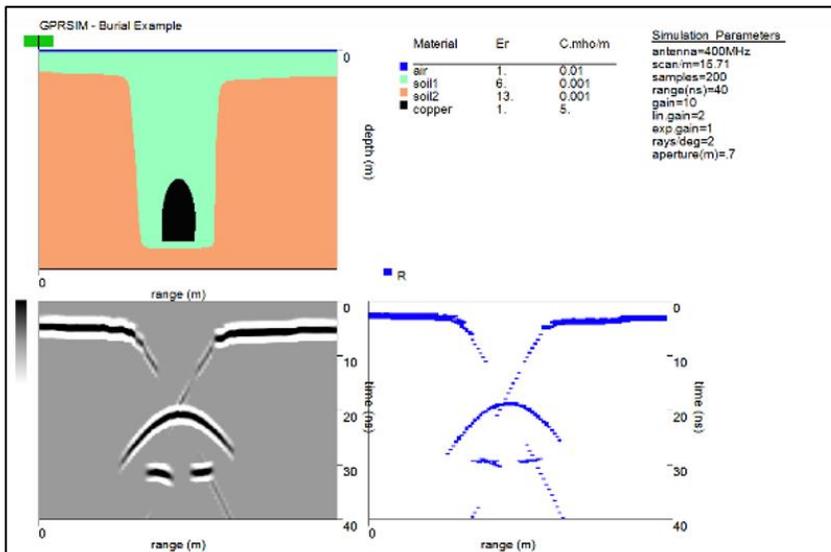
C. GPRSIM final simulation output which includes the model, material characteristics, radargram, and reflection characteristics

Figure 66: GPRSIM final simulation running process and output which includes the model, material characteristics, radargram, and reflection characteristics

gprMax is useful for creating simpler models and operates through a command prompt or Python interface. B-scans are compiled from multiple A-scans. However, the models are limited to representing simple targets in simple or somewhat complex matrices due to the range of parameters used to create the models the simulated radar responses are derived from. The parameters used to create materials within the model are described in Table 23.

Table 23: Parameters used to create materials in gprMax simulation models

Parameter	Description
ε_{rs}	DC (static) relative permittivity
$\varepsilon_{r\infty}$	Relative permittivity at theoretically infinite frequency
τ	Relaxation time of the medium in seconds
σ	DC (static) conductivity, measured in Siemens/m
μ_{τ}	Relative permeability
σ^*	Magnetic conductivity
Identifier	String identifier to define material

7.2.2. *Labelling Data*

Training data for the object detection task were labelled with rectangular bounding boxes around target objects. Data were labelled using Dataturks, an in-browser image labelling app with text, rectangle, and polygon labelling capabilities (shown in Figure 67) so that bounding box classes and coordinates could be exported in JSON and XML formats for easy import into machine learning models. Bounding boxes were assigned one of two classes – 'Grave' and 'Not Grave.' Where 'Not Grave' labels indicated responses similar to graves but proven not to result from graves through corroboration from the associated archaeological excavation report or an analysis of the timeslice data, as well as any examples of noise in data and other images features that should not be considered examples of graves.
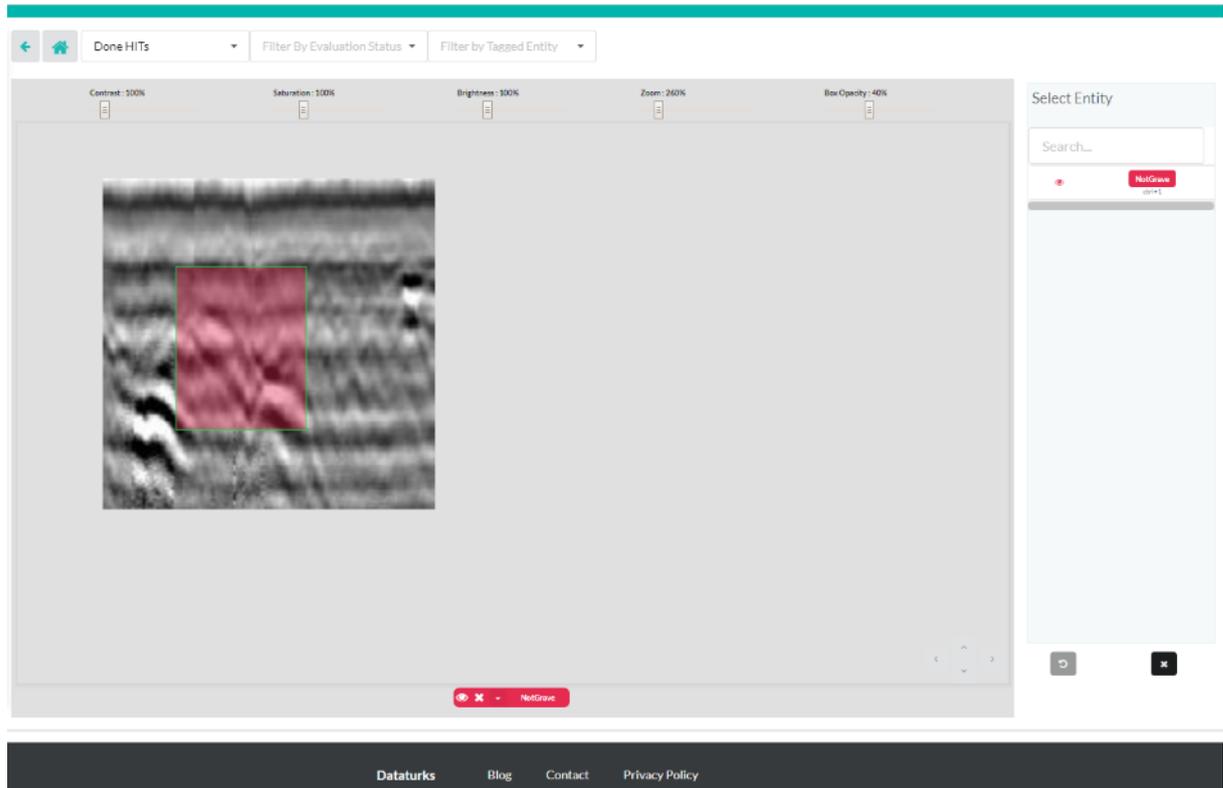
Figure 67: Example of labelling training data with bounding boxes in Dataturks

The bounding box coordinates were exported in JSON files and integrated into the model for training. The original images and labels, both class and bounding box, are extracted from the JSON files in the first running of the training script, as shown in Figure 68.

```
In [ ]: js = (path/'.json').open().readlines()
        bbs_g = [json.loads(x) for x in js]

        # non-graves / negative examples
        js = (path/'.json').open().readlines()
        bbs_ng = [json.loads(x) for x in js]

        # one bounding box per negative example
        for i in range(len(bbs_ng)):
            if bbs_ng[i]['annotation'] is not None:
                bbs_ng[i]['annotation'] = [bbs_ng[i]['annotation'][0]]
            else:
                bbs_ng[i]['annotation'] = [{'label': ['other'], 'points': [[0., 0.], [0., 0.], [0., 0.], [0., 0.]]}]

        bbs = bbs_g + bbs_ng

        dest = path/'Downloaded'
        urls = [bb['content'] for bb in bbs]
```

```
In [ ]: # download the data
        from fastai.vision.data import download_image

        url2name = lambda x: '_'.join(x.split('___')[-1].split('_')[1:])

        def _download_image_inner(dest, url, i):
            dest_file = dest/url2name(url)
            if not dest_file.exists():
                download_image(url, dest_file)

        dest.mkdir(exist_ok=True)
        for url in urls: _download_image_inner(dest, url, 0)
```

Figure 68: Section of training script for extracting training images and labels from JSON files

7.3.   Classification

Initial tests were performed to assess classification accuracy, a simple approach to automated data interpretation, before initiating tests for object detection, a more sophisticated approach to automated data interpretation. Several approaches to classifying images were explored as more training data became available. All models used convolutional neural networks. The most appropriate model was determined by its accuracy, F_1 score, and loss (for assessing overfitting of the model), Maxwell Correlation Coefficient (MCC). Accuracy is defined by the number of correct predictions divided by the number of total predictions a model makes and is used as an indicator of how well a model can make correctly predict class assignments. The F_1 score accounts for the precision and recall metrics; where, precision measures how well a model can predict positives (e.g. image contains a grave), and recall measures the effect of false negatives (e.g. model predicted no grave in an image, but there is a grave in the image) on positive predictions. Loss is an indicator of a model's prediction capability. The MCC measures the correlation between labels and predictions in a binary classification problem using true positive, true negative, false positive, and false negative predictions. An MCC score of 1 indicates a good correlation and -1 indicates a weak correlation. Further discussion of these and other metrics is provided in the glossary in Appendix A

7.3.1.   *TensorFlow*

Machine learning models are readily available open-source through TensorFlow and Python's API. TensorFlow is an open-source machine learning framework from Google™ (Huang et al. 2017). A convolutional neural network (CNN) using TensorFlow models (Inception V3 and Inception Resnet V2) with real-valued feature vectors derived from tiff image processing and feature extraction from known true positive and true negative grave responses were tested initially.

The CNN uses a binary classifier to classify a response as either a grave or not a grave with varying degrees of confidence. A second script for a multiclassifier was developed to determine the viability and capability of classifying grave responses using established typologies (e.g. stone-lined, simple earthen, pillow stone) based on the most suitable training data available. The CNN uses Inception V3 and Inception Resnet V2 deep learning models available from TensorFlow. Both models can retrain the last layer of the network, but the models differ in their number of convolutional and pooling layers as well as their ability to assign classifiers. Inception V3 is used for the binary classification convolutional neural network. Figure 69 provides a representation of the algorithm's architecture to demonstrate the convolutional and pooling layers. The description of the layers in the Inception Resnet V2 architecture is included in Table 24.
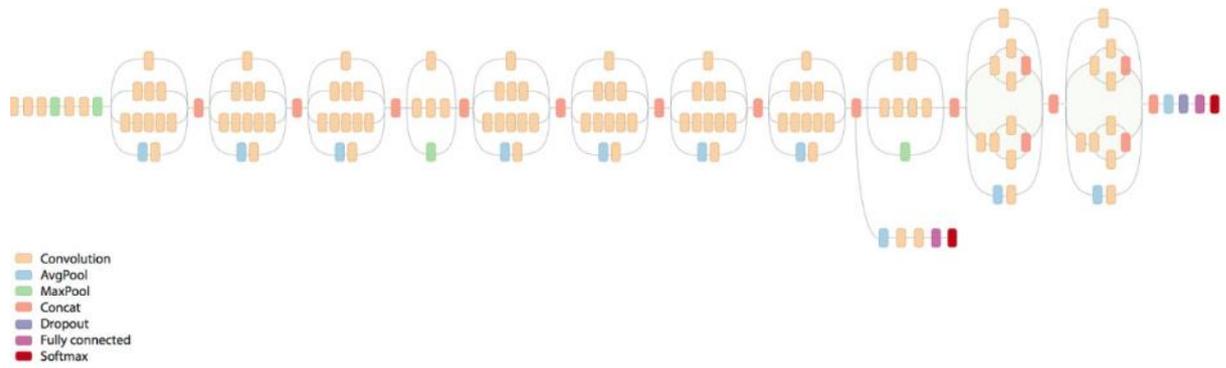
Figure 69: Schematic of Inception V3 architecture (Shlens 2016)

Inception Resnet V2 is used in conjunction with a single shot multibox detector (SSD) for object localisation and detection, and, unlike Inception V3, the model can handle multiple classifiers (Liu et al. 2016). Figure 70 provides a representation of the model's architecture to demonstrate the convolutional, pooling, and residual layers. The networks recognise minimum 32x32 pixel images in RGB or greyscale as an input and use multiple convolutional and pooling layers to inform the fully connected output layers. Inception Resnet V2 utilises residual connections between groups of convolutional and pooling layers to improve performance (Alemi 2016). The description of the layers in the Inception Resnet V2 architecture is provided in Table 24.

Attempts to create a classification model using the Inception ResNet V2 architecture pre-trained on ImageNet data (using ImageNet weights) and Keras library (see Table 21) proved to be unsuccessful as the overfitting of the model, demonstrated by the loss function and accuracy in Figure 71 and Figure 72, could not be adequately handled. However, in the future, this architecture may prove successful in a multiclassifier problem when more training data for each class is available.
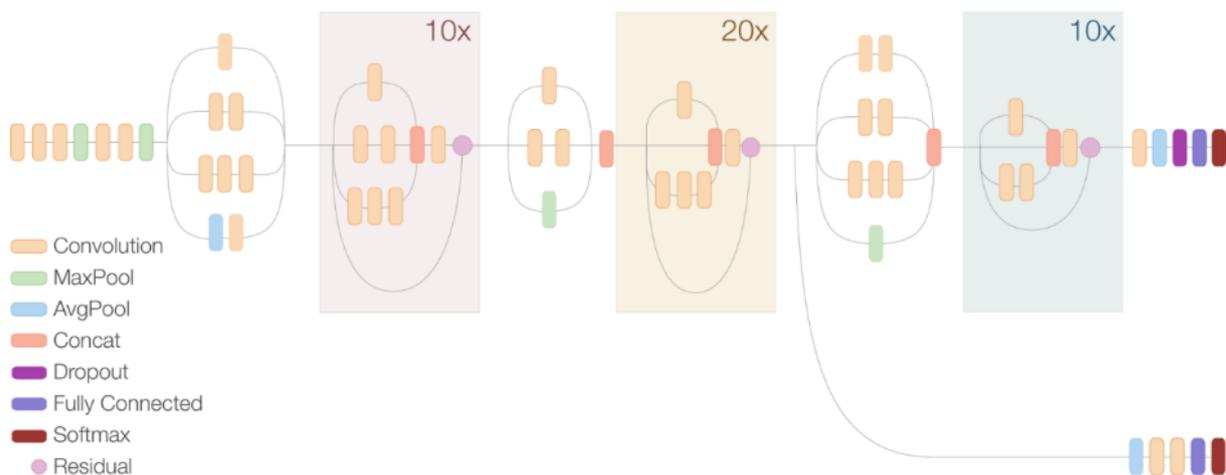
Figure 70: Schematic of Inception Resnet V2 architecture (Alemi 2016)

Table 24: Description of layers in the Inception Resnet V2 and Inception V3 architectures

Layer	Description
Convolutional	Filters images and extracts subregions – each layer creates a feature map comprised of individual values mathematically representative of each subregion
Average Pooling	Determines the average value for each subregion of the feature map in order to downsample, or reduce, the size of the feature map
Max Pooling	Determines the maximum value for each subregion of the feature map to downsample, or reduce, the size of the feature map
Concat	Concatenates the list of tensors along one dimension
Dropout	Prevents overfitting of the model by randomly setting a fraction of inputs to 0 during training
Fully Connected	Produces a tensor variable by creating a fully connected weight matrix and multiplying it by the input tensors
Softmax	Activation function which normalises the inputs into a probability distribution
Residual	Shortcut to building a deeper network which copies an activation layer and moves it deeper within the network

To address the current binary classification problem (grave vs no grave), a classification model using the Inception V3 architecture pre-trained on ImageNet data was also retrained using the GPR dataset. As is demonstrated by the accuracy and loss trends of the retrained Inception V3 presented in Figure 71 and Figure 72, this classification model also had a problem with overfitting. Additional Dropout layers and regularisers were added to the existing architecture, which led to modest improvement. The size of the training dataset was then increased, and higher quality examples were added. The model still achieves high accuracy and low loss in training but the opposite in testing, which is likely a result of overfitting. In the following images, epoch refers to the number of times the algorithm has gone through the entire training dataset.
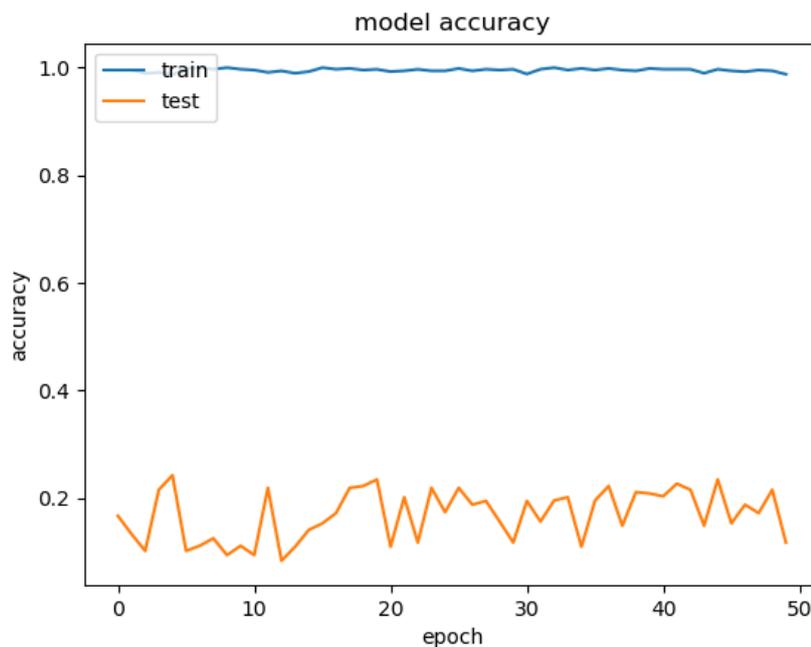
Figure 71: Classification accuracy of retrained Inception V3 model over 50 epochs
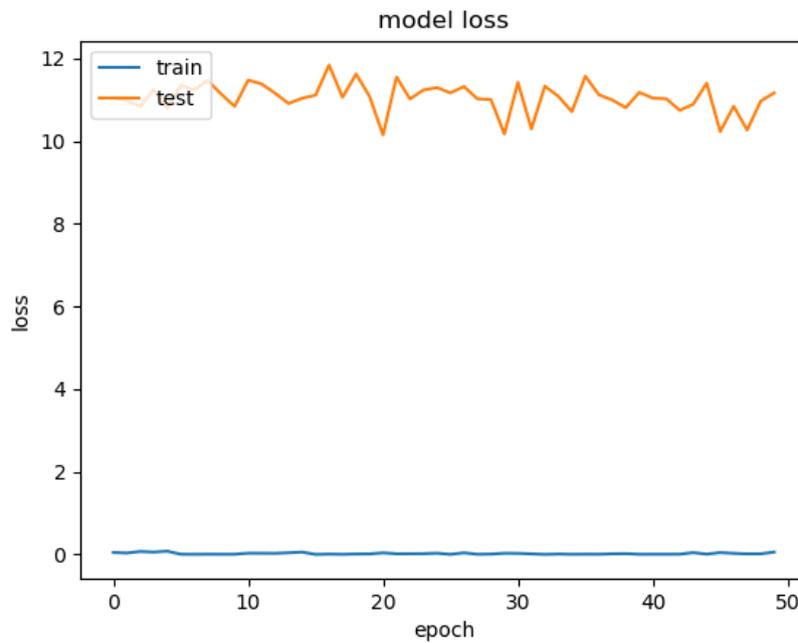
Figure 72: Loss of retrained Inception V3 model over 50 epochs. Low loss indicates a model is good at making predictions.

Because of these issues with the Inception models, classification models using ResNet and VGG models were also investigated to assess if either method of coping with overfitting would suit the data in this project (see Sections 7.3.2 and 7.3.3). The metrics for each model are derived from the number of true positive (TP), true negative (TN), false positive (FP), false negative (FN) classifications made during validation. Each metric is discussed in detail in Appendix A. The metrics of the tested models provided in Table 25 demonstrate ResNet architectures are more appropriate for this classification problem and their accuracy increases as the size of the architecture increases as it should according to He et al. (2016).

Table 25: Metrics of ResNet and VGG models tested

Model	TP	TN	FP	FN	Precision	Recall	F_1 Score	MCC
ResNet18	75	74	16	27	0.824176	0.735294	0.777202	0.557183
ResNet34	75	75	15	27	0.833333	0.735294	0.78125	0.568627
ResNet50	78	90	7	17	0.917647	0.821053	0.866667	0.753812
ResNet101	91	78	15	8	0.858491	0.919192	0.887805	0.761675
ResNet152	88	92	3	9	0.967033	0.907216	0.93617	0.87678
VGG19_bn	73	89	10	20	0.879518	0.784946	0.829545	0.689957

7.3.2. *VGG*

To achieve a more discriminative, accurate decision function than those configurations which use single, larger rectification (e.g. ReLU) layers, Simonyan & Zisserman (2015, p. 3) tested the use of stacked convolutional layers. Visual Geometry Group (VGG) configurations are convolutional neural networks which implement stacks of 3x3 convolutional layers rather than a single 7x7 convolutional layer, which "incorporate three non-linear rectification layers instead of a single one, which makes the decision function more discriminative" (Simonyan & Zisserman 2015, p. 3). VGG19 is the largest of the configurations provided by Simonyan & Zisserman (2015, p. 3), with 19 weight layers and 144 parameters, as shown in Figure 73. The difference in accuracy between VGG16 and VGG19 is negligible, but VGG19 performed faster (Simonyan and Zisserman 2015, pp. 6–8); however, further improvements were made to the VGG19 model by Simon et al. (2016, p. 1).
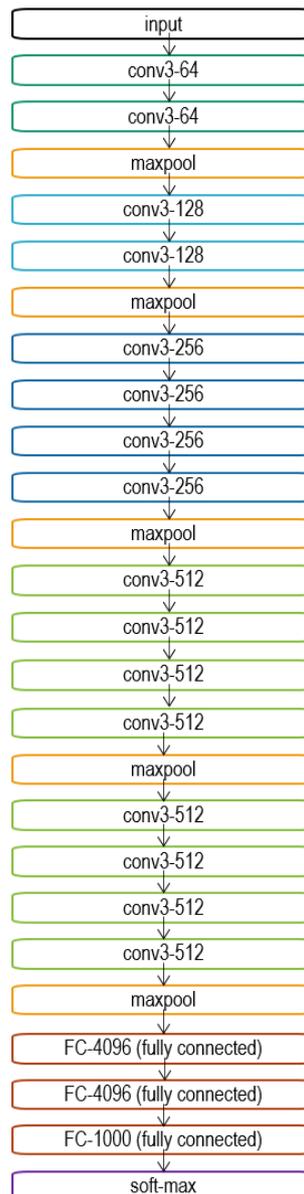
Figure 73: VGG19 architecture *(*adapted from Simonyan & Zisserman 2015, p. 3)

In adding batch normalization layers after the convolutional layers in the VGG19 model (shown in Figure 73), the VGG19_bn model can use a higher learning rate than VGG19, which leads to a lower error rate (Simon et al. 2016, p. 1).

Tests using the VGG19_bn model on GPR data were implemented using the PyTorch and fastai libraries as the model is available pre-trained in the library and ready to use for transfer learning. The existing model was retrained in the last layer using the same training data as in the TensorFlow and ResNet models.
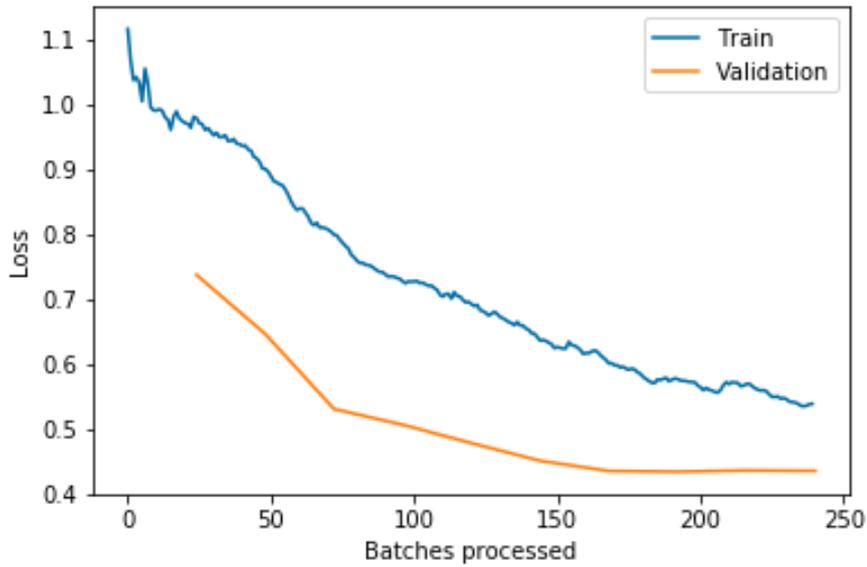
Figure 74: Training and validation loss of VGG19_bn model across ten epochs, where batches are a portion of the training dataset and the batch size is defined in the training script. Low loss indicates a model is good at making correct predictions, with a perfect model at a loss value of 0.
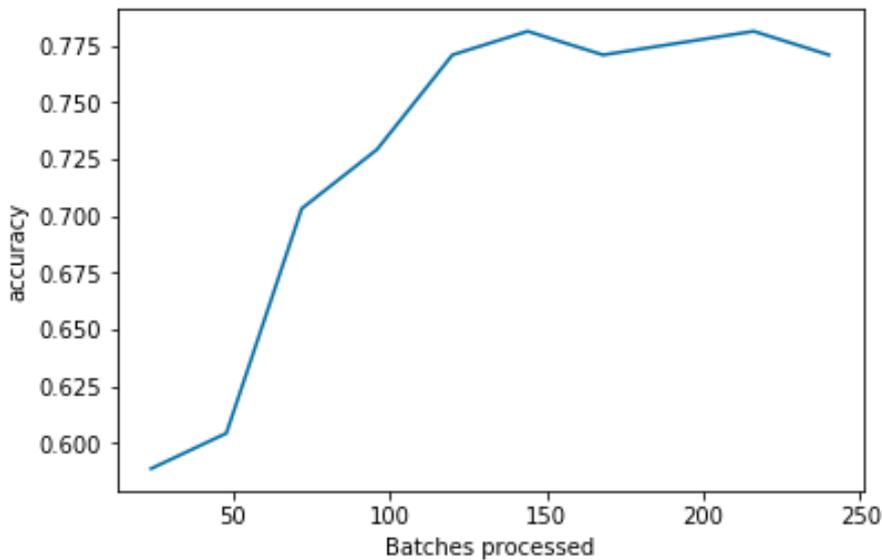
Figure 75: Accuracy of VGG19_bn model across ten epochs, where batches are a portion of the training dataset, and the batch size is defined in the training script

Figure 74 shows the model's training and validation loss decreased through the ten epochs and had a difference of *c*. 0.15 on the final epoch. These trends indicate that the model fitted well; however, the final accuracy (84.3%), shown in Figure 75, and MCC (0.689957) were still much lower than the larger ResNet models.

The confusion matrix in Figure 76 shows the number of true positive, false positive, true negative, and false negative predictions made. Aside from poor accuracy, the VGG19_bn model also was not chosen as the final model because it had a higher rate of false negatives than false positives.

Figure 76: Confusion matrix describing the accuracy of predictions made using the VGG19_bn model which shows the number of correct predictions and incorrect predictions

The class activation maps, shown in Figure 77, demonstrate that the model learned to identify and localise the graves with high confidence without being provided with this information during training. In these heatmaps, or activation maps, the highlighted areas indicate images features the model is using to make predictions. These results show that even simple models with little training can be readily applied to the complex task of classifying graves. However, for data interpretation in archaeogeophysics, further localisation of the target is necessary.

prediction/actual/loss/probability

Figure 77: Heatmaps indicating the most accurate predictions made using the VGG19_bn model, where the highlighted area indicates a prediction

7.3.3. *ResNet*

Generally, larger architectures outperform smaller architectures, as demonstrated by He et al. (2016, p. 7-8), unless there is a risk for overfitting the model. A brief discussion on overfitting is provided in Chapter 6.3.1. Five sizes of ResNet architectures were tested to ensure that the largest architecture, ResNet152, would perform as well as expected. While the training/test data was consistently split 80/20, the random selection of the images for the training and test datasets will affect the accuracy of the model, ResNet models generally outperformed VGG and TensorFlow models. ResNet50 and ResNet152 performed the best within the residual networks.
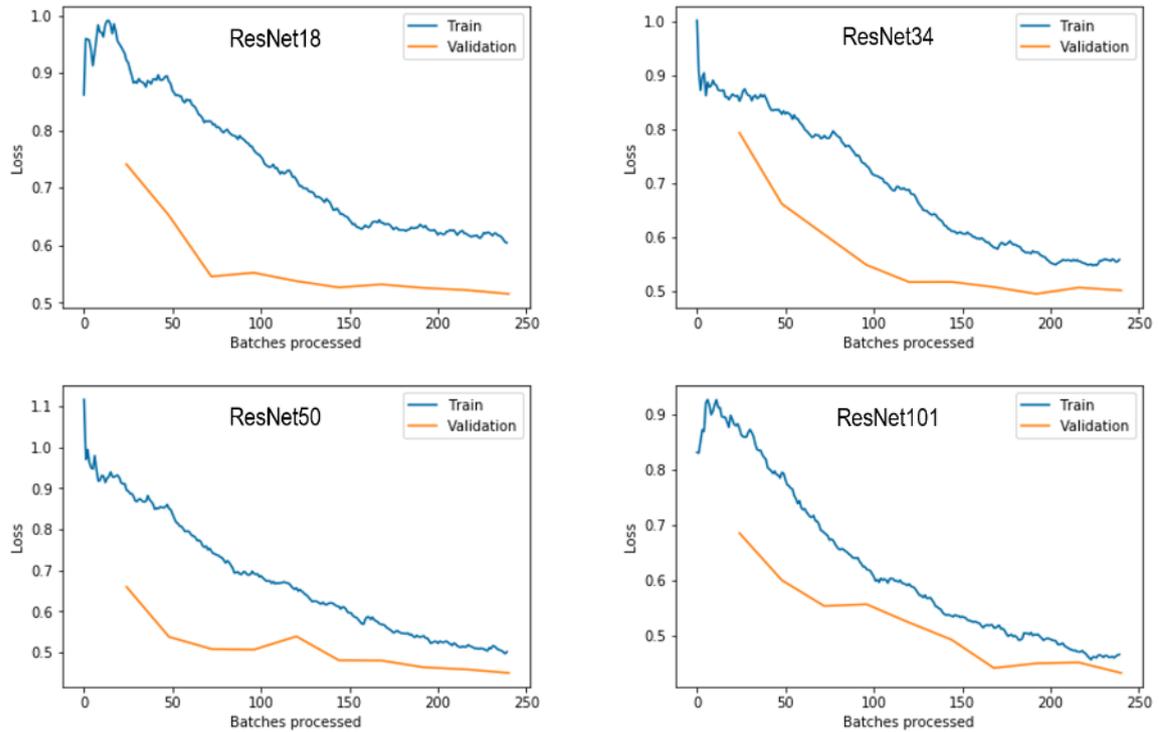
Figure 78: Training and validation losses for ResNet18, ResNet34, ResNet50, and ResNet101 models across ten epochs, demonstrating that training and validation loss trends become more similar as architecture size increases. Low loss values indicate a model is good at making correct predictions, with a perfect model at a loss value of 0.

The near convergence of the training and validation loss is indicative of a well-fitted model. The closer to zero and the smallest difference between training and validation loss demonstrates a model is training well. Accuracy of the model also improved with architecture size as expected. Figure 78 illustrates the improvement in loss, while Figure 79 illustrates the improvement in accuracy as the architecture size increases, with ResNet101 having a near-perfect loss.
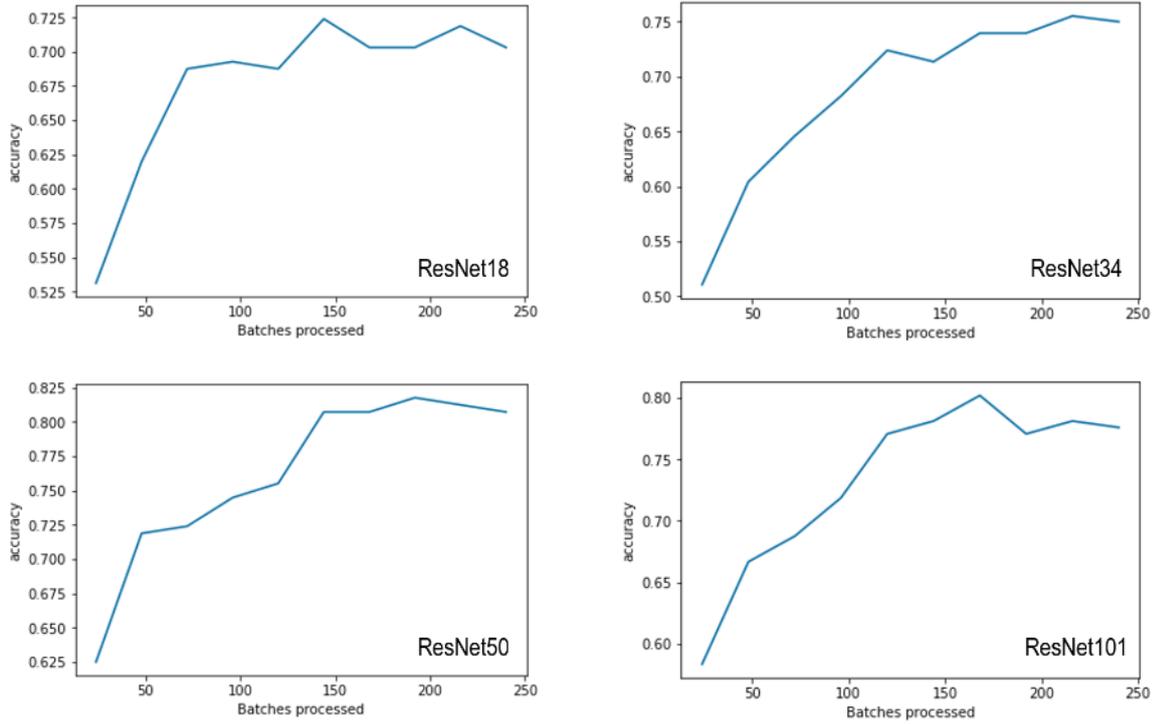
Figure 79: Accuracy of ResNet18, ResNet34, ResNet50, and ResNet101 models across 10 epochs
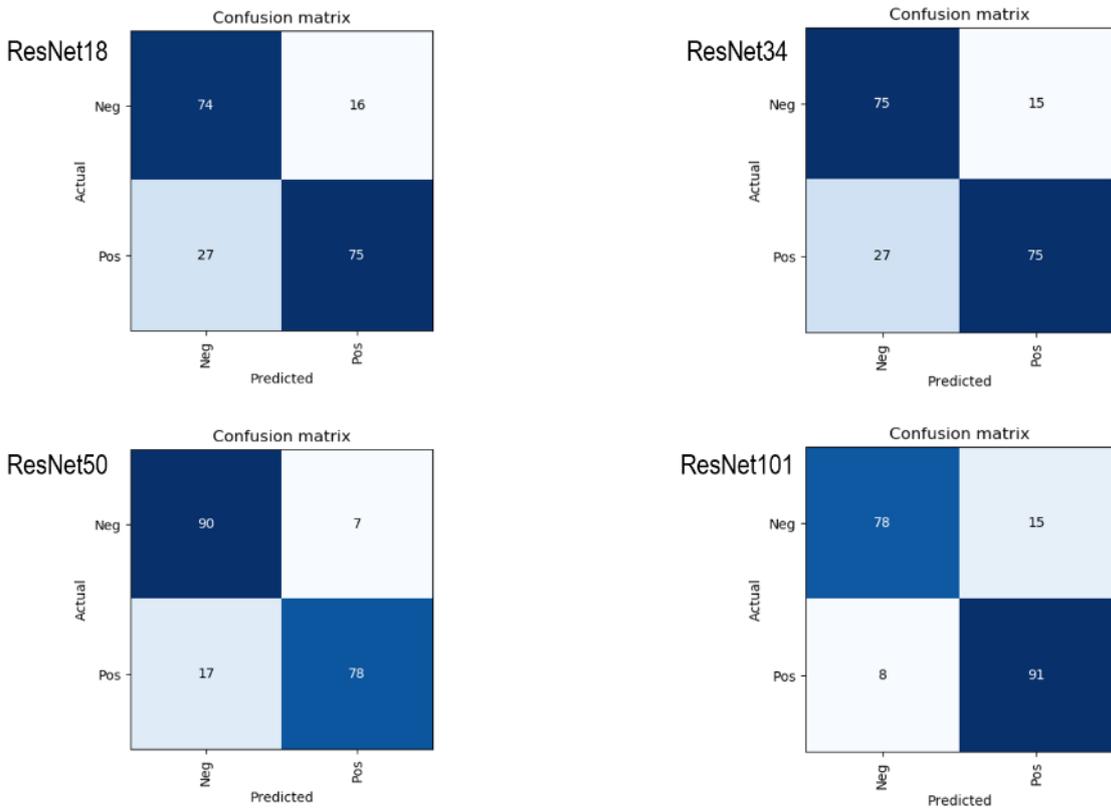
Figure 80: Confusion matrices describing the accuracy of predictions made using ResNet18, ResNet34, ResNet50, and ResNet101 models with the number of correct predictions and incorrect predictions, particularly highlighting the rate of false negative predictions

144

Interestingly, ResNet101 is the only model producing a higher false positive (Type I error) rate than false negative (Type II error) rate, as illustrated in Figure 80. While a perfect model is preferable, this is unachievable. Rather, to be cautious in the interpretation of archaeogeophysical data, false positives are preferred over false negatives. Thus, a model with high accuracy and low rate of false negative predictions, as in ResNet101, is preferred.

The accuracy achieved during tests with the ResNet50 and ResNet101 models were promising. The continuing increase in accuracy and improvement in loss during both training and validation justified the testing of the ResNet152 model as well.
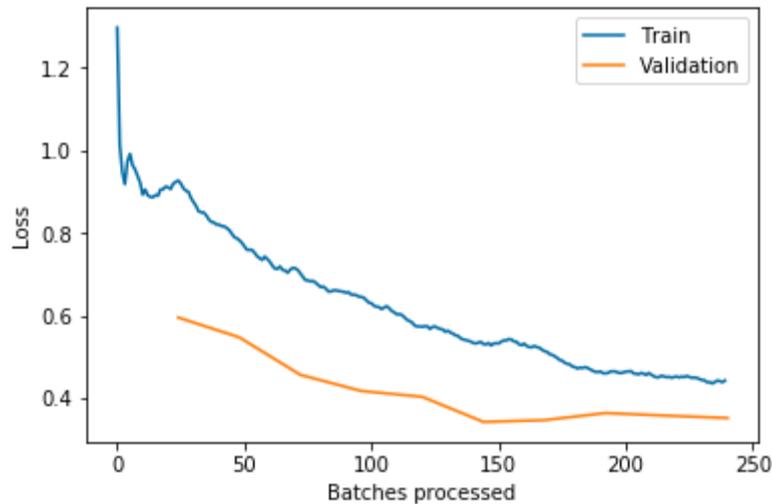
Figure 81: Training and validation loss of ResNet152 model across ten epochs, where the low loss values indicate the model is good at making correct predictions
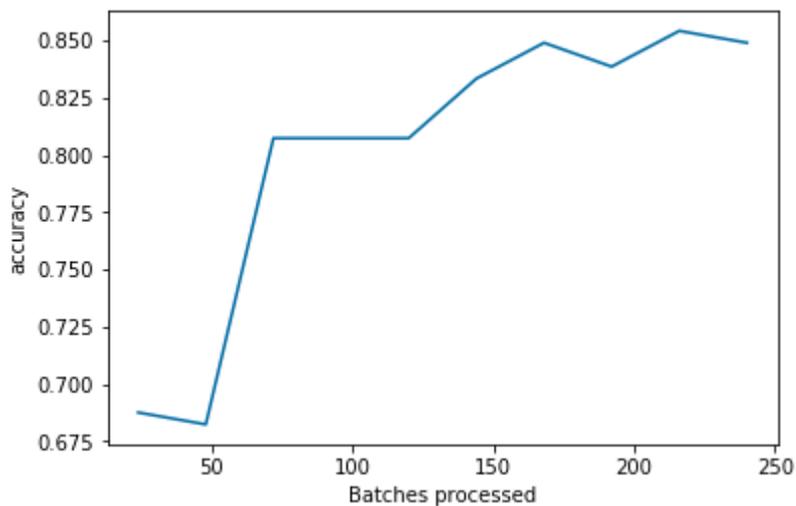
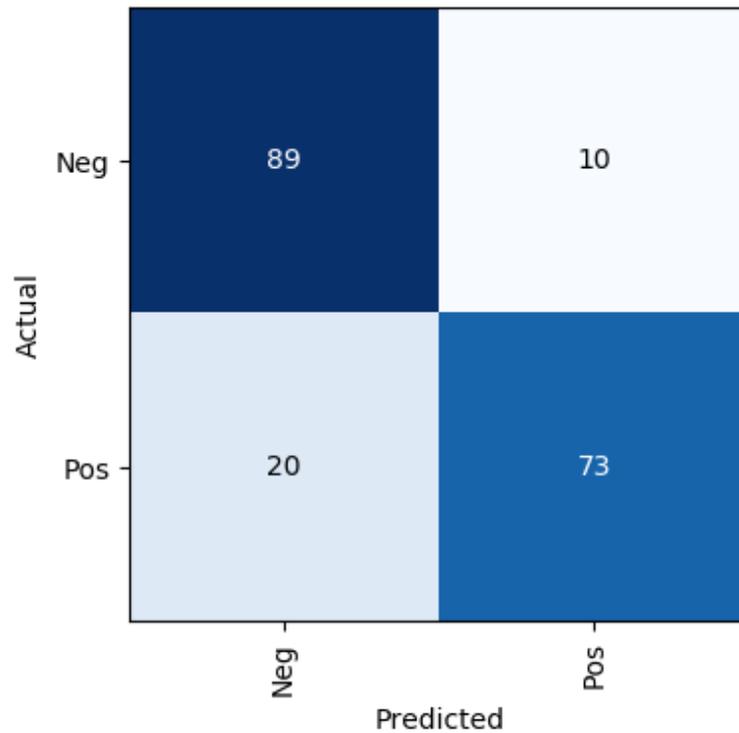
Figure 82: Accuracy of ResNet152 model across ten epochs
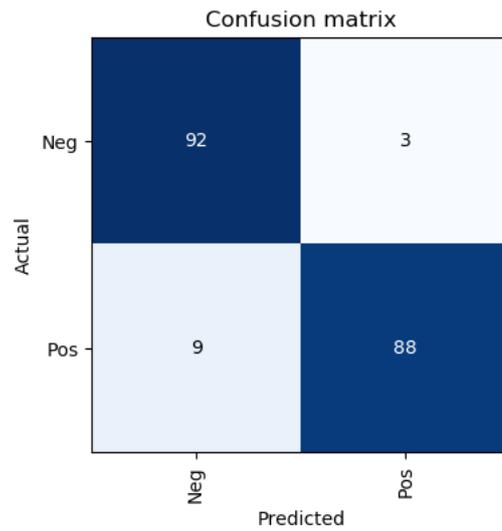
Figure 83: Confusion matrix describing the accuracy of predictions made using the ResNet152 model and shows the number of correct predictions and incorrect predictions
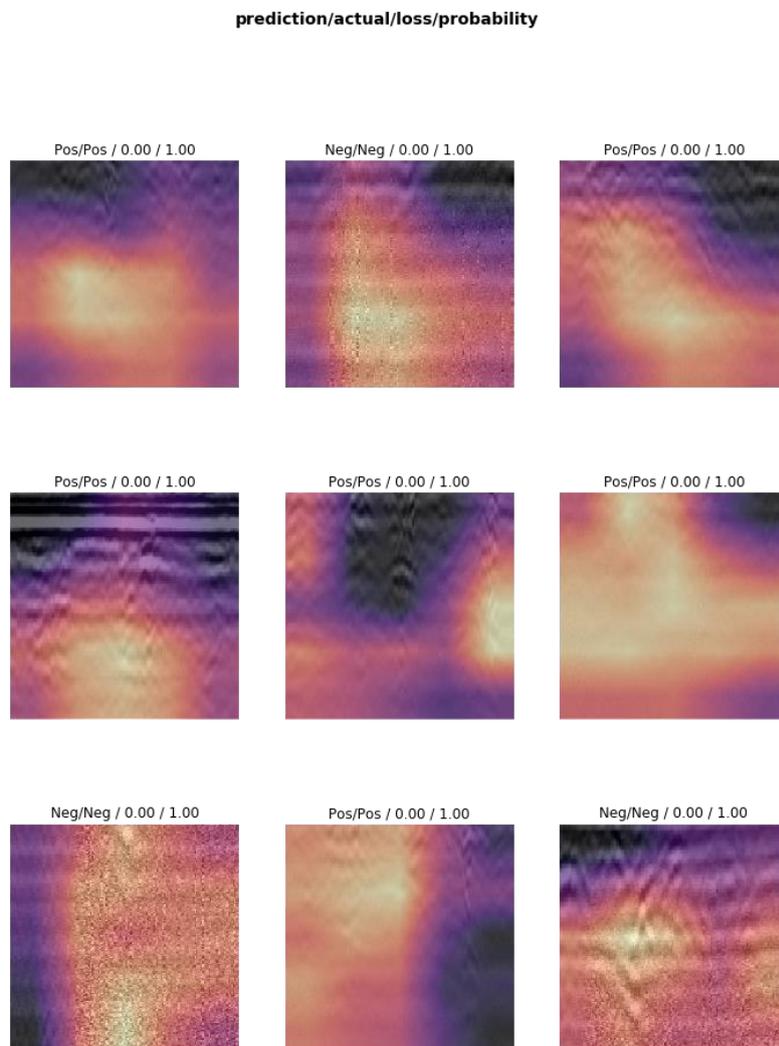
Figure 84: Heatmaps indicating the most accurate predictions made using the ResNet152 model, where the highlighted area indicates a prediction

prediction/actual/loss/probability
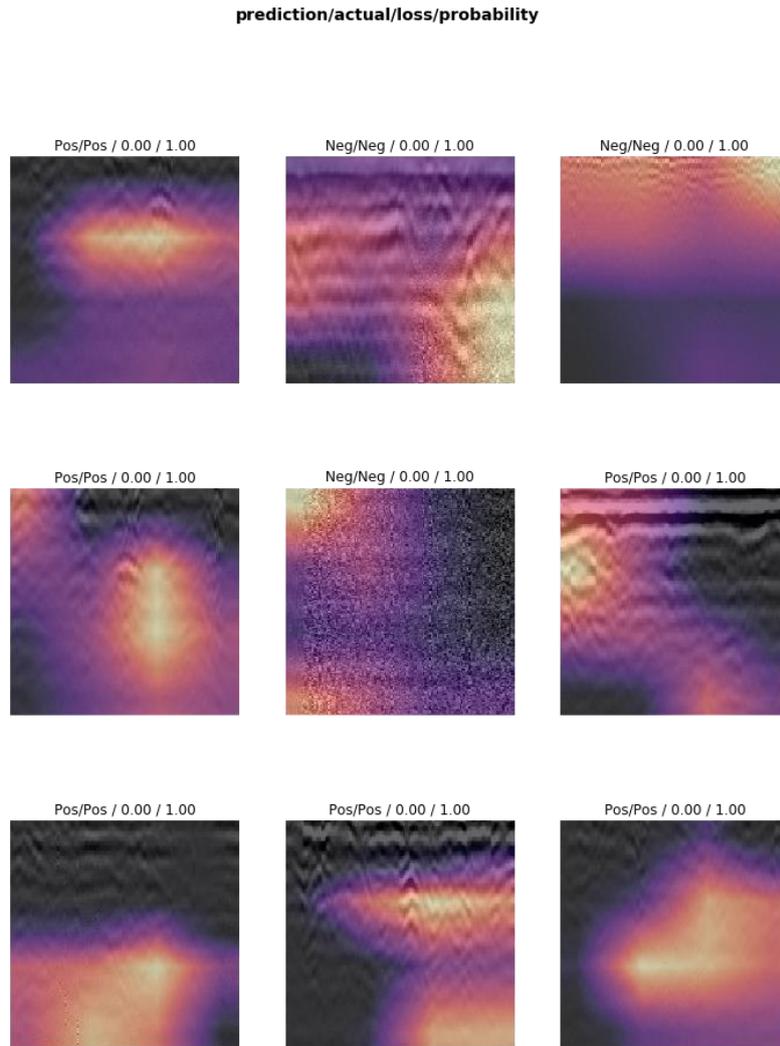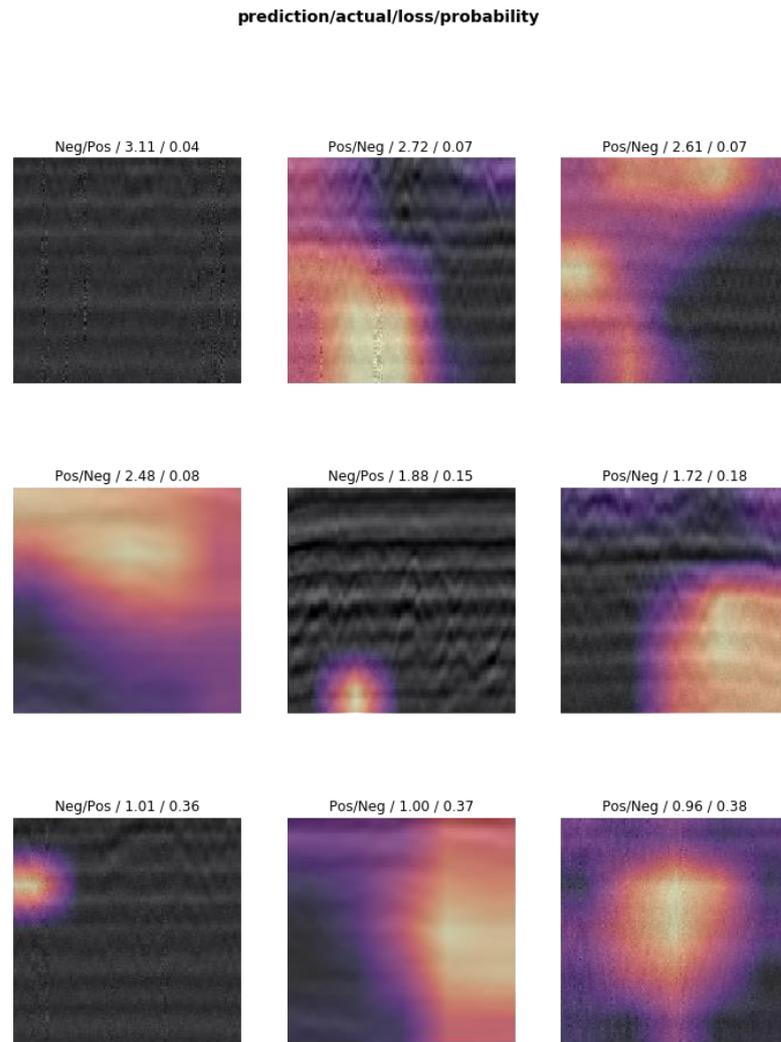

Figure 85: Heatmaps indicating the least accurate predictions made using the ResNet152 model, where the highlighted area indicates a prediction

When isolating the models which achieved an accuracy above 0.8, the Matthew's Correlation Coefficient (MCC) and F_1 score are good indicators of a suitable model. The metrics for the best performing ResNet models provided in Table 26 justify the use of the ResNet152 model even though the ResNet101 model had a preferable ratio of false positives and false negatives. ResNet152 achieved a near-perfect F_1 score and high MCC, and while it favours false negatives over false positives, the accuracy and F1 score are close to perfect, making it less of a concern.

Table 26: Metrics for the ResNet models with accuracy above 0.8

Model	TP	TN	FP	FN	Accuracy	Precision	Recall	F_1 Score	MCC
ResNet50	78	90	7	17	0.875	0.917647	0.821053	0.866667	0.753812
ResNet101	91	78	15	8	0.880208	0.858491	0.919192	0.887805	0.761675
ResNet152	88	92	3	9	0.9375	0.967033	0.907216	0.93617	0.87678

Based on the results from initial tests of the CNNs, ResNet152 was chosen as the base model for the image classification and object detection tasks.

7.4.   Object Detection

Following the results of the classification tests, a single object detection model was created to provide the end-user with a localised detection. The same training data from the classification tasks were also used for object detection. However, prior to training, the images were manually labelled with bounding boxes that indicate regions of interest. Rectangular bounding boxes were drawn around responses from two classes – graves and 'non-graves.' 'Non-graves' refer to those responses that are not a result of graves, including responses that are visually similar to graves but are known not to the result of graves, background geology, and signal noise. An example of the training data is presented in Figure 86. A JSON file containing the training data and bounding box coordinates is provided with the digital content of this thesis.
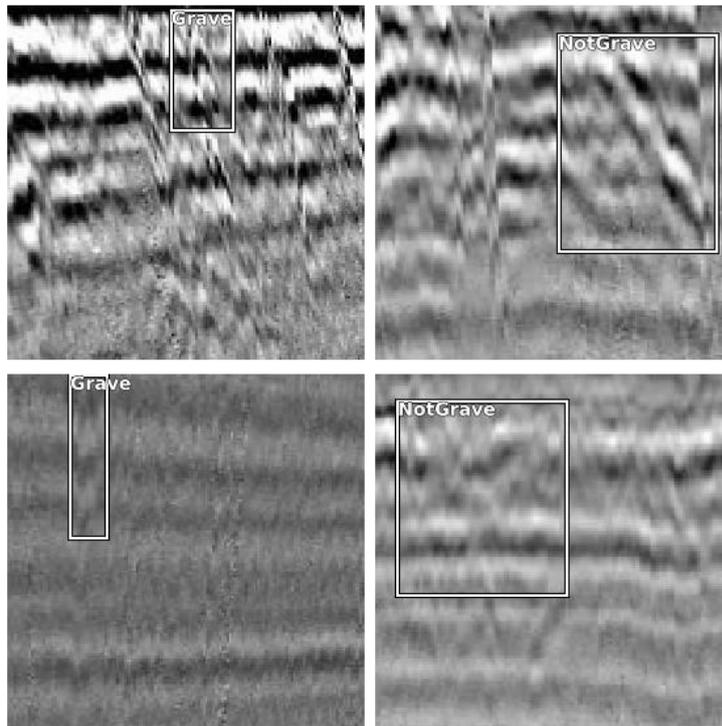
Figure 86: Example of bounding box training data for object detection

ResNet152 was used as a base model with a custom head (top of the network), batch normalisation and dropout, and loss function added to manage the input of and training with a single bounding box per image. The additions to the pre-trained ResNet152 model are presented in Figure 87. The custom head uses a dropout layer, three batch normalisation layers, and two 2D convolutional layers

instead of 2D convolutional layers, 2D batch normalisation layers, and ReLU layers in order to handle to the bounding box input. The model was trained using the fast.ai library.

```python
def l1_loss(input, tar_regr, tar_clsf):

    # mask inputs and targets to ignore bounding boxes for non-graves
    mask = tar_clsf[:,:,None].float().repeat(1, 1, 4)
    inp_masked = input[0] * mask
    tar_masked = tar_regr * mask
    return F.l1_loss(inp_masked, tar_masked, reduction='mean')

def cross_entropy(input, tar_regr, tar_clsf): return F.binary_cross_entropy(input[1], tar_clsf.float())

def accuracy(input, tar_regr, tar_clsf): return ((input[1] > 0.5).int() == tar_clsf.int()).double().mean()


class MyLoss(nn.Module):
    def forward(self, input, tar_regr, tar_clsf):
        tar_clsf[tar_clsf==2] = 0
        loss_regr = l1_loss(input, tar_regr, tar_clsf)
        loss_clsf = cross_entropy(input, tar_regr, tar_clsf)
        return loss_regr + loss_clsf


class ConvBNDrop(nn.Module):
    def __init__(self, nin, nout, stride=2, drop=0.1):
        super().__init__()
        self.conv = nn.Conv2d(nin, nout, 3, stride=stride, padding=1)
        self.bn = nn.BatchNorm2d(nout)
        self.drop = nn.Dropout(drop)

    def forward(self, x):
        return self.drop(self.bn(F.relu(self.conv(x))))
#            return self.drop(self.bn(F.leaky_relu(self.conv(x))))


class MyHead(nn.Module):
    def __init__(self):
        super().__init__()
        self.drop = nn.Dropout(0.1)
        self.sconv0 = ConvBNDrop(2048, 1024, stride=1)
        self.sconv1 = ConvBNDrop(1024, 512, stride=1)
        self.sconv2 = ConvBNDrop(512, 256)
        self.oconv_regr = nn.Conv2d(256, 4, 3, stride=1, padding=0)
        self.oconv_clsf = nn.Conv2d(256, 1, 3, stride=1, padding=0)

    def forward(self, x):
        x = self.drop(F.relu(x))
        x = self.sconv0(x)
        x = self.sconv1(x)
        x = self.sconv2(x)

        o_regr = torch.tanh(self.oconv_regr(x))
        o_clsf = torch.sigmoid(self.oconv_clsf(x))

        return [o_regr.squeeze(dim=-1).squeeze(dim=-1)[:,None,:], o_clsf.squeeze(dim=-1).squeeze(dim=-1)]
```

Figure 87: Modifications to the pre-trained ResNet152 model

The model was trained for ten epochs in the initial stage to determine the learning rate. The model was reset and then learned for ten epochs while recording the training loss, validation loss, L1 loss, cross-entropy, accuracy, and time per epoch. Then the model training continued with all layers unfrozen/trainable, first determining the learning rate and learner object. Unfreezing the model makes all parameters trainable, but for transfer learning, it is customary to retrain the last layer first and then unfreeze preceding parameters as making more layers trainable increases the amount of computation needed and increases training time. The model was the retrained using all trainable layers for ten epochs.

The metrics of the object detection model were recorded across 100 training sessions. The classification accuracy of the ResNet152 object detection model is summarised in Figure 88 - Figure 89. The training, L1, and validation losses are also summarised in Figure 90 - Figure 93.

Figure 88: Classification accuracy of the object detection model across 100 training sessions
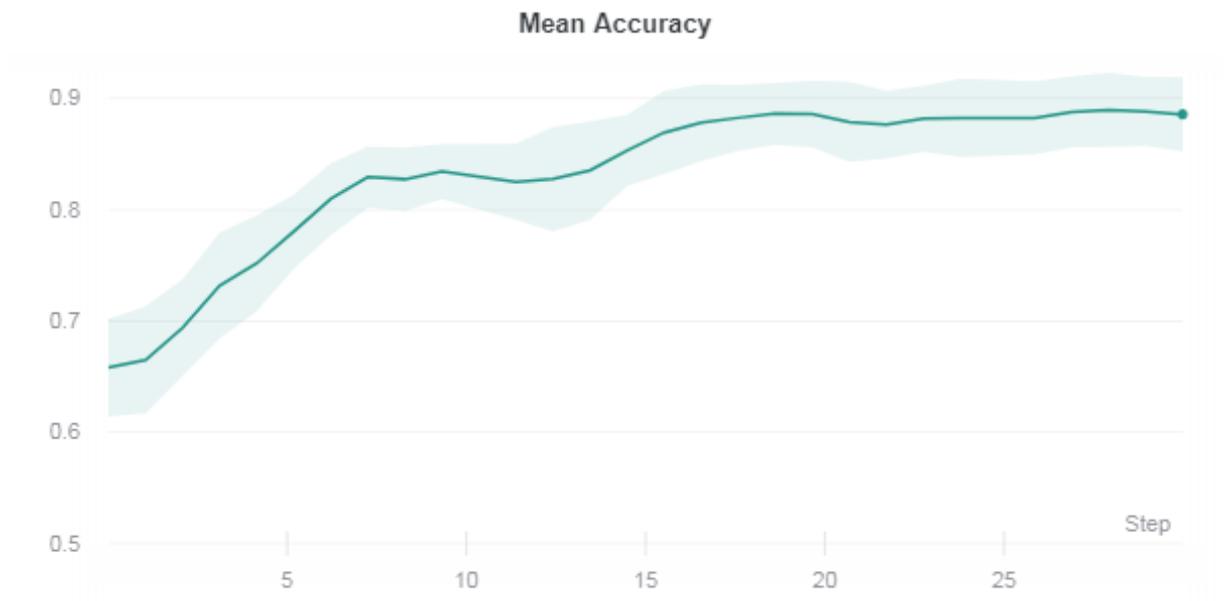
Mean Accuracy

Figure 89: Mean classification accuracy and standard deviation of the object detection model across 100 training sessions

Mean Loss
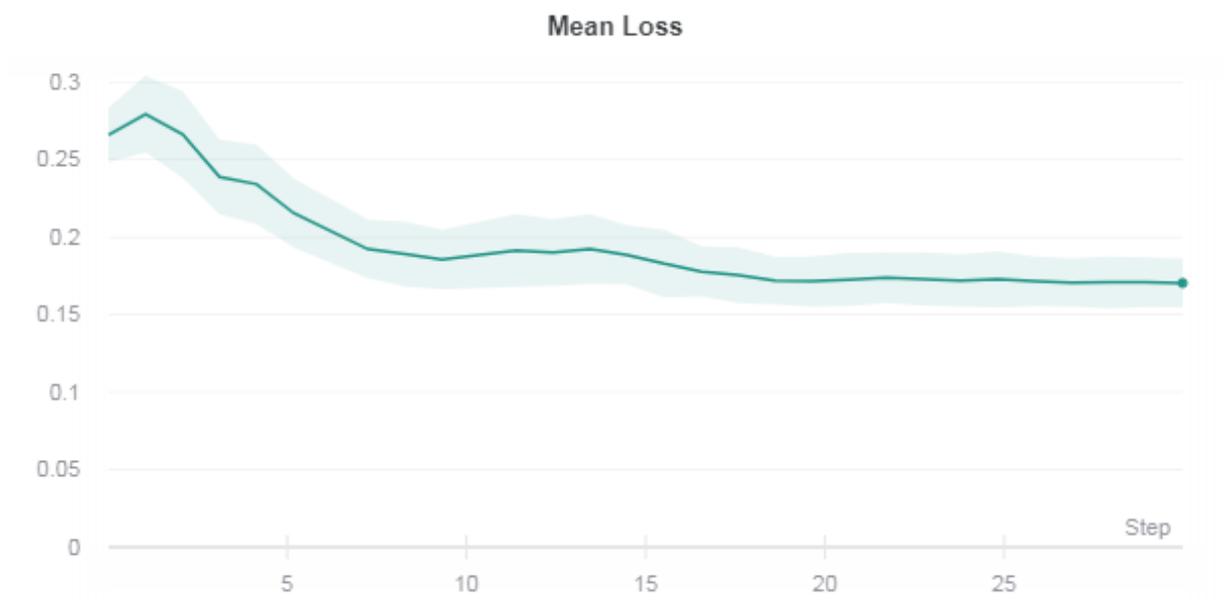
Figure 90: Mean training loss and standard deviation of the object detection model across 100 training sessions. Low loss values indicate a model is good at making correct predictions, with a perfect model at a loss value of 0.
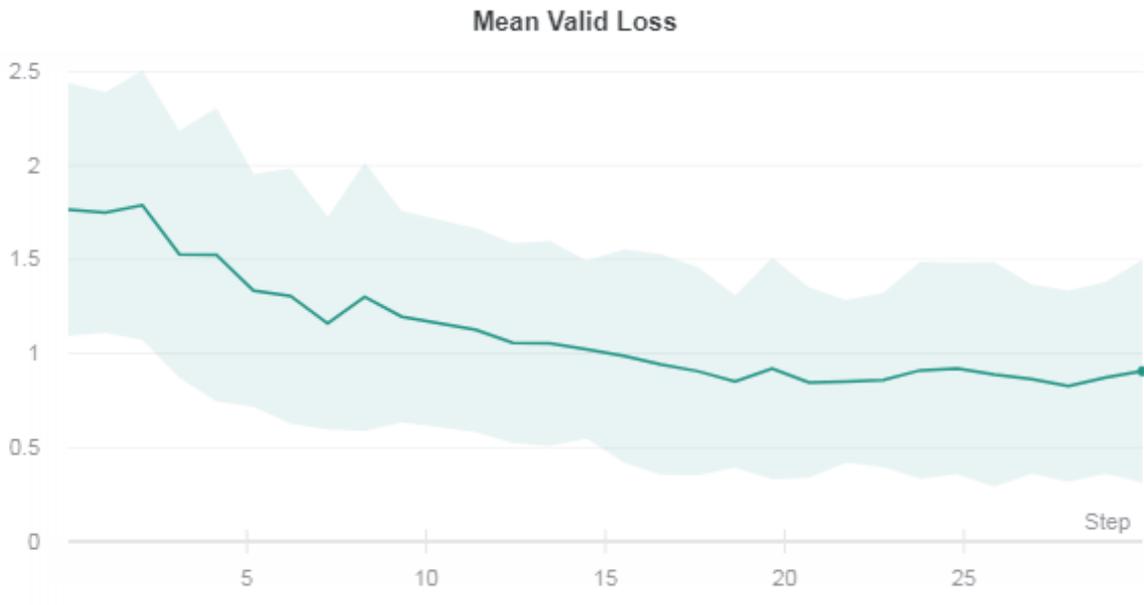
Mean Valid Loss

Figure 91: Mean validation loss and standard deviation of the object detection model for 100 training sessions. This trend indicates the model is becoming better at making correct predictions throughout training.

Mean Training Loss | Valid Loss | L1 Loss
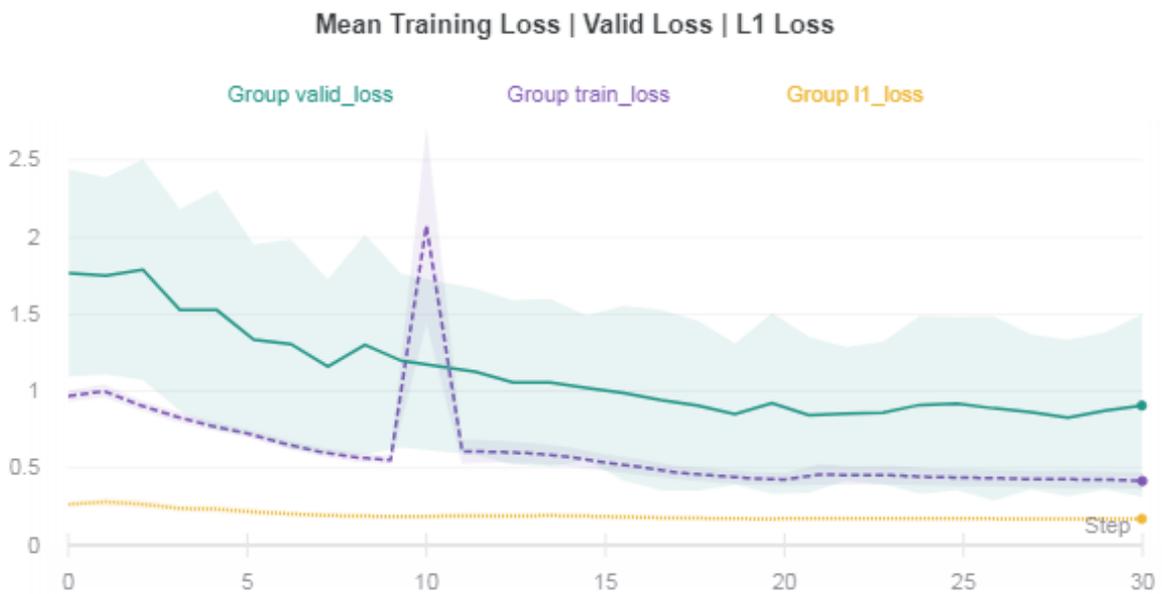
Figure 92: Mean training, L1, and validation losses and standard deviations of the object detection model across 100 training sessions. L1 loss, or mean absolute error, is the sum of all absolute differences. A loss value of 0 indicates a perfect model.
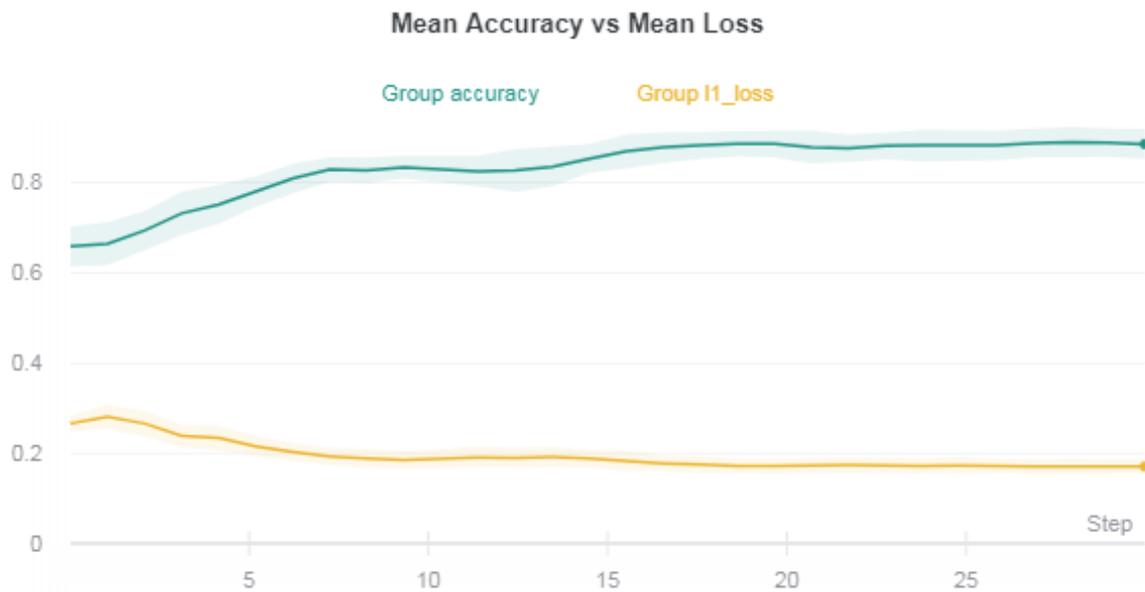
Figure 93: Mean training loss compared with classification accuracy and standard deviations of the object detection model across 100 training sessions

The final model was packaged and exported using pickle for integration in the web app and future implementation in ArcGIS.

7.5. System Properties and Training Time

The training was carried out on a workstation with the following system properties:

- Windows 7;
- Intel Xeon CPU E5-1650 3.20 GHz processor;
- Nvidia Tesla K40 12 GB GDDR5 GPU;
- 32 GB RAM.

Using this basic setup, a full training cycle of 10 epochs lasted a minimum of 1 hour 55 minutes 46 seconds, and a maximum of 2 hours 32 minutes 46 seconds. Over 100 complete training cycles were carried out using the ResNet152 architecture for assessing the capabilities of object detection on radargrams.

7.6. Deployed Classification API

Reilig for Radargrams, an input-output web app with a graphical user interface (GUI) is provided as a user-friendly medium for commercial geophysicists and volunteer groups to use the automatic object detection functionality of the model without prior knowledge of programming languages. The

final trained model was put in evaluation mode and packaged and linked to Python and JavaScript scripts to create the app.

Following training, the final trained model was packaged as a pickle (.pkl) file. Packaging the model in evaluation mode allows it to be used outside of the training script and keeps the learned weights and features associated with the model. A python script manages and routes the detection requests from end-users. Interaction with the inference script is implemented through a POST request using the code shown in Figure 94 below.

```python
@app.route('/analyze', methods=['POST'])
async def analyze(request):
    img_data = await request.form()
    img_bytes = await (img_data['file'].read())
    img = open_image(BytesIO(img_bytes))

    prediction = predict(learn, img)
    # arr = (img.data.numpy().transpose(1,2,0) * 255).astype('uint8')

    # if prediction[1][1] > 0.5:
    #     bbox, text = show(prediction[0], noshow=True)
    #     arr = cv2.rectangle(arr.copy(), tuple(bbox[:2]), tuple(bbox[:2]+bbox[2:]), (255, 0, 0), thickness=1)
    # else:
    #     arr = cv2.putText(arr.copy(), 'No detection', (10, 40), cv2.FONT_HERSHEY_SIMPLEX, 1.5, (252, 0, 0), thickness=2)

    # retval, buffer = cv2.imencode('.png', arr[:,:,::-1])
    # img_as_text = base64.b64encode(buffer)
    # return PlainTextResponse('data:image/png;base64,%s' % img_as_text.decode())

    bbox, text = show(prediction[0], noshow=True)
    out = {
        'prob': str(prediction[1][1].numpy()[0]),
        'bbox': str(bbox.astype('int').tolist()),
        'size': str(list(img.shape[1:]))
    }
    return JSONResponse(out)
```

Figure 94: Excerpt of code to interact with the trained model for inference in the web app

End-users interact with the machine learning model via a web interface. The web page was designed and created using HTML and CSS coding to maximise the functionality of the input and prediction requests and create a useable project output aimed at a target audience of users with no programming skills. The final static web app, which is comprised of the packaged trained model, JavaScript and python scripts, and HTML and CSS files for web page formatting, was packaged and deployed using Docker and Render and can be accessed at reilig.onrender.com. Render is a cloud provider that deploys and hosts Docker files as static websites and web apps, allowing for the constant update of backend and frontend code directly from GitHub. As such, all code associated with the web app interfaces created during this project is hosted on GitHub. The user interface is shown in Figure 95 and an example of detections being made in real-time is provided in the digital appendices.

154

In the Reilig for Radargrams web app, the end-user is prompted to upload a JPEG image which is at least 150x150 pixels. If the uploaded image is smaller than 150x150 pixels, the model has an increased error in identifying an accurate extent and location for the bounding box. There is no noticeable error when an image larger than 150x150 pixels is uploaded. When an appropriate image is uploaded, the end-user will initialise the object detection via the click of an 'Analyse' button. Once analysed, the original uploaded image is returned with either a single detection indicated by a blue bounding box or no bounding box and text below the image indicating no graves were detected. Both positive and negative detections are returned with a confidence score ranging from 0 to 1, 1 being the most confident. If the end-user would like to save only the bounding box, they can right-click on the returned image and save the image in PNG format. However, if the end-user would like to save the bounding box along with the image and corresponding text, the most appropriate methods for this are to screenshot the extent of the relevant section of the web page or to save the entire web page as a PDF. Either format is suitable for integration into an existing CAD file for GIS. The user guide containing instructions on how to format data and save the outputs from the object detector is provided in a link on the web app and is included in Appendix F. This guide also provides users with a brief overview of machine learning and an introduction to this research project.
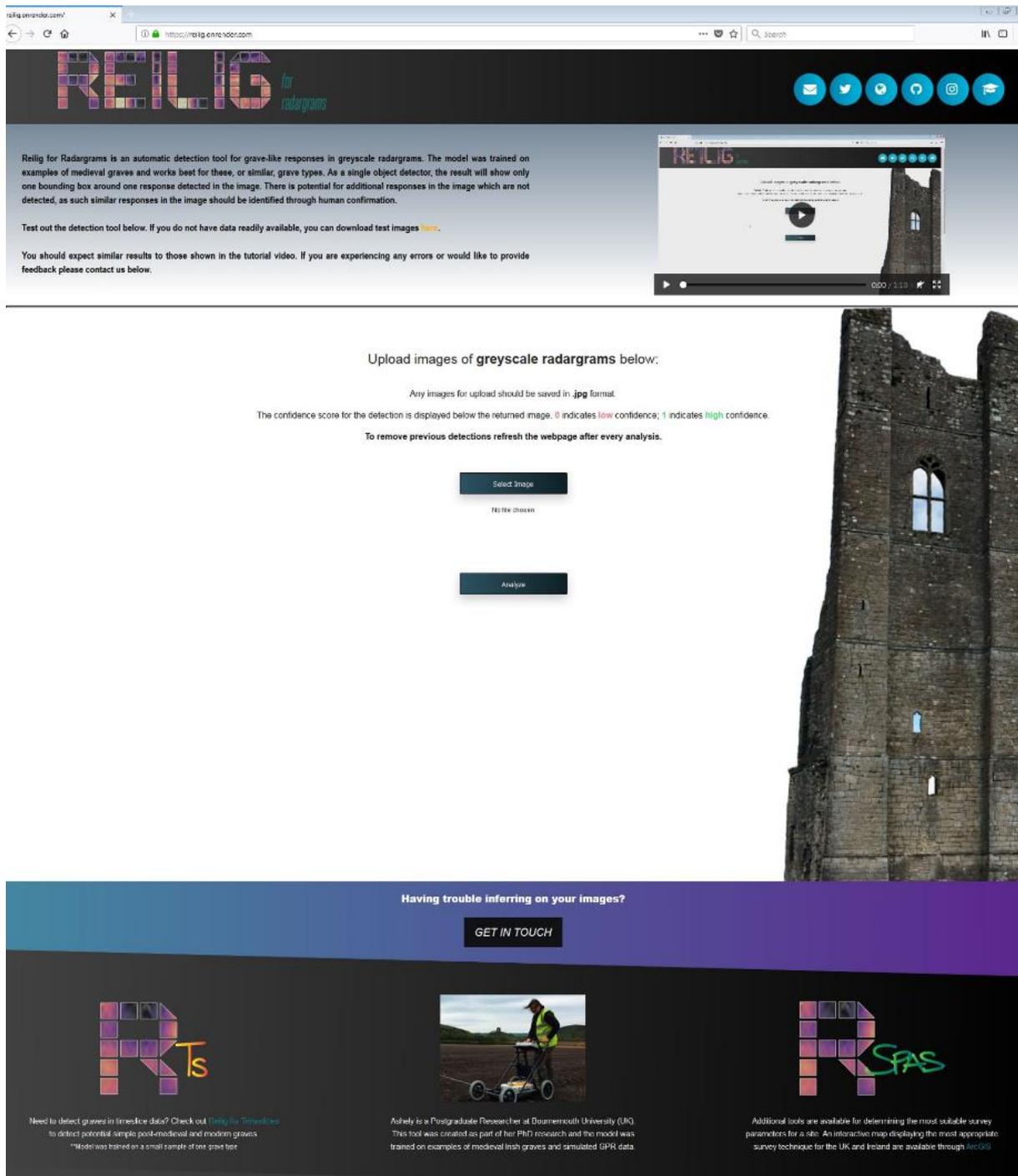
Figure 95: The layout of the Reilig for Radargrams web app user interface. The end-user is prompted with two buttons to upload and analyse their images, as well as a tutorial and multiple methods to contact the author for assistance with detecting on their images

8. INITIAL TESTS ON "REAL-WORLD" DATA

8.1. Introduction

The data provided to experts in the questionnaire discussed in Chapter 1.2 were also interpreted using the single object detector. The results of this automatic interpretation, their confidence, and the implications on the future viability of machine learning interpretation aides are presented in Chapter 8.2 and Chapter 10.

Data from the case study sites detailed in Chapter 8.6 - 8.12 were interpreted using the classification and object detection models. A visual comparison of the technical reports (included in digital appendices) with human-led interpretations of the data and the machine-led interpreted data was carried out. Focus was also paid to the time spent completing the human-led and machine-led interpretations and how this may impact on commercial surveys. These surveys provided data to test the viability of machine-led interpretation in a commercial setting, but in some ways were limited by site accessibility and landowner permissions. All data were compared to the manual interpretations and historic records accounted for when assessing the accuracy of the object detection results where there no was appropriate excavation data available. Where excavations are ongoing or could not be completed within the timescales of this project, excavation data will be feedback to the training process and new, more accurate models created. The models trained as part of this project will serve as a starting model for further transfer learning projects on radargram data.

Case studies also form an important part of this research as they serve as the assessment of the viability of these approaches adopted in a "real-world" scenario. Figure 100 illustrates the approximate location of the "real-world" test surveys included in the secondary validation dataset. The GPR data from these surveys were interpreted using the object detection model. Results from both methods of interpretation, manual and automatic, were compared as a secondary validation method and to determine if using an automatic classification method with human confirmation improves the speed and ease of technical report production. Testing the use of the web app and object detection model on data from these surveys will demonstrate whether the machine-led interpretation is more accurate (or quantifiable) and faster than human-led interpretation, and whether the output of the web app could be presented in a format appropriate for a technical report within a reasonable timescale.

8.2. Single Object Detector vs. Experts

The radargrams presented to experts in the GPR data questionnaire (see Appendix D) were also interpreted using the web apps discussed in Chapter 7.5. The outputs from the web app are presented in Figure 96 - Figure 99. The only incorrect detections made by the machine are Quadrant B in Radargram 1 and Quadrants A and C in Radargram 2.

No grave detected. Confidence of detection is 0.9993.　　No grave detected. Confidence of detection is 1.0000.

No grave detected. Confidence of detection is 1.0000.　　No grave detected. Confidence of detection is 0.9998.

Figure 96: Interpretations made by the single object detector on Radargram 1 presented to experts in the questionnaire

The machine did not detect a grave in Quadrant B in Radargram 1, where there should be one and had 100% confidence in the interpretation. This may be a result of the model not being trained on enough examples of this grave type or the response from the grave being too similar to 'non-grave' responses included in the training dataset.

158

No grave detected. Confidence of detection is 0.9975.　　No grave detected. Confidence of detection is 0.9975.

No grave detected. Confidence of detection is 0.9984.　　No grave detected. Confidence of detection is 1.0000.
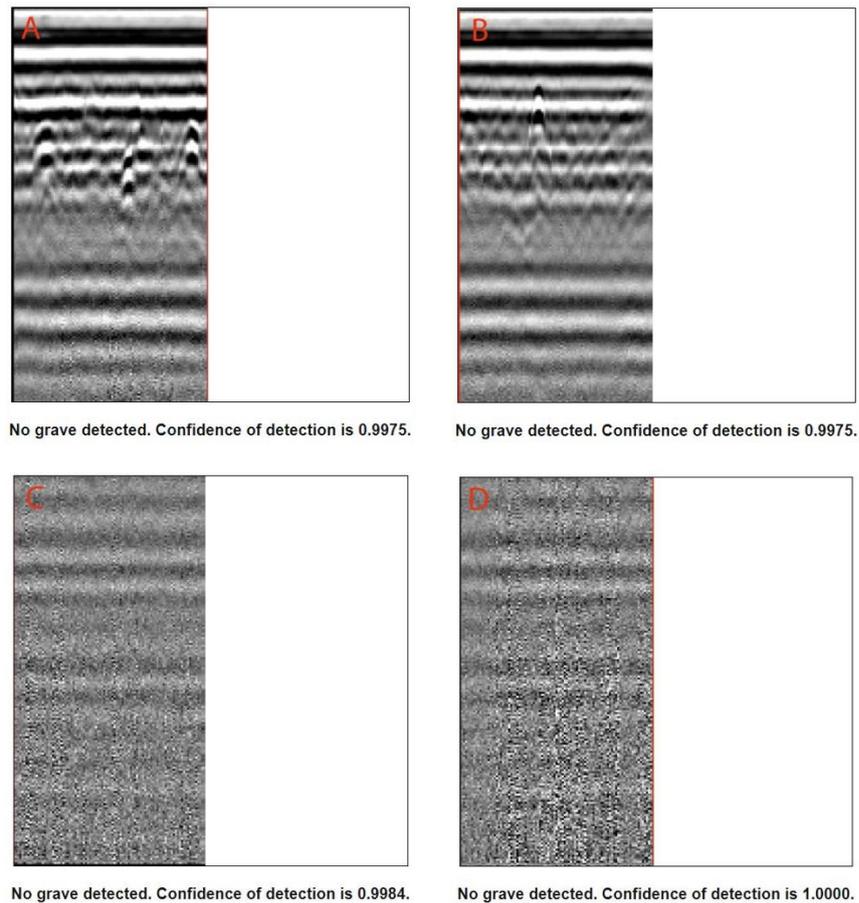
Figure 97: Interpretations made by the single object detector on Radargram 2 presented to experts in the questionnaire

The machine also incorrectly interpreted Quadrants A and C in Radargram 2. Due to the similarity of the responses to that in Quadrant B (Radargram 1), the incorrect interpretation is also likely to be a result of the range of training data supplied to the model. These incorrect interpretations show that machine learning on these types of data still has limitations and cannot be used as a sole method of data interpretation; instead, it should be considered as an initial step that can be carried out quickly. It is therefore essential to maintain a level of human verification of the machine learning results and continue to feedback and re-train the model as more information becomes available. Further monitoring of the prevalence of false negative responses must be carried out by human interpreters, and the level of acceptable false negative responses must be weighed against the increased speed of interpretation as determined by the surveyors. This is further discussed in Chapters 10 and 11 with relation to the impact of these results on the viability of machine learning for commercial geophysics.

No grave detected. Confidence of detection is 0.7172.

No grave detected. Confidence of detection is 0.6772.

No grave detected. Confidence of detection is 0.9909.

No grave detected. Confidence of detection is 0.9870.
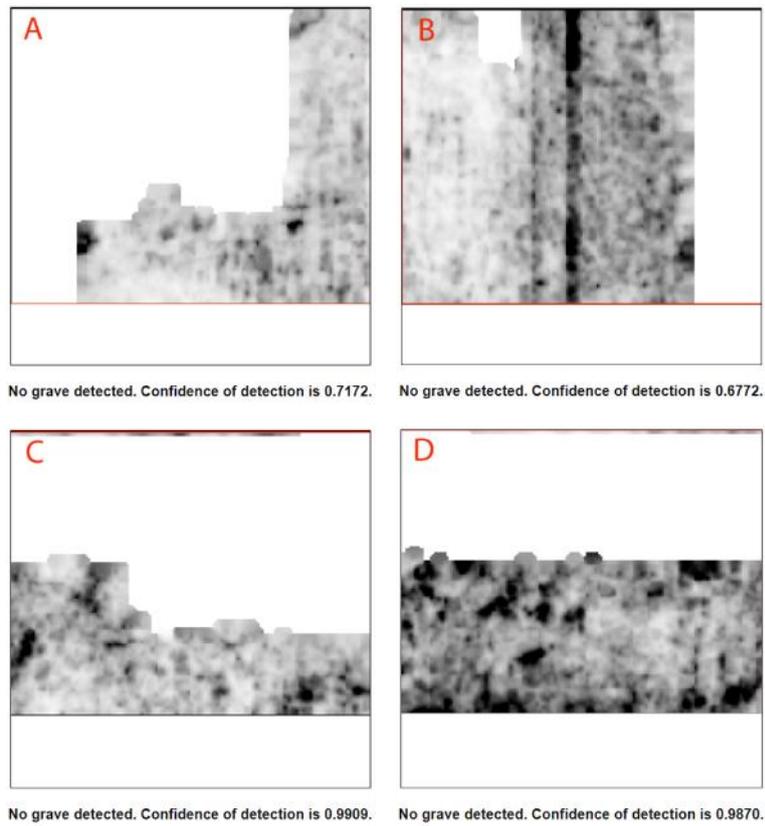
Figure 98: Interpretations made by the single object detector on Timeslice 1 presented to experts in the questionnaire

No grave detected. Confidence of detection is 0.8046.

No grave detected. Confidence of detection is 0.8477.

No grave detected. Confidence of detection is 0.7619.
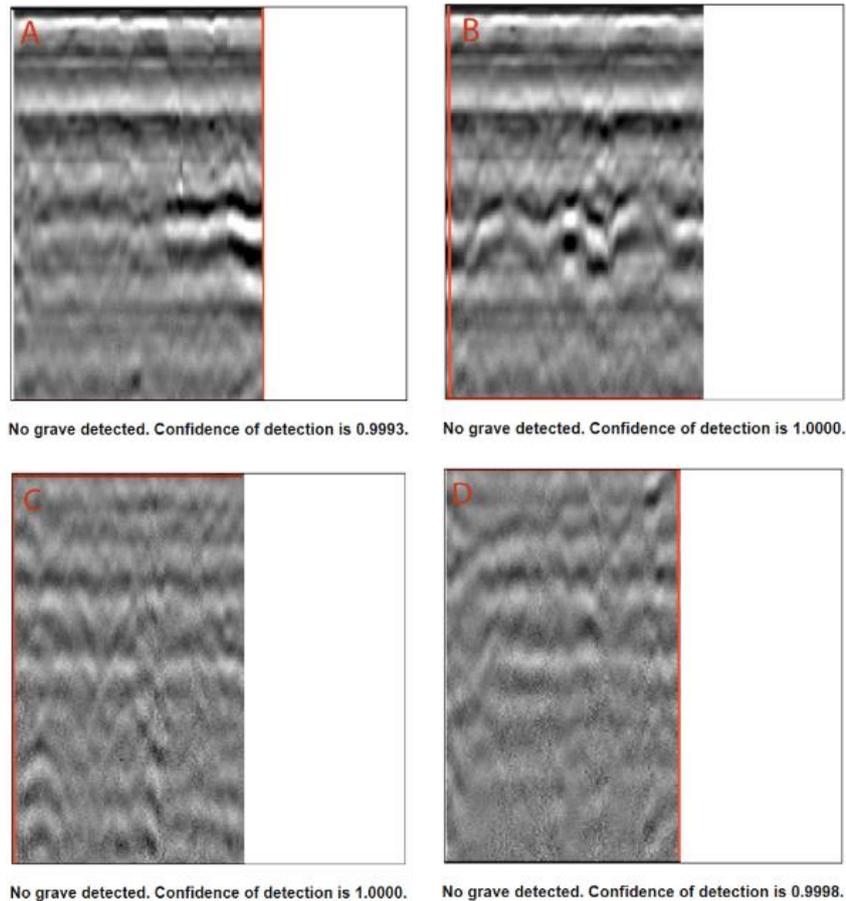
No grave detected. Confidence of detection is 0.9818.
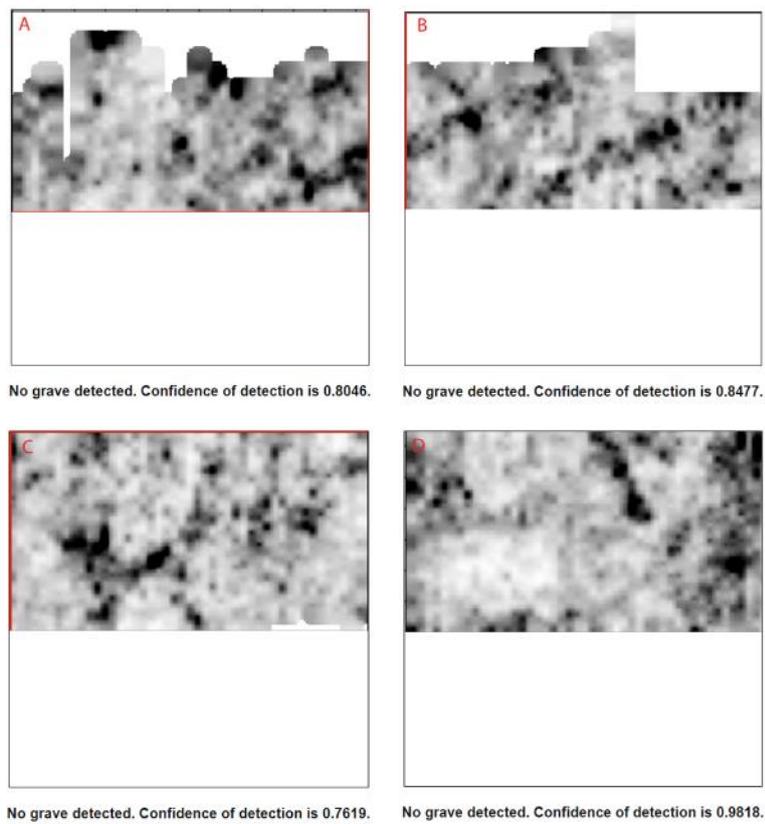
Figure 99: Interpretations made by the single object detector on Timeslice 2 presented to experts in the questionnaire

160

The machine correctly interpreted all quadrants in Timeslice 1 (Figure 98) and Timeslice 2 (Figure 99). This is also reflected in the experts' interpretation of the timeslices, which was more accurate than their interpretation of the radargrams. However, the results may be solely a result of the grave being more clearly defined, in both colour and shape, in timeslices.

Even though the human-led interpretation of the timeslices was more accurate than the radargrams, this should not be an indicator that only timeslices should be used during interpretation. Instead, the timeslice and radargram data for a site should be viewed together for a well-informed interpretation. The results from the interpretation questionnaire (see Appendix D for the full questionnaire) also further demonstrates the need for additional machine learning support in the data interpretation stage of a project.

8.3. Case Study Data

The case study sites were selected based on their accessibility, time scale, suitability for GPR survey, and potential for excavation. The survey of the study sites used a collaborative approach and was facilitated by the directors of the Irish Heritage School (trading as Irish Archaeology Field School) and the Blackfriary Archaeology Field School. As a 'real-world validation,' only sites with presumed medieval activity and known, recorded burials were included.

The five case study areas are located in the Republic of Ireland and constrained to three counties – Offaly, Meath, and Wexford. While the superficial drift geologies of each site are a similar glacial till with silty clay to clay loam soils, the bedrock geologies differ slightly. Each site also has recorded medieval activity, though this varies from excavation evidence to historical evidence between sites.
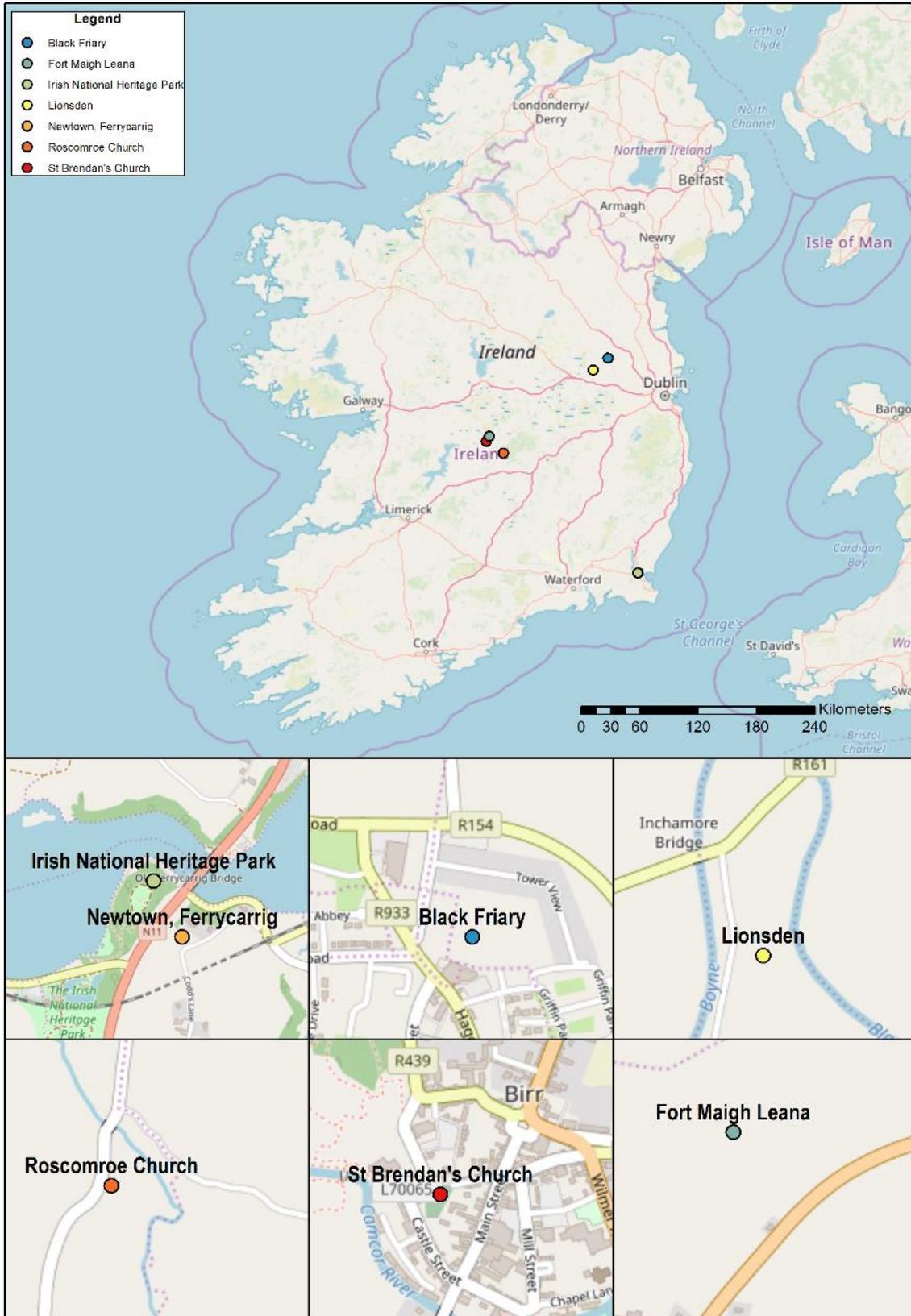
Figure 100: Approximate location of surveyed sites included in the secondary validation dataset

8.4. Survey Details

Full technical reports for each case studies are provided in digital appendices.

GPR survey, accompanied by EMI, earth resistance, and magnetic surveys where necessary to identify areas to target with high-resolution GPR survey, were conducted on known and potential burial sites using the parameters detailed in Table 27. Survey grids were established using an RTK GPS with minimum 0.05m accuracy or a self-tracking total station if necessary and marked with non-permanent markers to allow for the accurate relocation of survey grids.

Table 27: Details of survey parameters to be used in test surveys

Task	Instrument Model	Traverse Interval	Sampling Interval
GPR Survey	MALÅ RAMAC X3M 500MHz antenna	0.25m, 0.5m	0.02m, 0.05m
Resistance Survey	Geoscan Research RM85 Multiplexed 0.25m-1.5m Twin Aligned centres	0.5m	0.5m
Magnetic Survey	Bartington Grad601-2	0.5m, 1m	0.125m
EMI Survey	Geonics EM38B CMD MiniExplorer	0.5m	Minimum 0.25m

Geophysical data were processed in GPR-Slice™, TerraSurveyor™, and Geoplot4™ where appropriate, and georeferenced and plotted in AutoCAD software (following the guidelines outlined in David et al. 2008).

Before interpreting the collected GPR data, all survey data from each site were interpreted manually. Results and interpretation of the GPR data are briefly discussed hereafter as an introduction to the comparison of human- and machine-led interpretations and their feasibility in commercial archaeology. Detailed results from each geophysical survey are provided as technical reports in digital appendices.

8.5. Automatic Data Interpretation

A total of 687 radargrams from the case study sites detailed in Chapter 8 were interpreted using the Reilig for Radargrams web app. Following the workflow detailed in Figure 101, it took approximately 6 hours and 40 minutes to process the 687 radargrams, convert them to a suitable format for the web app, analyse the radargram images in the web app, and save the interpretations. The time spent on the processing and interpretation of data from each site is detailed in Table 28.
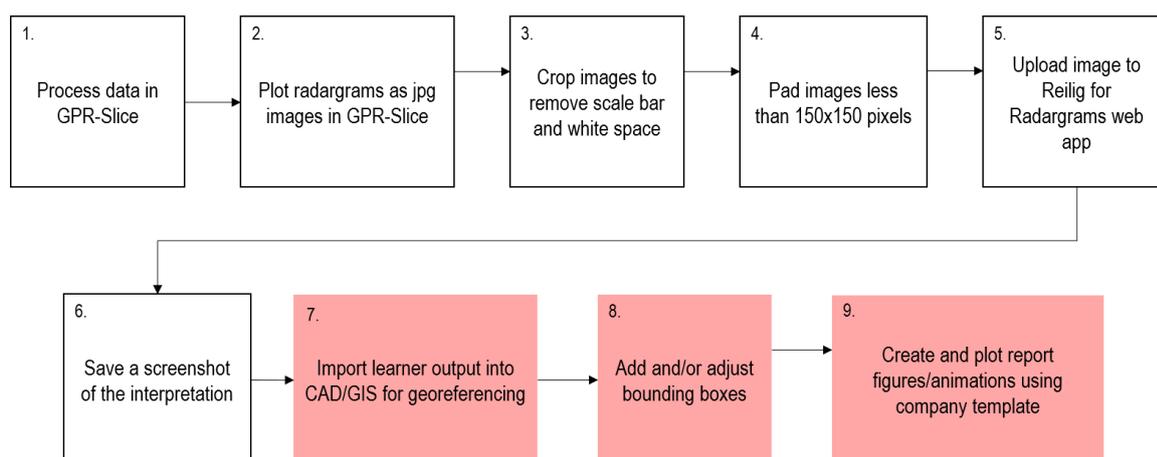
Figure 101: Workflow followed to achieve machine-led interpretation of the GPR data from the seven case study sites (indicated in white). The red hatched steps indicate further steps that could be taken depending on the quality of the output required for a technical report.

The time required to process and save the radargrams was dependent on the length of the traverse. Those sites with radargrams longer than 10m were analysed more slowly by the Reilig app than the shorter radargrams. Some background noise was intentionally left in the radargrams from Lionsden, Newtown, and Roscomroe to assess how the model would handle inadequately processed data. The detections made by the Reilig model for each site and their validity are discussed hereafter. The full set of interpretation outputs from the Reilig web app are provided in the digital appendices.

Table 28: Time allocated to the processing and machine-led interpretation of GPR data from each case study site using the Reilig for Radargrams web app

Site	Number of Radargrams	Time Spent to Achieve Output from Reilig
Black Friary	47	33 minutes
Fort Maigh Leana	12	12 minutes
Irish National Heritage Park	215	106 minutes
Lionsden	200	120 minutes
Newtown, Ferrycarrig	90	49 minutes
Roscomroe Church	26	27 minutes
St Brendan's Church	97	53 minutes
TOTAL	*687*	*400 minutes (6 hours 40 minutes)*

8.6. Fort Maigh Leana

Dr Denis Shine (Director, Irish Heritage School) facilitated and supported a programme of geophysical survey at Fort Maigh Leana, County Offaly. Fort Maigh Lena, also known as Fort Moyleana or Maigh Luaghne, is listed as a bivallate ringfort (O'Brien and Sweetman 1997). However, its dimensions (77m diameter circular landscape enclosed by a double bank and fosse (ditch) approximately 4.5m wide x 3.4m deep) exceed that of typological Irish ringforts. The site, shown in

Figure 102 at the time of survey, is situated in Ballindown, County Offaly, near Clonmacnoise and Birr town, within a rich landscape of medieval and later Georgian activity. The site itself is potentially of regional and national significance as its name, Maigh Leana is derived from the 'Battle of Maigh Leana' that is believed to have taken place near Tullamore in the second century AD. Eugene O'Curry's (1855) translation of the *Cath Maige Leana* manuscript provides understandings of the events of the battle. The text details the battle between Eoin Mor (also Eugene the Great or Mug Nuadat) and the High King 'Conn of the Hundred Battles' as they fought over the two 'halves' of Ireland – *Leith Cuinn* (Conn's half) and *Leth Moga* (Mug's Half) because Eoin erected three mounds around which to gather his armies thereby breaking the peace treaty he had with Conn. Locally, Fort Maigh Leana is thought to be one of the three mounds.

Figure 102: NW facing panoramic view of the survey area at Fort Maigh Leana

There is no excavation evidence from this site as landowners were not able to facilitate test excavations within the timeframe of this project. However, based on the historic evidence for the site, it is unlikely that there are any graves present within the area surveyed. As excavation data becomes available, this will be fed back into the training process to improve the accuracy of detections.

8.6.1. *Manual Interpretation Results*

GPR, EMI, and gradiometer surveys were conducted on the interior of Fort Maigh Leana. The surveys delineated areas of archaeological potential, geological variation, and modern activity. For a detailed interpretation, see the technical report included in the digital appendices.
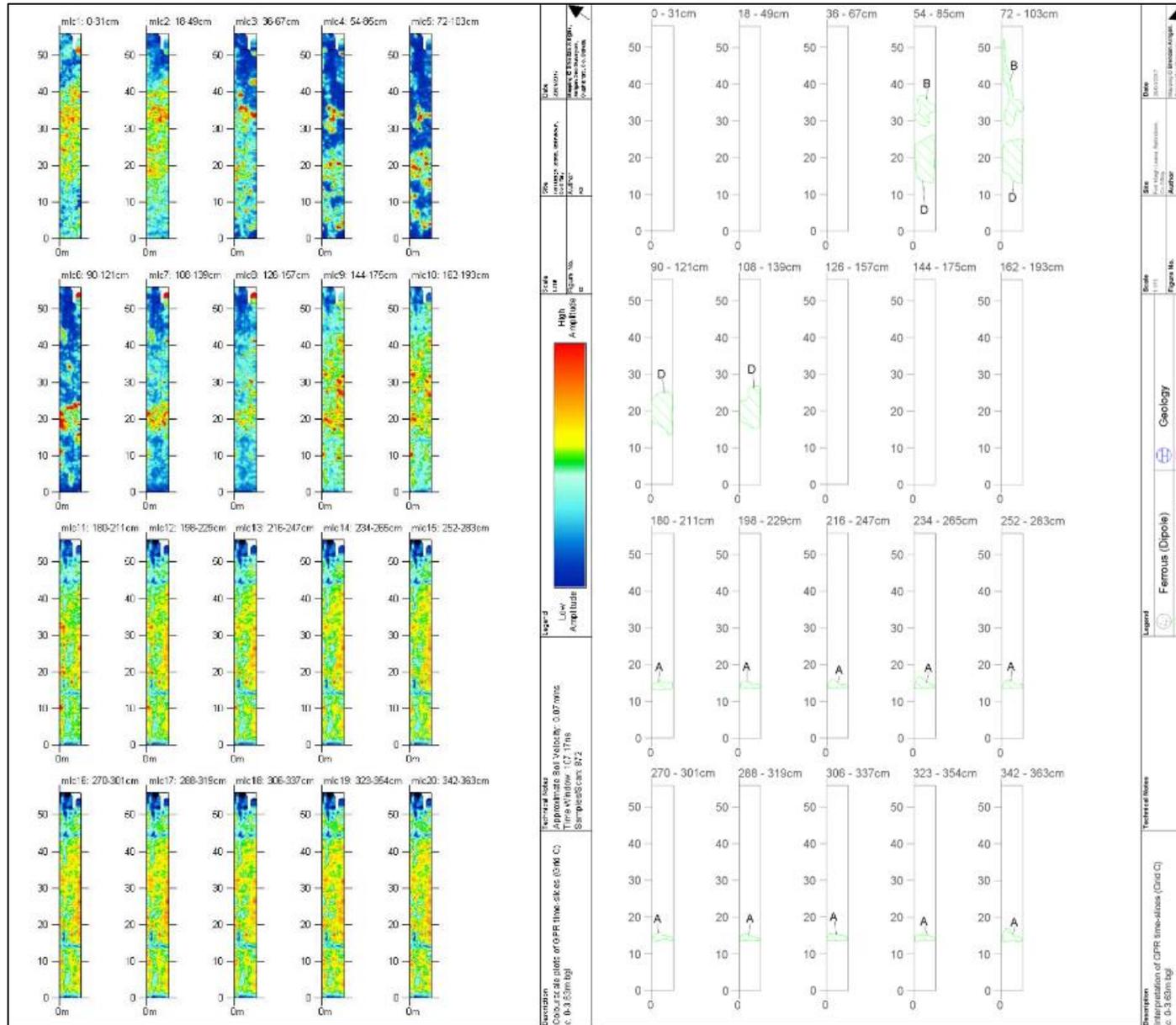
Figure 103: Manual interpretation of all GPR timeslices from Fort Maigh Leana, Ballindown. The survey delineates areas of archaeological potential.

8.6.2. *Object Detection Results*

Of the 12 radargrams which comprise Area C of the Fort Maigh Leana dataset, there were no graves detected. This was the expected result as the data from other geophysical techniques indicated that a souterrain or similar feature was present in Area C and it is unlikely that individuals would have been buried between the top of the souterrain and the original topsoil layer.

8.7. St Brendan's Church

The Birr 2020 group commissioned a programme of geophysical survey at St Brendan's Church and the surrounding area in Birr, County Offaly, Ireland. A GPR survey was conducted within the churchyard at St Brendan's, along the paved pedestrian access route north of the churchyard, and across the 'old pig market' south of the churchyard.

Caimin O'Brien (An Roinn Cultúir Oidhreachta agus Gaeltachta 2018) describes St Brendan's Church as an:

> "Ancient monastic site founded by St. Brendan in the seventh century associated with the illuminated manuscript known as the 'Gospels of Mac Regol' ... Present remains consist of poorly preserved remains of a multiperiod square shaped church (ext. dims. 16.65m N-S; 16.95m E-W; wall T 0.7m) built with roughly coursed limestone with large sandstone quoin stones with 17th-century tower (ext. dims. 6.6m E-W; 6.5m N-S; wall T 1.7m) three storeys high added to the S end of W wall. All floors were wooden carried in the thickness of the wall with evidence of internal wooden stairs."

Test trench evaluations and monitoring were undertaken along Main Street by Michael Tierney (2008) and Daniel Noonan (2003). Both excavations noted human remains which may indicate the original church precinct boundary extended outside the current boundary. The aim of the survey, within the scope of this research project, was to identify geophysical responses from potential unmarked graves and sections of the original precinct boundary that had not already been noted during excavations. Known sections of the precinct boundary are currently located underneath the carpark for local shops.

No excavations were planned for St Brendan's Church as the graveyard is no longer in use but still holds importance within the local community. There are, however, several marked graves visible throughout the graveyard.

8.7.1. *Manual Interpretation Results*

A manual interpretation of the GPR data delineated areas of archaeological potential, geological variation, and modern activity. The interpretation of a selected timeslice is provided in Figure 104. For a detailed interpretation, see the technical report included in the digital appendices.
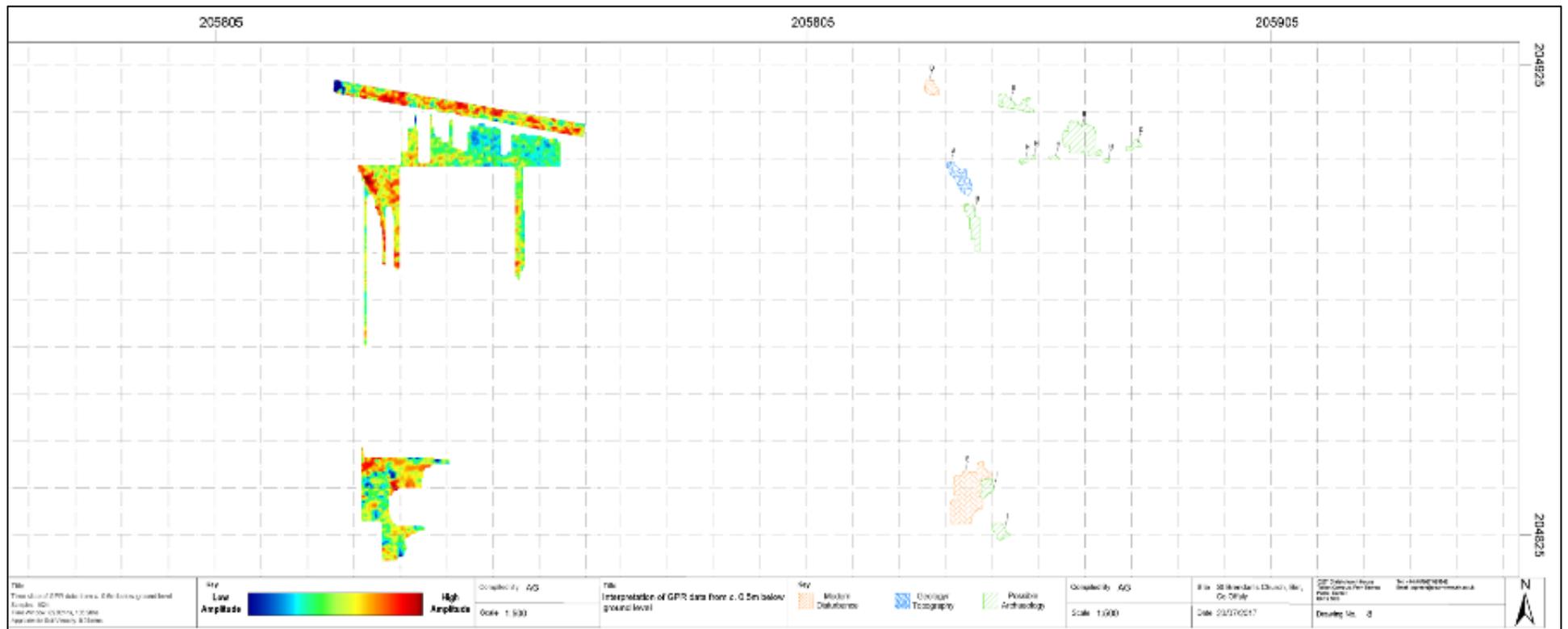
Figure 104: Manual interpretation of selected GPR timeslices from *c.* 0.5m below ground level at St Brendan's Church, Birr. The survey delineates areas of archaeological potential as well as modern anthropogenic activity.

8.7.2. *Object Detection Results*

From the 97 radargrams in the St Brendan's Church dataset, seven radargrams were returned with positive detections for graves as shown in Figure 105.

Figure 105: Approximate outline of the survey area and outline of the radargrams included in the St Brendan's Church dataset indicating whether a grave was detected in the radargram (shown in blue) or no grave was detected in the radargram (shown in grey)

The detections in adjacent traverses align, indicating multiple 'hits' on an east-west aligned grave. An example of this type of detection is provided in Figure 106.

Grave detected. Confidence of detection is 0.9487.   Grave detected. Confidence of detection is 0.9195.
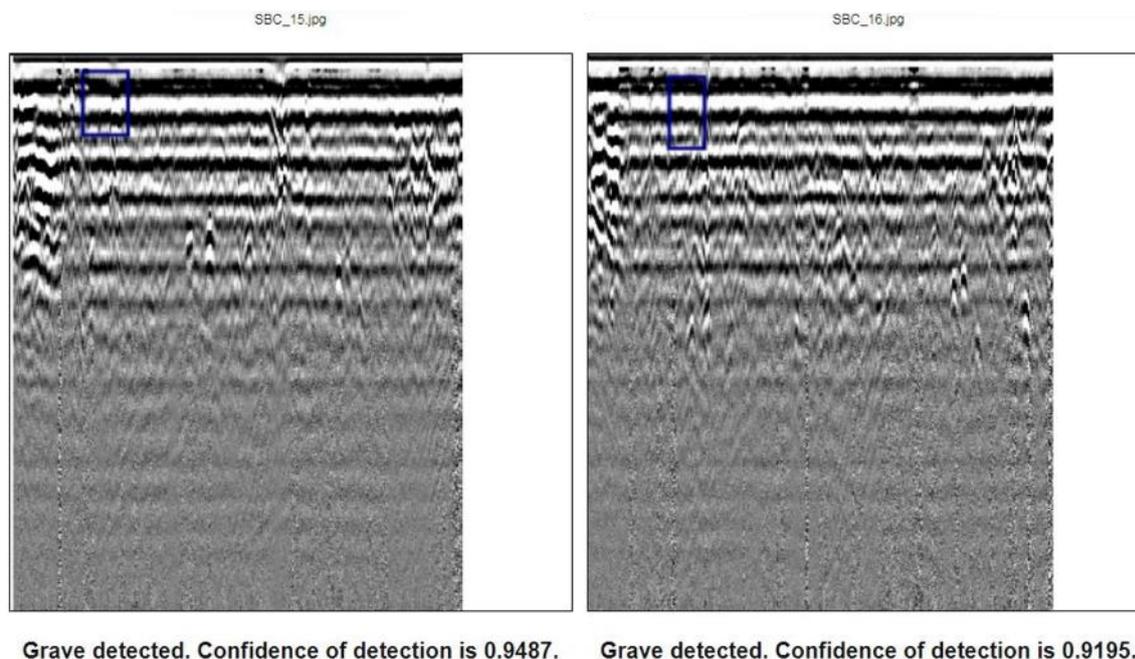
Figure 106: Example of detections aligned in adjacent traverses, increasing the likelihood that these responses relate to an east-west aligned grave

8.8. Roscomroe Church

Amanda Pedlow (Heritage Officer, Offaly County Council) commissioned a programme of geophysical survey at Roscomroe Church, Roscrea, The Leap, County Offaly. This survey aimed to inform on the slippage and potential damage to the exterior church walls.

Roscomroe Church itself is noted in the Archaeological Inventory of County Offaly (O'Brien and Sweetman 1997, p. 109) as a church and graveyard:

> "Situated on a slight rise W of a tributary of the Camcor River. Traditionally cited as an Early Christian foundation attributed to St Molua and to whom a nearby holy well is dedicated. Present remains comprise a late medieval rectangular church (ext. dims. 18.6 E-Q x 7.8m N-S; wall T 0.8m) constructed of roughly coursed shale, standing within an irregularly shaped graveyard enclosed by a stone wall the West gable with partial remains of a bellcote and a breach indicating that former presence of a window, together with the W ends of both the N and S walls survive to full height. Elsewhere no more than wall-footings remain. The splay of a doorway is visible at the W end of the S wall."

No excavations were planned for Roscomroe Church outside of the church exterior, as the graveyard contains modern graves, as well as the possible ?medieval graves, and still holds importance within the local community. There are, however, several marked graves visible throughout the graveyard, particularly to the southeast of the church. In Figure 107, there are several recent modern grave markers visible next to an area of historic and presumed ?medieval stone grave markers.

170

Figure 107: NNE facing shot of Roscomroe Church which demonstrates to the phasing of grave markers and ?grave markers south of the church

8.8.1. *Manual Interpretation Results*

The surveys delineated areas of archaeological potential indicated by a possible enclosure, geological variation, and modern activity, as demonstrated by Figure 108. For a detailed interpretation, see the technical report included in the digital appendices.
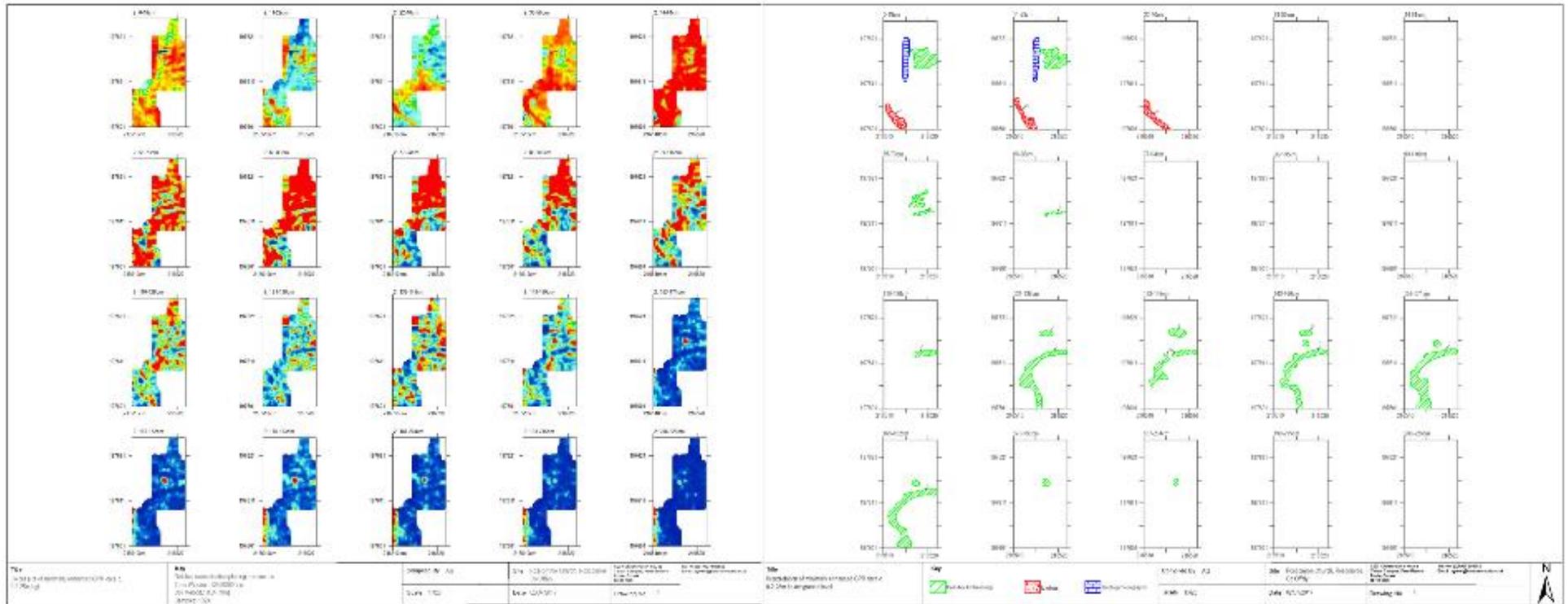
Figure 108: Manual interpretation of all GPR timeslices from Roscomroe Church, Roscomroe. The survey delineates areas of archaeological potential as well as modern anthropogenic activity.

8.8.2. *Object Detection Results*

Of the 26 radargrams included in the Roscomroe Church dataset, six were returned with positive detections for graves, as shown in Figure 109. The detections were of mixed accuracy, where the size and location of several bounding boxes (shown in Figure 110) were highly accurate, and others expressed a similar issue to the INHP example shown in Figure 118.

Figure 109: Outline of the radargrams included in the Roscomroe Church dataset indicating where a grave was detected in the radargram and the confidence of the detection (indicated by the colour scale)

Roscomroe_6.jpg

Grave detected. Confidence of detection is 0.8274.
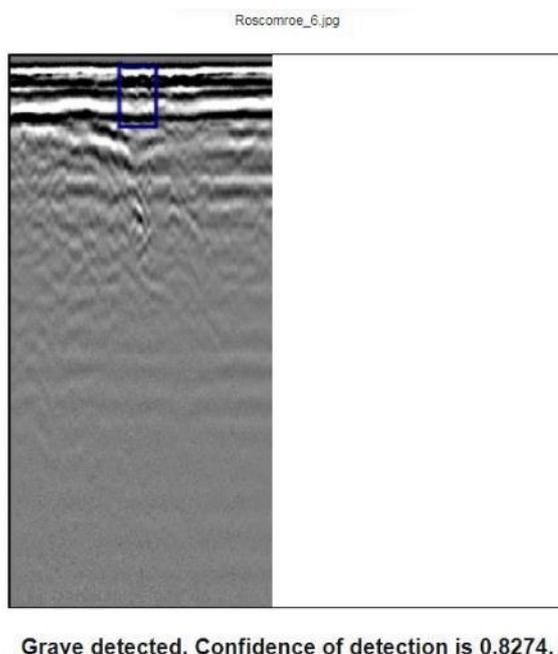
Figure 110: Example of an accurate and high confidence detection in the Roscomroe Church dataset

8.9.   The Black Friary

The Irish Archaeology Field School facilitated and supported a programme of geophysical survey at the Dominican Friary in Trim, County Meath. The Black Friary, a Dominican Friary founded in 1263, is situated in the Boyne Valley landscape in Trim, County Meath, Ireland. After the dissolution of religious orders by Henry VIII in 1536, the Friary was demolished and quarried for limestone. The historic quarrying has produced a thick (*c*. 40-60 cm) rubble layer across most of the site which is overlain by modern dumping. Despite the destruction of the Friary, it continued to hold significance within the community, as evidenced by its continued use as a burial ground throughout the post-medieval period. The Friary is situated in a semi-urban setting outside the northern medieval boundary of Trim town.

Burials are expected in the survey area, shown in Figure 112, as it lies between the nave of the church and the cloister garth. Ongoing excavations on the site have recovered many burials to the south and east of the survey area. These burials are often simple extended inhumations with a loose grave fill surrounded by a compact clay/silty clay layer, sometimes with stone lining, disturbed building materials (ashlar limestone blocks), intercutting other burials, or evidence for coffin burials (O'Carroll, Scott, et al. 2017, O'Carroll, Shine, et al. 2017). In more disturbed areas of the site, disarticulated human remains are more common. It is postulated that the presence of disarticulated remains is a result of burials being disturbed during the robbing out of the friary following the

174

dissolution of religious orders in the 16th century. Excavations have also identified a ditch, likely relating to the southern edge of the friary, indicating the extent of burials within the friary boundaries (O'Carroll, Scott, et al. 2017, O'Carroll, Shine, et al. 2017). To date, excavations have recovered burials Cuttings 1, 2, 3, 5, 6, 7, 8, 9, 10, and 12 as outlined in Figure 111. This excavation evidence can be used to verify the likelihood of automatic detections being of interest and any future analysis of data from the site. Excavation is still ongoing on the site, and as more data becomes available this will be used to inform and retrain a new model.
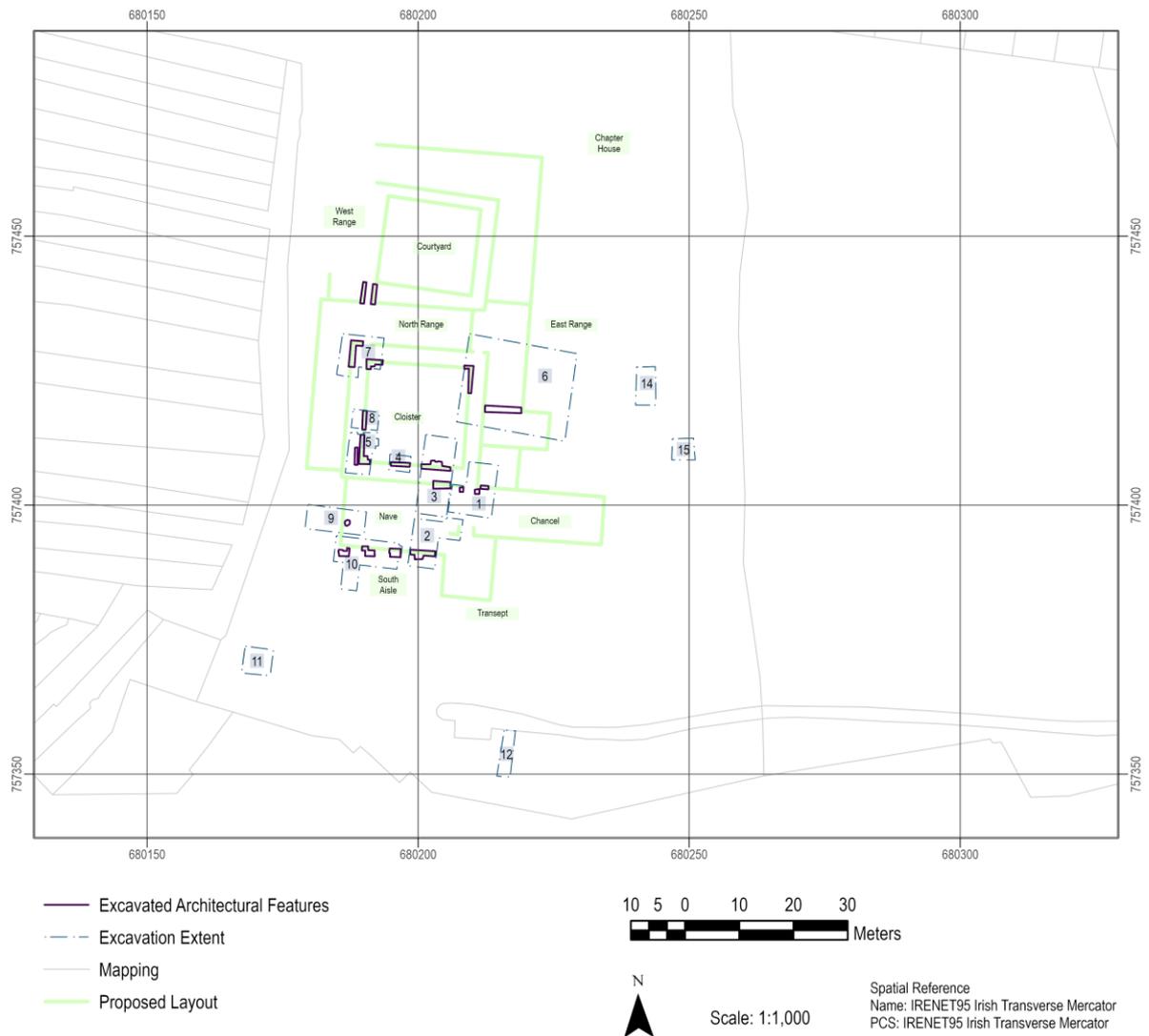
Figure 111: Plan of the excavation areas and proposed extent of the friary buildings as of 2018 (adapted from O'Carroll et al. 2017a, b)

8.9.1. *Manual Interpretation Results*

A series of multi-method geophysical surveys were conducted over three years at the Black Friary. The survey data included in this project are from those surveys of the cloister and within the possible extent of the cemetery. The survey identified potential archaeological features likely relating to the

original monastery structure as well as post-medieval and modern anthropogenic activity. For a detailed interpretation, see the technical report included in the digital appendices.

8.9.2. *Object Detection Results*

From the 47 radargrams in the Black Friary 2017 dataset, 26 radargrams were returned with positive detections for graves, as shown in Figure 112.
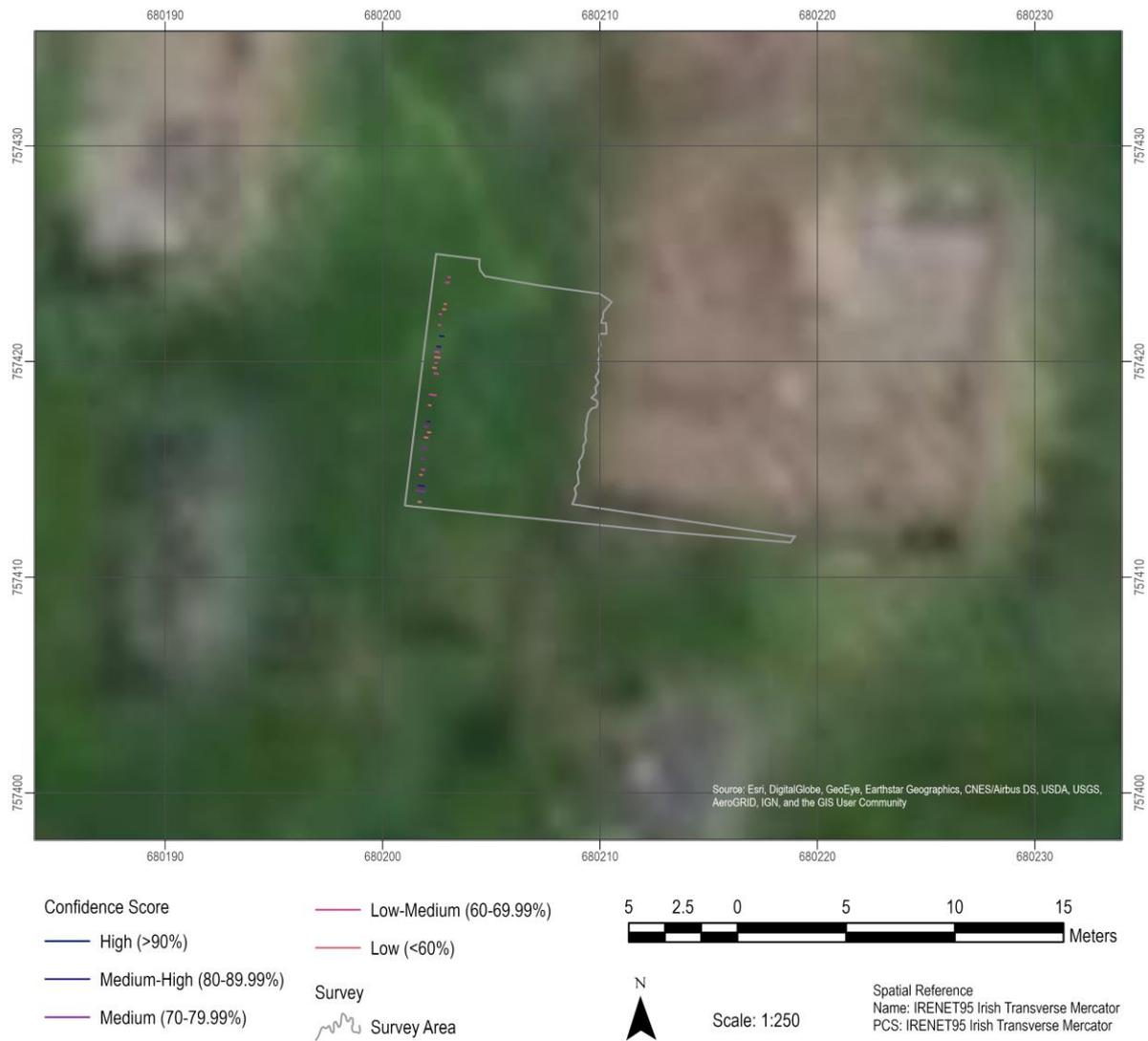
Figure 112: Outline of the radargrams included in the Black Friary 2017 dataset indicating where a grave was detected in the radargram and the confidence of the detection (indicated by the colour scale)

While the grave detections were all reasonably group and spatially aligned (indicative of multiple 'hits' on an east-west aligned grave), the model struggled with detecting an appropriately sized extent of the anomaly. For example, the detection shown in Figure 113 which had a high confidence rate (0.9004), has a bounding box equivalent to less than 0.5m wide. Whereas, an example of a less

confident detection (0.6028), shown in Figure 113, is more similar to the size that is expected for a north-south transect across an east-west aligned grave.
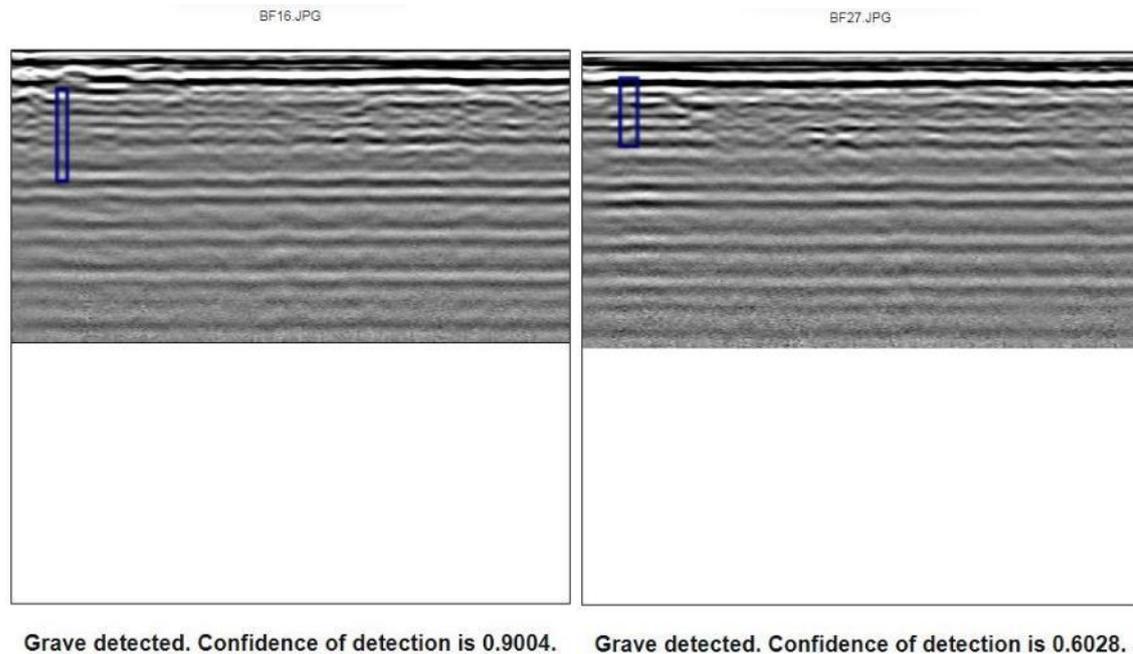
Figure 113: An example of a highly confident grave detection (left) and a less confident grave detection (right) from the Black Friary 2017 dataset

8.10. Lionsden

The Blackfriary Archaeology Field School facilitated and supported a programme of geophysical survey on an area of land *c.* 460m south of the point where the Rivers Boyne and Blackwater meet, in Lionsden, County Meath. Lionsden, County Meath, is located approximately 11.5km southwest of Trim, County Meath, and 27.75km east of Mullingar, County Westmeath.

Human remains were first discovered during digging on the site in 1986. The remains consisted of human lower limbs with a minimum number of individuals (MNI) of three. The burials appear to have been unaccompanied inhumations in unprotected pits. The burials had been extended, aligned west/east, and lay underneath the wall of a historic mill (M. Gunn, pers. comm.). The burials were preserved *in situ* by a layer of concrete.

In 2006, the discovery of additional burials was reported to the National Museum. Heavily disturbed human bone was found throughout the topsoil surrounding the house built on the site. Two perforated scallop shells were also found in the disturbed topsoil, potentially indicating that burials took place here in the medieval period.

There is no current excavation evidence from archaeological excavations for this site as landowners were not able to facilitate test excavations within the timeframe of this project. However, based on the historic evidence for the site, there is potential for graves to be present within the area surveyed, but their extent is unknown. As excavation data becomes available, this will be fed back into the training process to improve the accuracy of detections.

8.10.1. *Manual Interpretation Results*

The surveys delineated areas of archaeological potential and modern services, as demonstrated by Figure 114. Features of archaeological potential that are morphologically similar to structural foundations are indicated in the survey data. For a detailed interpretation, see the technical report included in the digital appendices.

Figure 114: Manual interpretation of a selected GPR timeslice from *c.* 0.5-0.6m below ground level at Lionsden, Co Meath. The survey delineates areas of archaeological potential as well as modern anthropogenic activity.

8.10.2. *Object Detection Results*

The Lionsden dataset is an example of the faults of machine-led interpretation. All 200 radargrams in the dataset returned positive detections for grave-like anomalies. Based on the location of the bounding box within the images, as shown in Figure 115, additional tests were run with various upscaling and downscaling on the images. The bounding boxes were continually in the same location within the image with minor variations in the confidence of the detection. This suggests that the detector is being 'fooled' by the stoniness of the ground surface along the millrace as well as the modern noise along the fence line. As with the INHP example, based on prior knowledge of training machine learning models (discussed in Chapters 6.4, 6.5, and 7), this error may be reduced with the addition of similar examples to the training dataset once they become available. Processing the data with a background removal filter also had no impact on the bounding box detection. The machine should be trained on images that include whitespace, padding, and scale bars as examples of non-graves for the machine to learn not to detect them as graves. This process was not carried out during the training for this project as the addition of these examples of non-graves would skew the balance of the positive and negative classes in the training dataset.

Grave detected. Confidence of detection is 0.9170.   Grave detected. Confidence of detection is 0.9946.
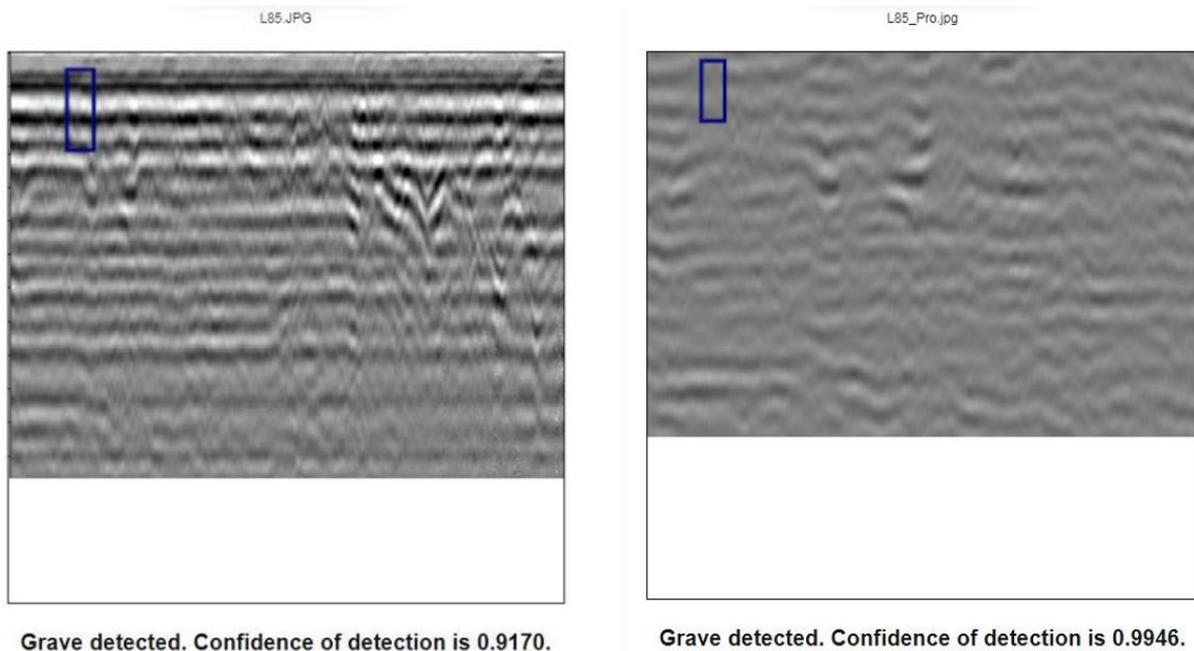
Figure 115: Example of the bounding box region detected on all of the Lionsden radargrams, shown in both the raw data (left) and processed data with background removal (right)

8.11. Irish National Heritage Park

Two programmes of geophysical survey were commissioned by the Irish Archaeology Field School (IAFS), one at the Irish National Heritage Park and one on private land in Newtown, Ferrycarrig. For a detailed interpretation, see the technical report included in the digital appendices.

The Irish National Heritage Park is located in Ferrycarrig, approximately 4km northwest of Wexford, County Wexford, and 17.4km south of Enniscorthy, County Wexford. A programme of geophysical survey was carried out on an area of land surrounding the round tower where IAFS excavations are in progress. The survey area encompasses two monuments – an imitation round tower and a castle ringwork. The imitation round tower (SMR No. WX037-028001-) made from shale was erected in 1857-8 to commemorate the Wexford residents who died during the Crimean War (Bennett 1989). The tower lies on a promontory overlooking the River Slaney (An Roinn Cultúir Oidhreachta agus Gaeltachta 2018). The ringwork (SMR No. WX037-028002-) is an oval area (*c*. 40m E-W; *c*. 27m N-S) defined by a cliff edge and an earthen bank (*c*. 15m wide) with an external fosse which has been recently cleared. Activity surrounding the monument is noted throughout the medieval period:

> "When the Anglo-Normans captured Wexford town in 1169, Dermot Mac Murrough granted it and the surrounding area to Maurice FitzGerald and Robert FitzStephen, and the latter fortified the steep rock at 'Karrech'. On the death of Dermot in May 1171, the Irish rebelled, captured Wexford town and besieged FitzStephen behind the 'earthen rampart' at Carraig. FitzStephen surrendered to trickery and was handed over to King Henry II at Waterford later in the year. It, or perhaps an unlocated castle (WX037-049----) at Carrick became the centre of the manor of Carrick, and a settlement of over a hundred burgesses became associated with it by 1307. However, in 1323-4 the castle was described as vacant and consisting of an unroofed hall and chapel. In 1420 the castle is described as burnt. References to the manor and castle continue into the late 16th century, and in time it came to be known as Shanacourt (Old Castle)" (An Roinn Cultúir Oidhreachta agus Gaeltachta 2018).

The monument (SMR No. WX037-028002-) and surrounding area were excavated in 1984-5 (Bennett 1984). The excavation identified the rock-cut fosse and internal stone revetment of the earthen bank. The interior had been disturbed by the construction of the round tower (WX037-028001-). Further excavation in 1987 recorded evidence (13th century pottery and three silver pennies) for at least two structures (Cotter 1987).

Previous and current excavations at the Irish National Heritage Park and environs are detailed in an edited volume on Carrick, County Wexford (Shine et al. 2019).

8.11.1. *Manual Interpretation Results*

The manual interpretation of the GPR survey data delineated areas of archaeological potential and modern activity, as shown in Figure 116. Features of archaeological potential that are morphologically similar to structural foundations are visible in the survey data. For a detailed interpretation, see the technical report included in the digital appendices.

Figure 116: Manual interpretation of a selected GPR timeslice from *c*. 0.5m below ground level at the Irish National Heritage Park. The survey delineates areas of archaeological potential as well as modern anthropogenic activity.

8.11.2. *Object Detection Results*

Of the 215 radargrams included in the INHP dataset, 29 had detections positive for grave-like anomalies. The location of these 29 radargrams in relation to the negative radargrams is shown in Figure 117.

Figure 117: Approximate location of the survey areas and outline of the radargrams included in the Irish National Heritage Park dataset indicating where a grave was detected in the radargram and the confidence of the detection (indicated by the colour scale)

In the noisier radargrams, like that shown in Figure 118, the location accuracy of the bounding box is poor while the confidence of the detection is still high (0.9930). The radargram was processed with a background removal filter and rerun through the object detector. The bounding box detection was more accurate in the processed data (see Figure 118). Due to this error in the bounding box locations, a guide was written for users of the web app to ensure they process their data where

possible. While an experienced surveyor would be able to locate the true grave-like response, an inexperienced surveyor may have trouble correcting the location of the bounding box if using raw data. It is possible, however, that this bounding box error can also be reduced by increasing the number of 'noisy' radargrams in the training dataset and retraining the model when these radargrams become available. Increasing the number of examples similar to the radargram in Figure 118 will make the model more robust and improve the model's ability to differentiate the 'noise' in the background from the target features.
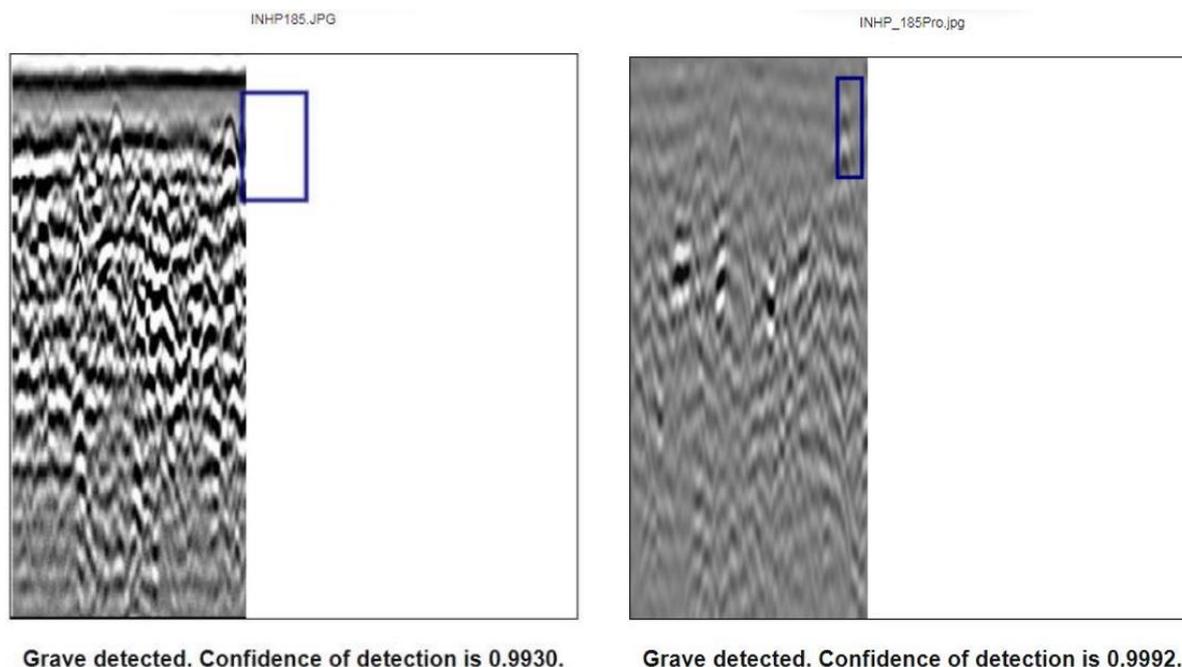
Figure 118: Example of a 'noisy' radargram in the INHP dataset where the accuracy of the location of the bounding box is reduced but the confidence in the detection is still high (left) and the same radargram processed with a background removal filter which shows a more accurate bounding box detection (right)

8.12. Newtown

Two programmes of geophysical survey were commissioned by the Irish Archaeology Field School (IAFS), one at the Irish National Heritage Park and one on private land in Newtown, Ferrycarrig. For a detailed interpretation, see the technical report included in the digital appendices.

Newtown, County Wexford, is located approximately 3.75km northwest of Wexford, County Wexford, and 17.9km south of Enniscorthy, County Wexford. A programme of geophysical survey was carried out on an area of land *c*. 125m south of the junction of the N11 and R730. The survey area comprises the accessible farmland at Tower Hill.

The survey area encompasses the findspot of a pit burial (SMR No. WX037-029). The pit burial is described as the "rim of a collared urn and a small sample of cremated bone, probably representing an older adolescent" (Sikora 2011, An Roinn Cultúir Oidhreachta agus Gaeltachta 2018).

Archaeological testing was carried out on the site in 2006 by Stafford McLoughlin Archaeology (McLoughlin 2006). Five trenches were machine excavated using a 1.8m wide toothless bucket. Four trenches contained definite archaeological features dating to the medieval period, and the fifth contained potential archaeological features (McLoughlin 2006). Of these features, the most notable are a medieval ditch of unknown extent and several deposits containing medieval pottery. It was, however, noted that no prehistoric features relating to the pit burial (SMR No. WX037-029) were identified (McLoughlin 2006).

The site in Newtown was chosen to be part of the testing phase of this project because previous excavations of the site noted the presence of ditches and it was pertinent for this project to determine the extent to which an object detector would be able to differentiate between ditch fill and graves.

8.12.1. *Manual Interpretation Results*

The manual interpretation of the GPR survey data delineated areas of archaeological potential, previous trial trenches, and modern activity, as shown in Figure 119. Features of archaeological potential that are morphologically similar to settlement activity are shown in the survey data. For a detailed interpretation, see the technical report included in the digital appendices.

Figure 119: Manual interpretation of a selected GPR timeslice from *c*. 0.5m below ground level at Newtown, Ferrycarrig. The survey delineates areas of archaeological potential, backfilled archaeological trial trenches, and modern anthropogenic activity.

8.12.2. *Object Detection Results*

From the 90 radargrams in the Newtown dataset, 71 radargrams were returned with positive detections for graves, as shown in Figure 120.
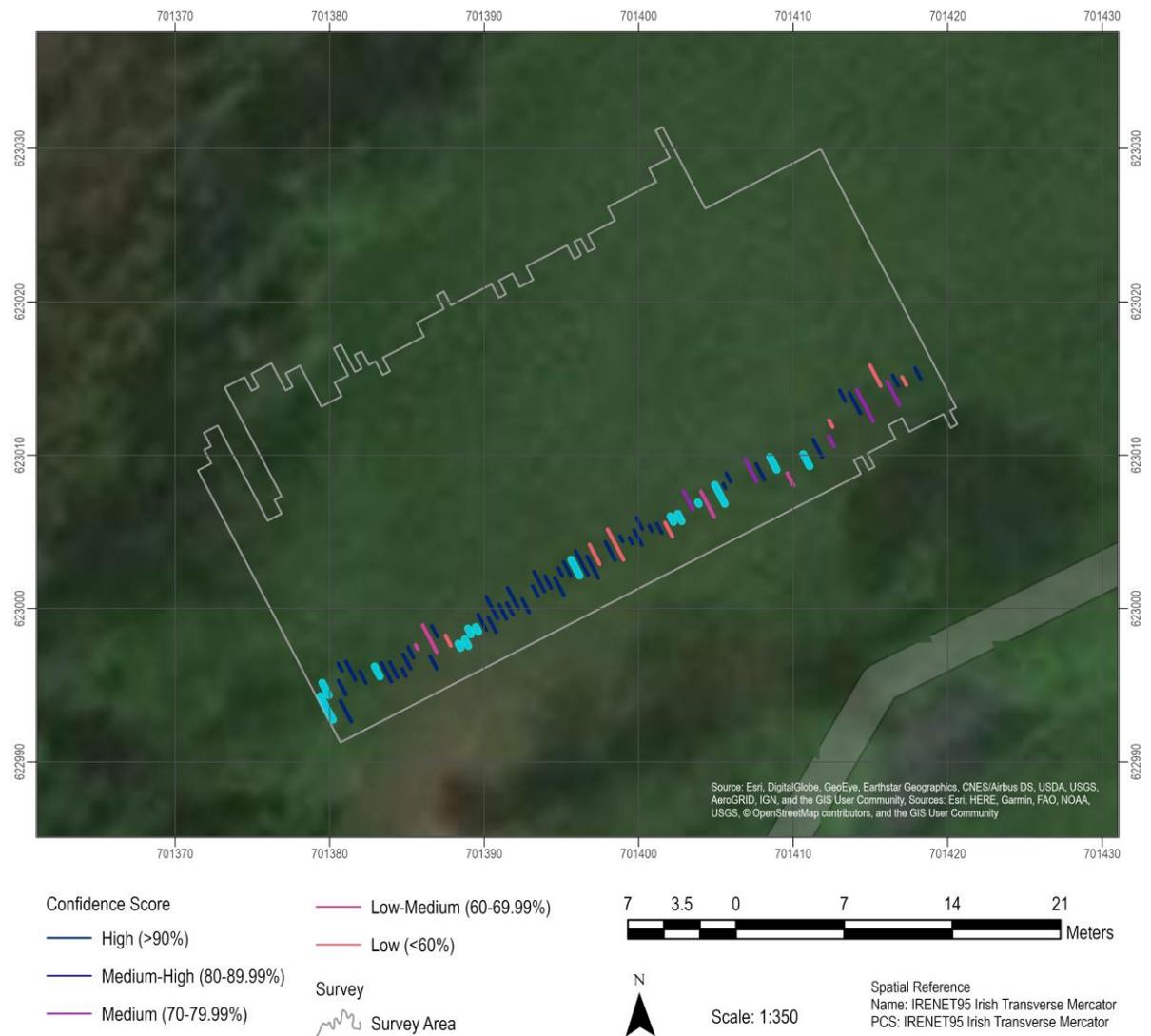
Figure 120: Outline of the radargrams included in the Newtown dataset indicating where a grave was detected in the radargram and the confidence of the detection (indicated by the colour scale)

While a large number of these detections are valid, there are again instances where the machine is 'fooled' by background noise in the radargram or minor DC drift from the GPR. With the training data that is available, it is unlikely that examples of all background variations and instrument fluctuations can be accounted for. Instead, in the instances where it is more likely the background variations that are being detected, not potential graves, the detections will require further validation by the surveyor.

8.12.3. *Summary of Interpretation and Plotting Timescales*

Overall, the speed at which these detections can be made is a marked improvement on the time required for manual interpretation, especially where the surveyor is inexperienced. There are, however, some instances where the detector is unable to differentiate the target response from the general background 'noise.' As the background noise and instrument noise become more prominent, the accuracy and validity of the predictions made by the detector decrease. However, it is likely, based on prior knowledge of machine learning model training, that the prevalence of false negative detections can be reduced with the addition of examples of 'noisy' data and more examples of positive responses to the training dataset once they become available. Even with the need for human validation of all detections, the process of saving the interpretations from the web app reduces the time spent on drawing the dataset and manual interpretations in CAD or a GIS. Taking these factors into account, a machine-led interpretation approach is suitable for commercial and research geophysics and can only become more accurate and reliable with the addition of varied, representative training data.

9. PRELIMINARY TESTS ON ADDITIONAL DATASETS AS AN INDICATOR OF COMMERCIAL VIABILITY

A preliminary test to determine the viability of future applications of the methodology discussed herein to modern clandestine graves was carried out. A pre-trained ResNet152 model was retrained on timeslices across modern clandestine burials proxies, using the same methods presented in Chapter 7. The training dataset consisted of 642 positive examples and 1442 negative examples of woodland burials from Marsh's (2013) survey data. The training data were created from 500MHz GPR data across marked modern earthern dug woodland burials with wicker coffins (Marsh 2013). Raw data files were provided by Marsh (2013) and these data were processed using GPR-Slice and compared with the cemetery layout to identify the known grave responses. The complete training dataset is provided in Appendix G.
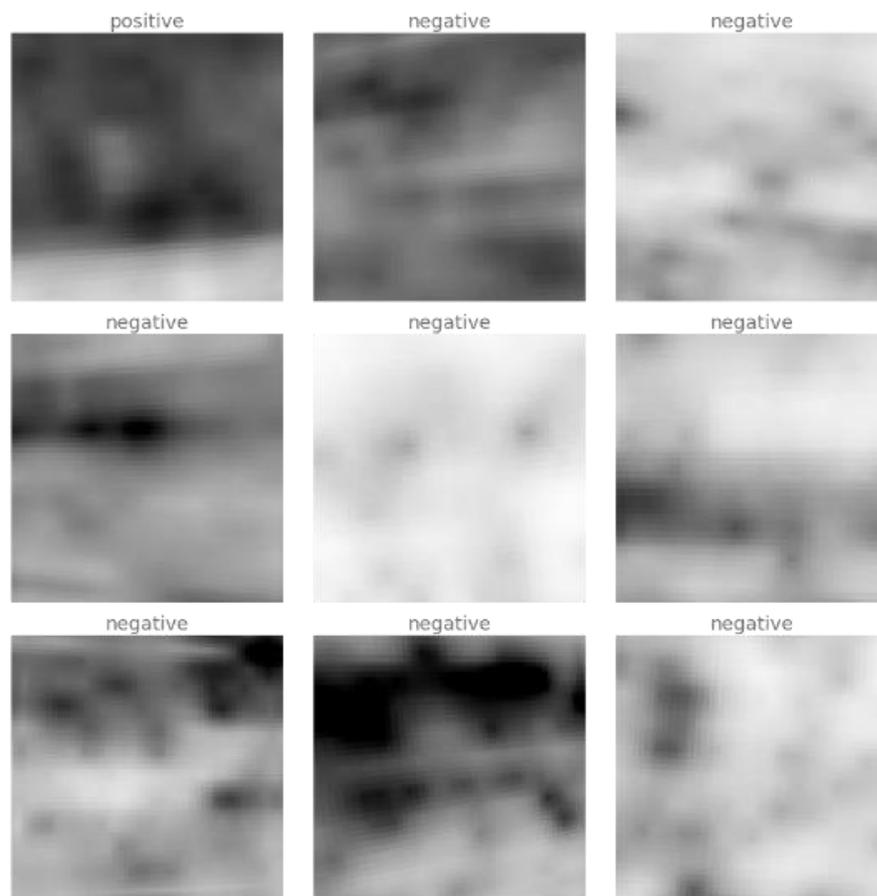
Figure 121: Sample of 150x150 pixel images included in the modern clandestine grave training dataset
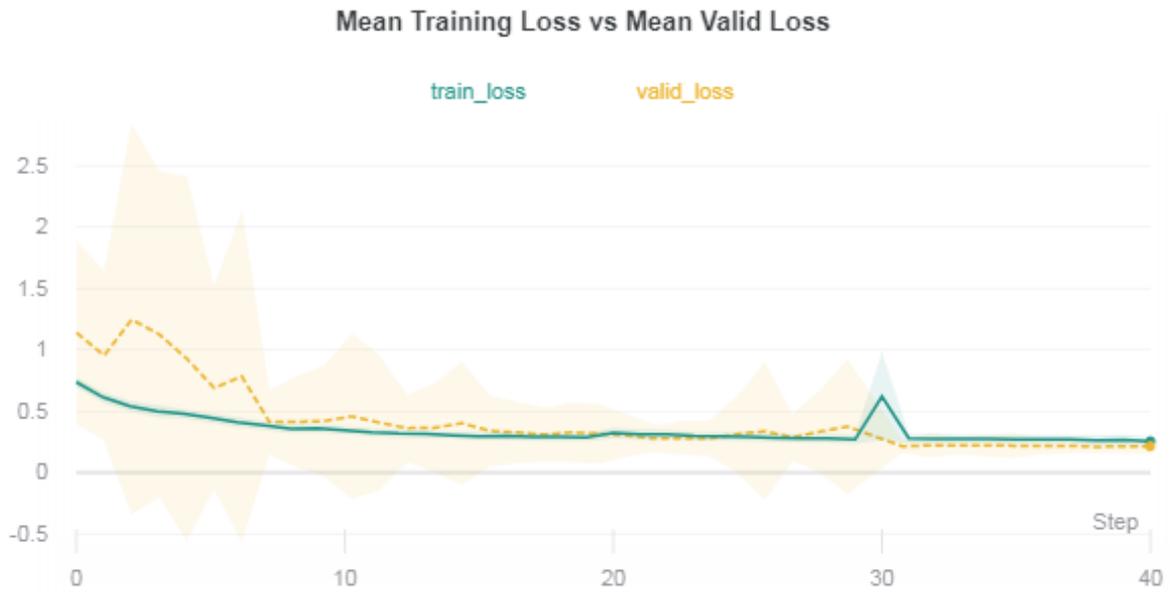
Figure 122: Mean training and validation losses and standard deviation across 100 training sessions for the modern grave dataset. A loss value of 0 indicates a perfect model.
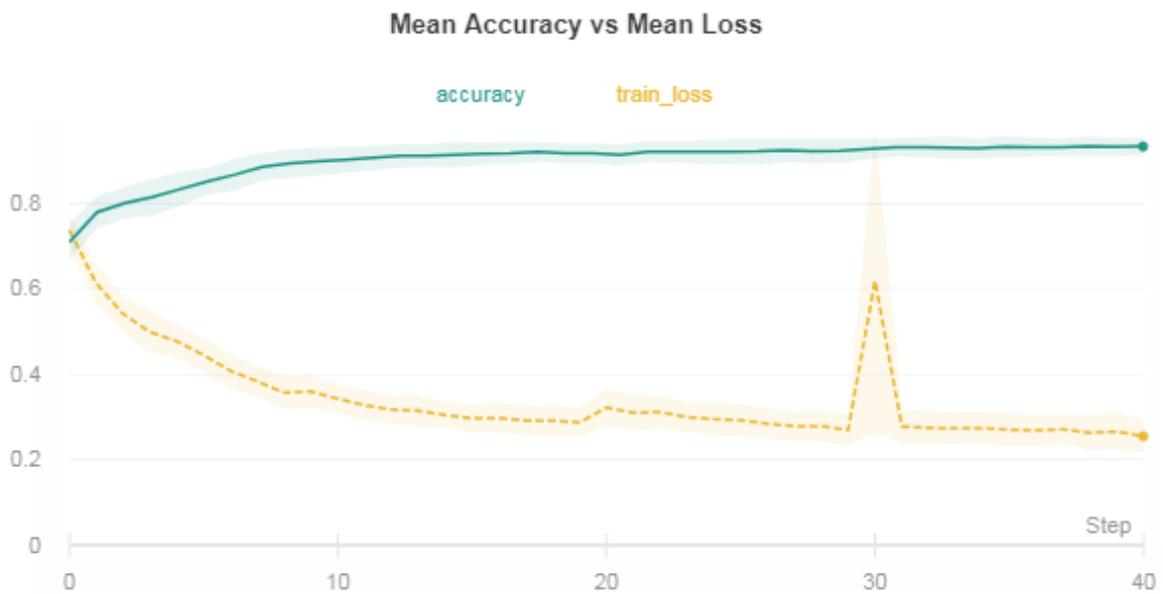
Figure 123: Mean training loss compared with classification accuracy and standard deviation across 100 training sessions for the modern grave dataset
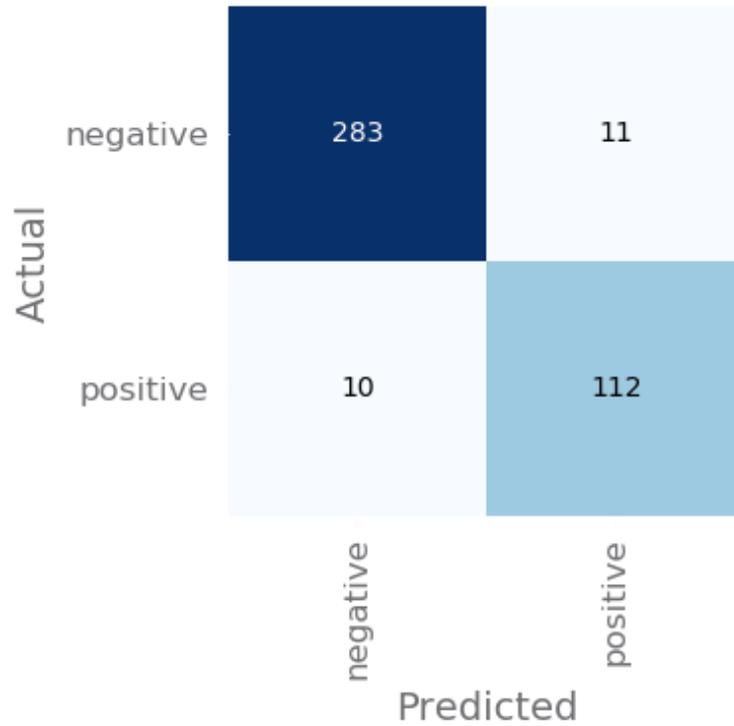
Figure 124: Confusion matrix for the ResNet152 model trained on timeslice data which shows the number of correct predictions and incorrect predictions

The maximum accuracy achieved during training was 94.95%. This high accuracy is most likely a result of the robustness of the training dataset.
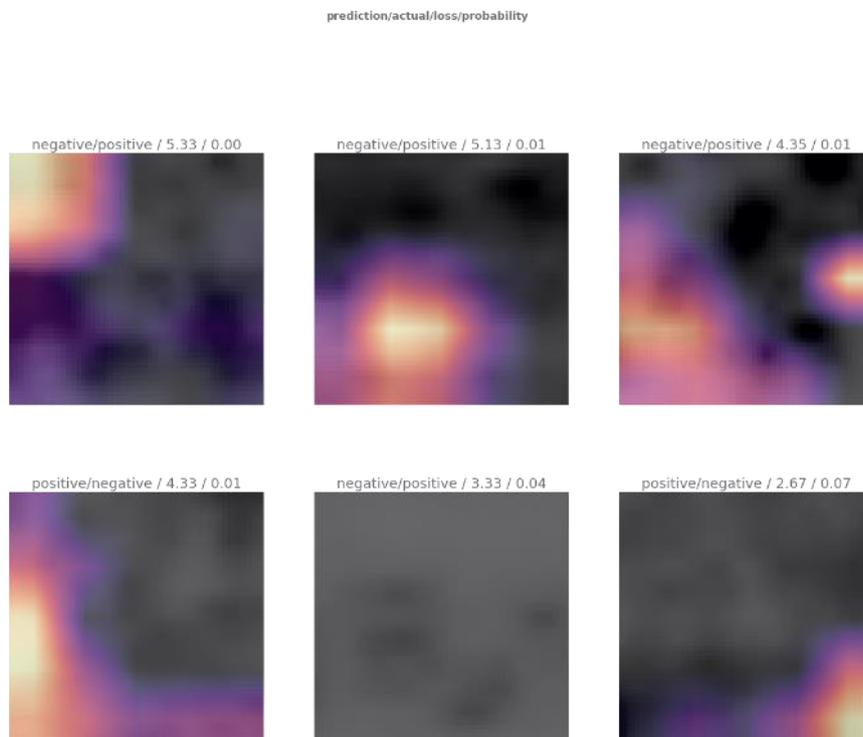
Figure 125: Heatmaps indicating the least accurate predictions made using the ResNet152 model for the modern clandestine grave dataset, where the highlighted area indicates a prediction
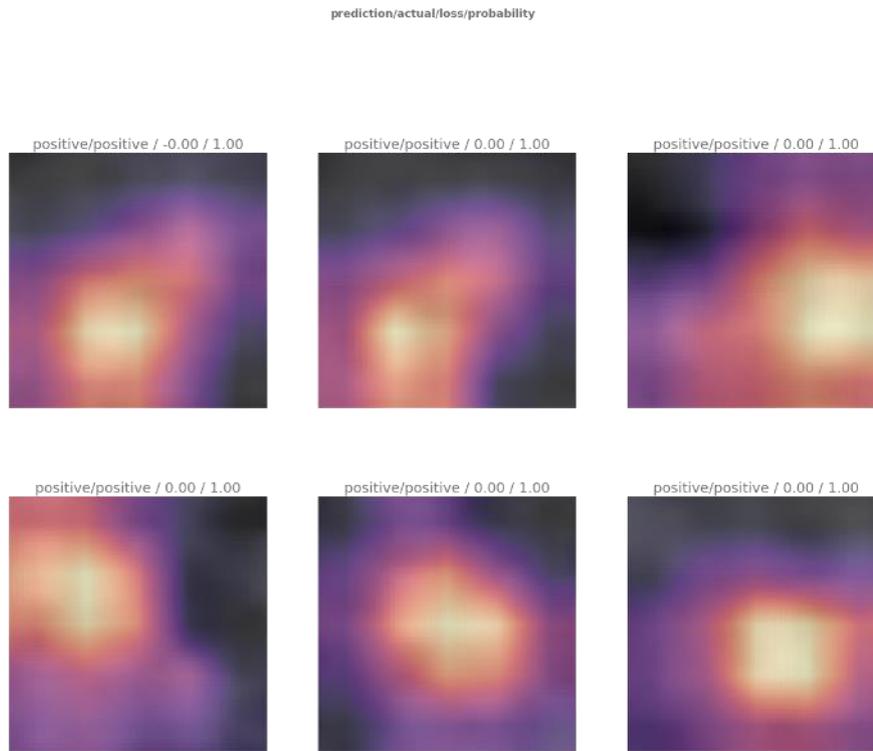
Figure 126: Heatmaps indicating the most accurate predictions made using the ResNet152 model for the modern clandestine grave dataset, where the highlighted area indicates a prediction

While the model was only trained on one type of grave from one site, the high accuracy of the model justifies further investigation of the applicability of machine-led interpretation to modern clandestine graves. Initial results from the woodland burials dataset are promising and demonstrate that, with additional data made available, in addition to radargrams, timeslices can be used for machine-led interpretation and there is potential to train suitable models for the detection of multiple types of clandestine graves.

Using this retrained ResNet152 model, a limited web app, Reilig for Timeslices, was developed to perform the same interpretations as Reilig for Radargrams on greyscale timeslice data. The app is only able to accurately detect cut-earth post-medieval and modern graves until it is trained on additional grave types. The Reilig for Timeslices web app can be accessed at reiligts.onrender.com. The functionality of the app is similar to Reilig for Radargrams, with one upload button and one 'Analyze' button, and returns the interpretations to the end-user on the original image uploaded. Examples of the app making detections on images are provided in the digital appendices.

10. DISCUSSION

The work undertaken in this project examined the feasibility of integrating machine learning and interactive assets into a GPR survey workflow in order to improve the pre-excavation localisation of graves which may impact commercial archaeology projects and the individual graves. The initial step in improving the rate of detection of graves is improving the survey methodologies applied on site. In order to make survey guidance accessible to persons of all skill levels who are wishing to undertake geophysical survey (from volunteer groups to experienced professionals), an interactive format which provides both prescriptive and discursive approaches to guidance will increase the rate at which the correct methodologies are applied to a site.

Following the implementation of appropriate methodologies for a survey, the interactive web apps for object detection demonstrate that the reporting process for GPR surveys can be made more efficient. A new workflow for the interpretation and reporting of GPR data is proposed hereafter, with the aim that the implementation of or integration of parts of this workflow will increase interpretation accuracy and decrease reporting time. While the primary focus of this project was the detection of archaeological graves, given the appropriate training data, the same workflow can also be applied to modern clandestine and mass graves.

10.1. Real-World Applications and Implications

Based on the results of the GPR data interpretation questionnaire, there is a need for an accurate, quantifiable method for the interpretation of GPR data. Part of this improvement in interpretation is the need for the improvement of the quality of data collected during survey so that humans and machines can both maximise their potential to interpret the data accurately.

Currently, the accepted workflow for interpreting and reporting data involves manually drawing polygons around responses of interest to create 2D figures. The research presented here suggests that implementing a new workflow which integrates machine-led interpretation of GPR data would be beneficial to surveyors and commercial clients. Figure 127 compares the current and proposed workflows for interpreting and reporting GPR data, with four minor adjustments to the existing workflow that will improve the speed and accuracy of reporting interpretations of data.

A. Current Workflow
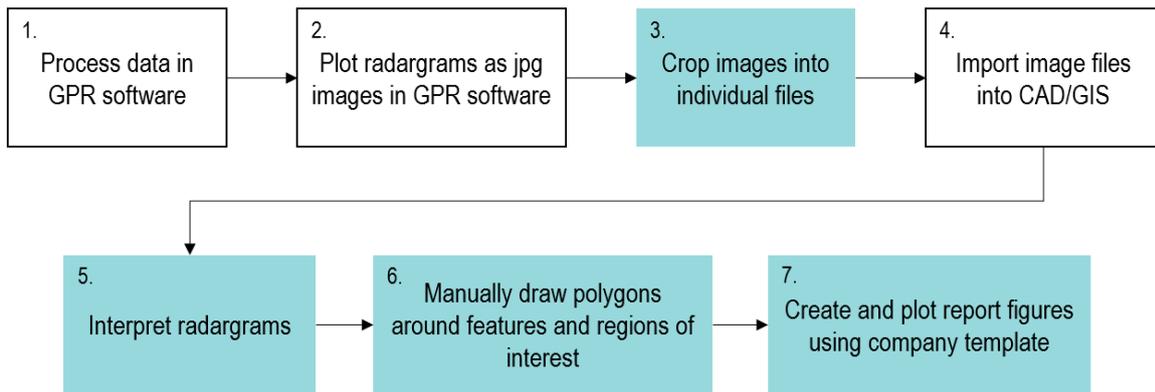
B. Proposed Workflow
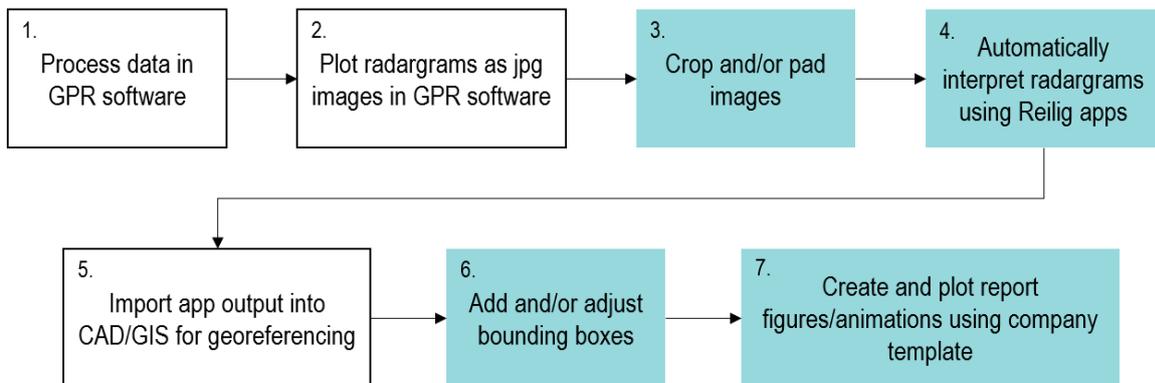
Figure 127: Comparison of the existing (top) and proposed (bottom) workflows for interpreting and reporting GPR. The proposed changes are highlighted in blue.

10.2. Fulfilment of the Research Aim

Digital tools for survey guidance and data interpretation were produced and provided in a format that is accessible to all surveyors, from community groups to commercial archaeologists. All tools are currently available as freely available web apps.

10.3. Fulfilment of the Research Objectives

1. *Assess the appropriateness of current practices for the identification and interpretation of graves in GPR data*

A questionnaire was developed to identify the differences in surveyors' abilities to interpret grave-like responses in GPR data. It aimed to determine if there are differences in the methods of interpretation between individuals and if interpreters focus on the same anomalies in a dataset. It also examined whether they were differences in the interpreters' ability to identify complex and ephemeral responses based on their experiences with GPR.

194

2. Develop an interactive decision-making tool for determining appropriate survey parameters for surveys in Ireland and the UK from existing guidelines

A review of the current survey guidance available identified that they were suitable for collecting interpretable data but often lacked clarity. The guidelines for Ireland and the UK were collated and plotted in decision trees. This format allowed for the consideration of both controllable and uncontrollable factors that can impact on the quality of survey data, such as sampling resolution, geophysical technique, and land cover. The outcomes of this basic decision tree were converted to a field in ArcGIS, plotted in the WGS84 coordinate system to make the results searchable, and provided online in an interactive, user-friendly format.

3. Develop input-output user-friendly automatic classification software for detecting and probability-scoring grave-like responses in GPR data

Machine learning models were developed for classification and object detection tasks. The object detection model provided better results that the classification model so it was hosted online for easy access by the end-user, while the classification model is available as a series of scripts that can be executed through a command prompt.

4. Apply machine learning to a range of sites in Ireland (five study areas)

The final trained object detector was tested on data collected from five datasets from surveys on a range of sites in Ireland to assess the detector's usability and effect on the timescales of a commercial project. It was shown that for a commercial project on a short timescale, knowledge of Python is necessary to batch process and infer on all images at once. For smaller-scale or non-commercial projects, a web app is suitable for inferring on single images.

5. Assess the potential impact of the developed tools on commercial archaeogeophysics

A questionnaire distributed to experienced GPR surveyors was used to assess the potential for inter-observer error in human-led interpretation. Results from the questionnaires demonstrated that there was a wide range of responses given and there was little agreement between participants. Participant's confidence in their interpretations was primarily low. Machine-led interpretation will be able to create a systematic, accurate interpretation workflow that can be implemented in all GPR surveys. The time necessary for machine-led and human-led interpretations were also compared to determine the monetary impact on commercial companies. The time difference between to two interpretation methods was negligible. By implementing a machine-led interpretation workflow companies could be at less risk for recovering unexpected graves.

6.   Test the ability to develop a training dataset for common modern clandestine burial practices

A small sample of modern woodland burials surveyed by Marsh (2013) and simulated clandestine graves held by Bournemouth University was used as a training dataset for modern graves. The ResNet152 model trained on the data achieved a maximum accuracy of 94.95%, indicating that suitable training datasets can be compiled for defined, simple graves with successful implementation in a machine learning model.

7.   Assess the potential impact of the computational tools on geophysics applied in a forensic context

The tests of classification of responses from modern graves were limited but could be successful with additional data. In forensic scenarios, there is an added element of the uncertainty of grave size and orientation that has to be addressed in the training dataset. If a suitable training dataset can be collated the probability associated with inferences could be a vital element in forensic searches and presenting evidence in court cases.

10.4. Project Limitations

The project was primarily limited by the volume and quality of training data available. Radargrams were used in this project because there were enough available so that the models did not require the data to be augmented to increase the size of the training dataset and counteract overfitting. It was hoped that models could also be developed for object detection in timeslices and other visualisation formats of GPR data; however, there was not enough suitable data available to train a model on such complex objects as graves.

11. CONCLUSION

11.1. Outcomes of the Research

The tangible outputs of the project include:

- A web app for determining the most appropriate survey parameters for archaeogeophysical surveys in the UK and Ireland
- A web app for the detection of single graves in radargrams
- A preliminary web app for the detection of single post-medieval – modern graves in timeslices
- A labelled dataset of proxy modern clandestine graves in 150x150 pixel crops of timeslices

The primary tasks of image classification and single object detection assessed within this project have proven to be a useful addition to the GPR data interpretation workflow. The identification of potential graves in GPR data is improved and provided a measure of accuracy. If implemented prior to human-led interpretation, automatic interpretations of data will reduce interpretation and reporting timescales for commercial geophysics.

11.2. Remarks on Future Work

While the results achieved with the existing machine learning models and web apps are promising, there is still additional data that can be added to the training dataset in order to increase and encourage more extensive usage of the tools presented herein.

11.2.1. *Medieval Graves in Western Europe*

As in Ireland, the popularisation of Christianity encouraged the use of west-east extended supine inhumations. Variations between the two countries become apparent when considering the traditional burial practices of population sub-groups (e.g. Jewish cemeteries, Leper cemeteries, Black Death cemeteries). The most common burial practice from the 13th century onwards was a simple earthen extended supine inhumation within an enclosed monastic cemetery; however, burial within these cemeteries required permissions from ecclesiastical authorities (Gilchrist and Sloane 2005). These burials were on average 1.8m L x 0.4m W x 0.4-0.7m D. There is evidence of some stone-lined, wood-lined, cist graves, and chalk-lined burials (Gilchrist and Sloane 2005).

Similar to monastic cemeteries in Ireland, there is a distinct separation of the laity from members of a religious community. In monastic settlements where there was no cloister, the laity was most often buried alongside the church close to the precinct gate; while the religious personnel were buried

alongside the east range and church (Gilchrist and Sloane 2005). Where there was a formal cloister plan, the laity were buried opposite the cloister, such that if the cloister were on the south side of the church, the laity would be buried on the north side of the church the religious community to the south and east, and vice versa for a mirrored cloister plan (Gilchrist and Sloane 2005).

Ornate burials such as stone coffins have been noted within church interiors. Such burials are often indicative of spatial hierarchies, whereby nobility, patrons, benefactors, high-status monks, and, less commonly, ordinary monks were interred with the church. Contrary to Irish burial practices, where there were many inhumations within the cloister garth and ambulatory, burials within the cloister were rare in medieval England (Gilchrist and Sloane 2005).

Geophysical survey is often used to help locate these burials and associated archaeological features before commercial archaeological excavations. Due to their similarity to medieval Irish graves, a machine learning approach should be readily applicable to datasets from Western Europe with the addition of sample data to the existing training dataset.

11.2.2. *Modern Clandestine Graves*

The location and recovery of forensic evidence and clandestine graves is a problem that law enforcement is still facing. Recently, Murray et al. (2018) applied deep learning to the detection of clandestine graves and surface remains in thermal imagery, hyperspectral imagery, and Structure from Motion 3D imagery. An untrained YOLO detector produced promising preliminary results on greyscale thermal imagery (Murray et al. 2018, p. 56-57). The YOLO detector was successful in detecting surface remains as features like the head and body were included in the pre-trained YOLO detector (Murray et al. 2018, p. 56). The authors did not provide information on the accuracy of their model, so it is unclear how successful their preliminary application is. However, they rightfully justify the need for methods to detect clandestine graves accurately and surface remains in surveys of geographically large search areas (Murray et al. 2018, p. 45). The same justification can also be applied to the geophysical survey of such areas.

With the integration of forensic archaeologists into investigations, geophysical methods have been used with increasing frequency to locate clandestine graves and related evidence (Schultz 2007, Ruffell et al. 2009, Dupras 2012). Research case studies by Schultz and others (Schultz 2007, 2008, 2012, Schultz and Martin 2012, Schultz et al. 2012), as well as Pringle et al. (2012, 2016), have demonstrated that clandestine and recent unmarked burials can be detected using geophysical methods. However, it is important to note that law enforcement may not possess the necessary

training to carry out geophysical surveys or interpret the data, and therefore still require forensic archaeologists and anthropologists to assist with investigations.

With the wide variety of controlled research on the geophysical responses of clandestine graves containing human cadaver proxies (i.e., bear, goat, kangaroo, and pig carcasses) or human cadavers, forensic archaeologists can geophysically characterise some of the most common burial scenarios (France et al. 1992, 1997; Strongman 1992; Freeland et al. 2003; Schultz 2008; Schultz and Martin 2011; Pringle et al. 2012; Powell 2003; Molina et al. 2016). Such knowledge and an extensive database of legacy data would make machine learning and automation suitable tools for overcoming issues around law enforcement's inexperience with geophysical survey. To continue with this research project's assessment of the viability of machine learning for improving grave detection, a compilation of a clandestine burial training dataset and retraining of the neural network is necessary. While this application of machine learning could prove crucial to forensic investigations, such examples of modern clandestine burials were excluded from this project in order to narrow the focus of the project and assess the viability of such tools on a simpler, yet still complex, dataset. As clandestine and mass graves are often situational, their geophysical characteristics and signatures can be incredibly varied and will require a large training dataset containing simple and complex representations of clandestine and mass graves at various intervals of time-since-death and in a range of soil types.

11.3. Concluding Remarks

Machine learning proved to be a viable solution to improving the accuracy of interpretations of GPR data, especially concerning ephemeral features like graves. The models are reasonably accurate, but outputs still require human confirmation. There are still improvements to be made in the usability of such machine learning approaches and reduction of human confirmation; however, these improvements cannot be made until a larger, more robust dataset becomes available. Object detection should be implemented where data are difficult to interpret or the survey covers an area where graves are likely to be present (as determined by HER data). However, the first step in improving the interpretability of data and identification of graves before excavation is implementing the most appropriate parameters during data collections. The existing guidance for geophysical survey provides appropriate national and international guidance. Implementing this guidance in an interactive digital format will increase engagement with these resources and improve the data that is collected. As archaeogeophysics moves towards digital applications and novel approaches to data collection and interpretation with the growth of landscape and high-density surveys, tools like

machine learning will become the preferred data interpretation aid. This work forms a strong base for justifying the use of object detectors in geophysics data interpretation as this work can only advance with the availability of larger suitable training datasets.

12. REFERENCES

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., Zheng, X., Brain, G., Osdi, I., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., Kudlur, M., Levenberg, J., Monga, R., Moore, S., Murray, D. G., Steiner, B., Tucker, P., Vasudevan, V., Warden, P., Wicke, M., Yu, Y., and Zheng, X., 2016. TensorFlow: A system for large-scale machine learning. *In*: *12th USENIX conference on Operating Systems Design and Implementation*. Savannah, GA: USENIX Association.

Advisory Panel on the Archaeology of Burials in England, 2017. *Guidance for Best Practice for the Treatment of Human Remains Excavated from Christian Burial Grounds in England* [online]. Available from: http://www.archaeologyuk.org/apabe/pdf/APABE_ToHREfCBG_FINAL_WEB.pdf.

Al-shabeeb, A. R., 2016. The Use of AHP within GIS in Selecting Potential Sites for Water Harvesting Sites in the Azraq Basin—Jordan. *Journal of Geographic Information System*.

Alemi, A., 2016. *Research Blog: Improving inception and image classification in TensorFlow* [online]. Available from: https://research.googleblog.com/2016/08/improving-inception-and-image.html [Accessed 10 Nov 2017].

An Roinn Cultúir Oidhreachta agus Gaeltachta, 2018. *Historic Environment Viewer* [online]. Available from: http://webgis.archaeology.ie/historicenvironment/ [Accessed 4 Jan 2019].

Armstrong, K. and Cheetham, P., 2008. Archaeological Geophysical Prospection in Peatland Environments: Locating the Sweet Track at Canada Farm, Shapwick Heath (Somerset). *In*: *EIGG 8th Meeting on Recent Work in Archaeological Geophysics, 16 December 2008*. London: EIGG.

Asadi, P., Gindy, M., and Alvarez, M., 2019. A Machine Learning Based Approach for Automatic Rebar Detection and Quantification of Deterioration in Concrete Bridge Deck Ground Penetrating Radar B-scan Images. *KSCE Journal of Civil Engineering*.

Aspinall, A., Gaffney, C. F., and Schmidt, A., 2008. *Magnetometry for Archaeologists*. Plymouth: AltaMira Press.

BABAO Working-group for ethics and practice, 2010. *British Association of Biological Anthropology and Osteoarchaeology - Code of Practice* [online]. Available from: https://babao.org.uk/assets/Uploads-to-Web/code-of-practice.pdf.

Ball, J., Anderson, D., and Chan, C. S., 2017. Comprehensive Survey of Deep Learning in Remote Sensing: Theories, tools, and challenges for the community. *Journal of Applied Remote Sensing* [online], 11 (4). Available from: https://www.spiedigitallibrary.org/journals/journal-of-applied-remote-sensing/volume-11/issue-04/042609/Comprehensive-survey-of-deep-learning-in-remote-sensing--theories/10.1117/1.JRS.11.042609.full?SSO=1.

Barry, T., 2003. *Medieval Ireland: An archaeology, 1100-1600*. The English Historical Review.

Bartlett, T. and Jeffery, K., 1997. *A Military History of Ireland*. Cambridge: Cambridge University Press.

Bennett, I., 1984. Preliminary Archaeological Excavations at Ferrycarrig Ringwork, Newtown Td., Co. Wexford. *Journal of the Wexford Historical Society*, 10, 25–43.

Bennett, I., 1989. The Crimean War Memorial, Ferrycarrig, Co. Wexford - A precisely dated round

tower. *Archaeology Ireland*, 3 (2), 58–60.

Bevan, B. W., 1983. Electromagnetics for Mapping Buried Earth Features. *Journal of Field Archaeology* [online], 10 (1), 47–54. Available from: http://www.jstor.org/stable/529747%5Cnhttp://www.tandfonline.com/doi/full/10.1179/0093469 83791504354.

Bevan, B. W., 1991. The Search for Graves. *GEOPHYSICS* [online], 56 (9), 1310–1319. Available from: http://library.seg.org/doi/10.1190/1.1443152.

Bevan, B. W. and Kenyon, J., 1975. Ground-Penetrating Radar for Historical Archaeology. *MASCA Newsletter*, 11 (2), 2–7.

Bhreathnach, E., 2010. From Fert(ae) to Relic: Mapping death in early sources. *In*: Corlett, C. and Potterton, M., eds. *Death and Burial in Early Medieval Ireland: In the light of recent archaeological excavations*. Bray: Wordwell, 316.

Bigman, D. P., 2012. The Use of Electromagnetic Induction in Locating Graves and Mapping Cemeteries: An example from Native North America. *Archaeological Prospection* [online], 19 (1), 31–39. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/arp.1416.

De Boer, A., 2007. Using Pattern Recognition to Search LiDAR Data for Archaeological Sites. *In*: *The World is in your Eyes: Proceedings of the XXXIII Computer Applications and Quantitative Methods in Archaeology Conference* [online]. Tomar: CAA Portugal, 245–254. Available from: https://proceedings.caaconference.org/files/2005/36_Boer_CAA_2005.pdf.

Bonsall, J., 2014. A Reappraisal of Archaeological Geophysical Surveys on Irish Road Corridors 2001-2010. [online]. University of Bradford. Available from: file:///C:/Users/Student/Downloads/Bonsall Thesis Volume 2 of 2 (5).pdf.

Bonsall, J., Fry, R., Gaffney, C., Armit, I., Beck, A., and Gaffney, V., 2013. Assessment of the CMD Mini-Explorer, a New Low-frequency Multi-coil Electromagnetic Device, for Archaeological Investigations. *Archaeological Prospection* [online], 20 (3), 219–231. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/arp.1458.

Bonsall, J., Gaffney, C., and Armit, I., 2013. Back and Forth: Paving the way forward by assessing 10 years of geophysical surveys on Irish road schemes. *In*: Kelly, B., Roycroft, N., and Stanley, M., eds. *Futures and Pasts: Archaeological Science on Irish Road Schemes* [online]. 158. Available from: https://wordwellbooks.com/index.php?route=product/product&product_id=1571 [Accessed 10 Nov 2017].

Bonsall, J., Gaffney, C., and Armit, I., 2014. *Preparing for the Future: A reappraisal of archaeo-geophysical surveying on Irish National Road Schemes 2001-2010* [online]. Dublin: Transport Infrastructure Ireland. Available from: http://www.tii.ie/technical-services/archaeology/resources/NRA-Guidelines-Bonsall-et-al-2014-Preparing-for-the-Future.pdf [Accessed 10 Nov 2017].

Boon, P., Maaten, L. van Der, Paijmans, H., Postma, E., and Lange, G., 2009. Digital Support for Archaeology. *Interdisciplinary Science Reviews* [online], 34. Available from: https://www.tandfonline.com/doi/abs/10.1179/174327909X441108.

Bournemouth University, 2019. *New Forest Archaeological Metal Detecting and Geophysics Survey Report: Turf Hill Inclosure*. Bournemouth.

Breiner, S., 1981. Magnetometers for Geophysical Prospection. *In*: Weinstock, H. and Overton, W. C., eds. *SQUID Applications to Geophysics* [online]. Tulsa, OK: Society of Exploration

Geophysicists, 3–12. Available from: http://library.seg.org/doi/book/10.1190/1.9781560802518 [Accessed 10 Nov 2017].

British Geological Survey, 2019. *BGS Geology 625k | DiGMapGB-625* [online]. Available from: https://www.bgs.ac.uk/products/digitalmaps/DiGMapGB_625.html [Accessed 10 Jan 2018].

Cahill, M. and Sikora, M., 2011. *Breaking Ground, Finding Graves: Reports on the excavation of burials by the National Museum of Ireland, 1927-2006*. Bray: Wordwell Ltd., in association with the National Museum of Ireland.

Calder, M. and Kennedy, D. M., 2013. The Application of Ground Penetrating Radar in Delineating Shore Platform Morphology: A case study from Wellington, New Zealand. *Journal of Coastal Research* [online], 291, 226–234. Available from: http://www.bioone.org/doi/abs/10.2112/JCOASTRES-D-13-00004.1.

Cerrillo-Cuenca, E., 2017. An Approach to the Automatic Surveying of Prehistoric Barrows through LiDAR. *Quaternary International* [online], 435 (Part B), 135–145. Available from: https://www.sciencedirect.com/science/article/pii/S1040618216002159.

Chandio, I. A., Matori, A. N. B., WanYusof, K. B., Talpur, M. A. H., Balogun, A. L., and Lawal, D. U., 2013. GIS-based Analytic Hierarchy Process as a Multicriteria Decision Analysis Instrument: A review. *Arabian Journal of Geosciences* [online]. Available from: https://link.springer.com/article/10.1007/s12517-012-0568-8.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., and Wirth, R., 2000. *CRISP-DM 1.0 - Step-by-Step Data Mining Guide* [online]. Available from: https://the-modeling-agency.com/crisp-dm.pdf.

Chaudhary, P., Chhetri, S. K., Joshi, K. M., Shrestha, B. M., and Kayastha, P., 2016. Application of an Analytic Hierarchy Process (AHP) in the GIS Interface for Suitable Fire Site Selection: A case study from Kathmandu Metropolitan City, Nepal. *Socio-Economic Planning Sciences* [online]. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0038012115000452.

Cheetham, P., 2005. Forensic Geophysical Survey. *In*: Hunter, J. and Cox, M., eds. *Forensic Archaeology* [online]. Abingdon-on-Thames: Routledge, 62–95. Available from: https://www.taylorfrancis.com/books/9780203970300 [Accessed 10 Nov 2017].

Chen, T., Li, M., Cmu, U. W., Li, Y., Lin, M., Wang, N., Wang, M., Xu, B., Zhang, C., Zhang, Z., and Alberta, U., 2015. MXNet: A flexible and efficient machine learning library for heterogeneous distributed systems. *eprint arXiv:1512.01274* [online]. Available from: https://arxiv.org/abs/1512.01274.

Chollet, F., 2015. Keras: The Python deep learning library. *keras.io*.

Conyers, L. B., 2012. *Interpreting Ground-Penetrating Radar for Archaeology*. New York: Left Coast Press.

Conyers, L. B., 2013. *Ground-Penetrating Radar for Archaeology*. 3rd ed. Walnut Creek, CA: AltaMira Press.

Conyers, L. B., 2016. *Ground-Penetrating Radar for Geoarchaeology*. Ground-penetrating Radar for Geoarchaeology. Chichester: Wiley-Blackwell.

Conyers, L. B. and Goodman, D., 1997. *Ground-Penetrating Radar: An introduction for archaeologists*. 1st ed. Walnut Creek, CA: AltaMira.

Corlett, C. and Potterton, M., eds., 2010. *Death and Burial in Early Medieval Ireland: In the light of recent archaeological excavations*. Bray: Wordwell.

Cotter, C., 1987. 'Ferrycarrig', Newtown. *In*: Bennett, I., ed. *Excavations 1987: summary accounts of archaeological excavations in Ireland*. Dublin: Wordwell.

Creamer, R., Simo, I., O'Sullivan, L., Reidy, B., Schulte, R., and Fealy, R., 2016. *Soil Property Maps* [online]. Johnstown Castle, Co Wexford. Available from: https://www.epa.ie/pubs/reports/research/land/EPA RR 204 final web.pdf.

Crook, K., Bissell, H., Beck, K., Haaf, E. ten, Folk, R., and Schwab, F., 2018. Sedimentary Rock. *Encyclopedia Britannica* [online]. Available from: https://www.britannica.com/science/sedimentary-rock.

Davenport, G. C., 2001. Remote Sensing Applications in Forensic Investigations. *Historical Archaeology* [online], 35 (1), 87–100. Available from: http://phrtoolkits.org/wp-content/uploads/2012/06/forensics_davenport.pdf.

David, A., Linford, N., and Linford, P., 2008a. *Geophysical Survey in Archaeological Field Evaluation*. Wetlands. London.

David, A., Linford, N., and Linford, P., 2008b. Geophysical survey in archaeological field evaluation, (1), 1–40.

Davis, J. L., Heginbottom, J. a, Annan, A. P., Daniels, R. S., Berdal, B. P., Bergan, T., Duncan, K. E., Lewin, P. K., Oxford, J. S., Roberts, N., Skehel, J. J., and Smith, C. R., 2000. Ground Penetrating Radar Surveys to Locate 1918 Spanish Flu Victims in Permafrost. *Journal of Forensic Sciences* [online], 45 (1), 68–76. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10641921.

Delaney, S., 2017. Ranelagh, Co. Roscommon: Preliminary observations. *In*: *Strata: Uncovering layers of archaeology* [online]. Dublin: Institute of Archaeologists of Ireland. Available from: http://www.iai.ie/wp-content/uploads/2016/03/IAI-Programme-2017.pdf.

Dinh, K., Gucunski, N., and Duong, T. H., 2018. An algorithm for automatic localization and detection of rebars from GPR data of concrete bridge decks. *Automation in Construction*.

Duc, T. T., 2006. Using GIS and AHP Technique for Land-Use Suitability Analysis. *International Symposium on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences* [online]. Available from: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.468.4502&rep=rep1&type=pdf.

Duffy, S., 2005a. *The Concise History of Ireland*. 2nd ed. Dublin: Gill & Macmillan Ltd.

Duffy, S., 2005b. *Medieval Ireland*. New York: Routledge.

Dupras, T. L., 2012. *Forensic Recovery of Human Remains: Archaeological approaches*. Boca Raton: CRC Press.

Eaton, T., 1999. Grave Site at Juarez Searched. *Dallas Morning News* [online], 1999, p. 1A. Available from: www.elsevier.nlrlocaterjappgeo [Accessed 10 Nov 2017].

Environmental Protection Agency Ireland, 2019. *EPA GeoPortal* [online]. Data Download. Available from: http://gis.epa.ie/GetData/Download [Accessed 10 Jan 2018].

Ernenwein, E. G., 2009. Integration of Multidimensional Archaeogeophysical Data using Supervised and Unsupervised Classification. *Near Surface Geophysics*, 7 (3), 147–158.

Fergus, R., 2015. *Neural Networks: MLSS 2015 Summer School* [online]. Available from: http://mlss.tuebingen.mpg.de/2015/slides/fergus/Fergus_1.pdf [Accessed 16 Jul 2019].

France, D., Griffin, T., Swanburg, J., Lindemann, J., Davenport, G. C., Trammell, V., Travis, C., Kondratieff, B., Nelson, A., Castellano, K., Hopkins, D., and Adair, T., 1997. NecroSearch

Revisited: Further multidisciplinary approaches to the detection of clandestine graves. *In*: Haglund, W. D. and Sorg, M. H., eds. *Forensic Taphonomy: The postmortem fate of human remains*. Boca Raton: CRC Press, 497–509.

France, D. L., Griffin, T. J., Swanburg, J. G., Lindemann, J. W., Clark Davenport, G., Trammell, V., Armbrust, C. T., Kondratieff, B., Nelson, A., Castellano, K., and Hopkins, D., 1992. A Multidisciplinary Approach to the Detection of Clandestine Graves. *Journal of Forensic Sciences* [online], 37 (6), 1445–1458. Available from: http://www.astm.org/doiLink.cgi?JFS13337J.

Freeland, T., Heung, B., Burley, D. V., Clark, G., and Knudby, A., 2016. Automated Feature Extraction for Prospection and Analysis of Monumental Earthworks from Aerial LiDAR in the Kingdom of Tonga. *Journal of Archaeological Science* [online], 69, 64–74. Available from: https://www.sciencedirect.com/science/article/pii/S0305440316300206.

Fry, R., 2014. Time-Lapse Geophysical Investigations Over Known Archaeological Features using Electrical Resistivity Imaging and Earth Resistance. [online]. University of Bradford. Available from: https://bradscholars.brad.ac.uk/handle/10454/6901.

Gaffney, C. F. and Gater, J., 2003. *Revealing the Buried Past: Geophysics for archaeologists*. Stroud: Tempus.

Gaffney, C. and Gaffney, V., 2011. Through an Imperfect Filter: Geophysical techniques and the management of archaeological heritage. *In*: Cowley, D., ed. *Remote Sensing for Archaeological Heritage Management*. Brussels: Europae Archaeologiae Consilium, 117–128.

Gaffney, C., Harris, C., Newman, M., Langton, M., and Walker, R., 2014. Fountains Abbey: A re-survey and new interpretations within the world heritage site. *In*: *Recent Work in Archaeological Geophysics* [online]. London: Near-Surface Geophysics Group. Available from: http://www.nsgg.org.uk/meetings/nsgg2014/NSGGAbstracts2014_sml.pdf.

Gansell, A. R., Meent, J. W. van de, Zairis, S., and Wiggins, C. H., 2014. Stylistic Clusters and the Syrian/South Syrian Tradition of First-Millennium BCE Levantine Ivory Carving: A machine learning approach. *Journal of Archaeological Science* [online], 44, 194–205. Available from: https://www.sciencedirect.com/science/article/pii/S0305440313003919.

Geological Survey Ireland, 2018. *Geological Survey Ireland Spatial Resources* [online]. Available from: https://dcenr.maps.arcgis.com/apps/MapSeries/index.html?appid=a30af518e87a4c0ab2fbde2aaac3c228 [Accessed 4 Jan 2019].

Geological Survey Ireland, 2019. *Data and Maps* [online]. Available from: https://www.gsi.ie/en-ie/data-and-maps/Pages/default.aspx [Accessed 10 Jan 2018].

Geonics Ltd., 2003. *EM38B Ground Conductivity Meter Operating Manual*. Mississauga, Ontario: Geonics Ltd.

Gilchrist, R. and Sloane, B., 2005. *Requiem: The medieval monastic cemetery in Britain*. London: Museum of London Archaeology Service.

Goodman, D., Nishimura, Y., and Rogers, J. D., 1995. GPR Time Slices in Archaeological Prospection. *Archaeological Prospection*, 2 (2), 85–89.

Goodman, D., Nishimura, Y., Uno, T., and Yamamoto, T., 1994. A Ground Radar Survey of Medieval Kiln Sites in Suzu City, Western Japan. *Archaeometry*, 36 (2), 317–326.

Goodman, D. and Piro, S., 2013. *GPR Remote Sensing in Archaeology*. GPR Remote Sensing in Archaeology. Berlin: Springer-Verlag.

Gosar, A. and Čeru, T., 2016. Search for an Artificially Buried Karst Cave Entrance Using Ground Penetrating Radar: A successful case of locating the S-19 cave in the Mt. Kanin massif (NW Slovenia). *International Journal of Speleology*, 45 (2), 135–147.

Green, A., 2015. Detecting Burials and Buried Structures: A comparison of parameters for ground-penetrating radar surveys in forensic and archaeological investigations of clay environments. Bournemouth University.

Green, A. and Holmes, R., 2017. Accounting for Environmental and Anthropogenic Factors: Approaches to enhancing horizontal resolution and interpretability in geophysical surveys. *In*: *CAA International* [online]. Atlanta: Georgia State University. Available from: http://eprints.bournemouth.ac.uk/28204/1/CAA Poster Edit.pdf [Accessed 10 Nov 2017].

Guillen, P., Larrazabal, G., González, G., Boumber, D., and Vilalta, R., 2015. Supervised learning to detect salt body. *In*: *SEG Technical Program Expanded Abstracts*.

Hagan, J., Shaw, L., and Brown, J., 2017. *Report of Results from Geophysical Survey at Buckland Rings* [online]. Lymington. Available from: https://nfknowledge.org/wp-content/uploads/2018/11/Report_Buckland1.pdf.

Harrison, M. and Donnelly, L. J., 2009. Locating Concealed Homicide Victims: Developing the role of geoforensics. *In*: *Criminal and Environmental Soil Forensics* [online]. Amsterdam: Springer Netherlands, 197–219. Available from: https://link.springer.com/book/10.1007/978-1-4020-9204-6.

He, K., Zhang, X., Ren, S., and Sun, J., 2016. Deep Residual Learning for Image Recognition. *In*: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* [online]. Washington, DC: IEEE Computer Society. Available from: https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/He_Deep_Residual_Learning_CVPR_2016_paper.pdf.

Historic England, 2019. *Glossary* [online]. Available from: https://historicengland.org.uk/services-skills/education/glossary/#index [Accessed 18 May 2019].

Horning, A., Ó Baoill, R., Donnelly, C., and Logue, P., eds., 2007. *The Post-Medieval Archaeology of Ireland, 1550–1850.* Dublin: Wordwell.

Hörr, C., Lindinger, E., and Brunnett, G., 2014. Machine Learning Based Typology Development in Archaeology. *Journal on Computing and Cultural Heritage* [online], 7 (1). Available from: http://delivery.acm.org/10.1145/2540000/2533988/a2-horr.pdf?ip=194.80.64.254&id=2533988&acc=ACTIVE SERVICE&key=BF07A2EE685417C5.ADA9075A69F4291D.4D4702B0C3E38B35.4D4702B0 C3E38B35&__acm__=1561047253_2ba13f73eac365754364f139c73dc315.

Howard, B., 2018. *Buckland Rings: Archaeological GPR survey report* [online]. Southampton. Available from: https://nfknowledge.org/wp-content/uploads/2018/11/Buckland-Rings-GPR-Report-V4.pdf.

Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S., and Murphy, K., 2017. Speed/accuracy trade-offs for modern convolutional object detectors. [online]. Available from: http://arxiv.org/abs/1611.10012 [Accessed 10 Nov 2017].

Hunter, J. and Martin, A. L., 2002. Studies in Crime: An introduction to forensic archaeology. *In*: Hunter, J. R., Roberts, C. A., and Martin, A. L., eds. *Studies in Crime: An introduction to forensic archaeology* [online]. London: Routledge. Available from: http://www.jstor.org/stable/506107?origin=crossref.

Imai, T., Saskayama, T., and Kanemori, T., 1987. Use of Ground-Probing Radar and Resistivity Surveys for Archaeological Investigations. *GEOPHYSICS* [online], 52 (2), 137–241. Available from: https://library.seg.org/doi/abs/10.1190/1.1442290.

Institute of Archaeologists of Ireland, 2016. *IAI Code of Conduct for the Archaeological Treatment of Human Remains - in the context of an archaeological excavation* [online]. Dublin. Available from: http://www.iai.ie/wp-content/uploads/2016/03/IAI-Code-of-Conduct-for-the-Archaeological-Treatment-of-Human-Remains.pdf.

Ishitsuka, K., Iso, S., Onishi, K., and Matsuoka, T., 2018. Object Detection in Ground-Penetrating Radar Images Using a Deep Convolutional Neural Network and Image Set Preparation by Migration. *International Journal of Geophysics*.

Johansson, M., 1999. The Hilbert Transform. [online]. Vaxjo University. Available from: http://fuchs-braun.com/media/d9140c7b3d5004fbffff8007ffffff0.pdf.

Jol, H. M. and Bristow, C. S., 2003. GPR in Sediments: Advice on data collection, basic processing and interpretation, a good practice guide. *Geological Society, London, Special Publications* [online], 211, 9–27. Available from: https://sp.lyellcollection.org/content/211/1/9.

Karpathy, A., 2019a. *Understanding and Visualizing Convolutional Neural Networks* [online]. Available from: http://cs231n.github.io/understanding-cnn/.

Karpathy, A., 2019b. *Convolutional Neural Networks: Architectures, Convolution / Pooling Layers* [online]. Available from: http://cs231n.stanford.edu/ [Accessed 16 Jul 2019].

Kaur, P., Dana, K. J., Romero, F. A., and Gucunski, N., 2016. Automated GPR Rebar Analysis for Robotic Bridge Deck Evaluation. *IEEE Transactions on Cybernetics*.

Kenyon, J. L. and Bevan, B., 1977. Ground-Penetrating Radar and Its Application to a Historical Archaeological Site. *Historical Archaeology*.

Killam, E. W., 1990. *The Detection of Human Remains*. 1st ed. Springfield, Illinois: Thomas.

Koppenjan, S., Streeton, M., Lee, H., Lee, M., and Ono, S., 2004. Advanced Signal Analysis for Forensic Applications of Ground Penetrating Radar. *In*: *Proceedings of the Tenth International Conference on Ground Penetrating Radar, Vols 1 and 2*. Washington, DC: IEEE, 443–446.

Kramer, I., 2015. An Archaeological Reaction to the Remote Sensing Data Explosion. University of Southampton.

Kramer, I., Hare, J., and Prugel-Bennett, A., 2017. A Future Perspective for Automated Detection of Archaeology Using Deep Learning with Remote Sensor Data. *In*: *ACM WomENcourage: Celebration of Women in Computing* [online]. Barcelona. Available from: https://eprints.soton.ac.uk/416400/.

Krizhevsky, A., Hinton, G. E., Sutskever, I., and Geoffrey E., H., 2012. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems 25 (NIPS2012)* [online], 1–9. Available from: https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf.

Küçükdemirci, M. and Sarris, A., 2020. Deep Learning Based Automated Analysis of Archaeo-geophysical Images. *Archaeological Prospection* [online]. Available from: https://onlinelibrary.wiley.com/doi/10.1002/arp.1763.

Kvamme, K. L., 2007. Integrating Multiple Geophysical Datasets. *In*: Wiseman, J. and El-Baz, F., eds. *Remote Sensing in Archaeology*. New York: Springer, 345–374.

De Laet, V., Paulissen, E., and Waelkens, M., 2007. Methods for the Extraction of Archaeological

Features from Very High-Resolution Ikonos-2 Remote Sensing Imagery, Hisar (southwest Turkey). *Journal of Archaeological Science*, 34 (5), 830–841.

Lambers, K., Verschoof-van der Vaart, W. B., and Bourgeois, Q. P. J., 2019. Integrating Remote Sensing, Machine Learning, and Citizen Science in Dutch Archaeological Prospection. *Remote Sensing* [online], 11 (7). Available from: https://www.mdpi.com/2072-4292/11/7/794.

Lameri, S., Lombardi, F., Bestagini, P., Lualdi, M., and Tubaro, S., 2017. Landmine detection from GPR data using convolutional neural networks. *In*: *25th European Signal Processing Conference, EUSIPCO 2017*.

Lanzarone, P., Garrison, E., Bobe, R., and Getahun, A., 2016. Examining Fluvial Stratigraphic Architecture Using Ground-Penetrating Radar at the Fanta Stream Fossil and Archaeological Site, Central Ethiopia. *Geoarchaeology*, 31 (6), 577–591.

Leckebusch, J., 2003. Ground-Penetrating Radar: A modern three-dimensional prospection method. *Archaeological Prospection* [online], 10 (4), 213–240. Available from: http://doi.wiley.com/10.1002/arp.211.

Leckebusch, J., Weibel, A., and Bühler, F., 2008. Semi-Automatic Feature Extraction from GPR Data for Archaeology. *Near Surface Geophysics*, 6 (2), 75–84.

LeCun, Y., Bengio, Y., and Hinton, G., 2015. Deep Learning. *Nature* [online], 521, 436–444. Available from: https://www.nature.com/articles/nature14539.pdf.

LeCun, Y., Bottou, L., Bengio, Y., and Ha, P., 1998. LeNet. *Proceedings of the IEEE*.

Lin, M., Chen, Q., and Yan, S., 2013. Network in Network. *arXiv e-prints* [online], eprint arX, 1–10. Available from: https://arxiv.org/abs/1312.4400.

Linford, N., Linford, P., Persico, R., and Piro, S., 2018. The Application of Semi-Automated Vector Identification to Large Scale Archaeological Data Sets Considering Anomaly Morphology. *EGU General Assembly 2018* [online], 20. Available from: https://meetingorganizer.copernicus.org/EGU2018/EGU2018-4919.pdf.

Linford, N. T., 2004. Magnetic Ghosts: Mineral magnetic measurements on Roman and Anglo-Saxon graves. *Archaeological Prospection* [online], 11 (3), 167–180. Available from: https://onlinelibrary.wiley.com/doi/epdf/10.1002/arp.232.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Laak, J. A. W. M. van der, Ginneken, B. van, and Sánchez, C. I., 2017. A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis* [online], 42, 60–88. Available from: https://www.sciencedirect.com/science/article/pii/S1361841517301135.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., and Berg, A. C., 2016. SSD: Single shot multibox detector. *In*: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 21–37.

van der Maaten, L., Boon, P., Lange, G., Paijmans, H., and Postma, E., 2006. Computer Vision and Machine Learning for Archaeology. *In*: Clark, J. T. and Hagemeister, E. M., eds. *Digital Discovery. Exploring New Frontiers in Human Heritage. CAA2006. Computer Applications and Quantitative Methods in Archaeology. Proceedings of the 34th Conference, Fargo, United States, April 2006.* [online]. Budapest: Archaeolingua, 476–482. Available from: http://explore.bl.uk/primo_library/libweb/action/display.do?tabs=detailsTab&gathStatTab=true&ct=display&fn=search&doc=ETOCCN075830853&indx=1&recIds=ETOCCN075830853.

Malczewski, J. and Rinner, C., 2015. Multiattribute Decision Analysis Methods. *In*: *Advances in Geographic Information Science*.

Marinoni, O., 2004. Implementation of the Analytical Hierarchy Process with VBA in ArcGIS. *Computers and Geosciences*.

Marinoni, O., 2017. *extAhp20 - Analytic Hierarchy Process for ArcGIS* [online]. Available from: https://www.arcgis.com/home/item.html?id=bb3521d775c94b28b69a10cd184b7c1f [Accessed 15 Dec 2019].

Marquardt, D. W., 1963. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics* [online], 11 (2), 431–441. Available from: https://epubs.siam.org/doi/10.1137/0111030.

Marsh, T., 2013. Assessing the Accuracy of Ground Penetrating Radar for the Detection of Deep Clandestine Graves. Unpublished MSc Dissertation, Bournemouth University.

Maulik, U. and Chakraborty, D., 2017. Remote sensing image classification: A survey of support-vector-machine-based advanced techniques. *IEEE Geoscience and Remote Sensing Magazine*.

McLoughlin, C., 2006. *Newtown, Wexford* [online]. Enniscorthy. Available from: https://excavations.ie/report/2005/Wexford/0014677/.

McNeill, J. D., 1980. *Electrical Conductivity of Soils and Rocks. Technical Note 5* [online]. Technical Notes. Mississauga, Ontario. Available from: http://www.geonics.com/pdfs/technicalnotes/tn5.pdf.

McNeill, J. D., 1985. EM34-3 Measurements at Two Inter-Coil Spacings to Reduce Sensitivity to Near-Surface Material. *Geonics Limited* [online], Technical, 1–3. Available from: http://www.geonics.com/pdfs/technicalnotes/tn19.pdf.

Mellett, J., 1992. Location of Human Remains with Ground-Penetrating Radar. *In*: Hanninen, P. and Autio, S., eds. *Fourth International Conference on Ground Penetrating Radar*. 359–365.

Mendoza, G., 2000. A GIS-Based Multicriteria Approach to Land Use Suitability Assessment and Allocation. *United States Department Of Agriculture Forest Service General Technical Report NC*.

Miller, P. S., 1996. Disturbances in the Soil: Finding buried bodies and other evidence using ground penetrating radar. *Journal of Forensic Sciences* [online], 41 (4), 648–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8754575.

Milosavljevic, A., 2020. Identification of salt deposits on seismic images using deep learning method for semantic segmentation. *ISPRS International Journal of Geo-Information*.

Moffat, I., 2015. Locating Graves with Geophysics. *In*: Sarris, A., ed. *Best Practices of GeoInformatic Technologies for the Mapping of Archaeolandscapes* [online]. Oxford: Archaeopress, 45–53. Available from: https://www.academia.edu/35804173/Locating_Graves_with_Geophysics [Accessed 3 Jan 2019].

Monteith, J. and Green, A., 2019. *Buckland Rings Hillfort (Scheduled Monument 21843) - Geophysical Survey Report*. Bournemouth.

Moore, G. M. and Ryder, C. M., 2015. The Use of Ground-Penetrating Radar to Locate Tree Roots. *Arboriculture & Urban Forestry*, 41 (5), 245–259.

Moysey, S., Knight, R. J., and Jol, H. M., 2006. Texture-Based Classification of Ground-Penetrating Radar Images. *GEOPHYSICS* [online], 71 (6). Available from: https://library.seg.org/doi/abs/10.1190/1.2356114.

Murray, B., Anderson, D. T., Wescott, D. J., Moorhead, R., and Anderson, M. F., 2018. Survey and

Insights into Unmanned Aerial-Vehicle-Based Detection and Documentation of Clandestine Graves and Human Remains. *Human Biology* [online], 90 (1), 45–61. Available from: https://bioone.org/journals/Human-Biology/volume-90/issue-1/humanbiology.90.1.03/Survey-and-Insights-into-Unmanned-Aerial-Vehicle-Based-Detection-and/10.13110/humanbiology.90.1.03.short.

Murray, N., 2017. 800 Medieval Bodies Found During Hillfort Dig. *Irish Examiner* [online], 1 April 2017. Available from: https://www.irishexaminer.com/viewpoints/analysis/800-medieval-bodies-found-during-hillfort-dig-446670.html.

Najafabadi, M. M., Villanustre, F., Khoshgoftaar, T. M., Seliya, N., Wald, R., and Muharemagic, E., 2015. Deep Learning Applications and Challenges in Big Data Analytics. *Journal of Big Data* [online], 2 (1), 1–21. Available from: https://journalofbigdata.springeropen.com/track/pdf/10.1186/s40537-014-0007-7.

Nattkemper, T. W., Twellmann, T., Ritter, H., and Schubert, W., 2003. Human vs. Machine: Evaluation of fluorescence micrographs. *Computers in Biology and Medicine* [online], 33 (1), 31–43. Available from: https://www.sciencedirect.com/science/article/pii/S0010482502000604.

Neal, A., 2004. Ground-Penetrating Radar and Its Use in Sedimentology: Principles, problems and progress. *Earth-Science Reviews*, 66 (3–4), 261–330.

Noonan, D., 2003. *24-26 Main Street, Birr, Offaly* [online]. Birr. Available from: https://excavations.ie/report/2003/Offaly/0010565/.

Núñez-Nieto, X., Solla, M., Gómez-Pérez, P., and Lorenzo, H., 2014. GPR signal characterization for automated landmine and UXO detection based on machine learning techniques. *Remote Sensing*.

O'Brien, C. and Sweetman, P. D., 1997. *Archaeological Inventory of County Offaly*. Dublin: Stationery Office.

O'Carroll, F., Scott, R., Corrway, L., and Kinch, I., 2017. *Black Friary Archaeological and Community Report for Season 8, Spring/Summer 2017* [online]. Available from: http://bafs.ie/wp-content/uploads/2018/12/Black-Friary-Report-summer-2017.pdf.

O'Carroll, F., Shine, D., and Scott, R., 2017. *Black Friary Archaeological and Community Report 2017* [online]. Available from: http://bafs.ie/wp-content/uploads/2017/12/Black-Friary-Report-2017-final.compressed-1.pdf.

O'Curry, E., 1855. *Cath Mhuighe Léana or The Battle of Magh Leana, Together with Tochmarc Moméra or the Courtship of Moméra*. Dublin: The Celtics Society.

O'Shea, K. and Nash, R., 2015. An Introduction to Convolutional Neural Networks. *arXiv preprint* [online]. Available from: https://arxiv.org/pdf/1511.08458.pdf.

O'Sullivan, A., McCormick, F., Kerr, T. R., and Harney, L., 2014. *Early Medieval Ireland, AD 400-1100: The evidence from archaeological excavations* [online]. Dublin: Royal Irish Academy. Available from: http://www.jstor.org/stable/j.ctt14jxtqz [Accessed 10 Nov 2017].

Olczak, J., Fahlberg, N., Maki, A., Razavian, A. S., Jilert, A., Stark, A., Sköldenberg, O., and Gordon, M., 2017. Artificial Intelligence for Analyzing Orthopedic Trauma Radiographs. *Acta Orthopaedica* [online], 88 (6). Available from: https://www.tandfonline.com/doi/full/10.1080/17453674.2017.1344459.

Orlando, L. and Slob, E., 2009. Using Multicomponent GPR to Monitor Cracks in a Historical Building. *Journal of Applied Geophysics* [online], 67 (4), 327–334. Available from: https://www.sciencedirect.com/science/article/pii/S0926985108001092.

Osman, K. T., 2013. *Soils: Principles, properties and management*. Soils: Principles, Properties and Management.

OSSAFreelance, 2012. A Basic Overview for the Recovery of Human Remains from Sites Under Development. *BAJR Practical Guide Series* [online], Guide 13, 1–20. Available from: http://www.bajr.org/BAJRGuides/13. A Basic Overview of the Recovery of Human Remains from Sites Under Development/13HumanRemainsDevelopment.pdf.

Pan, S. J. and Yang, Q., 2010. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* [online], 22 (10), 1345–1359. Available from: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5288526.

De Paor, L., Hunt, J., Plenderleith, H. J., and Dolley, M., 1969. *Excavations at Mellifont Abbey, Co. Louth* [online]. Dublin: Royal Irish Academy. Available from: https://books.google.co.uk/books?id=rxNQGwAACAAJ.

Pasolli, E., Melgani, F., and Donelli, M., 2009. Automatic Analysis of GPR images: A pattern-recognition approach. *IEEE Transactions on Geoscience and Remote Sensing*, 47 (7), 2206–2217.

Pasolli, E., Melgani, F., Donelli, M., Attoui, R., and de Vos, M., 2008. Automatic Detection and Classification of Buried Objects in GPR Images Using Genetic Algorithms and Support Vector Machines. *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium* [online], (1), II-525-II–528. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4779044.

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., and Lerer, A., 2017. Automatic Differentiation in PyTorch. *In*: *31st Conference on Neural Information Processing Systems* [online]. Long Beach, CA: NIPS, 1–4. Available from: https://openreview.net/pdf?id=BJJsrmfCZ.

Pringle, J. K., Holland, C., Szkornik, K., and Harrison, M., 2012. Establishing Forensic Search Methodologies and Geophysical Surveying for the Detection of Clandestine Graves in Coastal Beach Environments. *Forensic Science International*, 219 (1–3).

Pringle, J. K., Jervis, J. R., Roberts, D., Dick, H. C., Wisniewski, K. D., Cassidy, N. J., and Cassella, J. P., 2016. Long-Term Geophysical Monitoring of Simulated Clandestine Graves Using Electrical and Ground Penetrating Radar Methods: 4-6 years after burial. *Journal of Forensic Sciences*, 61 (2), 309–321.

Qiao, L., Qin, Y., Ren, X., and Wang, Q., 2015. Identification of buried objects in gpr using amplitude modulated signals extracted from multiresolution monogenic signal analysis. *Sensors (Switzerland)*, 15 (12), 30340–30350.

Raina, R., Ng, A., and Koller, D., 2006. Constructing informative priors using transfer learning. *Proceedings of the 23rd international conference on Machine learning* [online], 713–720. Available from: http://dl.acm.org/citation.cfm?id=1143934.

Rao, K. R., Kim, D. N., and Hwang, J. J., 2011. *Fast Fourier Transform - Algorithms and applications*. 1st ed. Amsterdam: Springer Netherlands.

Rawat, W. and Wang, Z., 2017. Deep Convolutional Neural Networks for Image Classification: A comprehensive review. *Neural Computation*, 29 (9), 2352–2449.

Richards, J., Tudhope, D., and Vlachidis, A., 2015. Text Mining in Archaeology: Extracting information from archaeological reports. *In*: Barcelo, J. A. and Bogdanovic, I., eds. *Mathematics and Archaeology*. Boca Raton: CRC Press.

Richardson, T. and Cheetham, P. N., 2013. The Effectiveness of Geophysical Techniques in Detecting a Range of Buried Metallic Weapons at Various Depths and Orientations. *Geological Society, London, Special Publications* [online], 384. Available from: http://sp.lyellcollection.org/content/early/2013/08/30/SP384.18.abstract.

Ruffell, A. and McAllister, S., 2015. A RAG System for the Management of Forensic and Archaeological Searches of Burial Grounds. *International Journal of Archaeology International Journal of Archaeology. Special Issue: Archaeological Sciences* [online], 3 (1), 1–8. Available from: http://www.sciencepublishinggroup.com/j/ija.

Ruffell, A., McCabe, A., Donnelly, C., and Sloan, B., 2009. Location and Assessment of an Historic (150-160 years old) Mass Grave Using Geographic and Ground Penetrating Radar Investigation, NW Ireland. *Journal of Forensic Sciences*, 54 (2), 382–394.

Ruffell, A. and McKinley, J., 2005. Forensic Geoscience: Applications of geology, geomorphology and geophysics to criminal investigations. *Earth-Science Reviews*, 69 (3–4), 235–247.

Ruffell, A., Pringle, J. K., and Forbes, S., 2014. Search Protocols for Hidden Forensic Objects Beneath Floors and Within Walls. *Forensic Science International* [online], 237, 137–145. Available from: https://www.sciencedirect.com/science/article/pii/S0379073814000048.

Saaty, T. L., 1980. *The Analytic Hierarchy Process*. New York: McGraw-Hill.

Saey, T., Van Meirvenne, M., De Smedt, P., Neubauer, W., Trinks, I., Verhoeven, G., and Seren, S., 2013. Integrating Multi-Receiver Electromagnetic Induction Measurements into the Interpretation of the Soil Landscape Around the School of Gladiators at Carnuntum. *European Journal of Soil Science*, 64 (5), 716–727.

Saha, S., 2018. *A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way* [online]. Towards Data Science. Available from: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53 [Accessed 30 Mar 2020].

Schmidhuber, J., 2014. Deep Learning in Neural Networks: An overview. *Neural Networks*, 61, 85–117.

Schmidt, A., 2001. *Geophysical Data in Archaeology: A guide to good practice*. 2nd ed. Oxford: Oxbow Books.

Schmidt, A., 2013. *Earth Resistance for Archaeologists*. Plymouth: AltaMira Press.

Schmidt, A., 2019. Guidelines for the Use of Geophysics in Archaeology: Should they be prescriptive? *In*: Meylemans, E. and De Smedt, P., eds. *The Use of Geophysical Prospection Methods in Archaeology* [online]. Brussel: agentschap Onroerend Erfgoed, 23–34. Available from: https://www.academia.edu/39278580/Guidelines_for_the_use_of_geophysics_in_archaeology_should_they_be_prescriptive.

Schmidt, A., Linford, P., Linford, N., David, A., Gaffney, C., Sarris, A., and Fassbinder, J., 2015. *EAC Guidelines for the Use of Geophysics in Archaeology: Questions to ask and points to consider* [online]. 2nd ed. Budapest: Archaeolingua. Available from: https://bradscholars.brad.ac.uk/bitstream/handle/10454/8129/eac_guidelines_2_final.pdf?sequence=1&isAllowed=y [Accessed 10 Nov 2017].

Schneider, A., Takla, M., Nicolay, A., Raab, A., and Raab, T., 2015. A Template-matching Approach Combining Morphometric Variables for Automated Mapping of Charcoal Kiln Sites. *Archaeological Prospection* [online], 22 (1), 45–62. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/arp.1497.

Schultz, J. J., 2007. Using Ground-Penetrating Radar to Locate Clandestine Graves of Homicide Victims: Forming forensic archaeology partnerships with law enforcement. *Homicide Studies* [online], 11 (1), 15–29. Available from: http://hsx.sagepub.com/cgi/doi/10.1177/1088767906296234.

Schultz, J. J., 2008. Sequential Monitoring of Burials Containing Small Pig Cadavers Using Ground Penetrating Radar. *Journal of Forensic Sciences*, 53 (2), 279–287.

Schultz, J. J., 2012. The Application of Ground-Penetrating Radar for Forensic Grave Detection. *In*: Dirkmaat, D., ed. *A Companion to Forensic Anthropology*. Hoboken, NJ: Wiley-Blackwell, 85–100.

Schultz, J. J., Grasmueck, M., Weger, R., Muztaza, N. M., Saidin, M. M., Azwin, I. N., Saad, R., Kofun, H., Leucci, G., Giacomo, G. Di, Conyers, L. B., Leckebusch, J., Roman, A., Lukyanov, S. P., Stepanov, R. A., Stukach, O. V, Urxqg, V. D. Q. G. F., Fxkn, P., Kn, H. G. X., Lukyanov, S. P., Stukach, O. V, Barone, P. M., and Altunel, E., 2012. Detecting Buried Remains Using Ground-Penetrating Radar. *Journal of Archaeological Science* [online], 36 (April), 235. Available from: http://dx.doi.org/10.1016/j.jas.2008.12.016.

Schultz, J. J. and Martin, M. M., 2012. Monitoring Controlled Graves Representing Common Burial Scenarios with Ground Penetrating Radar. *Journal of Applied Geophysics*, 83, 74–89.

Shine, D., Potterton, M., Mandal, S., and McLoughlin, C., eds., 2019. *Carrick, County Wexford: Ireland's first Anglo-Norman Stronghold*. Dublin: Four Courts Press.

Shlens, J., 2016. *Research Blog: Train your own image classifier with Inception in TensorFlow* [online]. Available from: https://research.googleblog.com/2016/03/train-your-own-image-classifier-with.html [Accessed 10 Nov 2017].

Sikora, M., 2011. 3.96 Newtown, Co. Wexford, E1189. *In*: *Breaking ground, finding graves - reports on the excavations of burials by the National Museum of Ireland, 1927-2006, vol. 1*. Dublin: Wordwell, 596–598.

Simon, M., Rodner, E., and Denzler, J., 2016. ImageNet Pre-Trained Models with Batch Normalization. *arXiv e-prints* [online], eprint arX, 1–4. Available from: https://arxiv.org/pdf/1612.01452.pdf.

Simonyan, K. and Zisserman, A., 2015. Very Deep Convolutional Networks for Large-Scale Image Recognition. *International Conference on Learning Representations,* [online], 2015, 1–14. Available from: https://arxiv.org/pdf/1409.1556.pdf.

Singh, A., Thakur, N., and Sharma, A., 2016. A Review of Supervised Machine Learning Algorithms. *IEEE* [online]. Available from: http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7724478 [Accessed 10 Nov 2017].

Singh, N. P. and Nene, M. J., 2013a. Buried Object Detection and Analysis of GPR Images: Using neural network and curve fitting. *In*: *2013 International Conference on Microelectronics, Communications and Renewable Energy* [online]. Annual International Conference on Emerging Research Areas, 1–6. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6576024.

De Smedt, P., Saey, T., Lehouck, A., Stichelbaut, B., Meerschman, E., Islam, M. M., Van De Vijver, E., and Van Meirvenne, M., 2013. Exploring the Potential of Multi-Receiver EMI Survey for Geoarchaeological Prospection: A 90ha dataset. *Geoderma* [online], 199, 30–36. Available from: https://www.sciencedirect.com/science/article/pii/S0016706112002893.

Somers, L. E., Hargrave, M. L., and Simms, J. E., 2003. *Geophysical Surveys in Archaeology: Guidance for surveyors and sponsors* [online]. Arlington, VA. Available from:

https://apps.dtic.mil/dtic/tr/fulltext/u2/a429240.pdf.

Stout, M., 2018. *Early Medieval Ireland: 431-1169*. Dublin: Wordwell.

Stove, G. C. and Addyman, P. V., 1989. Ground Probing Impulse Radar: An experiment in archaeological remote sensing at York. *Antiquity* [online], 63 (239), 337–342. Available from: https://www.cambridge.org/core/journals/antiquity/article/ground-probing-impulse-radar-an-experiment-in-archaeological-remote-sensing-at-york/4DEAE0857E1F74FC71ABDD3404C8616F.

Teagasc, 2017. *Soil Information System* [online]. Available from: http://gis.teagasc.ie/soils/map.php [Accessed 4 Jan 2019].

TensorFlow, 2019. *Build a Convolutional Neural Network using Estimators* [online]. Available from: https://www.tensorflow.org/tutorials/estimators/cnn [Accessed 20 Jun 2019].

Tierney, M., 2008. *28 Main Street, Birr, Offaly* [online]. Birr. Available from: https://excavations.ie/report/2008/Offaly/0020016/.

Toms, C., Rogers, C. B., and Sathyavagiswaran, L., 2008. Investigation of Homicides Interred in Concrete - The Los Angeles experience. *Journal of Forensic Sciences* [online], 53 (1), 203–207. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/j.1556-4029.2007.00600.x.

Traviglia, A., Cowley, D., and Lambers, K., 2016. Finding Common Ground: Human and computer vision in archaeological prospection. *AARGNEWS* [online], 53, 11–24. Available from: https://openaccess.leidenuniv.nl/handle/1887/43751.

Trier, Ø. D., Cowley, D. C., and Waldeland, A. U., 2018. Using Deep Neural Networks on Airborne Laser Scanning Data: Results from a case study of semi-automatic mapping of archaeological topography on Arran, Scotland. *Archaeological Prospection* [online], 26 (2), 165–175. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/arp.1731.

Trier, Ø. D., Larsen, S. Ø., and Solberg, R., 2009. Automatic Detection of Circular Structures in High-Resolution Satellite Images of Agricultural Land. *Archaeological Prospection*, 16 (1), 1–15.

Trier, Ø. D., Salberg, A. B., and Pilø, L. H., 2016. Semi-Automatic Mapping of Charcoal Kilns from Airborne Laser Scanning Data Using Deep Learning. *In*: Matsumoto, M. and Uleberg, E., eds. *Oceans of Data. Proceedings of the 44th Conference on Computer Applications and Quantitative Methods in Archaeology* [online]. Oxford: Archaeopress, 219–231. Available from: https://app.cristin.no/results/show.jsf?id=1448317.

Tschandl, P., Codella, N., Akay, B. N., Argenziano, G., Braun, R. P., Cabo, H., Gutman, D., Halpern, A., Helba, B., Hofmann-Wellenhof, R., Lallas, A., Lapins, J., Longo, C., Malvehy, J., Marchetti, M. A., Marghoob, A., Menzies, S., Oakley, A., Paoli, J., Puig, S., Rinner, C., Rosendahl, C., Scope, A., Sinz, C., Soyer, H. P., Thomas, L., Zalaudek, I., and Kittler, H., 2019. Comparison of the Accuracy of Human Readers Versus Machine-Learning Algorithms for Pigmented Skin Lesion Classification: An open, web-based, international, diagnostic study. *The Lancet Oncology* [online]. Available from: https://www.sciencedirect.com/science/article/pii/S147020451930333X.

Unterberger, R. R., 1992. Ground Penetrating Radar Finds Disturbed Earth Over Burials. *In*: Hänninen, P. and Autio, S., eds. *Fourth International Conference on Ground Penetrating Radar*. Espoo, Finland: Geologian Tutkimuskeskus, 351–357.

Vaidya, O. S. and Kumar, S., 2006. Analytic Hierarchy Process: An overview of applications. *European Journal of Operational Research*.

Vedaldi, A. and Lenc, K., 2014. *MatConvNet - Convolutional neural networks for MATLAB* [online].

arXiv e-prints. Available from: https://arxiv.org/pdf/1412.4564.pdf.

Verdonck, L., 2016. Detection of Buried Roman Wall Remains in Ground-Penetrating Radar Data using Template Matching. *Archaeological Prospection* [online], 23 (4), 257–272. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/arp.1540.

Verdonck, L., Launaro, A., Millett, M., Vermeulen, F., and Bellini, G., 2017. Semi-Automated Object Detection in GPR Data Using Morphological Filtering. *In*: Jennings, B., Gaffney, C., Sparrow, T., and Gaffney, S., eds. *AP2017: 12th International Conference of Archaeological Prospection*. Oxford: Archaeopress, 261–263.

Verdonck, L., De Smedt, P., and Verhegge, J., 2019. Making Sense of Anomalies: Practices and challenges in the archaeological interpretation of geophysical data. *In*: Persico, R., Piro, S., and Linford, N., eds. *Innovation in Near-Surface Geophysics*. Amsterdam: Elsevier, 151–194.

Verdonck, L., Taelman, D., Vermeulen, F., and Docter, R., 2015. The Impact of Spatial Sampling and Migration on the Interpretation of Complex Archaeological Ground-Penetrating Radar Data. *Archaeological Prospection*, 22 (2), 91–103.

Verschoof-van der Vaart, W. B. and Lambers, K., 2019. Learning to Look at LiDAR: The use of R-CNN in the automated detection of archaeological objects in LiDAR data from the Netherlands. *Journal of Computer Applications in Archaeology* [online], 2 (1), 31–40. Available from: https://journal.caa-international.org/articles/10.5334/jcaa.32/.

Vickers, R., Dolphin, L., and Johnson, D., 1976. Archaeological Investigations at Chaco Canyon Using Subsurface Radar. *In*: Lyons, T. R., ed. *Remote Sensing Experiments in Culture Resource Studies*. Albequerque: University of New Mexico, 81–101.

Waddell, J., 2010. *The Prehistoric Archaeology of Ireland*. 3rd ed. Dublin: Wordwell.

Warren, C., Giannopoulos, A., and Giannakis, I., 2016. gprMax: Open source software to simulate electromagnetic wave propagation for Ground Penetrating Radar. *Computer Physics Communications*, 209, 163–170.

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H., 2014. How Transferable Are Features in Deep Neural Networks? *Advances in Neural Information Processing Systems*, 3320–3328.

Zaremba, N. J., Smith, C. G., Bernier, J. C., and Forde, A. S., 2016a. Application of Ground Penetrating Radar for Identification of Washover Deposits and Other Stratigraphic Features: Assateague Island, MD. *Journal of Environmental & Engineering Geophysics* [online], 21 (4), 173–186. Available from: http://jeeg.geoscienceworld.org/lookup/doi/10.2113/JEEG21.4.173.

Zeiler, M. D. and Fergus, R., 2014. Visualizing and Understanding Convolutional Networks. *Computer Vision–ECCV 2014* [online]. Available from: https://cs.nyu.edu/~fergus/papers/zeilerECCV2014.pdf.

Zhao, W., Forte, E., Levi, S. T., Pipan, M., and Tian, G., 2015. Improved High-Resolution GPR Imaging and Characterization of Prehistoric Archaeological Features by Means of Attribute Analysis. *Journal of Archaeological Science* [online], 54, 77–85. Available from: https://www.sciencedirect.com/science/article/pii/S0305440314004464.

Appendix A.   Glossary

Abbey	Monastery run by an abbot/abbess
Accuracy	$$\frac{\text{True Positive + True Negative}}{\text{True Positive + True Negative + False Positive + False Negative}}$$
A-scan	Trace view of individual reflections in GPR data
Áth Cliath	Town
Burial	The practice of burying a dead body (see also *Inhumation*)
B-scan	Combined reflections in a GPR profile (see also *Radargram*)
Cillíní	Children's burial ground (plural)
Cóiceda	Province
Conductivity	The degree to which a material conducts electricity
Confusion Matrix	Table which describes the performance of a classification model by comparing the actual and predicted values in order to calculate a model's recall, accuracy, precision, and F_1 score.
Cross-Entropy Loss	Measure of the performance of a classification model where the output is a probability between 0 and 1. A high loss is indicative of the predicted value diverging from the actual value. Also called log loss. Where M=2: $-(y \log(p) + (1- y) \log(1-p))$ Where M>2: $\sum_{c=1}^{M} y_{o,c} \log(p_{o,c})$ M = number of classes; y = binary indicator; p = predicted probability; o = observation; c = class
C-scan	Combined B-scans in a survey to create a 3D dataset (see also *Time slice*)
Cúigí	Province
Damliac	House of stones

Dielectric permittivity	The degree of electrical polarisation a material experiences under the influence of an external electric field
Domnach	The Lord's Place
Epoch	A full pass of the training dataset through the algorithm during training. Sometimes equal to the batch size if the batch size is set to the entire training dataset in the training script.
F_1 score	$2 \cdot \dfrac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$
False negative	1. All features recorded during archaeological evaluation that did not produce a correlative geophysical response 2. Predicted value is negative, actual value is positive
False positive	1. Identified geophysical responses that do not correlate to an archaeological feature recorded during excavation 2. Predicted value is positive, actual value is negative
Feature map	The output produced after a convolutional filter is applied to merge data
Geinti	Pagan, Non-Christian
Grave	A hole dug in the ground to receive a coffin or dead body, typically marked by a stone or mound
Grave cut	The boundary between the redeposited soil in the grave shaft and the surrounding undisturbed soil matrix
Grave fill	The material redeposited in the grave shaft once the body has been interred
Inhumation	The practice of burying a dead body (see also *Burial*)
Land cover	Natural and manufactured surface cover indicative of the physical land types (e.g. infrastructure, vegetation, woodland)
Land use	How people use the land and its purpose (e.g. farming, timber production)

Leth Cuinn	Conn's Half
Leth Moga	Mug's Half
Loss	An integer indicative of a model's ability to predict or classify an example. A perfect model has a loss value of 0, unless the model has overfitted.
Matthew's Correlation Coefficient	A correlation coefficient between targets and predictions in a binary classification problem, where 1 is correct, 0 is random, and -1 is incorrect classification. $$\frac{TP \cdot TN\text{-}FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$
Mean Absolute Error	Metric often used to assess the quality of regression models $$\frac{1}{n}\sum_{j=1}^{n}\left|y_j\text{-}\hat{y}_j\right|$$ Where the prediction error is the actual value – the predicted value and the mean absolute error is the average of all absolute prediction errors.
Monastery	Building occupied by monks or canons (people who dedicate their lives to God)
Ogam, Ogham	Pre-Christian Irish alphabet
Precision value	A measure of how precise a model is for positive predictions $$\frac{True\ Positives}{(True\ Positives + False\ Positives)}$$
Radargram	A radar image of a profile of subsurface material(s) with X or Y, and Z values (see also *B-scan*)
Recall	A measure of the effect of false negatives on positive predictions $$\frac{True\ Positives}{True\ Positives + False\ Negative}$$
Reilig	Churchyard

Sampling strategy	Systematic sampling of a survey area
Spatial resolution	The intervals at which data are collected within a survey area, often given in the "sampling interval x traverse interval" format (e.g. 0.125m x 1m)
Stride	Distance (in pixels) the filter moves between applications, usually 1 or 2 but can overlap when using a 3x3 filter size.
Thermoremanence	Phenomenon that occurs when materials are heated and subsequently cool, creating the permanent thermoremanent magnetism that is detected through geophysical survey
Time slice	A radar image of multiple combined profiles of subsurface material(s) with X, Y, and Z values (see also *C-scan*)
True negative	1. An area of no geophysical anomalies corroborated through archaeological evaluation 2. Predicted value is negative, actual value is negative
True positive	1. Positive correlation between a geophysical response identified in a dataset and an excavated feature 2. Predicted value is positive, actual value is positive

Appendix B. GPR Surveys in Ireland, 1997-2006

Table 29: Details of the detection device consent licences from the National Monuments Service, Ireland (1997-2006). Post-2006 licences were not available.

Site	Townland	County	Year	License #	Applicant	Coordinates	NGR
Clonmacnoise Bridge	Clonmacnoise	Offaly	1997	97R0005	Aidan O'Sullivan/Donal Boland	2000962E, 230644N	N 00962 30644
Clonmacnoise	Clonmacnoise	Offaly	1997	97R0017	Harold Mytum	200810E, 230579N	N 00810 30579
Mullaghkeeran	Banagher	Offaly	1997	97R0018	Harold Mytum	200978E, 215491N	N 00978 15491
Clonmacnoise	Clonmacnoise	Offaly	1999	99R011	Kevin Barton/Deirdre O'Hara	200924E, 230618N	N 00924 30618
Rathcroghan and Carnfree	Various	Roscommon	1999	99R014	Joe Fenwick	183330E, 281049N	M 83330 81049
Rathcroghan and Carnfree	Various	Roscommon	2000	99R015	Joe Fenwick	182453E, 278459N	M 82453 78459
Newtown Deserted Settlement	Newtown	Cork	1999	99R015	Kevin Barton	173964E, 064353N	W 73864 64353
Caherquin	Caherquin	Kerry	1999	99R041	GeoArc Ltd	035847E, 105047N	Q 35847 05047
Churchyard of St Finbarre's Cathedral		Cork	1999	99R043	Kevin Barton	167006E, 071424N	W 67006 71424
Clonmacnoise	Clonmacnoise	Offaly	2000	00R021	Deirdre O'Hara	200924E, 230618N	N 00924 30618
Tankardstown	Tankardstown	Waterford	2000	00R022	Kevin Barton/Shane Rooney	245050E, 099184N	X 45050 99184
Ballykilcline	Ballykilcline	Roscommon	2000	00R032	Kevin Barton/Louise Geraghty	199911E, 286422N	M 99911 86422
Dookinelly (Calvy)	Dookinelly	Mayo	2000	00R033	Kevin Barton/Louise Geraghty	065767E, 304147N	F 65767 04147
Slievemore 'Famine Village'	Slievemore	Mayo	2000	00R034	Kevin Barton/Louise Geraghty	065040E, 307821N	F 65040 07821
Lough Corrib		Galway/Mayo	2000	00R060	Deirdre O'Hara	115354E, 250331N	M 15354 50331
Castletown Harbour	Drishane/Castletownend	Cork	2001	01R041	Connie Kelleher	067404E, 045639N	V 67404 45639
Caherquin	Caherquin	Kerry	2001	01R082	Louise Geraghty	035736E, 104949N	Q 35736 04949
Trim	Trim	Meath	2001	01R101	Hartmut Krahn	280752E, 256656N	N 80752 56656
Rathfarnham Church & Graveyard	Rathfarnham	Dublin	2002	02R044	Chris Gaffney	314369E, 228675N	O 14369 28675
Pollagh, Roscahill	Pollagh	Galway	2002	02R083	Colin Brown	124309E, 228856N	M 24309 28856

Templemichael, Longford	Abbeycarton	Longford	2002	02R189	Shane Rooney	213334E, 276135N	N 13334 76135
Maynooth Castle	Maynooth	Kildare	2003	03R001	Paul Gibson/John Bradley	293576E, 237639N	N 93576 37639
Barnasrahy, Killaspugbrone	Barnasrahy	Sligo	2003	03R029	Joe Fenwick	165660E, 335136N	G 65660 35136
Newcastle West	Newcastle West	Limerick	2003	03R057	GeoArc Ltd	128280E, 133668N	R 28280 33668
Barnagore	Barnagore	Cork	2003	03R082	GeoArc Ltd	166241E, 099801N	W 66241 99801
Derragh, Abbeylara	Derragh	Longford	2003	03R101	GeoArc Ltd	239965E, 278772N	N 39965 78772
Earthwork at Tulsk	Tulsk	Roscommon	2003	03R103	Niall Brady	183413E, 281079N	M 83413 81079
Moated site at Cloonfree	Cloonfree	Roscommon	2003	03R104	Niall Brady	190381E, 279990N	M 90381 79990
Foreshore at Ard West Pier	Carna	Galway	2003	03R136	Kevin Barton	075744E, 230984N	L 75744 30984
Dowth	Dowth	Meath	2004	04R001	Paul Gibson/George Eogan	302364E, 273776N	O 02364 73776
Anneville	Anneville	Meath	2004	04R022	Paul Gibson	266090E, 246027N	N 66090 46027
Gardens	County Kilkenny	Kilkenny	2004	04R026	John Nicholls	250779E, 156105N	S 50779 56105
Lemanaghan	Lemanaghan	Offaly	2004	04R088	Paul Gibson	216289E, 226791N	N 16289 26791
Killeigh	Killeigh	Offaly	2005	05R015	Paul Gibson	236523E, 218197N	N 36523 18197
Ballyboggan	Ballyboggan	Meath	2005	05R019	Paul Gibson	263914E, 241933N	N 63914 41933
Donaghmore	Donaghmore	Kildare	2005	05R020	Paul Gibson	296260E, 237082N	N 96260 37082
Taghadue	Taghadue	Kildare	2005	05R021	Paul Gibson	292432E, 234619N	N 92432 34619
Kilmainham	Dublin	Dublin	2005	05R034	Heather Gimson	312610E, 233711N	O 12610 33711
Kilskeer	Kilskeer	Meath	2006	06R022	Paul Gibson	265834E, 271623N	N 65834 71623
Rahan	Rahan	Offaly	2006	06R023	Paul Gibson	225655E, 225437N	N 25655 25437
Not Available	Magheraboy	Sligo	2006	06R044	Hartmut Krahn	168039E, 335059N	G 68039 35059
Centre & Customs House	Waterford City	Waterford	2006	06R049	Heather Gimson	260955E, 112484N	S 60955 12484

Appendix C. Reports for Archaeological Excavations where Early Medieval – Later Medieval Period Burials Were Recovered

Table 30: Records of the archaeological excavation licences granted for evaluations on Irish road schemes (2001-2013), where burials dating to the Early Medieval – Later Medieval periods were recovered

Site	Period(s) of Activity	Consultancy	Year	Licence No.	Reference
Platin Fort Report on resolution of Site, County Meath	Early Medieval	Archaeological Consultancy Services Ltd.	2001	01E0044	https://repository.dri.ie/catalog/5x226w173
Claristown 2, County Meath	Early Christian Medieval	Archaeological Consultancy Services Ltd.	2001	01E0039	https://repository.dri.ie/catalog/5425zr452
Johnstown 1 Vol 7 Appendices 7-16, County Meath	Early Medieval	Archaeological Consultancy Services Ltd.	2002	02E0462	https://repository.dri.ie/catalog/pg15qv68w
Johnstown 1 Vol 6 Appendices 2-6, County Meath	Early Medieval	Archaeological Consultancy Services Ltd.	2002	02E0462	https://repository.dri.ie/catalog/p841p300v
Johnstown 1 Vol 5 Appendix 1 Part 2, County Meath	Early Medieval	Archaeological Consultancy Services Ltd.	2002	02E0462	https://repository.dri.ie/catalog/p267m932n
Johnstown 1 Vol 4 Appendix 1 Part 1, County Meath	Early Medieval	Archaeological Consultancy Services Ltd.	2002	02E0462	https://repository.dri.ie/catalog/nv93jh644
Johnstown 1 Vol 3 Plates & Illustrations, County Meath	Early Medieval	Archaeological Consultancy Services Ltd.	2002	02E0462	https://repository.dri.ie/catalog/np19gq96z
Johnstown 1 Vol 2 Figures, County Meath	Early Medieval	Archaeological Consultancy Services Ltd.	2002	02E0462	https://repository.dri.ie/catalog/ng45dz288
Johnstown 1 Vol 1 Final Report, County Meath	Early Medieval	Archaeological Consultancy Services Ltd.	2002	02E0462	https://repository.dri.ie/catalog/n871c560f
Site 120 Fort Hill, County Louth	Early Medieval	Irish Archaeological Consultancy Ltd	2002	02E1326	https://repository.dri.ie/catalog/v6936m97g

Site 121 Balriggan Volume 1, County Louth	Early Medieval	Irish Archaeological Consultancy Ltd	2002	02E1325	https://repository.dri.ie/catalog/v1194v293
Site 121 Balriggan 1 Volume 2, County Louth	Early Medieval	Irish Archaeological Consultancy Ltd	2002	02E1325	https://repository.dri.ie/catalog/tt4532617
Harlockstown Site 19, County Meath	Early Medieval	CRDS Ltd	2003	03E1526	https://repository.dri.ie/catalog/5425zr495
Cookstown Site 25 Vol 3 Specialist Reports, County Meath	Early Medieval High Medieval	CRDS Ltd	2003	03E1252	https://repository.dri.ie/catalog/1r66xf93b
Cookstown Site 25 Vol 2, County Meath	Early Medieval High Medieval	CRDS Ltd	2003	03E1252	https://repository.dri.ie/catalog/1j92vp25n
Cookstown Site 25 Vol 1, County Meath	Early Medieval High Medieval	CRDS Ltd	2003	03E1252	https://repository.dri.ie/catalog/1c18sw57g
Raystown Site 21, County Meath	Early Medieval	CRDS Ltd	2003	03E1229	https://repository.dri.ie/catalog/hh6469688
Sites 2,5,6,12,13,14 Morett, County Laois	Early Christian Early Medieval	Archaeological Consultancy Services Ltd.	2003	03E1367	https://repository.dri.ie/catalog/9g559z42b
Site 109 Cloghvally Upper 1, County Monaghan	Early Medieval	Irish Archaeological Consultancy Ltd	2003	03E1255	https://repository.dri.ie/catalog/8p592t70d
Site 110 Monanny 1, County Monaghan	Early Medieval	Irish Archaeological Consultancy Ltd	2003	03E0888	https://repository.dri.ie/catalog/6w92hs58z
Site D Morett, County Laois	Early Medieval Late Medieval	Valerie J Keeley Ltd	2003	03E0461	https://repository.dri.ie/catalog/0z70pb18m
Ballydavis Site 1, County Laois	Early Medieval	Valerie J Keeley Ltd	2003	03E0151	https://repository.dri.ie/catalog/zs269q217
26 Sites, County Laois	Medieval	Archaeological Consultancy Services Ltd.	2003	03E0623	https://repository.dri.ie/catalog/pk02rs957
Kilshane Site 5 Vol 3 Figs and Plates, County Dublin	Medieval	CRDS Ltd	2003	03E1359	https://repository.dri.ie/catalog/4b29qm777
03E1359 Kilshane Site 5 Vol 2 Appendices, County Dublin	Medieval	CRDS Ltd	2003	03E1359	https://repository.dri.ie/catalog/4455nv09j
03E1359 Kilshane Site 5 Vol 1 Text, County Dublin	Medieval	CRDS Ltd	2003	03E1359	https://repository.dri.ie/catalog/3x81m241p
Site 126 Carn More 4, County Louth	Medieval	Irish Archaeological Consultancy Ltd	2003	03E0872	https://repository.dri.ie/catalog/6970bf548

Morett 15, County Laois	Early Christian	Archaeological Consultancy Services Ltd.	2003	03E1624	https://repository.dri.ie/catalog/b851k314f
Kilcloghans, County Galway	Early Medieval	Headland Archaeology Ltd.	2006	06E1139	https://repository.dri.ie/catalog/0c48h401z
Marlhill, County Tipperary	Early Medieval	Margaret Gowan & Co. Ltd	2007	E2124	https://repository.dri.ie/catalog/g445rv013
Ross 2, County Meath	Early Medieval	Archaeological Consultancy Services Ltd.	2008	E3381	https://repository.dri.ie/catalog/6h44d7249
Loughbown 1, County Galway	Medieval	Eachtra Archaeological Projects	2008	E2442	https://repository.dri.ie/catalog/0000cd82v
Ballybar Lower 3, County Carlow.	Early Medieval	Headland Archaeology Ltd.	2009	E2620	https://repository.dri.ie/catalog/mp495z408
Ballybar Upr 2, County Carlow.	Early Medieval	Headland Archaeology Ltd.	2009	E2588	https://repository.dri.ie/catalog/r2087792c
Borris AR 33 Vol 2, County Tipperary	Early Medieval	Valerie J Keeley Ltd	2009	E2376	https://repository.dri.ie/catalog/8623xc627
Borris AR 33 Vol 1, County Tipperary	Early Medieval	Valerie J Keeley Ltd	2009	E2376	https://repository.dri.ie/catalog/8336wg78w
Ballykilmore 6 Volume 3, County Westmeath	Early Medieval Medieval	Valerie J Keeley Ltd	2009	E2798	https://repository.dri.ie/catalog/vd678d713
Ballykilmore 6 Volume 2, County Westmeath	Early Medieval Medieval	Valerie J Keeley Ltd	2009	E2798	https://repository.dri.ie/catalog/v6936n03d
Ballykilmore 6 Volume 1, County Westmeath	Early Medieval Medieval	Valerie J Keeley Ltd	2009	E2798	https://repository.dri.ie/catalog/v4065r193
Clonfad 3 Volume 2, County Westmeath	Early Medieval Late Medieval	Valerie J Keeley Ltd	2009	E2723	https://repository.dri.ie/catalog/cj82zp112
Clonfad 3 Vol 1, County Westmeath	Early Medieval Late Medieval	Valerie J Keeley Ltd	2009	E2723	https://repository.dri.ie/catalog/cc08ww43w
Clonfad 3 Plates, County Westmeath	Early Medieval Late Medieval	Valerie J Keeley Ltd	2009	E2723	https://repository.dri.ie/catalog/c534v375k

Clonfad 3 Figures, County Westmeath	Early Medieval Late Medieval	Valerie J Keeley Ltd	2009	E2723	https://repository.dri.ie/catalog/bz60sb07v
E2220 Bushfield or Maghernaskeagh or Lismore 1, County Laois	Early Medieval	Archaeological Consultancy Services Ltd.	2009	E2220	https://repository.dri.ie/catalog/7w62tq037
Parknahown 2, County Laois	Early Medieval	Archaeological Consultancy Services Ltd.	2009	E2196	https://repository.dri.ie/catalog/js95kw61s
Parknahown 5 Vol I Text., County Laois	Early Medieval	Archaeological Consultancy Services Ltd.	2009	E2170	https://repository.dri.ie/catalog/d21844297
Parknahown 5 Vol IV Images, County Laois	Early Medieval	Archaeological Consultancy Services Ltd.	2009	E2170	https://repository.dri.ie/catalog/cv442b616
Parknahown 5 Vol III Osteo Rpt, County Laois	Early Medieval	Archaeological Consultancy Services Ltd.	2009	E2170	https://repository.dri.ie/catalog/cn700j93q
Parknahown 5 Vol II Append 1_17, County Laois	Early Medieval	Archaeological Consultancy Services Ltd.	2009	E2170	https://repository.dri.ie/catalog/cf95xs255
Treanbaun, County Galway	Early Medieval	CRDS Ltd	2009	E2123	https://repository.dri.ie/catalog/9p29cr13n
Cross, County Galway	Early Medieval	CRDS Ltd	2009	E2069	https://repository.dri.ie/catalog/7w62tq01p
Carrowkeel Vol 2, County Galway	Early Medieval Late Medieval	Headland Archaeology Ltd	2009	E2046	https://repository.dri.ie/catalog/6970bf573
Tinryland 1, County Carlow	Medieval	Headland Archaeology Ltd	2009	E2589	https://repository.dri.ie/catalog/qf861w88z
Busherstown 4, County Carlow	Medieval	Headland Archaeology Ltd	2009	E2581	https://repository.dri.ie/catalog/pn89ss16p
Russellstown 1, County Carlow	Medieval	Headland Archaeology Ltd	2009	E2571 E2572	https://repository.dri.ie/catalog/n2979h72k
Killeany 1, County Laois	Medieval	Archaeological Consultancy Services Ltd.	2009	E2171	https://repository.dri.ie/catalog/d7925w97f
Lismullin 1 Vol 3 Images, County Meath	Medieval	Archaeological Consultancy Services Ltd.	2009	E3074	https://repository.dri.ie/catalog/5h743983m

Lismullin 1 Vol 2 Spec Rpts, County Meath	Medieval	Archaeological Consultancy Services Ltd.	2009	E3074	https://repository.dri.ie/catalog/5b001j15z
Lismullin 1 Vol 1 Text, County Meath	Medieval	Archaeological Consultancy Services Ltd.	2009	E3074	https://repository.dri.ie/catalog/5425zr47m
Mullagh Site 2, County Longford	Medieval Late Medieval	CRDS Ltd	2009	09E0314	https://repository.dri.ie/catalog/k069np260
Owenbristy, County Galway	Early Medieval	Eachtra Archaeological Projects	2010	E3770	https://repository.dri.ie/catalog/6970bf68m
Rossbrien Site 1, County Limerick	Early Medieval	Irish Archaeological Consultancy Ltd	2010	E3933	https://repository.dri.ie/catalog/fq97n866t
Borris AR 36, County Tipperary	Early Medieval	Valerie J Keeley Ltd	2010	E2491	https://repository.dri.ie/catalog/dr27bc267
Gortmakellis AR 1, County Tipperary	Early Medieval	Valerie J Keeley Ltd	2010	E2356	https://repository.dri.ie/catalog/5x226w30n
Richhill Site 2, County Limerick	Early Medieval	Headland Archaeology Ltd	2010	E2311	https://repository.dri.ie/catalog/0k22jr93z
Lowpark, County Mayo	Early Medieval	Mayo County Council	2010	E3338	https://repository.dri.ie/catalog/gf06vh467
Carrowkeel Vol 1, County Galway	Early Medieval Late Medieval	Headland Archaeology Ltd	2010	E2046	https://repository.dri.ie/catalog/63968n89w
Cappydonnell Big 1, County Offaly	Medieval	Irish Archaeological Consultancy Ltd	2010	E2653	https://repository.dri.ie/catalog/2z119502c
Ardagawna 1, County Roscommon	Medieval	Valerie J Keeley Ltd	2010	E3270	https://repository.dri.ie/catalog/0k22jr840
Ballintotty Site 2, County Tipperary	Early Medieval	Aegis Archaeology Ltd	2011	E2935	https://repository.dri.ie/catalog/6t05gw85n
Camlin 3, County Tipperary	Early Medieval	Valerie J Keeley Ltd	2011	E3580	https://repository.dri.ie/catalog/q237x674t
Faughart Lower Vol 2 Area 15 site 116, County Louth	Early Medieval	Archaeological Development Services Ltd.	2011	E3801	https://repository.dri.ie/catalog/bn99pn52z
Faughart Lower Vol 1 Area 15 Site 116, County Louth	Early Medieval	Archaeological Development Services Ltd.	2011	E3801	https://repository.dri.ie/catalog/bg25mv84h
Kilmainham 1C Vol 1 Text, County Meath	Early Medieval Medieval	Irish Archaeological Consultancy Ltd	2011	E3140	https://repository.dri.ie/catalog/nv93jh63v

Kilmainham 1C Images, County Meath	Early Medieval Medieval	Irish Archaeological Consultancy Ltd	2011	E3140	https://repository.dri.ie/catalog/np19gq95p
Kilmainham 1C Figure 61-83, County Meath	Early Medieval Medieval	Irish Archaeological Consultancy Ltd	2011	E3140	https://repository.dri.ie/catalog/ng45dz270
Kilmainham 1C Figure 41-60, County Meath	Early Medieval Medieval	Irish Archaeological Consultancy Ltd	2011	E3140	https://repository.dri.ie/catalog/n871c559p
Kilmainham 1C Figure 21-40, County Meath	Early Medieval Medieval	Irish Archaeological Consultancy Ltd	2011	E3140	https://repository.dri.ie/catalog/n2979c91z
Kilmainham 1C Figure 1-20, County Meath	Early Medieval Medieval	Irish Archaeological Consultancy Ltd	2011	E3140	https://repository.dri.ie/catalog/mw237m238
Kilmainham 1C Volume 2, County Meath	Early Medieval Medieval	Irish Archaeological Consultancy Ltd	2011	E3140	https://repository.dri.ie/catalog/mp495t552
Gardenhill Site 1, County Limerick.	Medieval	Aegis Archaeology Ltd	2011	E2320	https://repository.dri.ie/catalog/1g05ts491
Aghnaskeagh Area 12 Site 111, County Louth	Early Christian	Archaeological Development Services Ltd.	2011	E3793	https://repository.dri.ie/catalog/8w334m405
Rathmorrissy, County Galway	Early Medieval	Headland Archaeology Ltd	2012	E4024	https://repository.dri.ie/catalog/2z119503n
Busherstown, County Offaly	Medieval	Eachtra Archaeological Projects	2012	E3661	https://repository.dri.ie/catalog/vt15d0061
Kellymount 4, County Kilkenny	Early Medieval	Irish Archaeological Consultancy Ltd	2013	E3857	https://repository.dri.ie/catalog/dz01d826b

Appendix D. Grave-like Responses in GPR Data Questionnaire

Identification of Grave Responses in GPR Data

1. How many years experience do you have in GPR survey? Has your experience with GPR survey primarily been in research or commercial settings?

☐ Less than 1 year

☐ 1 - 3 years

☐ 3 - 5 years

☐ 5 - 7 years

☐ 7 - 10 years

☐ More than 10 years

☐ Research

☐ Commercial

2. When interpreting GPR data, you:

☐ Use radargrams

☐ Use timeslices

☐ Use animations

☐ Use isosurfaces

☐ Use automatic feature detection

☐ Other (please specify)

3. In the image provided which quadrant, if any, contains grave-like response(s)? How confident are you in your interpretation?

Notes:
Black indicates high amplitude reflections.
Depth below ground level is provided in ns and cm.
The bedrock geology is White Chalk.
Data were collected using a MALA RAMAC X3M with 500MHz central frequency antenna.

	Grave Present	Multiple Graves Present	High Confidence	Medium Confidence	Low Confidence
A	☐	☐	☐	☐	☐
B	☐	☐	☐	☐	☐
C	☐	☐	☐	☐	☐
D	☐	☐	☐	☐	☐

4. In the image provided which quadrant, if any, contains grave-like response(s)? How confident are you in your interpretation?

Notes:
Black indicates high amplitude reflections.
Depth below ground level is provided in ns and m.
The bedrock geology is Lincolnshire Limestone.
Data were collected using a MALA RAMAC X3M with 500MHz central frequency antenna.

	Grave Present	Multiple Graves Present	High Confidence	Medium Confidence	Low Confidence
A	☐	☐	☐	☐	☐
B	☐	☐	☐	☐	☐
C	☐	☐	☐	☐	☐
D	☐	☐	☐	☐	☐

5. In the image provided which quadrant, if any, contains grave-like response(s)? How confident are you in your interpretation?

Notes:
Black indicates high amplitude reflections.
Depth below ground level is provided in ns and cm.
The bedrock geology is dark limestone and shale.
Data were collected using a MALA RAMAC X3M with 500MHz central frequency antenna.

	Grave Present	Multiple Graves Present	High Confidence	Medium Confidence	Low Confidence
A	☐	☐	☐	☐	☐
B	☐	☐	☐	☐	☐
C	☐	☐	☐	☐	☐
D	☐	☐	☐	☐	☐

6. In the image provided which quadrant, if any, contains grave-like response(s)? How confident are you in your interpretation?

Notes:
Black indicates high amplitude reflections.
Depth below ground level is provided in ns and cm/m.
The bedrock geology is grey-green metagreywackes and slates.
Data were collected using a MALA RAMAC X3M with 500MHz central frequency antenna.

	Grave Present	Multiple Graves Present	High Confidence	Medium Confidence	Low Confidence
A	☐	☐	☐	☐	☐
B	☐	☐	☐	☐	☐
C	☐	☐	☐	☐	☐
D	☐	☐	☐	☐	☐

Appendix E.   Individual Questionnaire Responses

Q1	Respondent 1	3 - 5 years
How many years' experience do you have in GPR survey? Has your experience with GPR survey primarily been in research or commercial settings?		Research
	Respondent 2	More than 10 years
		Research
	Respondent 3	More than 10 years
		Research
		Commercial
	Respondent 4	More than 10 years
		Commercial
	Respondent 5	More than 10 years
		Research
		Commercial
	Respondent 6	7 - 10 years
		Research
		Commercial
	Respondent 7	More than 10 years
		Research
		Commercial
	Respondent 8	1 - 3 years
		Research
	Respondent 9	More than 10 years
		Commercial
	Respondent 10	1 - 3 years

Q2	Respondent 1	Use radargrams
When interpreting GPR data, you:		Use timeslices
		Use animations
	Respondent 2	Use timeslices
		Use animations
		Other (please specify):
		Envelope of timeslices, and usually no more sophisticated processing than that.
	Respondent 3	Use radargrams
		Use timeslices
	Respondent 4	Use radargrams
		Use timeslices
		Use animations
	Respondent 5	Use radargrams
		Use timeslices
	Respondent 6	Use radargrams
		Use timeslices
		Use animations
	Respondent 7	Use radargrams
		Use timeslices
		Use animations
	Respondent 8	Use radargrams
		Use timeslices
	Respondent 9	Use radargrams
		Use timeslices
	Respondent 10	Use radargrams
		Use timeslices
		Use animations

| Q3 | Respondent 1 | A Grave Present, Medium Confidence |
| | | B Grave Present, Low Confidence |

233

In the image below which quadrant(s), if any, contains grave-like response(s)? How confident are you in your interpretation? **Notes: Depth below ground level is provided in ns and cm. The bedrock geology is White Chalk. Data were collected using a MALA RAMAC X3M with 500MHz centre frequency antenna.**	Respondent 2	A Grave Present, Low Confidence B Multiple Graves Present, Medium Confidence C Grave Present, Low Confidence D High Confidence
	Respondent 3	D Multiple Graves Present, Low Confidence
	Respondent 4	C Grave Present, Low Confidence
	Respondent 5	A Low Confidence B Low Confidence C Grave Present, High Confidence D Low Confidence
	Respondent 6	B Grave Present, Low Confidence C Grave Present, Low Confidence
	Respondent 7	A Grave Present, Low Confidence B Multiple Graves Present, Low Confidence
	Respondent 8	A Low Confidence B Low Confidence C Low Confidence D Low Confidence
	Respondent 9	A Grave Present, Medium Confidence B Multiple Graves Present, Medium Confidence
	Respondent 10	A Multiple Graves Present B Grave Present, High Confidence, Medium Confidence C High Confidence D High Confidence
Q4 **In the image below which quadrant(s), if any, contains grave-like response(s)? How confident are you in your interpretation?** **Notes: Depth below ground level is provided in ns and m. The bedrock geology is limestone. Data were collected using a MALA RAMAC X3M with 500MHz centre frequency antenna.**	Respondent 1	A Multiple Graves Present, Medium Confidence B Medium Confidence, Grave Present
	Respondent 2	A Multiple Graves Present, Medium Confidence B Grave Present, High Confidence C High Confidence D High Confidence
	Respondent 3	A Grave Present, Medium Confidence
	Respondent 4	A Grave Present, Multiple Graves Present, Medium Confidence B Grave Present, Low Confidence
	Respondent 5	A Grave Present, Low Confidence
	Respondent 6	A Multiple Graves Present, Medium Confidence B Grave Present, Low Confidence
	Respondent 7	B Grave Present, Low Confidence
	Respondent 8	A Low Confidence B Low Confidence C Low Confidence D Low Confidence
	Respondent 9	A Medium Confidence, Multiple Graves Present B Grave Present, Medium Confidence
	Respondent 10	A Multiple Graves Present B High Confidence, Medium Confidence, Grave Present C Multiple Graves Present D High Confidence
Q5 **In the image below which quadrant(s), if any, contains grave-like response(s)? How confident are you in your interpretation?**	Respondent 1	A Grave Present, Low Confidence D Multiple Graves Present, Medium Confidence
	Respondent 2	A Multiple Graves Present, Medium Confidence B Medium Confidence C Multiple Graves Present, Medium Confidence D High Confidence, Multiple Graves Present
	Respondent 3	A Grave Present, Medium Confidence

Notes: Black indicates high amplitude reflections. Depth below ground level is provided in ns and cm. The bedrock geology is dark limestone and shale. Data were collected using a MALA RAMAC X3M with 500MHz centre frequency antenna.		D Medium Confidence, Multiple Graves Present
	Respondent 4	A Grave Present, Low Confidence
		D Multiple Graves Present, Low Confidence
	Respondent 5	A Grave Present, Low Confidence
		C Grave Present, Medium Confidence
		D Multiple Graves Present, Low Confidence
	Respondent 6	D Multiple Graves Present, Medium Confidence
	Respondent 7	D Multiple Graves Present, Low Confidence
	Respondent 8	A Low Confidence
		B Low Confidence
		C Medium Confidence
		D Low Confidence
	Respondent 9	Multiple Graves Present, Low Confidence
		C Multiple Graves Present, Low Confidence
		D Multiple Graves Present, Medium Confidence
	Respondent 10	A Grave Present
		B Multiple Graves Present
		C High Confidence
		D Medium Confidence
Q6 In the image below which quadrant(s), if any, contains grave-like response(s)? How confident are you in your interpretation? **Notes:** Black indicates high amplitude reflections. Depth below ground level is provided in ns and cm. The bedrock geology is grey-green metagreywackes and slates. Data were collected using a MALA RAMAC X3M with 500MHz centre frequency antenna.	Respondent 1	A Low Confidence
		B Low Confidence
		C Low Confidence
		D Low Confidence
	Respondent 2	A Medium Confidence
		B Medium Confidence
		C Grave Present, Low Confidence
		D Medium Confidence, Grave Present
	Respondent 3	Respondent skipped this question
	Respondent 4	A Grave Present, Low Confidence
	Respondent 5	A Multiple Graves Present, Low Confidence
	Respondent 6	Respondent skipped this question
	Respondent 7	A Multiple Graves Present, Low Confidence
		B Multiple Graves Present, Low Confidence
		C Multiple Graves Present, Low Confidence
		D Multiple Graves Present, Low Confidence
	Respondent 8	A Low Confidence
		B Low Confidence
		C Low Confidence
		D Low Confidence
	Respondent 9	Respondent skipped this question
	Respondent 10	A Multiple Graves Present
		B Grave Present
		C High Confidence, Medium Confidence
		D High Confidence

Appendix F.   User Guide for Object Detection Web Apps

March 2020

Machine Learning for Geophysics Data
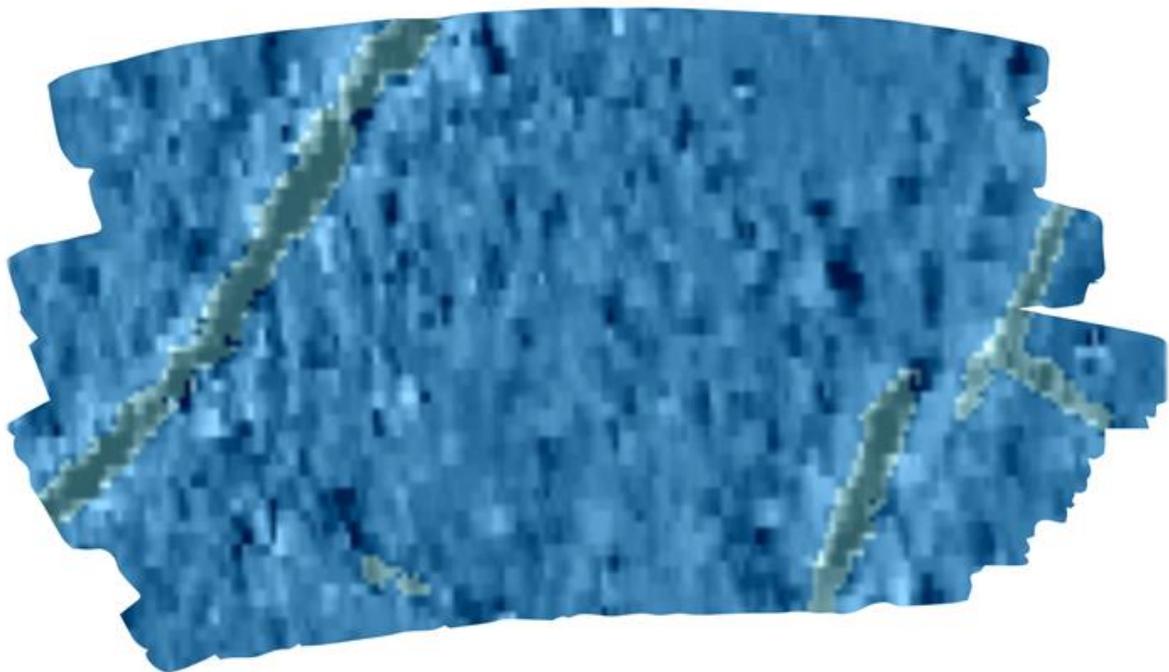
Ashely Green
Postgraduate Researcher
Department of Archaeology & Anthropology
Bournemouth University

Bournemouth University

Using the Reilig Web App for Inference

Requirements:

- Data that has been processed with bandpass filtering, background removal, and migration, if possible

- Data plotted on the custom template size derived from GPR-Slice:
  - **Origin** X=100, Y=240
  - **Length** X=544, Y=385
  - **Shift** X=1000, Y=500

- Image of data plotted in greyscale (with white being low amplitude responses and black being high amplitude responses)

- Image at least 150 pixels wide without padding, scale, or border around the data

Instructions:

1. Save jpeg images from processing software (GPR-Slice was used in this project) with all whitespace and scales surrounding the data removed.
2. Navigate to reilig.onrender.com.
3. Upload a single image using the '**Select Image**' button.
4. Run image through the machine learning model by clicking the '**Analyze**' button.
5. The result will return a single bounding box if a possible grave is detected in the image. No bounding box will be returned if there is no possible grave detected in the image. The confidence score of the detection is displayed below the returned image and uses a scale of 0-1.
6. If you would like to save the returned image and bounding box, you should screenshot the webpage or save it as a pdf.
7. If you would like to save only the bounding box interpretation, you can right-click on the image to save it.
8. Refresh the webpage to clear any previous detections, and repeat Steps 3-8 on any additional images.

**If your image is less than 150 pixels wide, you may wish to try mirroring the data rather than including whitespace or excluding the image.

The following is a brief introduction to the doctoral research carried out in order to create the web app and a general introduction to machine learning and its application to archaeological data.

1. Outline

1. Intro to machine learning and learning tasks

2. Training

3. Classification and Object Detection examples

4. Segmentation examples

5. Inferring with a trained model

6. Machine learning workflow

2. Machine Learning

Supervised Learning

Learns patterns from labelled data

Uses learned patterns to predict future outcomes

Semisupervised Learning

Trains on a small labelled dataset in order to label large training datasets

Uses learned patterns to predict future outcomes

Unsupervised Learning

Identifies hidden patterns in unlabelled data

Clusters data based on the degree of similarity or dissimilarity

In supervised learning the learning machine is provided with labelled data (input data with associated target value(s), such as an image broken down into labelled contents) and is expected to learn the data for target mapping, in order to predict (or estimate) target values for previously unobserved input data. The dataset (described as $D\{x, y\}_{n=1}^{N}$) includes input features (x) and labels (y), which are often images and class labels or bounding boxes with associated class labels for classification or object detection tasks (Litjens et al. 2017, p. 62). The following work was achieved using supervised transfer learning.

3. Classification vs Detection vs Segmentation

Image Classification
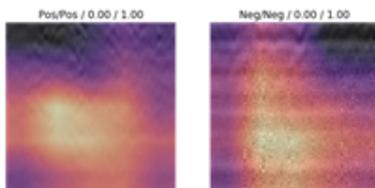
- Simplest Task
- Only determines if there a specified object in an image.
- The heatmap/activation map on the left demonstrates the image features the machine is interpreting.

Object Detection
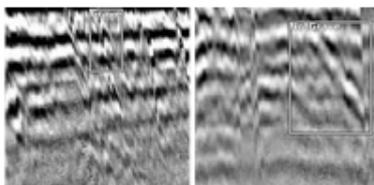
- Object Localisation
- Classifies a feature and determines a bounding box (location) for a specified feature in an image.
- The bounding boxes on the left demonstrate the image features the machine is interpreting.

Image Segmentation
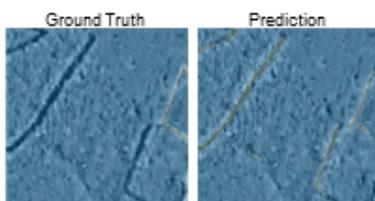
- Pixel-level Classification
- Classifies a pixels in a image into a specified number of classes (1…N). Instance segmentation will classify multiple objects of the same class as separate features.
- The segmentation maps on the left demonstrate the image features the machine is interpreting.

4. Generating Training Data

Bounding Boxes for Classification and Object Detection
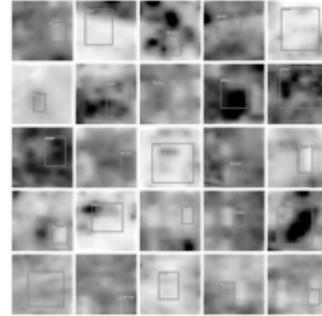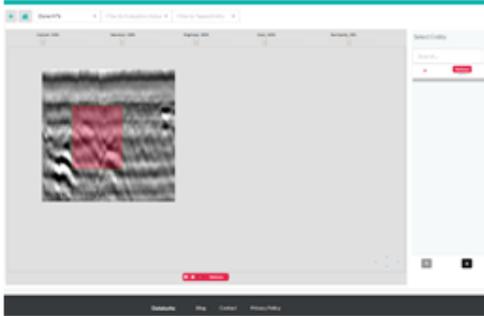
Training data for the object detection task were labelled with rectangular bounding boxes around target objects using Dataturks, a web browser based image labelling tool (shown above). This allowed bounding box classes and coordinates to be exported in json and xml formats for easy import into machine learning models. In the following examples, bounding boxes were assigned one of two classes – 'Grave' and 'Not Grave'.

Sketch Tool/Polygons for Segmentation

Image   Original Label   Converted Label (0...N)
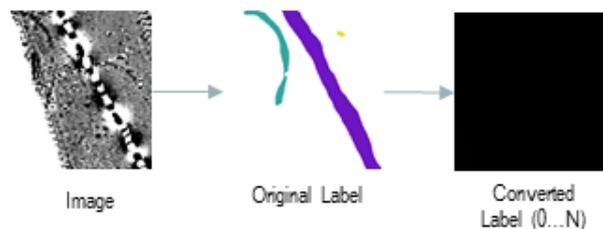
Training data for segmentation tasks were labelled with polygons around target objects. Data were labelled using Labelbox, a web browser based image labelling tool with text, rectangle, and polygon labelling capabilities (shown above) so that class and label information could be exported in json and image formats. In the segmentation example following, polygons were assigned one of three classes – 'Archaeology', 'Dipole' and 'Modern'.

The strings in the exported json file were used to download the original images and label masks and convert them to a scale of 0..N, corresponding to the number of classes used.

5. Convolutional Neural Networks for the Detection of Graves in Radargrams

Models were trained on common medieval grave types derived from:

Geophysical surveys in Ireland
Irish excavation reports

___

Training dataset consisted of real and simulated radargrams with:

537 Positive Examples
680 Negative Examples
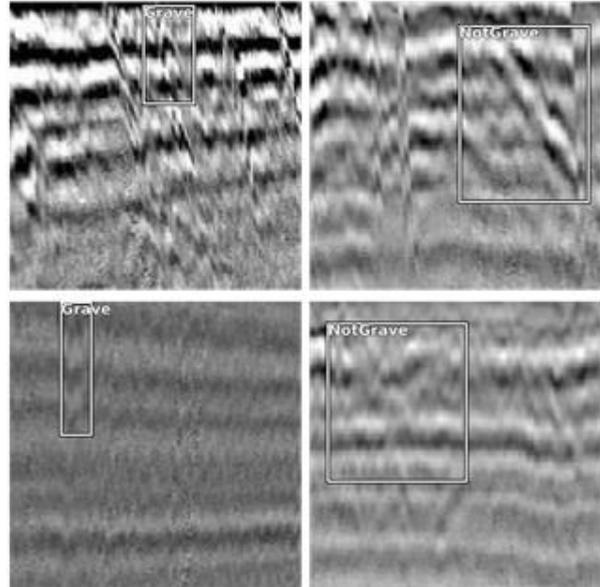
Radargram Metrics

	Accuracy	F$_1$ Score	MCC
VGG19_BN	0.8430	0.8295	0.6899
ResNet50	0.8750	0.8667	0.7538
ResNet101	0.8802	0.8878	0.7617
ResNet152	*0.9375*	*0.9362*	*0.8768*

94% Accuracy

6. Convolutional Neural Networks for the Detection of Graves in Timeslices

Models were trained on common proxies for modern clandestine graves derived from:

BU grey literature

Training dataset consisted of real timeslices with:

531 Positive Examples
480 Negative Examples

Ground Truth / Prediction

Timeslice Metrics

Mean Training Loss vs Mean Valid Loss

Mean Accuracy vs Mean Loss
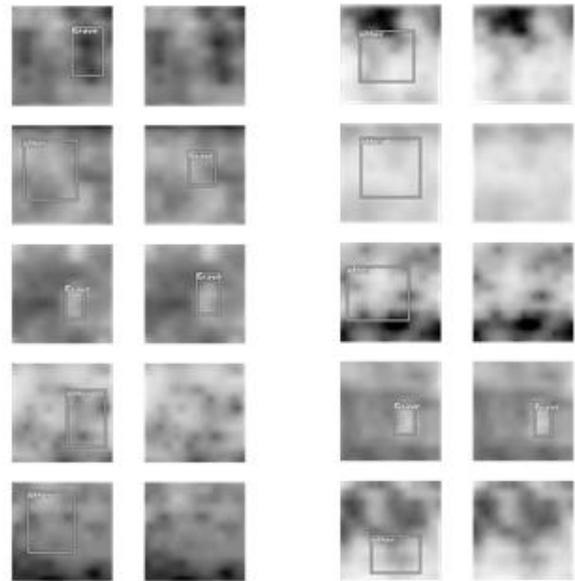
Prediction Samples
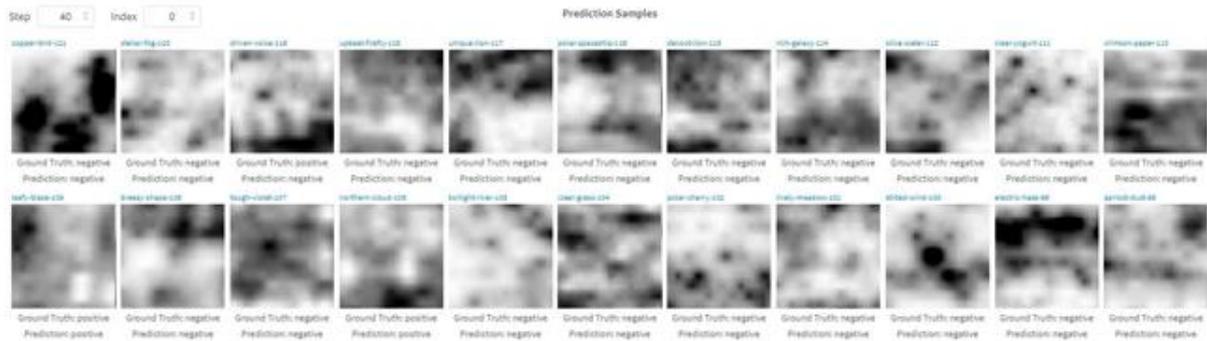
96% Accuracy

7. Convolutional Neural Networks for the Detection of Archaeological Features in Magnetometry Data

U-Net for Segmentation

- conv 3x3, ReLU
- copy and crop
- max pool 2x2
- up-conv 2x2
- conv 1x1
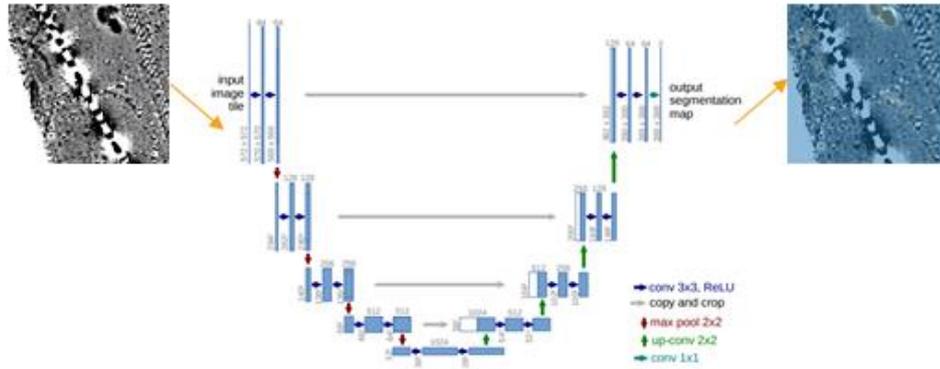
(Ronneberger et al. 2015)

Models were trained on archaeological and modern features in magnetometry data derived from:

BU grey literature

Training dataset consisted of magnetometry data with:

1148 Dipole Examples
1006 Archaeology Examples
 732 Modern Examples

Image   Label Mask

Segmentation Metrics

Training Loss vs Valid Loss

train_loss   valid_loss

Mean Accuracy

Ground Truth   Prediction   Ground Truth   Prediction   Ground Truth   Prediction
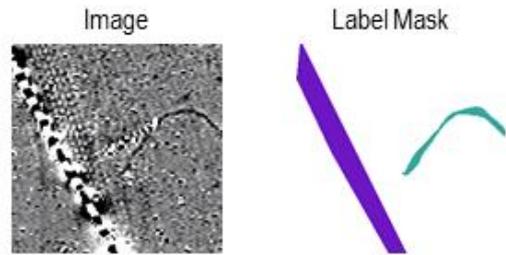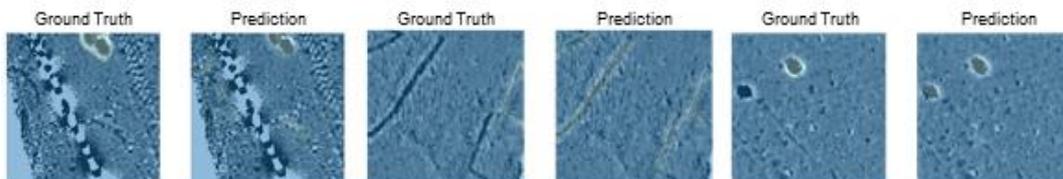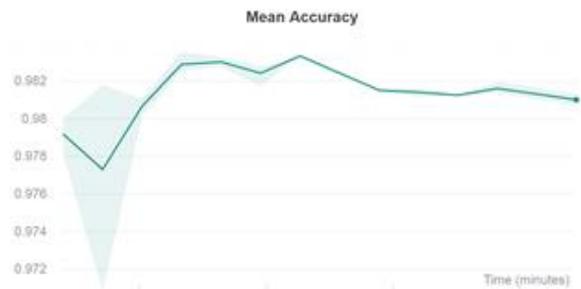
98% Accuracy

8. Inferring with Trained Models

Web App
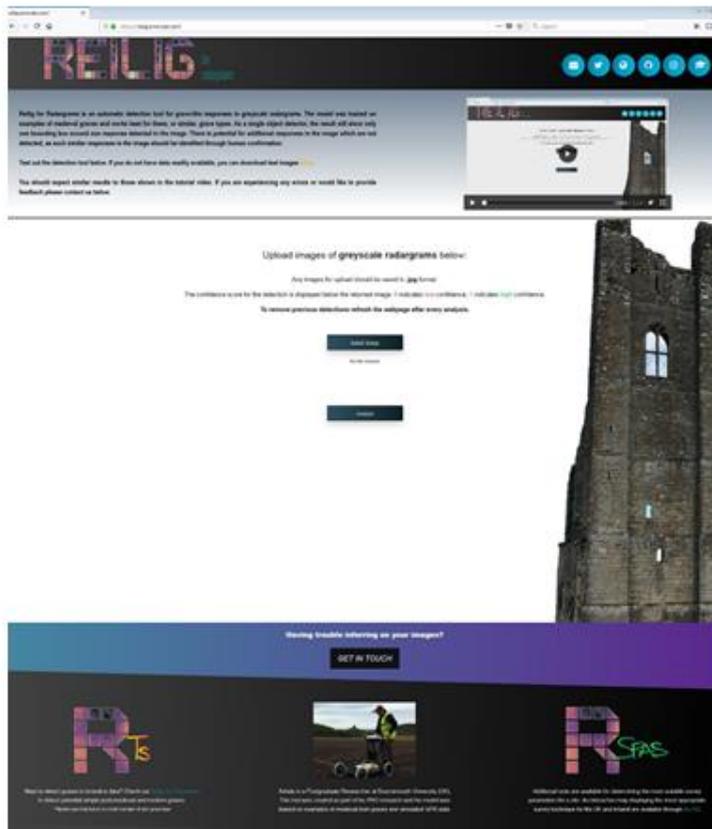
reilig.onrender.com   reiligts.onrender.com
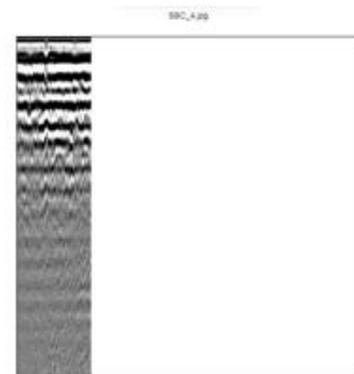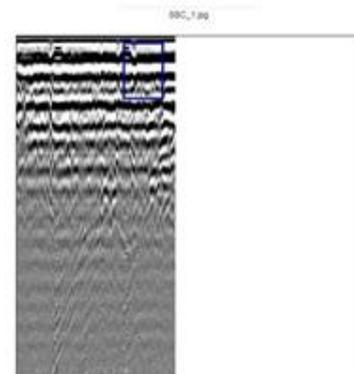
The web app, written with Javascript and hosted on Render, allows the user to upload a single image for prediction. The output is the original image with no bounding box output is there is not a grave in the image and a single bounding box if there is a grave in the image. The prediction is also returned with a confidence score on a scale of 0 -1.

Other Options

- ArcGIS Pro
- Collab Notebook – GPU and Cloud capabilities
- Jupyter Notebook
- Python executable

9. Proposed Workflow for Data Interpretation
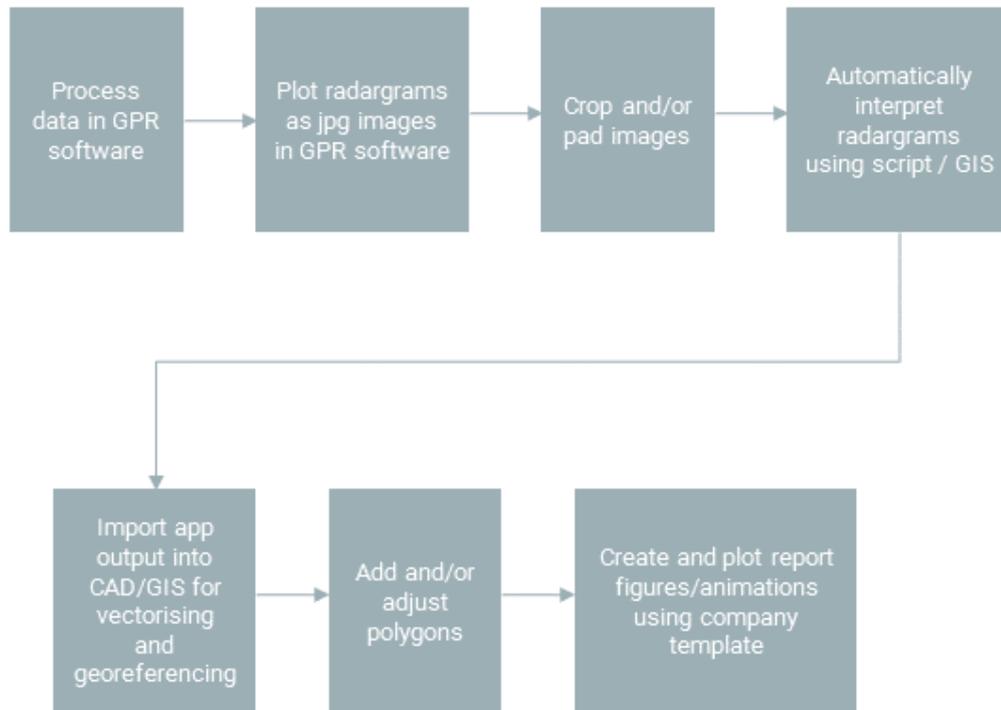
10. References

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Laak, J. A. W. M. van der, Ginneken, B. van, and Sánchez, C. I., 2017. A Survey on Deep Learning in Medical Image Analysis. *Medical Image Analysis* [online], 42, 60–88. Available from: https://www.sciencedirect.com/science/article/pii/S1361841517301135.

Ronneberger, O., Fischer, P., and Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015 [online]. Available from: https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28.

Appendix G.   Training Data

The following images are the training data for the object detection and classification models. The section headers indicate whether a grave is present in the image and bounding box data can be downloaded separately from the json file included in the digital appendices.

Radargram Training Data: Non-grave Images
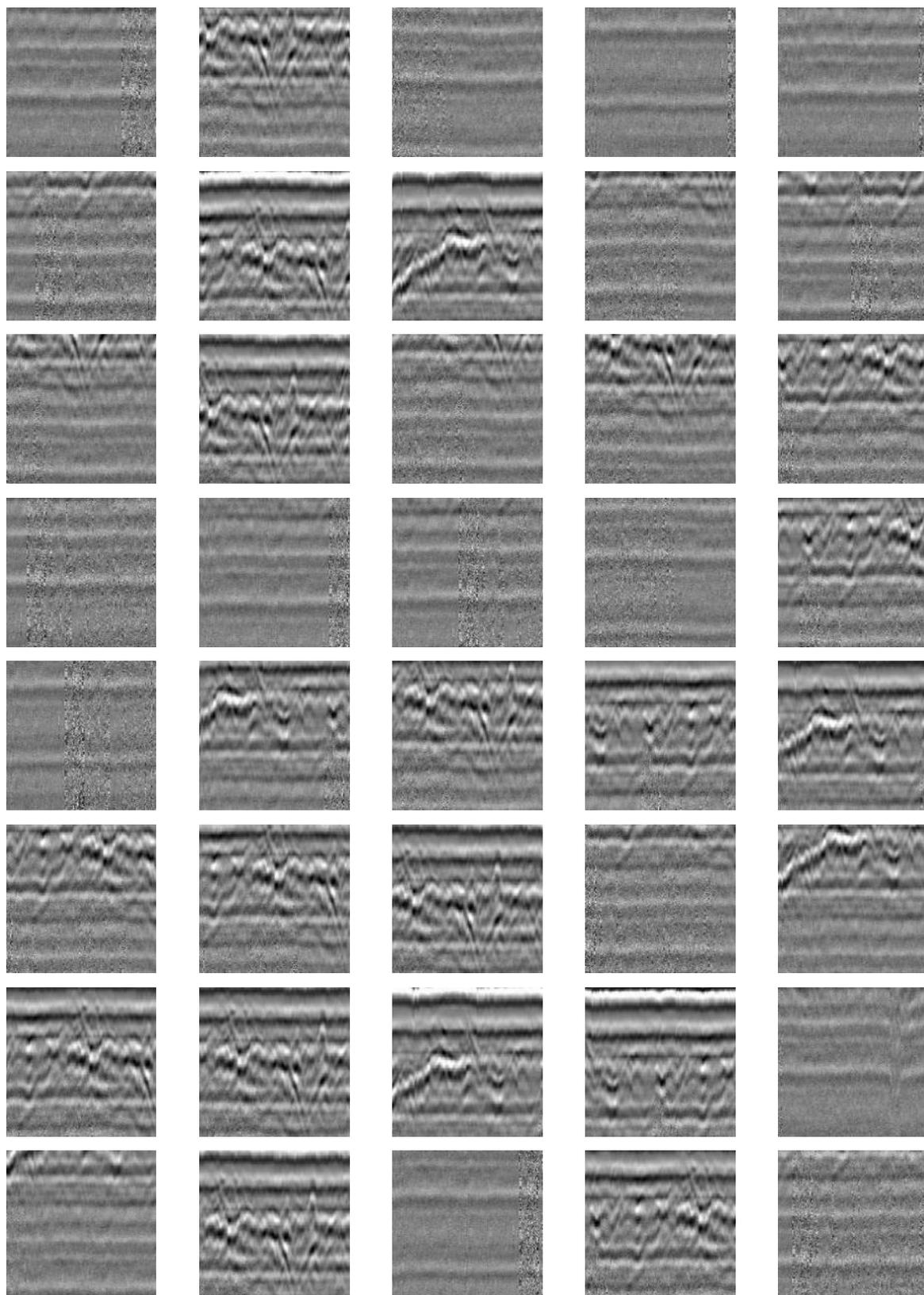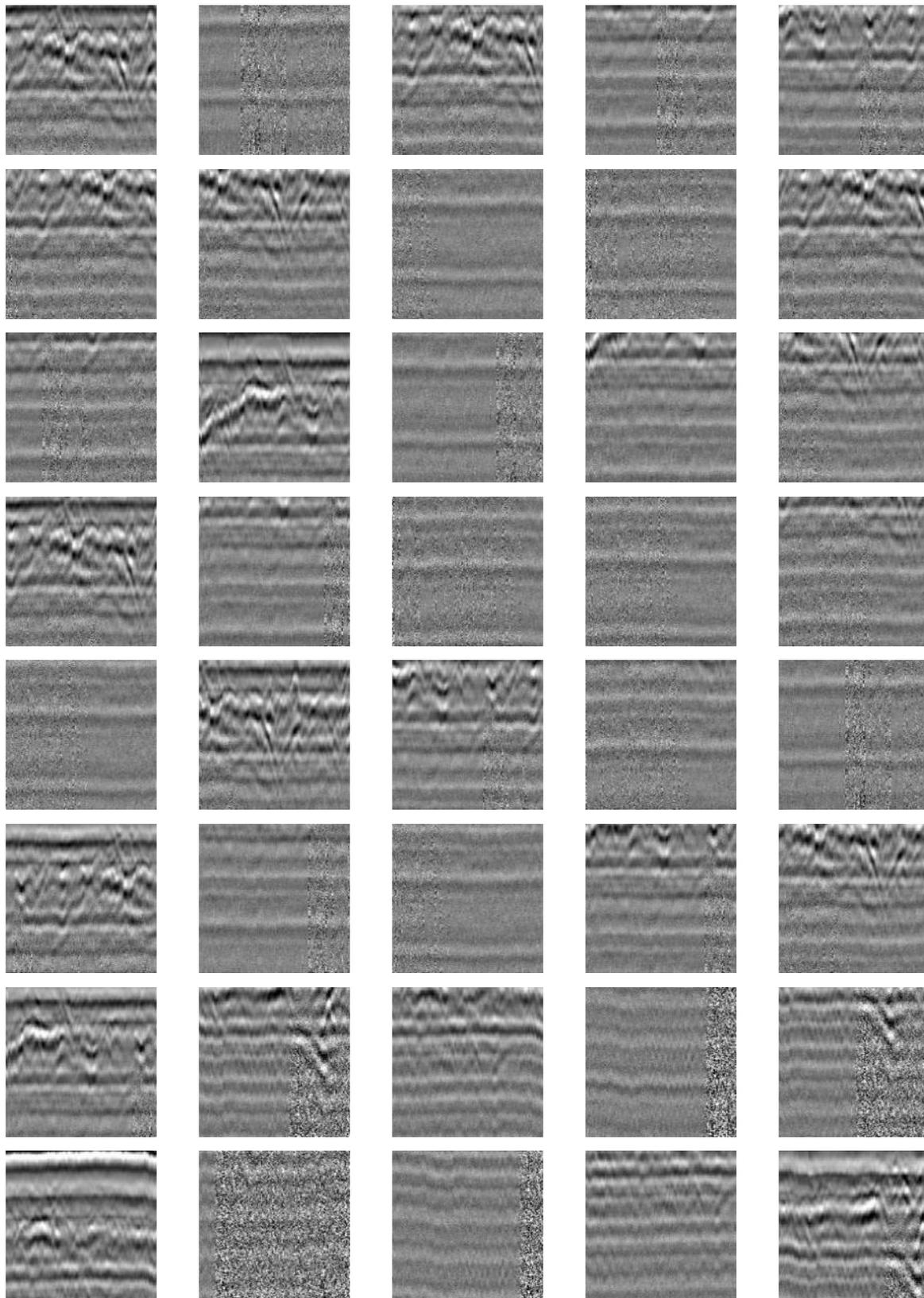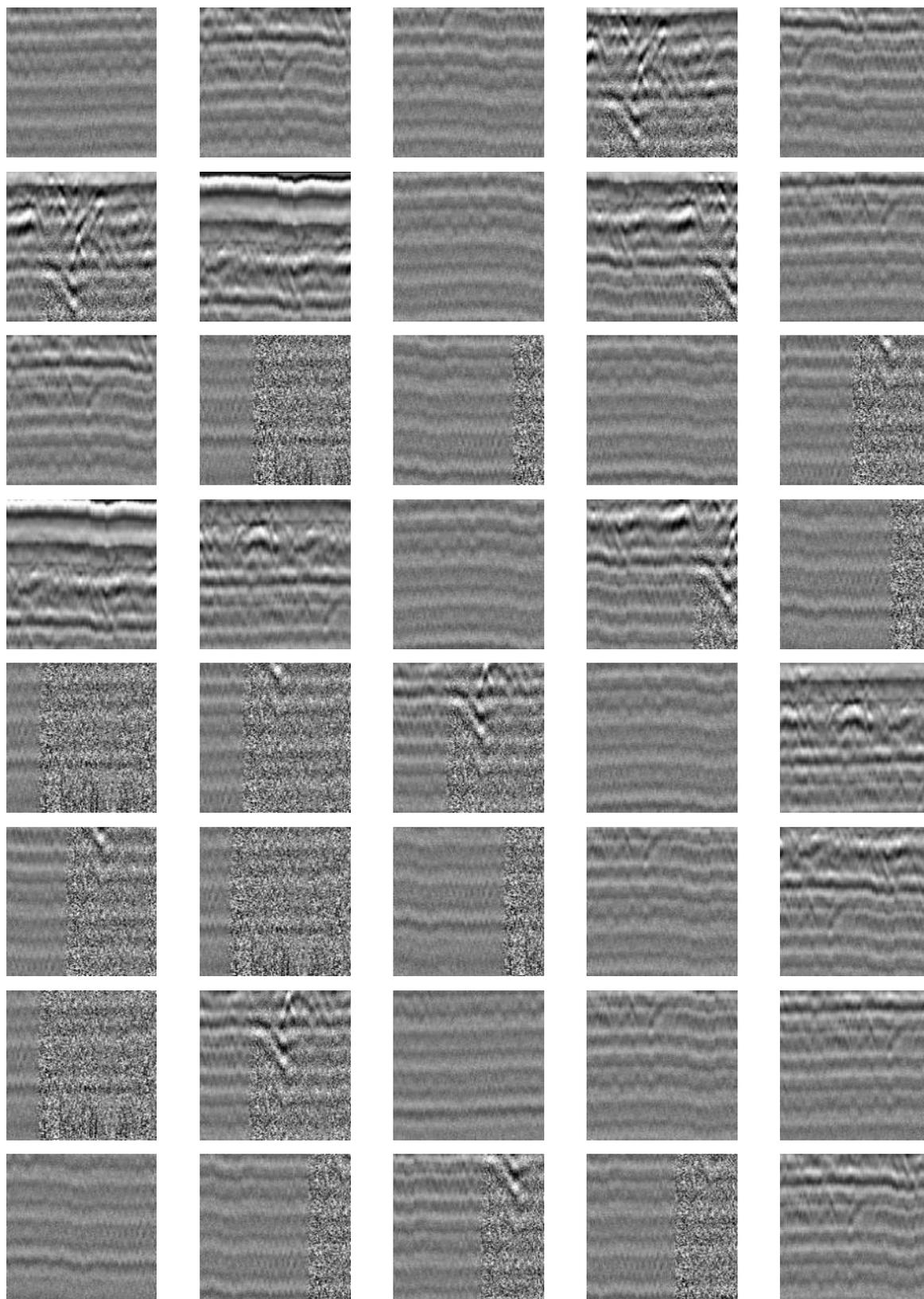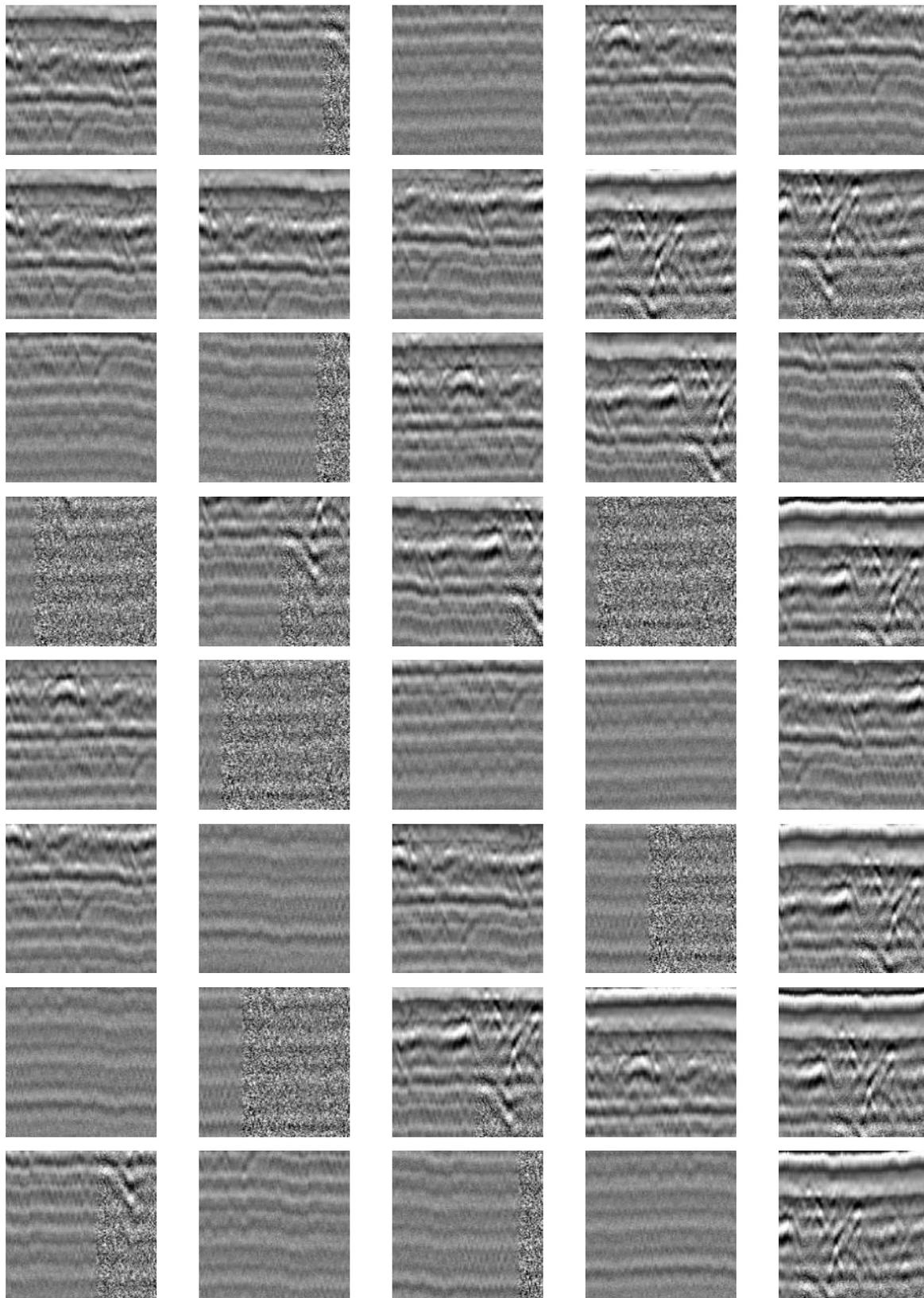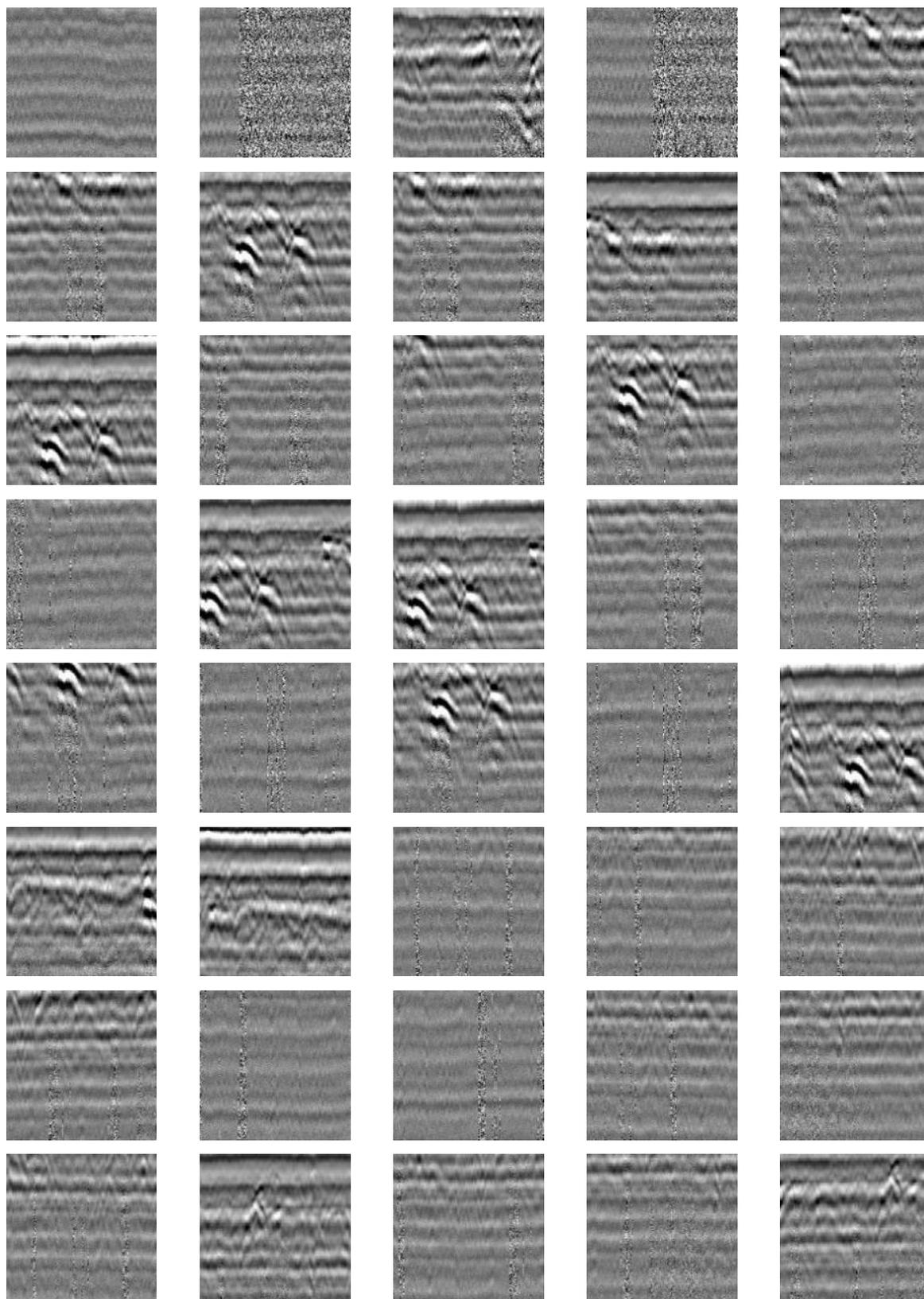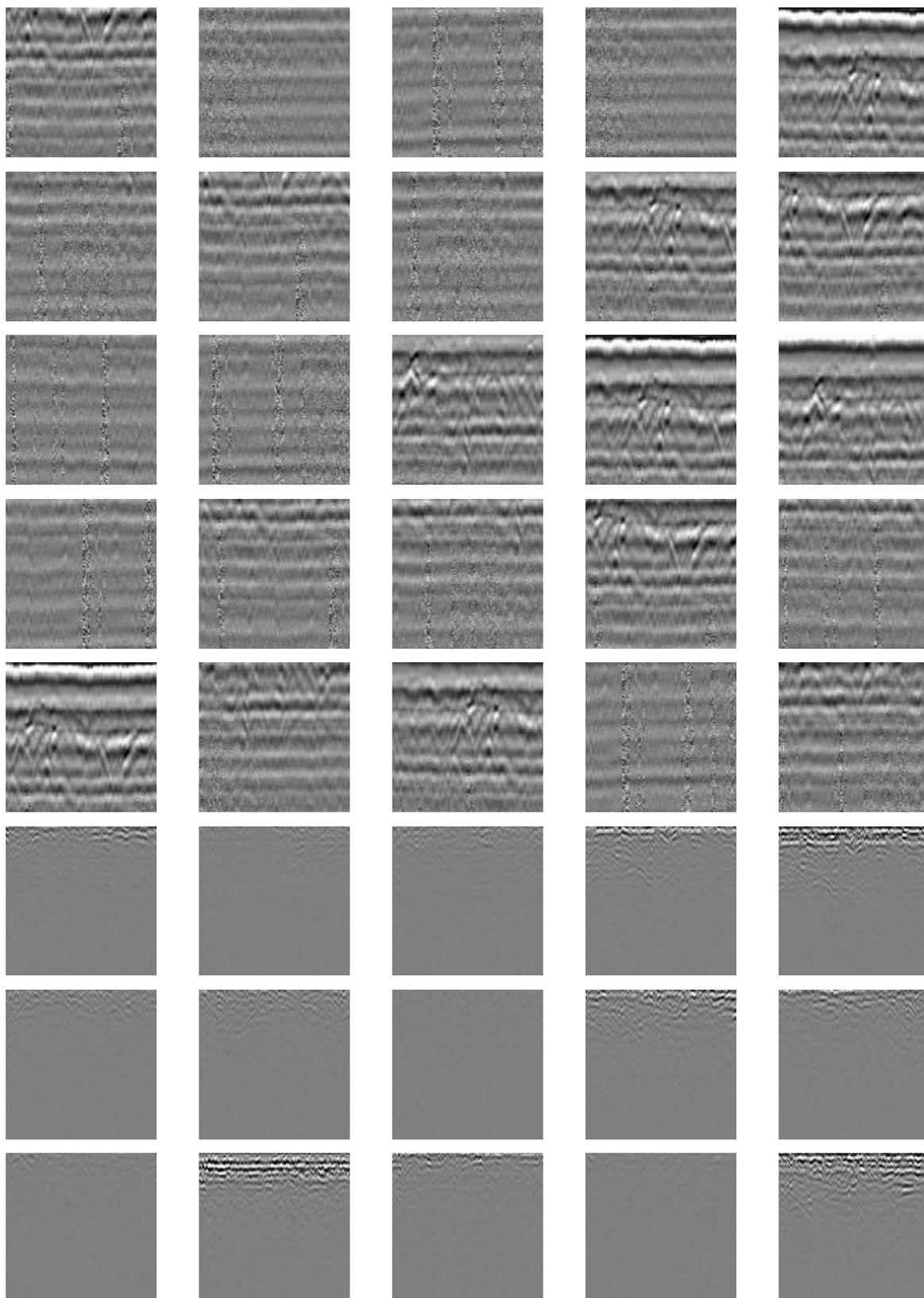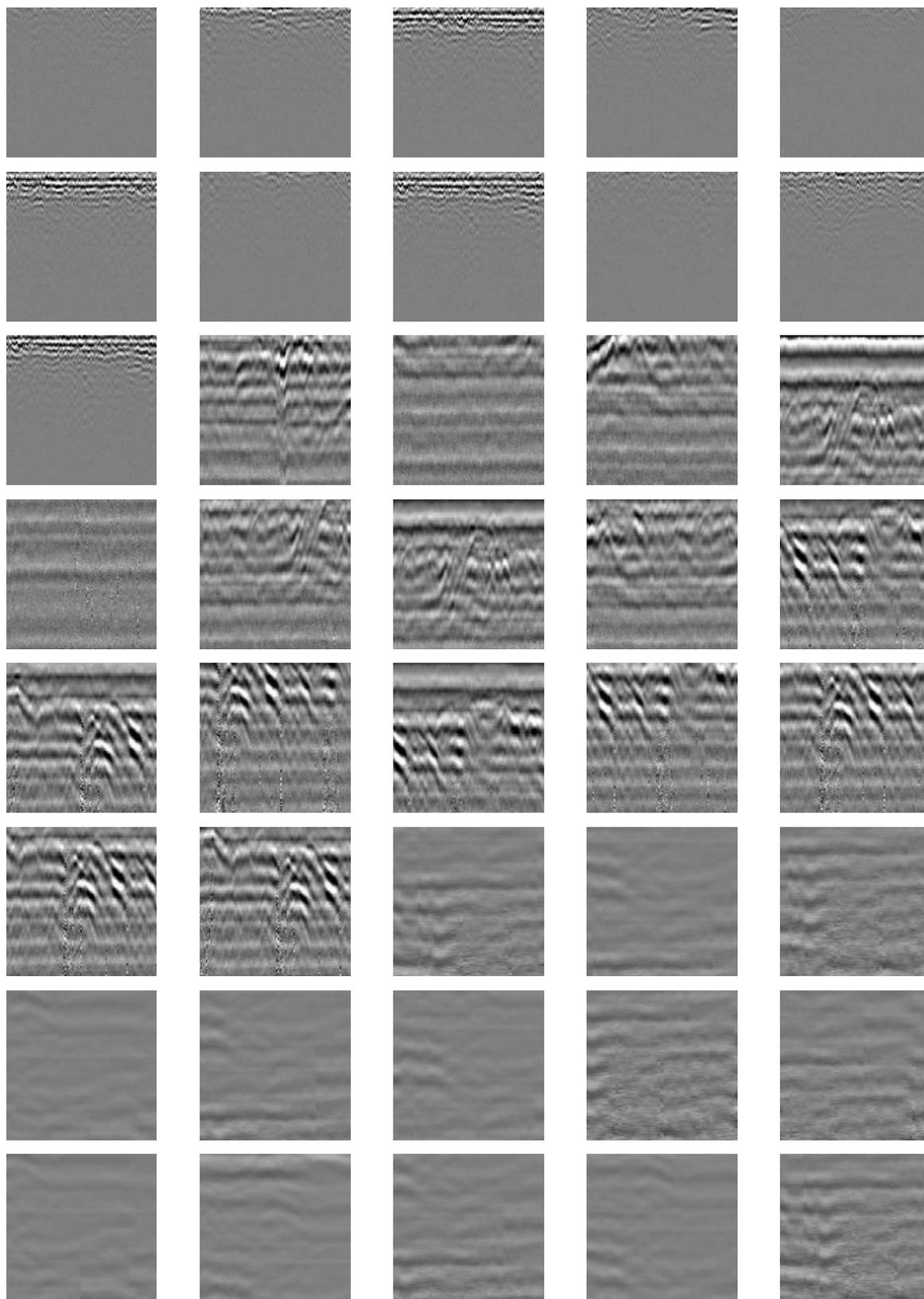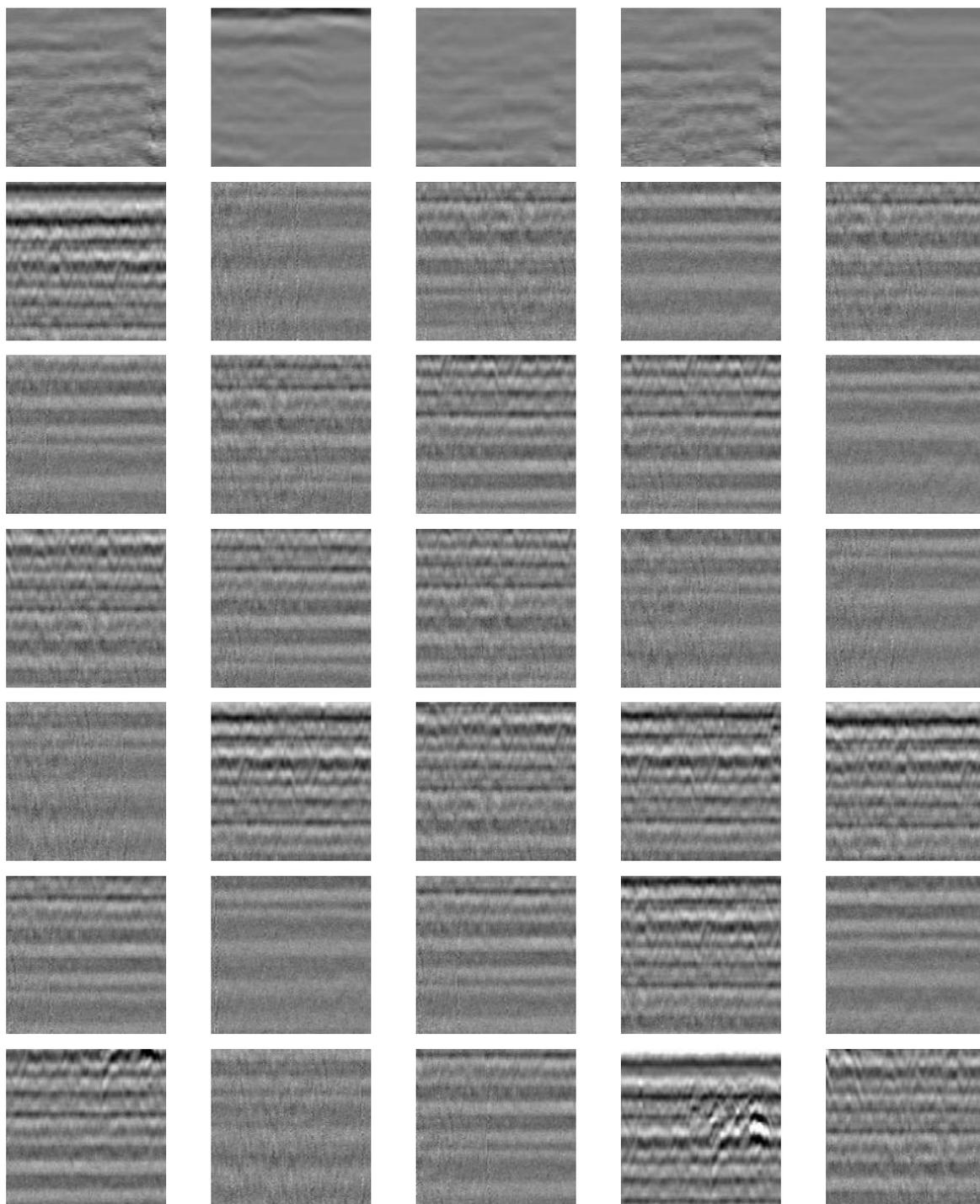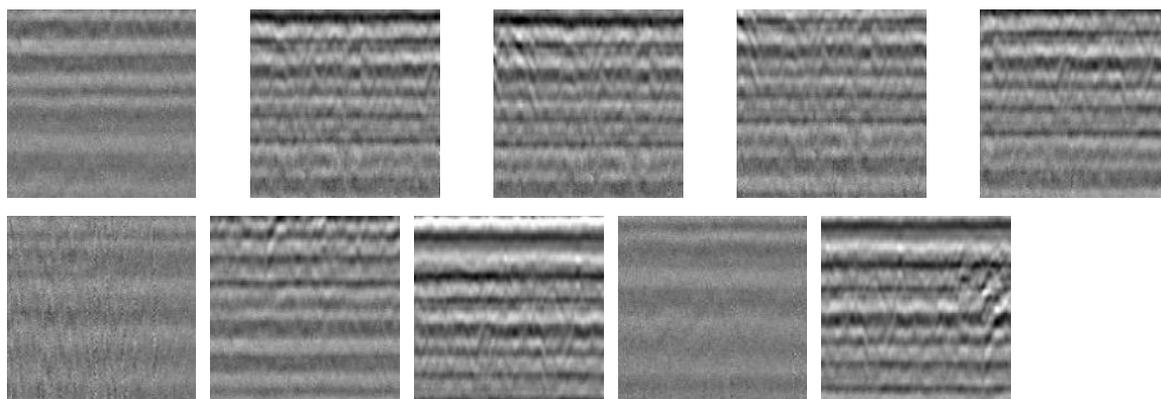
Radargram Training Data: Grave Images
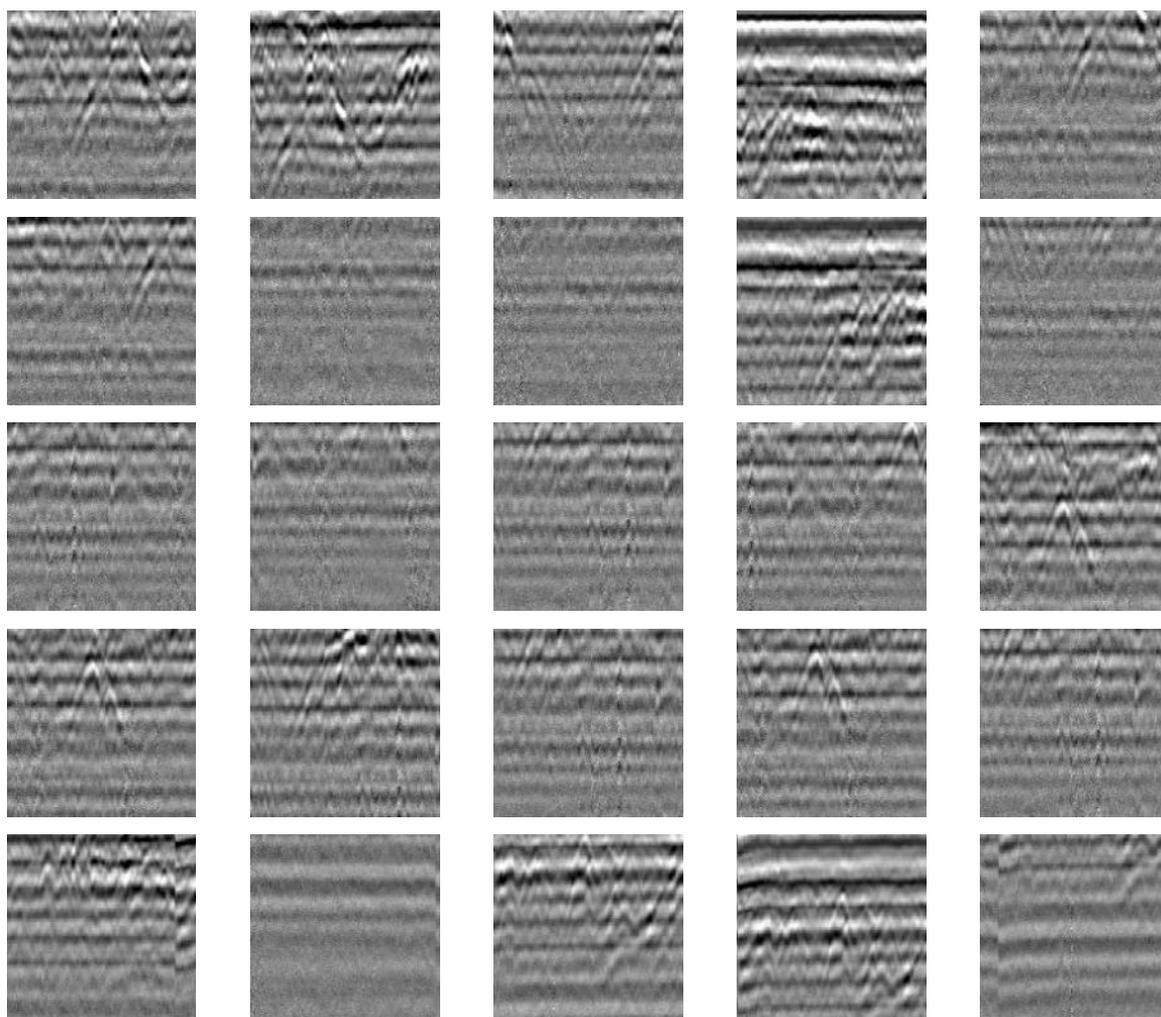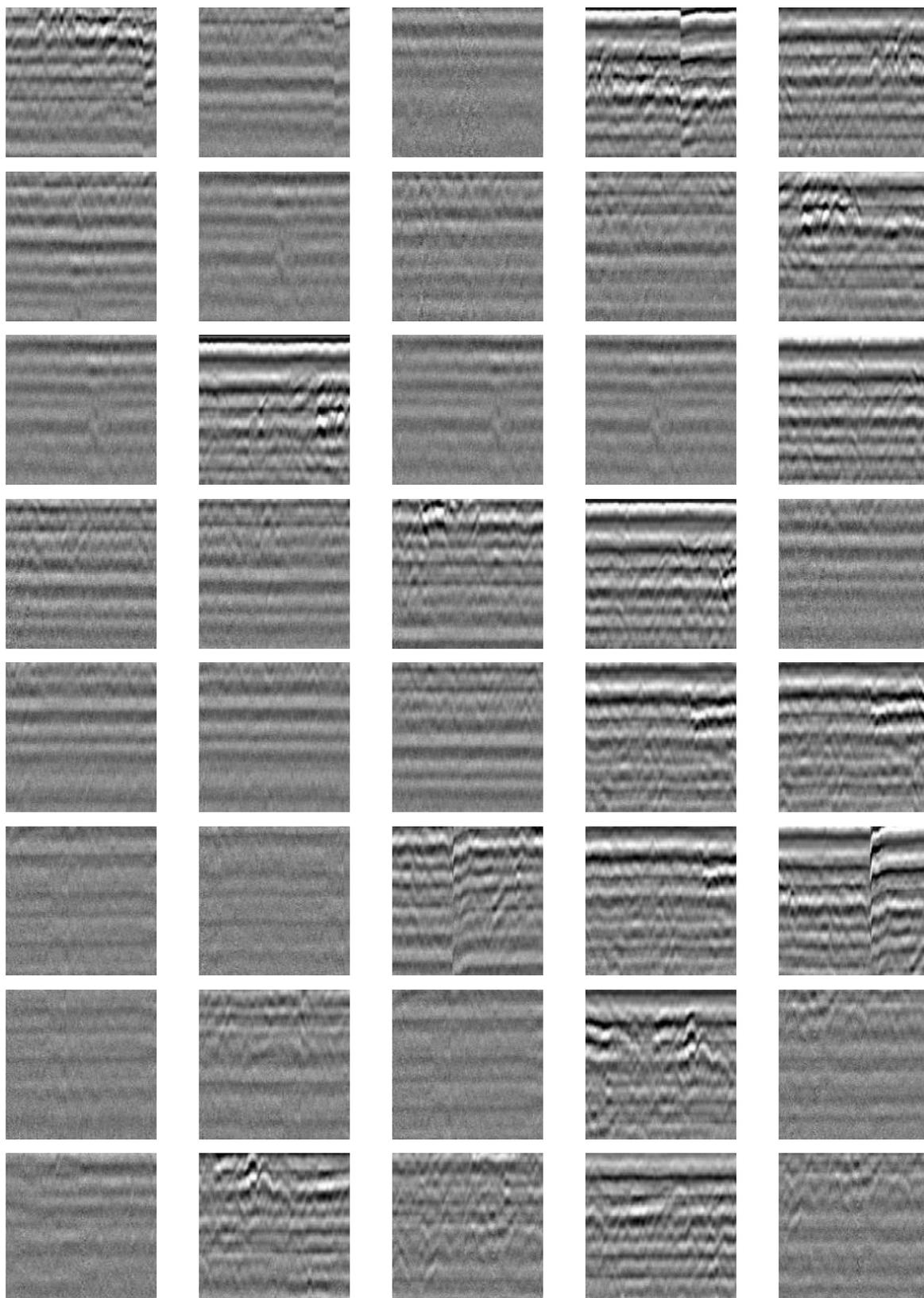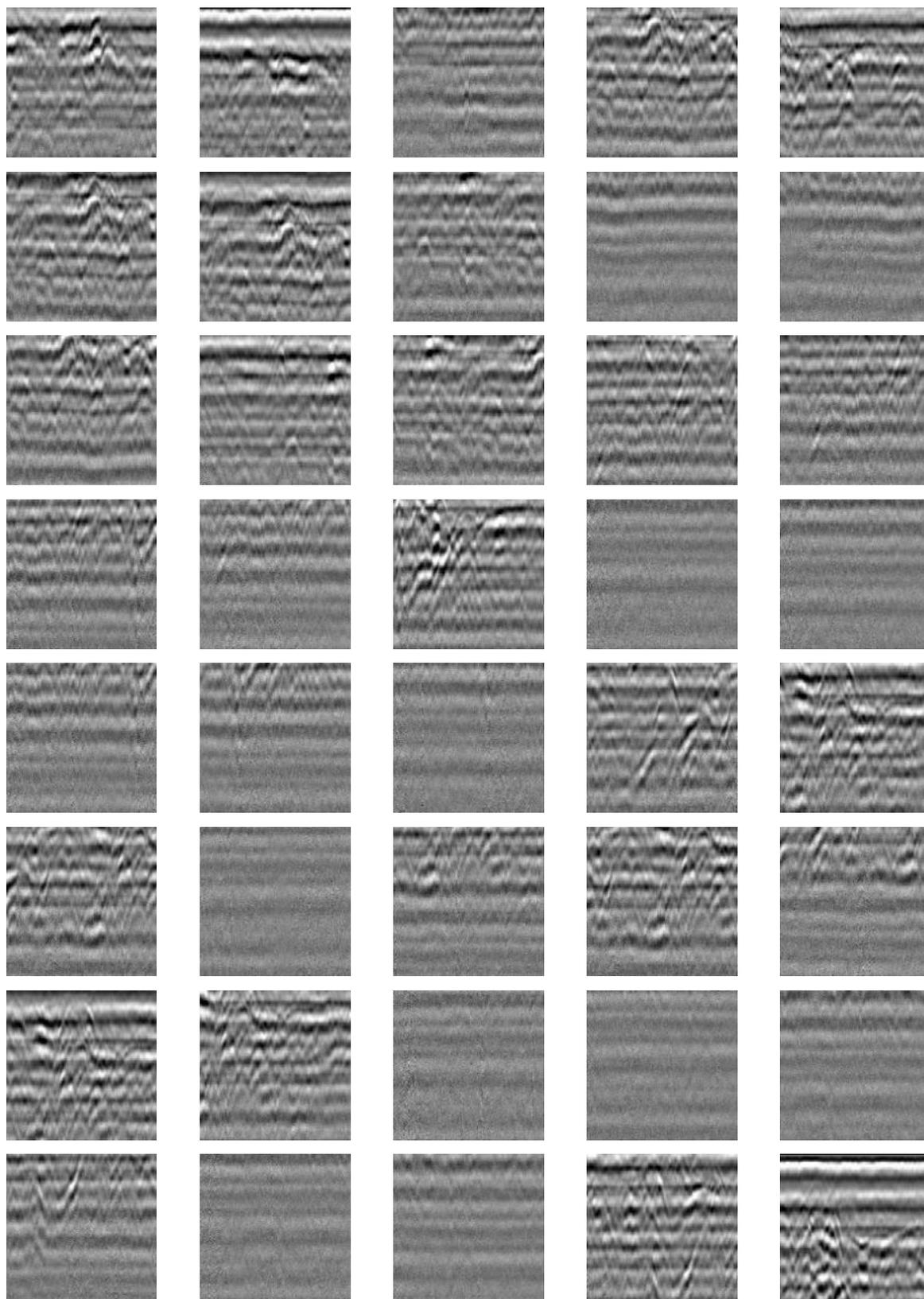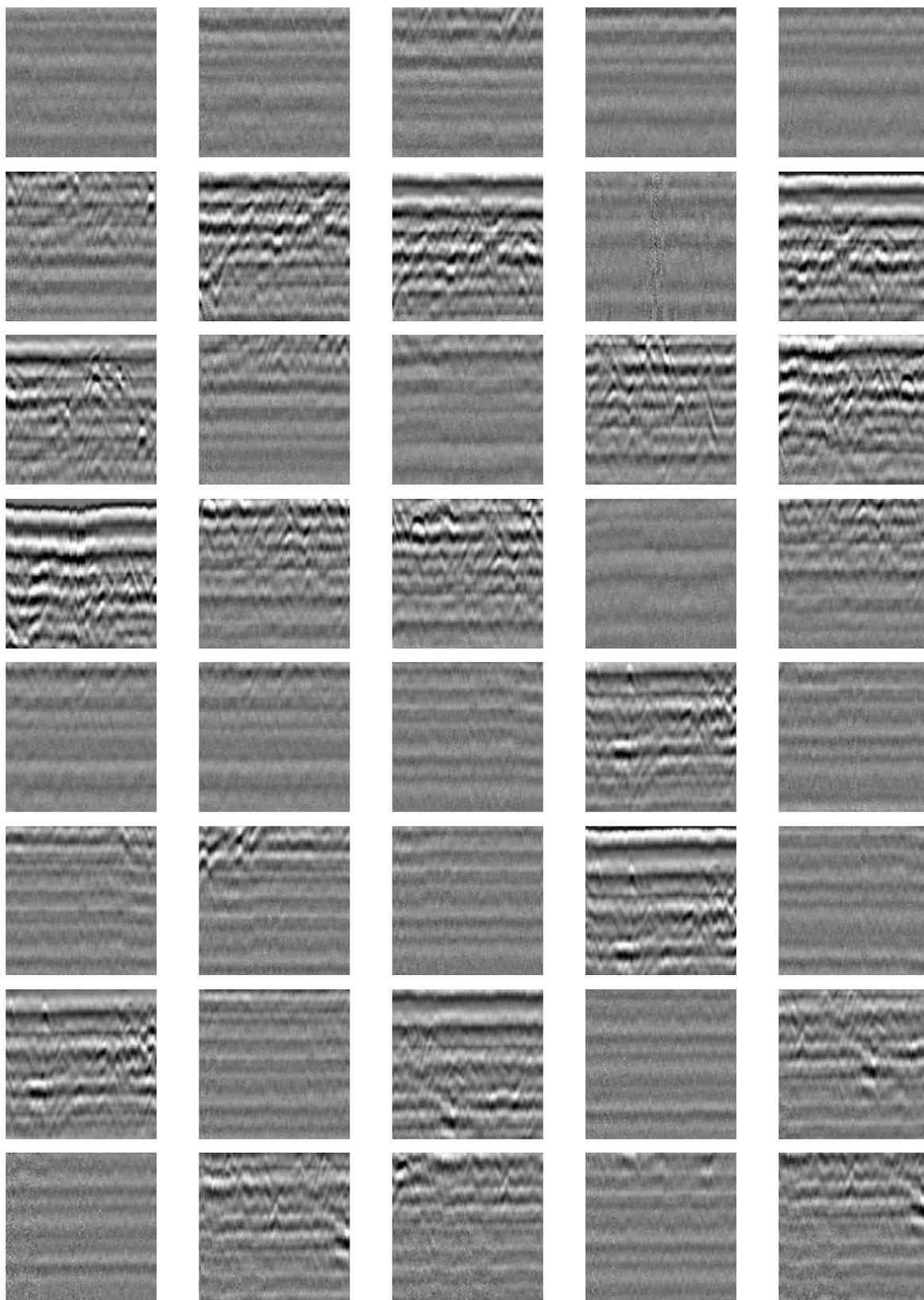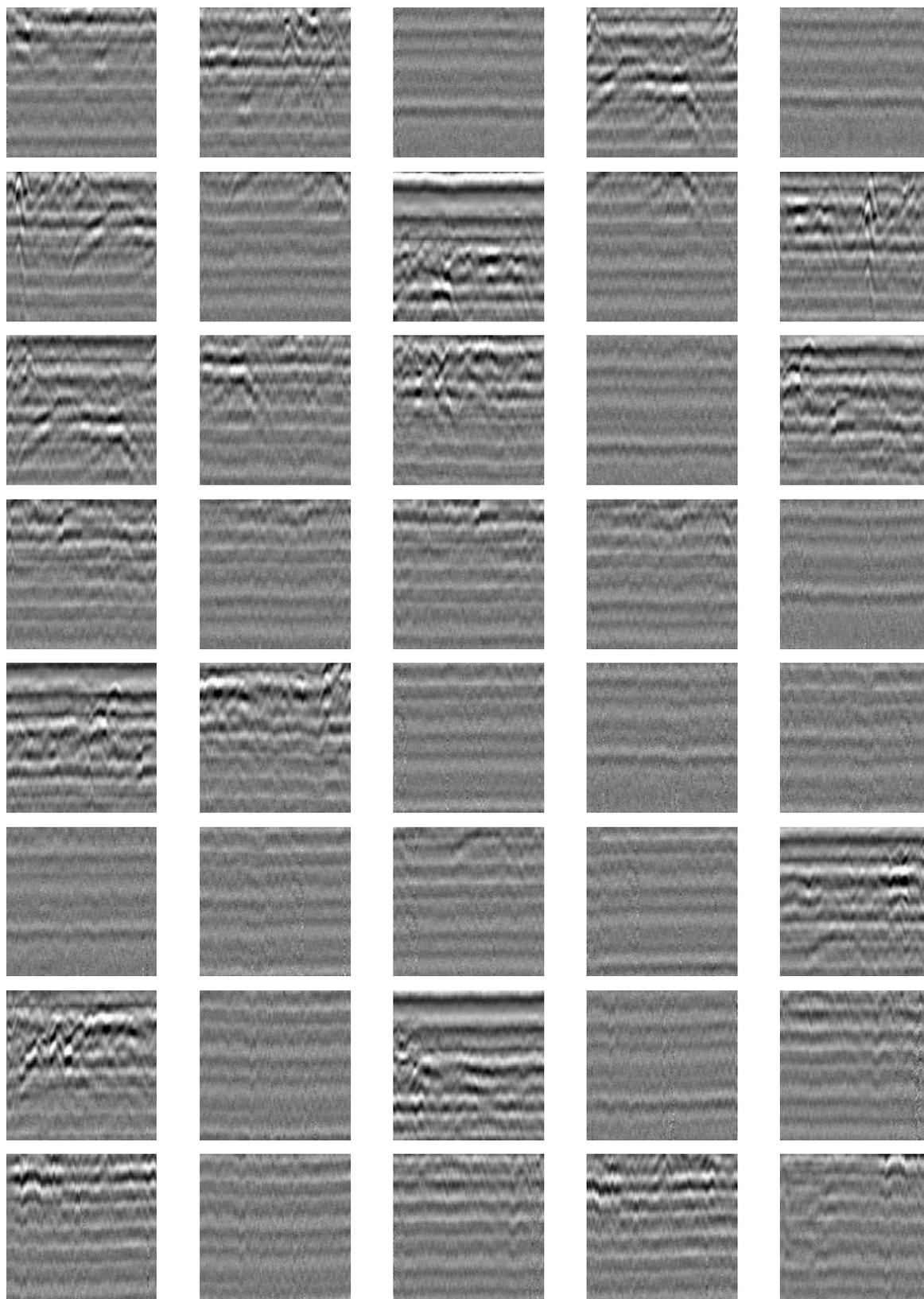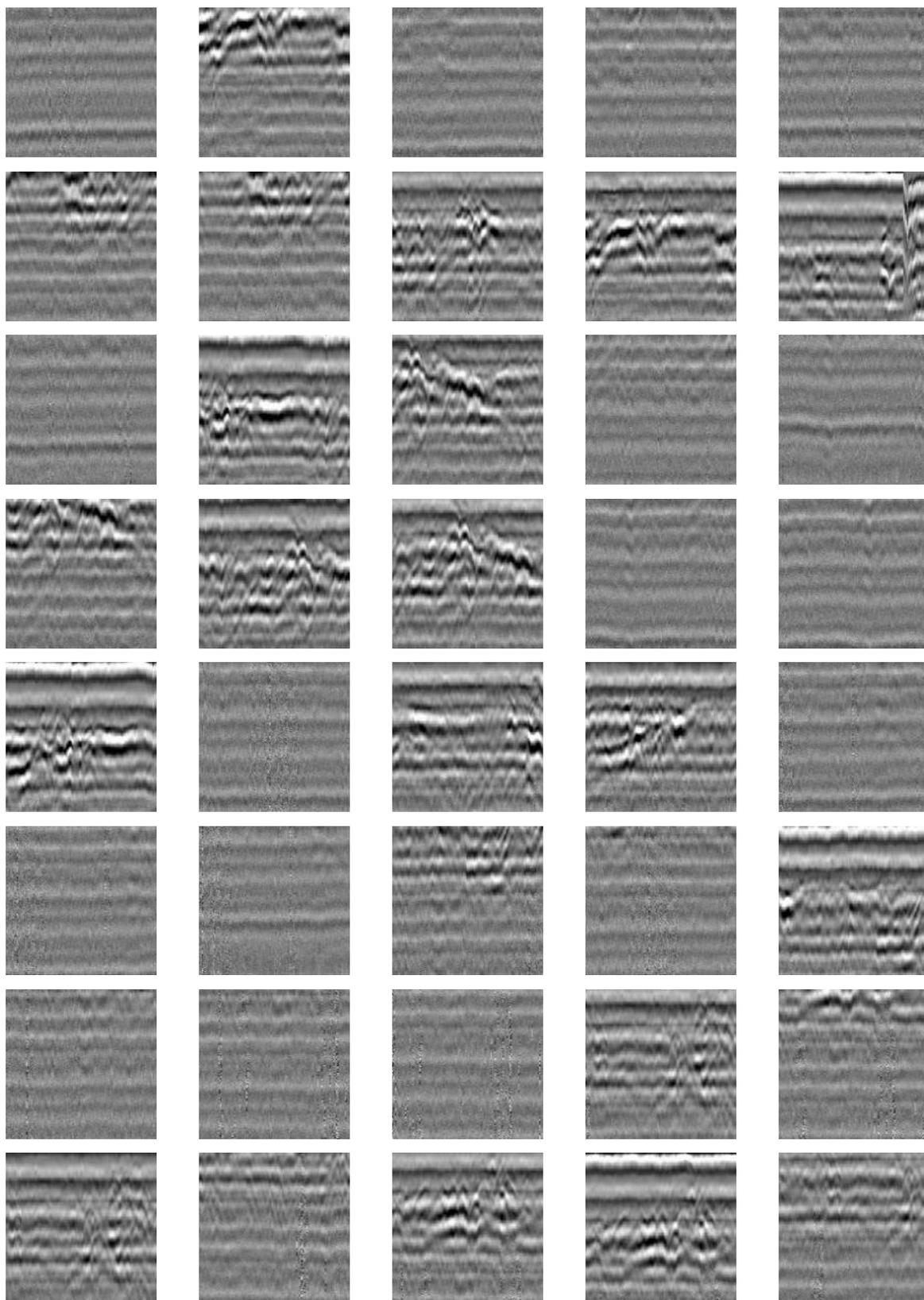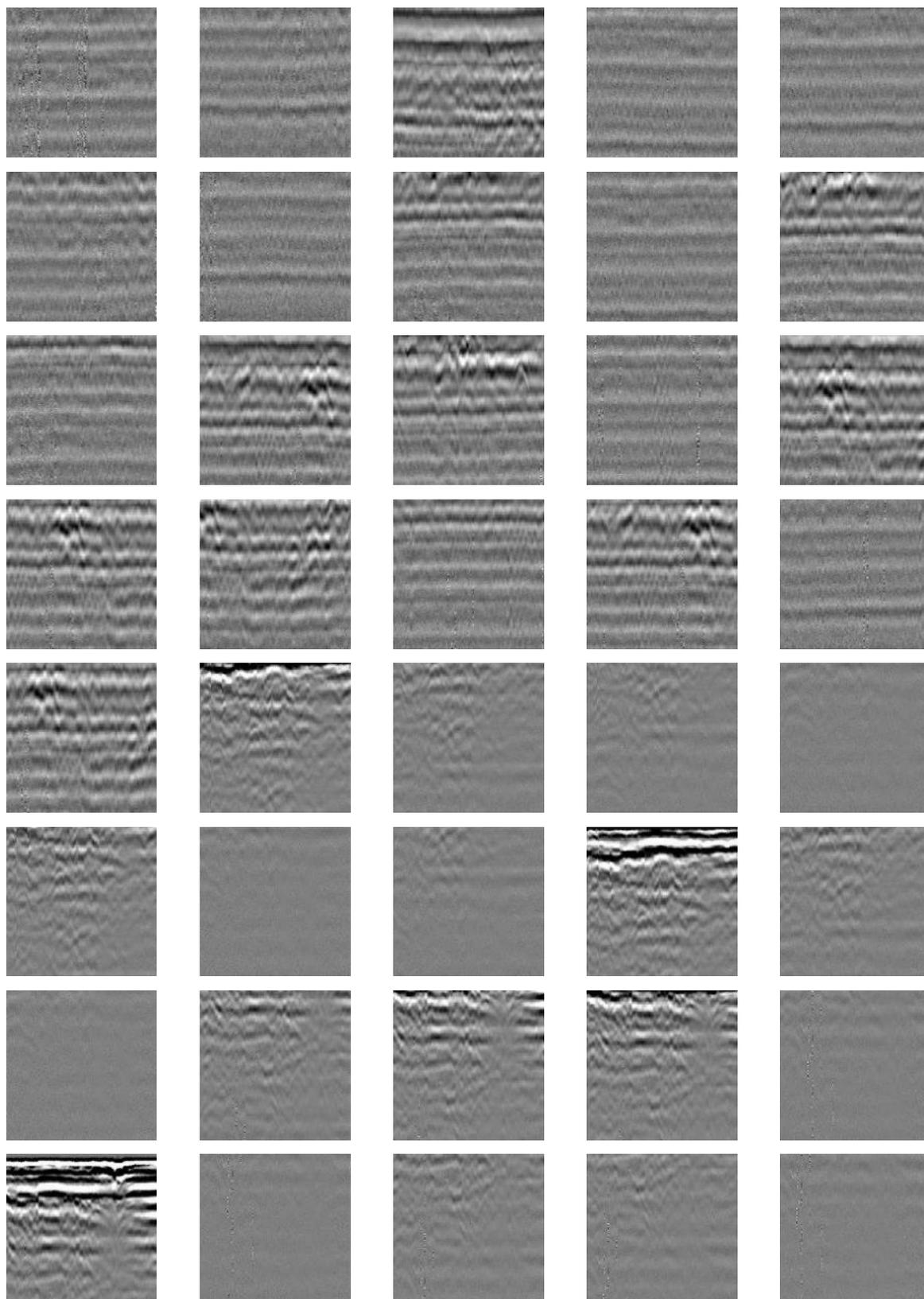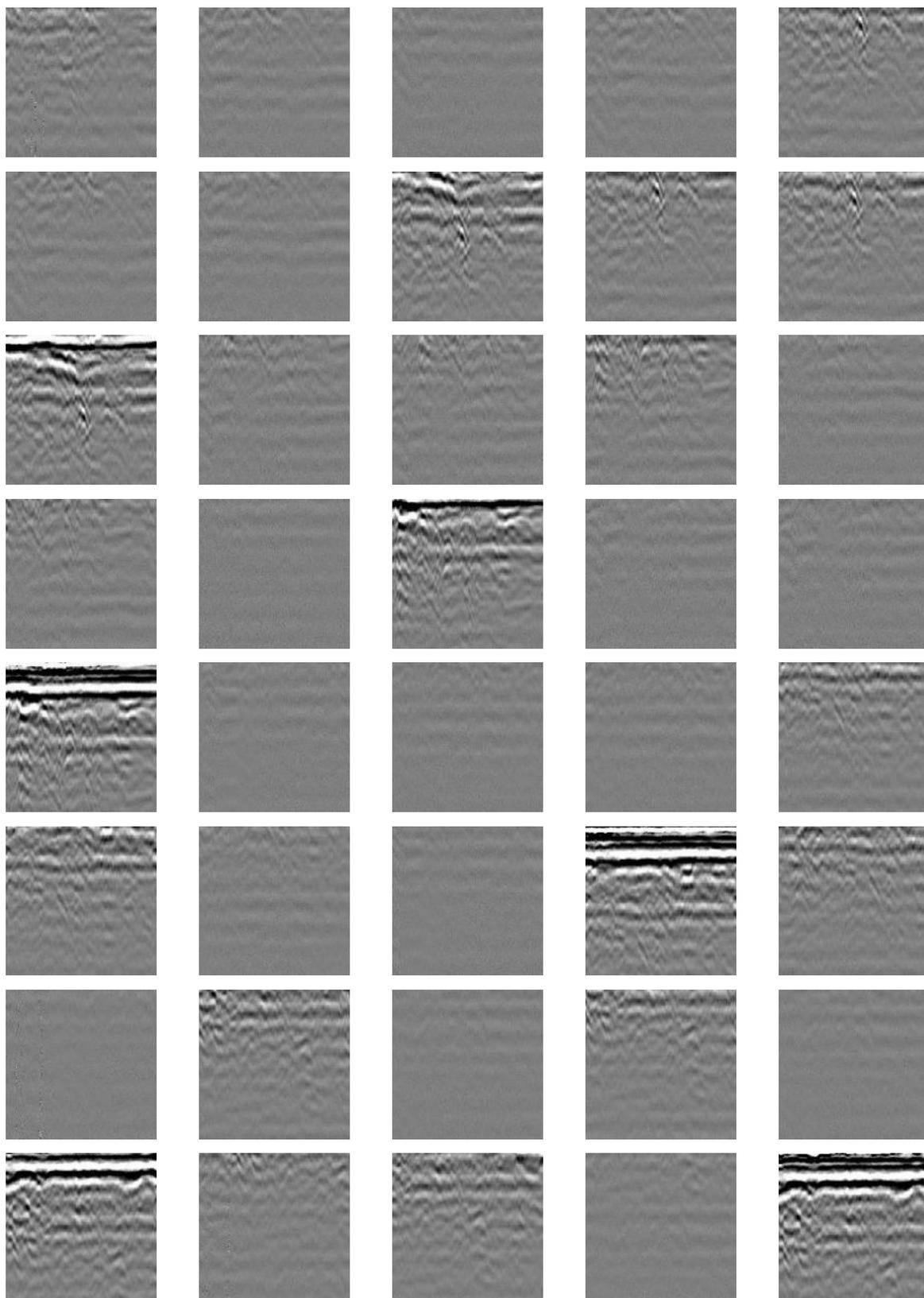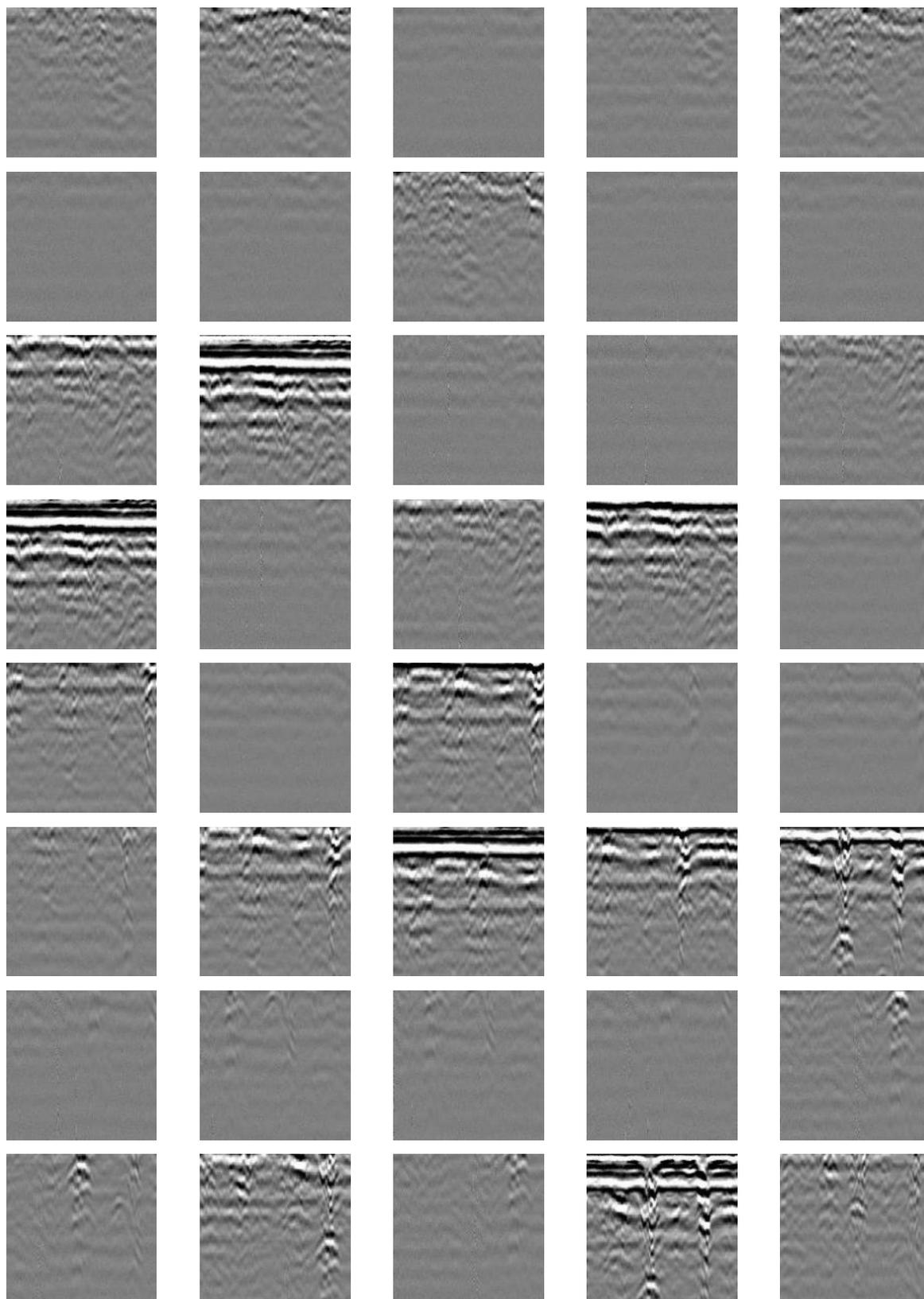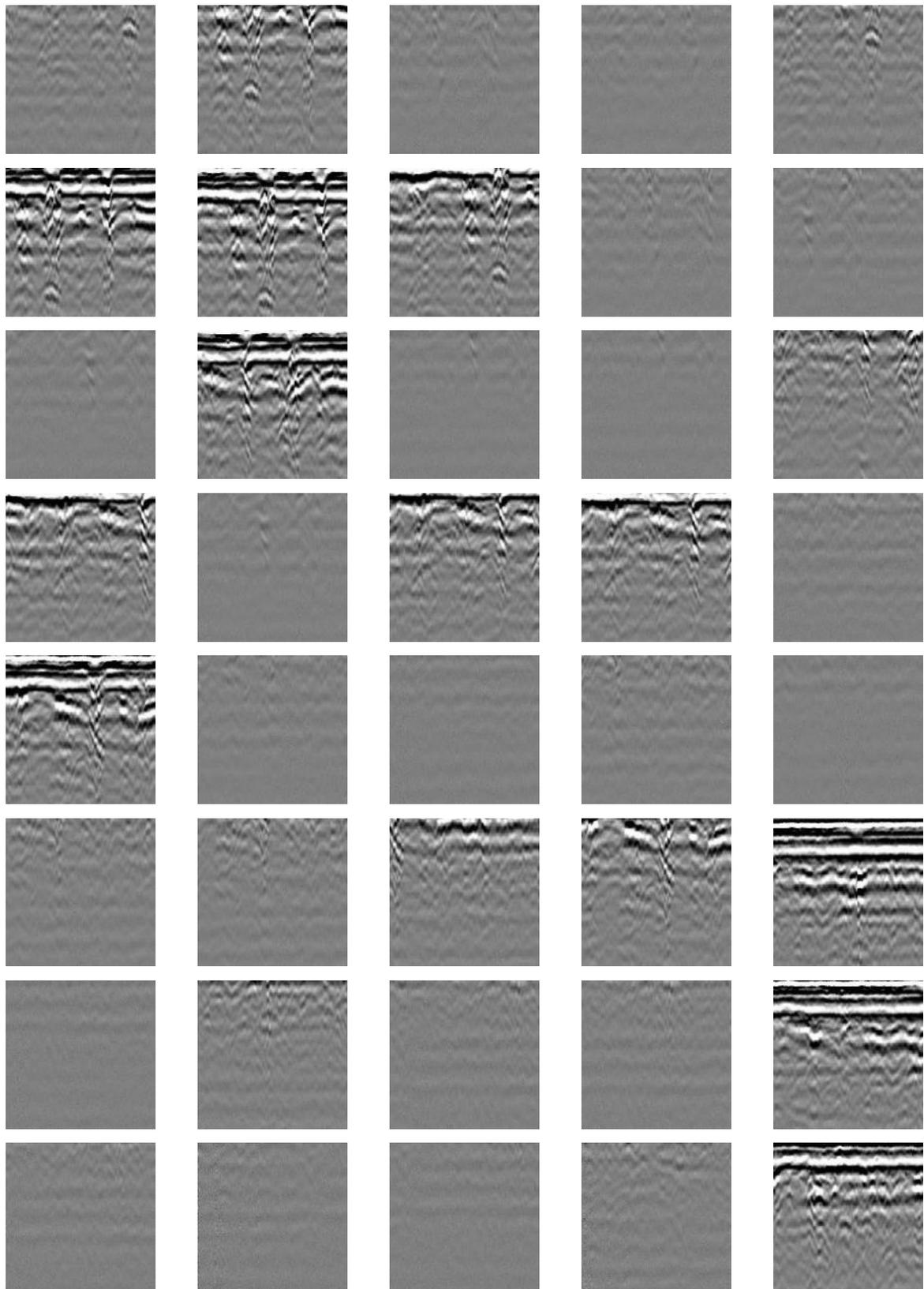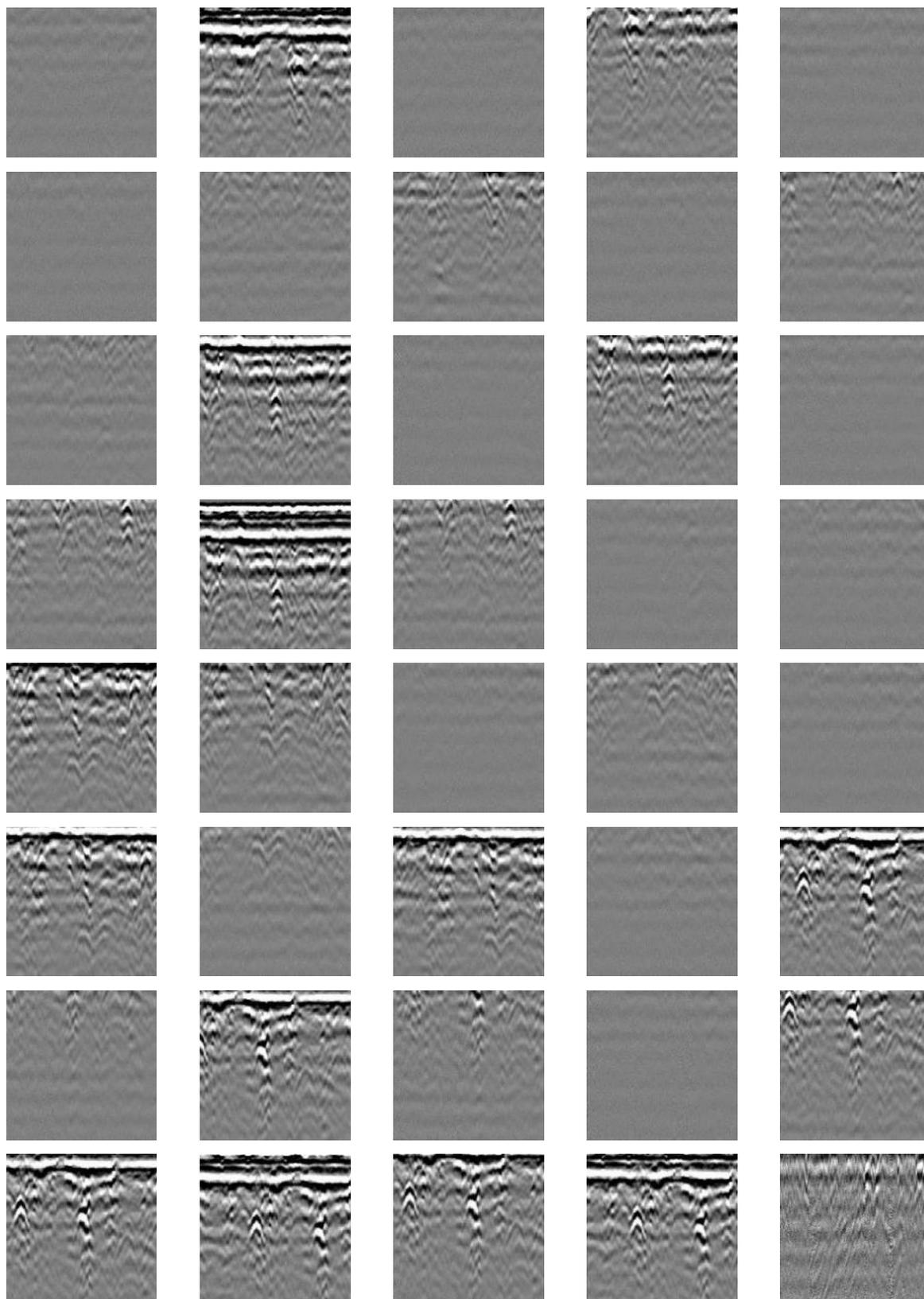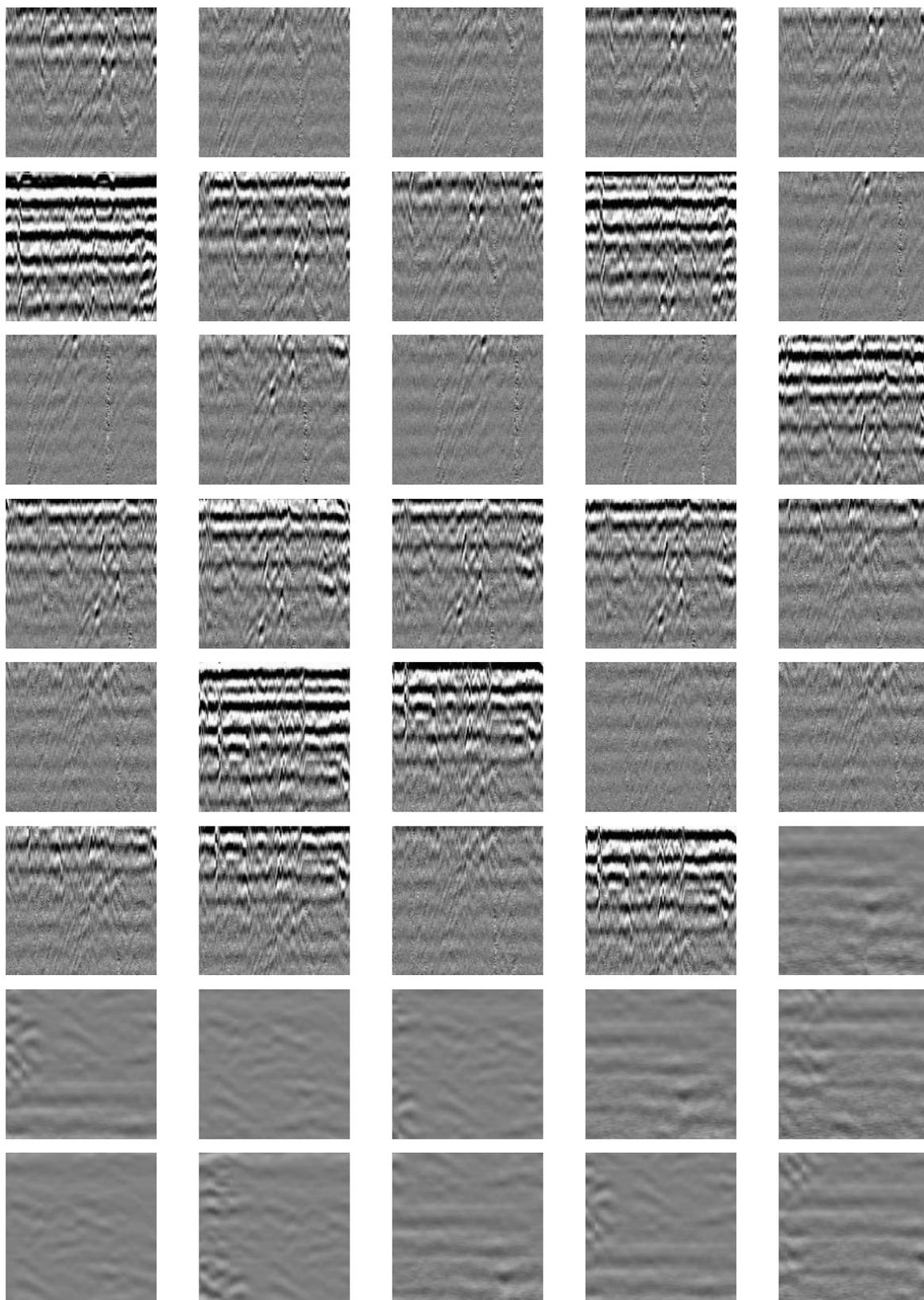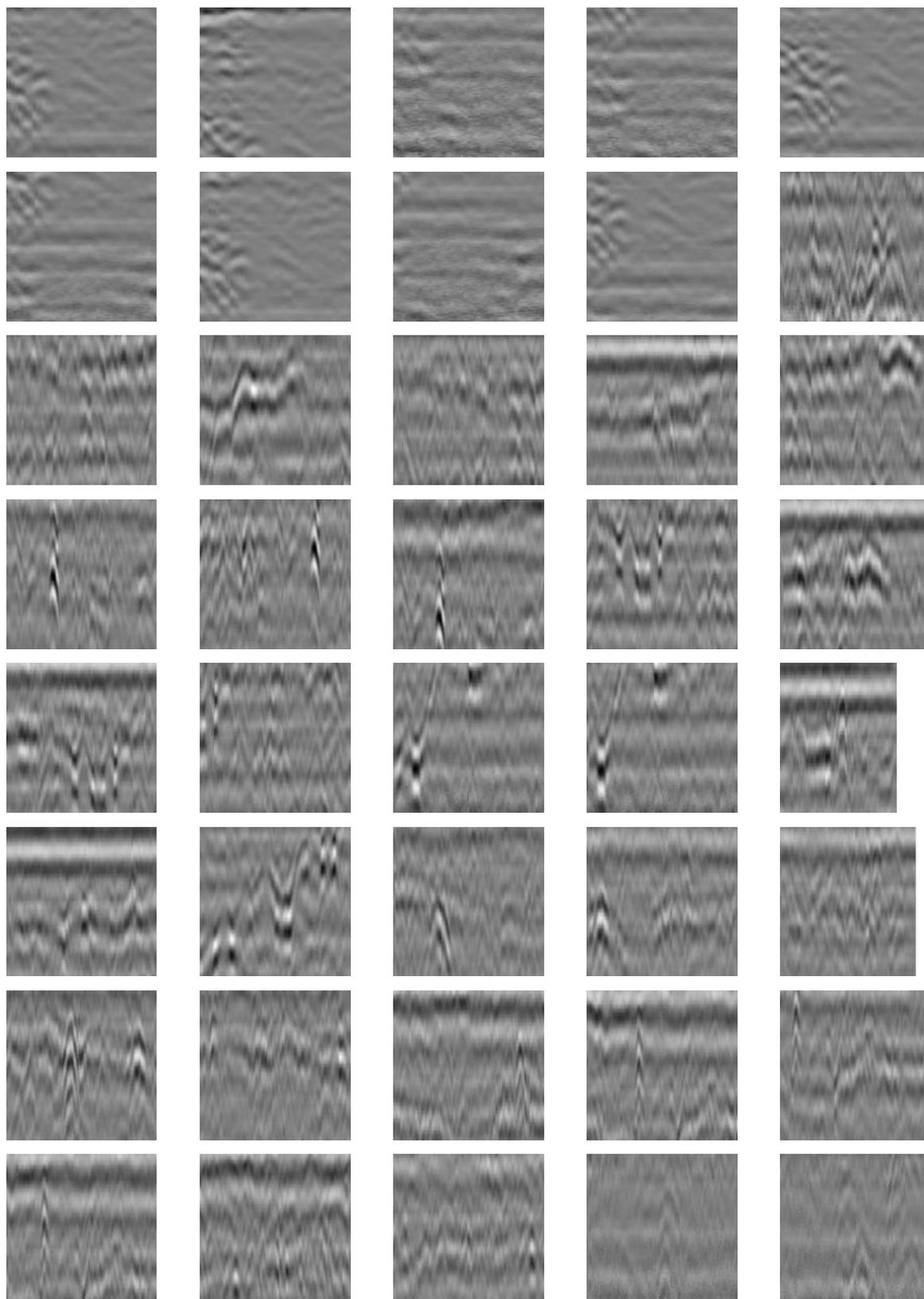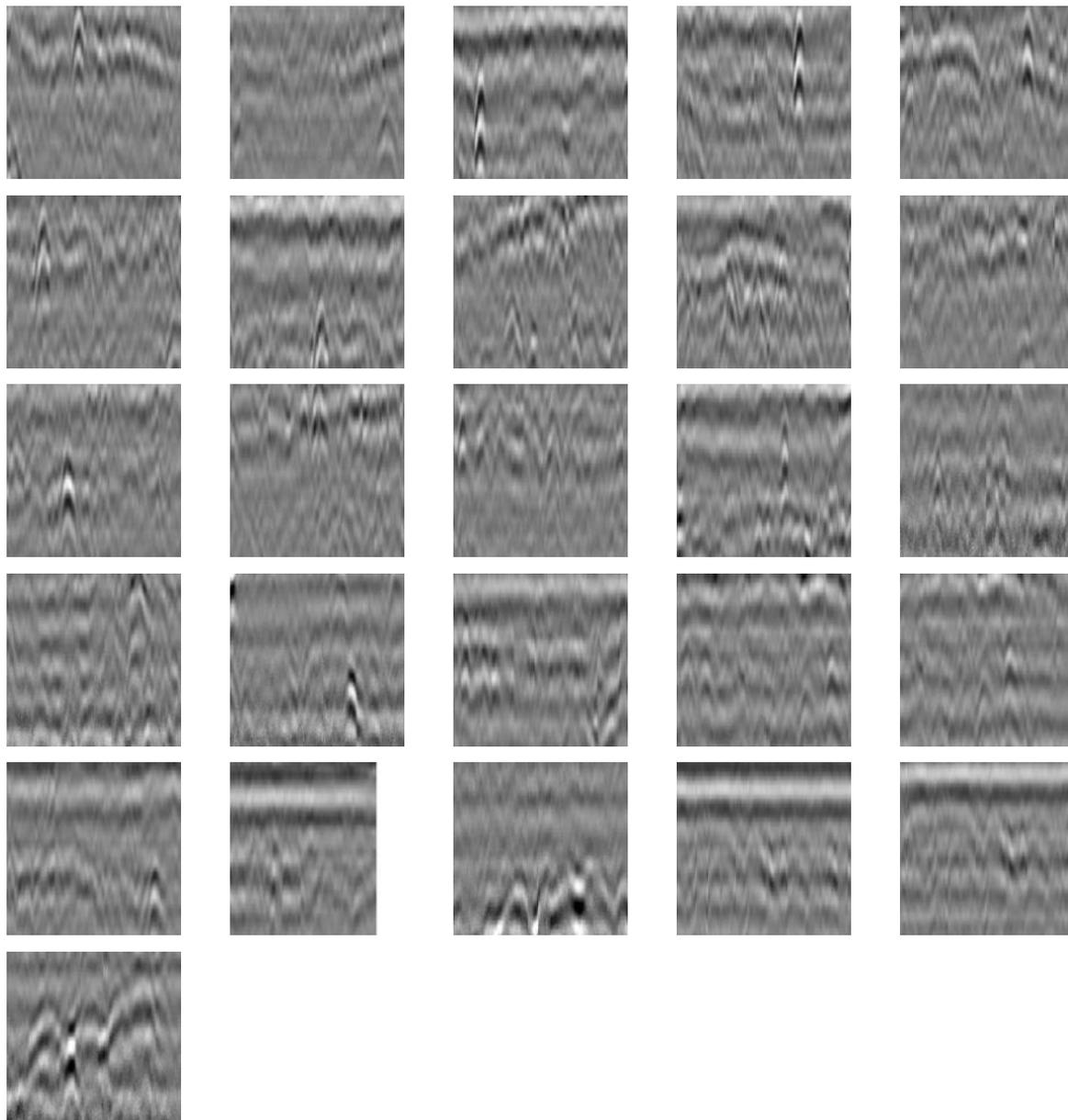
Timeslice Training Data: Non-grave Images
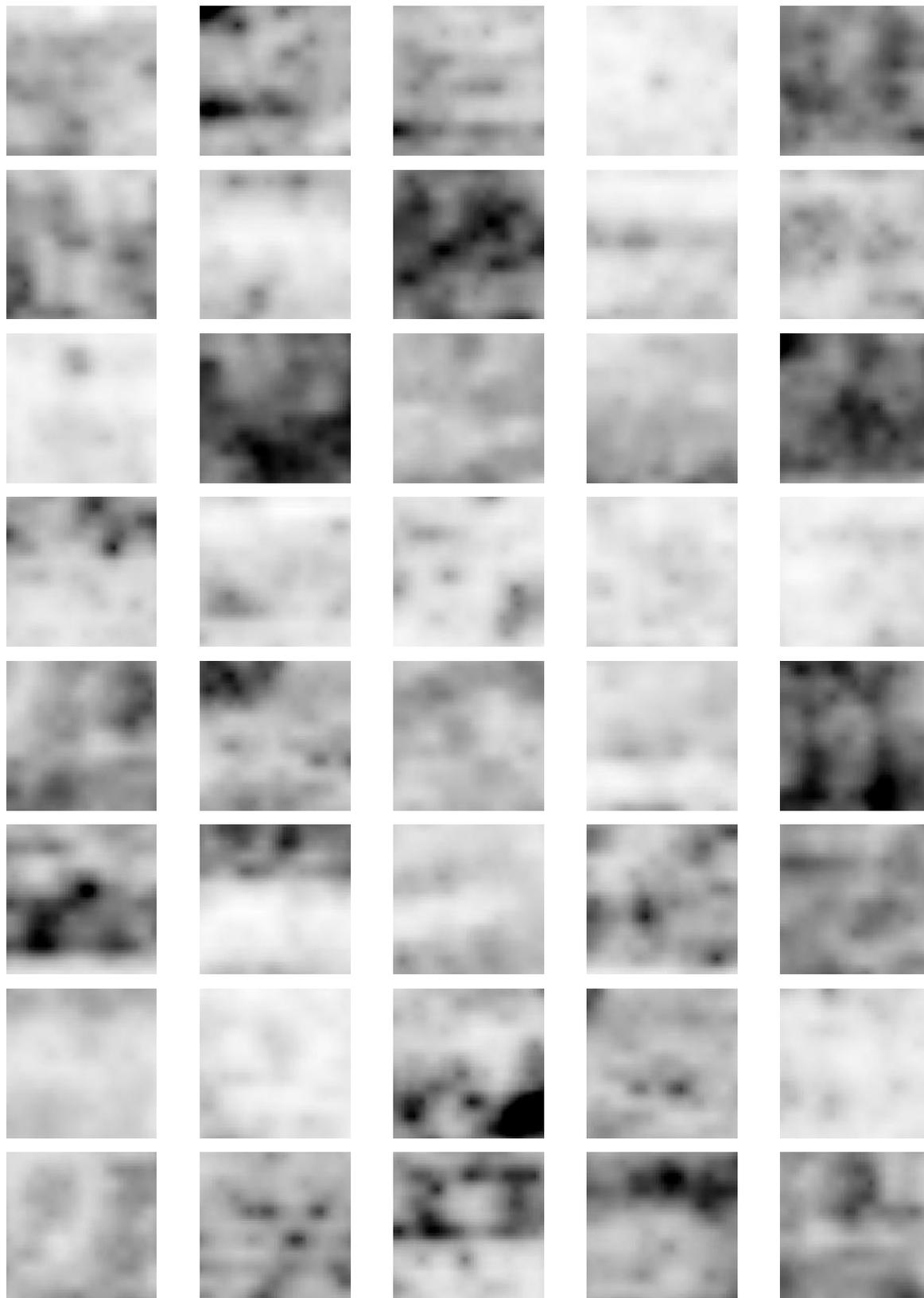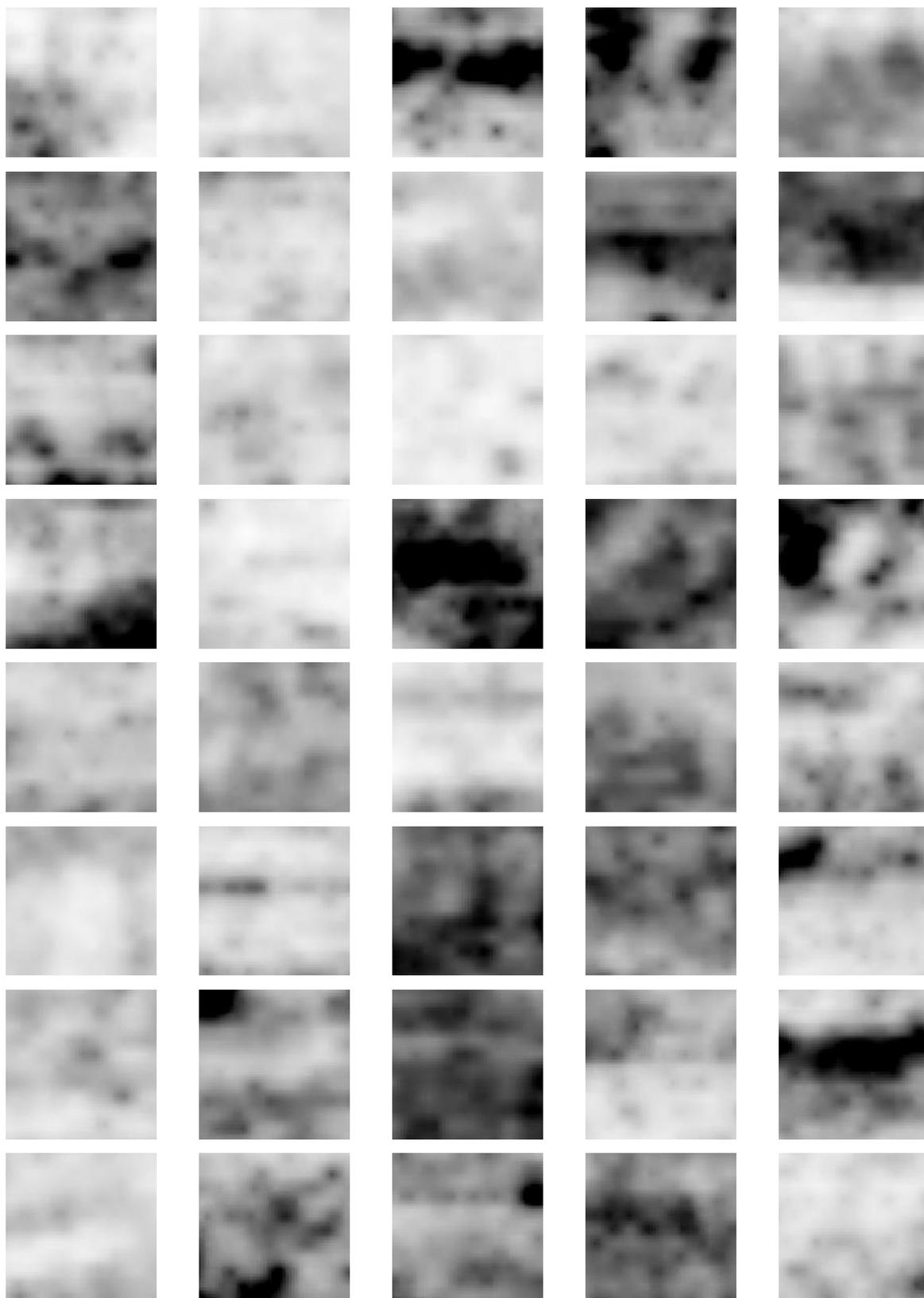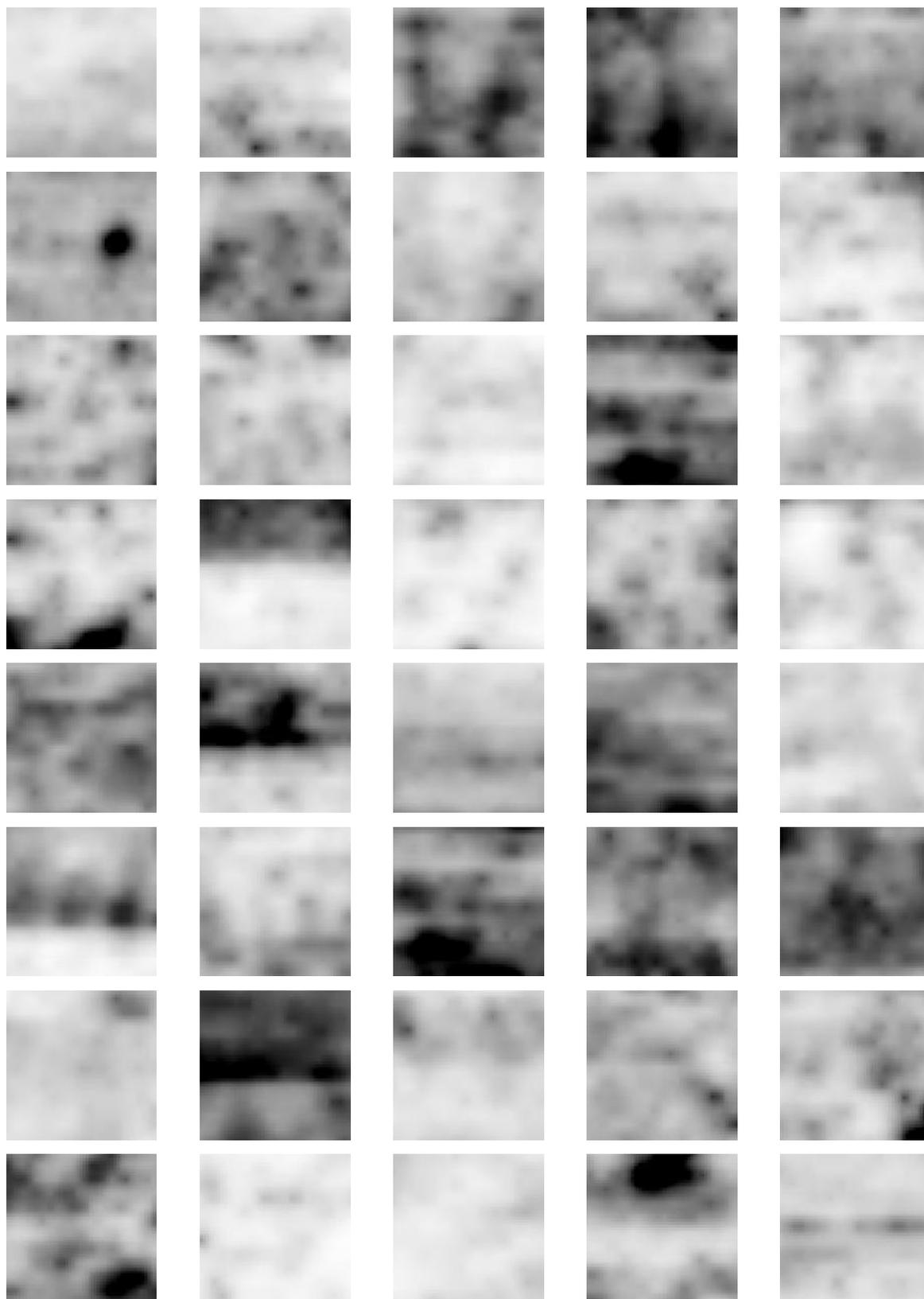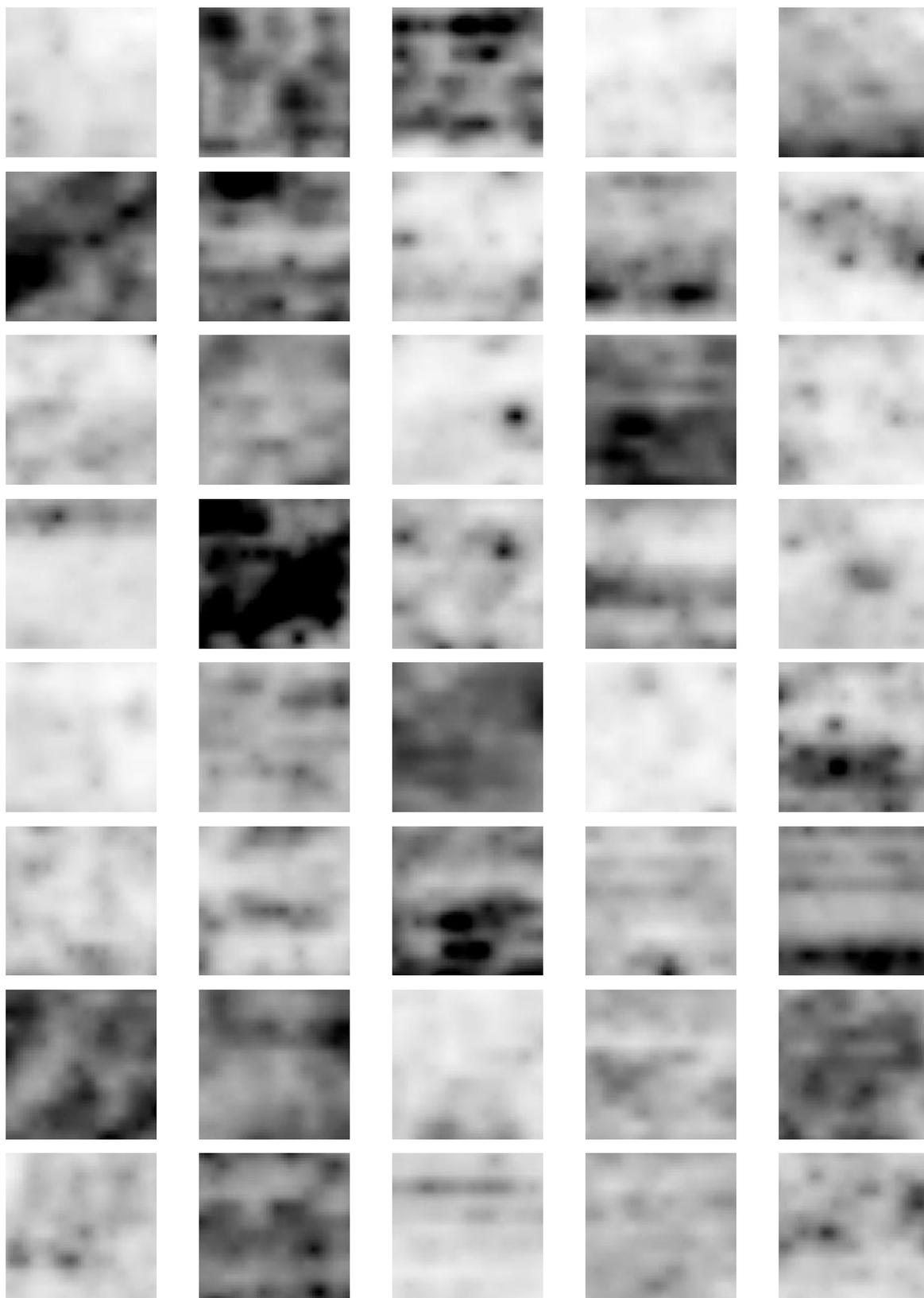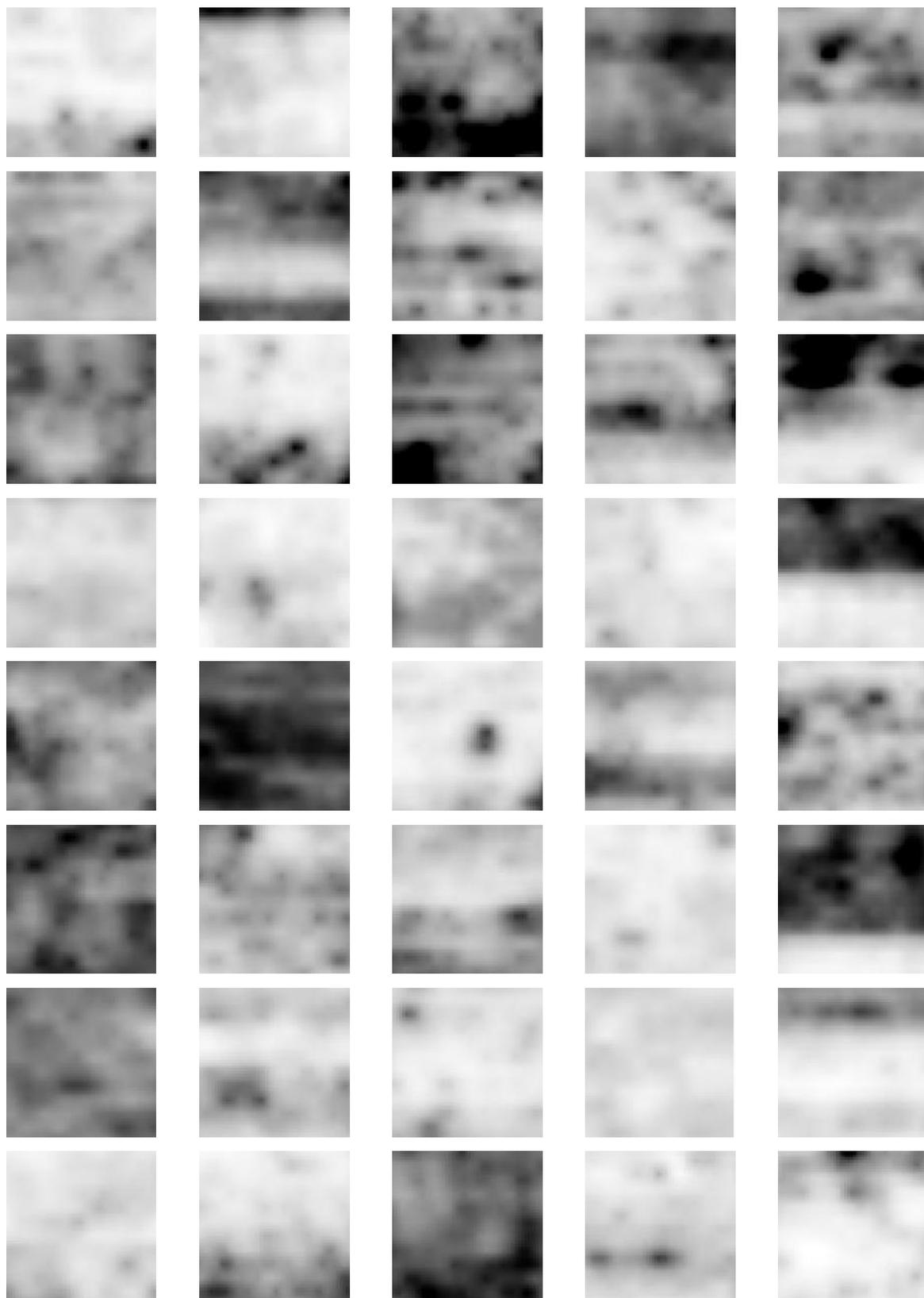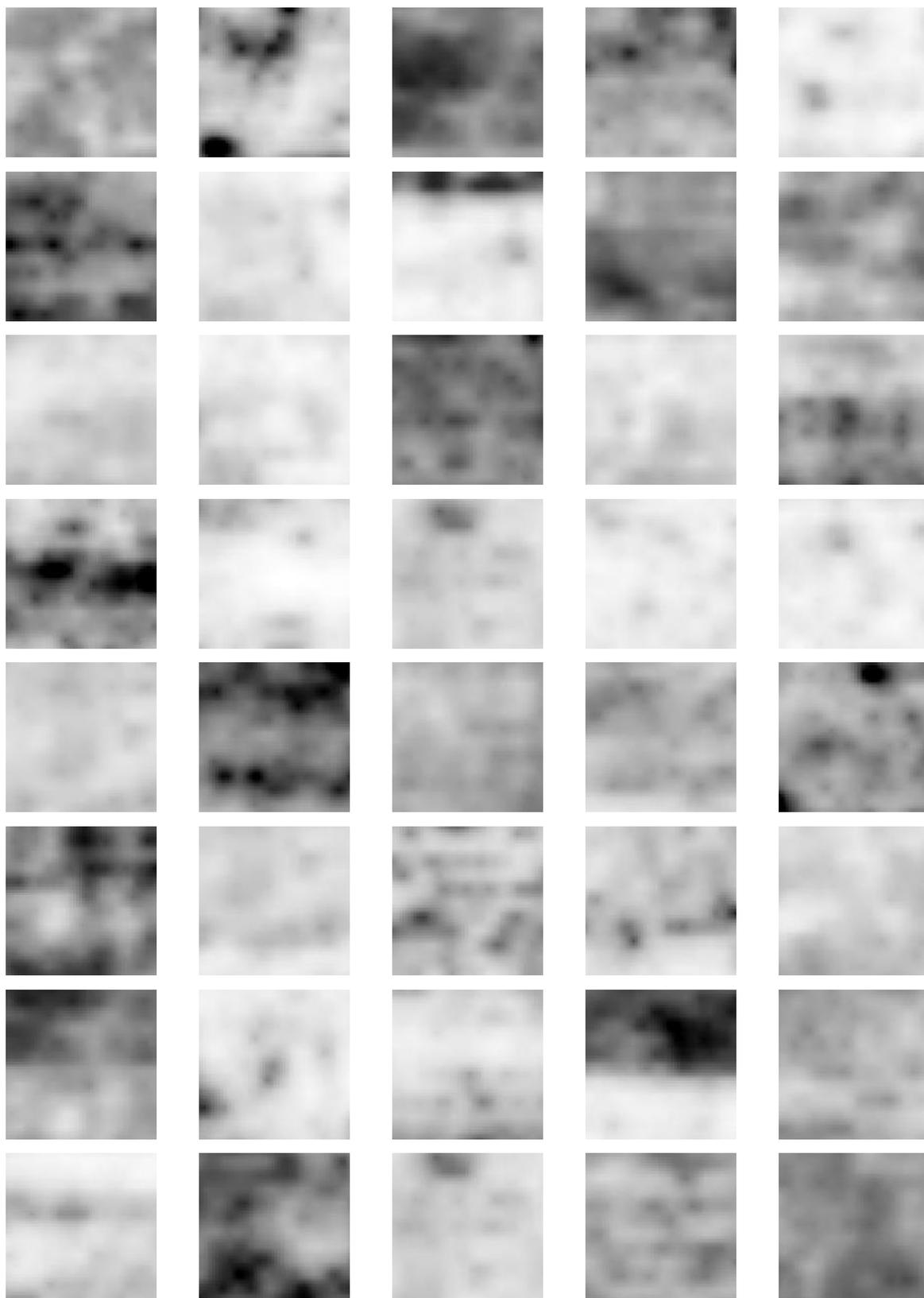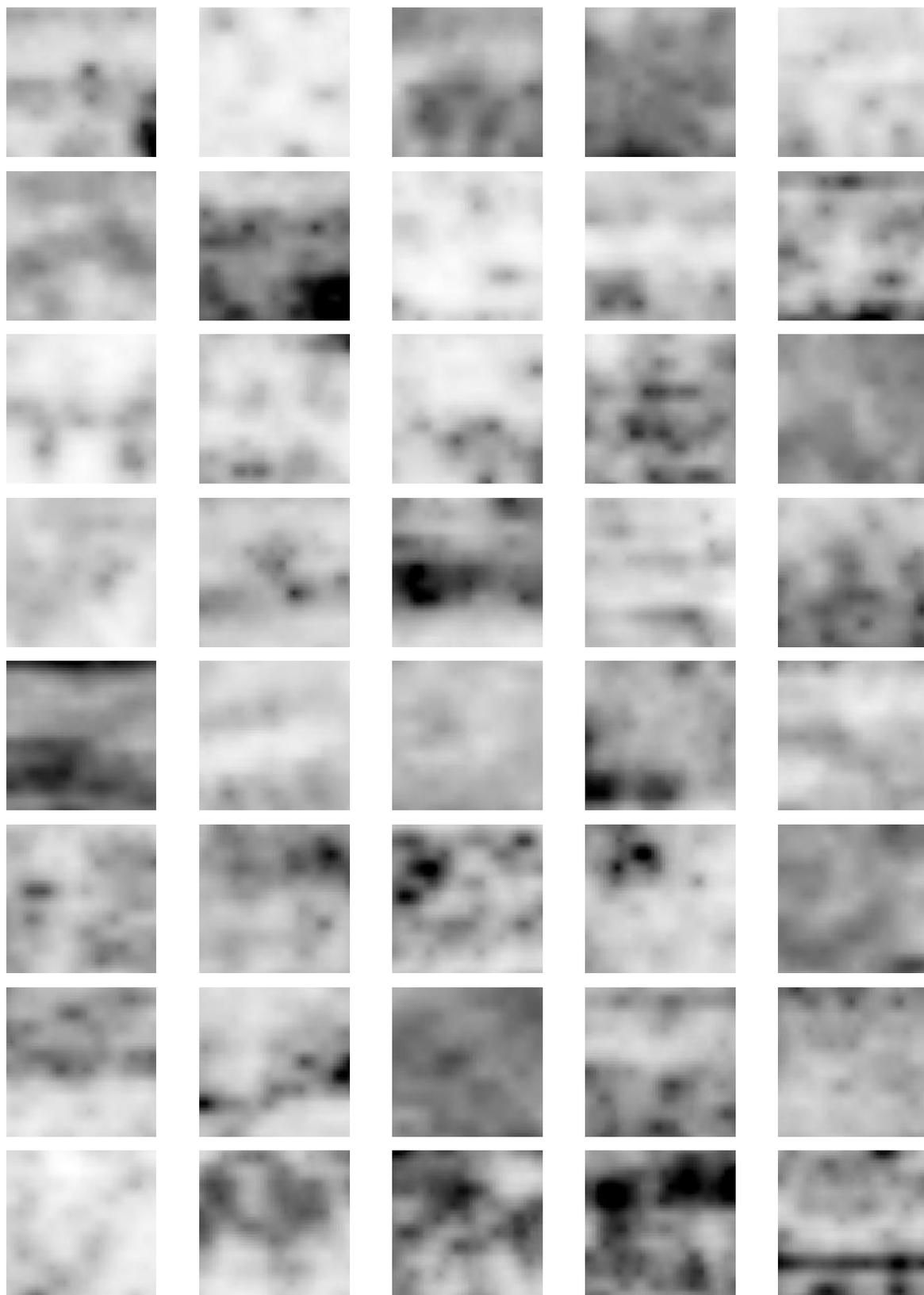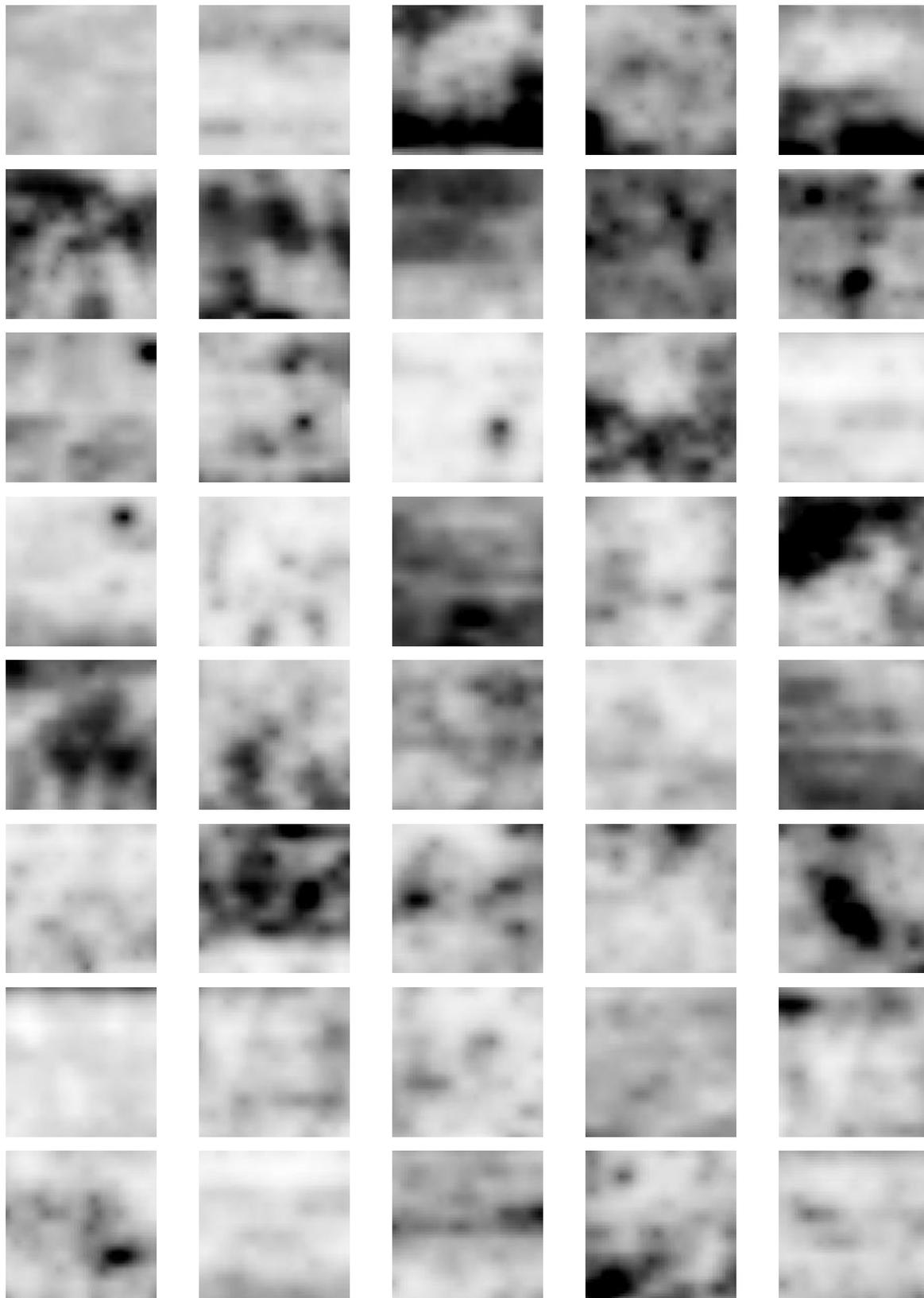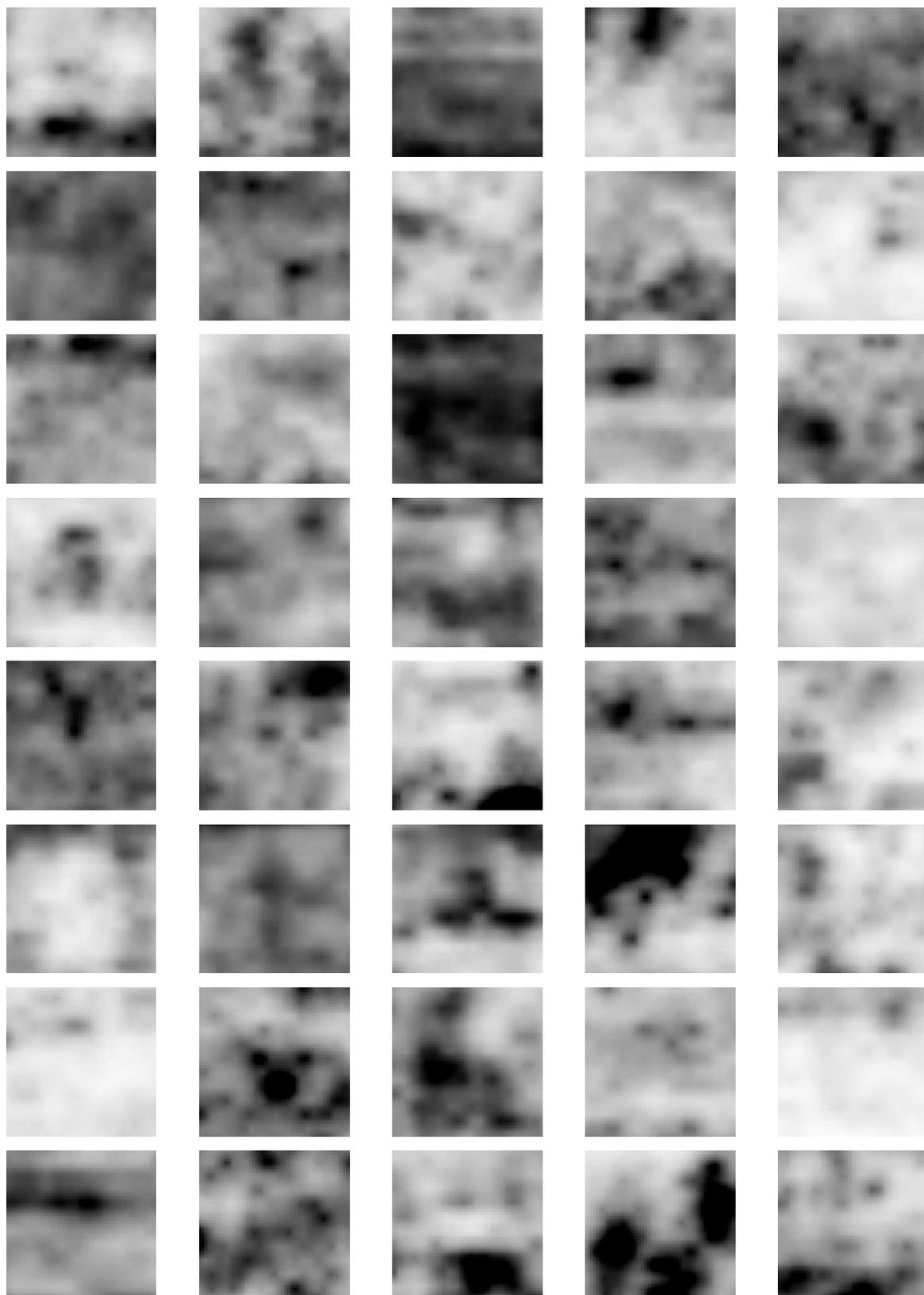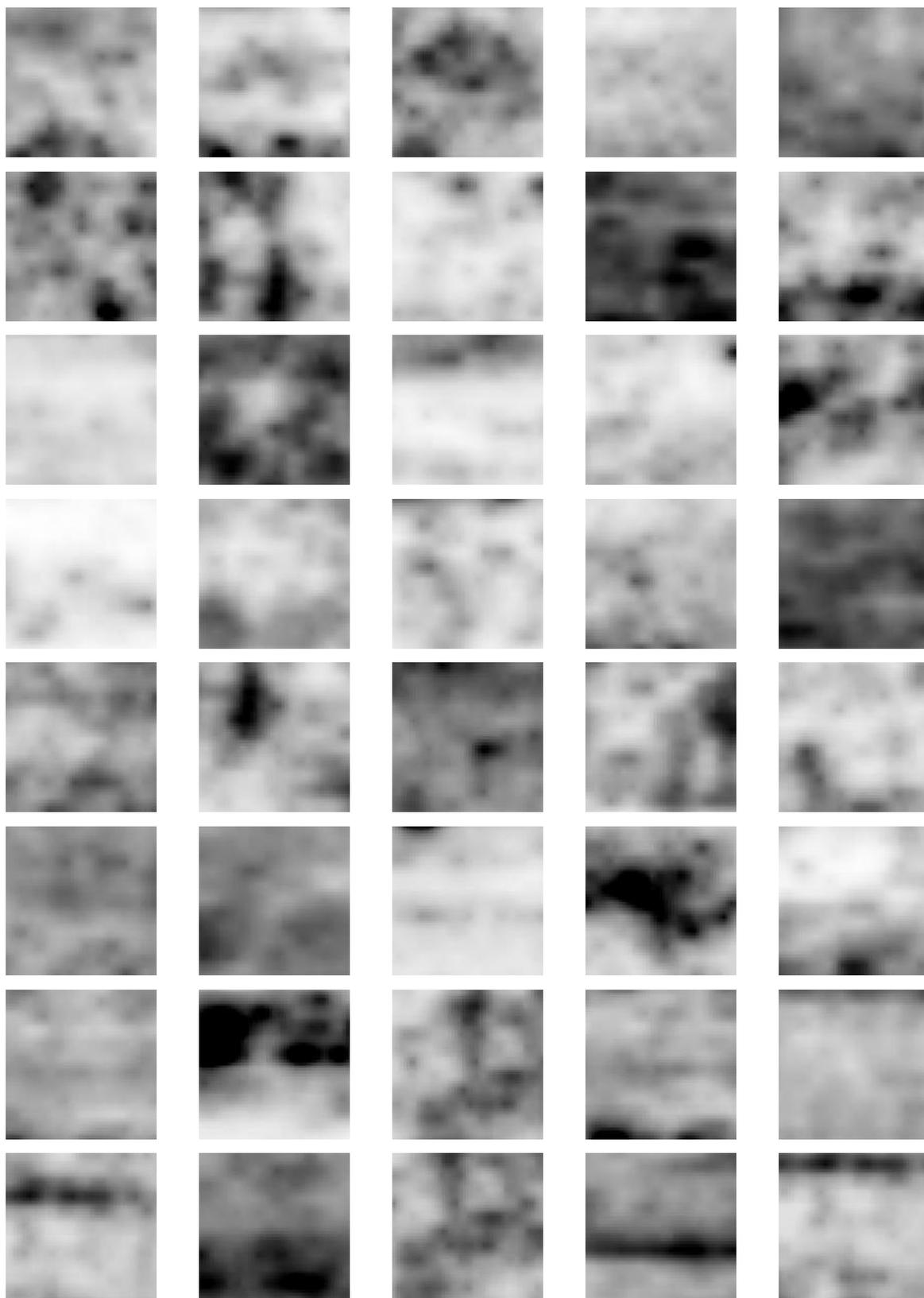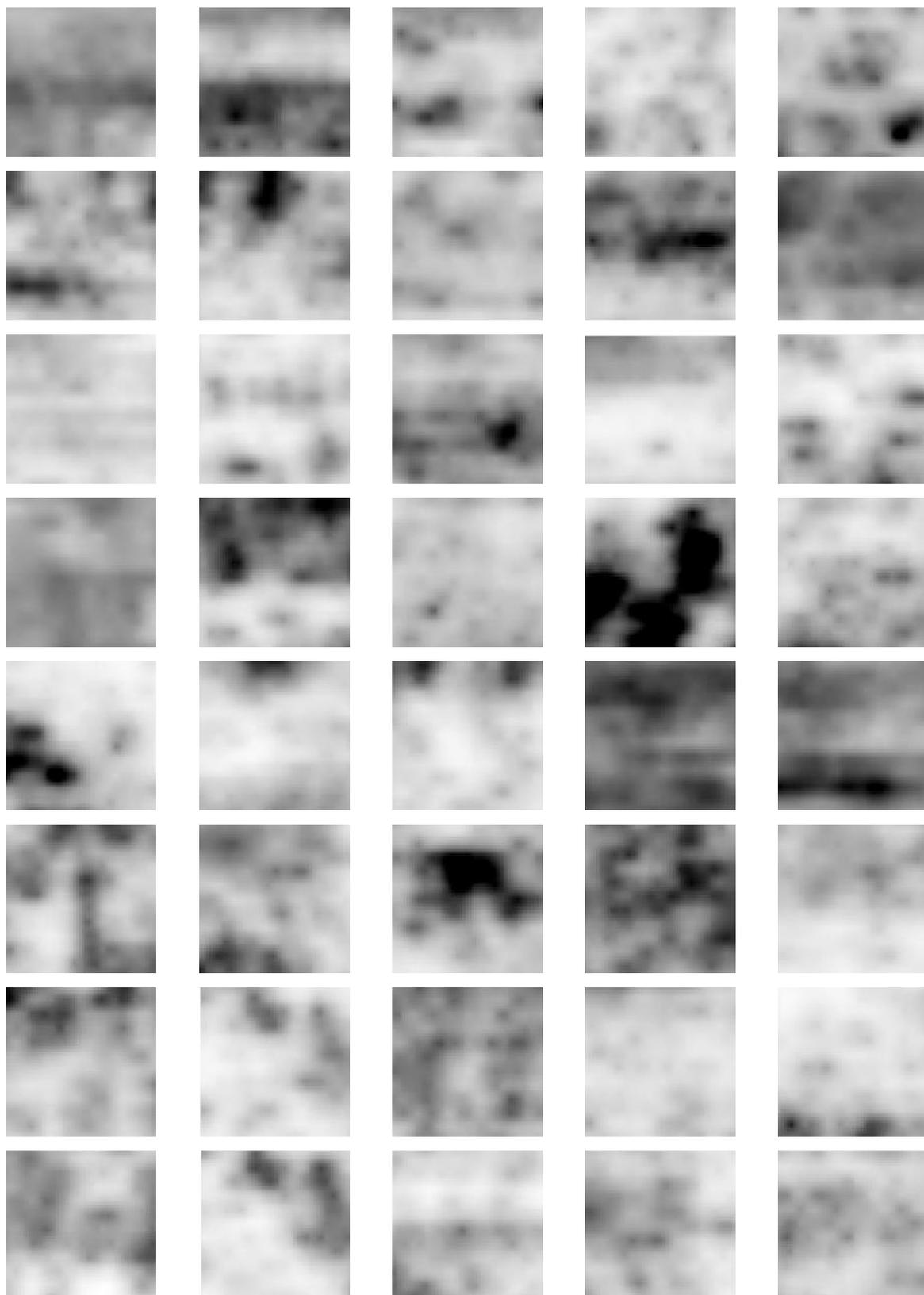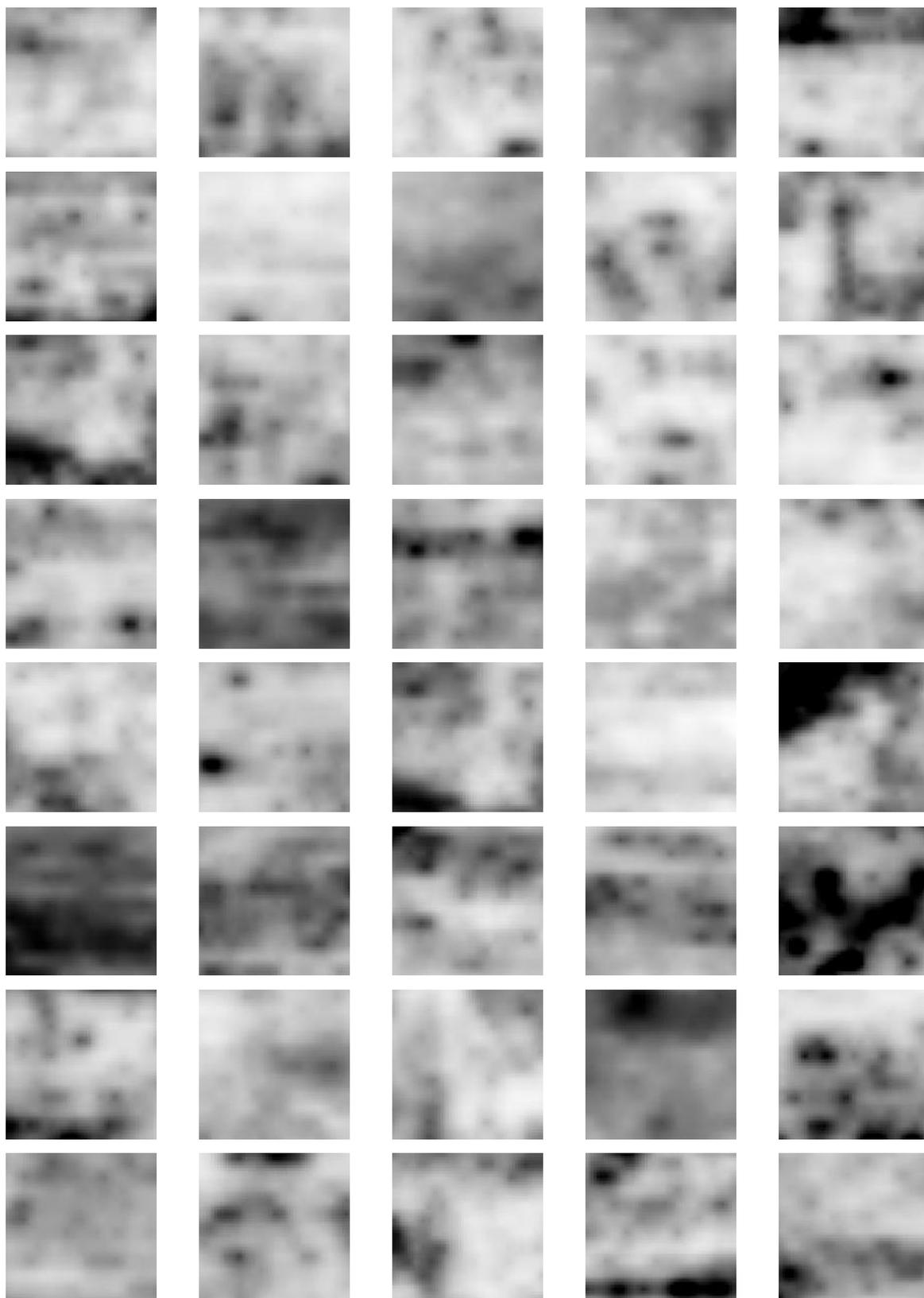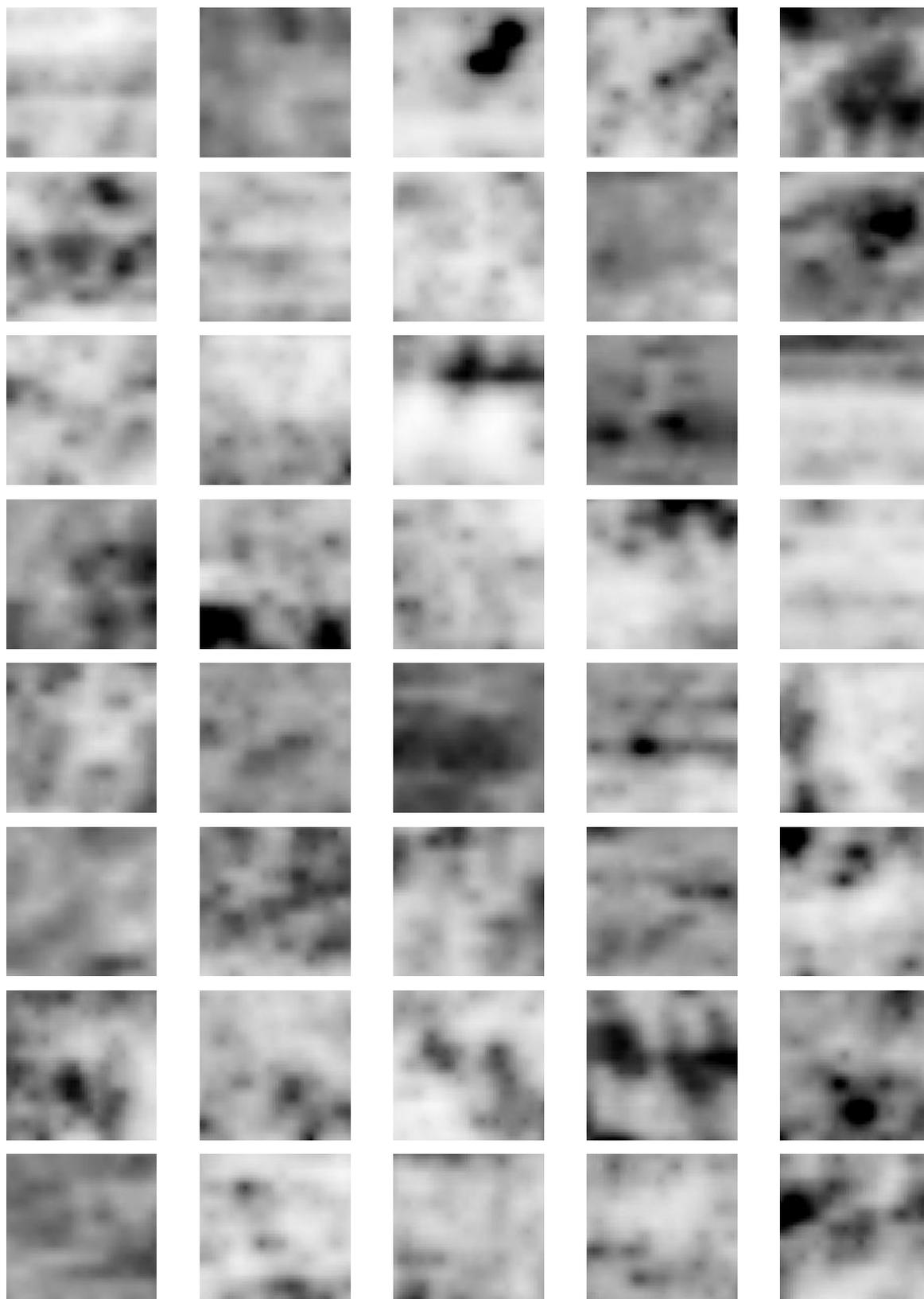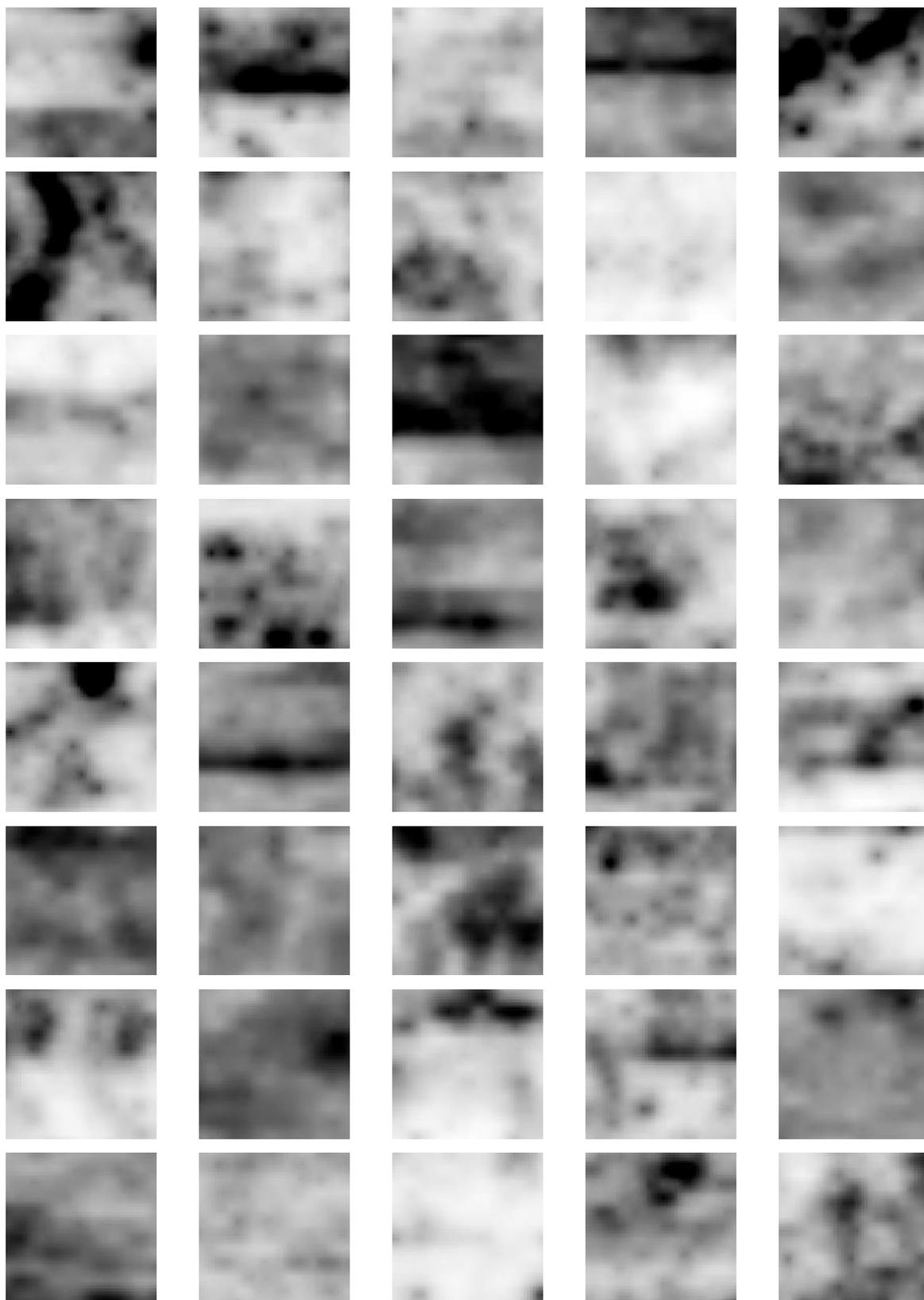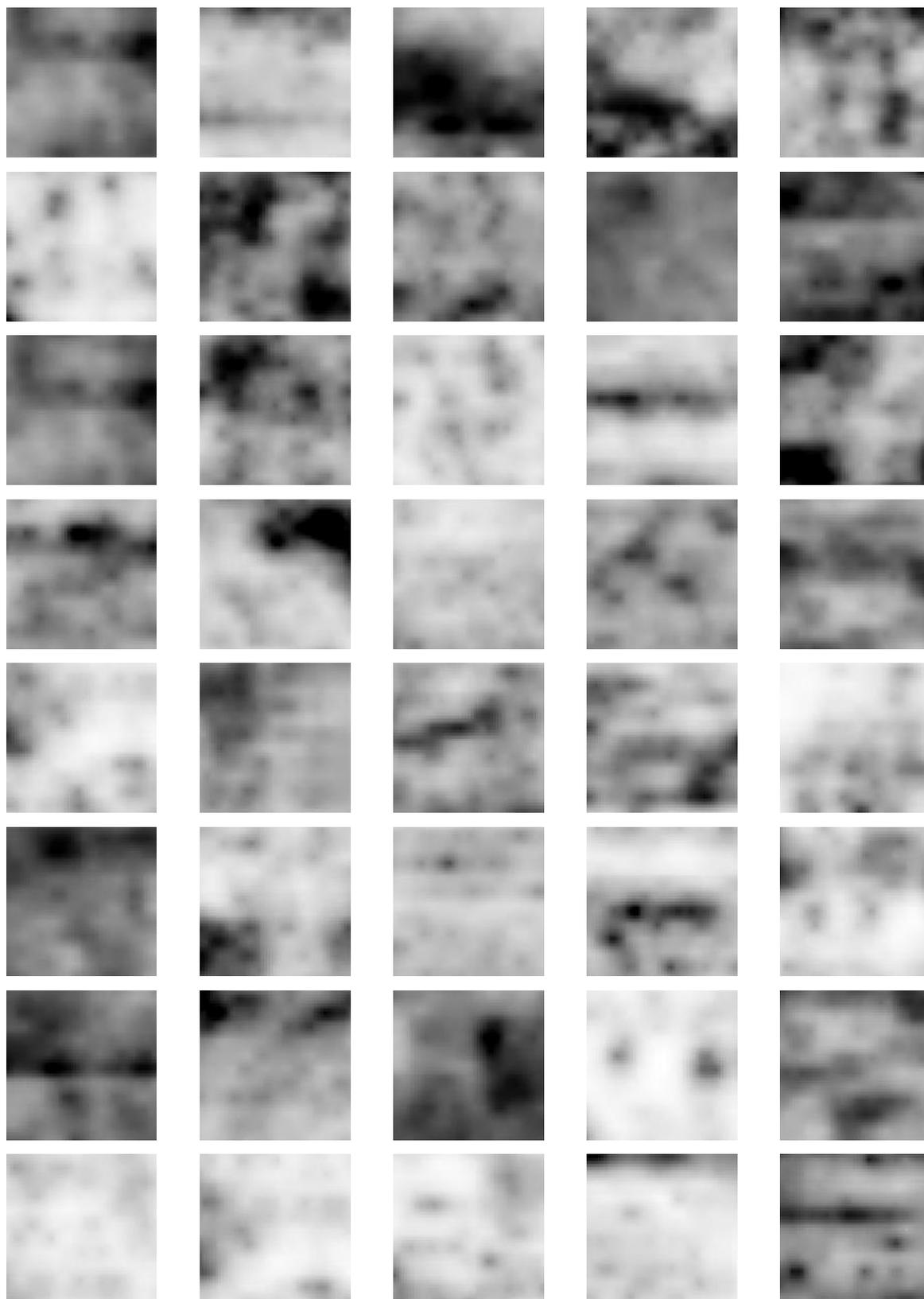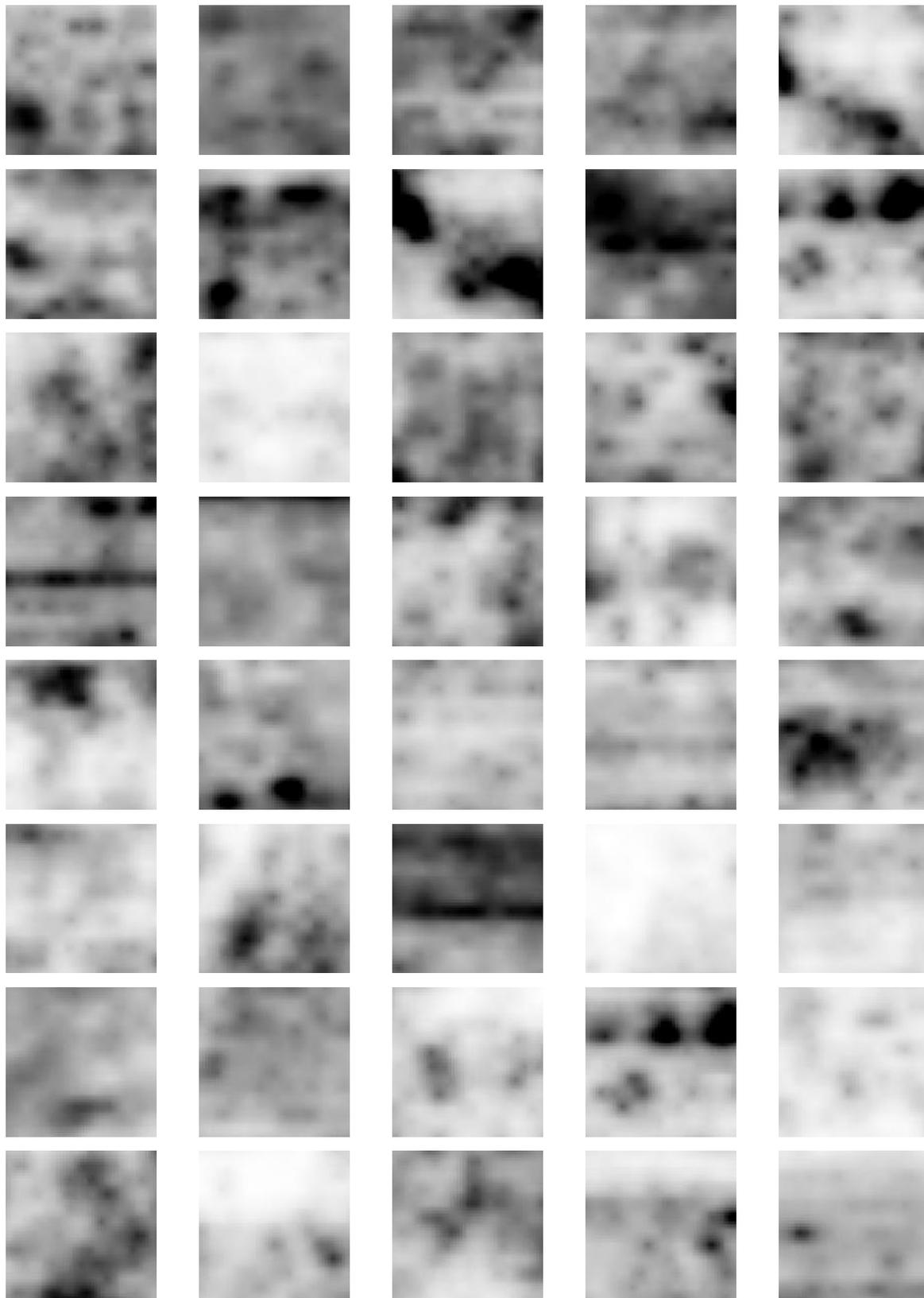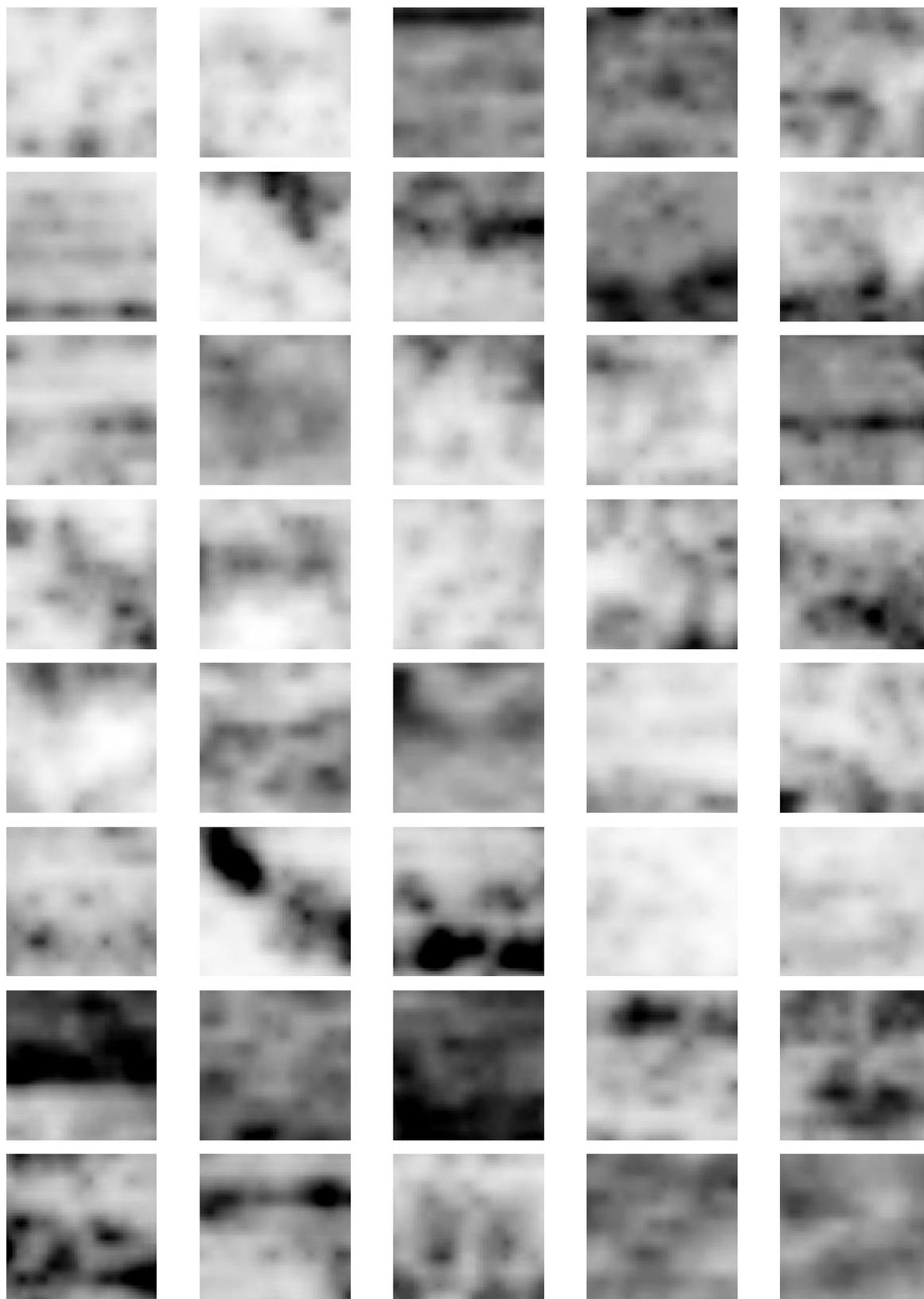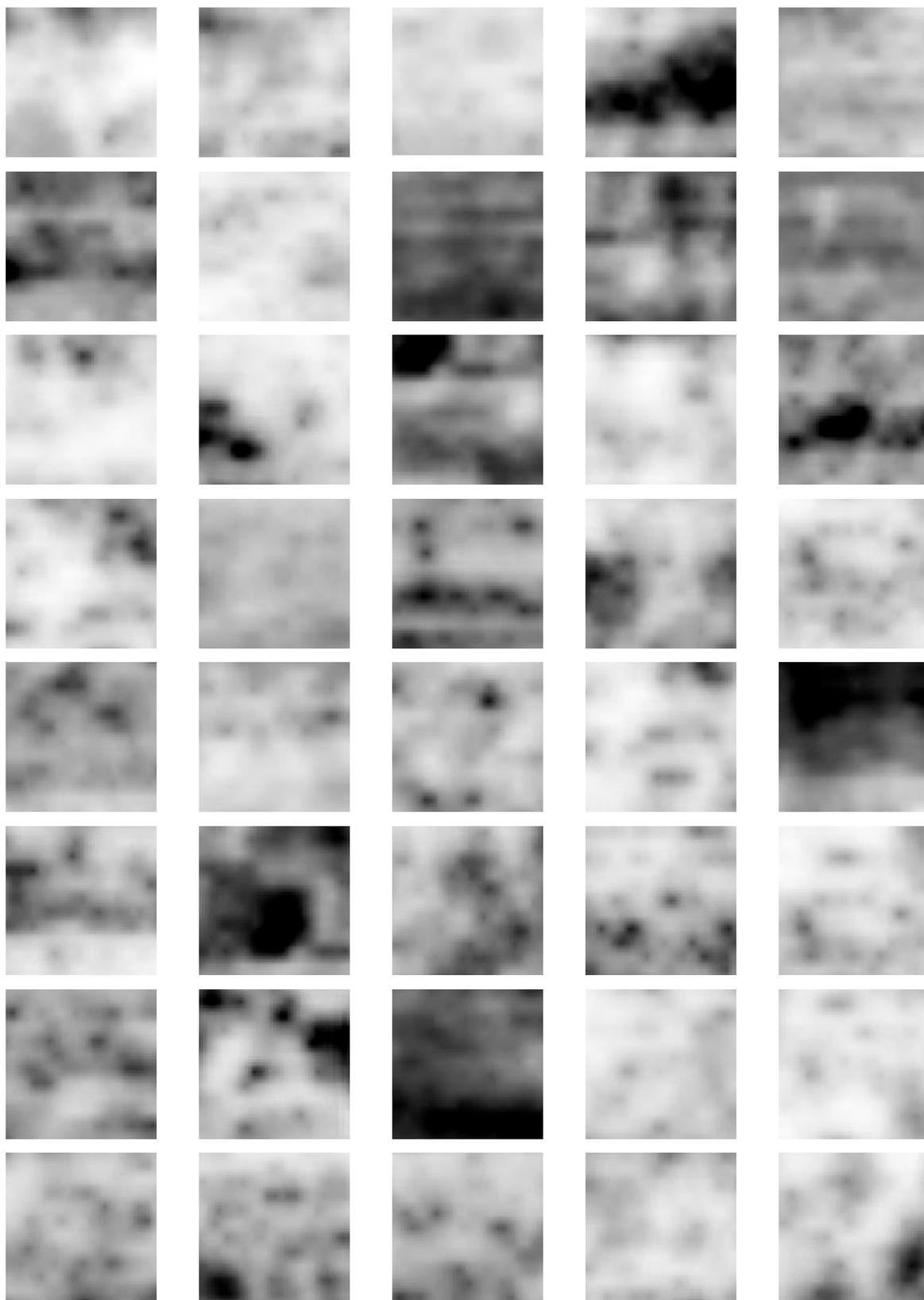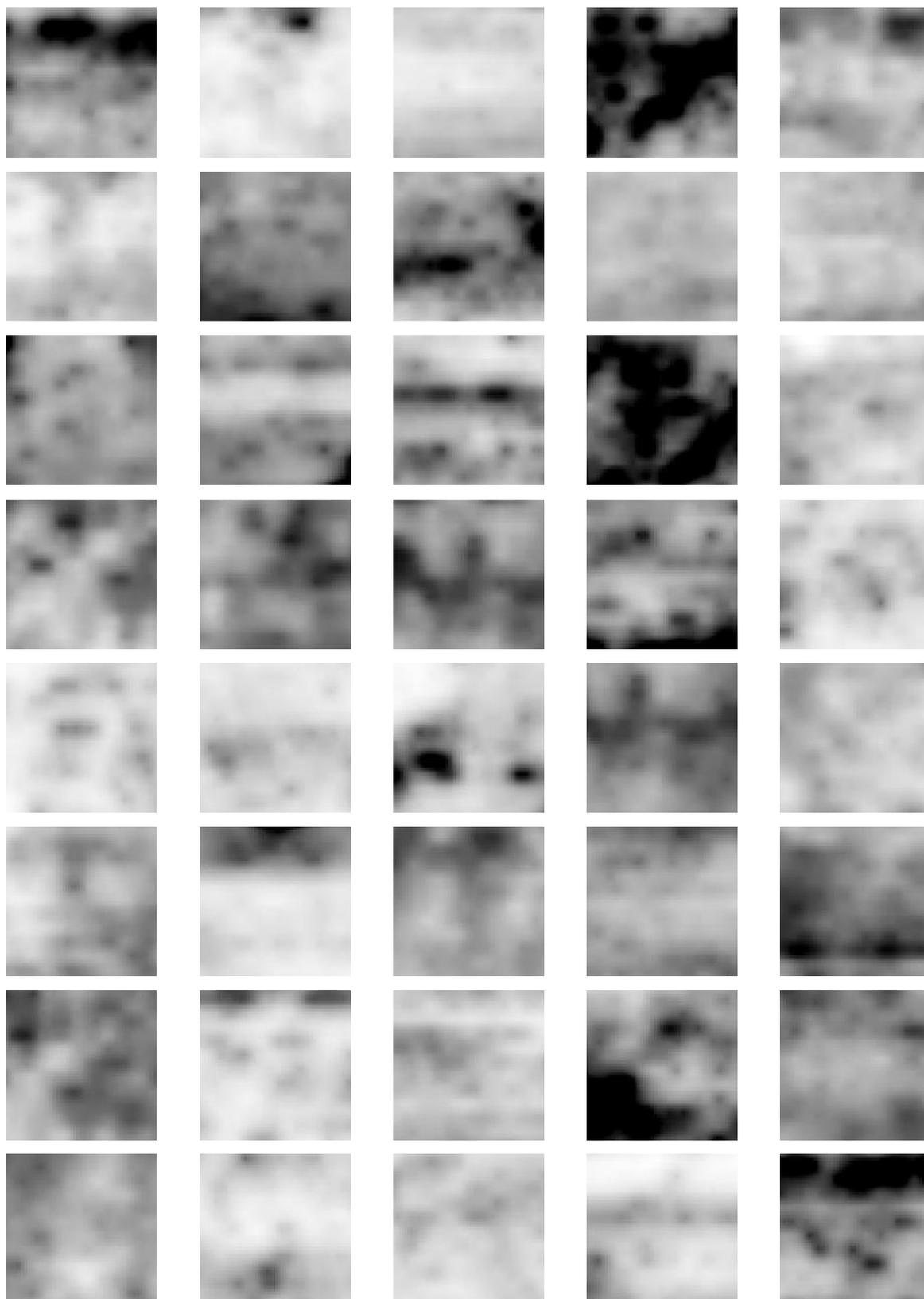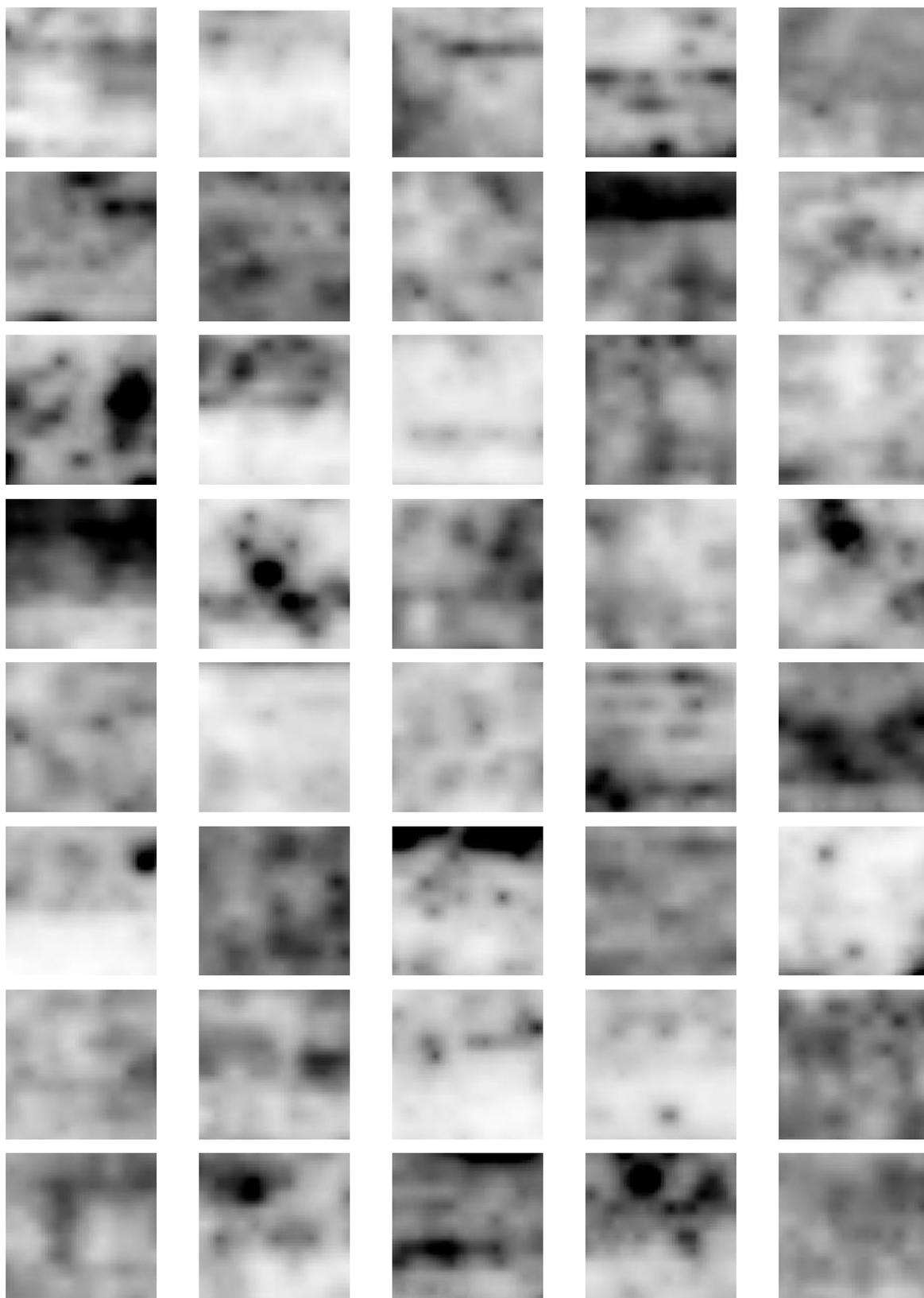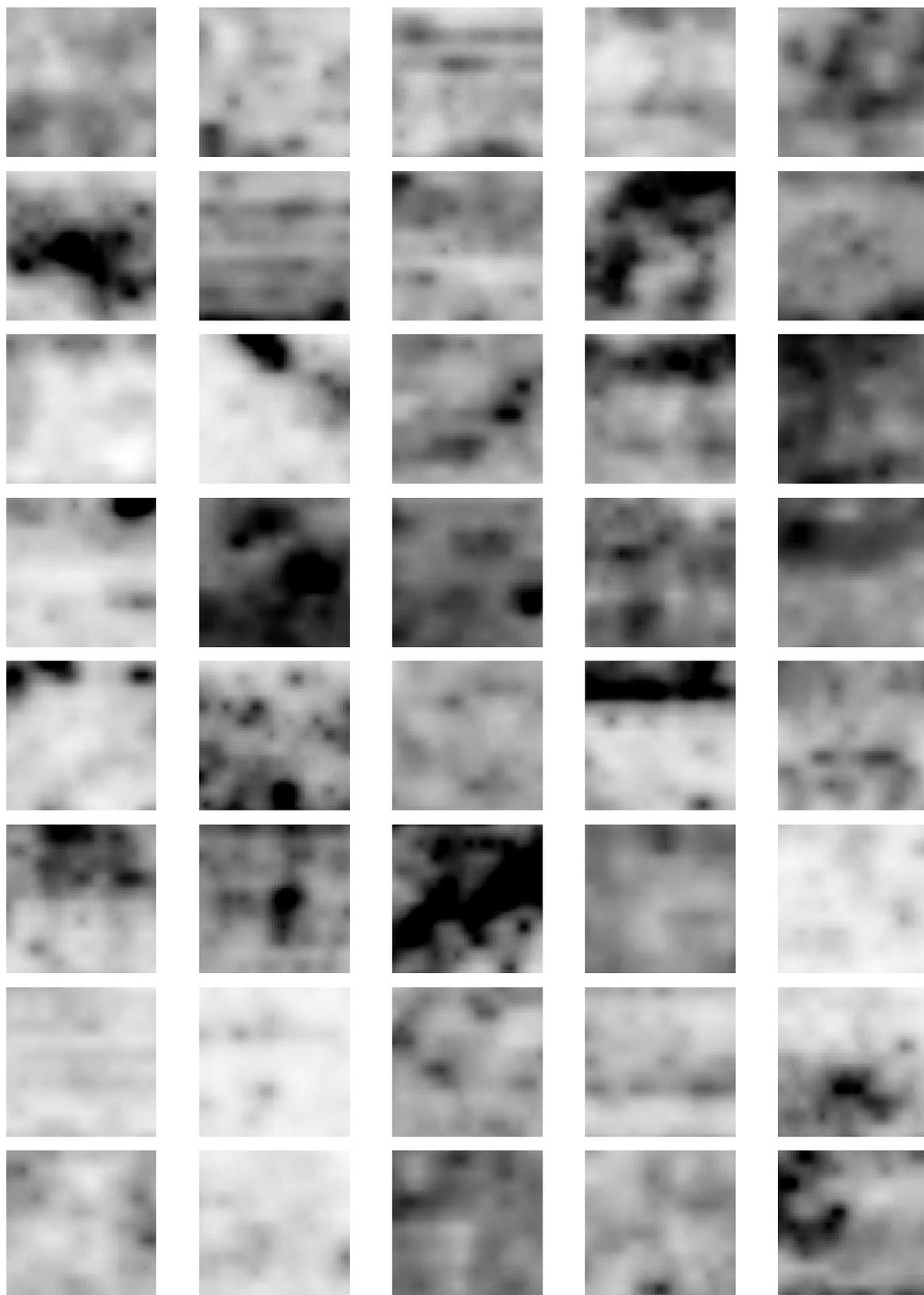
Timeslice Training Data: Grave Images
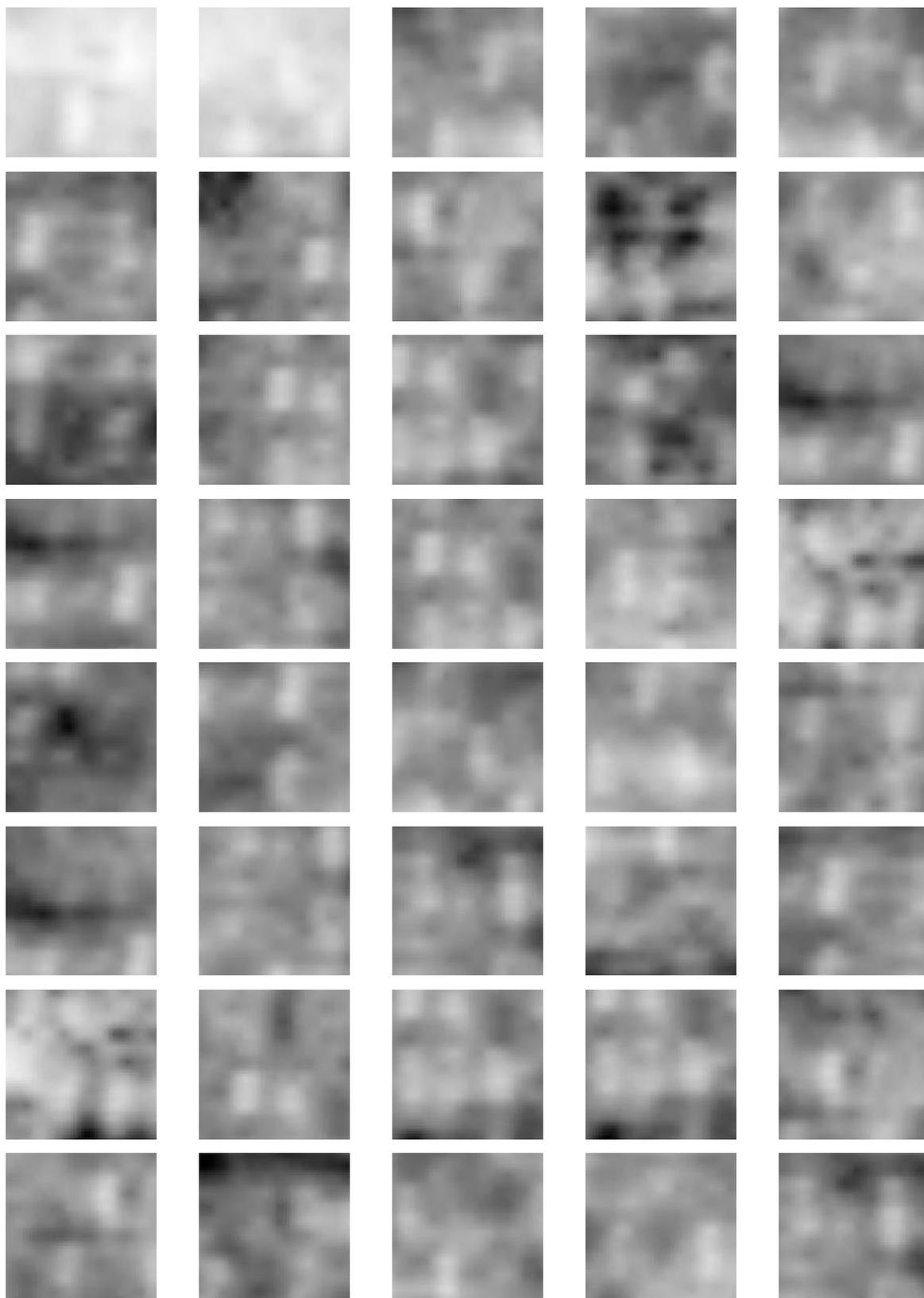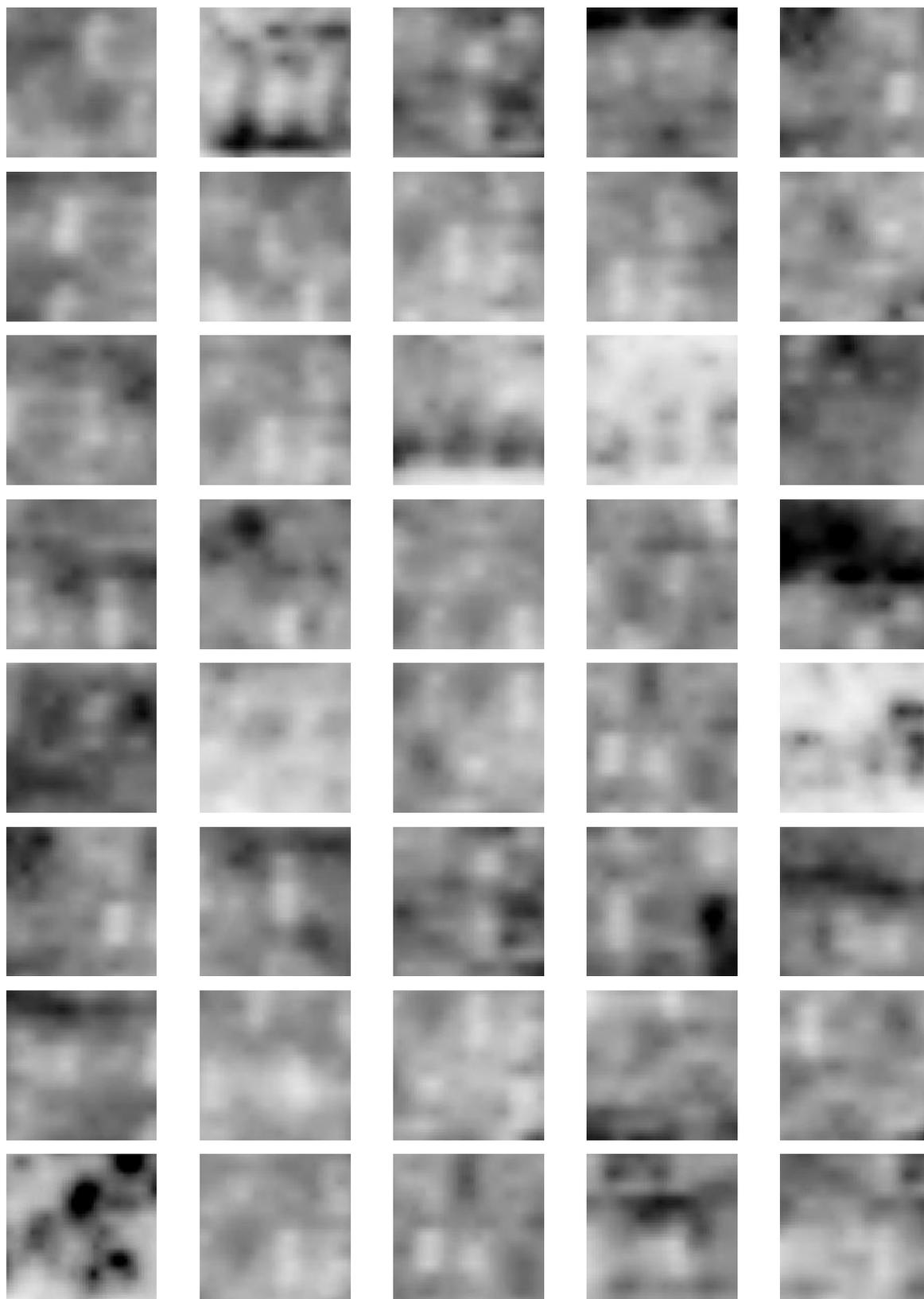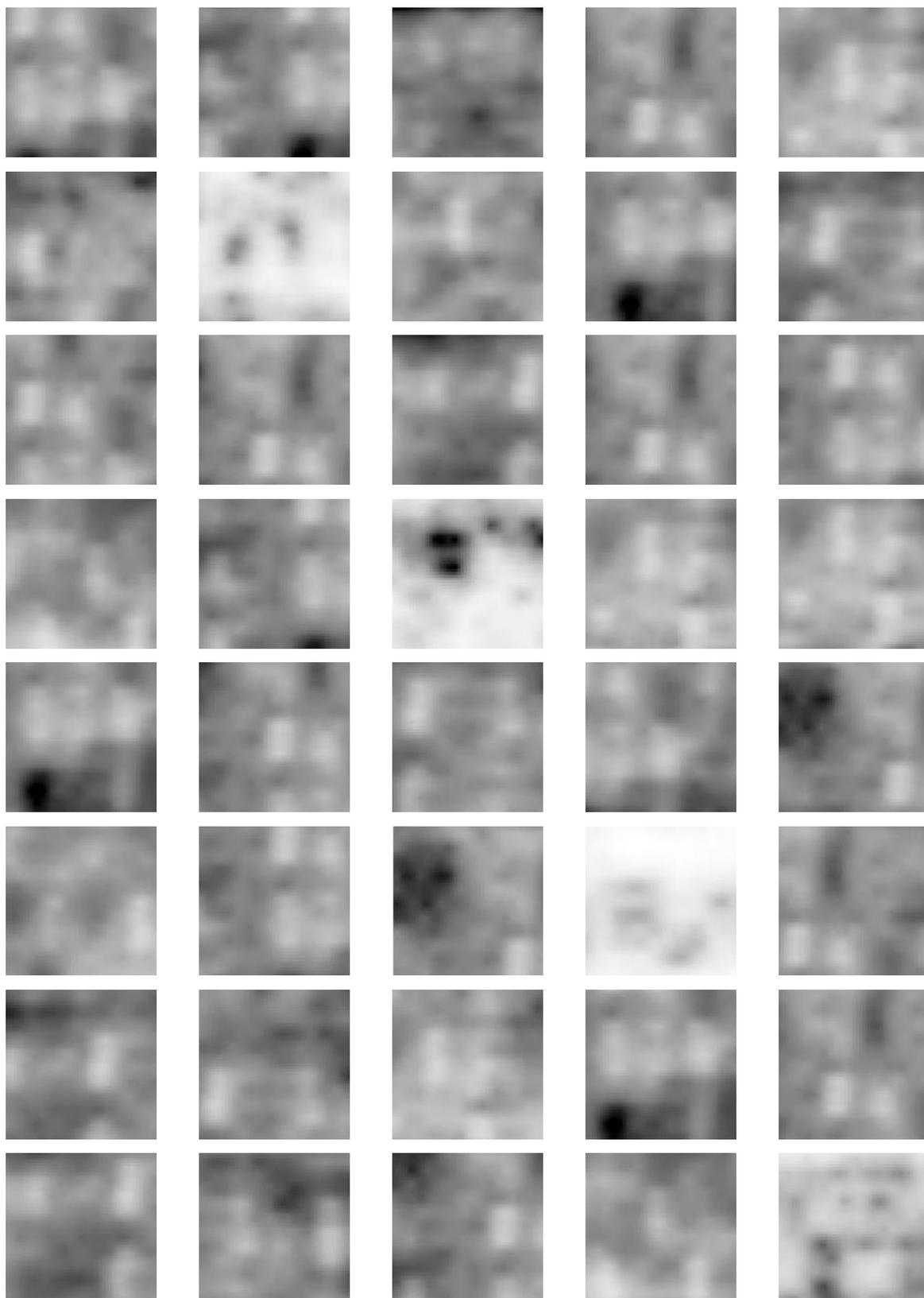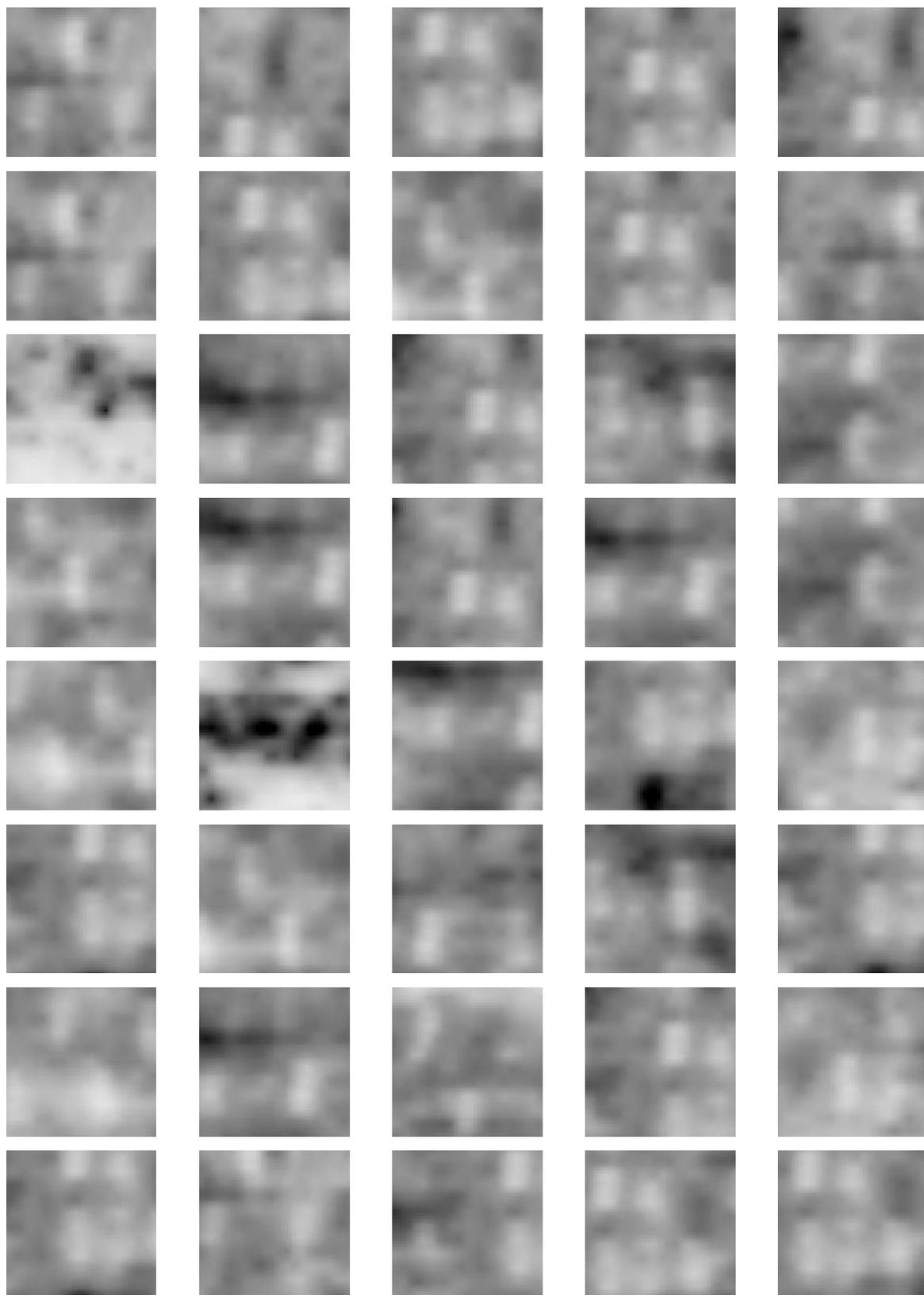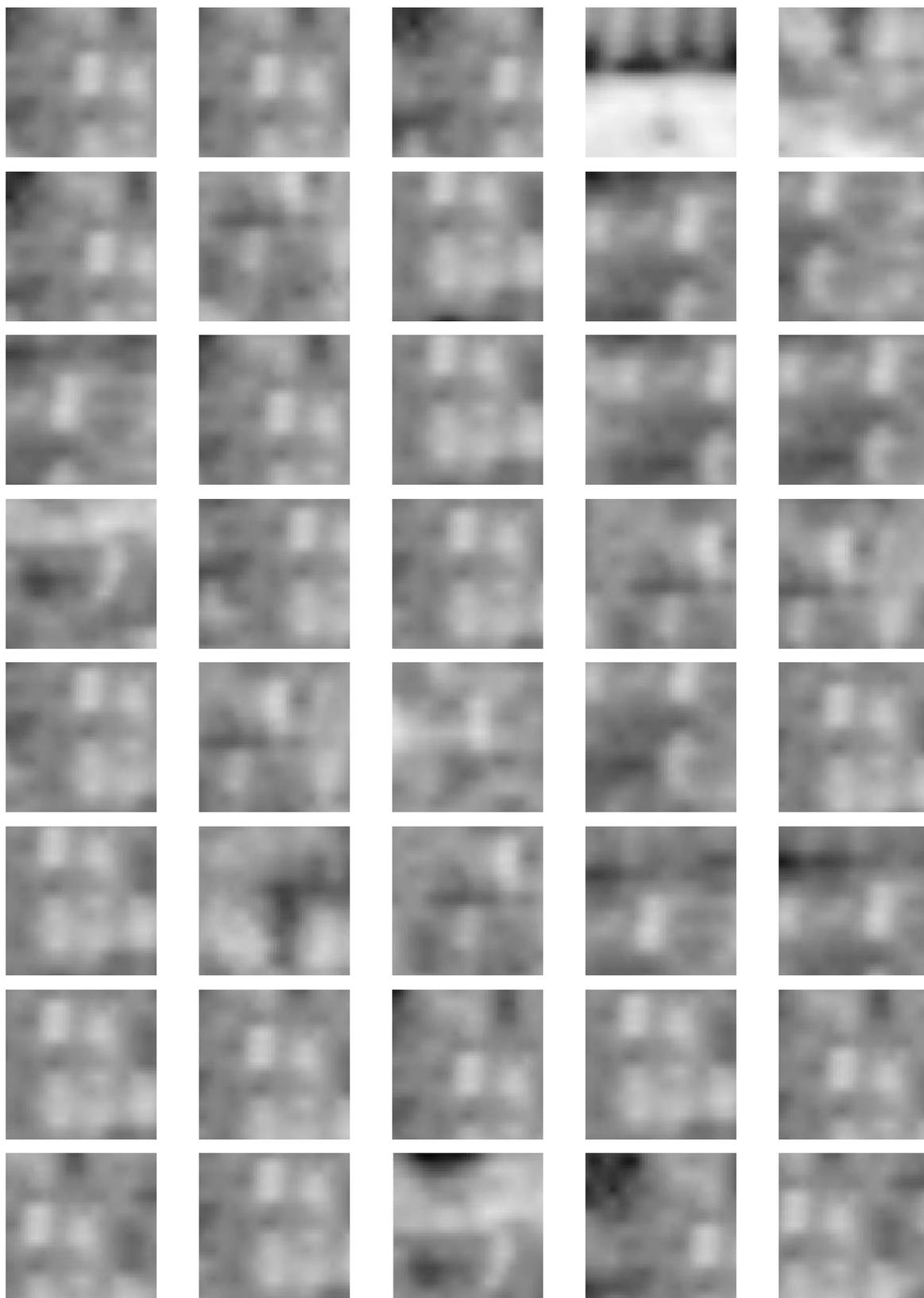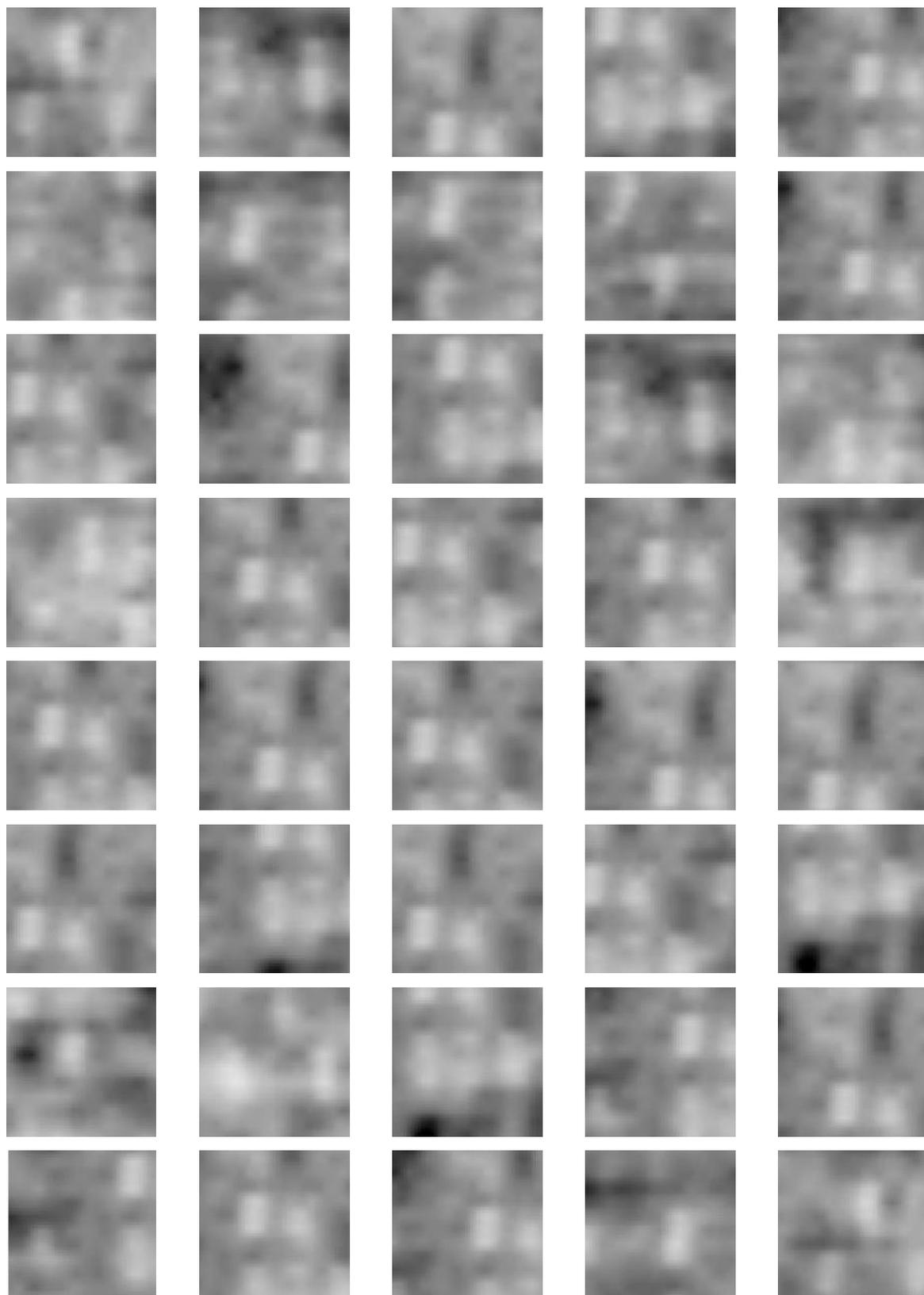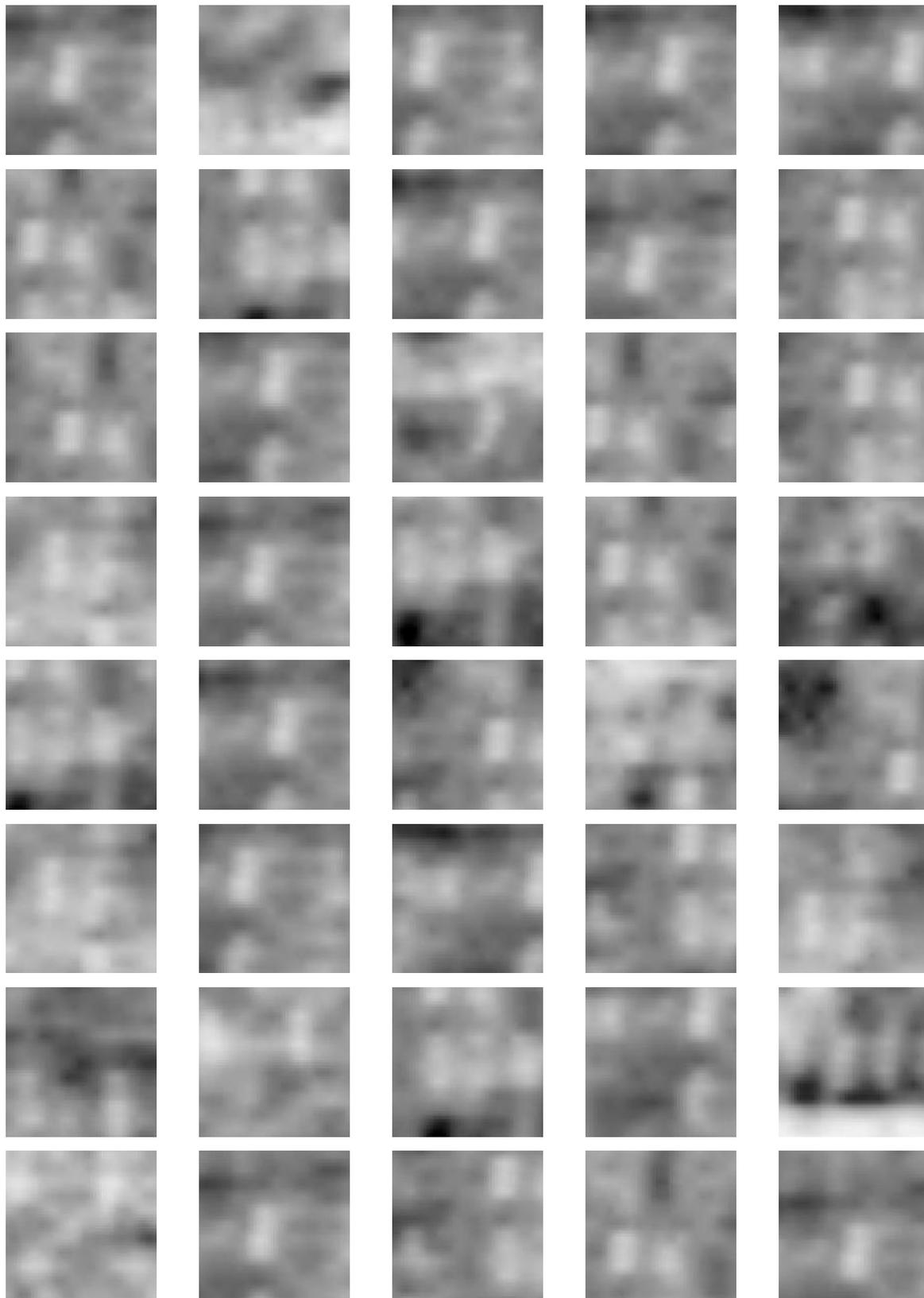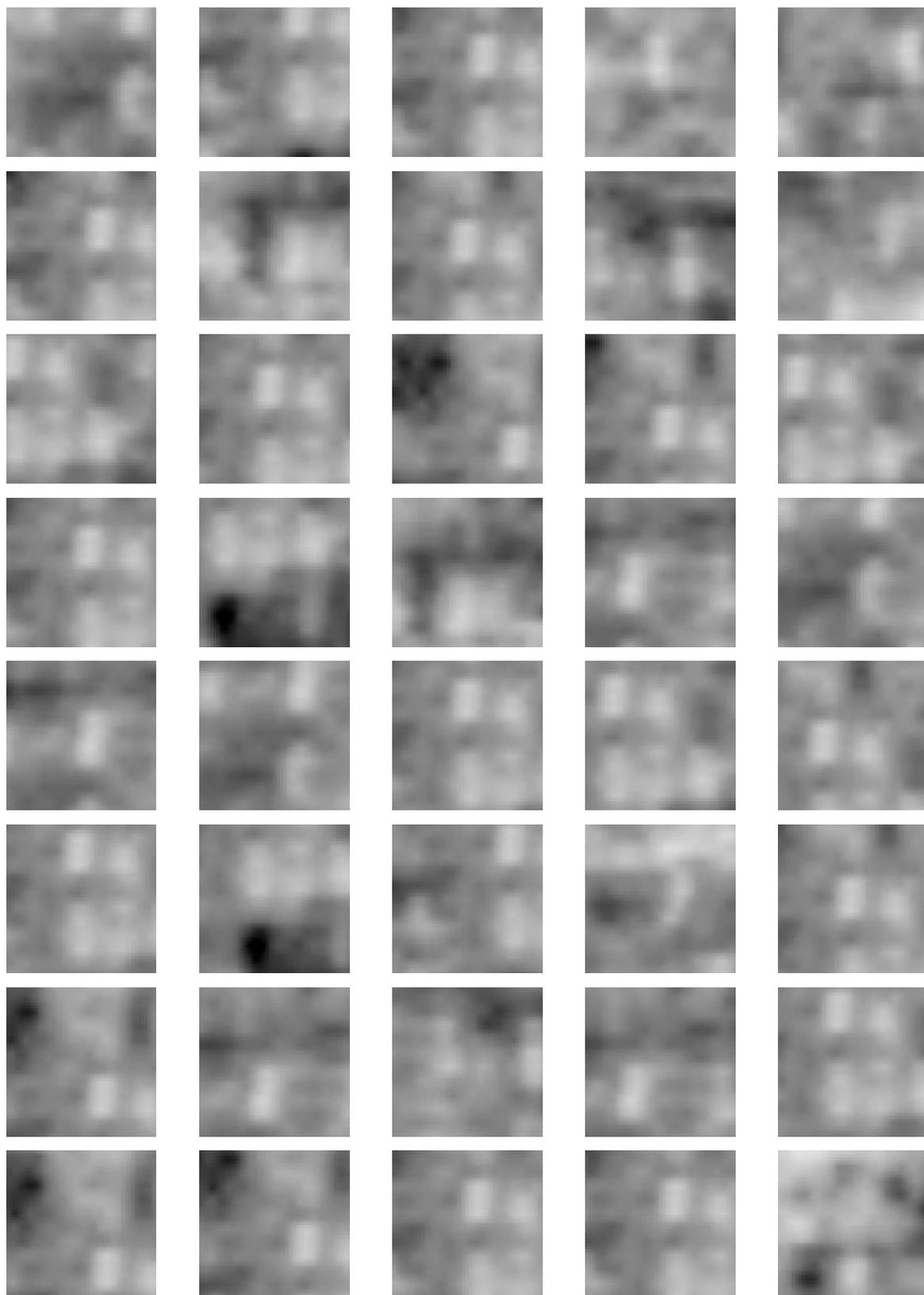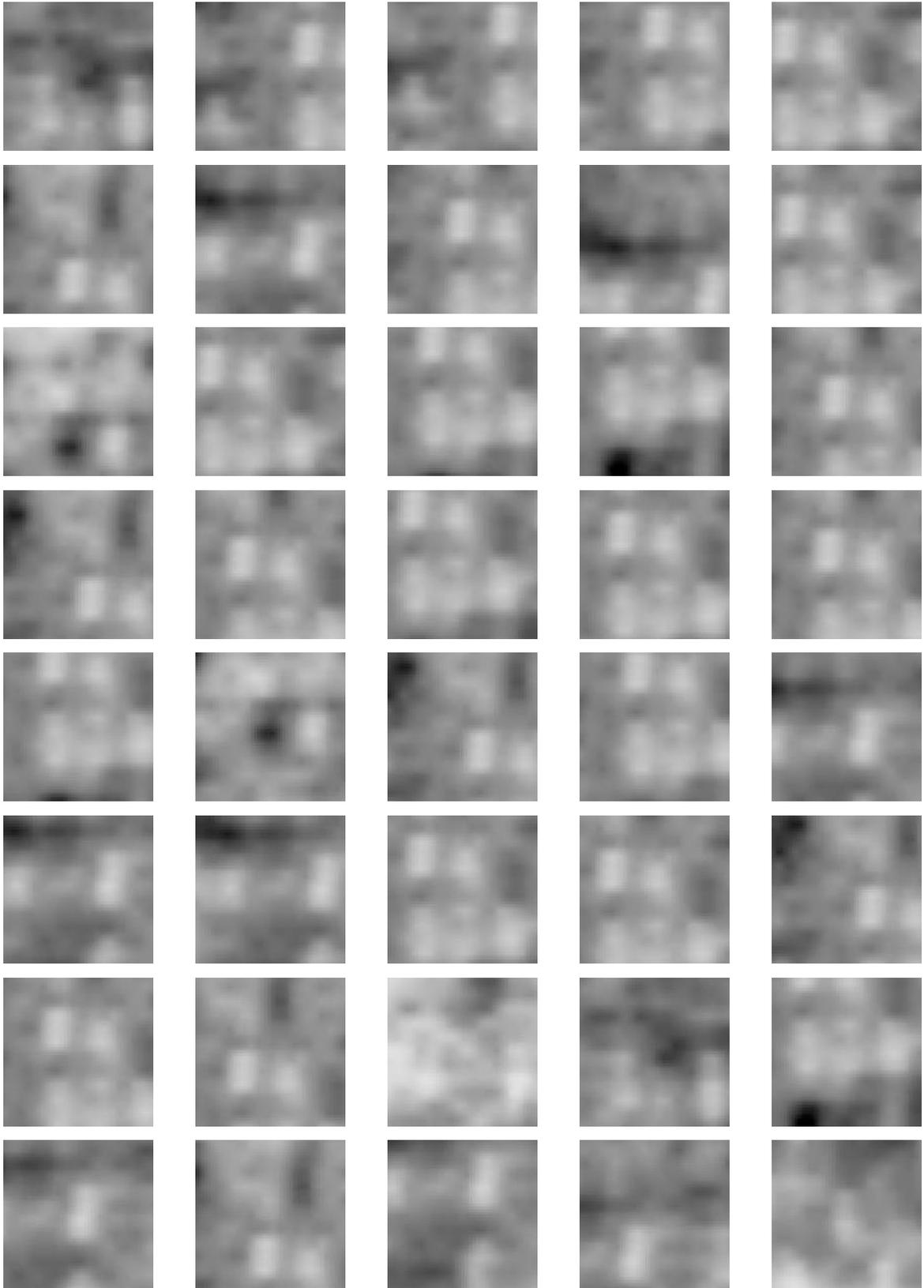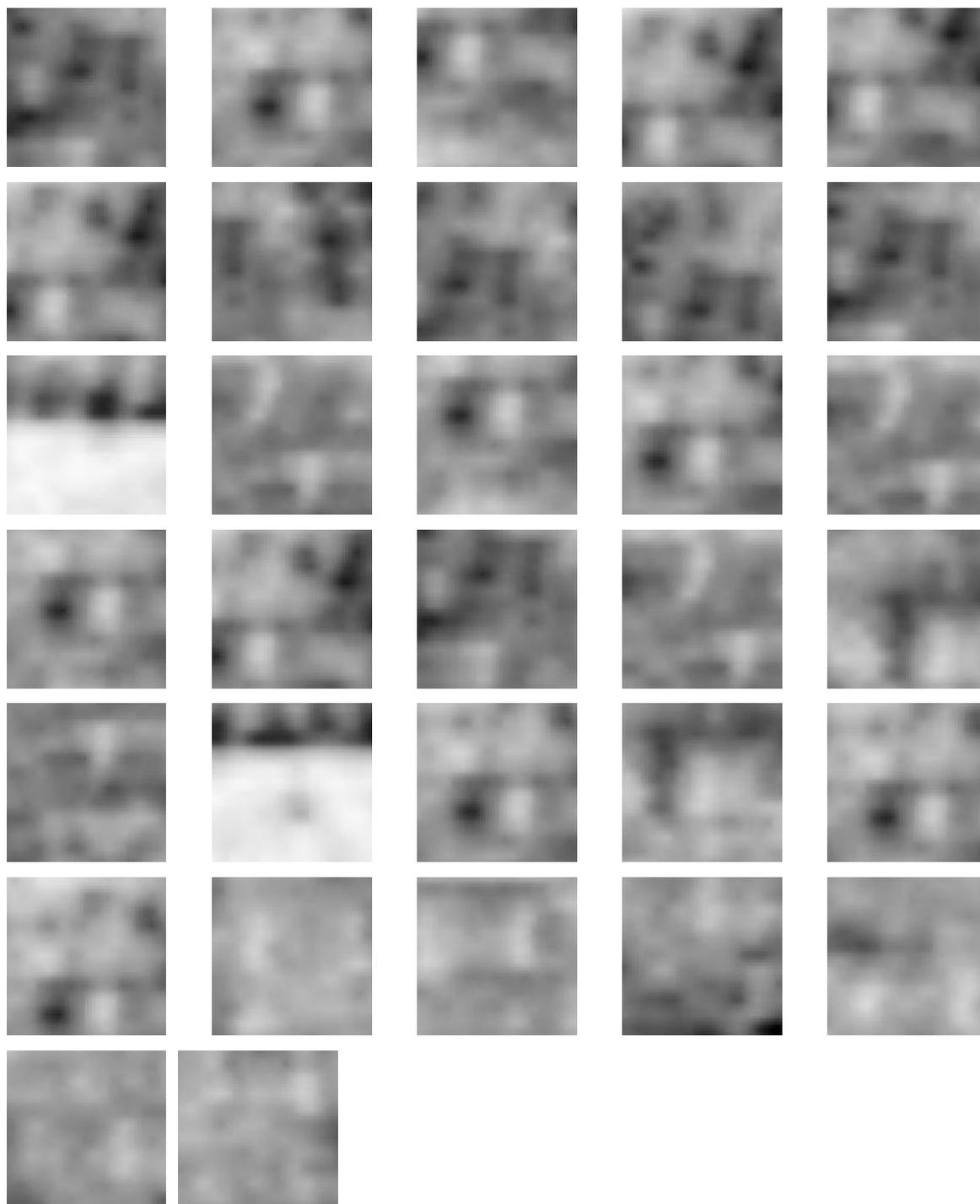
Appendix H.　　Rank Values Used in Suitability Analysis

As the salinity of waterlogged areas could not be determined from opensource data, the suitability of techniques for these sites are largely ranked with a 0 score below. It should, however, be noted that GPR and additio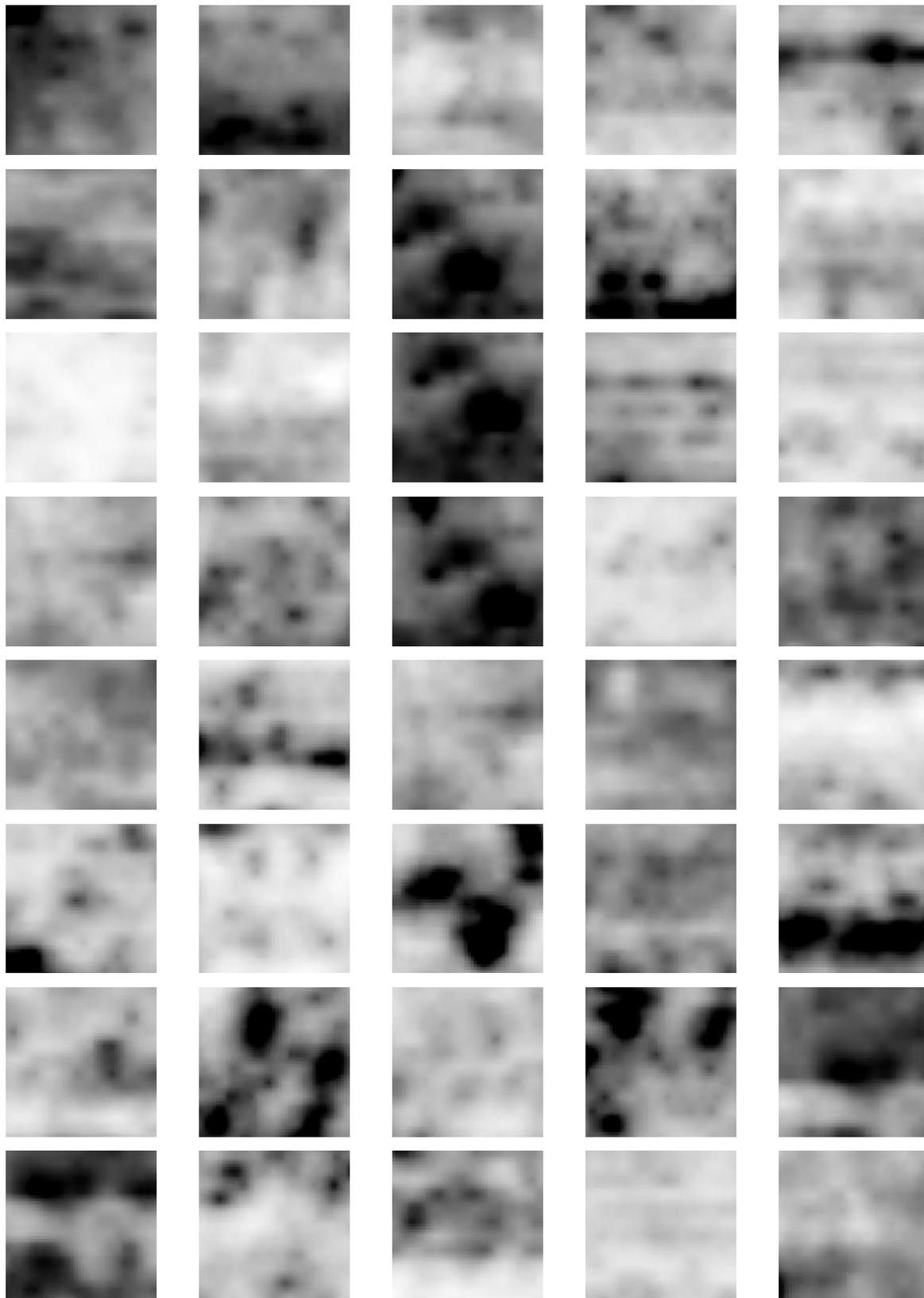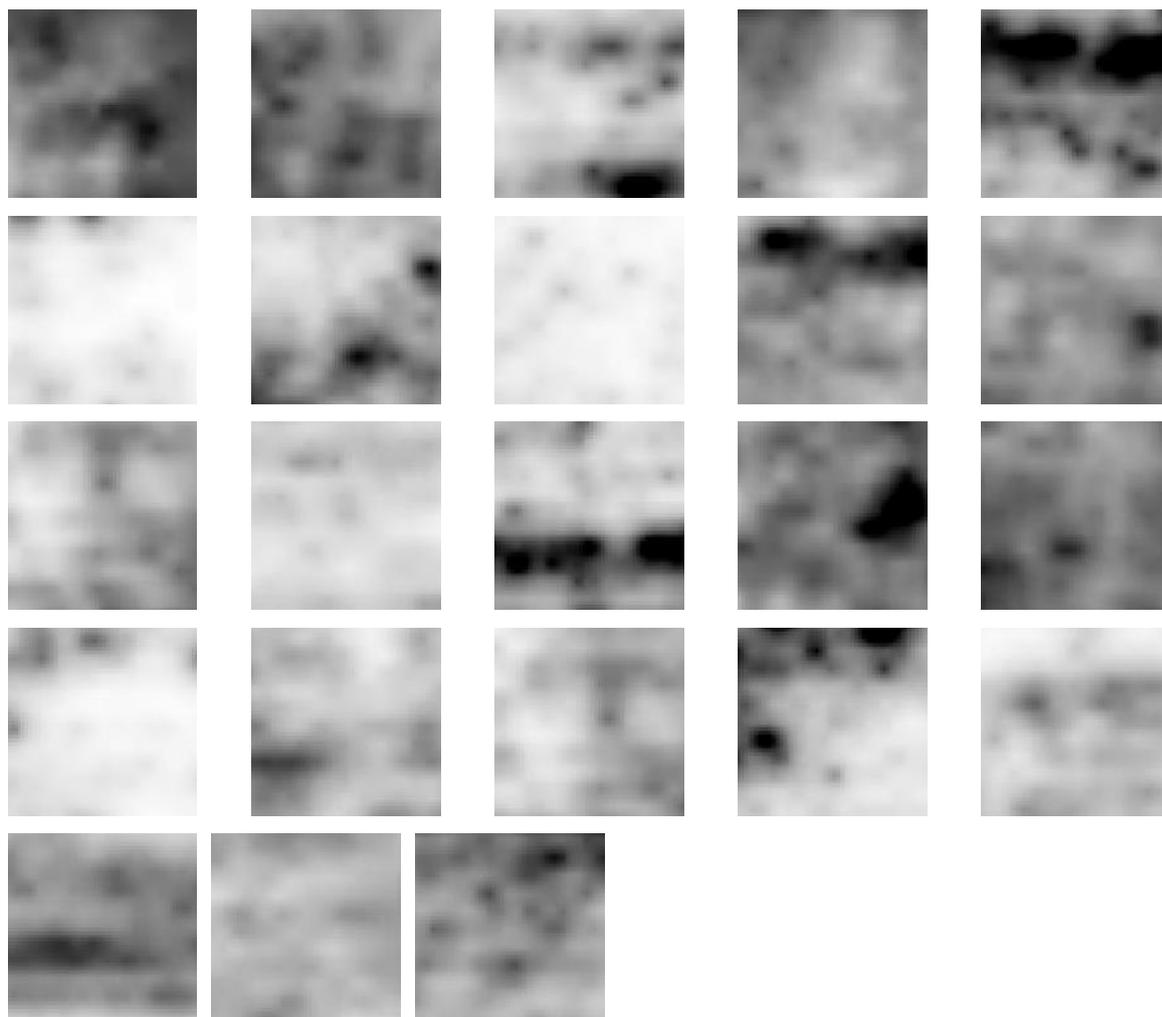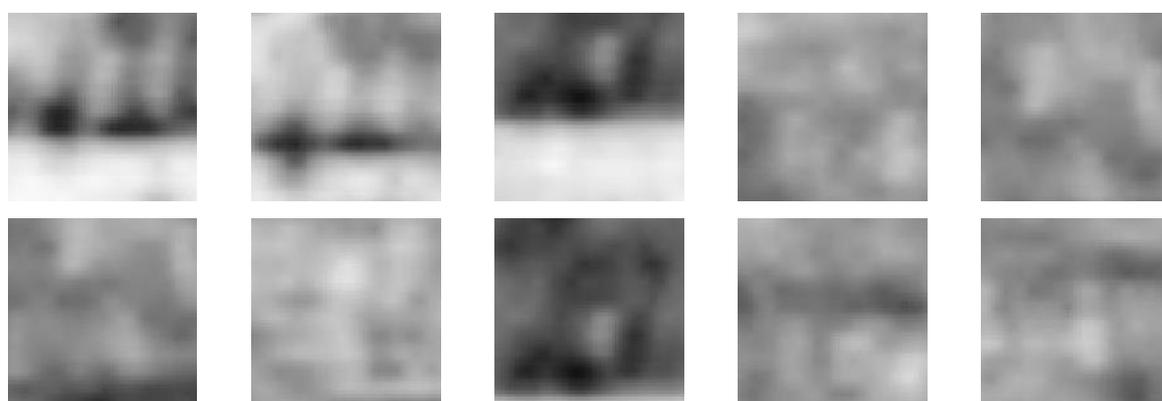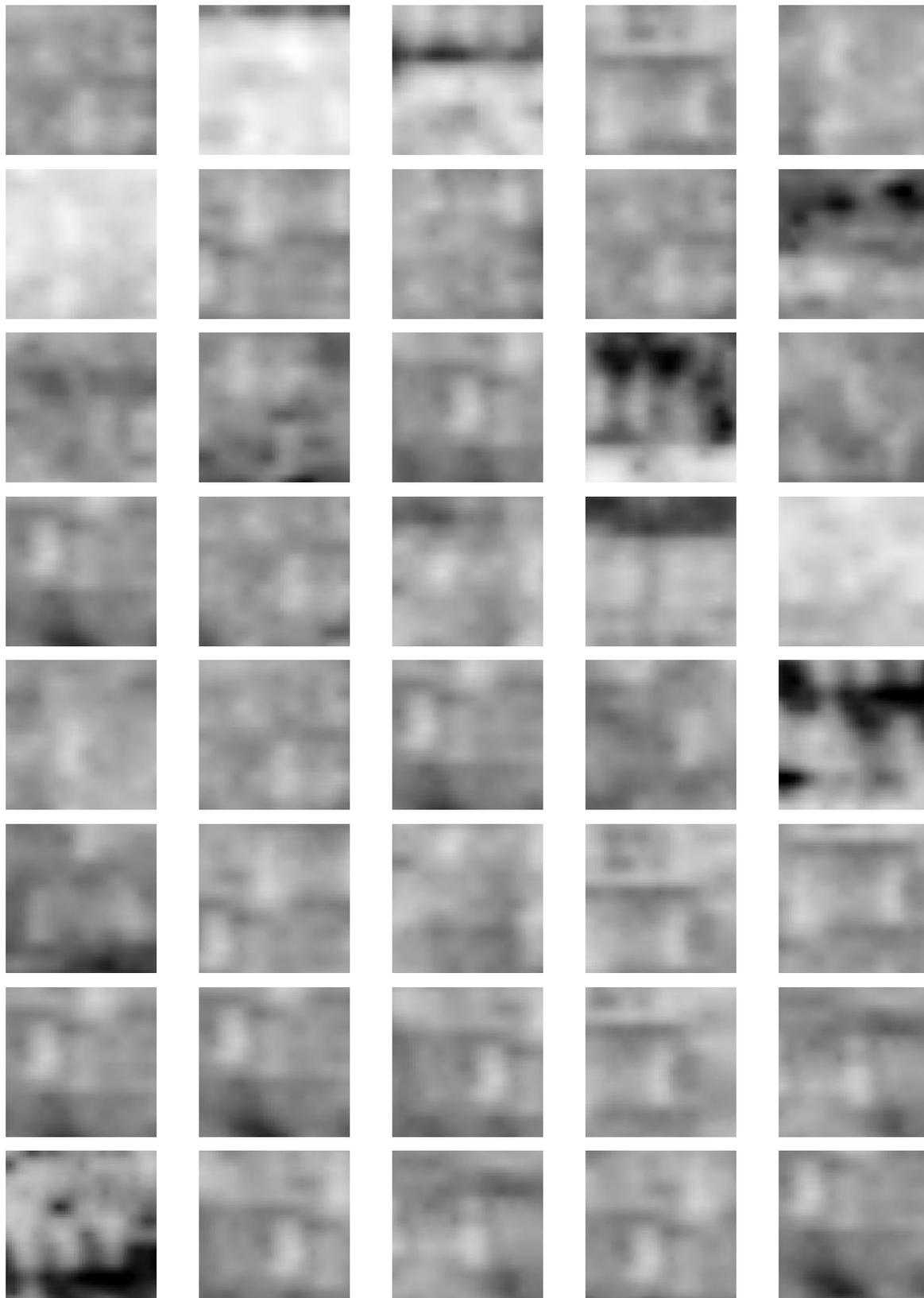nal techniques not thoroughly discussed herein (e.g. electrical resistivity tomography) may still be suitability in these areas and the suitability maps instruct users to seek further specialist advice for non-terrestrial surveys.

Table 31: Rank values for bedrock geologies used in the suitability analysis

Bedrock Geology	Magnetometry Suitability	EMI Suitability	Earth Resistance Suitability	GPR Suitability
Anorthosite	1	2	5	2
Appinite	4	3	4	3
Basalt	1	2	5	2
Breccia and Metabreccia	3	4	3	4
Breccia, Conglomerate and Sandstone	3	4	3	4
Chalk	5	4	5	5
Chalk and Sandstone	3	4	5	5
Clay	4	3	4	2
Clay and Lignite	3	4	4	3
Clay, Silt and Sand	3	3	4	2
Clay, Silt, Sand and Gravel	3	3	3	2
Conglomerate and [Subequal/Subordinate] Sandstone, Interbedded	2	4	4	4
Conglomerate, Sandstone, Siltstone and Mudstone	4	3	2	3
Diamictite	3	4	3	4
Diorite	1	2	5	2

Dolerite	1	3	5	3
Dolerite and Tholeiitic Basalt	1	3	5	2
Dolomitised Limestone and Dolomite	3	4	4	4
Dolostone	3	4	4	4
Felsic Lava	1	2	5	2
Felsic Lava and Felsic Tuff	1	3	5	2
Felsic Tuff	2	3	5	2
Felsic-Rock	1	2	5	2
Gabbro	1	3	5	3
Gneiss	2	4	5	4
Gneiss and Granite	1	2	5	2
Gneissose Psammite and Gneissose Semipelite	2	4	5	4
Gneissose Semipelite and Gneissose Psammite	2	4	5	4
Granite	1	2	5	2
Graphitic Pelite, Calcareous Pelite, Calcsilicate-Rock and Psammite	2	4	5	4
Gravel, Sand, Silt and Clay	2	3	3	3
Greywacke	2	2	3	2
Hornblende Schist	4	3	4	3
Lava and Tuff	1	3	5	2
Lava, Tuff, Volcaniclastic Rock and Sedimentary Rock	1	3	5	2
Limestone	3	5	5	5
Limestone and Calcareous Sandstone	3	5	5	5
Limestone and Mudstone, Interbedded	3	5	5	5
Limestone with Subordinate Sandstone and Argillaceous Rocks	3	5	5	5
Limestone, Argillaceous Rocks and Subordinate Sandstone, Interbedded	3	5	5	5
Limestone, Mudstone and Calcareous Mudstone	3	4	4	4

Limestone, Mudstone, Sandstone and Siltstone, with Subordinate Chert, Coal and Conglomerate	3	5	5	5
Limestone, Sandstone, Siltstone and Mudstone	3	5	5	5
Mafic Gneiss	2	3	5	3
Mafic Igneous-Rock	1	3	5	2
Mafic Lava	1	3	5	2
Mafic Lava and Mafic Tuff	1	3	5	2
Mafic Tuff	1	3	5	2
Mafite	1	3	5	2
Metalimestone	3	5	5	5
Metasedimentary Rock	3	4	4	4
Metavolcaniclastic Igneous-Rock and Metavolcaniclastic Sedimentary-Rock	1	3	5	2
Mica Schist	3	3	2	3
Migmatitic Rock	2	4	5	4
Mudstone	4	3	3	3
Mudstone, Chert and Smectite-Claystone	4	3	2	3
Mudstone, Sandstone and Conglomerate	3	3	2	3
Mudstone, Sandstone and Limestone	4	3	2	3
Mudstone, Siltstone and Sandstone	4	3	3	3
Mudstone, Siltstone, Limestone and Sandstone	3	3	2	3
Mudstone, Siltstone, Sandstone, Coal, Ironstone and Ferricrete	3	3	2	3
Mylonitic-Rock and Fault-Breccia	3	4	3	4
Pelite	3	2	3	2
Psammite	2	4	5	4
Psammite and Pelite	2	4	5	4
Psammite and Semipelite	2	4	5	4

Psammite, Pelite, Semipelite and Calcsilicate-Rock	2	4	5	4
Psammite, Semipelite and Pelite	2	4	5	4
Pyroclastic-Rock	1	3	5	2
Quartz-Arenite	2	4	5	4
Quartzite	3	4	5	4
Rhyolite	3	3	5	3
Sand, Silt and Clay	3	3	4	2
Sandstone	1	5	3	4
Sandstone and [Subequal/Subordinate] Argillaceous Rocks, Interbedded	3	4	4	4
Sandstone and [Subequal/Subordinate] Limestone, Interbedded	2	5	4	5
Sandstone and Conglomerate, Interbedded	2	5	3	5
Sandstone and Mudstone	3	4	4	4
Sandstone and Siltstone, Interbedded	2	4	4	4
Sandstone and Subordinate Breccia	2	4	4	4
Sandstone with Subordinate Argillaceous Rocks and Limestone	2	5	3	5
Sandstone with Subordinate Conglomerate and Siltstone	3	4	4	4
Sandstone with Subordinate Conglomerate, Siltstone and Mudstone	3	4	4	4
Sandstone, Breccia and Conglomerate	2	4	4	4
Sandstone, Conglomerate and [Subordinate] Argillaceous Rocks	2	4	4	4
Sandstone, Limestone and Argillaceous Rocks	2	5	3	5
Sandstone, Mudstone, Siltstone and Conglomerate	3	3	2	3
Sandstone, Siltstone and Mudstone	2	4	4	4
Schist	4	3	4	3
Sedimentary Rock Cycles, Clackmannan Group Type	2	4	4	4
Sedimentary Rock Cycles, Strathclyde Group Type	3	4	4	4
Semipelite	2	4	5	4

Semipelite and Pelite	2	4	5	4
Serpentinite	2	4	5	4
Serpentinite, Metabasalt, Metalimestone and Psammite	1	3	5	2
Shale	5	5	5	5
Siltstone	3	4	4	4
Siltstone and Sandstone with Subordinate Mudstone	3	4	4	3
Slate	5	5	5	5
Syenitic-Rock	1	2	5	2
Ultramafitite	1	3	5	2
Volcanic	1	2	5	2
Wacke	2	2	3	2

Table 32: Rank values for superficial/drift geologies used in the suitability analysis

Superficial/Drift Geology	Magnetometry Suitability	EMI Suitability	Earth Resistance Suitability	GPR Suitability
Airfield/Airport	1	4	3	4
Alluvium	3	2	4	3
Alluvium (Clayey)	3	2	4	3
Alluvium (Gravelly)	3	2	4	3
Alluvium (Sandy)	3	2	4	3
Alluvium (Silty)	3	2	4	3
Bedrock Outcrop or Subcrop	0	0	0	0
Blanket Peat	2	4	3	4
Blown Sand	2	4	3	2

Brickearth	3	4	3	4
Causeway	0	0	0	0
Clay with Flints	5	3	3	2
Crag Group	2	4	3	4
Crannog	0	0	0	0
Cut Over Raised Peat	2	4	3	4
Dam	0	0	0	0
Drift Geology Not Mapped	0	0	0	0
Embankment	2	4	3	4
Eskers Comprised of Gravels of Acidic Reaction	2	3	3	3
Eskers Comprised of Gravels of Basic Reaction	2	4	3	4
Estuarine Silts and Clays	2	4	3	3
Fen Peat	2	4	3	4
Glacial Sand and Gravel	2	4	3	4
Glaciomarine Sediments	2	4	3	4
Gravels Derived from Basic Igneous Rocks	2	4	3	4
Gravels Derived from Cambrian Sandstones and Shales	2	4	3	4
Gravels Derived from Carboniferous Sandstones and Shales	2	4	3	4
Gravels Derived from Chert	2	4	3	4
Gravels Derived from Devonian and Carboniferous Sandstones	2	4	3	4
Gravels Derived from Devonian Sandstones	2	4	3	4
Gravels Derived from Granite	2	4	3	4
Gravels Derived from Limestones	2	4	3	4
Gravels Derived from Lower Palaeozoic and Devonian Sandstones	2	4	3	4
Gravels Derived from Lower Palaeozoic Sandstones	2	4	3	4
Gravels Derived from Lower Palaeozoic Sandstones and Shales	2	4	3	4

Gravels Derived from Lower Palaeozoic Sandstones and Shales	2	4	3	4
Gravels Derived from Lower Palaeozoic Shales	2	4	3	4
Gravels Derived from Metamorphic Rocks	2	4	3	4
Gravels Derived from Namurian Sandstones and Shales	2	4	3	4
Gravels Derived from Quartzite	2	4	3	4
Gravels Derived from Silicified Limestones and Chert	2	4	3	4
Hardstand	0	1	0	3
Industrial	2	4	3	4
Irish Sea Till Derived from Acidic Volcanic Rocks	1	2	5	2
Irish Sea Till Derived from Basic Igneous Rocks	5	3	3	2
Irish Sea Till Derived from Cambrian Sandstones and Shales	5	3	3	2
Irish Sea Till Derived from Devonian Sandstones	5	3	3	2
Irish Sea Till Derived from Limestones	5	3	3	2
Irish Sea Till Derived from Lower Palaeozoic Sandstones and Shales	5	3	3	2
Kartsified Bedrock Outcrop or Subcrop	0	0	0	0
Lacustrine Clays	2	4	3	3
Lacustrine Deposits (Undifferentiated)	2	4	3	4
Lacustrine Gravel	2	4	3	4
Lacustrine Sands	2	4	3	3
Lacustrine Sediments	2	4	3	4
Lacustrine Silts	2	4	3	3
Lake Marl	2	4	3	4
Landfill	1	4	2	4
Landslip	2	4	3	4
Made Ground	2	4	3	4
Marine Beach Sands	2	4	3	1

Marine Gravel and Sands (Often Raised)	2	4	3	4
Peat	2	4	3	4
Pier	0	0	0	0
Raised Marine Deposits (Undifferentiated)	2	4	3	4
Raised Peat (Intact)	2	4	3	4
River Terrace Deposits (Undifferentiated)	2	4	3	4
Sand and Gravel of Uncertain origin	2	3	3	3
Scree	2	4	3	4
Spoil Heap	2	4	3	4
Tailings Pond	0	0	0	0
Tidal Marsh	0	2	0	1
Till	5	3	3	2
Till Derived from Acidic Volcanic Rocks	1	2	5	2
Till Derived from Basic Igneous Rocks	5	3	3	2
Till Derived from Cambrian Sandstones and Shales	5	3	3	2
Till Derived from Carboniferous Sandstones	5	3	3	2
Till Derived from Carboniferous Sandstones and Cherts	5	3	3	2
Till Derived from Carboniferous Sandstones and Shales	5	3	3	2
Till Derived from Cherts	5	3	3	2
Till Derived from Devonian and Carboniferous Sandstones	5	3	3	2
Till Derived from Devonian and Carboniferous Sandstones and Shales	5	3	3	2
Till Derived from Devonian Sandstones	5	3	3	2
Till Derived from Granites	5	3	3	2
Till Derived from Limestones	5	4	4	3
Till Derived from Lower Carboniferous Sandstones and Shales	5	3	3	2

Till Derived from Lower Palaeozoic and Carboniferous Sandstones and Shales	5	3	3	2
Till Derived from Lower Palaeozoic and Devonian Sandstones	5	3	3	2
Till Derived from Lower Palaeozoic Sandstones	5	3	3	2
Till Derived from Lower Palaeozoic Sandstones and Shales	5	3	3	2
Till Derived from Lower Palaeozoic Shales	5	3	3	2
Till Derived from Metamorphic Rocks	5	3	3	2
Till Derived from Namurian and Carboniferous Sandstones and Shales	5	3	3	2
Till Derived from Namurian Sandstones and Shales	5	3	3	2
Till Derived from Quartzites	5	3	3	2
Till Derived from Silicified Limestone and Cherts	5	3	3	2
Urban	2	3	3	4
Water	0	0	0	0
Windblown Sands	2	4	3	2
Windblown Sands and Dunes	2	4	3	2

Table 33: Rank values for soils used in the suitability analysis

Soil	Magnetometry Suitability	EMI Suitability	Earth Resistance Suitability	GPR Suitability
Tidal Marsh	0	2	0	1
Rock	1	3	1	1
Island	0	1	0	0
Loamy	3	3	3	3
Alluvium	2	3	4	2

Peat	1	4	3	4
Water Body	0	0	0	0
Clay	3	3	3	3
Urban	1	2	1	3
Salt Marsh	0	2	0	1
Sandy	2	3	2	2

Table 34: Rank values for landcovers used in the suitability analysis

Landcover	Magnetometry Suitability	EMI Suitability	Earth Resistance Suitability	GPR Suitability
Arable	5	5	5	4
Bog	1	4	1	3
Burnt Areas	1	2	3	3
Coastal	1	3	1	2
Cultivated	5	5	3	3
Grassland	5	5	5	4
Heather	5	5	4	3
Marsh	0	3	0	2
Mineral Extraction	1	2	3	4
Moors and Heathland	5	5	3	4
Outcrop	2	2	0	1
Pasture	5	5	5	5
Rock	2	2	0	0
Sediment	5	5	5	5
Transportation	1	3	2	4

	1	2	3	4
Urban	1	2	3	4
Water	0	1	0	1
Wetland	0	3	0	2
Woodland	3	3	3	2