Intelligence-Based Medicine 7 (2023) 100091

Contents lists available at ScienceDirect

Intelligence-Based Medicine

journal homepage: www.sciencedirect.com/journal/intelligence-based-medicine

Prediction of Alzheimer's disease from magnetic resonance imaging using a convolutional neural network

Kevin de Silva, Holger Kunz *

Institute of Health Informatics, University College London, UK

A B S T R A C T

Objectives: The primary goal of this study is to examine if a convolutional neural network (CNN) can be applied as a diagnostic tool for predicting Alzheimer's Disease (AD) from magnetic resonance imaging (MRI) using the MIRIAD-dataset (Minimal Interval Resonance Imaging in Alzheimer's Disease) from one single central slice of the brain.

Methods: The MIRIAD dataset contains patients' health records represented by a set of MRI scans of the brain and further diagnostic data. Hyperparameters and configurations of CNNs were optimized to determine the best-performing model. The CNN was implemented in Python with the deep learning library 'Keras' using Linux/Ubuntu as the operating system.

Results: This study obtained the following best performance metrics for predicting Alzheimer's Disease from MRI with Matthew's Correlation Coefficient (MCC) of 0.77; accuracy of 0.89; F1-score of 0.89; AUC of 0.92. The computational time for the training of a CNN takes less than 30 sec. s with a GPU (graphics processing unit). The prediction takes less than 1 sec. on a standard PC.

Conclusions: The study suggests that an axial MRI scan can be used to diagnose if a patient has Alzheimer's Disease with an AUC score of 0.92.

1. Introduction

1.1. Alzheimer's diseases

Alzheimer's Disease (AD) is associated with the progressive accumulation of abnormal proteins in the brain, which leads to progressive synaptic, neuronal, and axonal damage [1]. ICD-11 (eleventh revision of the International Classification of Diseases) from the WHO codifies Alzheimer's with 6D80* and 8A20* as a disorder with neurocognitive impairment as a major feature [2]. Clinical symptoms include loss of memory, linguistic and cognitive degradation, and personality and mood changes [3]. The number of people with AD worldwide is estimated at 50 million in 2017, growing to 132 million by 2050, while the total cost associated with AD worldwide as of 2018 is estimated at 1 trillion dollars [3]. Although these costs and prevalence numbers appear high, they may represent a substantial underestimate of the true figures since undiagnosed AD can be as high as 80% of all cases worldwide [3].

Although currently, no drugs can cure AD early diagnosis and treatment of AD has substantial benefits, both in terms of personal wellbeing and societal cost [3]. A class of drugs, cholinesterase inhibitors, are effective at slowing down the progression of AD [3]. Given the advantages of early-stage diagnosis of Alzheimer's disease, any methodology that improves early detection is beneficial. There is no specific biomarker for AD and diagnosis relies on a range of tests which include one or more of the following: cognitive assessment tests, blood tests, Computerized Tomography (CT), Magnetic Resonance Imaging (MRI), Single Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET) [4]. Of relevance to this study are MRI and the cognitive assessment test – Mini-Mental State Examination (MMSE). MRI scans show the atrophy of certain brain regions that are indicative of Alzheimer's [5,6]. MMSE is a quick, inexpensive test, scoring from 0 to 30, where higher scores are indicative of better cognitive functioning [7].

1.2. Convolutional neural networks to detect Alzheimer

Studies have been using CNNs (Convolutional Neural Networks) to diagnose Alzheimer's [8–26] using data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) or Open Access Series of Imaging Studies (OASIS). The ADNI-dataset has 1455 participants with five diagnosis groups [27]. The OASIS-dataset is composed of 193 participants aged 62 years or more [27]. The primary goal of this study is to examine if convolutional neural networks can also be applied as a diagnostic tool using the MIRIAD-dataset (Minimal Interval Resonance Imaging in Alzheimer's Disease).

* Corresponding author.
E-mail address: h.kunz@ucl.ac.uk (H. Kunz).

https://doi.org/10.1016/j.ibmed.2023.100091
Received 29 April 2022; Received in revised form 8 November 2022; Accepted 4 January 2023
Available online 8 January 2023

2. Methods

2.1. Data and material

MIRIAD (Minimal Interval Resonance Imaging in Alzheimer's Disease) is a series of longitudinal volumetric T1-MRI scans of mild-moderate Alzheimer's subjects and controls [28]. An overview of the MIRIAD demographics and publications is published in Malone et [28]. The dataset consists of scans with the same scanner with accompanying information on gender, age, and Mini-Mental State Examination (MMSE) scores [28]. The data used in this study contains two class labels: Alzheimer's Disease (AD) if a patient has an MMSE score of 26 or under at baseline and healthy control (HC) with an MMSE of 27 or above [28]. This is also the cut-off point to describe the class label for each feature vector. Each patient has multiple MRI scans from different time points. Many scans were collected of each participant at intervals from two weeks to two years, the study was designed to investigate the feasibility of using MRI as an outcome measure for clinical trials of Alzheimer's treatments [28]. Table 1 shows the demographics of the included patients.

Each scan is provided in *NIfTI*-format (Neuroimaging Informatics Technology Initiative) [29]. It is an open file format for volumetric images with a size of 256 x 256 x 124.

Fig. 1 shows a sample of the MRI dataset. An axial, sagittal, and coronal view is displayed. The raw dataset still contains bone structures. The bone structures are not relevant for the diagnosis of Alzheimer's and were removed in the pre-processing steps.

2.2. Feature engineering and pre-processing

Pre-processing is an important step to prepare the dataset for the following training of the classification algorithm. The MIRIAD dataset is pre-processed by applying spatial normalization, bias correction, and grey matter segmentation. Spatial normalization is the process of mapping images from different scans onto a single template. There are two steps to this: linear transformation (e.g. translation, rotation, shear) and non-linear transformation (e.g. warping). This results in all images referencing the same coordinate space [30] and should adjust, for example, for different subject positioning when the MRI was recorded.

The ratio of MRI scans of AD subjects to healthy controls is approximately 2:1. To mitigate this imbalance, data augmentation is performed by creating copies and flipping them. This results in almost the same number of instances labelled for AD and non-AD subjects. This can also be considered a specific type of oversampling in medical imaging.

Finally, grey matter segmentation is performed, and grey matter is extracted from the raw data. This excludes features that are unlikely to be discriminative in the classification task e.g., skull bones. The Python 'Nipype' library interface is used, allowing all processing to be done in Python [31]. An axial MRI scan of the central part of the brain for each patient was used as an input for the following classification algorithm.

2.3. Convolutional neural network

Convolutional Neural networks (CNNs) are a specialized kind of neural network for processing data that has a grid-like topology [32]. A CNN consists of several layers: convolutional, pooling, and fully connected layers. Each convolutional layer consists of a certain number of trainable parametric filters. Each convolutional layer is typically followed by a pooling layer which reduces the feature space. Finally, the data is passed to one or more fully connected layers and the predicted output is produced. A further description of the basic ingredients of a convolutional neural network can be derived from a textbook in deep learning [32] and are not further explained.

The applied CNN to distinguish between Alzheimer's and non-Alzheimer patients is used as a classification algorithm. Classification is to learn a mapping from inputs x to output y, where $y \in \{1,.., C\}$ with C being the number of classes [33]. If $C = 2$, this is called binary classification [33]. In our study, a binary classification task is performed to distinguish between patients with Alzheimer's and patients who do not show signs of Alzheimer's. The current configuration for a convolutional neural network is a bespoke solution. Pre-trained models from Keras (ResNet-50 and Inception_V3) were tested as a baseline method. The baseline models (ResNet-50 and Inception_V3) achieved low accuracies of 0.72 and 0.55 on a test dataset and were disregarded for further analysis. We converted our greyscale images to RGB images with three channels and rescaled them for the pre-trained models. The number of layers and convolutional filters per layer were varied. The

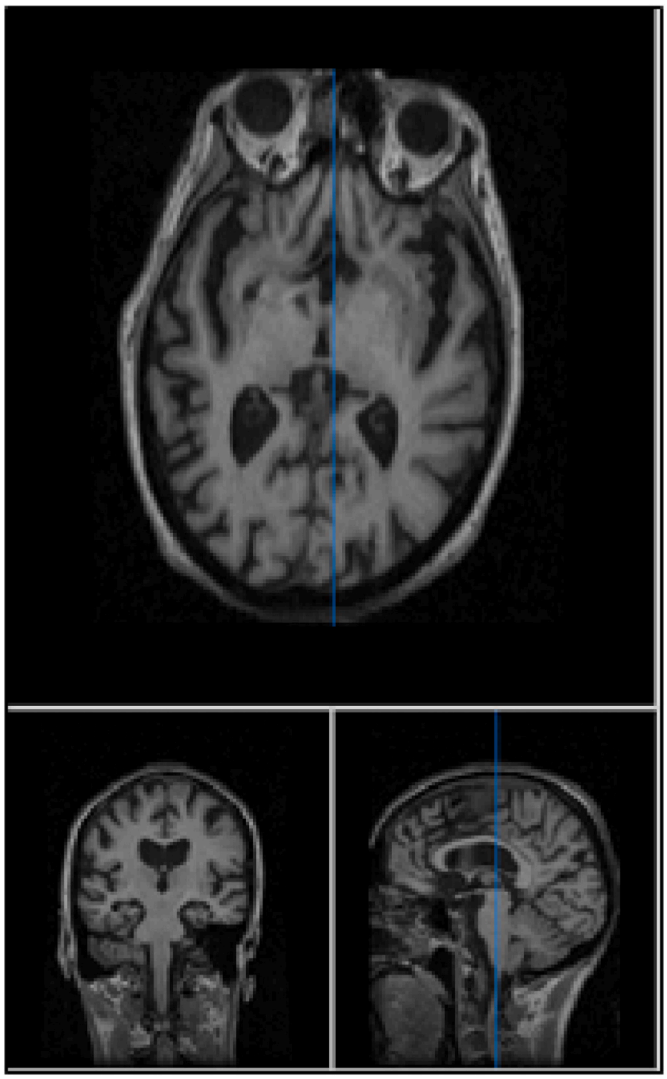

Fig. 1. Sample of raw MRI-data – axial, sagittal, and coronal views of the MIRIAD-dataset.

Table 1
MIRIAD demographic information.

	Alzheimer's Disease (n = 46 (number of patients), Total MRI-scans = 465)	Healthy Controls (n = 23 (number of patients), Total MRI-scans: 243)
Age at study entry	69.4 ± 7.1	69.7 ± 7.2
Men	41%	52%
Mean (SD) baseline MMSE	19.2 ± 4	29.4 ± 0.8

hyperparameter tuning used GridSearch with an exhaustive search over specified parameter values. For example, a 3-layer or 4-layer setting was defined and for each layer a range of different input size of 32, 64, 128, or 256. The following sections describe parameters for loss the loss function, max-pooling, dropout layer, activation function, and regularization.

2.3.1. Loss function

As a loss function for the convolutional neural network, the binary cross-entropy was chosen [34,35]. Every training epoch of the CNN has the aim to reduce the loss function (binary cross-entropy). RMSprop is a gradient-based optimization technique used in training neural networks. It has also been applied in deep learning for MR images by Medina et al. [36].

2.3.2. Convolutional filter size and max-pooling

For a two-dimensional image I as our input (from an MRI scan), a two-dimensional kernel K can be used. In this study, the convolutional filter size was set to (3,3). In convolutional network terminology, the output is referred to as a feature map [32]. The convolutional operation can be described as follows [32]:

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(m,n)K(i-m, j-n) \tag{1}$$

A pooling function replaces the output with a summary statistic. For example, the max-pooling operation reports the maximum output within an area [32]. Different layers of the CNN, specifically pooling, include the two following important aspects, namely that: (i) they introduce nonlinearity into the neural network to achieve improved function approximation ability, and (ii) they help CNNs to achieve spatial invariance, capable of recognizing features in an image regardless of location. The Max-pooling filter size of the final configuration after hyperparameter tuning was set to (2,2).

2.3.3. Dropout layer

Dropout provides a computationally inexpensive method for regularizing a model and preventing overfitting [32,37]. During training, units get randomly get removed [37]. The randomly selected unit is removed from the network, along with all its incoming and outgoing connections [37]. It prevents overfitting and provides a way of approximately combining exponentially many different neural network architectures efficiently [37]. Dropout introduces an extra hyperparameter—the probability of retaining a unit [37]. A value of $p = 1$ implies no dropout, and low values of p mean more dropout [37]. The dropout rate was set to 0.4 in our configuration to avoid overfitting.

2.3.4. Activation function

Neurons in the activation map pass through a non-linear function [38]. There are different activation functions. For example, the sigmoid function, the rectified linear unit (ReLU), and the leaky rectified linear unit (leaky ReLU). The logistic sigmoid function can be defined as the following [12]:

$$f_{sigmoid}(x) = \frac{1}{1+\exp(-x)} \tag{2}$$

Another activation-function is the ReLu-function [12]:

$$f_{ReLu}(x) = \max(0,x) = \begin{cases} 0, x < 0 \\ x, x \geq 0 \end{cases} \tag{3}$$

whenever the activation values are zero, the ReLu-function cannot learn in a gradient-based learning method [12]. Therefore, a leaky ReLu-function can be used.

$$f_{Leaky\ Relu}(x) = \begin{cases} x, x \geq 0 \\ \alpha x, x < 0 \end{cases} \tag{4}$$

In our study, the parameter alpha was set to 0.1 and a leaky rectified linear unit was used.

2.3.5. Regularization

To prevent overfitting a regularization method can be used to train the neural network. L1-Regularization is also known as Lasso-Regularization [39]. L2-Regularization, also known as Ridge Regularization [39]. L1+L2 Regularization is also known as Elastic Net Regularization [39]. A small value for the regularization parameters for $L1 = 0.001$ and $L2 = 0.002$ has been added to prevent overfitting.

Table 2 contains the settings of the final CNN model.

2.4. Performance metrics

The evaluation of model performance is an essential step in understanding and developing a machine learning algorithm. Definitions of conventional performance metrics such as accuracy, precision, specificity, recall, and F1-score are not further described. The definition can be obtained from textbooks in machine learning such as Goodfellow et al. [32], Murphy [33], and Hastie et al. [39]. This study used as an additional metric Matthew's Correlation Coefficient (MCC) [40]. The following abbreviations have been used TP = True Positives, TN = True Negatives, FP=False Positive, FN=False Negative.

The MCC is defined as [40]:

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{5}$$

The MCC metric is more balanced than metrics like accuracy and F1-score because its score is high only if the classifier is good on both positive and negative predictions [41]. The MCC is calibrated so that it ranges from −1 to +1. A value of 0 indicates a result close to chance, the closer to +1 the score is, the better the result [41]. Receiver Operating Characteristic (ROC) curves have also been plotted for the best outcome. The data are split into training, validation, and test-dataset. Table 3 shows the exact split of the data. The best configuration of the CNN was determined with the highest MCC on unseen medical images of a set of AD and non-AD patients.

2.5. Implementation

The used hardware: Intel Core i9 with 64 GB of memory, hard disk: Samsung 1 Terabyte SSD, GPU: NVIDIA GE Force RTX 2080 TI. The implementation used as software: 'Statistical Parametric software', Matlab, Python, Keras, and as operating system Linux/Ubuntu.

Table 2
Configuration of the CNN.

Setting/Parameter	Values in Keras
Loss Function	Binary Cross-Entropy
Optimiser Function	RMSprop(lr = 0.001)
Convolutional filter (kernel) size	(3, 3)
Padding for all convolutional layers	"Same" (results in output feature map being the same size as input)
Padding for max-pooling layers	"Same" (results in output feature map being half the size of the input)
Max-pooling filter size	(2, 2)
Activation function for all layers	Leaky ReLU (alpha = 0.1)
Weight regularization added to all models to mitigate overfitting	L1 = 0.001,L2 = 0.002
Dropout layer added to all models to mitigate overfitting	0.4
Batch size	100
Number of epochs	20

Table 3
Data augmentation, training, validation, test.

Class Label	Number of MRI-scans	Total slices after data augmentation	Slices in training split	Slices in validation split	Slices in test split
AD	465	465	326	39	100
non-AD	243	486	342	42	102

3. Results

3.1. Pre-processed medical images

Fig. 2 shows the axial, sagittal, and coronal views in 2D as well as a 3D surface rendering of a typical pre-processed scan. Extraneous features (for example the skull and bone structures) have been removed.

3.2. Results from the optimized configuration of the CNN

Table 4 shows for the best model with optimized MCC the associated confusion metric (TN = 91, FN = 12, FP = 11, TP = 88).

The obtained best performance metric for MCC = 0.77.

$$MCC = \frac{88 \times 91 - 11 \times 12}{\sqrt{(88+11)(88+12)(91+11)(91+12)}} = 0.77 \quad (6)$$

Accuracy = 0.89=(88 + 91)/(91 + 12+11 + 88); Precision = 0.89 = 88/(11 + 88); Specificity = 0.89 = 91/(91 + 12); Recall = 0.88 = 88/(12 + 88); F1 = 0.88 = 88/(88 + 0.5(11 + 12)).

Table 4
Performance for the optimized CNN-model.

Confusion matrix			
0: negative diagnosis Alzheimer			
1: positive diagnosis Alzheimer			
true label		0	1
	0	91	11
	1	12	88
	predicted label		

The final architecture for the model was guided by the best MCC-score of 0.77 in 3-layer CNN with 64 convolutional filters in each layer, represented as (64, 64, 64) and an associated AUC of 0.92. Fig. 3 shows AUC scores for different configurations.

The computational time for the training using a CNN takes less than 30 sec.s with a GPU (graphics processing unit) using NVIDIA GE Force RTX 2080 TI. The prediction or classification using the trained algorithm takes less than 1 sec. on a standard PC.

4. Discussion and conclusion

This study shows that convolutional neural networks for pattern recognition of neurological conditions such as Alzheimer's can be used. A CNN was proposed to distinguish between patients having

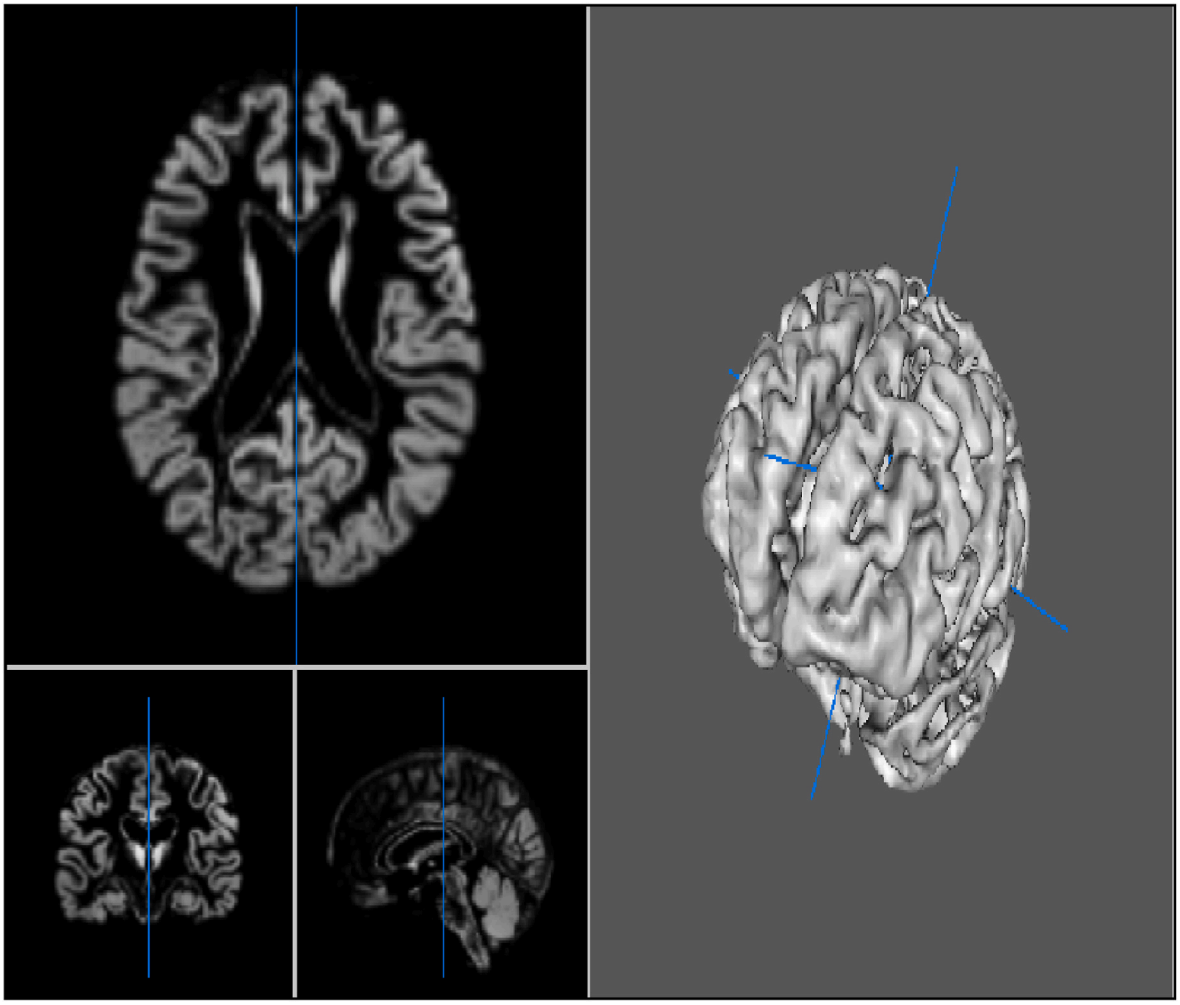

Fig. 2. Pre-processed data – axial, sagittal, coronal, and 3D surface view.

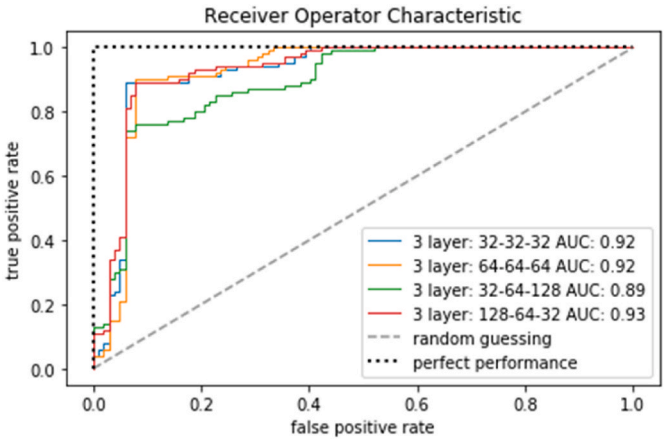

Fig. 3. ROC-plot for different configurations of the CNN.

Alzheimer's Disease (AD) and patients who have not been diagnosed with AD. Medical images of the brain have been used as input for CNN. The CNN and the number of layers and convolutional filters per layer were varied and optimized based on Matthew's Correlation Coefficient (MCC). This study obtained the following performance metrics for predicting Alzheimer's Disease from MRI scans of the brain; MCC: 0.77, Accuracy: 0.89, F1: 0.89, AUC: 0.92. An AUC>0.90 is seen as an excellent diagnostic test [42]. A potential interpretation of AUC values is 1.0 for a perfect test, 0.9–0.99 for an excellent test, 0.8–0.89 for a good test, 0.7–0.79 for a fair test, 0.51–0.69 for a poor test, and 0.5 of no value [42]. This suggests that potentially a diagnostic tool could be developed based on the provided methodology. A comparison with the baseline models (ResNet-50 and Inception_V3) using transfer learning revealed a low performance. The pre-trained models in Keras have been trained on natural images from ImageNet. The modality of the images is different, and this could affect the effectiveness of pre-trained models and transfer learning in Keras [43].

Minimizing the risk of overfitting was done with three techniques: Firstly, the dataset was randomly split into a training set, validation set, and test set. The general assumption is that the instances are randomly selected. This ensures that the trained CNN will be tested on previously unseen instances. Secondly, regularization parameters were set to non-zero values. This ensures that the loss function considers regularization. This is a common method in machine learning to avoid overfitting. Thirdly, a specified dropout rate is a safeguard that the neural network at the dense layer (fully connected deep neural network) does not overlearn presented instances. Random units of the dense layer get removed [37]. It can empirically be shown that a neural network that makes use of a dropout rate reduced the risk of overfitting [37]. These three methods 1) random split, 2) regularization, and 3) dropout helped to minimize the potential risk of overfitting and generally avoid that a complex function is perfectly fitted to the provided training set.

The underlying problem is a binary classification problem to diagnose Alzheimer vs. non-Alzheimer. The MCC-score is a more reliable statistical rate that produces a high score only if the prediction obtained good results in all four confusion matrix categories (true positives, false negatives, true negatives, and false positives) [40]. For this reason, the optimization of our study used the MCC score. The performance metrics such as F1-score or AUC are a by-product of this process. The MCC score is useful for imbalanced datasets where the AUC might be less useful [41]. This imbalance was reduced by using methods of augmentation which can be considered as a type of oversampling in machine learning for medical images. The confusion matrix shows that overall few instances have been misclassified as false negatives or false positives.

Analysis of the literature identifies the usage of convolutional neural networks for AD-classification from MRI [16]. A direct comparison of the performance is limited as the studies used different datasets such as the ADNI or OASIS-dataset for Alzheimer's disease. For example, different demographics, origins, and sample sizes have an impact on the performance metric. Additionally, the hyperparameters were tuned on different performance metrics and the CNNs used different configurations regarding kernels and the number of layers. Another factor is different cross-validation methods across the studies. Since random seed values are used to split data in different folds or to split into training sets and test set a meaningful comparison is limited, too. The ADNI-dataset was used by Aderghal et al. [8], Taqi [21], Cheng [13], Lian et [19], Farooq et al. [14], Senanayake [20], Gunawardena [15], Hosseini et al. [16], Korolev [18], Bäckström et al. [9], Folego et al. [26], Feng et al. [25], Wu [23]. The OASIS dataset has been used by Islam et al. [17], Wang et al. [12], Hon et al. [44], Ebrahimi and Luo [45]. The models in Hosseini et al. [16] and Bäckström et al. [9] used accuracy as a performance metric for optimization. Accuracy, sensitivity, specificity have been used in Farooq et al. [14] and Wang et al. [12]. Also, a multi-class classification process was used in Islam et al. [17] as opposed to binary classification in this study.

A key question towards a potential clinical deployment is the reliability of such as clinical decision support system. Even though the AUC is above 0.90 it is crucial to indicate that the performance metrics must be seen in the context of the specific clinical application. The therapeutic consequences for false positive and false negative subjects must be carefully considered. As a rule of thumb, diagnostic tools having an AUC>0.9 could potentially be candidates for a clinical decision support system. However, a decision for a potential deployment cannot be based on a fixed threshold but also need an extended qualitative study that considers the expert opinions of clinicians to provide further insights whether such a decision support system is fit for purpose. The practical advantage of the developed CNN lies in the fact that one axial scan provides sufficiently enough information to achieve high performance. One immediate potential usage of a deployed system could be in a low-resource setting or where clinical consultants are not readily available.

A recent study confirms that Alzheimer's disease can be diagnosed from a single brain scan [46]. The model was trained on T1-weighted MRI scans obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) [46]. The applied method uses a two-stage process including least absolute shrinkage and selection operator (LASSO) [46]. The study achieved an accuracy of 0.92 to detect AD [46]. However, the results are not directly comparable due to different datasets, performance scores, different objectives, and settings. However, the study further proves that a single slice from MRI can be used to diagnose Alzheimer's disease. The focus of this study is directed towards the use case of predicting Alzheimer's from a single MRI slide using a CNN.

Ethical approval and consent to participate

Ethical approval for the data (and subsequently its release) was received from the local MIRIAD research ethics committee, and written consent was obtained from all participants (see Malone et al. [28]).

Consent for publication: Consent for publication using the dataset is provided by Malone et al. (see Ref. [28]).

Availability of data and materials

The MIRIAD (Minimal Interval Resonance Imaging in Alzheimer's Disease) dataset is publicly available. Data are here made publicly available as a common resource for researchers to develop, validate and compare techniques, particularly for measurement of longitudinal volume change in serially acquired MR (see Ref. [28]). By registering and agreeing to the data use agreement the data can be downloaded. Datasets are available in the MIRIAD database for research, which is accessible after registration from a public repository using the following URL: https://www.ucl.ac.uk/drc/research/research-methods/minimal-interval-resonance-imaging-alzheimers-disease-miriad.

Funding

The original data collection was funded through an unrestricted educational grant from GlaxoSmithKline (Grant 6GKC) and funding from the UK Alzheimer's Society (Grant RF116) and the Medical Research Council (to Professor Fox). The Dementia Research Centre is an Alzheimer's Research UK (ARUK) Coordinating Centre [28]. The Wellcome Trust Centre for Neuroimaging is supported by core funding from the Wellcome Trust [grant number 091593/Z/10/Z] (see Ref. [28]). The implementation of the algorithms of this work makes use of the provided MIRIAD-data and has been conducted as a postgraduate dissertation without dedicated funding.

Authors' contribution

The first author (KdS) conducted the implementation of the algorithms and prepared the draft of the manuscript. The co-author (HK) provided academic guidance, revised the manuscript, and verified the scientific robustness of the applied methods. The authors confirm that all methods were carried out in accordance with relevant guidelines and regulations. MIRIAD investigators did not participate in analysis or writing of this report.

Declaration of competing interest

No competing interests declared.

Acknowledgement

The authors would like to acknowledge the contribution of Dr Ian Malone from the Dementia Research Centre (DRC) at UCL for the provision of the MIRIAD dataset.

References

[1] Ramani A, Jensen JH, Helpern JA. Quantitative MR imaging in Alzheimer disease. Radiology; 2006. https://doi.org/10.1148/radiol.2411050628.
[2] WHO. ICD-11 eleventh revision of the international classification of diseases. 2021. 2021.
[3] ADI. Alzheimer's disease international. 2020, https://www.alz.co.uk/; 2020.
[4] NHS. Overview - Alzheimer's disease. 2020, https://www.nhs.uk/conditions/alzheimers-disease/; 2020.
[5] Frisoni GB, Fox NC, Jack Jr CR, Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol 2010;6:67–77. https://doi.org/10.1038/nrneurol.2009.215.
[6] Waldemar G, Dubois B, Emre M, Georges J, McKeith IG, Rossor M, Scheltens P, Tariska P, Winblad B, Efns. Recommendations for the diagnosis and management of Alzheimer's disease and other disorders associated with dementia: EFNS guideline. Eur J Neurol 2007;14:e1–26. https://doi.org/10.1111/j.1468-1331.2006.01605.x.
[7] Folstein MF, Folstein SE, McHugh PR. Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189–98. https://doi.org/10.1016/0022-3956(75)90026-6.
[8] C. Aderghal, K., Benois-Pineau, J., Afdel, K., Gwenaëlle, FuseMe: Classification of sMRI images by fusion of Deep CNNs in 2D + projections, in: 15th International Workshop on Content-Based Multimedia Indexing, n.d. https://doi.org/10.1145/3095713.3095749.
[9] Backstrom K, Nazari M, Gu IYH, Jakola AS. An efficient 3D deep convolutional network for Alzheimer's disease diagnosis using MR images. In: Proceedings - international symposium on biomedical imaging. IEEE Computer Society; 2018. p. 149–53. https://doi.org/10.1109/ISBI.2018.8363543.
[10] Basaia S, Agosta F, Wagner L, Canu E, Magnani G, Santangelo R, Filippi M. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. Neuroimage Clin 2019;21:101645. https://doi.org/10.1016/j.nicl.2018.101645.
[11] Basaia S, Agosta F, Wagner L, Canu E, Magnani G, Santangelo R, Filippi M. Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. Neuroimage Clin 2019;21:101645. https://doi.org/10.1016/j.nicl.2018.101645.
[12] Wang SH, Phillips P, Sui Y, Liu B, Yang M, Cheng H. Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. J Med Syst 2018;42:85. https://doi.org/10.1007/s10916-018-0932-7.
[13] Cheng D, Liu M. CNNs based multi-modality classification for AD diagnosis. in: 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), IEEE; 2017. p. 1–5. https://doi.org/10.1109/CISP-BMEI.2017.8302281.
[14] A.F.S.M.A.M.A.S. Rehman. A deep CNN based multi-class classification of Alzheimer's disease using MRI. in: IEEE International Conference on Imaging Systems and Techniques (IST); 2017. https://doi.org/10.1109/IST.2017.8261460.
[15] Gunawardena KANNP, Rajapakse RN, Kodikara ND. Applying convolutional neural networks for pre-detection of alzheimer's disease from structural MRI data. in: 2017 24th International Conference on Mechatronics and Machine Vision in Practice (M2VIP), IEEE; 2017. p. 1–7. https://doi.org/10.1109/M2VIP.2017.8211486.
[16] Ehsan Hosseini Asl Ayman S El-Baz RSK. Alzheimer's disease diagnostics by adaptation of 3D convolutional network. IEEE; 2016. https://doi.org/10.1109/ICIP.2016.7532332.
[17] Islam J, Zhang Y. Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. Brain Inform 2018;5:2. https://doi.org/10.1186/s40708-018-0080-3.
[18] Korolev S, Safiullin A, Belyaev M, Dodonova Y. Residual and plain convolutional neural networks for 3D brain MRI classification. in: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), IEEE; 2017. p. 835–8. https://doi.org/10.1109/ISBI.2017.7950647.
[19] Lian C, Liu M, Zhang J, Shen D. Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. IEEE Trans Pattern Anal Mach Intell 2020;42:880–93. https://doi.org/10.1109/TPAMI.2018.2889096.
[20] Senanayake U, Sowmya A, Dawes L. Deep fusion pipeline for mild cognitive impairment diagnosis. in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE; 2018. https://doi.org/10.1109/ISBI.2018.8363832. 1394–1997.
[21] Taqi AM, Awad A, Al-Azzo F, Milanova M. The impact of multi-optimizers and data augmentation on TensorFlow convolutional neural network performance. in: 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), IEEE; 2018. p. 140–5. https://doi.org/10.1109/MIPR.2018.00032.
[22] Valliani A, Soni A. Deep residual nets for improved alzheimer's diagnosis. New York, NY, USA: in: Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology,and Health Informatics, ACM; 2017. https://doi.org/10.1145/3107411.3108224. 615–615.
[23] Wu C, Guo S, Hong Y, Xiao B, Wu Y, Zhang Q. Discrimination and conversion prediction of mild cognitive impairment using convolutional neural networks. Quant Imag Med Surg 2018;8:992–1003. https://doi.org/10.21037/qims.2018.10.17.
[24] Ebrahimi A, Luo S. For the A. Disease Neuroimaging Initiative, Convolutional neural networks for Alzheimer's disease detection on MRI images. J Med Imaging 2021;8. https://doi.org/10.1117/1.JMI.8.2.024503.
[25] Feng W, Van Halm-Lutterodt N, Tang H, Mecum A, Mesregah MK, Ma Y, Li H, Zhang F, Wu Z, X G, Yao Erlin. Automated MRI-based deep learning model for detection of Alzheimer's disease process. Int J Neural Syst 2020;30. https://doi.org/10.1142/S012906572050032X.
[26] R P, Guilherme Folego A, Weiler Marina, Casseb Raphael F. Alzheimer's disease detection through whole-brain 3D-CNN MRI. Front Bioeng Biotechnol 2020;8.
[27] Wen J, Thibeau-Sutre E, Diaz-Melo M, Samper-González J, Routier A, Bottani S, Dormont D, Durrleman S, Burgos N, Colliot O. Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation. Med Image Anal 2020;63. https://doi.org/10.1016/j.media.2020.101694.
[28] Malone IB, Cash D, Ridgway GR, MacManus DG, Ourselin S, Fox NC, Schott JM. MIRIAD–Public release of a multiple time point Alzheimer's MR imaging dataset. Neuroimage 2013;70:33–6. https://doi.org/10.1016/j.neuroimage.2012.12.044.
[29] Larobina M, Murino L. Medical image file formats. J Digit Imag 2014;27:200–6. https://doi.org/10.1007/s10278-013-9657-9.
[30] Ashburner J, Friston KJ. Unified segmentation. Neuroimage 2005;26:839–51. https://doi.org/10.1016/j.neuroimage.2005.02.018.
[31] Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front Neuroinf 2011;5:13. https://doi.org/10.3389/fninf.2011.00013.
[32] Ian Goodfellow AC. Yoshua bengio, deep learning, adaptive c. Cambridge, Mass. ; London: MIT Press; 2016.
[33] Murphy KP. Machine learning : a probabilistic perspective, Adaptive c. Cambridge, Mass. ; London: MIT Press; 2012.
[34] Chollet F. Deep learning with Python. Manning Publications; 2018.
[35] Sánchez Fernández I, Yang E, Calvachi P, Amengual-Gual M, Wu JY, Krueger D, Northrup H, Bebin ME, Sahin M, Yu K-H, Peters JM. Deep learning in rare disease. Detection of tubers in tuberous sclerosis complex. PLoS One 2020;15:e0232376. https://doi.org/10.1371/journal.pone.0232376.
[36] Medina G, Buckless CG, Thomasson E, Oh LS, Torriani M. Deep learning method for segmentation of rotator cuff muscles on MR images. Skeletal Radiol 2021;50:683–92. https://doi.org/10.1007/s00256-020-03599-2.
[37] K S, Nitish Srivastava S, Hinton Dropout. A simple way to prevent neural networks from over tting. J Mach Learn Res 2014;15:1929–58.
[38] Li X, Pang T, Xiong B, Liu W, Liang P, Wang T. Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification. in: Proceedings - 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2017; 2018. https://doi.org/10.1109/CISP-BMEI.2017.8301998.
[39] Friedman THRTJ. The elements of statistical learning. New York, NY: Springer New York; 2009. https://doi.org/10.1007/b94608.

[40] Jurman DCG. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genom 2020;21.
[41] Chicco D. Ten quick tips for machine learning in computational biology. BioData Min 2017;10:35. https://doi.org/10.1186/s13040-017-0155-3.
[42] Carter JV, Pan J, Rai SN, Galandiuk S. ROC-ing along: evaluation and interpretation of receiver operating characteristic curves. Surgery (United States) 2016;159:1638–45. https://doi.org/10.1016/j.surg.2015.12.029.
[43] He K, Girshick R, Dollar P. Rethinking ImageNet pre-training. in: 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE; 2019. p. 4917–26. https://doi.org/10.1109/ICCV.2019.00502.
[44] Hon M, Khan NM. Towards Alzheimer's disease classification through transfer learning. Proceedings - 2017 IEEE International Conference on Bioinformatics and Biomedicine, BIBM; 2017. p. 1166–9. https://doi.org/10.1109/BIBM.2017.8217822. 2017. 2017-January.
[45] Ebrahimi A, Luo S. For the A. Disease Neuroimaging Initiative, Convolutional neural networks for Alzheimer's disease detection on MRI images. J Med Imaging 2021;8:1–18. https://doi.org/10.1117/1.JMI.8.2.024503.
[46] Inglese M, Patel N, Linton-Reid K, Loreto F, Win Z, Perry RJ, Carswell C, Grech-Sollars M, Crum WR, Lu H, Malhotra PA, Silbert LC, Lind B, Crissey R, Kaye JA, Carter R, Dolen S, Quinn J, Schneider LS, Pawluczyk S, Becerra M, Teodoro L, Dagerman K, Spann BM, Brewer J, Vanderswag H, Fleisher A, Ziolkowski J, Heidebrink JL, Zbizek-Nulph, Lord JL, Zbizek-Nulph L, Petersen R, Mason SS, Albers CS, Knopman D, Johnson K, Villanueva-Meyer J, Pavlik V, Pacini N, Lamb A, Kass JS, Doody RS, Shibley V, Chowdhury M, Rountree S, Dang M, Stern Y, Honig LS, Mintz A, Ances B, Morris JC, Winkfield D, Carroll M, Stobbs-Cucchi G, Oliver A, Creech ML, Mintun MA, Schneider S, Geldmacher D, Love MN, Griffith R, Clark D, Brockington J, Marson D, Grossman H, Goldstein MA, Greenberg J, Mitsis E, Shah RC, Lamar M, Sood A, Blanchard KS, Fleischman D, Arfanakis K, Samuels P, Duara R, Greig-Custo MT, Rodriguez R, Albert M, Varon D, Onyike C, Farrington L, Rudow S, Brichko R, Greig MT, Kielb S, Smith A, Raj BA, Fargher K, Sadowski M, Wisniewski T, Shulman M, Faustin A, Rao J, Castro KM, Ulysse A, Chen S, Sheikh MO, Singleton-Garvin J, Doraiswamy PM, Petrella JR, James O, Wong TZ, Borges-Neto S, Karlawish JH, Wolk DA, Vaishnavi S, Clark CM, Arnold SE, Smith CD, Jicha GA, el Khouli R, Raslau FD, Lopez OL, Zmuda M, Butters M, Oakley M, Simpson DM, Porsteinsson AP, Martin K, Kowalski N, Martin KS, Keltz M, Goldstein BS, Makino KM, Ismail MS, Brand C, Reist C, Thai G, Pierce A, Yanez B, Sosa E, Witbracht M, Kelley B, Nguyen T, Womack K, Mathews D, Quiceno M, Levey AI, Lah JJ, Hajjar I, Cellar JS, Burns JM, Swerdlow RH, Brooks WM, Silverman DHS, Kremen S, Apostolova L, Tingus K, Lu PH, Bartzokis G, Woo E, Teng E, Graff-Radford NR, Parfitt F, Poki-Walker K, Farlow MR, Hake AM, Matthews BR, Brosch JR, Herring S, van Dyck CH, Mecca AP, Good SP, MacAvoy MG, Carson RE, Varma P, Chertkow H, Vaitekunis S, Hosein C, Black S, Stefanovic B, Heyn CC, Hsiung G-YR, Kim E, Mudge B, Sossi V, Feldman H, Assaly M, Finger E, Pasternak S, Rachinsky I, Kertesz A, Drost D, Rogers J, Grant I, Muse B, Rogalski E, Mesulam JRM-M, Kerwin D, Wu C-K, Johnson N, Lipowski K, Weintraub S, Bonakdarpour B, Pomara N, Hernando R, Sarrael A, Rosen HJ, Mackin S, Nelson C, Bickford D, Au YH, Scherer K, Catalinotto D, Stark S, Ong E, Fernandez D, Miller BL, Rosen H, Perry D, Turner RS, Johnson K, Reynolds B, MCCann K, Poe J, Sperling RA, Johnson KA, Marshall GA, Yesavage J, Taylor JL, Chao S, Coleman J, White JD, Lane B, Rosen A, Tinklenberg J, Belden CM, Atri A, Spann BM, Zamrini KACE, Sabbagh M, Killiany R, Stern R, Mez J, Kowall N, Budson AE, Obisesan TO, Ntekim OE, Wolday S, Khan JI, Nwulia E, Nadarajah S, Lerner A, Ogrocki P, Tatsuoka C, Fatica P, Fletcher E, Maillard P, Olichney J, DeCarli C, Carmichael O, Bates V, Capote H, Rainka M, Borrie M, Lee T-Y, Bartha R, Johnson S, Asthana S, Carlsson CM, Perrin A, Burke A, Scharre DW, Kataki M, Tarawneh R, Kelley B, Hart D, Zimmerman EA, Celmins D, Miller DD, Ponto LLB, Smith KE, Koleva H, Shim H, Nam KW, Schultz SK, Williamson JD, Craft S, Cleveland J, Yang M, Sink KM, Ott BR, Drake J, Tremont G, Daiello LA, Drake JD, Sabbagh M, Ritter A, Bernick C, Munic D, Mintz A, O'Connelll A, Mintzer J, Wiliams A, Masdeu J, Shi J, Garcia A, Sabbagh M, Newhouse P, Potkin S, Salloway S, Malloy P, Correia S, Kittur S, Pearlson GD, Blank K, Anderson K, Flashman LA, Seltzer M, Hynes ML, Santulli RB, Relkin N, Chiang G, Lin M, Ravdin L, Lee A, Sadowsky C, Martinez W, Villena T, Peskind ER, Petrie EC, Li G, Aboagye EO. A predictive model using the mesoscopic architecture of the living brain to detect Alzheimer's disease. Commun Med 2022;2:70. https://doi.org/10.1038/s43856-022-00133-4.