Computers in Biology and Medicine 171 (2024) 108168

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

A novel radiological software prototype for automatically detecting the inner ear and classifying normal from malformed anatomy

Abdulrahman Alkojak Almansi [b], Sima Sugarova [c], Abdulrahman Alsanosi [d], Fida Almuhawas [d], Louis Hofmeyr [e], Franca Wagner [f], Emerencia Kedves [g], Kiran Sriperumbudur [a], Anandhan Dhanasingh [a,**], Andras Kedves [a,b,*]

[a] MED-EL Medical Electronics GmbH., Department of Research and Development, Innsbruck, Austria
[b] University of Pecs, Faculty of Engineering and Information Technology, Institute of Information and Electrical Technology, Pecs, Hungary
[c] St. Petersburg ENT and Speech Research Institute, St. Petersburg, Russia
[d] King Saud University, King Abdullah Ear Specialist Center (KAESC), Department of Otolaryngology, Riyadh, Saudi Arabia
[e] Dr Loius Hofmeyr's workplace to Stellenbosch University Division of Otorhinolaryngology, Stellenbosch, South Africa
[f] University Hospital Bern, University Institute for Diagnostic and Interventional Neuroradiology, Switzerland
[g] University of Sopron, Doctoral School of Wood Sciences and Technologies, Sopron, Hungary

ARTICLE INFO

Keywords:
Artificial intelligence
Deep learning
Radiology
Digital health

ABSTRACT

Background: To develop an effective radiological software prototype that could read Digital Imaging and Communications in Medicine (DICOM) files, crop the inner ear automatically based on head computed tomography (CT), and classify normal and inner ear malformation (IEM).

Methods: A retrospective analysis was conducted on 2053 patients from 3 hospitals. We extracted 1200 inner ear CTs for importing, cropping, and training, testing, and validating an artificial intelligence (AI) model. Automated cropping algorithms based on CTs were developed to precisely isolate the inner ear volume. Additionally, a simple graphical user interface (GUI) was implemented for user interaction. Using cropped CTs as input, a deep learning convolutional neural network (DL CNN) with 5-fold cross-validation was used to classify inner ear anatomy as normal or abnormal. Five specific IEM types (cochlear hypoplasia, ossification, incomplete partition types I and III, and common cavity) were included, with data equally distributed between classes. Both the cropping tool and the AI model were extensively validated.

Results: The newly developed DICOM viewer/software successfully achieved its objectives: reading CT files, automatically cropping inner ear volumes, and classifying them as normal or malformed. The cropping tool demonstrated an average accuracy of 92.25%. The DL CNN model achieved an area under the curve (AUC) of 0.86 (95% confidence interval: 0.81-0.91). Performance metrics for the AI model were: accuracy (0.812), precision (0.791), recall (0.8), and F1-score (0.766).

Conclusion: This study successfully developed and validated a fully automated workflow for classifying normal versus abnormal inner ear anatomy using a combination of advanced image processing and deep learning techniques. The tool exhibited good diagnostic accuracy, suggesting its potential application in risk stratification. However, it is crucial to emphasize the need for supervision by qualified medical professionals when utilizing this tool for clinical decision-making.

1. Introduction

According to reports from all over the world, children with congenital hearing loss have an inner ear malformation (IEM) rate of 20%–30% [1,2]. Gross malformations (cochlear aplasia, cochlear nerve aplasia, Michel's deformity) are a contraindication to cochlear implantation surgery and major deformities may cause surgical issues. Since computed tomography (CT) provides important information on the state of the inner ear and other abnormalities both before and after surgery, it has been the main method of diagnosis [3]. It can be challenging for non-specialists to make an accurate diagnosis of inner ear pathologies

* Corresponding author. MED-EL Medical Electronics GmbH., Department of Research and Development, Innsbruck, Austria.
** Corresponding author.
E-mail addresses: anandhan.dhanasingh@medel.com (A. Dhanasingh), andras.kedves@medel.com (A. Kedves).

https://doi.org/10.1016/j.compbiomed.2024.108168
Received 4 December 2023; Received in revised form 29 January 2024; Accepted 15 February 2024
Available online 16 February 2024
0010-4825/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Abbreviations

(AUC)	Area under curve
(3D)	Three-dimensional
(CT)	Computed tomography
(CNN)	Convolutional neural network
(CI)	Confidence interval
(DL)	Deep learning
(DICOM)	Digital Imaging and Communications in Medicine
(ENT)	Ear, nose, and throat
(HU)	Hounsfield unit
(IPIII)	Incomplete partition type III
(IEM)	Inner ear malformation
(MWU)	Mann-Whitney U rank test
(Px)	Px:pixels
(ROC)	Receiver operating characteristic
(RCNN)	Region-Based Convolutional Neural Network

due to their complex anatomical structure [4]. However, automatic detection of IEMs using deep learning (DL) could lead to the development of new methods for diagnosing all types of IEMs. DL is a type of artificial intelligence (AI) that can learn to recognize patterns in data, and it has been shown to be effective for automatically detecting anatomical structures in medical images.

A tool for automatic identification of the malformed inner ear does not exist to our best knowledge up to know, however, there are attempts to diagnose it semiautomatically (with manually adjusted inner ear localization) [5] or using methods developed for normal anatomy and finding the inner ear with unknown accuracy [6,7].

Despite the fact that DL has become more common in the medical community [8], in particular when contrasting human and machine performance [9], and DL is employed to evaluate temporal bone CTs [5], no study on fully automated inner ear crop and classification has been published before.

To analyse the complex structure of the temporal bone, 3 dimensional - convolutional neural network (3D-CNN) would be required. The recent development of CNN architectures and powerful computer graphics processing units with a lot of memory has made 3D-DL useful for medical image analysis [10].

Additionally, it was mentioned that, when compared to unsupervised techniques, supervised abnormality detection with only 64 abnormal samples was excellent. In supervised deep learning, training data with labels like “normal” or “abnormal” are used to build a standard classification CNN that classifies test data [11].

To the best of our knowledge, a fully automatized DL workflow for inner ear malformation diagnosis has not yet been developed and validated in any of the available studies. Our AI technology can help professionals evaluate CT images more accurately. The tool might also be able to provide crucial second opinion information needed for patient-tailored diagnosis and surgery prior to or after the implantation of a cochlear implant (CI). One study only covers congenital inner ear malformations, missing ossification [6], or does not have diverse (based on 1 country), and robust number of input data [7]. Another study classified inner ear with outstanding accuracy, however it provided a semi-automatic way of inner ear localization, which made it time-consuming and impossible to implement in real-life workflow and was restricted to using a single institute technique [5]. Present study focused on software development for the fully automatized diagnosis of wide range of inner ear malformations, by evaluating the effectiveness of supervised DL on CT images and whether abnormal regions were accurately detected. Deep learning-based categorization needs data from multicentre studies and large-scale evaluations, but none of these studies have been carried out for any IEM. To fill this unmet need, we gathered information from three cohorts and countries, created and evaluated an inner ear cropping tool, and assessed the diagnostic performance of a deep learning-based classifier. As a result, we developed an AI based software for inner ear diagnosis.

2. Material and methods

2.1. Patient cohorts, image analyses, pre-processing

All experiments were conducted in strict accordance with the Declaration of Helsinki, and the results were reported in full compliance with the Standards for Reporting Diagnostic Accuracy. The retrospective study to use the anonymized images for research was approved by the Institutional Review Board and ethics committees of King Saud University, Saudi Arabia (IRB No. 22/0084/IRB), St. Petersburg ENT and Speech Research Institute, Russia (IRB No. 23_001/IRB), and Stellenbosch University, South Africa (IRB No. S_23_001/IRB). The ethics committees granted their approval on the grounds that the study was retrospective and did not require informed patient consent.

CT scans of the head of 2053 potential CI candidates with various inner ear anatomical types were retrieved from the three cohorts retrospectively between January 1, 2016, and December 31, 2021. The CT images were acquired using different vendors and scanning parameters, based on the provided Digital Imaging and Communications in Medicine (DICOM) files. Non-enhanced axial reconstructed images with fields of view that encompassed both ears or one depicted side (Fig. 1) were included. In general, 256-slice (range: 128–512) multidetector-row CT scanners were used. The following scanning parameters were used in the majority of cases: axial plane, 0.625 mm (range: 0.48–1.5) slice thickness, auto-mAs (range: 120–145 mAs), 140 kV, 484 to 1024 matrix, bone kernel, and rotation time 1 s with 0.3 mm reconstruction in the axial and coronal views. The cases included in the cohorts were selected randomly. Based on the radiological reports, the images were enrolled into six defined categories: (1) normal anatomy (defined as the “normal” group), (2) common cavity (CC), (3) cochlear hypoplasia (CH), (4) enlarged vestibular aqueducts (EVAS), (5) incomplete partition (IP) type I, (6) IP type III, and (7) ossification (defined as the “abnormal” group) (Fig. 5). Pre- and post-operative head CTs of implanted (post-op) and non-implanted ears were also included.

The following conditions were excluded: (1) a history of a cranial injury; (2) not sufficient slice thickness (>1 mm) and (3) resolution with more than 1.5 mm/pixel.

2.1.1. Cropping algorithms

We developed different algorithms for cropping the inner ear in a volume and chose the top two for final test;

Hounsfield units (HU) and distance-based cropping: The fundamental principle behind this approach lays on identifying high Hounsfield unit values within the intricate bone structure of the inner ear. The method initiates by applying an averaging technique using a window filter, typically set at dimensions around 5 mm × 5 mm x 5 mm, in order to merge the inner ear as one blob presented in the whole volume. It loops through each voxel, starting from the highest HU value voxel in descending order. It calculates the distance between each consecutive voxels, if this computed distance exceeds a predetermined threshold, typically set at 20 mm, which corresponds to the average longest distance between two points within a normal inner ear, the previous voxel is chosen as the reference point, then a cropping with considering the selected voxel as the center will be applied.

Two-staged spatial auto k-means clustering cropping: This methodology focuses on the interrelations among voxels in addition to their HU values, the first stage starts by averaging the volume with 5 mm × 5 mm x 5 mm kernel, then it takes only the highest 0.5% voxels from the whole volume. The algorithm following this step automatically determines the optimal number of clusters using the elbow method, the next step is to take the closest cluster to the volume center, presuming the inner ear is

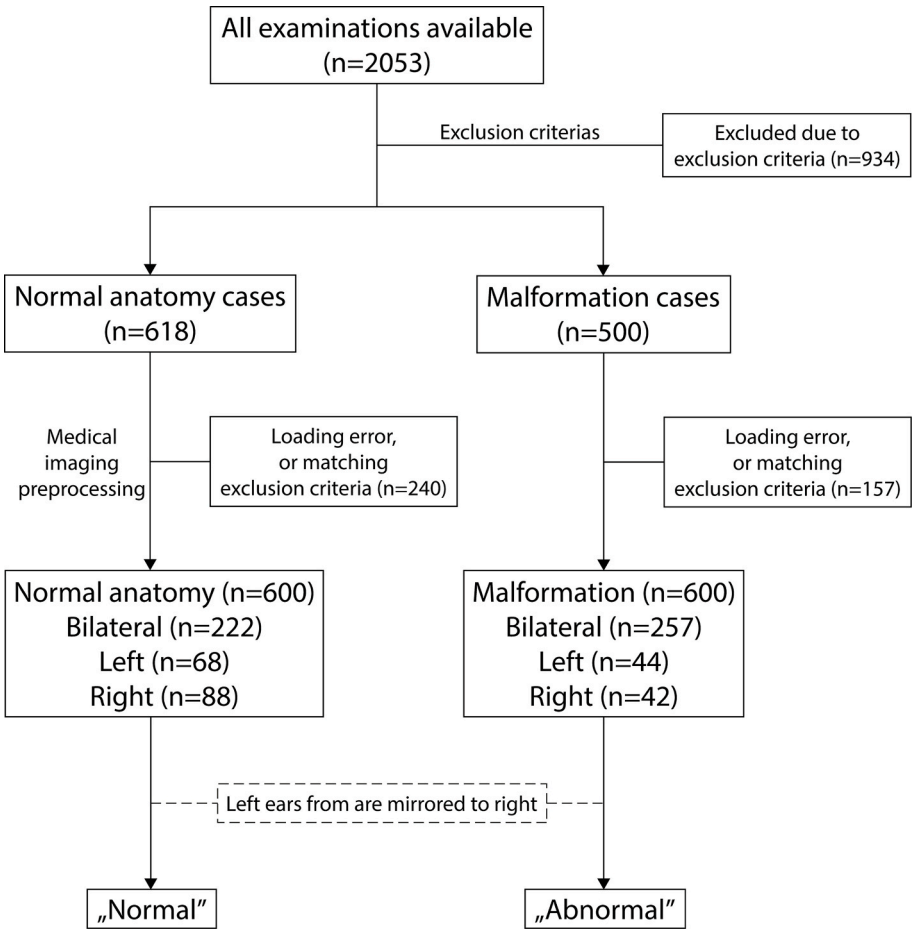

Fig. 1. Flowchart outlining the selection of cases. If there were multiple examinations for a patient, the oldest evaluation for abnormalities was included, and the contralateral normal ear was eliminated. Images of the left ear, derived from bilateral images, were all horizontally mirrored to the right.

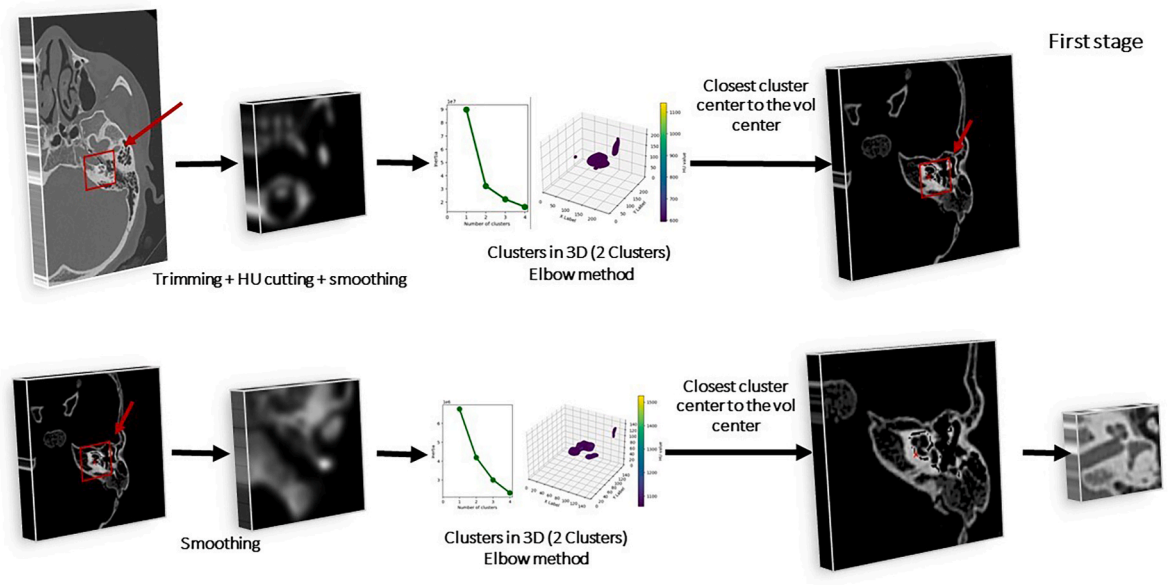

Fig. 2. Cropping algorithm of the inner ear.

typically around the volume center, cropping with 45 mm × 45 mm x 45 mm dimensions will be applied, the second stage follows with same procedure but with different parameters namely 3 mm × 3 mm x 3 mm smoothing filter and 20 mm × 20 mm x 18 mm cropping dimensions.

We used elbow method to increase the accuracy. This approach examines the fraction of variance or inertia plotted against cluster count, where the variance is determined based on the calculation of the within-cluster sum of squares value (Euclidean distance). that is predicated on the notion that improving data separation does not result from adding clusters above a particular threshold. The initial clusters will provide a lot of information, but eventually the marginal gain will sharply decline, creating an angle in the graph. This is the point where the "elbow criterion" is applied, since the right "k," or number of clusters, is selected [12]. Algorithm is summarized in Fig. 2.

2.2. Dataset, model developments, settings, and evaluation

Based on the cropped clinical CT series, 600 malformation inputs were selected. 8.33% of the data (n = 50/600) was saved for secondary validation. Of the remaining 91.67% (550/600), 70% (385/550) was used to train the model and 30% (165/550) was used for testing and validation. The same data pipeline was created for normal anatomy (n = 600). In total, 1200 CT series were selected as input.

Two biomedical engineers with 5 and 13 years of experience, respectively, and a neuroradiologist with 10 years of experience in head and neck imaging collaboratively enrolled cases into two labeled datasets: "Normal" and "Malformed." The evaluation of the test data for supervised deep learning is summarized in Fig. 3. We performed a 5-fold cross-validation.

The DL CNN model was created with internal parameters from the ground up. The architecture of this supervised deep learning was as follows: The neural network comprises:

(1) Input layer: The input layer has a shape of (width = 128, height = 128, depth = 64, 1), means that the input data is a 3D image with 1 channel.(2) Convolutional layers: The convolutional layers use filters of size (3, 3, 3) to extract spatial features from the input data. The filters are applied to the input data in a sliding window fashion, and the outputs of the filters are combined to produce a feature map. The feature maps are then passed to the next convolutional layer, where more complex features are extracted. (3) Max pooling layers: The max pooling layers down sample the feature maps by taking the maximum value from a pool of pixels. This helps to reduce the computational complexity of the model as well helps to prevent overfitting. (4) Batch normalization layers: The batch normalization layers normalize the outputs of previous layers, which helps to improve the stability of the training process. (5) Global average pooling layer: The global average pooling layer averages the outputs of all the convolutional layers, which produces a single feature vector. (6) Dense layers: The dense layers project the feature vector onto a new space. The first dense layer has 512 units, and the second dense layer has 1 unit. (7) Output layer: The output layer produces a single output value, which is the probability of the input data belonging to a certain class. The parameters of the neural network are optimized and compiled. The architecture of the model can be seen on Supplementary Fig. 1.

After loading the data, it was segregated into defined "Normal" and "Malformed" groups. Then, augmented images for training the CNN were generated. Data augmentation and auto-contrast functions were used each time the training and test data were read, respectively.

The following data augmentation methods were used: random rotation range (<10°), random horizontal shifts (−40px-+40px for three axes), random zoom (65–145%), and random rotation (<15°). Additionally, contrast correction was performed by randomly removing the lowest and highest values within 0–2%. Image contrast was altered for the auto-contrast function by removing the 0.7% of minimum and maximum values. All parameters were used to apply random transformations to the images to increase the size of the training dataset and improve the robustness of the model. Batch size has been set to 50. For each dataset, the maximum number of epochs was set at 100, and training was terminated early if the validation loss did not decrease after 10 epochs. Structure of the model could be seen in Supplementary Fig. 1. The models trained on an NVIDIA GeForce RTX 4080 Laptop GPU with 12 GB GDDR6 VRAM. We tested the model after training it to see how well it performed.

2.3. Validation, statistic evaluation

To validate the developed decision support system, 100 images (50 "Normal" and 50 "Malformed") images randomly selected (and separated before model development, and these images have never seen by the model during the model training). AI performance for the full dataset (except 100 images for interobserver analysis) was also measured with standard Area Under the Curve of the Receiver Operating Characteristic

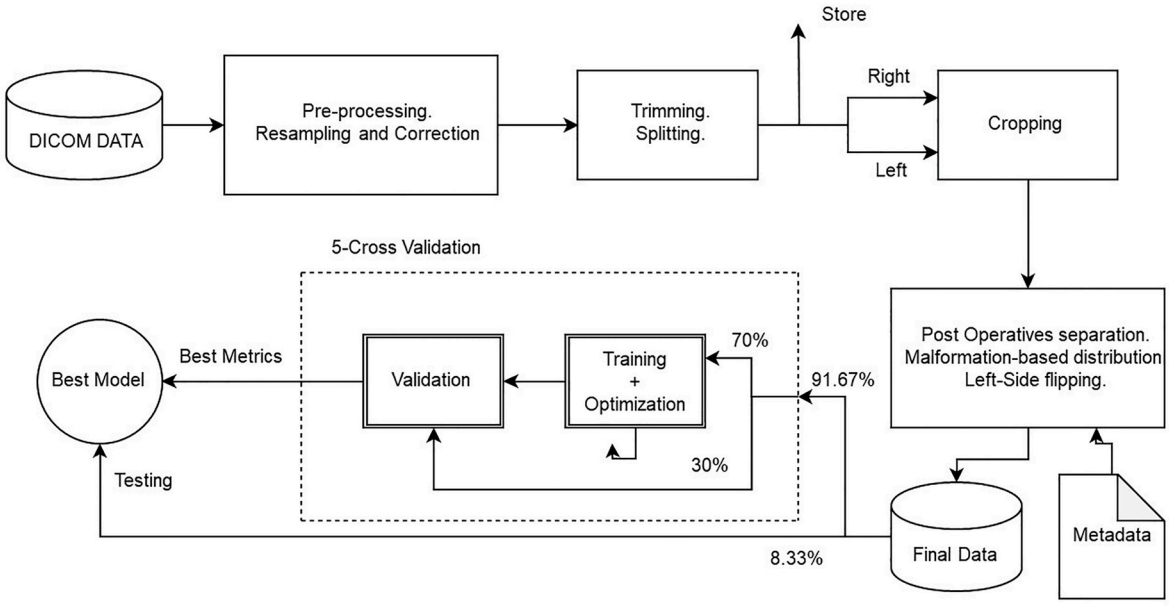

Fig. 3. Supervised deep learning model architecture.

curve (AUC-ROC) curve.

Statistical analyses: The Spearman rank-order correlation coefficient was used to calculate the correlation between two variables. Analyses were carried out using Python 3.12.0.

3. Results

The analysis included 600 "Normal" anatomy and 600 "Malformed" CT series, as can be seen in the flowchart (Fig. 1.) and evaluated. The most prevalent abnormality was cochlear hypoplasia.

Cropping tools accuracy: Hounsfield units and distance-based cropping algorithm achieved the following mean accuracy of 87.2% with the malformations. The determined point is more likely to be located around the semi-circular canals. Two-staged spatial auto k-means clustering cropping algorithm achieved 92.5% mean accuracy with the malformations. Determined point is closer to the basal turn center and Internal auditory canal. Detailed accuracies per malformation types, "Normal" and "Malformed" groups, and for pre and postoperative groups could be seen in Table 1., Table 2., and Table 3.

During the development of AI classifier, accuracy, precision, recall, and F1 scores are 0.812, 0.791, 0.8, and 0.766, respectively. The supervised DL learning provided accurate categorization in each of the four situations. Performance shown by the area under the curve (AUC) is 0.86 (95% CI: 0.81–0.91). On top of application programming interfaces (API), graphic user interface (GUI) was built (Fig. 4, Video 1).

Supplementary video related to this article can be found at https://doi.org/10.1016/j.compbiomed.2024.108168

4. Discussion

In this study, we developed a method for getting fully automatized and immediate response for differentiating normal from malformed pathologic inner ear anatomy, as well a decision support tool for a complex aspect of inner ear surgery. Two automatized cropping tools developed for cropping the inner ear volume, performances have been measured and compared. Additionally, we created a DL model using a multicentric, multi-national CT series, and we assessed the models' accuracy, precision, recall, and f1 score as well as their overall performance (AUC). By using cross-sectional medical imaging, our goal was to create an AI-based efficient software with GUI for inner ear diagnosis. Radiology was revolutionized by deep learning, which made it possible to identify anatomical variation [13]. Smaller datasets were used in earlier deep learning-based studies on the diagnosis of malformations in head and neck/temporal bone CT scans [5,6,14].

The differences in accuracy between the HU-based and K-means based algorithms are not generally attributable to the type of anatomy. For instance, the HU-based algorithm incorrectly identified teeth and skull fragments in a small number of instances, whereas the K-means algorithm did not exhibit these errors. The differences in accuracy for different anatomical types are likely attributable to the variability in the input data. For example, in the case of cochlear hypoplasia (CH), the accuracy can vary depending on the specific subcategory (CH I, CH II, CH III, or CH IV) [1], which may cause the HU-based algorithm to perform less consistently than the K-means algorithm (84% and 95%, respectively).

We assumed that a completely self-developed, easily structured would be suitable for classification, which was also confirmed by the fact in another study with a lot more complex architecture they reached almost the same accuracy [6]. In our study, we created a unified method, which uses directly CT series as an input, therefore we do not lose any data during the training of the model, which could lead to a model that is more trustful compared to a model that uses a segmentation, then a point cloud. Moreover, in our case, any clinical CT can be used as an input regardless of the input matrices which makes it applicable worldwide, especially in developing countries with older CT machines and a lack of professional staff. Ogawa et al. reported applying unsupervised DL using 3D variational autoencoder in detecting inner ear abnormalities on CT images with fewer malformation types and was limited to a single institute [5]. Overall, we ended up with a solution that is completely automatized and not limited by a semi-automatized method [5] or lack of variance of input data during the training [6]. We developed a model with good supervised learning accuracy [15] however, Ogawa et al. found unsupervised learning to be more promising and achieved the highest accuracy when compared to other supervised learning based studies. This is noteworthy due to high variance between the two unsupervised learning based models' performance for the same purpose, which raises questions [5,6]. Moreover, the software could be used for primary screening of inner ear malformations in primary care settings or for incidental CT examinations of newborns.

A systematic review by Huang et al., in 2023 found that deep learning models have achieved diagnostic accuracy comparable to that of human experts in narrow clinical tasks for several medical domains and imaging modalities [16]. Given the complexity of inner ear anatomical variations, it would be advantageous for end-users to have an assistive AI tool in their decision-making, especially during the preparation of cochlear implant surgeries. This tool could aid in the selection of the appropriate electrode for a patient-tailored setting and in taking precautionary steps to seal the cochlea well stopping the cerebrospinal fluid gusher, a complication that occurs in most incomplete partition type III cases. If unnoticed by the ENT surgeon prior to surgery can lead to interoperative complications. Fujima et al. also were of the opinion that deep learning techniques are a beneficial adjunct to the diagnostic process for specific diseases on head CT scans and the performance of AI is not significantly different from that of human experts [17]. While the goal of this tool is to assist the average user in the absence of a radiological specialist, it is worth conducting studies comparing DL AI model and a group of radiologists and experts.

In our study, we used diversification approach to reduce risk based

Table 2
Accuracy comparison between algorithms on post and pre-operative scans. Abbreviation: Hounsfield unit (HU).

Algorithm	Post operative	Pre-operative
HU-based	41%	87.2%
K-means-based	88%	92.5%

Table 3
Overall cropping accuracies for normal and abnormal anatomies. Abnormal group consists of Common cavity (CC), Incomplete partition type I, Incomplete partition type III (IPIII), Enlarged vestibular aqueduct syndrome, Ossification (OS), Cochlear hypoplasia (CH). Abbreviation: Hounsfield unit (HU).

Algorithm	Normal anatomy	Abnormal
HU-based	83%	88%
K-means-based	90%	93%

Table 1
Accuracy table for the proposed algorithms on the pre-operative data. Abbreviations: Hounsfield unit (HU), Normal anatomy (NA), Common cavity (CC), Incomplete partition type I, Incomplete partition type III (IPIII), Enlarged vestibular aqueduct syndrome, Ossification (OS), Cochlear hypoplasia (CH).

Algorithm	NA	CC	IPI	IPIII	EVAS	OS	CH	Overall
HU-based	83%	93%	94%	76%	86%	88%	84%	87.2%
K-means-based	90%	88%	98%	93%	88%	95%	95%	92.5%

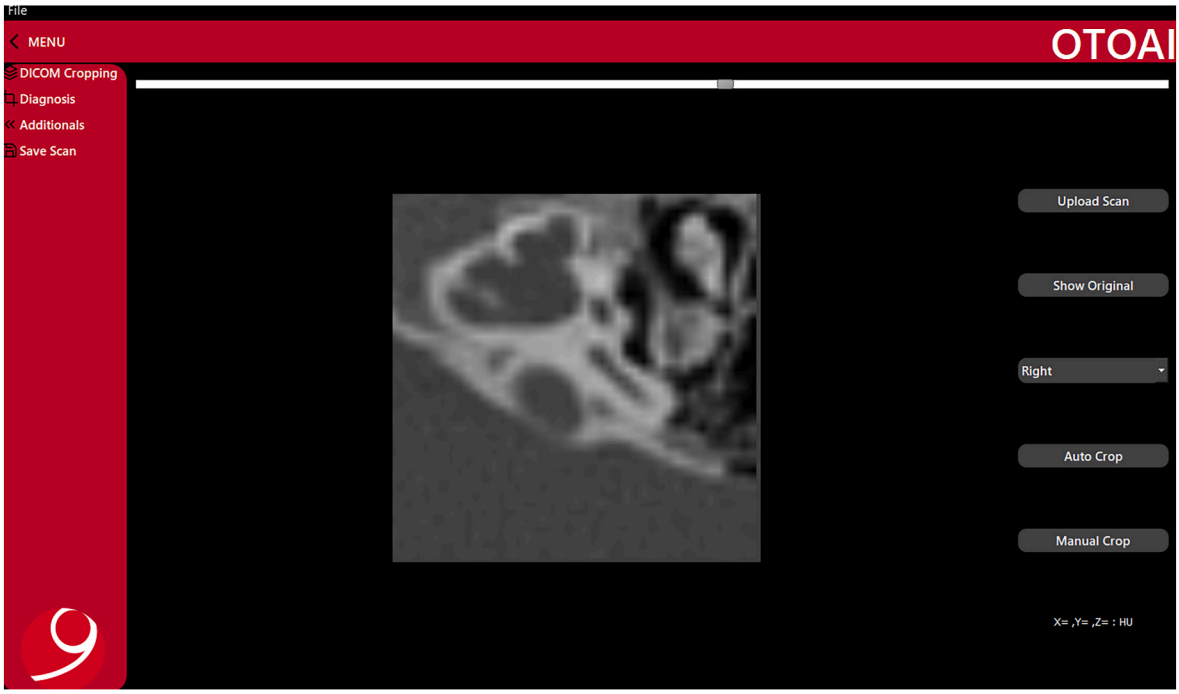

Fig. 4. Software prototype GUI along with functions and the classification at the time of submission of the draft. Abbreviations: graphical unit interface (GUI).

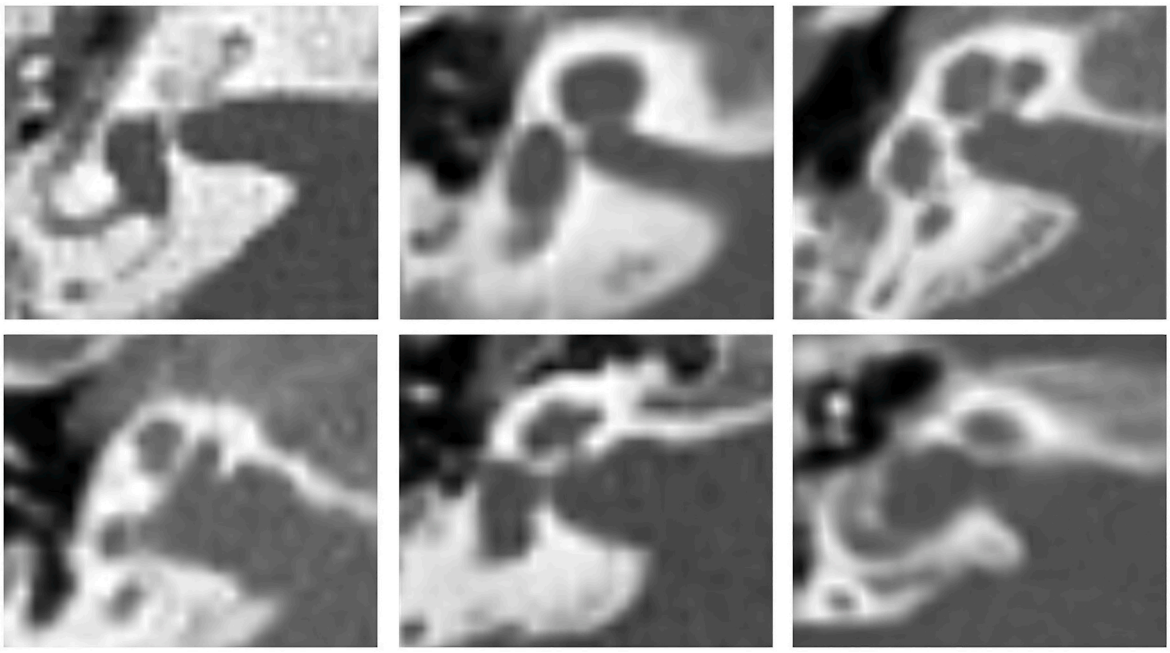

Fig. 5. Mid-modiolar section of an ossification, incomplete partition type I, normal anatomy (upper line), incomplete partition type III, cochlear hypoplasia, common cavity (bottom line).

on the Nobel Prize-winning study by Markowitz, who first described that returns can be increased while keeping risk low by diversifying a portfolio [18]. Thus, we involved input images from different institutes and countries covering wide variations in inner ear anatomy making our model strong thereby benefitting young residents and physicians prior to CI surgery. However, the strict requirements of the General Data Protection Regulation (GDPR) make it difficult for other parties to provide images, so using the tool locally may be a viable approach to obtaining additional expert opinions from outside of the institute. Another possible implementation is to support the additional post-processing of the temporal bone, based on the caution of software, to help radiologists report faster.

One of the limitations of our software is that we had used only 1200 cases from three countries to train and test our AI model which certainly needs improvisation involving more sites and several images to make a robust model. The limitations of the cropping algorithms are the following: HU-algorithm detected the teeth, and parts of the skull in a few cases. This would have caused the lower overall accuracy of the HU-algorithm based crops compared the K-means solution. K-means algorithm detected the skull in those cases where the inner ear is far from the

volume center. Perhaps we could have achieved higher specificity and sensitivity if we had used unsupervised 3D techniques or 3D-Region-Based Convolutional Neural Network (RCNN) as reported by Ogawa et al. [5] However, AUC values were, varied in different research that used the same unsupervised methodology. In the next study, our software should be validated by human experts. Future research needs to focus on the development of multi-class classifiers for additional categories of malformations, auto-3D segmentation of the complete inner ear, and/or AI-based methods for cochlear nerve research using magnetic resonance imaging diffusion tensor imaging (MRI DTI). A notable limitation of our study is that we were unable to include all types of malformations, including vestibular malformations, due to the balanced distribution of data.

Deep learning methods have the potential to be a valuable tool for the identification of inner ear malformations on head CT scans. This technology could serve as an effective adjunct to the diagnostic process.

5. Conclusions

This research shows the development and validation of a possible fully automatized workflow for classifying normal versus abnormal inner ear anatomy. The tool could carry good diagnostic accuracy during risk stratification, however, must be supervised by the decision-maker.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data sharing

We cannot make any individual patient-level data available to others, but these data can be requested from the respective institutions.

Contributors

AD, and AK conceptualised and designed the study. SS, AA, FA, and LH contributed associated radiological data. AD and AK pre-processed the data. FW was responsible for quality control of the radiological data. AD and AK provided computing resources and contributed to the clinical interpretation of the data. AK created the concept for the cropping tool and the deep learning architecture. AAA and AK developed the final form of the cropping tool. AK created the final form of the deep learning convolutional neural network. AAA and AK developed the software. AA and AK analysed the data. FW verified the underlying data. EK, AAA and AK made the visualisation. All authors had access to the underlying data. All authors contributed to interpretation of the results. KS, AK wrote the manuscript, and all authors critically reviewed the manuscript.

Evidence before this study

We conducted a systematic literature review to identify studies published in English or German up to Oct 15, 2023, that used the terms “inner ear,” “deep learning,” “radiology,” and “algorithm” in PubMed, Google Scholar, and ScienceOpen. We found only two prior study that used deep learning (DL) algorithms to detect inner ear abnormalities, but none of them used image inputs from various geographical locations. Deep learning has the potential to enable the development of efficient image analysis algorithms for medical imaging of the inner ear. However, to our knowledge, the fully automatised “1-click-solution”, using supervised deep learning with image (pre)processing instead of segmentation, to distinguish inner ear abnormalities from normal anatomy in computed tomography (CT) scans has not been previously investigated.

Added value of this study

To accurately distinguish between classes of inner ears in temporal bone CTs, we created cropping algorithms, and presented the performance of them. Using three separate cohorts from different institutes, we developed a DL convolutional neural network (CNN) for classification. We measured the accuracies, and the stability of each tool (cropping and classifier) and provided an efficient graphical user interface (GUI), a software.

Implications of all the available evidence

Our AI tool may enable specialists to assess CT scans more accurately. Although further prospective validation is required, the tool may also provide critical second-opinion information essential for patient-tailored diagnosis and surgery before or after CI placement. The integration of our implemented tool with an ontological planning software or other environments would significantly accelerate the development of digital radiology for inner ear malformations, for which, to our knowledge, no solution was previously available.

CRediT authorship contribution statement

Abdulrahman Alkojak Almansi: Writing – original draft, Software, Methodology. Sima Sugarova: Data curation. Abdulrahman Alsanosi: Data curation. Fida Almuhawas: Data curation. Louis Hofmeyr: Data curation. Franca Wagner: Writing – review & editing, Validation, Methodology. Emerencia Kedves: Visualization. Kiran Sriperumbudur: Writing – review & editing. Anandhan Dhanasingh: Writing – review & editing, Resources, Project administration, Funding acquisition. Andras Kedves: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank all contributors and investigators from MED-EL Medical Electronics GmbH, Department of Research & Development, Innsbruck, Austria. We thank King Saud University, St. Petersburg ENT and Speech Research Institute, Stellenbosch University, University Institute for Diagnostic and Interventional Neuroradiology, University Hospital Bern, University of Pecs, and University of Sopron supporting this study. We would like to thank everyone who helped to make this research possible.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2024.108168.

References

[1] L. Sennaroğlu, M.D. Bajin, Classification and current management of inner ear malformations, Balkan Med. J. 34 (2017) 397–411, https://doi.org/10.4274/balkanmedj.2017.0367.

[2] B. Sun, P. Dai, C. Zhou, [Study on 2,747 cases of inner ear malformation for its classification in patient with sensorineural hearing loss], Lin chuang er bi yan hou tou jing wai ke za zhi = J. Clin. Otorhinolaryngol. head, neck Surg. 29 (2015) 45–47.

[3] G. Widmann, D. Dejaco, A. Luger, J. Schmutzhard, Pre- and post-operative imaging of cochlear implants: a pictorial review, Insights Imaging 11 (2020).

[4] M. Bassiouni, H.C. Bauknecht, G. Muench, H. Olze, J. Pohlan, Missed radiological diagnosis of otosclerosis in high-resolution computed tomography of the temporal bone—retrospective analysis of imaging, radiological reports, and request forms, J. Clin. Med. 12 (2023).
[5] M. Ogawa, et al., Utility of unsupervised deep learning using a 3D variational autoencoder in detecting inner ear abnormalities on CT images, Comput. Biol. Med. 147 (2022).
[6] L. Diez, R. Reinhold, Unsupervised Classification of Congenital Inner Ear Malformations Using DeepDiffusion for Latent Space Representation Unsupervised Classification of Congenital Inner Ear Malformations Using DeepDiffusion for Latent Space Representation, 2023.
[7] A.T. Radutoiu, F. Patou, J. Margeta, R.R. Paulsen, P. López Diez, Accurate localization of inner ear regions of interests using deep reinforcement learning, in: Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 13583, LNCS, 2022, pp. 416–424.
[8] G. Litjens, et al., A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.
[9] X. Liu, et al., A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis, Lancet. Digit. Heal. 1 (2019) e271–e297.
[10] S.P. Singh, et al., 3d deep learning on medical images: a review, Sensors 20 (2020) 1–24.
[11] L. Ruff, R.A. Vandermeulen, B.J. Franks, K.-R. Müller, M. Kloft, Rethinking Assumptions in Deep Anomaly Detection, 2020.
[12] Purnima Bholowalia, Arvind Kumar, EBK-means: a clustering technique based on elbow method and K-means in WSN, Int. J. Comput. Appl. 105 (2014) 975–8887.
[13] Z. Hasan, et al., Convolutional neural networks in ENT radiology: systematic review of the literature, Ann. Otol. Rhinol. Laryngol. 132 (2023) 417–430.
[14] D. Zhang, J.H. Noble, B.M. Dawant, Automatic detection of the inner ears in head CT images using deep convolutional neural networks, Proc. SPIE-Int. Soc. Opt. Eng. 64 (2018) 78.
[15] A.A.H. de Hond, E.W. Steyerberg, B. van Calster, Interpreting area under the receiver operating characteristic curve, Lancet Digit. Heal. 4 (2022) e853–e855.
[16] S.C. Huang, et al., Self-supervised learning for medical image classification: a systematic review and implementation guidelines, npj Digit. Med. 6 (2023).
[17] N. Fujima, et al., Utility of deep learning for the diagnosis of otosclerosis on temporal bone CT, Eur. Radiol. 31 (2021) 5206–5211.
[18] T. Journal, N. Mar, Portfolio Selection Harry Markowitz 7 (2007) 77–91.