The Journal of International Advanced Otology

J Int Adv Otol 2023; 19(5): 360–367 • DOI: 10.5152/iao.2023.231073

Original Article

A Deep Learning Algorithm to Identify Anatomical Landmarks on Computed Tomography of the Temporal Bone

Zubair Hasan[1,2], Seraphina Key[3], Michael Lee[1], Fiona Chen[4], Layal Aweidah[2], Aaron Esmaili[5], Raymond Sacks[1,6], Narinder Singh[1,2]

[1]University of Sydney, Faculty of Medicine and Health, New South Wales, Australia
[2]Department of Otolaryngology - Head and Neck Surgery, Westmead Hospital, New South Wales, Australia
[3]Monash University, Faculty of Medicine, Nursing and Health Sciences, Victoria, Australia
[4]Department of Otolaryngology, Royal Children's Hospital Melbourne, Melbourne, Australia
[5]Department of Otolaryngology, Sir Charles Gairdner Hospital, Nedlands, Australia
[6]Department of Otolaryngology - Head and Neck Surgery, Concord Hospital, New South Wales, Australia

ORCID iDs of the authors: Z.H. 0000-0002-5346-578X, S.K. 0000-0002-1019-5986, M.L. 0000-0003-4117-0404, F.C. 0000-0002-9169-2617, L.A. 0000-0002-8989-6602, A.E. 0000-0002-8878-7658, R.S. 0000-0002-9260-5175, N.S. 0000-0002-7719-1832.

Cite this article as: Hasan Z, Key S, Lee M, et al. A deep learning algorithm to identify anatomical landmarks on computed tomography of the temporal bone. *J Int Adv Otol.* 2023;19(5):360-367.

BACKGROUND: Petrous temporal bone cone-beam computed tomography scans help aid diagnosis and accurate identification of key operative landmarks in temporal bone and mastoid surgery. Our primary objective was to determine the accuracy of using a deep learning convolutional neural network algorithm to augment identification of structures on petrous temporal bone cone-beam computed tomography. Our secondary objective was to compare the accuracy of convolutional neural network structure identification when trained by a senior versus junior clinician.

METHODS: A total of 129 petrous temporal bone cone-beam computed tomography scans were obtained from an Australian public tertiary hospital. Key intraoperative landmarks were labeled in 68 scans using bounding boxes on axial and coronal slices at the level of the malleoincudal joint by an otolaryngology registrar and board-certified otolaryngologist. Automated structure identification was performed on axial and coronal slices of the remaining 61 scans using a convolutional neural network (Microsoft Custom Vision) trained using the labeled dataset. Convolutional neural network structure identification accuracy was manually verified by an otolaryngologist, and accuracy when trained by the registrar and otolaryngologist labeled datasets respectively was compared.

RESULTS: The convolutional neural network was able to perform automated structure identification in petrous temporal bone cone-beam computed tomography scans with a high degree of accuracy in both axial (0.958) and coronal (0.924) slices ($P < .001$). Convolutional neural network accuracy was proportionate to the seniority of the training clinician in structures with features more difficult to distinguish on single slices such as the cochlea, vestibule, and carotid canal.

CONCLUSION: Convolutional neural networks can perform automated structure identification in petrous temporal bone cone-beam computed tomography scans with a high degree of accuracy, with the performance being proportionate to the seniority of the training clinician. Training of the convolutional neural network by the most senior clinician is desirable to maximize the accuracy of the results.

KEYWORDS: Otolaryngology, radiographic image interpretation, artificial intelligence, neural networks

INTRODUCTION

Surgery of the temporal bone and mastoid is often required for the clearance of chronic ear diseases such as chronic suppurative otitis media or cholesteatoma, for complications of acute mastoiditis and prior to implanting hearing devices such as cochlear implants.[1] Between July 2017 and June 2022, 4356 patients underwent mastoidectomy in Australia).[2] Recognizing intraoperative landmarks such as the mastoid air cells, sigmoid sinus, tegmen, facial nerve, and ossicular chain is critical to avoiding complications of surgery such as hearing loss, vertigo, facial paralysis, or CSF leak.[1,3-5] As a result, accurate identification of these landmarks is a crucial requirement for radiologists and otolaryngologists.[3,6]

Corresponding author: Zubair Hasan, e-mail: zubairhasan112@gmail.com
Received: January 22, 2023 • **Accepted:** May 18, 2023 • **Publication Date:** September 29, 2023
Available online at www.advancedotology.org

Petrous temporal bone (PTB) cone-beam computed tomography (CBCT) is performed to aid pre-operative diagnosis and identification of landmarks with 262 364 PTB CBCT scans performed in Australia between July 2017 and June 2022.[2] Conventional interpretation of CT imaging is subject to several human factors such as clinical experience, fatigue, and time pressures.[7] Artificial intelligence (AI) may augment pre-operative identification by automated labeling and identification of important anatomical structures and landmarks on CT imaging in a consistent and identifiable manner and be utilized as a valuable teaching instrument.

There is an expanding interest in utilizing deep learning in radiological analysis given the data-centric nature of radiologic imaging. Deep learning may facilitate identification of these structures and has previously been utilized for structure and region identification on chest x-rays.[8] Convolutional neural networks (CNNs) is a promising methodological technique which can be applied to a range of computer vision tasks. Recent advances in hardware and software in the past 2 to 3 decades have expanded the role of CNNs which have been applied increasingly to real-world computer-vision problems, including in medical imaging.[9] The utility of deep learning in temporal bone imaging to date is sparse, although applications in other domains of radiology such as chest x-rays (CXR) and CT are more well developed.[10-14] Radiology studies in CNN typically approach image analysis as a classification-based problem,[6] although object detection by bounding boxes is also a well-established technique, popularized in facial detection and self-driving cars.[15,16] In object detection by bounding boxes, a training data set "trains" the algorithm to localize pre-labeled named objects, to which a class label is attached, within the overall image. Subsequently, a test set is presented and the algorithm is challenged to localize the object in question in previously "unseen" images.[17] As opposed to the binary approach of classification-based computer vision tasks, bounding boxes have increased complexity. They seek to not only identify if a structure is present in an image but also where in the image it is located.[18] Bounding boxes have been used in radiology to identify anatomy and pathology on CXR[14] and have proposed uses elsewhere such as identifying brain tumors/edema on MRI and identifying fractures on pelvic and wrist x-rays.[8,19,20]

Structure identification on temporal bone imaging is an area of paucity in the AI literature[21] and hence may be a good test of the ability of CNNs in identifying fine anatomical structures within a confined space. The aim of this study is to determine the accuracy of using a deep-learning CNN algorithm to identify critical structures in temporal bone CT imaging through object detection with bounding boxes. A secondary aim is to compare the accuracy of the CNN in identifying these structures when trained separately by an otolaryngology registrar as compared to a board-certified otolaryngologist.

MATERIAL AND METHODS

Institutional ethics approval was obtained from the Western Sydney Local Health District Ethics Committe (2021/PID03049), and the study was conducted in accordance with the principles of the Declaration of Helsinki. Verbal informed consent was obtained from the patients who agreed to take part in the study.

De-identified retrospective PTB CBCT scans were obtained from a large Australian public tertiary hospital radiology Picture Archiving and Communication System (PACS), with a total of 129 scans extracted. Petrous temporal bone cone-beam computed tomography scans were obtained from a scanner with capability for fine slice image acquisition (0.3 mm). Axial and coronal slices at the level of the malleoincudal joint were chosen for the purposes of this study, as these are the views an otolaryngologist would look through pre-operatively to identify important landmarks. Exclusion criteria were previous temporal bone surgery based upon radiological appearance. Clinical information was not collected.

Images were divided into training and test sets. Azure's Custom Vision (Microsoft Corporation, Redmond, Washington, USA) platform was utilized to draw bounding boxes around the relevant structures using the test set where the structures were present. Custom Vision allows application of machine learning (ML) algorithms to perform classification tasks as well as identification tasks using bounding boxes on custom datasets as small as 50 images.

Bounding box mode was utilized on 68 axial and 66 coronal training images. Images were interpreted in Joint Photographic Experts Group (JPEG) format in their original resolution on high-resolution computer monitors with bounding boxes placed around structures by a board-certified otolaryngologist (ZH) and independently by an otolaryngology registrar (FC). Important temporal bone structures of interest to an operating surgeon were chosen for identification. On axial imaging, they were mastoid air cells, sigmoid sinus, internal acoustic meatus, facial nerve, ossicles, cochlea, and the vestibule. On coronal imaging, they were mastoid air cells, internal acoustic meatus, facial nerve, cochlea, the vestibule, carotid canal, ossicles, semi-circular canal, and tegmen. Training images were uploaded to the platform and subsequently the algorithm was trained on this image set using the 1-hour training mode.

A second unseen test set composed of 61 axial and 61 coronal sequences was uploaded to the Custom Vision platform and the trained algorithm was tested in automated structure identification on the training performed by both the otolaryngology registrar and board-certified otolaryngologist (test 2). Algorithm performance in correctly identifying anatomical structures of interest was recorded. Responses were recorded, including the number of instances the structure of interest was identified manually in the test set (ground truth) and the number of instances the algorithm correctly identified the structure of interest when trained on the training set labeled by the otolaryngology registrar (test 1) and otolaryngologist (test 2).

Statistical analysis was performed using MedCalc 2011 (Ostend, Belgium). Sensitivity and specificity were calculated with 95% confidence intervals. Receiver-operating characteristic (ROC) curves were generated with area under the curve calculations and 95% confidence intervals based on the methodology described by DeLong et al.[22]

RESULTS

The final dataset included 129 axial images and 127 coronal images. In 2 CT series, an adequate coronal slice where all the necessary structures were visible could not be identified. The axial images were split into 68 training images and 61 test images and the coronal images were divided into 66 training images and 61 test images.

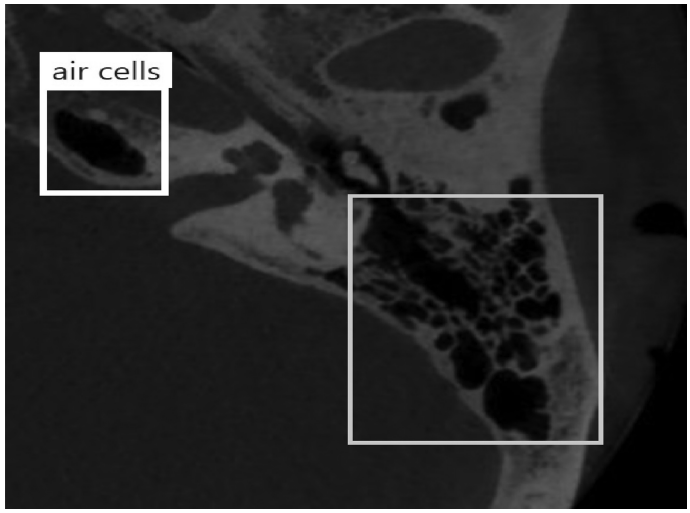

Figure 1. Pneumatized petrous apex misclassified as mastoid air cells.

Axial Imaging

Mastoid Air Cells

Mastoid air cells were present in all 61 test images. The mastoid air cells were identified reliably on the test set in 100% (61/61) of images when the dataset was trained based on labeling both by the otolaryngology registrar (train 1) as well as the otolaryngologist (train 2). One image had a pneumatized mastoid apex, and this was also labeled by the algorithm as "mastoid air cells" (Figure 1).

Sigmoid Sinus

The sigmoid sinus was present in 57 test images. In train 1, it was identified accurately in 55/57 (96.49%) images and in all 57 (100%) images in train 2.

Internal Acoustic Meatus

The internal acoustic meatus was present in 60 test images. In train 1, it was identified accurately in 53/60 (88.33%) images and in all 60 (100%) images in train 2.

Facial Nerve

The facial nerve was present in 52 test images. In train 1, it was identified accurately in 39/52 (75%) images and in 51/52 (98.08%) images in train 2.

Cochlea

The cochlea was present in 54 test images. In train 1, it was identified accurately in 33/54 (61.11%) images and in 53/54 (98.15%) images in train 2.

Vestibule

The vestibule was present in 55 test images. In train 1, it was identified accurately in 23/54 (41.82%) images and in 53/54 (96.36%) images in train 2.

Total Structures Identified

A total of 400 structures were present in the test set (see Table 1 – ground truth) and a total of 27 structures were not present across the test set images, producing a total of 427 identification tasks for the algorithm (Figure 2). Receiver operating characteristic curve was generated with area under the curve of 0.851 (95% CI, 0.813-0.883) when the algorithm was trained by the registrar compared to 0.958 (95% CI, 0.934-0.975) when trained by an ENT surgeon (see Figure 3). Sensitivity was 81.25 (95% CI, 77.1-85.0) and specificity was 88.89 (95% CI, 70.8-97.6) when trained by the registrar and 99.00 (95% CI, 97.5-99.7), and 92.59 (95% CI, 75.7-99.1), respectively, when trained by the surgeon.

Coronal Imaging

Mastoid Air Cells

The mastoid air cells were present in 61 test images. In train 1, it was identified accurately in all 61 (100%) images in train 1 and train 2.

Internal Acoustic Meatus

The internal acoustic meatus was present in 61 test images. In train 1, it was identified accurately in all 61 (100%) images in train 1 and train 2.

Facial Nerve

The facial nerve was present in 61 test images. In train 1, it was identified accurately in all 61 (100%) images in train 1 and train 2.

Cochlea

The cochlea was present in 61 test images. In train 1, it was identified accurately in all 61 (100%) images in train 1 and train 2.

Vestibule

The vestibule was present in 61 test images. In train 1, it was identified accurately in all 61 (100%) images in train 1 and train 2.

Carotid Canal

The carotid canal was present in 60 test images. In train 1 it was identified accurately in 52/60 (86.67%) images and in 55/60 (91.67%) images in train 2.

Table 1. Ground Truth and Percentage Accuracy for Anatomical Landmarks in Axial CT

	Anatomical Structure						
	Mastoid Air Cells	Sigmoid Sinus	IAM	Facial Nerve	Ossicles	Cochlea	Vestibule
Ground truth	61	57	60	52	61	54	55
ENT registrar	61	55	53	39	61	33	23
CNN train 1 (%)	100.00	96.49	88.33	75.00	100.00	61.11	41.82
ENT surgeon	61	57	60	51	61	53	53
CNN train 2 (%)	100.00	100.00	100.00	98.08	100.00	98.15	96.36

CNN, convolutional neural network.

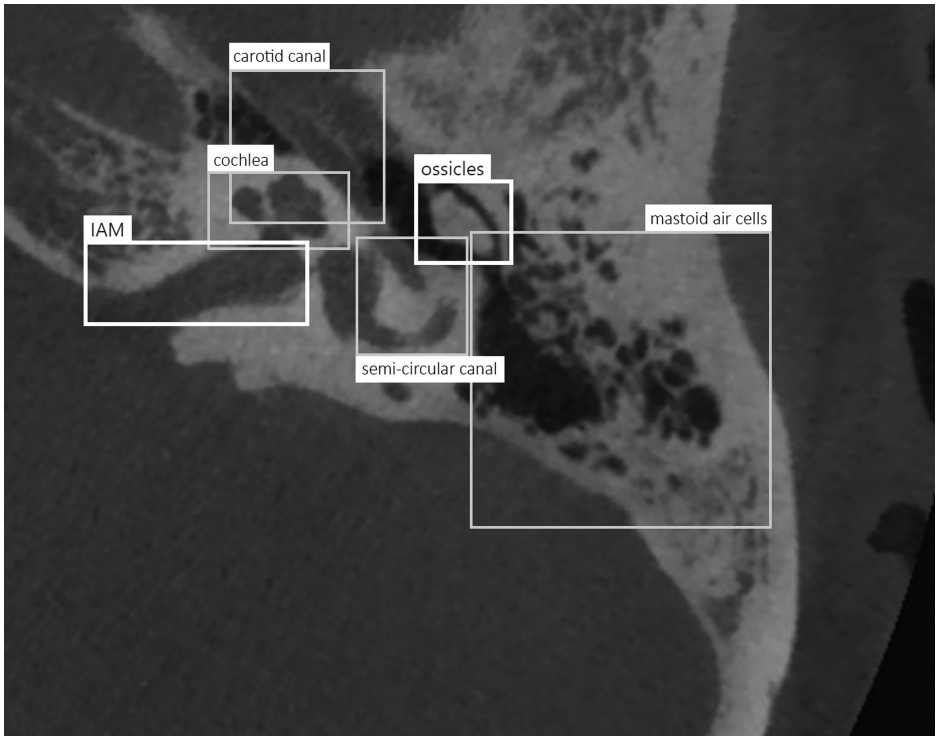

Figure 2. Labeled structures on axial CT imaging. CT, computed tomography.

Ossicles
The ossicles were present in 61 test images. In train 1, it was identified accurately in all 61 (100%) images in train 1 and train 2.

Semi-circular Canal
The semi-circular canal was present in 61 test images. In train 1, it was identified accurately in all 61 (100%) images in train 1 and train 2.

Tegmen
The tegmen was present in 61 test images. In train 1, it was identified accurately in all 61 (100%) images in train 1 and train 2.

Total Structures Identified
A total of 548 structures were present in the test set (see Table 2 – ground truth) and 1 structure (carotid canal not clearly seen on 1

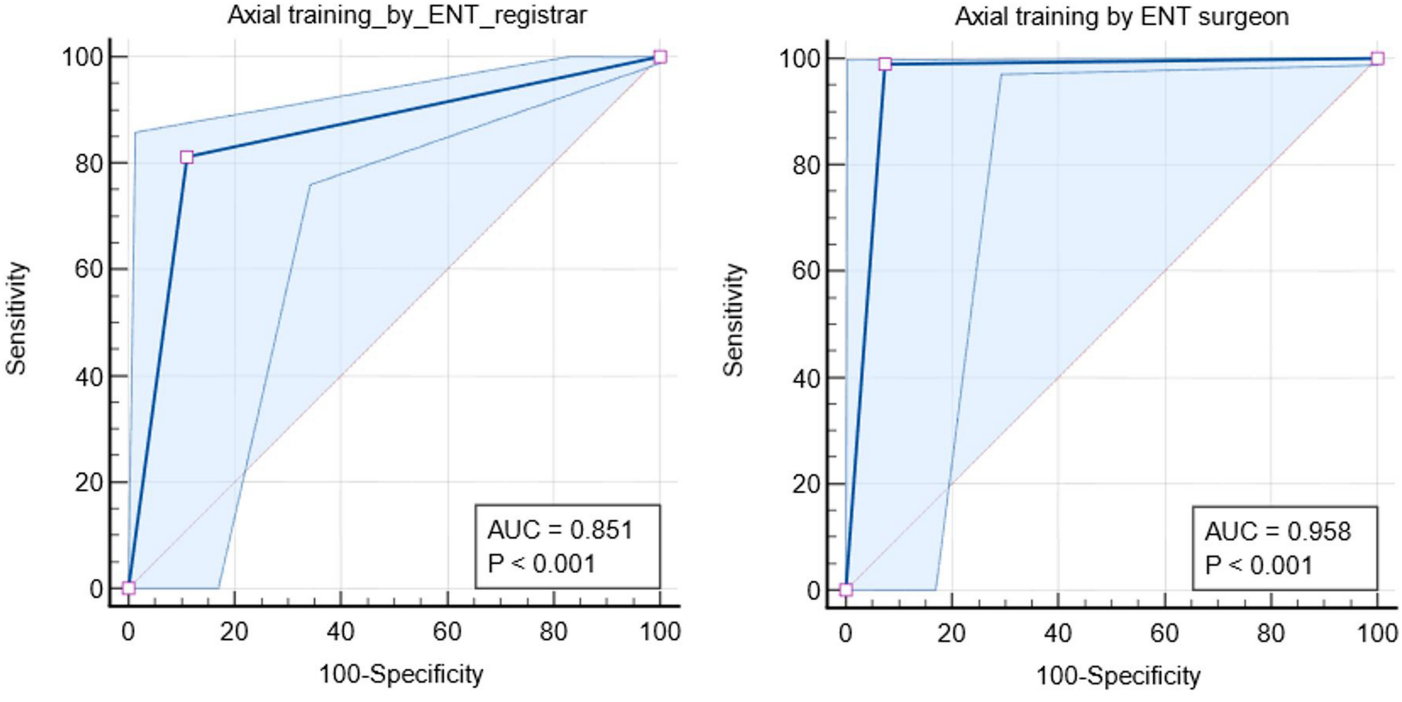

Figure 3. Receiver operator curve (ROC) for axial CT training by ENT registrar (left) and ENT surgeon (right). CT, computed tomography.

Table 2. Ground Truth and Percentage Accuracy for Anatomical Landmarks in Coronal CT

	Anatomical Structure								
	Mastoid Air Cells	**Internal Acoustic Meatus**	**Facial Nerve**	**Cochlea**	**Vestibule**	**Carotid Canal**	**Ossicles**	**Semi-circular Canal**	**Tegmen**
Ground truth	61	61	61	61	61	60	61	61	61
ENT registrar	61	61	61	61	58	52	61	61	61
Train 1 (%)	100.00	100.00	100.00	100.00	95.08	86.67	100.00	100.00	100.00
ENT surgeon	61	61	61	61	61	55	61	61	61
Train 2 (%)	100.00	100.00	100.00	100.00	100.00	91.67	100.00	100.00	100.00

CT, computed tomography.

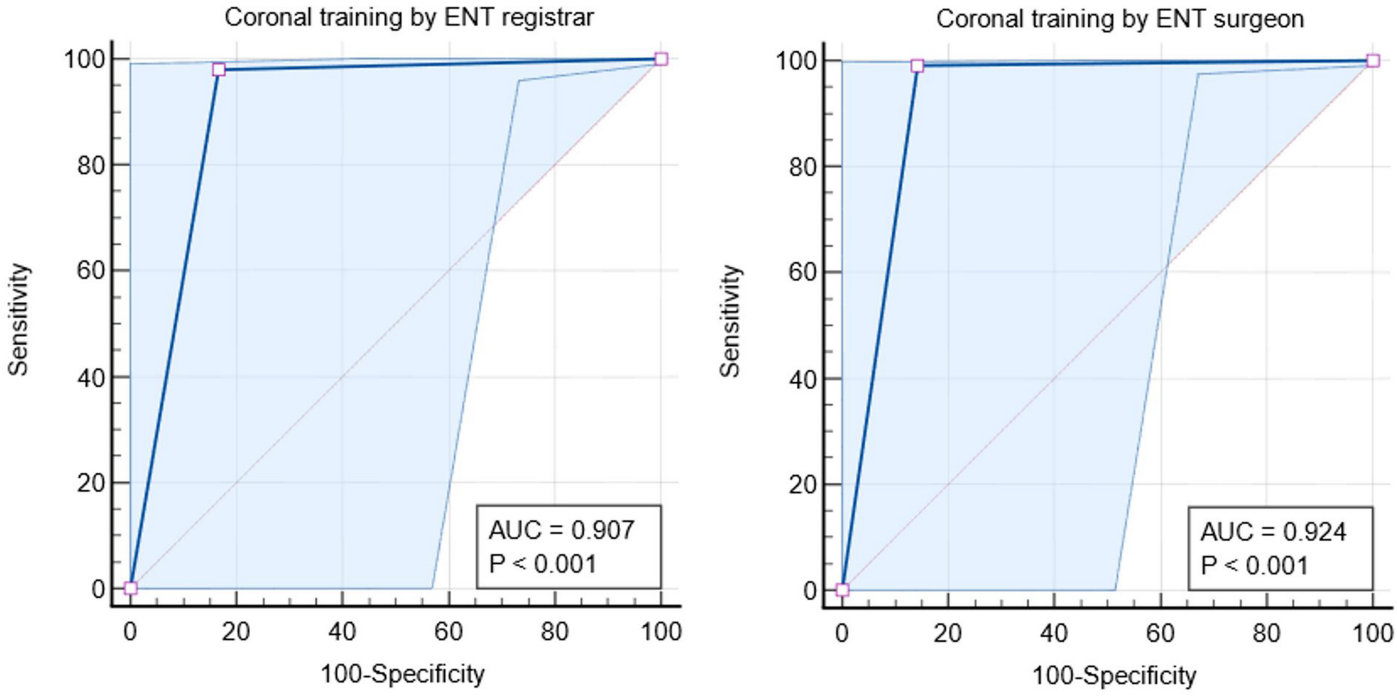

Figure 4. Receiver operator curve (ROC) for coronal CT training by ENT registrar (left) and ENT surgeon (right). CT, computed tomography.

image) was absent across the test set, producing a total of 549 identification tasks for the algorithm. Receiver operating characteristic curve was generated with area under the curve of 0.907 (95% CI, 0.879-0.930) when the algorithm was trained by the registrar compared to 0.924 (95% CI, 0.899-0.945) when trained by an ENT surgeon (see Figure 4). Sensitivity was 97.99 (95% CI, 96.4-99.0) and specificity 83.33 (95% CI, 35.9-99.6), respectively, when trained by the registrar; sensitivity was 99.09 (95% CI, 97.9-99.7) and specificity was 85.71 (95% CI, 42.1-99.6) when trained by the surgeon.

DISCUSSION

This study demonstrates that CNNs can perform structure identification in temporal bone imaging with a high degree of accuracy with a sample size as little as 60 images in both axial and coronal planes. Receiver operating characteristic curves achieved up to 0.958 on axial and 0.924 on coronal in the test 2 testing set, indicating excellent diagnostic accuracy of the testing model, with statistically significant *P* value of <.001 for both curves. Sensitivity and specificity were similarly high for both axial and coronal test sets.

The CNN was trained on structures in this study which are of relevance to the operating surgeon and are paramount to safe mastoid surgery. Simple structures such as the mastoid air cells and ossicles were identified with a high degree of accuracy. In one instance, a pneumatized petrous apex was incorrectly identified by the algorithm as mastoid air cells (Figure 1). This reassuringly indicates that the algorithm is identifying cells based on their appearance rather than other features such as location or by coincidence. The ossicles were also identified consistently by the algorithm, namely at the level of the malleoincudal joint. Fine structures such as the tympanic portion of the facial nerve (Figure 5) as well as the superior semi-circular canal (Figure 6) are also identified with a high degree of accuracy including in coronal slices where the structure may only be a few pixels wide. Although it is unclear whether it is the size, shape, or position of the structure which allows the algorithm to identify the structure, it is postulated that this may be a combination of these factors as observed in Figure 5 where the facial nerve is correctly identified based potentially on its ovoid shape, position, and density, whereas other similar structures such as turns of the cochlea are

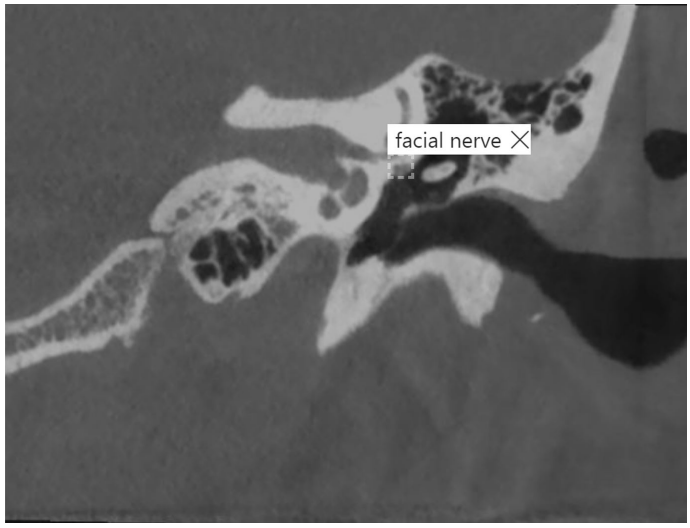

Figure 5. Tympanic segment of facial nerve identification by CNN model in test set. CNN, convolutional neural network.

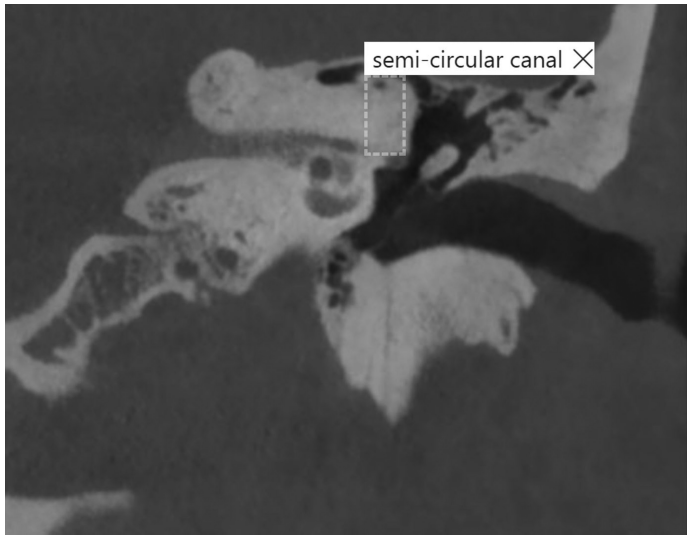

Figure 6. Lateral semi-circular canal identification by CNN model in test set. CNN, convolutional neural network.

correctly not assigned the same label. Position may be an additional factor as observed in Figure 6 where an incomplete superior canal limb is still correctly identified based on its expected position. One common concern surrounding CNNs is that the method of deriving conclusions is a "black box," wherein the convolutions prior to the final output are not able to be examined.[23] Class attribution maps have been utilized in other studies for example in lung nodule detection which can increase confidence in how the structure is being identified as belonging to the correct class.[24] Although beyond the scope of the current study, this technique poses an avenue of future research in temporal bone imaging and may increase confidence that the algorithm is identifying structures correctly rather than by chance, promoting confidence in future clinical applications.

Additionally, CNN training is improved proportionately to the expertise of the training clinician. The algorithm's performance was weaker when trained by a less senior clinician (train 1), in particular with features that may be difficult to identify clearly on single axial CT images such as the cochlea or the vestibule on axial imaging and the carotid canal and vestibule in the skull base on coronal imaging. When trained by a board-certified otolaryngologist, the accuracy of CNNs in structure identification of temporal bone imaging is high, approaching 100% for most structures. In our study, structure identification in coronal CT was 100% in 6 of 7 structures for train 2 (ENT consultant), compared to 100% in 5 of 7 structures in train 1. Similarly, although overall accuracy was lower in axial CT for most structures, there was a higher accuracy for train 2.

The greatest difficulty in automated structure identification was with the carotid canal especially where the dome of the canal was not well defined and identification was not performed by the algorithm or where other canals in the base of the skull such as the hypoglossal canal were incorrectly labeled as the carotid canal. In a few instances, the canal was more widely labeled than the structure was present and although correctly labeled the precision of labeling was not as tight as when trained by the less senior clinician (n = 52/60 images, 86.67%), compared to a more senior clinician (n = 55/60, 91.67%).

One limitation of this study is the limited number of frames per patient incorporated into the training and test sets. In our model, only one axial and coronal image from each patient wasutilized, forming a 2D sample. In clinical practice, CT scans may be scrutinized through several slices to form a 3D image for the observer. Training the algorithm on entire CT sequences to pick out relevant slices and label structures and pathology is a difficult task for the CNN model used. With future developments in CNN technology, the ultimate aim would be to produce this 3D-based CNN model, which is most relevant to clinical practice. This study explored the use of a pre-existing CNN for classifying image datasets, which limited our ability to introduce a validation set for hyperparameter optimization. Future studies using customizable CNN technology could introduce a validation set to potentially further optimize the accuracy of structure identification.

Additionally, the malleoincudal joint was used to select the frames for training and testing the CNN model, as it is the level at which most of these structures are present. It is unclear how the algorithm would perform with anatomical variation such as ossicular abnormalities or whether finer anatomical detail can be discriminated such as identifying individual ossicles. Temporal bone images with post-operative change as well as significant pathology (other than mastoids with effusions or sclerosis) were not included as a uniform test and training set was required for this index CNN on temporal bone structure identification. Future studies may be able to further investigate how this model performs when images with pathology or post-operative changes are included.

Another limitation of structure identification by bounding box is the requirement to constrain 3D and irregular nature of structures into rectangular boxes. While automated segmentation of structures can be performed with ML, these processes are semi-automated, time-consuming, and may require manual refinement. Convolutional neural networks segmentation has been performed for larger, more gross anatomical structures such as whole-body adipose tissue[25] and complicated anatomical structures in the maxillofacial complex.[26] Segmentation utilizing CNNs of fine temporal bone anatomy is worthy of investigation and may optimize temporal bone surgery.

Clinically relevant usage of AI is increasingly prominent in the medical sphere. Of AI applications approved for clinical use by the FDA, radiology-based applications are enjoying the highest quantity of approved software devices.[27] Areas of utility in the broader medical landscape include radiological identification, augmentation of surgical instrumentation, and real-time identification of intraoperative structures.[27] Although wider usage in otolaryngology is still developing,[21] image-navigated systems in particular is an avenue AI may augment.

The use of CNNs can also be used to increase efficiency, and therefore, better utilization of the workforce in radiology practice.[28] Similarly, preparative planning of mastoid surgery such as cochlear implantation requires cognitive load and time for identification of relevant structures for planning of the surgical dissection. With the use of CNNs, this cognitive load can be reduced and efficiency improved, allowing the surgeon to focus on other tasks. The use of AI may have the capacity to enhance medical education,[29] and in this setting, the use of AI can be introduced in the education of trainee otolaryngologists and trainee radiologists as well as non-specialty services in the identification of important temporal bone structures, either in isolation or in relation to a disease process. In the future, the use of AI temporal bone imaging in identifying pathology such as ossicular chain erosion of cholesteatoma, especially when combined with AI technology used for video otoscopes,[30] may render itself a useful tool in teleradiology and telehealth medicine. Given the accuracy of identifying structures such as the facial nerve, CNNs may help in differentiating pathological dilemmas on temporal bone CT, for example, differentiating an early-stage glomus tympanicum from a facial nerve schwannoma at the tympanic segment on coronal slices.[31]

Robotic mastoidectomy is increasing in popularity. Phantom and cadaver studies have been performed with the view of preparing the mastoid bed for more advanced otologic surgery or for more direct access to the round window with image-guided techniques.[32,33] Currently, preliminary investigations in using CNNs for mastoidectomy in recognizing surgeon movements and intraoperative landmarks have been successful.[17,34] The authors anticipate that such applications will only increase as technology and expertise becomes more widespread.

Artificial intelligence is an exciting technology that has the potential to significantly change clinical practice including in structure identification in otologic radiology. This study demonstrates CNNs can have a high degree of accuracy at identifying structures, with the performance correlating with the accuracy of the data labeling during the training phase. This suggests the seniority of the training clinician is paramount to ensuring the accuracy of the results. Further investigation would be desirable to investigate the accuracy of identification of more complex structures such as individual ossicles as well as in other radiological planes.

Data Sharing: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics Committee Approval: Institutional ethics approval was obtained from the Western Sydney Local Health District Ethics Committee (Approval No: 2021/PID03049).

Informed Consent: Verbal informed consent was obtained from the patients who agreed to take part in the study.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – Z.H.; Design – Z.H., F.C.; Supervision – R.S., N.S.; Resources – Z.H., F.C.; Materials – Z.H., F.C.; Data Collection and/or Processing – Z.H., F.C.; Analysis and/or Interpretation – Z.H., S.K., M.L.; Literature Search – N/A; Writing – Z.H., S.K., M.L., L.A., A.E.; Critical Review – Z.H., S.K., M.L.

Declaration of Interests: The authors have no conflict of interest to declare.

Funding: The authors declared that this study has received no financial support.

REFERENCES

1. Prakash B, Aggarwal N, Babu A, Sandhya D, Amulya T. A study on surgical implications and variations of suprameatal spine and other landmarks on the lateral surface of temporal bone. *Indian J Otolaryngol Head Neck Surg*. 2021:1-6.
2. ServicesAustralia. *Medicare ITEM Reports*; 2022. Available at: http://medicarestatistics.humanservices.gov.au/statistics/mbs_item.jsp. Updated October 25, 2022.
3. George AP, De R. Review of temporal bone dissection teaching: how it was, is and will be. *J Laryngol Otol*. 2010;124(2):119-125. [CrossRef]
4. Harkness P, Brown P, Fowler S, Grant H, Ryan R, Topham J. Mastoidectomy audit: results of the Royal College of Surgeons of England comparative audit of ENT surgery. *Clin Otolaryngol Allied Sci*. 1995;20(1):89-94. [CrossRef]
5. Migirov L, Eyal A, Kronenberg J. Intracranial complications following mastoidectomy. *Pediatr Neurosurg*. 2004;40(5):226-229. [CrossRef]
6. Ahuja AT, Yuen HY, Wong KT, Yue V, Van Hasselt AC. Computed tomography imaging of the temporal bone—normal anatomy. *Clin Radiol*. 2003;58(9):681-686. [CrossRef]
7. Brady AP. Error and discrepancy in radiology: inevitable or avoidable? *Insights Imaging*. 2017;8(1):171-182. [CrossRef]
8. Wu J, Gur Y, Karargyris A, et al., eds. Automatic bounding box annotation of chest x-ray data for localization of abnormalities 17th international symposium on biomedical imaging (ISBI). IEEE; 2020. [CrossRef]
9. Dey S. *Python Image Processing Cookbook: Over 60 Recipes to Help You Perform Complex Image Processing and Computer Vision Tasks with Ease*. Packt Publishing; 2020.
10. Suzuki K, Abe H, MacMahon H, Doi K. Image-processing technique for suppressing ribs in chest radiographs by means of massive training artificial neural network (MTANN). *IEEE Trans Med Imaging*. 2006;25(4):406-416. [CrossRef]
11. Suzuki K, Armato III SG, Li F, Sone S, Doi K. Massive training artificial neural network (MTANN) for reduction of false positives in computerized detection of lung nodules in low-dose computed tomography. *Med Phys*. 2003;30(7):1602-1617. [CrossRef]
12. Suzuki K, Doi K. How can a massive training artificial neural network (mtann) be trained with a small number of cases in the distinction between nodules and vessels in thoracic ct? 1. *Acad Radiol*. 2005; 12(10):1333-1341. [CrossRef]
13. Suzuki K, Li F, Sone S, Doi K. Computer-aided diagnostic scheme for distinction between benign and malignant nodules in thoracic low-dose CT by use of massive training artificial neural network. *IEEE Trans Med Imaging*. 2005;24(9):1138-1150. [CrossRef]
14. Milam ME, Koo CW. The current status and future of FDA-approved artificial intelligence tools in chest radiology in the United States. *Clin Radiol*. 2023;78(2):115-122. [CrossRef]
15. Farfade SS, Saberian MJ, Li L-J, eds. Multi-view face detection using deep convolutional neural networks. In: Proceedings of the 5th ACM on

International Conference on Multimedia Retrieval; 2015:643-650. [CrossRef]
16. Simhambhatla R, Okiah K, Kuchkula S, Slater R. Self-driving cars: evaluation of deep learning techniques for object detection in different driving conditions. *SMU Data Sci Rev*. 2019;2(1):23.
17. Choi J, Cho S, Chung JW, Kim N. Video recognition of simple mastoidectomy using convolutional neural networks: detection and segmentation of surgical tools and anatomical regions. *Comput Methods Programs Biomed*. 2021;208:106251. [CrossRef]
18. Szegedy C, Toshev A, Erhan D. Deep neural networks for object detection. *Adv Neural Inf Process Syst*. 2013;26.
19. Krogue JD, Cheng KV, Hwang KM, et al. Automatic hip fracture identification and functional subclassification with deep learning. *Radiol Artif Intell*. 2020;2(2):e190023. [CrossRef]
20. Thian YL, Li Y, Jagmohan P, Sia D, Chan VEY, Tan RT. Convolutional neural networks for automated fracture detection and localization on wrist radiographs. *Radiol Artif Intell*. 2019;1(1):e180001. [CrossRef]
21. Hasan Z, Key S, Habib AR, et al. Convolutional neural networks in ENT radiology: systematic review of the literature. *Ann Otol Rhinol Laryngol*. 2023:132(4):417-430. [CrossRef]
22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-845. [CrossRef]
23. Sheu YH. Illuminating the black box: interpreting deep neural network models for psychiatric research. *Front Psychiatry*. 2020;11:551299. [CrossRef]
24. Venugopal VK, Vaidhya K, Murugavel M, et al. Unboxing AI-radiological insights into a deep neural network for lung nodule characterization. *Acad Radiol*. 2020;27(1):88-95. [CrossRef]
25. Küstner T, Hepp T, Fischer M, et al. Fully automated and standardized segmentation of adipose tissue compartments via deep learning in 3D whole-body MRI of epidemiologic cohort studies. *Radiol Artif Intell*. 2020;2(6):e200010. [CrossRef]
26. Preda F, Morgan N, Van Gerven A, et al. Deep convolutional neural network-based automated segmentation of the maxillofacial complex from cone-beam computed tomography:A validation study. *J Dent*. 2022; 124:104238. [CrossRef]
27. Benjamens S, Dhunnoo P, Meskó B. The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. *npj Digit Med*. 2020;3(1):118. [CrossRef]
28. van Leeuwen KG, de Rooij M, Schalekamp S, van Ginneken B, Rutten MJ. How does artificial intelligence in radiology improve efficiency and health outcomes? *Pediatr Rad*. 2021:1-7.
29. Duong MT, Rauschecker AM, Rudie JD, et al. Artificial intelligence for precision education in radiology. *Br J Radiol*. 2019;92(1103):20190389. [CrossRef]
30. You E, Lin V, Mijovic T, Eskander A, Crowson MG. Artificial intelligence applications in otology: a state of the art review. *Otolaryngol Head Neck Surg*. 2020;163(6):1123-1133. [CrossRef]
31. Wiggins RH, Harnsberger HR, Salzman KL, Shelton C, Kertesz TR, Glastonbury CM. The many faces of facial nerve schwannoma. *AJNR Am J Neuroradiol*. 2006;27(3):694-699.
32. Yoo MH, Lee HS, Yang CJ, et al. A cadaver study of mastoidectomy using an image-guided human–robot collaborative control system. *Laryngoscope Investig Otolaryngol*. 2017;2(5):208-214. [CrossRef]
33. Zagzoog N, Yang VXD. State of robotic mastoidectomy: literature review. *World Neurosurg*. 2018;116:347-351. [CrossRef]
34. Korte C, Schaffner G, McGhan CLR. A preliminary investigation into the feasibility of semi-autonomous surgical path planning for a mastoidectomy using LSTM-recurrent neural networks. *J Med Devices*. 2021;15(1). [CrossRef]