The Journal of International Advanced Otology

J Int Adv Otol 2023; 19(3): 217–222 • DOI: 10.5152/iao.2023.22958

Original Article

A Computer Vision Algorithm to Classify Pneumatization of the Mastoid Process on Temporal Bone Computed Tomography Scans

Zubair Hasan[1,2], Michael Lee[2], Fiona Chen[3], Seraphina Key[4], Al-Rahim Habib[1,2], Layal Aweidah[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]Department of Otolaryngology Royal Children's Hospital Melbourne, Victoria, Australia
[4]Monash University, Faculty of Medicine, Nursing and Health Sciences, Victoria, Australia
[5]University of Notre Dame, Faculty of Medicine, New South Wales, 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, M.L. 0000-0003-4117-0404, F.C. 0000-0002-9169-2617, S.K. 0000-0002-1019-5986, A.-R.H. 0000-0001-6327-4648, L.A. 0000-0002-8989-6602, R.S. 0000-0002-9260-5175, N.S. 0000-0002-7719-1832.

Cite this article as: Hasan Z, Lee M, Chen F, et al. A computer vision algorithm to classify pneumatization of the mastoid process on temporal bone computed tomography scans. *J Int Adv Otol.* 2023;19(3):217-222.

BACKGROUND: Pneumatization of the mastoid process is variable and of significance to the operative surgeon. Surgical approaches to the temporal bone require an understanding of pneumatization and its implications for surgical access. This study aims to determine the feasibility of using deep learning convolutional neural network algorithms to classify pneumatization of the mastoid process.

METHODS: De-identified petrous temporal bone images were acquired from a tertiary hospital radiology picture archiving and communication system. A binary classification mode in the pretrained convolutional neural network was used to investigate the utility of convolutional neural networks in temporal bone imaging. False positive and negative images were reanalyzed by the investigators and qualitatively assessed to consider reasons for inaccuracy.

RESULTS: The overall accuracy of the model was 0.954. At a probability threshold of 65%, the sensitivity of the model was 0.860 (95% CI 0.783-0.934) and the specificity was 0.989 (95% CI 0.960-0.999). The positive predictive value was 0.973 (95% CI 0.904-0.993) and the negative predictive value was 0.935 (95% CI 0.901-0.965). The false positive rate was 0.006. The F1 number was 0.926 demonstrating a high accuracy for the model.

CONCLUSION: The temporal bone is a complex anatomical region of interest to otolaryngologists. Surgical planning requires high-resolution computed tomography scans, the interpretation of which can be augmented with machine learning. This initial study demonstrates the feasibility of utilizing machine learning algorithms to discriminate anatomical variation with a high degree of accuracy. It is hoped this will lead to further investigation regarding more complex anatomical structures in the temporal bone.

KEYWORDS: Otolaryngology, radiology, artificial intelligence, algorithms, machine learning

INTRODUCTION

The mastoid process makes up part of the temporal bone, a pyramidal shaped bone of the skull base. One of the hallmarks of mastoid anatomy is variable pneumatization, the development of which progresses from the time of birth to around age 10.[1] Surgical approaches to the temporal bone require an understanding of the significance of this pneumatization and its implications for surgical access.[2,3] Degree of pneumatization may also be predictive of surgical success in middle ear surgery.[4] Pneumatization is variable and while the mastoid antrum is uniformly present from 22 weeks of gestation, the remaining mastoid and temporal bone pneumatizes variably in 3 phases to an average adult volume of 8 mL.[5,6] In addition to the implications for surgical ease of access to the temporal bone, the degree of pneumatization may also have implications on disease processes such as otitis media, mastoiditis, and cholesteatoma through its role in buffering the middle ear and mastoid air system. Sade and Fuchs[7,8] found that pneumatization

Corresponding author: Zubair Hasan, e-mail: zhas3802@uni.sydney.edu.au
Received: October 08, 2022 • **Accepted:** March 14, 2023 • **Publication Date:** May 31, 2023
Available online at www.advancedotology.org

Content of this journal is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

less than 6 cm^3 confers an 18 times higher risk of secretory otitis media and subsequent chronic otitis media compared to patients with normal volumes of pneumatization. The alternate view is an environmental theory arguing that the disease processes in early development lead to the arrest of pneumatization.[5]

Due to the importance of mastoid pneumatization, the degree of pneumatization is routinely reported in computed tomography (CT) scans of the temporal bone for diagnosis, prognosis, and surgical planning. The reporting process can be augmented with machine learning (ML) algorithms. Machine learning is a discipline of artificial intelligence that uses computers to make analyses based on pre-programmed training, usually using structured data. Deep learning is a subset of ML, where complex analyses are made often on unstructured data such as images and videos using multi-layered "neural network" architectures which are modeled on the human visual cortex and brain.[9]

While early ML models have been utilized for decades, advances in both hardware and software have allowed machine and deep learning to permeate into our technology and way of life.[10] In computer vision, convolutional neural networks (CNNs) use iterative "kernels" or filters, a matrix of numbers that aim to detect features of an image such as edges and lines, to analyze a second matrix, the image itself, and are therefore well suited to the analysis of radiological images.[10]

Arguably, deep learning is part of the future of radiology,[11] and already, several applications of algorithms to body parts such as the lung and prostate,[12-16] have been described. Even in the head and neck, several studies in thyroid and head and neck oncology have utilized artificial intelligence to classify and identify radiological features such as on paranasal sinus CT scans[17,18] and thyroid ultrasound.[19-24] Deep learning algorithms have yet to be applied to discriminate mastoid pneumatization or other anatomical variants in the temporal bones thus far in the reported literature. Determining anatomical variation in the temporal bone is a complex task even for specialist otolaryngologists and head and neck radiologists. This study aims to determine the feasibility of using deep learning CNN algorithms to classify pneumatization of the mastoid process.

METHODS

Institutional ethics approval was obtained from the Western Sydney Local Health District Ethics Committee (2021PID03049) and the study was conducted in line with the principles of the Declaration of Helsinki.

De-identified retrospective petrous temporal bone cone-beam computed tomography (CBCT) images were obtained at random from a tertiary hospital radiology picture archiving and communication system. Temporal bone images were obtained from a CBCT with the capability for fine slice image acquisition (0.3 mm). Axial slices at the level of the antrum, at the approximate level of the incudomalleal joint, in a bone window were chosen for this study. Scans were excluded if the patient had undergone previous temporal bone surgery.

Digital Imaging and Communications in Medicine (DICOM) images were exported in JPEG format and divided in half to extract left and right temporal bones from each patient's CT scan. Where the scan did not align at the exact level of the incudomalleal joint on both sides, the level which best approximated this joint on at least 1 side was taken. The dataset was further augmented by flipping images left to right and vice versa, as described by Bloice et al.[25] to create 4 images from each patient scan.

A board-certified otolaryngologist (ZH) and an otolaryngology resident (FC) independently divided the extracted images into a binary classification (pneumatized or sclerosed). The images were considered pneumatized if there was a well-developed air cell system around the mastoid antrum and sclerosed if there were minimal cells around the mastoid antrum. Further review was undertaken by a board-certified head and neck radiologist with agreement by consensus (ground truth). Images were divided into training and test sets in a 2:1 ratio, maintaining the proportion of sclerosed and pneumatized images. The distribution of partially and fully opacified mastoids was analyzed between the 2 groups using the chi-square test for homogeneity using MedCalc Statistical Software© (v20.022; Ostend, Belgium).

Microsoft Azure's Custom Vision (Redmond, Washington, USA) platform was utilized for the purposes of training and testing the algorithm. Custom Vision is based on a pretrained CNN and uses transfer learning to perform image analysis on an unseen dataset. Custom Vision allows the application of ML algorithms to perform binary and multiclass classification tasks on custom datasets as small as 50 images. For this experiment, the binary classification mode was utilized. Training images were uploaded to the platform and subsequently, the algorithm was trained on this image set using the 1-hour training mode.

Performance metrics included precision, recall, and mean average precision based on the initial training set. The test set was subsequently manually uploaded onto the Custom Vision platform to test accuracy following training and class label probability recorded in MedCalc. The probability threshold >65% was chosen to represent positivity for the purposes of this study. This allowed the calculation of performance metrics including sensitivity, specificity, positive and negative predictive values, and F1 score. A receiver operating characteristic curve was generated using MedCalc software using the methodology of DeLong et al.[26] to model the accuracy of the model. False positive and negative images were reanalyzed by the investigators and qualitatively assessed to consider reasons for inaccuracy.

RESULTS

A total of 800 images were obtained from 200 patients. Sixteen images were excluded due to previous mastoidectomy. The final dataset contained 260 sclerosed temporal bones and 524 pneumatized temporal bones. After division into sclerosed versus pneumatized image sets, these sets were further divided randomly in a 2:1 ratio into training and test sets such that the sclerosed set included 174 training images and 86 test images and the pneumatized set included 348 training images and 176 test images (Tables 1 and 2). The chi-square test was applied to the training and test sets to compare the homogeneity of partially and fully opacified mastoids in both the pneumatized and sclerosed cohorts. For the training set the chi-square statistic was 43.729 with a *P*-value < .00001 and for the test set the chi-square statistic was 29.212 with a *P*-value of <.00001 indicating non-homogeneity between the 2 sets.

Table 1. Distribution of Dataset into Training and Test Sets

	Training (%)	Test (%)	Total
Pneumatized	348 (66.4)	176 (33.6)	524
Sclerosed	174 (66.9)	86 (33.1)	260
Total	522 (66.5)	262 (33.5)	784

Table 2. Confusion Matrix Demonstrating Performance of the Model at Probability Threshold of 65%

Mastoid Sclerosis		Ground Truth		
		Sclerosed (%)	Pneumatized (%)	Total
Algorithm	Sclerosed	74 (97.4)	2 (2.6)	76
	Pneumatized	12 (6.5)	174 (93.5)	186
	Total	86	176	262

The overall accuracy of the model was 0.954. At a probability threshold of 65%, the sensitivity of the model was 0.860 (95% CI 0.783-0.934) and the specificity was 0.989 (95% CI 0.960-0.999). The positive predictive value was 0.973 (95% CI 0.904-0.993) and the negative predictive value was 0.935 (95% CI 0.901-0.965). The false positive rate was 0.006. The F1 number was 0.926 demonstrating a high accuracy for the model.

The receiver operator characteristic (ROC) curve with 95% CI is demonstrated in Figure 1, with the area under the curve of 0.968 (95% CI 0.939-0.986; SE 0.0132) indicating a very good fit of the model for distinguishing sclerosed from pneumatized mastoid processes. This curve demonstrates the variance of sensitivity and specificity as the probability threshold is adjusted.

Pneumatized Test Set

One hundred seventy-six images that were labeled "pneumatized" (ground truth) were tested following the initial training phase (example in Figure 2). Seven of these images contained partially opacified mastoids, and none of the images contained fully opacified mastoids (Table 3). Of these images, 174 were correctly identified by the model (true positives). Only 2 images were identified incorrectly (false negatives) by the algorithm as sclerosed in this set. These images were reviewed by the investigators. The reason for misclassification is likely the presence of sparse periantral cells which may have led the algorithm to classify these images as sclerosed (Figure 3). Of the 2 false negative images in the pneumatized test set, 1 mastoid was partially opacified in the lateral epitympanum (see Figure 3).

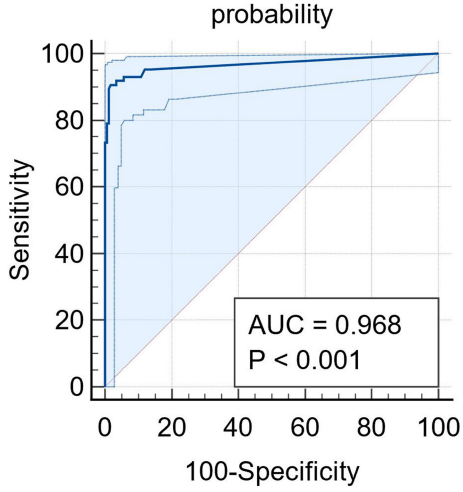

Figure 1. ROC curve for mastoid pneumatization model. ROC, receiver operator characteristic.

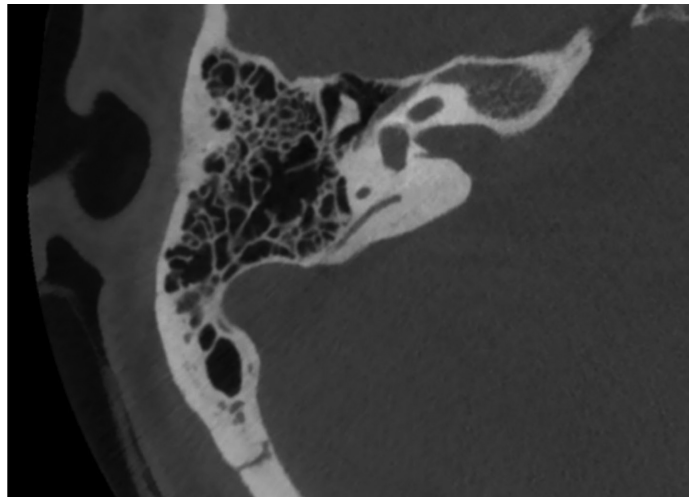

Figure 2. Example of an image from the pneumatized test set demonstrating a well-developed mastoid antrum and numerous and well aerated periantral cells.

Sclerosed Test Set

Eighty-six images that were labeled "sclerosed" (ground truth) were tested following the initial training phase (example in Figure 4). Twelve of these images contained partially opacified mastoids, and 10 of the images contained fully opacified mastoids (Table 3). Of these images, 12 images (false negatives) were categorized pneumatized by the algorithm and the remaining 74 were correctly categorized by the model (true positives). On review of false negative images, they were largely sclerosed but did have the presence of some periantral cells similar to the misclassification of the 2 images from the pneumatized test set (Figure 5). Of the 12 false negative images in the sclerosed test set 4 mastoids were partially opacified and none were fully opacified.

DISCUSSION

This is the first study utilizing deep learning techniques to classify pneumatized and sclerosed mastoids. The study demonstrates that computer vision has a high degree of accuracy (95.4%) in performing this binary classification task, with a moderate sample size of 784 images. Microsoft estimates that for most classification tasks

Table 3. Distribution of Partially and Fully Opacified Mastoids in Training and Test Sets

	Pneumatized		Sclerosed	
	Training	Test	Training	Test
Partially opacified	9/348 (2.6%)	7/176 (4.0%)	13/174 (7.5%)	12/86 (14.0%)
Fully opacified	3/348 (0.9%)	0/176 (0%)	22/174 (12.6%)	10/86 (11.6%)
Non-opacified	336/348 (96.5%)	169/176 (96.0%)	139/174 (79.9%)	64/86 (74.4%)

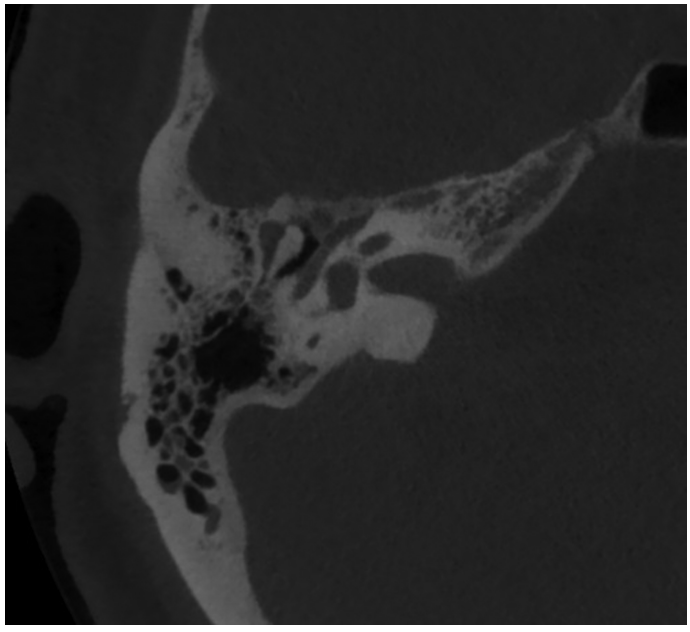

Figure 3. Temporal bone image from pneumatized test set incorrectly labeled by the algorithm as sclerosed.

using the Custom Vision platform a sample size of 50 is adequate. However, algorithms do risk perpetuating biases.[27] Researchers and developers need to consider where and how data is collected to augment, validate and test an algorithm, to prevent reinforcing biases. Larger datasets are more generalizable and avoid the problem of overfitting where the algorithm learns features intrinsic to the training dataset rather than features that may be more generalizable for real-world examples.[28]

The overall accuracy of the algorithm was 95.4%. This is consistent with other studies which have utilized ML algorithms for classification tasks in the head and neck region such as concha bullosa (83%) (18), anterior ethmoidal artery (81%)[17] and osteomeatal complex on paranasal sinus CT (85%).[29] The specificity of the algorithm was higher than the sensitivity indicating the algorithm had greater certainty in the identification of pneumatized mastoids than those

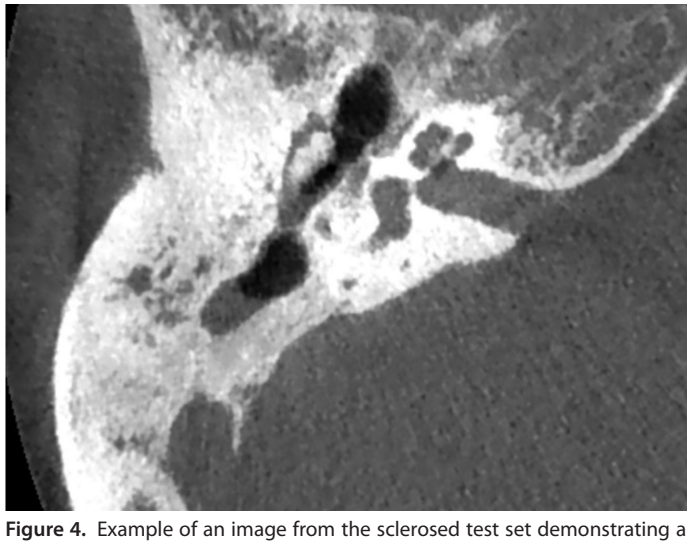

Figure 4. Example of an image from the sclerosed test set demonstrating a contracted mastoid antrum and sparse periantral cells.

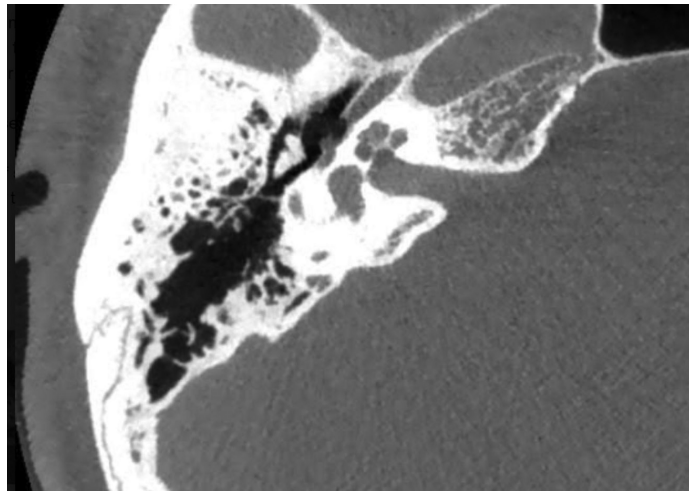

Figure 5. Temporal bone image from sclerosed test set incorrectly labeled by the algorithm as pneumatized.

which are sclerosed. The likely reason for this is that although binary classification was performed, mastoid pneumatization, and sclerosis are not entirely binary categories. The mastoid antrum is always developed to a varying degree and the presence of periantral cells also demonstrates a wide degree of variability. Mastoids with a varying degree of pneumatization were more likely to be misclassified, with the false negatives from the sclerosed test set being mastoids with a large antrum and sparse periantral cells, whereas those correctly identified had a smaller antrum and minimal to no periantral cells. This was the basis for clinicians' categorization as sclerosed whereas the algorithm may have classified according to the overall volume of non-bone density in the mastoid process.

The difficulty in developing robust classification schemes for mastoid pneumatization is well reflected in the literature and a myriad of classification schemes are present, many ill-defined. These classification schemes include descriptions of the various cell tracts. One of the earliest descriptions of mastoid pneumatization was by Allam et al[30] in 1969. They divided pneumatization into 3 regions—mastoid, petrous, and accessory tracts. The mastoid portion is divided by the petrosquamous fissure externally and Koerner's septum internally from the petrous portion. The cell tracts of the petrous portion were further divided into the petrous apex and perilabyrinthine tracts. Finally, accessory tracts included zygomatic, squamous, occipital, and styloid.

In addition to the various tracts described above, the degree of pneumatization has also been variably described. Some authors favor a 3-tiered classification—sclerotic, diploic, and pneumatized, although the precise definitions with respect to the number of air cells present and degree of pneumatization have not been well described. Virapongse et al[5] utilized a 4-tiered system where the number of periantral cells was counted—grades 1 to 4 from <10 cells to >50 cells. This system has not been widely adopted. Han et al[1] utilized a system relating pneumatization and the presence of cells to the anterior, middle, and posterior aspects of the sigmoid sinus. Although these complex descriptions of pneumatization exist, the authors felt a binary classification was appropriate as it is widely used in clinical practice, has the greatest surgical relevance in most procedures on the temporal bone, and is more amenable to an image classification algorithm.

For the purposes of this study, the axial slices analyzed were taken at the incudomalleal joint. The level of this joint has been previously used for pneumatization classification studies[1] as this is the most surgically relevant region of the mastoid being the level of the mastoid antrum and entry to the antrum and thereafter the epitympanum and middle ear in mastoid surgery. Additionally, pathologies such as cholesteatoma often occur here and, accordingly, this level is of surgical relevance. Although other levels of the mastoid were not used in this study, it is anticipated that similar levels of performance would be obtained.

On the basis of our simple binary classification, a high degree of accuracy was obtained demonstrating this model may be suitable for future investigation of more complex anatomical variations including the presence/absence of dehiscence of critical structures, aberrant anatomy of the major vessels, and slope of the tegmen tympani/mastoideum. The investigators intend to use the same model to test the hypothesis that artificial intelligence and computer vision may be harnessed to help surgeons identify important anatomical variants relevant to surgical access to the temporal bone and improve teaching and surgical safety.

When comparing the training and test groups with respect to the degree of opacification, there was a non-homogenous distribution of non-opacified, partially opacified, and fully opacified mastoids as evidenced by the chi-square metric. However, the investigators feel this is unlikely to have confounded the results significantly as both groups did include opacified mastoids and the vast majority of both sets were classified accurately (Table 4). Nevertheless, this is a limitation of the methodology as this could potentially confound the ability of the CNN in correctly differentiating between pneumatized or sclerosed mastoids. In some CNN models, class attribution maps or "heat maps" can be generated to identify regions which the algorithm is "seeing" to perform its classification task.[12] In future papers this may increase confidence that classification is being performed based on the air cells themselves rather than another undefined feature of the image.

Although the algorithm has performed well in our data population, external validation of this algorithm has not yet been tested for specific demographic groups. Temporal bone anatomy can differ with the patient's age, race, and gender. A limitation of our study is that we did not gather demographic information, and therefore it remains unclear if differing demographics would affect the validity of the algorithm. This is also an opportunity for future application, where a refined algorithm could be developed for specific groups (e.g. pediatrics). Additionally, future experiments the investigators seek to conduct will include performing more complex multitask classification, object detection, and segmentation.

Potential applications of this technology include simplification of mundane lower-level tasks for reporting radiologists allowing them to dedicate time to more complex interpretation. Additionally, for surgeons, identification and classification of variation in mastoids including pneumatization, dehiscence of critical structures, aberrant anatomy of major vessels, low-lying tegmen tympani, and mastoideum will improve the teaching of junior surgeons by directing attention to these critical variants. It is possible, if intraoperative navigation takes increasing prominence to augment operative otology, ML, and computer vision can be incorporated to direct the surgeon's attention to these structures and variations, thus improving operative safety.

Table 4. Distribution of Partially and Fully Opacified Mastoids by Identification Accuracy

	Pneumatized		Sclerosed	
	True Positive	False Negative	True Positive	False Negative
Partially opacified	6/174 (3.4%)	1/2 (50%)	8/74 (10.8%)	4/12 (33.3%)
Fully opacified	0/174 (0%)	0/2 (0%)	10/74 (13.5%)	0/12 (0%)
Non-opacified	166/174 (95.4%)	1/2 (50%)	56/74 (75.7%)	8/12 (66.7%)

CONCLUSION

The temporal bone is a complex anatomical region of interest to otolaryngologists, and mastoid pneumatization has significant implications on pathology and surgical access. Surgical planning requires high-resolution CT scans, the interpretation of which can be augmented with ML. This initial study demonstrates the feasibility of utilizing ML algorithms to discriminate temporal bone pneumatization with a high degree of accuracy—a sensitivity of 0.860 and a specificity of 0.989. It is hoped this will lead to further investigation regarding more complex anatomical structures in the temporal bone.

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

Ethics Committee Approval: Ethical committee approval was received from the Western Sydney Local Health District Ethics Committee (2021/PID03049) and the study was conducted in line with the principles of the Declaration of Helsinki.

Informed Consent: Informed consent was obtained from all patients included in this study.

Peer-review: Externally peer-reviewed.

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

Declaration of Interests: The authors declare that they have no competing interest.

Funding: The authors declare that this study had received no financial support.

REFERENCES

1. Han SJ, Song MH, Kim J, Lee WS, Lee HK. Classification of temporal bone pneumatization based on sigmoid sinus using computed tomography. *Clin Radiol*. 2007;62(11):1110-1118. [CrossRef]
2. Woolley AL, Oser AB, Lusk RP, Bahadori RS. Preoperative temporal bone computed tomography scan and its use in evaluating the pediatric cochlear implant candidate. *Laryngoscope*. 1997;107(8):1100-1106. [CrossRef]

3. Lim AE, Hurley R, Slim MAM, Locke R. SCOTLAND preoperative tympanomastoid CT temporal bone mnemonic system. *ENT & Audiol News*. 2020;3(2):1017.
4. Hussl B, Welzl-Mueller K. Secretory otitis media and mastoid pneumatization. *Ann Otol Rhinol Laryngol Suppl*. 1980;89(3 Pt 2):79-82. [CrossRef]
5. Virapongse C, Sarwar M, Bhimani S. Computed tomography of temporal bone pneumatization: 1. Normal pattern and morphology. *Am J Neuroradiol*. 1985;6(4):551-559.
6. Cinamon U. The growth rate and size of the mastoid air cell system and mastoid bone: a review and reference. *Eur Arch Otorhinolaryngol*. 2009;266(6):781-786. [CrossRef]
7. Sadé J, Fuchs C. Secretory otitis media in adults: I. The role of mastoid pneumatizationas a risk factor. *Ann Otol Rhinol Laryngol*. 1996;105(8):643-647. [CrossRef]
8. Sadé J, Fuchs C. Secretory otitis media in adults. 2. The role of mastoid pneumatization as a prognostic factor. *Ann Otol Rhinol Laryngol*. 1997;106(1):37-40. [CrossRef]
9. Fukushima K, Neocognitron MS. *A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition. Competition and Cooperation in Neural Nets*. Springer; Berlin; 1982:267-285.
10. Suzuki K. Overview of deep learning in medical imaging. *Radiol Phys Technol*. 2017;10(3):257-273. [CrossRef]
11. Saba L, Biswas M, Kuppili V, et al. The present and future of deep learning in radiology. *Eur J Radiol*. 2019;114:14-24. [CrossRef]
12. Iino Y, Imamura Y, Hiraishi M, Yabe T, Suzuki J. Mastoid pneumatization in children with congenital cholesteatoma: an aspect of the formation of open-type and closed-type cholesteatoma. *Laryngoscope*. 1998;108(7):1071-1076. [CrossRef]
13. 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]
14. 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]
15. 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]
16. 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]
17. Huang J, Habib AR, Mendis D, et al. An artificial intelligence algorithm that differentiates anterior ethmoidal artery location on sinus computed tomography scans. *J Laryngol Otol*. 2020;134(1):52-55. [CrossRef]
18. Parmar P, Habib AR, Mendis D, et al. An artificial intelligence algorithm that identifies middle turbinate pneumatisation (concha bullosa) on sinus computed tomography scans. *J Laryngol Otol*. 2020;134(4):328-331. [CrossRef]
19. Abdolali F, Kapur J, Jaremko JL, Noga M, Hareendranathan AR, Punithakumar K. Automated thyroid nodule detection from ultrasound imaging using deep convolutional neural networks. *Comput Biol Med*. 2020;122:103871. [CrossRef]
20. Keramidas EG, Maroulis D, Iakovidis DK. TND: a thyroid nodule detection system for analysis of ultrasound images and videos. *J Med Syst*. 2012;36(3):1271-1281. [CrossRef]
21. Li X, Zhang S, Zhang Q, et al. Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images: a retrospective, multicohort, diagnostic study. *Lancet Oncol*. 2019;20(2):193-201. [CrossRef]
22. Liu T, Guo Q, Lian C, et al. Automated detection and classification of thyroid nodules in ultrasound images using clinical-knowledge-guided convolutional neural networks. *Med Image Anal*. 2019;58:101555. [CrossRef]
23. Ma J, Duan S, Zhang Y, et al. Efficient deep learning architecture for detection and recognition of thyroid nodules. *Comp Intell Neurosci*. 2020;2020:1242781. [CrossRef]
24. Ma J, Wu F, Zhu J, Xu D, Kong D. A pre-trained convolutional neural network based method for thyroid nodule diagnosis. *Ultrasonics*. 2017;73:221-230. [CrossRef]
25. Bloice MD, Roth PM, Holzinger A. Biomedical image augmentation using Augmentor. *Bioinformatics*. 2019;35(21):4522-4524. [CrossRef]
26. 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]
27. Nagendran M, Chen Y, Lovejoy CA, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ*. 2020;368:m689. [CrossRef]
28. Ying X, eds. An overview of overfitting and its solutions. *J Phys Conf S*. IOP Publishing. 2019.
29. Chowdhury NI, Smith TL, Chandra RK, Turner JH. Automated classification of osteomeatal complex inflammation on computed tomography using convolutional neural networks. *Int Forum Allergy Rhinol*. 2019;9(1):46-52. [CrossRef]
30. Allam AF. Pneumatization of the temporal bone. *Ann Otol Rhinol Laryngol*. 1969;78(1):49-64. [CrossRef]