

Review

Not peer-reviewed version

Artificial Intelligence in Otitis Media Diagnosis: A Review of Diagnostic Accuracy, Limitations, Clinical Integration and Future Directions

[Sulymon Saka](#) *

Posted Date: 6 October 2025

doi: 10.20944/preprints202510.0453.v1

Keywords: Otitis media; deep learning; artificial intelligence; deep metric learning; otoscopic images; diagnostic accuracy



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Review

Artificial Intelligence in Otitis Media Diagnosis: A Review of Diagnostic Accuracy, Limitations, Clinical Integration and Future Directions

Sulymon A. Saka^{1,2}

¹ Department of Otolaryngology, Irrua Specialist Teaching Hospital, Irrua, Edo State, Nigeria; sakasulymon@gmail.com

² Ambrose Alli University, Ekpoma, Edo State, Nigeria

Abstract

Otitis media is a common pediatric ear infection that, if left undiagnosed or misdiagnosed can lead to complications such as hearing loss. Traditional diagnostic methods rely on subjective clinical assessments which can result in variability in accuracy. The integration of artificial intelligence, particularly deep learning offers a promising approach for automated and objective diagnosis. This study reviews the application of deep learning algorithms in the automatic detection of otitis media using otoscopic images, with a focus on deep metric learning techniques and their diagnostic performance. An evaluation of deep learning-based models including convolutional neural networks (CNNs) and deep metric learning, was conducted. Performance metrics such as sensitivity, specificity, accuracy, and the area under the receiver operating characteristic curve (AUC-ROC) were analyzed. A comparative evaluation of traditional CNN-based models and deep metric learning techniques was done to assess their relative strengths in diagnostic accuracy and generalization to diverse datasets. The advantages of deep metric learning in improving model generalization and robustness were also discussed. AI-driven models demonstrated high accuracy in detecting otitis media, with deep metric learning enhancing feature differentiation and classification performance. Several studies reported sensitivity and specificity values exceeding 90% with AUC-ROC values approaching 1.0, indicating strong diagnostic capability. However, variations in dataset quality, image preprocessing, and model interpretability remain key challenges. Additionally, this review explores the clinical feasibility of AI-based otoscopic analysis by assessing the integration of AI into routine otolaryngological practice. Deep learning, particularly deep metric learning holds significant potential for enhancing the automated diagnosis of otitis media in pediatric patients. Future research should focus on dataset standardization, model transparency, and real-world clinical validation to ensure widespread adoption in healthcare settings.

Keywords: otitis media; deep learning; artificial intelligence; deep metric learning; otoscopic images; diagnostic accuracy

Introduction

Otitis media (OM) is a prevalent pediatric condition, particularly affecting children aged 6 months to 2 years and it is one of the leading causes of healthcare visits and antibiotic prescriptions worldwide [1]. It is classified into acute otitis media (AOM), otitis media with effusion (OME), and chronic suppurative otitis media (CSOM) [2]. The most common bacteria implicated include *Streptococcus pneumoniae*, followed by *Haemophilus influenzae* and *Moraxella catarrhalis*. However, following the introduction of the conjugate pneumococcal vaccines, the pneumococcal organisms have evolved to non-vaccine serotypes. The most common viral pathogens implicated in otitis media are the respiratory syncytial virus (RSV), coronaviruses, influenza viruses, adenoviruses, human metapneumovirus, and picornaviruses [1,2]. Epidemiological variations in OM incidence,

driven by socioeconomic status, healthcare access and environmental exposure further complicate early detection and intervention strategies. Failure to diagnose OM accurately can lead to hearing impairment, speech delays, tympanic membrane (TM) perforation, mastoiditis, labyrinthitis, petrositis, meningitis, brain abscess, hearing loss, lateral and cavernous sinus thrombosis [1]. It is one of the most common infections in early childhood, with studies indicating that approximately 85% of children experience at least one episode by the age of three [3]. The incidence of OM varies globally, with higher rates observed in Low-and Middle-Income Countries (LMICs) due to factors such as poor hygiene, inadequate access to healthcare, malnutrition, and recurrent upper respiratory tract infections [1,2,4]. Environmental and genetic factors, including exposure to tobacco smoke, daycare attendance, and anatomical variations in the eustachian tube, also contribute to susceptibility. Other risk factors include ciliary dysfunction, cochlear implants, vitamin A deficiency, allergies and lack of breastfeeding [1]. Despite these known risk factors, a significant gap remains in predictive modelling for early OM identification which artificial intelligence (AI) could help bridge.

Symptoms of OM include but not limited to ear pain, fever, hearing loss, ear fullness, ear discharge, ringing in the ear (tinnitus) and sometimes balance disturbances [1]. Traditional diagnostic methods of otitis media (OM) rely on clinical assessment using otoscopy, pneumatic otoscopy, tympanometry and acoustic reflectometry. The gold standard for diagnosing OM includes the combination of clinical symptoms and otoscopic findings. However, these techniques often require specialized training and experience, limiting their accessibility in primary healthcare settings. Otoscope is the primary tool for assessing the tympanic membrane (TM), looking out for bulging, erythema, opacity or loss of mobility of the TM. In cases of Otitis Media with Effusion (OME), there may be retracted, dull, cloudy TM or bubbles behind the TM (air-fluid level) and in some cases of OME, amber or greyish fluid may be seen behind the TM. Pneumatic otoscope is used to assess the TM mobility, helping to distinguish acute otitis media (AOM) from OME. Tympanometry is an objective test used to measure middle ear pressure and tympanic membrane compliance, it is useful for diagnosing OME while acoustic reflectometry is a non-invasive method that detects middle ear effusion based on sound reflection [1,5]. However, these approaches are subjective and depend on the examiner's expertise, leading to significant inter-observer variability and frequent misdiagnoses. Studies have shown that general practitioners often misdiagnose OM in up to 50% of cases, resulting in either unnecessary antibiotic use or delayed treatment [6,7].

Recent advancements in artificial intelligence (AI) have introduced automated diagnostic tools that improve accuracy and reduce subjectivity [5,8–10]. Deep metric learning (DML) is a specialized deep learning technique and has shown promise in medical imaging by learning feature representations that improve differentiation between similar and dissimilar images. Unlike conventional classification models, DML focuses on similarity-based learning which is particularly useful in detecting OM given the subtle variations in otoscopic images [11]. Additionally, AI-based models have demonstrated potential in training medical professionals through automated diagnostic feedback, further enhancing their clinical applicability.

This review aims to provide an in-depth analysis of AI-based otoscopic image analysis, with a particular focus on deep metric learning as a novel approach for automating the detection of OM. It explores conventional deep learning techniques, discusses the advantages of DML, and highlights key challenges and future directions in the field. Furthermore, this study assesses the feasibility of integrating AI models into real-world clinical workflows including consideration for regulatory approval and physician acceptance.

Artificial Intelligence in Otitis Media (OM) Diagnosis: Current Approaches and Limitations

The integration of AI into otologic diagnosis has primarily focused on automating image analysis of tympanic membrane (TM) abnormalities. The majority of studies have utilized CNNs, which are well-suited for image classification tasks [9,10,12–27]. However, emerging techniques such as deep metric learning (DML), which learns an embedding space for distinguishing pathological

from normal images have demonstrated superior performance in small and imbalanced datasets [11]. The use of hybrid models combining CNNs and DML has recently been explored, yielding promising results in both sensitivity and specificity.

The most widely adopted AI models for otologic image classification are CNNs. Despite their success, CNNs often struggle with generalization to diverse datasets due to over-fitting, an issue that DML seeks to address through improved feature extraction and similarity-based learning. A study employed a CNN-based deep learning model using CNNs to classify pediatric otitis media (OM) from otoscopic images. The models, based on Xception and mobileNet-V2 architecture were trained on 10,703 images and tested on 1,500 high-quality images and 102 smartphone-captured, the model achieved accuracy of 97.45% for Xception and 95.72% for MobileNet-V2, demonstrating their abilities to classify Acute Otitis Media (AOM), Otitis Media with Effusion (OME) and normal TM. On smartphone-captured images, accuracy remained at 90.66% for Xception and 88.56% for MobileNet-V2, demonstrating that the performance on smartphone images was slightly lower hence the need for robust preprocessing [27]. Similarly, a study analyzed 214 images using CNN-based automatic diagnostic algorithm for pediatric otitis media, reporting a diagnostic accuracy of 91.4%, sensitivity of 89.5% and specificity of 93.3% [28]. These studies underscore the efficacy of CNNs in otoscopic image interpretation. However, CNNs rely on large, well-labeled datasets and often struggle with class imbalance. Another study applied AI for image-based triage, analyzing a total of 6,527 otoscopic images from Aboriginal and Torres Strait Islander children. Their AI model categorized images into normal, abnormal, or non-diagnostic, achieving 99.3% accuracy for acute otitis media, 96.3% for chronic otitis media, 77.8% for otitis media with effusion (OME), and 98.2% to classify wax/obstructed canal. [9]. This study demonstrates the potential of AI to assist in large-scale screening programs, particularly in pediatric and underserved populations. However, challenges such as image quality and model bias need to be addressed to enhance reliability.

A study utilized a two-stage attention-aware convolutional neural network (CNN) to classify otoscopic images of otitis media. Their model, based on ResNet50 demonstrated high accuracy (93.4%) with impressive F1 scores across different conditions—94.3% for normal images, 96.8% for otitis media with effusion and 91.7% for active chronic suppurative otitis media [15]. These findings indicate that AI can achieve near-expert-level performance in image-based diagnosis. However, further validation in diverse populations is needed to confirm the model’s generalizability. Similarly, Noda et al. investigated the potential of GPT-4 Vision (GPT-4V) in classifying middle ear diseases. Their study revealed variable accuracy across conditions, with 89.2% accuracy for acute otitis media, 76.5% for chronic otitis media and 85.7% for otitis media with effusion. While these results suggest that large vision-language models like GPT-4V may be useful for otologic diagnosis, their lower accuracy in chronic conditions highlights the need for fine-tuning and disease-specific optimizations [29]. Another study by Afify et al. optimized deep learning approaches for otoscopic diagnosis using CNN architecture optimised through Bayesian hyperparameter to classify otoscopic images into normal, myringosclerosis, ear wax and chronic otitis media, this model achieved 98.10% accuracy, 98.11% sensitivity, 99.36% specificity and positive predictive value (PPV) of 98.10% [12]. Mohammed et al. further refined CNN-based classification (CNN-LSTM model) using Bayesian optimization, demonstrating an increase in diagnostic accuracy to 100% compared to conventional CNNs. This model analyzed 880 otoscopic images into four categories as stated above and achieved 100% accuracy, 100% sensitivity, 100% specificity and 100% PPV [18]. While this reported perfect performance metrics, it is essential to consider dataset’s size and diversity. Further validation on larger and more varied datasets is necessary to ensure model’s generalizability and robustness in real world clinical settings. Table 1 summarizes studies on AI models for Otitis Media (OM) diagnosis, evaluates dataset size, image population, limitations and performance metrics like sensitivity, specificity, accuracy and AUC.

Table 1. Findings from studies on Otitis Media utilizing AI model: A summary of model performance, dataset size, image population limitations and key metrics.

Study Detail	AI Model	Dataset	Image	Sensitivity	Specificity	Accuracy	AUC-	Limitations
		Size	Population	(%)	(%)	(%)	ROC	
Tran et al., 2018 [26]	CNN-based AI	214	Pediatrics	89.5	93.3	91.4	0.92	Limited generalizability; potential bias; Small dataset; Lack of external validation; Limited scope of classification; dependence on image quality; no LMIC representation
Cai et al., 2021 [12]	Two-stage attention CNN	6,066	Mixed	NR	NR	93.4	0.98	Single- centre data collection; no LMIC representation; Lack of pediatric-specific preprocessing techniques
Afify et al., 2024 [9]	CNN architecture optimised through Bayesian hyperparameter	880	NR	98.10	98.11	99.36	NR	Homogeneous dataset; risk of overfitting; no external validation; Limited dataset size and diversity; single centre data collection; lack of external validation; dependence on image quality
Wang et al., 2022 [7]	SCAD supervised deep	(Self- 100	Pediatrics	83.3	80.0	81.7	0.88	Single centre data collection; Extremely small dataset; limited dataset diversity; potential overfitting; absence of external validation; dependence on

learning)								image quality; no comparison to clinician performance.
[11]	DML	1,336	Pediatrics	NR	NR	86.0	NR	Limited dataset size; potential selection bias; lack of external validation; dependence on image quality; Single-rater labels (potential bias); no LMIC data; limited subtype differentiation
Tsutsumi et al., 2021 [21]	CNN (MobileNetV2-based model)	400	Mixed	70.0	84.4	77.0	0.902	Limited dataset size and diversity; use of publicly available images; lack of external validation; dependence on image quality; potential overfitting; poor multiclass accuracy (66%); lacks pediatric-specific analysis
Byun et al., 2021 [10]	CNN (ResNet)	2272	Mixed	NR	NR	97.2	NR	Single-centre data collection; Limited dataset diversity, potential overfitting; lack of external validation; dependence on image quality; dataset bias towards adult populations; no real-world clinical validation
Fang et al., 2024 [28]	CNN	1,137	Mixed	80.0	96.0	94.0	NR	Limited dataset diversity potential selection bias (higher proportion of normal images); single-centre data collection; absence of real-world clinical validation; limited pediatric data; Limited to telemedicine use cases
Habib et al., 2022 [27]	ML	6,527	Pediatrics	NA	NA	91.1	0.997	Focus on triage (normal/abnormal) rather than OM subtypes; dataset



									restricted to Australian indigenous children; no external validation	
Khan et al., 2020 [14]	CNN			2,484	Mixed	95.0	NR	87.0	0.99	Lack pediatric-specific validation; no ethical consideration for data use; no external validation; limited generalizabilty; lack of real-world clinical testing
Mohammed et al., 2022 [16]	CNN-LSTM			880	Mixed	100.0	100.0	100.0	NR	Unrealistically high metrics suggesting overfitting; small, non-diverse dataset; no external validation; limited generalizabilty; lack of external validation; lack of pediatric-specific validation; no ethical consideration for data use
Noda et al., 2024 [29]	GPT-4 vision (AI model)			190	Mixed	NR	NR	82.1	NR	Very small dataset; lower accuracy for chronic OM subtypes; no pediatric-specific analysis; potential bias in training; absence of real-world clinical testing; dependence on quality images
Sandström et al., 2022 [18]	CNN			273	Mixed	93.0	100.0	90.0	NR	Limited dataset size; no external validation, potential overfitting; limited consideration of clinical variability; ethical consideration
Viscaino et al., 2020 [23]	SVM (ML)			880	Mixed	87.8	95.9	93.9	1.00	Traditional ML (non-deep learning) approach; lacks generalizability to complex OM subtypes; lack of external validation; limited dataset and



diversity; dependence on image quality; clinical integration challenges									
Viscaino et al., 2022 [24]	CNN (single green channel model)	22,000	Mixed	85.0	95.0	92.0	NR	Focus on colour-channel dependency; no pediatric subgroup analysis; unclear clinical utility; lack of external validation; single-centre data collection	
Wu et al., 2021 [25]	CNN (Xception)	12,203	Pediatrics	NR	NR	97.5	NR	Limited diagnostic scope, single centre data collection; dependence on image quality; lack of external validation; lower performance on smartphone-captured images; no LMIC validation	
Wu et al., 2021 [25]	CNN (MobileNet-V2)	12,203	Pediatrics	NR	NR	95.7	NR	Limited diagnostic scope, single centre data collection; dependence on image quality; lack of external validation; lower performance on smartphone-captured images; no LMIC validation	

*NR - Not Reported; SCAD- Shift contrastive anomaly detection; Mixed - Both adults and pediatrics; NA-Not accessible; SVM - Support Vector Machine (SVM); CNN - Convolutional Neural Network. Although, these advancements have improved AI performance, CNNs still lack the ability to generalize well across varied clinical settings. Unlike traditional CNNs, which classify images based on learned hierarchical features, deep metric learning (DML) optimizes a distance function that directly maps similar images closer together while pushing dissimilar images apart. This approach is particularly advantageous in low-data environments, where limited annotated images are available. Sundgaard et al. explores the use of deep metric learning (DML) to classify OM from 1,336 otoscopic images categorized into three diagnostic groups (AOM, OME and No Effusion). This DML techniques included the contrastive loss, triplet loss and multi-class N-pair, compared their performance to traditional cross-entropy and class-weighted cross-entropy methods. Among these, triplet loss emerged as the most effective, achieving a high accuracy of 85% on a highly imbalanced dataset. This demonstrates its ability to capture the subtle variations in otoscopic images, even when some classes like AOM are underrepresented. This study also noted that recall may be lower with DML, as this model classifies images based on their proximity to cluster centres and if recall is more important for a specific application, cross-entropy loss might be a better choice. Also, noted that poor images from pediatric patients may hinder accurate diagnosis as high quality images could not be captured due to crying and moving during examination hence developing a technique to enhance image quality will be of importance and incorporating multimodal data, that is combining images with other data such as patient history or tympanometry results will help improve diagnostic accuracy. Author also acknowledged that the ground truth labels used for training and evaluation were provided by a single experienced ENT specialist, which introduces potential subjectivity and bias [11]. To ensure reliability, multiple ENT specialist should provide diagnoses for the same cases.

Comparative Analysis of Diagnostic Performance

Despite these limitations, deep metric learning (DML) demonstrates superiority over conventional Convolutional Neural Networks (CNNs) by addressing key challenges such as dataset bias, generalization, and classification accuracy. Traditional CNNs rely heavily on large, diverse datasets for training and tend to struggle with overfitting when the dataset is limited or lacks diversity. They also depend on predefined classification categories, which can be restrictive when dealing with complex medical conditions that exhibit subtle variations. DML, on the other hand, enhances the learning process by focusing on feature similarity rather than absolute classification. Unlike CNNs, which map inputs directly to predefined labels, DML employs techniques such as contrastive loss and triplet loss to learn meaningful feature representations. This allows the model to group similar cases together and better differentiate between visually similar but clinically distinct conditions. In the context of Sundgaard et al., this hybrid approach improved classification accuracy in pediatric datasets by mitigating dataset size limitations and selection bias, two major challenges faced by standalone CNN models. Another significant advantage of DML over CNNs is its robustness to variations in image quality and patient demographics [11]. Many CNN-based models in the Table 1, such as those by Mohammed et al. and Wang et al. exhibited poor generalizability due to their dependence on high-quality, homogenous datasets. In contrast, DML improves model performance by learning from relative distances between cases rather than absolute pixel values, making it more adaptable to real-world clinical variability. Also, DML achieved 85% accuracy in imbalanced dataset while CNNs achieved 77-93% accuracy (Table 1) [7,11,18].

Furthermore, AI models, particularly those using deep learning and deep metric learning, have consistently achieved high diagnostic accuracy, often surpassing the performance of general practitioners (GPs) and even some ENT specialists. For example, Wu et al. developed a deep learning model that achieved an overall accuracy of 93.7% in classifying pediatric OM, with sensitivity and specificity exceeding 93% for all three diagnostic categories (AOM, OME, and NE) [27]. Similarly, Sundgaard et al. reported an accuracy of 85% using deep metric learning, which was comparable to the performance of expert clinicians. In contrast, studies have shown that the diagnostic accuracy of GPs and ENTs typically ranges from 50% to 73%, as reported by Pichichero & Poole (2001) [11,31]. Noda et al. reported that GPT-4V (82.1% accuracy) outperformed certified pediatricians who had an accuracy of 70.6% while otolaryngologists achieved an accuracy exceeding 95% surpassing GPT-4V's performance [29]. This significant gap highlights the potential of AI to enhance diagnostic precision, particularly in primary care settings where access to otologists may be limited. Also, this variability is often due to differences in experience, training, and the subjective nature of otoscopic image interpretation. Additionally, physicians often rely on their experience and contextual knowledge to interpret low-quality images, but this can lead to inconsistent diagnoses. For instance, Pichichero & Poole found that even experienced clinicians agreed on the diagnosis in only 64% of AOM cases, highlighting the subjectivity and variability in manual diagnosis [31]. This variability underscores the need for objective diagnostic tools that can provide consistent and reliable results.

One of the key challenges in OM diagnosis is the class imbalance in datasets, where certain conditions (e.g., AOM) are underrepresented compared to others (e.g., OME or NE). AI models, particularly those using deep metric learning (e.g., triplet loss), have shown promise in addressing this issue. For instance, Sundgaard et al. found that deep metric learning improved precision for underrepresented classes like AOM, although recall was slightly lower due to the clustering-based approach [11]. In contrast, traditional diagnostic methods often struggle with class imbalance, leading to misdiagnosis or over-reliance on clinical judgment, which can vary significantly among physicians.

Clinical Integration and Future Direction

The transition of AI-driven otitis media (OM) diagnosis from research to clinical practice demands a multifaceted strategy that addresses technical, ethical, and infrastructural barriers. While deep metric learning (DML) has demonstrated superior performance in pediatric cohorts—achieving 85% accuracy in imbalanced datasets compared to conventional CNNs—its real-world adoption hinges on overcoming dataset limitations and ensuring equitable access. A critical priority is the establishment of globally representative datasets, particularly from low-resource settings where OM prevalence is highest. Collaborative initiatives, such as federated learning frameworks could enable multi-centre data pooling while preserving patient privacy, mitigating the homogeneity bias seen in studies like . [12,18]. To enhance clinical utility, AI models must evolve beyond static image analysis. Multimodal integration, combining otoscopy with tympanometry, patient history, or even audiometric data could address diagnostic ambiguities, such as distinguishing chronic OM subtypes. For instance, hybrid architectures merging DML’s feature similarity with temporal modeling (e.g., LSTM networks) may better capture dynamic disease progression. Furthermore, vision-language models like GPT-4V, though currently limited by variable accuracy, could be refined through clinician-in-the-loop training, where real-time feedback from otolaryngologists fine-tunes AI outputs [29].

Telemedicine platforms present a pragmatic avenue for deployment, particularly in underserved regions. Algorithm-driven devices, such as smartphone-compatible otoscopes, could democratize access but require optimization for low-quality images—a challenge highlighted by Wu et al. [27,30]. Edge computing solutions, which process data locally on devices, may alleviate bandwidth constraints in remote areas while ensuring compliance with data privacy regulations like general data protection regulation (GDPR). Ethical and regulatory frameworks must parallel technical advancements. Transparent AI, incorporating explainability tools such as Grad-CAM heatmaps (Gradient-weighted Class Activation Mapping) is essential to foster clinician trust. Additionally, bias mitigation strategies—reweighting loss functions for underrepresented LMIC pathologies or auditing training data for demographic diversity—are critical to prevent algorithmic disparities. Regulatory bodies should prioritize pediatric-specific validation protocols, ensuring AI tools meet stringent safety standards for children who constitute the majority of OM cases.

Finally, implementation science must bridge the gap between innovation and practice. Training programs for frontline healthcare workers, interoperable electronic health record (EHR) integration and cost-effectiveness analyses in LMIC settings will determine AI’s scalability. By fostering partnerships among AI developers, clinicians, and policymakers, the next generation of OM diagnostics can transcend academic benchmarks, delivering equitable, human-centric care globally.

Conclusions

AI shows great promise in improving otitis media (OM) diagnosis and management, with deep learning models achieving high accuracy in some studies. However, challenges like multiclass classification, generalizability, and image quality must be addressed. Future AI systems must integrate DML, federated learning and clinician-in-the-loop frameworks to democratize OM diagnosis globally.

Acknowledgements: The author acknowledges Professor Joseph Kei for inspiring the idea for this literature review and for his helpful suggestions during its conceptualization.

Author’s contribution: SAS conceived the idea and wrote the whole manuscript.

Funding: None

Data availability: No datasets were analyzed during the current study

Ethical approval and consent to participate: Not applicable

Clinical trial number: Not applicable

Consent for publication: Not applicable

Competing interests: The authors declare no competing interests

References

1. Danishyar A (2023) Acute Otitis Media. In: Acute Otitis Media, Updated 2023 Apr 15. StatPearls Publishing
2. Morris PS, Leach AJ (2009) Acute and Chronic Otitis Media. *Pediatr Clin North Am* 56:1383–1399. <https://doi.org/10.1016/j.pcl.2009.09.007>
3. Gaddey HL et al. (2019) Otitis Media: Rapid Evidence Review. *Am Fam Physician* 100:350–356
4. Mukara KB, Lilford RJ, Tucci DL, Waiswa P (2017) Prevalence of Middle Ear Infections and Associated Risk Factors in Children under 5 Years in Gasabo District of Kigali City, Rwanda. *Int J Pediatr* 2017:1–8. <https://doi.org/10.1155/2017/4280583>
5. Pichichero ME (2021) Can Machine Learning and AI Replace Otoscopy for Diagnosis of Otitis Media? *Pediatrics* 147:e2020049584. <https://doi.org/10.1542/peds.2020-049584>
6. Legros J-M, Hitoto H, Garnier F, Dagorne C, Parot-Schinkel E, Fanello S (2008) Clinical qualitative evaluation of the diagnosis of acute otitis media in general practice. *Int J Pediatr Otorhinolaryngol* 72:23–30. <https://doi.org/10.1016/j.ijporl.2007.09.010>
7. Wang W, Tamhane A, Santos C, Rzasz JR, Clark JH, Canares TL, Unberath M (2022) Pediatric Otoscopy Video Screening With Shift Contrastive Anomaly Detection. *Front Digit Health* 3:810427. <https://doi.org/10.3389/fdgth.2021.810427>
8. Cao Z, Chen F, Graiss EM, Yue F, Cai Y, Swanepoel DW, Zhao F (2023) Machine Learning in Diagnosing Middle Ear Disorders Using Tympanic Membrane Images: A META-ANALYSIS. *The Laryngoscope* 133:732–741. <https://doi.org/10.1002/lary.30291>
9. Habib A-R, Crossland G, Patel H, Wong E, Kong K, Gunasekera H, Richards B, Caffery L, Perry C, Sacks R, Kumar A, Singh N (2022) An Artificial Intelligence Computer-vision Algorithm to Triage Otoscopic Images From Australian Aboriginal and Torres Strait Islander Children. *Otol Neurotol Off Publ Am Otol Soc Am Neurotol Soc Eur Acad Otol Neurotol* 43:481–488. <https://doi.org/10.1097/MAO.0000000000003484>
10. Habib A-R, Kajbafzadeh M, Hasan Z, Wong E, Gunasekera H, Perry C, Sacks R, Kumar A, Singh N (2022) Artificial intelligence to classify ear disease from otoscopy: A systematic review and meta-analysis. *Clin Otolaryngol Off J ENT-UK Off J Neth Soc Oto-Rhino-Laryngol Cervico-Facial Surg* 47:401–413. <https://doi.org/10.1111/coa.13925>
11. Sundgaard JV, Harte J, Bray P, Laugesen S, Kamide Y, Tanaka C, Paulsen RR, Christensen AN (2021) Deep metric learning for otitis media classification. *Med Image Anal* 71:102034. <https://doi.org/10.1016/j.media.2021.102034>
12. Afify HM, Mohammed KK, Hassanien AE (2024) Insight into Automatic Image Diagnosis of Ear Conditions Based on Optimized Deep Learning Approach. *Ann Biomed Eng* 52:865–876. <https://doi.org/10.1007/s10439-023-03422-8>
13. Byun H, Yu S, Oh J, Bae J, Yoon MS, Lee SH, Chung JH, Kim TH (2021) An Assistive Role of a Machine Learning Network in Diagnosis of Middle Ear Diseases. *J Clin Med* 10:3198. <https://doi.org/10.3390/jcm10153198>
14. Byun H, Lee SH, Kim TH, Oh J, Chung JH (2022) Feasibility of the Machine Learning Network to Diagnose Tympanic Membrane Lesions without Coding Experience. *J Pers Med* 12:1855. <https://doi.org/10.3390/jpm12111855>
15. Cai Y, Yu J-G, Chen Y, Liu C, Xiao L, M Graiss E, Zhao F, Lan L, Zeng S, Zeng J, Wu M, Su Y, Li Y, Zheng Y (2021) Investigating the use of a two-stage attention-aware convolutional neural network for the automated diagnosis of otitis media from tympanic membrane images: a prediction model development and validation study. *BMJ Open* 11:e041139. <https://doi.org/10.1136/bmjopen-2020-041139>
16. Khan MA, Kwon S, Choo J, Hong SM, Kang SH, Park I-H, Kim SK, Hong SJ (2020) Automatic detection of tympanic membrane and middle ear infection from oto-endoscopic images via convolutional neural networks. *Neural Netw* 126:384–394. <https://doi.org/10.1016/j.neunet.2020.03.023>

17. Liu Y, Gong Q (2024) Deep Learning Models for Predicting Hearing Thresholds Based on Swept-Tone Stimulus-Frequency Otoacoustic Emissions. *Ear Hear* 45:465–475. <https://doi.org/10.1097/AUD.0000000000001443>
18. Mohammed KK, Hassanien AE, Afify HM (2022) Classification of Ear Imagery Database using Bayesian Optimization based on CNN-LSTM Architecture. *J Digit Imaging* 35:947–961. <https://doi.org/10.1007/s10278-022-00617-8>
19. Rony MdAH, Fatema K, Raiaan MAK, Hassan MdM, Azam S, Karim A, Jonkman M, Beissbarth J, Boer FD, Islam SMS, Leach A (2024) Artificial Intelligence-Driven Advancements in Otitis Media Diagnosis: A Systematic Review. *IEEE Access* 12:99282–99307. <https://doi.org/10.1109/ACCESS.2024.3428700>
20. Sandström J, Myburgh H, Laurent C, Swanepoel DW, Lundberg T (2022) A Machine Learning Approach to Screen for Otitis Media Using Digital Otoscope Images Labelled by an Expert Panel. *Diagnostics* 12:1318. <https://doi.org/10.3390/diagnostics12061318>
21. Taleb A, Leclerc S, Hussein R, Lalande A, Bozorg-Grayeli A (2024) Registration of preoperative temporal bone CT-scan to otoendoscopic video for augmented-reality based on convolutional neural networks. *Eur Arch Oto-Rhino-Laryngol Off J Eur Fed Oto-Rhino-Laryngol Soc EUFOS Affil Ger Soc Oto-Rhino-Laryngol - Head Neck Surg* 281:2921–2930. <https://doi.org/10.1007/s00405-023-08403-0>
22. Tseng CC, Lim V, Jyung RW (2023) Use of artificial intelligence for the diagnosis of cholesteatoma. *Laryngoscope Investig Otolaryngol* 8:201–211. <https://doi.org/10.1002/lio2.1008>
23. Tsutsumi K, Goshtasbi K, Risbud A, Khosravi P, Pang JC, Lin HW, Djalilian HR, Abouzari M (2021) A Web-Based Deep Learning Model for Automated Diagnosis of Otosopic Images. *Otol Neurotol Off Publ Am Otol Soc Am Neurotol Soc Eur Acad Otol Neurotol* 42:e1382–e1388. <https://doi.org/10.1097/MAO.0000000000003210>
24. V AJ, S J, Kovilpillai J JA, G A, K A (2024) Otoscopy Image Classification Using Embedded AI. In: 2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS). IEEE, Bangalore, India, pp 1–6
25. Viscaino M, Maass JC, Delano PH, Torrente M, Stott C, Auat Cheein F (2020) Computer-aided diagnosis of external and middle ear conditions: A machine learning approach. *PLOS ONE* 15:e0229226. <https://doi.org/10.1371/journal.pone.0229226>
26. Viscaino M, Talamilla M, Maass JC, Henríquez P, Délano PH, Auat Cheein C, Auat Cheein F (2022) Color Dependence Analysis in a CNN-Based Computer-Aided Diagnosis System for Middle and External Ear Diseases. *Diagnostics* 12:917. <https://doi.org/10.3390/diagnostics12040917>
27. Wu Z, Lin Z, Li L, Pan H, Chen G, Fu Y, Qiu Q (2021) Deep Learning for Classification of Pediatric Otitis Media. *The Laryngoscope* 131:E2344–E2351. <https://doi.org/10.1002/lary.29302>
28. Tran T-T, Fang T-Y, Pham V-T, Lin C, Wang P-C, Lo M-T (2018) Development of an Automatic Diagnostic Algorithm for Pediatric Otitis Media. *Otol Neurotol Off Publ Am Otol Soc Am Neurotol Soc Eur Acad Otol Neurotol* 39:1060–1065. <https://doi.org/10.1097/MAO.0000000000001897>
29. Noda M, Yoshimura H, Okubo T, Kosu R, Uchiyama Y, Nomura A, Ito M, Takumi Y (2024) Feasibility of Multimodal Artificial Intelligence Using GPT-4 Vision for the Classification of Middle Ear Disease: Qualitative Study and Validation. *Jmir Ai* 3:e58342. <https://doi.org/10.2196/58342>
30. Fang T-Y, Lin T-Y, Shen C-M, Hsu S-Y, Lin S-H, Kuo Y-J, Chen M-H, Yin T-K, Liu C-H, Lo M-T, Wang P-C (2024) Algorithm-Driven Tele-otoscope for Remote Care for Patients With Otitis Media. *Otolaryngol--Head Neck Surg Off J Am Acad Otolaryngol-Head Neck Surg* 170:1590–1597. <https://doi.org/10.1002/ohn.738>
31. Pichichero ME, Poole MD (2001) Assessing Diagnostic Accuracy and Tympanocentesis Skills in the Management of Otitis Media. *Arch Pediatr Adolesc Med* 155:1137. <https://doi.org/10.1001/archpedi.155.10.1137>

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.