Journal of Al-Qadisiyah for Computer Science and Mathematics **Vol.17.(2) 2025,pp.Comp 18–28**

University Of AL-Qadisiyah

Available online at www.qu.edu.iq/journalcm

JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS

ISSN:2521-3504(online) ISSN:2074-0204(print)

Artificial Intelligence-Based Diagnostic Methods for Otitis Media: A Review Paper

Mohammed Mahmoud Hussein[a], Salwa Khalid Abdulateef[b,*], Khalid Khalis Ibrahim[c]

[a,b,c]Computer Science Department, Computer Science and Mathematics College, Tikrit University, Tikrit, Iraq.

E-mail: mohammed.m.hussein4233@st.tu.edu.iq, khalid.salwa@tu.edu.iq, khalid.kh.ibrahim@tu.edu.iq

ARTICLE INFO

Article history:
Received: 9 /3/2025
Rrevised form: 3 /4/2025
Accepted : 20 /4/2025
Available online: 30 /6/2025

Keywords:

Otitis Media,

Artificial Intelligence,

Deep Learning,

Machine Learning,

Diagnostic Accuracy,

Computer-Aided Diagnosis,

ABSTRACT

Otitis media (OM) is a common inflammatory condition, particularly in children, and poses significant diagnostic and treatment challenges. It also significantly affects adults, especially those with atopic conditions. OM can lead to complications such as hearing loss and speech delays. The advent of electronic health records, big data, and artificial intelligence (AI) offers transformative opportunities in OM diagnosis. AI plays a crucial role in enabling early detection and enhancing diagnostic precision through advanced image analysis and predictive modeling. This review paper explores modern techniques used for the diagnosis of OM, with an emphasis on both traditional and machine learning approaches. A wide range of studies has been evaluated, demonstrating the application of AI in improving diagnostic accuracy and treatment planning. Notably, AI approaches—particularly deep neural networks—have shown remarkable success in otoscopy image analysis. Additionally, recently developed hybrid models that combine multiple techniques have outperformed individual approaches. Despite these advancements, challenges remain, including limited dataset standardization and issues with image quality.

https://doi.org/10.29304/jqcsm.2025.17.22174

1. Introduction

Otitis Media (OM) is one of the most common inflammatory diseases worldwide, particularly affecting children. Before the age of seven, nearly every child has at least one bout of an ear infection [1]. Nowadays, doctors usually prescribe antibiotics after performing a visual examination. In many cases, inappropriate treatment can lead to serious bacterial infections [2]. As a result, developing and investigating cutting-edge techniques is essential for accurate diagnosis.

Unfortunately, physicians often misdiagnose OM due to subtle symptoms, the young age of patients, and limited knowledge about the patient and family history. A delayed diagnosis of OM can lead to hearing loss, speech difficulties, and serious complications [3]. Studies show that widespread antibiotic use for OM is concerning, as it offers limited improvement while contributing to antibiotic resistance [4]. AI involves developing systems that

*Corresponding author: Salwa khalid Abdulateef

Email addresses: khalid.salwa@tu.edu.iq.

Communicated by 'sub etitor'

replicate human intelligence, enabling automation of tasks that usually require human cognition. Contemporary technology is unable to equal or exceed general human intelligence. The AI has exhibited the capability to accomplish clearly defined subtasks independently, without requiring external (human) assistance. AI, combined with e-health records, has the potential to transform the care of OM patients. Using extensive long-term data, deep learning models can detect underlying diseases [5], classify disease types, forecast risks [6], analyze medical images [5], detect anatomical features [6], and assist in localization. These capabilities can improve disease management, support surgical planning, and personalize treatment strategies [7].

This review will concentrate on machine learning (ML) and deep learning (DL), which are among the most frequently utilized. ML includes a set of algorithms that require feature engineering and structured data to perform effectively. In contrast, DL relies on artificial neural networks to automatically extracts features from raw input, making it extremely successful for difficult applications like picture and speech recognition. This paper reviews the latest studies using AI techniques in diagnosing OM, outlining diagnostic methods and clinical trials using them. The review also summarizes key findings from recent studies, focusing on challenges such as image quality, clinical implementation, and dataset standardization. The structure of the paper is as follows: An introduction is given in Section 1, and Section 2 talks about otitis media, Section 3 presents recent reviews and explains the contributions, Section 4 lists the diagnostic methods used in otitis media, Section 5 discusses the topic, Section 6 discusses challenges and suggestions for future research, and Section 7 provides the conclusion of the review paper.

2. Overview of Otitis Media

As seen in Fig. 1, the human ear is made up of three parts: the inner ear, middle ear, and outer ear. The membrane of the tympanic (eardrum) defines the boundary between the middle and outer ears [8].

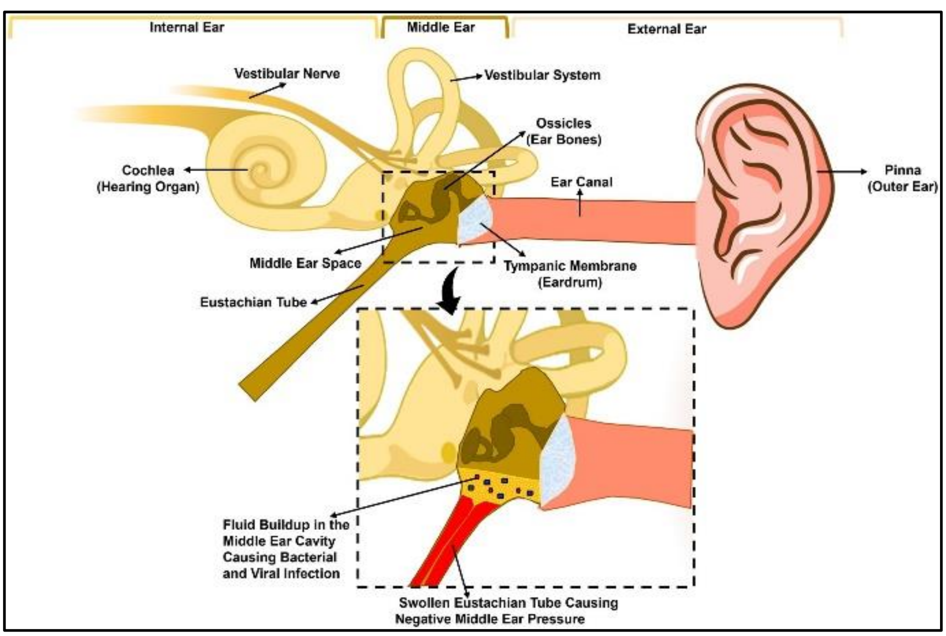

Fig. 1 - Human ear anatomy and the causes of acute otitis media [8].

A viral or bacterial infection of the upper respiratory tract can inflame the nasopharynx and Eustachian tube, leading to fluid retention in the middle ear leading to fluid retention in the middle ear, which promotes bacterial adhesion and colonization. Negative pressure in the middle ear is another effect of eustachian tube dysfunction, allowing germs or viruses from the nasopharynx to enter and cause infection. For every Tympanic Membrane (TM) condition, representative photographs are displayed in Fig. 2.

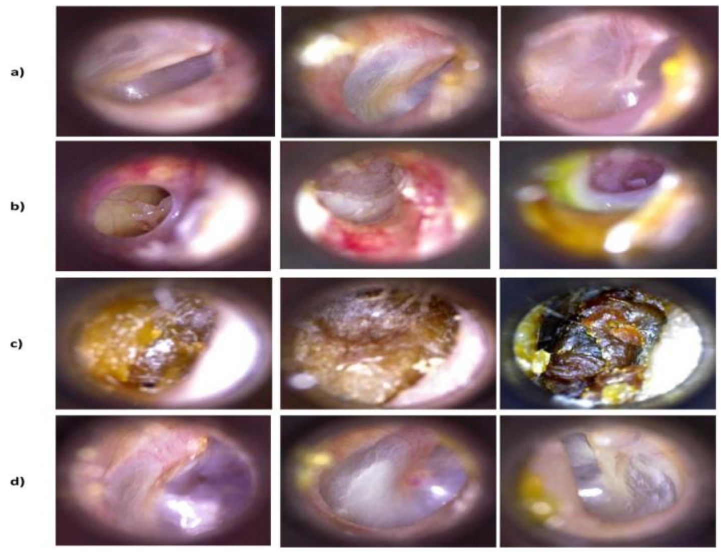

Fig. 2 - Tympanic membrane (TM) conditions. (a) Normal (b) Chronic Otitis Media (c) Earwax plug (d) Myringoesclerosis.

The three main forms of OM are acute otitis media (AOM) [9, 10], otitis media with effusion (OME) [11, 12], and chronic otitis media (COM) [13], each with unique pathological characteristics. AOM is caused by a bacterial infection leading to fluid accumulation and inflammation in the middle ear, often resulting in pain and fever. OME is more severe than AOM and involves persistent fluid accumulation in the middle ear caused by inflammation, even without an active infection. Both AOM and OME present with TM swelling, making clinical differentiation between them challenging. Chronic otitis media (COM) [14, 15] involves prolonged infection or inflammation in the middle ear, often accompanied by otorrhea (ear discharge) and a perforated tympanic membrane. Without treatment, COM can lead to worsening hearing loss, repeated infections, and severe complications like mastoiditis, cholesteatoma, or brain infections. Chronic suppurative otitis media (CSOM) is a severe, persistent form of COM associated with ongoing ear discharge and significant auditory impairment. CSOM is the leading cause of hearing loss among children in low-income regions worldwide [16, 17]. Accurate diagnosis and proper management of OM—especially distinguishing its subtypes—are crucial to prevent complications and ensure effective treatment. AI-based diagnostic models are being actively explored to enhance the accuracy and speed of differentiating AOM, OME, and COM. These tools support clinicians in early intervention and improved patient outcomes.

Delayed or incorrect diagnosis of OM can lead to persistent inflammation, hearing loss, and damage to the eardrum. In children, it can negatively affect the development of speech and language [18, 19, 20]. The most severe complications include mastoiditis, meningitis, and brain abscesses. Misdiagnosis may also lead to unnecessary antibiotic use, contributing to hard-to-treat antibiotic resistance. Therefore, prompt and precise diagnosis is critical in avoiding complications and employing appropriate strategies.

3. Literature Review

Although there aren't many, some recent publications have examined the developments in the field of otitis media diagnosis using AI. In [21], the most common challenges limiting the accuracy of diagnosing acute otitis media are discussed. New methods and techniques are also introduced to improve detection. To decrease errors in the diagnosis of otitis media, pediatricians and otolaryngologists should actively promote teleotology and the use of artificial intelligence. In [22], artificial intelligence techniques used in medical imaging in otolaryngology were reviewed. Studies were extracted from five databases without date restrictions. The review found that 26 out of the 32 studies classified data using tympanic membrane images, with an average diagnostic accuracy of 86%. Three additional studies reported an average diagnostic accuracy of 90.8% using a combination of segmentation and classification techniques. In [23], key ideas in artificial intelligence and machine learning are introduced, the ways in which these methods are currently used to diagnose, treat, and manage otitis media are explained, and the difficulties in developing machine learning methods in the future with AI assistance are examined.

In order to categorize middle ear diseases using TM images, the development of ML models was methodically assessed and their diagnostic accuracy was compared in [24]. There were 20,254 TM photos from 16 studies (7,025 normal TM images and 13,229 TM images). Each study's sample size varied from 45 to 6,066. The machine learning techniques that were included have accuracy ranging from 76.00% to 98.26%. In [25], a historical review of the application of ML in medical research and clinical use, particularly in otology, was conducted. Applications of machine learning in this field were reviewed, including diagnosis, effective identification, and prediction of surgical outcomes.

While previous work has often focused on a limited number of individual studies or techniques, this paper systematically classifies and evaluates a wide range of AI methods, including traditional machine learning, deep learning, and hybrid approaches. Previous reviews have focused primarily on image-based diagnosis, while this review will also include analysis of tympanogram and audio frequency data. It also identifies unexplored areas such as automated tympanogram analysis and the use of transfer learning in otoscopy. Furthermore, it highlights the benefits of ensemble learning and hybrid models, which have rarely been addressed in previous surveys. By structuring the findings based on diagnostic methods, data types, and model performance, this review provides a more comprehensive and updated synthesis of AI-based diagnostic methods for OM than previous reviews.

4. Methodology

A thorough literature search was carried out with the aid of internet databases such as PubMed, MEDLINE, EMBASE, Scopus, and ResearchGate, to identify relevant studies published between 2016 and 2024. The focus was on AI-based approaches for diagnosing, treating, or managing otitis media (OM) patients. Only publications in English that involved human subjects were considered. To ensure broad coverage of recent advancements, the review includes articles, abstracts, and conference proceedings. To discover additional relevant studies, the references of the selected papers and key review articles were cross-referenced. Data extraction included methodology, sample size, AI models used, and key findings. The synthesis of these studies aimed to highlight trends, challenges, and future directions in AI applications for OM management.

One of the most popular applications of AI and ML is the classification algorithm [21, 22]. It has been demonstrated in recent years that artificial intelligence significantly improves the ability to diagnose otitis media by utilizing patient data obtained with endoscopy, imaging, or various diagnostic devices [23, 24]. These advances are particularly exciting, as they offer the potential to improve diagnostic accuracy and guide more effective therapy.

The majority of computer-assisted techniques in this sector rely on the analysis of two-dimensional pictures taken with otoendoscopes or conventional otoscopes [30]. The quality of the otoscope recordings and frames chosen from those videos has a significant impact on how well these image analysis applications function; as a result, image analysis apps may exhibit low reproducibility between and between readers. A tool that gives quantitative data on the presence of fluid in the middle ear and the function of TM and middle ear structures is the tympanometry. The tympanogram serves as the data's visual representation. The machine learning literature contains few research for the automated analysis of tympanometry, in contrast to the otoscope [31]. In this research, we will review the latest literature on the use of artificial intelligence in the diagnosis of otitis media. We categorize the literature according to the use of traditional ML, DL, or hybrid methods, while addressing the methodologies used in each study, as well as the most important findings and challenges.

4.1. Machine Learning

Machine learning algorithms have been used to diagnose diseases by analyzing large data sets to discover patterns and predict outcomes with high accuracy. Supervised models are used to classify diseases based on medical imaging, lab results, and patient history while unsupervised learning techniques help identify hidden patterns in patient data, aiding in early detection of disease. There are some machine learning experiments done for automated otoscopic image-based ear condition diagnosis. An auto-diagnosis of otitis media using a decision tree was performed in [32], which classified the tympanic membrane into five categories with an accuracy of 81·58%. But crucial diagnoses like attic retraction are absent from the classification categories. The dataset includes 389 tympanic membrane images collected using various rigid endoscopes and video-otoscopes. Two independent experts confirmed the diagnoses. The image resolution varies across the dataset. There were 60 W/O, 51 AOM, 69 OME, 86 CSOM, and 123 normal in the dataset.

In order to facilitate the automatic diagnosis of middle ear disorders, Wideband Absorbance Immittance (WAI) was characterized across several frequency pressure zones in both normal middle ear and ears with OME using machine learning techniques in [33]. The highlighted main areas offer direction to practitioners to enhance their comprehension and interpretation of WAI data, potentially leading to prompt and precise diagnostic choices. The data type of the dataset is wideband acoustic immittance with total of 672 records with lebels of Normal and OME.

High quality preprocessed images of eardrums taken with digital video of otoscopes are used to train the algorithm in [34], which then uses predetermined indicators to categorize images that are not diagnosed into five OM groups. Images obtained with commercial video-otoscopes had an accuracy of 80.6%, whereas photos taken on-site with a cheap, custom-made video-otoscope had an accuracy of 78.7%. The dataset includes otoscopy images different cases of tympanic membrane where the authors removed 73 images due to insufficient image quality.

In [35], the impact of middle-ear effusion volume on children's wideband acoustic immittance with effusion-associated OM was examined. It was discovered that absorbance, a specific measure of acoustic immittance, decreased systematically as effusion volume increased, especially in the 1–5 kHz frequency range. A multivariate logistic regression technique revealed high accuracy when identifying OM based on effusion occurrence and volume.

OME detection was developed and verified in [36] with the aid of in-ear microphones and a machine learning model. Two commercial microphones were inserted into each ear canal to record the sound produced by participants as they continuously uttered five three-vowel vowels. Table 1 shows a comparison of literature that uses machine learning in realm of diagnose method, dataset, diagnose categories, models, main results, and limitations.

Table 1 - Comparison of literature in diagnosing otitis media that used machine learning.

Ref.	Diagnose Method	Dataset	Diagnose Categories	Models	Main Results	Limitations
[32] 2018	Smartphone Otoscopy	Tympanic membrane dataset, Video-Otoscope tympanic membrane dataset	Normal, OME, CSOM, W/O, AOM	Decision Tree, Neural Network	Accuracy of Decision Tree: 81.58%; Accuracy of Neural Network: 86.84%	Low-resolution images
[33] 2021	2D Frequency-Pressure WAI Images	672 WAI records from 405 participants across five hospitals in China	Normal vs. OME	Random Forest	Accuracy: 75%	Exclusion of other middle ear conditions
[34] 2016	Digital video-otoscopes	Using various commercially accessible video-otoscopes, 562 TM pictures were taken	CSOM, O/W, n TM, OME, AOM	Decision Tree	Accuracy: 78.7%	If further training photos are added, the final decision tree will need to be modified because it is fixed
[35] 2021	Tympanometry	49 children ages 9 months to 11 years had their wideband acoustic immittance	OME, AOM, No Effusion	Logistic Regression	Improved diagnostic certainty from 57% (otoscopy only) to 77%	Model explainability remains low; requires clinical validation

		tested			with WBT integration	
[36] 2023	Ear microphones	The mean age of the 31 adults diagnosed with OME was 60 years, with 20 being male and 11 being female.	OME, Normal	SVM, Gaussian Naive Bayes, AdaBoost, and Random Forest	Accuracy: 80.65%	Variations in pronunciation among individuals, gender, and age cause great variation and variation in the vocal recording

4.2. Deep Learning

The rapid development of deep learning methods has transformed diagnostic prediction and electronic health records, particularly enhancing image-based disease recognition in fields like radiology and dermatology. Managing the heterogeneity included in varied datasets has proven to be extremely challenging for medical data analysis when employing traditional methodologies [37]. However, the introduction of deep learning has replaced traditional approaches in the administration of electronic health information, including medical imaging and their use for diagnostic prediction, indicating a paradigm change. Several medical domains have effectively incorporated DL or deep neural networks. Convolutional neural networks (CNNs) are supervised DL techniques used in the majority of these investigations. In [38], a CNN model based on the EfficientNet-B4 architecture was presented. It predicted both primary classes (OME, COM, and normal) and secondary conditions such as attic cholesteatoma, meningitis, ventilating tube insertion, and otomycosis. Furthermore, with a dice similarity coefficient (DSC) of 95.19%, the model correctly predicted the principal class; in secondary classes, the DSCs for the diagnosis of meningitis and cholesteatoma were 88.37% and 88.28%, respectively.

In [39], a deep learning model was developed and validated to use multi-center otoscopic images to detect attic retraction pocket and atelectasis in OME patients. Threefold random cross-validation has been utilized with 6393 OME otoscopic images. The model exhibited a detection accuracy of 79% for atelectasis and 89% for attic retraction pocket. A large-scale tympanic measurement-based automated diagnostic technique for otitis media detection was presented in [40]. CNNs for otitis media classification was developed using a wide-band tympanogram analysis and saliency maps were calculated to get knowledge about the decision-making process of the CNNs. The method shows high performance in comprehensive detection of OM with an accuracy of 92.6%. The dataset includes wideband tympanometry measurements for diagnosing OME and AOM with total of 1014 images. In [41], a DL methodology was used to autonomously separate tympanic membranes (TMs) from videoendoscopic pictures. A hybrid loss function that combines active contour loss with dice loss is described for a fully convolutional network, and the method is tested on a dataset of 1139 endoscopic pictures. In [42], the effectiveness of five loss functions was compared with an automated classification of otitis media. The findings demonstrate that deep metric loss functions sacrifice recall in order to gain high precision on the underrepresented class. Comparing the triplet loss to the class-level weighted cross-entropy loss, the former produced the maximum precision in the AOM class without significantly lowering recall. In [43], a CNN-based method was introduced for automatic segmentation of video-otoscopic images. The model uses residual blocks for decoding, attention gates for skip connections, and EfficientNet as the encoder. In order to help neural networks with segmentation tasks, the research also presents a new loss function term.

In [44], a deep learning method was created to use computed tomography (CT) imaging of the middle ear to diagnose a variety of chronic middle ear conditions, including middle ear cholesteatoma and chronic suppurative OM. The final dataset was created using the labeling of a professional otolaryngologist and consisted of 973 ears divided into three conditions: normal, CSOM, and MEC. The foundation for diagnosis was made up of two DL networks that were used for different tasks: a classification network to finish the diagnosis and a "region of interest" area search network to extract the unique image of the middle ear anatomy. The effectiveness of a convolutional neural network in otitis media screening was examined in [45] utilizing digitized otoscopic images with total of 347 tympanic membrane images that were labeled by a panel of experts. The CNN was trained then tested using the same photos that were divided into three screening groups. These samples had a majority expert diagnosis. DL is used in [46] to analyze eardrum and ear canal images in order to detect biomedical ear infections. The method uses a noise removal procedure based on Wiener filtering (WF) to get rid of the noisy data. In [47], a neural network algorithm using a library of children's eardrum images brought to the surgical suite with the goal of executing myringotomy and possibly placing a recurring tube AOM or OME was compared to human physicians' capacity for diagnosis.

Large datasets and processing power are needed to build deep neural networks from scratch, which is impractical in many application domains. Transfer learning, on the other hand, is the process of reusing and optimizing publicly available CNN models that have been pre-trained for natural pictures for a particular use. Those features are categorized into a new set of classes by a new fully-connected layer comes after the majority of the network layers in a public network model are moved to a new model in transfer learning. Research on medical imaging using transfer learning demonstrated great classification accuracy that was on par with or even superior to creating CNN from scratch [43, 44].

In [50], nine convolution-based deep neural networks using transfer learning were trained. The models classify ear conditions into six diagnostic categories. The models were combined into an ensemble classifier, where their classification scores are aggregated to improve diagnostic accuracy. This approach mimics a multi-specialist opinion system to enhance reliability and accuracy for practical clinical use. Pre-trained CNNs were utilized for eardrum categorization in [51], whereas a new diagnostic model based on the Faster Regional Convolutional Neural Network (Faster R-CNN) was employed for eardrum detection. The model's performance was assessed when various noise effects were present. The data was set using simple image augmentation techniques like flip and rotate. By using two popular CNN networks, Xception and MobileNet-V2, an otoscopic image classifier for pediatric OM was developed in [52] based on the concepts of DL and transfer learning. AOM, OME, and normal ears were among the otoscopic images used with total of 12,203 images. Additionally, a prospective test set of 102 images taken with a smartphone and a WI-FI-connected otoscope was utilized to assess the paradigm for monitoring and screening at home. Table 2 shows a summary of the papers that uses deep learning in realm of diagnose method, dataset, diagnose categories, models, main results, and limitations.

Table 2 - Comparison of literature in diagnosing otitis media that used deep learning.

Ref.	Diagnose Method	Dataset	Diagnose Categories	Models	Main Results	Limitations
[38] 2022	Otoscopy Images	1,630 OME, 1,534 COM, 3,466 none	OME, COM, None	EfficientNet-B4	Primary class DSC: 95.19%; Secondary classes DSC: 88.37%	Requires broader dataset for generalization.
[39] 2023	Otoscopy Images	6,393 OME images	Atelectasi, Retraction Pocket	CNN with CAM	AUC: 0.87	Limited multi-center datasets and external validation
[40] 2022	Tympanometry	Include 1014 WBT readings that were taken in the Kamide ENT clinic.	OME, AOM	CNN with saliency map integration	Accuracy: 92.6%	Incapable of differentiating between subtypes of otitis media (e.g., acute versus effusion).
[41] 2021	Otoscopy Images	1,139 images	Normal, OM	Fully Convolutional Network (FCN)	DSC: 89.5%; HD: 19.189	Segmentation needs further testing
[42]	Otoscopy Images	1336 images assessed by a medical	AOM, OME, and No Effusion	Deep metric learning	Accuracy: 85%	Lack of objective

2021		specialist				measurements
[43] 2021	Otoscopy Images	1012 otoscopic pictures are included in the TM dataset.	Normal, AOM, OME, COM	Three primary paradigms: Attention gate for skip connection; EfficientNet for encoder; ResNet for decoder	Average DSC: 92.9%	Image artifacts include homogeneous intensity and weak borders.
[44] 2022	Computed Tomography (CT)	973 ears labeled by otolaryngologist	MEC, CSOM, Normal	VGG16	F1-score: 87.2% Precision: 90.1% Recall: 85.4%	Need to increase the speed of frame and reduce the size of the network
[45] 2022	Otoscopy Images	347 eardrum pictures taken with a digital otoscope	Normal, Pathological, Wax	CNN	Accuracy: 90%	Using only one CNN model
[46] 2024	Otoscopy Images	Large labeled dataset of ear images	Infected vs. not infected	Fuzzy Restricted Boltzmann Machine (FRBM)	Improved noise reduction	Based on high-quality input images
[47] 2023	Otoscopy Images	639 images of eardrum	Normal, OME, AOM	Neural network, Commercial Model (Google)	Accuracy: 80.8%	Unusually high-quality clinical photos were included in the training data.
[50] 2019	Otoscopy Images	10,544 otoendoscopic images	Tympanic perforation, Attic retraction, Otitis externa ± myringitis, Tumor, Normal	Inception-V3 and ResNet101, combined into an ensemble classifier	Average accuracy: 93.67%	Dataset dependency may limit generalization
[51] 2020	Otoscopy Images	1692 augmented otoscope images (raw image 282)	Normal vs. Abnormal	Faster R-CNN for eardrum identification and pre-trained CNNs for eardrum categorization	Accuracy: 90.48% VGG-16	A few raw images
[52] 2021	Otoscopy images from smartphone-enabled wireless otoscope	12,203 otoendoscopic images 102 images captured by smartphone	AOM, OME, Normal	Xception, MobileNet-V2	Accuracy: 95.72%	Insufficient to assess the severity of OM

4.3. Hybrid Approaches

Some studies have used a combination of several methods to improve diagnostic results. In [53], digital otoscopy images and tympanoplasty measurements were combined for accurate diagnosis of tympanic abnormalities. By using a decision fusion mechanism, predictions from both methods were combined, significantly improving the diagnostic accuracy to 84.9%, compared to the individual methods. This approach addresses the limitations of subjective visual examinations by incorporating complementary biophysical data.

The CNN model, along with its deep characteristics and the images captured by the otoscope device, were utilized in [54]. In the initial step of the experiment, the VGG16 model was used effort to distinguish between these images as Normal and Abnormal. The second step included obtaining the activation maps. After that, it was fed into Support Vector Machines (SVM). In [55], a novel deep learning-based voting ensemble framework was proposed for diagnosing middle ear diseases using otoscopic images. By combining the strengths of multiple pre-trained CNN models, the soft voting ensemble achieved state-of-the-art performance with high accuracy, sensitivity, and specificity. The dataset includes 880 otoscopy pictures of 180 patients, ages 7 to 65, with a resolution of 420 × 380. Four categories—chronic otitis media, tympanosclerosis, earwax plug, and normal otoscopy—are used to group the images in the dataset. In [56], a method using the EfficientNet-B7 backbone to classify tympanic membrane diseases and hearing decline in children is presented. The model's ability to integrate classification and regression tasks makes it a versatile tool for tympanic membrane diagnosis and hearing assessment in children. The dataset used is named SCH Tympanic Membrane Dataset which includes 23,302 JPG image files with distinct training tasks for each of the two datasets—one for classification and one for regression. When a patient visits, they often have an otoendoscopy to take pictures of their eardrums. These images have a resolution of 1280 x 1350 pixels. Table 3 shows a summary of the papers that uses hybrid approaches in realm of diagnose method, dataset, diagnose categories, models, main results, and limitations.

Table 3 - Comparison of literature in diagnosing otitis media that used hybrid approaches.

Ref.	Diagnose Method	Dataset	Diagnose Categories	Models	Main Results	Limitations
[53] 2020	Tympanometry and Otoscopy videos	ImageNet Large Scale Visual Recognition Challenge database that has more than 1.2 M images from 1,000 classes	Normal vs. Abnormal	Random forest for tympanometry, Inception-ResNet-v2 for otoscopy, and majority voting for fusion	Accuracy: 84.9%	Fusion accuracy depends on the quality of each individual modality
[54] 2022	Tympanometry	956 middle ear images	Normal vs. Abnormal	VGG16 + SVM	Accuracy: 82.17% for specific feature layers	Lack of multi-center testing
[55] 2024	Otoscopy Images	Public Ear Imagery dataset (880 otoscopy images)	Normal, Earwax plug, Myringosclerosis, Chronic otitis media	Deep learning-based ensemble method (Soft and hard voting ensembles of pre-trained CNNs)	Accuracy: 98.8%, Sensitivity 97.5%, Specificity 99.1%	Depend on high-quality otoscopic images

[56] 2024	Otoscopy Images	Open-access eardrum dataset, SCH eardrum dataset	Eardrum diseases (4–5 classes) and pediatric hearing (regression)	EfficientNet-B7-based convolutional neural network with drop connect for generalization and multi-layer perceptron in the decoder	Accuracy: 93.59% on open-access dataset, Accuracy: 98.28% on SCH dataset	Performance degradation caused by this data imbalance

5. Discussion

Deep learning approaches generally outperform traditional machine learning in terms of accuracy. For instance, ensemble methods like Inception-V3 and ResNet101 achieved 93.67% accuracy [50], while MobileNet-V2 and Xception reported 95.72% [52]. CNN-based models also performed well, such as one using tympanometry data with 92.6% accuracy [40], and another ensemble model reaching 98.8% [55]. In contrast, traditional models like Decision Trees and Random Forests achieved 75–86% accuracy [32], [33], [34]. Hybrid models show promise, with an ensemble approach combining CNNs and majority voting reaching 84.9% [53], and VGG16+SVM scoring 82.17% [54]. The highest performance came from newer hybrid CNN models, like EfficientNet-B7, achieving up to 98.28% [56].

Deep learning models also demonstrated strong performance across other evaluation metrics. For instance, VGG16 applied to CT scans achieved a precision of 90.1%, recall of 85.4%, and an F1-score of 87.2% [44], indicating balanced classification capabilities. Ensemble models like those using EfficientNet-B7 and CNN-based soft/hard voting approaches reported high sensitivity (97.5%) and specificity (99.1%) [55], highlighting their effectiveness in minimizing false negatives and false positives.

Using AI-based methods has shown better results in classification, image analysis, and decision-making. Research has demonstrated that CNNs are capable of distinguishing between OM subtypes, which improves the accuracy of otoscopic image evaluation. Furthermore, hybrid models that combine AI-driven image recognition and tympanometry have produced better diagnostic accuracy.

6. Challenges and Future Considerations

Despite considerable progress in AI-driven OM diagnosis, several key challenges persist. One major issue is the lack of lightweight and efficient AI models that can be seamlessly integrated into real-time clinical workflows. Many current algorithms are computationally intensive and not optimized for practical deployment in point-of-care settings. Additionally, there is a lack of standardized datasets, which hampers model generalizability and consistent performance across diverse clinical environments. Real-time monitoring remains a significant hurdle, especially in dynamic scenarios like surgeries, where speed and precision are critical.

The review also highlights considerable heterogeneity among existing studies in terms of data quality and sources. Many studies deliberately exclude difficult cases, such as blurred or defocused otoscopic images, which reduces the robustness of their models. This selective inclusion increases reported accuracy but undermines real-world applicability. To address these limitations, future research should prioritize data diversity and resilience. Solutions such as dataset augmentation, synthetic data generation, and Federated Learning can help mitigate bias and enhance model robustness. Future efforts should focus on the development of diverse datasets that incorporate multiple diagnostic features to improve AI models' accuracy and generalizability. Multimodal approaches that combine otoscopy images, tympanometry measurements, and acoustic analysis can provide a more comprehensive evaluation of middle ear inflammation. Otoscopy images offer visual insights into tympanic membrane abnormalities, while tympanometry assesses middle ear pressure and fluid presence. Acoustic analysis can detect subtle auditory changes associated with otitis media, improving early detection. Integrating these diagnostic modalities can enhance classification accuracy and reduce false positives and negatives.

Conclusion

This review highlights the efficacy of AI methodologies in automating and assisting with the diagnosis of OM. Despite being in the development and testing phases, AI can enhance the practices of otolaryngologists and primary care clinicians by augmenting the effectiveness and precision of diagnosis through inclusion of ML and deep learning techniques has further enhanced diagnostic accuracy, paving the way for automated and accessible solutions. Diverse datasets and advanced analytics improve generalizability and early detection. Future research should prioritize data fusion and real-time clinical applicability for better outcomes.

References

[1] A. Prasad, S. M. A. Hasan, and M. R. Gartia, "Optical identification of middle ear infection," Molecules, vol. 25, no. 9, p. 2239, Sep. 2020. doi: 10.3390/molecules25092239.

[2] D. M. Spiro, K. Y. Tay, D. H. Arnold, J. D. Dziura, M. D. Baker, and E. D. Shapiro, "Wait-and-see prescription for the treatment of acute otitis media: a randomized controlled trial," Jama, vol. 296, no. 10, pp. 1235-1241, Sep. 2006. doi: 10.1001/jama.296.10.1235.

[3] C. G. Brennan-Jones, A. J. Whitehouse, S. D. Calder, C. Da Costa, R. H. Eikelboom, and S. E. Jamieson, "Does otitis media affect later language ability? A prospective birth cohort study."

[4] J. H. Wolleswinkel-van den Bosch, E. A. Stolk, M. Francois, R. Gasparini, and M. Brosa, "The health care burden and societal impact of acute otitis media in seven European countries: results of an Internet survey," Vaccine, vol. 28, pp. G39-G52, Jun. 2010. doi: 10.1016/j.vaccine.2010.06.014.

[5] A. T. Khalaf and S. K. Abdulateef, "Ophthalmic diseases classification based on YOLOv8," Journal of Robotics and Control (JRC), vol. 5, no. 2, pp. 408-415, 2024. doi: 10.18196/jrc.v5i2.21208.

[6] S. K. Abdulateef, A. N. Ismael, and M. D. Salman, "Feature weighting for Parkinson's identification using single hidden layer neural network," International Journal of Computing, vol. 22, no. 2, pp. 225-230, 2023. doi: 10.47839/ijc.22.2.3092.

[7] S. A. Qureshi et al., "Intelligent ultra-light deep learning model for multi-class brain tumor detection," Applied Sciences, vol. 12, no. 8, p. 3715, Apr. 2022. doi: 10.3390/app12083715.

[8] A. Zahid, J. C. Wilson, I. D. Grice, and I. R. Peak, "Otitis media: recent advances in otitis media vaccine development and model systems," Frontiers in Microbiology, vol. 15, p. 1345027, 2024. doi: 10.3389/fmicb.2024.1345027.

[9] R. P. Venekamp, S. L. Sanders, P. P. Glasziou, and M. M. Rovers, "Antibiotics for acute otitis media in children," Cochrane database of systematic reviews, no. 11, 2023.

[10] S. Mohanty et al., "Incidence of acute otitis media from 2003 to 2019 in children ≤ 17 years in England," BMC Public Health, vol. 23, no. 1, p. 201, 2023. doi: 10.1186/s12889-023-14982-8.

[11] H. Hidaka, M. Ito, R. Ikeda, Y. Kamide, H. Kuroki, A. Nakano, et al., "Clinical practice guidelines for the diagnosis and management of otitis media with effusion (OME) in children in Japan–2022 update," *Auris Nasus Larynx*, vol. 50, no. 5, pp. 655-699, 2023. doi: 10.1016/j.anl.2022.12.004.

[12] J. H. Shim, W. Sunwoo, B. Y. Choi, K. G. Kim, and Y. J. Kim, "Improving the accuracy of otitis media with effusion diagnosis in pediatric patients using deep learning," *Bioengineering*, vol. 10, no. 11, p. 1337, 2023. doi: 10.3390/bioengineering10111337.

[13] H. Zhang, J. Huang, T. Li, S. Svanberg, and K. Svanberg, "Optical detection of middle ear infection using spectroscopic techniques: phantom experiments," *J. Biomed. Opt.*, vol. 20, no. 5, pp. 057001-057001, 2015. doi: 10.1117/1.JBO.20.5.057001.

[14] G. Brescia, A. Frosolini, L. Franz, A. Daloiso, F. Fantin, A. Lovato, et al., "Chronic otitis media in patients with chronic rhinosinusitis: a systematic review," *Medicina*, vol. 59, no. 1, p. 123, 2023. doi: 10.3390/medicina5901012.

[15] E. M. Schouwenaar, C. A. Hellingman, and J. J. Waterval, "Health-related quality of life after otologic surgical treatment for chronic otitis media: systematic review," *Front. Neurol.*, vol. 14, p. 1268785, 2023.

[16] E. Heward, H. Saeed, S. Bate, A. Rajai, J. Molloy, R. Isba, et al., "Risk factors associated with the development of chronic suppurative otitis media in children: systematic review and meta-analysis," *Clin. Otolaryngol.*, vol. 49, no. 1, pp. 62-73, 2024.

[17] M. Khairkar, P. Deshmukh, H. Maity, and V. Deotale, "Chronic suppurative otitis media: a comprehensive review of epidemiology, pathogenesis, microbiology, and complications," *Cureus*, vol. 15, no. 8, 2023.

[18] J. Pitaro, S. Waissbluth, M. C. Quintal, A. Abela, and A. Lapointe, "Characteristics of children with refractory acute otitis media treated at the pediatric emergency department," *Int. J. Pediatr. Otorhinolaryngol.*, vol. 116, pp. 173-176, 2019.

[19] C. Conlon, B. Zupan, E. Pirie, and C. Gupta, "The impact of otitis media on speech production in children: a systematic review," *J. Commun. Disord.*, p. 106490, 2024.

[20] M. Savaş, "The effect of otitis media with effusion on language and cognitive skills in school age children," *Experimed*, vol. 13, no. 2, pp. 156-162, 2023.

[21] S. Esposito, S. Bianchini, A. Argentiero, R. Gobbi, C. Vicini, and N. Principi, "New approaches and technologies to improve accuracy of acute otitis media diagnosis," *Diagnostics*, vol. 11, no. 12, p. 2392, 2021.

[22] D. Song, T. Kim, Y. Lee, and J. Kim, "Image-based artificial intelligence technology for diagnosing middle ear diseases: a systematic review," *J. Clin. Med.*, vol. 12, no. 18, p. 5831, 2023.

[23] X. Ding, Y. Huang, X. Tian, Y. Zhao, G. Feng, and Z. Gao, "Diagnosis, treatment, and management of otitis media with artificial intelligence," *Diagnostics*, vol. 13, no. 13, p. 2309, 2023.

[24] Z. Cao, F. Chen, E. M. Grais, F. Yue, Y. Cai, D. W. Swanepoel, and F. Zhao, "Machine learning in diagnosing middle ear disorders using tympanic membrane images: a meta-analysis," *Laryngoscope*, vol. 133, no. 4, pp. 732-741, 2023.

[25] H. Koyama, "Machine learning application in otology," *Auris Nasus Larynx*, vol. 51, no. 4, pp. 666-673, 2024.

[26] M. G. Crowson, J. Ranisau, A. Eskander, A. Babier, B. Xu, R. R. Kahmke, et al., "A contemporary review of machine learning in otolaryngology–head and neck surgery," *Laryngoscope*, vol. 130, no. 1, pp. 45-51, 2020. doi: 10.1002/lary.27850.

[27] Q. An, S. Rahman, J. Zhou, and J. J. Kang, "A comprehensive review on machine learning in healthcare industry: classification, restrictions, opportunities and challenges," *Sensors*, vol. 23, no. 9, p. 4178, 2023.

[28] M. Viscaino, J. C. Maass, P. H. Delano, M. Torrente, C. Stott, and F. Auat Cheein, "Computer-aided diagnosis of external and middle ear conditions: a machine learning approach," *PLoS One*, vol. 15, no. 3, p. e0229226, 2020. doi: 10.1371/journal.pone.0229226.

[29] H. Byun, S. Yu, J. Oh, J. Bae, M. S. Yoon, S. H. Lee, et al., "An assistive role of a machine learning network in diagnosis of middle ear diseases," *J. Clin. Med.*, vol. 10, no. 15, p. 3198, 2021.

[30] C. Senaras, A. C. Moberly, T. Teknos, G. Essig, C. Elmaraghy, N. Taj-Schaal, et al., "Autoscope: automated otoscopy image analysis to diagnose ear pathology and use of clinically motivated eardrum features," in *Proc. SPIE Med. Imaging: Computer-Aided Diagnosis*, vol. 10134, pp. 500–507, Mar. 2017.

[31] M. Moein, M. Davarpanah, M. A. Montazeri, and M. Ataei, "Classifying ear disorders using support vector machines," in *Proc. 2010 Second Int. Conf. Comput. Intell. Nat. Comput.*, vol. 1, pp. 321–324, Sep. 2010. doi: 10.1109/CINC.2010.5643830.

[32] H. C. Myburgh, S. Jose, D. W. Swanepoel, and C. Laurent, "Towards low cost automated smartphone-and cloud-based otitis media diagnosis," *Biomed. Signal Process. Control*, vol. 39, pp. 34–52, 2018.
[33] E. M. Grais et al., "Analysing wideband absorbance immittance in normal and ears with otitis media with effusion using machine learning," *Sci. Rep.*, vol. 11, no. 1, p. 10643, 2021.
[34] H. C. Myburgh, W. H. Van Zijl, D. Swanepoel, S. Hellström, and C. Laurent, "Otitis media diagnosis for developing countries using tympanic membrane image-analysis," *EBioMedicine*, vol. 5, pp. 156–160, 2016.
[35] G. R. Merchant, S. Al-Salim, R. M. Tempero, D. Fitzpatrick, and S. T. Neely, "Improving the differential diagnosis of otitis media with effusion using wideband acoustic immittance," *Ear Hear.*, vol. 42, no. 5, pp. 1183–1194, 2021.
[36] K. C. Ting et al., "Detection of Otitis Media With Effusion Using In-Ear Microphones and Machine Learning," *IEEE Sensors J.*, 2023.
[37] S. M. F. Malik, M. T. Nafis, M. A. Ahad, and S. Tanweer, "A Multi-Dataset Classification-Based Deep Learning Framework for Electronic Health Records and Predictive Analysis in Healthcare," *arXiv preprint*, arXiv:2409.16721, 2024.
[38] Y. Choi et al., "Automated multi-class classification for prediction of tympanic membrane changes with deep learning models," *PLoS One*, vol. 17, no. 10, p. e0275846, 2022. doi: 10.1371/journal.pone.0275846.
[39] J. Zeng et al., "A deep learning approach to the diagnosis of atelectasis and attic retraction pocket in otitis media with effusion using otoscopic images," *Eur. Arch. Otorhinolaryngol.*, vol. 280, no. 4, pp. 1621–1627, 2023.
[40] J. V. Sundgaard et al., "A deep learning approach for detecting otitis media from wideband tympanometry measurements," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 7, pp. 2974–2982, 2022. doi: 10.1109/JBHI.2022.3159263.
[41] V. T. Pham, T. T. Tran, P. C. Wang, and M. T. Lo, "Tympanic membrane segmentation in otoscopic images based on fully convolutional network with active contour loss," *Signal Image Video Process.*, vol. 15, no. 3, pp. 519–527, 2021. doi: 10.1007/s11760-020-01772-7.
[42] J. V. Sundgaard et al., "Deep metric learning for otitis media classification," *Med. Image Anal.*, vol. 71, p. 102034, 2021.
[43] V. T. Pham, T. T. Tran, P. C. Wang, P. Y. Chen, and M. T. Lo, "EAR-UNet: A deep learning-based approach for segmentation of tympanic membranes from otoscopic images," *Artif. Intell. Med.*, vol. 115, p. 102065, 2021.
[44] Z. Wang et al., "Structure-aware deep learning for chronic middle ear disease," *Expert Syst. Appl.*, vol. 194, p. 116519, 2022.
[45] J. Sandström, H. Myburgh, C. Laurent, D. W. Swanepoel, and T. Lundberg, "A machine learning approach to screen for otitis media using digital otoscope images labelled by an expert panel," *Diagnostics*, vol. 12, no. 6, p. 1318, 2022.
[46] I. M. Mehedi, M. S. Hanif, M. Bilal, M. T. Vellingiri, and T. Palaniswamy, "Artificial Intelligence with Deep Learning Based Automated Ear Infection Detection," *IEEE Access*, 2024.
[47] M. G. Crowson, D. W. Bates, K. Suresh, M. S. Cohen, and C. J. Hartnick, "“Human vs Machine” Validation of a Deep Learning Algorithm for Pediatric Middle Ear Infection Diagnosis," *Otolaryngol. Head Neck Surg.*, vol. 169, no. 1, pp. 41–46, 2023.
[48] H. C. Shin et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.
[49] D. S. Kermany et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.
[50] D. Cha, C. Pae, S. B. Seong, J. Y. Choi, and H. J. Park, "Automated diagnosis of ear disease using ensemble deep learning with a big otoendoscopy image database," *EBioMedicine*, vol. 45, pp. 606–614, 2019.
[51] E. Başaran, Z. Cömert, and Y. Çelik, "Convolutional neural network approach for automatic tympanic membrane detection and classification," *Biomed. Signal Process. Control,* vol. 56, p. 101734, 2020. doi: 10.1016/j.bspc.2019.101734.
[52] Z. Wu et al., "Deep learning for classification of pediatric otitis media," *Laryngoscope*, vol. 131, no. 7, pp. E2344–E2351, 2021. doi: 10.1002/lary.29302.
[53] H. Binol et al., "Decision fusion on image analysis and tympanometry to detect eardrum abnormalities," in *Med. Imaging 2020: Comput.-Aided Diagnosis*, vol. 11314, pp. 375–382, SPIE, Mar. 2020.
[54] A. Çalışkan, "Classification of tympanic membrane images based on VGG16 model," *Kocaeli J. Sci. Eng.*, vol. 5, no. 1, pp. 105–111, 2022. doi: 10.34088/kojose.1081402.
[55] K. Akyol, E. Uçar, Ü. Atila, and M. Uçar, "An ensemble approach for classification of tympanic membrane conditions using soft voting classifier," *Multimedia Tools Appl.*, pp. 1–22, 2024. doi: 10.2139/ssrn.4020054.
[56] H. Lee, H. Jang, W. Jeon, and S. Choi, "Diagnosis of Tympanic Membrane Disease and Pediatric Hearing Using Convolutional Neural Network Models with Multi-Layer Perceptrons," *Appl. Sci.*, vol. 14, no. 13, p. 5457, 2024. doi: 10.3390/app14135457.