21. Yüzyılda Fen ve Teknik Dergisi
ISSN: 2587-0327
2025, Cilt/Volume: 12, Sayı/Issue: 24, 76-86
Research Article / Araştırma Makalesi
Received: 16.11.2025; Accepted: 25.12.2025

Detection of Tinnitus from Audiogram Images Using Machine Learning Methods

Remzi Gürfidan[1*], Yasin Tatli[2]

Abstract: Tinnitus is an important symptom of hearing problems, the diagnosis of which is generally carried out through subjective clinical interviews and audiological tests; this may lead to variability in its diagnosis and follow-up. The aim of this study is to develop an objective decision-support approach for tinnitus detection by classifying audiogram images with state-of-the-art machine learning methods. This study used a publicly accessible audiogram dataset with standard clinical hearing threshold information and tinnitus labels. These records were converted into image representations and split into training, validation, and test sets. A conventional convolutional neural network and four transfer-learning–based deep architectures, namely InceptionV3, DenseNet121, VGG19, and Xception, were trained and tested. Model performance was evaluated based on metrics such as validation accuracy, validation loss, test accuracy, precision, recall, and F1-score. Considering the models above, the best overall results were obtained using the Xception model with 0.9387 validation accuracy and 91.67% test accuracy, together with rather balanced precision, recall, and F1-score values of 0.928. It was followed by models InceptionV3 and DenseNet121, which attained 0.9202 validation accuracy and over 90% test accuracies, while the CNN baseline resulted in 74.02% test accuracy. VGG19 showed markedly lower performance, indicating that not all deep architectures are equally suitable for this task. The findings demonstrate that modern deep learning architectures, particularly Xception, can successfully extract discriminative patterns from audiogram images and provide reliable, objective support for tinnitus detection. These results suggest that integrating such models into audiology workflows could contribute to faster, more standardised, and reproducible clinical decision-making.

Keywords: Tinnitus, audiogram, deep learning, medical image classification

[1]**Address:** Isparta University of Applied Science, Isparta Vocational School of Information Technologies, Database, Network Design and Management; Isparta, Türkiye
[2]**Address:** Avrasya University, Vocational School of Health Services, Department of Medical Services and Techniques, Audiometry Programme, Trabzon, Türkiye

***Corresponding author**: remzigurfidan@isparta.edu.tr

Citation: Gürfidan, R., Tatlı, Y. (2025). Detection Of Tinnitus from Audiogram Images Using Machine Learning Methods. 21. Yüzyılda Fen ve Teknik Dergisi, 12(24): 76-86.

INTRODUCTION

Tinnitus is a highly prevalent auditory symptom characterised by the perception of sound in the ears or head despite the absence of any external sound stimulus. Clinically referred to as ringing in the ears, this condition can manifest as a humming, buzzing, whistling, machine noise, sounds synchronised with the pulse, or more complex tonal perceptions (Savage and Waddell, 2012). Rather than being a disease in itself, tinnitus is considered a symptom of various underlying pathologies of ontological, neurological, vascular, or psychiatric origin. Prevalence studies reveal that the proportion of individuals who experience tinnitus at least once in their lifetime is quite high, and that chronic tinnitus, which is significantly distressing, has reached the level of a major public health problem (Baguley et.al., 2013, Pryce et.al., 2023).

The diagnosis of tinnitus primarily involves a detailed clinical history and a comprehensive ear, nose, and throat examination as the fundamental steps. During this process, the duration of the tinnitus, whether it is unilateral or bilateral, whether it is pulsatile or not, and accompanying symptoms such as hearing loss, vertigo, balance disorders, anxiety, depression, or sleep problems are systematically investigated. Standard audiological tests such as pure tone audiometry, speech audiometry, tympanometry, and acoustic reflex measurements are commonly used in the diagnostic evaluation (Tang et.al. 2019). In addition, tinnitus matching (frequency and intensity determination), masking threshold measurements, and various questionnaire scales (e.g., Tinnitus Handicap Inventory, Tinnitus Functional Index, etc.) are used to quantitatively determine the subjective characteristics and severity of tinnitus. where necessary, advanced imaging methods such as magnetic resonance imaging or computed tomography and neurological assessments are used to investigate underlying organic pathologies (Kleinjung et.al., 2024; Ogawa et.al., 2020). Today, the significance of tinnitus extends beyond being merely a complaint limited to the auditory system, and its multifaceted effects on an individual's quality of life are increasingly emphasised. Chronic tinnitus can lead to loss of productivity in work and social life, difficulty concentrating, sleep disorders, irritability, anxiety and depression; in this respect, it creates a significant economic and psychosocial burden on both an individual and societal scale (Watts et.al., 2018). In particular, modern working environments with increased noise exposure, an ageing population, the use of ototoxic medications, and an increase in cardiovascular risk factors are contributing to an expanded risk profile for tinnitus. This situation makes the early detection, systematic assessment, and management of tinnitus a priority area in contemporary health policies (Meller, 2007). In parallel, the increase in neuroimaging, neurophysiological recording, and cognitive neuroscience-based studies aimed at explaining the pathophysiology of tinnitus suggests in the current literature that this symptom may be associated with a complex process of restructuring between auditory and non-auditory networks in the central nervous system (Simoes et.al., 2021). Within this framework, tinnitus is considered not only as a result of damage to the peripheral auditory pathway, but also as a clinical problem requiring a multidisciplinary approach, closely related to plasticity processes, mood regulation, attention, and perception mechanisms. Therefore, the diagnosis, monitoring, and treatment of tinnitus are increasingly being approached holistically, with contributions from various disciplines such as audiology, otology, neurology, psychiatry, psychology, and public health (Couth et.al., 2019).

2. RELATED WORKS

The primary aim of Huang and colleagues' study is to identify objective, tinnitus-related features in the Ensemble Spontaneous Activity (ESA) signal recorded from the cochlear surface in experimental animals with a noise-induced tinnitus model, and to assess the reversibility of these features using extracochlear electrical stimulation (EES). Although tinnitus affects 10–15% of the global population, clinical assessment relies on subjective reports, limiting diagnostic accuracy and treatment development. Researchers recorded ESA from groups of animals exposed to noise, with tinnitus, without tinnitus, and with tinnitus symptoms alleviated after EES. Spectral distribution characteristics such as centre frequency, bandwidth, skewness, and kurtosis were extracted using Fast Fourier Transform (FFT) and sliding window analysis. To classify tinnitus status, five machine learning models (Logistic Regression, Random Forest, Support Vector Machine, XGBoost, Gradient Boosting Machine) were trained, and the key predictors were determined using SHAP analysis. The results obtained indicate that spectral changes associated with tinnitus occur in frequency bands related to the noise frequency and its harmonic/sub-harmonic ranges (particularly 1400–1600 Hz, 8000–9000 Hz, and 16,000–

17,000 Hz). These spectral changes were reversed by EES, and this situation showed parallelism with the improvement in behavioural tinnitus assessment (GPIAS). In machine learning classification, Support Vector Machine (SVM) achieved the highest performance (AUC = 0.962) and SHAP analysis identified the centre frequency (1400–1600 Hz) and distortion (8000–9000 Hz; 16,000–17,000 Hz) as the most influential features. These findings suggest that specific ESA spectral features may serve as objective and reversible biomarkers with potential for clinical diagnosis and treatment monitoring (Huang et al., 2025).

Zhang et al. addressed the challenges of extracting high-dimensional tinnitus features from electroencephalography (EEG) signals with low signal-to-noise ratio (SNR) and the fragility of existing models against erroneous (noisy) labels in training data. They developed a robust, multi-view model ensemble called XTinnitusNet for EEG-based tinnitus diagnosis. Methodologically, XTinnitusNet features a dual architecture integrating two main components: a Graph Attention Neural Network (GANN) that analyses brain functional connectivity features and a Multi-Scale Convolutional Neural Network (MSCNN) that extracts multi-scale time series features. This integration aims to enhance the ability to capture complex EEG patterns. To improve performance under noisy labelling conditions, a co-teaching plus mechanism was incorporated into the training process; this mechanism enabled cross-updating between MSCNN and GANN using disagreement data with minimal loss. The results obtained demonstrate that XTinnitusNet outperforms baseline models in terms of diagnostic accuracy and robustness against noise levels. The model achieved an AUC success rate of 88.33% at a 0.1% noise rate and maintained a high AUC performance of 84.02% even when the noise rate increased to 0.3%. Feature interpretability analysis revealed that XTinnitusNet effectively extracted meaningful functional connectivity features, particularly in the left temporal and frontal cortices (Zhang et al., 2026).

Allgaier et al. investigated whether raw EEG data could be used for tinnitus prediction and classification using deep neural networks, addressing the issues of general treatment deficiency stemming from the significant heterogeneity of tinnitus patients and the difficulty of objectively measuring symptoms. Based on the assumption that tinnitus originates from processes in the brain, the researchers explored the potential of deep learning. As a method, 16,780 raw EEG samples, each lasting one second, collected from 42 subjects (29 tinnitus patients, 13 controls) were analysed. An end-to-end approach was adopted, meaning the system was allowed to automatically learn the features necessary to solve the task. Four different preprocessing scenarios were compared: with or without noise reduction (using the Auto reject library), and with or without up-sampling or down-sampling methods. A fully connected deep neural network architecture comprising five dense layers, a dropout layer, and a flatten layer was employed for classification. The results obtained demonstrated that a maximum accuracy score of 75.6% was achieved on the test set when noise reduction and down-sampling procedures were applied. Noise reduction (Auto reject) increased accuracy by 10 points in the down-sampling approach and by 2 points in the up-sampling approach. This study confirmed that deep learning approaches hold potential for detecting EEG patterns specific to tinnitus patients, which is difficult for humans (Allgaier et al., 2021).

Mai et al. aimed to investigate differences in the specificity of resting-state EEG signals in tinnitus patients by combining time-frequency domain and non-linear power analysis (e.g., entropy) techniques, thereby providing clinicians with an objective approach to tinnitus diagnosis. As a method, resting-state EEGs collected from tinnitus patients and healthy subjects, each comprising 10 individuals, were analysed. Based on the assumption that the brains of tinnitus patients are chaotic and dynamic systems, the non-linear dynamic analysis

method was selected. Features in the d, h, a1, a2, b1, b2, b3, and c rhythms were extracted using Wavelet Packet Transform (WPT) and Sample Entropy. Support Vector Machines (SVM) were used for objective classification, combined with optimal feature combinations. The findings revealed that the Sample Entropy values of the d, a2, and b1 bands were higher in tinnitus patients than in healthy subjects, and this difference was highly significant ($p < 0.01$). These differences were particularly concentrated in the d rhythm in the right parietal region, the a2 rhythm in the central region, and the c rhythm in the left prefrontal region. SVM classification achieved a high average accuracy rate of 91.38% using optimal feature combinations. It was interpreted that the increase in entropy values reflects an increase in chaotic behaviour in the brain and may characterise the onset of tinnitus (Mai et al., 2023).

Awais et al. seek to develop the understanding and methods regarding tinnitus treatment by analyzing EEG signals recorded during treatment; hence, they represent a new methodology. As a method, five different acoustic therapies (Relaxing Music, Tinnitus Retraining Therapy-TRT, Enriched Acoustic Environment Therapy-EAET, Binaural Beats Therapy-BBT, and Auditory Discrimination Therapy-ADT). Pre-processing steps included noise elimination-0.5-1 Hz and 40-50 Hz bandpass filtering, removal of bad channels, and artefact elimination using the RANSAC algorithm. Several time, frequency, and time-frequency features were extracted, including Power Spectral Density (PSD). The novelty in this study is to transform the feature vectors of EEG channels into a new Graph Network structure where the nodes represent the feature vectors, and the edges represent the relations between them-euclidean distance less than 0.2. Next, this graph structure was analyzed using the GCN-LSTM model, which combines Graph Convolutional Network with Long Short-Term Memory networks. These results obtained then provide evidence that the GCN-LSTM model achieved remarkable success in tinnitus detection: 99.41% accuracy, 0.991 AUC, and 0.9700 F1 score. High accuracy provides evidence of the appropriateness of such a model in effectively discerning tinnitus by capturing complex EEG patterns and temporal dependencies of Awais et al. (2024).

Hong et al. support the feasibility of the diagnosis of tinnitus symptoms by deep learning analysis (EEGNet) of electroencephalographic signals recorded while performing auditory cognitive tasks. Tinnitus is assumed to be related to maladaptive neuroplasticity caused by disturbances in the auditory system. As a method, 11 patients with chronic tinnitus and 11 healthy controls underwent EEG recording during an active oddball task as a top-down cognitive task and a passive oddball task as a bottom-up sensory task. A compact convolutional neural network architecture called EEGNet was used for automatic learning of discriminative features from raw EEG data. To assess the performance of this model, an 11-layer, leave-one-out cross-validation method, along with the AUC score, was applied. The main findings indicated that the deep learning model (EEGNet) successfully identified tinnitus patients with an AUC value of 0.886 using target stimuli during the active oddball task. Time-frequency analysis of EEG signals demonstrated that there was a significant decrease in pre-stimulus alpha activity in the tinnitus group as compared to the healthy group during both the active and passive oddball tasks ($p < 0.005$). Also, the analysis of convolutional layer filters of EEGNet demonstrated the alpha band (8 - 13 Hz) as the most decisive feature for the classification of tinnitus patients. These results support the concept that task-related EEG features are considered the neural signature of tinnitus symptoms (Hong et al., 2023).

3. MATERIAL AND METHOD

This section will provide detailed and technical information on the dataset used in the study and the evaluation metrics of the trained machine learning models.

3.1. Dataset

In this study, machine learning models for the objective detection of tinnitus from audiogram data were developed using the open dataset "Audiogram Analysis for Tinnitus Detection," which was made available by Kaveh Kavianpour on the Kaggle data science platform for researchers (Kaggle, 2024). This dataset comprises tabular records of multidimensional auditory features, viz. hearing thresholds at different frequency bands for both ears, which were obtained from standard clinical audiogram measurements of hundreds of individuals, along with classification labels reflecting the presence/absence of tinnitus and the related clinical condition of the participants. Tinnitus graphs for the right and left ear included in the data set are shown in Figure 1. The dataset was split into training, validation, and test sets in the order of 70%, 20%, and 10% respectively, and each model was trained on the same data split. The model hyperparameters were kept fixed for Xception, ResNet50, and EfficientNet-B0; all models were trained for 30 epochs, with a batch size of 32, using the Adam optimiser and a learning rate of 0.0001.

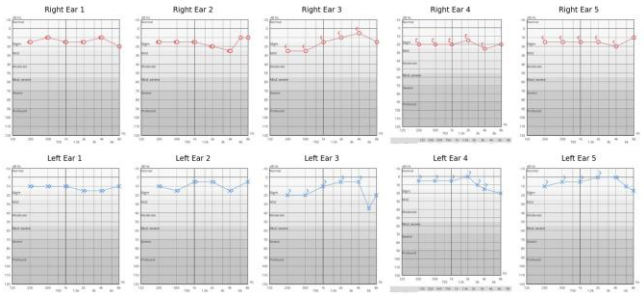


Figure 1. Dataset samples

The audiogram data presented in tabular format on Kaggle was obtained by converting each frequency–threshold (dB HL) pair into a graphical representation in accordance with the standard clinical audiogram format. For this purpose, the Matplotlib library was used in the Python environment. All generated audiogram images were rescaled to a resolution of 224×224 pixels to be compatible with the model input size.

3.2. Evaluation Metrics

TP (True Positive), FP (False Positive), FN (False Negative), TN (True Negative) values obtained from the confusion matrix. The correct classification ability and misclassification status of our model are calculated with Equation 1.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

The Accuracy value in Equation 1 shows the accuracy of our model. However, accuracy alone is not enough in cases where there is no equal distribution. Precision, Recall and F1-Score

values must also be considered for the accuracy value to make sense. Equation 2 shows the calculation of Precision.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

As can be seen from Equation 2, the Precision value allows us to find out how many of the values predicted as positive are actually positive. Precision is especially important when False Positive values are high.

Recall is a metric that shows how many of the values that we should have predicted as positive we actually predicted as positive. Equation 3 shows the formula used to calculate the Recall value.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

High Recall and Precision values are extremely important for our model to produce accurate results. The F1-Score value is obtained by taking the harmonic mean of these two values. The calculation of the F1-Score value is shown in Equation 4.

$$F1 - Score = 2 * \frac{precision * recall}{precision + recall} \tag{4}$$

4. RESULTS

In this section, the evaluation metric results obtained from machine learning models, training and loss graphs, and confusion matrices are presented and interpreted. The numerical results of the performance evaluation metrics are shown in Table 1. The Inceptionv3, DenseNet121, VGG19, Xception, and CNN algorithms, which have proven their success on images, were selected for training.

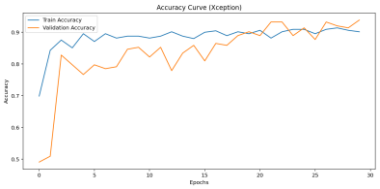
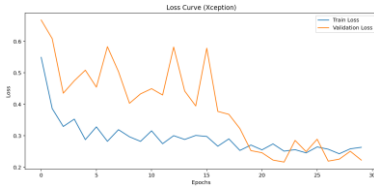
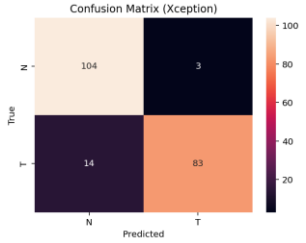
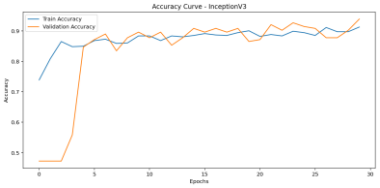
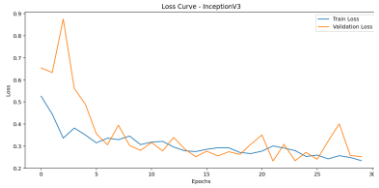
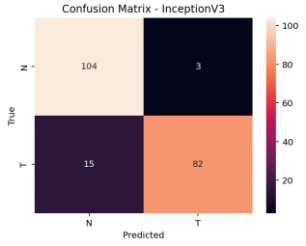
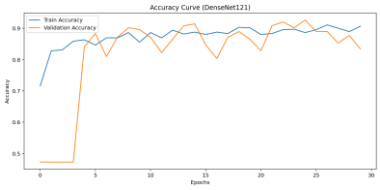
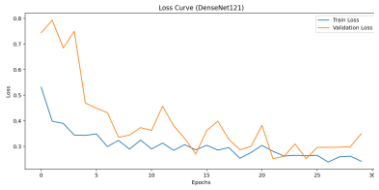
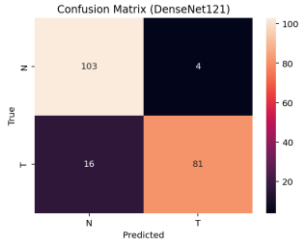
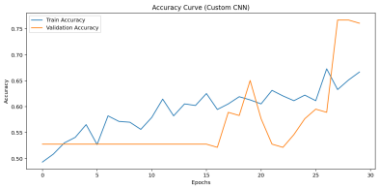
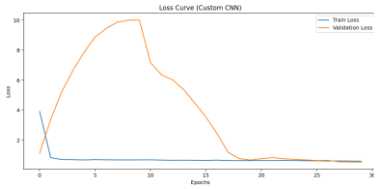
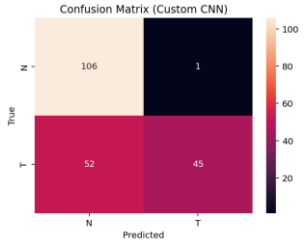
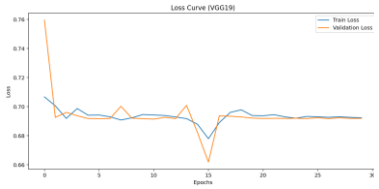
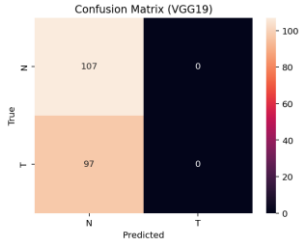
Precision Recall, F1-Score, Test Accuracy, Validation Accuracy, and Validation Loss were selected as evaluation metrics. These metrics are quite sufficient and provide objective results for observing the success of classification outcomes.

Table 1. Machine learning evaluation metric values

Algorithm & Metrics	Val_Accuracy	Val_Loss	Test Accuracy	Precision	Recall	F1-Score
Xception	0.9387	0.2221	91.6667	0.9280	0.9280	0.9280
Inceptionv3	0.9202	0.2445	91.1765	0.9200	0.9120	0.9120
DenseNet121	0.9202	0.2618	90.1961	0.9100	0.9000	0.9000
CNN	0.7607	0.5325	74.0196	0.8225	0.7450	0.7250
VGG19	0.5276	0.6919	52.4510	0.2800	0.5200	0.3600

When the results in Table 1 are evaluated overall, the validation and test performance of transfer learning-based deep architectures (InceptionV3, DenseNet121, Xception, VGG19) and a classical CNN model are compared. By considering both validation performance (Val_Accuracy) and test set performance (Test Accuracy) together, it is possible to comment on the generalisation capabilities of the models and potential overfitting/underfitting situations. The Xception architecture stands out as the most successful model in the table. The validation accuracy of 0.9387 and validation loss of 0.2221 indicate that the model was optimised during the training process with both high accuracy and relatively low error. The test accuracy of 91.6667 suggests that the model experienced only a limited performance loss compared to the validation set and therefore has strong generalisation capabilities. When examining the classification metrics, the fact that the Precision, Recall, and F1-Score values are all at the 0.9280 level and very close to each other indicates that the model performs consistently between false positive and false negative errors and that its ability to distinguish between classes is consistent. The InceptionV3 and DenseNet121 architectures can also be considered high-performance models very close to Xception. Both models have an accuracy of 0.9202, which is only approximately 1.8% lower than Xception. InceptionV3's validation loss is 0.2445, while DenseNet121's is 0.2618, indicating that both models demonstrate remarkably similar performance during the training process. When examining the test accuracies, InceptionV3 achieves 91.1765% accuracy and DenseNet121 achieves 90.1961% accuracy; this indicates that the validation–test gap is limited and that there is no significant overfitting in the models. In terms of the Precision–Recall–F1 triplet, values of 0.9200–0.9120–0.9120 were obtained for InceptionV3 and 0.9100–0.9000–0.9000 for DenseNet121. This indicates that both models have a balanced and satisfactory performance in terms of capturing true positives (Recall) and limiting false positives (Precision), while Xception has a marginal but significant advantage over the other two. The classic CNN model exhibits lower performance compared to higher-level architectures based on transfer learning. With a validation accuracy of 0.7607 and a validation loss of 0.5325, these values suggest that the CNN cannot learn the complex patterns in the dataset as effectively as deep and wide architectures. The fact that the test accuracy remains at 74.0196% also supports this interpretation. On the other hand, the relatively high Precision value of 0.8225 indicates that the model is more cautious when predicting the positive class (relatively limiting false positives), while the lower Recall (0.7450) and F1-Score (0.7250) values indicate that some true positive examples are missed. The VGG19 architecture, however, stands out significantly from other models in terms of both validation and test performance, yielding poor results. The validation accuracy of 0.5276 and validation loss of 0.6919 indicate that the model exhibits behaviour close to random classification on the dataset. The test accuracy remaining around 52.4510% also confirms this situation. When examining the classification metrics, Precision is seen to be at a very low level of 0.2800, while the Recall value is 0.5200. This table shows that the model has a high number of false positives but nevertheless fails to capture a significant portion of the true positives; consequently, the F1-Score remains at a low level of 0.3600. Table 2 shows the accuracy, loss and confusion graphs of the machine learning models.

Table 2. Machine learning evaluation metric values

Algs.	Accuracy	Loss	Confusion M.
Xception			
InceptionV3			
DenseNet121			
CNN			
VGG19			

When the accuracy, loss, and complexity matrix graphs presented for each model in Table 2 are evaluated together, it is observed that the Xception, InceptionV3, and DenseNet121 architectures exhibit a stable learning trend throughout the training process. In Xception, the training and validation accuracy curves rise rapidly after a few epochs, settling above 90%, and remain quite close to each other as the epochs progress. In parallel, the training and validation loss decreased steadily, with only minor fluctuations observed. The confusion matrix also shows that correct classifications are concentrated on the diagonal, while incorrect classifications remain quite limited in number.

A similar picture emerges for InceptionV3 and DenseNet121: in both models, the accuracy curves show a rapid increase in the early epochs, then reach relative saturation; the loss curves, on the other hand, show a steady decrease, indicating successful optimization. Although the complexity matrices of these two models show slightly more errors in the off-diagonal cells compared to Xception, overall, the positive and negative classes are largely correctly separated.

In contrast, it is noteworthy that both the learning dynamics and final classification performance of the classical CNN and VGG19 architectures are more problematic. In the accuracy graph for CNN, while training accuracy increases, validation accuracy remains more limited and fluctuates in places; in the loss graph, training loss rapidly drops to very low values, while validation loss remains high for a long time. This suggests that the model overfits the training data and has weaker generalization capacity.

The confusion matrix for the CNN also reveals that correct classifications are fewer compared to Xception InceptionV3 DenseNet121, while incorrect classifications are significantly more numerous. For VGG19, the accuracy curves exhibit high-amplitude fluctuations, i.e., a highly irregular trajectory, and show no clear improvement trend; the loss curves are similarly unstable. The confusion matrix for VGG19 shows that a significant portion of the classes are mislabeled and that the model's discriminative power remains quite low; this indicates that architecture fails to learn the problem under the current dataset and training settings.

4. DISCUSSION AND CONCLUSIONS

In this study, tinnitus detection was investigated using audiogram images and multiple machine learning architectures, and the results clearly indicate the superiority of transfer-learning–based deep models over a conventional CNN approach. Xception, followed closely by InceptionV3 and DenseNet121, provided high validation and test accuracies together with balanced precision, recall, and F1-score values. This pattern shows that these architectures can generalize well to unseen audiogram data and capture clinically meaningful differences between tinnitus and non-tinnitus cases. In contrast, the weak and unstable performance of VGG19 underlines that model choice must be problem-specific and that deeper networks do not automatically guarantee better results.

Compared with many studies in the literature that rely on EEG-based biomarkers or complex neurophysiological recordings for tinnitus detection, the present work focuses on standard audiogram information routinely collected in clinical practice. This constitutes an important practical advantage: the proposed approach can be implemented without additional hardware, does not change existing diagnostic workflows, and can be integrated into hospital information systems as a low-cost decision-support module. In future work, the robustness of the models can be increased by enlarging the dataset with multi-centre audiogram records, applying more advanced data augmentation strategies, and exploring ensemble approaches that combine audiogram-based predictions with EEG or questionnaire-based information. Overall, the findings suggest that deep learning applied to audiogram images is a promising direction for developing objective, reproducible, and scalable tools to support clinicians in tinnitus diagnosis and follow-up.

Acknowledgements
It should be written as short as possible and expressing the contribution made without giving the number.

Ethics Committee Approval
N/A

Peer-review
Externally peer-reviewed.

Author Contributions
Conceptualization: R.G., Y.T..; Investigation: R.G., Y.T..; Material and Methodology: R.G.; Supervision: R.G.; Visualization: R.G.; Writing-Original Draft: Y.T., R.G.; Writing-review & Editing: R.G., Y.T. Other: All authors have read and agreed to the published version of manuscript.

Conflict of Interest
The authors have no conflicts of interest to declare.

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

REFERENCES / KAYNAKLAR

Allgaier, J., Neff, P., Schlee, W., Schoisswohl, S., & Pryss, R. (2021). Deep Learning End-to-End Approach for the Prediction of Tinnitus based on EEG Data. In *Proceedings of the 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*

Awais, M., Kassoul, K., El Omri, A. E., Aboumarzouk, O. M., Abdulhadi, K., & Brahim Belhaouari, S. (2024). Graph-Based Electroencephalography Analysis in Tinnitus Therapy. *Biomedicines*, *12*(7), 1404

Baguley, D., McFerran, D., & Hall, D. (2013). Tinnitus. *The Lancet*, *382*(9904), 1600-1607.

Couth, S., Mazlan, N., Moore, D. R., Munro, K. J., & Dawes, P. (2019). Hearing difficulties and tinnitus in construction, agricultural, music, and finance industries: contributions of demographic, health, and lifestyle factors. *Trends in hearing*, *23*, 2331216519885571.

Hong, E.-S., Kim, H.-S., Hong, S. K., Pantazis, D., & Min, B.-K. (2023). Deep learning-based electroencephalic diagnosis of tinnitus symptom. *Frontiers in Human Neuroscience*, *17*, 1126938

Huang, L., Chen, M., Fan, S., Tursun, N., He, X., Li, W., & Li, S. (2025). Machine learning identification of tinnitus-related features in auditory peripheral spontaneous activity in a guinea pig noise-induced tinnitus model. *Hearing Research*, *465*, 109371

Kaggle, 2024. Audiogram Analysis for Tinnitus Detection. *https://www.kaggle.com/datasets/kavehmcsd/audiogram-analysis-for-tinnitus-detection/data. Accessed: 25-11-16.*

Kleinjung, T., Peter, N., Schecklmann, M., & Langguth, B. (2024). The current state of tinnitus diagnosis and treatment: a multidisciplinary expert perspective. *Journal of the Association for Research in Otolaryngology*, *25*(5), 413-425.

Mai, J., Wang, X., Li, Z., Liu, J., Zhang, Z., & Fu, H. (2023). EEG signal classification of tinnitus based on SVM and sample entropy. *Computer Methods in Biomechanics and Biomedical Engineering*, *26*(5), 580–594.

Møller, A. R. (2007). Tinnitus: presence and future. *Progress in brain research*, *166*, 3-16.

Ogawa, K., Sato, H., Takahashi, M., Wada, T., Naito, Y., Kawase, T., ... & Kanzaki, S. (2020). Clinical practice guidelines for diagnosis and treatment of chronic tinnitus in Japan. *Auris Nasus Larynx*, *47*(1), 1-6.

Pryce, H., Dauman, N., & Burns-O'Connell, G. (2023). What is the burden of tinnitus?. *Frontiers in Psychology*, *13*, 981776.

Savage, J., & Waddell, A. (2012). Tinnitus. *BMJ clinical evidence*, *2012*, 0506.

Simoes, J. P., Daoud, E., Shabbir, M., Amanat, S., Assouly, K., Biswas, R., ... & Genitsaridi, E. (2021). Multidisciplinary tinnitus research: Challenges and future directions from the perspective of early-stage researchers. *Frontiers in aging neuroscience*, *13*, 647285.

Tang, D., Li, H., & Chen, L. (2019). Advances in understanding, diagnosis, and treatment of tinnitus. *Hearing loss: Mechanisms, prevention and cure*, 109-128.

Watts, E. J., Fackrell, K., Smith, S., Sheldrake, J., Haider, H., & Hoare, D. J. (2018). Why is tinnitus a problem? A qualitative analysis of problems reported by tinnitus patients. *Trends in hearing*, *22*, 2331216518812250.

Zhang, C., Wang, F., Ding, Z., Liu, P., Xue, X., Wang, L., Jiang, Y., Zhang, Z., Guo, X., Lu, Q., Liu, J., Peng, X., Ma, Y., Chen, J., Shen, W., & Yang, S. (2026). XTinnitusNet: Multi-view robust model ensemble for EEG-based tinnitus diagnosis with noisy labels. *Biomedical Signal Processing and Control*, *112*, 108434