International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

AI-POWERED OTOSCOPIC IMAGE CLASSIFICATION FOR EAR DISEASE DETECTION.

Ms. A.Suhana [a] , M.Bavatharani [b] , V.Keerthana [c] , P.Kavibharathi[d]

[a] Assistant Professor, Department Of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, India.
[b, c, d]Student, Department Of Computer Science and Engineering, Dhirajlal Gandhi College of Technology, India.

ABSTRACT:

Otoscopy is a diagnostic procedure used to examine the external auditory canal and the eardrum (tympanic membrane) using an otoscope. This handheld instrument consists of a light source and a magnifying lens or camera that allows healthcare professionals to visualize the structures of the ear canal and the eardrum. Otoscopy is commonly performed to assess ear health, identify abnormalities such as infections, inflammation, wax buildup, foreign bodies, structural defects, tumors, or signs of trauma. It is an essential part of the evaluation process for patients with ear-related symptoms like pain, hearing loss, or discharge. Otoscopy is typically performed as part of a comprehensive ear examination and may be accompanied by other tests such as tympanometry or audiometry to further assess ear function. Otoscopy plays a crucial role in diagnosing various ear pathologies, yet accurate and timely classification of otoscopy images remains a challenging task. In this study, we propose a deep learning-based approach for automated otoscopy classification. We curated a diverse dataset comprising otoscopy images encompassing normal anatomy and a spectrum of pathologies, including infections, inflammations, tumours, and structural abnormalities. We employed Convolutional Neural Networks (CNNs) with Pre trained model named as DenseNet framework, a powerful class of deep learning models, for feature extraction and classification. The dataset was pre-processed to enhance uniformity and augment diversity, and the model was trained with categorical cross-entropy loss. Hyperparameters were fine-tuned to optimize performance, and the model was evaluated using standard metrics including accuracy, precision, recall, F1-score, and AUC-ROC. Our results demonstrate the effectiveness of the proposed approach in accurately classifying otoscopy images, showcasing its potential as a valuable tool in clinical practice for aiding in the diagnosis of ear pathologies.

Keywords: Artificial neural networks, Dense network, Deep Learning, Machine learning, Otoscopy images.

1. INTRODUCTION

Ear-related disorders, particularly those affecting the middle and inner ear, are among the most prevalent health concerns worldwide. Otoscopy, a fundamental diagnostic tool used by clinicians to examine the ear canal and eardrum, plays a crucial role in identifying infections, perforations, and other abnormalities. However, accurate diagnosis often requires specialized expertise, and misinterpretations can lead to incorrect treatments, worsening patient conditions. With the rapid advancement of artificial intelligence (AI) and deep learning, automated diagnostic systems have emerged as promising solutions to support healthcare professionals in achieving higher diagnostic accuracy and efficiency. Deep learning, particularly Convolutional Neural Networks (CNNs), has shown remarkable success in medical image analysis by learning complex features and patterns from large datasets. By leveraging CNN-based models, an automated otoscopy classification system can be developed to classify otoscopic images into different categories of ear conditions. Such a system can aid in early detection, reduce human diagnostic errors, and ensure faster clinical decision-making. Additionally, image preprocessing techniques, including noise reduction and contrast enhancement, can further improve model robustness, making AI-driven diagnostics a reliable complement to traditional otoscopic examination. The integration of AI into otoscopy not only enhances diagnostic precision but also expands accessibility to quality healthcare, particularly in remote and underserved areas where specialized medical expertise may be limited. A user-friendly interface incorporating the trained model can assist general practitioners, telemedicine platforms, and even individuals in performing preliminary assessments before consulting an ear specialist.

2. LITERATURE SURVEY

2.1EAR DISEASE DETECTION USING R-CNN

This study explores the application of Region-based Convolutional Neural Networks (R-CNN) for automated ear disease detection. The authors (Anandamurugan et al., 2022**)** propose a deep learning model trained on a dataset of otoscopic images to identify various ear conditions. The paper discusses the preprocessing techniques used to enhance image quality and segment the tympanic membrane region. The study evaluates the performance of R-CNN in detecting common ear diseases such as otitis media, cholesteatoma, and tympanic membrane perforation. The research highlights the advantages of R-CNN in extracting spatial features and localizing disease-specific regions in images. The model's accuracy is validated against expert-labeled datasets, showcasing competitive results compared to traditional diagnostic methods. The authors compare R-CNN with other object detection frameworks, such as Faster R-CNN and YOLO, demonstrating its efficiency in precise region-based classification. The study also explores the integration of AI-assisted diagnostic tools with telemedicine applications, allowing for remote disease screening. Performance metrics such as accuracy, precision, recall, and F1-score are used to evaluate the system.

2.2. EVALUATION OF HUMAN EAR ANATOMY AND FUNCTIONALITY BY AXIOMATIC DESIGN

This study presents an axiomatic design framework for evaluating human ear anatomy and functionality. The research (Sundar et al., 2021) aims to create a structured approach for analyzing the relationship between ear morphology and its function. The authors employ a biomimetics approach to study the ear's structural design and its role in sound transmission. The study introduces a functional decomposition model, breaking down ear components such as the pinna, tympanic membrane, ossicles, and cochlea to assess their contribution to auditory processing. The authors analyze biomechanical properties, including sound wave propagation, vibration response, and fluid dynamics in the cochlea. The paper discusses how structural abnormalities and deformities can impact hearing efficiency and contribute to auditory disorders. The study incorporates computational modeling techniques to simulate ear mechanics and predict functional impairments. The research also explores the application of machine learning algorithms for automated analysis of anatomical variations and their correlation with hearing disorders. The findings highlight the significance of design principles in medical diagnostics, particularly in developing hearing aids and prosthetic ear implants.

3. SYSTEM STUDY

3.1. EXISTING SYSTEM

The current approach to diagnosing ear-related disorders primarily relies on manual visual inspection conducted by healthcare professionals using a traditional otoscope. This method involves direct examination of the ear canal and tympanic membrane to detect abnormalities such as infections, perforations, wax accumulation, and structural issues. While this process is fundamental in clinical settings, its effectiveness depends heavily on the practitioner's experience and skill. Inconsistent diagnostic skills among practitioners can lead to varying interpretations of similar symptoms, which increases the risk of misdiagnosis and improper treatment plans for patients. In rural and remote areas, the challenge becomes even more significant due to the limited availability of specialized ear, nose, and throat (ENT) doctors. General practitioners, who may not have extensive training in otoscopic evaluation, are often the only point of contact for patients. As a result, delays in identifying and treating ear conditions can occur, leading to complications such as chronic infections, persistent hearing loss, or unnecessary medication usage. The lack of access to high-quality diagnostic tools and professional expertise in these regions further exacerbates the problem, highlighting the limitations of the traditional manual otoscopy system. Some existing computer-aided diagnostic (CAD) systems have attempted to support clinicians by using traditional machine learning algorithms like Support Vector Machines (SVM) and decision trees. These systems typically rely on handcrafted features for analysis, which makes them less adaptable to real-world variability in otoscopic images. Factors such as image noise, lighting inconsistencies, and anatomical variations often affect the accuracy of these models. Moreover, the absence of large, well-labeled datasets further limits their training potential and generalization capabilities. Consequently, the existing systems struggle to deliver reliable results, emphasizing the need for a more advanced, automated, and intelligent diagnostic framework.

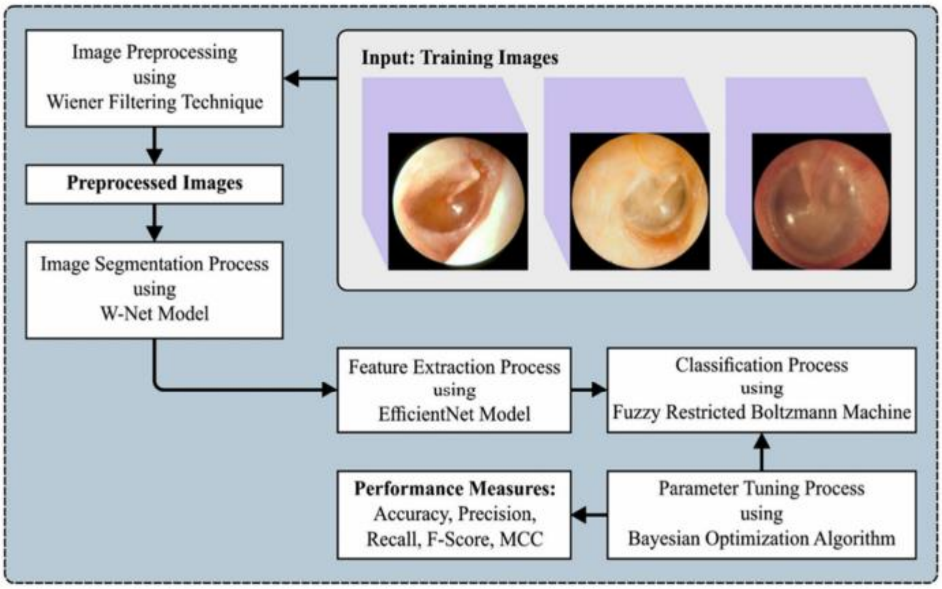

Figure 3.1.1: Existing system

3.2. PROPOSED SYSTEM

The proposed system introduces an AI-powered otoscopic image classification approach leveraging deep learning, specifically Convolutional Neural Networks (CNNs), to enhance the accuracy and reliability of ear disease diagnosis. Unlike conventional diagnostic methods, which depend on clinician expertise or handcrafted features, CNNs can automatically learn and extract meaningful patterns from raw otoscopic images. These models are trained on a diverse and labeled dataset, enabling them to accurately identify various ear conditions, such as infections, inflammations, structural abnormalities, and tumors. By utilizing advanced feature extraction and classification capabilities, the proposed system significantly reduces human error and improves diagnostic consistency. To ensure optimal model performance, image pre-processing techniques such as noise reduction, contrast enhancement, and artifact removal are applied. These processes help standardize the input images and mitigate the impact of lighting variations, wax accumulation, and other visual obstructions that typically hinder classification accuracy. Hyperparameter tuning and performance evaluation using metrics like accuracy, F1-score, AUC-ROC, precision, and recall ensure the robustness and reliability of the model in diverse clinical scenarios. The system is designed to generalize well across different populations, making it suitable for real-world application in both urban and rural healthcare settings. Furthermore, the system is implemented as a user-friendly web or mobile application, allowing healthcare practitioners and remote users to upload otoscopic images and receive immediate classification results. Along with the diagnosis, the model provides confidence scores to support clinical decision-making. A secure backend database stores patient records and diagnostic outcomes, enabling continuous monitoring and follow-up. This proposed system not only facilitates early detection and timely treatment but also empowers telemedicine platforms and under-resourced healthcare providers to deliver high-quality care efficiently.

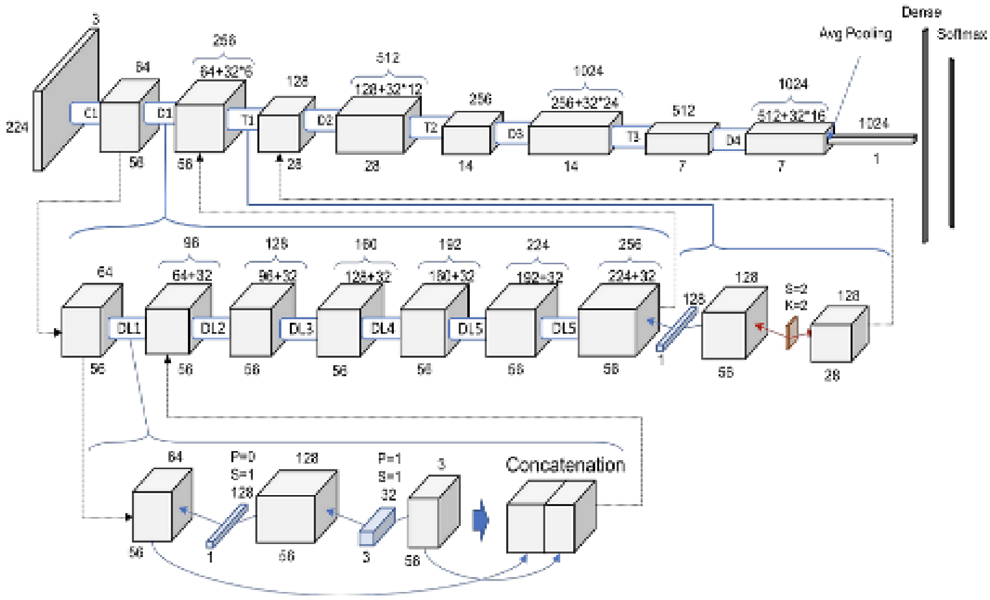

Figure 3.2.1: Proposed system

Category	Existing System	Proposed System
Diagnosis Accuracy	Manual, Error-Prone	AI-Based, Accurate
Automation Support	No Automation, Manual Process	Automated, CNN-Driven

4. METHODOLOGY

4.1 DATA COLLECTION

Otoscopic images were collected from publicly available datasets and verified clinical repositories. The dataset includes a wide range of ear conditions such as:

- Normal tympanic membranes
- Acute Otitis Media (AOM)
- Chronic Otitis Media
- Tympanosclerosis
- Foreign objects
- Inflammations and tumors

This ensures a diverse and balanced dataset, crucial for training a generalized deep learning model.

4.2 PREPROCESSING

To enhance image quality and model input uniformity, the following preprocessing techniques were applied:

- Resizing all images to a fixed resolution.
- Noise reduction using median filtering to eliminate random variations.
- Contrast enhancement via histogram equalization.
- Normalization of pixel values to a standard range.
- Data augmentation (rotation, flipping, brightness variation) to expand the dataset and prevent overfitting.

4.3 MODEL SELECTION: DENSENET FRAMEWORK

The DenseNet architecture was selected due to its ability to reuse features and improve gradient flow through dense connections. Each layer receives input from all preceding layers, promoting efficient feature extraction.
Key components of the architecture include:

- Dense Blocks: Multiple convolution layers connected to each other to capture intricate patterns.
- Transition Layers: Employed for down-sampling while retaining essential information.
- Fully Connected Layers: Serve as classifiers for determining the disease category.

MODEL TRAINING

- The model was trained using the categorical cross-entropy loss function.
- Adam optimizer was used for weight updates.
- Hyperparameters such as learning rate, batch size, and number of epochs were fine-tuned using grid search**.**
- The dataset was split into training (70%), validation (15%), and testing (15%) subsets.

EVALUATION METRICS

Model performance was evaluated using the following metrics:

- Accuracy
- Precision
- Recall
- F1-Score
- AUC-ROC Curve

These metrics ensure a comprehensive understanding of the model's classification capabilities, especially for imbalanced classes.

4.6 REAL-TIME INFERENCE SYSTEM

A user-friendly interface was developed for real-time inference. Users can upload otoscopic images, and the system classifies them into the appropriate disease category, displaying a confidence score and storing results for follow-up.

5. MODULES IMPLEMENTATION

5.1 LIST OF MODULES

- Image Collection and Preprocessing
- Build the CNN Model
- Evaluation Metrics
- Disease Classification
- Real-Time Interference

5.2 MODULES DESCRIPTION

IMAGE COLLECTION AND PREPROCESSING

In this module, otoscopic images are collected from various sources, including clinical settings and public datasets, to form a diverse and comprehensive dataset. These images may contain noise, varying lighting conditions, and artifacts like wax or reflections, which can affect classification accuracy. To ensure consistent input for the model, image preprocessing techniques are applied. These include resizing the images to a uniform dimension, normalizing pixel values to a standard range, and enhancing contrast to improve image clarity. Noise reduction algorithms are also used to eliminate irrelevant information, while techniques like histogram equalization may be applied to improve visual quality. The goal of this module is to prepare high-quality images that allow the deep learning model to extract relevant features efficiently, reducing the impact of distortions and improving classification performance.

BUILD THE CNN MODEL

The CNN model is built to automatically learn complex patterns and features from the otoscopic images. It is designed with multiple convolutional layers followed by pooling layers to extract hierarchical features from the images. The convolutional layers apply various filters to detect edges, textures, and other features in the images, while the pooling layers reduce the dimensionality and retain the most significant features. The final layers of the CNN model include fully connected layers that perform the classification task, mapping the extracted features to specific ear conditions, such as infections, inflammations, or normal anatomy. During training, the model uses labeled data to optimize its parameters through backpropagation, employing categorical cross-entropy loss as the objective function. The CNN's architecture is fine-tuned to maximize accuracy and ensure generalization across diverse otoscopic images.

.

EVALUATION METRICS

To assess the performance of the CNN model, various evaluation metrics are employed to quantify its accuracy and reliability. Key metrics include accuracy, precision, recall, F1-score, and AUC-ROC. Accuracy measures the overall correctness of the model's predictions, while precision and recall evaluate its ability to correctly identify true positive cases and avoid false negatives. F1-score provides a balance between precision and recall, ensuring the model maintains both high sensitivity and specificity. The AUC-ROC curve evaluates the model's performance across different classification thresholds, measuring its discriminative ability. These metrics are calculated on a validation dataset during the training process to track model progress and prevent overfitting. The results from these evaluation metrics guide the adjustment of hyperparameters and help refine the model for optimal performance.

DISEASE CLASSIFICATION

The disease classification module is responsible for categorizing otoscopic images into different ear conditions, such as infections, tumors, inflammations, and normal anatomy. Once the CNN model has been trained on the pre-processed dataset, it is used to classify new otoscopic images based on the learned features. The model processes the input image and outputs a predicted diagnosis, typically in the form of a probability distribution over the possible classes. The class with the highest probability is chosen as the predicted condition. In cases where the model's prediction is uncertain, the confidence score associated with the prediction helps healthcare professionals assess the likelihood of the diagnosis. This module plays a critical role in supporting clinical decision-making by providing fast, reliable, and consistent classifications, particularly for conditions that may be difficult to diagnose visually.

REAL-TIME INTERFERENCE

The real-time inference module enables the system to process and classify otoscopic images instantly as they are uploaded by users or healthcare professionals. Through a web or mobile interface, users can upload otoscopic images, which are then pre-processed, passed through the trained CNN model, and classified into specific ear conditions. The system provides immediate diagnostic results along with confidence scores, helping clinicians make quick decisions. This module is designed to work efficiently in clinical settings, reducing the time required for diagnosis and enabling faster patient care. It also ensures accessibility in remote areas where specialists may not be readily available, providing valuable diagnostic support to general practitioners.

6. SYSTEM ARCHITECTURE

The architecture of the proposed AI-powered otoscopic image classification system begins with the collection and preprocessing of otoscopic images, where noise is reduced and image quality is enhanced to improve model input. These pre-processed images are then fed into a Convolutional Neural Network (CNN) model, which automatically extracts key features and patterns relevant to different ear conditions. The CNN model classifies the images into categories such as normal, infection, inflammation, or tumor. The system provides the predicted diagnosis along with confidence scores and stores the results in a secure database, enabling real-time inference, continuous patient monitoring, and accessibility through a web or mobile interface for healthcare professionals and remote users.

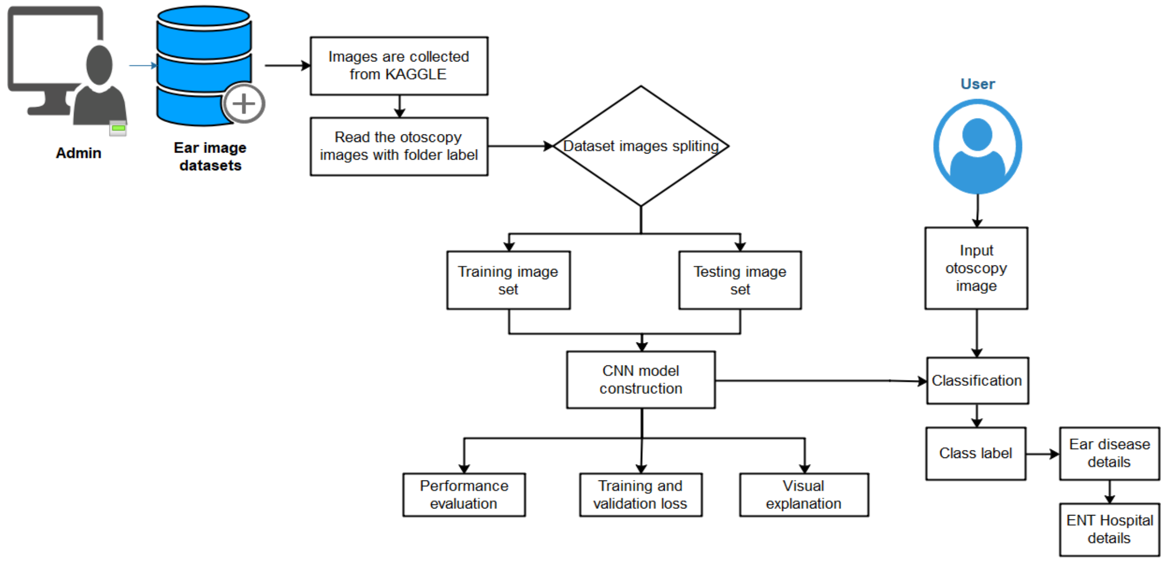

Figure 4.1: System Architecture

7. EXPERIMENTAL RESULTS

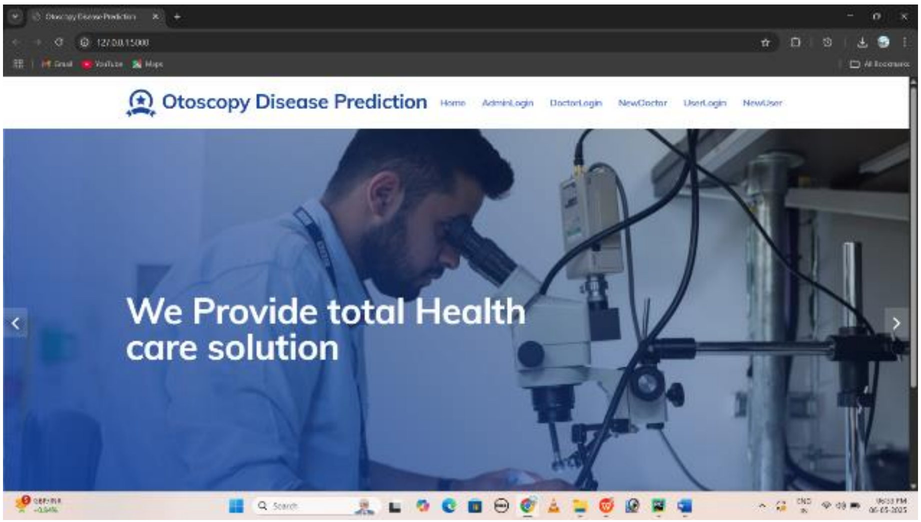

Figure 6.1: Home Page

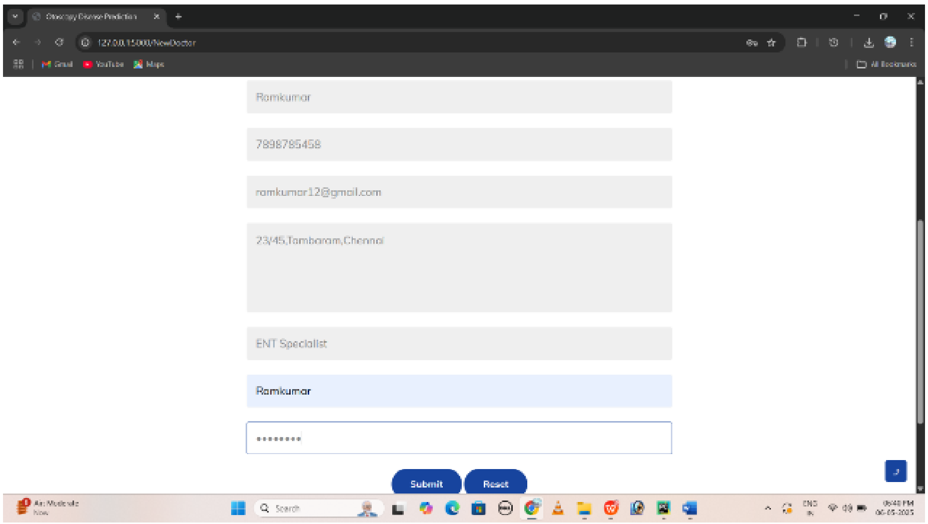

Figure 6.2: Doctor Registration Page

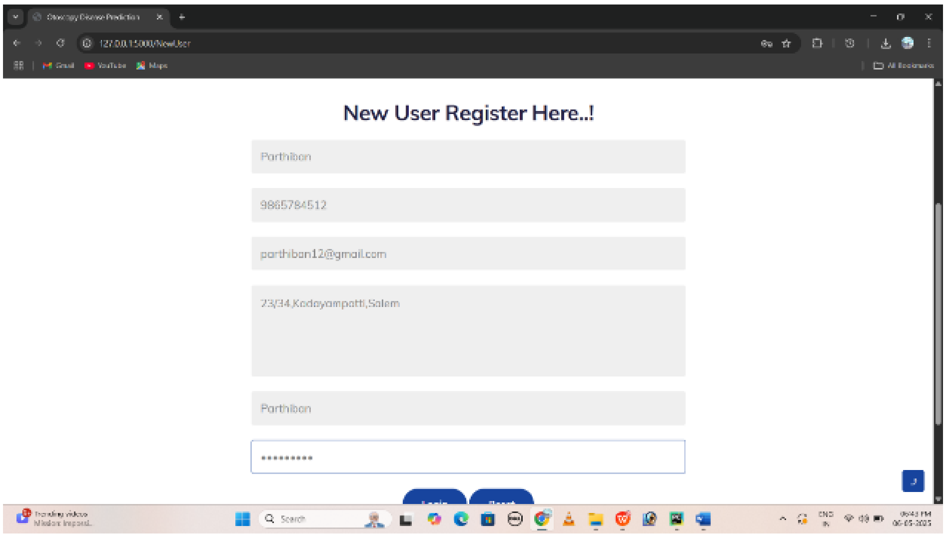

Figure 6.3: User Registration Page

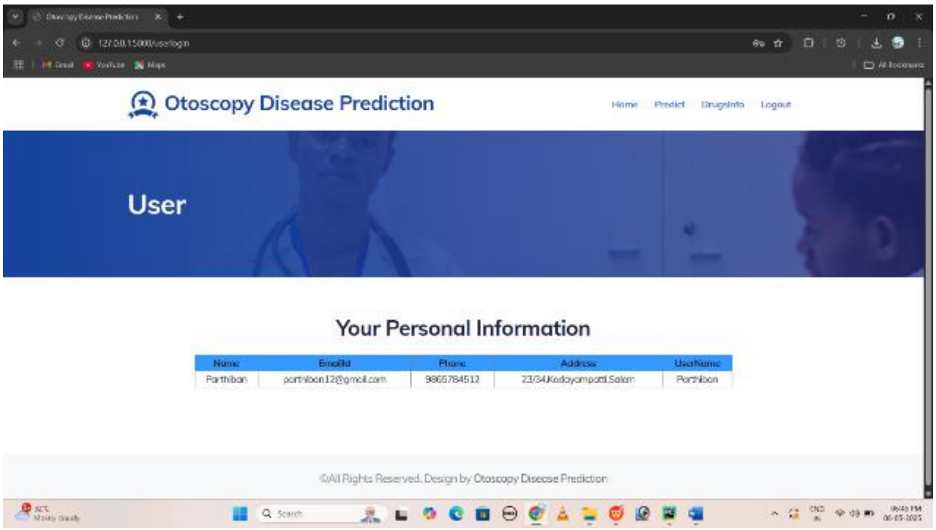

Figure 6.4: User Information Page

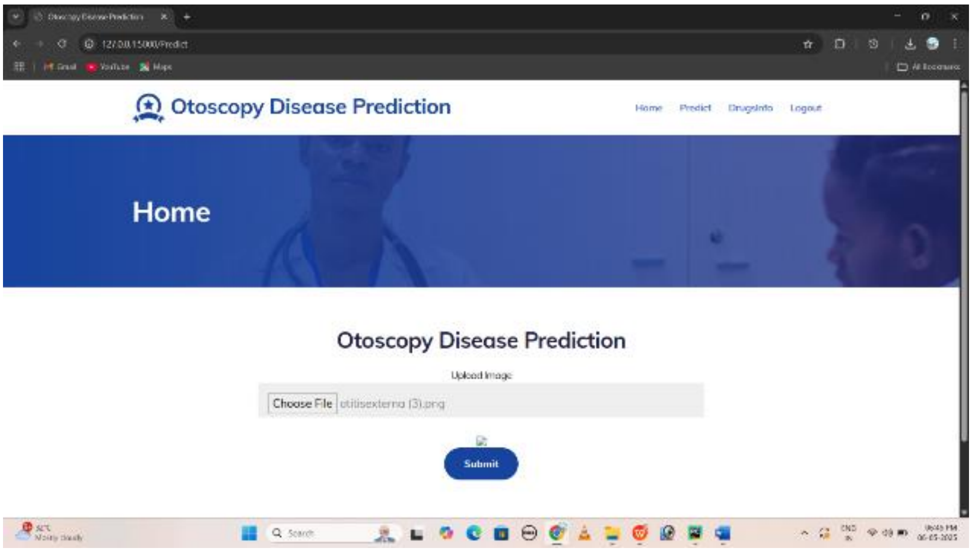

Figure 6.5: Image Uploading Page

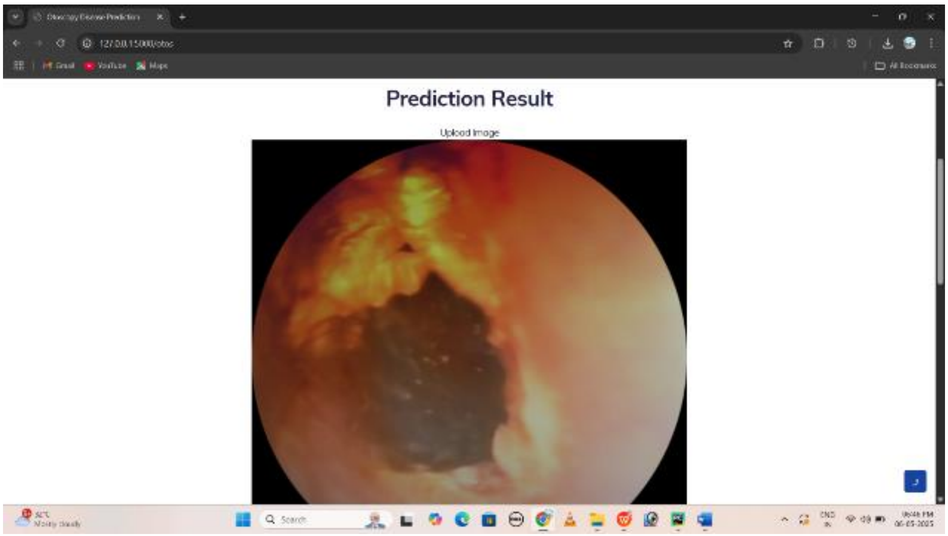

Figure 6.6: Prediction Result Page

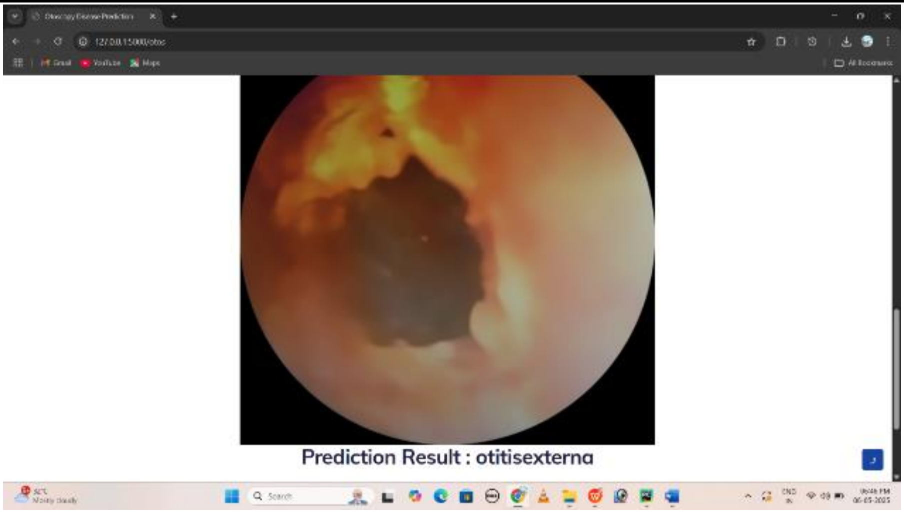

Figure 6.7: Disease Prediction Page

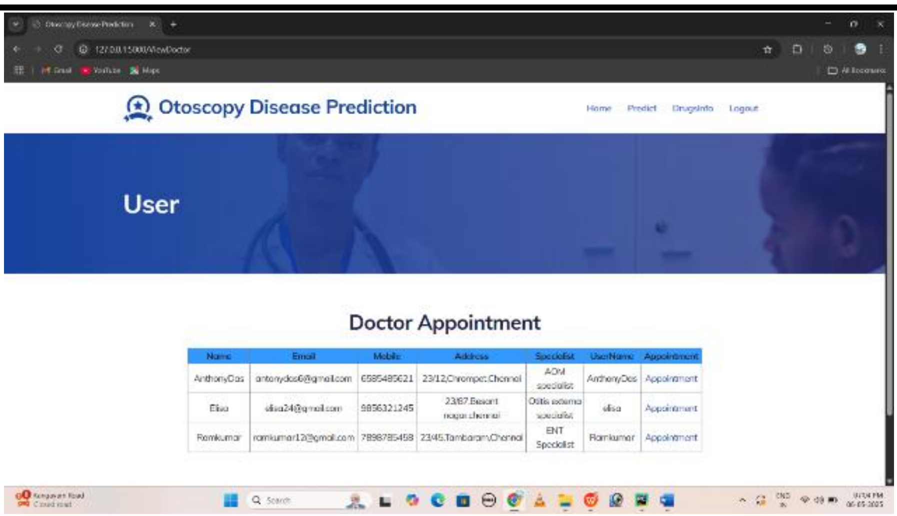

Figure 6.8: Doctor Appointment Page

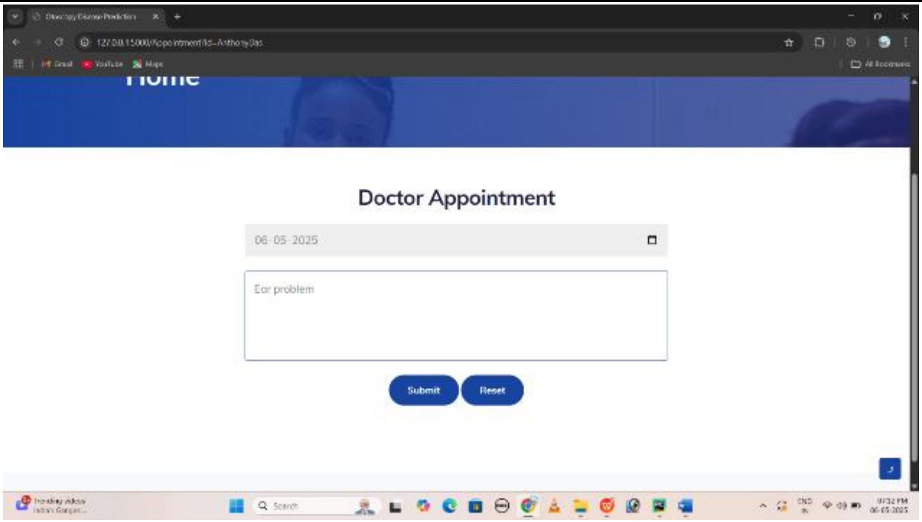

Figure 6.9 Doctor Appointment Booking Page

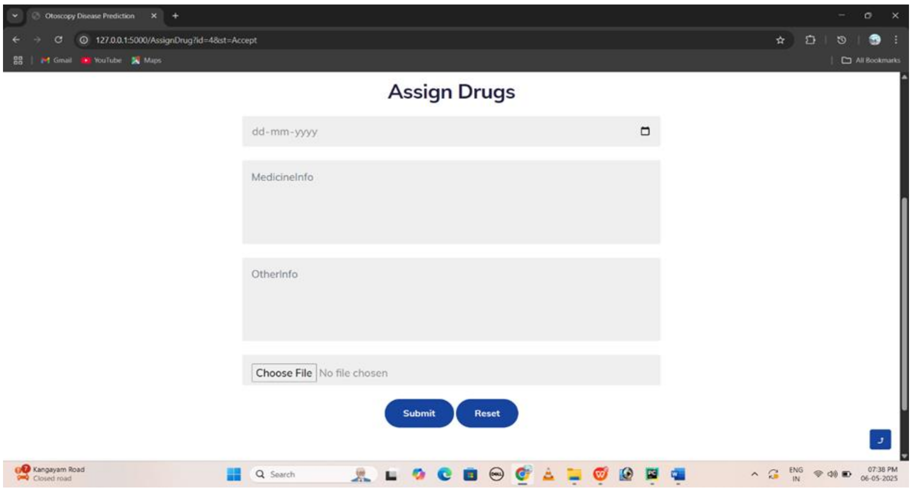

Figure 6.10 Drugs Assigning page

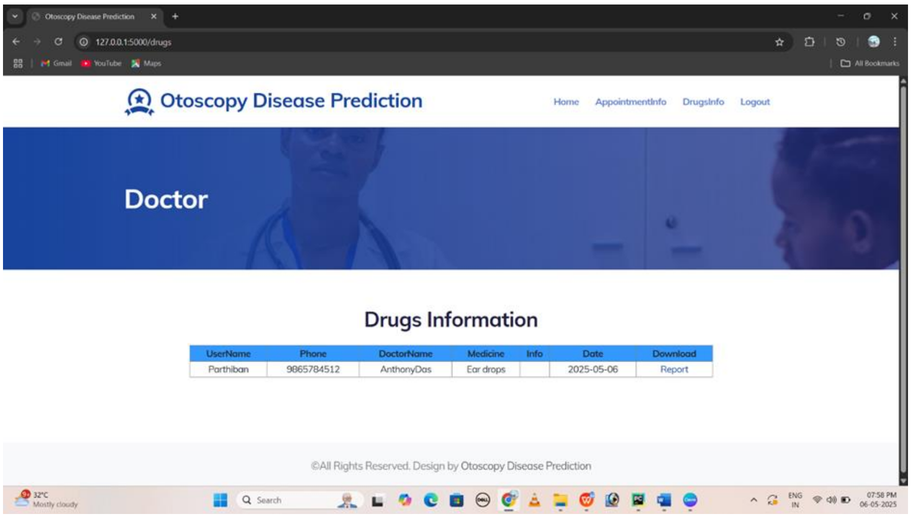

Figure 6.11 Drugs Information page

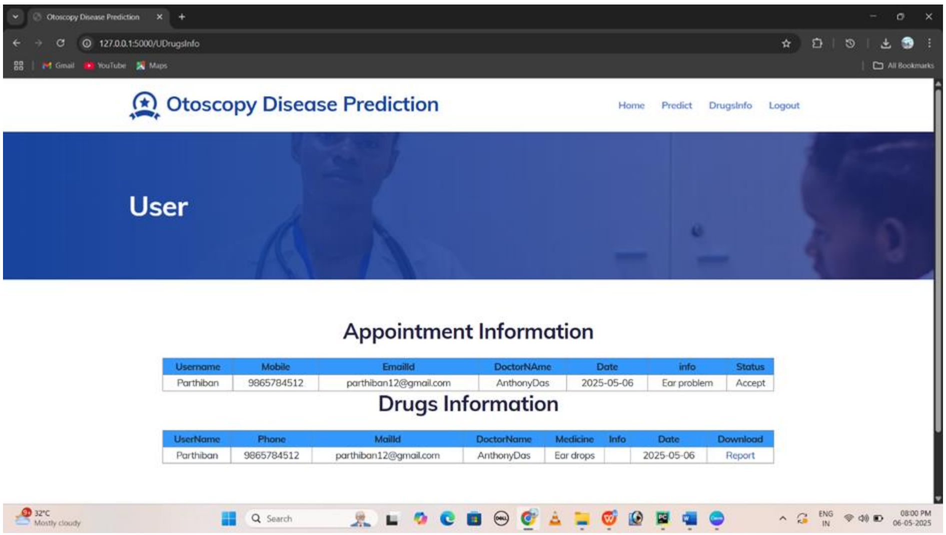

Figure 6.12 Appointment and Drugs information

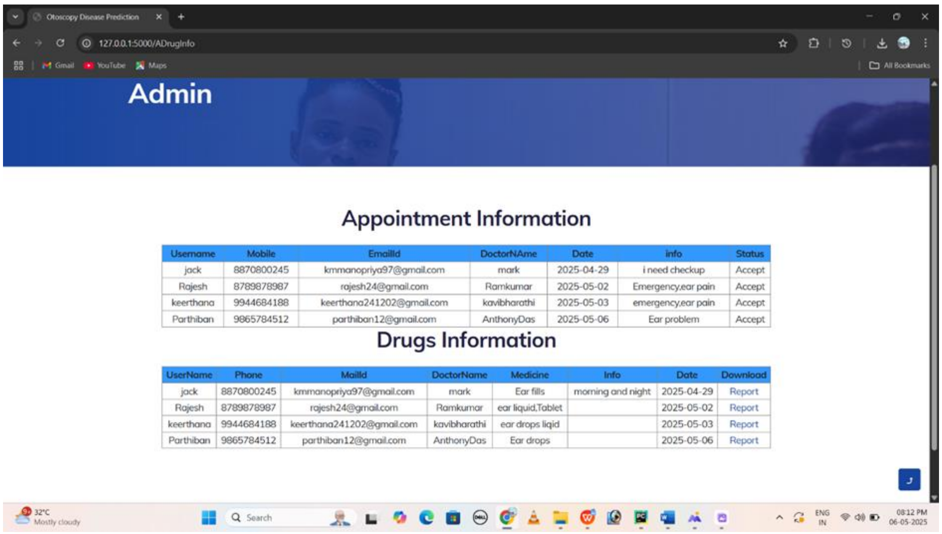

Figure 6.13 Admin Page

CONCLUSION AND FUTURE ENHANCEMENTS

CONCLUSION

The proposed AI-powered otoscopic image classification system demonstrates significant potential in enhancing the diagnostic accuracy and efficiency of ear disease detection. By leveraging Convolutional Neural Networks (CNNs) for feature extraction and classification, the system can process otoscopic images and reliably identify a range of ear conditions, including infections, inflammations, tumors, and normal anatomy. This approach significantly reduces the reliance on manual interpretation, which is often subject to human error, and addresses the limitations of traditional otoscopy methods. With image preprocessing techniques enhancing the quality of input data, the system provides more consistent and accurate results, ultimately supporting healthcare professionals in making quicker, more informed decisions. Moreover, the integration of this system into clinical practice, particularly in remote or underserved areas, offers a promising solution to bridge the gap in healthcare accessibility. By providing real-time image classification through a user-friendly interface, the system can aid both general practitioners and specialists in diagnosing ear conditions swiftly, thereby improving patient outcomes. The ability to store patient records securely and track diagnostic progress over time further enhances the utility of the system. In conclusion, this AI-driven solution has the potential to transform the diagnosis of ear diseases, making it more reliable, accessible, and efficient, thereby contributing to the overall improvement of ear healthcare.

FUTURE ENHANCEMENTS

[1] Anandamurugan, S., et al. "Ear disease detection using R-CNN." 2022 Fifth International Conference on Computational Intelligence and Communication Technologies (CCICT). IEEE, 2022.
[2] Sundar, Pratap Sriram, Chandan Chowdhury, and Sagar Kamarthi. "Evaluation of human ear anatomy and functionality by axiomatic design." Biomimetics 6.2 (2021): 31.
[3] Cai, Yuexin, et al. "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 11.1 (2021): e041139.
[4] Camalan, Seda, et al. "OtoPair: Combining right and left eardrum otoscopy images to improve the accuracy of automated image analysis." Applied Sciences 11.4 (2021): 1831.
[5] Mohammed, Kamel K., Aboul Ella Hassanien, and Heba M. Afify. "Classification of ear imagery database using Bayesian optimization based on CNN-LSTM architecture." Journal of Digital Imaging 35.4 (2022): 947-961.
[6] Song, Dahye, et al. "Image-based artificial intelligence technology for diagnosing middle ear diseases: a systematic review." Journal of Clinical Medicine 12.18 (2023): 5831.
[7] Zeng, X. "Efficient and accurate identification of ear diseases using an ensemble deep learning model. Sci. Rep. 11 (1), 10839 (2021)."
[8] Metcalfe, Christopher, et al. "A systematic review of remote otological assessment using video-otoscopy over the past 10 years: reliability and applications." European Archives of Oto-Rhino-Laryngology 278 (2021): 4733-4741.
[9] Viscaino, Michelle, et al. "Computer-aided diagnosis of external and middle ear conditions: A machine learning approach." Plos one 15.3 (2020): e0229226.
[10] Jin, Yincheng, et al. "Earhealth: an earphone-based acoustic otoscope for detection of multiple ear diseases in daily life."

REFERENCES

[1] Roy, Pradeep Kumar, et al. "A framework for hate speech detection using deep convolutional neural network." IEEE Access 8 (2020): 204951-204962.

[2] Aluru, Sai Saketh, et al. "Deep learning models for multilingual hate speech detection." arXiv preprint arXiv:2004.06465 (2020).

[3] Mullah, NanlirSallau, and Wan Mohd Nazmee Wan Zainon. "Advances in machine learning algorithms for hate speech detection in social media: a review." IEEE Access 9 (2021): 88364-88376.

[4] Cao, Rui, Roy Ka-Wei Lee, and Tuan-Anh Hoang. "DeepHate: Hate speech detection via multi-faceted text representations." Proceedings of the 12th ACM Conference on Web Science. 2020

[5] Khan, Shakir, et al. "BiCHAT: BiLSTM with deep CNN and hierarchical attention for hate speech detection." Journal of King Saud University-Computer and Information Sciences 34.7 (2022): 4335-4344.

[6] Mozafari, Marzieh, Reza Farahbakhsh, and Noel Crespi. "A BERT-based transfer learning approach for hate speech detection in online social media." Complex Networks and Their Applications VIII: Volume 1 Proceedings of the Eighth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2019 8. Springer International Publishing, 2020.

[7] Rabiul Awal, Md, et al. "AngryBERT: Joint Learning Target and Emotion for Hate Speech Detection." arXiv e-prints (2021): arXiv-2103.

[8] Alkomah, Fatimah, and Xiaogang Ma. "A literature review of textual hate speech detection methods and datasets." Information 13.6 (2022): 273.

[9] Malik, Jitendra Singh, Guansong Pang, and Anton van den Hengel. "Deep learning for hate speech detection: a comparative study." arXiv preprint arXiv:2202.09517 (2022).

[10] Khan, Shakir, et al. "HCovBi-caps: hate speech detection using convolutional and Bi-directional gated recurrent unit with Capsule network." IEEE Access 10 (2022): 7881-7894.

[11] Toraman, Cagri, Furkan Şahinuç, and Eyup Halit Yilmaz. "Large-scale hate speech detection with cross-domain transfer." arXiv preprint arXiv:2203.01111 (2022).

[12] Patil, Hrushikesh, Abhishek Velankar, and Raviraj Joshi. "L3cube-mahahate: A tweet-based marathi hate speech detection dataset and bert models." Proceedings of the Third Workshop on Threat, Aggression and Cyberbullying (TRAC 2022). 2022.

[13] Gravano, Agustín, et al. "Assessing the Impact of Contextual Information in Hate Speech Detection." IEEE Access, vol. 11, pp. 30575-30590, 2023, doi: 10.1109/ACCESS. 2023.3258973. (2023).

[14] Velankar, Abhishek, Hrushikesh Patil, and Raviraj Joshi. "A review of challenges in machine learning based automated hate speech detection." arXiv preprint arXiv:2209.05294 (2022).

[15] Akuma, Stephen, TyosarLubem, and Isaac Terngu Adom. "Comparing Bag of Words and TF-IDF with different models for hate speech detection from live tweets." International Journal of Information Technology 14.7 (2022): 3629-3635.