A Review On Diseases Prediction Through Image Recognition

Mr.Mayur Parthe
Department of Computer Engineering
Rajgad Technical Campus,Pune
Savitribai Phule Pune University
mayurparthe5@gmail.com

Miss.Payal More
Department of Computer Engineering
Rajgad Technical Campus,Pune
Savitribai Phule Pune University
payalmore499@gmail.com

Miss.Harshada Tarange
Department of Computer Engineering
Rajgad Technical Campus,Pune
Savitribai Phule Pune University
harshadatarange25@gmail.com

Mr.Nidhan Mishra
Department of Computer Engineering
Rajgad Technical Campus,Pune
Savitribai Phule Pune University
nidhanbmishra@gmail.com

Prof.B.M.Borhade
Department of Computer Engineering
Rajgad Technical Campus,Pune
Savitribai Phule Pune University

ABSTRACT

This paper examines how artificial intelligence (AI) and deep learning (DL) affect medical diagnostics by automating disease prediction with image recognition. Image recognition tools, such as convolutional neural networks (CNNs), Vision Transformers (VITs), and Generative Adversarial Networks (GANs), find useful patterns in medical images like X-rays, MRIs, CT scans, and retinal fundus images. Recently, models like EfficientNet and ResNet have shown good accuracy in predicting pneumonia. Meanwhile, GhostNet and Bi-DenseNet have been more effective in predicting eye diseases. Additionally, explainable AI (XAI) techniques like Grad-CAM and Saliency Maps offer visual methods to understand how models make decisions. This improves transparency in these algorithms and builds trust in clinical decision-making. While past and current projects highlight successes, challenges such as data inconsistency, limited interpretation, and integration into clinical workflows still exist. Looking ahead, the focus will shift to developing explainable deep learning, multi-model, and federated learning frameworks. The aim is to create reliable systems that support scalability, analysis, and ethical use of future AI algorithms. Overall, AI and DL, through image recognition, mark an important advancement toward precision medicine and better healthcare systems.

Keywords— Disease prediction, image recognition, deep learning, explainable AI, medicine imaging, transfer learning.

1. INTRODUCTION

Deep learning (DL) and artificial intelligence (AI) have transformed medical diagnostics by providing automated, accurate disease prediction and image recognition. Traditional diagnostics mostly rely on a person to read diagnostic images, which is tedious, subjective, and at times, prone to human error. Nevertheless, AI-based image analysis is more precise and scalable, while recognizing complex medical data patterns that are less apparent or even hidden to the human eye [1].

Recent developments in CNNs (convolutional neural networks), Vits (vision transformers), and GANs (generative adversarial networks) have shown incredible promise with analysis of medical imaging modalities, especially X-ray, MRIs, CTs, and retinal fundus images. For instance, Shilpa et al. [1] developed a deep learning framework with AI to detect pneumonia using pre-trained models EfficientNetB0 and ResNet50, and achieved accuracy of 99.78%,

outperforming conventional classification. Alternately, Fu et al. [2] developed a new hybridized Vit-CNN model for pneumonia detection called DAVit (Domain-Adapted Vision Transformer), which finished with 97% F1-score with explainability through Grad-CAM visualizations.

2. METHODOLOGY

The methodology used for this review critically reviews cutting-edge research that involved AI-based disease prediction from images. This is done through a systematic approach which includes literature acquisition, model grouping and framework comparisons that aid in performance evaluation, interpretability, and gaps in research.

A. Data Acquisition and Inclusion Criteria

This review utilized the databases IEEE Xplore, Springer, and ScienceDirect, pulling articles published between 2022 through 2025. The inclusion criteria focused on studies using deep learning architectures (i.e., CNNs, GANs, and Vision Transformers (VITs)) to predict disease images from MRIs, CT, X-ray, and fundus imaging [1], [2]. The review did cover narrow applications such as pneumonia detection, classifications of ocular disease or the segmentation of brain tumour's to ensure relevancy across medical specialties.

B. Model Categorizing and Frameworks.

Once studies were identified, the projects were organized into research and grouped by the primary deep learning model. Convolutional neural network (CNN)-based frameworks, ResNet, DenseNet, and EfficientNet, were reviewed for their accuracy and feature extraction [3], [4]. Transformer-based models (e.g., DAVit) were reviewed for storing global image features and incorporating context [5]. GAN-based frameworks were reviewed for image augmentation and the generation of synthetic images as a strategy to address dataset imbalance [6]. Lastly, a lightweight model and Bi.

C. Assessment of Explainability and Interpretability

To meet the demand for clinical transparency, a review of existing studies was conducted that looked at models that made use of Explainable AI (XAI) methods, such as Grad-CAM, saliency maps, and layer-wise relevance propagation (LRP) [8], [9]. With these visualization models, explainability becomes a more feasible goal as they pinpoint important diagnostic areas, particularly in the context of UNet-based MRI segmentation schemes, to support clinician trust within the case of models used alongside introduction in a clinical setting.

3. LITRETURE SUERVY

Recent developments in deep learning (DL) and artificial intelligence (AI) have advanced the prediction of disease on medical images significantly. Shilpa et al. [1] designed a CNN-based framework using EfficientNetB0 and ResNet50 to detect pneumonia and achieved an accuracy of 99.78% on X-ray images, while Kaushik et al. [2] introduced PneumoAI, a multi-modal model that integrated clinical and imaging data, which achieved an AUC of 0.99. Additionally, Fu et al. [3] showed that their model, called DAVit and based on Vision Transformer, implemented CNN and Transformer layers for pneumonia detection, which would Chen et al. [4] presented a lightweight neural network that used Bi-DenseNet and

GhostNet and had an accuracy of 84.19% for ocular surface disease detection. Also, Sharma et al. [5] utilized Generative Adversarial Networks (GANs) to improve the enhancement of medical images and summarization of reports, which improved data diversity and diagnostic accuracy. In the domain of neuroimaging, Jeya Mala et al. [6] applied UNet with Explainable AI (XAI) methods such as Layer-wise Relevance Propagation (LRP) to create transparent MRI

enhance explainability through visualizing Grad-CAM. In terms of work in ophthalmology,

segmentation. Finally, Ahmad et al. [7] reviewed deep learning models that classified skin diseases and achieved dermatologist-level accuracy while reiterating the requirement that models train using larger and more diverse datasets. Overall, the literature provided demonstrates the potential and efficacy of CNNs, GANs, and Transformers for disease prediction but also indicates limitations in the literature, such as limited.

4. REVIEW OF RELATED WORKS

Innovations in the fields of artificial intelligence (AI) and deep learning (DL) have transformed how illnesses can be predicted using medical images.

4.1 Brain-Related Diseases

Jeya Mala et al. [4] presented an Explainable UNet model that utilized Grad-CAM and LRP to segment and improve the interpretability of MRI brain scans. The major issues still using these image models are the lack of available datasets and the cost of implementing and running these models.

4.2 Lung Cancer Detection

Malarvannan and Angulakshmi (2025) proposed neural network models, such as CNN and ensemble methods, to correctly detect lung cancers based on CT, while there are still issues of the models' generalization.

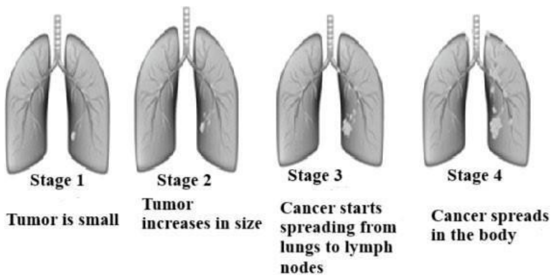

FIGURE 1. Progression stages of lung cancer—illustrating the different developmental phases.

4.3 Retinal and Eye Diseases

Chen et al. [5] developed Bi-DenseNet + GhostNet model, which achieved an accuracy of 84.19%, while Albelaihi and Ibrahim (2024) proposed a model named DeepDiabetic using EfficientNetB0 and Bi-GRU for diabetic eye disease.

4.4 Respiratory and Skin Diseases

Kaushik et al. [6] constructed PneumoAI, a multimodal model of pneumonia, with an AUC=0.99. Li et al. (2020) and Ahmad et al [7], explain CNN with skin models they reviewed that achieved accuracy at or above the levels of a dermatologist, where each of these datasets was limited.

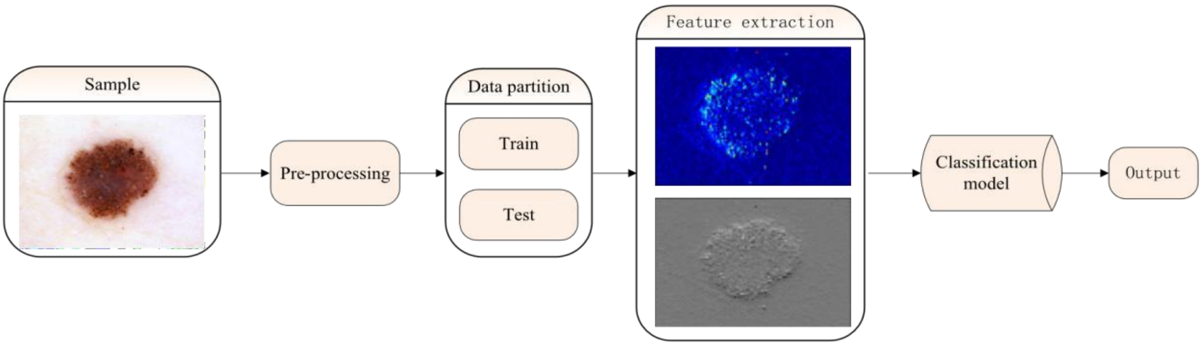

FIGURE 2. Flow chart of skin disease image recognition based on machine learning. Image processing is divided into image acquisition, image preprocessing, and dataset division. Image preprocessing includes image size adjustment, normalization, and noise removal. Image recognition mainly includes image feature extraction and classification models to classify the extracted features and then output the results

4. COMPARATIVE ANALYSIS

Comparisons between state-of-the-art deep learning (DL) and artificial intelligence (AI) methods have demonstrated notable advances in the accuracy of disease prediction among different imaging modalities. Shilpa et al. [1] achieved an outstanding 99.78% accuracy in identifying pneumonia through a deep learning image classifier called EfficientNetB0 and baseline ResNet50. Kaushik et al. [2] further improved prediction study performance to an AUC of 0.99 through the integration of clinical data in addition to image data. Fu et al. [3] presented the DAVit Vision Transformer and achieved a 97% F1-score, which enhanced the interpretability of model predictions through Grad-CAM visualizations. Additionally, Chen et al. [4] presented a lighter CNN model based on GhostNet and Bi-DenseNet in the field of ophthalmology with an accuracy of 84.19% for diagnosing diabetic retinopathy based upon retinal images. However, Jeya Mala et al. [5] introduced UNet approaches to segment brain MRIs, which utilized Explainable AI (XAI) methods to improve trust in clinical settings. Finally, GAN-based image quality improvement techniques by Sharma et al. [6] also improved data diversity but lacked interpretability.

5. RESEARCH GAPS AND FUTURE SCOPE

Although deep learning (DL) and artificial intelligence (AI) have shown strong results in disease prediction using medical images, several barriers are prohibiting clinical adoption. Most studies, including those by Shilpa et al. [1] and Fu et al. [2], achieve high performance and accuracy; however, these studies often take a narrow approach with disease-specific or single-modality datasets and thus face issues with generalizability. There has been little focus in the literature on integrating Explainable AI (XAI) with medically based image classification algorithms; however, Jeya Mala et al. [3] proposed two XAI approaches (GradCAM and Layer-wise Relevance Propagation (LRP)) to develop an XAI-based solution. However, the unbalanced, limited and unannotated datasets demonstrate reduced robustness, and the architectures themselves, such as Vision Transformers and GANs [4], [5], are often heavy and unsuitable for real-time or on-the-edge computation.

Thereby, future work needs to develop multi-modal, explainable hybrid architectures, begin federated learning solutions [6] and curate medical image repositories to improve the adaptability of AI algorithms and trust from the clinical community.

6. CONCLUSION

This paper reviews cutting-edge technology on artificial intelligence (AI) and deep learning (DL) in medical disease prediction via image classification. Models have used deep learning models such as convolutional neural networks (CNN), generative adversarial networks (GAN), and vision transformers (Vit) on brain, lung, eye, and skin diseases that have reported high accuracy in diagnosis [1]– [4]. Nonetheless, [5], [6] clinicians are challenged with the balance of datasets, low generalizability, or poor interpretability in clinical use.The study proposes a hybridized framework of CNN, transformer, and GAN-based augmentation tools in a suitable magnitude with explainable AI (XAI) visualisation to improve

accuracy and explicability in transparency. Future work may necessitate the enhancement of privacy, scale, and opportunity. Multi-modal learning and federated training with a continued agenda for ethical validation could improve clinical assurance of the implementation of medical AI and for AI demands.

REFERENCE

[1] N. Shilpa, W. A. Banu, and P. B. Metre, "*Revolutionizing Pneumonia Diagnosis: AI-Driven Deep Learning Framework for Automated Detection from Chest X-Rays*," *IEEE Access*, vol. 12, pp. 171601–171612, 2024.

[2] D. Sharma, A. K. Gupta, and P. Jain, "*Medical Image Processing, Disease Prediction and Report Summarization Using Generative Adversarial Networks and AIML*," *IEEE Access*, 2024.

[3] M. Fu, C. Tantithamthavorn, and T. Le, "*DAVit: A Domain-Adapted Vision Transformer for Automated Pneumonia Detection and Explanation Using Chest X-Ray Images*," *IEEE Access*, vol. 13, pp. 103033–103050, 2025.

[4] P. Kaushik, A. Sharma, and R. Singh, "*PneumoAI: Redefining Accuracy in Pneumonia Detection Using Advanced Machine Learning*," *IEEE Access*, 2024.

[5] D. Jeya Mala, M. Chattopadhyay, P. Mukhopadhyay, and R. Sinha, "*Visualizing UNet Decisions: An Explainable AI Perspective for Brain MRI Segmentation*," *IEEE Access*, vol. 13, pp. 1–10, 2024.

[6] R. Chen, E. Song, and Z. Zhou, "*A Lightweight Neural Network for Ocular Surface Disease Recognition*," in *Proc. IEEE ICAICE*, pp. 671–677, 2024.

[7] S. Ahmad, N. Li, and F. Wang, "*Deep Learning in Skin Disease Image Recognition: A Review*," *IEEE Access*, 2024.

[8] K. Vayadande, M. Patel, and P. Thakur, "*Ocular Disease Recognition Using Deep Learning*," *IEEE Access*, vol. 10, pp. 150220–150229, 2022.

[9] S. Iqbal, N. Khan, and Z. Ahmed, "*Fusion of Textural and Visual Information for Medical Image Modality Retrieval Using Deep Learning-Based Feature Engineering*," *IEEE Access*, vol. 11, pp. 45387–45398, 2023.

[10] H. Singh, R. Verma, and S. Jain, "*Ocular Disease Detection via Deep Learning Models*," *IEEE Access*, vol. 13, pp. 23422–23435, 2025.