”Skin Disease Detection Using Customized CNN layer and Binary Tree”

Satyajit Suhas Patil
Computer Science and Engg (of AIML.)
KIT College of Engineering
Kolhapur
satyajitsp26@gmail.com

Ayush Sangram Patil
Computer Science and Engg (of AIML.)
KIT College of Engineering
Kolhapur
ayushpatil182k4@gmail.com

Sandesh Mahadeo Lavhate
Computer Science and Engg (of AIML.)
KIT College of Engineering
Kolhapur
sandeshlavhte191@gmail.com

Om Tejas Shah
Computer Science and Engg (of AIML.)
KIT College of Engineering
Kolhapur
omshah12@gmail.com

Shrivardhan Suryaji Patil
Computer Science and Engg (of AIML.)
KIT College of Engineering
Kolhapur
patilparth390@gmail.com

Dr.Uma Gurav
Computer Science and Engg (of AIML.)
KIT College of Engineering
Kolhapur
uma.gurav@kitcoek.in

Dr.Akshay Chinchanikar
Computer Science and Engg (of AIML.)
KIT College of Engineering
Kolhapur
akshaychinchnikar@kitcoek.in

***Abstract*—Skin diseases represent one of the most prevalent health issues globally, affecting millions across different age groups and geographic regions. The burden of dermatological conditions is particularly severe in low-resource settings, where access to trained dermatologists is limited. Traditional diagnostic methods are often time-consuming, costly, and require specialized medical expertise, which may not always be readily available. In this research, we propose an automated, deep learning-based system for the classification of skin diseases using a VGG16 model with transfer learning.**

The VGG16 architecture, pre-trained on the ImageNet dataset, has been fine-tuned on a domain-specific dermatological image dataset to detect and classify common skin conditions such as eczema, psoriasis, acne, and various types of skin cancer. This transfer learning approach significantly reduces the need for large, labeled datasets while leveraging the powerful feature extraction capabilities of deep convolutional neural networks. By freezing the initial layers and retraining the final classification layers, the model effectively learns to identify subtle patterns and textures in skin lesion images.

Our dataset comprises high-quality clinical images annotated by dermatological experts. To enhance model generalization and mitigate overfitting, various data augmentation techniques such as rotation, flipping, and zooming were applied. The model is trained and validated using a stratified train-test split to ensure balanced representation of all disease classes. Evaluation metrics such as accuracy, precision, recall, and F1-score were employed to measure the performance of the model, with the system achieving high classification accuracy across multiple disease categories.

This approach demonstrates the viability of using deep learning for real-time, low-cost, and efficient skin disease diagnosis, making it especially valuable in telemedicine applications and rural healthcare settings. Furthermore, our system can be integrated into mobile or web platforms, enabling non-expert users to receive preliminary diagnostic feedback, thus promoting early detection and timely medical intervention.

In conclusion, this work highlights the transformative potential of artificial intelligence in medical diagnostics. By augmenting the capabilities of healthcare professionals and reducing diagnostic errors, AI-powered tools like the proposed VGG16-based model can play a crucial role in improving public health outcomes and making dermatological care more accessible and equitable.

I. Introduction

Skin diseases are a global health concern, affecting millions of people annually across all age groups and demographics. These conditions encompass a wide spectrum, ranging from mild and non-contagious issues like eczema and acne to severe and life-threatening illnesses such as melanoma and other forms of skin cancer. Accurate and early diagnosis of these conditions is critical to preventing disease progression, minimizing healthcare costs, and improving patient outcomes. However, the diagnostic process is often hampered by several challenges, including the shortage of dermatologists, particularly in rural and underdeveloped regions, as well as the high variability in the appearance of skin lesions across different skin tones, lighting conditions, and anatomical locations.

In recent years, rapid advancements in artificial intelligence (AI), particularly in computer vision and machine learning, have paved the way for automated diagnostic systems capable of assisting clinicians in medical image analysis. Deep learning models—especially Convolutional Neural Networks (CNNs)—have achieved remarkable success in image classification tasks due to their ability to automatically learn hierarchical feature representations from raw input images. Among these, the VGG16 architecture has emerged as a widely adopted and effective CNN model, known for its depth and simplicity, making it well-suited for transfer learning applications in the medical domain.

This project explores the implementation of VGG16 for the classification of skin diseases through a transfer learning approach. By leveraging a model pre-trained on the extensive ImageNet dataset, we capitalize on the general visual features it has already learned, such as edge detection and texture recognition. These foundational features are then adapted to the specific domain of dermatology by fine-tuning the upper layers of the network with a labeled dermatological image dataset. This method not only reduces the need for large volumes of medical training data—often difficult to obtain due to privacy and annotation constraints—but also significantly accelerates model convergence and enhances performance.

The proposed system undergoes rigorous training and validation on a curated set of skin disease images that represent various categories such as dermatitis, fungal infections, vitiligo, and malignant lesions. Preprocessing steps, including normalization and data augmentation (e.g., rotation, flipping, brightness adjustment), are applied to improve the model's robustness and generalization capabilities. Performance is evaluated using key metrics such as accuracy, sensitivity, specificity, precision, and F1-score, ensuring a comprehensive assessment of the model's diagnostic reliability.

By integrating this deep learning-based approach, we aim to build a practical, scalable, and cost-effective diagnostic tool that can be deployed in mobile health applications, teledermatology platforms, and primary care settings. This not only aids dermatologists in making quicker and more informed decisions but also empowers non-specialist healthcare providers and patients in remote areas with early access to reliable diagnostic support. Ultimately, this work demonstrates how AI can bridge critical gaps in healthcare delivery and contribute to more equitable and accessible dermatological care.

II. Problem Statement

Traditional diagnosis of skin diseases relies on clinical examination and biopsy, which can be expensive, slow, and heavily dependent on expert interpretation. There is a need for an automated system that can provide accurate preliminary diagnoses to assist medical professionals and reach underserved populations. The objective of this research is to develop an AI-driven model using VGG16 and transfer learning techniques to classify images of skin lesions into various disease categories with high accuracy and minimal computational cost.

III. Related Work

Numerous studies in recent years have successfully applied deep learning techniques to the classification of dermatological images, highlighting the transformative potential of artificial intelligence in medical diagnostics. A landmark study by Esteva et al. (2017) demonstrated that convolutional neural networks (CNNs) could achieve dermatologist-level accuracy in detecting skin cancer from clinical images. This study marked a pivotal moment in medical AI research, setting the stage for a wide range of applications across diagnostic imaging domains. Similarly, Rajpurkar et al. introduced CheXNet, a 121-layer DenseNet model trained to detect pneumonia from chest X-rays. This work not only underscored the efficacy of deep CNN architectures in clinical settings but also showcased the powerful utility of transfer learning in enabling high performance on limited medical datasets.

Building on these foundations, recent research efforts have explored advanced architectures such as ResNet, Inception, and DenseNet for skin disease classification. These models offer deep and complex network structures capable of learning rich and abstract representations. For instance, ResNet's residual connections mitigate the vanishing gradient problem, allowing for the training of extremely deep networks. Inception modules enable multi-scale feature extraction within the same layer, and DenseNet facilitates feature reuse across layers, improving parameter efficiency. Despite their strong performance, these architectures are computationally intensive and typically require significant computational resources and large annotated datasets for effective training—requirements that are often impractical in many real-world clinical environments, especially in low-resource settings.

In contrast, transfer learning using the VGG16 architecture provides a practical compromise between performance and computational feasibility. VGG16 is relatively shallow compared to more modern architectures, yet it remains highly effective due to its consistent use of small (3x3) convolutional filters and deep stacked layers. When pre-trained on large-scale datasets like ImageNet, VGG16 captures a rich hierarchy of visual features that can be repurposed for specific tasks, such as skin disease classification. By freezing the lower layers and fine-tuning only the deeper layers or classifier components, the model can be adapted to a new domain with minimal data and reduced training time, while still achieving high accuracy.

This balance between simplicity and efficacy makes VGG16 particularly advantageous in domains where labeled medical data is scarce, and where model interpretability and deployment efficiency are key concerns. Moreover, its architecture is well-supported in most deep learning frameworks, and it is less prone to overfitting when trained on small-to-moderate datasets due to its manageable number of parameters. Therefore, VGG16-based transfer learning emerges as a robust and scalable solution for skin disease classification, providing an effective path forward in the development of AI-driven diagnostic tools.

IV. technical Aproach

Data collection, data preprocessing, machine learning models, clinical knowledge integration, and user interface design are some of the technical elements that go into creating a disease and medication recommendation system. By considering patient-specific information and the most recent medical research, these systems seek to offer individualized and evidence-based therapy suggestions. An outline of the technological work required to create such a system may be found below. 1. Information Gathering and Combination

Building an illness and medication recommendation system begins with gathering and combining different types of data. This comprises: Patient Information: Important patient data, including medical history, age, gender, lifestyle choices, and test results, are stored in electronic health records, or EHRs. EHR data must be integrated into the system, guaranteeing its completeness and currentness. Clinical Data: This include prescription databases, guidelines, and treatment plans. Medical terminology and medication information are standardized using medical ontologies such as RxNorm and SNOMED CT (Systematized Nomenclature of Medicine). Pharmacological Data: It is necessary to record and retrieve information about medications, including their dosages, side effects, interactions, and indications and contraindications. Drug reference databases like the FDA database may provide this. The technical pipeline of this study includes the following components:

Dataset Collection: A publicly available dataset (e.g., HAM10000 or ISIC 2018) was used, containing high-resolution images of various skin diseases. Data Preprocessing: Images were resized to 224x224 pixels, normalized, and augmented (rotation, flipping, zoom) to increase diversity and robustness. Model Architecture: The VGG16 model was imported with pre-trained weights from ImageNet. The top classification layers were removed and replaced with new dense layers tailored to the number of skin disease categories. Training Strategy: Only the final few layers of VGG16 were unfrozen and fine-tuned, while the rest were kept frozen to retain generalized features. Loss Function Optimizer: Categorical cross-entropy loss with the Adam optimizer was used for model training. 5. Feedback Loop and User Interface (UI) For healthcare practitioners to effectively engage with the system, an intuitive user interface is necessary. The user interface must be made to provide concise, useful suggestions with little cognitive strain. Important characteristics include: Dashboard: Current treatments, lab results, and pertinent patient data should all be shown via the system. Suggestion Output: The suggested drug, together with any potential dosages and adverse effects, and the disease diagnosis should be prominently highlighted. Mechanism of Feedback: In order to improve the model, healthcare providers can submit input on the recommendations' usefulness and accuracy. By using this data to enhance the machine learning models, the system would be able to continuously adjust to new developments in patient outcomes and healthcare trends. 6. System Evaluation and Validation Before deployment, the system must undergo extensive validation to ensure its clinical efficacy and safety. This involves: Cross-validation: Machine learning models are validated using techniques like k-fold cross-validation to assess their generalizability to new data. Clinical Trials: In many cases, disease and medicine recommendation systems need to be tested in clinical environments to ensure that they provide valid and effective recommendations. Performance Metrics: Metrics such as accuracy, precision, recall, F1 score, and Area Under the ROC Curve (AUC) are used to evaluate the system's performance in diagnosing diseases and recommending treatments.

V. Knowlwdge Representation

In this project, knowledge is represented in multiple forms, each contributing to the system's ability to understand, classify, and evaluate skin disease images effectively:

Image Features: The primary form of knowledge representation stems from the hierarchical feature extraction process performed by the convolutional layers of the VGG16 model. These layers progressively learn and encode essential visual characteristics from the input images, such as edges, textures, shapes, and color gradients. Lower layers typically capture basic features like lines and contours, while deeper layers extract more abstract patterns relevant to specific skin lesions. These learned features form a multidimensional representation of each image, which serves as the foundation for accurate classification.

Class Labels: Each image in the dataset is annotated with a corresponding skin disease category, such as eczema, psoriasis, melanoma, or tinea. These labels act as the ground truth during supervised training, guiding the model to learn discriminative features that distinguish between different skin conditions. The class labels are an explicit form of knowledge that provides semantic meaning to the visual patterns learned by the model.

Model Weights: The knowledge acquired during training is encoded within the weights and biases of the neural network. In the context of transfer learning, the pre-trained VGG16 weights already contain general-purpose visual knowledge gained from the ImageNet dataset. During fine-tuning, the model adjusts these weights to specialize in the domain of dermatology. These learned parameters encapsulate the mappings between complex image features and the corresponding disease classes, enabling the model to make accurate predictions on unseen data.

Performance Metrics: Quantitative evaluation metrics—such as accuracy, precision, recall, F1-score, and the confusion matrix—represent the effectiveness and reliability of the system's learned knowledge.

Accuracy measures the overall correctness of predictions.

Precision indicates the proportion of true positive predictions among all positive predictions.

Recall reflects the model's ability to identify all relevant cases of a particular class.

F1-score balances precision and recall to provide a single performance measure.

Confusion Matrix provides a detailed breakdown of classification outcomes across all classes, helping identify patterns of misclassification and areas for improvement. These metrics not only validate the performance of the trained model but also offer insights into the quality of its internal representations and decision-making capabilities.

Together, these layers of knowledge representation—from visual features to model parameters and evaluation outcomes—create a comprehensive framework that enables the system to perform accurate and intelligent skin disease classification. This multi-level knowledge structure is essential for developing a reliable and interpretable AI-based diagnostic tool suitable for clinical and real-world deployment.

VI. Implementation and Results

The system was developed using Python, leveraging the TensorFlow and Keras deep learning frameworks, which provide high-level APIs and GPU acceleration capabilities for efficient model development and training. To optimize training performance and reduce computational time, all experiments were conducted on a GPU-enabled system, ensuring faster convergence and better handling of the computationally intensive operations involved in training convolutional neural networks.

The key implementation steps are summarized below:

Dataset Loading and Preprocessing: The dermatological image dataset, which includes labeled examples of various skin conditions, was loaded and preprocessed before being fed into the model. Preprocessing involved:

Resizing images to 224×224 pixels to match the input shape required by VGG16.

Normalizing pixel values to a [0, 1] range.

Data augmentation techniques such as random rotations, flips, zooms, and shifts were applied to enhance generalization and mitigate overfitting due to limited dataset size.

Model Construction with Transfer Learning: The VGG16 model pre-trained on the ImageNet dataset was used as the base model. The convolutional layers of VGG16 were frozen to preserve the general visual features already learned, and custom dense layers were stacked on top to tailor the network for the specific task of skin disease classification. The modified architecture included:

Global Average Pooling to reduce dimensionality.

Fully connected (Dense) layers with ReLU activation functions.

Dropout layers to randomly deactivate neurons during training, reducing the risk of overfitting.

Batch Normalization layers to stabilize learning by normalizing layer inputs.

Softmax output layer to predict probabilities across multiple skin disease classes.

Model Training and Validation: The dataset was split into an 80/20 ratio for training and testing, respectively. The model was compiled using the Adam optimizer with a suitable learning rate, and categorical cross-entropy loss was used due to the multi-class nature of the classification task. Early stopping and model checkpointing were employed to avoid overfitting and retain the best-performing model during training.

Results and Evaluation The model achieved strong performance metrics on the test dataset, confirming the effectiveness of VGG16 with transfer learning for skin disease classification:

Accuracy: The model achieved approximately 89perc accuracy on the test data, indicating a high overall correctness in classification decisions.

Precision and Recall: Averaged between 87perc and 91perc across all classes, reflecting the model's robustness in distinguishing true positives and minimizing false positives and negatives. These metrics are particularly important in medical applications where both false negatives (missed diagnoses) and false positives (unnecessary anxiety or treatment) carry serious implications.

F1-Score: The harmonic mean of precision and recall also demonstrated strong values, indicating a balanced performance across different skin disease categories.

Confusion Matrix: The confusion matrix revealed high true positive rates across most classes, with some minor confusion between visually similar conditions. This analysis helped identify areas for potential improvement, such as using ensemble models or increasing training data diversity.

VII. Future Scope

The suggested condition and medication recommendation system has a number of exciting directions for further study and advancement that might greatly improve its functionality and increase its range of applications. Using cutting-edge deep learning methods to enhance feature representation and prediction accuracy is one crucial avenue. Convolutional neural networks (CNNs) for image data and recurrent neural networks (RNNs) or long short-term memory (LSTM) networks for time-series data are two examples of deep learning models that are better able to capture complex relationships in patient data, including vital signs over time, clinical notes, and medical imaging. The system can learn more complex patterns and enhance its predictive abilities by utilizing deep learning, particularly in situations when conventional machine learning models might not be able to perform well. The recommendation system can develop into a more complete tool for handling a greater range of healthcare requirements by broadening its coverage to include these intricate and chronic illnesses. Overall, these developments would enhance the system's predictive precision and flexibility while also broadening its application to a wider range of patient care, which would ultimately result in improved healthcare outcomes. Future enhancements to this project include:

Dataset Expansion: Incorporate more diverse and balanced datasets covering rare skin diseases. Multilingual Mobile App Integration: Build a real-time diagnostic app for global access. Explainability: Use Grad-CAM or SHAP to visualize model decision regions. Hybrid Models: Combine CNNs with NLP for patient history-based predictions. Teledermatology Integration: Connect the model with remote diagnosis platforms to assist healthcare professionals.

VIII. Conclusion

This research presents an efficient and accurate method for automated skin disease detection using the VGG16 model with transfer learning. The proposed system leverages the strengths of deep learning to achieve dermatologist-comparable performance, making it highly valuable in clinical and remote settings. The use of a pre-trained VGG16 model accelerates development, reduces computational cost, and maintains high accuracy even with limited labeled data. This system has the potential to revolutionize dermatological diagnosis by offering scalable, accessible, and real-time support to healthcare providers worldwide.

REFERENCES

[1] Misdiagnosis in America: Shocking Statis- tics, Mercola.com. [Online]. Available: https://articles.mercola.com/sites/articles/archive/2015/09/30/diagnostic-errors.aspx. [Accessed: 03-Dec-2017].

[2] P. Bedi, H. Kaur, and S. Marwaha, Trust Based Recommender System for the Semantic Web, in Proceedings of the 20th International Joint Conference on Artifical Intelligence, San Francisco, CA, USA, 2007, pp. 26772682.

[3] What is a Knowledge-Based System (KBS)? - Definition from Techopedia, Techopedia.com. [Online]. Available: https://www.techopedia.com/definition/7969/knowledge-based-system- kbs. [Accessed: 03-Dec-2017].

[4] E. Coiera, Guide to Health Informatics, 2Ed. CRC Press, 2003.

[5] G. M. Kahn, M. L. Fagan, and H. E. Shortliffe, Time in Clinical Decision Support Systems: Temporal Reasoning in ONCOCIN and ONYX, SIGBIO Newsl, vol. 8, no. 1, pp. 1316, Mar. 1986.

[6] M. K. Goldstein et al., Implementing clinical practice guidelines while taking account of changing evidence: ATHENA DSS, an easily mod- ifiable decision-support system for managing hypertension in primary care, Proc. AMIA Symp., pp. 300304, 2000.

[7] M. Popescu and M. Khalilia, Improving disease prediction using ICD-9 ontological features, 2011, pp. 18051809.

[8] Recommender System for Diagnosis and Treatment of Tinnitus - PDF. [Online]. Available: http://docplayer.net/59629697-Recommender- system-for-diagnosis-and-treatment-of-tinnitus.html. [Accessed: 03-Dec-2017].

[9] User Interface Design Basics, 21-May-2014. [Online]. Available: /what- and-why/user-interface-design.html. [Accessed: 07-Dec-2017].

[10] Grampurohit, Sneha, and Chetan Sagarnal. ”Disease Prediction using Machine Learning Algorithms.” 2020 International Conference for Emerging Technology (INCET). IEEE, 2020.

[11] Verma, Anurag Kumar, Saurabh Pal, and Surjeet Kumar. ”Prediction of skin disease using ensemble data mining techniques and feature selection method—a comparative study.” Applied biochemistry and biotechnology 190.2 (2020): 341-359.

[12] Ramalingam, V. V., Ayantan Dandapath, and M. Karthik Raja. ”Heart disease prediction using machine learning techniques: a survey.” International Journal of Engineering Technology 7, no. 2.8 (2018): 684-687.

[13] Uddin, Shahadat, et al. ”Comparing different s upervised machine learning algorithms for disease prediction.” BMC medical informatics and decision making 19.1 (2019): 1-16.

[14] Ramalingam, V. V., Dandapath, A., Raja, M. K. (2018). Heart disease prediction using machine learning techniques: a survey. International Journal of Engineering Technology, 7(2.8), 684-687.

[15] Chen, Min, et al. ”Disease prediction by machine learning over big data from healthcare communities.” Ieee Access 5 (2017): 8869-8879.

[16] Mrs. Shital. P. Chattar, Prathamesh Wadile , Bhushan Saraf , Omkar Kardile and Aditya Pawar. ” IOT Based Heart Disease Prediction.” International Journal for Research in Applied Scien ce Engineering Technology (IJRASET) (2021):2321-9653.

[17] Megha Vyas , Priya Kunwar, Vinod Joshi, Mihir Maheshwari and Mr. Girish Ameta. “Software for Disease Predic tion.” International Journal for Research in Applied Science Engineering Technology(IJRASET) :2321-9653,2020.

[18] Yuvraj Nikhate and M. V. Jonnalage dda. “Survey On Heart Disease Prediction using Machine Learning.” International Journal of Creative Research Thoughts(2020) :2320-2882.

[19] Raj H . Chau han, Daksh N. Naik, Rinal A. Halpati, Sagarkumar J. Patel4 and Mr. A.D.Prajapati. “ Diseas e Prediction Using Machine Learning.”International Research Journal Of Engineering Technology(IRJET):23950056,2020.

[20] Sandhya B , Monika K. J , Vyshnavi and Usman K. “ Disease Prediction and Hospital Recommendation using Machine Learning Algorithms.” International Journal for Research in Applied Science Engineering Technology(IJRASET): 2321-9653,2020.

[21] Lamido Yahaya , Nathaniel David Oye and Etemi Joshua Garba. “A Comp rehensive Review on Heart Disea se Prediction Using Data Mining and Mach ine Learning Techniques. ”American Journal of Artificial Intelligence:2639-9717 ,vol 4, 2020.

[22] Neelaveni, J., and MS Geetha Devasana. ”Alzheimer disease prediction using machine learning algorithms.” 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, 2020.