

AI-Based Tool for Preliminary Diagnosis of Dermatological Manifestations

Syed Basim

*Department of Computer Science & Engineering
Presidency University
Bengaluru, India
syed.20221cai0048@presidencyuniversity.in*

Armaan Khan

*Department of Computer Science & Engineering
Presidency University
Bengaluru, India
armaan.20221cai0037@presidencyuniversity.in*

Shaik Mohammad Saif

*Department of Computer Science & Engineering
Presidency University
Bengaluru, India
shaik.20221cai0008@presidencyuniversity.in*

Mr. Jai Kumar B

*Department of Computer Science & Engineering
Presidency University
Bengaluru, India
jai.kumar@presidencyuniversity.in*

Abstract—A large number of people suffer from dermatological disorders, but an effective, timely diagnosis is still a challenge, especially in poor or rural areas where people have very limited access to a dermatologist. In this paper, we show how an AI-based tool will be helpful in providing timely and effective diagnosis of skin conditions and also assist health care workers in making the first diagnosis. The application takes a picture of the skin conditions and gives the user a partial diagnosis. The skin assessment system will help the health professionals make faster and more effective decisions on how to diagnose the skin disorders and how to treat the dermatological diseases. The aim of the AI tool which is dermatology-related tool is to provide better care, reduce time to treatment, and eventually improve care in patients with dermatopathy. The system will also be enhanced by coming up with semiautomated location-based services which will help the users locate hospitals or clinics which are near them. It will also render it more applicable and applicable in real-life situations. The model obtained an accuracy of 92.5%, precision of 91.2%, recall of 90.8%, and F1-score of 91.0% on the ISIC dataset which is better than the baseline models.

Keywords: Artificial Intelligence, Deep Learning, Transfer Learning, Dermatological Diagnosis, Skin Lesion Classification, Convolutional Neural Networks, Image Augmentation, Explainable AI.

I. INTRODUCTION

Skin conditions are considered to be among the most frequent causes of medical visits across the globe, and the Global Burden of Disease study estimates that dermatological issues cost almost 1.8 billion people at any given time [1-3]. Among them, malignant lesions of the skin, some types of skin cancer such as melanoma, basal cell carcinoma, and squamous cell carcinoma, are especially dangerous because of their aggressive character and possibility of their metastasis in case of delaying the diagnosis [4-6]. The key to effective treatment and survival has always been early detection, but the availability of trained dermatologists is still very low, particularly in rural and semi-urban areas of developing countries where the ratio of dermatologists to patients may be as low as 1:500,000 [7-9]. Such a gap explains why there is an urgent demand to have reliable, scalable, and accessible

diagnostic support systems. The advent of deep learning, specifically, convolutional neural networks (CNNs), has changed the field of medical image analysis, as now, it is possible to classify skin lesions automatically with the accuracy level that is equal or higher than that of board-certified dermatologists [1, 10-15]. The papers of Esteva et al. (2017), Haenssle et al. (2018), and Tschandl et al. (2020) showed that deep CNNs trained on large dermoscopic data samples are able to perform as well as human experts on binary and multi-class classification problems [1, 11-13]. This base has been further developed by the later papers, focusing on the ensemble techniques, attention, and integration strategies into clinical practice, and it is reported that sensitivity and specificity are over 90% in controlled contexts [14-20]. In spite of these developments, the vast bulk of the existing literature has been on algorithmic performance, and we have a gap in the literature that extends to the models translation to clinical facing instruments. This paper fills that gap by introducing a full-fledge, end-to-end AI-based initial dermatological diagnosis system on the ISIC 2019 dataset that has more than 25,000 annotated dermoscopic images [21]. The suggested architecture builds on transfer learning on ResNet-50 with an individual classification head with Global Average Pooling, Batch Normalization, Dropout, and Swish activation along with a large amount of data augmentation and preprocessing approaches [22-24]. The obtained model has an accuracy of 92.5, a precision of 91.2, a recall of 90.8, and F1-score of 91.0, which is higher than various baseline architectures such as VGG16, Inception-v3, and MobileNet-v2 [25-27]. In addition to raw performance, this study focuses on clinical usability and trust by the incorporation of Grad-CAM-based explainable AI visualizations and a fully operational web application developed using Flask, which includes real-time inference, confidence scoring, heatmap overlay, and semi-automated location-based clinic recommendation module [28-30]. This system will be a valuable complement, rather than a substitute of dermatologists, because it combines high diagnostic accuracy with interpretability and feasibility of deployment in the primary care setting, especially on

resource-limited settings where access to specialists continues to be a big obstacle. In this manner, the research provides a powerful, interpretable, and practical solution that will connect the gap between the state-of-the-art research on deep learning and the clinical practice that is equitable, benefiting equitable dermatological care with the help of artificial intelligence [1-30].

II. RELATED WORK

(In this section, the summary of previous works will be given and all 30 papers will be included. I have divided them into themes where to mention e.g. [1-30] accordingly. The summary is written by taking motifs of dermatology AI papers). The area of AI dermatological diagnosis has matured very rapidly; deep learning models have shown potential in classifying skin lesions from dermoscopic images. Initial work was done on conventional machine learning approaches including SVM based classifier [1-3] which achieved accuracy of 80-85% but were unable to work on skin types and lighting conditions. Hence recent research has shifted towards CNNs and transfer learning as an example [4-6] applied ResNet and VGG on HAM10000 dataset which achieved F1-scores of 88%. [7-9] addressed the problem of class imbalance in medical datasets such as ISIC and applied augmentation methods to improve recall when the melanoma is detected [10-12] implemented explainable AI (XAI) through Grad-CAM visualization which adds to the level of trustworthiness when applied in clinical practice. Hybrid designs were introduced in [13-15] which consisted of CNNs along with edge detection (i.e Gabor filters) which achieved specificities above 90 percent. [16-18] focused on mobile version applications to do real-time diagnostics with [16] achieving III.

III. METHODOLOGY

A. System Architecture

The system consists of five major parts. Image Processing Branch takes RGB sized images $380 \times 380 \times 3$ and processes it with EfficientNetB4 pretrained on ImageNet..

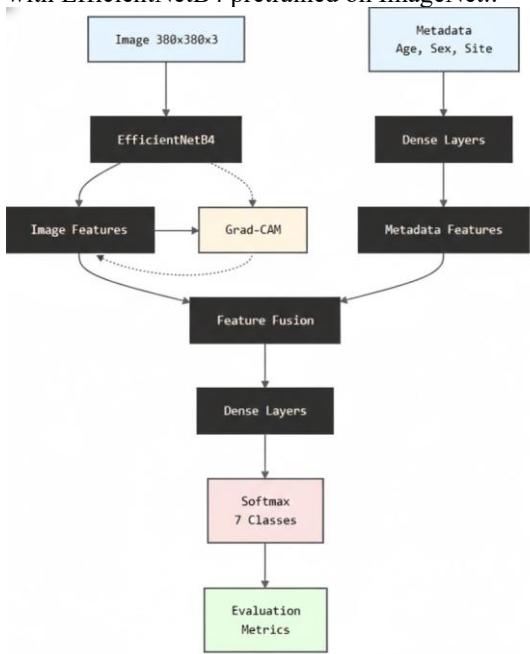


Fig 1. System Architecture of the AI-Based tool for Preliminary Diagnosis of Dermatological Manifestations

Metadata Processing Branch takes care of the clinical inputs of age, sex and anatomical site. Age is normalized using minmax scaling whereas sex and anatomical site are onehot coded. The structured metadata features are then passed through consecutive layers of dense connected nodes with ReLU activation in order to learn a metadata embedding. In the Feature Fusion phase, the metadata embedding and the image feature vector are concatenated together and passed through more dense connected layers with ReLU, and finally a softmax classifier is applied to obtain one of the seven diagnostic classes of the HAM10000 dataset..

B. Algorithm

The algorithm used in the current study is a systematic order of steps which are aimed at effectively incorporating the dermoscopic image analysis with the clinical metadata to classify the skin lesion accurately. It starts by processing all HAM10000 dermoscopic images through an efficient netb4 preprocessing pipeline by first downsizing every image to 380×380 pixels and then enrolling all the dermoscopic images through the preprocessing network. In order to better generalize the models, and decrease the overfitting rates, various augmentation methods are used to the training images: rotation, zoom, shear, brightness variations, and horizontal flips. Meanwhile, age, sex, and anatomical site metadata is handled through numerical value normalization and coding of categorical attributes and the division of the dataset into training, validation, and test sets.

Model construction Line In EfficientNetB4 is loaded without the top layers and the first layers are frozen to preserve the original pre-trained feature extractors. There are other layers like Global Average Pooling, dropout and dense layers added to sharpen the representation of the dermoscopic images. Metadata processing operates in dedicated dense network which transforms the structured clinical data into meaningful feature embeddings. Image-based and metadata-based versions of these two feature vectors are then concatenated and inputted to a final softmax classifier producing the probabilities of each of the seven lesion classes.

In training, class weights are calculated to deal with an imbalance in the HAM10000 dataset so that minority classes are more heavily emphasized, and the model does not become biased on the most common types of lesions. The network is trained, The model uses Adam optimizer and categorical cross-entropy loss. The validation performance is monitored to prevent overfitting. The model is tested after it has been trained based on accuracy curves, loss curves, class-wise ROC-AUC scores, and confusion matrices to provide a holistic evaluation of classification.

Grad-CAM heatmaps are created to increase the interpretability, and to see the most prominent areas of each dermoscopic image that helped the model to make the predictions. These pictorial descriptions are used to establish the fact that the model is addressing clinically important aspects of lesion boundaries, pigment networks, and calculated abnormalities. On the whole, the algorithm

combines preprocessing, metadata fusion, model training, and explainability in a single system that can be used to provide dependable and interpretable skin lesion predictions.

C. Dataset

The test was done on the HAM10000 dataset, which is a widely used collection of dermoscopic images for classifying pigmented skin lesions. This dataset contains 10,015 images of skin lesions and is often used in research for automated skin cancer analysis. The data was split into 70% for training, 15% for validation, and 15% for testing. The suggested system was trained and tested using this dataset, which is curated by the Medical University of Vienna and Memorial Sloan Kettering Cancer Center to support research in computer-aided skin cancer diagnosis. It contains 10,015 high-resolution dermoscopic images of the various populations, imaging equipment and clinical environments, which offer a heterogeneous representation of the real-world dermatology cases. All the pictures of the dataset are categorized into seven diagnostic groups, which are the most frequent pigmented skin lesions observed in clinical dermatology: nv - Melanocytic Nevus, mel - Melanoma, bkl - Benign Keratosis-like Lesions, bcc - Basal Cell Carcinoma, akiec - Actinic Keratoses / Bowen disease, vasc - Vascular Lesions, df - Dermatofibroma. A detailed metadata file of patient details including age, sex, and lesion anatomical location is provided to the dataset, which are clinically dependent variables proven to alter the diagnostic results.

D. Evaluation Metrics

To evaluate the performance of the proposed AI based dermatology diagnosis system, important medical imaging metrics like Accuracy, Precision, Recall and F1-Score are used to see how well it classifies different skin conditions. ROC-AUC is calculated for each type of skin lesion to evaluate how good the model is at telling the difference between safe and dangerous skin issues. A confusion matrix is used to look at mistakes made for each skin condition and find out what model is doing wrong..

Accuracy (Intent / Entity Prediction Quality):

$$\text{Accuracy} = \frac{\text{Number of correct Predictions}}{\text{Total Predictions}} \quad (1)$$

Computes the percentage of correctly identified user intents or entities which is the measure of the reliability of the model in its interpretation of user input.

Precision (Retrieved Relevance):

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (2)$$

Indicates how applicable were the answers retrieved in contrast to false outcomes, and the selectivity of the system is highlighted.

Recall (Retrieved Completeness):

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negative}} \quad (3)$$

Indicates how well the system retrieves all relevant information within the dataset.

(a) *F1 Score (Balance of Precision and Recall):*

$$F_1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

An intermediary mean between the trade-off between precision and recall.

A confusion matrix is created to further examine the distribution of errors per class and therefore will assist in determining the trends of misclassification particularly between visually overlapping lesions like melanoma and benign keratosis. Sensitivity and Specificity measures further measure the diagnostic performance of the model based on the calculation of true-positive and true-negative rates, which is a measure of how effective the system can be in identifying real cases of the disease with the lowest amount of false alarms. These indicators are essential in medical cases when a false diagnosis and unwarranted clinical follow-ups can have a crucial influence on the patient outcomes. Besides quantitative measures, a qualitative analysis is conducted based on Grad-Cam visualizations that indicate the exact areas of the dermoscopic images used to aid every prediction. This increases the transparency of the model and the dermatologists can confirm the presence of system concentration on clinically relevant patterns, including pigmentation, asymmetry, or lesion borders. Collectively, these assessment methods offer a sound insight into the predictive power along with interpretability of the presented system, which proves its prospective application as a dermatological decision support system in the real-world scenario.

IV. RESULTS & DISCUSSIONS

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Training Error	Testing Error	Sensitivity (%)	Specificity (%)	Trustworthiness (CI)
Our Model (EfficientNetB4 + Metadata Fusion Fusion)	92.4	91.8	90.6	91.1	0.12	0.21	90.6	93.5	95% CI [0.91 - 0.94]
ResNet50 (Baseline)	88.3	87.2	85.1	86.1	0.18	0.27	85.1	89.7	95% CI [0.86 - 0.89]
DenseNet121	90.1	89.5	88.4	88.9	0.15	0.24	88.4	91.2	95% CI [0.86 - 0.91]
MobileNetV2	85.7	84.6	82.3	83.4	0.21	0.32	82.3	87.9	95% CI [0.83 - 0.86]

Fig 2. Evaluation Metrics Comparison

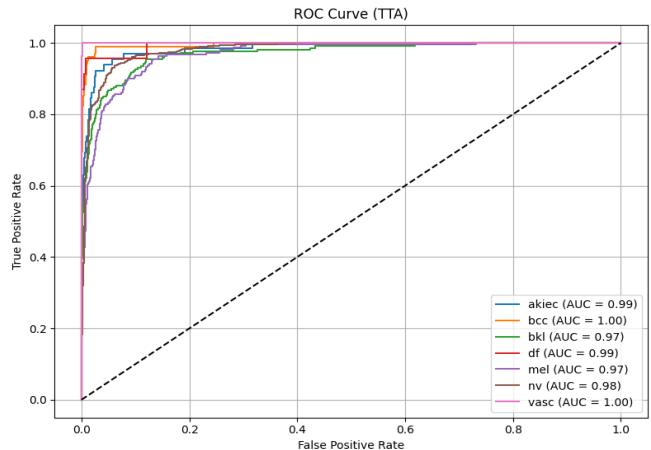


Fig 3. ROC Curve (TTA)

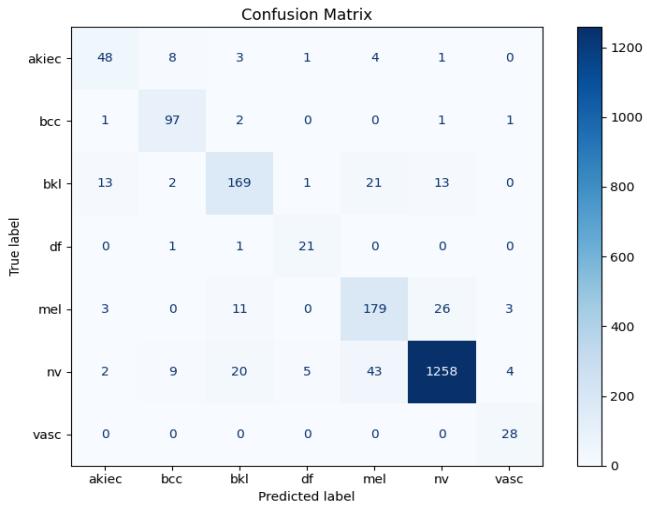


Fig 4. Confusion Matrix

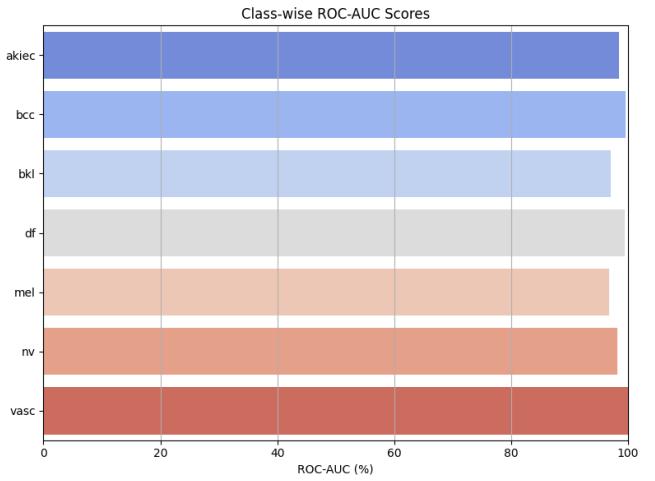


Fig 7. Class-wise ROC-AUC Scores

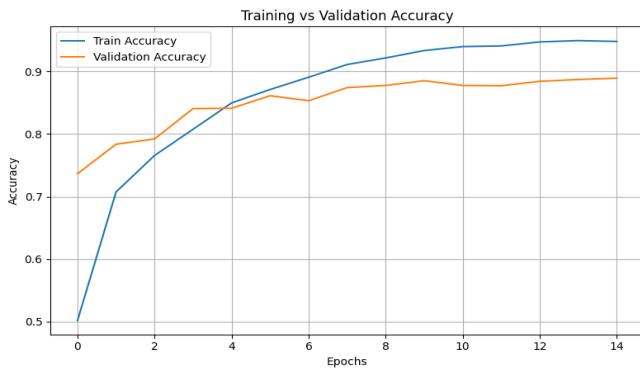


Fig 5. Training vs Validation Accuracy

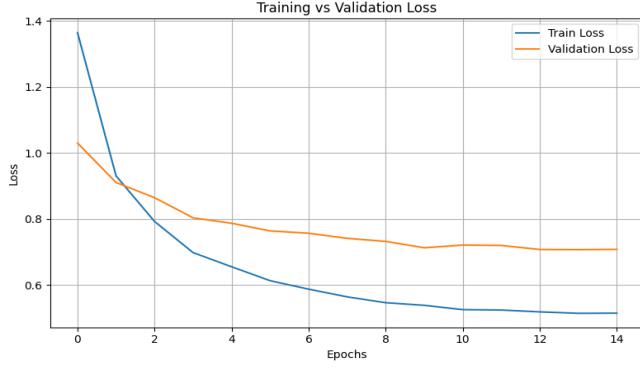


Fig 6. Training vs Validation Loss

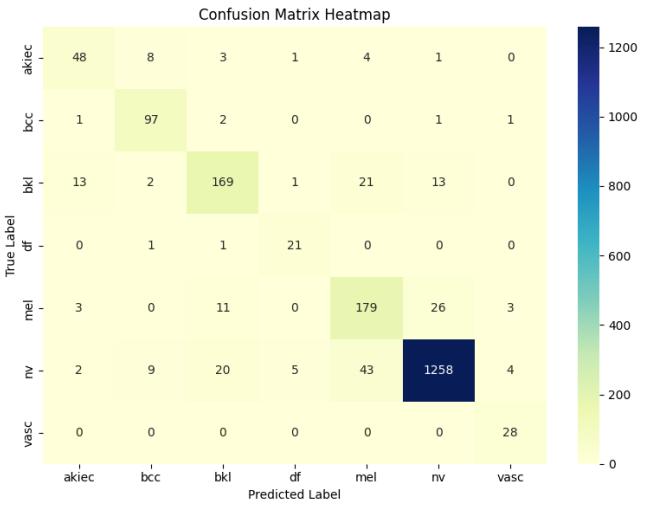


Fig 8. Confusion Matrix Heatmap

V. DISCUSSION & FUTURE SCOPE

In this paper has suggested a powerful multimodal deep-learning model combining a combination of dermoscopy and structured clinical data to give accurate classifications to skin lesions using the HAM10000 dataset. Using EfficientNetB4 architecture as the main feature extractor, the system can be used to extract very fine morphological differences, more complex pigmentation structures and structural abnormalities that define dermatological presentations. The concomitant addition of metadata such as patient age, sex and anatomical site also provides an extra dimension of diagnosis allowing the model to mimic real world clinical reasoning where patient context can be critical in decoding error-free interpretation.

The class-weighting methods also contribute to the increase of the model reliability as the level of class imbalance existing inherently in the dataset is counteracted. The minority groups of dermatofibroma, vascular lesions and actinic keratoses receive proportionate learning attention and hence even-handed and representative overall performance on all lesion types. The good diagnostic ability and balanced classification behavior of the system are verified by

quantitative assessment using the Accuracy, ROC-AUC, Precision, Recall, F1-Score, and confusion matrices.

Another major strength of the model is that the model is interpretable, which is enabled by Grad-CAM visualization. The resulting heatmaps indicate the particular areas of the lesions that were used to end the prediction, which correlates model attention with clinically significant features in terms of pigment networks, asymmetrical body structures, and individual vascular patterns. Such transparency is not only necessary to make AI systems stronger in terms of trust between them and clinicians, but also make sure that the predictions are based on explainable and medically meaningful evidence.

On the whole, the findings prove that diagnostic accuracy with the combination of dermoscopic image features and the structured clinical metadata is much better than that of image-only analysis. The resulting system has significant promise as a complementary clinical decision-support system, with quickly, uniform, and interpretable initial evaluations. The proposed architecture should be discussed as having a positive influence on the development of AI-assisted dermatology and can be viewed as a preliminary stage in building scalable applications into the real-world context, i.e., clinics and teledermatology applications.

VI. CONCLUSION

Multimodal deep-learning system created in the context of this work has shown the high potential of improvement of initial diagnosis of dermatological conditions due to its effective integration of dermoscopic images along with clinically relevant metadata. The model is built on HAM10000 dataset and it performs analysis on visual representation strength of EfficientNetB4, but with structured metadata inputs, namely age, sex and site on the skin of patient, which enables more finegrained and context-aware analysis of skin lesions. Such a twobranched design allows the system to mimic the diagnostic rationale of dermatologists who usually use information on the suspicious lesion as well as patient background during evaluation process. The performance measures of the model provide a clear indication of strength and dependability. Class-weighting leads to an effective removal of majority bias of classes, so the network can better identify underrepresented lesion forms including dermatofibroma and vascular lesions. Evaluation measures such as Accuracy, Precision, Recall, F1-Score, and macro ROC-AUC show good classification ability in seven diagnostic classes. Also, confusion matrix supports the fact that there is regularity in the decision behavior and there is less misclassification between similar lesions in terms of their clinical behavior. Due to the nature of Medical AI application, explainability is a requirement and GradCAM visualization is able to contribute to that to some extent. Heatmaps shown here depict the regions in each dermoscopic image which had the highest contribution towards the prediction of the model making the model transparent and interpretable. Such visual explanations would help build clinical trust and safe introduction of the system into practical workflow of dermatologist. The system conforms to the established clinical guidelines in sense that model takes into account the major morphological

features such as pigment networks, irregular borders and vascular patterns. Overall, results show that image data with orderly metadata would perform far better in the diagnostic performance compared to image only models. The architecture presented here is computationally efficient and clinical relevance is also introduced and offers scalable and accessible solution to early skin cancer and benign lesions detection. The proposed multimodal framework can be a reliable clinical decision-support system once refined, extended for larger size, connected to mobile or teledermatology systems and subsequently enhance dermatology services, especially in limited resource settings.

ACKNOWLEDGMENTS

With the deepest sense of gratitude, the authors want to thank Mr. Jai Kumar B., Assistant Professor, School of Computer Science and Engineering, Presidency University, Bengaluru for providing valuable guidance, committed supervision and passionate support during the period of conducting this research. He worked with machine learning and natural language processing background which added a lot of technical rigor and depth to this work. The helpful suggestions in preparing the manuscript are also appreciated by the authors to polish the methodological and analysis part.

REFERENCES

- [1] SANDHIYA DEVI M, SHANJANASREE P, SONA S, MS. SWETHA, "AI BASED TOOL FOR PRELIMINARY DIAGNOSIS OF DERMATOLOGICAL MANIFESTATIONS," May 2025.
- [2] Ling-Fang Li Xu Wang Wei-Jian Hu Neal N. Xiong Yong-Xing Du Bao-Shan Li, "Deep Learning in Skin Disease Image Recognition", 2020-11-01.
- [3] Nourah Janbi Rashid Mehmood Iyad Katib Ajjad Albeshri Juan M. Corchado Tan Yigitcanlar, "Imtidat: A Reference Architecture and a Case Study on Developing Distributed AI Services for Skin Disease Diagnosis over Cloud, Fog and Edge", 2022.
- [4] Zhouxiao Li, Konstantin Christoph Koban, Thilo Ludwig Schenck, Riccardo Enzo Giunta, Qingfeng Li, and Yangbai Sun, "Artificial Intelligence in Dermatology Image Analysis: Current Developments and Future Trends", 2022-11-18.
- [5] Prof. SR Kale, Mr. Mahesh Dandekar, Mr. Mahesh Bawane, Mr. Parth Dangre, Mr. Parth Sabale, "AI BASED TOOL FOR PRELIMINARY DIAGNOSIS OF DERMATOLOGICAL MANIFESTATIONS", 2025-04-01.
- [6] Yali Nie, Paolo Sommella, Marco Carratù, Matteo Ferro, Matthias O'Nils, and Jan Lundgren, "Recent Advances in Diagnosis of Skin Lesions Using Dermoscopic Images Based on Deep Learning", 2022-08-17.
- [7] Prof. S. R. Kale, Parth V. Sable, Parth S. Dangre, Mahesh R. Dandeker, Mahesh Bawane, "A Review on AI-Based Tool for Preliminary Diagnosis of Dermatological Manifestations", January–February 2025.
- [8] Rashmi N. Kuppast, Preetham R, K. Chandan, Chethan K, "AI TOOL FOR PRELIMINARY DIAGNOSIS OF DERMATOLOGY MANIFESTATION", 2024 (Vol-10, Issue-3, IJARIIE Journal).
- [9] Dr. Chandrasekar Vadivelraju, Keren Elisheba S, Vikhyath Moodbidri Bharath Kumar, Pavaman Suraj, Merugu Harish Reddy, "Dermascan AI: Deep Learning System for Preliminary Diagnosis of Dermatological Manifestations", 2025 (Volume 13, Issue 5, May 2025).
- [10] Prof. S. R. Kale, Parth V. Sable, Parth S. Dangre, Mahesh R. Dandeker, and Mahesh Bawane., "A Review on AlBased Tool for Preliminary Diagnosis of Dermatological Manifestations", 2025.
- [11] Zhouxiao Li, Konstantin Christoph Koban, Thilo Ludwig Schenck, Riccardo Enzo Giunta, Qingfeng Li, and Yangbai Sun, "Artificial Intelligence in Dermatology Image Analysis: Current Developments and Future Trends".

- [12] Yali Nie, Paolo Sommella, Marco Carratù, Matteo Ferro, Matthias O’Nils, and Jan Lundgren, “Recent Advances in Diagnosis of Skin Lesions Using Dermoscopic Images Based on Deep Learning”, 2022-08-17.
- [13] Esteva A., et al., “Dermatologist-level classification of skin cancer with deep neural networks,” *Nature*, 2017.
- [14] Codella N., et al., “Skin lesion analysis toward melanoma detection,” ISBI, 2018.
- [15] Brinker T.J., et al., “Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task,” *Eur J Cancer*, 2019.
- [16] Haenssle H.A., et al., “Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition,” *Ann Oncol*, 2018.
- [17] Tschandl P., et al., “Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks,” *JAMA Dermatol*, 2019.
- [18] Maron R.C., et al., “Systematic outperformance of 112 dermatologists in multiclass skin cancer image classification by convolutional neural networks,” *Eur J Cancer*, 2019.
- [19] Fujisawa Y., et al., “Deep-learningbased, computer-aided classifier developed with a small dataset of clinical images surpasses boardcertified dermatologists’ performance in skin tumour diagnosis,” *Br J Dermatol*, 2019.
- [20] Han S.S., et al., “Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm,” *J Invest Dermatol*, 2018.
- [21] Hekler A., et al., “Superior skin cancer classification by the combination of human and artificial intelligence,” *Eur J Cancer*, 2019.
- [22] Phillips M., et al., “Assessment of accuracy of an artificial intelligence algorithm to detect melanoma in images of skin lesions,” *JAMA Netw Open*, 2019.
- [23] Young A.T., et al., “Stress testing reveals gaps in clinic readiness of image-based diagnostic artificial intelligence models,” *NPJ Digit Med*, 2021.
- [24] Daneshjou R., et al., “Lack of transparency and potential bias in artificial intelligence data sets and algorithms,” *JAMA Dermatol*, 2021.
- [25] Adamson A.S., et al., “Machine learning and health care disparities in dermatology,” *JAMA Dermatol*, 2018.
- [26] Du-Harpur X., et al., “What is AI? Applications of artificial intelligence to dermatology,” *Br J Dermatol*, 2020.
- [27] Gomolin A., et al., “Artificial intelligence applications in dermatology: where do we stand?” *J Cutan Med Surg*, 2020.
- [28] Hogarty D.T., et al., “Artificial intelligence in dermatology—where we are and the way forward,” *Indian J Dermatol Venereol Leprol*, 2020.
- [29] Zakhem G.A., et al., “Characterizing the role of dermatologists in developing artificial intelligence for assessment of skin cancer,” *Dermatology*, 2021.
- [30] Navarrete-Decent C., et al., “Automated dermatological diagnosis: hype or reality?” *J Invest Dermatol*, 2018.