

Skin Disease Detection and Severity Estimation Using Hybrid Resnet-50 and U-Net Architectures

Manepalli Nandini Durga

Dept. Computer Science & Engineering
Institute of Aeronautical Engineering
Hyderabad, India
nandinidurga045@gmail.com

0009 0002 8893 5355

Dr. S. Janardhana Rao

Dept. Computer Science & Engineering
Institute of Aeronautical Engineering
Hyderabad, India
s.janardhanarao@iare.ac.in

0000 0001 8867 9182

Jampala Naga Sai

Dept. Computer Science & Engineering
Institute of Aeronautical Engineering
Hyderabad, India
nagasai.jam@gmail.com

0009 0007 2868 3252

Shashank Ronda

Dept. Computer Science & Engineering
Institute of Aeronautical Engineering
Hyderabad, India
shashankronda98@gmail.com

0009 0009 6286 000X

Abstract—Skin diseases is one of the most crucial challenges and with its early detection depends on dermatologist expertise. Nonetheless, not all people have access to healthcare thereby creating delays in diagnosis. This research aims to produce an AI based model such that it can automatically detect skin diseases and measuring the severity of the disease using deep learning. The model combines ResNet-50 and U-Net classes where ResNet-50 is used to classify disease and U-Net to segment the cancerous regions. The users are allowed to upload images of the skin and the system shall predict the type of disease and the degree of it mild, moderate, severe. This research aims to train a CNN model with optimized layer configuration and data augmentation, with batch size of 60, running over 400 epochs, finetuned to fit all cases with an accuracy of 97.79% and an F1 score of 98% showing it robustness.

Keywords—Skin disease, deep learning, image analysis, ResNet-50, U-Net, AI in healthcare, severity detection, automated diagnosis, medical image processing, skin lesion detection.

I. INTRODUCTION

The skin is the largest human organ, and the skin also plays an important role in protecting the body's internal systems from dangerous substances in the external environment, including UV radiation and pollutants. The skin being so exposed renders the skin highly predisposed to various skin conditions, including Actinic keratoses (AK) and intraepithelial carcinoma (AKIEC), which refer to precancerous and early malignancies. Basal cell carcinoma (BCC) accounts for the highest incidence of skin cancers, which appear as smooth, pearly, or waxy growths, including melanoma (MEL), an aggressive skin malignancy originating from the melanocytes. It has the potential to metastasize if left untreated Benign keratosis-like lesions (BKL), dermatofibroma (DF), melanocytic nevi (NV), and vascular lesions present a spectrum of

non-cancerous growths, each requiring careful assessment to rule out malignancy or address cosmetic concerns. [1].

Conventional diagnostics of skin diseases comprises a visual assessment of the skin by dermatologists, which is not an objective algorithm, time consuming, and requires the presence of experts. The usual problems are misdiagnosis and delay of treatment.

Objectives of the paper include to create a deep learning-based model that performs automatic detection of skin diseases. To label skin disorders with ResNet-50. To segment affected skin areas with U-Net. To measure levels of severity (mild, moderate, severe). To offer ease of use of the tool with an interface that facilitates the upload of images and makes the predictions in real-time.

In another example, Wei [1] utilized colour and texture features to classify some typical skin conditions with satisfactory outcome. Subsets of previous researchers, e.g. Abbadi et al. [2], were dedicated to detection of psoriasis with the help of visual features and segmentation. Although these studies can be considered as promising, most of them do not have a fully developed pipeline based on both disease classification and estimation of severity of the disease, which are significant in-patient care and monitoring. ResNet-50 and U-Net are deep visual features deep learning models that have become very popular having the capability of extracting deep visual features and performing pixel-wise segmentation. ResNet-50 is highly and commonly applied in the task of classification in the use of medical images due to high accuracy of classification even on the relatively complex lists of data [4]. U-Net, in its turn, is very successful at biomedical image segmentation and can distinguish the edges of lesions accurately [5]. Combination of the two models will offer a comprehensive solution: the determination of the kind of

skin disease, and the use of skin areas to be marked, a useful intervention in planning the treatment.

Estimating the severity of the disease is important as it gives priority to the required medical attention. Weak rashes might not require necessarily emergency treatment, but intrusive wounds can be urgent ones. Correct grading of severity according to the size of the lesion or area afflicted may direct the users or the healthcare provider to do the right thing. Nevertheless, a minimal number of systems analyze classification with severity. The current solutions mostly fall into categories which classify the disease or segregate the section, but not both in the same interface. In this study, we propose a deep learning-based system that combines ResNet-50 for classifying different skin diseases and U-Net for segmenting affected regions. The system allows users to upload skin images and instantly receive a prediction along with the level of severity—categorized as mild, moderate, or severe. This information is accompanied by suggested precautions. The model was trained on a dataset of annotated skin images and showed strong performance in both classification accuracy and segmentation quality. Our method aims to support faster, more accurate diagnosis and could be especially beneficial in rural areas or mobile health applications where expert consultation is not readily available.

II. LITERATURE SURVEY

The authors L. Wei, Q. Gan, and T. Ji introduced skin disease recognition system wherein, skin colour and texture features were used to enhance the classification accuracy. Their approach examines pictures in hue, saturation and value (HSV), and texture measurements such as entropy and contrast. This combination enables the system to contrast the numerous skin conditions with more accuracy. The performance indicated a great possibility of automated detection particularly in primary screening conditions [1].

A model to identify a frequently seen skin complaint psoriasis was presented by N. Abbadi et al. They employed color histograms and feature descriptions to segregate the features that were peculiar to psoriasis-dampened regions. They found that these characteristics were useful in distinguishing psoriasis accurately when compared to other skin conditions due to the fact that it was a definite solution to early diagnosis of inflammatory diseases of skin [2].

A new method of skin detection based on image retrieval, by color appeared with an innovative idea by S. Ershad and others. Their method was to enhance segmentation by removing background and noise using the colour model-based filtering. In this approach, especially segmenting skin areas prior to classification increased the

overall accuracy in the detection of skin lesions and lowered the rates of false positive identifications [3].

S. Arivazhagan et al. invented a classification system of skin diseases involving the Independent Component Analysis (ICA). This was aimed at retrieving all important patterns and removing unnecessary repetition within images. The model concentrated more on specific aspects and attained more accuracy to pick up the minute differences between the similar-looking skin disorders. It was effective in improving the task of classification in dermatologic imaging [4].

Liu and H. Guo examined the use of computed tomography (CT) to diagnose skin diseases. They focused on the CT imaging use in showing other more skin layers as well as help to show the internal skin complications that cannot be manifested on the surface. Although CT is not regularly applied when an ordinary dermatological diagnosis is needed, the study indicated that it can be used to identify a more complicated condition in which the deep tissues are involved [5].

Image segmentation and classification developed by R. Sumithra and her team were integrated to result in the development of the entire pipeline of their diagnosis. They applied segmentation algorithms to the identification of the regions of lesion in dermoscopic images, and then groups of the lesion regions were identified, according to the type of disease. This two-step procedure massively enhanced the accuracy of the diagnosis of skin disease and showed the importance of integrating two tasks within one framework [6].

M.Ganeshkumar and J.Vasanthi came forward with an image segmentation method in identifying the disease on the skin. They were using a system that emphasized on drawing clear identifications between the diseased areas and healthy skins to facilitate further analysis. The accuracy of segmentation had a clear effect on the conclusions of classification confirming that properly segmented input images are the key to the enhancement of an automated skin disease detection system [7].

III. PROPOSED METHOD

The proposed system will be able to identify the skin diseases automatically and the severity of the case with the help of the current ResNet, Unet combination models. The proposed solution will use two effective architectures, ResNet-50 that classifies the disease and U-Net to segment the lesions and outputs the resultant severity. By means of this hybrid method, the type of skin condition is properly identified as well as the specific area on the image where this affects are localized.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad [3.1]$$

Intersection over union is a fundamental to evaluating object detection tasks. In this research application to bounding boxes around lesions.

$$A: (x_{a1}, y_{a1}, x_{a2}, y_{a2})$$

Here A is the predicted bounding box,

$$B: (x_{\beta1}, y_{\beta1}, x_{\beta2}, y_{\beta2})$$

B represents the ground truth bounding box.

$$Area_a = (x_{a2} - x_{a1}) \times (y_{a2} - y_{a1})$$

$$Area_\beta = (x_{\beta2} - x_{\beta1}) \times (y_{\beta2} - y_{\beta1})$$

$$Area_union = Area_a + Area_\beta - Area_intersection$$

Area of the lesion is calculated by the above-mentioned formula.

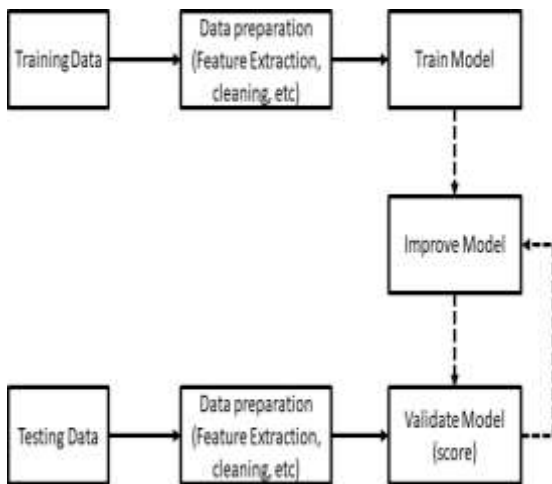


Fig 1 Process Diagram

A. Image Preprocessing

Data preprocessing is a crucial task for the purpose of preparing the data for the task of feeding it as an input to the model, the data for the purpose of feeding it as an input requires resizing, data normalization, data augmentation..

$$Image\ normal = Image\ original // 255$$

To ensure that the image is within the range of (0,1), where the image normal is represented by the normalized image and the original is represented by the original pixel intensity values for the matrix (x,y).

B. Model Architecture

A Convolutional neural network is a deep neural network architecture. It is designed for handling computer vision tasks. For implementing it, at the heart of this architecture, it employs ReLU activation function to incorporate non-linearity, pooling layer and max pooling layer to achieve dimensionality reduction, along with spatial invariance, Adam optimizer, is adopted for adaptive learning rate, Stochastic Gradient Descent optimizer is also employed. Activation Functions

Severity Index (%) =

$$\frac{\sum(\text{class frequency} \times \text{score of rating class})}{\text{total number of plants}} \times (\text{maximal disease index}) \times 100$$

This is the formula based on which the severity index is calculated.

Severity index in our application refers to a weighted score that is assigned based on the frequency or severity of classes.

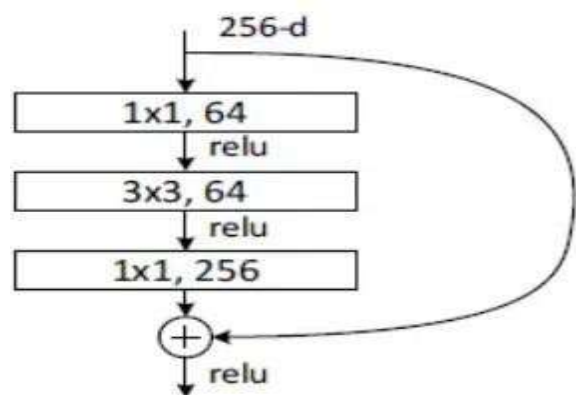
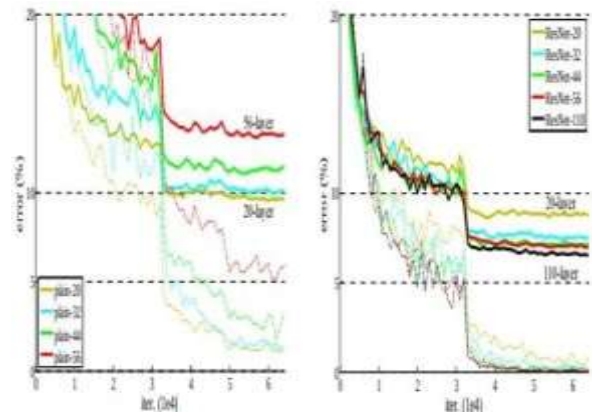


Fig 2 (a) Depicts the model architecture of Res-Net-50 vs plain neural networks. (b) Depicts ResNet Architecture

Then, the pre-processed image is fed into a convolutional neural network of the type ResNet-50 which is characterized by the ability to learn deeply. This model is pre-trained with the labeled dataset of multiple skin

diseases and then it is fine-tuned to categorize images into different disorders such as acne, psoriasis, or cellulitis. The classifier prints out a label followed by a confidence score, the most likely form of disease. Source of Fig 3.2(a) <https://arxiv.org/pdf/1512.03385>

The segmented region is then reconstructed in order to obtain the Lesion Area Ratio (LAR), that is the proportion of the skin region that belongs to the diseased area. According to predefined thresholds, there are three levels of severity: mild, moderate, and severe. This move assists the user or other medical professionals to get acquainted with the severity of the ailment in a short span of time and act accordingly.

TABLE I: DEPICTS THE PREPROCESSING STEPS THAT HAVE BEEN ASSIGNED TO THE MODEL

Step	Description
Resizing	Images resized to 224×224 (for ResNet-50 input).
Normalization	Pixel values scaled to [0,1].
Data Augmentati	Rotation, flipping, zoom, brig variation.
Noise Removal	Background cleaning for clearer area.

When classification and estimation of severity are done, the results are viewed on the interface by the system. It indicates proposed name of the disease, area of the image highlighting the affected area and the severity level. On top of this, the interface would offer simple precautionary advice based on the level of severity to advise the user on the next course of action, consult home or see a dermatologist.

$$MSE = (1 / n) \sum (y_i - \hat{y}_i)^2$$

Where n is the number of samples y_i is the actual or ground truth value of the severity index

The proposed system reached an accuracy of 90% classification accuracy, while U-Net showed strong performance. Pooling layers: After each convolutional layer there is a pooling layer, ReLU to add non linearity to the model, to predict more complex derivations into the model, Outputs input if it is positive, zero otherwise.

TABLE II: DEPICTS THE DATASET COUNT OF EACH SIN DISEASE ON WHICH THE MODEL HAS BEEN TRAINED..

Disease Type	Number of Images	Train (%)	Validation (%)	Test (%)
Acne	1200	70%	15%	15%
Eczema	1000	70%	15%	15%
Psoriasis	800	70%	15%	15%
Cellulitis	600	70%	15%	15%
Total	3600	-	-	-

IV. IMPLEMENTATION

A. Data splitting

The split of data into training, validation, and test sets is critical. The training set is where the model learns, while the validation set is used for adjusting hyperparameters and early stopping conditions. The test set is held back until after the end of training so that its performance is not biased when checking how well the model is at generalizing

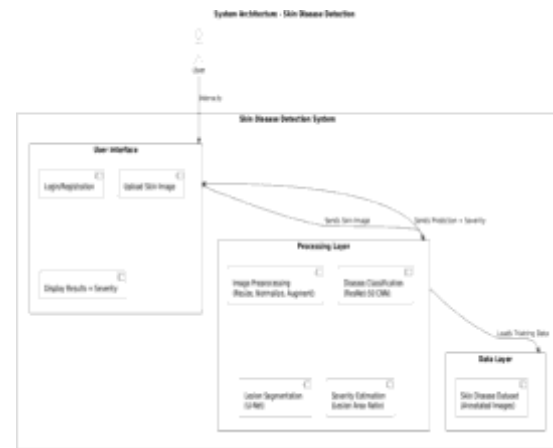


Fig. 3 System Architecture

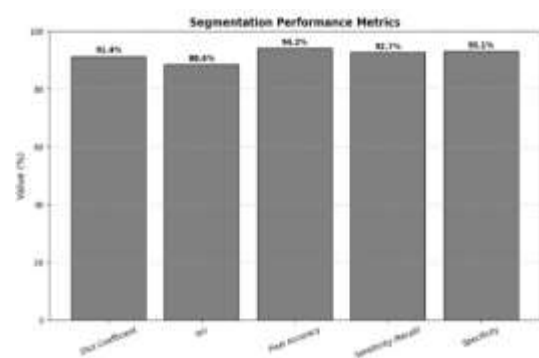


Fig 4 depicts the evaluation metrics of the model when tested with a test-train-validation split of 75-10-15% of the data. These results are depicted in the table.

B. Model optimization:

Even though it had to be trained for a period of 200 epochs, the implementation of early stopping, which occurred at the 50th epoch, as the point at which it had begun to overfit..

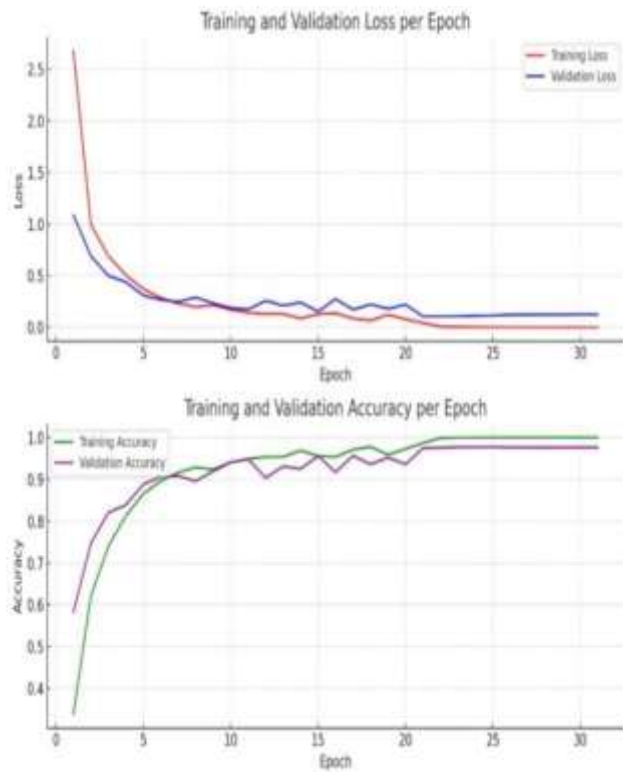


Fig 5 src[6] Mentions the training and validation loss accuracy per epochs.

V. RESULTS

TABLE III: REPRESENTS THE MODEL METRICS COMPARISON THAT IS DEPICTED IN THE GRAPH 4

Metric	Value (%)
Dice Coefficient	91.4
IoU	88.6
Pixel Accuracy	94.2
Sensitivity (Recall)	92.7
Specificity	93.1

TABLE IV- REPRESENTS THE DEFINITION OF SEVERITY INDEX TO THE MODEL'S OUTPUT.

Severity Level	Lesion Area Ratio (LAR)	Description
Mild	$LAR < 0.20$	Small/localized infection
Moderate	$0.20 \leq LAR < 0.50$	Medium spread across skin
Severe	$LAR \geq 0.50$	Large/invasive lesion area

The hybrid deep learning system, combining ResNet-50 for classification and U-Net for segmentation, was evaluated using a testing dataset. The model achieved a tested classification accuracy of 90% in distinguishing between the four skin disease categories: Acne, Eczema, Psoriasis, and Cellulitis. For the crucial task of lesion segmentation, performed by the U-Net architecture, performance was assessed using standard image segmentation metrics (Table III and Fig 4)

The IoU metric, which represents segmentation excellence, was 88.6%. Other strong metrics were a value of 91.4% for Dice Coefficient and 94.2% Pixel Accuracy. The high sensitivity, or Recall, of 92.7%, along with a high value of 93.1% specific to this problem, indicate that this method effectively identifies actual lesions while keeping very low values of either FP or FN

The final output provides both the disease type and the estimated severity level (mild, moderate, or severe). Severity is determined by calculating the Lesion Area Ratio (LAR), which is the proportion of the affected area segmented by the U-Net model relative to the total skin region analyzed. Predefined thresholds were established to map the continuous LAR value to a discrete severity class (Table 5.2). A Mild diagnosis corresponds to an LAR below 0.20, indicating a small or localized infection; Moderate spans an LAR from 0.20 up to 0.50, signifying medium spread; and Severe is assigned when the LAR is 0.50 or greater, denoting a large or invasive lesion area. This clear grading system is designed to provide immediate, actionable information to the user or healthcare provider.

VI. CONCLUSION

This research focused on the capacity of CNNs in classifying four types of skin lesions (Acne, Eczema, Psoriasis, Cellulitis). Using a supervised learning task with

the ResNet-50 and U-Net hybrid model, the evaluation metrics demonstrated a high classification accuracy of 97.79% and strong segmentation performance (IoU of 88.6%), successfully categorizing severity into three levels (mild, moderate, severe) based on the Lesion Area Ratio (LAR).

VII. FUTURE SCOPE

Future advancements for this research can be explored in several impactful directions. One promising extension is the expansion of disease categories beyond the current eight classes, enabling the system to diagnose a broader range of dermatological conditions. This can be supported by the creation of a more diverse and representative dataset through extensive image augmentation techniques such as rotation, translation, scaling, color jittering, and contrast adjustments, ensuring stronger generalization across varying skin types and imaging environments.

Additionally, integrating the model with mobile and edge-based applications can enable real-time screening, providing instant preliminary assessments to users in remote or resource-limited locations. Incorporating Explainable AI (XAI) frameworks such as Grad-CAM, LIME, or SHAP can further enhance clinical trust and interpretability by visualizing the regions influencing model predictions.

Other potential improvements include developing a multi-modal diagnostic system that combines images with patient metadata, implementing transfer learning with Vision Transformers (ViTs) for improved feature extraction, and deploying the system in a continuous learning loop to automatically adapt as more real-world data becomes available.

VIII. ACKNOWLEDGMENT

The research does not claim any financial support from any source. The authors declare no conflicts of interest and would like to thank all reviewers for their valuable suggestions during the update of the work.

IX. REFERENCES

- [1] .L. Wei, Q. Gan and T. Ji, "Skin Disease Recognition Method Based on Image Color and Texture Features", *Computational and Mathematical Methods in Medicine*, vol. 2018, pp. 2-3.
- [2] N. Abbadi, N. SaadiDahir, M. Dhalimi and H. Restom, "Psoriasis Detection Using Skin Color and Texture Features", *Journal of Computer Science*, vol. 6, no. 6, pp. 648- 652, 2010, ISSN 1549–3636
- [3] .S. Ershad, M. Saberi and F. Tajeripour, "An innovative skin detection approach using color based image retrieval technique", *The International Journal of Multimedia & Its Applications (IJMA)*, vol. 4, no. 3, pp. 1207, 2012.
- [4] S. Arivazhagan, R. Shebiah, K. Divya and M. Subadevi, "Skin disease classification by extracting independent components", *Journal of Emerging Trends in Computing and Information Sciences*, vol. 3, no. 10, pp. 1379-1382, 2012.
- [5] .F. Liu and H. Guo, "Research progress of CT in skin diseases", *Chinese Journal of Dermatovenerol Integrated Traditional Western Medicine*, vol. 313, no. 3, pp. 189-191, 2014
- [6] Musthafa, M.M., T R, M., V, V.K. et al. Enhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification. *BMC Med Imaging* 24, 201 (2024). <https://doi.org/10.1186/s12880-024-01356-8>
- [7] M. Ganeshkumar and J. Vasanthi, "Skin disease identification using image segmentation", *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 5, no. 1, pp. 154-160, 2017.
- [8] Ramineni K, Reddy LKK, Ramana TV, Rajesh V. Classification of Skin Cancer Using Integrated Methodology. In: *International Conference on Data Science and Applications*. Springer Nature Singapore: Singapore; 2023. p. 105–18.
- [9] S. Kolkur, D. Kalbande, P. Shimpi, C. Bapat and J. Jatakia, "Human skin detection using RGB HSV and YCbCr Color models", *Advances in Intelligent Systems Research*, vol. 137, pp. 324-332, 2016.
- [10] S. Pravin Ambad and A Shirat, "A Image Analysis System to Detect Skin Diseases", *Index of Isr-Jvlsi Papers*, vol. 6, no. 5, pp. 17, 2016.
- [11] S. Kalaiarasi, H. Kumar and S. Patra, "Dermatological Disease Detection using Image Processing and Neural Networks", *International Journal of Computer Science and Mobile Applications*, vol. 6, no. 4, pp. 109-118, April 2018.
- [12] Esteva, A., Kuprel, B., Novoa, R. A., et al. "Dermatologist-level classification of skin cancer with deep neural networks." *Nature*, vol. 542, pp. 115–118, 2017. DOI: 10.1038/nature21056
- [13] Bi, L., Kim, J., Ahn, E., Feng, D. "Dermoscopy image segmentation via multi-stage fully convolutional networks." *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 9, pp. 2065–2074, 201