Intelligent Skin Disease Detection Using Ensemble Convolutional Neural Networks with EfficientNet and ResNet

[1]Ms. Abhilasha S Chougale,[2]Mr. Rajwardhan S. Todkar

[1]PG Scholar,[2]Assistant Professor
[1]CSE- Data Science, [2]CSE- AIML
[1]D. Y. Patil Agriculture & Technical UniversistyTalsande, Kolhapur,India.

Abstract:Skin disorders are a worldwide problem which affects millions of people and are a major health concern because of the wide range of appearance they may exhibit and the different severity that they may have. Prompt and correct diagnosis is essential to stop the progression of disease, lower treatment costs and enhance patient outcomes. In recent times, deep learning has progressed to the point where it is possible to build intelligent diagnostic systems that can precisely analyze dermatological images. This research introduces an ensemble classification approach by utilizing both EfficientNet and ResNet architecture to increase the effectiveness of automatic skin disease prediction. By balancing the network scaling, EfficientNet helps in optimized feature extraction, and ResNet's residual connections help to overcome the gradient degradation problem, enabling deeper feature learning. The proposed ensemble approach is a fusion of the prediction power of both models, which gives better classification results than the individual networks. The system developed in this work includes a detailed image preprocessing, data augmentation and normalization methods to enhance the model generalization and to reduce overfitting. A curated dataset of images of various types of skin diseases is used for training and assessment. Transfer learning and fine-tuning techniques are implemented to make use of the weights of the pre-trained model to efficiently adapt the model to the characteristics of dermatological images and decrease the complexity of training. The ensemble framework combines the predictions of EfficientNet and ResNet together via a model fusion mechanism, enabling complementary feature representations to help with final disease predictions. The performance of the proposed solution is evaluated by commonly used evaluation metrics like accuracy, precision, recall, F1 score and confusion matrix analysis. According to the experimental results, the ensemble model has a better classification accuracy and robustness, and it is highly reliable to distinguish the visually similar skin conditions, compared with the individual deep learning architectures. The system that is proposed is scalable and practical in assisting computer-aided dermatological screening and clinical decision support. The framework helps healthcare providers with the quick and accurate diagnosis of various diseases, promotes AI-powered healthcare applications and enables the creation of accessible and efficient dermatological diagnostic services.

Keywords -Skin Disease Prediction, Deep Learning, Ensemble Learning, EfficientNet, ResNet, Convolutional Neural Networks, Dermatology Image Classification, Medical Imaging, Transfer Learning, Feature Extraction, Hybrid Model, Classification Accuracy, Computer-Aided Diagnosis, Model Fusion, Artificial Intelligence, Healthcare Automation, Disease Detection.

I. INTRODUCTION

Skin diseases are one of the most common medical conditions globally and impact people of all ages, geographical regions, and socioeconomic status. The rise in dermatological diseases has raised the need to detect the problem in time and treat it appropriately to avoid its complications and enhance patient quality of life. The diagnosis of skin diseases is traditionally based on visual examination, clinical experience, dermoscopic examination and laboratory exams. The diagnostic process is sometimes difficult, however, because there are visually similar symptoms in many skin disorders. Manual diagnosis may be complicated and time consuming because of variations in lesion color, texture, shape, and size. Also, the increasing number of patients and the scarcity of skilled dermatologists have led to the need for automated diagnostic systems for assisting clinical decision making. Therefore, computer-aided diagnosis using image information has become a potential solution to improve the efficiency and uniformity of dermatological evaluation. Medical image analysis has been revolutionized by recent advances in artificial intelligence, especially deep learning. One of the most successful deep learning techniques has been Convolutional Neural Networks (CNNs) that have achieved very good performance in learning hierarchical visual features directly from image data without the need for feature engineering. CNN-based approaches have successfully been applied to medical imaging, disease detection and classification. Some of the latest CNN architectures that have achieved outstanding results in high-level image recognition applications include efficientNet and ResNet. To optimally extract features, EfficientNet uses a compound scaling approach that allows for network depth, width and resolution to be optimised for computational effectiveness and feature extraction efficiency. ResNet, by contrast, adds residual learning mechanisms that let deeper neural networks to be successfully trained to address the vanishing gradient problem in learning. Such properties render both architectures well suited to skin pattern recognition of various skin diseases.

CNNs might experience some restrictions if faced with a wide variety of medical data sets that are heterogeneous in nature. The classification of skin diseases is a challenging classification task because of its high intra-class variability and low inter-class distinguishability, meaning that different diseases can have similar visual appearance and the same disease can look different on different patients. These complexities can limit the generalization power of the single-models and raise the concerns about misclassification. This is where ensemble learning can help overcome these limitations by combining the strengths of several deep learning models. Ensemble methods, which integrate complementary feature representations and prediction capabilities, can lead to higher robustness, higher accuracy and higher stability. The proposed research employs an ensemble framework of efficientnet and resnet to benefit from efficient feature scaling and deep residual learning to deliver more accurate prediction of skin disease.

To implement an effective skin disease prediction system, a coordinated workflow of preparing the dataset, preprocessing the images, training the system, and evaluating the performance of the system needs to be created. The study includes extensive preprocessing techniques in order to enhance the image quality and data consistency. To simulate real-world variations and improve model generalization, data augmentation techniques like rotation, flipping, brightness adjustment, scaling, and noise injection are used. In addition, transfer learning methods are implemented to leverage the pre-trained weights from the large-scale image datasets, which leads to a mechanism to efficiently learn even with the relatively small medical datasets. Procedures for fine-tuning are then carried out to adjust the pretrained models to the properties of dermatological images. The outputs from EfficientNet and ResNet are combined using an ensemble fusion approach, providing a combined classification system with enhanced diagnostic performance. The efficiency of the framework is rigorously tested with various evaluation metrics, such as accuracy, precision, recall, F1-score, confusion matrix analysis to comprehensively assess model behaviour.

II. LITERATURE REVIEW

The automated skin disease classification has witnessed a large boost in research in recent years with the availability of large-scale dermatological image repositories like HAM10000, ISIC, PH2, and SIIM-ISIC. These are publicly available data sets and are used as standard reference for test the performance of deep learning models in dermatological image analysis. During 2022, there were numerous studies that extensively used transfer learning methods using ImageNet pretrained CNN models such as ResNet, EfficientNet, DenseNet or MobileNet architectures. The researchers also strengthened the robustness of the model by implementing various image pre-processing steps like hair artefact removal, color normalization, lesion segmentation and image enhancement. Multiple studies also delved into the self-supervised learning methods and contrastive learning for overcoming the issues of scarcity of annotated data and the large class imbalance. Furthermore, the multimodal learning frameworks that incorporated dermoscopic images, clinical photographs, and patient information, combined with powerful augmentation strategies like GAN-based image synthesis and test-time augmentation, showed greater generalization potential. In general, the research direction in 2022 moved away from the traditional single model training, to more complex learning paradigms, such as contrastive representation learning, fairness-aware optimization and multimodal fusion, addressing real-world clinical variability.

SuperCon: Supervised Contrastive Learning for Imbalanced Skin Lesion Classification, Chen, K.; Zhuang, D.; Chang, J. M. (2022). This study proposed a two-stage learning framework to tackle the class imbalance problem in the skin lesion datasets. Supervised contrastive learning was first used to train discriminative feature representations, by encouraging samples with the same class label to cluster together and samples with different class labels to move apart. The learned representations subsequently underwent a classification layer that was fine-tuned. On the experimental results with ISIC and HAM10000 datasets, we achieved much better F1-score and AUC, especially for the underrepresented classes. It can be well adapted to the EfficientNet and ResNet architectures and is very useful for ensemble-based classification systems since it can produce better and more robust feature representations. [1]

CIRCLe: Color Invariant Representation Learning for Unbiased Classification of Skin Lesions.Pakzad A, Kumar A, Hamarneh G (2022). The CIRCLe framework introduced a colour-invariant learning strategy to tackle the issue of skin-tone-related bias in the classification of dermatological images. The technique was based on a novel regularization strategy which aimed to ensure that the latent feature representations of images with the same diagnosis were similar for images with different skin tones. Testing on over sixteen thousand images across various Fitzpatrick skin categories found enhancements in fairness measures with high classification performance. The method provides a feasible approach to mitigate demographic bias in EfficientNet–ResNet ensemble models designed for clinical use in real-world settings. [2]

Alam, T. M., et al. – An Efficient Deep Learning Based Skin Cancer Classifier (2022). This work proposed a transfer learning based approach using the fine-tuned CNN approach for skin cancer classification on the HAM10000 dataset. After classifying the images, they were preprocessed using image preprocessing techniques like hair artifact removal and color normalization to enhance the image quality for classification. The EfficientNet variants and augmentation and cross validation strategies resulted in high classification accuracy of more than ninety per cent. The research revealed that ensemble-based models offer better uniformity and robustness compared to single models, and emphasized the significance of training configuration and optimization techniques. [3]

Alwakid, G. et al. – "Melanoma Detection using Deep Learning-Based Classifications" (2022). In this work, the authors have explored the detection of melanoma using image enhancement and deep learning classification. Prior to classification, ESRGAN was used to enhance the image quality with a modified ResNet-50 network architecture. Experimental results demonstrated that it was able to achieve better accuracy, F1-score, and AUC than the classic ResNet implementations. The results highlight the significance of image enhancement methods, especially when dealing with poor-quality images produced by mobile phones and in teledermatology systems. [4]

V. Anand, et al. – “Multi-class Skin Disease Classification Using Transfer Learning” (2022). The research in this work was done on the classification of skin disease of multiple classes using pre-trained CNN architectures like ResNet and MobileNet. To tackle dataset imbalance, the researchers added class-aware augmentation, focal loss functions, and balanced sampling. They found that transfer learning and augmentation proved to be effective for boosting the classification accuracy in the absence of sufficient training data. The study gives useful insights into the model building approaches in the practical field of image classification systems for skin problems. [5]

Skin Lesion Classification of Dermoscopic Images Using Deep Convolutional Networks.B. Shetty, M. Gowda, M. Anitha, R. Priya, and S. Shekhar. (2022). Skin Lesion Classification of Dermoscopic Images Using Deep Convolutional Networks. The authors suggested that a hybrid classification approach that incorporates deep feature extraction and traditional machine learning methods could be employed. Features extracted from the transfer-learned ResNet architectures were cleaned up using Kernel Principal Component Analysis (KPCA) and then classified using Support Vector Machines (SVM). Experimental results showed high accuracy on seven categories of skin lesions, suggesting that the feature-level fusion can be a useful alternative to an end-to-end deep learning pipeline in ensemble systems. [6]

Wang, Y., Feng, Y., et al. – Adversarial Multimodal Fusion with Attention Mechanism for Skin Lesion Classification using Clinical and Dermoscopic Images (2022). In this research, an adversarial multimodal fusion method was proposed, which aims to learn the representations of both modalities and the shared representations from clinical and dermoscopic images. They were then incorporated with an attention mechanism and an adversarial learning component to help the model learn complementary diagnostic information between modalities. The proposed approach outperformed the single-modality systems, and emphasized the importance of multimodal integration in enhancing the classification accuracy and robustness. [7]

Du, S.; et al. – "FairDisCo: Fairer AI in Dermatology via Disentanglement and Contrastive Learning" (2022). FairDisCo proposed a fairness-informative learning framework, using disentanglement and contrastive learning methods to filter out the skin tone information without compromising the disease-relevant information. Experimental results indicated a decrease in performance differences between the various skin-tone groups with good diagnostic performance. The study offers valuable insights for creating fair and clinically responsible dermatological AI systems. [8]

Y. Wu, et al. - Federated Contrastive Learning for Dermatological Image Analysis (2022). In this study, two types of federated and self-supervised learning were investigated to train dermatological classification models in multiple healthcare institutions. The framework's generalization capability was enhanced using the contrastive learning objectives, and patient privacy was maintained while using heterogeneous datasets. This proposed strategy is useful especially for creating a diverse set of pretrained models to be used in ensemble learning frameworks. [9]

III. PROPOSED SYSTEM

The architecture is proposed to be a structured and sequential flow consisting of different computational stages starting from the acquisition of skin disease images, followed by various computational stages, and ending with the prediction of skin disease and recommendations for treatment. The framework shows how dermatological images are processed by a network of modules in order to identify the diseases with accuracy. The individual parts of the architecture have particular roles and work to improve the system's effectiveness. Modularity allows for the replacement, optimization and update of individual components without impacting the entire flow process. This flexibility enables scalability and the ability to handle larger datasets and more complex deep learning models in future iterations of the framework. The left-to-right operational flow depicts the progressive transformation of raw dermatological images to significant clinical insights. The different processing steps are clearly separated to enhance interpretability and to show the contributions of the different modules in the ensemble classification framework. In summary, the architecture depicts a complete picture of the automated skin disease prediction capability of the proposed ensemble deep learning system.

In the first phase of the framework, dermatological images will be gathered from publicly available datasets, clinical repositories, and mobile platforms for image acquisition. Such images are frequently of different resolutions, illumination, image quality and acquisition environments, necessitating preprocessing before they can be used for modelling. The preprocessing module includes image resizing, normalization, enhancement, and augmentation, among others, to ensure uniformity of input samples. Various techniques, including rotation, flipping, scaling, brightness and contrast changes, are used to make the datasets more diverse and for greater robustness of the model. These procedures minimize the effects of noise and irrelevant visual artifacts and maintain the diagnostically important lesion features. Moreover, augmentation strategies alleviate overfitting by introducing model with more visual variations. This, in turn, provides a basis for reliable learning and enhances the generalization performance of the model in terms of environmental conditions, imaging devices, and skin types.

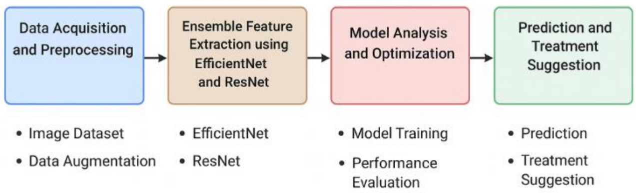

Fig.1 – Architecture Diagram

The second key part of the architecture includes "ensemble feature extraction", where EfficientNet and ResNet are two very effective architectures of convolutional neural networks. EfficientNet benefits from capturing fine-grained spatial and textural information, with its compound scaling approach, while ResNet allows for deeper feature learning by using residual connections to address the gradient degradation issue. In the proposed configuration, two networks are used to process the preprocessed images concurrently, to achieve complementary feature extraction from the same input data. The resulting feature representations are then fused to form a more descriptive and discriminative skin lesion representation. This fusion approach improves the ability of the system to differentiate image sets that are very similar in terms of appearance and may not be easily categorized by a single model. The ensemble feature extraction module is the brain of the system and is the main method used to enhance the classification accuracy, robustness and diagnostic reliability over the accuracy of any individual CNN architecture.

The third stage of the framework deals with model training, analysis, and optimization. The fully connected layers are used for the classification of the diseases, and thefeature representations are fused before passing them. There are several optimization techniques that are used during training, such as hyper parameter tuning, loss function evaluation, gradient optimization, etc., and performance monitoring. Other feature enhancements like dropout regularization, learning rate scheduling and weight optimization are added during training to improve convergence and avoid overfitting. The architecture also incorporates a detailed evaluation mechanism that monitors the performance of the model continuously, capturing metrics like accuracy, precision, recall, F1-score and confusion matrix analysis. These procedures assure effective model learning and stable and reliable predictive performance. The system is continually improved and validated to achieve the highest possible classification accuracy while maintaining good generalization performance.

The last step in the architecture is prediction and recommendation of treatment for the disease. Once the classification is finished, the ensemble model will give the result of the skin disease category predicted and the percentage of confidence of the diagnosis. From those outputs, a rule-based knowledge module generates preliminary suggestions for treatment and healthcare advice. Depending on the condition identified, the system might offer preventive strategies, self-care tips, and referrals for medical consultation to a dermatologist if needed. The recommendation mechanism is not meant to take the place of clinical expertise, but rather to help assist in the decision-making process. The final module turns the deep learning predictions into actionable health care information, adding to the system's value for research and clinical applications. Therefore, the entire system can be considered as a smart diagnostics assistance system that can support the early detection of diseases and encourage timely medical treatment.

IV. METHODOLOGY

4.1. Data Acquisition and Preprocessing

The proposed methodology starts by acquiring a comprehensive dermatological image set from public image repositories and clinical images. The dataset aims to contain categories of skin diseases, varying characteristics of the diseases, and various skin tone representations to strengthen the model generalization capacity. Each image is preprocessed to standardize the size of the images, crop the images, correct illumination, normalize the images, and reduce the noise. To enhance the model's robustness and add diversity to the dataset, data augmentation methods like horizontal and vertical flipping, rotation, brightness adjustments, contrast adjustments, and zooming are used. All images are converted to fit into EfficientNet and ResNet input requirements. In addition, label verification is performed to reduce error in the labelling process and enhance the reliability of training. The data is then randomly split into training, validation and test sets, with stratified sampling to ensure class balance. The preprocessing steps guarantee the consistency, balance, and appropriateness of the input data for effective ensemble learning.

4.2. Ensemble Feature Extraction Using EfficientNet and ResNet

After pre-processing, a combination of EfficientNet and ResNet architectures is used for feature extraction. Through efficientNet, the compound scaling principle is used to provide detailed spatial and texture information, while keeping computational cost low. ResNet, however, employs the mechanisms of residual learning which enable deeper feature extraction and mitigate the vanishing gradient problem. The two architectures are pre-trained with the ImageNet weights to take advantage of transfer learning and to speed up learning for small medical datasets. The extracted features produced by the two networks are then fused together using feature concatenation or weighted fusion methods to obtain a more complete representation of dermatological features. The model's ability to recognize subtle patterns in the lesions, color changes, and structural abnormalities captured through this integrated feature space allows it to distinguish between various types of patterns and detect even the tiniest differences that might be missed by a single architecture. Thus, the ensemble feature extraction process provides better diagnostic accuracy and the robustness of the classification.

4.3. Model Analysis and Optimization

The feature vectors from the ensemble networks are then combined and sent to the fully connected layers that perform multiclass disease classification. Softmax activation is used at the output layer to produce the probability distributions over all the disease categories. To improve training stability and prevent overfitting, a number of optimization methods are used during model training, such as learning rate scheduling, dropout regularization, batch normalization, etc. Adam or RMSProp are some of the optimization algorithms used to facilitate efficient gradient updates and speed up convergence. Validation accuracy, training and validation loss curves, precision, recall, F1-score, and confusion matrix analysis are used to continuously monitor model performance. Optimization of learning rate, batch size, dropout rate and number of training epochs is done using Hyperparameter tuning procedures. Furthermore, cross validation methods are employed to evaluate the consistency and reliability of the model by testing it on various sub-sets of data. This comprehensive optimization process enhances the model's power of generalization and reduces classification error rate.

4.4. Prediction and Treatment Suggestion

After the training and optimization process, the ensemble model is deployed and it can predict the skin disease categories in unseen images. The model produces a disease label prediction and confidence values indicating how sure the model is about its prediction for every image in the input images. Based on the identified condition, a rule-based recommendation module give some initial treatment suggestions, preventive measures and direction to further medical consultation. The recommendations are based on dermatological guidelines and medical knowledge sources which are freely accessible. The aim of this module is not to provide a medical diagnosis but to help the user get a quick and first-hand idea and the right advice for healthcare. The proposed system is designed to be a holistic framework for disease diagnosis, offering automated identification and initial treatment recommendations, which can assist in early detection and support healthcare decisions

V. RESULTS AND DISCUSSION

5.1. Performance Analysis of EfficientNetB0

The EfficientNetB0 model was trained for 10 epochs to test its performance in the automated skin disease classification task. In the training stage, the model obtained a training accuracy of 83.39% and a validation accuracy of 62.94%, which means that there is a need for more epochs to train the model to gain discriminatory knowledge about the lesion features. The model converged quickly and showed considerable performance boost as training continued. In the second epoch, the accuracy rose to 93.36%, showing the effectiveness of transfer learning and image preprocessing to speed up feature extraction.

It was seen that with every epoch, the amount of loss during training is reduced, reducing from 0.5082 in the 1st epoch to 0.0396 in the last epoch. At the same time, the accuracy of the training procedure also improved continuously and ultimately showed 98.85%, which proves that the network parameters are optimized. The highest validation accuracy (98.25%) was obtained at the 9th epoch and the lowest validation loss (0.0640) was obtained at the 8th epoch. The outcome suggests that EfficientNetB0 was able to acquire meaningful dermatological patterns and retain a high level of generalization ability.

The classification report also confirms the success of the model. On the test data set, 96% overall accuracy was obtained. The weighted average precision, recall and F1 score were obtained as 0.96, 0.96 and 0.95, respectively. The classification results were excellent with some classes having a perfect precision and recall. The relatively lower recall (0.69) however for Class 4 indicates that there is overlap in the visual characteristics of certain lesion categories, and these are difficult to distinguish. However, the overall results show that the classification of skin diseases using EfficientNetB0 is reliable and accurate.

Observation from EfficientNetB0 Accuracy

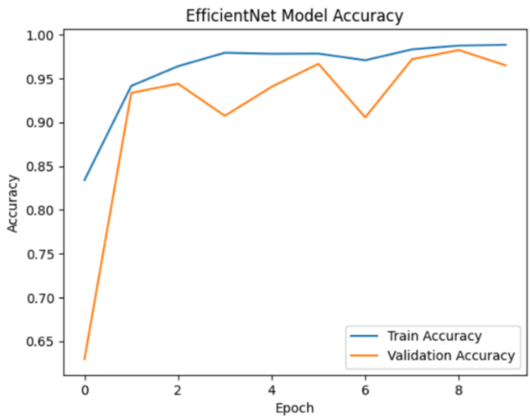

Fig.2 - Observation from EfficientNetB0 Accuracy Graph

The training accuracy curve shows a gradual rise over time, suggesting that accuracy is being steadily improved as the training progresses, with features being refined and learned continuously. The curve for the accuracy on the validation data set is very similar to the curve for the training data set, indicating good generalization and little overfitting. The temporary fluctuations that are seen in epochs 4 and 7 is probably related to changes in lesion complexity and class distribution in the validation samples. However, the model training accuracy remains stable at more than 90% even with these variations, showing stable learning behavior.

Observation from EfficientNetB0 Loss

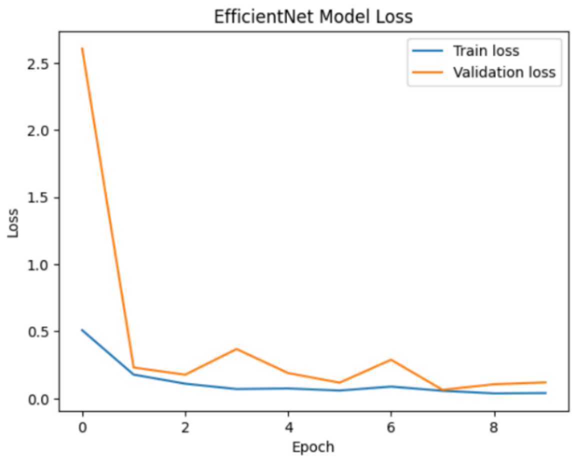

Fig.3 - Observation from EfficientNetB0 Loss Graph

The graph of the training loss shows a steady decline, which is a sign of successful optimization and error reduction in the classification task. The loss value on the validation set drops rapidly in the first few epochs, and stays fairly flat during training. The minor variations in the validation loss are temporary fluctuations in the accuracy of the validation set, but the general trend suggests that there is convergence. The difference between training loss and validation loss is relatively small, suggesting that the model is not overfitting to the training data.

5.2. Performance Analysis of ResNet50

The ResNet50 was trained on the same data set and preprocessing steps to make a comparison of the effectiveness of residual learning for skin disease prediction. The model had a lower learning rate, with training accuracy of 69.57% and validation accuracy of 55.07%, during the first epoch compared to the learning rate of EfficientNetB0. There is a very high validation loss at the start of the epochs, which indicates that more epochs were needed for the model to converge feature representations.

The accuracy of classification increased significantly as training went on. At the end of the epoch, the accuracy of the training was 96.50% and validation was 90.03%. The training loss was reduced from 0.9936 to 0.1116, which shows the network is well optimized. Compared to EfficientNetB0, ResNet50 showed larger variations in the validation loss, characterised by periods of extreme loss at epochs 8 and 9. These differences suggest that there is some sensitivity to certain lesion categories, and that mild overfitting and/or feature learning instability may have occurred during certain parts of the training process.

The classification report shows that the performance of the test is 89% overall accuracy. The weighted precision, recall and f1 values were 0.90, 0.89, and 0.88, respectively. The model was successful for most disease categories, but not for Class 4, which had a recall value of only 0.47, suggesting a lack of success in the correct recognition of this disease type. The total averaged F1 score of 0.85 also indicates non-uniform performance of the classifiers across the classes.

Observation from ResNet50 Accuracy

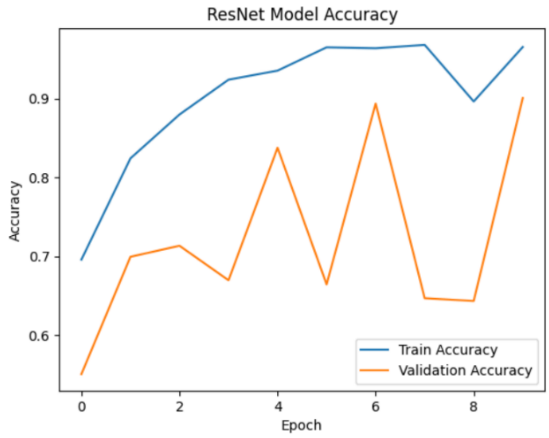

Fig.4 - Observation from ResNet50 Accuracy Graph

The training accuracy graph shows a steady rise in accuracy as the number of epochs increases, suggesting that the model is learning the disease-specific features. The validation accuracy curve, however, shows significant fluctuations in the accuracy of the validation during the training. The reduction in the number of classifications during epochs 6, 8, and 9 are significant, indicating instability with complex lesion patterns. The final validation accuracy is above 90%, but there is a lot of variance during training, suggesting that the generalization is less than that of EfficientNetB0.

Observation from ResNet50 Loss

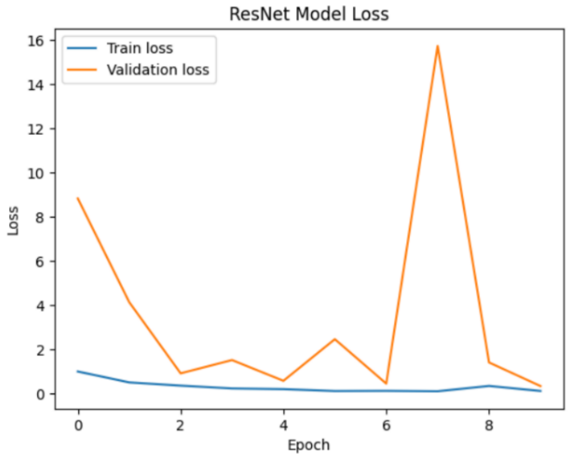

Fig.5 - Observation from ResNet50 Loss Graph

The loss during training keeps decreasing over the course of learning, thus showing effective optimization. By contrast, the loss on validation shows significant variability, particularly in the form of very high peaks during some epochs. The spikes are evidence of occasional lack of representativeness of the model for unseen samples. There is an even more pronounced difference between the training loss and the validation loss in the case of EfficientNetB0 than with the larger model, indicating a higher risk of overfitting.

5.3. Performance Analysis of Hybrid EfficientNetB0–ResNet50 Ensemble Model

This section presents the performance analysis of the Hybrid EfficientNetB0–ResNet50 Ensemble Model. A hybrid ensemble architecture was proposed to take advantage of EfficientNetB0's and ResNet50's complementary strengths. Ensemble model: The model uses the efficient feature extraction mechanism of EfficientNet along with the deep residual learning mechanism of ResNet. The goal was to get more detailed representations of the features and increase the robustness of the classification in different categories of skin diseases.

Overall, the ensemble model performed very well throughout the training process. The accuracy of the model during training rose from 82.63% at the first epoch to 99.36% at the twentieth epoch. Likewise, the accuracy of validation has significantly increased and has reached the maximum of 98.25% in the last epoch. The loss values for training kept decreasing across the epochs, ranging from 0.5822 to 0.0164, and the loss value for the validation reached a minimum value of 0.0571, which shows that the optimization and convergence were quite effective.

The classification report is in line with the superior performance of the ensemble architecture. The model's test accuracy was 98%, a higher accuracy than EfficientNetB0 and ResNet50. All the four weighted precision, recall and F1 score values were around 0.98, suggesting an even distribution of classification performance across all the disease categories. The precision and recall for most classes were above 95% and a recall of 92% was obtained for Class 4 even though the class was challenging for the individual models. This result illustrates how feature fusion resulted in the increased performance of the model in discriminating between similar skin lesions in visual space.

Observation from Hybrid Model Accuracy

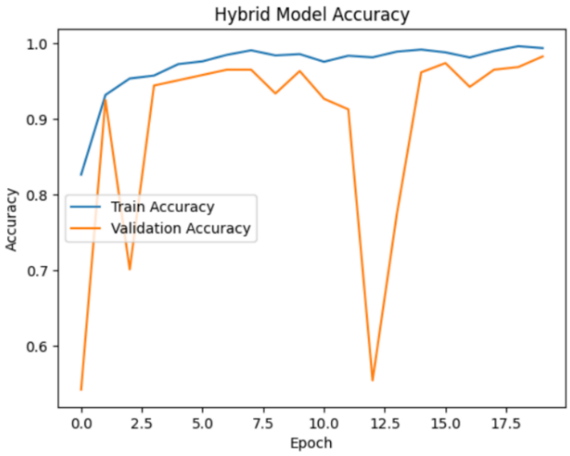

Fig.6 - Observation from Hybrid Model Accuracy Graph

The training accuracy curve shows a gradual rise during training which ultimately reaches a high level (close to 100%). The trend of validation accuracy is similar to training accuracy and is always above 92% since the second epoch. The model initially has some fluctuations in the epochs 13 and 14 but recovers almost immediately and gets the best validation accuracy in the last epochs. This behavior reflects high learning stability and a better generalization ability.

Observation from Hybrid Model Loss

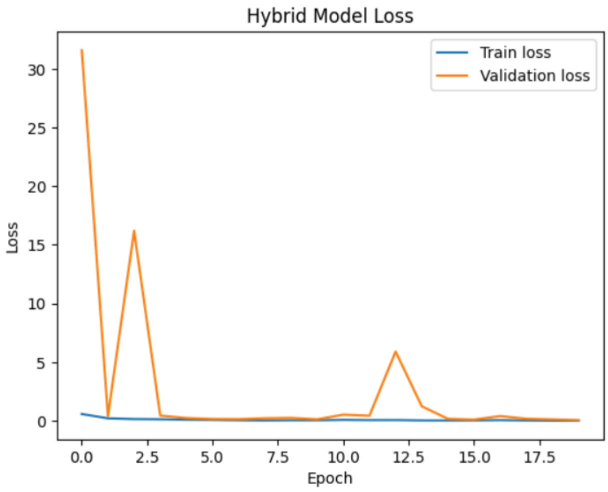

Fig.7 - Observation from Hybrid Model Loss Graph

The training loss is reduced gradually with each epoch, indicating successful parameter optimization. The loss on validation set first fluctuates as a result of the complexity of stacking two deep architectures but later it seems to be settled and is at its minimum value in the last epoch. Both the curves meet at the point that the convergence of both curves shows that the ensemble model is able to reduce the classification error and has a good performance on unseen data.

5.4. Comparative Analysis of All Models

The effectiveness of the proposed ensemble approach can easily be seen by comparing the three models. We got 96% accuracy for EfficientNetB0 and 89% for ResNet50 for the overall. The hybrid model EfficientNetB0-ResNet50 gave the highest accuracy of 98% compared to the individual models.

EfficientNetB0 was found to be more stable and converged better than ResNet50. While ResNet50 achieved good deep hierarchical features extraction performance, it suffered from invalidation instability and high sensitivity to class imbalance problem. The ensemble model gave better accuracy, recall and F1 scores and the misclassification rate decreased significantly, combining the best of both architectures.

The results from the experiments confirm the advantages of ensemble learning, in terms of feature diversity, generalization ability and reliability of the classification. The proposed hybrid architecture is appropriate to fully capture both fine-grained texture information and deep semantic representation, which is particularly suitable for skin disease automatic prediction system.

VI. FUTURE WORK

Future research can focus on expanding the dataset to include additional skin disease categories, rare dermatological conditions, and images representing diverse skin tones, age groups, and geographic populations. Increasing dataset diversity will enhance model generalization capabilities and reduce potential prediction bias. Furthermore, incorporating real-time clinical images captured through smartphones and portable imaging devices can improve the practical applicability of the system in real-world healthcare environments. Another promising direction involves integrating Explainable Artificial Intelligence (XAI) techniques into the diagnostic framework. Methods such as Grad-CAM, LIME, and saliency map visualization can highlight the lesion regions that contribute most significantly to model predictions. Such interpretability mechanisms can assist dermatologists in validating automated decisions and increase trust in AI-assisted diagnostic systems. Additionally, deploying the model as a mobile or web-based healthcare application would significantly improve accessibility, particularly in rural and underserved regions where dermatological expertise may be limited. Edge computing technologies could further support offline inference and resource-efficient deployment.

Future studies may also explore multimodal learning frameworks that combine image-based features with patient-specific information such as medical history, symptoms, demographic attributes, and clinical observations. Integrating visual and non-visual information has the potential to improve diagnostic accuracy and enable more personalized treatment recommendations. The inclusion of clinical metadata may further reduce false predictions and enhance the overall effectiveness of intelligent skin disease diagnosis systems.

VII. CONCLUSION

This research presented an intelligent skin disease prediction framework based on an ensemble deep learning architecture that integrates EfficientNetB0 and ResNet50 for automated dermatological image classification. The primary objective of the study was to improve diagnostic accuracy and classification reliability by combining the complementary strengths of two state-of-the-art convolutional neural network architectures. EfficientNetB0 contributed efficient feature scaling and detailed texture extraction, while ResNet50 provided deeper hierarchical feature learning through residual connections. The fusion of these architectures enabled the proposed system to capture a broader range of lesion characteristics and improve disease discrimination capability.

Experimental evaluation demonstrated that EfficientNetB0 individually achieved a classification accuracy of 96%, whereas ResNet50 achieved an accuracy of 89%. Although both architectures produced satisfactory results, EfficientNetB0 exhibited superior convergence behavior and stronger generalization performance. ResNet50 successfully learned complex lesion representations but experienced greater fluctuations in validation accuracy and loss, indicating reduced stability during training. These findings suggest that relying on a single architecture may limit classification performance when dealing with highly diverse dermatological datasets.

The proposed hybrid EfficientNetB0–ResNet50 ensemble model significantly outperformed the individual architectures. The ensemble achieved an overall classification accuracy of 98%, along with weighted precision, recall, and F1-score values of approximately 98%. The classification report demonstrated balanced performance across nearly all disease categories, including those that posed challenges for the standalone models. Furthermore, the ensemble exhibited lower classification error, improved recall for minority classes, and enhanced robustness against visually similar lesion patterns. These improvements confirm the effectiveness of feature fusion and ensemble learning in medical image classification applications.

Analysis of training and validation curves revealed that the ensemble model maintained strong learning stability and convergence characteristics throughout the training process. Although temporary fluctuations were observed during intermediate epochs, the model consistently recovered and achieved superior validation performance. The reduction in validation loss and improvement in classification metrics indicate that the proposed architecture successfully generalized to unseen samples while minimizing overfitting.

Overall, the findings demonstrate that combining EfficientNetB0 and ResNet50 creates a powerful and reliable diagnostic framework for automated skin disease prediction. The proposed system has the potential to support dermatologists by providing accurate, consistent, and rapid diagnostic assistance. In addition to improving classification performance, the framework contributes to the advancement of artificial intelligence in healthcare by offering a scalable solution for computer-aided dermatological diagnosis. Future enhancements involving larger datasets, multimodal learning, explainable artificial intelligence techniques, and real-time clinical deployment may further improve the effectiveness and practical applicability of the proposed system in healthcare environments

VIII. REFERENCES

[1] Chen, K., Zhuang, D., & Chang, J. M., "SuperCon: Supervised Contrastive Learning for Imbalanced Skin Lesion Classification," arXiv Preprint, arXiv, 2022.
[2] Pakzad, A., Kumar, A., &Hamarneh, G., "CIRCLe: Color Invariant Representation Learning for Unbiased Classification of Skin Lesions," ISIC Workshop Proceedings / arXiv Preprint, ISIC Foundation, 2022.
[3] Alam, T. M., et al., "An Efficient Deep Learning-Based Skin Cancer Classifier," PubMed Central (PMC), National Center for Biotechnology Information (NCBI), 2022.
[4] Alwakid, G., et al., "Melanoma Detection Using Deep Learning-Based Classifications," Healthcare Journal, MDPI Publishing, 2022.
[5] Anand, V., et al., "Multi-class Skin Disease Classification Using Transfer Learning," International Journal of Pattern Recognition and Artificial Intelligence, World Scientific Publishing, 2022.
[6] Shetty, B., et al., "Skin Lesion Classification of Dermoscopic Images Using Deep Convolutional Networks," Scientific Reports, Springer Nature, Vol. 12, 2022.
[7] Wang, Y., Feng, Y., Zhen, L., et al., "Adversarial Multimodal Fusion with Attention Mechanism for Skin Lesion Classification Using Clinical and Dermoscopic Images," Medical Image Analysis, Elsevier, Vol. 78, 2022.
[8] Du, S., et al., "FairDisCo: Fairer AI in Dermatology via Disentanglement and Contrastive Learning," arXiv Preprint, arXiv, 2022.
[9] Wu, Y., et al., "Federated Contrastive Learning for Dermatological Applications," arXiv Preprint, arXiv, 2022.
[10] Shetty, B., & Fernandes, R., "Skin Lesion Classification of Dermoscopic Images Using Machine Learning and Convolutional Neural Networks," Nature Portfolio Research Collection, Springer Nature, 2022.
[11] Mohammed, A. S., & Mohammed Ali, M. S., "Contrastive Self-Supervised Ensemble Transfer Learning for Robust Skin Cancer Classification and Early Detection," PubMed Central (PMC), National Center for Biotechnology Information (NCBI), 2022.
[12] Debelee, T. G., "Skin Lesion Classification and Detection Using Machine Learning Techniques: A Systematic Review," Applied Sciences, MDPI Publishing, 2022–2023.