Empirical Study on the working of different Deep Learning Models on Skin Disease Classification on Dermoscopic Images

Uday Tripathi
Student, School of Computer Science and Engineering
Lovely Professional University)
Phagwara, India
Email – uday007tripathi@gmail.com Orcid id- 0009-0001-6572-2691

***Abstract*—Skin disorders are a major global public health concern that require prompt and precise diagnosis in order to effectively treat. In this work, I Investigate the automated classification of skin diseases using dermoscopic pictures and deep learning algorithms. To create reliable classification models, I make use of state-of-the-art architectures like ResNet, DenseNet, and MobileNet in conjunction with convolutional neural networks (CNNs). Preprocessing the picture data, training the models on a carefully selected dataset, and assessing the models' performance using accuracy metrics comprise our methodology.Deep learning has the potential to accurately classify skin conditions, which could benefit dermatologists and enhance patient outcomes. Our work highlights the significance of thorough assessment of deep learning models in clinical applications and advances the field of computer-aided diagnosis in dermatology.**

***Index Terms*—CNN, DenseNet, RestNet, MobileNet, DeepLearning**

I. Introduction

Millions of people worldwide are impacted by skin diseases, which frequently provide diagnostic hurdles for medical experts. For these diseases to be effectively treated and patients to get quality care, an accurate diagnosis must be made quickly. Skin diseases can be contagious (like scabies and lice), non-contagious (like medication allergies and rosacea), chronic (like psoriasis and atopic eczema), rare (like sweet syndrome and ofuji disease), or none at all. The prevailing perspective of society downplays the significance of skin diseases when they manifest and favors skipping medical appointments. Global statistics [4] indicate that 1.79 percent of human physical disability worldwide are caused by skin disorders. In different parts of the world, between 30 percent and 70 percent of people suffer from skin disorders [9]. Deep learning techniques have been applied in dermatology and the results have been promising in automating the classification of skin conditions based on clinical photos or dermoscopic images. Convolutional Neural Networks (CNNs) and their variants, such as ResNet, DenseNet and MobileNet, are among the deep learning models that have shown impressive performance in extracting intricate patterns and characteristics from vast skin image datasets. These models offer a potential remedy for the difficulties involved with manual diagnosis by efficiently identifying and categorizing a variety of skin disorders. Beyond technical advancements, precise deep learning classification of skin diseases is crucial. Because it makes early identification, objective diagnosis and effective triage of skin problems possible. it has the potential to revolutionize the practice of dermatology. Furthermore, computerized classification of skin diseases can help with ongoing dermatological research and learning as well as remote diagnosis in underserved areas. In this empirical study, I compare the effectiveness of several deep learning architectures—CNN, ResNet, DenseNet, and MobileNet—for the categorization of skin diseases. I assess their efficacy using important measures like accuracy and loss in an effort to determine which model is best suited for an automated, trustworthy diagnosis of skin conditions. The findings of this analysis are crucial for developing the use of deep learning in dermatology, which will eventually lead to better patient outcomes, more accurate diagnosis, and more efficient healthcare delivery in the area of skin diseases management.

II. Related Works

Deep learning algorithms have the ability to solve many problems in medical science. [1] trie to use different CNN algos to classify face skin diseases. Deep learning algorithms have been used in numerous research to treat skin conditions [3] [4] [5] [6]. For instance, the Inception-v3 network can now categorize skin malignancies with performance comparable to that of professional dermatologists for nine kinds of tumors, the computer achieved an accuracy of 55.4 percent whereas, the accuracy of two doctors was 53.3percent and 55.0 percent [3]. [4] obtained an accuracy of 87.25±2.24percent on the dermoscopic pictures for four common skin conditions including SK, BCC, psoriasis and melanocytic nevus using the same network structure. These studies demonstrate the potential for dermatoses to benefit from the use of existing deep learning techniques. Deep learning's potential applicability to face-related diseases is also encouraging at the same time. developed a deep learning system known as DeepGestalt. They used over 17000 real facial photos of genetic abnormalities to train their model, which is now able to identify over

200 genetic syndromes with a comparatively high degree of accuracy. [8] investigated the use of CNNs to categorize acne into several severity classifications, from clear to severe. Their findings indicate that their strategy exceeded skilled medical professionals in terms of accuracy. There is also a research [2] that shows the classification of skin diseases in which they plot the frequency graph by measuring the reading on both normal and affected skin of different skin diseases.

III. Methodology

In this paper I have discussed the working of mainly four deep learning models named, CNN, Resnet, DenseNet and MobileNet. I applied these models to find the training and validation accuracy and loss values on the same train-test-val split on an image dataset I found on Kaggle.

A. Data collection and Preprocessing

The study made use of a Kaggle dataset that included dermoscopic photos of different skin conditions. Preprocessing was done on the dataset to guarantee consistency and quality. To improve model generalization, images were enhanced, normalized, and scaled to a consistent resolution. Techniques for augmenting data included flipping, rotating, zooming, and adjusting brightness.

B. Model Development

Four Deep Learning architectures were implemented and evaluated for skin disease classification:

1) Convolutional Neural Network (CNN): A baseline CNN model consisting of multiple convolutional layers followed by fully connected layers [10] .

2) ResNet (Residual Network): A deeper architecture utilizing residual connections to facilitate training of deeper networks [12] .

3) DenseNet (Densely Connected Convolutional Network): A network architecture that maximizes information flow between layers by connecting each layer to every other layer in a feed-forward fashio [11].

4) MobileNet: A lightweight CNN architecture optimized for mobile and embedded devices, utilizing depth-wise separable convolutions [13] .

C. Traning and Evaluation

The preprocessed dataset with a split between training and validation was used to train each model. Using the Adam optimizer, the training method involved minimizing the loss due to category cross-entropy. Metrics like accuracy, precision, recall, and F1-score were used to track the model's performance throughout training.

D. Performance Comparision

A hold-out test set was used to evaluate the trained models' performance in terms of accuracy, loss and other pertinent metrics. To determine the best architecture for classifying skin diseases, the data were examined and contrasted.

IV. Results

Each deep learning architecture (CNN, ResNet, DenseNet, MobileNet) was trained and evaluated on the same split of the dataset on skin disease classification task. The following summarizes the key findings from the experimental results:

1) CNN: The baseline CNN achieved a validation accuracy of approximately 56.4percent after 10 epochs of training. The accuracy steadily increased throughout training, indicating potential for further improvement with longer training or model adjustments. The following figures show you the training accuracy and loss of CNN

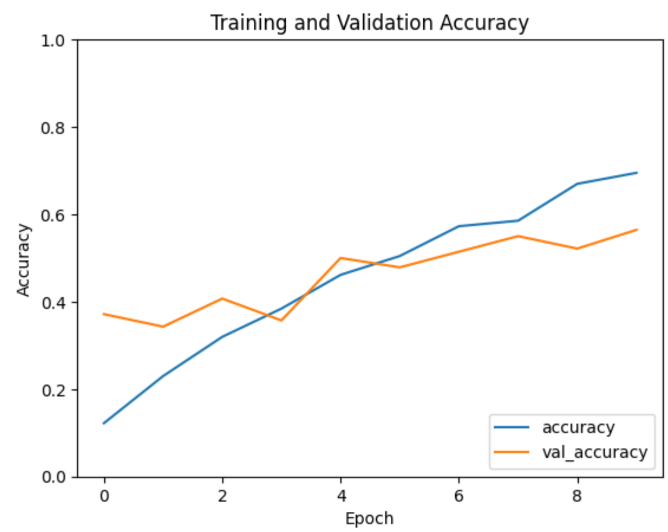

Fig. 1. Training and Validation Accuracy (CNN)

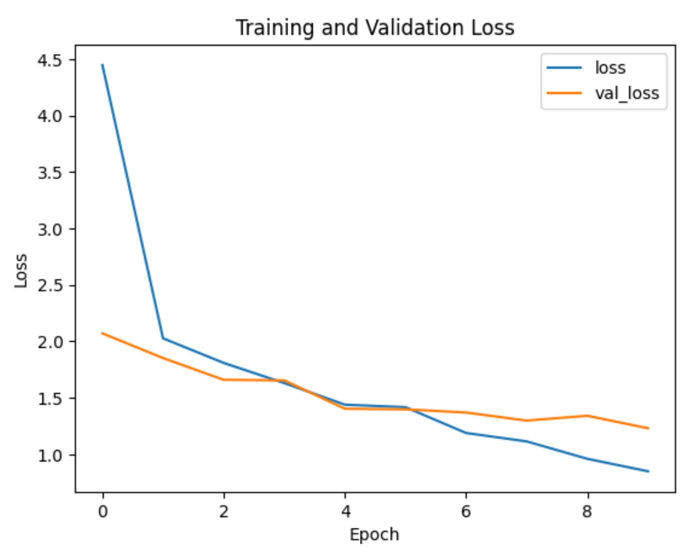

Fig. 2. Training and Validation Loss (CNN)

2) ResNet: ResNet performed poorly with a validation accuracy of around 7.1persent after 10 epochs. This may suggest challenges in training deeper architectures or issues with model configuration for the given task. The following figures show you the training accuracy and loss of ResNet.

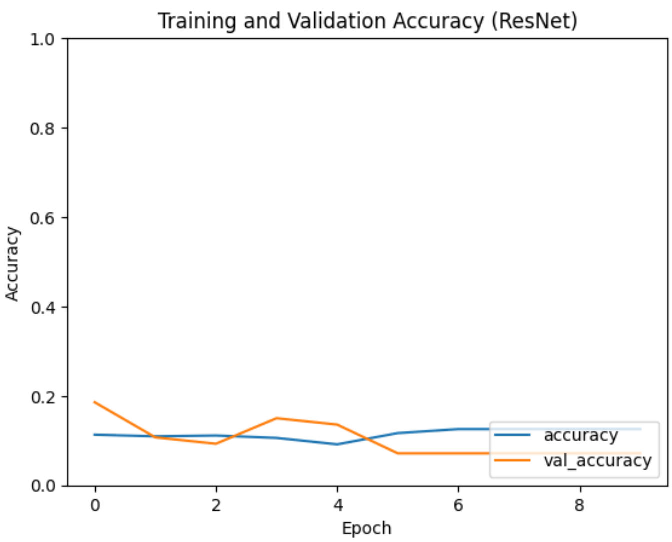

Fig. 3. Training and Validation Accuracy (ResNet)

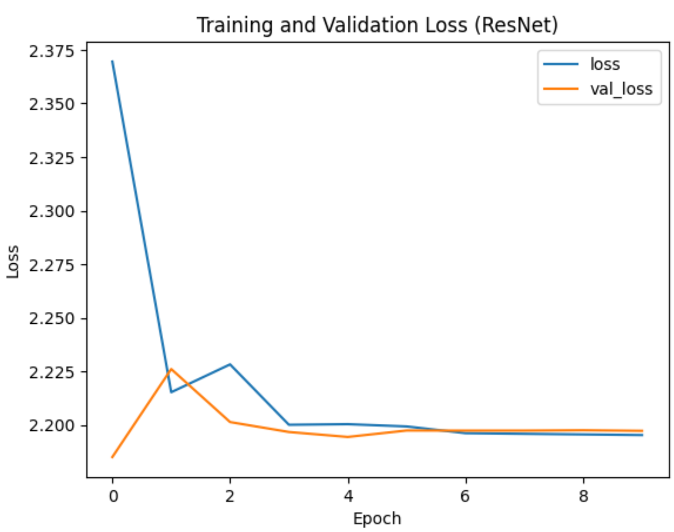

Fig. 4. Training and Validation Loss (ResNet)

3) DenseNet: DenseNet showed promising results with a validation accuracy of about 66.4percent after 10 epochs. The accuracy consistently improved over epochs, indicating effective information propagation across densely connected layers. The following figures show you the training accuracy and loss of DenseNet.

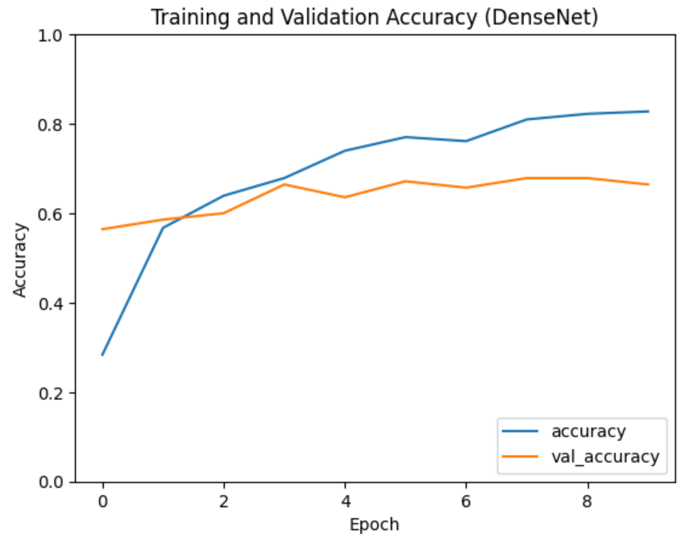

Fig. 5. Training and Validation Accuracy (DenseNet)

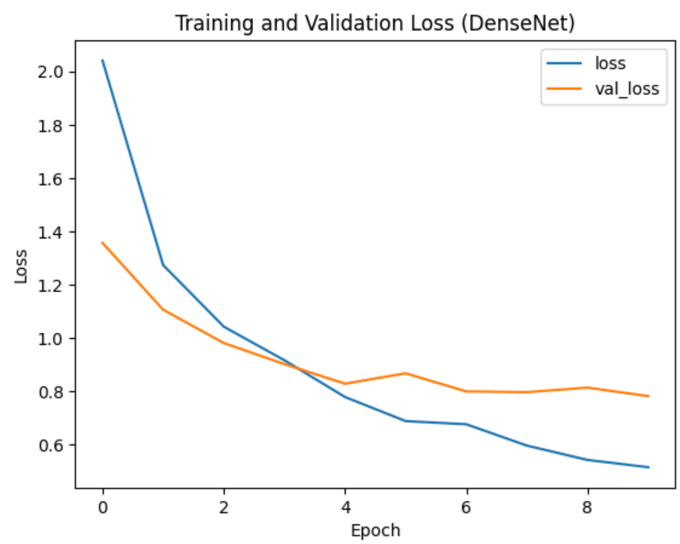

Fig. 6. Training and Validation Loss (DenseNet)

4) MobileNet: MobileNet achieved a validation accuracy of approximately 44.2percent after 10 epochs. The model's lightweight architecture may have contributed to faster training but with slightly lower accuracy compared to DenseNet. The following figures show you the training accuracy and loss of MobileNet.

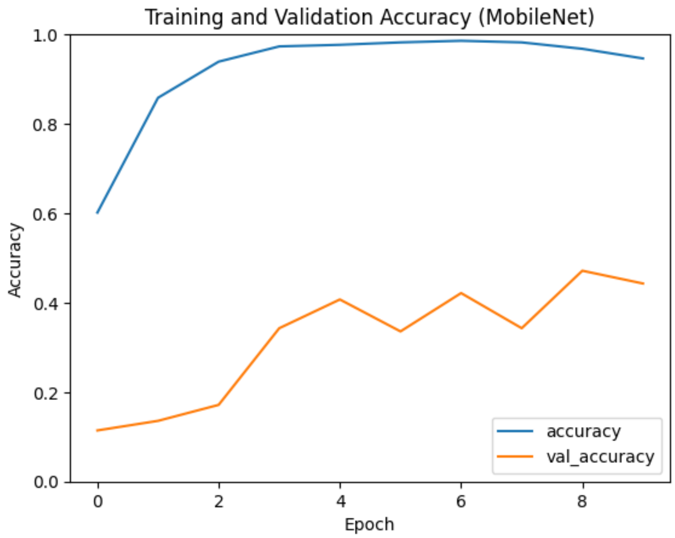

Fig. 7. Training and Validation Accuracy (MobileNet)

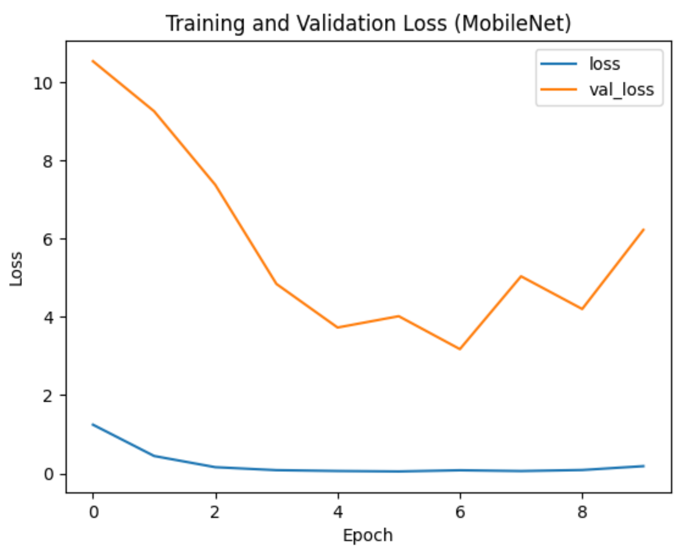

Fig. 8. Training and Validation Loss (MobileNet)

V. Comparision and Analysis

On the basis of validation accuracy, loss, and convergence behavior, each model's performance was compared. DenseNet surpassed CNN and MobileNet as the best-performing architecture in terms of accuracy. ResNet performed the worst, maybe as a result of difficulties with deep network training or problems with model complexity.

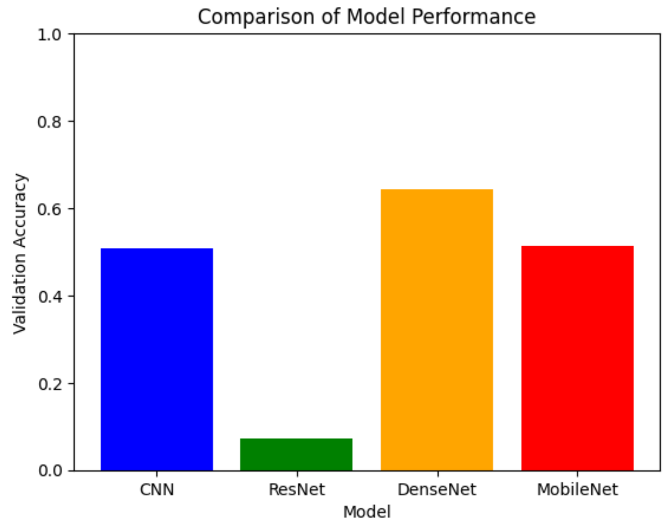

Fig. 9. Performance comparison of the models

I provided the confusion matrix to evaluate the performance and robustness of our skin disease classification system using data from the top-performing deep learning architecture, DenseNet. An essential technique for assessing a classification model's predicted performance is the confusion matrix. It offers a thorough analysis of the model's capacity to accurately categorize various skin disease classifications and highlights any possible misclassification hotspots. Class labels are represented by the rows of the confusion matrix, and the model's anticipated class labels are represented by the columns. We may learn more about the advantages and disadvantages of the DenseNet model's ability to differentiate between different categories of skin diseases by examining this matrix. Now let's examine the confusion matrix produced by our trained DenseNet model, which demonstrated a high degree of accuracy in our tests.

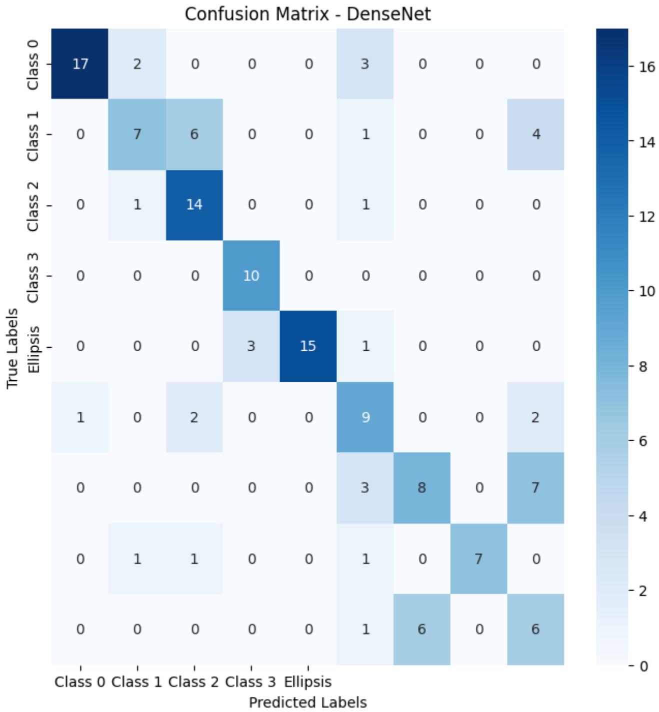

Fig. 10. Confusion Matrix of DenseNet

VI. Challenges and Limitations

The study encountered several challenges and limitations during experimentation. These include class imbalance in the dataset, limited computational resources for training complex models like ResNet, and potential overfitting issues with deeper architectures.When debating the difficulties and constraints associated with using deep learning to the classification of skin diseases, it's critical to take into account a number of variables that may affect how useful and feasible it will be to use these models in actual situations. An explanation of the difficulties and constraints is provided below:

A. Data Quality and Quantity

Acquiring broad and high-quality datasets that accurately depict different skin disorders is one of the main issues. Models developed from little or biased datasets may not be sufficiently resilient to make a good generalization to new data.

B. Class Imbalance

Skin disease datasets frequently exhibit class imbalance, with some diseases having an excess of representation and others having an underepresentaion. Biased models that perform well on majority classes but badly on minority classes may result from this imbalance.

C. Interpretability of Deep Learning Models

Deep learning models especially complex architectures like ResNet, DenseNet and MobileNet. Are often considered black-box models making it challenging to interpret how they arrive at their predictions. This lack of interpretability can be a barrier in medical applications where transparency and trust are crucial.

D. Generalization Across Diverse Populations

The manifestation of skin diseases can vary among different demographic groups (such as age and ethnicity) thereby influencing the generalizability of deep learning models that are trained on particular populations.

E. Computational Resources

Deep learning model training demands a lot of processing power, particularly for large architectures like ResNet and DenseNet (e.g., GPU accelerators). Insufficient availability of those resources may impede the advancement and implementation of sophisticated models.

F. Overfitting and Model Tuning

Overfitting is a common problem with deep learning models, especially when they are trained on limited datasets. Although they can be time-consuming and demand knowledge proper regularization techniques and hyperparameter adjustment are necessary.

VII. Future Direction

As we look to the future, there are a number of developments and paths that could improve the effect and usefulness of deep learning-based skin disease classification in medical practice. Here are a few possible avenues for future research:

- Integration of Multi-Modal Data: Investigate the merging of many modalities of data sources like genetic data, clinical notes and dermoscopic pictures. This all-encompassing strategy can enhance model performance and offer more thorough diagnostic information.
- Transfer Learning and Model Generalization: Examine transfer learning strategies that allow pre-trained deep learning models to be applied to new datasets related to skin diseases, enhancing model generalization and minimizing the requirement for sizable annotated datasets.
- Real-Time Deployment and Edge Computing: Put your attention on deep learning model optimization for edge device real-time inference which will enable telemedicine and point-of-care applications. This calls for effective deployment techniques, quantization and model compression.
- Collaboration with Dermatology Experts: To jointly create solutions that tackle clinically pertinent problems and give priority to patient outcomes, encourage multidisciplinary cooperation between dermatologists, data scientists and AI researchers.
- Global Health and Accessible Healthcare: Encourage the creation of scalable and affordable deep learning solutions to tackle the difficulties associated with diagnosing skin diseases in underprivileged areas hence promoting globally accessible healthcare solutions.

VIII. Conclusion

In this work I looked into how well different deep learning models classified skin conditions using dermoscopic picture data. I was able to get important information on how well Convolutional Neural Networks (CNN), ResNet, DenseNet, and MobileNet performed in this difficult task through empirical analysis and evaluation. I found that DenseNet is the best in classifying the skin condition amongst all.My study uses dermoscopic image data to evaluate and compare deep learning models, which advances the field of computer-aided diagnosis for skin diseases. The results highlight how deep learning can help dermatologists and other medical professionals classify diseases accurately and quickly which can lead to better patient care and diagnostic results. Also, our study tells about the limitations and challenges faced and also talk about some solution on how we can face the challenges and work beyond our limitations. All in all our research helps in the advancement of medical science and on how deep learning architectures can help improve diagnosis and healthcare as a whole.

References

[1] Z. Wu et al., "Studies on Different CNN Algorithms for Face Skin Disease Classification Based on Clinical Images," in IEEE Access, vol. 7, pp. 66505-66511, 2019, doi: 10.1109/ACCESS.2019.2918221.

[2] P. B. Manoorkar, D. K. Kamat and P. M. Patil, "Analysis and classification of human skin diseases," 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), Pune, India, 2016, pp. 1067-1071, doi: 10.1109/ICACDOT.2016.7877750.

[3] Esteva, A., Kuprel, B., Novoa, R. et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115–118 (2017). https://doi.org/10.1038/nature21056.

[4] Zhang X, Wang S, Liu J, Tao C. Towards improving diagnosis of skin diseases by combining deep neural network and human knowledge. BMC Med Inform Decis Mak. 2018 Jul 23;18(Suppl 2):59. doi: 10.1186/s12911-018-0631-9.

[5] Masood A, Al-Jumaily AA. Computer aided diagnostic support system for skin cancer: a review of techniques and algorithms. Int J Biomed Imaging. 2013;2013:323268. doi: 10.1155/2013/323268.

[6] Esteva A, Kuprel B, Novoa R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639):115-118.

[7] Gurovich, Y., Hanani, Y., Bar, O. et al. Identifying facial phenotypes of genetic disorders using deep learning. Nat Med 25, 60–64 (2019). https://doi.org/10.1038/s41591-018-0279-0

[8] Melina A, Dinh NN, Tafuri B, Schipani G, Nisticò S, Cosentino C, Amato F, Thiboutot D, Cherubini A. Artificial Intelligence for the Objective Evaluation of Acne Investigator Global Assessment. J Drugs Dermatol. 2018 Sep;17(9) 1006-1009. PMID: 30235389.

[9] J. Yang, X. Sun, J. Liang and P. L. Rosin, "Clinical Skin Lesion Diagnosis Using Representations Inspired by Dermatologist Criteria," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 1258-1266, doi: 10.1109/CVPR.2018.00137.

[10] Alzubaidi, L., Zhang, J., Humaidi, A.J. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8, 53 (2021). https://doi.org/10.1186/s40537-021-00444-8

[11] Gao Huang and Zhuang Liu and Kilian Q. Weinberger, "Densely Connected Convolutional Networks", CoRR, abs/1608.06993, 2016, http://arxiv.org/abs/1608.06993

[12] Kaiming He and Xiangyu Zhang and Shaoqing Ren and Jian Sun, "Deep Residual Learning for Image Recognition", CoRR, abs/1512.03385, 2015, http://arxiv.org/abs/1512.03385.

[13] Howard, Andrew and Zhu, Menglong and Chen, Bo and Kalenichenko, Dmitry and Wang, Weijun and Weyand, Tobias and Andreetto, Marco and Adam, Hartwig. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.

[14] Saja Salim Mohammed and Jamal Mustafa Al-Tuwaijari 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1076 012045DOI 10.1088/1757-899X/1076/1/012045.

[15] Nazia Hameed, Antesar M. Shabut, Miltu K. Ghosh, M.A. Hossain,Multi-class multi-level classification algorithm for skin lesions classification using machine learning techniques,Expert Systems with Applications,Volume 141,2020,112961,ISSN 0957-4174, doi: 10.1016/j.eswa.2019.112961.