Efficient Multi-Class Classification Of Skin Diseases Using Convolution Neural Network and Convolution Block Attention Module

K. Kumaran
Assistant Professor
Department of Information Technology
Easwari Engineering College
Chennai, India
kumaran.me.cse@gmail.com

Rachana A
Department of Information Technology
Easwari Engineering College
Chennai, India
rachana.aravindan@gmail.com

Priyadharshini T K
Department of Information Technology
Easwari Engineering College
Chennai, India
pd021tk@gmail.com

Raghul S
Department of Information Technology
Easwari Engineering College
Chennai, India
raghulselvaraj01@gmail.com

Abstract— Skin diseases are a common eventuality amongst humans. Diseases such as Melanoma has dangerous impacts and have high potential to infect other parts of the body. Due to the differences in their skin texture, presence of skin hair, and color, diagnosis is very difficult. To improve accuracy of diagnosis for various skin conditions, techniques like machine learning must be developed. The Medical Decision Support System will assist in accurately detecting the various kinds of skin diseases using a range of images that have been trained on datasets. Because manual diagnosis can be labor-intensive and time-consuming, the system uses Convolutional Neural Networks, with a variety of models, and feeds the most accurate model into the Convolution Block Attention Model (CBAM). This technique's CBAM module makes the system operate to provide high accuracy with effective computing. By deploying a web application, the technology also makes it possible for physicians all over the world to utilize for analysis.

Keywords—Skin Diseases, Medical Decision Support System, CNN, CBAM, Machine Learning with Deep Learning, Squeeze and Excitation Network, Datasets, Kaggle API, HAM10000

I. Introduction

The skin is a complex active organ which, when fails, can result in serious consequences. Skin conditions that are prone to spread are dangerous while begin diagnosed or under treatment. Early automated identification and analysis of the various skin diseases provide challenges as they are uncommon with their symptoms, and early identification. Determining the rate of disease growth can be extremely difficult, even for highly specialized clinics. which can cause delay in treatment and therapy. Convolutional Neural Networks (CNN) have been one of the revolutionizing Deep Learning Networks that are found used in pharmaceutics, life science and engineering. In traditional networks, using a series of automatic coding stages, the size of their available training sets have limited their ability to perform in the applications, along with the densitivity of the neural network structure. Comparatively to other conventional models, the model we present in this paper with newly added classifier layers performs better at image segmentation and classification with less training data. This is because it is based on various neural networks. The accuracy and quantitative effectiveness of the network can be enhanced through the training of the network with various models. This will ensure that the highest level of accuracy is obtained in categorizing the different classes of skin.

A. Problem Description

The skin is an active, complex organ that, when it fails, can have serious consequences. Skin conditions that are prone to spreading can have serious complications at the early stages of diagnosis or treatment when they are still in their early stages of diagnosis or treatment. In addition to the fact that skin diseases are uncommon with their symptoms, identifying and analysing them early on in their development can be extremely challenging for specialized clinics, which can lead to delays in the treatment and therapy process. In order to understand the rate of disease progression, it is very challenging to identify and analyse skin diseases early in their development. To be able to obtain a high level of accuracy in its classification of skin classes, the network must be trained with a variety of models in order to achieve the highest level of accuracy in its classifications. As a result, diagnostic precision and computational efficiency are both improved as a result.

II. RELATED WORKS

Metin Akay et al., "Deep learning classification of systemic sclerosis skin using the MobileNetV2 model." doi:10.1109/OJEMB.2021.3066097. IEEE Open Journal of Engineering in Medicine and Biology 2 (2021): 104–110.

The author proposes a model using a novel mobile Deep Learning Network for Systemic Sclerosis Skin. The proposed method uses a dense-connectivity Convolution Neural Network, or UNet, that has a higher number of classifier layers than a conventional Convolution Neural Network. When combined with a small training module, it produces more accurate segmentations of images, as well as a mobile training module, which would assist with the training process in a more efficient manner. The journal made use of a very effective training module known as "MobileNetV2" which was designed for embedded and mobile application development. The current method has done better than other methods with more than 85% accuracy using the dataset. It requires very little computational effort because of its robustness in identifying the impacted region significantly faster and with almost a 2-fold reduction in computations compared to the standard model. Due to the smaller training datasets and absence of faster output processors, this model performed better than its earlier iterations. Only one disease could be classified using this methodology. Despite the fact that this model was able to provide an accuracy of 85%, it would still be possible to improve accuracy and processing speed if alternative CNN models were used, such as the CBAM model.

Wu ZH, Zhao S, Peng Y, He X, Zhao X, Huang K, Wu X, Fan W, Li F, Chen M, and Li J on various CNN algorithms for categorising facial skin diseases from clinical photos. 2019 May 22;7:66505-11.PY - 2020/04/01 T1 - IEEE Access

It is the purpose of this paper to study the various CNN models and algorithms that can be used to correctly classify facial skin diseases based on clinical dermatological images, using the various CNN models and algorithms. As a result of using Xiangya Derm, one of China's biggest datasets for clinical images of skin diseases, the author used a dataset made up of 2656 images of faces with six common skin diseases, including rosacea (ROS), actinic keratosis (AK), and other major diseases. According to the study, five widely used network algorithms were used to classify these illnesses in the dataset, and their outcomes were compared. In terms of the performance of the best model for testing the dataset that contains about 388 dermatoscopic images, it achieved about 92% accuracy as well as 77.0% and 70.8% respectively in terms of mean and recall precision. Although this model was able to provide an considerable accuracy, further implementations can be done to increase

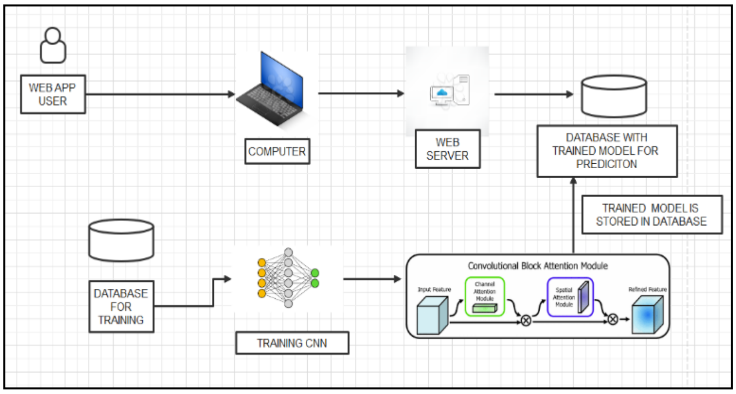
III. SYSTEM DESIGN

A. OBJECTIVE

In this project, we aim to design and implement a Medical Decision Support System to be used to diagnose seven different classes of skin diseases through images using various Convolutional Neural Networks combined with CBAM. Using an efficient convolutional network, we are able to use multiclass classification of skin diseases with high accuracy by using an efficient convolutional network. As a result of the Decision Support System, physicians will be able to use it for free and with ease.

B. PROPOSED SYSTEM

The Medical Decision Support System will support diagnosing multiple skin diseases through various images, trained by datasets to accurately diagnose the different classes to the skin diseases. As manual diagnosis can be time-consuming and requires more human effort, the system uses Convolutional Neural Network, with different models, and using the model with the highest accuracy will be fed into the Convolution Block Attention Model (CBAM). The CBAM module used in this technique makes the system work to produce a high accuracy with efficient computation. The system also enables physicians all around to be able to connect to the system for analysis, by the deployment of a web application. The datasets used for training is fed into the CNN Training Module. After processing the trained model is stored into the database for



access by the User.

Fig 1. Architecture Diagram of Proposed System

The architecture for the Proposed system is explained elaborately here. With proper connectivity, the user here is first using the Application via his phone or other electronic. The image of the dermatoscopic region that required attention, is captured through the web application and is sent to the computer for processing. The computer, in turn, will process this dermatoscopic image, and will try to find a similar image for it to compare. The database is a trained model, and it has. It has been built to differentiate the 7 different class of dermatoscopic data. On the other hand, the database is put into the CNN training model, where it encounters the CBAM and the Channel Attention Module and the Spatial Attention Module.

The various skin diseases that are being classified in the proposed structure are:

1. Melanocytic Nevi
2. Melanoma
3. Benign Keratosis-like Lesions
4. Dermatofibroma
5. Basal cell carcinoma
6. Vascular lesions
7. Actinic keratoses

Target Users: Physicians (more specifically, Dermatologists)

Technology Used: Data Analysis, Deep Learning. Cloud Storage

IV. METHODOLOGY

As a part of the Decision Support System, there is a dataset containing a substantial number of multi-source dermatoscopic images that illustrate pigmented lesions from various different sources. Using the HAM10000 (Human Against Machine with 10000 training images) dataset. With a total of 10015 dermatoscopic images were included in the final dataset, consisting of various diagnostic categories such as Dermatofibroma (df), Melanoma (mel), Actinic Keratoses and intraepithelial carcinoma / Bowen's disease (akiec), Benign Keratosis-like lesions such as solar lentigines / seborrheic keratoses and lichen-planus like keratoses (bkl), Bass Cell Carcinoma (bcc), melanocytic nevi (nv) and vascular lesions such as angiokeratomas, angiomas pyogenic granulomas and hemorrhage(vasc).Data analysis is done experimentaly to understand the preprocessing and analysis of data better. Separate dashboard will be created for the user to upload the image of the captured skin disease and will be created using ReactJS.

A. MODULES

Module 1: Data Analysis and Pre-processing.

The input data are skin disease images from HAM10000 dataset. Input images are of seven types as listed in table 01. These images are usually taken in different sizes and resolution. The images are reshaped to a fixed size. Since the dataset is imbalanced, some images of minority classes are oversampled using random oversampling technique. Random oversampling involves adding replacement examples drawn at random from the minority class to the training dataset. The images are then converted into array format as suited for input to the model. The values in the array are normalized to range between 0 to 1. The train-val-test split is 75-20-5. Both the train and validation set are oversampled.

TABLE 1. TYPES OF SKIN DISEASES IN RH CLASSIFICATION

Skin disease	Number of images
Melanocytic nevi	1205
Melanoma	1143
Benign keratosis-like lesions	1009
Basal cell carcinoma	614
Actinic keratoses	346
Vascular lesions	258
Dermatofibroma	152

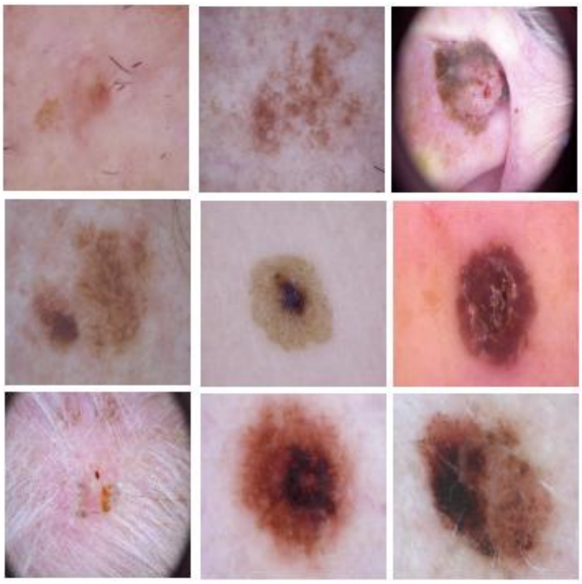


Fig 2. Images of skin diseases in input dataset.

Data augmentation techniques such as image rotation, image flipping, image zooming is implemented.

Module 2 : Base model implementation.

There are multiple Convolutional Neural Networks popularly used in the image processing field and computer vision, specifically for object recognition along with detection. CNN models that are considered as the first sequential block of the proposed model is termed as base model. Base models we used are MobileNet, VGG16, VGG19, InceptionV3, Xception, EfficientNet, DenseNet, ResNet. The proposed model is the base model is followed by an pooling layer that uses averaging pooling with pool size of 2. This is then followed by a flattening layer, two dense layers of 64 neurons with ReLU activation with a dropout layer in between. The last layer will be the output layer that consists of 7 neurons corresponding to 7 different types of skin diseases with softmax activation. The model is compiled with SGD (Stochastic Gradient Descent) optimizer with Categorical cross entropy. The metric evaluated is the accuracy of the model.

Module 3: Base model integration with Squeeze and Excitation network along with CBAM. Base Model is chosen to be ResNet50. Squeeze and excitation modules are provided between every two blocks of ResNet as shown in Fig 3. FC stands for Fully Convolutional Neural Network.

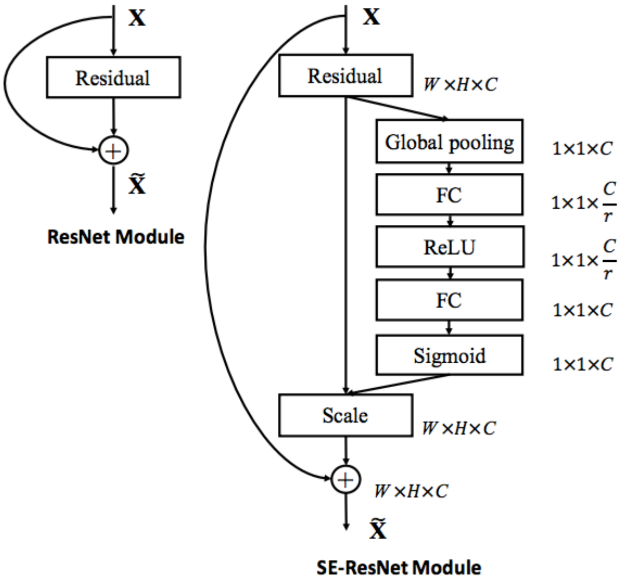


Fig 3. SE-ResNet Module.

To have both channel wise attention and spatial attention, we choose to integrate CBAM with the model. An MLP (for channel attention) and a sigmoid function are used to create a mask from the input feature map in the attention

$$\begin{aligned}\mathbf{M_c}(\mathbf{F}) &= \sigma(MLP(AvgPool(\mathbf{F})) + MLP(MaxPool(\mathbf{F}))) \\ &= \sigma(\mathbf{W_1}(\mathbf{W_0}(\mathbf{F^c_{avg}})) + \mathbf{W_1}(\mathbf{W_0}(\mathbf{F^c_{max}}))),\end{aligned}$$

module. Taking a CxHxW feature map as input, it outputs a 1xHxW attention map (for a 3D attention map, CxHxW attention map is used). The output is then refined and highlighted by elementally multiplying this attention map by the input feature map. Attention mechanisms are typically used in relation with dimensions that are spatial and channel. The generation of these the spatial attention map and the channel attention map can be done concurrently or consecutively.

CBAM sequentially derives a single dimensional channel attention map Mc ∈ C×1×1 with a second dimensional spatial attention map Ms ∈ 1× H×W through an intermediate feature map F∈C×H×W as input. As a general outline of the attention process, we can conclude with the following:

$$\begin{aligned}\mathbf{F}' &= \mathbf{M_c}(\mathbf{F}) \otimes \mathbf{F}, \\ \mathbf{F}'' &= \mathbf{M_s}(\mathbf{F}') \otimes \mathbf{F}',\end{aligned}$$

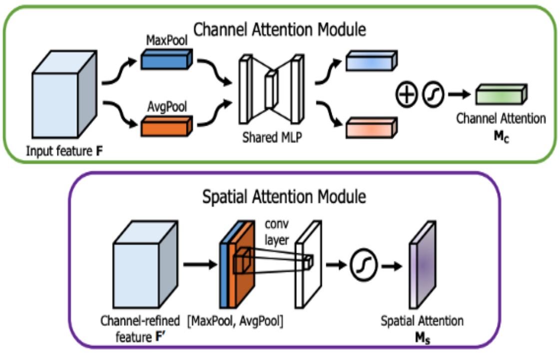


Fig 4. Convolutional Block Attention Module

The symbol ⊗ indicates element-wise multiplication. The attention values are copied in accordance with the multiplication process: channel attention values are broadcast along the spatial dimension, and vice versa. The result of final refinement is F". The computation behind each attention map is illustrated in Fig.3. In order to create the module for directing attention, the channel vector Fc Cx1x1 should be constructed by pooling the feature map F. Send this FC to a very small, C/r-hidden MLP. Here, r represents the hidden channel's reduction ratio (For instance, 64 neurons will be present in the hidden layer if the channel vector length is 1024 and the reduction ratio is 16). The Batch Normalization Layer should be added before this MLP. This also applies to Max Pooling.

Here, W0 ∈ C/r×C and W1∈ C×C/r denote the sigmoid function (σ). Notably, W0 and W1 of the MLP weights follow the ReLU activation function and are shared between the two inputs.

Spatial Attention Module:
Using the input feature map F, produce the two intermediate feature maps Fsavg and Fsmax 1xHxW. Combine these two outputs, Global Average Pool and MP (Max Pooling), and then pass the combined product through a convolutional block with a size of 7x7 kernels. Here, CBAM makes use of large kernel sizes to achieve the same goal as BAM without reducing receptive field. Additionally, this has a d = 1 simple convolutional block.

$$\begin{aligned}\mathbf{M_s}(\mathbf{F}) &= \sigma(f^{7\times7}([AvgPool(\mathbf{F}); MaxPool(\mathbf{F})])) \\ &= \sigma(f^{7\times7}([\mathbf{F^s_{avg}}; \mathbf{F^s_{max}}])),\end{aligned}$$

Sequentially arranged feature maps of this generation. After the Channel attention map has been applied to the input map, the spatial map is then applied. Finally, this F" is updated on the previous input convolutional layer.

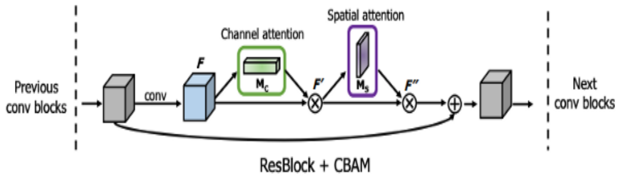


Fig 5. RESBlock and CBAM

Module 4: Training and Testing

ResNet is chosen as base model. CBAM is integrated on top on ResNet. For training, the ImageNet dataset's pretrained weights are utilized. The model is configured with a batch size of 32 and a learning rate of 0.0001 to run for 20 epochs. Using a regularizer of 0.1 and a dropout rate of 0.15, the model is kept from overfitting. Model is made to run with early stopping monitoring the validation loss. If validation loss does not decrease for 5 contiguous epochs, then training halts. Test photos are fed into the model after model training to verify model accuracy.

V. RESULTS

The effective multiclassing of skin diseases made possible by CNN with CBAM has inspired us to employ deep learning for picture classification and segmentation. In theis The lack of medical data and requirement for a GPU supercomputer prevents the basic task from being processed more quickly. Close to 10,015 skin images were used in this study for training the system, validation and for the testing studies environment. Our unique deep learning network architecture, which consists of several trained CNN modules integrating with CBAM to provide us a higher accuracy with a well-trained dataset, has been proposed as a solution to the problems. It was discovered that Xception provided the highest accuracy of 95% of all CNN models employed to get the highest level of accuracy, followed by VGG16 with an accuracy of 90% and VGG19 and DenseNet with accuracy of 83% and 82%, respectively. The same datasets run more efficiently on our suggested network architecture when using the same laptop.

Fig 6. Performance Graph charted with various Models

VI. FUTHER ENHANCEMENTS AND CONCLUSION

Our ultimate objective is to use this method as a rapid and effective method to differentiate between the various categories of skin illnesses and have it function as a Medical Decision Support System to aid clinicians. But with the limited medical resources and applications, they have been slowed down due the lack of data size for training the models and for the use of a supercomputer with can enable faster compute and processing. To find a solution to these prevailing issues, we proposed a Neural Network that consists of various Convolution Neural Networks, which when working together with the Attention Block Module, would give out a higher accuracy rate and within a low computational period. 7 hours were taken to train the system. Our further research allowed us to examine and understand the efficiency of each Convolution Neural Network to assess the various classes of skin diseases.

Further enhancements can be done to this approach to make this Decision Support system a quicker and more efficient method to analyze and classify the different types of skin diseases with minimal contact to the infected area. We believe that, if more work is put into the Decision Support System, we will also be able to assess the extremity of the skin diseases to help dermatologists take the right call and began procedure. More than just 7 different skin diseases can be classified with no compromise in the accuracy of compute power of the system, After the skin condition has been ruled out as a possible diagnosis, a punch biopsy of the afflicted skin will be performed. The suggested system architecture will be able to provide a high-accuracy diagnosis in a matter of minutes once the biopsy has been processed and photographed. Comparing this to manual analysis will result in significant time and financial savings.

VII. REFERENCES

[1] Akay YM, Akay M, Du Y, Sershen CL, Wu M, Chen TY, Assassi S. using the MobileNetV2 model to classify systemic sclerosis skin using deep learning. 2021 Mar 17;2:104-10 IEEE Open Journal of Engineering in Medicine and Biology

[2] Wu ZH, Zhao S, Peng Y, He X, Zhao X, Huang K, Wu X, Fan W, Li F, Chen M, and Li J on various CNN algorithms for categorising facial skin diseases from clinical photos. 2019 May 22;7:66505-11.PY - 2020/04/01 T1 - IEEE Access

[3] Park J., Woo S., Lee J. Y., and Kweon I. For convolutional neural networks, a straightforward and compact attention module. The International Journal of Computer Vision, April 2020;128(4):783-98 M. Akay et al.

[4] Zhao, S., Peng, Y., He, X., Zhao, X., Huang, K., Wu, X., Fan, W., Li, F., Chen, M., and Li, J. studies on various CNN algorithms for classifying facial skin

diseases from clinical images. 66505–66511 in IEEE Access, 7.

[5] Waghmode, V., Sodha, A., and Bhavathankar, P. Rathod et al., March 2018. Convolutional neural networks for skin disease diagnosis. 2018 saw the second ICECA (Electronics, Communications, and Aerospace Technology) international conference (pp. 1048-1051). IEEE.

[6] Ji B., Liu Y., Kang J., Li Y. a network intrusion detection method based on CNN and CBAM The IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS) 2021 will take place on May 10. (pp. 1-6). IEEE.

[7] Yang, N., and He, C. July 2022 Based on the ResNet+ CBAM attention mechanism, malaria detection. The third international conference on information science, parallel computing, and distributed systems (ISPDS) will take place in 2022. (pp. 271-275). IEEE.

[8] Jiang, L., Jia, L., Wang, Y., Zang, Y., Li, Q., Leng, H., Xiao, Z., Long, W., and Jiang, 2022. CBAM and MobileNetV3 for counting bamboo sticks. IEEE Access, 10(5), pages 5396–53971.

[9] Hui L. and Yu-Jie S. the creation of an improved facial recognition algorithm based on convolution neural networks. 2018 May 31 ICIEA 13th IEEE Conference on Industrial Electronics and Applications (pp. 2802-2805). IEEE.

[10] D. Vaghjiani, S. Saha, Y. Connan, S. Frost and Y. Kanagasingam, "Visualizing and Understanding Inherent Image Features in CNN-based Glaucoma Detection," 2020 Digital Image Computing: Techniques and Applications (DICTA), Melbourne, Australia, 2020, pp. 1-3, doi: 10.1109/DICTA51227.2020.9363369.