Multiple Skin-Disease Classification Based on Machine Vision Using Transfer Learning Approach

Conference Paper · November 2023

DOI: 10.1109/ICCCNT56998.2023.10307018

Md Abrar Hamim

University of Wollongong

Shahadat Hossain Sajim

Daffodil International University

Fahim Ur Rahman

Daffodil International University

Multiple Skin-Disease Classification Based on Machine Vision Using Transfer Learning Approach

Md Abrar Hamim
Computer Science and Engineering
Daffodil International University
Dhaka,Bangladesh
abrar15-11821@diu.edu.bd

Shahadat Hossain Sajim
Computer Science and Engineering
Daffodil International University
Dhaka,Bangladesh
shahadate15-2691@diu.edu.bd

Fahim Ur Rahman
Computer Science and Engineering
Daffodil International University
Dhaka,Bangladesh
fahim15-13674@diu.edu.bd

F.M. Tanmoy
Computer Science and Engineering
Daffodil International University
Dhaka,Bangladesh
tanmoy15-13672@diu.edu.bd

***Abstract*—Covering the majority of our body parts, the outer shell-like structure that protects the human body from any outcoming harms, is perhaps the skin. Being the most exposed part, it also suffers from different infectious diseases that causes the inside organs to be vulnerable too. Although it is pretty common to be affected by several skin diseases, identifying the disease flawlessly is often seen to be confusing as the diseases tend to be hard to distinguish between. Applying computer vision with a decent trained classification model can come in really useful in such scenarios. Among vastly available classification models, not every model can perform similarly in terms of identifying the precise disease category. To solve this concern, a custom collected dataset has been gathered, processed according to needs and afterwards, a transfer learning model known as "MobileNet-v2" has been trained and tested. The testing accuracy as demonstrated by the model was 83% in terms of both the testing dataset and unseen images. The study reflects that, if flawless dataset is ensured and the training parameters are maintained, accurate skin disease detection can be automated and at the same time it can be lightweight that reduces resource usages being a light model. The trained model can also be useful in medical implementation by taking further improving techniques.**

Keywords— Multiple skin disease, Machine vision, Transfer learning.

I. INTRODUCTION

The bleaker and more morbid the climate and environment around us become, the greater the proportion of skin disorders we encounter. Undoubtedly, the skin is the largest organ in the human body, and it is also the most exposed. A great many of us are unaware that it is susceptible to the elements in a variety of ways. The skin consists of two layers: the epidermis (the top layer, which we think of as the "skin") and the dermis (the lower layer, which gives the skin its strength and elasticity). According to a Food and Drug Administration survey, the average American is subjected to nearly 300 different chemicals through their skin each day. Environmental Working Group research revealed that more than fifty percent of lotions, shower gels, and beauty products on the market involve skin-irritating chemicals that can trigger acne, hives, breathing difficulties, or allergies. According to the National Library of Medicine, skin diseases account for 1.79 percent of the global disease burden measured in disability-adjusted life years (DALYs). Here is a list of skin diseases and their associated global DALYs, from highest to lowest: atopic, contact, and seborrheic dermatitis; acne vulgaris; urticaria; psoriasis; viral skin diseases; fungal skin diseases; scabies; melanoma; pyoderma; cellulitis; keratinocyte carcinoma; decubitus ulcer; and alopecia areata. All of this concern calls for an efficient way of detecting the most common skin disease in the US and the world. Because most skin diseases are not lethal as long as they are diagnosed early, they can be cured. If doctors can determine the most prevalent illnesses, they can treat patients more efficiently and limit their spread. But as skin is a visual marker, patients need a large amount of blood drawn to test for diseases. A substantial percentage of false-positive findings, lengthy turnaround times for testing results, and inability to be performed in a clinical context are further issues. So the scope of image processing can be a huge step in the correct diagnosis of many skin diseases. Using image processing and neural networks, we can discover the patterns of a disease and create software that accurately recognizes the patterns of a range of diseases in a timely manner. This can be done with classification algorithms. In the fields of deep learning and machine learning, "classification algorithms" are used to teach a computer a higher-level abstraction from the examples it is given in order to predict which class a picture belongs to. The categorization algorithms mobilnetV2, VGG19, and VGG16 are among the well-known and potent ones we employ. These models are well-tested and well-trained in several disciplines of image recognition, as evidenced by public benchmarks like MNIST, CIFAR-10, SVHN, and ImageNet.

Later in this research, we analyzed and compared these three classification models and chose the mobilnetV2 model, which delivers the greatest accuracy at the lowest computing cost of the three classifiers. In medical centers and modern healthcare facilities, health data is growing, which may lead to difficulties in extracting key or most crucial info for a decision support system as a result of these intricacies, inconsistencies, and time consumption for a single operation for a dermatologist to make a diagnosis of internal infections using such lengthy reports, especially on skin due to its severe impact on lives and health. Our AI algorithms might be used to analyze massive datasets and retrieve salient characteristics to allow rapid medical diagnosis through the dermatological information management system (DI-IS) software. This is a potential solution for patients without access to dermatologists, decreasing their access costs and boosting the quality of dermatology care provided. The use of automated systems like this can greatly reduce human errors and speed up the processing of health data, leading to faster and more precise treatment. We are attempting to extend artificial intelligence's comprehension of skin color, texture, and other traits in order to boost the accuracy of illness identification. Additional research is necessary to determine the scope of AI-driven dermatology applications for the medical community. As it is a subfield of data mining, further applications include developing automatic tools for the detection and treatment of skin cancer, which affects over 5.4 million Americans every year. And the applications keep growing, and the opportunities are endless.

II. Literature Review

For the purpose of identifying skin diseases, we are developing an image processing-based technique. After classifying the skin region feature picture of the illness result, the disease type may be determined using image analysis. For feature extraction, we use a pretrained CNN, and for the proper purpose of categorizing the pictures, we use a multiclass SVM. The method detects three different forms of skin illnesses with a 100% accuracy rate and operates with single-color inputs [1]. In this paper, we have presented a way for identifying various dermatological skin conditions utilizing computer vision-based methodologies. For feature extraction, we employed a variety of image processing techniques, and for both training and testing, we used feed-forward artificial neural networks. The technology operates in two stages, first extracting important information from color photographs of the skin, and then diagnosing disorders. The technology effectively diagnoses nine distinct forms of dermatological skin problems with a 90% accuracy rate [2]. In this work, five diverse system study methods were selected and fine-tuned on specific skin infection registries in order to estimate the precise level of skin illnesses. We focused on random forest, Naive Bayes, logistic regression, SVM kernel, and CNN as system study methods. A comparative analysis based on the characteristics of the confusion matrix and the accuracy of the instruction was produced and delineated using graphs. Among the selected media outlets, CNN was shown to be the most effective in training people to identify skin conditions [3]. The primary objective of this endeavor is to enhance the precision of diagnostic structures by employing image processing and typing approaches. MSE and PSNR are retrieved as image quality evaluation skills. The derived texture capabilities may be utilized to identify the existence of pore and skin disorders and, if present, categorize illnesses such as melanoma, leprosy, and eczema using a decision tree [4].The technique of the gray-degree co-prevalence matrix, also known as GLCM, became applied to phase images of skin illness so that the texture and shading functions of numerous skin ailment photographs may be received precisely [5].The objective of this project is to classify various diseases based on the photographs supplied as input. This project utilizes the Python software platform exclusively. Images are taken from several public databases, such as DermWeb and DermNet. Using machine learning with TensorFlow to train the dataset and the SVM technique to categorize the five skin disorders in a Python program are the methods proposed here. You can identify the type of condition, such as psoriasis, melanoma, rosacea, vitiligo, or xanthoma, and the Python program output will display the result as the name of the ailment [6]. The reddening approach is utilized in this study to identify red items on the skin of the face. By selecting features based on gamut, average RGB color space intensity, and average tonal intensity, the output of the reddening technique is maximized. This study's data set comprises 35 facial photos. Utilizing sensitivity, specificity, and accuracy to quantify detection performance. The sensitivity, specificity, and accuracy, according to a dermatologist, were 54%, 99.1%, and 96.2%, respectively. In contrast, PT. For AVO personnel, sensitivity, specificity, and accuracy reached 67.4%, 99.1%, and 97.7%, respectively. According to the results, the system is highly capable at identifying face redness [7]. Using a deep learning neural net and image analysis to diagnose skin cancer, an attempt was made to circumvent this issue. In this article, the authors discuss the most recent breakthroughs in authoritative deep learning principles for skin cancer diagnosis and classification [8]. Researchers found that melanoma was both malignant and benign when detected at an early stage. Data from dermoscopy were obtained from the HP2 and ISIC in 2018. Blended feature isolation. Big transfer learning models like AlexNet and ResNet50 were used to pre-train the CNNs. On PH2 (97.50%) and ISIC 2018 (98.35%), the ANN model outperformed the CNN model. In order to categorize skin illnesses, we used CNN transfer learning, namely AlexNet and ResNet50 [9]. This study provides a technique that assesses the form and size features of linked components and utilizes these qualities to categorize bacterial and viral tropical skin diseases. Compare the results of the gradient boosting machine approach to those of the deep learning strategy to demonstrate its effectiveness. The results demonstrate that when trained on 1460 photos, our method's performance is equivalent to that of a convolutional neural network (CNN)-based method. CNN was also pre-trained and required augmentation to accomplish this accomplishment. In contrast, our method is at least 56 times greater than CNN's [10]. This project will design and construct a system to combine Pigmented Skin Lesions (PSL) images, their analysis, and related observations and conclusions. Medical experts prototype This database is a library. A section of the system will use AI to evaluate, analyze, and categorize photo library data based on textures and morphological features. Medical staff at a remote location can use portable data collecting devices (such as cell phones) to take photographs of PSL and send them to the recommendation system, which can subsequently decide if the PSL is cancerous or benign (non-threatening) [11]. The skin gray color profile was used in the work that was proposed to be done in order to fulfill the role of an input parameter. The skin profile was able to be established as a result of this. In this article, an ABCD approach is proposed as a means of melanoma, a kind of skin cancer, detection. The techniques and methods of image processing and soft computing technologies that are used to detect melanoma more accurately are broken down and presented in this study. The study was carried out by the National Cancer Institute [12]. This research study explores ways of improving machine learning algorithms using three distinct skin melanoma datasets, each of which presents a unique set of challenges. It evaluates several preprocessing steps, machine learning models, and feature extraction techniques to discover the optimal way for each data set. In addition, the temporal and spatial complexity of the method is assessed in an effort to decrease consumption of resources without significantly degrading the model's ability to perform [13]. The suggested approach is effective in preserving context for reliable forecasting. Disease development may be evaluated with the use of a grayscale co-occurrence matrix. FineTuned Neural Networks (FTNN), Convolutional Neural Networks (CNN), Very Deep Convolutional Networks for large-scale image identification created by Visual Geometry Group (VGG), and the Convolutional Neural rice field were all compared to see how they performed. Upgraded structure of the communication network. Using the HAM10000 dataset, the suggested technique achieves a rate of accuracy more than 85%, making it the most effective of its kind [14]. Using this technology, dermatologists can diagnose eczema, psoriasis, lichen planus, benign tumors, fungal infections, and viral infections more rapidly and precisely increase. This technique may be utilized to identify the causes of skin problems and treat them accordingly. Python was used to produce a functional version of the expert system recommendation. This approach utilizes a dataset provided from DERMNET to train a 50-layer residual neural network. ResNet's precision increased to 95%. Use 10 epochs to train and make predictions with your model [15]. This study developed a system based on deep learning to identify common skin conditions. Using 1067 images, our dermoscopic method achieved an accuracy of 87.25 2.24 percent. Situations involving results-based diagnosis and classification demonstrate the need of merging dermatologists' expertise with computational methodologies. The method for future computer-assisted decision making can be improved by the semantic synthesis of diagnostic scenarios [16]. Utilizing the suggested method, the model's efficacy is evaluated. When applied to the optimum subset of erythemato-squamous disease, correlation and heat map feature selection procedures are proven to be efficient. To demonstrate the effectiveness of the proposed model, we compute a variety of statistical parameters, including the mean, standard deviation, root mean square error, kappa statistical error, area under the receiver operating characteristics, and accuracy. The results of the feature selection methods used in the staking ensemble methodology are superior than those of the individual machine learning methods. The obtained data indicates that the proposed model performs better than its predecessors [17].

III. METHODOLOGY

In this part, we have demonstrated our approach towards solving the identification problem of different skin diseases using different deep learning models with the help of transfer learning to ensure lower resource usage and better preciseness while detection. A short overview of our conducted experimental process is visualized in the figure below:

A. Data Collection:

The implemented system tends to identify skin diseases from a multi-class dataset. The dataset used to propose is a custom one which adds to its uniqueness from other available state-of-the art online accessible datasets. Using a custom dataset has its own benefits although in certain circumstances it seems to be hard enough to put it together precisely. The primary source where we have looked for useful data is "Kaggle". Kaggle provided datasets that were humongous and irrelevant to some extent with our goal. So, selecting the most appropriate ones became necessary. Upon selection, it was found that the numbers of classes of different skin diseases were not up to our primary expectations. 8 different classes containing images were found beneficial to our target and these were added first. The other required classes required us to search individually on the internet and scrap from there manually.

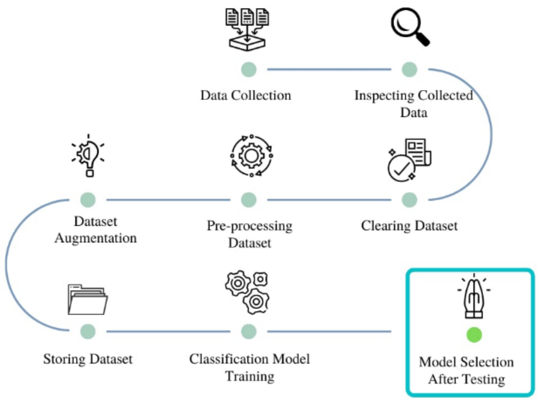

Fig1: Methodology Flowchart

B. Inspecting Collected Data:

Data inspection is the systematic process of viewing the collected datasets and verifying manually if the collected images are relevant to each portion or not. Although data inspection can be done in any stage of dataset preparation, it is most suitable to review the dataset before applying any sorts of pre-process or editing techniques. Through a proper inspection of our collected data, it seems useful regarding the outcome of useful and beneficial information about the dataset. Inspecting the dataset also allows us to find out an overall state and provides a clear idea on which we are going to work later. In the process of dataset inspection, it was evident that the primary collected dataset consists of several unwanted and irrelevant data that would cause training loss to increase without actual noticing. For ensuring the best validation for the model, these unwanted pieces of images were marked for clearing in the following step.

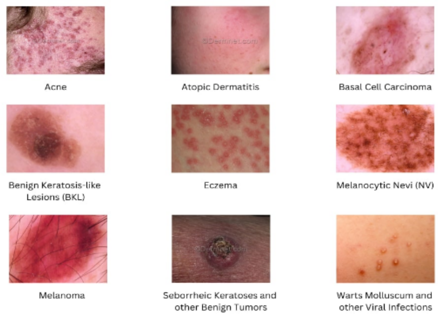

Fig2: Dataset Example

C. Clearing Dataset:

The corrupted and inaccurate data that were marked in the inspection phase, were carefully removed from the corresponding dataset while making sure that no useful data is hampered in the process as clearing dataset tends to indicate the cleaning procedure for the dataset we are working with. After a successful clearing attempt, the dataset state consisting 9 classes of different numbers of images can be demonstrated below.

Table1: Dataset Images

D. Pre-processing Dataset:

Image classification is a remarkable process by itself. But achieving that positive remark is quite a difficult task for the computer by default. Just plain images will not suffice in this constituent of image processing. There are a significant number of fundamental restraints that hold us back here. If the image size is inconsistent for the whole dataset, it will be a noticeable hindrance to the classification model, as it is principally faster for the model to train on smaller and even image sizes. Preprocessing is a process of altering the dataset of images in different ways to make it easy for the model to calculate those images efficiently. It is essentially the steps that are taken to format the images before they are fed to the model to train on.

For our dataset we first collected the dataset from different sources, and that's why the images are of fundamentally different formats, the most important one is size, first we needed to change the size of our images to a fixed and consistent size, the reason for this is simply because, any deep learning model will expect the array representation of the images to be of same size, otherwise the calculations will not yield positive result. Because we used different models to comparatively get our result, we needed different image sizes accordingly. Every model expects different sizes of image resolution for the most effective outcome. So, we used the desired image sizes for each model respectively. We mainly resized our images to 448x448 for VGG16 and VGG19 models and for Mobilnet_V2 we also retained the similar size as it expects any image size greater than 32x32. We also rescaled the images for efficient training and processing of our models. Rescaling it with the scaling ratio of (1./255), this essentially rescales our image pixels values to the range of [0, 1]. This normalization is important for any machine learning algorithm, because rescaling our data helps the training process much faster than it is without rescaling. And the accuracy also increased because of it.

Class Name	**Consisting Images**
Melanocytic Nevi (NV)	4500
Melanoma	3100
Warts Molluscum and other Viral Infections	2080
Benign Keratosis-like Lesions (BKL)	2080
Seborrheic Keratoses and other Benign Tumors	1848
Eczema	1534
Atopic Dermatitis	1189
Basal Cell Carcinoma	1125
Acne	1094
Total:	**18550**

E. Dataset Augmentation:

All image dataset is not built ground up. Some of the dataset is collected from different sources. And these dataset does not contain the same amount of data 100% of the time, and some data classes contain less data to problematically impact the accuracy of the model. This happens because the biases of our deep learning model do not remain in equilibrium, the bias of the model shifts toward one type of data or one or more classes and remains poor on the classes with low number of data, this results in inaccurate predictions. For that Augmentation is needed, augmentation is the process of modifying existing images slightly to generate more dataset. There are many types of augmentation methods that are used in a dataset, there are zoom augmentation, rotation augmentation, hue shift, hue blocks, shearing etc.

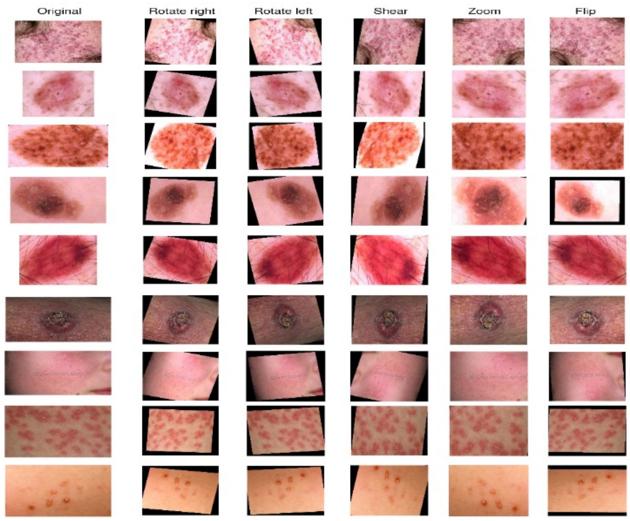

Fig3: Dataset Augmentation

Augmentation is also necessary for introducing variation in any dataset, sometimes the dataset we are using is overfitted to the model and the model becomes less capable to distinguish the classes in the validation and test dataset. In these scenarios we need to create augmented data so that the model can accurately distinguish new data from the real world. In our skin disease dataset, we used augmentation to teach the model to know the skin disease types in any orientation and shot variety possible. Also, to even out the amount of data in each class as our data is collected from different sources.

In order to enrich and diversify our dataset such that deep learning models could train without suffering significant loss, we applied a variety of augmentation techniques. When we intended to expand the model's data, we kept the most authenticity.

F. Storing Dataset:

The more novel and unique the research, the harder it is to find data for. Some datasets are usually generated from scratch, but fortunately for us, we didn't need to do so because the data that we worked with is much more gathered than built. Our dataset is substantially larger than text or numeric data because we are working with visual data. Image datasets are considerably more difficult to find and larger in size. That brings us to the dataset's storage security. To train the dataset, we combined data from several sources into a local directory and fed it into the deep learning model. We can utilize the same folder structure for all of the transfer learning models we used in this research because they all use the same ImageDataGenerator API from Tensorflow Keras. We have developed a cloud solution as a data backup measure since it is much more accessible and secure, which eliminates the possibility of data loss and corruption. We used a dataset that was 3 Gigabytes in size and had about 18552 photos. For our cloud solution, we used Google Drive, and for local computation, we used our own personal PCs.

G. *Classification Model:*

1) *Model Training:*

Model training is the process of training our machine learning or deep learning model with our selected and curated training data. Training data is the data that the ML model learns on. In supervised or unsupervised models there are mainly three categories of data that we work with, Training, testing and Validation data. And they are quite self explanatory by their name.it is also known by the term model fitting or data fitting. Since we are comparing 3 models in this paper, we are effectively fitting three models with the same data. We are using all the three dataset types, with a validation split of 80/20, 80% for training and 20% for validation and testing. Below we are including the brief definition and description of the models that we trained for our problem at hand.

2) *VGG 16 :*

VGG16 is an object detection and classification model that is widely used in the professional industry. The transfer learning approach made this model very famous for its efficiency and its use of the famous "ImageNet" dataset. It is a profound model that by default can classify 1000 classes of images with an astounding 92.7% accuracy. It consists of 21 layers in its architecture and can work with higher quality images with all 3 color channels. Trained on our dataset, the VGG16 model gave us ___ accuracy and ___ validation loss which is quite good considering the difficulty and intricacy of our problem. The feature extraction capability helped us get a grasp of the approachability and potential problems that can arise from using these models.

3) *VGG 19:*

VGG19 is a one of a kind state of the art Convolutional Neural Network that is one of the latest deep learning models for object detection and classification among the VGG architecture. In the eyes of Simonyan and Zisserman VGG19 is a significant improvement from VGG16. It consists of 19 layers in total, a common 16 convolution layer and 3 fully connected layers. Consisting of 19.6 billion FLOP's is also uses the ImageNet weights and is a substantially better performer in the field of object recognition. In our dataset the VGG19 model performed comparatively better than the VGG16 model in terms of accuracy and faster training time, the validation loss also signifies improvements. We acquired a total accuracy of __ and validation loss of __ which is better in terms of performance and proof of concept in the improvement of the VGG-Net architecture in transfer learning for our problem and dataset.

4) *MobilNet-v2:*

Efficiency is a must in object detection and recognition. And MobilNet architecture is a perfect candidate for that. MobileNet-v2 is also a Convolutional Neural Network that specializes in image processing, but the improvement comes from the efficiency of the model. Because it is developed mainly with mobile devices in mind, it is highly optimized and that is a great opportunity for our dataset. Since we are working with such a large dataset of around 20k data, and also with the potential of using this model with mobile devices to detect skin disease on the fly with our smart phone, we are using it for our problem. The dataset we are using and the approach of transfer learning that we are using is very compatible with this model and gives us an astounding accuracy of __ and a validation loss of __. We can see that amongst the three models we are getting the most performance in case of accuracy and also loss from this model in particular.

Because of its inverted residual structure the skin disease feature map extraction is much more logical than random and much more robust in nature.

After analyzing the accuracy and losses of all three deep learning models. We came to a conclusion that MobilNet-v2 gives us the best accuracy and performance overall. We got significantly better training time and training accuracy too. In the classification report we can see that the average accuracy of predicting the right class for each of the classes is much better in this model than VGG16 and VGG19. This makes MobilNet-v2 much more robust and generalized for our dataset.

5) *Proposed Model Architecture:*

A convolutional neural network with 53 layers is known as MobileNet-v2. The ImageNet database contains the network's pretrained version, which has already been tested on more than a million images [19]. The transfer learning model we use in this work is MobileNet-v2. MobileNet-v2 uses the Depthwise Separable Convolutions (DSC) methodology for scalability and introduces a brand-new structure called reversed residuals to preserve the information in addition to tackling the issue of data lost in non-linear layers in "connected" by employing linear bottlenecks.

a) *Depthwise Separable Convolutions:*

In addition to the Depthwise Separable Convolutions used in MobileNet-v1, this version also uses them. By merging the Depthwise convolution and the Pointwise convolution, the total amount of parameters and processing costs can be reduced from the normal convolution's 32 to about 18 for the same level of performance.

b) *Linear Bottlenecks:*

Two characteristics serve as the basis for the linear bottleneck idea. If a layer's result comes in the form of ReLU (Bx) and the outcome remains non-zero, it is obvious that the corresponding piece of the input space (x) is confined to a linear transformation (Bx). In other words, only the non-zero volume region of the output domain can be used by deep networks as a linear classifier. To the contrary hand, data in a channel is always irrecoverably destroyed when ReLU collapses it. Whenever there is a pattern in the collection of kernel functions and there are multiple channels, the data can still be stored in those channels. In this instance, this kind of structure serves as a layer that can incorporate the input into a reduced-dimensional region of the activating space. The foregoing arguments show that if the input space could be embedded into a low-dimensional environment while using ReLU as an activation function, a linear transformation is enough to store all of the pertinent data. Convolutional blocks may use linear bottleneck layers to achieve this.

c) Inverted Residuals:

Nearly all of the data is held in the bottlenecks, which leads to the introduction of shortcuts between different bottlenecks to improve gradient propagation capabilities in multiplier layers. The expansion layer can be viewed as an implementation specific with a non-linear tensor transformation. Additionally, using the inverted design (Inverted Residual) is more computationally simple than using the standard structure.

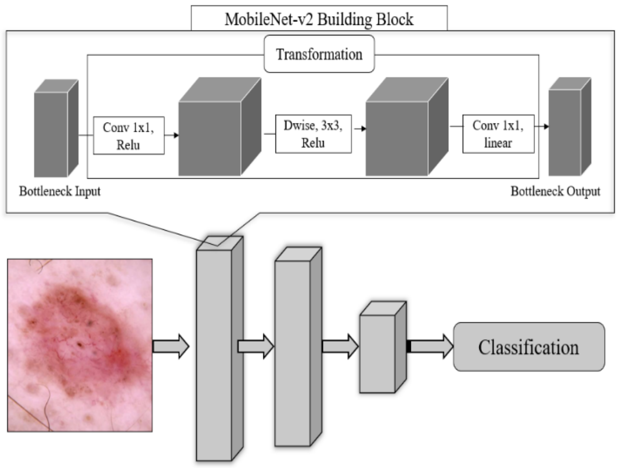

Fig4: Model Architecture of MobileNet-v2

d) Model Architecture:

The core building block is a bottleneck depthwise separable convolution with residuals. By applying transfer learning, MobileNet-v2 can be seen as the head, while other layers can be seen as the tail for certain classification tasks. Here, a 448 by 448 image serves as the input for a MobileNet-v2 variation. In order to filter the input, the model first transforms the input's condensed low-dimensional representation into high dimensions. The attributes are then projected using a linear convolution to return to a low-dimensional form. This model's organizational structure may simultaneously preserve the data, deal with MobileNet-fixed v1's quantity of filters, and maintain the block size modestly. The MobileNet-v2 building block is depicted in Figure () along with the specifics of transformation, such as bottleneck input, output, and operations.

IV. Result Evaluation

Having a diverse and large amount of image data, we had to implement three different deep learning algorithms with the approach of transfer learning so that resource usage and time to execute the model could be lessened. In order to achieve the highest accuracy possible, three DL algorithms, VGG16, VGG19 and MobileNet-v2 have been implemented and tested over our custom dataset. As the results and outputs are in statistical format straight out of the model, we can convert the numbers into a table for comparison among them. The comparison table among the three models with their respective evaluations scores is as followed:

Table3: Model Efficiency

DNN Model	**Accuracy (Testing)**
1. VGG16	59.33%
2. VGG19	62.24%
3. MobileNet-v2	83.00%

A. Training and Validation Curves:

Among three different DL algorithms, comparing results with each other, a conclusion can be formed that the "MobileNet-v2" is performing with the highest accuracy on the test dataset portion. The training and validation loss and accuracy curves of the best accuracy model MobileNet-v2 are as followed:

Fig5:Training loss and Validation Loss Curve

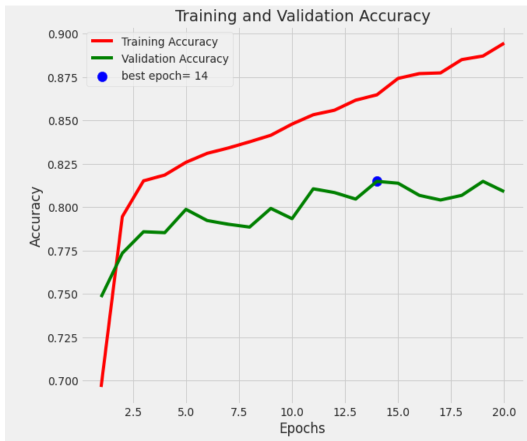

Fig6: Training Accuracy and Validation Accuracy Curve

Here, training loss indicated the frequency of errors while training the model. Also, validation loss is showing the performance of the training on the validation subset of the dataset, and validation loss is slightly higher than the training loss. From the starting epoch, both the training and the validation loss is seen to decrease as the number of epochs progresses. The best epoch is recorded as epoch 49 where both the losses were at the minimum state. The overall curves of these two combined, demonstrates a good fit for our dataset eventually indicating that no over-fitting or under-fitting is seen.

Another noticeable improvement is seen in the curves of training and validation accuracy. Starting from the initial epoch, the best epoch is stated 23 indicating highest validation accuracy sitting just over 82.5%.

B. Classification Report: of MobileNet-v2:

Table4: Classification Report of MobileNet-v2

Class Labels	Precision	Recall	F1-Score	Support	Accuracy
Acne	64%	71%	68%	108	83.00%
Atopic Dermatitis	58%	54%	56%	123	
Basal Cell Carcinoma	99%	99%	99%	118	
Benign Keratosis-like Lesions (BKL)	91%	89%	90%	178	
Eczema	78%	70%	74%	166	
Melanocytic Nevi	93%	95%	94%	460	
Melanoma	96%	95%	95%	333	
Seborrheic Keratoses and other Benign Tumors	64%	69%	66%	173	
Warts Molluscum and other Viral Infections	68%	67%	67%	200	

C. Generating Confusion Matrix:

Code-generated heatmap of the confusion matrix can be shown in fig: . The confusion matrix visualizes predictions and actual labels of the test images in which it is seen that the number of wrong predictions is significantly lower than the true predictions that match with the actual labels of the images. Confusion matrix is extremely useful while measuring recall, precisions or accuracy of trained models. The lower the wrong prediction ratio gets, the better accuracy we achieve in terms of calculating accuracy of our model which is almost 83%.

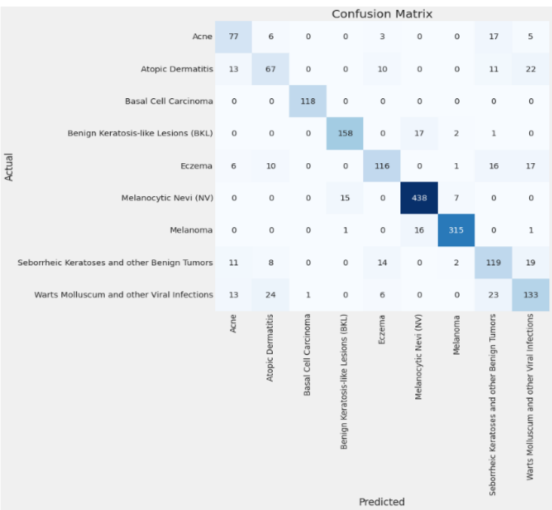

Fig7: Confusion Matrix

V. Conclusion

Being the most exposed and outer layer of the body, our skin protects the inner organs from being harmed through outside harms. As every other organ of the body, it also suffers from different sicknesses during this tiresome uninterrupted process, the concern regarding well-being of the skin is crucial. In most of the cases, skin diseases do not usually hamper our everyday movements at large amount, but the least infectious attack from any disease could lead to something very dangerous. Unlike other procedure, the sensitivity of skin is also a concern. As a result, we need to be very cautious and responsive at the same time for avoiding unwanted dilemmas. Here comes the usefulness of a machine-based detection program that can drastically reduce the anxiety of whether we are affected or not. Keeping this in mind, a proposal that includes automatic detection of skin disease and successfully identifying the kind of it may be advantageous in terms of early finding. The preciseness in finding out the disease with lesser times also contributes to a better health inclusive of other benefits. Although a decent amount of dataset is used in training the proposed model, there remains some lacking in terms of variety of image sets. In future, we will be focusing on the dataset image variety along with ensuring more focused images of particular diseases. With that, we have also planned for implementation of this on mobile application for greater use case scenarios and easy access to our system.

References

[1] G ALEnezi, Nawal Soliman ALKolifi. "A method of skin disease detection using image processing and machine learning." Procedia Computer Science 163 (2019): 85-92.

[2] Yasir, Rahat, Md Ashiqur Rahman, and Nova Ahmed. "Dermatological disease detection using image processing and artificial neural network." In 8th International Conference on Electrical and Computer Engineering, pp. 687-690. IEEE, 2014.

[3] Bhadula, Shuchi, Sachin Sharma, Piyush Juyal, and Chitransh Kulshrestha. "Machine learning algorithms based skin disease detection." International Journal of Innovative Technology and Exploring Engineering (IJITEE) 9, no. 2 (2019): 4044-4049.

[4] Pugazhenthi, V., Sagar K. Naik, Amruta D. Joshi, Shreya S. Manerkar, Vinita U. Nagvekar, Kalpita P. Naik, Chinmay G. Palekar, and K. Sagar. "Skin disease detection and classification." International Journal of Advanced Engineering Research and Science (IJAERS) 6, no. 5 (2019): 396-400.

[5] Wei, Li-sheng, Quan Gan, and Tao Ji. "Skin disease recognition method based on image color and texture features." Computational and mathematical methods in medicine 2018 (2018).

[6] Muthalagu, R., M. Mohammed Musheer, K. Nandhini, and N. Saranraj. "Type of Skin Disease Identification by Machine Learning Using Python."

[7] Muhimmah, Izzati, Nurul Fatikah Muchlis, and Arrie Kurniawardhani. "AUTOMATIC FACIAL REDNESS DETECTION ON FACE SKIN IMAGE." IIUM Engineering Journal 22, no. 1 (2021): 68-77.

[8] Saeed, Jwan, and Subhi Zeebaree. "Skin lesion classification based on deep convolutional neural networks architectures." Journal of Applied Science and Technology Trends 2, no. 01 (2021): 41-51.

[9] Alsaade, Fawaz Waselallah, Theyazn HH Aldhyani, and Mosleh Hmoud Al-Adhaileh. "Developing a recognition system for diagnosing melanoma skin lesions using artificial intelligence algorithms." Computational and mathematical methods in medicine 2021 (2021).

[10] Kiwanuka, Fred N., Omar Eltaher Abuelmaatti, Anang Hudaya Muhamad Amin, and Brian J. Mukwaya. "Tropical Skin Disease Classification using Connected Attribute Filters." In VISIGRAPP (5: VISAPP), pp. 338-345. 2021.

[11] Okuboyejo, Damilola A., Oludayo O. Olugbara, and Solomon A. Odunaike. "Automating skin disease diagnosis using image classification." In proceedings of the world congress on engineering and computer science, vol. 2, pp. 850-854. 2013.

[12] Muralidhar, V., and R. Satya Prasad. "ANALYSES OF SKIN IMAGES USING IMAGE PROCESSING ABCD TECHNIQUES."

[13] Alasandagutti, Akhil Reddy. "Using Deep Learning to Automate the Diagnosis of Skin Melanoma." (2021).

[14] Srinivasu, Parvathaneni Naga, Jalluri Gnana SivaSai, Muhammad Fazal Ijaz, Akash Kumar Bhoi, Wonjoon Kim, and James Jin Kang. "Classification of skin disease using deep learning neural networks with MobileNet V2 and LSTM." Sensors 21, no. 8 (2021): 2852.

[15] Sharma, Moolchand, Bhanu Jain, Chetan Kargeti, Vinayak Gupta, and Deepak Gupta. "Detection and diagnosis of skin diseases using residual neural networks (ResNet)." International Journal of Image and Graphics 21, no. 05 (2021): 2140002.

[16] Zhang, Xinyuan, Shiqi Wang, Jie Liu, and Cui Tao. "Towards improving diagnosis of skin diseases by combining deep neural network and human knowledge." BMC medical informatics and decision making 18, no. 2 (2018): 69-76.

[17] Verma, Anurag Kumar, and Saurabh Pal. "Prediction of skin disease with three different feature selection techniques using stacking ensemble method." Applied biochemistry and biotechnology 191, no. 2 (2020): 637-656.