International Journal of Engineering Science Invention (IJESI)
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org || PP. 202-209

Identification and classification of skin diseases usingDeep learning techniques

Venkatesh R Pai[1], Soujanya G Pai[2], Srivathsa K[3], Dr. Rekha PM[4]
[1]*(Student, Information Science and Engineering, JSS Academy of Technical Education,Bangalore, Karnataka, India)*
[2]*(Student, Information Science and Engineering, JSS Academy of Technical Education,Bangalore, Karnataka, India)*
[3]*(Student, Information Science and Engineering, JSS Academy of Technical Education,Bangalore, Karnataka, India)*
[4]*(Associate professor, Dept of Information Science and Engineering, JSS Academy of TechnicalEducation, Bangalore, Karnataka, India)*

Abstract: *Health is the most important aspect of a human's life. Humans are covered by the largest organof their body, the skin. Protecting the body from germs and regulating body temperature are the basic functions of the skin. If not protected, then it may lead to skin diseases caused by fungal infections, allergies, bacteria, etc.The cost of diagnosing skin diseases may not seem cost-effective to everyone. Moreover, dermatologists may haveto invest more time and use their professional experience in diagnosing skin diseases. The paper aims to understand the various methods for building a neural network model that helps in the identification and classification of skin diseases caused by bacteria and fungus. The problem statement is an image classification task. Since the task requires an image as input, the commonly used image preprocessing techniques are presented. The Convolutional Neural Network (CNN) is known to be the most suitable deep learning algorithm for image analysis. The functions of the various layers in the CNN architecture and the various approaches in transfer learning are studied. Also, a solution that leverages the strength of combining multiple CNN is considered.*
Keywords - *Convolutional Neural Network, Deep learning, Skin disease, Image processing, transferlearning.*

I. Introduction

The human skin is the largest and outermost organ of the human body. In addition to water and proteins, the skin also contains fats and minerals. The skin primarily protects the human body from harmful substances existing outside the body and also prevents the outflow of various nutrients present inside the human body. The epidermis, dermis, and hypodermis make up the three main layers of the skin, and all three layers are susceptibleto skin diseases like acne, rashes, wrinkles, and even cancer that are life-threatening. The skin, along with the hair, nails, oil glands, and sweat glands, is part of the integumentary system.

Skin is a very sensitive part of the body. Various internal and external factors harm the functionality of the skinwhich leads to various skin diseases. The problem of skin diseases is a global concern that affects people of all ages. Symptoms such as burning, redness, itching, and swelling can be caused by anything that harms the skin. Allergic reactions can irritate the genetic structure of the skin which causes rashes and other skin conditions. Several skin conditions, such as acne, can also negatively impact the appearance of the skin. Skin diseases are chronic and may grow into malignant tissues. Hence, to minimize the major damage, it should be treated immediately. Several skin diseases may not exhibit any symptoms; they should be diagnosed and treated before they could cause serious problems.

Dermatology is an area of medicine that deals with issues related to hair, nails, and skin. Traditionally, the diagnosis of a skin disease usually involves clinical screening and then a dermoscopic analysis. Dermatologists often have to perform laboratory tests before concluding the type and stage of skin disease. Despite advances in medical equipment giving accurate and quick results, the cost of such a diagnosis is usually expensive for many people to consult a dermatologist. Thus, there is a need for providing a quick and cost-effective solution. Researchers have explored many algorithms and methods to diagnose skin diseases.

II. Literature Survey

The most common form of the disease in humans is skin disease. The causes of skin diseases include fungal infections, viruses, bacteria, etc. Dermatologists play a crucial role in the traditional method of skin disease identification. A dermatologist observes the patient in the first instance to gather the skin condition based on hisknowledge and experience and it is followed by a skin imaging process known as dermoscopy for observing theskin structure.

Traditional approaches to skin disease diagnosis have been replaced with machine learning methods to overcome their limitations. It requires the use of manual extractors for the extraction of features of skin diseases and then using machine learning algorithms for classification. Further, since this process is performed manually, it requires professional medical knowledge as well as the ability to conduct deep exploratory data analysis to reduce dimensions that limit its capability when it comes to recognizing skin disease images. [1].

A lot of researchers have been interested in using image recognition-based deep learning techniques to diagnose skin diseases due to the advancement of deep learning technology.

In [2] authors have used a hybrid approach that involves deep feature fusion, and several SVM classifiers. It combines the probabilities from multiple classifiers to obtain the final classification. In the researchers' study, three pre-trained deep learning models were used as deep feature generators, including VGG16, AlexNet, and ResNet-18. For extracting the features, the last fully connected layers of the pre-trained AlexNet and pre-trained VGG16. Since ResNet-18 has only one fully connected layer, the features were extracted from the last convolutionlayer of this model. ROC curve was used as an accuracy metric. The combination of all networks resulted in 97.55 % and 83.83% for classifying melanoma and seborrheic keratosis diseases respectively.

The paper [3] focuses on computer vision algorithms and various image processing algorithms for feature extraction, and ANNs are used to train and test the algorithms. It first preprocesses the images of the skin to extractspecific features, then determines the disease type. Eight image processing algorithms were used. They were the median filter, binary mask, smooth filter, sobel operator, gray image, histogram, sharpening filter, and YCbCr. Ten different features were used for modeling the data. Using them a test accuracy of 90%, 85%, and 88% was achieved for supervised, unsupervised semi-supervised systems respectively.

For identifying histopathological characteristics of clinically evaluated samples, the authors of [4] used computer vision and machine-learning methods. There were 6 diseases that they explored, which were: pityriasis rosea, chronic dermatitis, pityriasis rubra pilaris, seborrheic dermatitis, psoriasis, and lichen planus. They used decision trees, KNN for classification, and the ANN model which gave an accuracy of 95%.

The authors of [5], AlexNet which is the oldest CNN architecture for feature extraction, and SVM classifiers were used for classification. ANN and CNN are both widely used for diagnosing skin diseases. Skin disease diagnosis using the CNN approach showed that the results are promising [6].

The authors of [7] presented a comparative analysis of different CNN architectures to detect skin bacterial infections. This paper consists of ResNet50, Xception, InceptionV3, VGG16, and VGG19. They also initially used the K-Fold cross-validation technique. Their network involved Adam optimizer and binary cross-entropy which is a loss function. The dropout rate was set to 25%. An accuracy of 91.303% and 84.545% was achieved using VGG16 and VGG19 architectures.

The authors of the paper [8] developed a web application system for classifying skin diseases using CNN architecture and the TensorFlow framework. Acne vulgaris and atopic dermatitis were some of the diseases considered in their study.

The authors of [9] presented the mathematical formulas behind the working of image segmentation and feature extraction of the image. For Feature Extraction various parameters are calculated such as mean, Variance, Energy,and Entropy from the image.

A study described in [10] combined InceptionResnetV2, MobileNet, and InceptionV3. They modified the layers of these architectures to suit their dataset for skin disease classification. The maximum voting rule was used for attaining the classification result. With an accuracy level of 88%, the authors were able to predict 20 diseases.

The authors of [11], have used a multi-model approach for skin lesion classification. They have used the transfer learning technique to utilize both ResNet-101 and ResNet-50 for feature extraction. For reducing the dimensions of the feature an algorithm called KcPCA (Kurtosis controlled Principal Component Analysis) was developed. This process ensured the selection of optimal features. For classification, SVM with a radial basis function is utilized.

Mondal et al., [12] discussed the Generative Adversarial Network (GAN) which can automatically discover and learn the patterns in input data. This way that the model can generate new samples from the original samples. GANs architecture consisting of CNNs can generate simulated data that is nearly a match for real data distribution. Features are extracted using MobileNet, GoogLeNet, ResNet, and DenseNet but DenseNet-121 and DenseNet- 169 have produced the best accuracy of 94.25% and 93.67%.

In [13], the authors used Deep Convolutional Neural Networks (DCNN) for image classification and feature extraction processes. The layers consisted of three convolutional and pooling layers, followed by a fully connected layer for feature extraction. Their work can be viewed as two sub-tasks. The first task involves the extraction of features from DCNN, and the extracted information is fed into a neural network of two layers for classification. The second task involves classification by SVM. The results of their experiment show the SVM classification on DCNN features is better than the performance of the neural network alone for the classification task. From their experimentation is understood that having several feature maps helps in discovering different patterns from the input image at several locations. The error on testing data starts to increase slowly once the number of feature maps and convolutional layers exceeds a certain limit. The training error increases after several iterations with a slightly high learning rate.

For image classification, [14] proposes a method for creating automatic CNN architectures using genetic algorithms (CNN-GA). Because of their algorithm, users with little to no experience in tuning CNN architectures can still find an appropriate CNN architecture from the provided images. For producing deeper CNNs, the authors have developed an encoding technique for the genetic algorithms that can encode CNNs having any depth. Two components for CNN-GA are designed to expedite analysis and save a lot of computing resources. Considering the computing resources consumed, the number of parameters required and the classification accuracy, they found that the proposed algorithm outperforms the existing automatic CNN architectures.

The authors of [15], have developed a new deep learning model that is named Convolutional eXtreme GradientBoosting (ConvXGB) for classification problems. Their work was carried out on several image datasets and general datasets available in the UCI repository. From an initial point of view, they considered that a single model didn't provide the amount of accuracy required for various models. As the name suggests, their model combines the Convolutional Neural Networks and the XGBoost classifier. ConvXGB also includes a module for preprocessing data. The ConvXGB model learns the input features using several convolutional layers, and XGBoost is used for class label prediction in the final layer. It lacks a pooling layer or a Fully Connected layer as opposed to the traditional CNN. This way the number of parameters is simplified in this model because the weight values do not have to be readjusted in the backpropagation cycle. As a result of the experiment, their ConvXGB model generally outperformed CNN and XGBoost alone for all data sets and, in some cases, by a significant margin.

From the literature survey, it is inferred that deep learning techniques have a greater impact on the diagnosis of skin diseases. The deep learning methods outperform the machine learning techniques for image recognition and classification since the former automates the feature engineering process. In addition, CNN architectures arepopular for the task of image processing. Also, the transfer learning technique allows using the existing architectures of CNN by either fine-tuning the layers or using the same layers with their previously learned weightsfor the classification task.

III. Preprocessing Step

Image preprocessing

All the images in the dataset need to undergo some kind of preprocessing step before it is used for training/testing the neural network model. This process simplifies the feature extraction ability of the neural network model by eliminating the irrelevant information in the image. The preprocessing consists of data cleaning and data integration. Preprocessing is an essential step that performs contrast enhancement of images. It is also important for highlighting the lesion region in comparison to the background/ healthy region.

Data Cleaning

It ensures the integrity of the data features of the images. For the skin disease images, the process involves removing the noise to reduce hair and shadow influence on skin disease identification.

Data conversion

It is a data transformation technique that converts the format of the data to a structure that is required by a deep learning model. Some of the data conversion techniques are size adjustment, normalization, grayscale conversion,and data augmentation.

Size adjustment

It modifies the size of the image as per the requirements of the model. The process ensures there is uniformity inthe size of images that are provided as input to the model.

Normalization

All the images will be put into a common statistical distribution such as [0, 1] or [-1, 1] in terms of the image sizeand pixel values. This technique enables the neural network model to learn the features faster.

Grayscale conversion

This step improves the contrast of the image and threshold processing. Unlike RGB images which have 3 channels,the Grayscale images have just one channel. If the colour of the image has no significance for the task in consideration, then it is better to convert the images to grayscale which reduces the complexity while processing the image to extract the features. For bacterial and fungal skin diseases that show colour-based symptoms, RGB images are necessary.

Image augmentation

Augmentation techniques add slightly modified copies of the existing images from the dataset. This not only increases the size of the dataset but also helps in training the neural network model better. Some of the image augmentation techniques are Image rotating, Image Flipping, and Image Shifting.

IV. Convolutional Neural Network

Designing a CNN from scratch is one of the methods that involve adding layers as per the requirements of the classification task.

Training a neural network model from scratch means that the weights of the models are randomly initialized andthe layers of the model should undergo a training process for finding the suitable weights.

A Convolutional Neural Network (CNN) is a deep learning algorithm for image classification and recognition. ACNN is designed to extract the features from the preprocessed images and then predict the class of the image through classification. CNN is suitable because it learns the important features of an image automatically through the use of filters and also captures the spatial features i.e., it captures the arrangement of pixels and theirrelationship in an image.

Convolutional Layer

This layer performs a convolution operation that consists of filters for detecting the edges or specific shapes of an image. It can also blur the unwanted noise in the image. After the convolution operation, an activation function isused to improve the nonlinearity of the model. The size of the input image is reduced after the convolutional operation. After a convolution operation, a padding operation can be performed for preserving the size of the inputimage. A hyperparameter called stride decides the amount by which the filter slides over the images during the convolution operation. Both padding and striding control the shape of the output obtained from the convolutional layers. The number of filters and the size of each filter needs to be specified for each convolutional layer. Any number of convolutional layers can be added to the network.

Pooling Layer

Pooling is an optional layer. When the size of the images is large, it may be required to reduce the number of trainable parameters for reducing the amount of computational power and time needed. Thus, the pooling operation reduces the spatial size of the convoluted image. The pooling layer detects the features that are invariantto scale or rotational changes and also prevents overfitting. Depending on the size of the image, pooling layers need to be periodically introduced between subsequent convolutional layers.

Both the convolutional layers and the pooling layers come under the process of feature extraction. The extracted features are known as feature maps. The size of the feature maps increases with the increase in the number of convolutional layers. The initial convolutional layers detect only the general features from the images such as edges, and lines. The convolutional layers present deeper in the network extract features that are more specific tothe image.

Fully Connected Layer

The feature maps obtained are passed to a flatten layer that converts the features as a 1D array and then the obtained 1D array is fed to the fully connected layer. In a fully connected layer, each neuron in the previous layer is connected to every neuron in the following layer. There can be many fully connected layers in the architectureof the neural network. A fully connected layer is nothing but the traditional neural network structure. After one ormore fully connected layers, an output layer is required to predict the output. The number of neurons in the output layer will be equal to the number of output classes. It makes use of an activation function to generate the output. For a multiclass classification problem, SoftMax is the most widely used activation function. A suitable loss function is used for computing the error in prediction. Once the forward pass is completed, the backpropagation process begins to update the weight and bias parameters for error and loss reduction. The weights are learned suchthat the loss function is minimized.

Fig. 1 shows the general architecture of CNN consisting of three layers namely: convolutional layer, pooling layer, and fully connected layer.

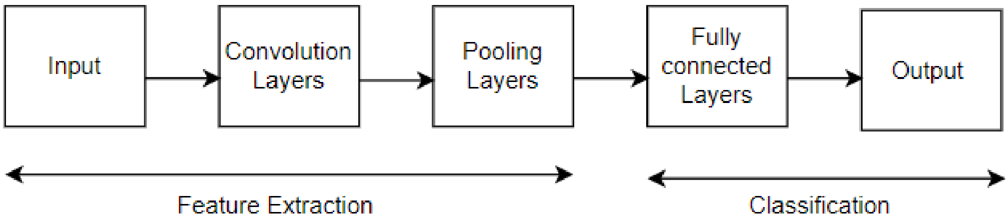

Fig 1: general architecture of cnn

V. Transfer Learning

Transfer Learning is the process of transferring the feature maps learned by one neural network that was trained on a task such as an image classification containing enough data, to a new task where the amount of data is insufficient for training. The neural network from which feature maps are utilized to solve another task is pre- trained. It means that transfer learning involves using neural networks trained on one task as the initial state for working on a related task. The transfer learning should be applied to the task in consideration only if both the tasks are similar to each other.

Transfer learning offers flexibility in the following four ways:

- The pre-trained network can be applied to a classification task without making any changes to the pre-trained network.
- The pre-trained network can be used only for the process of feature extraction.
- Some layers of the pre-trained model can be integrated into a new neural network but the layers of thepre-trained model can be frozen during the training process.
- Finally, some portion of the pre-trained network can be integrated into a new neural network, and then those layers can be retrained during the training process.

Each of the above four approaches to the transfer learning technique saves a significant amount of time in developing and training a convolutional neural network model.

VGG16 is model is one of the most widely used pre-trained CNN architecture. VGG16 consists of 21 layers but only 16 layers of those are trainable. The architecture includes 13 convolutional layers, 5 max-pooling layers, and 3 fully connected layers. Several small filters are used throughout the network with max-pooling operations being performed periodically. Instead of using large filters, VGG16 architecture considers smaller filters with more depth in the network. These small-sized filters have a size of 3x3 and 1x1 with a stride of one. Max pooling layers are used with a size of 2x2 and stride size 2. The number of filters increases as the model deepens. The model starts with 64 filters and gradually expands to 128 and 256 as the depth of the model is increased, and finally up to 512 filters when the process of feature extraction ends.

The VGG16 model was trained on a subset of the ImageNet database. Originally, ImageNet was a project designed to classify and label images into more than 20,000 distinct categories for computer vision research. For CNN anddeep learning, ImageNet refers to the Image Large Scale Visual Recognition Challenge (ILSVRC). The objectiveof ILSRVC is to build a model that can accurately classify input images into 1,000 categories. ILSVRC is a subsetof ImageNet. From each of the 1000 categories, ILSVRC contains around 1000 images. These 1000 categories represent objects that we encounter in daily life such as vehicles, household objects, cats, dogs, etc.

The task of identifying and classifying skin diseases consists of medical images. These images are quite different from the images present in the ImageNet database. Accordingly, a pre-trained network such as VGG16 is neitherappropriate for entirely using it for classification nor using the same weights for the feature extraction process. Instead, the layers of the VGG16 architecture can be fine-tuned by freezing the initial convolutional layers that extract the general features and then by retraining the remaining layers with the skin disease images.

VI. Combining CNN

A large number of experiments have shown that in deep learning, the multi-model fusion approach is better thanthe single model approach [16] where the size of the dataset was huge or the single model approach resulted in unsatisfactory results. A multi-model fusion approach learns through multiple learners and all the intermediate results are integrated to provide a result.

The Fig. 2 represents the multi-model fusion approach by combining various CNN architectures. These CNN architectures are used only as feature extractors and they can be a new neural network trained from scratch or a fine-tuned network of an existing pre-trained network. When the features of the various CNN architectures are combined, using a dimensionality reduction technique like PCA helps in reducing the dimensions. The combinedfeatures are then fed to the classifier for performing the classification task. The classifier can be a fully connectednetwork with an output layer containing as many neurons as the number of output classes or machine learning classifiers such as SVM classifier, RandomForest classifier, and XGBoost.

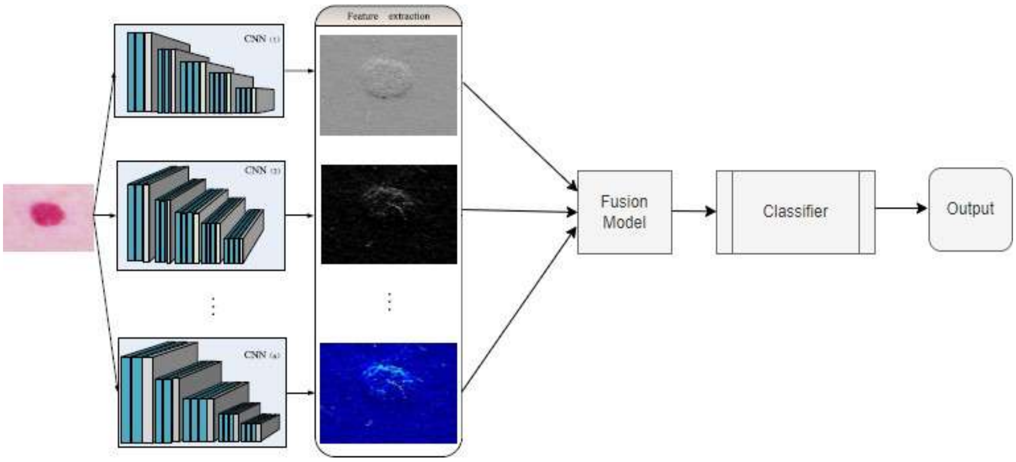

Fig 2: multi-model feature fusion

VII. General Procedure

A web app can be developed to display the predictions for providing a smooth user interface.
The overall procedure for identifying and classifying the skin diseases caused by bacteria and fungus is depicted in Fig. 3

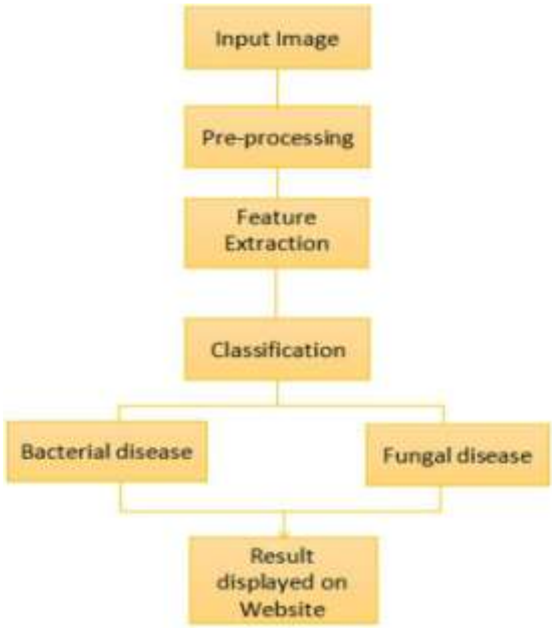

Fig 3: General Procedure

A model trained with a suitable dataset should be identifying the disease as bacterial or fungal. After identification, the model should classify the disease by specifying the name of the disease. Fig. 4 represents an expected output.

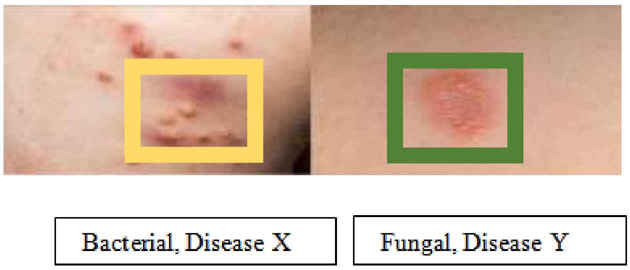

Fig 4: Expected Output

VIII. Conclusion

The various preprocessing methods and deep learning approaches for classifying skin diseases are identified through the Literature Survey. Selecting a suitable method and approach depends on the size and nature of the images in the dataset. Developing an efficient model for identifying and classifying skin diseases assists a dermatologist in decision making. Such a model reduces the time required for diagnosis and also minimizes the cost involved by a significant amount.

Acknowledgment

It is our pleasant duty to place on record our deepest sense of gratitude to our respected guide Dr. Rekha P M, who is an Associate Professor and Head of the ISE department at JSS Academy of Technical Education, Bengaluru for the constant encouragement, valuable help, and assistance in every possible way. We would like to thank our external project guides Mr. Prateek Vashistha, Subject Matter Expert, FICE, and Mr. Mohammed Ameer, SubjectMatter Expert, FICE for providing the problem statement and guiding us.

References

[1] Li, L. F., Wang, X., Hu, W. J., Xiong, N. N., Du, Y. X., & Li, B. S. (2020). Deep Learning in Skin Disease Image Recognition: A Review. IEEE Access.

[2] Mahbod, A., Schaefer, G., Wang, C., Ecker, R., & Ellinge, I. (2019, May). Skin lesion classification using hybrid deep neural networks. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1229-

1233). IEEE.

[3] Yasir, R., Rahman, M. A., & Ahmed, N. (2014, December). Dermatological disease detection using image processing and artificial neural network. In 8th International Conference on Electrical and Computer Engineering (pp. 687-690). IEEE

[4] Kumar, V. B., Kumar, S. S., & Saboo, V. (2016, September). Dermatological disease detection using image processing and machine learning. In 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR) (pp. 1-6). IEEE.

[5] ALEnezi, N. S. A. (2019). A method of skin disease detection using image processing and machine learning. Procedia Computer Science, 163, 85-92.

[6] Harangi, B. (2018). Skin lesion classification with ensembles of deep convolutional neural networks. Journal of biomedical informatics, 86, 25-32.

[7] Nayyef, S. A., Hasan, L. M., & Ibrahim, A. A. (2020, October). Using Deep Learning and Medical Data Classification for Predicting Bacterial Skin Infection. In 2020 4th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT) (pp. 1-9). IEEE.

[8] Akyeramfo-Sam, S., Philip, A. A., Yeboah, D., Nartey, N. C., & Nti, I. K. (2019). A web-based skin disease diagnosis using convolutional neural networks. International Journal of Information Technology and Computer Science, 11(11), 54-60.

[9] Manish Kumar and Rajiv Kumar, An intelligent system to diagnosis the skin disease, ARPN Journal of Engineering and Applied Sciences VOL. 11, NO. 19, OCTOBER 2016 ISSN 1819-6608.

[10] Patnaik, S. K., Sidhu, M. S., Gehlot, Y., Sharma, B., & Muthu, P. (2018). Automated skin disease identification using deep learning algorithm. Biomedical & Pharmacology Journal, 11(3), 1429.

[11] Khan, M. A., Javed, M. Y., Sharif, M., Saba, T., & Rehman, A. (2019, April). Multi-model deep neural network based features extraction and optimal selection approach for skin lesion classification. In 2019 international conference on computer and information sciences (ICCIS) (pp. 1-7). IEEE.

[12] Mondal, B., Das, N., Santosh, K. C., & Nasipuri, M. (2020, July). Improved Skin Disease Classification Using Generative Adversarial Network. In 2020 IEEE 33rd International Symposium on Computer-Based Medical Systems (CBMS) (pp. 520-525). IEEE.

[13] Bodapati, J. D., & Veeranjaneyulu, N. (2019). Feature extraction and classification using deep convolutional neural networks.Journal of Cyber Security and Mobility, 261-276.

[14] Sun, Y., Xue, B., Zhang, M., Yen, G. G., & Lv, J. (2020). Automatically designing CNN architectures using the genetic algorithm for image classification. IEEE transactions on cybernetics, 50(9), 3840-3854.

[15] Thongsuwan, S., Jaiyen, S., Padcharoen, A., & Agarwal, P. (2021). ConvXGB: A new deep learning model for classification problems based on CNN and XGBoost. Nuclear Engineering and Technology, 53(2), 522-531.

[16] XU, M., GUO, L., SONG, P., CHI, Y., DU, S., GENG, S., & ZHANG, Y. (2019). Skin Disease Recognition Method Based on Multi-Model Fusion of Convolutional Neural Network. Journal of Xi'an Jiaotong University, 53(11), 125-130.