

SKIN CANCER CLASSIFICATION: A DEEP LEARNING APPROACH

N. Nigar^{1*}, A. Wajid¹, S. Islam¹ and M. K. Shahzad²

¹Department of Computer Science, Rachna College of Engineering and Technology, Gujranwala, Pakistan

²Power Information Technology Company (PITC), Ministry of Energy, Power Division, Government of Pakistan, Lahore, Pakistan

*Corresponding author(s). E-mail(s): natasha@uet.edu.pk

ABSTRACT: Skin diseases are common in human beings because of significant changes in surrounding environments. The most of these diseases are curable if diagnosed at initial stages. Therefore, early diagnosis can spare people's precious lives. To address these issues, we proposed a novel model based on deep learning to diagnose the skin disease at a preliminary stage using classification. The developed model correctly identifies six different skin diseases namely, actinic keratosis, benign keratosis, melanoma, basal cell carcinoma, insects bite and skin acne. Several state-of-the-art algorithms are examined on benchmark datasets (International Skin Imaging Collaboration (ISIC) 2019 dataset and UCI Data Center) for accuracy, precision, recall and F1-score metrics. The results show that convolutional neural network (CNN) has a distinct superiority over its peers with accuracy rate of 97%, precision 91%, recall 91% and F1-score 91%. This system will provide skin care handling services that are precise and accurate and help the dermatologist in early diagnosis of skin diseases.

Keywords: Skin Cancer, Deep Learning, CNN, Region of Interest

(Received 14.03.2023

Accepted 17.06.2023)

INTRODUCTION

Skin diseases are common in human beings because of significant changes in surrounding environments[1]. It is the most infectious dermatological ailment prevailing worldwide. Different skin disease happens among humans due to abnormal development of tissues, or some happens genetically. If these minor ailments are not treated in time, it may lead to skin cancer. For example, most people have moles and occasionally get new ones. Most of the moles are harmless but if not treated in time, they can lead to diseases like melanoma and acne. To conduct the appropriate treatment, it is essential to identify a skin lesion correctly.

In contrast to the surrounding normal skin, the lesion is the abnormal region of skin. Its distribution, kind, colour, and form can all vary. There are 2,032 distinct lesions, of which melanocytic and non-melanocytic are most common. A lesion is classified as melanocytic or non melanocytic based on whether it contains melanocytes and melanin pigment[2].

A skin imaging method namely, dermoscopy is extensively used to increase the diagnostic performance with maximum accuracy and thereby lower the number of skin cancer fatalities [3]. With this approach, the lesion area may be seen and understood well using a magnified and illuminated skin image [4]. This technique is used to improve the medical professionals' capacity to identify skin cancer in its early stages. Typically, dermatologists

examine the dermoscopic pictures with their naked eyes, which is laborious, error-prone, and needs a high level of competence and focus [5]. The difficulty in making an appropriate selection is due to the considerable resemblance between diseased skin and normal moles. To address such issues, there is need to design an efficient computer-aided system to assist dermatologist.

In this work, a deep learning driven novel model is developed that aims early diagnosis of skin diseases. For this purpose, various state-of-the-art algorithms are compared. The developed model is trained intensively to address the problem of imbalance dataset and how they could affect the model's accuracy. In summary, the key contributions are presented below:

- i. **Design of Novel System:** A deep learning based novel system is designed for diagnosis of 6 different skin diseases in an accurate and quick manner.
- ii. **State-of-the-art Algorithms Comparative Analysis:** To diagnose skin diseases, a comparative analysis of different state-of-the-art algorithms is performed.
- iii. **Data Set:** The developed approach is evaluated on benchmark data instances i.e., ISIC 2019 [6] and UCI Data Center [7].

The rest of the paper is structured into six sections. Section 2 overviews the literature

review. The materials and methods are in section 3. The experimental analysis and results are detailed in section 4. Section 5 presents the study limitations whereas section 6 concludes this paper with possible future developments.

LITERATURE REVIEW

Background: The late diagnosis of skin diseases may lead to skin cancer which is a rapidly prevailing disease all over the world, involving multiple factors from environmental effect to genetic susceptibility. Skin cancerous cells are life-threatening abnormal regions that frequently develop in any part of the human body [5]. The recent advancements in machine learning (ML) field have allowed researchers for recent breakthroughs that results in accurate and more efficient algorithms to predict the skin disease. This study is making a substantial impact on the healthcare sector by lowering death rates through the early identification of skin cancer[8]. In this context, we evaluate five different state-of-the-art algorithms for diagnosis of skin diseases. These

algorithms were chosen because most of them are open source and are easy to implement.

A popular algorithm for analysing medical images is the CNN [9], which belongs to a class of deep neural networks that can identify and categorise significant elements from the images. The CNN basic structure is shown in Fig. 1. The VGG16 [10] is CNN extension used for object detection and classification. It is among one of the most efficient image classification algorithms, and comprised of 16 layers with weights. The major draw back associated is that it is very slow to train. Resnet-50 is an extended version of CNN with 50 layers comprising of 48 convolutional layers, one average pool layer, and one layer named as MaxPool. It is one of the best deep neural networks for classification problems with excellent performance [11].

Decision Tree (DT) [12] is the most widely used non-parametric supervised learning technique for classification and prediction. They are simple to use, offer scalability for huge datasets, and manage imbalanced datasets. However, they can be difficult to train computationally and have considerable with high variance.

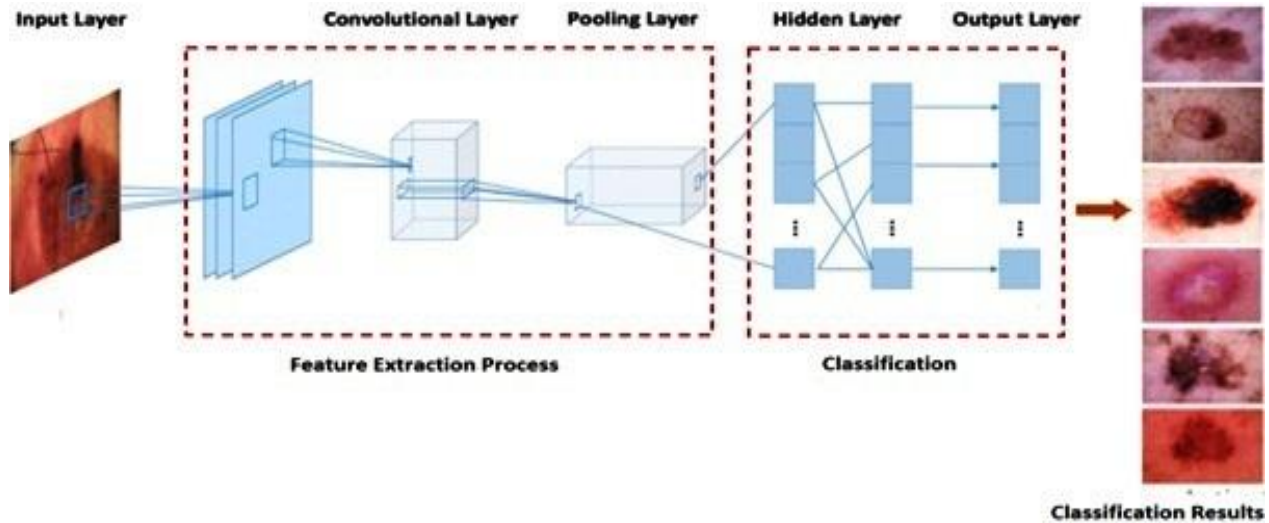


Fig. 1: The basic structure of CNN model

Random forest (RF) [13] is a popular ML algorithm and is extensively used. It is attributed with flexibility and ease of use. The resulting outcome of several decision trees is merged to provide a single value and, may deal with both classification and regression problems. Additionally, it allows feature selection so that users may see immediately which features are important.

Related Works: In this section, different studies have been presented with focus on skin disease detection. To address this, ALEnezi [14] developed an image processing-based method for the detection of skin diseases. The digital image of effected skin area is taken as input, and then image analysis is employed to identify

disease type. Their system was capable to detect 3 types of different skin diseases namely, eczema, melanoma, and psoriasis, with an accuracy rate of 100%. However, their dataset size was 100 images with few taken from the web. Also, this study limits to the detection of only 3 diseases.

Kolkur *et al.* [15] proposed a new system incorporating top 5 features i.e., contrast, correlation, energy, entropy, and homogeneity based on the findings. Bhadula *et al.* [16] used different ML algorithms i.e., Naive Bayes, Random Forest, Kernel SVM, Logistic Regression, and CNN for the detection of 3 skin diseases i.e acne, lichen planus, and Sjs Ten. Their results show

that among all the algorithms CNN results in best training precision for the selected skin diseases.

Chakraborty *et al.* [17] proposed a neural based method using skin imaging for the detection of basal cell carcinoma and skin angioma as different skin diseases, with achieved accuracy rate of 86.67%. They applied non-dominated sorting genetic algorithm - II (NSGA-II) for the training purposes of artificial neural network (ANN) by using ISIC dataset. Haddad *et al.* [18] work is based on computer aided application that detects the skin diseases based on the digital skin image and, then, filter is applied to remove noise or unwanted things for image analysis. After that image is gray scaled to get the useful information.

Ajith *et al.* [19] proposed work is based on image processing techniques for 6 different types of skin diseases diagnosis i.e., warts, tinea corporis, acne, vitiligo, nail psoriasis and eczema. The developed mobile application is also accessible in remote areas. The efficiency of this system is upto 80%. Chakraborty *et al.* [20] developed a ANN based on meta-heuristic to classify the images. They considered 3 skin diseases namely angioma, basal cell carcinoma, and lentigo simplex. They also used ISIC dataset. Their proposed method results inaccuracy rate of 87.92% along with 94.2% precision. However, it can result in a more accurate prediction model when a larger training dataset is used.

Rathod *et al.* [21] proposed an automated system for skin diseases detection using images and ML classification. Their system incorporates a computational technique to analyze, process, and relegate the image data predicated on various features of the images. There is limitation to their work that using advanced computational techniques and large dataset, results can be improved. Shanthi *et al.* [22] proposed approach utilized CNN for skin diseases recognition. Their method employs computer vision technique and identifies four types of skin diseases. The results achieved the accuracy rate of 98.6% to 99.04%. Bajwa *et al.* [23] work used deep neural net works for the identification of skin diseases. The evaluated their work using two largest skin datasets that are also publicly available i.e., Derm Net and ISIC Archive. For Derm Net, their proposed method for 23 diseases classification achieves accuracy rate of 80% and area under the curve (AUC) as 98%. For ISIC archive, achieved average accuracy of 93% average and 99% AUC to classify 7 diseases.

Rimi *et al.* [24] developed a system to diagnose the skin diseases using CNN. They achieved 73% precision of 500 pictures of various diseases using dermnet dataset. Iqbal *et al.* [25] proposed an advanced deep CNN model with the aim of skin lesions multi-class classification. The dataset is acquired from ISIC-17, ISIC-18, and ISIC-19. Their results demonstrate precision, sensitivity, and specificity rate of 94%, 93%, and 91% respectively. Mendes *et al.* [26] developed model

classifies the 12 lesions, incorporating Melanoma and Basal Cell Carcinoma (BCC). ResNet-152 architecture was trained over 3,797 images. They performed test with 956 images and achieved 0.96 AUC for Melanoma and 0.91 for BCC. In our work, we employ CNN to detect the 6 different diseases to help the dermatologist in early detection of these diseases.

MATERIALS AND METHODS

This section presents the proposed methodology to detect the six skin diseases as shown in Fig. 2. It is divided into 5 parts and explained below:

- i. Data pre-processing
- ii. Data Augmentation
- iii. Segmentation
- iv. Feature Extraction
- v. Model Training

Data Pre-processing: The dermoscopic images are subjected to noise in the form of hairs and bubbles, causing inaccuracies in classification. In this step, we pre-process the images to remove the noise, fine hair, and bubbles. For smoothing image from noise, region of interest (ROI) Open CV is used. ROI is commonly used technique in image processing to minimize the impact of small structures like thin hairs and small air bubbles. Moreover, contrast enhancement is applied to sharpen image border and for enhancing the segmentation accuracy. The standard image size is taken as 224x224 pixels by resizing the images in data set.

Each image in the training set is labelled according to respective class. For this, the data set is classified into ranges and then, label encoding and one hot encoding is applied. As we have multiple skin diseases images in our dataset. So, we have converted them into multiple classes (0 to 5) denoting 0 as actinic keratosis, 1 as melanoma, 2 as basal cell carcinoma, 3 as skin acne, 4 as benign keratosis, and 5 as insects bite. This step has a reasonable effect to enhance the classifier's correct prediction performance and reduces computation time.

Data Augmentation: To address the issue of limited dataset, data augmentation is a technique to add up in the data by adding slightly changes in the existing data set. These changes may be of height, weight, focus and angle of images. This is useful in training to avoid overfitting and improves the model's accuracy. To augment the dataset, Image Generator Python library is used to generate and balance the data classes.

Segmentation: In this step, the image is partitioned into segments to locate the infected area. Segmentation finds the ROI and removes the healthy skin from the image. We employ threshold segmentation technique for the segmentation process. In this technique, segmentation is

performed based on different extremes or colors in the foreground and background regions of the image. The input image is either grey scale or colored. Image segmentation results in a binary image, and it is

completed by scanning and labeling each pixel either as an object or background based on its binary or grey level (Fig. 3).

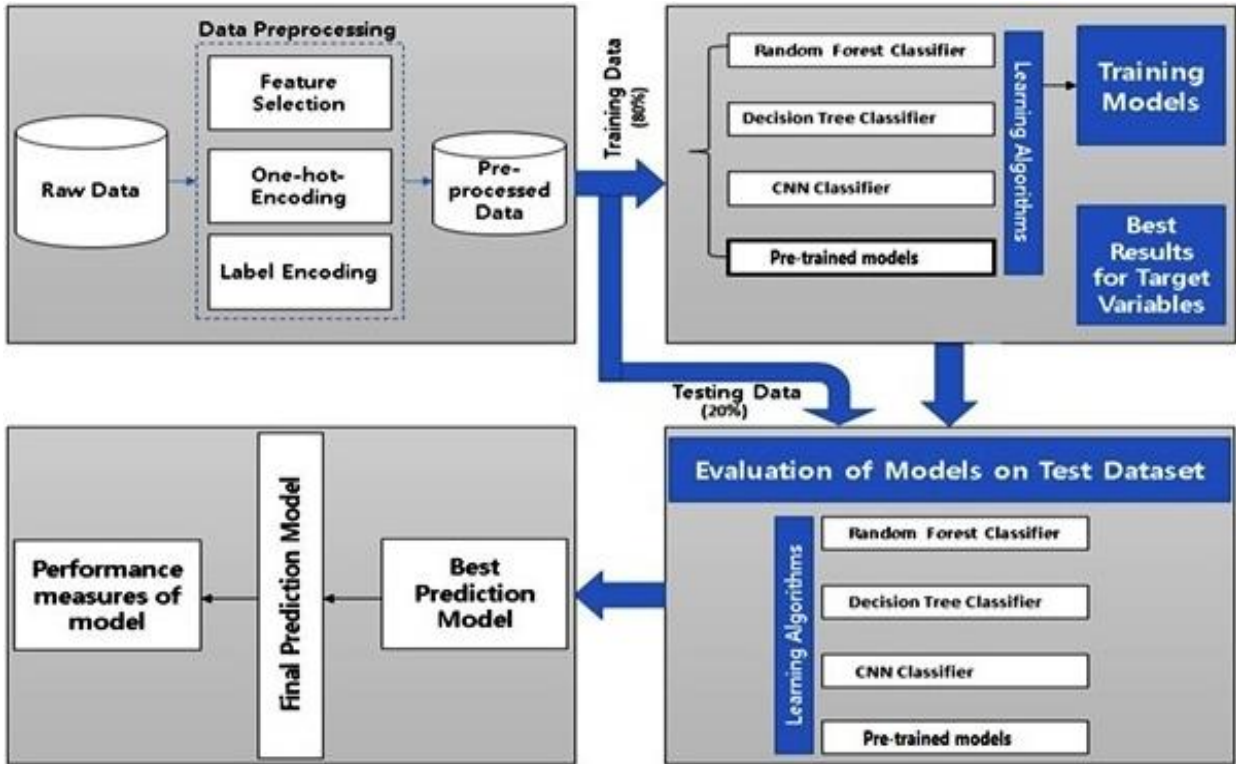


Fig. 2: The workflow of proposed methodology.

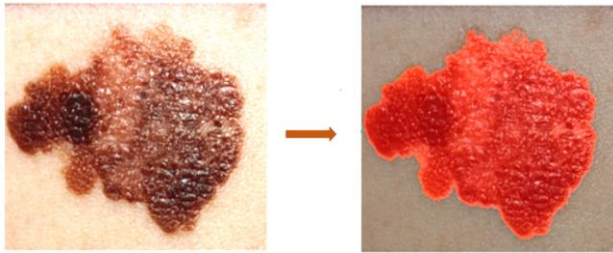


Fig. 3: Image segmentation process

Feature Extraction: In this step, we employ the feature extraction technique to reduce the background noise and some extra features of face from the image that are not needed for detection process. The feature extraction is accomplished into two processes. In first process, background noise is removed and detect the skin in the image. In this process color-based segmentation and region-based segmentation is used to extract the skin samples from the image. In second process, after the skin detection, the extra facial features i.e., eyes, lips and eyebrows from the image are removed. A deep learning library of face recognition is used for the extraction of the

facial features. In this way, features are extracted that are used in the training and testing process.

Model Training: For training purpose, the proposed system incorporates a feedforward multilayer network and backpropagation algorithm. The input layer is given input image size. Relu activation function is used in convolutional and dense layers. Pooling layers are used with (2,2) pool size. For stabilizing the learning process, batch normalization technique is employed. To avoid model over fitting, regularization functions are employed. In the final layer, softmax activation function is employed to predict multinomial probability distribution. Convolutional layers and Pooling layers are used in the feature extraction process. Hidden and output layers take part in the classification (Fig. 1). Nodes on hidden and output layers adjust the weights that depends on the error and loss rate in the classification. The back propagation process, propagates the error back, and new weights are adjusted to minimize the error rate. Gradient descent is used for the modification of weights on the neural network nodes. Weights are updated according to the gradient of the error curve which makes it more efficient in prediction as well as classifying task.

EXPERIMENTAL ANALYSIS

Dataset: In this study, skin diseases dataset is collected from multiple sources including ISIC dataset [6]. However, the whole dataset is not available publicly due to patient’s privacy. So, in this study, only those images are used which are publicly available. Some images are collected from the UCI Data Center [7]. The collection of the 30,000 images is based on 6 classes named as actinic keratosis (AKIEC), basal cell carcinoma (BCC), benign keratosis (BKL), insects bite, melanoma (MEL), and skin acne.

Performance Measures: In classifications problems, the results are categorized either positive class or negative class; the resulting quadrant is known as confusion matrix [27] as shown in Fig. 4. The results’ are evaluated using below stated performance measures:

$$Accuracy = \frac{TP + TN}{FP + TN + TP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1score = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (4)$$

		Predicted Class	
		True Positive (TP)	False Negative (FN)
True Class	False Positive (FP)	True Negative (TN)	
	True Negative (TN)		

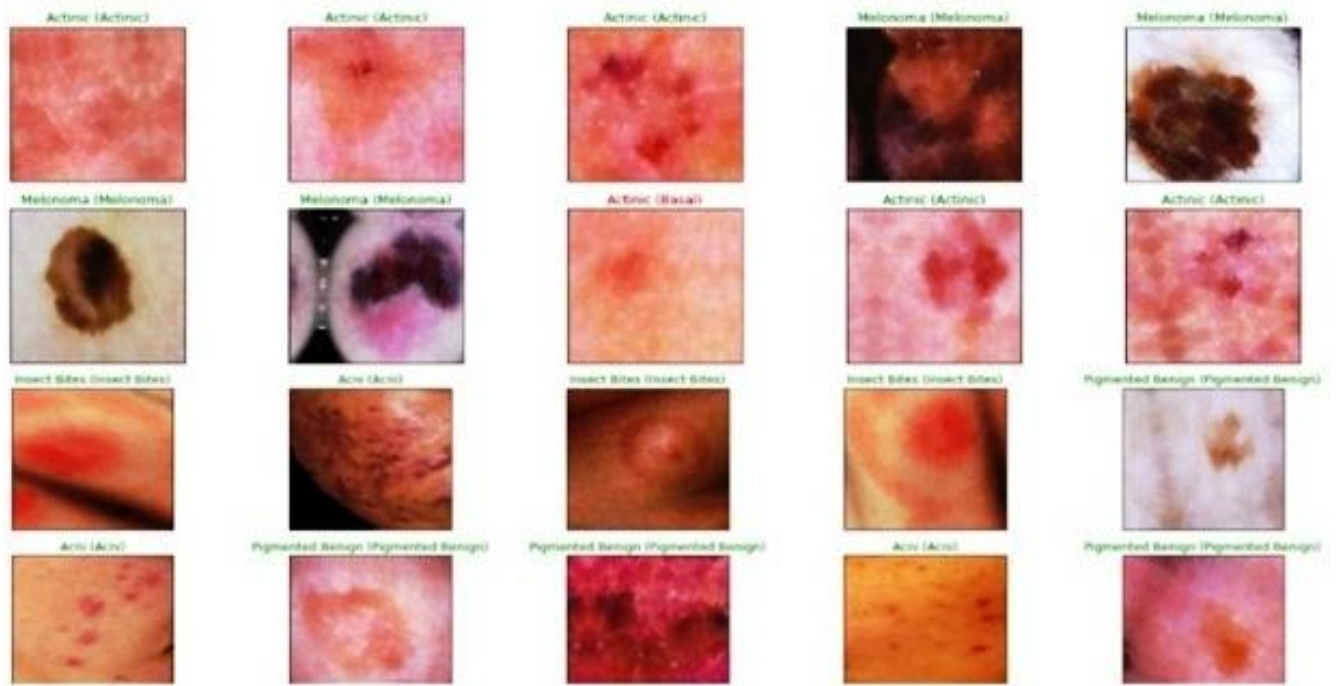


Fig. 4: Confusion matrix

Parameters: The hyperparameters values for CNN, VGG16 and Resnet-50 used for skin diseases detection are presented in Table 1. For Decision Tree and Random Forest, we employ default parameters settings using the Ginicriterion.

Table 1: Hyperparameters for the CNN and its variants classifiers.

Parameters	Values
Input size	224*224*3
Batch size	32
Error Function	Categorical Cross Entropy
Activation	ReLU
Optimizer	SDGM
Learning Rate	0.01
Dropout	0.05
Training/ Validation	&0% / 30%

Table 2: Comparison of performance metrics of ML models

	Accuracy	Precision	F1-score	Recall	Support
Random Forest	0.99	0.69	0.68	0.68	900
Decision Tree	0.51	0.49	0.5	0.5	900
Resnet-50	0.81	0.76	0.76	0.76	600
VGG16	0.33	0.16	0.16	0.16	1050
CNN	0.97	0.91	0.91	0.91	450

Table 3: Comparison of training and testing accuracy of ML models

Classifier	Training Accuracy	Testing Accuracy
Random Forest	99%	67.75%
Decision Tree	100%	49.90%
Resnet-50	81%	76.89%
VGG16	33.33%	33.33%
CNN	97%	91%

Table 4: Comparison of ML model training time

Classifier	Training Time
CNN	457 – 572 (seconds) /epoch
VGG16	2850 – 3024 (seconds) /epoch
Resnet-50	1155s – 1521 (seconds) /epoch
Decision Tree	1800 (seconds)
Random Forest	2250 (seconds)

RESULTS AND DISCUSSION

The performance of five pre-trained ML models namely, Random Forest, Decision Tree, Resnet-50, VGG16, and CNN to detect the skin disease is presented in this section. The comparative analysis of these models using performance metrics is shown in Table 2. It is

obvious from results that CNN model significantly outperforms other four models returning better values of accuracy, precision, recall and F1-score. It can be obvious from results that Random Forest returns the highest accuracy of 99%, however, other performance metrics values are less as compared to CNN. VGG16 classifier performs least as compared to its peer algorithms. We also show the training and testing accuracy achieved by each model (Table 3). Training accuracy is drawn during the model training while testing accuracy is achieved when testing the trained model. The difference between the training and testing accuracy shows about the training performance. The results depict that all the classifiers show the overfitting or under fitting behaviour except CNN. CNN shows good training as well as testing accuracy with normal overfitting.

We also compare the performance of each model based on the training time. Table 4 shows the performances of ML classifiers on chosen data set with respect to their time taken during training the model. It can be observed that CNN gives the least training time whereas Random Forest returns the highest time.

Next, we present the accuracy and loss graph of ML models to show the overfitting or under fitting behaviour on every epoch during training. Figure 5a shows that VGG16 is highly under fit on the given data during training on every epoch, whereas Resnet-50 (Fig. 5b) is highly over fit on starting epochs, training and testing accuracy is low, and loss is high at initial epoch. CNN at initial epoch over fit while as the epochs increase model’s behaviour is in the best way (Fig. 5c). We increased the number of epochs and get the accuracy up to 91% with normal overfitting.

The above comparative results indicate that CNN has the distinct superiority over its peers in predicting the skin cancer with highest accuracy. The accuracy of the selected model is 97% on training and 91% on test images. The results of predicted classes with actual classes are shown in Fig. 6.

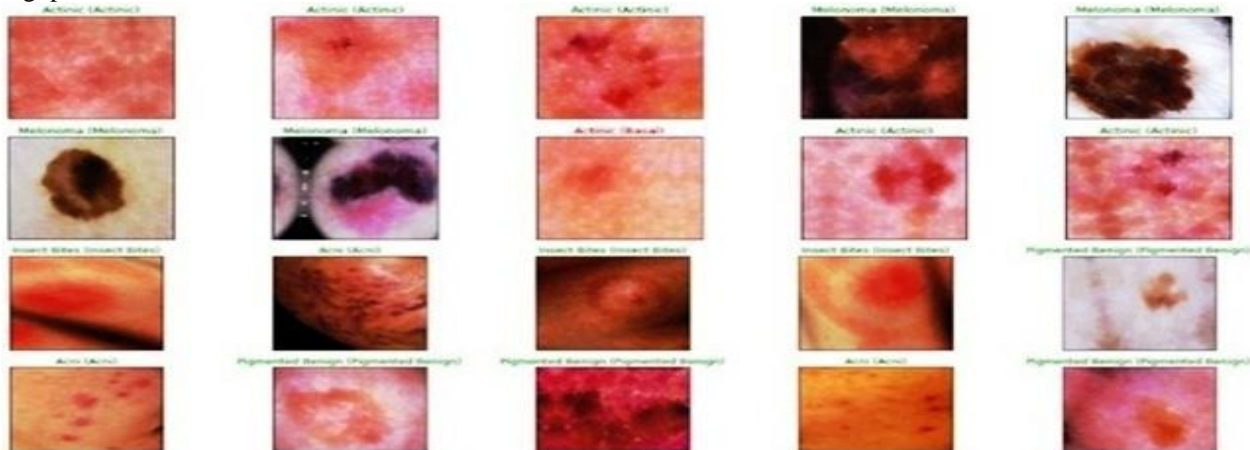


Fig. 5: Accuracy and loss graphs of ML models (VGG16, Resnet-50, CNN) for skin detection

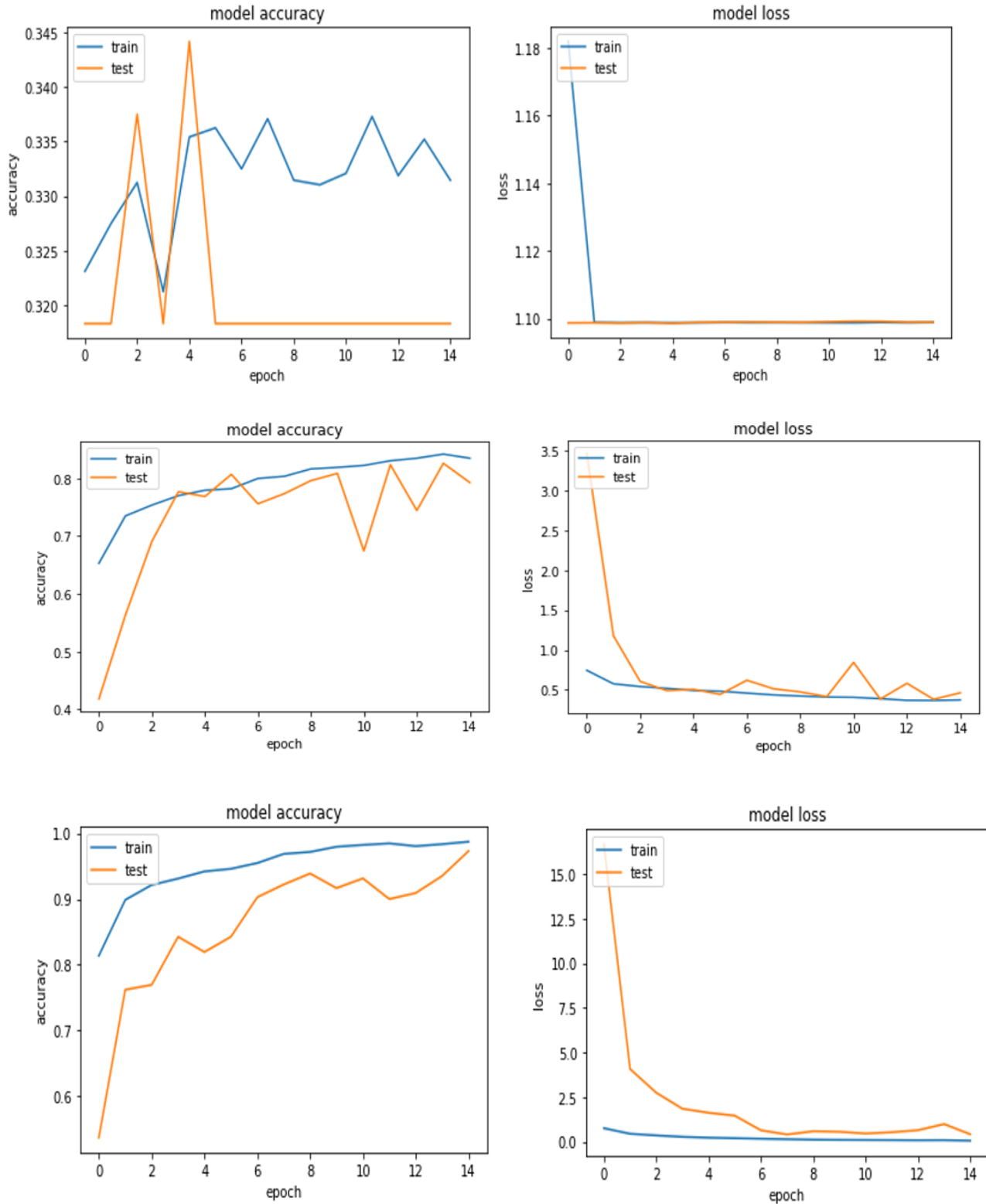


Fig. 6: CNN model predicted results

Limitations: This study helps in diagnosing the skin cancer at early stages in an efficient and accurate manner. However, this work is associated with few limitations. First, adding more real-world images in the data set could lead to better diagnosis accuracy. Second, the use of real-

world dataset may increase the developed approach's scalability and validity. In the dataset, the problem of class imbalancing by reducing the dataset size, could also affect the classifier's performance.

Conclusion: The most prevalent cancer is skin cancer, which is a serious health issue. Due to advancement in the field of ML, ML-based solutions aid in the early detection of skin cancer, reduce diagnostic costs and, thus, save many human lives. This is a great help to the both the doctors and patients using smartphone apps or websites. This study aims for innovative and creative solution for patients and skin disease situations that affect in different ways in daily life.

In this study, deep learning model CNN is trained using different datasets to diagnose the skin diseases. This model diagnoses the six different skin diseases namely, actinic keratosis, basal cell carcinoma, benign keratosis, melanoma, insects bite and skin acne efficiently. This study genuinely brings about a significant change in the lives of people. In future, the functionality could be expanded to include more skin diseases and their detection and treatment.

REFERENCES

- [1] Siegel RL, Miller KD, Goding Sauer A, Fedewa SA, Butterly LF, Anderson JC, *et al.* Colorectal cancer statistics, 2020. CA: a cancer journal for clinicians. 2020;70(3):145–164.
- [2] Barata C, Marques JS, Emre Celebi M. Deep attention model for the hierarchical diagnosis of skin lesions. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops; 2019.
- [3] Li X, Wu J, Chen EZ, Jiang H. From Deep Learning Towards Finding Skin Lesion Biomarkers. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2019.p. 2797–2800.
- [4] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. Journal of Big Data. 2019.
- [5] Thanh D, Dvoenko S. A denoising of biomedical images. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. 2015;40(5):73.
- [6] 2019 I.: Available from: <https://www.isic-archive.com/#!/topWithHeader/wideContentTop/main>.
- [7] Dua D, Graff C.: UCI Machine Learning Repository. Available from: <http://archive.ics.uci.edu/ml>.
- [8] Nigar N, Umar M, Shahzad MK, Islam S, Abalo D. A Deep Learning Approach Based on Explainable Artificial Intelligence for Skin Lesion Classification. IEEE Access. 2022;10:113715–113725.
- [9] O’Shea K, Nash R. An introduction to convolutional neural networks. arXiv preprint arXiv:151108458. 2015.
- [10] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556. 2014.
- [11] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016.
- [12] Quinlan JR. Induction of decision trees. Machine learning. 1986;1:81–106.
- [13] Breiman L. Random forests. Machine learning. 2001;45:5–32.
- [14] ALKolifi ALEnezi NS. A Method of Skin Disease Detection Using Image Processing and Machine Learning. Procedia Computer Science. 2019;163:85–92. 16th Learning and Technology Conference 2019, Artificial Intelligence and Machine Learning: Embedding the Intelligence.
- [15] Kolkur S, Kalbande DR. Survey of texture based feature extraction for skin disease detection. In: 2016 International Conference on ICT in Business Industry & Government (ICTBIG); 2016. p. 1–6.
- [16] Bhadula S, Sharma S, Juyal P, Kulshrestha C. Machine learning algorithms based skindisease detection. International Journal of Innovative Technology and Exploring Engineering (IJITEE). 2019; 9(2):4044–4049.
- [17] Chakraborty S, Mali K, Chatterjee S, AnandS, Basu A, Banerjee S, *et al.* Image based skin disease detection using hybrid neural network coupled bag-of-features. In: 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON). IEEE; 2017.
- [18] Haddad A, Hameed SA. Image Analysis Model For Skin Disease Detection: Framework. ICCCE; 2018. p. 1–4.
- [19] Ajith A, Goel V, Vazirani P, Roja MM. Digital dermatology: Skin disease detection model using image processing. In: 2017 International Conference on Intelligent Computing and Control Systems (ICICCS); 2017. p. 168–173.
- [20] Chakraborty S, Mali K, Chatterjee S, Banerjee S, Mazumdar KG, Debnath M, *et al.* Detection of skin disease using metaheuristic supported artificial neural networks. IEMECON; 2017. p. 224–229.
- [21] Rathod J, Waghmode V, Sodha A, Bhavathankar P. Diagnosis of skin diseases using Convolutional Neural Networks. In: 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA); 2018. p. 1048–1051.
- [22] Shanthi T, Sabeenian RS, Anand R. Automatic diagnosis of skin diseases using CNN.

- Microprocessors and Microsystems. 2020; 76: 103074.
- [23] Bajwa MN, Muta K, Malik MI, Siddiqui SA, Braun SA, Homey B, *et al.* Computer-aided diagnosis of skin diseases using deep neuralnet works. *Applied Sciences*. 2020;10(7):2488.
- [24] Rimi TA, Sultana N, Ahmed Foysal MF. Derm-NN: Skin Diseases Detection Using Convolutional Neural Network. *ICICCS*; 2020.p. 1205–1209.
- [25] Automated multi-class classification of skin lesions through deep convolutional neural network with dermoscopic images. *Computerized Medical Imaging and Graphics*.2021;88:101843.
- [26] Mendes DB, da Silva NC. Skin lesions classification using convolutional neural networks in clinical images. *arXiv preprint arXiv:181202316*. 2018.
- [27] Visa S, Ramsay B, Ralescu AL, Van Der Knaap E. Confusion matrix-based feature selection. *MAICS*. 2011;710(1):120–127.