Melanoma Skin Lesions Classification using Deep Convolutional Neural Network with Transfer Learning

Md. Khairul Islam
Department of Biomedical Engineering
Islamic University,
khairul.ice06@gmail.com
Kushtia-7003, Bangladesh.

Md Shahin Ali
Department of Biomedical Engineering
Islamic University,
Kushtia-7003, Bangladesh.
shahinbme.iu@gmail.com

Md Mosahak Ali
Department of Biomedical Engineering
Islamic University,
Kushtia-7003, Bangladesh.
mdmosahakali1998@gmail.com

Mst. Farija Haque
Department of Biomedical Engineering
Islamic University,
Kushtia-7003, Bangladesh.
farijahaque@gmail.com

Abhilash Arjan Das
Department of Biomedical Engineering
Islamic University,
Kushtia-7003, Bangladesh.
abarjan00@gmail.com

Md. Maruf Hossain
Department of Biomedical Engineering
Islamic University,
Kushtia-7003, Bangladesh.
marufbme.iu@gmail.com

D S Duranta
Department of Biomedical Engineering
Islamic University,
Kushtia-7003, Bangladesh.
duranta03@gmail.com

Md Afifur Rahman
Department of Biomedical Engineering
Islamic University,
Kushtia-7003, Bangladesh.
arafif0@gmail.com

2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA) | 978-1-6654-1511-8/20/$31.00 ©2021 IEEE | DOI: 10.1109/CAIDA51941.2021.9425117

***Abstract*—Skin cancer is basically the unnatural growth of skin tissues and it can be fatal. Lately, it has evolved into one of the most perilous types of other cancers in the human body. Premature detection can help to endure the patient. Detection of skin cancer is quite difficult. At present in medical image diagnosis, the performance of computer vision is quite conducive. Together with the progress in technology and impetuous increment in computer provision, different types of machine learning techniques and deep learning models have arisen for the analysis of medical images particularly skin lesion images. In this study, we propose a deep learning model with some image pre-processing steps that help to categorize skin lesions with a better classification rate than other existing models. Normalization, data reduction, and data augmentation are used in pre-processing steps to classify benign and malignant cancer lesions from the HAM10000 dataset. From the experimental result, the proposed model gained an accuracy of 96.10% in training and 90.93% during testing. This model reduces the execution time and performs well-handled.**

***Index Terms*—Skin cancer, Deep neural network, Convolutional neural network, UV rays, Feature extraction.**

I. Introduction

A. Motivation

In health and medicine, skin cancer is one of the major types of cancer found in people aged 40-60 years especially in males rather than women [1–3]. Mainly the skin cancer occurs from Ultraviolet (UV) rays. People when do not protect their skin, UV rays can damage their skin's DNA, and it can't be controlled skin cell growth and lead to cancer. Malignant melanoma skin cancer is the most dangerous cancer with the highest mortality rate of 10,000 people die from this type of cancer every year in the U.S [4].

Only early-stage skin cancer detection can reduce the mortality rate of skin cancer because it is one of the most curable cancers [5]. According to the skin cancer foundation every hour, more than 2 people die of skin cancer and more than 9500 people are diagnosed every day in the U.S. According to the European cancer information system (ECIS), 2.7 million new cases and 1.3 million deaths occurred in 2020. In 2020, it reveals that 62% of new diagnoses and 76% of death occur in peoples over 65 years old. Previous observation from all perspectives, it is an emergency issue to detect and diagnose skin cancer early. Nowadays, there have been extensive research solutions by developing computer image analysis algorithms [5] and the deep learning method is playing a major role in the early detection of skin cancer. In our study, we develop the condition of the art of Convolutional Neural Network (CNN) technique for executing the classification of images of the skin lesion with training some existing models to compare into respective cancer for getting better results with a high classification rate [6].

B. Contribution

The main contribution of this paper is described as follows:

- We propose a deep convolutional neural network (DCNN) that will detect skin cancer with a better accuracy, even if the patients are in the early stages.
- The proposed model works more accurately and efficiently than other existing deep learning models with a large dataset.
- The required execution time of our proposed model is lower than other pre-trained models like ResNet, AlexNet, VGG-16, and InceptionV4 to execute the output results.

C. Organization

The rest of this paper is represented as follows: Section II narrates the related works. Section III narrates materials and method of our proposed work. The Proposed Method is described in Section IV. Section V represents Result and Discussion and section VI represents the conclusion and future works.

II. Related Work

Salian, Abhishek C., et al. [3] proposed a custom CNN model in their study and got 80.61% of testing accuracy. However, their classification rate might be improved with another deep CNN model.

To track out melanoma skin lesions, a computer-aided model is applied with an image processing tool and the achieved result from the proposed method is quite satisfying. Since the tool is made more user friendly and pithy for images gained, it can serve out the motive of automatic skin cancer diagnosis [7].

In paper [8], the authors proposed that supervised and unsupervised classification can detect skin cancer with some machine learning techniques. For example, Support Vector Machine, K-means clustering, and Neural Network. However, the accuracy could be improved if they used more data.

In [9], the authors used image processing techniques applying Deep Convolutional Neural Networks (DCNN) for skin lesion classification and trained and made out a resource-constrained mobile-ready deep neural network architecture selecting images from several dermoscopic libraries. They used a data-set consisting of 48,373 dermoscopic images assembled from three separated archives labeled and affirmed by expert dermatologists. In addition, they used to develop a complete mobile base classifier for the detection of a lesion. However, this result will be better if they used a confidence spectrum from the deep neural network model that relied on the certainty of the classification.

Ansari, Uzma Bano, et al. [10] proposed a skin cancer detection system using Support Vector Machine algorithms for early detection of skin cancer with Gray Level Co-occurrence Matrix (GLCM) with a good results. However, they could have improved their result using some fruitful techniques.

In [11], the authors used histogram of Oriented Gradient operator to represent skin lesion. Grey Level Run Length matrix, Local Binary Patterns are used to conquer stagnation. This feature extraction model applies re-initialization mechanisms, in depth sub-dimension feature search, adaptive acceleration coefficients and multiple remote leaders. To execute the feature optimization they used two enhanced Particle Swarm Optimization models applying sine and helix model and random acceleration coefficients.

Vidya, M., and Maya V. Karki [12] applied GLCM and HOG to extract textural feature which increases quality of the skin lesion images and alleviates the unwanted things like the color of the skin, hair etc. With the help of Geodesic Active Contour (GAC), the lesion portion was segmented and then classified into two classes, labelled as benign and melanoma using SVM, KNN and Naive Bayes.

Pham, Tri Cong, et al. [13] proposed a method that uses different techniques as Hue Saturation Value (HSV), Local Binary Pattern (LBP) Histogram of Oriented Gradients (HOG) and Scale-Invariant Feature Transform (SIFT) for feature extraction. The input images were pre-processed by the Linear Normalization, Gaussian Blur and a combination of both to remove the noises. the balanced Random Forest classifier was used as a classifier to classify benign and melanoma lesions with an accuracy of 74.75%.

III. Materials and Method

A. Dataset

In this paper, we used the HAM10000 dataset for our experiment that contains 10015 skin lesion images that we took for the pre-processing, feature extraction, and finally classification. Those images are classified into two sections labeled as Benign and Malignant based on the stages of cancer. The Benign types are the early stage of skin cancer which can be cured by proper treatment. On the other hand, malignant types are the critical stage in which the surviving rate is very low. Fig. 1 shows the benign and malignant skin lesions.

B. Data Preparation

Data preparation is an important step for our workflow. To diminish unwanted deformation and elevates some image features to make better the quality of image data, preparation is done which is significant for further image pre-processing [14]. HAM10000 dataset has consisted of three labeled skin lesions like as benign, malignant, and some images are unknown.

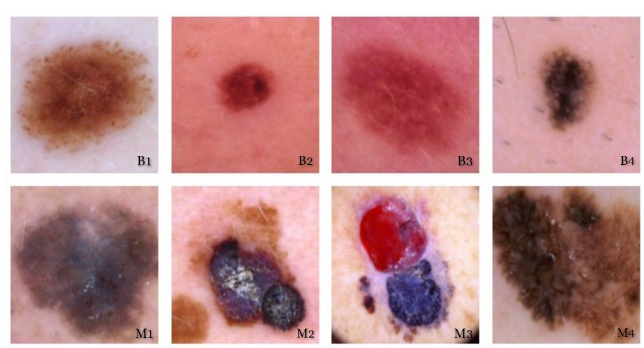

Fig. 1: B1, B2, B3, B4 are the benign and M1, M2, M3, M4 are the malignant skin lesions from the dataset respectively.

For our work purposes, we took two labels and reduced the rest. We have taken 6705 images of the benign class and 2135 images which contain malignancy. We filtered out all the images and reduced background unexpected noises and artifacts to get a better accuracy. We reduced some images that contain noise, some are lower contrast images from the main dataset to get a high classification rate. We have divided these images into three parts like 80% as training 10% as validation and 10% as testing. We have labeled images that are easily recognizable. Finally, our data is ready for the next step of data pre-processing.

C. Pre-processing

Image pre-processing is a crucial step to take aside background noises and abnormalities. Pre-processing of images also enhances the image quality [15] for a better experiment. The images are converted from benign to 0 and malignant to 1 using the label encoder function of the python library. Then we resized those images. It is easier to handle the resized images than the original images are taken from the dataset.

D. Data Normalization

Data normalization is an extremist important technique that is used to train deep neural network. It normalize inputs data and reduce undesirable characteristics and data redundancies. As a result, it is easier for us to pick the correct learning rate. However, it is difficult to pick the proper learning rate except for normalization. There are a few types of Normalization processes like first normalization, second normalization, third normalization, four normalization [16] and so on. Vanishing gradients we need to normalize the input data that will be optimized using the gradient descent function. We normalized the input data by dividing by 255 which is known as the gray scale value of an image.

E. Image Size Interpolation

Image interpolation illustrates an arbitrary continuously defined function as a discrete sum of shifted and weighted fundamental features. The perfect image interpolation model should maintain the characteristics of the image output since interpolated images passing through artifacts, as like blurring, edge discontinuation and effects in checkerboard [17]. Image size interpolation is required for the transfer learning with a DCNN on a computer vision motive. Transfer learning is a method by which another pertained deep neural network is loaded by us. In computer vision tasks, there are different types of pertained models that are trained for our purposed work. Besides, every pre-trained model has a special input vector associated with it. In this study, we took benign and malignant images with an input size of 96×96. Hence, different interpolation algorithms are used to resize or re-sample the dataset that are loaded lesion images and to fit the input size of that model [9].

F. Data Augmentation

In DCNN a large amount of images are needed for testing and training to gain prosperous classification rates. But executing this process is very ponderous especially in the case of skin cancer because it has a confined number of labeled images in the dataset [18]. The total dataset was separated into two categories labeled as Benign and Malignant. The parameters we used in our proposed model is alluded in Fig. 2

IV. Proposed Method

A. Convolutional Neural Network (CNN)

The convolution layer is the leading part of CNN which takes out the details from input images. It revealed tremendous performance over the last decades in a variety of pattern recognition concerned fields; starting from image processing to voice identification [19]. Convolution maintains the connection among pixels by realizing image aspects using tiny courts of input data. It is a mathematical strategy that requires two inputs such as an image matrix and a filter or kernel. TABLE I represents the architecture of our proposed model with its parameter values. In our study, an operation has enacted using 64 numbers of the kernel and at last, the assortment which utmost the accuracy was assigned.

B. Input Layer

This layer receives information (data), features, signals, or measurements from the exterior system. Those inputs (samples or patterns) are generally normalized into the limit values propagated by activation functions [20]. The input layer is the provider of the entire CNN. In CNN pre-analyzed images pertained to the system through the input layer. The input layer has 27648 nodes. Each node fits every segment of the image.

C. Activation Function

The activation function is significant in a CNN architecture that helps to decide which node would fire or not. In our paper, we used the Rectified Linear Unit (ReLU) activation function. The ReLU activation function has advance implementations like object finding in images, voice recognition, playing games etc. One of the specific function of a multi-layer neural network

using ReLU activation function (or ReLU network) is that every time the output will be a piece-wise input [21]. It helps us to converts all negative inputs to zero and remains positive values. It plays a major role in the execution by eliminating cancellation effects in subsequent layers. It used to separate specific stimulation and unspecific obstacle. It is faster and effective training of CNN architectures on large and complex data sets.

D. Pooling layer

Pooling is a significant concept in CNN's which avelliates spectral dissonance in the input features [22]. This layer gathers tiny rectangular blocks from the convolutional layer and specimens them to provide a single output from that block [23]. The Pooling layer is applied to make CNN faster

Data Augmentation Parameter	Parameter Value	Action
rotation_range	5	Input data generates with the rotation from -10 to 10
height_shift_range	0.3	Image is randomly shifted vertical direction by 0.2
width_shift_range	0.3	Image is randomly shifted in horizontal direction by 0.2
shear_range	0.3	Stretch the image angle slantly in degrees by a factor of 0.2
zoom_range	0.3	Zoom in or out 0.2 from the center
channel_shift_range	10	Randomly shifts channel values to variate the color
horizontal_flip	True	Flips the image horizontal direction randomly
fill_mode	nearest	Closest pixel value is chosen to fill the empty values

Fig. 2: Using Data Augmentation parameters and values

and reduces the memory size of the network. In our paper, we used a max-pooling operation that helps to diminish the image size via alleviating the number of pixels in the product from the preceding convolutional layer and it alleviates the resolution of the given output of a convolutional layer. We used 2×2 window sizes for max-polling operation.

E. Fully connected layer

In fully connected layer, all the information from one layer is united to every activation section of the subsequent layer. The features map generated by the max-pooling is maintained into the Fully Connected layer. It receives the output of the preceding layers and "flattens" them, and converts them into an individual vector that can be an input for the following layer [24].

F. Dropout Layer

During training time, the dropout layer is a technique that sets input units to 0 with a frequency of rate at each step in a random manner. Its main objectives are to prevent neural networks from over-fitting, reducing the capacity or thinning the network during training, and to do the training of a CNN faster. The accuracy will gradually be increased using dropout. Besides, the accuracy will gradually be decreased first. Finally, we noticed that accuracy is increased as we reduce the dropout parameter. So the performance will increase for lower values. The probability of dropout used in this study was 0.25.

TABLE I: Proposed CNN architecture with the parameter values

Layer Types	Size	Output Shape
Input	(96,96,3)	(94, 94, 64)
Convolution+ReLu	64(3×3) filter	(94, 94, 64)
MaxPooling	(2×2) filter	(47, 47, 64)
Convolution+ReLu	64(3×3) filter	(45, 45, 64)
MaxPooling	(2×2) filter	(22, 22, 64)
Convolution+ReLu	64(3×3) filter	(20, 20, 64)
MaxPooling	(2×2) filter	(10, 10, 64)
Convolution+ReLu	64(3×3) filter	(8, 8, 64)
MaxPooling	(2×2) filter	(4, 4, 64)
Fully-Connected+ReLu+Dropout	256 neuron	1
Sigmoid	2 way	1

G. Output Layer

The last layer of neurons in the output layer produces given outputs for the model. Output layer neurons may be built or observed in a different way and are made much like other artificial neurons in the neural network. Besides, they are the last "actor" nodes on the network. In the output layer, we used the sigmoid function because of binary classification. The sigmoid function has two important characteristics like the prediction of a model with a recognized probability and a simple derivative function. Consequently, it is used for models where we have to predict the probability as an output. So, the sigmoid function is a perfect output unit for the binary classification problems.

H. Training Details

Training in the deep neural network has consisted of two phases like forwarding phase which is passed through the network and the backward phase that is used for back-propagation to update the weights. Keras and TensorFlow are used to implement the proposed method and its related evolution. we select the binary cross-entropy cost function which is the key function. Adam Optimization with an elementary learning rate of 0.001 has been used as an optimizer to diminish the cost function. At the time of training, the rate of learning has been updated by half of its value after a sequent of a few epochs, and sometimes the improvement of our validation accuracy fluctuated. In our proposed model, we trained our model around 200 epochs including the batch size of 128 and 36 iterations per epoch. Our model has been performed on a PC under a 64-bit Windows 10 environment with an Intel Core i5-8250U CPU processor, 8 GB of RAM, and NVIDIA GeForce MX130 GPU card. Our proposed model took 9s per epoch in the training section. After finishing the training part, our model has been evaluated by using the remaining test dataset which is 20% of the whole dataset.

V. Results and Discussion

In our study, our model was trained with resized images but those were imbalanced.The results of our study are described and compared with other existing models that are trained before in TABLE II. We were successfully able to separate benign and malignant skin lesions with 96.10% of training and 90.93% of testing accuracy which is much satisfying than other schemes. Figure 4 and 5 represent the model accuracy and model loss respectively. To evaluate the performance like accuracy, F1 score, the precision and recall of each model confusion matrix is required. Accuracy estimates the proficiency of a strategy to compute the

TABLE II: Performance comparison between Proposed System and Existing deep Learning models

Model Type	Precision	Recall	F1 Score	Training Acc	Testing Acc
AlexNet	97.88	92.01	94.85	93.82	90.86
ResNet	96.00	93.63	94.80	91.25	90.93
InceptionV4	92.08	94.15	93.10	90.31	88.26
VGG-16	99.97	86.02	92.48	86.17	86.02
Proposed model	94.69	94.13	94.73	96.10	90.93

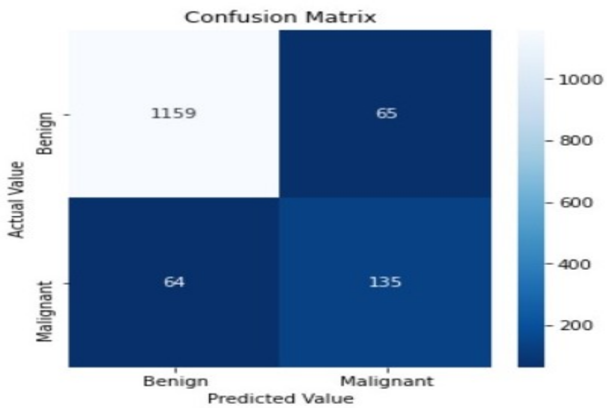

Fig. 3: Confusion matrix of our proposed model based on performance

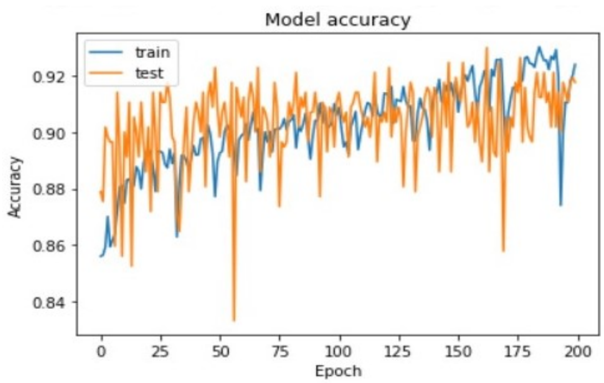

Fig. 4: Model accuracy of our proposed model

exact significance. Precision is the number of positive values achieved from the correctly predicted cases and the proportion of our model accurately distinguishing true positives is called recall. F-score is characterized as the consonant mean of the model's precision and recall. All those terms calculate through true positive(TP), true negative(TN), false positive(FP), and false-negative(FN). Where,

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

TABLE III: Comparison of computational time t over the existing models.

Models	Time per epoch	Total time
AlexNet	15-17 sec	25±1 minutes
ResNet	14-16 sec	24±1 minutes
InceptionV4	17-19 sec	29±1 minutes
VGG-16	15-16 sec	25±1 minutes
Proposed model	**9-10 sec**	**16±1 minutes**

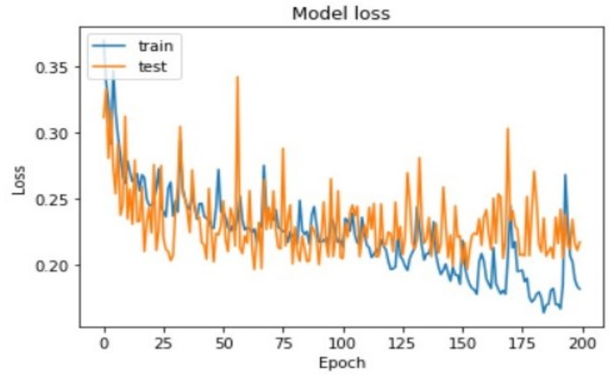

Fig. 5: Model loss of our proposed model

$$F1 = \frac{2TP}{(2TP + FP + FN)} \tag{3}$$

$$Accuracy = \frac{(TP + TN)}{(TP + FP + FN + TN)} \tag{4}$$

In this study, we compared our proposed model over the existing pre-trained models with the performance of precision, F1 score and Accuracy respectively. There are some good results from the existing models but the execution time of our proposed model is less than other existing schemes that are illustrated in TABLE III.

VI. Conclusion and Future Work

In this study, we used CNN to assemble a wide data set including benign and malignant lesions to figure out the issue of inaccurate-accurate conclusions in skin cancer detection. The recommended technique comprises images impressive for taking out the area of interest in the image itself and then expanding some images to yield an enormous dataset that encompasses a large number of images for each category. The arising dataset has been related to the CNN model to equip the model, which contains numerous layers such as convolution layers, fully connected layers, and a softmax layer. Screening the inapt state of the image, accuracy has been enhanced. Here, we got 96.10% of training and 90.93% of testing accuracy. Experimental results implied better accuracy and the mentioned procedure has an identical enactment to other skin cancer detection algorithm.

In future, we will try to examine and reformat the CNN architecture to enhance accuracy, attain better image data for training by applying recent augmentation algorithms. In order to train the model for attaining better data and the final strategy is to formulate this model as efficient and use-able for smartphone applications.

Acknowledgment

This work was supported by the Department of Biomedical Engineering (BME), Islamic University, Kushtia-7003, Bangladesh.

References

[1] N. Alfed, F. Khelifi, A. Bouridane, and H. Seker, "Pigment network-based skin cancer detection," in *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*. IEEE, 2015, pp. 7214–7217.

[2] P. Bumrungkun, K. Chamnongthai, and W. Patchoo, "Detection skin cancer using svm and snake model," in *2018 international workshop on advanced image technology (IWAIT)*. IEEE, 2018, pp. 1–4.

[3] A. C. Salian, S. Vaze, P. Singh, G. N. Shaikh, S. Chaparieri, and D. Jayaswal, "Skin lesion classification using deep learning architectures," in *2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA)*. IEEE, 2020, pp. 168–173.

[4] R. Mishra and O. Daescu, "Deep learning for skin lesion segmentation," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2017, pp. 1189–1194.

[5] M. A. Kadampur and S. Al Riyaee, "Skin cancer detection: applying a deep learning based model driven architecture in the cloud for classifying dermal cell images," *Informatics in Medicine Unlocked*, vol. 18, p. 100282, 2020.

[6] H. Nahata and S. P. Singh, "Deep learning solutions for skin cancer detection and diagnosis," in *Machine Learning with Health Care Perspective*. Springer, 2020, pp. 159–182.

[7] S. Jain, N. Pise *et al.*, "Computer aided melanoma skin cancer detection using image processing," *Procedia Computer Science*, vol. 48, pp. 735–740, 2015.

[8] H. Mhaske and D. Phalke, "Melanoma skin cancer detection and classification based on supervised and unsupervised learning," in *2013 International conference on Circuits, Controls and Communications (CCUBE)*. IEEE, 2013, pp. 1–5.

[9] A. Ech-Cherif, M. Misbhauddin, and M. Ech-Cherif, "Deep neural network based mobile dermoscopy application for triaging skin cancer detection," in *2019 2nd International Conference on Computer Applications & Information Security (ICCAIS)*. IEEE, 2019, pp. 1–6.

[10] U. B. Ansari and T. Sarode, "Skin cancer detection using image processing," *Int Res J Eng Technol*, vol. 4, no. 4, pp. 2875–2881, 2017.

[11] T. Y. Tan, L. Zhang, and C. P. Lim, "Intelligent skin cancer diagnosis using improved particle swarm optimization and deep learning models," *Applied Soft Computing*, vol. 84, p. 105725, 2019.

[12] M. Vidya and M. V. Karki, "Skin cancer detection using machine learning techniques," in *2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)*. IEEE, 2020, pp. 1–5.

[13] T. C. Pham, G. S. Tran, T. P. Nghiem, A. Doucet, C. M. Luong, and V.-D. Hoang, "A comparative study for classification of skin cancer," in *2019 International Conference on System Science and Engineering (ICSSE)*. IEEE, 2019, pp. 267–272.

[14] S. Jayade, D. Ingole, and M. D. Ingole, "Skin cancer detection using gray level co-occurrence matrix feature processing," in *2020 5th International Conference on Devices, Circuits and Systems (ICDCS)*. IEEE, 2020, pp. 49–53.

[15] A. N. Hoshyar, A. Al-Jumaily, and A. N. Hoshyar, "The beneficial techniques in preprocessing step of skin cancer detection system comparing," *Procedia Computer Science*, vol. 42, pp. 25–31, 2014.

[16] M. M. I. Rahi, F. T. Khan, M. T. Mahtab, A. A. Ullah, M. G. R. Alam, and M. A. Alam, "Detection of skin cancer using deep neural networks," in *2019 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*. IEEE, 2019, pp. 1–7.

[17] A. Amanatiadis and I. Andreadis, "A survey on evaluation methods for image interpolation," *Measurement Science and Technology*, vol. 20, no. 10, p. 104015, 2009.

[18] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Skin cancer classification using deep learning and transfer learning," in *2018 9th Cairo International Biomedical Engineering Conference (CIBEC)*. IEEE, 2018, pp. 90–93.

[19] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*. Ieee, 2017, pp. 1–6.

[20] I. N. Da Silva, D. H. Spatti, R. A. Flauzino, L. H. B. Liboni, and S. F. dos Reis Alves, "Artificial neural network architectures and training processes," in *Artificial neural networks*. Springer, 2017, pp. 21–28.

[21] K. Eckle and J. Schmidt-Hieber, "A comparison of deep networks with relu activation function and linear spline-type methods," *Neural Networks*, vol. 110, pp. 232–242, 2019.

[22] T. N. Sainath, A.-r. Mohamed, B. Kingsbury, and B. Ramabhadran, "Deep convolutional neural networks for lvcsr," in *2013 IEEE international conference on acoustics, speech and signal processing*. IEEE, 2013, pp. 8614–8618.

[23] U.-O. Dorj, K.-K. Lee, J.-Y. Choi, and M. Lee, "The skin cancer classification using deep convolutional neural network," *Multimedia Tools and Applications*, vol. 77, no. 8, pp. 9909–9924, 2018.

[24] I. Razzak, G. Shoukat, S. Naz, and T. M. Khan, "Skin lesion analysis toward accurate detection of melanoma using multistage fully connected residual network," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–8.