Machine Learning with Applications 5 (2021) 100036

Contents lists available at ScienceDirect

Machine Learning with Applications

journal homepage: www.elsevier.com/locate/mlwa

An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models

Md Shahin Ali [a], Md Sipon Miah [b,c], Jahurul Haque [a], Md Mahbubur Rahman [b], Md Khairul Islam [a,*]

[a] Department of Biomedical Engineering, Islamic University, Kushtia 7003, Bangladesh
[b] Department of Information and Communication Technology, Islamic University, Kushtia 7003, Bangladesh
[c] School of Computer Science, National University of Ireland Galway, H91 TK33 Galway, Republic of Ireland

ARTICLE INFO

Keywords:
Skin cancer
Pre-processing
Convolutional neural network
Classification
Transfer learning

ABSTRACT

Skin cancer is one of the top three perilous types of cancer caused by damaged DNA that can cause death. This damaged DNA begins cells to grow uncontrollably and nowadays it is getting increased speedily. There exist some researches for the computerized analysis of malignancy in skin lesion images. However, analysis of these images is very challenging having some troublesome factors like light reflections from the skin surface, variations in color illumination, different shapes, and sizes of the lesions. As a result, evidential automatic recognition of skin cancer is valuable to build up the accuracy and proficiency of pathologists in the early stages. In this paper, we propose a deep convolutional neural network (DCNN) model based on deep learning approach for the accurate classification between benign and malignant skin lesions. In preprocessing we firstly, apply filter or kernel to remove noise and artifacts; secondly, normalize the input images and extract features that help for accurate classification; and finally, data augmentation increases the number of images that improves the accuracy of classification rate. To evaluate the performance of our proposed, DCNN model is compared with some transfer learning models such as AlexNet, ResNet, VGG-16, DenseNet, MobileNet, etc. The model is evaluated on the HAM10000 dataset and ultimately we obtained the highest 93.16% of training and 91.93% of testing accuracy respectively. The final outcomes of our proposed DCNN model define it as more reliable and robust when compared with existing transfer learning models.

1. Introduction

1.1. Motivation

Cancer is an extremist life threat to human life. It may sometimes cause certain death to the human. Different types of cancer may exist in the human body and skin cancer is one of the fastest-growing cancers that can cause death. It is provoked by some factors like smoking, alcohol usage, allergies, infections, viruses, physical activity, environmental change, exposure to ultraviolet (UV) light, and so on. The DNA inside the skin cells can be annihilated by the radiation of UV rays from the sun. In addition, unusual swellings of the human body are also a cause of skin cancer.

There are four most frequent types of skin cancer like Actinic keratoses, Basal cell carcinoma, Squamous cell carcinoma, and Melanoma (Dorj, Lee, Choi, & Lee, 2018). The world health organization (WHO) predicts that one in every three cancers diagnosed is skin cancer (Pacheco & Krohling, 2019). The number of people who are diagnosed as skin cancer patients has been increasing at a fairly constant rate over the last decades in the USA, Canada, and Australia (Canadian Cancer Society's Advisory Committee on Cancer Statistics, 2014; Cancer Council Australia, 2018; World Health Organization (WHO), 2019).

It is assumed that approximately 5.4 million cases of skin cancer are to be diagnosed in the United States per year. The demand for quick and effective clinical screenings is arising day by day (Rogers, Weinstock, Feldman, & Coldiron, 2015; Stern, 2010). Malignant melanoma is a powerful type of malignancy, which starts from melanocytes situated in the epidermis of the skin (Monisha, Suresh, Bapu, & Rashmi, 2019). This kind of malignancy may have started to encompass rapidly and be harder to treat. As a result, the early detection of skin cancer may lead to diagnosis and treatment with increasing the chances of lives. Over the last decades, there are different types of computer-aided diagnosis

* Corresponding author.
E-mail addresses: shahinbme.iu@gmail.com (M.S. Ali), m.miah1@nuigalway.ie (M.S. Miah), jahurul@bme.iu.ac.bd (J. Haque), mrahman@ict.iu.ac.bd (M.M. Rahman), khairul.ice06@gmail.com (M.K. Islam).
URLs: https://www.iu.ac.bd (M.S. Ali), https://www.iu.ac.bd, http://www.nuigalway.ie (M.S. Miah), https://www.iu.ac.bd (J. Haque), https://www.iu.ac.bd (M.M. Rahman), https://www.iu.ac.bd (M.K. Islam).

https://doi.org/10.1016/j.mlwa.2021.100036
Received 19 February 2021; Received in revised form 16 April 2021; Accepted 19 April 2021
Available online 29 April 2021

(CAD) systems that are proposed to identify skin cancer. Traditional computer vision algorithms are mainly used as a classifier to extract a large number of features like shape, size, color, and texture in order to detect cancer. Nowadays, artificial intelligence (AI) has become an aptitude to face these problems. The most authorized deep-learning architectures such as deep neural networks (DNN), convolutional neural networks (CNN), long short term memory (LSTM), recurrent neural networks (RNN) are using in the medical field to detect cancer cell. These models are also successfully performed in classifying skin cancer. Moreover, CNN is a DNN particularly has already been achieved remarkable results in this field. CNN is the most popular model which is the collections of machine learning algorithms (Ray & Chaudhuri, 2021) used for feature learning and object classification (Nugroho, Slamet, & Sugiyanto, 2019). In addition, transfer learning is being used in these fields in large data sets to improve the results accurately.

1.2. Contribution

The main contribution of our paper is described as follows:

- We propose a deep convolutional neural network (DCNN) model that classifies skin cancer more accurately even if the patients are in an early stage.
- The result of our proposed DCNN model achieves much better accuracy than other existing deep learning (DL) models with a large dataset.
- The required execution time of our proposed DCNN model is very much lower than existing transfer learning models like AlexNet, ResNet, VGG-16, DenseNet, and MobileNet to execute the output results.
- Transfer learning models are also evaluated on the same dataset to compare our proposed DCNN model and ultimately we get the best classification accuracy from our proposed model.

1.3. Organization

The rest of this paper is represented as follows: Section 2 represents the literature review. Section 3 represents the challenges of skin cancer detection. Materials and methods are described in Section 4. Transfer learning is described with our proposed DCNN model in Section 5. Section 6 represents training and performance of our proposed DCNN model. Section 7 represents result and discussion and the conclusion and future works are summarized in Section 8.

2. Literature review

Melanoma is one type of skin cancer that are responsible to create malignant tumor on skin. Skin cancer is detected by using dermatological photographs. Machine learning based on the high-performance image is used to detect skin cancer that achieved good efficiency in identification (Srividhya, Sujatha, Ponmagal, Durgadevi, Madheshwaran, et al., 2020). However, the accuracy of the model can be increased by extracting more features and sensitivity is more deviated. In (Hoshyar, Al-Jumaily, & Hoshyar, 2014), the author proposed a method using image processing steps that helps to increase the detection accuracy of skin cancer. However, they could not narrate a specific model that can efficiently detect cancer. In an another study, the author proposed the architecture driven model that utilized a DL algorithm for the detection of skin cancer. DL based on model-driven architecture can be built so quickly that is why the model can predict the result as quickly. It acquired a better result in detecting of skin cancer (Kadampur & Al Riyaee, 2020). However, the methodology requires real-time interfacing with medical images so that it can improve the medical field.

In (Hasan, Barman, Islam, & Reza, 2019), the author proposed CNN based skin cancer detection where the feature is extracted from dermoscopic images using feature extracting techniques. They achieved an accuracy of detection 89.5% in the testing phase. However, the accuracy of detection was not sufficient that was needed to improve. In addition, over fitting was existed between their testing and training phase that was a weakness in that research. In (Li & Shen, 2018), the author proposed a lesion indexing network (LIN) based on DL to detect and classification of skin cancer. They acquired excellent result with DL based LIN by extracting more features. However, segmentation performance was needed to increase for further enhancing of result.

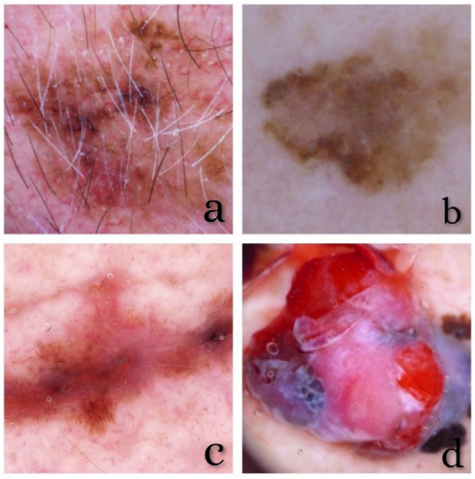

Fig. 1. Challenges of skin lesions detection: a: hair artifacts, b: low contrast, c: irregular boundaries, d: color illumination.. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In (Tschandl et al., 2019), the author used CNN to detect skin cancer from pigmented melanocytic lesions from dermoscopic images. However, it was difficult to screen non-melanocytic and non-pigmented skin cancer. Also, it acquired a lower accuracy of detection. In (Saba, Khan, Rehman, & Marie-Sainte, 2019), the author proposed a DCNN that incorporates three steps that perform excellently to detect skin lesions where firstly, color transformation is used to enhance contrast; secondly, CNN approach is used to extract lesion boundaries; finally, transfer learning is used to extract deep features. However, the method achieved a good result for some times of dataset but results may vary for a different dataset. In (Jafari et al., 2016), the author proposed a CNN based model to detect melanoma skin cancer where they used preprocessing and post-processing of image for the enhancement before and after segmentation, respectively. The model produced lesion regions by combining local and global contextual information. It acquired a good result for the prediction and classification. However, execution time is not mentioned which can increase the value of the results.

Le et al. proposed a transfer learning model which is known as ResNet50 without preprocessing steps or feature handcraft selection (Le, Le, Ngo, & Ngo, 2020) with better accuracy. Precision, recall, and F1 score are not enough high but by taking the preprocessing steps the results may be improved with a better classification rate of skin lesions images.

3. Challenges of skin cancer detection

There are some difficulties to detect skin cancer that can be attributed to variation in image types and sources (Al-Masni, Al-Antari, Choi, Han, & Kim, 2018; Aljanabi, Özok, Rahebi, & Abdullah, 2018). The variation in the appearance of human skin color makes skin cancer detection harder and complex (Naji, Jalab, & Kareem, 2019). These challenges are visually illustrated in Fig. 1 and the most visual characteristics of skin lesions images are described below:

1. The main difficulties in skin cancer are the several sizes and shapes of the images that cannot give an accurate result of identification. In this perspective, pre-processing is required for accurate analysis.
2. A few wasted signals are to be compromised that are not originally part of an image but can interrupt to get a good result. So, in the pre-processing steps, all these noise and artifacts should be removed.
3. Sometimes, low contrast from neighboring tissues poses additional difficulties and makes it harder to analyze skin cancer precisely.
4. Color illumination also makes some difficulties having its factors like color texture, light rays, and reflections.
5. The human body exists some moles which may never develop cancer cell but they create some difficulties to detect skin cancer accurately from cancerous images.
6. The current bias that distorts the performance of the models to achieve a better result is another challenge regarding skin cancer detection.

4. Materials and methods

In this section, we narrate the following steps and the flowchart of our methodology shown in Fig. 2. DCNN and transfer learning models are used for the classification task between benign and malignant lesions. Some additional layers are used in our proposed DCNN model. In addition, we use some transfer learning models like AlexNet, ResNet, VGG-16, DenseNet, and MobileNet on this same dataset for the performance comparison.

4.1. Dataset

Our dataset is simply a collection of different types of skin cancer images. To apply DL approaches it is necessary a large amount of data that makes a good result. However, the collection of skin cancer images are much critical. Besides, it is one of the main concerns to apply DL algorithms for the lack of training data. To overcome these problems, we utilized a freely accessible dataset of 10015 dermoscopy pictures acquired from Australian and Austrian patients named HAM10000. There exist 6705 of the images that are benign, 1113 are malignant, and 2197 images are unknown lesions in this dataset. The ground truth for this dataset was affirmed by one or the other pathology, master agreement, or confocal microscopy. The HAM10000 data utilized in our paper are made carefully out of melanocytic lesions that are biopsy-demonstrated and explained as malignant or benign (Esteva et al., 2017).

4.2. Data preparation

In data preparation, we transform raw data into a more suitable form to create an effective model. Our dataset contains sets of images comparing to similar lesions but from various perspectives or, various images of the same lesions on the same person. Some images were removed from the test and validation sets that are visually blurry and far-away but were still used in training. There are some factors in data preparation such as data cleaning, feature selection, data transforms, feature engineering, dimensionality reduction, etc that finalize the dataset for training.

4.3. Preprocessing

Data preprocessing is the first and crucial step in preparing the raw data and making it compatible with a machine learning model (Garg, Maheshwari, & Shukla, 2019). The main objective of preprocessing our dataset is to enhance the original medical images by removing air bubbles, noise, and artifacts which is caused by applied gel before capturing images (Alfed, Khelifi, Bouridane, & Seker, 2015). In our proposed study, we have removed noise and artifacts from the images to get a high classification rate. Besides, we have also applied data reduction, normalization of data, feature extraction, and finally, converted the string data of label into numerical data.

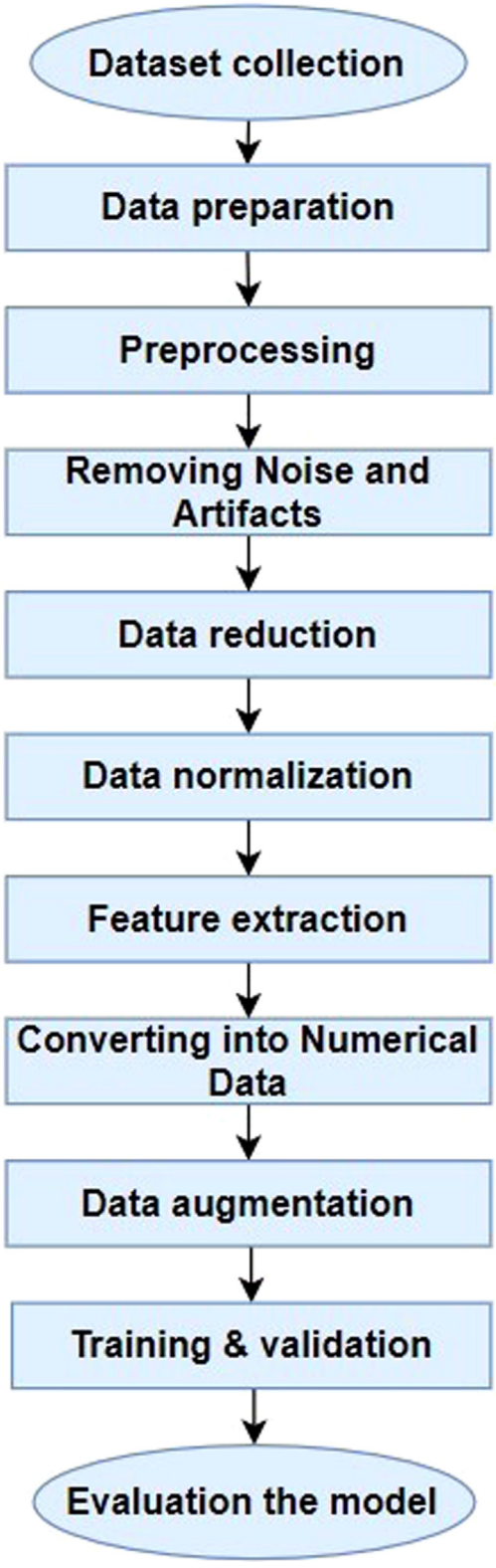

Fig. 2. The following steps of our proposed methodology.

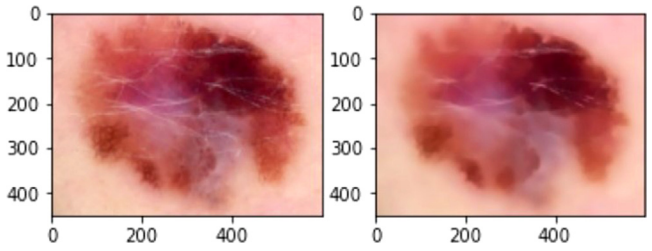

Fig. 3. Removing noise and artifacts from the input image.

4.4. Reflective noise and artifacts removal

There are various types of techniques for image smoothing like Gaussian blurring, median blurring and so on to remove even a small quantity of noise. For the removal of noises and artifacts reflection from the images, a simple thresholding algorithm is used. Fig. 3 represents the removal of noise from the original input image.

Each of the single-pixel (x, y) can be detected and classified as a reflection artifact according to the following condition:

$$\{I(x, y > T_{R1})\} and \{(I(x, y) - I_{avg}(x, y)) > T_{R2}\} \tag{1}$$

where I denotes the image with single-pixel (x, y), $I_{avg}(x, y)$ denotes the average intensity of the pixel's neighborhood which is computed by the local mean filter with dimensions i.e., 12×12 and the threshold values which are obtained by experimental i.e., $T_{R1} = 0.87$ and $T_{R2} = 0.096$, respectively.

When the areas of artifacts are recognized, an imprinting activity is applied likewise to the data of their identified pixels area, staying away from the impact of the artifacts. It may be accomplished in another approach. For this, consider a pixel that has noises, mathematically,

$$T_p = T_{p0} + n \tag{2}$$

where T_{p0} = true value of pixel and n= the noise in that pixel. We get $T_p = T_{p0}$ since the mean of noise and artifacts are zero.

4.5. Data reduction

In imaging dataset, data reduction is the process of reducing the number of images from the original data into a smaller volume. It is a great challenge to get the best classification rate from the overall dataset due to having many images with noise and artifacts, some blur images, some images that have low contrast, some images that have mole beside the wounds, and some images have color illumination (Jeong, Lu, Huo, Vidakovic, & Chen, 2006; Namey, Guest, Thairu, & Johnson, 2008). In our proposed DCNN model, we manually deducted some images that have these characteristics. After deducting the images we considered 6136 images of benign lesions and 979 images of malignant lesions.

4.6. Normalization of data

Data normalization is the way toward planning the database that diminishes data redundancy, data uprightness, and takes out unfortunate characteristics like insertion, update, and deletion anomalies. There are a few existing techniques of normalization such as min–max normalization, z-score normalization, and decimal scaling normalization (Rao et al., 2008). The main objectives of data normalization are-

- It is the association of data to seem comparable across all records and fields.
- It enhances the cohesion and coherence of entry types leading to purifying, lead generation, segmentation, and higher quality of data.

It can be normalized the dataset by dividing 255 which is known as the gray scale value of an image. But in our study, we normalized our dataset in z-score normalization (Patro & Sahu, 2015) as per given formula:

$$V_i' = \frac{V_i - \bar{E}}{std(E)} \tag{3}$$

where V_i'= z-score normalized values, v_i is the value of row E of ith column.

$$std(E) = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^{n} (V_i - \bar{E})^2} \tag{4}$$

and mean value,

$$\bar{E} = \frac{1}{(n)} \sum_{i=1}^{n} V_i \tag{5}$$

Generally, for the manipulation of data like scale down, it scales the values in the range of $[0 - 1]$ and ensures the data is stored logically.

4.7. Feature extraction

Feature extraction is a technique that plays a vital role in image processing to make it more manageable groups for further processing. In our research, we extract a huge amount of features that helps to identify and recognize the pattern of a large number of dataset (Hobbs, Hillson, Lawande, & Smith, 2011). In addition, it selects and combines variable to extract features that diminish the number of resources without any information loss of raw data (Kumar & Bhatia, 2014).

4.8. Working with numerical data

In machine learning techniques for the imputation, the most frequent types of data that are going to deal with are numerical values (Biessmann, Salinas, Schelter, Schmidt, & Lange, 2018). We maintain some procedure for getting numerical values that hold various scale for each features. Moreover, it needs to be normalized and standardized of that data for better processing in the training and model supports with a wide range of DL network typologies (Köster et al., 2017). In our experiment, the LabelEncoder function is used that is provided by the python standard library and converted the two labels of benign and malignant into 0 and 1.

4.9. Data augmentation

Data augmentation is a strategy that is used for artificially increasing the amount of data by adding slightly modified copies without actually collecting new data from existing training data. By either data warping or oversampling artificially enhance the training dataset size or it helps the model to protect from over-fitting from the root of the problem (Shorten & Khoshgoftaar, 2019). To overcome this over-fitting, we augmented our data by using different augmentation parameters such as rotation, random cropping, mirroring, and color-shifting using principle component analysis (Ech-Cherif, Misbhauddin, & Ech-Cherif, 2019). The parameters that we used for the data augmentation of our dataset are illustrated in Table 1.

5. Transfer learning with our proposed DCNN model

Transfer learning is a machine learning technique in where the pre-trained model can be reused on different but related problems (Mehrotra, Ansari, Agrawal, & Anand, 2020). A large number of images are required to train a new DNN model (Hosny, Kassem, & Foaud, 2018). Unfortunately, there are no available datasets with thousands of labeled skin lesion images. It is a challenging task to train a large size of medical datasets like ImageNet using all the parameters that are existed in the neural networks. However, we have trained AlexNet, ResNet,

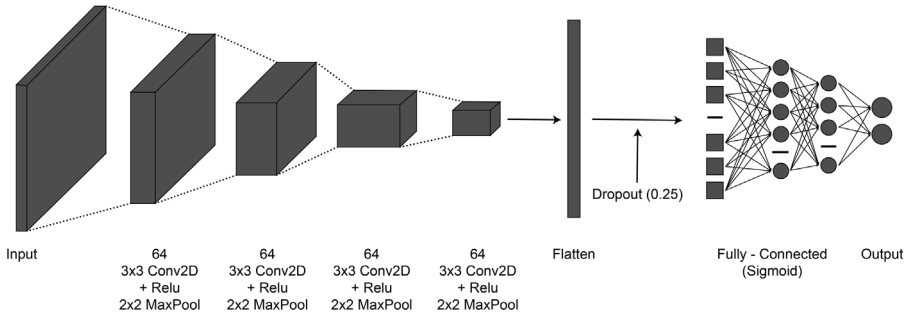

Fig. 4. Model architecture of our proposed DCNN model.

Table 1
Using data augmentation parameters and values.

Data augmentation parameter	Parameter value	Action
Rotation_range	10	Input data generates with the rotation from −10 to 10
Height_shift_range	0.2	Image is randomly shifted vertical direction by 0.2
Width_shift_range	0.2	Image is randomly shifted in horizontal direction by 0.2
Shear_range	0.2	Stretch the image angle slantly in degrees by a factor of 0.2
Zoom_range	0.2	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

VGG16, DenseNet, MobileNet, and our proposed DCNN model that we proposed in our study with a large medical dataset. In ImageNet, the value of the last layer which is used for the classification is 1000 whilst 10 in AlexNet, and 2 in our proposed DCNN model for the classification between malignant and benign skin lesions. So we used a sigmoid layer instead of a softmax. To gain the updated weights in a better way classification the back-propagation function is used for the fine-tuning. Adam optimizer is used as a gradient descent algorithm in our proposed DCNN model. At last, an augmentation step is applied to overcome the limitations of the labeled images from the dataset. Fig. 4 describes the architecture of our proposed DCNN model.

6. Training and performance

There are several parameters that are used to be tuned up in our proposed DCNN model. However, the principal ones we have used in our experiment (with their chosen values) are given underneath. This configuration is compared with several different types of configurations and deep neural network models that are tested before.

- The pooling layer (MaxPooling2D): is used to reduce the size of input images for faster training.
- Batch size (128): the number of processed images in every iteration.
- Initial learning rate (0.001): sets up with which rate will begin the learning procedure.
- Validation frequencies (36): the iterations number between evaluations of the validation metrics.
- Optimizer (Adam): Adam is a substitution optimization algorithm for stochastic gradient descent in order to minimize the loss function for training DL models. In our research, we selected the Standard Gradient Descent algorithm with a Momentum of 0.999.
- Loss function (binary cross-entropy): it is used to set up a binary classification problem from the dataset.
- Number of epochs (100): the number of times the using dataset is passed to the DNN.

Both the plain and the progressive models were recently trained and fine-tuning was performed on the HAM10000 dataset to move their insight into the skin diseases classification problems. In our research, we split the dataset into three sets such as training, validation, and testing in order to fine-tuning and evaluate the proposed DCNN model. The dataset is divided into two ways like 70% as the training, 20% as the validation, with 10% as the testing set and 80% as the training, 10% as the validation, with 10% as the testing set for the better comparison between transfer learning models and our proposed DCNN model. The model was implemented in jupyter notebook and Google Colab (12 GB RAM) on a system with Intel Core i5-8250U CPU processor, 8 GB of RAM, and NVIDIA GeForce MX130 GPU card. In this field, the performance of every model is getting compared with the existing neural network based on Accuracy, Precision, Recall, F-measure, and execution time. In addition, we have shown the confusion matrix in Fig. 5 which is a table of suggested output for measuring the model performances. There are four important terms in the confusion matrix:

- **True Positives (TP):** The occurrence in which we anticipated true and the real yield was likewise true.
- **True Negatives (TN):** The occurrence in which we anticipated false and the real yield was likewise false.
- **False Positives (FP):** The occurrence in which we anticipated true and the real yield was likewise false.
- **False Negatives (FN):** The occurrence in which we anticipated false and the real yield was likewise true.

Precision: It is known as the number of right positive results partitioned by the number of positive outcomes anticipated by the classifier.

$$Precision = \frac{TP}{TP + FP} \tag{6}$$

Recall: It is the quantity of right sure outcomes divided by the number of every conjugate samples (all samples that ought to have been identified as sure).

$$Recall = \frac{TP}{TP + FN} \tag{7}$$

F1-score: It is also known as harmonic mean that measures to seek a balance between precision and recall. It takes both false positives and false negatives for the computation and performs well on an imbalanced dataset.

$$F1 - score = \frac{2TP}{2TP + FP + FN} \tag{8}$$

Accuracy: It refer to the ratio of numbers of correct predictions to the total number of input samples.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

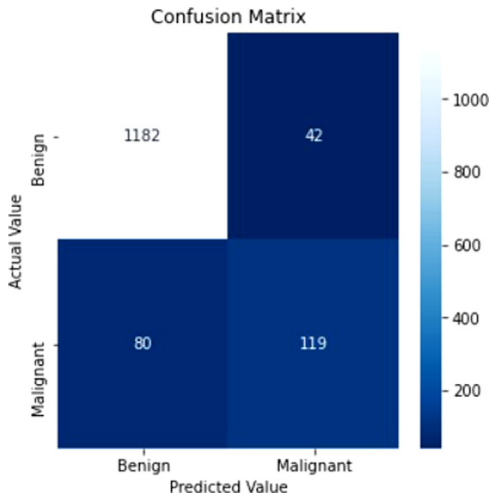

Fig. 5. Regarding performance, confusion matrix of the proposed DCNN model.

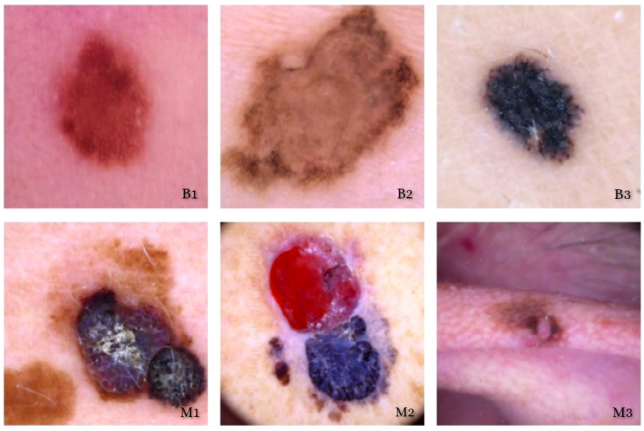

Fig. 6. B1, B2, B3 are the benign and M1, M2, M3 are the malignant skin lesions respectively.

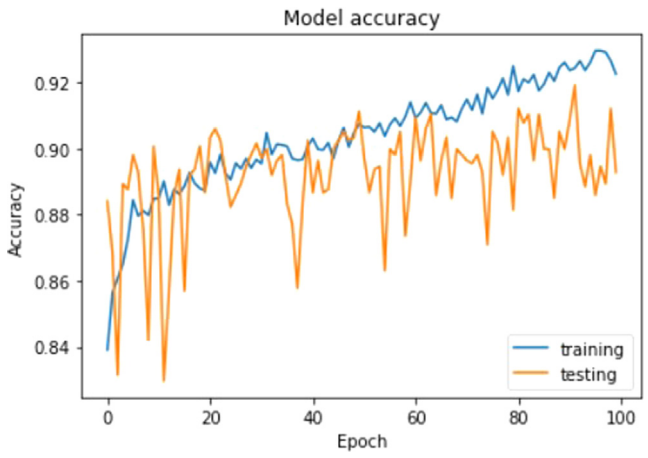

Fig. 7. Performance curve based on accuracy per epoch of our proposed DCNN model.

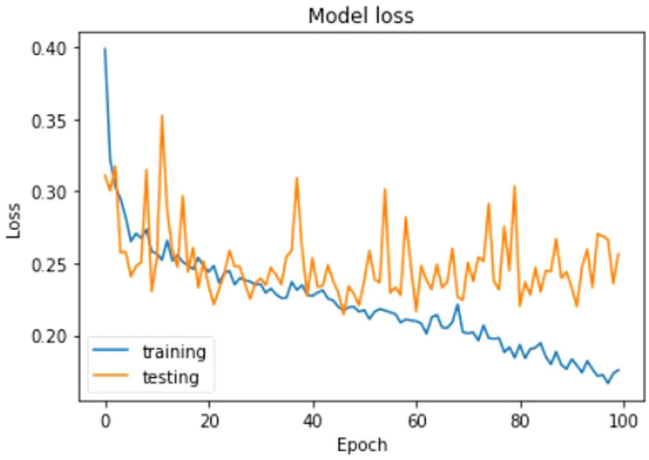

Fig. 8. Performance curve based on loss per epoch of our proposed DCNN model.

7. Result and discussion

The main purpose of our propose model is to classify benign and malignant skin lesions from the dataset that are extracted by using DCNN. The results are obtained from the total number of images that we took after the data reduction step and all the images are collected from the Internet (http://www.isic-archive.com). Some benign and malignant images are shown in Fig. 6.

In this study, we performed the experiment with our proposed DCNN model in two ways with 70% of training images, and another is aimed to evaluate the proposed model with 80% of training images that shows the best accuracy. In addition, we also performed some existing DNN models like AlexNet, ResNet, VGG-16, DenseNet, and MobileNet on the same dataset but got the best classification rate from our proposed DCNN model. Fig. 7, and Fig. 8 represent the accuracy and loss between training and testing steps respectively obtained from our proposed DCNN model.

The results of our proposed DCNN model are compared with some transfer learning models with two types of split data that are shown in the Tables 2 and 3.

The detailed discussion of the proposed DCNN model is conducted based on numerical performance and visual results. In this model, we got such a good classification rate with a big help of image processing techniques that are described in Section 4.3 to Section 4.9. We also performed our proposed model with 200 epochs but at that time the model was getting over-fitted. After several times of fine-tuning, we got the best results with 100 epochs and obtained the area under curve (AUC) value of 0.847 that is shown in Fig. 9. The proposed DCNN model can be deployed in a CAD system that will help effectively to classify the skin lesions at the early stage. In addition, the early detection of malignant growth in skin especially for those without admittance to doctors can significantly encourage them to get the treatment and enrich the survival possibility.

8. Conclusion and future work

In this paper, the proposed DCNN model achieves a better classification rate when compared to other transfer learning models. The ability of the proposed method is to classify benign and malignant skin lesions by replacing the output activation layer with sigmoid for the binary classification. Moreover, the proposed method was evaluated on a dataset named HAM10000 in where we obtained a better for training, and testing accuracy, respectively; than other existing transfer learning models. In addition, the imbalanced dataset and absences of a large number of images interrupted the model to acquire better accuracy. As a result, we balanced the dataset for both level that improves accuracy of classification. We also trained some transfer learning models on same dataset where the acquired result is not better than our proposed DCNN model. In some cases, transfer learning models were performed well but they took a high execution time per epochs than our proposed DCNN model that are illustrated in Table 4. To the best aspect of our

Table 2
The performance comparison between our proposed DCNN model and transfer learning models taking 70% of training, 20% of validation and 10% of testing data.

Model	Precision	Recall	F1 score	Training Acc	Testing Acc
AlexNet (Han, Zhong, Cao, & Zhang, 2017)	96.89	90.70	93.70	92.05	88.81
ResNet (Targ, Almeida, & Lyman, 2016)	84.94	97.50	90.79	92.78	85.20
VGG-16 (Guan et al., 2019)	88.27	95.18	91.59	88.83	86.09
DenseNet (Carcagnì et al., 2019)	91.00	91.75	91.37	91.36	85.25
MobileNet (Sinha & El-Sharkawy, 2019)	84.62	94.57	89.32	92.93	82.62
Proposed DCNN model	94.63	93.91	94.27	92.69	**90.16**

Table 3
The performance comparison between our proposed DCNN model and transfer learning models taking 80% of training, 10% of validation and 10% of testing data.

Model	Precision	Recall	F1 score	Training Acc	Testing Acc
AlexNet (Han et al., 2017)	97.88	92.01	94.85	93.82	90.86
ResNet (Targ et al., 2016)	96.00	93.63	94.80	91.25	90.93
VGG-16 (Guan et al., 2019)	1.00	86.02	92.48	86.17	86.02
DenseNet (Carcagnì et al., 2019)	91.42	94.83	93.09	91.64	88.33
MobileNet (Sinha & El-Sharkawy, 2019)	92.08	94.15	93.10	90.31	88.26
Proposed DCNN model	96.57	93.66	95.09	93.16	**91.43**

Table 4
Comparison of computational time t over the existing models.

Model	Time per epoch	Total time
AlexNet (Han et al., 2017)	14-16 s	24±1 min
ResNet (Targ et al., 2016)	13-15 s	23±1 min
VGG-16 (Guan et al., 2019)	14-15 s	24±1 min
DenseNet (Carcagnì et al., 2019)	16-18 s	28±1 min
MobileNet (Sinha & El-Sharkawy, 2019)	11-12 s	19±1 min
Proposed DCNN model	**9-10 s**	**16 ± 1 min**

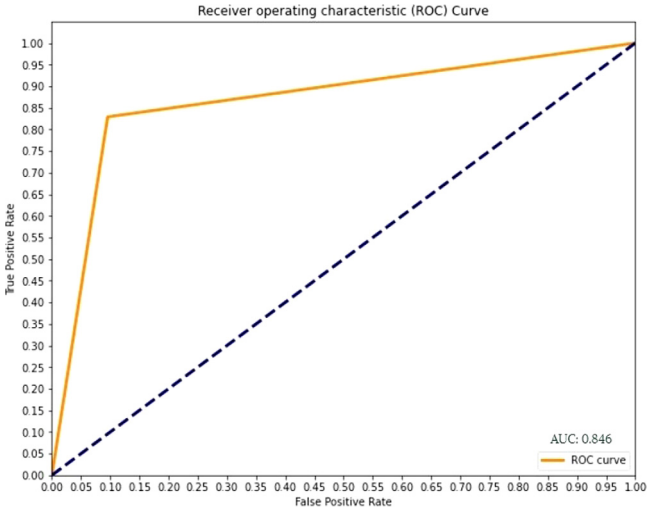

Fig. 9. Receiver Operating Characteristic curve of our proposed DCNN model.

knowledge, we are not acquainted with any previous work available that was done in the classification between benign and malignant skin lesions on same dataset.

In future, we will work on a large dataset with more labeled skin lesions with building up a successful DNN including pre-processing steps to gain the best prediction and classification accuracy. Skin cancer can also be diagnosed with the CAD which is user-friendly and robust for any conditions of acquired images. So we will also try to make a DNN which may detect various types of skin lesions using CAD systems.

CRediT authorship contribution statement

Md Shahin Ali: Conceptualization, Methodology, Software, Writing - original draft. **Md Sipon Miah:** Writing - review & editing. **Jahurul Haque:** Data curation, Writing - review & editing. **Md Mahbubur Rahman:** Writing - review & editing. **Md Khairul Islam:** Methodology, Software, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported in part by the Department of Biomedical Engineering (BME), Islamic University, Kushtia 7003, Bangladesh, and in part by the department of Information and Communication Technology (ICT), Islamic University, Kushtia-7003, Bangladesh.

References

Al-Masni, M. A., Al-Antari, M. A., Choi, M.-T., Han, S.-M., & Kim, T.-S. (2018). Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Computer Methods and Programs in Biomedicine*, *162*, 221–231.

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

Aljanabi, M., Özok, Y. E., Rahebi, J., & Abdullah, A. S. (2018). Skin lesion segmentation method for dermoscopy images using artificial bee colony algorithm. *Symmetry*, *10*(8), 347.

Biessmann, F., Salinas, D., Schelter, S., Schmidt, P., & Lange, D. (2018). " Deep" learning for missing value imputationin tables with non-numerical data. In *Proceedings of the 27th ACM international conference on information and knowledge management* (pp. 2017–2025).

Canadian Cancer Society's Advisory Committee on Cancer Statistics (2014). Canadian cancer statistics 2014 - special topic: Skin cancers. Retrieved from https://www.cancer.ca/statistics. (Accessed 30 December 2020).

Cancer Council Australia (2018). Understanding skin cancer - a guide for people with cancer, their families and friends. Retrieved from https://www.cancer.org.au/about-cancer/types-of-cancer/skin-cancer.html. (Accessed 30 December 2020).

Carcagnì, P., Leo, M., Cuna, A., Mazzeo, P. L., Spagnolo, P., Celeste, G., et al. (2019). Classification of skin lesions by combining multilevel learnings in a densenet architecture. In *International conference on image analysis and processing* (pp. 335–344). Springer.

Dorj, U.-O., Lee, K.-K., Choi, J.-Y., & Lee, M. (2018). The skin cancer classification using deep convolutional neural network. *Multimedia Tools and Applications*, *77*(8), 9909–9924.

Ech-Cherif, A., Misbhauddin, M., & Ech-Cherif, M. (2019). Deep neural network based mobile dermoscopy application for triaging skin cancer detection. In *2019 2nd international conference on computer applications & information security (ICCAIS)* (pp. 1–6). IEEE.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115–118.

Garg, R., Maheshwari, S., & Shukla, A. (2019). Decision support system for detection and classification of skin cancer using CNN. arXiv preprint arXiv:1912.03798.

Guan, Q., Wang, Y., Ping, B., Li, D., Du, J., Qin, Y., et al. (2019). Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study. *Journal of Cancer*, *10*(20), 4876.

Han, X., Zhong, Y., Cao, L., & Zhang, L. (2017). Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification. *Remote Sensing*, *9*(8), 848.

Hasan, M., Barman, S. D., Islam, S., & Reza, A. W. (2019). Skin cancer detection using convolutional neural network. In *Proceedings of the 2019 5th international conference on computing and artificial intelligence* (pp. 254–258).

Hobbs, L., Hillson, S., Lawande, S., & Smith, P. (2011). *Oracle 10g data warehousing*. Elsevier.

Hoshyar, A. N., Al-Jumaily, A., & Hoshyar, A. N. (2014). The beneficial techniques in preprocessing step of skin cancer detection system comparing. *Procedia Computer Science*, *42*, 25–31.

Hosny, K. M., Kassem, M. A., & Foaud, M. M. (2018). Skin cancer classification using deep learning and transfer learning. In *2018 9th Cairo international biomedical engineering conference (CIBEC)* (pp. 90–93). IEEE.

Jafari, M. H., Karimi, N., Nasr-Esfahani, E., Samavi, S., Soroushmehr, S. M. R., Ward, K., et al. (2016). Skin lesion segmentation in clinical images using deep learning. In *2016 23rd international conference on pattern recognition (ICPR)* (pp. 337–342). IEEE.

Jeong, M. K., Lu, J.-C., Huo, X., Vidakovic, B., & Chen, D. (2006). Wavelet-based data reduction techniques for process fault detection. *Technometrics*, *48*(1), 26–40.

Kadampur, M. A., & Al Riyaee, S. (2020). Skin cancer detection: applying a deep learning based model driven architecture in the cloud for classifying dermal cell images. *Informatics in Medicine Unlocked*, *18*, Article 100282.

Köster, U., Webb, T., Wang, X., Nassar, M., Bansal, A. K., Constable, W., et al. (2017). Flexpoint: An adaptive numerical format for efficient training of deep neural networks. In *Advances in neural information processing systems* (pp. 1742–1752).

Kumar, G., & Bhatia, P. K. (2014). A detailed review of feature extraction in image processing systems. In *2014 fourth international conference on advanced computing & communication technologies* (pp. 5–12). IEEE.

Le, D. N., Le, H. X., Ngo, L. T., & Ngo, H. T. (2020). Transfer learning with class-weighted and focal loss function for automatic skin cancer classification. arXiv preprint arXiv:2009.05977.

Li, Y., & Shen, L. (2018). Skin lesion analysis towards melanoma detection using deep learning network. *Sensors*, *18*(2), 556.

Mehrotra, R., Ansari, M., Agrawal, R., & Anand, R. (2020). A transfer learning approach for AI-based classification of brain tumors. *Machine Learning with Applications*, *2*, Article 100003.

Monisha, M., Suresh, A., Bapu, B. T., & Rashmi, M. (2019). Classification of malignant melanoma and benign skin lesion by using back propagation neural network and ABCD rule. *Cluster Computing*, *22*(5), 12897–12907.

Naji, S., Jalab, H. A., & Kareem, S. A. (2019). A survey on skin detection in colored images. *Artificial Intelligence Review*, *52*(2), 1041–1087.

Namey, E., Guest, G., Thairu, L., & Johnson, L. (2008). Data reduction techniques for large qualitative data sets. *Handbook for Team-Based Qualitative Research*, *2*(1), 137–161.

Nugroho, A. A., Slamet, I., & Sugiyanto (2019). Skins cancer identification system of HAM10000 skin cancer dataset using convolutional neural network. *AIP Conference Proceedings*, *2202*, Article 020039.

Pacheco, A. G., & Krohling, R. A. (2019). Recent advances in deep learning applied to skin cancer detection. arXiv preprint arXiv:1912.03280.

Patro, S., & Sahu, K. K. (2015). Normalization: A preprocessing stage. arXiv preprint arXiv:1503.06462.

Rao, Y., Lee, Y., Jarjoura, D., Ruppert, A. S., Liu, C.-g., Hsu, J. C., et al. (2008). A comparison of normalization techniques for microRNA microarray data. *Statistical Applications in Genetics and Molecular Biology*, *7*(1).

Ray, A., & Chaudhuri, A. K. (2021). Smart healthcare disease diagnosis and patient management: Innovation, improvement and skill development. *Machine Learning with Applications*, *3*, Article 100011.

Rogers, H. W., Weinstock, M. A., Feldman, S. R., & Coldiron, B. M. (2015). Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012. *JAMA Dermatology*, *151*(10), 1081–1086.

Saba, T., Khan, M. A., Rehman, A., & Marie-Sainte, S. L. (2019). Region extraction and classification of skin cancer: A heterogeneous framework of deep CNN features fusion and reduction. *Journal of Medical Systems*, *43*(9), 289.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of Big Data*, *6*(1), 60.

Sinha, D., & El-Sharkawy, M. (2019). Thin mobilenet: An enhanced mobilenet architecture. In *2019 IEEE 10th annual ubiquitous computing, electronics & mobile communication conference (UEMCON)* (pp. 0280–0285).

Srividhya, V., Sujatha, K., Ponmagal, R., Durgadevi, G., Madheshwaran, L., et al. (2020). Vision based detection and categorization of skin lesions using deep learning neural networks. *Procedia Computer Science*, *171*, 1726–1735.

Stern, R. S. (2010). Prevalence of a history of skin cancer in 2007: results of an incidence-based model. *Archives of Dermatology*, *146*(3), 279–282.

Targ, S., Almeida, D., & Lyman, K. (2016). Resnet in resnet: Generalizing residual architectures. arXiv preprint arXiv:1603.08029.

Tschandl, P., Rosendahl, C., Akay, B. N., Argenziano, G., Blum, A., Braun, R. P., et al. (2019). Expert-level diagnosis of nonpigmented skin cancer by combined convolutional neural networks. *JAMA Dermatology*, *155*(1), 58–65.

World Health Organization (WHO) (2019). How common is the skin cancer?. Retrieved from https://www.who.int/uv/faq/skincancer/en/index1.html. (Accessed 30 December 2020).