Journal of Theoretical and Applied Information Technology
15th May 2020. Vol.98. No 09

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

EMPIRICAL STUDY AND ENHANCEMENT ON DEEP TRANSFER LEARNING FOR SKIN LESIONS DETECTION

NOUR ELDEEN M. KHALIFA[1], MOHAMED LOEY[2], AHMED A. MAWGOUD[3], MOHAMED HAMED N. TAHA[4]

[1]Information Technology Department, Faculty of Computers and Artificial Intelligence, Cairo University, Giza, Egypt.

[2]Computer Science Department, Faculty of Computers and Artificial Intelligence, Benha University, Benha, Qalyubia, Egypt.

[3]Information Technology Department, Faculty of Computers and Artificial Intelligence, Cairo University, Giza, Egypt.

[4]Information Technology Department, Faculty of Computers and Artificial Intelligence, Cairo University, Giza, Egypt.

E-mail: [1]nourmahmoud@cu.edu.eg, [2]mloey@fci.bu.edu.eg, [3]aabdelmawgoud@pg.bu.edu.eg
[4]mnasrtaha@cu.edu.eg

ABSTRACT

Skin cancer is the most common type of cancer. One in every three cancers diagnosed is a skin cancer according to skin cancer foundation statistics globally. The early detection of this type of cancer would help in raising the opportunities of curing it. The advances in computer algorithms such as deep learning would help doctors to detect and diagnose skin cancer automatically in early stages. This paper introduces an empirical study and enhancement on deep transfer learning for skin lesions detection. The study selects different pre-trained deep convolutional neural network models such as resnet18, squeezenet, google net, vgg16, and vgg19 to be applied into two different datasets. The datasets are MODE-NODE and ISIC skin lesion datasets. Data augmentation techniques have been adopted in this study to enlarge the total number of images in the datasets to be 5 times larger than the original datasets. The adopted augmentation techniques make the DCNN models more robust and prevent overfitting. Moreover, seven accredited performance matrices in deep learning have been used to conclude an optimal selection of the most appropriate DCNN model that fits the nature of the skin lesions datasets. The study concludes that vgg19 is the most appropriate DCNN according to testing accuracy measurement and achieved 98.8%. The seven performance matrices strengthen this result. Also, a comparative result was introduced with related works. The vgg19 overcomes the related work in terms of testing accuracy and the performance matrices on both datasets. Finally, the vgg19 model was trained on a smaller number of images than the related work by 10 times, which proved that the choice of data aug-mentation techniques played an important role in achieving better results. That would reflect on reducing the training time, memory consumption and the calculation complexity.

Keywords: *Cancer, Skin Cancer, Melanoma, Deep Transfer Learning, Convolutional Neural Network.*

1. INTRODUCTION

Skin lesions cancer is one of the most fatal and malignancy types of cancer diseases [1]. Skin lesions cancer belongs to cancer family which includes breast and colon cancer [2] which are popular than it around the world. According to the American Cancer Society, about 99,550 new cases of estimated new skin cancer cases and deaths in the United States in 2018. Essentially, melanoma and other nonepithelial skin are the most known skin cancer types [3]. In Egypt, skin cancer is very common among elderly males due to sun exposure. Skin cancer represented 5% of the malignant tumors of the entire body. In Egypt, the most common skin cancer is Basal Cell Carcinoma BCCs (77%) was followed by 15% for Squamous Cell Carcinomas (SCC), and 8% for melanomas [4]. In the last years, the death rate increased significantly because of skin melanoma lesions. In the early stage of skin cancer detection by physicians reach over 90% will save patients' life. Moreover, visual examination of skin

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

melanoma cancer is hard and may lead to the misleading classification of skin lesions due to the high similarity between different kinds of skin lesions (melanoma and non-melanoma). Furthermore, the use of traditional machine learning and image processing techniques in skin lesion classification has achieved high accuracy solution of the visual examination [5][6].

Melanoma is a kind of skin cancer also called malignant melanoma, which colors the skin and produces pigment-containing cells. Melanoma is less common compared to other skin cancer types but it is very risky [7]. Early detection of Melanoma is very important as its tendency to spread to other parts of the body and the high levels of curability in the early stages. Traditional ways in melanoma detection require well-trained specialists to overcome variations of inter-observation. However, it will increase the efficiency and accuracy of the early classification of this kind of skin cancer, if the melanoma recognition system has been done automatically [1], [5], [6].

4.1 Deep Learning (DL)

Traditional image classification techniques provided reasonable outcomes and performance regarding medical image disease detection using infected and uninfected images, but it was limited to small data sets and academic results. Deep Learning, as a subfield of machine learning, is concerned with algorithms inspired by the structure of the brain called artificial neural networks [8]–[10]. It is now considered a promising tool specifically image detection and object classification and recognition. Deep learning improves such automated diagnosis systems to achieve higher results, widen disease scope, and implementing applicable real-time medical image [11]–[16] disease detection systems.

Deep Learning, as a branch of artificial intelligence and machine learning, depends on algorithms for data processing and thinking process simulation, or for developing abstractions [17]–[19]. Deep Learning (DL) maps inputs to outputs by using layers of methods to process and analyze hidden patterns in data and visually objects detection [20]–[22]. Data passed through each layer of a deep network, with the output of the previous layer providing input for the next layer. The input layer is the first layer in the deep neural network, while the output layer is the final layer in the deep network. All the hidden layers located between input layers and output layers [17], [23].

1.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are the most successful type of architecture for image classification and detection to date. A single CNN architecture contains many different layers of neural networks that work on classifying edges and simple/complex features on shallower layers and more complex deep features in deeper layers [24], [25]. In 2012, a paper introduced by Ciregan et al. at [26] showed how CNNs using max-pooling based on Graphics Processing Unit (GPU) can improve many vision benchmark records on MNIST [27], Latin letters, traffic signs [28], NORB (jittered, cluttered) [29], Chinese characters [30], and CIFAR10 benchmarks. In 2012, a CNN system introduced by Krizhevsky et. al. [31] won the large-scale ImageNet LSVRC-2010 [32] competition by achieving a significant classification accuracy margin over traditional machine learning methods. In 2012, Ciresan et al.'s [33] introduced a CNN system also that won the ICPR 2012 mitosis detection competition on analysis of large medical images for breast cancer detection.

In the following eight years, various advances in deep convolutional neural networks further reduced the error rate on the image classification competition tasks. CNN models demonstrated significant improvements in succeeding in the ImageNet Large Scale Visual Recognition Competition (ILSVRC) annual challenges. The Visual Geometry Group at Oxford (VGG) developed the VGG-16 and VGG-19 model for the ILSVRC-2014 competition with a 7.3% Top-5 error rate [34]. The winner of the ILSVRC 2014 competition was GoogleNet with a 6.7% Top-5 error rate [35]. In 2015, Residual Neural Network (ResNet) is the winner ILSVRC 2015 competition with a 3.6% Top-5 error rate. [36].A model called Xception [37] was introduced that uses depth-wise separable convolutions to outperform the Inception-V3 model [38] on the ImageNet [32] dataset classification task. Huang et. al introduced [39] A new CNN variant called Densely Connected Convolutional Networks (DenseNet) where each layer is connected directly to every later layer. The DenseNet has achieved considerable accuracy in classification, using significantly fewer computations and parameters, over the state-of-the-art.

The remaining of this paper is organized as follows. In section 2, explores related work and determines the scope of this works. In section 3,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

discusses the dataset used in our model. In section 4, the model's architecture will be presented while section 5 discusses our outcomes and discussion of the paper. Finally, section 6 provides conclusions and directions for further research.

2. RELATED WORKS

This section conducts a survey on the latest researches for applying deep learning and machine learning in the field of medical skin lesions. At present, researchers across the world have obtained significant results by are applying machine and deep learning tools in a wide diversity of medical, industrial and agricultural image analyses/understanding tasks. In [40], A set of Morphological high-level intuitive features (HLIF) has been introduced by Amelard et al. to provide automated tools for the detection of skin cancer. The experimental outcomes show that adding a small set of HLIFs to the large state-of-the-art low-level skin lesion feature set decreasing the cross-validation error while increases sensitivity, specificity, and accuracy. The proposed HLIF system achieved a classification rate of 87.4%.

Almaraz-Damian et al. [41] designed a new computer-aided diagnosis system to extract features of such type of Melanoma images, based on the designed algorithm that uses ABCD Rule and Textural Features. The proposed system achieved 75.1 % using a train/test split. Authors [42] present a decision support system, which we call MED-NODE. MED-NODE used a non-dermoscopic digital image of skin lesions from which it extracts the skin lesion regions and then the system computes descriptors regarding the color and texture. Their system achieved a classification rate of 81%. Jafari et al. [43] introduced a system for the detection of skin lesions in images captured by smartphones. The system reduced the noise in the input images by a guided filtering method, and then they applied the ABCD rule of dermatology for melanoma detection to extract the skin lesions features. The experimental outcomes showed that the proposed system was capable of achieving an accuracy of 79%.

Nasr-Esfahan et al. [44] applied a convolutional neural network with skin clinical images. They used illumination correlation, and then they mask generation and Gaussian filter to increase the accuracy of their model. The enhanced and segmented skin images were sent to the CNN model for feature extraction and classification. This method achieved an accuracy of 81% evaluated on a publicly available dataset of 170 skin lesion images. In paper [45], Kostopoulos et al. proposed a computer-based analysis for Plain photography from different image databases. The extraction of features was done by the Probabilistic Neural Network (PNN) to decide the type of skin lesion. The system was evaluated utilizing fourteen features for discriminating with high accuracy melanomas from moles captured using plain photography. The system achieved an accuracy rate equal to 76.2%.

Premaladha and Ravichandran [46] introduced a computer-aided diagnosis system (CAD) that equips efficient algorithms for skin lesion classification and prediction through deep learning and supervised algorithms. The skin images were enhanced using the contrast limited adaptive histogram equalization technique (CLAHE). Median filter with the Normalized Otsu's Segmentation (NOS) used to separate and segment the normal skin image. They utilized Hybrid Adaboost-Support Vector Machine (SVM) and Deep Learning-based Neural Networks to achieve classification rate that reached 93%. In [47], Pham et.al. used deep CNN and data augmentation to enhance the classification performance of skin images and overcome the problem of data limitation. They used Inception V4 architecture to implement the feature extraction process. They used three classifiers (NN, SVM, RF) to improve performances of classification. The best accuracy performance achieved is 89%.

Esteva et al. [48] demonstrate a trained end-to-end single CNN to classify clinical skin lesions. They classified 129,450 clinical images to three classes of skin lesions called melanomas, seborrheic keratosis, and benign/nevus. They used Google's Inception v3 CNN model and achieved an accuracy rate of 72.1%. Yu et al. [49] introduced an automated methodology for skin classification in dermoscopy images based on a full convolutional residual network (FCRN). The residual learning was used to deal with degradation and overfitting problems. Their study shows that very deep CNN's with effective training mechanisms can solve medical image analysis problems. The outcomes of experiments show that a recognition rate of 85.5%.

Finally, in paper [1] Hosny et al. proposed a deep transfer learning model using Alexnet to classify medical skin lesions images. The transfer learning weights are fine-tuned by replacing the classification layer with a softmax layer. Augmentation techniques were applied to each of the three datasets to overcome the overfitting problem.

Journal of Theoretical and Applied Information Technology
15th May 2020. Vol.98. No 09

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The performance of the proposed transfer learning model is tested using three skin lesions datasets, DermIS- DermQuest, MED-NODE, and ISIC. The average accuracy for the proposed method with the DermIS- DermQuest is 96.86%. For the MED-NODE dataset, the average value of the accuracy measure is 97.7%. Finally, the average accuracy measure with the ISIC dataset is 95.91%.

In this paper, we construct deep learning pre-train models to recognize normal and upnormal conditions. To input the skin lesions images to the convolutional neural network, we embedded the medical skin images using data augmentation to get rid of overfitting. After that, a classifier is used to ensemble the outputs of the final prediction outcomes. The proposed method was evaluated on two skin lesions datasets from MED-NODE and ISIC.

3. DATASET

This research conducted its experiments on two datasets. The first dataset is MED-NODE. The MED-NODE dataset consists of 170 images, 100 naevus and 70 melanoma images, from the digital image archive of the Department of Dermatology of the University Medical Center Groningen (UMCG) used for the development and testing of the MED-NODE system for skin cancer detection from macroscopic images[42]. The second dataset is ISIC. The ISIC dataset consists of 2000 images is provided as training data, including 374 "melanoma", 254 "seborrheic keratosis", and the remainder as benign nevi (1372) [50]. Examples for images for both datasets MED-NODE and ISIC are presented in Figure 1.

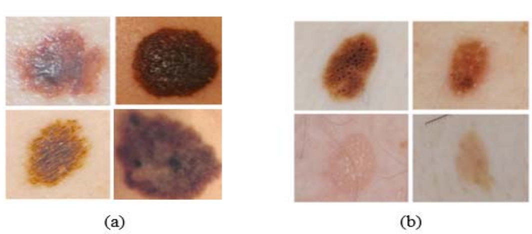

Fig. 1. Sample Images For (A) The MED-NODE Dataset, (B) The ISIC Dataset.

4. PROPOSED MODELS

The authors of research tried first to build their deep neural networks structure based on the works presented in [51]–[53] but the testing accuracy wasn't acceptable. So, The proposed methods used in this research relied on the pre-trained CNN architectures to transfer the learning weights to reduce the training time, mathematical calculations and the consumption of the available hardware resources which was adapted in similar research in [54], [55]. The used pre-trained CNN models are resnet18 [36], squeezenet [56], googlenet [57],vgg16 [58],vgg19 [50]. The mentioned CNN models had a quite few numbers of layers if it is compared to large CNN models such as xception, densenet, and inceptionresnet which consist of 71, 201 and 164 layers accordingly. Table 1 present the number of layers of the used CNN models.

Table 1. Number of layers for different CNN models

Model	Resnet18	Squeeze net	Google net	Vgg16	Vgg19
Number of Layers	18	18	22	16	19

The previous CNN models were customized in the last fully connected layer to match the number of classes of the datasets as illustrated in Figure 2.

4.2 Data Augmentation Techniques

The most common method to overcome overfitting is to increase the number of images used for training by applying label-preserving transformations [59]. Besides, data augmentation schemes are applied to the training set to make the resulting model more invariant for any kind of transformation and noise. The used augmentation techniques used in this research are:

- Reflection around the X-axis.
- Reflection around the Y-axis.
- Reflection around the X-Y axis.
- Pepper and salt Noise [60].
- Gaussian Noise [60].

The adopted augmentation techniques have raised the number of images on both datasets by 5 times larger than the original dataset. The MED-NODE dataset raised to 850 images and The ISIC dataset raised to 10000 images. This will lead to a significant improvement in CNN testing accuracy as it will be discussed in the following section. Also, will make the proposed methods immune to memorize the data and be more robust and accountable for the testing and verification phase.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

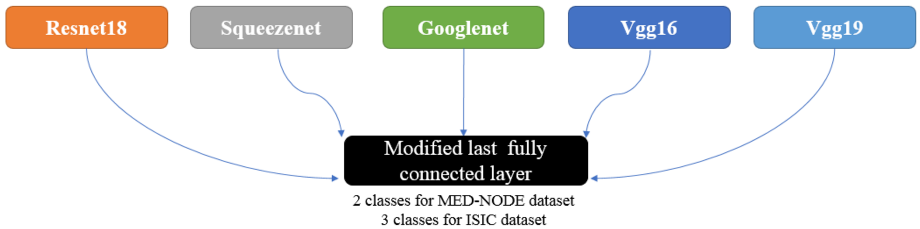

Fig. 2. Proposed Model's Customization For Skin Lesions Detection

5. EXPERIMENTAL RESULTS

The proposed architecture was developed using a software package (MATLAB). The implementation was CPU specific. All experiments were performed on a computer server with an Intel Xeon E5-2620 processor (2 GHz), 96 GB of RAM.

5.1 Testing Accuracy Measurement

Testing accuracy is one of the measurements which proves the accuracy of any proposed research. The confusion matrix also is one of the accurate measurements which give more insight into the achieved testing accuracy. Figures 3,4,5,6 and 7 presents the confusion matrix for the different CNN models used in this research.

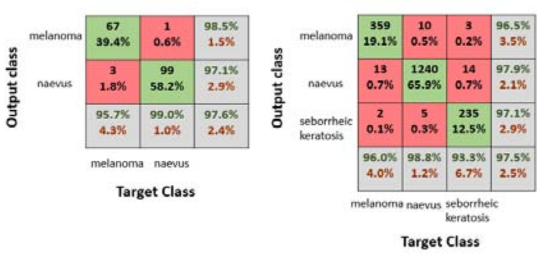

Fig. 3. Resnet18 confusion matrix for MED-NODE and ISIC datasets.

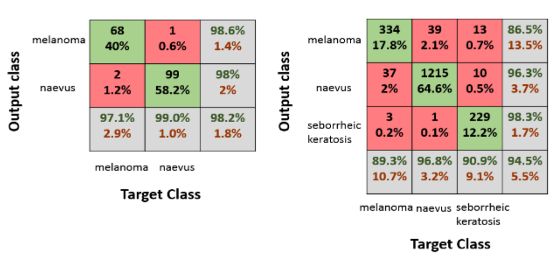

Fig. 4. Squeezenet confusion matrix for MED-NODE and ISIC datasets.

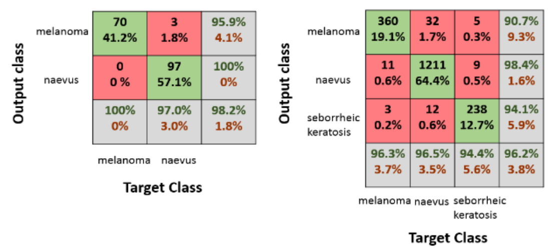

Fig. 5. Googlenet confusion matrix for MED-NODE and ISIC datasets.

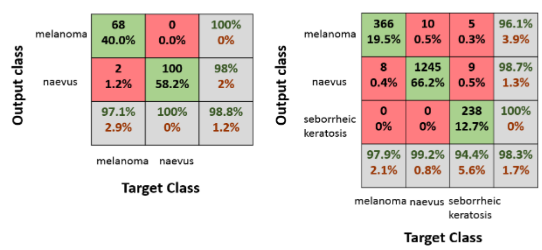

Fig. 6. Vgg16 confusion matrix for MED-NODE and ISIC datasets.

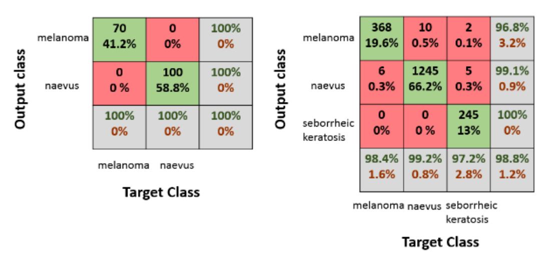

Fig. 7. Vgg19 confusion matrix for MED-NODE and ISIC datasets.

The summary of the testing accuracy is presented in Table 2.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2. Testing accuracy for the different CNN models

Model/Testing Accuracy	Resnet18	Squeezenet	Googlenet	Vgg16	Vgg19
MED-NODE	97.60 %	98.20 %	98.20 %	98.80 %	100 %
ISIC	97.50 %	94.50 %	96.20 %	98.30 %	98.80 %

It is clearly shown that the model Vgg-19 CNN model achieved the maximum testing accuracy. It achieved 100% for the MED-NODE dataset and 98.80% for the ISIC dataset. The testing accuracy measurement is not enough to prove the robustness of the proposed models, more measurements are indeed needed to draw a conclusion about the proposed model which fits more the dataset. The following section will discuss the other performance matrices.

1.1 Performance Evaluation and Discussion

To evaluate the performance of the proposed models, more performance matrices are needed to be investigated through this research. The most common performance measures in the field of deep learning are Precision, Recall, F1 Score, Selectivity, Negative Predictive Value, Informedness and Markedness [61], and they are presented from equation (1) to equation (7).

$$\text{Precision} = \frac{\text{TP}}{(\text{TP+FP})} \quad (1)$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP+FN})} \quad (2)$$

$$\text{F1 Score} = 2 * \frac{\text{Precision.Recall}}{(\text{Precision+Recall})} \quad (3)$$

$$Selectivity = \frac{TN}{TN+FP} \quad (4)$$

$$Negative\ Predictive\ Value = \frac{TN}{TN+FN} \quad (5)$$

$$Informedness = Precision + Selectivity - 1 \quad (6)$$

$$Markedness = Recall + Negative\ Predictive\ Value - 1 \quad (7)$$

Where TP is the count of True Positive samples, TN is the count of True Negative samples, FP is the count of False Positive samples, and FN is the count of False Negative samples from a confusion matrix.

Table 3 presents the performance matrices for the proposed deep learning models for MED-NODE dataset, while Table 4 illustrates the same matrices for ISIC dataset.

Table 3. Performance matrices for different CNN model on MED-NODE dataset

	Resnet18	Squeezenet	Googlenet	Vgg16	Vgg19
Precision	98.76 %	98.78 %	96.45 %	100 %	100 %
Recall	96.39 %	97.58 %	100 %	97.59 %	100 %
F1 Score	97.56 %	98.17 %	98.19 %	98.78 %	100 %
Selectivity	98.76 %	98.78 %	96.45 %	100 %	100 %
Negative Predictive Value	96.39 %	97.58 %	100 %	97.59 %	100 %
Informedness	97.53 %	97.55 %	92.89 %	100 %	100 %
Markedness	92.77 %	95.16 %	100 %	95.18 %	100 %

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 4. Performance matrices for different CNN model on ISIC dataset

	Resnet18	Squeezenet	Googlenet	Vgg16	Vgg19
Precision	93.52 %	86.06 %	89.6 %	94.26 %	95.92 %
Recall	95.10 %	90.21 %	93.77 %	97.99 %	98.51 %
F1 Score	94.30 %	88.07 %	91.63 %	96.08 %	97.19 %
Selectivity	97.40 %	94.04 %	95.65 %	97.71 %	98.38 %
Negative Predictive Value	98.06 %	95.96 %	97.48 %	99.22 %	99.42 %
Informedness	90.92 %	80.1 %	85.25 %	91.97 %	94.29 %
Markedness	93.15 %	86.16 %	91.25 %	97.21 %	97.93 %

The vgg19 model achieved the most stable and the maximum percentage according to the performance matrices on both datasets. The Vgg19 achieved testing accuracy and its performance matrices will be compared to one of related works in [1] which use the same datasets and similar data augmentation techniques but different CNN model (Alex-net).

5.2 Comparison with related works

Tables 5 and 6 present a comparative result with the related works on both MED-NODE and ISIC datasets. The work presented in [1] is the closest research to ours. It used the same dataset and similar data augmentation techniques but different CNN models (Alex-net). It is clearly shown in table 5 that the proposed model using vgg19 has overcome the related works in terms of testing accuracy, sensitivity and selectivity as a performance matrix. It achieved on MED-NODE and ISIC dataset an accuracy 100% and 98.8% respectively. Moreover, the proposed model has also a small number of training images. The related work in [1] trained on 9350 images for MED-NODE dataset while the proposed model trained only on 850 images on the same dataset. Also, the related work in [1] trained on 110000 images for ISIC dataset while the proposed model trained only on 10000 images on the same dataset. That means, it is not about enlarging the dataset number of images by hazard, the choice of the augmentation techniques play an important role in designing the deep CNN model. That will reflect on reducing the training time, memory consumption and the calculation complexity.

Table 5. MED-NODE dataset comparative outcomes.

Author	Accuracy	Sensitivity	Selectivity
Kostopoulos et al. [45]	76.20%	73.90%	77.80%
Jafari et al. [43]	79.00%	90.00%	72.00%
Giotis et al. [42]	81.00%	80.00%	81.00%
Esfahan et al. [44]	81.00%	81.00%	80.00%
Premaladha, and Ravichandran [46]	92.89%	94.83%	90.46%
Hosny et al. [1]	97.7%	97.34%	97.34%
Proposed Model	**100%**	**100%**	**100%**

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 6. ISIC dataset comparative outcomes.

Author	Accuracy	Sensitivity	Selectivity
Esteva et al. [48]	72.10%	-	-
Yu et al. [49]	85.50%	50.70%	94.10%
Pham et al. [47]	89.00%	55.60%	97.10%
Hosny et al. [1]	95.91%	88.47%	93.00%
Proposed Model	**98.80%**	**98.51%**	**98.38%**

6. CONCLUSION AND FUTURE WORKS

More people around the world are diagnosed with skin cancer than all other cancers combined each year globally. It affects more than 58 million people only in the united states of America. With the advances of computer algorithms, early detection of skin lesions with help to defeat this cancer in its early stages. This paper introduced an empirical study and enhancement on deep transfer learning for skin lesions detection. The presented study applied different pre-trained deep convolutional neural networks models such as resnet18, squeezenet, google net, vgg16, and vgg19. The selection of DCNN models due to the small number of layers on its architectures.

This study used two different datasets for skin lesions; MODE-NODE and ISIC datasets. Augmentation techniques have been adopted in this study to enlarge the total number of images in the dataset to be 5 times larger than the original datasets. The adopted augmentation techniques make the proposed models used in the study more robust and prevented overfitting. Moreover, seven accredited performance matrices in deep learning have been used to conclude a solid choice of the most appropriate DCNN model that fits the nature of the skin lesions. The study concluded that vgg19 is the most appropriate DCNN according to testing accuracy measurement and achieved 98.8% and all the seven performance matrices strengthen this result. Also, a comparative result was introduced with related works. The vgg19 overcome the related work in terms of testing accuracy and the performance matrices on both datasets.

Finally, the vgg19 model which was the result of the study was trained in a smaller number of images than the related work by 10 times time, which proved that the choice of data augmentation techniques played an important role in the achieved results. That would reflect on reducing the training time, memory consumption and calculation complexity. One of the potential future works is applying new architectures of deep neural networks such as Generative Adversarial Neural networks. GAN will be used before the proposed architecture. It will help in generating new images from the trained images, which will reflect on the accuracy of the different models.

REFRENCES

[1] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Classification of skin lesions using transfer learning and augmentation with Alex-net," *PLoS One*, vol. 14, no. 5, p. e0217293, May 2019.

[2] M. Loey, M. W. Jasim, H. M. El-bakry, M. H. N. Taha, and N. E. M. Khalifa, "Breast and Colon Cancer Classification from Gene Expression Profiles Using Data Mining Techniques," pp. 1–16, 2020.

[3] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA. Cancer J. Clin.*, vol. 68, no. 1, pp. 7–30, Jan. 2018.

[4] M. R. Hussein, "Skin cancer in Egypt: A word in your ear," *Cancer Biol. Ther.*, vol. 4, no. 5, pp. 593–595, May 2005.

[5] L. Ballerini, R. B. Fisher, B. Aldridge, and J. Rees, "A Color and Texture Based Hierarchical K-NN Approach to the Classification of Non-melanoma Skin Lesions BT - Color Medical Image Analysis," M. E. Celebi and G. Schaefer, Eds. Dordrecht: Springer Netherlands, 2013, pp. 63–86.

[6] N. Codella, J. Cai, M. Abedini, R. Garnavi, A. Halpern, and J. R. Smith, "Deep Learning, Sparse Coding, and SVM for Melanoma Recognition in Dermoscopy Images BT - Machine Learning in Medical Imaging," 2015, pp. 118–126.

[7] H. Mahmoud, M. Abdel-Nasser, and O. A. Omer, "Computer aided diagnosis system for skin lesions detection using texture analysis methods," in *2018 International Conference on Innovative Trends in Computer Engineering (ITCE)*, 2018, pp.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

140–144.

[8] D. Rong, L. Xie, and Y. Ying, "Computer vision detection of foreign objects in walnuts using deep learning," *Comput. Electron. Agric.*, vol. 162, pp. 1001–1010, 2019.

[9] A. Brunetti, D. Buongiorno, G. F. Trotta, and V. Bevilacqua, "Computer vision and deep learning techniques for pedestrian detection and tracking: A survey," *Neurocomputing*, vol. 300, pp. 17–33, 2018.

[10] J. Maitre, K. Bouchard, and L. P. Bédard, "Mineral grains recognition using computer vision and machine learning," *Comput. Geosci.*, 2019.

[11] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Z. Med. Phys.*, vol. 29, no. 2, pp. 102–127, 2019.

[12] A. Maier, C. Syben, T. Lasser, and C. Riess, "A gentle introduction to deep learning in medical image processing," *Z. Med. Phys.*, vol. 29, no. 2, pp. 86–101, 2019.

[13] J. Zhang, Y. Xie, Q. Wu, and Y. Xia, "Medical image classification using synergic deep learning," *Med. Image Anal.*, vol. 54, pp. 10–19, 2019.

[14] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.

[15] E. Gibson *et al.*, "NiftyNet: a deep-learning platform for medical imaging," *Comput. Methods Programs Biomed.*, vol. 158, pp. 113–122, 2018.

[16] S. Liu *et al.*, "Deep Learning in Medical Ultrasound Analysis: A Review," *Engineering*, vol. 5, no. 2, pp. 261–275, 2019.

[17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[18] G. Eraslan, Ž. Avsec, J. Gagneur, and F. J. Theis, "Deep learning: new computational modelling techniques for genomics," *Nat. Rev. Genet.*, 2019.

[19] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep Learning for Computer Vision: A Brief Review," *Comput. Intell. Neurosci.*, vol. 2018, p. 7068349, 2018.

[20] J. Riordon, D. Sovilj, S. Sanner, D. Sinton, and E. W. K. Young, "Deep Learning with Microfluidics for Biotechnology," *Trends Biotechnol.*, vol. 37, no. 3, pp. 310–324, 2019.

[21] J. You, R. D. McLeod, and P. Hu, "Predicting drug-target interaction network using deep learning model," *Comput. Biol. Chem.*, vol. 80, pp. 90–101, 2019.

[22] K. Jaganathan *et al.*, "Predicting Splicing from Primary Sequence with Deep Learning," *Cell*, vol. 176, no. 3, pp. 535-548.e24, 2019.

[23] C. Cao *et al.*, "Deep Learning and Its Applications in Biomedicine," *Genomics. Proteomics Bioinformatics*, vol. 16, no. 1, pp. 17–32, 2018.

[24] N. E. Khalifa, M. Hamed Taha, A. E. Hassanien, and I. Selim, "Deep galaxy V2: Robust deep convolutional neural networks for galaxy morphology classifications," in *2018 International Conference on Computing Sciences and Engineering, ICCSE 2018 - Proceedings*, 2018, pp. 1–6.

[25] N. Eldeen, M. Khalifa, M. Hamed, N. Taha, and A. E. Hassanien, "Aquarium Family Fish Species Identification System Using Deep Neural Networks," 2018.

[26] D. Ciregan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3642–3649.

[27] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[28] J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, "The German Traffic Sign Recognition Benchmark: A multi-class classification competition," in *The 2011 International Joint Conference on Neural Networks*, 2011, pp. 1453–1460.

[29] Y. LeCun, F. J. Huang, and L. Bottou, "Learning methods for generic object recognition with invariance to pose and lighting," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, 2004, vol. 2, pp. II-104 Vol.2.

[30] F. Yin, Q. Wang, X. Zhang, and C. Liu, "ICDAR 2013 Chinese Handwriting Recognition Competition," in *2013 12th International Conference on Document Analysis and Recognition*, 2013, pp. 1464–1470.

[31] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proceedings of the 25th International Conference on Neural Information*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Processing Systems, 2012, pp. 1097–1105.

[32] J. Deng, W. Dong, R. Socher, L. Li, L. Kai, and F.-F. Li, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[33] D. C. Cireşan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, 2013, pp. 411–418.

[34] S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," in *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, 2015, pp. 730–734.

[35] C. Szegedy *et al.*, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.

[36] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[37] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1800–1807.

[38] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[39] G. Huang, Z. Liu, L. v. d. Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269.

[40] R. Amelard, A. Wong, and D. A. Clausi, "Extracting morphological high-level intuitive features (HLIF) for enhancing skin lesion classification," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2012, pp. 4458–4461.

[41] J. A. Almaraz-Damian, V. Ponomaryov, and E. Rendon-Gonzalez, "Melanoma CADe based on ABCD Rule and Haralick Texture Features," in *2016 9th International Kharkiv Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves (MSMW)*, 2016, pp. 1–4.

[42] I. Giotis, N. Molders, S. Land, M. Biehl, M. F. Jonkman, and N. Petkov, "MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images," *Expert Syst. Appl.*, 2015.

[43] M. H. Jafari, S. Samavi, N. Karimi, S. M. R. Soroushmehr, K. Ward, and K. Najarian, "Automatic detection of melanoma using broad extraction of features from digital images," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 1357–1360.

[44] E. Nasr-Esfahani *et al.*, "Melanoma detection by analysis of clinical images using convolutional neural network," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 1373–1376.

[45] S. A. Kostopoulos *et al.*, "Adaptable pattern recognition system for discriminating Melanocytic Nevi from Malignant Melanomas using plain photography images from different image databases," *Int. J. Med. Inform.*, vol. 105, pp. 1–10, 2017.

[46] J. Premaladha and K. S. Ravichandran, "Novel Approaches for Diagnosing Melanoma Skin Lesions Through Supervised and Deep Learning Algorithms," *J. Med. Syst.*, vol. 40, no. 4, p. 96, 2016.

[47] T.-C. Pham, C.-M. Luong, M. Visani, and V.-D. Hoang, "Deep CNN and Data Augmentation for Skin Lesion Classification BT - Intelligent Information and Database Systems," 2018, pp. 573–582.

[48] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, p. 115, Jan. 2017.

[49] L. Yu, H. Chen, Q. Dou, J. Qin, and P. Heng, "Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks," *IEEE Trans. Med. Imaging*, vol. 36, no. 4, pp. 994–1004, 2017.

[50] N. C. F. Codella *et al.*, "Skin lesion analysis toward melanoma detection: A challenge at the 2017 International symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC)," in *Proceedings - International Symposium on Biomedical Imaging*, 2018.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[51] N. E. M. Khalifa, M. H. N. Taha, A. E. Hassanien, and A. A. Hemedan, "Deep bacteria: robust deep learning data augmentation design for limited bacterial colony dataset," *Int. J. Reason. Intell. Syst.*, 2019.

[52] N. E. M. Khalifa, M. H. N. Taha, D. Ezzat Ali, A. Slowik, and A. E. Hassanien, "Artificial Intelligence Technique for Gene Expression by Tumor RNA-Seq Data: A Novel Optimized Deep Learning Approach," *IEEE Access*, 2020.

[53] N. E. M. Khalifa, M. H. N. Taha, A. E. Hassanien, and I. M. Selim, "Deep Galaxy: Classification of Galaxies based on Deep Convolutional Neural Networks."

[54] N. Khalifa, M. Loey, M. Taha, and H. Mohamed, "Deep Transfer Learning Models for Medical Diabetic Retinopathy Detection," *Acta Inform. Medica*, vol. 27, no. 5, p. 327, 2019.

[55] N. E. M. Khalifa, M. Loey, and M. H. N. Taha, "Insect pests recognition based on deep transfer learning models," *J. Theor. Appl. Inf. Technol.*, vol. 98, no. 1, pp. 60–68, 2020.

[56] F. N. Iandola, M. W. Moskewicz, K. Ashraf, S. Han, W. J. Dally, and K. Keutzer, "SqueezeNet," *arXiv*, 2016.

[57] R. Shah and Y. Yang, "GoogLeNet," *Popul. Health Manag.*, 2015.

[58] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," pp. 1–14, 2014.

[59] N. Khalifa, M. Taha, A. Hassanien, and H. Mohamed, "Deep Iris: Deep Learning for Gender Classification Through Iris Patterns," *Acta Inform. Medica*, vol. 27, no. 2, p. 96, 2019.

[60] A. K. Boyat and B. K. Joshi, "A Review Paper : Noise Models in Digital Image Processing," *Signal Image Process. An Int. J.*, vol. 6, no. 2, pp. 63–75, Apr. 2015.

[61] C. Goutte and E. Gaussier, "A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation," 2010.