Convolution Neural Network Method for Skin Cancer Diagnosis: Comparison and Improvement

Xiang Li [1, †], Jiazhi Pang [2, †], Jiahui Meng [3, †], Jiahui Meng [4, †], Jiale Zhang [4, *, †]

[1] Sichuan University-Pittsburgh Institute, Sichuan University, Chengdu, China

[2] School of Mechanical Engineering, Hefei University of Technology, Hefei, China

[3.]Sichuan University-Pittsburgh Institute, Sichuan University, Chengdu, China

[4.]School of Statistics, Xi'an University of Finance and Economics, Xi'an, China

* Corresponding Author Email: ZJL_-@xaufe.edu.cn

† These authors contributed equally

Abstract. Skin cancer, the most common cancer in the world, has many detection steps and the detection process is easy to make mistakes. A detection method based on convolutional neural network (CNN) is proposed to assist doctors in the detection. Based on the development of CNN in the classification and diagnosis of skin cancer in recent years, this paper compares and summarizes the development of each step in this process. After reviewing previous papers, it can be concluded that the classification process is roughly divided into four parts. In addition, the evaluation indicators of the model are further analyzed. AUC Sen and SPE are the most basic evaluation indicators in the model evaluation. As a skin classifier, CNN improves the accuracy of classification and diagnosis results to a great extent. CNN model has also made progress in "lightweight" and "concise". However, there are few evaluation indicators available for different CNN methods, and the evaluation latitude is relatively single. In the future, the evaluation indicators should develop in more aspects, it will enable to better understand the personality of a CNN model.

Keywords: skin cancer detection; image processing; Convolutional Neural Networks.

1. Introduction

Skin cancer, as one of the most popular cancers over the world, accounts for at least 40% of cancer cases. According to the World Health Organization, between 2 and 3 million non-melanoma skin cancers and at least 132,000 malignant melanomas occur globally each year. There are three main types of skin cancers: basal-cell skin cancer (BCC), squamous-cell skin cancer (SCC) and melanoma. Among them, the first two types of nonmelanoma skin cancer are relatively safety; people can easily be cured by surgical removal. However, melanoma is more dangerous and deadliest. If melanoma cannot be found on time, it can easily transfer to other parts of human body, including the lungs, liver, spleen, or brain. Specially, melanoma is responsible for almost 10 percent of brain metastasis. However, melanoma can be cured in nearly 95 percent of cases if detected early enough. Therefore, the timely detection and cure of melanoma will be the most important task. Nowadays, the detection of skin cancer will include initial clinical screening and following dermoscopic analysis, a biopsy and histopathological examination. However, due to the limits of the dermatologists' experience of visual evaluation of dermoscopic images, this method has a probability of mistake. Hence, it will be essential to build an automated computational system to detect melanoma with a relative high accuracy among a large amount of skin cancer data.

According to previous review, there are many articles, which are related to skin cancer detection by Convolutional Neural Networks (CNN) method. By summarizing these articles, the utilization of CNN in skin cancer detection can be divided roughly into four main steps. Based on the review articles, this paper can be separated into the flowing parts. First one is image preprocessing, which aims to enhance the character of dermoscopic images for lesion segmentation, even to create different dataset by other devices like digital camera instead of dermoscopic. This part includes detection of dataset, building own dataset with different technique, enhance images with FILpFHSV, add the third

dimension deep into the image and so on. Second part is lesion segmentation; it is, more clearly, after separating skin cancer lesion with surrounding environment, more accurate the detection will be, which can be divided into XOA operation, Mask R-CNN and Retina-Deeplab. Third part is feature extraction as an input of the following CNN model, and the fourth part is the introduction of different CNN model, mainly including traditional CNN model and optimized CNN model. Among them, feature extraction method has been included. Finally, the evaluation methods of several skin cancer detection are further organized.

2. Image Preprocessing and lesion segmentation

The main challenge of skin cancer detection in image preprocessing and lesion segmentation includes the imbalance classification of images because of the small amount of melanoma types of images, which gives the probability of inaccuracy about melanoma detection. In addition, some artifacts like hair, illumination, veins, air bubbles, color calibration marks etc. will all affect the image segmentation and the accuracy of feature extraction [1]. Therefore, in this part, several authors utilized different approaches to enhance images and segment image precisely, as shown in Fig.1.

In detail, the study by Tanzila state: by rotating images to four angles in training process for data augmentation and following the fast local Laplacian filtering (FILpF) with HSV color transformation, the author will obtain ξR, ξG, and ξB as represent of three-color channel and ξLl [φx, y] represented images after FILpF for later lesion segmentation. Hence, for lesion segmentation, XOA operation will be performed to former computation to select a range of the most nearly lesion pixel points. After that, the lesion and healthy pixel will be separately loaded into pre-training CNN model for following procedures. At the end, they give the comment of images enhancement and segmentation by testing on dataset PH2 and ISIC 2017, the result shows an incredible increasing of accuracy, which is almost 1.5% [2].

Fatemeh proposed a new idea to obtain the final segmented images. In their opinions, the combination of different segmentation method results by appropriate combination strategies can improve the performance. Therefore, two different segmentation methods: Mask R-CNN and Retina-Deeplab are utilized to provide the segmentation result. Then, they will be combined by geodesic and graph-based method to obtain the final image [3].

Although most of the popular dataset of skin detection research are from ISIC, Shunichi create their own dataset, which is taken by digital, single-lens reflex cameras instead of dermoscopic image and trained by bounding-box annotation method. This article specially provides a different view of our paper, considering most of dataset of skin cancer detection research are dermoscopic image [4]. Furthermore, Ivan, Lyudmila etc. utilize the Raman spectra analysis to give a significant better result due to the much more accurate computation of low intensity Raman bands in the intense autofluorescence background [5].

Kassani conducted a comparative study of skin lesions. Dermatoscopic images are used to classify cancers. The pretreatment methods include illumination correction, horizontal flip, vertical flip, random contrast, and random brightness. The results were evaluated by four values: specificity, sensitivity, accuracy, and F value. The experimental results show that the classification accuracy rate of ResNet50 is 92.08% better than AlexNet, Xception, VGGNet16 and VGGNet19 ARCHi [6]. Nour Abuared explained and evaluated training classification strategies. CNN and TL based on Vgg19 turn out to be powerful tools to help diagnose skin cancer with a high degree of precision. The formation and accuracy of the network tests is 0.985 and 0.975, and the formation and test losses are 0.099 and 0.119, respectively, which is satisfactory. If required, additional preprocessing can be performed to improve training accuracy [7].

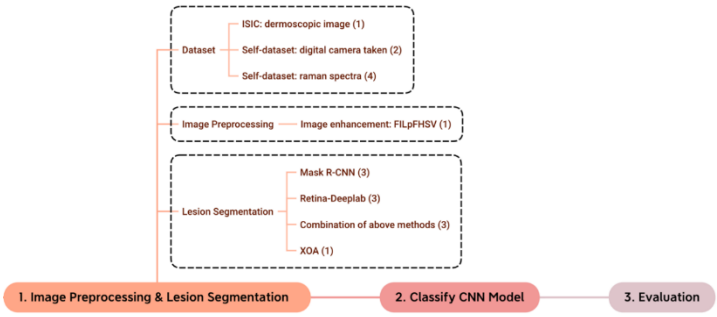


Figure 1. Structure of image preprocessing and lesion segmentation

Table 1. Conclusion of Part ii

Reference	Author	Main Objective	Main Method	Results
[1]	Tanzila,	Segmentation and enhancement	FILpFHSV and XOA	Accuracy increasing 1.5%
[2]	Shunichi	New dataset	Digital, single-lens reflex camera	Similar accuracy but easier to obtain dataset
[3]	Fatemeh	Segmentation and combination	Mask R-CNN and Retina-Deeplab combination	Accuracy increasing a lot
[4]	Ivan	Image enhancement	Raman spectra analysis	Accuracy increasing
[5]	Kassani	Image classification	illumination correction, horizontal flip, vertical flip, random contrast and random brightness	Accuracy increasing
[6]	Nour	Improvement of accuracy	Vgg19-based CNN and TL	Accuracy increasing a lot

3. TRADITIONAL Model

3.1. Store and Read Skin Cancer Data

Anthony Kioria Waweru et al. proposed a skin damage method called DenseNet201 to facilitate storage and reading of skin cancer related data. This traditional model is directly used for dermatoscopic feature extraction [8]. At the same time, DenseNet201 has the highest accuracy in the transfer learning model, and the experimental results show that this framework has a high balance accuracy. Rahman et al. improved DenseNet method and proposed a robust system model ResNet, DenseNet, and Xcepiton. The accuracy, precision, recall rate and specificity of the three training models were verified by weighted average integration technique. DenseNet model has 89% accuracy. The Accuracy of the Xception model was 87%, while the ResNet model performed poorly in balancing accuracy and precision [9]. Mina Al-Saad conceived, formed, and evaluated a DCNN after dealing with imbalances in the categories of datasets. The performance of DCNN was compared with VGG16 and VGG19. DCNN has advantages over VGG16 and VGG19. DCNN model which was processed was not overly broad. Using additional information on images potentially improves network performance and robustness [10]. Recently proposed a method of extracting boundaries and identifying lesions via deep neural networks. The clustering control entropy selection method was used to distinguish the features and the MLP method was used for classification [11], as illustrated in Fig. 2.

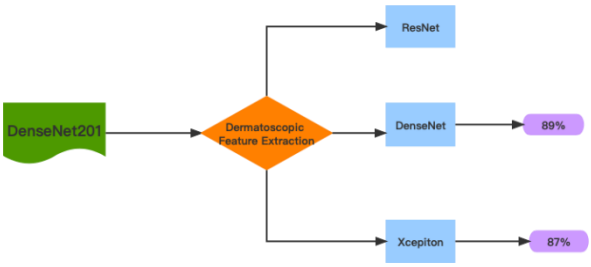


Figure 2. Correlation of traditional model accuracy graph.

3.2.Comparison of Model Performance

Yessi Jusman compared and analyzed the performance of each training model from two aspects of classification accuracy and calculation time. The results show that the VGG-16 model can set the best classification accuracy in the comparison network, and the VGG-16 and customized CNN model are much faster than the multilayered perceptron in test time. This study is helpful to systematically compare and analyze the application of several neural networks in skin cancer classification [12]. Masood proposed a semi-supervised self-suggested learning model, which used tagged and unlabeled data, and adopts support vector machine algorithm to enhance classification results and offset the influence of misclassified data. The classification method and deep neural processing technique of the model are superior to KNN, ANN, SVM and semi-supervised algorithm [13].

4. Optimized Model

4.1. Optimization Based on Traditional Architecture

Convolutional neural network (CNN) is always on its way to be optimized. CNN is mainly composed of input layer, convolutional layer, activation function, pooling layer, full connection layer and loss function. All components serve to two targets, feature extraction and then based on features decision or classification. So, its optimization mainly focuses on improving the performance of these two targets. For example, R-CNN model [14] was used by author to detect Melanoma. Quicker R-CNN has 3 main parts that are convolutional layers for feature extraction process. RPN is an object detector which locates the lesion area and segments the original image and then give proposal to pooling layer. And there are two completely connected layers which operate for object classification and produce object position coordinates. Comparing to traditional CNN model, Faster R-CNN can provide a strong computing capability, largely shortening training time. In the experiment executed by authors of [14] under a same circumstance where Windows 10 with hardware specifications of Nvidia GTX 1070 TI, 8 GB RAM, and i5 Processor are provided, the time training a same dataset with 600 images of Faster R-CNN is 1.5 hours and 8 hours for MobileNet, another optimized CNN. The illustration of Faster R-CNN is shown in Fig.3.

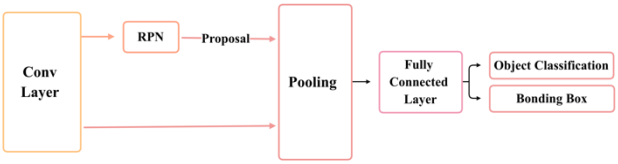


Figure 3. Structure of faster R-CNN.

There are other improvements made by acknowledging the components in CNN. [15] states a new method for feature extraction. The most important change is that they utilize scattered wavelet transform to decompose the segmented skin cancer images and extract the handcraft feature as the additional input of fully connected layer of CNN.

4.2. Light Weight CNN

Although the deep learning neural networks indeed solve the challenging problem for lack of professional doctors and performs as a reference for more sophisticated diagnoses. There is still a challenge that implementation of deep neural networks has high requirements for image processing with CNN models. CNN models include a complex architecture of usually exceeding hundreds of layers. Therefore, calculations of CNN models cover vast memory units and powerful processors. These limitations become the largest obstacle in the way of applying in remote and small clinic in countryside, where not only lacks doctors but also advanced processing devices.

Google develops famous MobileNet to solve this problem by applying depthwise separable convolution, which could shrink the parameters size but keep accuracy on some degree. MobileNet has been evolved from its original version, MobileNet-V1, to MobileNet-V3. MobileNet-V2 is improved based on MobileNet-V1, which solves the problem of nonlinearities in the narrow layers of the model. This progress is made by reducing parameters in Depthwise Separable Convolution [16]. MobileNet-V3 imports NAS (Neural architectural search) to construct its based model MnasNet, and a new activation function h-swish(x) is used [17]. Although MobileNet solves the problem of high requirement of calculating capability. The loss of accuracy is still there.

To compensate the loss of accuracy, a novel weight pruning training strategy is developed. [18] proposes a pruning strategy for lightening weight of CNN. This approach involves not only pruning the unimportant CAD (Computer - Aided Design) weights, but also retraining the model to achieve improved performance.

4.3. Transferring Learning

Transferring learning is a popular optimizing method for training model. It aims to process target data with a present model trained by the previous processing of another dataset. Authors imported MobileNet removes the last five layers, including the tight layer with SoftMax activation. Then, they froze all the layers of weight except for the last 25 layers of training [19]. Their final model reaches the accuracy of 0.90. In [20], authors add a Global average pooling layer, dropout, and dense layers on a pre-trained based model including ResNet50, InceptionV3, Xception, and MobileNet, to detect skin lesions. The architecture of after-transfer-learning model is shown in Fig. 4.

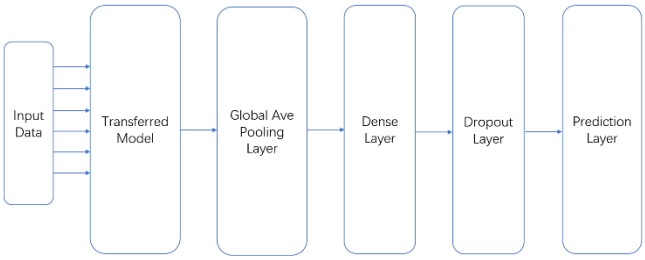


Figure 4. Architecture of after-transfer-learning.

They evaluated and compared these based models by accuracy, recall and precision. At last, ResNet50 was picked out and applied as based model for transfer learning. The result of ResNet50 transfer learning was compared with other transfer learning models like MobileNet, EfficientNet B1 and Inception ResNet.

4.4. Other Methods and Algorithms

There are some other special improvement method or algorithm like GWO (grey wolf optimization) in [21]. GWO is proposed by Seyedali Mirjalili in 2014. Its main hierarchal architecture is developed by mimicking the actual movement of hunting by group of wolves. Wolves are classed as alpha, beta, and omega, ranked by authority. Alpha wolf has highest power, while omega wolves obey the order by any means. Mathematically, regarding the position as the optimal solution for the question, and the movement of wolves as solution's search space, GWO focuses on dealing with the toughness

caused by manually establishing the hyper parameters. This paper also compared the results of GWO, PSO (Particle Swarm Optimization) and GA (Genetic Algorithm) based hyper parameter optimizing method. And [22] improves the performances of ResNet-50 model by applying two hierarchical strategies, dividing diagnosing procedure into two sequential parts in two different sequences. One is to identify the skin lesions between non-melanocytic and melanocytic lesions in hier1, and then if it's melanocytic, identify between melanoma and Nevi lesions. Another is to identify the skin lesions between benign and malignant lesions in hier1, and then if it's malignant, identify between non-melanocytic and melanocytic lesions. The architecture of two hierarchical strategies is shown in Fig.5.

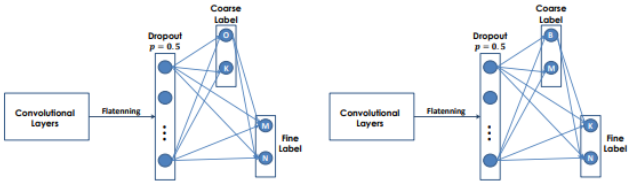


Figure 5. Architecture of hierarchical models.

4.5. evaluation

Model evaluation is an important part of model comparison and screening. Model evaluation and screening depend on model evaluation indicators. By consulting research papers in recent years, it can be concluded that the common indicators of evaluation models are sensitivity, specificity, accuracy, recall rate, dice similarity (DC), Jaccard index (Ji) and F1 score. For example, In Thanh, D. N. H.'s paper, A new image evaluation method (TDS) is proposed. When evaluating its model, they use these evaluation indexes: accuracy, sensitivity, specificity, the Sorensen-Dice and the Jaccard metrics to assess. In many medical images segmentation work, including skin injury segmentation, the Jaccard score is the most important index to measure the quality of segmentation results [23]. To evaluate the performance of skinner, Sulaiman Vesal, Nishant Ravikumar and Andreas Maier use two similarity measures: dice similarity (DC) and Jaccard index (Ji), the authors evaluated the lesion segmentation accuracy of our network according to the provided ground truth template. They also counted sensitivity, specificity, and accuracy metrics, directly comparing this method with the latest technology proposed in recent studies [24]. And in M. Rahman's paper, a skin cancer detection decision support system (DSS) based on comprehensive classification and retrieval is proposed. To evaluate the effectiveness of DSS proposed in this paper, the author puts forward that the classification accuracy of the system is measured by weighted average accuracy, recall rate and F1 score. Weighting by category frequency may better estimate overall performance because category frequencies are not consistent in the dataset. Retrieval effectiveness is measured by the precise recall (PR) graph commonly used in the field of information retrieval. Segmentation performance is measured by common segmentation metrics, such as pixel level sensitivity, pixel level specificity, Dice coefficient, and Jaccard index [9].

In addition to the above indicators, Pham, Tri Cong et al. Proposed a method to calculate a user-defined measure based on the basic indicators AUC, SEN and SPE, which is called Custom Balanced Accuracy (BACC) [25]. In addition, we can also calculate "YI" (Youden index) [26] and "Δ" (difference between Sen and SPE). And lift diagram and gain diagram also can be used to measure how "better" the prediction ability of the model is compared with that without using the model. ROC, lift, and gains are all based on the confusion matrix and several derived indicators (sensitivity and specificity, etc.). However, up to now, few people have applied lift graph and gain graph as performance indicators to the evaluation of CNN model.

5. Conclusion

In short, in the process of skin cancer diagnosis, the automatic discrimination system with the help of data can effectively reduce the workload of medical staff. At the same time, it also greatly reduces

the diagnostic error rate. Through many skin cancer data, the establishment of an automatic computing system to detect melanoma can aid with doctors' diagnoses. As mentioned above, most of today's processing systems can be divided into four steps: image preprocessing, model construction, model optimization, and evaluation. This paper compares and summarizes the commonly used computing systems in recent years, as well as the image processing, model comparison, model optimization, and model evaluation involved in the processing process. In the main model construction stage, according to the operation mechanism of different systems, the traditional CNN model is summarized in two aspects: improving the accuracy of the model through image preprocessing and storing and reading cancer data. Finally, the performance of various training models is compared and analyzed through classification accuracy and calculation time. The optimization stage of the model is divided into optimization based on traditional structure and more "lightweight" optimization.

In addition, in the future, image processing will be applied in the medical field, and the auxiliary doctor diagnosis will be further developed and improved. For example, it will develop in the direction of lighter proportion, faster operation, and higher accuracy. The computing system can also be combined with a natural language processing system or other systems to help doctors and patients realize a simple, fast, and efficient diagnosis.

References

[1] P. M. M. Pedro, A. T. Lucas, T. M. N. Luis, A. A. A. Pedro, M. Rui, F. Pinto, P. P. Rui, M. M. Sergio. “Melanoma classification using light-Fields with morlet scattering transform and CNN: Surface depth as a valuable tool to increase detection rate”, Medical Image Analysis 75, 2022.

[2] S. Tanzila, A. K. Muhammad, R. Amjad, L. M. Souad, “Region Extraction and Classification of Skin Cancer: A Heterogeneous framework of Deep CNN Features Fusion and Reduction”, Journal of Medical Systems, 2019.

[3] B. Fatemeh, J. T. Mohammad, Z. Majid, “Skin lesion segmentation from dermoscopic images by using Mask R-CNN, Retina-Deeplab, and graph-based methods”, Biomedical Signal Processing and Control, 2019

[4] S. Jinnai, N. Yamazaki, Y. Hirano, Y. Sugawara, Y. Ohe and R. Hamamoto, “The Development of a Skin Cancer Classification System for Pigmented Skin Lesions Using Deep Learning” Department of Dermatologic Oncology, 2021

[5] A. B. Ivan, A. B. Lyudmila, A. K. Yulia, A. M. Alexander, V. K. Sergey, P. Z. Valery, “Classification of skin cancer using convolutional neural networks analysis of Raman spectra” Computer Methods and Programs in Biomedicine 219, 2022

[6] Sara Hosseinzadeh Kassania, Peyman Hosseinzadeh Kassanib, “A comparative study of deep learning architectures on melanoma detection,” Department of Computer Science, University of Saskatchewan, Saskatchewan, Canada, Department of Biomedical Engineering, Tulane University, New Orleans, Louisiana, USA

[7] M. Al-Saad, S. A. Amin, W. Mansoor, “Deep Convolutional Neural Network (DCNN) for Skin Cancer Classification,” N. Aburaed, A. Panthakkan, College of Engineering and IT, University of Dubai, UAE. 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS).

[8] A. K. Waweru, K. Ahmed, Y. Miao, P. Kawan, “Deep Learning in Skin Lesion Analysis Towards Cancer Detection,” College of Engineering and Science, Victoria University Melbourne, Australia, 17 May.2021.

[9] O. Layode, T. Alam, M. Rahman, “Deep Learning Based Integrated Classification and Image Retrieval System for Early Skin Cancer Detection”, IEEE Applied Imagery Pattern Recognition Workshop (AIPR)., 2019

[10] N. Abuared; Alavikunhu Panthakkan; Mina Al-Saad; Saad Ali Amin; Wathiq Mansoor,“Skin Cancer Classification Model Based on VGG19 and Transfer Learning,” 9 Feb. 2021.

[11] Tanzila Saba, Muhammad Attique Khan, Amjad Rehman, Souad Larabi Marie-Sainte“Region Extraction and Classification of Skin Cancer: A Heterogeneous framework of Deep CNN Features Fusion and Reduction,” 3 July.2019.

[12] Yessi Jusman; Indah Monisa Firdiantika; Dhimas Arief Dharmawan; Kunnu Purwanto“Performance of Multi-layer Perceptron and Deep Neural Networks in Skin Cancer Classification” 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech).

[13] Adel Al-Jumaili; Hyrul Anam“Self-Supervised Learning Model for Skin Cancer Diagnosi,” 2 July.2015.

[14] C. A. Hartanto and A. Wibowo, "Development of Mobile Skin Cancer Detection using Faster R-CNN and MobileNet v2 Model," International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), 2020, pp. 58-63.

[15] K. S. R. Sekhar, T. R. Babu, G. Prathibha, K. Vijay and L. C. Au Ming, “Dermoscopic image classification using CNN with Handcrafted features”, Journal of King Saud University – Science

[16] C. A. Hartanto and A. Wibowo, "Development of Mobile Skin Cancer Detection using Faster R-CNN and MobileNet v2 Model," International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE), 2020, pp. 58-63.

[17] H. Andrew, S. Mark, C. Grace, C. Liang-Chieh, C. Bo, T. Mingxing, W. Weijun, Z. Yukun, P. Ruoming, V. Vijay, V. Le, A. Hartwig, “Searching for MobileNetV3”, Cornell University, Nov 2019.

[18] K. Xiang, L. Peng, H. Yang, M. Li, Z. Cao, S. Jiang, and G. Qu, “A novel weight pruning strategy for light weight neural networks with application to the diagnosis of skin disease”, Applied Soft Computing, 2021.

[19] H. Younis, M. H. Bhatti, and M. Azeem, “Classification of Skin Cancer Dermoscopy Images using Transfer Learning”, 2019 IEEE 15th International Conference on Emerging Technologies (ICET).

[20] H. K. Kondaveeti, and P. Edupuganti, “Skin Cancer Classification using Transfer Learning”, International Conference on Advent Trends in Multidisciplinary Research and Innovation (ICATMRI), 2020.

[21] R. Mohakud, and R. Dash, “Designing a grey wolf optimization based hyper-parameter optimized convolutional neural network classifier for skin cancer detection”, Journal of King Saud University - Computer and Information Sciences.

[22] C. Barata, and J. S. Marques, “Deep Learning for Skin Cancer Diagnosis with Hierarchical Architectures”, 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI).

[23] D. N. Thanh, V. Prasath, L. M. Hieu and N. N. Hien, "Melanoma skin cancer detection method based on adaptive principal curvature colour normalisation and feature extraction with the ABCD rule", Journal of Digital Imaging, 2020.

[24] S. Vesal, N. Ravikumar and A. Maier, SkinNet: A deep learning framework for skin lesion segmentation, 2018, [online] Available: https://arxiv.org/abs/1806.09522.

[25] T. C. Pham, V. D. Hoang, C. T. Tran, M. S. K. Luu, D. A. Mai, A. Doucet, et al., "Improving binary skin cancer classification based on best model selection method combined with optimizing full connected layers of deep CNN", Proc. Int. Conf. Multimedia Anal. Pattern Recognit. (MAPR), Oct. 2020.

[26] T. J. Brinker et al., “Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task,” Eur. J. Cancer, vol. 113, May 2019.