Convolutional Neural Network for Diagnosing Skin Cancer

Mohammad Ashraf Ottom[1]

Department of Computer Information Systems Yarmouk University Irbid, Jordan

Abstract—Diagnosis of melanoma (skin cancer disease) is a challenging task in medical science field due to the amount and nature of the data. Skin cancer datasets are usually comes in different format and shapes including medical images, hence, data require tremendous efforts for preprocessing before the auto-diagnostic task itself. In this work, deep learning (convolutional neural network) is used to build a computer model for predicting new cases of skin cancer. The first phase in this work is to prepare images data, this include images segmentation to find useful parts that are easier for analysis and to detect region of interest in digital images, reduce the amount of noise and image illumination, and to easily detect sharp edges (boundaries) of objects. Then, the proposed approach built a convolutional neural network model which consists of three convolution layers, three max pooling layers, and four fully connected layers. Testing the model produced promising results with accuracy of 0.74. The result encourages and motivates for future improvement and research on online diagnosing of melanoma in early stages. Therefore, a web application was built to utilize the model and provide online diagnosis of melanoma.

Keywords—*Convolutional neural network CNN; melanoma; skin cancer; image preprocessing*

I. INTRODUCTION

Skin cancers (lesion cancer) are the most common of all cancers, where abnormal growth of skin cells with the ability to spread to neighboring cells or other parts in the body. In most cases, skin cancers are caused due to exposure to ultraviolet radiation from the sunlight or from some cosmetic devices such as tanning beds. Skin Cancers can be categorized into three main types: Melanoma, Squamous Cell skin Cancer (SCC), and Basal Cell skin Cancer (BCC). Melanoma is the most damaging and common type of skin cancer and can appear on eyes or mouth, where SCC and BCC are less common and do not spread to other areas and unlikely to cause death, Fig. 1 shows samples of skin cancer images [1].

The main symptoms of melanoma cancer is the appearance of large brown spot, spread by other patches smaller and darker color, changes in color and size of moles in the body, or the descent of blood from them, as well as, the appearance of dark small and irregular edges lesion, usually on the hands and feet [2]. The American cancer society expect about hundred thousand new melanomas cases in United States in 2019, 60% are men and 40% are women, and about thousands of skin cancer patients are expected to die due to melanoma in US during 2019 [3], in addition, the American cancer society found that the percentage of diagnosed skin cancers increased 54% for the period between 2008 to 2019, and likely to diagnose 104,350 new cases in 2019 in U.S, and the estimated death is 11,650 where 8,030 are male and 3,620 are female. Therefore, there is a tremendous need for computer application to provide rapid and initial diagnoses advice for potential skin cancer [4].

Diagnosing skin cancers in early stages could increase the survival rate and save medical cost and treatment for potential patients. Computer-Aided application for diagnosing skin cancers can contribute to timely detection of skin cancers before it is too late, and encourage potential patients to seek further pathology test and physicians advices.

The purpose of this paper is to explore the efficiency of using Deep Learning Neural Network in medical images, practically, skin cancers images. The aim is to develop a computer application that receive captured image from end-user, then response to end-user with a set of recommendations such as the percentages of skin cancer in the captured image and a set of procedures to follow. Early diagnosis and detection of skin cancer can prevent patients from consequences such as surgical procedure to remove skin cancer as shown in Fig. 2. The rest of the paper is organized as the following: Section 2 express the literature review about the subject and previous attempts to utilize medical images and neural networks for the purpose of diagnosing melanoma, Section 3 describe the uses of neural network for medical images analysis and prediction, Section 4 is designated to present the methodology and material used in the research, Section 5 is to discuss the main results and finding obtained, and Section 6 is to conclude the main theme and results in this research.
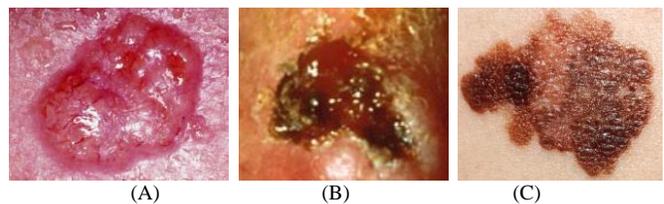

Fig. 1.   Sample Skin Cancer Images, (A) Basal Cell Skin Cancer, (B) Squamous Cell Skin Cancer, and (C) Melanoma.
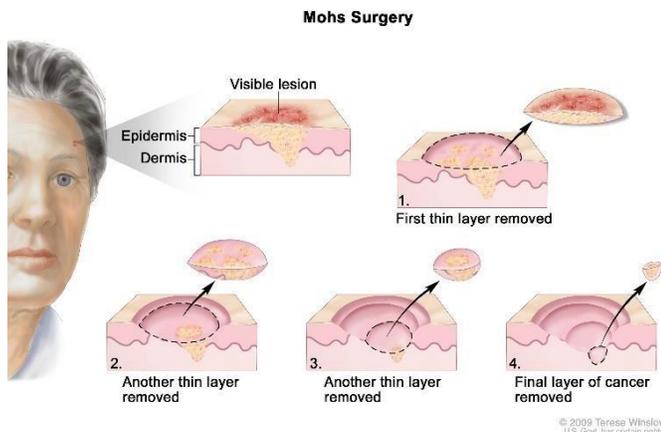
Fig. 2. A Surgical Procedure Steps to Remove Skin Cancer [2].

II. NEURAL NETWORK FOR MEDICAL ANALYSIS

Computer vision is a way to understand or simulate human vision of things (objects, events, motion, etc.) and attempt to make computer to do the same, by simulating brain images and videos processing that come from the retina. Computer vision can be defined as a branch of Artificial Intelligence (AI) that focus on using geometry to transform digital image or sequence of images (video) into high dimensional data, and then extract features, such as texture, points, lines, edges, shapes, motions, and ridges. A common application of computer vision is medical image processing, which aims to obtain significant information from images to automate patient's diagnosis. An example of this automation is computer-aided detection of breast cancer and skin cancer [7].

Artificial Neural Network (ANN) is the very fundamental of deep learning. ANN originally presented by the neuro physiologist Warren McCulloch and the mathematician Walter Pittsin 1943. They proposed the first neural network architecture (computational model) inspired from human biological neuron (Fig. 3) of how biological neurons are connected together in animals' brains to transmit data and accomplish complex computations tasks. In 1960, further ANN architectures contributed to believe in machine intelligent, then ANN entered forgetting era till 1980 when gained interest again, but 1990 shown newly invented alternative machine learning techniques much powerful than ANN such as SVM (Support Vector Machines). Recently, ANN gained another life chance due to the huge amount of quality data and significant increase in computers storage and power which assisted ANN to produce better results. Nowadays, ANN and its related techniques are considered as a powerful machine learning tool to tackle large and highly complex tasks, for example, Google uses deep learning to classify tremendous number of images, Apple's Siri speech recognition tool, and YouTube for suggesting best related videos for the big number of users every day [5], [6].

Computer vision and image classification problem were among the main factors of development of deep learning, because AI researchers aimed to learn features from raw of images which require higher computations and memory resources in compare to text and numerical learning models. In particular, the network called AlexNet has been used and modified to solve computer vision and image classification problems which draw the map for the Convolutional Neural Network (CNN) [8].

The increased popularity in Convolutional Neural Networks in medical analysis and computer vision is due to its outstanding performance in analyzing and classifying images. Consequently, CNN became one of the most popular models in deep learning and computer vision. The key idea behind convolutional neural networks is to build partially connected layers. For example, an image with shape 100×100 which form 10,000 pixels as input to the network and suppose the first layer consist of only 1000 neurons then the number of connections between input layer and first hidden layer will be about 10 million connections, which require huge computations and memory. However, CNN can resolve this issue using partially connected layers. In CNNs, there are receptive fields to connect the input layer to a feature map. Receptive fields can be defined as overlapping windows that travel over the pixels of an input image to create a feature map. The shifting length in input image window and the size of window itself is determined during model design and implementation. The process of creating the feature map is also called convolution. An example of such a convolutional layer, the layer that connects the input pixels to each unit in the feature map, is shown in Fig. 4 [9], [10].

The convolutional neural network is typically consisting of three layers, convolution layer, pooling layer, and fully connected layer. In addition, CNN may contain optional layers such as dropout layer, however, convolution, pooling, and fully connected layers are the most popular architecture for CNN.
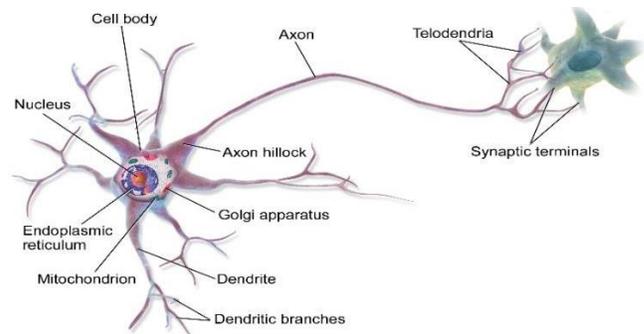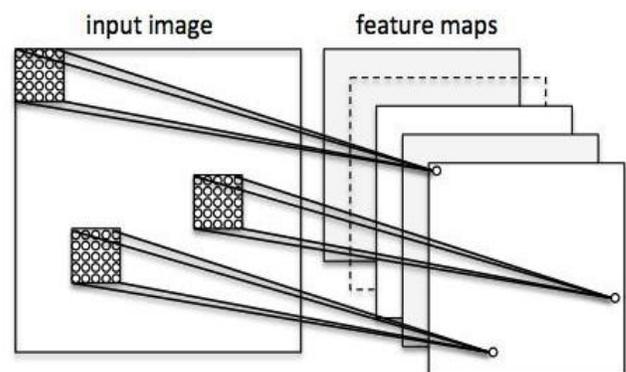

Fig. 3. Biological Neuron.

Fig. 4. An Example of a Convolutional Layer, the Layer that Connects the Input Pixels to Each unit in the Feature Map.

A. Convolution Layer

The convolution layer is the central structure for CNN and performs the most computational load. This layer performs a dot product between two matrices, the first matrix is the set filters also known as kernels and usually smaller in size than the image itself, for example, kernels for detecting vertical or horizontal edges. The second matrix is window matrix of the receptive field which is typically a portion of image with predefined size and sliding. Most of the images are composed of three channels (RGB), therefore, the kernel and receptive field will be also extending to three channels too. Initially, the kernel travel across input image with to produce smaller image that map input image with kernel, sample output of convolutional layer is shown in Fig. 5 [11].

B. Pooling Layer

The main purpose of pooling layer is to reduces the size of output which produced from previous layer (convolution layer), reducing the size of features will reduce the computation time and makes features robust against noise and outliers. There are several ways to do pooling such as max pooling and average pooling, however, the most popular is max pooling which calculate the maximum output from the neighborhood as shown in Fig. 6 [12]. Sample fully connected layer is shown in Fig.7.
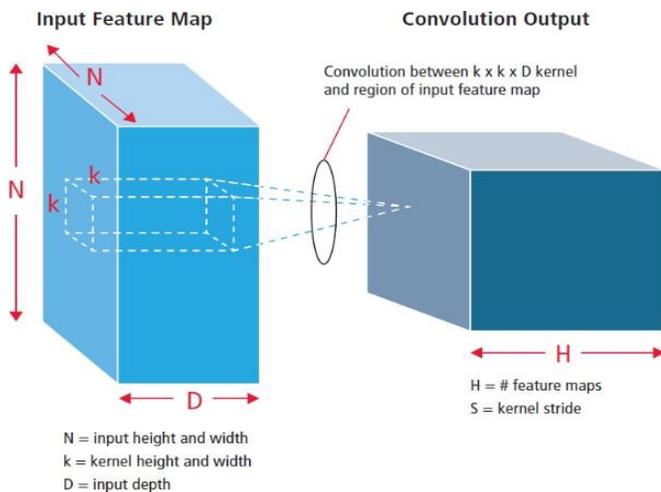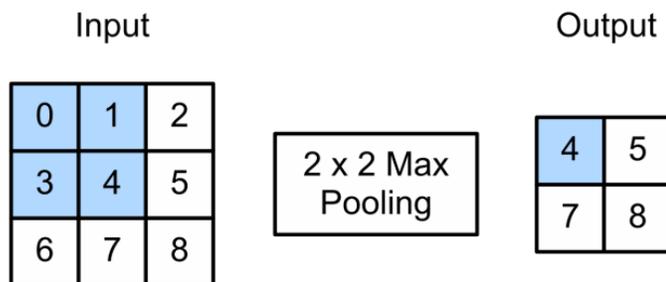

Fig. 5. Representation of Convolutional Process [11].

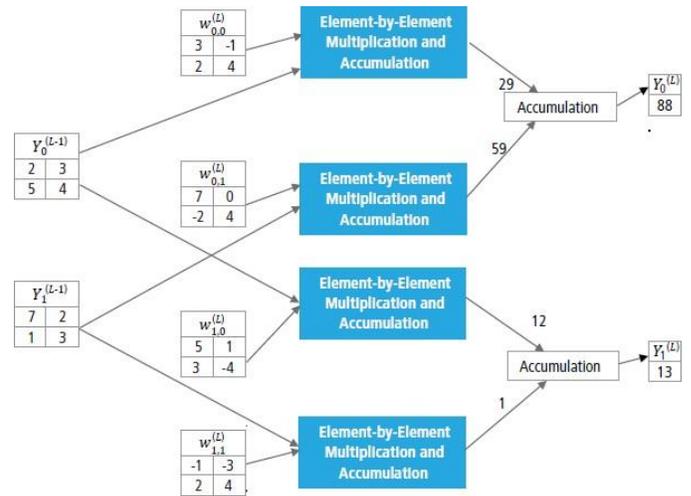Fig. 6. Maximum Pooling with a Pooling Window Shape of 2×2 [12].

Fig. 7. Sample Fully Connected Layer [12].

III. LITERATURE REVIEW

The literature shows several attempts to diagnose skin cancer cases using deep learning techniques, such as CNN. Esfahani et al. [13] proposed CNN architecture for diagnosing melanoma lesions, clinical images were preprocessed in order to reduce image illumination, then images fed to convolutional neural network models. The CNN model was successful to distinguish between malignant and benign images. Experimental results show that the proposed method was capable to diagnose melanoma lesions cases.

Giotis et al. [14] presented an expert system call MED-NODE to aid doctors in melanoma detection. The proposed system used extracted lesion regions in the image, then computes indicators such as the color and texture, along with visual attributes provided by experts. The proposed system performed comparably to state-of-the-art methods.

Mahbod et al. [15] shown that convolutional neural networks is superior over traditional methods. They proposed a a hybrid fully automatic computerized method for skin lesion classification, they used three pre-trained deep models (AlexNet, VGG16, ResNet-18) to extract features. The extracted features then are used to train SVM (support vector machine) classifiers and evaluated on the 150 validation images from the ISIC 2017 dataset, the proposed method is shown to achieve very good classification performance.

Jaisakthi, Chandrabose, and Mirunalini [16] proposed a method for skin lesion segmentation in images and to classify skin cancer types from images. The proposed method consists of preprocessing and segmentation using semi-supervised learning algorithm. The purpose of the first phase is noise removal using filtering technique, the second phase skin lesions are segmented based on clustering technique. The training images were downloaded from the ISIC 2017 challenge website, the experimental results shown low accuracy but will draw a map for future improvement.

IV. METHOD AND MATERIAL

A. Dataset

Dataset used in this manuscript was downloaded from the ISIC 2017: Skin Lesion Analysis Towards Melanoma Detection challenge website. This challenge provides a set of 2,000 publicly available dermoscopic images to participants. The data set also includes validation set with 150 images and test set with 600 images, Table I shows dataset statistics and Fig. 8 shows sample images from the dataset.

B. Methodology

Preprocessing data is among the most important phases in data science project, the idea is use a systemically approach to prepare data before feeding data into machine learning models. Segmentation is a popular preprocessing approach when the analysis and to detect objects and objects boundaries in digital images. In this work, Dataset images are RGB format (Red, Green, Blue) and since RGB images are more linked with the amount of light and illumination; it make it difficult to extract image features and boundaries, therefore, all images have been converted into HSV color space (Hue, Saturation, Value) which are more useful and relevant for objects detections in digital images. The second phase in dataset preprocessing is to apply Bilateral filter to all images, to keep sharp edges in the image because Bilateral filter substitutes the intensity of each pixel with a weighted average of intensity values from neighboring pixels [16]. The third step is to convert images into grayscale images to reduce the complexity and dimension of images then automatically detect edges of the objects in an image using Canny edge detection method. Final step of preprocessing is to create a mask for each image and then apply Bitwise method to extract the desired object in image; Fig. 9 shows an example of prepressing one image.

Since the dataset is not balanced (1626 benign and 374 malignant), augmentation technique used to generate and boost training images. Therefore, we used image zoom and image shift methods to produce 3000 images for each class label, Fig. 10 shows sample augmentation for one of the images.

After preparing the dataset, the next phase is to feed training dataset into Convolution Neural Network Model. Fig. 11 shows the architecture of CNN, which consist of three convolution layers, three max pooling layers, and four fully connected layers.

TABLE. I.   DATASET DESCRIPTION

Cancer Diagnosis	Count		Total
	Nevus	Seborrheic	
Benign			1626
	1372	254	
Malignant	374		374

Melanoma

Nevus

Seborrheic keratosis
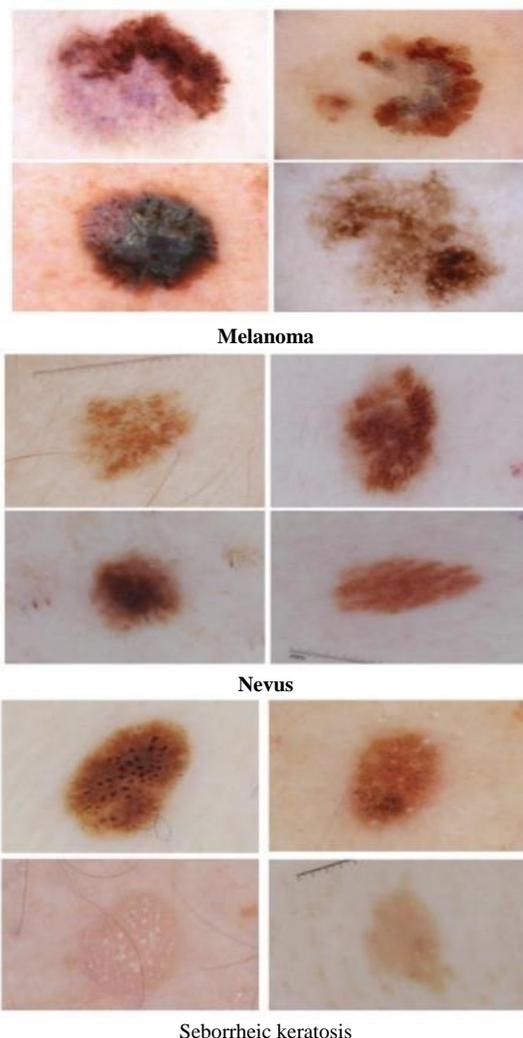
Fig. 8.   Sample Images from the Dataset.
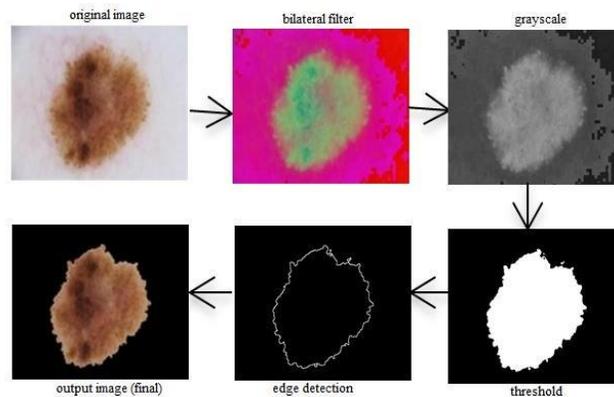
Fig. 9.   Dataset Preprocessing Steps.

V. RESULTS AND DISCUSSION

The available dataset was applied to the proposed model for diagnosing skin cancers using convolutional neural network. Number of epochs was set to 25. Accuracy on first epoch was and increased slightly in almost every following epoch to reach maximum accuracy of 0.74 on epoch number 25 which shows a satisfactory accuracy for diagnosing melanoma as shown in Fig. 12. In summary, our method shown pleasing but not superlative results, the importance of this work is to show the ability of early detection of melanoma using deep learning techniques such as convolutional neural network, which could draw the map for future research and investigation.

The trained CNN model has been saved in *h5* file format to store the weights and model configuration in a single file, which can be used in webserver to receive new unseen cases and produce an output or diagnosis online for users.

Fig. 10. Image Augmentations for Boosting Training Dataset.
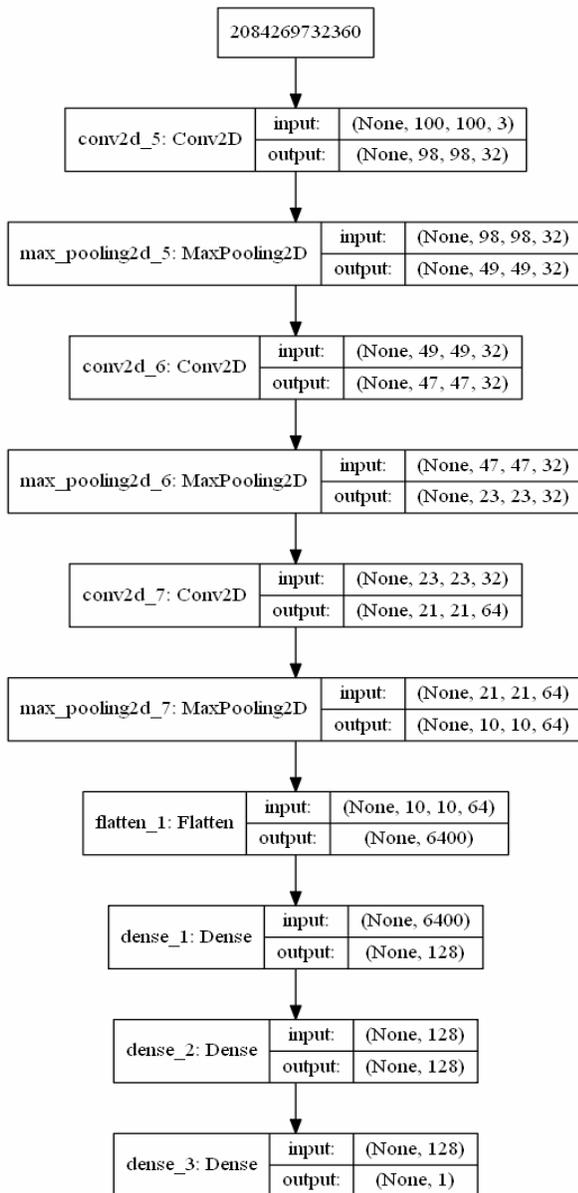
Fig. 11. CNN Architecture Model for Diagnosing Melanoma.
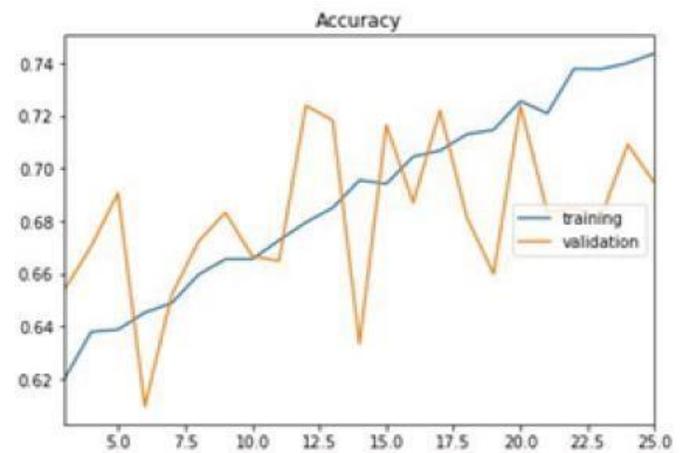

Fig. 12. Model Accuracy over 25 Epochs.

VI. CONCLUSION

In this paper, a computational method based on deep learning (Convolutional Neural Network) was implemented, which utilized 2000 images provided by ISIC (International Skin Imaging Collaboration). The proposed method includes images preprocessing for extracting the region of interest in the image itself, and then augmenting some images to produce a bigger dataset which contains 3000 images for each class. The resulting dataset have been applied into CNN model to train the model, which comprise several layers such as convolution layers, pooling layers, and fully connected layers. Testing the model produced promising results with accuracy of 0.74. The result encourages and motivate for future improvement and research for online diagnosing of melanoma in early stages before it is too late. Future work of current research is to investigate and restructure the CNN architecture to increase accuracy, obtain more images data for training, apply new augmentation algorithms to train the model using more data, and the ultimate future plan is to make this model accessible and useable using smart phones applications.

REFERENCES

[1] Dunphy, L. M., Winland-Brown, J., Porter, B., & Thomas, D. (2015). Primary care: Art and science of advanced practice nursing. FA Davis.

[2] National Cancer Institute (2013), Skin Cancer Treatment Health Professional Version, Retrieved 2019 from https://www.cancer.gov/types/skin/hp/skin-treatment-pdq

[3] Siegel, R. L., Miller, K. D., & Jemal, A. (2019). Cancer statistics 2019. CA: a cancer journal for clinicians, 69(1), 7-34.

[4] American Cancer Society (2019), Cancer Facts & Figures 2019.

[5] Géron, A. (2017). Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc.".

[6] McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, 5(4), 115-133.

[7] Nixon, M., & Aguado, A. S. (2012). Feature extraction and image processing for computer vision. Academic Press.

[8] Zhou, S. K., Greenspan, H., & Shen, D. (Eds.). (2017). Deep learning for medical image analysis. Academic Press.

[9] Raschka, S. (2015). Python machine learning. Packt Publishing Ltd.

[10] Géron, A. (2017). Hands-on machine learning with Scikit-Learn and TensorFlow: concepts, tools, and techniques to build intelligent systems. " O'Reilly Media, Inc."

[11] Hijazi, S.L., Kumar, R.R., & Rowen, C. (2015). Using Convolutional Neural Networks for Image Recognition By.

[12] A. Zhang, Z. Lipton, M. Li, A.J. Smola (2019), Dive into Deep Learning, https://www.d2l.ai/.

[13] Nasr-Esfahani, E., Samavi, S., Karimi, N., Soroushmehr, S. M. R., Jafari, M. H., Ward, K., & Najarian, K. (2016, August). 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) (pp. 1373-1376). IEEE.

[14] Giotis, I., Molders, N., Land, S., Biehl, M., Jonkman, M. F., & Petkov, N. (2015). MED-NODE: a computer-assisted melanoma diagnosis system using non-dermoscopic images. Expert systems with applications, 42(19), 6578-6585.

[15] Mahbod, A., Schaefer, G., Wang, C., Ecker, R., & Ellinge, I. (2019). Skin lesion classification using hybrid deep neural networks. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1229-1233). IEEE.

[16] Jaisakthi, S. M., Chandrabose, A., & Mirunalini, P. (2017). Automatic skin lesion segmentation using semi-supervised learning technique. arXiv preprint arXiv:1703.04301.