jet INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2019, VOL. 65, NO. 4, PP. 597-602
Manuscript received May 13, 2019; revised September, 2019. DOI: 10.24425/ijet.2019.129818

Skin Lesion Analysis Toward Melanoma Detection Using Deep Learning Techniques

Fatma Sherif, Wael A. Mohamed, and A.S. Mohra

***Abstract*—In the last few years, a great attention was paid to the deep learning Techniques used for image analysis because of their ability to use machine learning techniques to transform input data into high level presentation. For the sake of accurate diagnosis, the medical field has a steadily growing interest in such technology especially in the diagnosis of melanoma. These deep learning networks work through making coarse segmentation, conventional filters and pooling layers. However, this segmentation of the skin lesions results in image of lower resolution than the original skin image. In this paper, we present deep learning based approaches to solve the problems in skin lesion analysis using a dermoscopic image containing skin tumor. The proposed models are trained and evaluated on standard benchmark datasets from the International Skin Imaging Collaboration (ISIC) 2018 Challenge. The proposed method achieves an accuracy of 96.67% for the validation set .The experimental tests carried out on a clinical dataset show that the classification performance using deep learning-based features performs better than the state-of-the-art techniques.**

***Keywords*—melanoma, Skin Cancer, convolutional neural network, deep learning**

I. INTRODUCTION

Malignant Melanoma is consider the most deadly forms of skin cancer and accounts for about 75% of skin cancer-related deaths [1]. According to the statistical data from the World Health Organization [2], between two and three million non-melanoma skin cancers and 132000 melanoma skin cancers occur globally every year. Precise identification of skin cancer at an early stage can greatly increase the survival rate of patients.

The Dermatoscopy technique has been developed to enhance the diagnostic performance of skin cancer. Dermoscopy is an expanded skin imaging technique to get a magnifying and luminous image of the skin area to increase spot clarity [3] and enhancing the visual impact of the skin lesion by removing the surface reflection. There are many advanced dermoscopic approaches, like ABCD rule [4], and the 7-point checklist [5]. Among these clinical analysis approaches, studies have shown that pattern analysis yields higher diagnostic performance over alternative ways [6].

However, automatic identification of skin cancer from the examination images is still a difficult task, because it faces many challenges. First, the low contrast between the skin lesions and the normal skin area makes it difficult to divide the lesion areas. Second, the incidence of melanoma and melanoma lesions may have a high degree of visual similarity, making it difficult to distinguish between melanoma lesions and melanoma. Thirdly, variations in skin conditions, such as skin color and natural hair, among the patients produce different manifestations of malignant melanoma, in terms of color, texture, etc.

In the last few years deep learning had gained popularity in feature learning and object classification and detection. The deep learning has been used on biomedical database, like skin cancer detection [7]. The diversity of features which can be detected by the different convolutional layers enables the network to handle large variations in the dataset. It permits the feature detection to be handled automatically, therefore ameliorating the difficulties of feature detection inherent in conventional pattern analysis techniques.

II. RELATED WORK

Many research works have been done using Computer vision and Image processing to detect malicious melanoma. In most cases, the emergence of an automatic learning model called deep learning has enabled the development of medical image analysis approaches that can display remarkable accuracy, to some extent raise concerns about the future of the radiologist Human [8]. Convolutional neural networks created promising results in the classification of skin lesions.

Esteva et al. [7]: compared the accuracy of deep learning to many dermatologists when classifying images of skin lesions. A total of 129450 images were used to train the network particularly on skin lesions, after pre-training on 1.28 million images from the ImageNet dataset. Network accuracy reached 72.1%, which is at least as good as the average of 23 approved dermatologists.

Nylund [9]: had achieved 89.3% accuracy by using an ImageNet dataset pre-trained network. He used over 20000 images from many different datasets, were used to retrain the network.

Mirunaliniy [10]: had used an automatically classifying system which uses the image representation gained from the dermoscopic through Google inspection model. They had achieved 65.8% as an overall AUC score through the validation set provided in ISBI challenge.

Kawahara et al. [11]: used a pre-trained ConvNet as a feature extractor rather than training a CNN from beginning. It demonstrates the use filters from a CNN pre-trained on original images generalize to classifying ten classes of non-dermoscopic skin images. This method achieves an accuracy of 81.8%.

The authors are with Faculty of Engineering/Benha University, Qalyubia, Egypt (e-mail: fatmasherif4@yahoo.com, Wael.ahmed@bhit.bu.edu.eg, amohra@bhit.bu.edu.eg.).

Haenssle et al [12]: A GoogLeNet Inception v3 model was adapted for skin lesion classification with transfer learning, whereby the weights were fine-tuned in all layers. The analysis was limited to melanoma versus benign nevi and the AUC ROC achieved for this task was 0.86.

In the study by Ridell and Spett [13], a CNN was trained based on Google Inception v3 so as to detect melanoma. It was then investigated how the accuracy of classifying between benign nevus and melanoma is affected by the size of the training dataset. Multiple image sizes were tested starting from 200 to 1600 images. This method achieves accuracy between 70.8% and 77.5.

Codella et al. [14]: using CNN to extract image descriptors by using a pre-trained model from (ILSVRC) 2012 database. They also investigate the 50 most recent network structures to win the ImageNet recognition challenge known as (DRN). The proposed system produces accuracy (76%).

III. THEORETICAL BACKGROUND

Deep learning is a new field of Machine learning based on algorithms aimed at devising features of sensory signals and helping to understand data, such as images, text and speech [15]. Deep learning structures consist of multiple levels with many hidden layers. His goal is through algorithms to indicate the machine how to change its parameters used in each layer of those in the previous layer. In a simple case, a group of neurons receives an input signal and passes one layer, producing another set of output signal. In fact, there is usually more than one hidden layer and each output is used as inputs for the next layer. Several hidden layers offer the advantage of solving complex pattern recognition problems, but are often difficult to train [16]. Therefore, according to the problem, a different number of hidden layers are required.

Supervised learning is a technique used by deep learning to train the weights, where all the training samples are labeled. Unsupervised learning is another technique, wherever all the training samples are not labeled and primary objective is to find the structure in the data. Unsupervised learning is another way, where all training samples are not labeled.

Fine-tuning is another commonly used method for determining weights, where pre-trained weights are available and used as a starting point, and weights are then set to a new data set called Transfer learning. This method leads to the train faster than starting from a random starting point, achieving better accuracy.

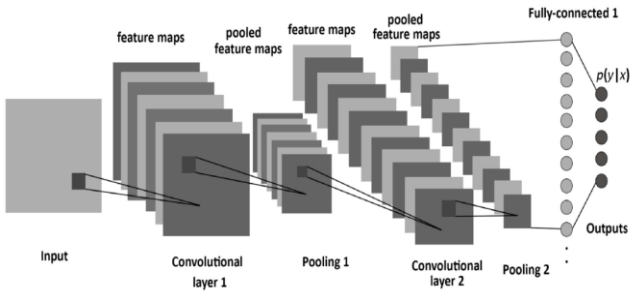

Fig. 1. Example of CNN architecture [17].

Convolutional Neural Networks (CNNs) a specific type of deep learning algorithm. They are very similar to traditional neural networks. But they arrange their neurons in 3 dimensions (width, height, and depth). The neuron inside the layer is also connected to a small area of the layer before it, and not just as connected to the traditional neural network. The structure of CNN networks consists of many different types of serial layers (convolution layers, pooling layers, non-linear layers, and fully connected layers), as shown in Fig. 1.

IV. PROPOSED SYSTEM

This section presents technique that was used in developing CNN model to classify skin legion as Melanoma and benign. The proposed System Architecture can be seen in Fig. 2.

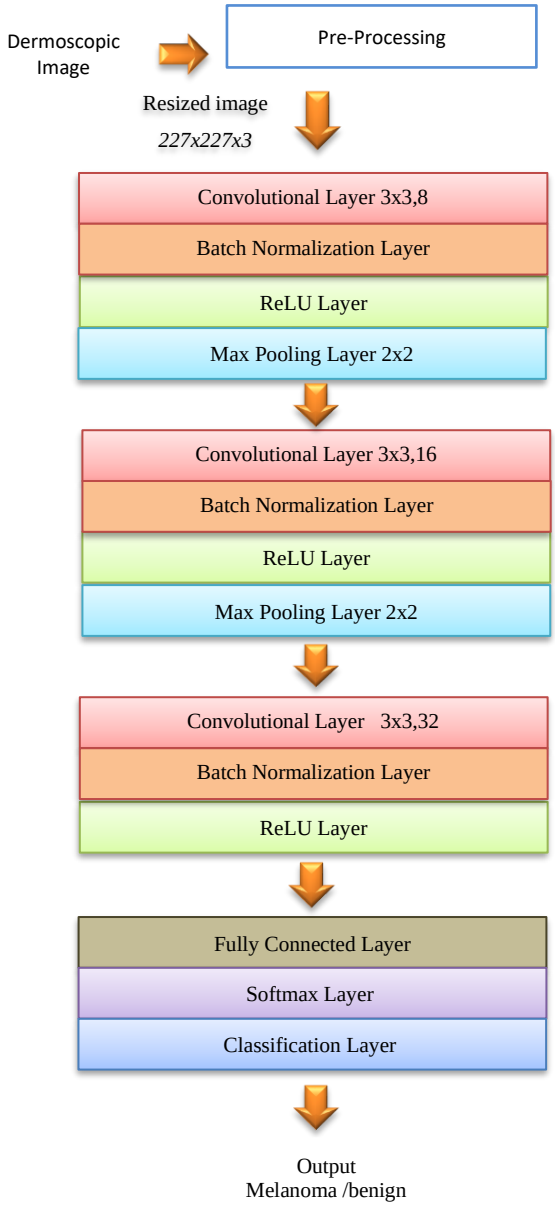

Fig. 2. The flowchart of proposed system.

A. Pre-Processing

The original training set contains images of skin lesions of various resolutions. Some of these images have a resolution higher than 900 × 750 that needs high-cost computation. Therefore, there is a need to rescale the lesion images for the deep learning network. To avoid the distortion in the form of the skin lesion caused by resizing the direct image, the central area

of the lesion image was first cropped and then proportionally size of the area was reduced to a lower resolution. All dataset images have been resized to a fixed value of 227 × 227.

B. CNN *Architecture*

In this work, CNN is used as a deep learning framework for the automated detection of malignant melanoma. CNN networks benefit from a range of convolves filters. They will examine different structures in the input images. Thus, when using CNN, the input is the image itself and also the network automatically extracts the suitable aspects of the image. The input to the proposed CNN network consists of RGB images of 227 × 227 × 3 size. These numbers correspond to the length, width, and channel size. The data set consists of color images, so the channel size is 3. Our CNN model consists of multiple layers.

Convolutional Layer

This layer is considered as the core building block of a CNN. The parameters of this layer consist of a group of learning filters. Every filter is small in spatial terms, it's filtered over the input size, and the point products between this size and the filter are calculated in any position, resulting in a two-dimensional activation map. In every convolutional layer, every filter will produce a unique activation map and all of them will be merged to produce output size. In our model, the convolutional layers have a filter size of 3x3 and have 8, 16, and 32 filters, respectively.

Batch Normalization Layer

It normalizes activation and gradients that spread across the network, making network training an easier improvement problem. Used to speed up network training and reduce network configuration sensitivity.

ReLU Layer

This layer is a nonlinear activation function. The most common activation function is the rectified linear unit (ReLU).

Max Pooling Layer

This layer performs a reduction operation along the weight and length resulting in a smaller representation and thus becoming more manageable, as seen in fig. 3. Max Pooling layers are usually between convolutional layers with the aim to reduce the number of parameters. The operation takes place in every activation map independently and by using the MAX function, it resizes it spatially. In this system, the Max Pooling Layers consists of size 2x2 and of strides 2.

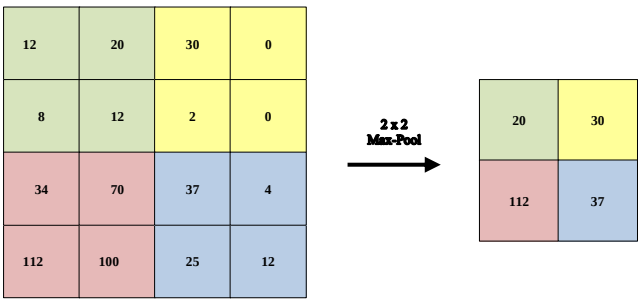

Fig. 3. Examples of how max pooling operates [18]

Fully connected layer

Is a layer in which the Neurons fully connected to all activation processes in previous layers. This layer combines all the features learned by the previous layers across the image to identify the larger patterns. In this work, The 2 layer feed forward neural network consists of 100 neurons in the hidden layer and 2 neurons in the output layer.

Softmax Layer

The softmax activation function normalizes the output of the fully connected layer.

Classification layer

It uses the possibilities returned by the Softmax activation function for every input to assign the input to a mutually exclusive category and the loss account. In this model, we use the SGDM optimizer to minimize the error function.

V. RESULTS AND DISCUSSION

A. Datasets and Evaluation

The proposed network was trained using 700 digital images (350 each for malignant and benign) as a dataset of skin lesion images from the ISIC 2018 Challenge [19] in JPEG format. To measure the performance of the proposed system, Classification accuracy was used as follow:

$$Classification\ Accuracy = \frac{\text{True detected cases}}{\text{All cases}} \quad (1)$$

B. Optimized Network Based on Quantitative Test Results

In this section, the best test results and performance evaluation created after training the optimized network structure are outlined, which were illustrated by the quantitative test results in the next section.

Experiments were performed on a core i5, 2.27GHz processor with 8GB RAM. We are using MATLAB® 2017b as programming language. The best results found in this study can be seen in Table I.

TABLE I
RESULTS FOR THE OPTIMIZED NETWORK, BASED ON RESULTS FROM QUANTITIES TESTING

Parameters	Value
Total Number of samples	700
(Train to test) ratio	(70% to 30%)
Number of epochs	40
Learning rate	0.001
Validation accuracy	96.67%

The training progress and results are shown in Fig. 4. The figure plots:

Training accuracy (Classification accuracy on each individual mini-batch**)**

Smoothed training accuracy (Smoothed training accuracy obtained by applying a smoothing algorithm to the training accuracy**)**

Validation accuracy (Classification accuracy on the entire validation set**)**

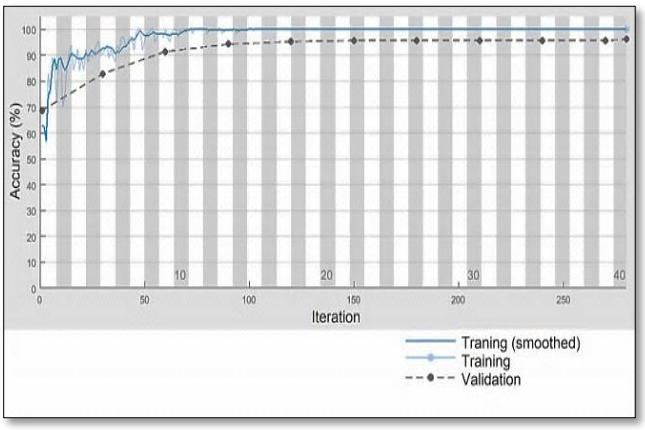

Fig. 4. CNN training run, a validation accuracy of 96.67% is reached.

C. Quantitative Test Results

This section displays results when we change the different hyper-parameters for the proposed CNN model. This part of the paper was done primarily to study the effect of hyper-parameters on the accuracy.

1) (Train and Test) ratio

Train and test ratio is an important aspect. Ensure that the training data set should include all possible patterns used to identify the problem, large enough to achieve statistically significant results, and represent the data set as a whole. In other words, do not choose a test group with different properties than the training group.

In Table II, the dataset of ISIC was used so as to show how the different values of train and test ratios affect the accuracy, where the dataset is randomly split. Our system achieved the best accuracy of 96.19% at the train to test ratio is (70 % to 30 %). At this stage our CNN model was trained for 50 epochs and learning rate of 0.01.

2) Epochs

Table III shows the results obtained using the CNN model during the learning procedure. Fifty epochs, which were the initial setting for all test cases, required longer time to train network Compared to 40 epochs (which have the same value of accuracy). Forty epochs were chosen because they have a good balance between accuracy and time spent on network training.

3) Learning Rate

Table IV shows the effect of alternating the learning rate. At this stage our CNN model was trained for 40 epochs and (Train to Test) ratio is (70 % to 30%). For this network, a learning rate of 0.001 achieved the best results.

TABLE II
PERFORMANCE OF DIFFERENT TRAIN TO TEST RATIOS

(Train to Test) ratio	Accurcy
(50% to50%)	96.13%
(60% to40%)	95%
(70% to30%)	96.19%
(80% to20%)	95%
(90% to10%)	92.88%

TABLE III
RESULTS OVER TRAINING FOR VARYING THE NUMBER OF EPOCHS IN PROPOSED SYSTEM, USING LEARNING RATE IS 0.01, AND (TRAIN TO TEST) RATIO IS (70 % TO 30 %)

Epoch	Accurcy	Training time
1	70.95%	5 min 56 sec
10	95.24%	20 min 10 sec
20	95.71%	31 min 19 sec
30	95.74%	51 min 14 sec
40	96.19%	73 min 28 sec
50	96.19%	92 min 3 sec
60	95.24%	179 min 4 sec

TABLE IV
TEST RESULTS FOR VARIOUS LEARNING RATES

learning rate	Accurcy	Training time
0.1	89.05%	61 min 23 sec
0.01	96.19%	73 min 28 sec
0.001	96.67%	95 min 15 sec

D. Comparison of our model with other deep learning algorithms

Our proposed system outperforms other challenging approaches and other deep learning systems. We achieved accuracy of 96.67% compared to 89.3% achieved by Nylund [8] using an Image-Net dataset pre-trained network. He used more than 20000 images from many different datasets to retrain the network. In Table V shows the comparison between previous research accuracy and accuracy in this paper. As shown, our proposed system achieved the highest accuracy with strong performance.

TABLE V
QUANTITATIVE COMPARISON OF SKIN LESION CLASSIFICATION USING DEEP LEARNING TECHNIQUES

Methods	Accurcy
Ramlakhan et al. [20]	66.7 %
Kawahara et al. [11]	81.8 %
Nylund [9]	89.3%
Menegola et al. [21]	79.2 %
Burdick et al [22]	69.3%
Proposed system	96.67%

E. Comparison of deep learning with conventional algorithm

Finally, our method is compared with STOLZ's method [4], which is a conventional method of melanoma classification. Comparison is done according to the same dataset (700 samples). In Fig. 5 showing Confusion Matrix of STOLZ's method, we achieved an accuracy of 76.60%.

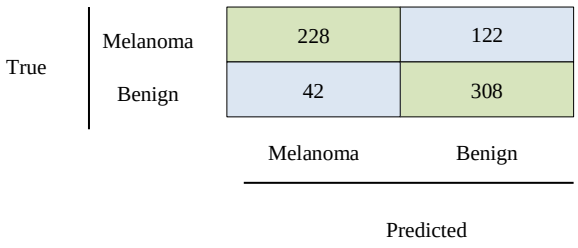

Fig. 5. Confusion Matrix for using STOLZ's method.

As can be seen in Table VI our method has a better Accuracy in classification. The segmentation algorithm can be misdirected by skin artifacts and the complex skin lesion pattern The STOLZ's method has a lower performance in extraction of lesions' borders and some pixels around the lesion's boundary are misclassified.

TABLE VI
QUANTITATIVE COMPARISON OF SKIN LESION CLASSIFICATION RESULTS

Classification Technique	Accuracy
STOLZ's method	76.60%
Proposed Method	96.67%

VI. CONCLUSION

For accurate detection of skin cancer images, accurate identification of the lesion area is of great importance. In this paper a method based on deep learning networks was presented for extraction of Melanoma in clinical images. The proposed deep learning structure showed the ability to detect melanoma cases from benign ones. In this work, a convolutional neural network is proposed to classify 2 types of the skin lesion in dermscopic images. Our approach is used the official ISIC 2018 dataset to train and validate the proposed deep learning model. Experimental results showed our better accuracy of 96.67%, as compared to other classification methods.

REFERENCES

[1] A.F. Jerant, J.T. Johnson, C.D. Sheridan, and T.J. Caffrey, "Early Detection and Treatment of Skin Cancer," Am. Fam. Physician, 62 (2): 357–68, 375–6, 381–2, 2000.
[2] "World Health Organization," Available: https://www.who.int/en/, [Accessed: 2018-09-10].
[3] M. Binder, M. Schwarz, A. Winkler, A. Steiner, A. Kaider, K. Wolff, and H. Pehamberger, "Epiluminescence microscopy. A useful tool for the diagnosis of pigmented skin lesions for formally trained dermatologists," Arch. Dermatol, 131(3):286-91, 1995.
[4] F. Nachbar, W. Stolz, T. Merkle, A. Cognetta, T. Vogt, and M. Landthaler, "The abcd rule of dermatoscopy. High prospective value in the diagnosis of doubtful melanocytic skin lesions," Journal of the American Academy of Dermatology, 30(4):551-9, 1994.
[5] G. Argenziano, G. Fabbrocini, P. Carli, V. De Giorgi, E. Sammarco, and M. Delfino, "Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: comp. of the abcd rule of dermatoscopy and a new 7-point checklist based on pattern analysis," Archives of Dermatology, 134(12):1563-70, 1998.
[6] P. Carli, E. Quercioli, S. Sestini, et al, " Pattern analysis, not simplified algorithms, is the most reliable method for teaching dermoscopy for melanoma diagnosis to residents in dermatology," Br. J. Dermatol, 148(5), 981–984, 2003.
[7] A. Esteva, B. Kuprel, R. Novoa, et al, "Dermatologist-level classification of skin cancer with deep neural networks," Nature, 542, 115–118, 2017.
[8] M. Walter, "Is this the end? machine learning and 2 other threats to radiologys future," goo.gIIM9X3SF, 2016, Available: https://www.radiologybusiness.com/topics/technology-management/end-machine-learning-and-2-other-threats-radiologys-future, [Accessed: 2018-09-08].
[9] A. Nylund, "To be, or not to be Melanoma: Convolutional neural networks in skin lesion classification," In: Dissertation (2016). Available: http://kth.diva-portal.org/smash/get/diva2:950147/FULLTEXT01.pdf. [Accessed: 2018-08-12].
[10] P. Mirunaliniy, A. Chandrabose, V. Gokul y, S. M. Jaisakthi." Deep Learning for Skin Lesion Classification", arXiv preprint arXiv: 1703.04364, 2017, Available: https://arxiv.org/pdf/1812.02316.pdf, [Accessed: 2018-09-15].
[11] J. Kawahara, A. BenTaieb, and G. Hamarneh, "Deep features to classify skin lesions," published in Conference: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague, Czech Republic, 2016. [DOI: 10.1109/ISBI.2016.7493528].
[12] H. Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl, and A. Blum, "Man against machine: Diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists," Ann Oncol: 29(8):1836-1842, 2018. [DOI: 10.1093/annonc/mdy166] [Medline: 29846502].
[13] P. Ridell, H. Spett, P. Herman, and Ö. Ekeberg, "Training Set Size for Skin Cancer Classification Using Google's Inception v3," 2017, Available,http://www.diva-portal.org/smash/get/diva2:1112097/FULLTEXT01.pdf, [Accessed: 2018-11-12].
[14] N. Codella, Q.-B. Nguyen, S. Pankanti, D. Gutman, B. Helba, A. Halpern, and J. R. Smith, "Deep learning ensembles for melanoma recognition in dermoscopy images," arXiv preprint arXiv: 1610. 04662, 2016.
[15] L. Deng, and D. Yu, "Deep Learning: Methods and Applications," Foundations and Trends in Signal Processing, 7, 3–4, 197–387, 2014.
[16] M. Nielsen, "Neural Networks and Deep Learning," Determination Press, 2015, Available: http://neuralnetworksanddeeplearning.com/, [Accessed: 2018-11-10].
[17] S. Albelwi and A. Mahmood, "A framework for designing the architectures of deep convolutional neural networks," Entropy, 19(6), 242, 2017.
[18] S. Hijazi, R. Kumar, and C. Rowen, "*Using Convolutional Neural Networks for Image Recognition*," Cadence Design Systems Inc., 2015.
[19] "International Skin Imaging Collaboration: Melanoma Project," Available: https://www.isic-archive.com, [Accessed: 2018-09-17].

[20] K. Ramlakhan, and Y. Shang, "A Mobile Automated Skin Lesion Classification System," Published in Conference: 2011 IEEE 23rd International Conference on Tools with Artificial Intelligence, Boca Raton, FL, USA, 2011. [DOI: 10.1109/ICTAI.2011.29].

[21] A. Menegola, M. Fornaciali, R. Pires, F. Bittencourt, s Avila, and E Valle, "Knowledge Transfer for Melanoma Screening with Deep Learning," Published in Conference: 2017 IEEE 14th International Symposium on Biomedical Imaging, Melbourne, VIC, Australia,2017. [DOI: 10.1109/ISBI.2017.7950523]

[22] J. Burdick, O. Marques, J. Weinthal, and B. Furht, "Rethinking Skin Lesion Segmentation in a Convolutional Classifier," Journal of Digit Imaging, 31(4), 435–440, 2018. [doi: 10.1007/s10278-017-0026-y].