Deep Learning for Two-Step Classification of Malignant Pigmented Skin Lesions

3 authors, including:

Parvaneh Esmaili
Near East University
19 PUBLICATIONS   **61** CITATIONS

Some of the authors of this publication are also working on these related projects:

Synchronous method for multiple robot manipulators View project

Wastewater View project

Deep Learning for Two-Step Classification of Malignant Pigmented Skin Lesions

Sertan Kaymak, Parvaneh Esmaili, and Ali Serener, *Member, IEEE*

Abstract—**Skin cancer is one of the most common types of cancer. Its early detection drastically improves outcomes and saves human lives. Well known skin cancer types are melanoma, basal cell carcinoma and squamous cell carcinoma. Melanoma is melanocytic malignant while basal cell carcinoma and squamous cell carcinoma are non-melanocytic malignant. Even though the diagnosis of these cancer types is done by a skin biopsy, automatic detection of skin cancer using computerized methods may lead to a faster and a more accurate diagnosis. The majority of automated skin cancer detection methods proposed by researchers so far concentrated only on melanocytic malignant type melanoma. Non-melanocytic malignant skin lesions could not be investigated in detail due to the lack of available datasets with different lesion classes. In this paper, an automatic detection of malignant pigmented skin lesions is investigated. For this, the two-step skin lesion diagnostic procedure of the dermatologists is followed. Using a deep learning model, the skin lesion is first classified as melanocytic or non-melanocytic and then malignant types are detected using other deep learning models. The performance evaluations show that melanocytic and non-melanocytic skin lesions are detected with the highest accuracy. They also show that melanocytic malignant skin lesions can be classified with a higher accuracy than non-melanocytic malignant skin lesions.**

Keywords—**Deep learning, dermoscopic skin images, malignant skin lesions, skin cancer, two-step diagnostic**

I. INTRODUCTION

Melanoma is one of the most deadly type of skin cancer. 75% of skin cancer related deaths in the world are caused by melanoma. As melanoma affects the melanocyte cells, it prevents the synthesis of the melanin [1]. Hence, detecting melanoma in the early stages becomes very important in combating cancer related deaths.

Skin lesions can be grouped into melanocytic and non-melanocytic types. Melanocytic lesions include melanoma and melanocytic nevi. Similarly, non-melanocytic lesions include basal cell carcinoma, actinic keratosis (earliest

Fig. 1. Melanoma skin lesion

form of squamous cell carcinoma and is also known as Bowen's disease), dermatofibroma and vascular lesion. Both melanocytic and non-melanocytic lesions have malignant and benign types. Malignant skin lesions are cancerous and include melanoma, basal cell carcinoma and actinic keratosis.

Dermoscopy is a non-invasive imaging tool widely used to analyze the skin lesions by morphological features. A sample dermoscopic image showing the melanoma skin lesion is given in Fig. 1. An important application of dermoscopy is the early diagnosis of melanoma, a melanocytic lesion that can occasionally appear non-melanocytic. In order to standardize the dermoscopic terminology, it was decided during the Consensus Net Meeting on Dermoscopy [2], [3] that dermoscopic classification of pigmented skin lesions should be carried out using a two-step diagnostic procedure. The first step of this procedure involves differentiating melanocytic lesions from non-melanocytic lesions and the second step includes differentiating melanoma from benign melanocytic lesions.

Due to some drawbacks, such as complexity, time consuming structure and variety in observing inter and intra features, this tool, however, may sometimes be inaccurate. Instead, computerized methods maybe used to improve the performance of dermoscopy [4]. A challenging example of this would be automated segmentation where computerized methods can be used to help categorize melanoma in the presence of different size, color, shape and variety in texture [5].

Computerized methods such as neural networks have indeed been used in the past for the classification of skin lesions using dermoscopic images. However, one of the main obstacles of such methods has been the limited availability of quality dermoscopic images involving different classes of skin diseases. As such, past research has mainly focused on melanocytic lesion detection [6]

Sertan Kaymak is with the Department of Electrical and Electronic Engineering, Near East University, Nicosia, Mersin 10, Turkey (phone: (+90) 392-223-6464 x5291; e mail: sertan.kaymak@neu.edu.tr).

Parvaneh Esmaili is with the Department of Electrical and Electronic Engineering, Near East University, Nicosia, Mersin 10, Turkey (phone: (+90) 392-223-6464 x5291; e mail: parvaneh.esmaili@neu.edu.tr).

Ali Serener is with the Department of Electrical and Electronic Engineering, Near East University, Nicosia, Mersin 10, Turkey (phone: (+90) 392-223-6464 x5291; e mail: ali.serener@neu.edu.tr).
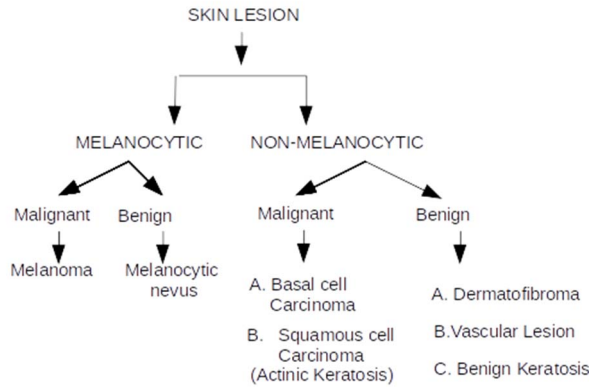
Fig. 2. 2018 ISIC challenge dataset skin lesion types

and avoided non-melanocytic lesions. It is well known that non-melanocytic lesions basal cell carcinoma and actinic keratosis are the other two main types of skin cancer besides melanoma. Therefore, it is crucial that non-melanocytic malignant legions are investigated using computerized methods as well.

The goal of this paper is to use AlexNet [7], one of the most well-known deep convolutional neural networks, along with the 2018 International Skin Imaging Collaboration (ISIC) Skin Lesion Analysis Towards Melanoma challenge dataset [8], the largest publicly available dataset of dermoscopic images, to detect malignant skin lesions using the two-step diagnostic procedure of expert dermatologists. While doing that, this research will take the second step of this procedure one step further and it will not only setup a deep learning architecture to detect melanocytic malignant lesions but it will use this architecture to detect non-melanocytic malignant lesions as well. Note that this dataset contains seven skin lesion types, namely melanoma, melanocytic nevus, basal cell carcinoma, actinic keratosis, benign keratosis, dermatofibroma and vascular lesion (see Fig. 2). This set of lesions cover more than 95% of the lesions faced in clinical settings.

The main contributions of this work are:

1. It improves upon the two-step dermoscopic procedure of dermatologists [3] by adding a supplemental step (non-melanocytic skin lesion classification) to it.
2. It uses dermoscopic images for malignant non-melanocytic pigmented skin lesion classification whereas past research used clinical images for such classification [9]. Clinical images may suffer from photographic variability.
3. It uses a much larger dermoscopic image dataset compared to past research (7285 images vs. 167 [10] or 107 [11] images) for more accurate non-melanocytic pigmented skin lesion classification.
4. It uses dermatofibroma as well as vascular lesion categories to more generalize the classification of non-melanocytic skin lesions. Even though these two benign non-melanocytic skin lesion categories are common in practice, [11] used a very few images of these skin lesions for classification (14 images out of 655 total). [10] and [12], on the other hand, have

totally omitted these categories in the binary non-melanocytic skin lesion classification.

The organization of this paper is as follows. Related work on skin lesion classification using machine learning techniques is given first. Then, the details of the research method and the database are described. Finally, the performance of the deep learning architecture is evaluated and discussed.

II. RELATED WORK

There are two main ways of detecting melanoma in dermoscopic images: conventional methods and deep learning methods. Among the conventional methods, early efforts to distinguish melanoma in the images considered low level based hand-crafted features. In [13], a gradient based algorithm and in [14], a wavelet network optimization algorithm were introduced based on feature selection. This algorithm attempted to find a proper number of features to improve the performance of melanoma identification. However, these features not only failed in identifying different varieties of melanoma but also differentiating the visual similarities between images with or without melanoma [1]. The research of [15] also focused on utilizing segmentation procedure first to identify melanoma by combining a neural network and a genetic algorithm. However, using low level based features lead to some limitations. Therefore, as an alternative solution during automation, border detection was proposed for dermoscopy. This method provided clinical features such as border irregularity and asymmetry information. The existence of low contrast among skin lesions and non-clear lesion borders makes the border detection method an open challenge for accurate identification of melanoma [16].

When compared with conventional methods, deep learning methods with hierarchical learning capabilities became a powerful tool in the detection of dermoscopic images [17]. Convolution neural networks (CNNs) are among the strongest methods in deep learning which are used for detection, classification and segmentation of melanoma. In [18], a hybrid method was proposed which consisted of a CNN and a support vector machine algorithm with a sparse coding method. This unsupervised hybrid method could help omit using annotated data for melanoma identification. An AlexNet based CNN was presented for feature extraction for melanoma in [7].

A fully convolutional deep learning algorithm with 19 layers was presented for segmentation of melanoma skin lesions in [4]. This work used a convolution operator in the last layers in order to achieve more accuracy. However, in some cases a fully convolutional network (FCN) has some drawbacks over CNN, such as a significant decrease in the number of parameters. It can help reduce memory usage but it will cause losing the flexibility. When compared with Maxpool operators, convolution operators run much slower. Hence, the training process will take much more time compare to a CNN network. In [4], two types of datasets, ISBI 2016 challenge and PH2, were used to analyze the proposed method. In [19], a fully convolutional residual network was proposed for melanoma detection utilizing low level hand crafted features. They investigated a fully convolutional recursive network (FCRN) for the

(a) Melanoma (b) Melanocytic Nevus (c) Benign Keratosis (d) Actinic Keratosis

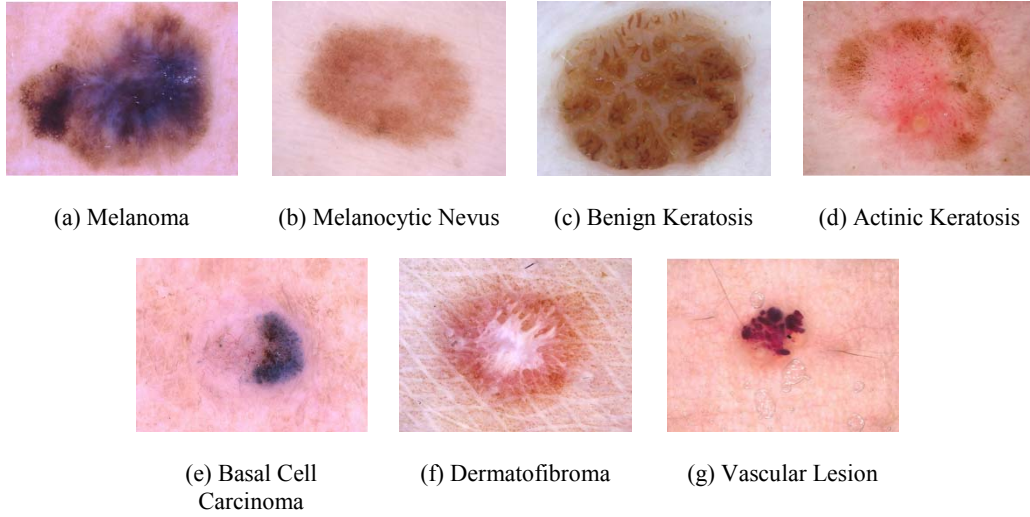(e) Basal Cell Carcinoma (f) Dermatofibroma (g) Vascular Lesion

Fig. 3. Skin lesions of 2018 ISIC challenge dataset

segmentation process and a deep neural network for classification. The experimental results were demonstrated on the 2016 ISBI dataset.

A deep convolution based neural network was investigated in [7] on a large dataset containing images with high resolution. In order to solve the over fitting problem, a network was introduced with fully connected layers. The experimental analysis was investigated on the ImageNet dataset.

A feature extraction method was proposed for melanoma detection which was based on wavelet and curvelet transformation techniques [20]. Also, a support vector machine classification was investigated by considering a random cross validation framework. The proposed framework was analyzed on the PH2 dataset.

Melanocytic and non-melanocytic skin lesions have been classified in [11] using a tumor area extraction algorithm. [9] used a deep learning algorithm for multi-class classification of 12 skin diseases using around 20000 clinical images. [12], on the other hand, carried out binary as well as multi-class classification of skin cancer using mostly clinical images and some dermoscopy images, making use of transfer learning to make the classification of clinical images more robust. In [10], a borderline detection algorithm was utilized first in order to classify melanocytic and non-melanocytic skin lesions using a layered and a flat model.

In order to validate the aforementioned methods, most of the researchers used the 2016 ISBI challenge dataset [21] (two classes, 900 dermoscopic images for training and 379 for testing) and PH2 dataset (three classes, 200 dermoscopic images). The 2017 ISBI challenge dataset [22] had about 2000 dermoscopic training images in three classes. This paper, on the other hand, uses the 2018 ISIC challenge dataset [8] with about 10000 dermoscopic images and seven different classes (Fig. 3). Hence, it is able to better utilize a deep learning architecture to apply the two-step diagnostic procedure of the dermatologists to detect melanocytic and non-melanocytic malignant skin lesions.

III. METHOD

The two-step deep learning classification method is achieved using the largest publicly available dermoscopic image dataset, 2018 ISIC challenge. The two-step method is based on generating three deep learning models. First, a deep learning model is generated to model melanocytic and non-melanocytic classes. The melanocytic class includes melanoma and melanocytic nevus images, whereas the non-melanocytic class includes basal cell carcinoma, actinic keratosis, benign keratosis, dermatofibroma and vascular lesion images. Table I lists the number of images for each class.

TABLE I: MELANOCYTIC AND NON-MELANOCYTIC CLASSES

Skin Lesion	Training images	Testing images	Total images
Melanocytic	6315	1503	7818
Non-Melanocytic	1696	501	2197
Total	8011	2004	10015

As there is imbalancing between the number of training images of each class, some of the images are then rotated to establish a class-balanced dataset. The number of images in this augmented dataset is given in Table II.

TABLE II: AUGMENTED MELANOCYTIC AND NON-MELANOCYTIC CLASSES

Skin Lesion	Training images	Testing images	Total images
Melanocytic	6315	1503	7818
Non-Melanocytic	6784	501	7285
Total	13099	2004	15103

Then, a second deep learning model is generated to model melanoma and melanocytic nevus. Table III gives the details of how many images there are for each class of this model. Again, due to imbalancing between the number of training images of each class, this dataset is then augmented by rotating some of the images. Table IV lists the number of images in this class-balanced dataset.

TABLE III: MELANOMA AND NEVUS CLASSES

Skin Lesion	Training images	Testing images	Total Images
Melanoma	800	313	1113
Melanocytic Nevus	5454	1251	6705
Total	6254	1564	7818

TABLE IV: AUGMENTED MELANOMA AND NEVUS CLASSES

Skin Lesion	Training images	Testing images	Total Images
Melanoma	6400	313	6713
Melanocytic Nevus	5454	1251	6705
Total	11854	1564	13418

Finally, a third deep learning model is generated to model non-melanocytic malignant and benign classes. Non-melanocytic malignant class includes basal cell carcinoma, actinic keratosis (Bowen's disease) whereas non-melanocytic benign includes benign keratosis, dermatofibroma and vascular lesion. The number of images for each class is listed in Table V.

TABLE V: NON-MELANOCYTIC MALIGNANT AND BENIGN CLASSES

Skin Lesion	Training images	Testing images	Total Images
Malignant	709	132	841
Benign	1048	308	1356
Total	1757	440	2197

During testing, the skin lesion is classified as melanocytic malignant using the first deep learning model generated and then melanoma is detected using the second deep learning model. Similarly, the skin lesion is classified as non-melanocytic malignant using the first deep learning model generated and then malignant types (basal cell carcinoma and squamous cell carcinoma) are detected using the third deep learning method.

AlexNet [7] architecture is selected for the deep learning method of the two-step diagnostic procedure with the following parameters: learning rate is 0.01, gamma is 0.1, momentum is 0.9 and weight decay is 0.0005.

IV. DATASETS

Publicly available 2018 ISIC challenge dataset is used for performance evaluation. This dataset contains 10015 dermoscopic skin images. These images are categorized as melanoma, melanocytic nevus, basal cell carcinoma, actinic keratosis (Bowen's disease), benign keratosis, dermatofibroma and vascular lesion. The total number of images in each category is 1113, 6705, 514, 327, 1099, 115 and 142, respectively (see Table VI).

V. PERFORMANCE EVALUATION

The two-step deep learning based classification method is evaluated using various classification tasks. Initially, the performance of the first deep learning model is evaluated for the classification of melanocytic and non-melanocytic skin lesions. Then, the performance of the

TABLE VI: SKIN LESIONS OF 2018 ISIC CHALLENGE DATASET

Skin Lesion	No. of images
Melanoma	1113
Melanocytic nevus	6705
Basal cell carcinoma	514
Actinic keratosis / Bowen's disease	327
Benign keratosis	1099
Dermatofibroma	115
Vascular lesion	142
Total	10015

second deep learning model is evaluated for the classification of melanoma and melanocytic nevus skin lesions. Finally, the performance of the third deep learning model is evaluated for the classification of non-melanocytic malignant and benign skin lesions.

A. Classification of Melanocytic and Non-Melanocytic Skin Lesions

In this investigation, the performance of the first deep learning model for melanocytic and non-melanocytic classification is performed. This is the first step of the two-step diagnostic procedure. The number of training and testing images used to evaluate the method performance of this investigation are listed in Table II.

The accuracy of the first step of the two-step deep learning based classification method is 78%, sensitivity is 83.9% and specificity is 63.5%, as given in Table VII.

TABLE VII: CLASSIFICATION PERFORMANCE OF MELANOCYTIC AND NON-MELANOCYTIC SKIN LESIONS

Method	Accuracy	Sensitivity	Specificity
AlexNet	78%	83.9%	63.5%

B. Classification of Melanoma and Melanocytic Nevus Skin Lesions

The performance of the second deep learning model for classification of melanoma and melanocytic nevus is carried out next. This is the second step of the two-step diagnostic procedure. Table IV lists the number of training and testing images used to evaluate the performance.

The accuracy of the second step of the two-step deep learning based method is 84%, sensitivity is 84.7% and specificity is 83.8%, as listed in Table VIII.

TABLE VIII: CLASSIFICATION PERFORMANCE OF MELANOMA AND MELANOCYTIC NEVUS SKIN LESIONS

Method	Accuracy	Sensitivity	Specificity
AlexNet	84%	84.7%	83.8%

C. Classification of Non-Melanocytic Malignant and Benign Skin Lesions

The performance of the third deep learning model for the classification of non-melanocytic malignant and benign skin lesions is observed next. The number of training and testing images used here are given in Table V. Note that this model is not a part of the two-step diagnostic procedure of the dermatologists and is done here as supplementary to the second step in order to be

able to detect non-melanocytic malignant skin lesions as well that are often faced in practice.

The accuracy of the supplementary AlexNet model is 58%, sensitivity is 60.6% and specificity is 57.8%, as tabulated in Table IX.

TABLE IX: CLASSIFICATION PERFORMANCE OF NON-MELANOCYTIC MALIGNANT AND BENIGN SKIN LESIONS

Method	Accuracy	Sensitivity	Specificity
AlexNet	58%	60.6%	57.8%

VI. DISCUSSIONS

A summary of the deep learning model performances is given in Table X. The highest accuracy and sensitivity values are of the second deep learning method. These values are lower for the first deep learning model. This shows that the classification performance of melanoma and melanocytic nevus classes is higher than the classification performance of melanocytic and non-melanocytic classes. It is also observed that the accuracy and sensitivity values of the third model are lower than the first and the second deep learning models.

From these results, we can see that melanoma and melanocytic nevus classes can be detected with the highest accuracy and that the detection of non-melanocytic malignant and benign class is the most difficult of the three models.

TABLE X: COMPARISON OF SKIN LESION CLASSIFICATION PERFORMANCES

Method	Accuracy	Sensitivity	Specificity
First Model	78%	83.9%	63.5%
Second Model	84%	84.7%	83.8%
Third Model	58%	60.6%	57.8%

VII. CONCLUSIONS

A two-step deep learning based classification method is implemented. This method is based on three deep learning models. A skin lesion can be classified as melanocytic using the first deep learning model. Then, the detection of melanoma skin lesion can be done using the second deep learning model. Similarly, a skin lesion can be classified as non-melanocytic using the first deep learning model and then malignant types (basal cell carcinoma, squamous cell carcinoma) can be detected using the third deep learning model.

The performance evaluation results show that the highest accuracy and sensitivity values are reported for the second deep learning model. They also show that the classification performance of melanoma and melanocytic nevus classes (second model) are higher that the classification performance of melanocytic and non-melanocytic classes (first model). Finally, they indicate that accuracy and sensitivity values of the third model are lower than the first and the second deep learning models.

It can be concluded from these results that since both melanocytic and non-melanocytic classes (first model) contain multiple skin lesion types, their detection is more difficult than melanoma and melanocytic nevus classes (second model). Furthermore, as the number of training images in the third model is not high enough, the deep

learning architecture is not able to fully differentiate between similar non-melanocytic malignant and benign skin lesion images. Hence, non-melanocytic malignant skin lesion detection (third model) is also concluded to be more difficult than melanocytic malignant skin lesion detection (second model).

REFERENCES

[1] O. Abuzaghleh, B. D. Barkana, and M. Faezipour, "Noninvasive real-time automated skin lesion analysis system for melanoma early detection and prevention," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 3, pp. 1–12, 2015.

[2] G. Argenziano, H. P. Soyer, S. Chimenti, R. Talamini, R. Corona, F. Sera, M. Binder, L. Cerroni, G. De Rosa, G. Ferrara, et al., "Dermoscopy of pigmented skin lesions: results of a consensus meeting via the internet," *Journal of the American Academy of Dermatology*, vol. 48, no. 5, pp. 679–693, 2003.

[3] H. Cabo, "A simple guide to dermatoscopic diagnosis of melanocytic and nonmelanocytic skin lesions," *Expert Review of Dermatology*, vol. 1, no. 3, pp. 333–339, 2006.

[4] Y. Yuan, M. Chao, and Y. Lo, "Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance," *IEEE Transactions on Medical Imaging*, vol. 36, no. 9, pp. 1876–1886, Sept. 2017.

[5] M. E. Celebi, H. A. Kingravi, B. Uddin, H. Iyatomi, Y. A. Aslandogan, W. V. Stoecker, and R. H. Moss, "A methodological approach to the classification of dermoscopy images," *Computerized Medical Imaging and Graphics*, vol. 31, no. 6, pp. 362 – 373, 2007.

[6] Y. Li and L. Shen, "Skin lesion analysis towards melanoma detection using deep learning network," *Sensors*, vol. 18, no. 2, pp. 556, 2018.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. 25th International Conference on Neural Information Processing Systems*, pp. 1097–1105, 2012.

[8] P. Tschandl, C. Rosendahl, and H. Kittler, "The HAM10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions," *arXiv preprint arXiv:1803.10417*, 2018.

[9] S. S. Han, M. S. Kim, W. Lim, G. H. Park, I. Park, and S. E. Chang, "Classification of the Clinical Images for Benign and Malignant Cutaneous Tumors Using a Deep Learning Algorithm.," *J. Invest. Dermatol.*, vol. 138, no. 7, pp. 1529–1538, 2018.

[10] K. Shimizu, H. Iyatomi, M. E. Celebi, K. A. Norton, and M. Tanaka, "Four-class classification of skin lesions with task decomposition strategy," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 1, pp. 274–283, 2015.

[11] H. Iyatomi, K. A. Norton, M. E. Celebi, G. Schaefer, M. Tanaka, and K. Ogawa, "Classification of melanocytic skin lesions from non-melanocytic lesions," in *Proc. 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5407–5410, 2010.

[12] A. Esteva *et al.*, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[13] H. Zhou, X. Li, G. Schaefer, M. E. Celebi, and P. Miller, "Mean shift based gradient vector flow for image segmentation," *Computer Vision and Image Understanding*, vol. 117, no. 9, pp. 1004–1016, 2013.

[14] A. R. Sadri, M. Zekri, S. Sadri, N. Gheissari, M. Mokhtari, and F. Kolahdouzan, "Segmentation of dermoscopy images using wavelet networks," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 4, pp. 1134–1141, 2013.

[15] F. Xie and A. C. Bovik, "Automatic segmentation of dermoscopy images using self-generating neural networks seeded by genetic algorithm," *Pattern Recognition*, vol. 46, no. 3, pp. 1012–1019, 2013.

[16] M. E. Celebi, Q. Wen, H. Iyatomi, K. Shimizu, H. Zhou, and G. Schaefer, "A state-of-the-art survey on lesion border detection in dermoscopy images," *Dermoscopy Image Analysis*, pp. 97–129, 2015.

[17] N. C. F. Codella, Q. Nguyen, S. Pankanti, D. Gutman, B. Helba, A. Halpern, and J. R. Smith, "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM Journal of Research and Development*, vol. 61, no. 4/5, pp. 5:1-5:15, 2017.

[18] N. C. F. 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," in *Proc. International Workshop on Machine Learning in Medical Imaging*, pp. 118-126, 2015

[19] L. Yu, H. Chen, Q. Dou, J. Qin, and P. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994–1004, 2017.

[20] F. Adjed, S. J. S. Gardezi, F. Ababsa, I. Faye, and S. C. Dass, "Fusion of structural and textural features for melanoma recognition," *IET Computer Vision*, vol. 12, no. 2, pp. 185–195, 2017.

[21] D. Gutman, N. Codella, E. Celebi, B. Helba, M. Marchetti, N. Mishra, and A. Halpern, "Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)," *arXiv preprint arXiv:1605.01397*, 2016

[22] N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. K. Mishra, H. Kittler, 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 *Proc. IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 168–172, 2018.