

CNN and Metadata for Classification of Benign and Malignant Melanomas

José-Sergio Ruiz-Castilla^(⋈), Juan-José Rangel-Cortes, Farid García-Lamont, and Adrián Trueba-Espinosa

Universidad Autónoma del Estado de México,
Centro Universitario UAEM Texcoco, Av. Jardín Zumpango s/n El Tejocote,
56259 Texcoco, Estado de México, Mexico
jsergioruizc@gmail.com, juan991029@gmail.com,
fglamont@gmail.com, atruebae@hotmail.com

Abstract. Skin cancer is detected in skin lesions. The most common skin cancer is melanoma. Skin cancer is increasing in several parts of the world. Due to the above, it is important to work on the classification of melanomas, in order to support the possible detection of malignant melanomas that cause skin cancer. We use Convolutional Neural Networks (CNN) for the classification of melanomas. We use images available from International Skin Imaging Collaboration (ISIC). We created a repository of 1000 images and did training with a sequential CNN to obtain two categories: benign and malignant melanomas. In the first instance we obtained results of 94.89% accuracy and 82.25% in validation. In the second instance we created another repository of 600 images for the method that we propose that consists in adding metadata within the same pixel matrix of the image in each RGB layer. The image was shown with a band of colors at the bottom. We made training with the CNN using images with metadata and achieved the results: 98.39% of accuracy and 79% of validation. Therefore, we conclude that adding the metadata repeatedly to the pixel matrix of the image improves the results of the classification.

Keywords: Melanomas · Convolutional neural networks · Metadata · Classification · Prediction

1 Introduction

The increase in the rate of cancer is a health problem in the world. The melanoma is among the most common types of cancer in the United States of America. In Mexico, skin cancer is becoming more frequent. This is due to different factors such as prolonged exposure to ultra violet rays or exposure to water contaminated by arsenic and other chemicals. In America, cancer is the second cause of death. It is estimated that 2.8 million people are diagnosed with cancer each year and 1.3 million people die from this disease annually. Approximately, 52% of new cancer cases occur in people 65 years old or younger. In case of not taking action, an increase of more than four million new cases and 1.9 million deaths due to cancer is expected for the year 2025 [1]. Melanomalike skin cancer can almost always be cured if discovered and treated promptly [2].

The National Health System in Mexico has three levels for the attention to the public. In the first level, basic health services are provided. It is the main scenario for preventive health and where 80% of the ailments are treated and solved. In the second level, patients referred from the first level are attended to perform diagnostic procedures, as well as therapeutic and rehabilitation processes. At the third level, diseases of low prevalence, high risk and more complex diseases are treated. At this level, patients sent from the first and second level are attended [3].

Therefore, a proposal is made to classify melanomas to help in the pre-diagnosis of melanoma skin cancer. The proposed solution is focused on the first level of the National Health System in Mexico. A convolutional neuronal network as well as metadata will be used from the image base offered by International Skin Imaging collaboration (ISIC).

1.1 Skin Cancer

Skin cancer represents the most common cancer in humans, and we are currently facing increasing rates of newly diagnosed cutaneous neoplasm each year. Knowledge has been gained concerning the biology of skin cancer and risk factors; diagnosis has been improved and novel therapeutic modalities have been developed. As UV-exposure is known as the major risk factor for skin cancer, prevention can be achieved through UV-protection and regular sunscreen use. In many countries primary care physicians are performing regular skin checks and initiate early treatment, working closely with a dermatologist. Therefore, knowledge regarding diagnosis, appropriate treatments including recommended excision margins, the performance of sentinel lymph nodes biopsies, and the ability to use novel topical treatment modalities are of importance for all physicians involved [4].

1.2 Melanomas

Melanomas are common cancers arising from the pigment cells of the skin. The principal environmental determinant of cutaneous melanoma is sunlight, with incidence rates varying more than tenfold between ethnically similar populations residing in environments with different levels of ambient sunlight [5]. Diagnosis should be based on a full-thickness excisional biopsy, with a small side margin. Wide local excision of a cuff of normal tissue after an excision biopsy reduces local recurrence rates [6]. Because melanoma in advanced stages is still incurable, early detection is indispensable to reduce mortality. With the introduction of dermoscopy into the clinical practice, the diagnostic accuracy of pigmented skin lesions can be improved. The use of dermoscopy allows the identification of many different structures and colors, not seen by the naked eye. Colors play an important role in dermoscopy. Common colors are light brown, dark brown, black, blue, blue-gray, red, yellow, and white [7].

1.3 Convolutional Neural Netware

CNNs, also known as convnets, are a type of deep learning model that is used almost universally in machine vision applications. Convnets can be applied to image

classification problems, particularly those involving small sets of training data. The convolution layers of a convent learn from local patterns (see Fig. 1): in the case of the images, the patterns are found in small 2D windows of the inputs.



Fig. 1. The images can be divided into local patterns, such as edges, textures, etc. [8]

The above feature is key and gives the convnets two interesting properties [8]:

- 1. The patterns they learn are invariant in translation. After learning a certain pattern in the lower right corner of an image, a *convnet* can recognize it anywhere: for example, in the upper left corner. A densely connected network would have to learn the pattern again if it appeared in a new location. This makes the *convnets* data efficient when processing images: they need fewer examples of training to learn representations that have power of generalization.
- 2. They can learn spatial hierarchies of patterns (see Fig. 2). A first convolution layer will learn small local patterns such as the edge; a second convolution layer will learn larger patterns made from the characteristics of the first layers, and so on. This allows social networks to efficiently learn increasingly complex and abstract visual concepts (because the visual world is fundamentally hierarchical spatial) [8].

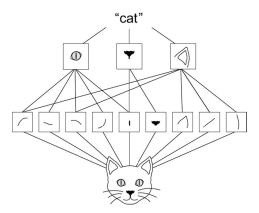


Fig. 2. The visual world forms a spatial hierarchy of visual modules: hyperlocal edges are combined in local objects such as eyes or ears, which are combined in high-level concepts such as "cat" [8].

2 Related Works

For the classification of images, it is necessary to apply techniques of image processing and extraction of characteristics such as color and borders, as well as, the segmentation of the image [9].

When looking to classify very similar images such as leaves of plants it is necessary to apply the extraction of characteristics such as texture, chromaticity and geometric shape and apply some combination of classification techniques [10, 11]. It is possible to classify fruits by extracting characteristics of fruit shape, texture and color (HSV). This classification can be very useful for supermarkets to handle fruit sales [12].

Mahbod among others [13] took available images of the ISIC in 2017. They applied three pre-processing steps in the images. First, they normalized the images by subtracting the average RGB value from the data set, since the pre-trained networks were optimized for those images. Next, the images were resized to the appropriate size using bi-cubic interpolation to feed the networks $(227 \times 227 \text{ and } 224 \times 224)$. Finally, the images increased to have a larger training set. Since the classifiers were trained for a problem of multiple classifications for the classes of: *melanoma*, *seborrheic keratosis* and classes of benign *nevi*. The performance of the classifier, in terms of accuracy (ACC) and area under the curve (AUC) of receiver operating characteristics were measured for 150 test images. The Achieving results of 79.9% to 89.2% of the value Accuracy.

Yang among others [14], *DCNN* (*Deep CNN*) multitasking was used. For the analysis of skin lesions. The *DCNN* multitasking model you implemented is based on the *GoogleNet* design architectures. The dermatoscopic training and evaluation data sets were taken from *ISIC* 2017. 2000 samples in the training data set and 150 samples in the validation data set. The results of the classification are evaluated by the area under the receiver operating characteristics curve (*AUC*). The *AUC* is a measure of how well a parameter can distinguish between two diagnostic groups (sick/normal). They compared the performance between *DCNN* multitasking and *GoogleNet*. The results obtained were 92.6% of the proposed model and 90.3% with *GoogleNet*.

Shoieb [15] among others took a pre-trained *CNN* in its last phase to classify the infected skin lesion according to the characteristics of *CNN* to train a *SVM* multiclass classifier. There are several pre-trained networks that have gained popularity. Most of these have been trained in the *ImageNet* data set, which has 1000 object categories and 1.2 million training images. Four classes of skin diseases have been selected for system validation. These classes are melanoma, basal cell carcinoma, eczema and impetigo. The data set comprises 134 images, 72 of them represent melanoma, 64 for basal cell carcinoma, 74 eczema and 31 for impetigo skin disease. The results of the classification of melanoma and non-melanoma lesions from the data obtained from the Dermatology Information System were 94% accuracy, 94% specificity and 94% sensitivity and for non-melanomas it was 94% accuracy, 94% of specificity and 94% sensitivity.

Codella among others [16] experimented with the double-cross-validation method and performed 20 times (40 total experiments) with 334 images of melanoma and 144 images of atypical nevi, as well as 2146 clearly benign lesions (2624 total) taken from *ISIC*. Two color dictionaries (*RGB*) and grayscale color spaces are built. The images

are resized to dimensions of 128×128 pixels before the extraction of 8×8 patches, to learn dictionaries of 1024 elements. To train classifiers we used a nonlinear *SVM* that uses a core intersecting histogram and sigmoid. *SVM* scores were assigned to probabilities by logistic regression in the training data. They used a 50% probability as a binary classification threshold. The fusion is carried out by means of an unweighted *SVM* average score (late fusion). The results obtained among others with *Caffe CNN* were 91.9% accuracy, 90.3% specificity and 92.10% sensitivity.

Dorj among others [17] did a classification of skin cancers using a deep convolutional neuronal network. The algorithm to obtain the general methodology that they used is the following:

Acquisition of image data (collect image data, cut the images, prepare training and test sets of images).

Classification into groups that use convolutional neuronal network functions (extract training functions using a convolutional neuronal network, train and test an ECOC-SVM classifier using convolutional neural network functions).

The data set in this study originated from related Internet sites. In this study, they used *RGB* images of skin cancer with 500–1000 pixels, .jpg and .tiff. With a total of 3753 images, which are collected from the Internet (google.com and naver.com, baidu.com and bing.com). The results of the four types of melanoma classified for the accuracy factor were: 92.3% for Actinic keratosis, 91.8% for Basal cell carcinoma, 95. 1% for Squamous cell carcinoma and 92.2% for Melanoma.

Liao [18] did a classification of skin diseases. Your data set of skin diseases from two different sources: *Dermnet* and *OLE*. *Dermnet* is one of the largest sources of photographic dermatology that is publicly available. It has more than 23,000 images of skin diseases in a wide variety of skin conditions. *Dermnet* biologically organizes skin diseases in two-tier taxonomy. In their approach, they transferred the learning of *ImageNet's* pre-trained modeling with *Caffe*, a deep learning framework that supports the deep and expressive training of *CNN*. They chose *VGG16*, *VGG19* and *GoogleNet* as their pre-trained models. The results obtained were 91% with *CGG16*, 90.9% with *VGG19* and 90.7% with *GoogleNet*.

Georgakopoulos among others [19] classified dermatoscopic images with two classes as "malignant" or "non-malignant". They tried to use an improved entry to *CNN* using the Gaussian filters, Gaussian Laplacian (*LoG*), Hessian matrix and Gabor. They investigated the value of increasing *CNN* entries with filters used from computer vision. The response of each available image with these filters is calculated and used as an additional input to *CNN*. They also used Transfer learning (*TL*) using two pre-trained network architectures.

Step 1: the set of malignant skin images is enlarged by applying image rotations in multiples of 90°, as well as mirror transformations in each rotated image.

Step 2: For file1, all images are resized to 256×256 . A pre-processing procedure extracts 224×224 image patches at random locations in the resized image, which are used as an input to CNN. Similarly, for CNN file2, the images are resized to 352×352 . The pre-processing stage extracts 320×320 image patches in random locations of the resized image, which are used as an input to CNN.

The expanded data set is used only for training, and the produced images did not need to be stored. When performing the training with the two sets of images. The results with CNN were of, 92.3% for the first set and 92.3% for the second set. While, for the second set with LoG were 93.9% and 88.6% for the first and second sets respectively [19].

Haenssle among others [20] used and trained a modified version of Google's CNN Inception v4 architecture. They created a test set of 300 images that included 20% of melanomas from all body sites and all frequent histotypes, and 80% of benign melanocytic nevi from different subtypes and body sites, including so-called "simulators" Melanoma. The images of the test set-300 were retrieved from the high-quality validated image library of the Department of Dermatology of the University of Heidelberg, Germany, and various camera/dermoscopy combinations were used for the acquisition of images the overlap between data sets for training/validation and testing.

In this research they focused on comparing the accuracy of the algorithm against experts in dermatology. Finding that, experts with more than five years of experience were better than the algorithm, whereas, experts with less than two years of experience were worse than the algorithm.

Moura among others [21], propose the classification of skin lesions using a descriptor formed by the combination of the characteristics of the ABCD rule (Asymmetry, Border, Color and Diameter) and pre-trained CNN. The characteristics were selected according to their gain ratios and used as inputs for the MultiLayer Perceptron (MLP) classifier.

Characteristics are properties that can be measured from an image, such as shape, color and texture. These attributes, grouped in a feature vector, are called an image descriptor. The descriptors used to extract characteristics of the lesion images of the skin and form the hybrid descriptor. The proposed ones are detailed below:

ABCD Rule: This rule is divided into four steps to describe the image and has five attributes: asymmetry (1), edge irregularity (2), color (1) and diameter (1). In the most successful tests, the proposed method achieved an accuracy rate of 94.9% using MPL.

Oliveira among others [22] comment that the extraction of characteristics to describe skin lesions is an area of research is challenging due to the difficulty to select significant characteristics. The computational diagnostic system developed for the skin lesion was applied to a set of 1104 dermoscopy images by a cross-validation procedure. The best results were obtained through an optimal trajectory forest classifier (OPF) with very promising results. The proposed system reached an accuracy of 92.3%, a sensitivity of 87.5% and a specificity of 97.1% when the complete set of characteristics was used. In addition, it achieved an accuracy of 91.6%, a sensitivity of 87% and a specificity of 96.2%, when 50 functions were chosen using a feature selection algorithm based on correlation. A total of 1104 images were selected from the original data set. Of these, 916 images were benign lesions and 188 images were malignant lesions. The images in the data set were resized at an average resolution of 400 × 299 pixels to simplify processing. The results of six different classifiers were: 75.8% for KNN, 68.2% for Bayes net, 86.9% for CA.5, 74.5% for MLP, 91.7% for SVM and 92.3 for OPF.

3 Proposed Method

3.1 Images and Images Metadata

For this work, we got images from ISIC a database of melanomas images and metadata able in "https://isic-archive.com/". Melanoma Project is an academia and industry partnership designed to facilitate the application of digital skin imaging to help reduce melanoma mortality. This archive serves as a public resource of images for teaching and for the development and testing of automated diagnostic systems. This is a site that store 23906 images from collaborators of area. We download the images and metadata from ISIC. We created two images packages. The characteristic of packages are shown on Table 1.

Table 1. Image packages from *ISIC*. 800 images for training and 400 images for validation. Also, 400 images with metdata for training and 200 images with metadata for validation

Packages	Set	Melanoma type	Images
Images without metadata	Training	Benign melanomas	600
		Malignant melanomas	600
	Validation	Benign melanomas	200
		Malignant melanomas	200
Images with metadata	Training	Benign melanomas	200
		Malignant melanomas	200
	Validation	Benign melanomas	100
		Malignant melanomas	100

3.2 Add Metadata to the Image

We duplicate the metadata repeatedly at the end of the matrix. The images have three channels (RGB), therefore, we duplicate the metadata in the three layers. Once the metadata is added repeatedly the image shows a band of color indicating that the metadata were added. This process was necessary to do it in the images of: training, validation and testing. See Fig. 3.

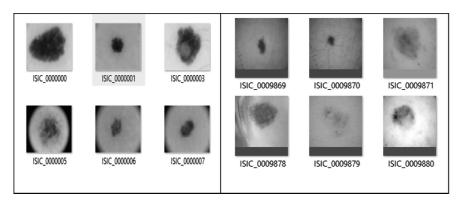


Fig. 3. Matrix of images (a) without metadata (b) with metadata

We choose three metadata: sex, age_approx and clin_size_long_diam_mm. We insert the gender with the values of 100 for feminine and 200 for masculine. In the same way, the size of the melanoma was inserted. Finally, the patient's age was inserted. The algorithm traverses the column from row 300 to 350 by inserting the metadata repeatedly. This process was applied to the three channels for each image of all image set.

4 Results

4.1 Instance 1 y 2. Results of Training and Validation

The first experiment was executed with CNN. With the next parameters: 8 Epochs, 350 of height, 350 of weight, 128 of batch_size, 2 classes, 1000 step_train, 300 step_validation, 32 filter_Conv1, 64 filter_Conv2, (3, 3) of size_Filter, (2, 2) of size_Filter2 and (2,2) of size_Pool. The values obtained are sown in the Table 2.

	Instance 1				Instance 2			
Epochs	ACC	VAL_ACC	LOSS	VAL_LOSS	ACC	VAL_ACC	LOSS	VAL_LOSS
1	0,7160	0,8100	0,6311	0,4821	0,8487	0,785	0,8487	0,6260
2	0,7701	0,8725	0,4622	0,3744	0,9393	0,795	0,1529	0,8093
3	0,8108	0,8650	0,4016	0,3667	0,9700	0,765	0,0856	0,8959
4	0,8517	0,8325	0,3301	0,4619	0,9826	0,795	0,0524	1,3705
5	0,8863	0,8525	0,2695	0,5041	0,9858	0,785	0,0421	1,2821
6	0,9152	0,8475	0,2097	0,6170	0,9830	0,820	0,0537	1,2562
7	0,9361	0,7825	0,1678	0,7229	0,9914	0,790	0,0248	1,3772
8	0,9489	0,8225	0,1313	0,7594	0,9890	0,790	0,0330	1,2230

Table 2. Results of accuracy and loss of Instance 1 and 2.

We executed the algorithms for instance 1 and instance 2, obtaining 94.89% and 89.90% in training, while 82.25% and 79.00% were obtained in the validation. As we can see the accuracy improved in the training, however it decreased in the validation.

In order to know the behavior of the accuracy and the loss in both scenarios, we create a graph for each case that we can see in Fig. 4.

In Instance 2 with images with metadata, better results were obtained in the accuracy, although not in the validation. In the Instance 2 we can see that, the validation accuracy was not ascending, but it remained stable.

The results were obtained from the two Instances. From the first instance that contains images without metadata, results of 94.89% accuracy and 82.25% validation were obtained. While, with the set of images with metadata, 98.90% accuracy and 79.00% validation was obtained. Finally, we did tests with images not included in the training or validation and the results were 61% in Instance 1 and 85% in instance 2. We did not find results of a test with new images in the related works. See the Table 3.

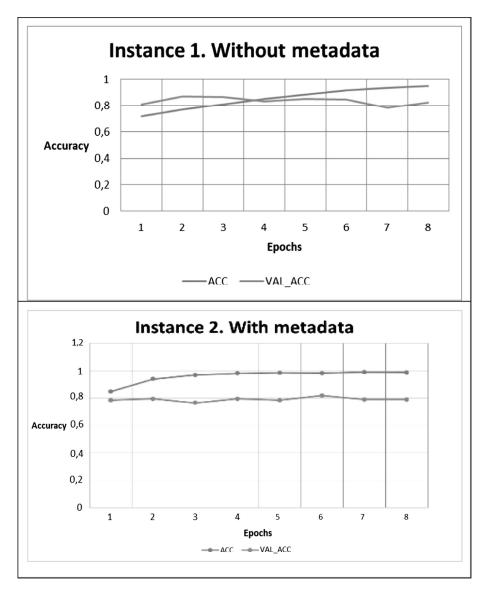


Fig. 4. Instance 1 y 2. Precision and loss behavior.

Our results and others research works | Accuracy Test **CNN** 94.89% 61% CNN & metadata 98.90% 85% Mahbod 79.90% Yang 92.60% 94.00% Shoieb Codella 91.90% Dorj 92.30% 91.00% Liao Georgakopoulos 93.90% Moura 94.90% Oliveira 92.30%

Table 3. Results from Instance 1 and Instance 2 ever test.

5 Conclusions

The Artificial neural networks can help us for the creation of applications for the classification of benign and malignant melanomas. The technique presented is authentic, because something similar was not found in the literature related. Including the metadata in a duplicate way within the same pixel matrix enriches the extraction of characteristics improving the results. We conclude that the proposal to add metadata directly to the matrix of the image improves the classification. We will continue doing tests with more metadata in order to find the most significant metadata. On the other hand, it could be a technique that could be used in the classification of other types of images.

References

- Pan American Health Organization, Pan American Health Organization (2017). http://www. paho.org. Accessed 14 Nov 2017
- The Skin Cancer Foundation, Skin Cancer Foundation (2013). http://www.cancerdepiel.org/ cancer-de-piel/melanoma. Accessed 14 Nov 2017
- P. A. Q. L. A. Burr Claudia, Association of Health and Social Welfare of Women and their Families, Mayo (2011). http://asbis.org.mx/PDF/guiapacienteparticipativo.pdf. Accessed 01 Nov 2017
- Stockfleth, E., Rosen, T., Shumack, S.: Managing Skin Cancer. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-540-79347-2
- Whiteman, D., Green, A.: Epidemiology of malignant melanoma. In: Dummer, R., Pittelkow, M., Iwatsuki, K., Green, A., Elwan, N. (eds.) Skin Cancer- A World-Wide Perspective, pp. 13–26. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-05072-5_2
- Guggenheim, M., et al.: Melanoma. In: Dummer, R., Pittelkow, M., Iwatsuki, K., Green, A., Elwan, N. (eds.) Skin Cancer - A World-Wide Perspective, pp. 307–342. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-05072-5_23

- Kolm, I., Dummer, R., Braun, R.P.: Dermoscopy. In: Dummer, R., Pittelkow, M., Iwatsuki, K., Green, A., Elwan, N. (eds.) Skin Cancer A World-Wide Perspective, pp. 373–378.
 Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-05072-5_26
- 8. Chollet, F.: Deep Learning with Python. Manning Publications Co., Shelter Island (2018)
- García-Lamont, F, Cervantes, J, López, A, Rodríguez, L: Segmentation of images by color features: a survey. Neurocomputing 292, 1–27 (2018). https://doi.org/10.1016/j.neucom. 2018.01.091, http://www.sciencedirect.com/science/article/pii/S0925231218302364. ISSN 0925-2312
- Cervantes, J., Taltempa, J., Lamont, F.G., Castilla, J.S.R., Rendon, A.Y., Jalili, L.D.: Análisis Comparativo de las técnicas utilizadas en un Sistema de Reconocimiento de Hojas de Planta. Revista Iberoamericana de Automática e Informática industrial 14(1), 104–114 (2017). Web. 15 May 2019
- Cervantes, J., Garcia Lamont, F., Rodriguez Mazahua, L., Zarco Hidalgo, A., Ruiz Castilla, J.S.: Complex identification of plants from leaves. In: Huang, D.-S., Gromiha, M.M., Han, K., Hussain, A. (eds.) ICIC 2018. LNCS (LNAI), vol. 10956, pp. 376–387. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-95957-3_41
- García-Lamont, F., Cervantes, J., López-Chau, A., Alvarado, M.: Fruit classification by extracting color chromaticity, shape and texture features: towards an application for supermarkets. IEEE Latin Am. Trans. 14, 3434–3443 (2016). https://doi.org/10.1109/TLA. 2016.7587652
- Mahbod, A., Ecker, R., Ellinger, I.: Skin Lesion Classification Using Hybrid Deep Neural Networks (2017). CoRR, abs/1702.08434
- 14. Yang, X., et al.: Cornell University, 04 March 2017 https://arxiv.org/abs/1703.01025
- Shoieb, D.A., Youssef, S., Aly, W.: Computer-aided model for skin diagnosis using deep learning. J. Image Graph. 116–121 (2016). https://doi.org/10.18178/joig.4.2.122-129, https://www.researchgate.net/publication/312188377_Computer-Aided_Model_for_Skin_ Diagnosis_Using_Deep_Learning
- Codella, N., Cai, J., Abedini, M., Garnavi, R., Halpern, A., Smith, J.R.: Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images. In: Zhou, L., Wang, L., Wang, Q., Shi, Y. (eds.) MLMI 2015. LNCS, vol. 9352, pp. 118–126. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24888-2_15
- Dorj, U.O., Lee, K.K., Choi, J.Y., et al.: Multimed. Tools Appl. 77, 9909 (2018). https://doi. org/10.1007/s11042-018-5714-1, https://link.springer.com/article/10.1007%2Fs11042-018-5714-1
- 18. Liao, H.: A Deep Learning Approach to Universal Skin Disease Classification (2015). https://www.semanticscholar.org/paper/A-Deep-Learning-Approach-to-Universal-Skin-Disease-Liao/af34fc0aebff011b56ede8f46ca0787cfb1324ac
- Georgakopoulos, S.V., Kottari, K., Delibasis, K., et al.: Neural Comput. Appl. (2018). https://doi.org/10.1007/s00521-018-3711-y, https://link.springer.com/article/10.1007/s00521-018-3711-y
- Haenssle, H.A., et al.: Reader study level-I and level-II groups, 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). https://doi.org/10.1093/annonc/mdy166, https://academic.oup.com/annonc/article/ 29/8/1836/5004443
- Moura, N., Veras, R., Aires, K., et al.: Multimed. Tools Appl. 78, 6869 (2019). https://doi. org/10.1007/s11042-018-6404-8, https://link.springer.com/article/10.1007%2Fs11042-018-6404-8
- Oliveira, R.B., Pereira, A.S. Tavares, J.M.R.S.: Neural Comput. Appl. (2018). https://doi.org/ 10.1007/s00521-018-3439-8. https://link.springer.com/article/10.1007/s00521-018-3439-8