Deep Learning Techniques Applied to Skin Lesion Classification: A Review

Giuliana M. Silva
Federal University of Technology Paraná
Curitiba, Brazil
giusil@alunos.utfpr.edu.br

André E. Lazzaretti
Federal University of Technology Paraná
Curitiba, Brazil
lazzaretti@utfpr.edu.br

Fernando C. Monteiro
Research Centre in Digitalization and Intelligent Robotics
Instituto Politécnico de Bragança
Bragança, Portugal
monteiro@ipb.pt

***Abstract*—Skin cancer is one of the most common cancers in the world. The most dangerous type of skin cancer is melanoma, which can be lethal if not treated early. However, diagnosing skin lesions can be a difficult task. Therefore, deep learning techniques applied to the diagnosis of skin lesions have been explored by researchers, given their effectiveness in extracting features and classifying input data. In this work, we present a review of latest approaches that apply deep learning techniques to skin lesion classification task. In addition, some datasets used for training and validating the models are introduced, informing their characteristics and specificities, as well as popular pre-processing steps and skin lesion segmentation approaches. Finally, we comment the effectiveness of the proposed models.**

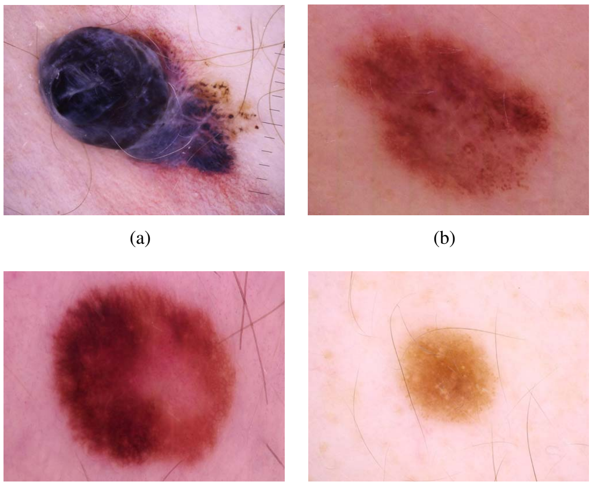
***Index Terms*—Deep learning, Skin lesion classification, Skin lesion segmentation, Skin diseases, Melanoma**

I. Introduction

Skin cancer is among the most common types of cancer in the world. It is mainly divided between malignant melanoma (MM) and non-melanoma (basal and squamous cell carcinoma). Melanoma is the most dangerous type of skin cancer and, when not treated early, it can be lethal [1]. Estimations from the Global Cancer Observatory show that melanoma was the 19th most incident cancer worldwide in 2020 having 324,635 new estimated cases [2] and, it was one of the most incident cancer in Oceania, North America, and Europe [3].

Usually, the diagnosis of pigmented skin lesions has a first assessment based on the ABCD criteria. In this task, the dermatologist detects four lesion features: (A) lesion asymmetry, (B) border irregularity, (C) colour variation and (D) lesion diameter greater than 6 mm [4]. However, the accuracy of diagnoses by observation relies on the dermatologist's experience. More experienced dermatologists can reach up to 80% of diagnostic accuracy, and, less experienced dermatologists, obtain an accuracy up to 62% [5]. Figure 1 shown three samples of melanoma and one common nevi sample.

To confirm the diagnosis, a biopsy can be performed, where a skin lesion sample is removed, analysed and classified [6]. Another technique is epiluminescence microscopy, also known as dermoscopy. It is a non-invasive technique that allows the visualization of the subcutaneous structures of the skin. Dermatologists trained with this technique increase their diagnostic skills, but those who are not trained may wrongly classify the samples [7].



(a) (b) (c) (d)

Fig. 1. Examples of melanoma lesions with (a) asymmetry, (b) border irregularity and (c) colour variation and a common nevi in (d) extracted from ISIC dataset.

Nowadays, new technological diagnosis methods namely computer-aided diagnosis (CAD) applied to dermatology has been researched, developed and commercialized [8]. Deep learning (DL) techniques are a promissory field of study to automatically detect and classify skin lesions. A deep neural network (DNN) is able to identify patterns, extract features and, based on these, performs data classification. Specifically, convolutional neural networks (CNN) are applied to computer vision and medical image analysis to recognize and classify different elements.

Two major tasks have been researched in the DL area related to skin lesion diagnosis: skin lesion segmentation and skin lesion classification. The skin lesion segmentation consists in automatically locate the region of the lesion. The skin lesion classification diagnosis the lesion based on the extracted features. This task can be preceded by a pre-processing step, a segmentation step, or both.

Another review papers addressing this subject have been published in recent years. Some of them have an extensive

2022 International Conference on Machine Learning, Control, and Robotics (MLCR) | 978-1-6654-5459-9/22/$31.00 ©2022 IEEE | DOI: 10.1109/MLCR57210.2022.00028

978-1-6654-5459-9/22/$31.00 ©2022 IEEE
DOI 10.1109/MLCR57210.2022.00028

explanation about the CNN architecture [9], [10] and others provide extra information about models, not CNN-based [8], [11], [12]. The proposed study aims to review the latests works of DL techniques applied to skin lesion classification, providing straightforward information about the diversity of CNN models that have been used for this purpose, and details about publicly available skin lesions datasets.

This paper is divided as follows: First, in section II, we present a set of skin lesion datasets that can be used to train the DNNs. In section III, we describe the image pre-processing methods. In sections IV, we provide an overview of skin lesion segmentation techniques. In section V, we detail some of the latest approaches used to perform the classification task. In section VI, we discuss and analyse the results of the works. In section VII, we provide a conclusion of the paper.

II. Skin lesion datasets

Training DNNs with large labelled datasets is an essential step to obtaining an efficient algorithm, and, nowadays, there are several skin lesions datasets publicly available for this purpose. Some databases are sectioned in training, validation and testing. This division can help to have a more reliable comparison between studies when using the same dataset. In this section, we present some of the most popular datasets for skin lesion diagnosis and their singularities.

The International Skin Imaging Collaboration (ISIC) archive is an open-source project that maintains some of the most popular datasets for skin lesion diagnosis and provides over 20000 images acquired with several devices [13]. Since 2016, ISIC sponsors challenges annually to promote the use of artificial intelligence algorithms to improve lesions diagnoses accuracy [14].

The ISIC 2016 dataset provides 900 dermatoscopic images for training and 379 images for testing with ground truth segmentation masks and gold standard diagnoses classification for benign and malignant lesions [15]. ISIC 2017 contains 2000, 150 and 600 dermatoscopic images for training, validation and testing, respectively. All the sets contain ground truth segmentation masks and gold standard diagnoses classified for melanoma, seborrhoeic keratosis and benign nevi [16]. The ISIC 2018 consists of 10015 images for training and 7 classification labels, 193 for validation and 1512 images for test. Only the training and validation sets contain the ground truth and segmentation masks [13], [17].

In ISIC 2019, we have 25331 images for training with metadata of gender, age and general anatomic site and gold standard lesion diagnoses classification. The training dataset is composed of HAM1000 and BCN20000 datasets. Additionally, it provides 8238 test images but without ground truth [16]–[18]. The ISIC 2020 provides 33126 images with metadata information and gold standard lesion diagnoses classification. It additionally provides 8238 test images without ground truth. Both ISIC 2019 and 2020 do not contain segmentation masks [19]. Although the large amount of data presented in these two datasets, both of them contain severe unbalance between classes that should be considered when using them as it can bias the trained model.

PH2 is a publicly available dataset obtained at the Dermatology Service of Hospital Pedro Hispano located in Matosinhos, Portugal, in collaboration with the Universidade do Porto. It has 200 dermatoscopic images classified in 80 common nevi, 80 atypical nevi, and 40 melanomas with a resolution of 768x560 pixels. For each image, a binary mask and medical annotation about diagnoses criteria are available [20].

Another publicly available dataset is the one obtained by the Department of Dermatology of the University Medical Center Groningen to develop their MEDNODE system. It consists in 170 non-dermoscopic images divided in 70 melanoma and 100 nevi [21].

DermQuest dataset contains 24082 images with 134 different diagnoses classification [22]. Although DermQuest was used in various researches, nowadays its website was deactivated in 2019.

PAD-UFES-20 is a dataset containing 2298 images with six diagnoses classification and metadata regarding the patient's age, lesion location and other nineteen clinical data. The images were collected using smartphones in eleven cities of Espírito Santo, state of Brazil [23].

The mentioned datasets are small when compared to the visual recognition ones for general purposes. To overcome this problem, techniques such as data augmentation and transfer learning can be applied. Data augmentation consists in rotating, translating, and rescaling images to obtain more data for training. Transfer learning consists in applying a pre-trained neural network to another set of training for the desired purposes.

III. Image pre-processing

Image pre-processing can be performed to obtain clearer images by removing noise and artefacts, correcting lighting issues, enhancing contrast, and also adapting the input data by resizing and cropping the images. Several authors use these methods to improve the accuracy of the segmentation algorithm.

Khouloud et al. [24] used a Gaussian filter to minimize the presence of artefacts from the skin lesion images. Akram et al. [25] utilized a software that performs a bilinear interpolation to localize the hair, remove it and apply a median filter where the hair was removed.

To enhance the image features, Farhat Afza et al. [26], proposed both local and global contrast enhancement. Khan et al. [27], applied an artificial bee colony algorithm to perform a contrast enhancement and improve the image contents. The datasets typically contain images taken from several different devices, which can cause lightning and colour variation between the images, Goyal et al. [28] applied colour and lightning normalization to decrease this difference.

Although some authors dismiss the image pre-processing step, others can give special attention to this. Sakar et al. [29] performed several tests using different pre-processing algorithms and compared them using accuracy (ACC) and

Area Under ROC Curve (AUC) metrics. This method led them to utilize a 4-step combination of image pre-processing algorithms consisting of (1) noising removal applying non-local means filter, (2) image enhancement applying CLAHE-DWT algorithm, (3) merging RGB colour space with the saturation channel, b* channel from CIELAB colour space and the inverted gray-scale channel, and (4) intensity normalization and image resizing.

The Shades of Grays method was used in three of the best ranked methods in ISIC 2019 Challenge [13], [30], [31]. This method is a colour constancy algorithm that aims to provide a better perception of the colours in the image by attenuating the impact of the light in the image.

IV. Skin lesion segmentation

Image segmentation in digital image processing that consists of segmenting the image into homogeneous regions. When applied to skin lesion images, these segmentation techniques can localize the lesion and define its boundaries. Once defined the Region of Interest (ROI), it can help to better extract features and classify the lesion, since the shape and border characteristics of the lesion are essential for the diagnosis [32].

However, skin lesion segmentation is a difficult task. The presence of artefacts, such as hair, skin texture, marks, reflections and shadows can interfere with accuracy [33]. To minimize this, image pre-processing is applied to reduce or erase the artefacts and enhance lesion features.

Several segmentation approaches can be applied in medical images as thresholding, region-based methods, edge detection, NNs and DNNs. The fully convolutional network (FCN) is one of the first semantic segmentation techniques proposed in the literature. There are several varieties of this architecture being applied for the skin lesion segmentation task.

Recent strategies fused FCNs with other models. One example is the FC-DPN (fully convolutional dual path network) proposed by Shan et al. [34], which consists of replacing the dense blocks in fully convolutional DenseNets (FC-DenseNets) with two types of dual path networks (DPN) blocks, projection and processing blocks. This approach aims to gather more features to improve the segmentation. Other related approach is the one proposed by Yu et al. [35] which consists of a fully convolutional residual network (FCRN) that can make pixel-wise predictions and is composed of residual blocks to avoid degradation problems.

Another conventional architecture built for medical image segmentation proposed by Ronneberger et al. [36] is the U-net. The U-net architecture was based on the FCN and can be trained with few images and result in more precise segmentation images. This model is composed of an encoder that operates as a regular CNN feature extractor, obtaining several feature maps in each layer, and a decoder that projects the features extracted by the encoder in the pixel space to obtain the segmented image.

Hasan et al. [33], proposed an inspired U-net architecture named DSNet. The DSNet encoder is a pre-trained CNN composed of feature layers, dense and transition blocks. The DSNet decoder is composed of a depth-wise separable convolution. Khouloud et al. [24], proposed a dual encoder-decoder architecture named W-net. The W-net is composed of a ResNet Encoder-Decoder with an attention mechanism followed by a ConvNet Encoder-Decoder. The first encoder-decoder is responsible for extracting the global shape of the lesion and the second sharps the boundaries. A Feature Pyramid is used in the last layer to reduce memory usage. Nguyen et al. [37] modified the U-net architecture by replacing the original encoder with an EfficientNet and the decoder is built with residual blocks from ResNet. This strategy aimed to use ResNet and EfficientNet architectures to avoid overfitting and maintain the efficient reception field size.

Other skin lesion segmentation methods based on different DL architectures can be found in the literature. One of them is proposed by Goyal et al. [28] where it was used a DeepLabv3+ and a Mask-RCNN to perform the segmentation task. The output of the two CNNs was fused using ensemble methods. The Ensemble-A method fuse the two outputs, the Ensemble-L method chooses the output with the larger area, and the Ensemble-S method selects the output with the smaller area. Post-processing approaches were used to fulfill the region and remove artefacts. Xie et al. [38] proposed a CNN composed of high-resolution feature blocks (HRFB) with an attention mechanism. This block is composed of 3 branches: main, spatial attention and channel-wise attention. The output of the HRFB block provides a feature map that preserves spatial details. At the end of the CNN, it is performed an up-sampling using bilinear interpolation.

Some works use non-DL approaches are also found in literature. Akram et al. [25] used contrast stretching to differentiate the background of the lesion with the foreground and mean segmentation and mean deviation based segmentation approaches. Meanwhile, Afza et al. [26] used superpixel computation employing simple linear iterative clustering technique, reconstruction and map generation as segmentation approach.

V. Skin lesion classification

Several researchers currently employ DNNs to improve the skin classification task. Some of them provide a binary classification (malignant or benign), others go further and classify the lesion by type of disease. Commonly, pre-processing and image segmentation are employed before the classification step. These steps are optional but can help the DNN classifier to better perform. The performance metrics most used to evaluate the algorithm are accuracy (ACC), Area Under ROC Curve (AUC) and precision (PR). In this section, a summary of methods for skin lesion classification is presented.

A. *Transfer-learning approaches*

Instead of training a DNN from scratch, you can use the knowledge of a network trained on a large dataset for one problem and use it to solve another problem. This technique is called transfer learning. Furthermore, an additional training set can be employed on the pre-trained network using another data set performing fine-tuning. These two techniques are often applied to overcome the problem of missing data, which is a prevalent obstacle in the skin lesion classification task.

Hasan et al. [39] used a framework that combines pre-processing approaches and a hybrid convolutional neural network (hybrid-CNN) applied to the skin lesion classification task. Three feature map generators initialized with pre-trained weights compose the hybrid-CNN: the first built with residual and convolutional blocks, the second consists of 3 convolutional blocks in sequence and the third is made with dense and transition blocks. Therefore, three feature maps are obtained and an extra one is built by fusing their outputs. The four feature maps are classified by different fully connected layers (FC). In the pre-processing step, a fine-tuned DSnet is used to obtain the lesion ROI, and then, data augmentation and class rebalancing are employed. The final result is an average of the FCs outputs. The training and the evaluation were performed using ISIC 2016, ISIC 2017 and ISIC 2018 datasets. The results obtained on ISIC 2016 were PR of 0.92 and AUC of 0.96, on ISIC 2017 were PR of 0.86 and AUC of 0.95, and, on ISIC 2018 were a PR of 0.85 and AUC of 0.97.

Khan et al. [27] proposed a model consisting of DenseNet for deep feature extraction with coupled Newton-Raphson driven iteration to select the most important features by removing redundant and irrelevant information. A multilayer feed-forward neural network operates as a classifier based on the selected features. Contrast stretching is performed and a faster RCNN is used for the segmentation before the feature extraction. The result for the ISIC 2016 was an ACC of 0.94 and an AUC of 0.98. For the ISIC 2017 an ACC of 0.934 and an AUC of 0.98 were obtained.

Hosny et al. [40] modified the AlexNet model by replacing the classification layer with a softmax layer and fine-tuning the weights. Before classification, data augmentation was performed and all images were segmented. The datasets DermIS-DermQuest, MEDNODE and ISIC 2017 were used for training and testing, with the first two divided into two classes and the last into three classes. The results of ACC were 0.969 for DermIS-DermQuest, 0.977 for MEDNODE and 0.959 for ISIC 2017.

Afza et al. [26] used a ResNet-50 for classification. Both local and global contrast enhancement were applied on the images as a first step and superpixel computation was used to segment the image. In the classification part, using pre-trained and fine-tuned ResNet-50, Grasshopper algorithm was employed to optimize the feature selection. At last, the ACC obtained for ISIC 2016 dataset was 0.911, on the PH2 dataset was 0.954, and on the HAM1000 dataset was 0.858.

To evaluate the accuracy of the transfer-learning against trained from scratch models, Lopez et al. [41] used the VGGNet in three different ways: training from scratch, using transfer-learning and using transfer-learning and fine-tuning the architecture. The input images were pre-processed by applying pixel normalization, image cropping and resizing, and data augmentation. The better-evaluated method was the fine-tuned VGGNet with ACC of 0.813. It obtained an ACC of 0.660 by training from scratch method, and, an ACC of 0.686 by transfer-learning method on ISIC 2016.

The efficiency of multiple pre-trained model combinations was also assessed in recent literature. Mahbod et al. [42] combined AlexNet, VGGNet16, ResNet-18 and ResNet-101 using inter and intra-network fusion and fine-tuning them. The fine-tuning was made several times with different settings, and a support vector machine (SVM) classifier is trained using the extracted features. colour normalization and resizing approaches were employed in this model. The training and testing were performed on ISIC 2016 and ISIC 2017. The obtained results were ACC of 0.877 and AUC of 0.914 on the ISIC 2017. The winner method for ISIC 2019 Challenge proposed by Gessert et. al. [30] also consists of the ensemble different architectures: a pre-trained EfficientNet, a SENet154 and two pre-trained ResNeXt-101 with weakly supervised learning. The metadata provided was processed by a two-layer dense NN. Data augmentation, cropping strategy and Shades of Gray method were applied as pre-processing steps. The datasets used for training were ISIC 2019, 7-point dataset and a private one. The final results obtained was and ACC of 0.926 and AUC of 0.923, and, by adding the metadata information, it was obtained ACC of 0.925 and AUC of 0.903.

Another leaderboarder of ISIC 2019 was the method proposed by Pacheco et al. [31] that used the ensemble of 13 pre-trained and fine-tuned models. In the pre-processment step data augmentation was performed. The final classifications consisted of the probabilities average given by each model. Training and testing were performed in ISIC 2019 dataset obtaining an ACC of 0.919 and an AUC of 0.892.

B. Hybrid approaches

Khouloud et al. [24], after employing image pre-processing and segment the images using the W-net, applied the Inception-Resnet for binary skin lesion classification. Inception-Resnet is a modified Inception model composed of residual blocks that combines the the residual block capability of solving the degradation problem with Inception concept that uses different sizes of kernels to gather both very apparent information and small details. Three convolutional layers are added before to reduce the computational cost. ISIC 2016, ISIC 2017 and PH2 datasets were used to train and evaluate the model. An ACC of 0.981, 0.969 and 0.985 was obtained on the ISIC 2016, ISIC 2017 and PH2 datasets respectively.

Jayapriya et al. [43] proposed a deep residual network (DRN) along with a handcrafted feature to provide better accuracy. The hybrid FCN for segmentation mentioned in the previous section is employed before the classification step. The DRN is composed of 50 layers of residual network and, along with it, a Local texton XOR patterns is added to form a new feature vector. The ISIC 2016 is used for training and ISIC 2016 and 2017 are used for testing. Finally, the ACC obtained is 0.889 for ISIC 2016 and 0.853 for ISIC 2017.

Sarkar et al. [29] proposed a model based on residual learning and separable convolutional approaches to perform binary skin lesion classification. Pre-processing algorithms were applied to perform noise removal and image enhancement. The model was trained on ISIC dataset and obtained an ACC of 0.995. Tests in other datasets were performed and the obtained results were 0.968 on PH2 dataset, 0.952 on MEDNODE dataset and 0.944 on DermIS dataset.

Ge et al. [44] employed a CNN model to extract different types of features from the segmented skin lesion image. The extracted features are convolutional features, location features, statistical parameters and gray-level co-occurrence matrix features. These features are used to classify the lesions from the ISIC 2016 between malignant or benign. The model obtained an ACC of 0.92.

Pollastri et al. [45] achieve the third best result int he first task of ISIC 2019 Challenge by ensembling different models and training them with different data augmentation approaches. ResNet-152, DenseNet-201 and SeResNext-50 compose the final approach with their outputs ensemble by an average of calibrated probabilities using Temperature Scaling. For training, only ISIC 2019 dataset was used obtaining an ACC of 0.924 and AUC of 0.886.

VI. Analysis and discussion

In this study, was proposed a review of works related to the tasks of skin lesions detection and classification. Most of the mentioned datasets are imbalanced, i. e., the number of images per class is uneven. In addition, the ISIC dataset, which is largely used in skin disease research, has several duplicated images. These problems can affect the model performance and influence the final result since the evaluation measures tend to bias towards majority classes. To address these issues, Cassidy et al. [46] proposed a curated ISIC dataset and a strategy for removing duplicated images. Even more, the volume of dermoscopy and clinical photos of skin lesions is small, which can obstruct the training and development of more efficient algorithms. To minimize the impact of the lack of available images, data augmentation and transfer learning can be employed [47].

For the skin lesion classification, the models that outperform the others are the ones proposed by Khouloud et al. [24], Khan et a. [27] and Sarkar et al. [29]. The common characteristic between this methods is the special attention given to pre-processing steps. Khouloud et al. [24] used the pre-processing steps mentioned before and W-net for segmentation. Khan et al. [27] used an artificial bee colony algorithm for contrast stretching and a Faster RCNN for lesion segmentation. Sarkar et al. [29] tested and compared different noise removal to pick the optimal one, applying image enhancement, and tested merging channels from different colour spaces into the RGB image to increase the accuracy of the model. Both models proposed by Khouloud et al. and Sarkar et al. use the deep residual concept but applyed it in two different ways. Khouloud et al. used this concept combined with Inception network and Sarkar et al. used a depthwise separable convolution to improve the efficiency and residual learning concepts to design of their models. Khan et al. used a fine-tuned DenseNet with a Newton Raphson-based method for feature selection in order to discard not relevant features.

The ensemble of different architectures in order to combine their outputs to obtain a final classification is a very popular approach found in the latest literature and very explored in ISIC 2019 Challenge [30], [31], [42], [45]. Moreover, as mentioned in Section 5.1, transfer-learning and fine-tuning techniques are largely used as a way to overcome the lack of data related to the skin lesion classification task and provide better results. Align to that, data augmentation is extensively used to improve the effectiveness of the models. The manipulation of the image in terms of position, contrast and saturation can provide robustness to the approach by mimicking the light and colour disparities that different devices can have when capturing these images [45].

VII. Conclusion

Skin cancer is one of the most incident cancers and, in the case of melanomas, it can be fatal if not treated early. The CAD systems are created to aid in the diagnosis of skin lesions by professionals. Among them, DL is a promising area of study for solving this type of problem.

In this study, a review and analysis of the works related to the tasks involved in the detection of skin lesions using DL. It was made an effort to present works with different approaches to promote a more diverse analysis. In addition, it was important to mention the datasets available for these purposes, and some image pre-processing methods that can help the models to better perform the tasks.

Acknowledgment

The authors are grateful to the Foundation for Science and Technology (FCT, Portugal) for financial support through national funds FCT/MCTES (PIDDAC) to CeDRI (UIDB/05757/2020 and UIDP/05757/2020) and SusTEC (LA/P/0007/2021).

References

[1] D. L. Narayanan, R. N. Saladi, and J. L. Fox, "Review: Ultraviolet radiation and skin cancer," *Int. J. of Derm.*, vol. 49, no. 9, pp. 978–986, 2010.

[2] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.

[3] J. Ferlay, M. Ervik, C. Lam, F., L. M., Mery, M. Piñeros, A. Znaor, I. Soerjomataram, and F. Bray, "Global cancer observatory: Cancer today. lyon, france: Int. ag. for research on cancer." last accessed 2 May 2022.

[4] N. R. Abbasi, H. M. Shaw, D. S. Rigel, R. J. Friedman, W. H. McCarthy, I. Osman, A. W. Kopf, and D. Polsky, "Early Diagnosis of Cutaneous Melanoma: Revisiting the ABCD Criteria," *JAMA*, vol. 292, no. 22, pp. 2771–2776, 2004.

[5] C. A. Morton and R. M. Mackie, "Clinical accuracy of the diagnosis of cutaneous malignant melanoma," *British Journal of Dermatology*, vol. 138, no. 2, pp. 283–287, 1998.

[6] S. Mane and S. Shinde, "A method for melanoma skin cancer detection using dermoscopy images," pp. 1–6, 2018.

[7] 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," *Archives of Derma.*, vol. 131, no. 3, pp. 286–291, 1995.

[8] K. Korotkov and R. Garcia, "Computerized analysis of pigmented skin lesions: A review," *Artificial Intell. in Medicine*, vol. 56, no. 2, pp. 69–90, 2012.

[9] V. Vipin, M. K. Nath, V. Sreejith, N. F. Giji, A. Ramesh, and M. Meera, "Detection of melanoma using deep learning techniques: A review," in *2021 International Conference on Communication, Control and Information Sciences (ICCISc)*. IEEE, 2021.

[10] H. Li, Y. Pan, J. Zhao, and L. Zhang, "Skin disease diagnosis with deep learning: A review," *Neurocomputing*, vol. 464, pp. 364–393, 2021.

[11] M. Dildar, S. Akram, M. Irfan, H. U. Khan, M. Ramzan, A. R. Mahmood, S. A. Alsaiari, A. H. M. Saeed, M. O. Alraddadi, and M. H. Mahnashi, "Skin cancer detection: A review using deep learning techniques," *International Journal of Environmental Research and Public Health*, vol. 18, no. 10, p. 5479, 2021.

[12] R. B. Oliveira, M. E. Filho, Z. Ma, J. P. Papa, A. S. Pereira, and J. M. R. Tavares, "Computational methods for the image segmentation of pigmented skin lesions: A review," *Computer Methods and Programs in Biomedicine*, vol. 131, pp. 127–141, 2016.

[13] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, H. Kittler, and A. Halpern, "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic)," 2019.

[14] M. S. K. C. Center, "The international skin imaging collaboration," last accessed 15 May 2022. [Online]. Available: https://www.isic-archive.com/

[15] D. Gutman, N. C. F. 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 int. skin imaging collaboration (isic)," 2016.

[16] N. C. F. Codella, D. Gutman, M. E. Celebi, B. Helba, M. A. Marchetti, S. W. Dusza, A. Kalloo, K. Liopyris, N. Mishra, H. Kittler, and A. Halpern, "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)," 2017.

[17] P. Tschandl, C. Rosendahl, and H. Kittler, "The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," *Scientific Data*, vol. 5, no. 1, 2018.

[18] M. Combalia, N. C. F. Codella, V. Rotemberg, B. Helba, V. Vilaplana, O. Reiter, C. Carrera, A. Barreiro, A. C. Halpern, S. Puig, and J. Malvehy, "Bcn20000: Dermoscopic lesions in the wild," 2019.

[19] V. Rotemberg, N. Kurtansky, B. Betz-Stablein, L. Caffery, E. Chousakos, N. Codella, M. Combalia, S. Dusza, P. Guitera, D. Gutman, A. Halpern, B. Helba, H. Kittler, K. Kose, S. Langer, K. Lioprys, J. Malvehy, S. Musthaq, J. Nanda, O. Reiter, G. Shih, A. Stratigos, P. Tschandl, J. Weber, and H. P. Soyer, "A patient-centric dataset of images and metadata for identifying melanomas using clinical context," *Scientific Data*, vol. 8, no. 1, 2021.

[20] T. Mendonca, P. M. Ferreira, J. S. Marques, A. R. S. Marcal, and J. Rozeira, "Ph2 - a dermoscopic image database for research and benchmarking," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2013.

[21] I. Giotis, N. Molders, S. Land, M. Biehl, M. F. Jonkman, and N. Petkov, "MED-NODE: A computer-assisted melanoma diagnosis system using non-dermoscopic images," *E. S. with A.*, vol. 42, no. 19, pp. 6578–6585, 2015.

[22] H. Liao, Y. Li, and J. Luo, "Skin disease classification versus skin lesion characterization: Achieving robust diagnosis using multi-label deep neural networks," in *2016 23rd International Conference on Pattern Recognition (ICPR)*, 2016, pp. 355–360.

[23] A. G. P. et al., "PAD-UFES-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones," *Data in Brief*, vol. 32, p. 106221, 2020.

[24] S. Khouloud, M. Ahlem, T. Fadel, and S. Amel, "W-net and inception residual network for skin lesion segmentation and classification," *Applied Intelligence*, vol. 52, no. 4, pp. 3976–3994, 2021.

[25] T. Akram, H. M. J. Lodhi, S. R. Naqvi, S. Naeem, M. Alhaisoni, M. Ali, S. A. Haider, and N. N. Qadri, "A multilevel features selection framework for skin lesion classification," *Human-centric Computing and Information Sciences*, vol. 10, no. 1, 2020.

[26] F. Afza, M. Sharif, M. Mittal, M. A. Khan, and D. Jude Hemanth, "A hierarchical three-step superpixels and deep learning framework for skin lesion classification," *Methods*, vol. 202, pp. 88–102, 2022, machine Learning Methods for Bio-Medical Img. and Sig. Processing: Recent Advances.

[27] M. A. Khan, M. Sharif, T. Akram, S. A. C. Bukhari, and R. S. Nayak, "Developed newton-raphson based deep features selection framework for skin lesion recognition," *Pattern Rec. Letters*, vol. 129, pp. 293–303, 2020.

[28] M. Goyal, A. Oakley, P. Bansal, D. Dancey, and M. H. Yap, "Skin lesion segmentation in dermoscopic images with ensemble deep learning methods," *IEEE Access*, vol. 8, pp. 4171–4181, 2020.

[29] R. Sarkar, C. Chatterjee, and A. Hazra, "Diagnosis of melanoma from dermoscopic images using a deep depthwise separable residual convolutional network," *IET Image Processing*, vol. 13, 2019.

[30] N. Gessert, M. Nielsen, M. Shaikh, R. Werner, and A. Schlaefer, "Skin lesion classification using ensembles of multi-resolution efficientnets with meta data," *MethodsX*, vol. 7, p. 100864, 2020.

[31] A. Pacheco, A.-R. Ali, and T. Trappenberg, "Skin cancer detection based on deep learning and entropy to detect outlier samples," 09 2019.

[32] 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.

[33] M. K. Hasan, L. Dahal, P. N. Samarakoon, F. I. Tushar, and R. Martí, "DSNet: Automatic dermoscopic skin lesion segmentation," *Computers in Biology and Medicine*, vol. 120, p. 103738, 2020.

[34] P. Shan, Y. Wang, C. Fu, W. Song, and J. Chen, "Automatic skin lesion segmentation based on FC-DPN," *Computers in Biology and Medicine*, vol. 123, p. 103762, 2020.

[35] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. 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.

[36] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention*. Springer International Publishing, 2015, pp. 234–241.

[37] D. K. Nguyen, T.-T. Tran, C. P. Nguyen, and V.-T. Pham, "Skin lesion segmentation based on integrating efficientnet and residual block into u-net neural network," in *2020 5th International Conference on Green Technology and Sustainable Development (GTSD)*, 2020, pp. 366–371.

[38] F. Xie, J. Yang, J. Liu, Z. Jiang, Y. Zheng, and Y. Wang, "Skin lesion segmentation using highresolution convolutional neural network," *Computer Methods and Programs in Biomedicine*, vol. 186, p. 105241, 2020.

[39] M. K. Hasan, M. T. E. Elahi, M. A. Alam, M. T. Jawad, and R. Martí, "DermoExpert: Skin lesion classification using a hybrid convolutional neural network through segmentation, transfer learning, and augmentation," *Informatics in Medicine Unlocked*, vol. 28, p. 100819, 2022.

[40] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Classification of skin lesions using transfer learning and augmentation with Alex-net," *PLOS ONE*, vol. 14, no. 5, p. e0217293, 2019.

[41] A. Romero-Lopez, X. G. i Nieto, J. Burdick, and O. Marques, "Skin lesion classification from dermoscopic images using deep learning techniques," in *Biomedical Engineering*. ACTAPRESS, 2017.

[42] A. Mahbod, G. Schaefer, I. Ellinger, R. Ecker, A. Pitiot, and C. Wang, "Fusing fine-tuned deep features for skin lesion classification," *Computerized Medical Imaging and Graphics*, vol. 71, pp. 19–29, 2019.

[43] K. Jayapriya and I. J. Jacob, "Hybrid fully convolutional networks-based skin lesion segmentation and melanoma detection using deep feature," *Int. Journal of Imaging Systems and Technology*, vol. 30, no. 2, pp. 348–357, 2019.

[44] Y. Ge, B. Li, Y. Zhao, E. Guan, and W. Yan, "Melanoma segmentation and classification in clinical images using deep learning," in *Proceedings of the 2018 10th International Conference on Machine Learning and Computing*. Association for Computing Machinery, 2018, p. 252–256.

[45] F. Pollastri, M. Parreño, J. Maroñas, F. Bolelli, R. Paredes, D. Ramos, and C. Grana, "A deep analysis on high-resolution dermoscopic image classification," *IET Computer Vision*, vol. 15, no. 7, pp. 514–526, 2021.

[46] B. Cassidy, C. Kendrick, A. Brodzicki, J. Jaworek-Korjakowska, and M. H. Yap, "Analysis of the isic image datasets: Usage, benchmarks and recommendations," *Medical Image Analysis*, vol. 75, p. 102305, 2022.

[47] F. Perez, C. Vasconcelos, S. Avila, and E. Valle, "Data augmentation for skin lesion analysis," in *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*. Springer Int. Publishing, 2018, pp. 303–311.