

Detection and Multiclass Classification of Skin Cancers Using Modified YOLOv8

Nathan Philip Bulla, Vellore Institute of Technology-Andhra Pradesh, Amaravati, Andhra Pradesh, India

Abstract

Skin cancer is the most widespread form of cancer, and prevention and early detection are key in tackling the disease. This paper details the implementation of the You Only Look Once version 8 (YOLOv8) model on the ISIC 2019 skin cancer dataset, to achieve a multi-class classification of the eight different skin conditions: Actinic keratosis, Basal cell carcinoma, Benign Keratosis, Dermatofibroma, Melanocytic nevus, Melanoma, Squamous cell carcinoma and Vascular lesions; and it's comparative performance with respect to other deep learning techniques. For preprocessing, Dull Razor algorithm is used, along with Contrast Limited Adaptive Histogram Equalisation (CLAHE) to improve visibility,

Keywords: skin cancer; YOLO; ISIC; skin lesion; black hat filter;

1. Introduction

1.1 Background

Cancer is a condition wherein the cells of the body lose their sensitivity to contact inhibition, causing them to proliferate rapidly while unchecked.

Skin cancer is the most common type of cancer worldwide, with increasing incidence rates over the years. Skin cancer develops typically as a result of prolonged exposure to ultraviolet (UV) radiation.

There are multiple different types of cancer, including those such as Basal Cell Carcinoma, Squamous Cell Carcinoma and Melanoma.

There can be several risk factors at play for the contraction of skin cancer. Prolonged exposure to ultraviolet radiation being the predominant cause; but can also be attributed to a family history of skin cancer, fair skin, and certain genetic conditions.

Prevention and early detection are crucial in combating skin cancer.

Prevention can be achieved by adopting sun-safe practices aimed at protecting the skin from UV radiation. These should include wearing protective clothing, using broad-spectrum sunscreen, avoiding direct sunlight and artificial tanning devices.

Early detection and treatment play a crucial role in reducing risk of complications.

Symptoms of skin cancer are varied, ranging from changes in appearance of moles and growths, itching or bleeding skin lesions, unusual sores, etc.

Regular self-examinations and dermatological check-ups are recommended to detect any suspicious skin lesions early.

Treatment options for skin cancers after detection may include surgical removal, radiation therapy, chemotherapy, immunotherapy, and also targeted therapy.

1.2 Contribution of proposed work

The paper uses a mix of various datasets, some of which has not been extensively explored.

The paper proposes the use of the YOLOv8 model to create an real-time automatic identifier for the type of skin cancer using dermoscopic skin lesion images.

1.3 Organisation of the paper

Section 2 consists of a literature survey done on related work in the domain of skin cancer classification using various techniques, models and approaches.

Section 3 illustrates a brief comparative analyses on previous literature based on the methods implemented, dataset worked on, diseases classified, and performance analyses of the approaches used.

Section 4 describes the proposed methodology, including the image preprocessing techniques, proposed model and model description.

Section 5 displays the results obtained by the proposed model on the dataset, with comparisons to other models.

Section 6 states the findings of the paper, along with the future scope for prospective work in this domain.

2. Literature Survey

Aishwarya et al.[1] used the YOLOv3 and YOLOv4 models to achieve multiclass classification of skin cancers into 9 classes, using the ISIC dataset. YOLOv3 and YOLOv4 achieve mean average precision scores of 88.03% and 86.52% respectively. YOLOv3 achieves accuracy of 0.981, precision of 0.921, recall of 0.921 and an F1 score of 0.921. YOLOv4 achieves an accuracy of 0.98, precision of 0.934, recall of 0.9, and an F1 score of 0.917.

Huang et al. [2] used the novel approach of converting the RGB images into Hyperspectral narrow-band images (HSI-NBI) and then subsequently fed to a YOLOv5 model, which is proven to have better detection speed to aid in real-time performance. Their model focused primarily on the classification of the three classes: Basal cell carcinoma, Squamous cell carcinoma, and Seborrheic keratosis. As for the RGB model, they achieved a precision of 0.888, recall of 0.758, specificity of 0.798, F1 score of 0.818, and an overall accuracy of 0.792. With the HSI model

they achieved a precision of 0.8, recall of 0.726, specificity of 0.786, F1 score of 0.761, and an accuracy of 0.787.

Karar et al. [3] worked with the HAM10000 dataset. They tested the performance of all EfficientNet models from B0 to B7. For preprocessing, they used Image inpainting to remove hair and other noise from the lesion images, and dataset size was increased using data augmentation. They used pretrained EfficientNet models, trained on ImageNet and fine-tuned. The best performance of the models was demonstrated by the B4 and B5 models with precision of 0.88 and 0.88, recall of 0.88 and 0.88, F1 score of 0.87 and 0.87, specificity of 0.88 and 0.88, and accuracy of 0.9753 and 0.9754 respectively.

Raghavendra et al. [4] proposed a novel Deep convolutional neural network (DCNN) using global average pooling along with preprocessing. They used the histogram equalization (CLAHE) to improve visibility and contrast, Black Hat filtering to remove hair and other noise, and data augmentation to balance the imbalanced dataset, along with Global average pooling. The proposed novel DCNN model achieved a precision of 0.97, recall of 0.97, F1 score of 0.97, ROC-AUC score of 0.997 and a testing accuracy of 0.972.

Venugopal et al. [5] proposed a deep neural network model consisting of a modified EfficientNetV2-M. They combined the ISIC 2018, 2019 and 2020 challenge datasets, which themselves were an amalgamation of the HAM10000, BCN20000, and ISBI images. Using the modified EfficientNetV2-M They achieved accuracy of 97.62% on the ISIC 2020 dataset, 95.49% on the ISIC 2019, and 94.80% on the HAM10000 dataset. With weighted precision, recall and F1 scores of 0.98, 0.96, 0.97 on the ISIC 2020, ISIC 2019, and HAM10000 datasets respectively. They achieved weighted recall scores of 0.98, 0.95, 0.95 on the ISIC 2020, ISIC 2019, and HAM10000 datasets respectively. As for the weighted F1 scores, they achieved 0.98, 0.95, 0.95 on the ISIC 2020, ISIC 2019, and HAM10000 datasets.

Magdy et al. [6] proposed two methods to classify the skin lesion images: KNN with pretrained DNN, and AlexNet with Grey Wolf Optimizer (GWO). For preprocessing, they resized the images, hair removal, cropping. Using KNN with pretrained VGG-16, they achieved Precision of 99.7481%, sensitivity of 99%, specificity of 99.75%, F1 score of 99.3726% and Accuracy of 99.375%. Using AlexGWO (AlexNet with Grey wolf Optimizer), they achieved precision of 99.47%, recall of 100%, specificity of 99.5%, F1 score of 98.63%, Accuracy of 99%.

Mehmood et al. [7] used the HAM10000 dataset for training. Their proposed modified model was the SBXception, which is a shallower and broader version of the Xception network. They achieved a precision score of 95.43%, recall score of 85.34%, and an accuracy of 96.97%.

J. M. Cadena et al. [8] proposed two Novel CNN models: CNN-1 and CNN-2, on a 2860 image sample of data from the ISIC archive. The better performing model, CNN-2, reached an AUC score of 0.915 ± 0.02 , and achieved and AUC score of 0.9626 on a test dataset.

Ogudo et al. [9] proposed the use of Optimized Stacked Sparse Autoencoder feature extractor with a Backpropagation Neural Network (OSSAE-BPNN). It uses the Sea Gull Optimization for the parameter tuning of the model. For preprocessing, the images were resized, and to remove

hair and noise, they used the Dull Razor algorithm. They achieved a maximum average accuracy of 94.7%, recall of 79.6%, specificity of 96.9%, precision of 80.6%, and F-score of 79.6%.

Zafar et al. [10] proposed segmentation via the use of DeepLabV3 model with the back pillar based on a pretrained MobileNetV2 model, and proposed the achievement of multiclass classification by feature extraction through the DenseNet201 and Slime mould algorithm (SMA), and the important extracted features are passed on to SVM and KNN classifiers. The proposed model gives an average accuracy of 91.7% on the ISIC 2019 dataset, 92.01% on HAM10000, 98.88% on the PH2 dataset, and 99.3% on the Med-Node dataset.

3. Comparative Analysis

S. No.	Authors	Year	Datas et	Method	Classification	Results
1	Aishwarya et al. [1]	2023	ISIC	YOLOv3 and YOLOv4	Actinic keratosis, Basal cell carcinoma, Dermatofibroma, Melanoma, Nevus, Pigmented benign keratosis, Seborrheic keratosis, Squamous cell carcinoma, Vascular lesion	YOLOv3: acc 0.981, prec 0.921, recall 0.921, F1 0.921; YOLOv4: acc 0.98, prec 0.934, rec 0.9, F1 0.917
2	Huang et al. [2]	2023	ISIC	YOLOv5 with HSI model	Basal cell carcinoma, Squamous cell carcinoma, Seborrheic keratosis	YOLOv5 with RGB: prec 0.888, rec 0.758, spec 0.798, F1 0.818; YOLOv5 with HSI: prec 0.8, rec 0.726, spec 0.786, F1 0.761, acc 0.787
3	Karar et al. [3]	2022	HAM 10000	EfficientNets B0-B7	Actinic keratosis, Basal cell carcinoma, Dermatofibroma, Benign keratosis, Melanoma, Nevus, Vascular lesions	EfficientNet B4: prec 0.88, rec 0.88, F1 0.87, spec 0.88, roc_auc 0.9753; EfficientNetB5: prec 0.88, rec 0.88, F1 0.87, spec 0.88, roc_auc 0.9754
4	Raghaven dra et al. [4]	2023	HAM 10000	Novel DCNN	Actinic keratosis, Basal cell carcinoma, Dermatofibroma, Benign keratosis, Melanoma, Nevus, Vascular lesions	prec 0.97, rec 0.97, F1 0.97, ROC-AUC 0.997, test acc 0.972

5	Venugopal et al.[5]	2023	ISIC 2020, 2019, 2018	Modified EfficientNetV2-M and EfficientNet-B4	Actinic keratosis, Basal cell carcinoma, Dermatofibroma, Benign keratosis, Melanoma, Nevus, Vascular lesions	Accuracy: 97.62% on the ISIC 2020 dataset, 95.49% on the ISIC 2019, 94.80% on the HAM10000 dataset. weighted precision scores of 0.98, 0.96, 0.97 on the ISIC 2020, ISIC 2019, and HAM10000 datasets respectively. weighted recall: 0.98, 0.95, 0.95, weighted F1 scores: 0.98, 0.95, 0.95
---	---------------------	------	-----------------------	---	--	--

4. Methodology

4.1. Dataset

The chosen dataset is the ISIC - 2019 challenge dataset that contains skin lesion images from various patients suffering from various malignant and benign skin cancers. The dataset contains The dataset in all contains 8 classes, namely Actinic Keratosis, Basal cell carcinoma, Benign keratosis, Dermatofibroma, Melanocytic nevus, Melanoma, Squamous cell carcinoma, and Vascular lesions.

The Dataset used is a mix of the ISIC 2019 dataset (which primarily consists of the HAM10000[11], BCN20000[12], MSK[13] dataset), ISIC 2020[14], and the PAD-UFES-20[15] dataset.

The PAD-UFES-20 dataset is used primarily for validation and testing purposes, to emulate real-time application of the model in regions where skin cancer diagnosis and research has not been carried out extensively, and is achieved by using the PAD-UFES-20 dataset which is not part of the ISIC Archive, and has not been used for training.

Dataset	ACK	BCC	BKL	DF	MEL	NEV	SCC	VAS
ISIC 2019	867	3323	2624	239	4522	12875	628	253
ISIC 2020	-	-	176	-	2738	5196	-	-
HAM10000	149	622	1338	160	1305	7737	229	180
BCN20000	737	2809	929	124	2857	4206	431	111

PAD-UFES-20	980	845	235	-	52	244	192	-
-------------	-----	-----	-----	---	----	-----	-----	---

Table 1

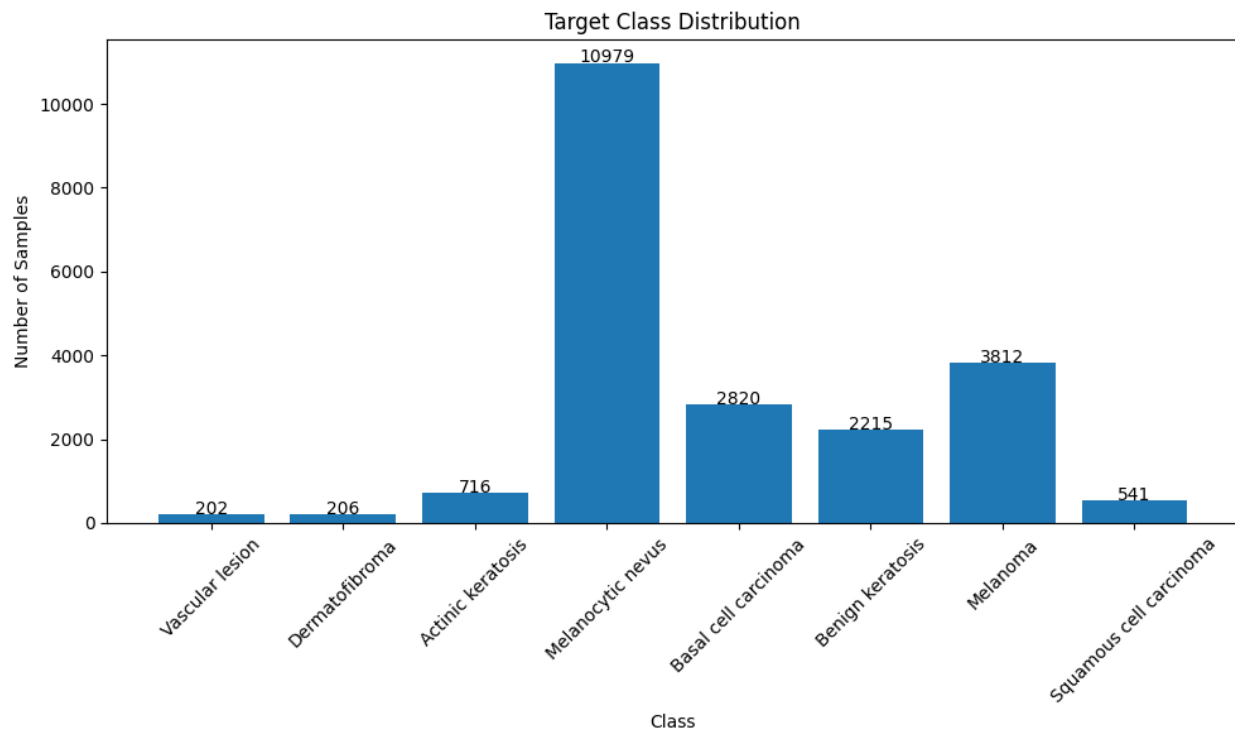


Fig. 1.0 Target class distribution showing imbalance in the ISIC 2019 training dataset



Fig. 1.1 Target class distribution showing imbalance in the total combined dataset.

4.2. Preprocessing

4.2.1 Data Augmentation

The dataset is quite imbalanced, therefore Data Augmentation is performed to boost the counts of the minority classes to achieve a balanced distribution, to achieve effective multiclass classification.

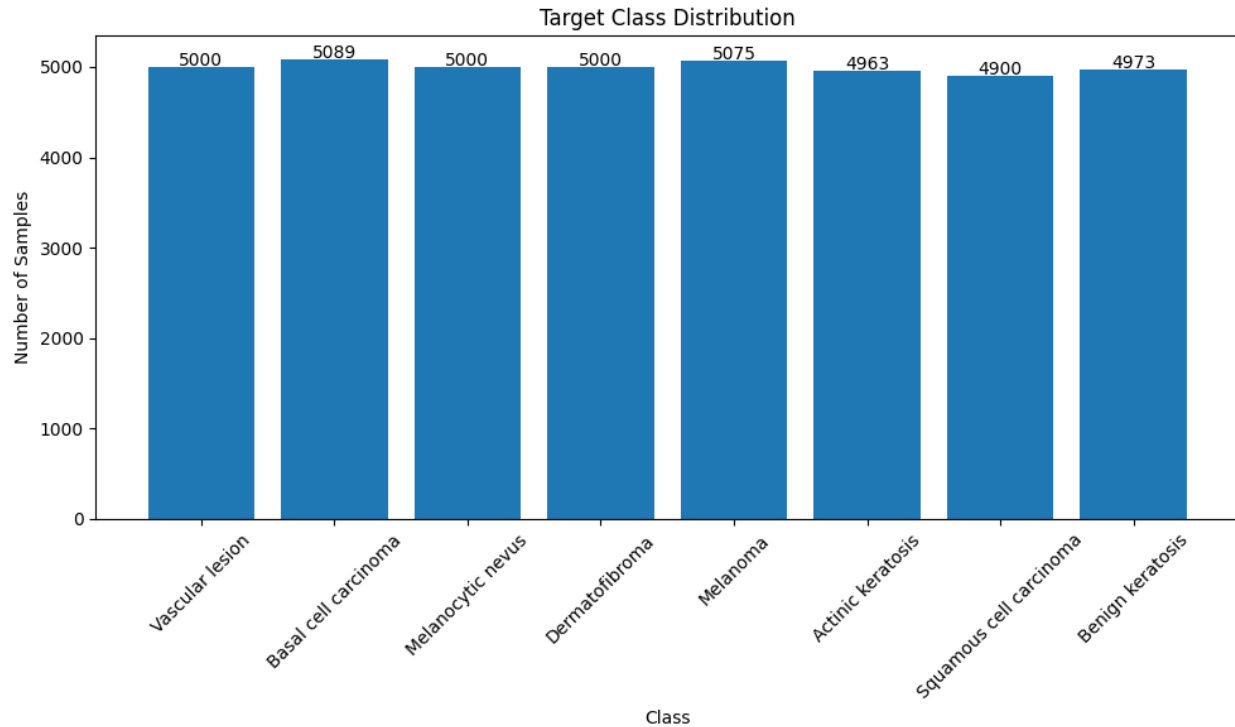


Fig. 1.2 Target class distribution of the ISIC - 2019 dataset post augmentation equalizing to around 5000 images per class

4.2.2 Hair Removal

Most of the images in the dataset have hair near, on or within the skin lesions, this will contribute to noise and may hinder the training. A common preprocessing technique used to solve the issue in similar studies [3][4] is the Black Hat filtering. This technique is used to remove the noise from the data.



(a) (b)
Fig. 1 (a) Image before hair removal (b) Image after hair removal

4.2.3 Contrast Limited Adaptive Histogram Equalization (CLAHE)

CLAHE is an algorithm used to boost the visibility of images. This makes it easier for feature extraction as the skin lesions are more pronounced and the ambiguity is reduced. Some patterns which might have been overlooked will become more prominent, and easier for the models to detect and process.

4.3. Proposed Model

The proposed model to be used is the You Only Look Once version 8 (YOLOv8), a State-of-the-art object detection algorithm that can perform a multitude of object detection, image classification, and image segmentation tasks.

The performance of the model is compared with pre-existing models, case in point, AlexNet, VGG19, MobileNetV3-Small, EfficientNet B5, YOLOv4, YOLOv5.

4.4. Model description

YOLOv8 is a complex model, consisting of the Backbone, a neck, and multiple heads.

It is reliable and fast, making it ideal for real time applications.

The backbone of the model is based on the CSPDarknet53 architecture, and is responsible for feature extraction at high resolution.

The neck connects the backbone to the heads. It uses a C2f module to improve the object detection prediction.

The heads perform the function of prediction of the bounding boxes, class labels and the confidence scores. The model has three heads, each predicting a bounding box of a different scale, enabling the detection of differently sized objects.

5. Result Analysis

The work is carried out on Google colab, using GPU runtime utilizing the Nvidia Tesla T4 GPU.

Upon using pretrained YoloV8n model on ISIC-2019 challenge dataset without augmentation, the observations were as follows. The training was performed for 10 epochs, batch size of 16, patience of 50, image size 64, using AdamW optimizer with learning rate 0.000714, momentum 0.9.



Fig. 3.1 Normalized Confusion matrix of YOLOv8n on ISIC 2019 without augmentation and without hair removal. Observed accuracies: top1-64.3%, top5-98.89%

Upon using pretrained YOLOv8m model on ISIC-2019 challenge dataset without augmentation or hair removal, the observations were as follows. The training was performed for 10 epochs, batch size of 16, patience of 50, image size 64, using AdamW optimizer with learning rate 0.000714, momentum 0.9.



Fig. 3.2 Normalized confusion matrix of YOLOv8m model on the ISIC-2019 challenge dataset without augmentation of hair removal.
Observed accuracies: top1-74.1%, top5-99.1%

6. Conclusion

The state-of-the-art model YOLOv8 was proposed to perform the task of multi-class classification of skin cancers using dermoscopic pictures of skin lesions.

Its performance was compared with pre-existing models such as AlexNet, VGG19, MobileNet, and also with older YOLO models.

The observations show that the model performs better than previous models across various metrics.

There is potential for creating and training a unified dataset consisting of all major and minor skin lesion datasets to achieve identification of an even greater number of skin diseases, not limited to just cancers.

7. References

- [1] N Aishwarya, K Manoj Prabhakaran, Frezewd Tsegaye Debebe, M Sai Sree Akshitha Reddy, Posina Pranavee, Skin Cancer diagnosis with Yolo Deep Neural Network, Procedia Computer Science, Volume 220, 2023, Pages 651-658, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2023.03.083>.
- [2] Huang, H.-Y., Hsiao, Y.-P., Mukundan, A., Tsao, Y.-M., Chang, W.-Y., & Wang, H.-C. (2023). Classification of Skin Cancer Using Novel Hyperspectral Imaging Engineering via YOLOv5. Journal of Clinical Medicine, 12(3), 1134, <https://doi.org/10.3390/jcm12031134>
- [3] Karar Ali, Zaffar Ahmed Shaikh, Abdullah Ayub Khan, Asif Ali Laghari, Multiclass skin cancer classification using EfficientNets – a first step towards preventing skin cancer, Neuroscience Informatics, Volume 2, Issue 4, 2022, 100034, ISSN 2772-5286, <https://doi.org/10.1016/j.neuri.2021.100034>.
- [4] Raghavendra, P.V.S.P., Charitha, C., Begum, K.G. et al. Deep Learning–Based Skin Lesion Multi-class Classification with Global Average Pooling Improvement. J Digit Imaging (2023). <https://doi.org/10.1007/s10278-023-00862-5>
- [5] Vipin Venugopal, Navin Infant Raj, Malaya Kumar Nath, Norton Stephen, A deep neural network using modified EfficientNet for skin cancer detection in dermoscopic images, Decision Analytics Journal, Volume 8, 2023, 100278, ISSN 2772-6622, <https://doi.org/10.1016/j.dajour.2023.100278>
- [6] A. Magdy, H. Hussein, R. F. Abdel-Kader and K. A. E. Salam, "Performance Enhancement of Skin Cancer Classification using Computer Vision," in IEEE Access, doi: 10.1109/ACCESS.2023.3294974
- [7] Mehmood, A., Gulzar, Y., Ilyas, Q. M., Jabbari, A., Ahmad, M., & Iqbal, S. (2023). SBXception: A Shallower and Broader Xception Architecture for Efficient Classification of Skin Lesions. Cancers, 15(14), 3604. <https://doi.org/10.3390/cancers15143604>
- [8] J. M. Cadena et al., "Melanoma Cancer Classification using Deep Convolutional Neural Networks," 2023 IEEE 13th International Conference on Pattern Recognition Systems (ICPRS), Guayaquil, Ecuador, 2023, pp. 1-7, doi: 10.1109/ICPRS58416.2023.10179049

[9] Ogudo, K. A., Surendran, R., & Khalaf, O. I. (2023). Optimal Artificial Intelligence Based Automated Skin Lesion Detection and Classification Model. *Computer Systems Science & Engineering*, 44(1).

[10] Zafar, M., Amin, J., Sharif, M., Anjum, M. A., Mallah, G. A., & Kadry, S. (2023). DeepLabv3+-Based Segmentation and Best Features Selection Using Slime Mould Algorithm for Multi-Class Skin Lesion Classification. *Mathematics*, 11(2), 364.
<https://doi.org/10.3390/math11020364>

[11] Tschandl, P., Rosendahl, C. & Kittler, H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci Data* 5, 180161 (2018).
<https://doi.org/10.1038/sdata.2018.161>

[12] arXiv:1908.02288
<https://doi.org/10.48550/arXiv.1908.02288>

[13] Noel C. F. Codella, David Gutman, M. Emre Celebi, Brian Helba, Michael A. Marchetti, Stephen W. Dusza, Aadi Kalloo, Konstantinos Liopyris, Nabin Mishra, Harald Kittler, Allan 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;
arXiv:1710.05006.

[14] Rotemberg, V., Kurtansky, N., Betz-Stablein, B., Caffery, L., Chousakos, E., Codella, N., Combalia, M., Dusza, S., Guitera, P., Gutman, D., Halpern, A., Helba, B., Kittler, H., Kose, K., Langer, S., Liopyris, K., Malvey, J., Musthaq, S., Nanda, J., Reiter, O., Shih, G., Stratigos, A., Tschandl, P., Weber, J. & Soyer, P. A patient-centric dataset of images and metadata for identifying melanomas using clinical context. *Sci Data* 8, 34 (2021).
<https://doi.org/10.1038/s41597-021-00815-z>

[15] Andre G.C. Pacheco, Gustavo R. Lima, Amanda S. Salomão, Breno Krohling, Igor P. Biral, Gabriel G. de Angelo, Fábio C.R. Alves Jr, José G.M. Esgario, Alana C. Simora, Pedro B.C. Castro, Felipe B. Rodrigues, Patricia H.L. Frasson, Renato A. Krohling, Helder Knidel, Maria C.S. Santos, Rachel B. do Espírito Santo, Telma L.S.G. Macedo, Tania R.P. Canuto, Luíz F.S. de Barros, PAD-UFES-20: A skin lesion dataset composed of patient data and clinical images collected from smartphones, *Data in Brief*, Volume 32, 2020, 106221, ISSN 2352-3409,
<https://doi.org/10.1016/j.dib.2020.106221>.

