Enhanced transfer learning model by image shifting on a square lattice for skin lesion malignancy assessment

Karl Thurnhofer-Hemsi, Rosa Maza-Quiroga, Enrique Domínguez, Miguel A. Molina-Cabello,
Ezequiel López-Rubio
Department of Computer Languages and Computer Sciences
University of Málaga
Bulevar Louis Pasteur, 35, 29071, Málaga, Spain,
The Biomedic Research Institute of Málaga (IBIMA)
C/ Doctor Miguel Díaz Recio, 28, 29010, Málaga, Spain
Emails: {karlkhader,rosammq,enriqued,miguelangel,ezeqlr}@lcc.uma.es

Abstract—Skin cancer is one of the most prevalent diseases among people. Physicians have a challenge every time they have to determine whether a diseased skin is benign or malign. There exist clinical diagnosis methods (such as the ABCDE rule), but they depend mainly on the physician's experience and might be imprecise. Deep learning models are very extended in medical image analysis, and several deep models have been proposed for moles classification. In this work, a convolutional neural network is proposed to support the diagnosis procedure. The proposed MobileNetV2-based model is improved by a shifting technique, providing better performance than raw transfer learning models for moles classification. Experiments show that this technique could be applied to the state-of-the-art deep models to improve their results and outperform the training phase.

Index Terms—image processing, deep learning, classification, skin lesion

I. INTRODUCTION

One of the several prevalent diseases among people globally is skin cancer. It is worth mentioning different types of skin diseases such as basal cell carcinoma, melanoma, and squamous cell carcinoma [1]. For years, researchers have predicted that 20% of American people will have skin cancer throughout their lives [2]. Therefore skin cancer the most common illness in the USA [3] and melanoma are the most lethal from skin cancer. Years ago, melanoma was unusual cancer, but nowadays the mortality ratio has increased substantially [4]. The American Cancer Society estimated about 6,850 deaths by melanoma for 2020. but there exist other skin cancers widespread in the population. Fortunately, they are generally no mortal, but they suppose considerable pressure on physicians and health care institutions.

The physicians have a challenge every time they have to diagnose a skin disease. The specialist analyzes various characteristics as individual lesional morphology, body place distribution, color, scaling, and lesions arrangement to determine if the diseased skin is benign or malign. There are several clinical diagnosis methods that use all these shreds of evidence to create the best possible diagnosis: ABCDE rules [5], 7-Point Checklist, Menzies method, etc. However, only experienced physicians have the necessary knowledge to analyze and diagnose skin disease by using these methods, and they can be imprecise and subject to many measurement errors. The visual examination of the lesion sampled from a suspicious lesion is the gold standard for skin disease diagnosis, so its correct reporting is essential. Since years ago is known that the diagnostic accuracy of melanoma is between 75% to 84% by an inexperienced specialist [6]. Also, the diagnosis can be affected by the mood and fatigue of the physician due to the large number of patients who have to treat as soon as possible.

Therefore, an automatic tool is necessary for physicians to have a first evaluation of their patients competitively, quickly, and that allows other physicians to make an easy consensus in the treatment. This motivates developing an automatic lesion classification system that can discriminate benign skin disease from non-benign. The method that can support specialists in their clinical procedure would be beneficial as a pre-diagnosis tool, providing additional accurate lesion information. Also, it will allow the early establishment of the type of lesion, essential to provide adequate treatment to the patient. Furthermore, the tool is of interest to hospitals to avoid problems with wrong diagnoses, physicians in underdeveloped environments where there are no specialists, and researchers who work to improve the diagnosis and treatment of the patients.

It is well known that to develop a prediction model valid for any image, it is necessary to have enough data to carry out a useful optimization. Furthermore, the training datasets need to be homogeneous on the acquisition, and they should not contain anomalous objects within the image. In that case, they may require to do some preprocessing steps and segment the region of interest [7]. The extraction of specific features from the images is another way to improve the classification accuracy [8]. The disadvantage is the requirement of specific knowledge to extract them and the extended time needed to choose the more suitable. Besides, many errors may be

introduced by preprocessing, or essential information might be lost. So these classical techniques are obsolete, and the arrival of deep learning has meant a colossal success, owning the capability to automatically learn the spatial features of each class of skin disease.

Nowadays, deep learning methods have become popular to resolve very complex tasks, including image classification [9] and segmentation [10]. Some of these systems do not use any image preprocessing method [11]. The architecture is behind the success of these networks, mostly based on convolutional layers. These layers have the ability to learn semantic features automatically from large-scale datasets [12]. They filter and extract basic shapes of the images in order to learn the different lesions. For example, in [13] is proposed a universal skin disease classification system based on a pre-trained convolutional neural network (CNN). They give an extraordinary performance in image and video processing [14]. CNN's require immense computational power to compute a large number of operations, which is solved by using graphic processing units (GPU). Thus, they allow the creation of reliable models for image classification and object segmentation, among others. Within the field of medical image analysis, there is evidence that deep learning has shown a significant improvement [15]. Deep models are very extended on health technology, perceiving blood flow from angiographies [16] or detecting vessel borders [17] with high precision. Nevertheless, there is a margin of improvement in skin lesion classification [18]. This work first segment and extract features with deep networks and then make the prediction. Commonly, different types of skin pathologies with similar characteristics are grouped into the same class, so the classification is made on two classes.

The existence of various and visually similar skin disease types makes it challenging to differentiate between them. It is also complicated to distinguish if a skin abnormality is a benign nevus, the most common type of mole, or another skin disease. Image shifting is a technique used to improve the training of the model using many variants of the feature position on an image [19]. Thus, this work proposes a deep convolutional neural network model fused with a shifting mechanism based on a square lattice to classify skin disease images on the two-classes: nevi or others.

Therefore, the main contributions of this work are:

1) The proposal a pre-diagnostic tool used to discriminate between benign and non-benign skin diseases.
2) The use image shifting techniques on a square lattice for skin lesion malignancy assessment using deep learning methods.

The remaining of the paper is structured as follows. A deep convolutional neural network with a regularly spaced shifting based on a square grid that distinguishes between classes with lesion and without it is outlined in Section 2. Next, Section 3 describes the dataset, the methods for the comparisons, as well as the results of the experiments, which are discussed. Finally, Section 4 exposes the conclusions and future trends in the detection of skin lesions.

II. METHODOLOGY

In this section the proposed methodology is detailed. The standard approach to the classification of skin lesion images by deep convolutional neural networks consists on applying a deep neural network F to an input image $\mathbf{X}$ in order to obtain a class scores vector $\mathbf{y}$:

$$\mathbf{y} = F(\mathbf{X}) \tag{1}$$

where $\mathbf{y} \in \mathbb{R}^C$, with C the number of skin lesion classes.

Now, let $\oplus$ be the image shift operator, so that $\mathbf{X} \oplus \mathbf{s}$ is the image $\mathbf{X}$ circularly shifted by the displacement vector $\mathbf{s} \in \mathbb{R}^2$. Please note that the components of the displacement vector can be negative, so that:

$$(\mathbf{X} \oplus \mathbf{s}) \oplus (-\mathbf{s}) = \mathbf{X} \tag{2}$$

for all images $\mathbf{X}$ and all displacement vectors $\mathbf{s}$. If $\mathbf{X}$ is shifted, its class should remain the same. We propose to combine the outputs of the deep classification network F for M shifted versions of the input image $\mathbf{X}$, so that the shift vectors $\mathbf{s}$ are chosen to lie on a square lattice on $\mathbb{Z}^2$, where the side of each square of the lattice is a fixed stride parameter $R \in \mathbb{N}^+$:

$$\mathbf{s} = (Rn, Rm) \tag{3}$$

where $n, m \in \mathbb{Z}$.

The combined class scores output vector is given by:

$$\hat{\mathbf{y}} = \frac{1}{N} \sum_{i=1}^{N} F(\mathbf{X} \oplus \mathbf{s}_i) \tag{4}$$

where $\mathbf{s}_i$ is the i-th displacement vector.

The estimation (4) of the combined class score vector $\hat{\mathbf{y}}$ can be regarded as a finite sample approximation of the average of the network outputs $F(\mathbf{X} \oplus \mathbf{s})$ for shifted inputs $\mathbf{X} \oplus \mathbf{s}$ over the possible displacement vectors $\mathbf{s}$:

$$\hat{\mathbf{y}} \approx \frac{1}{(H-1)(W-1)} \int_S F(\mathbf{X} \oplus \mathbf{s}) \, d\mathbf{s} \tag{5}$$

where the input image has size $H \times W$ pixels, and S is the set of possible displacement vectors:

$$S = [0, H-1] \times [0, W-1] \tag{6}$$

Please note that negative circular shifts are equivalent to highly positive ones:

$$\mathbf{X} \oplus (-\mathbf{a}) = \mathbf{X} \oplus ((H, W) - \mathbf{a}) \tag{7}$$

Therefore, S contains all possible circular shifts. The square lattice S should be constrained to avoid duplicity in the shifted images, and it can be personalized with a specific stride parameter R between its elements. Thus, the algorithm of the proposed model is as follows:

1) Generate a square lattice S over all the $H \times W$ pixels of the data input shape, using a square side equal to 1.
2) Center S in the middle of the image.
3) Multiply the elements of the lattice by the stride R.
4) Restrict the grid to a maximum value M of displacement (in pixels): $\{\mathbf{s} = (s_x, s_y) : |s_x| \leq M \wedge |s_y| \leq M\}$
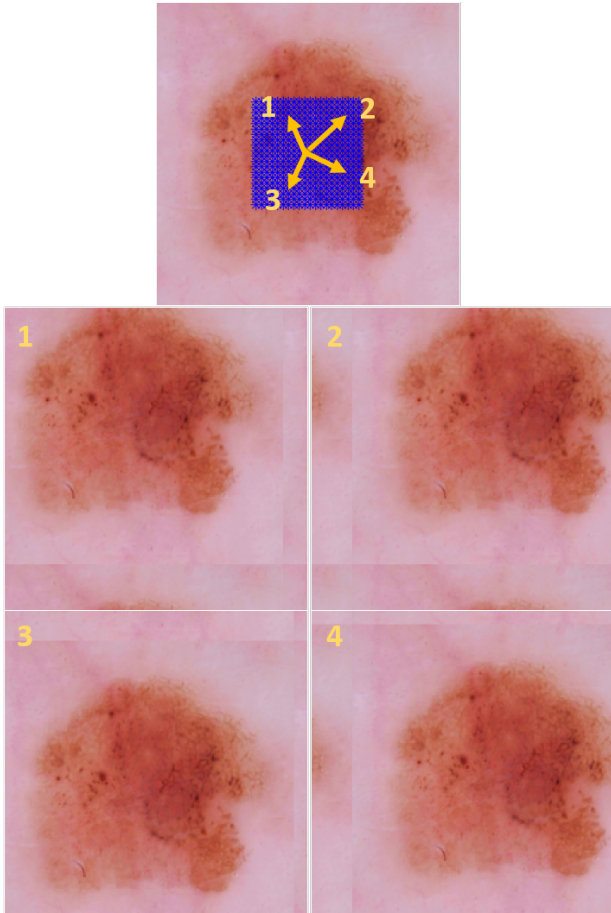
Fig. 1. Scheme of the square lattice model and some example of the displacements.

5) Shift the image by the selected shifts **s** and test it with the neural network.
6) Compute the combined output using Eq. (4).

In Fig. 1 is represented the square lattice over which the displacements vectors are selected. The grid is placed in the center of the image, and as many displacements as elements of the lattice are created. The picture shows four examples of image shifting.

III. EXPERIMENTAL RESULTS

This section describes the methods for the comparisons, the dataset, and the discussed results of the set of experiments.

A. Dataset

For the main performance evaluation of our proposal, we used the well-known HAM10000 dataset [20]. These data contain 10,050 dermoscopic images of benign and malignant moles, the latter divided into six classes: actinic keratosis, basal cell carcinoma, benign keratosis, dermatofibroma, melanoma, and vascular skin. This dataset, whose images were collected from two prestigious institutions in Austria and Australia, was employed for the International Skin Imaging Collaboration (ISIC) 2018 challenge, and it has become a benchmark for testing new classification models.

A disadvantage of this dataset is the unbalanced distribution of the classes. No more than 7,000 images are belonging to the nevi class, while the others contain no more than 1,000, gathering around 3,000 images of abnormal moles. Therefore, during the classification model training, it may be adequate to use data augmentation so that the data can be balanced and the network will learn better the features of the images.

Additionally, a second different dataset was used to study the accuracy of our proposal with completely unseen data. The PAD-UFES-20 dataset [21] is composed of a set of clinical smartphone images and patient data collected by the University of Espírito Santo (Brazil). It contains a total of 2,298 images of six different classes: actinic keratosis, basal cell carcinoma, seborrheic keratosis, squamous cell carcinoma, Bowen's disease, melanoma, and nevus. Therefore, this data is somehow similar to the one used for training the network, although the characteristics of the images are different. The problem here is that the number of elements of the nevi class is quite small, while our network is trained with a larger number of images of this type. Thus, the positive predictions might be imprecise.

In this work, we intend to classify between two classes, nevus, and non-nevus. This is for interest for dermatologists to have early predictions of the status of each patient. Later, they can determine the exact type of benign or malign mole in order to prescript an adequate treatment. Without loss of generality, we will name the benign class as *nv*, and the other class by *mel*.

B. The deep network: MobileNetV2

The proposed neural model is a MobileNetV2 [22], which is composed of 53 layers: an initial full convolutional layer followed by 19 residual bottleneck blocks. These blocks are connected by shortcut connections (similar to ResNet) in order to eliminate the non-linearity and maintain the representation of the data (see Fig. 2). This model is oriented to mobile devices with significant improvements to the previous version (MobileNetV1), and it has been trained on the ImageNet dataset. Comparative results on traditional image datasets (e.g., ImageNet, COCO, VOC,...) demonstrate MobileNetV2 is more effective, faster, and uses fewer parameters than MobileNetV1 with the same accuracy.

The deep network was trained in Matlab R2019b, with the following configuration of the hyper-parameters:

- Batch size = 16.
- Learning rate = 0.0001.
- Validation frequency = 10.
- Max. epochs = 10.

Moreover, the data augmentation transformations used in this work were horizontal and vertical flipping with a probability of 0.5, as well as random rotations between -90° and 90°, with a probability of 0.75.
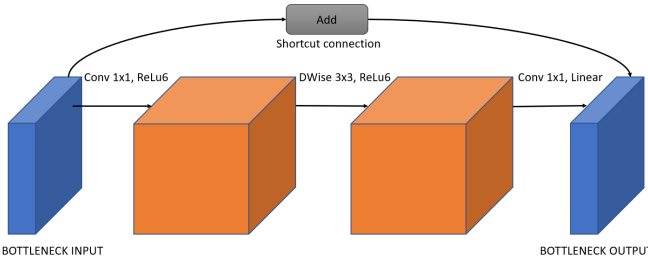
Fig. 2. Structure of the residual bottleneck blocks of the MobileNetV2 neural network.

Finally, we carried out 10-fold cross-validation to have a more precise comparison of the models, using 70% of samples for training, 10% for validation, and 20% for testing.

C. Square lattice parameters

The definition of the dimensions and spacing of the square lattice depends on the type of deep network and the images of the dataset.

On the one hand, the input layer of the convolutional neural network restricts the maximum values of the displacement vectors. MobileNetV2 requires an image resizing up to 224×224 pixels. On the other side, the characteristics of the images are important in order not to generate oversized shiftings that distort too much the shape of the moles. Therefore, we analyzed the data carefully, and we concluded that most of the moles are placed in the center of the image with a margin of $M = 40$ pixels around each side.

In addition, the grid can be generated with different spacing between its elements, i.e., the stride. A tiny stride (1 pixel) would create a very dense lattice, and probably many of the predictions would be identical. However, a very big stride would provoke the opposite effect, and also fewer score vectors would be present to compute the combined mean class output. In our work, we have adopted an intermediate position, taking a stride of $R = 4$ pixels.

D. Evaluation metrics

The typical classification measures were used to analyze the performance of the shifting model. That is, in addition to the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), we computed the following ones:

$$TPR = \frac{TP}{TP + FN} \tag{8}$$

$$TNR = \frac{TN}{TN + FP} \tag{9}$$

$$PPV = \frac{TP}{TP + FP} \tag{10}$$

$$NPV = \frac{TN}{TN + FN} \tag{11}$$

$$Acc = \frac{TP + TN}{TP + FP + FN + TN} \tag{12}$$

$$F1 = 2 \cdot \frac{PPV \cdot TPR}{PPV + TPR} \tag{13}$$

which are the True Positive Rate (sensitivity), True Negative Rate (specificity), Positive Predictive Value (precision), Negative Predictive Value, Accuracy, and F1-score, respectively. The higher the measures, the better, ranged in [0,1].

The TPR and the TNR provide a measure of how well the method is classifying the relevant instances, while the PPV and the NPV specify the proportion of true detections among the retrieved instances. The Accuracy and F1-score provide a general overview of the performance, taking into account the positive and negative samples. The latter gives equal importance to precision and recall.

E. Results

First, the performance obtained by the cross-validation procedure is analyzed in detail in Fig.s 3 and 4. The confusion matrices, as well as TPR and TNR (right column), and PPV and NPV (last column), are presented for each test set split.

Fig. 3 presents the results of the raw MobileNetV2 model. The first thing we realize is that the most accurate prediction is for the nevi class. This is logical due to the greater number of images of this type. Therefore, the number of false positives is also high, being the sixth split the worst with 83.3% of correct negative predictions. On the other hand, this split is the second-best focusing on the number of false negatives. The fourth one is the worst split since it only classifies adequately 71% of the non-nevi moles. In general, one can find some disparities among the split classifications, reaching differences between 70 and 120 images. This fact demonstrates that MobileNetV2 is highly dependent on the training set images, even when it was trained with data augmentation.

The outcomes of the proposed model are depicted in Fig. 4. Compared with the previous charts, it is clear that the number of true positives has increased significantly. Thus, the proportion of false negatives is lower, which is essential for the medical diagnosis. For example, the fourth split, which was bad, has improved by 1%. Other splits have had an improvement of almost 3%. Nevertheless, there are cases where the percentage of true negatives has decreased a bit. Summing up, the shifting model yielded excellent results detecting the *mel* class, and preventing from not detecting a malign mole is more critical than classifying a nevus as positive.

An overall summary of the cross-validation experiments is presented in TABLE I. The average measures among all splits were computed for both methods. While the number of TN and FP is quite similar, the TP and FP have increased and decreased, respectively, using our proposal. This is reflected in the average TPR, with an enhancement of 1.5%. The accuracy and F1-score, the two most important measures, also demonstrate that the shifting model is better since their scores have increased around 1%.

A detailed analysis of four images is also presented in TABLE II. The intention is to study the behavior of the lattice for the four possibles outcomes: true positive, true negative, false
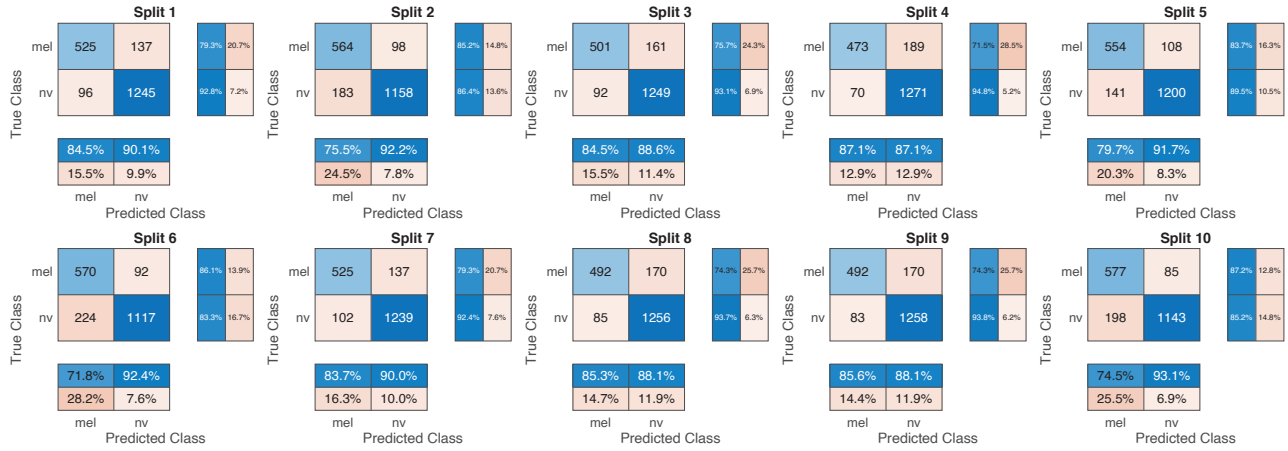
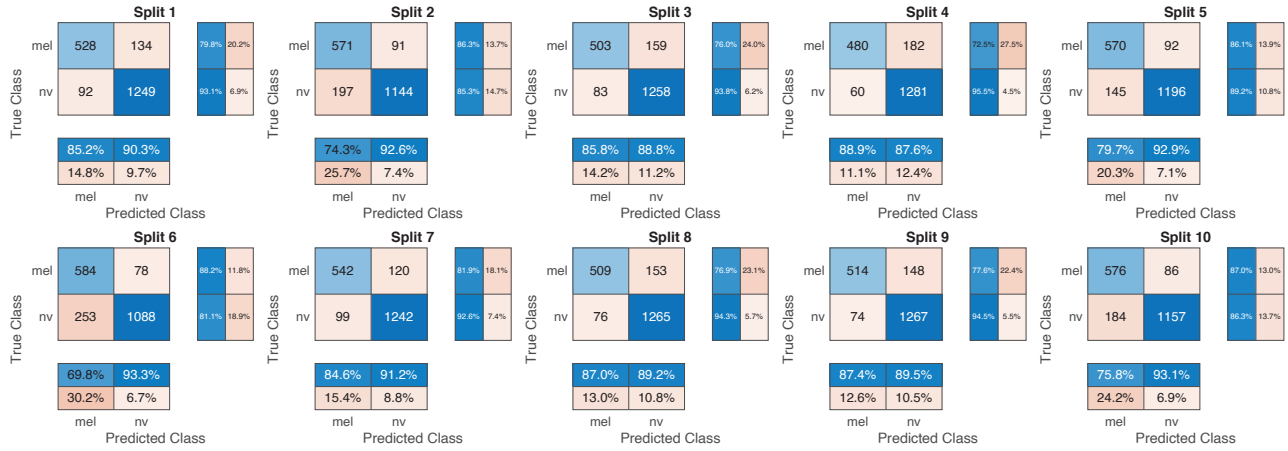Fig. 3. Confusion matrices of the 10-fold cross-validation for the raw MobileNetV2 model.

Fig. 4. Confusion matrices of the 10-fold cross-validation for the proposed shifting model with MobileNetV2.

TABLE I
AVERAGE RESULTS OF THE CROSS-VALIDATION WITH BOTH THE RAW AND
THE SHIFTING MODELS USING HAM10000 DATASET

Measures	Raw Model	Shifting Model
TP	527.3	537.7
TN	1213.6	1214.7
FP	127.4	126.3
FN	134.7	124.3
TPR	79.65%	81.22%
TNR	90.50%	90.58%
PPV	80.54%	80.98%
NPV	90.01%	90.72%
Acc	86.91%	87.49%
F1	80.09%	81.10%

positive, and false negative. For that purpose, we plotted the square grid points colored in red if the mole is classified as *mel* and in green if it is a nevus. The first column depicts an image labeled as a malign mole. Surprisingly, MobileNetV2 classifies it as a n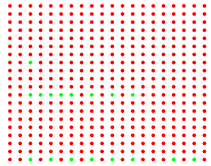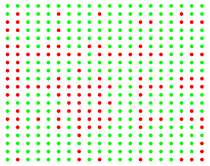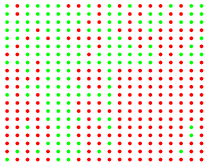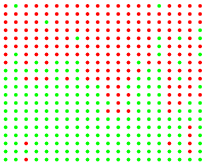evus, but the lattice shows that most displacements outputted the *mel* class. Only 17 shifts predict nevi, being some of them very close to the original image, that is, the center of the lattice (representing null displacement). The position of the mole within the image may be then essential in order to obtain an adequate prediction. Next, a true nevi image is analyzed. Even for the human eye, it is tough to distinguish the mole. Again the center of the grid outputted *mel*, although 280 of the 400 shifts said that the class was *nv*, so our model was able to identify the correct class.
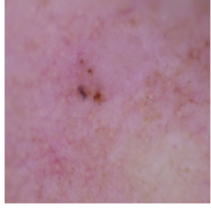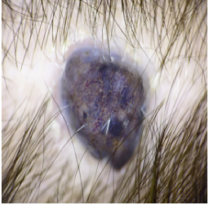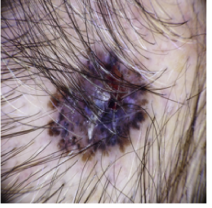
In the last two examples, both methods failed in the prediction, but it is interesting to show the detailed outcomes. The third row corresponds to a false positive case. Some of the skin hairs near the mole probably provoked confusion because the network might have considered it part of the mole and looked irregular. Nevertheless, our method was able to classify it as a nevus for 162 cases. Something similar occurs with the fourth example of a false negative. The balance of the lattice predictions is even closer (189 of 400 yielded the *mel* class). In this case, it is curious to study why the bottom side shifts

Case	TP	TN	FP	FN
Image	ISIC image nº 0024573	ISIC image nº 0024809	ISIC image nº 0024891	ISIC image nº 0025584
Lattice (red: *mel* green: *nv*)				
No. of mel/nv	383/17	120/280	238/162	189/211
GT	mel	nv	nv	mel
Raw	nv	mel	mel	nv
Proposed	mel	nv	mel	nv

classified the mole as benign but the upper side as abnormal. The reason is the number of images used for training the network that contains a similar mole on the upper part of the image. Therefore, the network specialized in detecting this type of images as *mel* because most of the mentioned set were malign moles.

TABLE III
QUANTITATIVE RESULTS WITH BOTH THE RAW AND THE SHIFTING MODELS USING PAD-UFES-20 DATASET

Measures	Raw Model	Shifting Model
TP	1129	1183
TN	224	221
FP	20	23
FN	925	871
TPR	54.97%	57.59%
TNR	91.80%	90.57%
PPV	98.26%	98.09%
NPV	19.50%	20.24%
Acc	58.88%	61.10%
F1	70.50%	72.58%

Finally, our method was also compared with the PAD-UFES-20 dataset, yielding the results of TABLE III. Recall that the number of element of the positive class with quite large compared with the HAM10000 datasets. This fact has produced inaccuracy in most of the measures, like TPR, NPV, or Acc. The best conclusion is that again the number of false negatives is reduced, and almost all the nevi have been correctly detected. The accuracy an F-measure reveal that our shifting method can improve around 2-2.5% the performance of the classification with unseen data.

Fig. 5 depicts two examples of this dataset where our shifting models classified correctly the image but the raw MobileNetV2 model not. The low scores obtained (62% and 71%, respectively) indicates that this images are very challenging
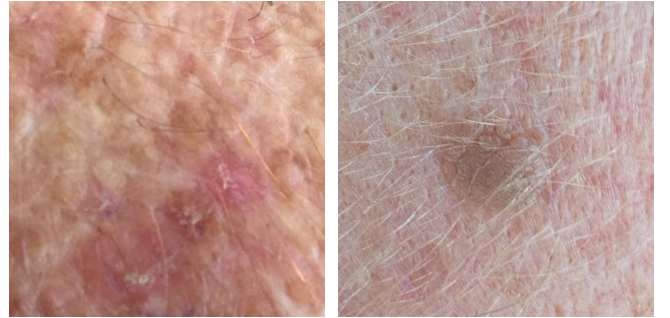

(a) 268 positives detected. Mean score = 62.26%

(b) 355 positives detected. Mean score = 71.54%

Fig. 5. Images *PAT_1070_304_845* and *PAT_1184_678_86* of PAD-UFES-20 dataset where the raw model yielded the incorrect nevi class.

for the image, due to the inappreciable shape of the mole. However, the shifting procedure allowed a better prediction. The number of positives is large enough in both examples, meaning that the probability of accepting the positive is high.

IV. CONCLUSIONS

In this paper, an improved convolutional neural network based on MobileNetV2 is proposed for moles classification. The deep model was improved by an enhanced transfer learning, which consists of a shifting technique. The vectors of displacements were created over a square lattice since regular shifts along all directions move the moles over all the image space. The proposed methodology was tested on the HAM10000 and PAD-UFES-20 datasets providing more accurate predictions, increasing the TPR (True Positive Rate), and decreasing the FNR (False Negative Rate). The outcomes demonstrate that the proposed shifting technique enhances the classification, avoiding a new re-training phase of the model.

Further works include applying the proposed shifting technique to other pre-trained neural models and comparing the results on other datasets. Moreover, the use of different topologies for the lattice generation depending on the type of data might improve the accuracy of the model. The consensus method is also an area of research, which might combine the patient information through another classification model to obtain the final output.

REFERENCES

[1] S. A. Gandhi and J. Kampp, "Skin cancer epidemiology, detection, and management," *The Medical clinics of North America*, vol. 99, no. 6, p. 1323—1335, November 2015. [Online]. Available: https://doi.org/10.1016/j.mcna.2015.06.002

[2] R. S. Stern, "Prevalence of a history of skin cancer in 2007: results of an incidence-based model," *Archives of dermatology*, vol. 146, no. 3, p. 279—282, March 2010. [Online]. Available: http://archderm.ama-assn.org/cgi/content/full/146/3/279

[3] G. P. Guy, C. C. Thomas, T. Thompson, M. Watson, G. M. Massetti, L. C. Richardson, and Centers for Disease Control and Prevention (CDC), "Vital signs: melanoma incidence and mortality trends and projections - united states, 1982-2030," *MMWR. Morbidity and mortality weekly report*, vol. 64, no. 21, p. 591—596, June 2015. [Online]. Available: https://europepmc.org/articles/PMC4584771

[4] N. H. Matthews, W.-Q. Li, A. A. Qureshi, M. A. Weinstock, and E. Cho, "Epidemiology of melanoma," *Exon Publications*, pp. 3–22, 2017.

[5] F. Nachbar, W. Stolz, T. Merkle, A. B. Cognetta, T. Vogt, M. Landthaler, P. Bilek, O. B.-Falco, and G. Plewig, "The abcd rule of dermatoscopy: high prospective value in the diagnosis of doubtful melanocytic skin lesions," *Journal of the American Academy of Dermatology*, vol. 30, no. 4, 1994.

[6] D. A.-R. Ali and T. M. Deserno, "A systematic review of automated melanoma detection in dermatoscopic images and its ground truth data," in *Medical Imaging*, 2012.

[7] M. H. Jafari, E. Nasr-Esfahani, N. Karimi, S. M. R. Soroushmehr, S. Samavi, and K. Najarian, "Extraction of skin lesions from non-dermoscopic images for surgical excision of melanoma," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 6, pp. 1021–1030, Jun 2017.

[8] B. Devassy, S. Yildirim-Yayilgan, and J. Hardeberg, "The impact of replacing complex hand-crafted features with standard features for melanoma classification using both hand-crafted and deep features," *Advances in Intelligent Systems and Computing*, vol. 868, pp. 150–159, 2019.

[9] C. Barata, J. S. Marques, and M. E. Celebi, "Deep attention model for the hierarchical diagnosis of skin lesions," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2019, pp. 2757–2765.

[10] H. M. Ünver and E. Ayan, "Skin lesion segmentation in dermoscopic images with combination of yolo and grabcut algorithm," *Diagnostics*, vol. 9, 2019.

[11] A. H. Shahin, A. Kamal, and M. A. Elattar, "Deep Ensemble Learning for Skin Lesion Classification from Dermoscopic Images," in *2018 9th Cairo International Biomedical Engineering Conference (CIBEC)*. IEEE, 2018, pp. 150–153.

[12] T. Zhou, K. Thung, X. Zhu, and D. Shen, "Effective feature learning and fusion of multimodality data using stage-wise deep neural network for dementia diagnosis," *Human Brain Mapping*, vol. 40, no. 3, pp. 1001–1016, feb 2019.

[13] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, p. 115—118, February 2017. [Online]. Available: https://doi.org/10.1038/nature21056

[14] J. Li, G. Zhou, Y. Qiu, Y. Wang, Y. Zhang, and S. Xie, "Deep graph regularized non-negative matrix factorization for multi-view clustering," *Neurocomputing*, dec 2019.

[15] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60 – 88, 2017.

[16] Z. Gao, X. Wang, S. Sun, D. Wu, J. Bai, Y. Yin, X. Liu, H. Zhang, and V. H. C. de Albuquerque, "Learning physical properties in complex visual scenes: An intelligent machine for perceiving blood flow dynamics from static CT angiography imaging," *Neural Networks*, vol. 123, pp. 82–93, mar 2020.

[17] Z. Gao, J. Chung, M. Abdelrazek, S. Leung, W. K. Hau, Z. Xian, H. Zhang, and S. Li, "Privileged Modality Distillation for Vessel Border Detection in Intracoronary Imaging," *IEEE Transactions on Medical Imaging*, pp. 1–1, 2019.

[18] N. Nida, A. Irtaza, A. Javed, M. Yousaf, and M. Mahmood, "Melanoma lesion detection and segmentation using deep region based convolutional neural network and fuzzy C-means clustering," *International Journal of Medical Informatics*, vol. 124, pp. 37–48, 2019.

[19] K. Thurnhofer-Hemsi, E. López-Rubio, E. Domínguez, R. M. Luque-Baena, and N. Roé-Vellvé, "Deep learning-based super-resolution of 3d magnetic resonance images by regularly spaced shifting," *Neurocomputing*, vol. 398, pp. 314–327, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231219314808

[20] 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, p. 180161, 2018.

[21] A. G. Pacheco and R. A. Krohling, "The impact of patient clinical information on automated skin cancer detection," *Computers in Biology and Medicine*, vol. 116, p. 103545, 2020.

[22] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.