2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Skin lesion classification using features of 3D border lines

Pedro M. M. Pereira[12], Lucas A. Thomaz[23], Luis M. N. Tavora[3], Pedro A. A. Assuncao[23], Rui Fonseca-Pinto[23], Rui Pedro Paiva[1], and Sergio M. M. Faria[23]

***Abstract*— Machine learning algorithms are progressively assuming important roles as computational tools to support clinical diagnosis, namely in the classification of pigmented skin lesions using RGB images. Most current classification methods rely on common 2D image features derived from shape, colour or texture, which does not always guarantee the best results. This work presents a contribution to this field, by exploiting the lesions' border line characteristics using a new dimension – depth, which has not been thoroughly investigated so far. A selected group of features is extracted from the depth information of 3D images, which are then used for classification using a quadratic Support Vector Machine. Despite class imbalance often present in medical image datasets, the proposed algorithm achieves a top geometric mean of 94.87%, comprising 100.00% sensitivity and 90.00% specificity, using only depth information for the detection of Melanomas. Such results show that potential gains can be achieved by extracting information from this often overlooked dimension, which provides more balanced results in terms of sensitivity and specificity than other settings.**

I. INTRODUCTION

Image processing methods have become important tools to assist professionals in their diagnostic decisions. One of their greatest contributions focuses on differentiating melanoma and nevus, respectively a life-threatening dermatological disease and a benign kind of skin lesion [1]. In the field of skin lesion classification, significant results exist using the dermoscopic image (2D) modality, while other image modalities or data dimensions are fairly unexplored. Nonetheless, the existing research on 3D information from stereo imaging indicates that depth features can lead to promising results by improving skin lesion classification performances [2], [3].

The 3D information of images can be obtained through different approaches and acquisition setups. In this work, depth maps are extracted from dense light-fields. Each light-field presents over $10,000,000$ pixels, yet if only a small line of less than $5,000$ pixels is extracted along the lesions' perimeter region, it might contain relevant information to classify the type of lesion. From such border line, features can be calculated using solely the depth information present along the set of connected pixels. Assuming that surface-level information (texture) differs from melanoma to nevus, the respective border lines are expected to have structural differences between them or a different overall geometric behaviour. Thus, discriminative features capable of extracting relevant information about such type of details must be used. Features extracted from electrocardiogram (ECG) signals seem specially suitable for this task, since in ECG classification problems, it is necessary to discriminate patterns from fine variations along a one-dimensional signal. Therefore, a similar type of features may be employed for classification of skin lesions, based on the depth values of the border lines.

Overall, the main contribution of this paper is the exploitation of 3D information from skin lesions, aiming to achieve high discrimination in the classification of melanoma versus nevus and, consequently, showing that this third dimension provides significant information for classification. Different from previous studies, this work investigates new 3D information from the segmentation mask border line to provide evidence that skin surface topology has potential discriminative information.

The remainder of the paper is organised as follows: Section II presents the state-of-the-art and other similar experiments, as well as details about the dataset. Further, Section III describes the proposed approach and relevant details about data pre-processing, feature extraction and classification. Finally, Section IV presents and discusses the experimental results and Section V exhibits the conclusions and future work.

II. BACKGROUND

In the field of skin lesion image processing, several processing steps have been established aiming to enhance melanoma classification [4], namely: image pre-processing, segmentation, and classification [4], [5], [6]. Algorithms employed in melanoma or skin lesion classification range from those using Deep Learning (DL), where the algorithm automatically selects which types of features are exploited for classification, to other classic Machine Learning algorithms that require hand-crafted features. Significant classification performance results have been achieved resorting to DL (e.g. [4]), however these algorithms require rich datasets with large amounts of balanced data. Such constraints might not be achievable and, even when they are, the image modality or information granularity might not be ideal for the problem.

This work was supported by the Fundação para a Ciência e a Tecnologia (FCT), Portugal, under PhD Grant SFRH/BD/128669/2017, Programa Operacional Regional do Centro, project PlenoISLA POCI-01-0145-FEDER-028325 and by FCT/MCTES through national funds and when applicable co-funded by EU funds under the project UIDB/EEA/50008/2020.

[1] P. M. M. Pereira, and R. P. Paiva are with CISUC, Department of Informatics Engineering, University of Coimbra, Pólo II - Pinhal de Marrocos, 3030-290 Coimbra, Portugal (e-mail: ruipedro@dei.uc.pt).

[2] P. M. M. Pereira, L. A. Thomaz, P. A. A. Assuncao, R. Fonseca-Pinto, and S. M. M. Faria are with Instituto de Telecomunicações, Morro do Lena - Alto do Vieiro, Leiria 2411-901, Portugal (e-mail: pedrommpereira@co.it.pt; lucas.thomaz@co.it.pt; amado@co.it.pt; rui.pinto@co.it.pt; sergio@co.it.pt).

[3] L. A. Thomaz, L. M. N. Tavora, P. A. A. Assuncao, R. Fonseca-Pinto, S. M. M. Faria are with ESTG, Polytechnic Institute of Leiria, Morro do Lena - Alto do Vieiro, Leiria 2411-901, Portugal (e-mail: luis.tavora@ipleiria.pt).

Existing datasets, described in the literature, revolve around the same modality, 2D dermoscopic images [4], [6], which are commonly used by field-experts. Despite the good classification results obtained using only dermoscopic information, they are rather limited, as only planar information can be retrieved from such data. To overcome this limitation, in [2], [3], a method using a stereoscopy technology is presented. Although the literature addressing 3D surface studies of melanoma and other skin lesions is almost nonexistent, some previous research indicate that improved results arise when using depth information (3D) [2], [3]. Even if real 3D information is absent, in [7], artificial 3D information was generated for datasets which do not provide it.

In order to fill the void of 3D skin lesion data, a dataset named Skin Lesion Light-fields (SKINL2) was created and made public to enable research over skin lesions 3D surface information [8]. This dataset provides 4-channel images (RGB+Z), where RGB corresponds to the 2D colour image and Z stands for the depth information. At the time of writing, to the authors best knowledge, there are no works published by other authors resorting to this recent dataset.

Similarly to [6], which uses dermoscopic images, in this work only depth information located at the lesions' border region is utilised for classification. The remaining image data of the lesion is discarded. This depth information in the border line is represented by a one-dimensional signal, from which a set of discriminative features is extracted.

In regard to feature extraction, one can consider two main approaches: either deep learning or handcrafted features. In this work, the latter option was used because the SKINL2 dataset is very small in comparison to what is normally necessary for deep learning approaches. Due to the reduced size of the dataset and the large amount of pixel data in each light field, the global depth map of the skin lesion is reduced to fewer data, yet representative of the 3D characteristics. This is done by only considering few border lines of the segmented lesion. Such data reduction is also necessary to avoid over-training as pointed out in [2], [3].

The depth information of a border line can be analysed as a time-series, like other types of known signals, such as ECG for instance. Thus, relevant characteristics can be discriminated by extracting the same type of features. Examples include regression/prediction coefficients as in [9], or localised entropy values as in [10], or some form of wavelet observation as in [11]. In [9], the authors present a feature extraction approach for reliable heart rhythm recognition. After data pre-processing and feature extraction steps, the classifier recognition of 6 types of heart rhythm reaches 99.68% by receiving two sets of features: the transform coefficients of a wavelet transform; and the values of auto regressive modelling applied to the temporal structures of ECG wave forms (model order selection is described by minimisation methods). In [10], the authors detail experiments about the influence, on the performance, of mother wavelets and level of decomposition for wavelet packet decomposition, type of entropy and the number of base learners in a random forest classifier. At the time, the authors stated that experimental results were superior to those of several state-of-the-art competing methods, showing that wavelet packet entropy had promising results for 1D signal classification, such as of ECG. In [11], the authors explain that such signals present complex irregular fluctuations. Hence, to extract information related with such fluctuations, the authors use multi-fractal analysis, specifically wavelet leader based multi-fractal analysis in short-time windows – which had already been proposed in [12] and achieved superior results.

III. CLASSIFICATION METHODOLOGY

In this work, the main goal of classification of skin lesions is to distinguish between melanomas and nevi. To this end the utilised methodology pipeline (depicted in Fig. 1) is comprised of three main steps: data preparation – comprised of extraction, preparation, and augmentation (described in subsection III-A); feature extraction (described in subsection III-B); and classification (described in subsection III-C).

A. Data Preparation

For each 4-channel image (RGB+Z) of the SKINL2 dataset (Fig. 2a), a lesion mask (Fig. 2b) is manually generated (using the colour RGB channels, Fig. 2c) so that the lesion perimeter pixels could be identified and their depth channel values (hereinafter – the "Z" channel, Fig. 2d) sequentially extracted from a random starting position (hereinafter – the "border line" vector, Fig. 2e).The Z-pixel values might not be all within acceptable ranges (mainly due to errors caused by light reflection), therefore all border line values higher than 10mm (chosen as arbitrary threshold) were replaced by previously valid values in the sequence. Afterwards, border line Z-values are normalised to a range of $[-0.5, 0.5]$. Note that RGB is only used to produce a segmentation mask (and then is then discarded), as all data used to train the model is from the Z channel.

Additionally, three supplementary border lines may be extracted depending on the experiment augmentation settings. If enabled, this step iteratively shrinks the lesion mask by 20 pixels until it produces 3 inner border lines, which are inside the lesion region. The belief is that melanoma and nevus surfaces are different, hence more information would allow to compensate any model overfitting and also reinforce a better comprehension about the problem dimension to the classifier.

Since not all border lines have the same length (as skin lesions come in all shapes and sizes), it was necessary to uniformise their size before the feature extraction process. Thus, four transformations were considered: **T1**) pad smaller lines with zeros; **T2**) repeat (by rotation) smaller lines; **T3**) linearly stretch smaller lines; or **T4**) cubicly interpolate smaller lines.

In addition to the already mentioned options, further data augmentation techniques were added to evaluate the generalisation of the model: by flipping/inverting all border lines (doubling the dataset size); and repeating data to balance-out the smaller class samples.

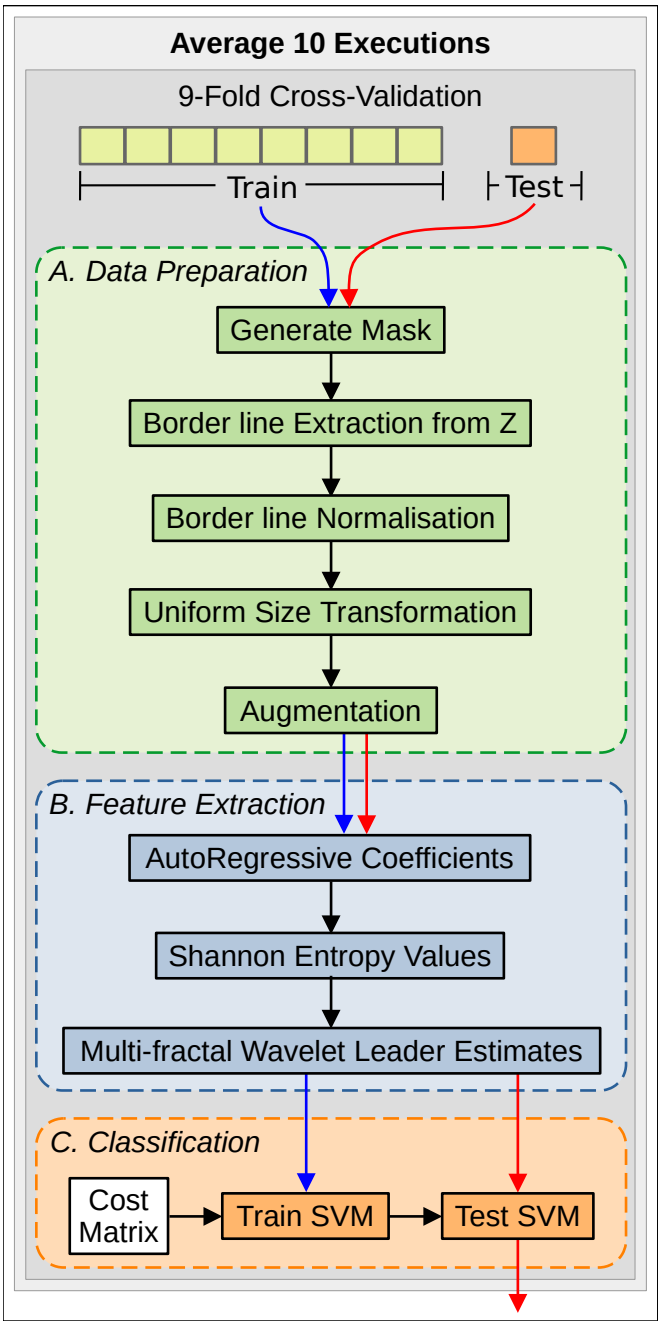
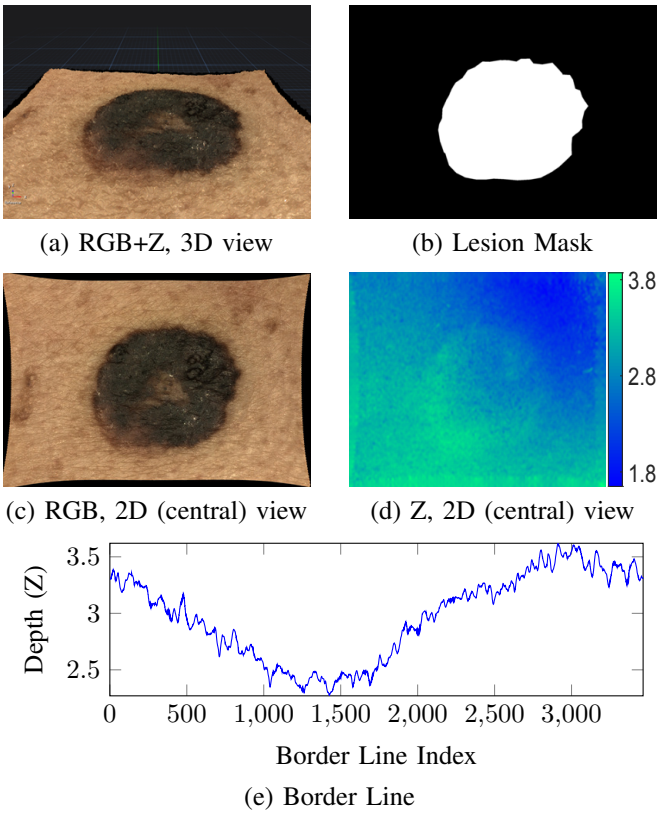


Fig. 1. Proposed methodology pipeline comprises three main blocks, which are executed inside a 9-Fold Cross-Validation scheme that is executed 10× for average calculations. Given train and test datasets, the first block prepares the data by: generating a lesion mask so that border lines may be extracted from the Z dimension; which are then normalised and transformed to a uniform size; and augmented in the end. Given the prepared data, the second block performs feature extraction by: computing named AutoRegressive Coefficients, Shannon Entropy Values, and Multi-fractal Wavelet Leader Estimates. Given the extracted features, the third block trains an SVM using the features generated from the fold-dependent train data and a defined cost matrix, and later tests the SVM model using the features generated from the fold-dependent test data. Blue arrows indicate the pipeline training sequence. Red arrows indicate the pipeline testing sequence. Black arrows indicate previous dependencies or common progressions through the pipeline.

B. Features Extraction

After having a set of equal-sized border lines, feature extraction takes place. Most of the features provide global



(a) RGB+Z, 3D view (b) Lesion Mask (c) RGB, 2D (central) view (d) Z, 2D (central) view (e) Border Line

Fig. 2. SKINL2 Dataset melanoma data sample *0059* (a) 3D visualisation, (b) lesion segmentation mask, (c) RGB central view, (d) depth channel "Z" central view with blue-to-light-blue colour bar, and (e) extracted border line Z-values in millimetres.

representations of the border line (which makes fine details along smaller sections become expressionless) or require continuous samples of the data to generate meaningful coefficients. Due to these cases, border line vectors were either empirically split into chunks of eight equal-sized windows or observed with windows of size 8. The selected list of features (186 in total) corresponds to a set including the more relevant features regarding their performance in classifying waveform signals with similarity to border lines. Their description is as follows:

- AutoRegressive model coefficients of order 4 [9] over eight equal-sized windows (producing 4×8 features). For each window, the model coefficients are estimated using the Burg method [13], which estimates the reflection coefficients and enables the reflection coefficients to estimate the AR parameters, recursively. Based on [9], where model order selection methods are used to determine the model order that provides the best fit in a similar classification problem, order 4 was selected for our model.
- Shannon entropy values [14] for the maximal overlap discrete wavelet packet transform at level 4 [10] applied to the signal divided into eight windows, resulting in $2^4 \times 8$ features.
- Multi-fractal wavelet leader estimates [12] (of the sec-

ond cumulant of the scaling exponents) and the range of its exponents, which quantify the local regularity, or singularity spectrum as in [11], creating 2×8 and 10 features, respectively. Wavelet variance measures the variability in a signal by scale (over octave-band frequency intervals), being extracted for each signal over the entire data length, as in [15], based on [14]. In order to have an unbiased estimate of the wavelet variance, it is necessary to use only levels with at least one wavelet coefficient unaffected by boundary conditions. Our signal (border line) length and the Daubechies 2 (db2) wavelet [16] result in the usage of 10 levels. Similarly to [11], the width of the singularity spectrum obtained from multi-fractal 1D wavelet leader estimates was selected as a measure of the multi-fractal nature of the border line signal. Note that the second order cumulants were selected because they broadly represent the departure of the scaling exponents from linearity – scaling exponents are scale-based exponents describing power-law behaviour in the signal at different resolutions.

C. Classification

The classification of the lesions is performed resorting to a Support Vector Machine (SVM) model with a polynomial kernel of second order and a box constraint of 1. Given that this is a binary classification problem, the SVM solver is the Iterative Single Data Algorithm, (which is optimised through a series of one-point minimisations and neither respects the linear constraint nor explicitly includes the bias term) [17].

In addition, due to the unbalanced nature of the dataset, adjustments to the classifier's cost matrix were also tested. In this matrix, each element consists of the cost of guessing that a sample belongs to class X (lines) when it belongs to class Y (columns), leaving all elements of its main diagonal equal to zero. Therefore, since the melanoma class is three times smaller than the nevus class, the experimental cost matrix adjustments were made to accommodate this discrepancy in unit steps: making a mistake in the melanoma classification have an importance equal to a that of a nevus ($[0\,1; 1\,0]$), or have twice ($[0\,1; 2\,0]$), or thrice ($[0\,1; 3\,0]$) said importance.

IV. EXPERIMENTAL ASSESSMENT

The proposed classification methodology was applied to the publicly available SKINL2 dataset [8]. The skin lesion light-fields were captured at a hospital facility (Centro Hospitalar de Leiria, Portugal), with a Raytrix R42 camera, from patients previously screened by a medical doctor during dermatology clinical appointments. Each image has 3858×2682 pixels of RGB channels, as well as the relative depth of each pixel. The skin lesions were classified and organised based on clinician diagnosis according to ICD10 (International Classification of Diseases) and on histopathological analysis. Procedures related with image acquisition were evaluated and approved by a health ethics committee. The procedure and purpose of the study was explained to all volunteers, who also signed an informed consent form. Particularly in this work, the second version of this dataset was used [18], due to its increase in lens magnification of $\approx 30\%$ (which means more detail) in comparison to its first version. This dataset currently comprises 9 melanomas and 27 nevi images, which undergo the pre-processing, feature extraction, and classification processes described in Section III. All experiments were performed using Leave-One-Out Cross-Validation (which effectively results in a 9-fold Cross-Validation because of the dataset size) and were executed 10 times to mitigate any biased or stochastic decision in the model on an Alienware M17xR3 Laptop (Intel® Core™ i7 2nd Generation 2670QM octa-core 64-bit CPU at 2.20GHz, with 8 GB of Memory RAM).

TABLE I
SUMMARY RESULTS FOR EACH EXPERIMENT.

#Lines	CostMatrix	Augmentation	Geometric-Mean (GM) T1	T2	T3	T4
1	[0 1 ; 1 0]	—	93.33	43.30	90.45	72.47
		Flip	87.66	46.56	78.06	84.98
		BalanceOut	94.87	50.55	78.06	84.98
		Flip+BalanceOut	87.66	50.55	81.91	68.77
	[0 1 ; 2 0]	—	94.87	63.25	72.47	81.91
		Flip	81.83	34.96	90.45	78.47
		BalanceOut	67.42	34.96	61.83	73.19
		Flip+BalanceOut	70.12	63.25	68.77	68.79
	[0 1 ; 3 0]	—	93.33	50.55	89.11	78.47
		Flip	81.83	55.65	69.56	81.91
		BalanceOut	64.98	64.09	62.50	78.47
		Flip+BalanceOut	67.42	60.18	64.09	68.77
1+3	[0 1 ; 1 0]	—	89.43	37.30	90.45	80.77
		Flip	84.60	58.79	82.73	88.87
		BalanceOut	68.71	0.00	70.57	69.86
		Flip+BalanceOut	70.09	51.70	68.77	77.56
	[0 1 ; 2 0]	—	86.36	37.30	73.46	81.83
		Flip	83.12	32.61	79.97	79.71
		BalanceOut	66.31	41.48	67.56	81.63
		Flip+BalanceOut	60.18	52.93	67.51	76.65
	[0 1 ; 3 0]	—	76.20	41.17	71.44	78.47
		Flip	66.69	54.72	68.77	87.66
		BalanceOut	66.69	0.00	67.52	66.15
		Flip+BalanceOut	62.18	40.72	77.98	72.16

The results achieved in these assessments are evaluated in terms of percentage of classification Accuracy (*Acc.*), Specificity (*SPE*), and Sensitivity (*SEN*). A more detailed description of these metrics can be found in [19]. In addition, because this is an unbalanced problem, the Geometric-Mean (*GM*), Eq. (1), is also used [20]. It corresponds to the geometric mean of sensitivity and specificity. To facilitate the reader's understanding, each experiment is first summarised into the *GM* metric (Table I) and then only the best results are detailed with the other three previous mentioned metrics (Table II). Metric results above 75% are highlighted with a colour gradient white-to-grey.

$$GM = \sqrt{SEN \times SPE} \tag{1}$$

Additionally, as mentioned in Section III-A, data augmentation variations are also shown in the results table

TABLE II

DETAILED METRIC RESULTS FOR THE BEST GROUPS.

#Lines	CostMatrix	Augmentation	T1 (Padding Zero) Acc.	SEN	SPE
1	[0 1 ; 1 0]	—	88.89	100.00	87.10
		Flip	88.89	85.71	89.66
		BalanceOut	91.67	100.00	90.00
		Flip+BalanceOut	88.89	85.71	89.66
	[0 1 ; 2 0]	—	91.67	100.00	90.00
		Flip	86.11	75.00	89.29
		BalanceOut	75.00	50.00	90.91
		Flip+BalanceOut	77.78	53.85	91.30

"Augmentation" column. Four variations are tested: no augmentation ("—-"), flipping the data ("Flip"), balancing-out ("BalanceOut"), or flipping and balancing-out data ("Flip+BalanceOut") during training. The experiments also include tests using a variable number of lines (column "#Lines"), as described in Section III-A (i.e. whether to use or not the 3 supplementary border lines). As mentioned in Section III-C, because of the unbalanced dataset distribution, balancing of the classifier's training was also performed resorting to its cost matrix, as expressed in the tables by column "CostMatrix". Furthermore, each combination of the previously mentioned variations is used four times, since there are four possible ways of adjusting the border lines length (that is T1/T2/T3/T4, as expressed in Section III-A).

The experimental setup takes, on average, 17 minutes to process the dataset, extract features, and classify using an SVM with 9-Fold Cross-Validation over 10 executions. There are $2 \times 3 \times 4 \times 4$ experiments of 9-Fold each in Table I (but only $1 \times 2 \times 4 \times 1$ in Table II), which are run 10 times for variance verification. The highest standard deviation is 4.3, and most executions have 0.0 (zero). Due to space constraints, standard deviations are not shown in tables.

From Table I, the highest and lowest results are achieved with experiments involving line size transformations T1 and T2, respectively, with an average difference of 30.30% *GM* between both. As expected, looping the extraction of Z values (along the mask perimeter until the recorded vector reaches the same length of the largest mask perimeter – T2) did not provide proper results because the repetition of information/details introduced large variations on the extracted features (which take into account the signal's structure). Experiment T3 has a similar behaviour to T1 but the linear interpolation seems to degrade the results by 2.99% *GM*, on average. Experiment T4, on the other hand (with cubic interpolation), only degrades 0.13%, on average. Therefore the scenario T1, which pads the remaining space with zeros, provides the overall best results.

Balancing the data by repeating samples of the smallest class (*BalanceOut* augmentation) provides a similar result to balancing the SVM cost matrix for the dataset unbalanced data ($[0\ 1; 2\ 0]$), which is also evident in Table II (91.67/100.00/90.00, rows 3 and 5).

Still in Table I, the presence of "normal" and "flipped" vectors (*Flip* augmentation) almost always provides worst results. From the observed experiments, the presence of flipped vectors greatly induced homogeneity across the features space, which makes it harder, or impossible, for the classifier to find a proper separation in the data. Therefore, using the two augmentation settings (*Flip+BalanceOut*) had similar outcomes. Depending on the configuration, one of the augmentation modes always induces worse results. Although this could be a misinterpreted effect due to the size of the dataset.

Regarding the number of extracted lines, the use of extra border lines generally degraded the results, except in T4. In 28 (out of 48) experiments, using more than one border line, provided worse results in comparison to using a single border line. This can be explained by the fact that, having used a small dataset, the classifier might not find the best hyper-plane separation and the added samples help nudge it, albeit the added samples are of a different origin thus can also decrease some of the data separability. Note, however, that in T4, in 9 out of 12 experiments the results improved in comparison to using only one border line. Thus, it is more evident that, despite of the data interpolation, the results are improved by using the additional border lines.

As can be seen in Table II, *SPE* values range from 87.10% to 91.30%, which are the best results. *SPE* indicates the ability to correctly reject healthy patients without a condition. This (high *SPE*) makes it useful for ruling in disease. However, given that the SKINL2 dataset is very unbalanced, it is not easy to obtain balanced *SPE* and *SEN* results at the same time. Even so, manually balancing the data through augmentation or repetition seems to help in certain cases. The accuracy performance is 91.67% when forcing sample balancing either by replicating the smaller class or adding more weight to said class misclassifications. It is must be noticed that the proposed method exploits only the lesion depth information for feature extraction and classification (for the melanoma vs nevus problem). However, while showing results similar to other methods reported in the literature, it is not directly comparable because the other methods are based on RGB colour images and, consequently, in other datasets. This approach relies on a less invasive technology (does not require physical contact with the patient) that uses light-field cameras instead of dermatoscopy devices.

Other top results, as the $89 - 90\%$ *GM* in experiment T3, could also be worth noticing. Because of the behaviour of the *GM* metric, factually, any result above 80% can only have *SEN* and *SPE* values in the $[64, 100]$ range. While results above 90% need *SEN*/*SPE* values in the $[81, 100]$ range. This means that other not-detailed results could compete with state-of-the-art skin lesion classification algorithms (if not for the data modality constraint, which makes direct comparison impossible). On a practical perspective, a passive mechanism even with results of 70% *GM* could be appealing, since false positive detection of nevus as melanoma is not completely undesirable. In T1, cost matrix $[0\ 1; 2\ 0]$, the highest *SPE* of 91.30 is achieved (70.12 *GM*), which means that a system would identify 7 out of 9 melanomas, albeit a

misclassification rate of 6 out of 27 nevi.

As previously mentioned, this work focuses on showing that skin surface topology has potential discriminative information for melanoma classification and, as such, comparisons with other literature works that resort to RGB only are neither possible or relevant. Additionally, at the time of writing, to the authors best knowledge, there are no works published by other authors resorting to the utilised dataset that provides RGB+Z images.

V. CONCLUSIONS AND FUTURE WORK

Automated melanoma detection is crucial to help dermatologists to improve their diagnostic accuracy. Among all skin lesion discriminations, classification of melanoma versus nevus is considered the most difficult, therefore a computer expert system is of utmost importance. As an alternative of recent works, where skin lesion classification is based on dermoscopic images (2D), this paper investigated other type of image information, which have been fairly unexplored – e.g. surface (3D) information; to find out whether it could potentially provide better discrimination of melanoma versus nevus. Taking advantage of the recently introduced technology of light-field cameras, this work provides a new insight on this domain, being the first one to demonstrate that classification of skin lesions, based on multi-dimensional imaging, is possible to achieve with quite good accuracy. To this end, the 3D border lines of the lesion were used to perform a classification with high discrimination. Due to its characteristics, the extracted signal, obtained from the border lines, can be classified using 1D features.

The achieved experimental results present a discrimination of melanomas against nevi of 94.87% *GM* (100.00% *SEN* and 90.00% *SPE*). Since these results are comparable with others available in the literature, they provide evidence that skin lesion classification (of melanoma and nevus) is possible using non-invasive techniques and avoiding the additional artifacts that the use of a dermatoscope (and gel) induces in algorithm pipelines.

Overall this work provides insight for further research in the field of skin lesion image detection, segmentation and classification to either improve existing methods/models that are lacking in performance or refine the existing top performers. It is also demonstrated that the extended 3D information enabled by the light-field cameras is useful, beyond conventional texture (2D), to improve lesion discrimination algorithms.

REFERENCES

[1] N. Linsangan, J. Adtoon, and J. Torres, "Geometric analysis of skin lesion for skin cancer using image processing," in *IEEE Int. Conf. on Humanoid, Nanotechnol., Information Technol., Commun. and Control, Environment and Management*, Baguio, Philippines, Nov. 2018, IEEE, pp. 1–5.

[2] S. McDonagh, R. Fisher, and J. Rees, "Using 3D information for classification of non-melanoma skin lesions," in *Proc. Med. Image Understanding and Anal.*, Dundee, United Kingdom, July 2008, BMVA Press.

[3] L. Smith, M. Smith, A. Farooq, J. Sun, Y. Ding, and R. Warr, "Machine vision 3D skin texture analysis for detection of melanoma," *Sensor Rev.*, Mar. 2011.

[4] S. Pathan, K. Prabhu, and P. Siddalingaswamy, "Techniques and algorithms for computer aided diagnosis of pigmented skin lesions—a review," *Biomed. Signal Process. and Control*, vol. 39, pp. 237–262, Jan. 2018.

[5] P. Pereira, R. Fonseca-Pinto, R. Paiva, P. Assuncao, L. Tavora, L. Thomaz, and S. Faria, "Dermoscopic skin lesion image segmentation based on local binary pattern clustering: Comparative study," *Biomed. Signal Process. and Control*, vol. 59, pp. 101924, May 2020.

[6] P. Pereira, R. Fonseca-Pinto, R. Paiva, P. Assuncao, L. Tavora, L. Thomaz, and S. Faria, "Skin lesion classification enhancement using border-line features–the melanoma vs nevus problem," *Biomed. Signal Process. and Control*, vol. 57, pp. 101765, Mar. 2020.

[7] T. Satheesha, D. Satyanarayana, M. Giri Prasad, and K. Dhruve, "Melanoma is skin deep: a 3D reconstruction technique for computerized dermoscopic skin lesion classification," *IEEE J. Transl. Eng. Health Med.*, vol. 5, pp. 1–17, Jan. 2017.

[8] S. Faria, J. Filipe, P. Pereira, L. Tavora, P. Assuncao, M. Santos, R. Fonseca-Pinto, F. Santiago, V. Dominguez, and M. Henrique, "Light field image dataset of skin lesions," in *Annu. Int. Conf. of the IEEE Eng. in Medicine and Biol. Soc.*, Berlin, Germany, July 2019, IEEE, pp. 3905–3908.

[9] Q. Zhao and L. Zhang, "ECG feature extraction and classification using wavelet transform and support vector machines," in *Int. Conf. on Neural Networks and Brain*, Beijing, China, Oct. 2005, IEEE, vol. 2, pp. 1089–1092.

[10] Taiyong L. and Min Z., "ECG classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, pp. 285, Aug. 2016.

[11] R. Leonarduzzi, G. Schlotthauer, and M. Torres, "Wavelet leader based multifractal analysis of heart rate variability during myocardial ischaemia," in *Annu. Int. Conf. of the IEEE Eng. in Medicine and Biol.*, Buenos Aires, Argentina, Aug. 2010, IEEE, pp. 110–113.

[12] S. Jaffard, B. Lashermes, and P. Abry, "Wavelet leaders in multifractal analysis," in *Wavelet Anal. and Appl.*, pp. 201–246. Springer, Dec. 2006.

[13] S. Kay, *Modern spectral estimation: theory and application*, Pearson Education India, 1988.

[14] A. Walden and A. Cristan, "The phase–corrected undecimated discrete wavelet packet transform and its application to interpreting the timing of events," *Proc. of the Royal Soc. of London. Series A: Math., Physical and Eng. Sciences*, vol. 454, no. 1976, pp. 2243–2266, Aug. 1998.

[15] E. Ann Maharaj and A. Alonso, "Discriminant analysis of multivariate time series: Application to diagnosis based on ECG signals," *Comput. Statist. & Data Anal.*, vol. 70, pp. 67–87, Feb. 2014.

[16] A. Cohen, "Ten Lectures on Wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 61, I. Daubechies, SIAM, 1992, xix+ 357 pp.," *J. of Approximation Theory*, vol. 78, no. 3, pp. 460–461, Sept. 1994.

[17] V. Kecman, T. Huang, and M. Vogt, "Iterative single data algorithm for training kernel machines from huge data sets: Theory and performance," in *Support vector machines: Theory and Appl.*, pp. 255–274. Springer, May 2005.

[18] S. Faria, M. Santos, P. Assuncao, L. Tavora, L. Thomaz, P. Pereira, R. Fonseca-Pinto, F. Santiago, V. Dominguez, and M. Henrique, "Dermatological imaging using a focused plenoptic camera: the SKINL2 light field dataset," in *Conf. on Telecommun.*, Lisbon, Portugal, June 2019.

[19] A. Baratloo, M. Hosseini, A. Negida, and G. Ashal, "Part 1: simple definition and calculation of accuracy, sensitivity and specificity," *Emergency*, Feb. 2015.

[20] A. Tharwat, "Classification assessment methods," *Appl. Comput. and Inform.*, Aug. 2020.