Automatic Detection of Dysplastic Nevi: a Multiple Instance Learning Solution.

(Discussion Paper)

Eugenio Vocaturo[1,2][0000−0001−7457−7118] and
Ester Zumpano[1][0000−0003−1129−3737]

[1] DIMES - Department of Informatics, Modelling, Electronic and System Engineering University of Calabria, Rende, Italy
{*e.vocaturo, e.zumpano*}*@dimes.unical.it*
[2] CNR-NANOTEC National Research Council, Rende, Italy
e*ugenio.vocaturo@cnr.it*

Abstract. Malignant melanoma is responsible for the highest number of deaths related to skin lesions. The similarities of melanoma with other skin lesions, such as dysplastic nevi, constitute a pitfall for computerized detection. The proposed algorithms and methods have had as main focus the dichotomous distinction of melanoma from benign lesions and they rarely focused on the case of melanoma against dysplastic nevi. Currently, there is a debate about *dysplastic nevi syndrome*, or rather about the number of moles present on the human body as potential melanoma risk factors. In this document, we consider the challenging task of applying a multi-instance learning (MIL) algorithm for discriminating melanoma from dysplastic nevi and outline an even more complex challenge related to the classification of dysplastic nevi from common nevi. Since the results appear promising, we conclude that a MIL technique could be at the basis of tools useful for skin lesion detection.

Keywords: Image Classification · Multiple Instance Learning · Dysplastic nevi Detection.

1 Introduction

According to the latest report of the World Health Organization (WHO), in 2018 *melanoma* has caused over 60.000 deaths and over 280.000 new cases have been diagnosed [1]. Despite the ever increasing diffusion and its aggressiveness, if melanoma is identified by an early diagnosis it is a type of curable cancer. Some clinical protocols such as the ABCDE rule [2] have been established to facilitate the task of specialists in identifying the lesion from its initial phase. These clinical

protocols take into consideration features lesion such as asymmetry, irregular edges, colors, diameters greater than 6 mm and evolving stages.

The importance of an early detection of melanoma have led research communities to develop automatic frameworks called Computer Aided Diagnosis (CAD) systems, for the analysis of skin lesions. CAD include steps such as image acquisition, pre-processing, segmentation, features extraction and selection and finally classification of lesions.

This work focuses on the classification task of discriminating melanoma from dysplastic nevi. Some studies have shown that specific ethnic groups present a great number of common and dysplastic nevi on their's bodies surface [3].

Individuals with *dysplastic nevi syndrome* or dysplastic nevi with family history of melanoma face a greater risk of developing melanoma [4]. These premises justify the perception that automatic diagnosis of skin lesions must consider, besides the distinction between melanoma and common nevi, also the one between melanoma and dysplastic nevi, that is more difficult due to the similarities of the two type of lesions [5].

Nowadays, few studies are available on this specific topic, therefore this paper is a contribution to this challenging task. Our basic idea takes its cue from the work [6] in which the authors have shown how the use of simple color features, on dermatoscopic images, lets to obtain satisfactory classification performances. In the present paper, we apply a recent MIL approach [7, 8] on dermatoscopic images using only color features to verify its effectiveness in the classification steps concerning Melanomas vs Dyspastic Nevi and Dysplastic Nevi vs Common ones. The paper is organized as follows. In the next section we focus on the role that the presence of dysplastic nevi and common nevi may imply in terms of risk of melanoma onset. In Section 3 we recall the peculiarities of a MIL approach, focusing on the Lagrangian relaxation type algorithm proposed in [9]. In Section 4 some numerical results on dermoscopic images are presented and finally brief conclusions are given in Section 5.

2 Dysplastic nevi

The term "dysplastic nevus" (DN) indicates a nevus with different histological and genetic characteristics compared to common nevus. More specifically, the term derives from the Greek "dis-" (bad or malfunction) and "-plasia" (development of growth or change) [10] and indicates a potentially dangerous lesion for his guest. Several studies have attempted to correlate the degree of dysplasia of melanocytes with the risk of melanoma [11, 12].

Syndrome of dysplastic nevus (DNS) refers to subjects who have a high number of benign moles and dysplastic nevi. Dysplastic nevi are more likely to undergo malignant transformation when they occur among members of melanoma families.

In [13], the authors indicate a cumulative lifetime risk of almost 100% in individuals who have dysplastic nevi and are related to melanoma; about 30% of melanomas occur within atypical moles. In 40-50% of cases, there is a genetic

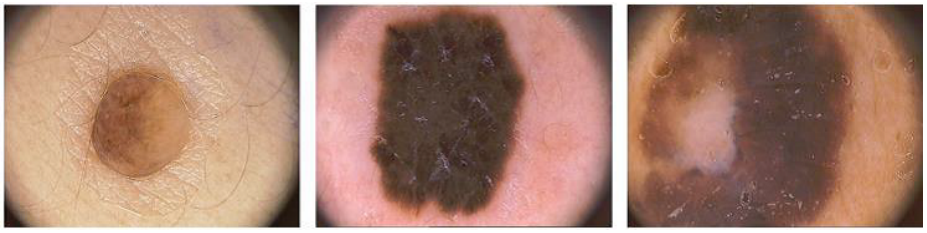

Fig. 1. Dermoscopic image of common nevus (a), dysplastic nevus (b) and melanoma (c)

predisposition for the formation of melanoma. The onset of this skin cancer has been associated with germline mutations in the CDKN2A gene, which encodes p16 (a regulator of cell division).

In [14], some studies based on histological analysis have correlated the presence of dysplastic nevi with melanoma. The caution that should be observed in correspondence with a diagnosis of a severe DNS should not be overlooked, as it could represent a miss-diagnosed in situ melanoma [15].

Basically, there are two objective criteria that have been shown to be related to the risk of melanoma:

- In [16], people with a number of nevi greater than 100 had a 7 times greater risk of melanoma than those with a count of less than 15.
- The presence of large nevi increases the relative risk of melanoma. If these nevi have a diameter less than 2.4 mm they have a relative risk of 1, while the relative risk progressively increases up to 5 if the lesion has a diameter greater than 4.4 mm [17].

Simultaneously with the definition of the exact cause-effect correlations, various solutions have been proposed over time for the automatic identification of skin lesions.

In [18] the authors, depicts a summary of many recent proposed methods by reporting the results in terms of sensitivity, specificity and dataset size. In the same work methods are categorized based on their classification scope: melanoma from benign ($M\ vs\ B$), melanoma from benign and dysplastic ($M\ vs\ (B + D)$) and melanoma versus only dysplastic nevi ($M\ vs\ D$).

The comparison among different approaches is far from being easy as each proposal has been applied on different datasets and adopts different features sets. As for the feature, an additional difference arise between global and local features. Global features are extracted taking the lesion as a whole, while local features are extracted from portions of the image. On a qualitative level, AdaBoost (AdB), artificial neural network (ANN), Support Vector Machines (SVM) appear to be the most effective methods [19]. To the best of our knowledge, nobody takes into consideration the classification task of dysplastic nevi against common nevi [20].

3 Multiple Instance learning

Machine Learning has become very important in medical image analysis. In fact, machine learning methods are currently used in the segmentation steps, in which each pixel of an image belongs to a particular tissue and in CAD systems to assign a category label to a whole image. Even, the availability of partially labeled data can be appropriately exploited using machine learning approaches [21], [22].

In particular, Multiple Instance Learning scenario is useful when disposing of local annotated labels is expensive, while global labels for whole images, such as the outcome of a diagnosis, are more readily available.

A MIL problem consists in the classification task of a set of items called *bags* and of the objects inside them called *instances*. The substantial difference compared to supervised classification consists in the fact that, in the learning phase, only the labels of the bags are known, and not those of the instances.

A natural application of the Multiple Instance Learning is in diagnostics by means of medical images, where we want to discriminate between non-healthy and healthy patients on the basis of their medical scan (bag). According to standard Mil assumption:

- if in the medical scan at least one sub-region (instance) is abnormal, then a patient is *positive*;
- if in the medical scan all the subregions (instances) are normal, then the patient is *negative*.

The classification of complex objects cannot be represented by single features vector, i.e. by single instance: this is another reason why we propose to use MIL techniques for melanoma detection.

With a MIL approach it is therefore possible to obtain global information from local one. For further details and general considerations on the MIL paradigm, we refer the reader to surveys [23]. In [24], a detailed review is given concerning Multiple Instance Learning applied for medical images and video analysis. The MIL approach, as far as we know, is still very rarely used for melanoma detection, and has never been used for the detection of dysplastic nevi.

4 Numerical experimentation

The MIL algorithm used for the classification task in this paper has been proposed in [9], and has been tested for the classification of both dermoscopic images [25, 26] taken from the PH2 database [27], and photographs datasets publicly available from two online databases https://www.dermquest.com and http://www.dermins.net [28]. The entire PH2 database contains 200 images of melanocytic lesions: 80 common nevi, 80 atypical nevi and 40 melanomas. All images were obtained using 8-bit RGB colors with a resolution of 768×560 pixels.

For the classification experiments we considered the 40 images of melanomas (1.a), the 80 of dysplastic nevi (1.b) and the 80 of common nevi (1.c), without taking into account the indications resulting from the manual analysis carried out by the specialists.

The only criterion is adopted at image level by considering: (i) positive the images related to melanomas and negative the ones related to dysplastic nevi and (ii) positive the images related to dysplastic nevi and negative the ones related to common nevi. Although pre-processing steps of the images allow for better performance [29, 30], the images we used were not pre-processed, i.e. they were not cleaned of the presence of noise, such as possible hair or halo left by the dermoscopic gel used to allow better illumination of the lesions.

The MIL algorithm, referred to as MIL-RL, has been re-implemented in Matlab and has been run on a Windows 10 system featuring a 2.21 GHz processor. We have duplicated all the images of melanomas, adding a Gaussian noise with zero mean with variance equal to 0.0001, as in the [6]. In this way we obtained a balanced dataset containing three classes of data, Melanomas (M), Dysplastic Nevi (DN) and Common Nevi (N) each with 80 images. The imbalance between the classes of training datasets should not be underestimated. The risk consists in undermining the classification performance of the models, which can manifest over-fitting, thus losing in generalization. For our experiments we considered the following two data set configurations:

- 160 images: 80 Melanomas vs 80 Dysplastic Nevi;
- 160 images: 80 Dysplastic Nevi vs 80 Common Nevi.

For each data set configuration, we performed two types of experiments, using a five fold and a ten fold cross-validation. The respective results are listed in Table 1 and Table 2, where we report the average of correctness, sensitivity, specificity, F-score and CPU time.

The proposed MIL optimization model is of SVM type; in order to appreciate the MIL classification paradigm, we report in the columns SVM and SVM-RBF the results obtained using a standard SVM approach [31] with linear and RBF kernels, respectively. For each data set configuration and for each experiment (5-CV and 10-CV), the best results in Table 1 and Table 2 have been underlined.

From numerical experiments it emerges that, in general, MIL-RL overcomes the SVM technique (with both linear and RBF kernels) in terms of correctness and sensitivity. Whenever accuracy is not 100%, low specificity values are a consequence of high sensitivity values. In medical fields, sensitivity plays a more important role than specificity since it is a measure of the ability to identify un-healthy patients. The F-score values show the good performance of the MIL approach in classifying melanoma from dysplastic nevi against the classic SVM technique.

We observe that the CPU times of MIL-RL algorithm are better than those recorded by linear SVM. Concerning the classification of dysplastic nevi against common nevi, the performances of all three methods appear unsatisfactory. The MIL-RL algorithm records the best values of accuracy and specificity, but overall

	5-CV			10-CV		
	MIL-RL	SVM	SVM-RBF	MIL-RL	SVM	SVM-RBF
Correctness (%)	87.50	72.50	85.63	86.25	69.38	86.25
Sensitivity (%)	92.56	77.21	87.06	91,08	69.65	87.88
Specificity (%)	81.50	67.51	85.51	82.12	69.87	85.95
F-score (%)	88.31	74.96	85.84	87.01	68.68	87.52
CPU time (secs)	0.90	1.84	0.04	1.20	2.05	0.03

Table 1. Average testing values: 80 melanomas and 80 dysplastic nevi

	5-CV			10-CV		
	MIL-RL	SVM	SVM-RBF	MIL-RL	SVM	SVM-RBF
Correctness (%)	60.63	60.00	49.38	59.38	58.13	51.88
Sensitivity (%)	35.07	54.91	53.58	31.77	43.67	58.92
Specificity (%)	84.86	67.58	46.97	87.06	73.48	46.47
F-score (%)	44.76	56.79	50.09	42.77	48.57	53.74
CPU time (secs)	1.38	1.95	0.01	1.71	2.13	0.03

Table 2. Average testing values: 80 dysplatic nevi and 80 common nevi

it is not effective to solve the proposed task. Better results could be obtained using images pre-processing steps and by using further useful features [33, 34].

5 Conclusions

In this paper we present an application of a MIL approach for the detection of melanoma by dysplastic nevi and of dysplastic nevi by common ones. These two issues are not widespread in the literature. Anyhow, pathologies such as the DNS require tools to support phisicians in diagnostic process and mobile applications useful for promoting self diagnosis. To this end, we point out that it is under implementation a module of the software Simpatico 3d that is in charge of allowing self diagnosis [35, 36, 38, 40, 41]. The obtained results show that in the first case the MIL approach is very promising, even using only color features and with no pre-processing step.

In the second case, the MIL approach as well as the SVM in the linear and Kernel RBF version, do not give satisfactory results. The excessive similarity of the lesions is not properly discriminated with approaches aimed at identifying linear separation surfaces. One way we intend to pursue is to apply MIL approaches that use spherical separation surfaces. In particular, the algorithm [32] seems to be an interesting proposal for the development of applications in which positive and negative elements have strong similar characteristics.

Future research could include the design of more sophisticated segmentation techniques in order to further improve classification results, as well as the application of the proposed method in other medical fields [39, 37] to identify other types of injuries.

References

1. http://gco.iarc.fr/today/explore.
2. R. Sanghera and P. S. Grewal, "Dermatological symptom assessment", in Patient Assessment in Clinical Pharmacy, p. 133–154, Springer, 2019.
3. Silva JH, et al., "Atypical mole syndrome and dysplastic nevi: identification of populations at risk for developing melanoma," Clinics (Sao Paulo), v.66(3):493–9, 2011.
4. A.C. Society, Melanoma skin cancer. http://www.cancer.org/acs/groups/cid/documents/webcontent/003120-pdf.pdf
5. Burroni M, et al., "Dysplastic naevus vs. in situ melanoma: digital dermoscopy analysis", Br J Dermatol, v.152(4), pp:679–84, 2005.
6. Barata C, et al. "Two systems for the detection of melanomas in dermoscopy images using texture and color features", IEEE Syst J 8(3), pp: 965–79, 2014.
7. Astorino, A., Fuduli, A., Veltri, P., and Vocaturo, E.,"On a recent algorithm for multiple instance learning. Preliminary applications in image classification". In IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1615-19, 2017.
8. Astorino, A., Fuduli, A., Gaudioso, M., and Vocaturo, E., "A multiple instance learning algorithm for color images classification". Proceedings of the 22nd International Database Engineering and Applications Symposium (pp. 262-266), 2018.
9. Astorino, A., Fuduli, A., and Gaudioso, M., "A Lagrangian relaxation approach for binary multiple instance classification". IEEE transactions on neural networks and learning systems, 2019.
10. Duffy K, Grossman D., "The dysplastic nevus: from historical perspective to management in the modern era: part I. Historical, histologic, and clinical aspects". J Am Acad Dermatol., 67(1):1.e1-1.e16. quiz 17-8, 2012.
11. Greene, M. H., et al., "High risk of malignant melanoma in melanoma-prone families with dysplastic nevi". Annals of internal medicine, 102(4), pp: 458-465, 1985.
12. Jeong, D. K., et al., "A case of malignant melanoma after repeated recurrent dysplastic nevi". Archives of craniofacial surgery, 20(4), 260, 2019.
13. R. Pampena, A. Kyrgidis, "A meta-analysis of nevus-associated melanoma: Prevalence and practical implications". In: Journal of the American Academy of Dermatology. Band 77, N. 5, pp. 938–945, November 2017.
14. Arumi-Uria M, McNutt NS, Finnerty B., "Grading of atypia in nevi: correlation with melanoma risk". Mod Pathol., 16(8):764-771, 2003.
15. Reddy KK, et al., "Atypical (dysplastic) nevi: outcomes of surgical excision and association with melanoma". JAMA Dermatol.,149(8): 928-934, 2013.
16. Gandini S, Sera F, Cattaruzza MS, et al., "Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical nevi". Eur J Cancer.,41(1):28-44, 2005.
17. Xiong MY, Rabkin MS, Piepkorn MW, et al., "Diameter of dysplastic nevi is a more robust biomarker of increased melanoma risk than degree of histologic dysplasia: a case-control study". J Am Acad Dermatol.,71(6):1257-1258, 2014.
18. Rastgoo, M., et al., "Automatic differentiation of melanoma from dysplastic nevi". Computerized Medical Imaging and Graphics, 43, 44-52, 2015.
19. Vocaturo, E., Perna D., and Zumpano E., "Machine Learning Techniques for Automated Melanoma Detection", IEEE International Conference on Bioinformatics and Biomedicine, (BIBM), pp. 2310-17, 2019.
20. Vocaturo, E., and Zumpano E., "Dangerousness of dysplastic nevi: a Multiple Instance Learning Solution for Early Diagnosis", IEEE International Conference on Bioinformatics and Biomedicine, (BIBM), pp: 2318-23, 2019.

21. Litjens, G., et al., "A survey on deep learning in medical image analysis". Med. Image Anal. 42, 60–88, 2017.
22. de Bruijne, M., "Machine learning approaches in medical image analysis: from detection to diagnosis". Med. Image Anal. 33, 94–97, 2019.
23. Carbonneau MA, et al., "Multiple instance learning: a survey of problem characteristics and applications". Pattern Recogn 77:329–353, 2018.
24. Quellec G, Cazuguel G, Cochener B, Lamard M, "Multiple instance learning for medical image and video analysis", IEEE Rev Biomed Eng 10, pp: 213–234, 2017.
25. Astorino, A., Fuduli, A., Veltri, P., and Vocaturo, E. (2020). "Melanoma Detection by Means of Multiple Instance Learning". Interdisciplinary Sciences: Computational Life Sciences, 12, 24–31, 2020.
26. A. Astorino, A. Fuduli, M. Gaudioso, and E. Vocaturo, "Multiple Instance Learning Algorithm for Medical Image Classification", Proceedings of the 27th Italian Symposium on Advanced Database (SEDB), 2019.
27. Mendonça T, et al., "Ph2 —a dermoscopic image database for research and benchmarking". In: 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 5437–5440, 2013.
28. A. Fuduli, P. Veltri, E. Vocaturo, E. Zumpano, "Melanoma detection using color and texture features in computer vision systems", Advances in Science, Technology and Engineering Systems Journal, vol. 4, no. 5, pp. 16-22 (2019).
29. E. Vocaturo, and E. Zumpano, and P. Veltri, "Image preprocessing in computer vision systems for melanoma detection," IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2117-24, 2018.
30. E. Vocaturo, and E. Zumpano, and P. Veltri, "On the Usefulness of Pre-Processing Step in Melanoma Detection Using Multiple Instance Learning", International Conference on Flexible Query Answering Systems, Springer, pp. 374-382, 2019.
31. Vapnik V., The nature of the statistical learning theory, Springer, New York 1995.
32. Gaudioso, M., Giallombardo, G., Miglionico, G., and Vocaturo, E., "Classification in the multiple instance learning framework via spherical separation". Soft Computing, pp:1-7, 2019.
33. E. Vocaturo, E. Zumpano, and P. Veltri, "Features for Melanoma Lesions Characterization in Computer Vision Systems", 9th International Conference on Information, Intelligence, Systems and Applications, (IISA) 2018, pp. 1–8, 2018.
34. Vocaturo E., Caroprese L., Zumpano E., "Features for Melanoma Lesions: Extraction and Classification", IEEE/WIC/ACM International Conference on Web Intelligence-Companion Volume,(pp.238-243), 2019.
35. E. Zumpano, et al., "SIMPATICO 3D: A Medical Information System for Diagnostic Procedures". BIBM pp. 2125-2128, 2018.
36. E. Zumpano, et al., "SIMPATICO 3D Mobile for Diagnostic Procedures", Proceedings of the 21st International Conference on Information Integration and Web-Based Applications and Services (IIWAS), pp. 468-472, Munich, Germany 2019.
37. E. Vocaturo, E. Zumpano and P. Veltri, "On discovering relevant features for tongue colored image analysis", IDEAS 2019, (pp.1-8), ACM 2019.
38. P. Iaquinta, et al.: "eIMES 3D: An innovative medical images analysis tool to support diagnostic and surgical intervention". FNC/MobiSPC 2017: 459-464.
39. E. Vocaturo, P. Veltri: "On the use of Networks in Biomedicine". FNC/MobiSPC 2017: 498-503.
40. P. Iaquinta, et al.:. "eimes 3d mobile: A mobile application for diagnostic procedures". BIBM, 1634–1641, 2018 DOI: 10.1109/BIBM.2017.8217904.
41. L. Caroprese, et al.:. "Software Tools for Medical Imaging". ADBIS (Short Papers and Workshops) 2018: 297-304