A Multiple Instance Learning Approach for the Automatic Classification of Skin Lesions

(Discussion Paper)

Eugenio Vocaturo[1,2], Ester Zumpano[1], Giovanni Giallombardo[1] and Giovanna Miglionico[1]

[1]*DIMES - University of Calabria, Rende, Italy*
[2]*CNR-NANOTEC National Research Council, Rende, Italy*

Abstract
The number of deaths linked to skin cancers has malignant melanoma as the main culprit. Early diagnosis helps manage this terrible form of cancer, but the similarity of melanoma to other skin lesions is an obstacle to effective detection. The scientific community is proposing different solutions to support the computerized analysis of skin lesions mainly focused on the dichotomous distinction of melanoma from benign lesions. The *dysplastic nevi syndrome* (DNS) correlates the number of moles present in the human body with an increased risk of melanoma development. Nowadays, the classification task concerning the differentiation of dysplastic nevi from common ones is still very little explored. In this paper, we explore the possibility of applying multiple instance learning (MIL) approaches to discriminate melanoma from dysplastic nevi and outline the even more complex challenge of discriminate between dysplastic and common nevi. The obtained results confirm that MIL techniques are useful for the automatic detection of skin lesions are promising, and give hope MIL techniques can be useful for solutions aiming at automatic detection of skin lesions.

Keywords
Dermoscopy imaging Classification, Multiple Instance Learning, Dysplastic nevi Detection

1. Introduction

The World Health Organization certifies that, in 2020, more than 57,000 people died of melanoma and that there were more than 320,000 new cases. The reported data testify that melanoma affects the populations of all geographical areas of the world and in particular those of Europe (50.1 % of total cases) and North America (27.7 % of total cases). Melanoma ranks 5th for age-standardized (World) incidence and mortality rates in 2020, for both males and females, considering all ages [1]. Despite the worrying scenario in terms of both new cases and deaths, if melanoma is identified by early diagnosis it is a treatable type of cancer. Specific clinical protocols such as the ABCDE [2] rule and the 7-PCL [3] are adopted as a guideline for identifying lesions from an early stage. The ABCDE rule, which is the most commonly adopted, suggests monitoring

SEBD 2021: The 29th Italian Symposium on Advanced Database Systems, September 5-9, 2021, Pizzo Calabro (VV), Italy
e.vocaturo@dimes.unical.it (E. Vocaturo); e.zumpano@dimes.unical.it (E. Zumpano); g.giallombardo@dimes.unical.it (G. Giallombardo); g.miglionico@dimes.unical.it (G. Miglionico)

symmetry, irregularity of the edges, colors of the lesion, its extension and evolution over time. Our proposal was applied to skin lesion images detected through dedicated instrumentation.

In particular, the used dataset contains dermatoscopic images: this particular type is widely used in Computer Aided Diagnosis (CAD) systems to support the diagnosis.

Considering that higher risk of developing melanoma pertains to individuals with *dysplastic nevi syndrome* and/or with family history of melanoma, our research focuses on the application of DC-SMIL [4], a multiple instance learning algorithm, on the challenging tasks of classifying melanoma vs dysplastic nevi and dysplastic nevi vs common ones [19, 24].

The first task results to be difficult for the great similarity of the two types of lesions [5]. Even more complex is the classification of dysplastic nevi from common ones: this issue is completely new and has not been addressed in the literature. Our goal is to verify how the MIL approaches are of interest when applied on binary classification tasks in which the images are very similar to each other.

The paper is organized as follows. In the next section we put in evidence that the presence of dysplastic nevi and common nevi may imply risk of melanoma onset. In Section 3 we introduce the Multiple Instance Learning approach, focusing on DC-SMIL a new MIL algoritmh that adopt spherical separation surfaces [4]. In Section 4 we describe the dermoscopic dataset used to test DC-SMIL reporting some preliminary results. Finally some conclusions are given.

2. Dysplastic nevi

The *Syndrome of Dysplastic Nevus (DNS)* refers to individuals that present a high number of both benign moles and dysplastic nevi. Individuals with dysplastic nevi are more likely to develop melanoma if familiarly conditions exists. In [6], a cumulative lifetime risk of almost 100% is reported for individuals who have dysplastic nevi and are related to melanoma; about 30% of melanomas occur within atypical moles. A genetic predisposition for the formation of melanoma is present in 40-50% of cases. The correlation between the presence of dysplastic nevi and the melanoma has been also investigated in [7]. The diagnosis of a severe DNS cannot be overlooked, as it could state for a miss-diagnosed in situ melanoma [8], it may reflect the dermatopathological uncertainty related to a wrong diagnosis. Figure 1 reports a dermoscopic image of common nevi, dysplastic nevi and melanoma.

Figure 1: Dermoscopic image of common nevus, dysplastic nevus and melanoma

Basically, the risk of melanoma is related to two different objective criteria:

- An increased risk of melanoma is related to a high number of nevi [9]. Individuals with a number of nevi greater than 100 have a risk of melanoma 7 times greater than those with a count of less than 15 [10].

- An increased risk of melanoma is related to the presence of large nevi. A histological study of nevi has shown that higher is the extension of the mole, greater is the risk of turning into melanoma: the relative risk of 1 for nevi with a diameter less than 2.4 mm, while the relative risk progressively increases up to 5 if the lesion has a diameter greater than 4.4 mm [11].

Fewer attentions have been given to the discrimination of melanoma from dysplastic nevi [12]. The topic investigated in this paper is the classification task of dysplastic nevi against common nevi, which, to the best of our knowledge, has never been taken into consideration.

3. Multiple instance classification via spherical separation

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.

Multiple Instance Learning scenario is particularly 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. MIL is an extension of supervised learning that can train classifiers using weakly labeled data. The goal is therefore to exploit the labels of the weaker bags for training. 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.

The MIL paradigm is particularly well suited to image classification, given that to classify an image, it is necessary to examine only some sub-regions. With MIL approaches it is therefore possible to obtain global information from local one. For general considerations on the MIL paradigm, we refer the reader to surveys [13, 14]. In [15], a detailed review is given concerning MIL applied for medical images and video analysis. MIL approaches, as far as we know, are still very rarely used for melanoma detection, and has never been used for the detection of dysplastic nevi.

In [16] we applied MIL-RL algorithm to discriminate melanoma from benign lesion. The results demonstrate the goodness of the proposed approach.

In a data driven way, we have therefore presented a new algorithm named DC-SMIL [4], which is suitable for image classification. DC-SMIL adopt spherical separation as a classification tool and come out with an optimization model which is of DC (Difference of Convex) type. In particular the adopted classification error function depend on center and radius of the sphere and we come out with an optimization model to minimize a combination of the volume of the sphere and of the classification error.

Our aim is to find a sphere $S(w, r) \subset \mathbb{R}^n$, of center $w \in \mathbb{R}^n$ and radius $r \in \mathbb{R}$, separating the two classes of bags. In order to separate the positive bags $X_1^+, \ldots, X_m^+$ from the negative ones $X_1^-, \ldots, X_k^-$, a sphere must have a nonempty intersection with each positive bag, while leaving outside all the instances belonging to negative bags.

A pictorial example of spherical separation is presented in Figure 2, where the sphere $S(w,r)$ separates the negative bags X_1^-, X_2^-, and X_3^- from the positive bags X_1^+ and X_2^+. In particular, we remark that while the bags depicted in Figure 2 are spherically separable, they are not separable by any hyper-plane.

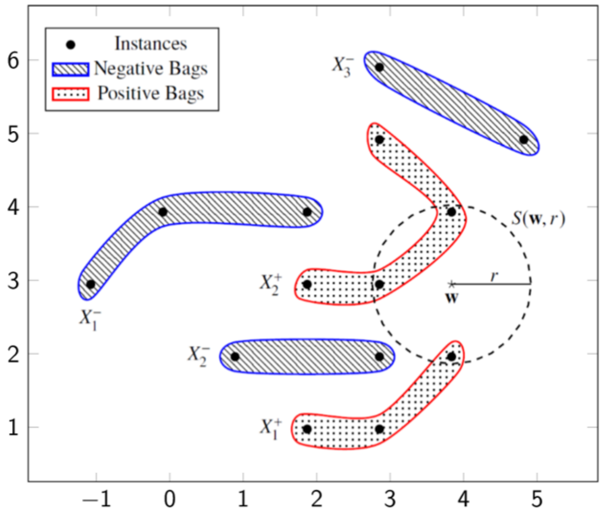

Figure 2: Spherical separation with three negative bags and two positive bags [4]

Based on the latter remark an optimization model was obtained with the aim to look for a separating sphere, if any, by minimizing a measure of all the classification errors of both the negative and the positive bags, that is

$$\min_{(w,r)\in\mathbb{R}^{n+1}} f(w,r) \tag{1}$$

where the loss function f is defined as

$$\begin{aligned} f(w,r) \quad &\triangleq \quad r^2 + C\sum_{i=1}^{k} \max\left\{0, \max_{j\in J_i^-}\{r^2 - \|x_j - w\|^2\}\right\} \\ &\qquad +C\sum_{i=1}^{m} \max\left\{0, \min_{j\in J_i^+}\{\|x_j - w\|^2 - r^2\}\right\} \end{aligned} \tag{2}$$

In particular, such loss function accounts for three contributions:

- the first term accounts for the volume of the sphere;
- the second one accounts for the misclassification error of the negative bags;
- the last term accounts for misclassification error of the positive bags.

Hence, the Spherical MIL program (SMIL) follows as the unconstrained optimization problem

$$\min_{(w,r)\in\mathbb{R}^{n+1}} f(w,r) \triangleq r^2 + C\mathcal{E}(w,r), \tag{3}$$

which combines, by introducing a trade-off parameter $C > 0$, the two objectives of minimizing the radius of the sphere and the classification errors of all the negative and positive bags. Here the radius minimization is aimed at reducing the false positive phenomenon when the calculated sphere is used as a classification tool.

4. Numerical results and final remarks

We have performed experiments applying DC-SMIL on various data sets to evaluate the goodness of the proposed technique and to compare the obtained results with those of other MIL methods. In particular, we applied DC-SMIL on a real dermatoscopic dataset (PH^2), with the aim of verifying that MIL spherical separation approach may be of interest in classification tasks in which the data to be classified have extreme similarity.

The entire PH^2 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 images without taking into account the indications resulting from the manual analysis carried out by the specialists.

In [17] the authors demonstrated how, by adopting only color features, satisfactory classification performances can be obtained using dermatoscopic images. Starting from this assumptions, we used a 30-dimensional vector for the representation of each sub-regions of each image. For further details please see [16] and [4]. To avoid the problems related to the use of datasets with unbalanced classes, we have duplicated all the images of melanomas, adding to the repeated ones a Gaussian noise with zero mean with variance equal to 0.0001, as in [17]. 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. For each data set configuration, we performed a ten fold cross-validation. The respective results are listed in Tables 1 and 2, where we report the average of correctness, sensitivity, specificity, F score and CPU time.

In order to appreciate the MIL classification paradigm, we report in the columns MIL-RL, SVM and SVM-RBF the results obtained using MIL-RL algorithm and standard SVM approach [18] with linear and RBF kernels, respectively. The best results in Tables 1 and 2 have been underlined.

4.1. Melanomas vs Dysplastic Nevi

From numerical experiments it emerges that, in general, MIL-RL overcomes DC-SMIL and SVM technique (with both linear and RBF kernels) in terms of accuracy 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.

	10-CV			
	DC-SMIL	MIL-RL	SVM	SVM-RBF
Correctness (%)	70.00	86.25	69.38	86.25
Sensitivity (%)	69.30	91,08	69.65	87.88
Specificity (%)	71.81	82.12	69.87	85.95
F-score (%)	69.09	87.01	68.68	87.52
CPU time (secs)	0.66	1.20	2.05	0.03

Table 1
80 melanomas and 80 dysplastic nevi

4.2. Dysplastic Nevi vs Common Nevi

With regard to the experimental section on the classification of dysplastic nevi against common nevi, the performances of MIL-RL and of SVM tecniques appear totally unsatisfactory. This is obvious because the images that were separated are very similar. MIL-RL registers the worst value of F-score and sensitivity, and overall it is not effective to solve the proposed task.

	10-CV			
	DC-SMIL	MIL-RL	SVM	SVM-RBF
Correctness (%)	59.38	59.38	58.13	51.88
Sensitivity (%)	59.73	31.77	43.67	58.92
Specificity (%)	59.88	87.06	73.48	46.47
F-score (%)	57.61	42.77	48.57	53.74
CPU time (secs)	0.58	1.71	2.13	0.03

Table 2
80 dysplatic nevi and 80 common nevi

The use of spherical separating surfaces, provided by DC-SMIL algorithm, allows significant improvements in the extremely difficult task of classify dysplastic nevi from common ones.

As shown in [19, 20] better results could be obtained in case of images pre-processing aimed at eliminating the presence of possible noises, such as possible hair. Even the adoption of further useful features extracted from blob is a possibility that would allow to improve the classification performances [21, 29]. Pre-processing steps and the adoption of a more numerous set of features appear to be an obligatory step when considering non-dermatoscopic images [22, 23].

The obtained results show that in the first case MIL-RL is very promising, even in the conditions in which we performed the experiments, i.e. with only color features and without using pre-processing steps.

In the second case, MIL-RL algorithm 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. On the other hand DC-SMIL, thanks to the use of spherical separation surfaces, seems to be an interesting proposal for the development of applications in contexts in which positive and negative elements have similar characteristics.

Our proposal based on MIL approaches, among the various proposals of artificial intelligence in this specific domain, constitutes an element of novelty [26]. Our goal is to set propose a framework for supporting diagnostics both for specialists and for patient self-diagnosis examination via mobile applications. In this way, modular solution which can be incorporated into integrated diagnostic systems [27, 28] would increase the value of the proposal.

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 [29, 30] to identify other types of injuries.

References

[1] http://gco.iarc.fr/today.

[2] R. Sanghera and P. S. Grewal, Dermatological symptom assessment, in Patient Assessment in Clinical Pharmacy, p. 133–154, Springer, 2019.

[3] G. Argenziano et al., Epiluminescence microscopy for the diagnosis of doubtful melanocytic skin lesions: Comparison of the abcd rule of dermatoscopy and a new 7-point checklist based on pattern analysis, Archives of Dermatology, v. 134, n. 12, pp. 1563–70, 1998.

[4] M. Gaudioso, et al., Classification in the multiple instance learning framework via spherical separation, Soft Computing, v.24, n.7, pp: 5071-5077, 2020.

[5] M. Burroni, et al., Dysplastic naevus vs. in situ melanoma: digital dermoscopy analysis. Br J Dermatol, v.152(4), pp:679–84, 2005.

[6] R. Pampena, A. Kyrgidis, A meta-analysis of nevus-associated melanoma: Prevalence and practical implications, American Academy of Dermatology. v. 77, n.5, pp. 938–945, 2017.

[7] M. Arumi-Uria, NS. McNutt, B. Finnerty, Grading of atypia in nevi: correlation with melanoma risk". Mod Pathol., 16(8):764-771, 2003.

[8] KK. Reddy, et al., Atypical (dysplastic) nevi: outcomes of surgical excision and association with melanoma. JAMA Dermatol.,149(8): 928-934, 2013.

[9] E. Rieger, et al., Overall and site-specific risk of malignant melanoma associated with nevus counts at different body sites: a multicenter case-control study of the German Central Malignant-Melanoma Registry. Int J Cancer. 62(4): 393-7, 1995.

[10] S. Gandini, F. Sera, M.S.Cattaruzza , et al., "Meta-analysis of risk factors for cutaneous melanoma: I. Common and atypical nevi". Eur J Cancer.,41(1):28-44, 2005.

[11] M.Y. Xiong, M.S. Rabkin,M.W. Piepkorn, 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.

[12] M. Rastgoo, et al., Automatic differentiation of melanoma from dysplastic nevi. Computerized Medical Imaging and Graphics, 43, 44-52, 2015.

[13] J. Amores, Multiple instance classification: review, taxonomy and comparative study. Artificial Intelligence 201:81–105, 2013.

[14] M.A. Carbonneau, et al., Multiple instance learning: a survey of problem characteristics and applications. Pattern Recogn 77:329–353, 2018.

[15] G. Quellec, G. Cazuguel, B. Cochener, M. Lamard, Multiple instance learning for medical image and video analysis, IEEE Rev Biomed Eng 10, pp:213–234, 2017.

[16] A. Astorino, et al., Melanoma Detection by Means of Multiple Instance Learning. Interdiscip Sci Comput Life Sci 12, 24–31, 2020.
[17] C. Barata, M. Ruela, M. Francisco, A.T. Mendonc, J. Marques. Two systems for the detection of melanomas in dermoscopy images using texture and color features. IEEE Syst J 8(3), pp: 965–79, 2014.
[18] V. Vapnik, The nature of the statistical learning theory, Springer, 1995.
[19] E. Vocaturo, E. Zumpano, 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.
[20] E. Vocaturo, E. Zumpano, P. Veltri, Image preprocessing in computer vision systems for melanoma detection, 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2117-24, 2018.
[21] E. Vocaturo, E. Zumpano, P. Veltri, Features for Melanoma Lesions Characterization in Computer Vision Systems, 9th International Conference on Information, Intelligence, Systems and Applications(IISA), pp. 1–8, 2018.
[22] A. Astorino, A. Fuduli, M. Gaudioso, E. Vocaturo, Multiple Instance Learning Algorithm for Medical Image Classification, Proceedings of the 27th Italian Symposium on Advanced Database (SEDB), 2019.
[23] 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.
[24] E. Vocaturo, E. Zumpano, Dangerousness of dysplastic nevi: a Multiple Instance Learning Solution for Early Diagnosis, 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2318-23, 2019.
[25] E. Vocaturo, E. Zumpano, G. Giallombardo, G. Miglionico: DC-SMIL: a multiple instance learning solution via spherical separation for automated detection of displastyc nevi, Proceedings of the 24th Symposium on International Database Engineering & Applications (IDEAS), pp. 4:1-4:9, 2020.
[26] Vocaturo, E., Perna D., and Zumpano E., Machine Learning Techniques for Automated Melanoma Detection, 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2310-17, 2019.
[27] E. Zumpano, et al., SIMPATICO 3D: A Medical Information System for Diagnostic Procedures. 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2125-2128, 2018.
[28] E. Zumpano, P. Iaquinta, F. Dattola, L. Caroprese, G. Tradigo, P. Veltri, E. Vocaturo, SIMPATICO 3D Mobile for Diagnostic Procedures,Proceedings of the 21st International Conference on Information Integration and Web-based Applications & Services (IIWAS), pp. 468-472, 2019.
[29] E. Vocaturo, E. Zumpano, P. Veltri, On discovering relevant features for tongue colored image analysis, Proceedings of the 23rd International Database Applications & Engineering Symposium, IDEAS, pp. 1-8, 2019.
[30] E. Vocaturo, E. Zumpano, The contribution of AI in the detection of the Diabetic Retinopathy, 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 1516-1519, 2020.