Decision Support Software for Melanoma Skin Cancer Detection (DECIME)

Carlos E. B. Sousa and **Leonardo B. G. Nascimento** and **Cláudio. M. S. Medeiros** and **Mauro Oliveira**[1] and **Iago A. Trajano**[2]

Abstract. In this paper is proposed a software to decision support for the detection of melanoma cancer. This approach is proposed, as this type of skin cancer is the only one that can metastasize, that is, proliferate to other organs such as lungs, liver, etc. The proposed system is elaborated through the acquisition of a set of images with 17805 samples, extractors of attributes gray level co-occurrence matrix (GLCM), Local Binary Pattern (LBP) and Central Moments. For the training and sample classification process, the Single Layer Perceptron (SLP), Multilayer Perceptron (MLP) and Support Vector Machine (SVM) classifiers are used. Subsequently, the best-evaluated model is implanted in a Raspberry Pi computer that, together with a webcam and a computer screen, allows the capture and classification of skin lesions in real-time. In the applied methodology, the authors obtained the best result of a 93 % accuracy from the use of Central Moments extractor and MLP classifier. In contrast to the state of the art of this problem, a high level of similarity is found between the accuracy rates of the Single Layer Perceptron (SLP) and the Multilayer Perceptron (MLP), demonstrating a possible resolution of the problem in a linear format.

1 INTRODUCTION

Skin cancer is currently presented as the most common illness in the world among fair-skinned people, the non-melanoma type being the most common kind of skin cancer worldwide. There are two major types of skin cancer: melanoma and non-melanoma [18][7].

Melanoma skin cancer is a malign tumor that stems from the uncontrolled proliferation of melanocytes (pigment producing cells)[2][1][11][16]. There are recent records of increase in mortality rates recurring from melanoma. The costs for melanoma treatment can be evaluated in several billions to countries that display a larger incidence of the disease. Several preventive strategies have been implemented in various high risk regions, with different success rates. Over the last four decades, melanoma incidence has increased worldwide, the highest of which being in Australia, where there are records of 40 new cases for every 100.000 citizens per year [13].

In this paper, therefore, is proposed a decision support software tool that employs techniques from Computer Vision and Artificial Intelligence using the classifiers Single Layer Perceptron (SLP) [19], Multilayer Perceptron (MLP) [14] and Support Vector Machine (SVM) [22] to make decision, providing healthcare professionals with a technological tool and giving the patients the possibility of a faster treatment.

The paper has the following organization: the Section 2 presents the related works that helped in the construction of this work. In Section 3 it is described the proposed method, that is, the acquisition of the set of images, elaboration of data sets, parameterization of the machine learning algorithms and implementation of the solution in a Raspberry Pi. The results obtained are presented in Session 4 and, finally, Section 5, the final considerations are presented.

2 RELATED WORKS

Mariam A. Sheha, Mai S. Mabrouk and Amr Sharawy [12] propose an automated method applied to a set of dermoscopic images for the diagnosis of melanoma. The extracted resources are based on Gray Level Co-occurrence Matrix (GLCM) and use of the MultiLayer Perceptron classifier (MLP) to classify melanoma cancer. The classifier chosen was used with two different techniques in the training and testing process: traditional and automatic. Traditional MLP obtained a superior performance to automatic MLP. While automatic MLP obtained a 93.4 % accurac y of the training set and 76 % on the testing set, traditional MLP obtained 100 % accuracy on the training set and 92 % on the testing set.

On another approach, H. Alquran et al [4] propose a method for detection, presentation, extraction and classification to detect melanoma, using image processing techniques that were applied on dermoscopic image samples suspect of melanoma. The classification system uses a SVM to classify the lesions. The study's SVM classifier obtained 92,1% accuracy in the classification, and appears to be a promising approach on the distinction of skin lesions of both benign and malign melanoma.

Ansari, U. B. and Sarode, T. [6] propose a system for early detection of skin cancer. The diagnosis methodology uses image processing techniques associated with a SVM classifier. The dermoscopic skin cancer images are obtained and submitted to several preprocessing techniques for noise removal and image enhancement. They are subsequently submitted to image segmentation by thresholding. The authors obtained an accuracy of 95 % with the aforementioned techniques.

3 THE PROPOSED METHOD

As previously mentioned, the authors of this work propose a decision support software for melanoma skin cancer detection. That said, as

[1] Federal Institute of Education, Science and Technology of Ceará, Brazil, email: carlosestevaobs, leonardobrendoti,claudiosa1965, amauroboliveira@gmail.com

[2] Federal University of Ceará, Brazil, email: iagoatrajano@gmail.com

shown in Fig. 1, the process is divided into 6 steps. The initial 4 stages consist of implementing the computational tool, while the last two stages comprehend practical use of the software.

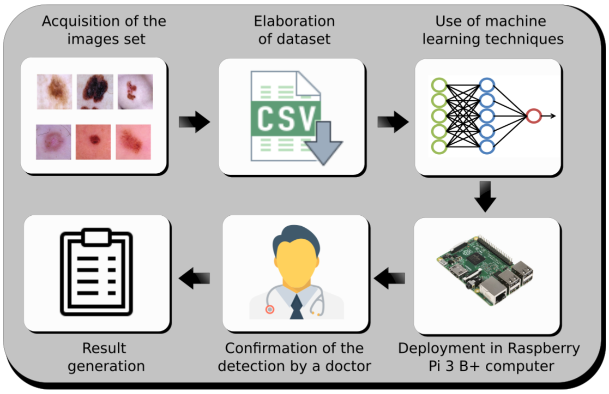

Figure 1. Steps of Implementation and use for the proposed software.

As shown on Fig. 1, the implementation of the tool can be divided into 4 stages: the acquisition of the image set, the elaboration of the data set, the use of machine learning techniques and the deployment on Raspberry Pi. The methodologies used for each illustrated step are presented next.

3.1 Acquiring the image set

The information is extracted from an image set named "dermoscopic pigmented skin lesions from HAM10k"[3]. This set contains 8903 samples of melanoma skin cancer images and 8902 samples of non-melanoma skin cancer images (Fig. 2) , with a total of 17805 samples.

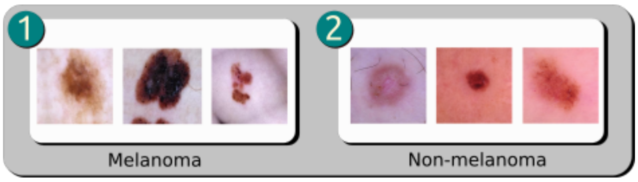

Figure 2. Example of samples used.

As for the ABCDE rule, professionals use The first five letters of the alphabet a guide to help people recognize the warning signs of melanoma. They are: Asymmetry (A) - wounds or stains are presented asymmetrically; Borders (B) - tend to be irregular; Colors (C) - they present different colors in the same wound; D (Diameter) - equal to or greater than 6mm; and Evolution (E) - whether it evolves in shape, color or elevation [21].

In this sense, based on the aforementioned rule, the authors propose the use of attribute techniques based on textures and shapes, as they are similar to the way of detecting the problem in a practical way. The techniques used are presented next.

[3] The data set is made available by Alexander Scarlat through the link: https://www.kaggle.com/drscarlat/melanoma

3.2 Elaborating the dataset

The authors propose the use of the extractors Gray Level Co-occurrence Matrix (GLCM) [8], Local Binary Pattern (LBP) [17] and Central Moments. The GLCM describes texture through a set of characteristics for the occurrences of each level of gray in the pixels of the image considering multiple directions [5]. In this way, the size of the matrix is determined from the distinct number of levels of pixels contained in the original image [20]. Haralick, K. Shanmugam and Dinstein [8] proposed a method to which 14 statistical measures of texture can be obtained from its use, however, only 13 are actually used, since the latter presents computational instability.

D. Huang et al. [9] state that in recent years, LBP has sparked a growing interest in image processing and computer vision. This method efficiently summarizes local image structures, comparing each pixel with its neighbor pixels. P. Mohanaiah and P. Sathyanarayana and L. GuruKumar [15] cite that one of the most important properties of this operator is computational simplicity, it makes it possible to analyze images in challenging configurations in real-time.

Moments are the statistical expectation of certain power functions of a random variable. The central moment is widely used in pattern recognition because of their discrimination power and robustness [3]. The main advantage of this extractor is their invariances to translations of the object. Therefore they are suited well to describe the shape of the object [10], which can become a strong tool for the elaboration of a data set that represents well the problem of the detection of melanoma cancer.

3.3 Applying machine learning techniques

Regarding the artificial intelligence algorithms, the single-layer Perceptron algorithm is implemented to verify the possibility of dealing with linear problems and, subsequently, the Multilayer Perceptron (MLP) and the Support Vector Machine, as they present themselves as robust algorithms to obtain better results, in case of a nonlinear problem.

To verify the efficiency of the proposed methodology, experiments are carried out with the Single Layer Perceptron (SLP) classifier, using learning rates of 0.01, 0.05, 0.1 and 0.5, signal function and reduction linear or exponential learning rates.

For the MLP classifier, we implemented several learning rates that vary between 0.1 and 0.5, linear or exponential reduction of these rates, logistic function or hyperbolic tangent. The values referring to the number of neurons in the hidden layer of the MLP are entered empirically, ranging from 2 to 100 neurons.

For the SVM classifier, 4 types of kernels are used; linear, RBF, polynomial and sigmoid. The value of C, that is, the penalty parameter is entered from the following values 0.01, 0.1, 1.0 and 10.

For both algorithms, the training is conditioned to end when it reaches an error rate of 10^{-5} and there are 2000 iterations, in which the weights with the best accuracy are chosen as the ones that best represent the results of the training.

As mentioned, the authors used 17.805 images in total, which are divided into 8.903 samples of the class "melanoma" and 8.902, of "non-melanoma". In this sense, 7.122 samples selected at random for each type of class were used to compose the training/tests set. The rest of the data, that is, 1.781 samples of the class "melanoma" and 1.780 samples of the class "non-melanoma", make up the test set.

For all classifiers, the K-Fold cross-validation technique is used, where a value of K equals 10. The application of this method consists of dividing the total set of training data into K subsets of the same size (Fig. 3). Thus, each subset is used for validation, while the rest of the set is applied to estimate the parameters.

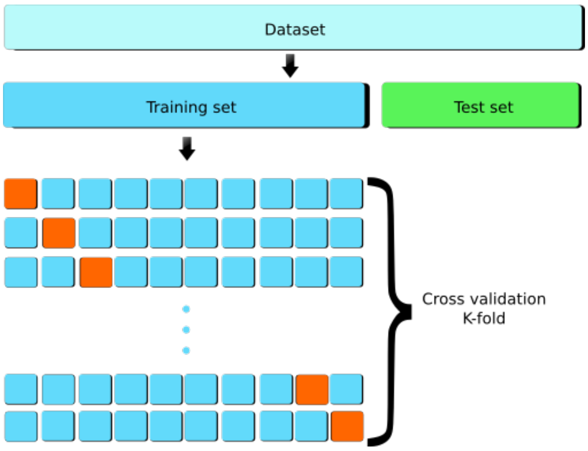

Figure 3. Training and testing processes.

As shown in Fig. 3, the data set is divided into 2 parts. The first is divided into 10 subsets that are used in the training and validation process. Finally, the second part of the set, unknown by the classifier, is used for the test process.

3.4 Deploying the software in Raspberry Pi 3 B+ computer

In order to make this application usable to healthcare professionals as well as to attain low acquisition costs, the computer program (Fig. 4) was deployed to a Raspberry Pi 3B+ computer.

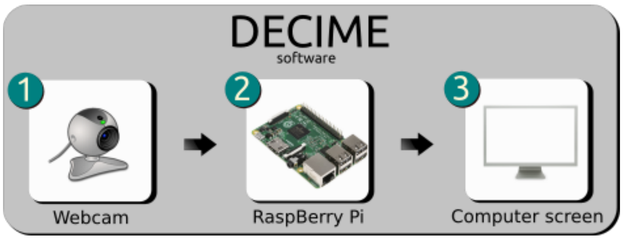

Figure 4. Deploying the program to a Raspberry Pi 3 B+ computer with a webcam.

After installing the classification software on the raspberry computer, we can use the webcam to obtain images of the skin for analysis, and the computer screen for verifying the suggested results.

3.5 Confirming the detection with a doctor and generating results

As previously mentioned, the authors propose a decision support software. Thus, the results generated by it must be analyzed by a doctor for proper conclusions.

4 RESULTS OBTAINED

In the Table 1 shows the best results obtained in this study. This solution is obtained with the Central Moments extractor and Multilayer Perceptron classifier. It should be noted that, in order to obtain the best results, the authors use various combinations between their attributes generated by the extractors. Thus, were obtained the best results (Table 1) by applying the following Central Moments: mu11, mu21, mu12, mu30, mu03.

Table 1. Results obtained from the extraction of attributes with Central Moments and classification with Multilayer Perceptron.

Stage	N.	Acc	Prec.	Rec.	F1
Train	Mel	1.00	1.00	0.88	0.94
	Non-mel.	0.89	0.90	1.00	0.94
Test	Mel	0.88	0.99	0.87	0.93
	Non-mel.	0.99	0.89	0.99	0.94

Although the Multilayer Perceptron classifier offers a better solution from the data obtained by the Central Moments extractor, the authors perceive similar results from the same attributes used with the Single Layer Perceptron classifier. This perception demonstrates that the problem of detecting melanoma cancer can be linearly separable, making it an easily resolvable problem. Table 2 illustrates the best results obtained from the classifier SLP.

Table 2. Results obtained from the extraction of attributes with Central Moments and classification with Single Layer Perceptron.

Stage	N.	Acc	Prec.	Rec.	F1
Train	Mel	0.97	0.97	0.89	0.93
	Non-mel.	0.90	0.90	0.98	0.94
Test	Mel	0.89	0.97	0.89	0.93
	Non-mel.	0.98	0.90	0.98	0.93

It is important to emphasize that by using the Central Moments extractor, one can divide the data almost linearly. Therefore, the decision limits for both classifiers have similarities and simplicity in their structures.

5 CONCLUSIONS

In view of the literature involved, this work was differentiated by using less robust techniques and achieving similar or even superior results. Providing a lower computational cost in the early detection of melanoma cancer.

Although most Artificial Intelligence models require powerful processing and extensive memory resources, there are still methods that, when applied correctly, produce satisfactory results. Therefore, the exact application of an attribute extractor related to the binary detection of melanoma skin cancer produces a good solution to the problem without requiring much processing power or memory. To that extent, the use of attribute extractors suitable for a problem can simplify its classification. In this scenario, a Perceptron can be a significant resource for achieving satisfactory results.

It should be noted that when deploying the software on the raspberry pi board, a simple, fast and portable solution is created, which enables support the decision to health care professionals regarding melanoma skin cancer detection. The software also helps improve accessibility to patients who live in regions of difficult access, since the device is portable enough for professionals to provide home care when the patient is unable to reach the care unit.

REFERENCES

[1] Z Abdel-Malek, V B Swope, I Suzuki, C Akcali, M D Harriger, S T Boyce, K Urabe, and V J Hearing, 'Mitogenic and melanogenic stimulation of normal human melanocytes by melanotropic peptides.', *Proceedings of the National Academy of Sciences*, **92**(5), 1789–1793, (1995).

[2] Tada A Im S Akcali C. Abdel-Malek Z, Suzuki I, *The melanocortin-1 receptor and human pigmentation*, Acad Sci., Ann N Y, 2nd edn., 1999.

[3] Sundos A Hameed Al_azawi, 'Eyes recognition system using central moment features', *Engineering and Technology Journal*, **29**(7), 1400–1407, (2011).

[4] H. Alquran, I. A. Qasmieh, A. M. Alqudah, S. Alhammouri, E. Alawneh, A. Abughazaleh, and F. Hasayen, 'The melanoma skin cancer detection and classification using support vector machine', in *2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, pp. 1–5, (2017).

[5] S Annadurai, *Fundamentals of digital image processing*, Pearson Education India, 2007.

[6] Uzma Bano Ansari and Tanuja Sarode, 'Skin cancer detection using image processing', *Int Res J Eng Technol*, **4**(4), 2875–2881, (2017).

[7] Skin Cancer Foundation. Melanoma warning signs - what you need to know about early detection, 20-.

[8] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein, 'Textural features for image classification', *IEEE Transactions on systems, man, and cybernetics*, (6), 610–621, (1973).

[9] Di Huang, Caifeng Shan, Mohsen Ardabilian, Yunhong Wang, and Liming Chen, 'Local binary patterns and its application to facial image analysis: a survey', *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, **41**(6), 765–781, (2011).

[10] Yaser Daanial Khan, Nabeel Sabir Khan, Shoaib Farooq, Adnan Abid, Sher Afzal Khan, Farooq Ahmad, and M Khalid Mahmood, 'An efficient algorithm for recognition of human actions', *The Scientific World Journal*, **2014**, (2014).

[11] Aaron B. Lerner and Joseph S. McGuire, 'Melanocyte-stimulating hormone and adrenocorticotrophic hormone', *New England Journal of Medicine*, **270**(11), 539–546, (1964). PMID: 14091524.

[12] Mai Mabrouk, Mariam Sheha, and Amr Sharawi, 'Automatic detection of melanoma skin cancer using texture analysis', *International Journal of Computer Applications*, **42**, 22–26, (03 2012).

[13] Vishal Madan, John T Lear, and Rolf-Markus Szeimies, 'Non-melanoma skin cancer', *The Lancet*, **375**(9715), 673 – 685, (2010).

[14] Warren S. McCulloch and Walter Pitts, 'A logical calculus of the ideas immanent in nervous activity', *The bulletin of mathematical biophysics*, **5**(4), 115–133, (Dec 1943).

[15] P Mohanaiah, P Sathyanarayana, and L GuruKumar, 'Image texture feature extraction using glcm approach', *International journal of scientific and research publications*, **3**(5), 1, (2013).

[16] K. Nouri, *Skin Cancer*, MGraw Hill professional, McGraw-Hill Education, 2007.

[17] Timo Ojala, Matti Pietikainen, and Topi Maenpaa, 'Multiresolution gray-scale and rotation invariant texture classification with local binary patterns', *IEEE Transactions on pattern analysis and machine intelligence*, **24**(7), 971–987, (2002).

[18] D.S. Rigel, R. Friedman, J.K. Robinson, M.I. Ross, C.J. Cockerell, H. Lim, E. Stockfleth, and J.M. Kirkwood, *Cancer of the Skin E-Book: Expert Consult*, Elsevier Health Sciences, 2011.

[19] F. Rosenblatt, 'The perceptron: a theory of statistical separability in cognitive systems (project para)', (1958).

[20] William Robson Schwartz and Hélio Pedrini, 'Método para classificação de imagens baseada em matrizes de coocorrência utilizando características de textura', *III Colóquio Brasileiro de Ciências Geodésicas, CuritibaPR, Brasil*, 1, (2003).

[21] Marina Tsatmali, Janis Ancans, and Anthony J Thody, 'Melanocyte function and its control by melanocortin peptides', *Journal of Histochemistry & Cytochemistry*, **50**(2), 125–133, (2002).

[22] V. N. Vapnik, 'The nature of statistical learning theory', *Spring - Verlag New York*, (1995).