Embedded System to Support Skin Cancer Recognition

Gabriel de A. Batista[1(✉)], Marcelo Nogueira[1,2], Nuno Santos[2,3], and Ricardo J. Machado[2,3]

[1] Software Engineering Research Group, Paulista University, UNIP, Campus Tatuapé, São Paulo, Brazil
[2] ALGORITMI Centre, School of Engineering, University of Minho, Guimarães, Portugal
[3] CCG/ZGDV Institute, Guimarães, Portugal

Abstract. Skin cancer is the most common among all cancers and its early diagnosis increases the patient's chances of healing. One of the ways to make this diagnosis is through dermatoscopy. Dermatoscopy is a technique that consists of recognizing structures present in the skin, not visible to the naked eye. Therefore, for assisting the use of dermatoscopy by health professionals, this work presents a device to support skin cancer recognition using the histogram of oriented gradients and machine learning, based on the ABCDE rule.

Keywords: Machine learning · Skin cancer · Histogram of Oriented Gradients · Gaussian Naive Bayes · K Neighbors Classifier

1 Introduction

Nowadays, there are great technological advances, in which it is possible to count on the support of intelligent systems that are increasingly present in commerce, industry, medicine, finance, etc. One of the great advances is the computer vision that is related to image analysis, which has been developing a lot in recent years. This area deals with the extraction of information from images and the identification and classification of objects present in them. Computer vision systems have been used to recognize people, signatures and objects; inspection of parts on assembly lines; orientation of robot movements in automated industries etc. They involve image analysis and artificial intelligence or decision-making techniques, which allow the identification and classification of objects or images.

The technology has brought many health benefits, such as electronic devices (ultrasound, defibrillator, pulse oximeter, etc.), applications, expert systems to aid decision making and even artificial neural networks for pre-diagnosis of diseases such as breast cancer or skin cancer. Dermatologists use equipment that makes it possible to scan images of skin lesions, allowing for clinical skin evaluation and monitoring of the development of the disease. The advent of large collections of medical images brought with it the need to use computational techniques for efficient processing, analysis and retrieval of the information contained in the image.

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12254, pp. 725–740, 2020.
https://doi.org/10.1007/978-3-030-58817-5_52

Skin cancer is by far the most common type of cancer, with basal cell carcinoma and squamous cell carcinoma being the most frequent [1], and melanoma the least frequent [2].

The world estimate shows that, in 2018, there were 287,723 new cases of melanoma skin cancer [3] and 1,042,056 non-melanoma skin cancer [4], totaling 1,329,779. There were 125,867 deaths, 65,155 from non-melanoma [4] and 60,712 from melanoma [3]. Male gender registered the highest number of occurrences, both in incidence and mortality (Fig. 1). In general, the highest incidence and mortality rates were observed in North America, Europe and Asia, and the lowest in Latin America, the Caribbean, Oceania and Africa (Fig. 2).

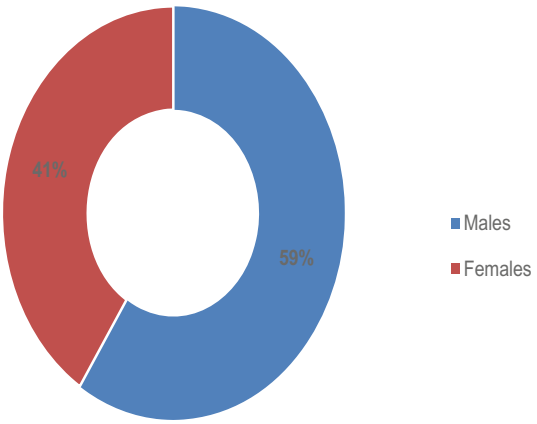


Fig. 1. Skin cancer occurrence rate.

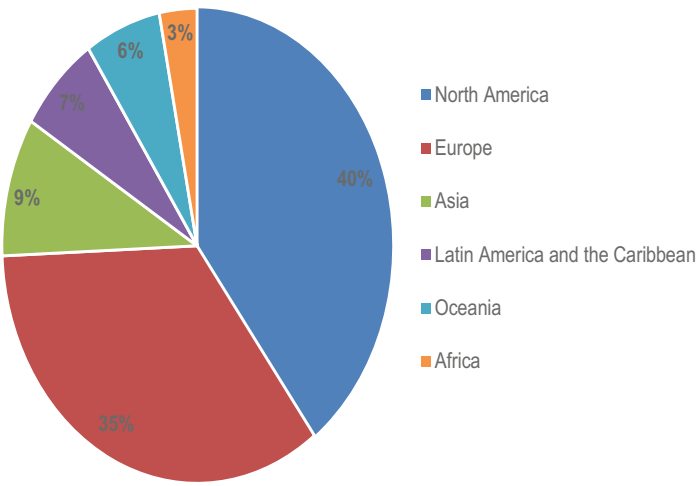
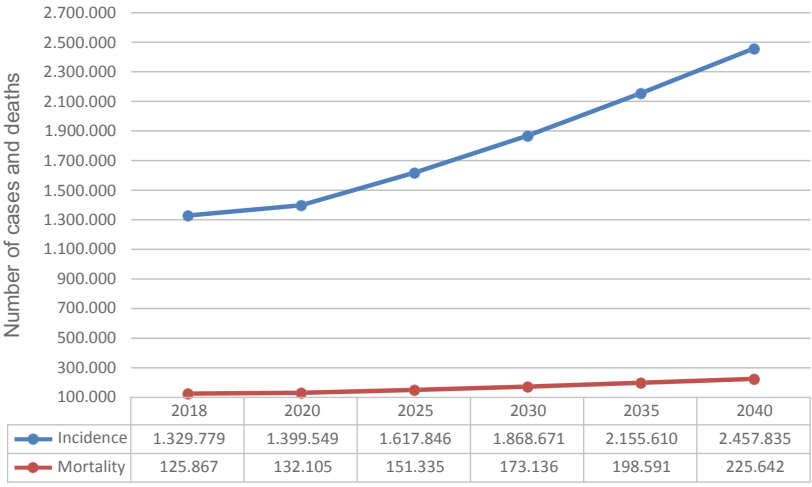


Fig. 2. Skin cancer incident and mortality rates, both sexes.

It is estimated that by the year 2040, incidents of melanoma skin cancer will increase by 62.3% and non-melanoma skin cancer will increase by 91.1%. The mortality rate will follow the same rate of increase, 74.4% for melanoma skin cancer and 83.8% for non-melanoma skin cancer (Fig. 3) [5].



	2018	2020	2025	2030	2035	2040
Incidence	1.329.779	1.399.549	1.617.846	1.868.671	2.155.610	2.457.835
Mortality	125.867	132.105	151.335	173.136	198.591	225.642

Fig. 3. Perspective of skin cancer incidents and mortality from 2018 to 2040.

The highest incidence rates worldwide are found in populations with a predominance of lighter skin color, such as Australia and New Zealand. Despite being the most frequent cancer, non-melanoma skin cancer is difficult to estimate, since not all cases are registered [6].

This work presents a device for supporting dermatoscopy by health professionals with the recognition of skin cancer using the histogram of oriented gradients and Gaussian Naive Bayes and K Neighbors classifiers, based on the ABCDE rule.

This paper is structured as follows: Sect. 2 presents the related work, encompassing topics from descriptors of characteristics, solutions for skin cancer treatments, BI-RADS and machine learning algorithms; Sect. 3 presents the methods applied for developing the device; Sect. 4 presents the results registered; Sect. 5 presents the research discussions; and Sect. 6 presents the conclusions.

2 Related Work

2.1 Descriptors of Characteristics

Most of the time, computer vision applications involve computationally complex tasks, i.e., object tracking, object identification, optical flow, among others. The first steps of all these applications are the detection, description and matching of the characteristics of high qualities, the descriptors being the most complicated and slow.

Descriptors focus on abstracting information from images that are associated with points of interest detected by the feature detector, however descriptors should avoid being complex or using too many math operations. A high-quality feature descriptor only describes a feature point, correctly identifying it in subsequent images [7].

The descriptor based on Histogram of Oriented Gradients (HOG) was proposed in 2005 by researchers Dalal and Triggs as part of a pedestrian detection algorithm in images. Generally used in pattern recognition and image processing to detect or recognize objects. This method aims to extract information regarding the orientation of the existing edges in an image, these edges being calculated through edge detection methods such as Sobel [8].

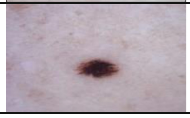
2.2 Treatments for Skin Cancer and the ABCDE Rule

Excessive and chronic exposure to the sun is the main risk factor for the onset of non-melanoma skin cancers, in relation to melanoma, in general, the greatest risk factor includes a personal or family history, in addition to sporadic and intense exposure in the sun with consequent sunburn in more than one episode. Other risk factors for all types of skin cancer include skin sensitivity to the sun and its color [9].

Surgery is the indicated treatment for melanoma skin cancer. Other forms of treatment that can be successful would be radiotherapy and chemotherapy depending on the stage of the cancer. When metastasis has already occurred (the cancer has already spread to other organs), melanoma is incurable in most of its cases. At this stage, a treatment strategy would only be to relieve symptoms and improve the patient's quality of life [10]. For non-melanoma skin cancer, surgery would be the most indicated treatment for both basal cell and epidermoid carcinoma. Basal cell carcinoma, when of low extent, can be treated with a topical medication (ointment) or radiation therapy, while epidermoid carcinoma, the usual treatment combines surgery and radiation therapy [11].

Skin spots or stains can be classified in a rule called ABCDE, which consists of evaluating five distinct characteristics. The same spot can have one or more of these characteristics and the higher the number, the greater the degree of suspicion of being a skin tumor. Some malignant skin tumors, however, escape this description and it is best to see a specialist if you suspect something different [12]. The ABCDE rule can be verified as in Table 1.

Table 1. ABCDE rule. [13]

ABCDE Rule	Benign	Malignant
A = Asymmetry, melanoma is suspected: the tumor is divided in half, and the halves are not similar.		
B = Irregular border, suspected melanoma: uneven or irregular borders.		
C = Color variation, suspected melanoma: there is more than one color of pigment.		
D = Diameter, suspected melanoma: if the diameter is greater than 6 mm.		

2.3 BI-RADS

The BI-RADS (Breast Imaging Reporting and Data System) is a system considered the greatest reference for standardization and uniformity of mammography. It was proposed by the American College of Radiology, with a focus on assisting and standardizing mammography so that the best approach can be defined, being defined in 6 levels, with zero being undetermined [14].

The annual hematological screening in women over 40 years old identifies 100 to 200 new cases of suspicious lesions in every 20,000 mammograms, with BI-RADS being a way to standardize and designate corresponding examinations [14].

2.4 Machine Learning

Machine learning is a field of study that gives computers the ability to learn without being explicitly programmed. Machine learning is the ability to improve performance in performing a task through experience [15]. It is an extremely important segment in artificial intelligence.

Informally, an algorithm is any well-defined computational procedure that takes some value or set of values as an input and produces some value or set of values as an output. Therefore, an algorithm is a sequence of computational steps that transform input into output.

There are several machine learning algorithms, but some were used in this work the MLP Classifier, Random Forest Classifier, AdaBoost Classifier, K Neighbors Classifier, Support Vector Machines, Gaussian Process Classifier, Quadratic Discriminant Analysis, Gaussian NB and Decision Tree Classifier.

During the process of creating a machine learning model, we need to measure its quality according to the objective of the task. There are mathematical functions that help us to evaluate the error and correctness of our models.

The metrics used in this work were ROC Curve, Confusion Matrix, Accuracy, Precision, Log Loss, Sensitivity and F1-Score. Since this research focused in using the Confusion Matrix, it is further detailed.

A confusion matrix is one of the easiest and most intuitive metrics to find the accuracy and precision of a model. It is used as a classification for problems where the output can be of two or more types of classes.

The confusion matrix has the following terms:

- TP (True Positives): the true positives are the cases where the true class is 1 (true) and the predicted also 1 (true). An example of a situation would be a patient having cancer (1) and the model classifies the case as cancer (1).
- TN (True Negatives): true negatives are the cases where the true class is 0 (false) and the predicted one is also 0 (false). A situation that applies to this case would be where a person does not have cancer (0) and the model classifies it as not being cancer (0).
- FP (False Positives): false positives are the cases where the true class is 0 (false) and the predicted is 1 (true). False being where the model predicted incorrectly and positively because the predicted class was positive (1). The case that applies situation where the person does not have cancer and the model classifies it as if it did.
- FN (False Negatives): False negatives are cases where the true class was defined as 1 (true) and the predicted one as 0 (false). False being where the model incorrectly predicted and negative because the predicted class was negative (0). An applicable situation would be where a person has cancer and the model defines it as if they did not have cancer.

3 Methods

The proposed system aims to provide support in decision making to diagnose melanoma skin cancer. For this purpose, some requirements were foreseen, which will be explained in the following sections.

3.1 Project Set-Up

Materials:

The components used in the project were a Raspberry Pi 3 model B, a 3.5" TFT LCD touch screen display, a V2 camera, a pair of sinks, a 16 GB SanDisk Class 10 card, a case and a 5V3A power supply (bivolt source). They were chosen after an abstraction of objects related to the constitution of the embedded system proposed in the planned architecture. An individual practical study was carried out of each component that played an important role in the functionality as a whole.

For each component that presents implementation complexity, practical tests were carried out separately, that is, one component was tested at a time, starting with the next only after having achieved a good result in the test of the previous one.

Figure 4 shows the prototype.



Fig. 4. Protype.

Database:

The ISIC Archive project database was used for the development of the software and for carrying out the planned tests.

1642 (one thousand six hundred and forty-two) images with skin tumor and 1689 (one thousand six hundred and eighty-nine) images with benign skin spots/stains were removed, both with their metadata (diagnostics) from the database for the realization the training and validation of the algorithm. The training phase used 70% of the total base and 30% were used for tests.

3.2 Software Development and Management Methodology

For good management and software development, the Scrum agile management and development methodology was used in conjunction with the RUP, and the unified modeling language (UML).

The project was divided into thirteen iterations. Each iteration goes through the four phases of the software development process used (conception, elaboration, construction and transition).

At the beginning of the project, the team used the Kanban agile methodology together with the RUP, but with the increase in complexity, we changed to the Scrum agile methodology together with the RUP. In the beginning, Trello was used, later Azure DevOps from Microsoft, always working with sprints of fifteen days.

3.3 Image Processing

This step aims to improve the original image, removing any other elements that may be present in it, such as hair, skin and even some possible noise from the image environment.

The locked system resizes an image to a size of 204 pixels wide by 204 pixels high for reduced processing. After resizing, the system can apply a smoothing filter, aiming to reduce the number of derived images, such as hair. With a smoothed image, the system can create a grayscale copy for binaries in order to apply an opening morphological filter to reduce some noise that appear after binarization and pass an original to the YUV encoding system, with the purpose of equalizing and converting the RGB system to calculate the tone limit. With a morphological filter application, the system can use a copy of the image as a mask for an original image, resulting in an image with only one skin spot highlighted.

Figure 5 depicts the image processing structure.

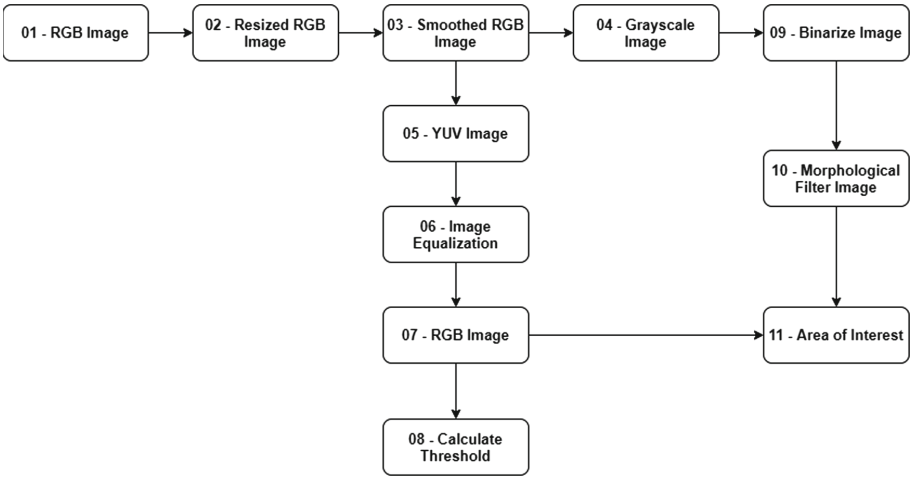


Fig. 5. Diagram of how image processing works.

Figure 6 shows an image of a spot before and after processing.

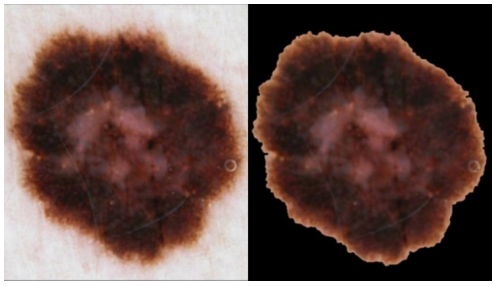


Fig. 6. Result of image processing.

Color Variation Analysis

To quantify the color variation of the skin spot, it is necessary to find the region of interest in the image resulting from the pre-processing in order to calculate the standard deviation and the number of points greater than the threshold of the histograms of each channel (RGB). This procedure represents rule C of the melanoma detection method.

Figure 7 shows the steps that were taken to develop this stage.

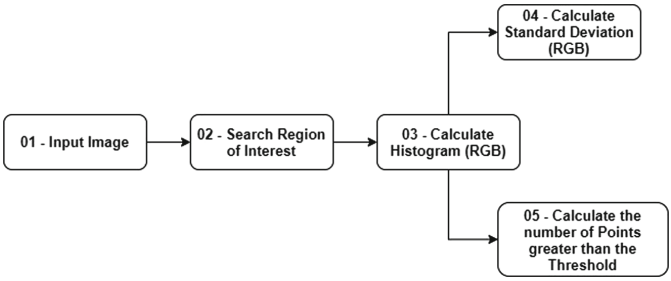


Fig. 7. Diagram of how color variation analysis works.

Table 2 lists the results found for each histogram belonging to a malignant and benign skin spot/stain.

Table 2. Standard deviation and points above the threshold in histograms

Channel	Standard deviation	Points above threshold
Benign		
Blue	353,48193	84
Green	464,20898	80
Red	145,35013	121
Malignant		
Blue	239,77455	64
Green	189,1244	78
Red	125,30375	129

If we look at the three histograms in each image and compare them, we will see that the histograms belonging to the malignant skin spot/stain are more “aggressive” than those of the benign skin spot/stain.

Edge Variation Analysis

To quantify the edge variation, it is necessary to find the region of interest in the image resulting from the processing, divide it into four parts, based on the center of the ROI, select the first part to calculate the edge in order to highlight the pixels brackets surrounded by darker pixels for a vector to result, where it will be calculated to identify the amount of local highs and lows contained in its vector from the horizontal and

vertical histograms. This procedure represents rule B of the melanoma detection method.

Figure 8 shows the steps that were taken to develop this stage.

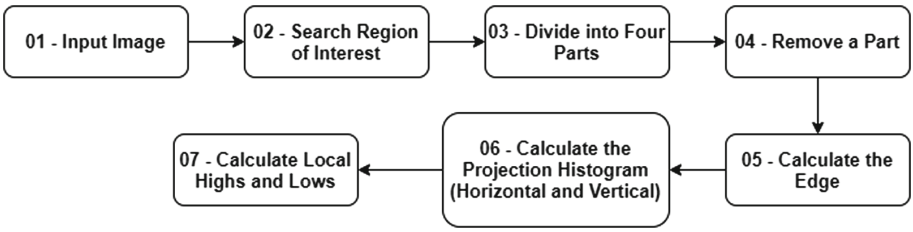


Fig. 8. Diagram of the operation of the edge variation analysis.

Table 3 lists the results obtained from each histogram belonging to a malignant and benign skin spot/stain.

Table 3. Maximum and minimum amounts of histograms

Quantities	Benign	Malignant
Maximum	236	189
Minimum	21984	22044

Diameter Analysis

For the extraction of the diameter of the skin stain, it is necessary to find the region of interest in the image resulting from the processing, so that it will be possible to find the contours of the ROI that will be used to extract the moments of the image, giving the ability to calculate some characteristics such as the center of the skin stain, the area, the radius and consequently the diameter. This procedure represents rule D of the melanoma detection method. Figure 9 shows the steps necessary to develop this step.

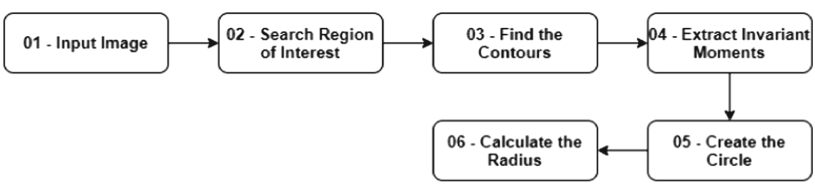


Fig. 9. Diameter analysis diagram.

Figure 10 shows the result obtained from the set of operations performed on the benign image, on the left side, and on the malignant image, on the right side.

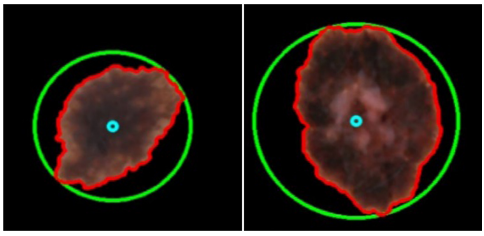


Fig. 10. Circumference of benign and malignant skin spot.

Table 4 lists the results obtained.

Table 4. Result of the diameter analysis

Benign (mm)	Malignant (mm)
5,1	6,7

Asymmetry Analysis

This step aims to compare two halves of the image, for this it is necessary to find the ROI and divide it in half, selecting the midpoint of the image width. This procedure represents rule A of the melanoma detection method.

Figure 11 shows the steps that were taken to develop this procedure.

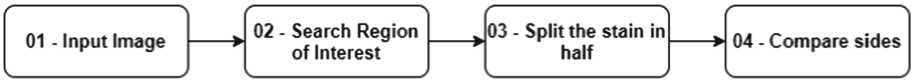


Fig. 11. Asymmetry analysis process.

These were the results obtained:

- **Benign:** 1,22036539; 0,76417758; −0,52827106; −0,16928178; 1,65588505; 0,07144739; −0,0779571;
- **Malignant:** −1,94320367; −1,24052967; −1,00843259; −1,6367098; −1,47681334; 0,72248248; −0,07596904;

The results described show us that the malignant skin spot/stain has more negative moments than the benign one and that the negatives of the benign skin spot/stain do not reach −1, while 5 of the 6 negative skin spots/malignant skin stain exceed −1.

Recognition of Standards

Standards are understood to mean properties that make it possible to group similar objects within a given class or category, through the interpretation of input data, which allow the extraction of the relevant characteristics of these objects.

This step aims to extract the characteristics of the image resulting from the processing. These characteristics are usually grouped into a scalar vector, called an image descriptor.

Figure 12 illustrates the steps that were necessary to develop the extraction of the patterns.



Fig. 12. HOG extraction process.

Figure 13 shows the result obtained in a malignant skin spot. The HOG managed to extract as many characteristics of the skin spots as possible, but the side effect was the increase in computational cost. However, the increase in confidence compensates for the loss of speed and the increase in cost.

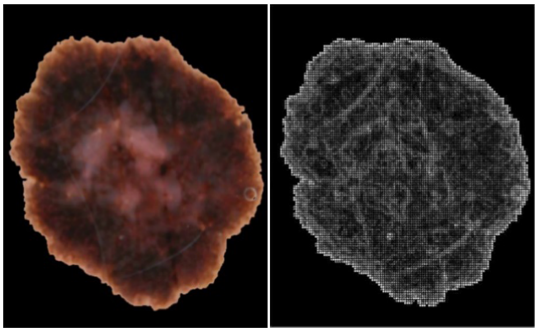


Fig. 13. HOG of the malignant spot.

Controlled Environment

The image acquisitions were performed in a controlled environment, so that a more accurate analysis of the risk scale classification algorithm for skin cancer tumor is possible.

Figure 14 shows the environment prepared for the tests.

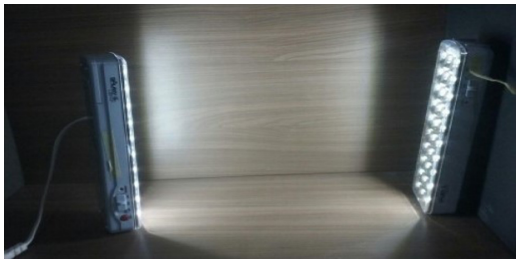


Fig. 14. Test environment.

4 Results

With the result returned by the Oriented Gradients Histogram method, it was possible to obtain a vector that represents the necessary characteristics for the performance of the probabilistic calculation, image classification and the degree of risk of the skin spot/stain being a skin tumor. For the validation of the classification algorithms, the confusion matrix was used.

Table 5 shows the amounts of VP (True Positive), VN (True Negative), FP (False Positive) and FN (False Negative) of the nine classifiers.

Table 5. Confusion matrix classifiers

Classifier	VP	VN	FP	FN
Ada Boost Classifier	57%	70%	30%	43%
Decision Tree Classifier	64%	60%	40%	36%
Gaussian NB	74%	44%	56%	26%
Gaussian Process Classifier	60%	71%	29%	40%
KNeighbors Classifier	73%	50%	50%	27%
Quadratic Discriminant Analysis	17%	85%	15%	83%
MLPClassifier	70%	60%	40%	30%
Random Forest Classifier	61%	71%	29%	39%
SVC	67%	62%	38%	33%

As can be seen in Table 5, the classifiers that had the highest indexes of true positives (skin cancer patient and skin cancer model) with 74%, 73% and 70% were Gaussian NB, KNeighbors Classifier and MLPClassifier, respectively, but at the start they were the ones with the lowest cancer incidence rates (unused patient or skin cancer and model classified as not skin cancer) with 44%, 50% and 60%, respectively. This means that they were the best at identifying malignant skin spots, but were not as good at identifying benign skin spots.

The classifiers that had the highest rate of true negatives (patient does not have skin cancer and the model classifies it as not skin cancer) with 85%, 71% and 71% were Quadratic Discriminant Analysis, Random Forest Classifier and Gaussian Process

Classifier, respectively. Even though Quadratic Discriminant Analysis was the best at detecting benign skin spots, it was the worst, with 17%, at detecting malignant skin spots, and, consequently, having the highest balance of false negatives.

Decision Tree Classifier was the classifier that obtained the smallest difference between true positives and false positives with 4%, SVC was the second with 5% difference between VP and FP and the third parties were Random Forest Classifier and MLPClassifier with 10% difference, each. However, the classifiers that managed to reach or exceed 70% of VP or FP and maintain a small difference, between 10% and 11%, were the Random Forest Classifier, Gaussian Process Classifier and MLPClassifier classifiers.

5 Discussion

When we are working with health, what we should take into account in the confusion matrix is the FN column (patient has skin cancer and the model classifies it as not being skin cancer), because it is better to refer the patient who does not have skin cancer for a battery of tests that will prove the inexistence, than making the mistake and discharging a patient with skin cancer. Knowing this, the two algorithms that obtained the lowest false negative rates were the Gaussian NB and KNeighbors Classifier.

With the two classifiers, Gaussian NB and KNeighbors Classifier, implemented in the risk grade recognition and classification algorithm, it was possible to perform the classification of the spots.

The spots are classified into four levels, zero being indeterminate, based on the likelihood that the skin spot is melanoma. The levels were based on BI-RADS.

Table 6 shows the possible responses returned by the risk scale classification algorithm and recognition of melanoma skin cancer.

Table 6. Levels of the algorithm

Category	Classification	Probability	Conduct
1	Very low risk	$\geq 0\%$ and $\leq 25\%$	It is advisable for the patient to continue doing the monitoring annually
2	Low Risk	>25% and $\leq 50\%$	It is advisable that the patient be referred to dermatoscopy
3	Medium Risk	>50% and $\leq 75\%$	It is advisable that the patient is referred to the confocal microscopy
4	High Risk	>75% and $\leq 100\%$	It is advisable to collect a little of the patient's tissue for a biopsy

Table 7 shows the results obtained in each analysis, where the first column is the results of the image without the melanoma and the second the results of the image with the melanoma. The first line is the result of the Gaussian NB algorithm and the second is the result of the KNeighbors Classifier.

Table 7. Results of the analysis

Algorithm	Benign	Malignant
Gaussian NB	The skin spot analyzed has a low risk level with 49,086% veracity of being melanoma! Is advisable that the patient be forwarded to dermatoscopy	The analyzed spot has a medium risk degree with 69.38% veracity of being melanoma! Is advisable that the patient be referred for confocal microscopy
KNeighbors Classifier	The skin spot analyzed has a low risk level with 43,148% veracity of being melanoma! It is advisable that the patient be referred for dermatoscopy	The analyzed skin spot has a medium risk degree with 74.98% veracity of being melanoma! Is advisable that the patient be forwarded to confocal microscopy

As can be seen in Table 7, in addition to the algorithm returning to the user the degree of risk of and the percentage of veracity, the necessary exam is returned to prove that the tissue is cancerous. If the skin spot has a low degree of veracity, the device recommends an examination, as only a specialist has the ability to state this proposition.

6 Conclusions

In view of the large number of skin cancer cases, several researchers have been trying to develop techniques to improve and speed up the diagnosis, since, when detected early, the chances of curing this disease increase considerably. One of the ways to make the diagnosis is through dermatoscopy. In this technique, the doctor has the help of a dermatoscopy to analyze the lesions based on some characteristics. Despite analyzing the injury in a broader way, and the extracted characteristics are well-founded, this diagnosis is subjective, as it is affected by some factors. Thus, the utility of computational analysis has been researched in helping professionals to carry out this type of diagnosis.

In order to improve the accuracy of this diagnosis, the development of this study enabled the research and development of a device capable of providing statistical support to melanoma skin cancer specialists using a classification and recognition algorithm. In addition, it also allowed an analysis of how it can improve the reliability and accuracy of diagnoses performed in hospital environments, and an assessment of the data treatment process and of some classifiers available to perform such a task.

The initial proposal was to use one analysis, instead of two, in the classification and recognition algorithm. However, with studies and research it was found the need to use two analyzes and, consequently, two classifiers. In view of this modification, it was implemented, along with the ABCDE rule, an oriented gradient histogram technique, which aims to detect, describe and recognize patterns and characteristics.

When doing the tests in a controlled environment and with some images of the test base, it was found that the need for a good camera and capturing an image of the skin

spot are fundamental for the performance of the processing and consequently of the analysis. However, even with a low-quality image, the device was able to analyze and classify it as expected. Thus, allowing the proposed objectives to be really achieved.

The agile process was a key contributor to the completion of this MVP, because with the thirteen iterations, we are able to check every fifteen days if we were heading in the right direction and, otherwise, it is pivoting.

Acknowledgements. This research is sponsored by the Portugal Incentive System for Research and Technological Development PEst-UID/CEC/00319/2020 and University Paulista – Software Engineering Research Group by Brazil.

References

1. American Cancer Society. About Basal and Squamous Cell Skin Cancer. https://www.cancer.org/cancer/basal-and-squamous-cell-skin-cancer.html. Accessed 29 Feb 2020
2. American Cancer Society. About Melanoma Skin Cancer, https://www.cancer.org/cancer/melanoma-skin-cancer.html. Accessed 29 Feb 2020
3. Globo Can. Melanoma of Skin, International Agency for Research on Cancer, http://gco.iarc.fr/today/data/factsheets/cancers/16-Melanoma-of-skin-fact-sheet.pdf. Accessed 29 Feb 2020
4. Globo Can. Non-melanoma Skin Cancer. International Agency for Research on Cancer. http://gco.iarc.fr/today/data/factsheets/cancers/17-Non-melanoma-skin-cancer-fact-sheet.pdf. Accessed 29 Feb 2020
5. Globo Can. Cancer Tomorrow. International Agency for Research on Cancer. http://gco.iarc.fr/tomorrow/. Accessed 29 Feb 2020
6. Stewart, B.W., Wild, C.P.: World Cancer Report: 2014, 1st edn. IARC, Lyon (2014)
7. Desai, A., Lee, D.J., Wilson, C.: Using affine features for an efficient binary feature descriptor. In: 2014 Southwest Symposium on Image Analysis and Interpretation (SSIAI), pp. 49–52. IEEE, San Diego (2014)
8. Panceri, J.A.C., Pinto, L.A., Pereira, F.G., Cavalieri, D.C., Komati, K.S.: Facial recognition based on HOG and PCA: an invariance to illumination-based comparison. Ifes Ciência **1**, 41–62 (2015)
9. Costa, C.S.: Epidemiologia do câncer de pele no Brasil e evidências sobre sua prevenção. Diagn Tratamento **4**, 206–208 (2012)
10. José Alencar Gomes da Silva National Cancer Institute – INCA, Câncer de pele melanoma. https://www.inca.gov.br/tipos-de-cancer/cancer-de-pele-melanoma. Accessed 29 Feb 2020
11. José Alencar Gomes da Silva National Cancer Institute – INCA. Câncer pele não melanoma. https://www.inca.gov.br/tipos-de-cancer/cancer-de-pele-nao-melanoma. Accessed 29 Feb 2020
12. Oncoguia Institute. Sobre o Câncer de Pele Basocelular e Espinocelular. http://www.oncoguia.org.br/conteudo/sobre-o-cancer/751/146/. Accessed 29 Feb 2020
13. ISIC Archive. https://isic-archive.com/. Accessed 29 Feb 2020
14. Teixeira, M.B.R.: Avaliação dos achados mamográficos classificados na categoria 4 do sistema BI-RADS® e sua correlação histopatológica. Master's thesis, Botucatu Medical School (2011)
15. Mitchell, T.M.: Machine Learning, 1st edn. McGraw-Hill Science, New York (1997)