

Automatic Classification of Images with Skin Cancer Using Artificial Intelligence

Israel Gaytán Campos¹, Wendy Morales Castro², Belem Priego Sánchez³,
Efrén Fitz Rodríguez¹, Rafael Guzmán Cabrera²

¹ Universidad Autónoma Chapingo,
Posgrado de Ingeniería Agrícola y Uso Integral del Agua,
México

² Universidad de Guanajuato,
Departamento de Estudios Multidisciplinarios,
México

³ Universidad Autónoma Metropolitana,
División de Ciencias Básicas e Ingeniería,
México

al19127213@chapingo.mx, {w.moralescastro, guzmanc}@ugto.mx, abps@azc.uam.mx,
efitzr@taurus.chapingo.mx

Abstract. Skin cancer is the most frequent malignant neoplasm in the world, it is a public health problem, which has increased in recent years due to environmental changes, different lifestyles, sun exposure, among others. One way to detect skin cancer is through the analysis of medical images, the analysis of these images can allow the detection of any abnormality. In this work, several block programming models with classifiers for the recognition of medical images of skin cancer are implemented. Pre-processing, manipulation, and machine vision to extract relevant features from images are the starting point for obtaining proper classification values. The main objective of this project is to carry out the analysis of a set of classification techniques, as well as verify that the combination of image processing operations and classification tools provide better performance compared to the classification values of the original images. Images of three types of skin cancer were used: melanocytic nevus, melanoma, and benign keratosis lesions. Each category contains 200 images. The images were subjected to a set of filters, to later use different classification algorithms. Six types of filters and five different classification techniques were implemented. The results obtained allow us to see the viability of the proposed method.

Keywords: Skin cancer, image processing, machine learning, artificial intelligence.

1. Introduction

Currently digital image processing is used in different disciplines [8]. Image processing includes a set of techniques applied to an image to identify, characterize, classify, and reconstruct images of objects, people, phenomena or signs, highlighting information of interest for further analysis [19]. In recent years it has had a satisfactory application, this is due to the development of new very useful tools, among which are medical diagnoses [12], space exploration [28], computer vision [22] among others. In medical diagnostics, image processing has been used to detect anomalies such as skin cancer, as is the specific case presented in this work.

The detection of severity in skin cancer lesions establishes a threshold of opportunity for the correct care and treatment of this disease [24]. Skin cancer represents one of the causes of death with the highest incidence in recent years worldwide. Statistically, worldwide cases of diagnoses in certain types of skin cancer are on the rise, as reported each year approximately 5.4 million new cases of skin cancer are registered in

the United States alone, despite the alarming figures that this represents, early detection of skin abnormalities shows a high survival rate [4,20]. Therefore, being able to detect skin cancer at an early stage becomes a task that must be performed accurately when diagnosing skin conditions. The rigorous inspection carried out by a specialist dermatologist when making a diagnosis would be time consuming in advancing to later stages of the disease. Accurate diagnoses often have a certain degree of subjectivity depending on the experience and criteria of the doctor in question. With few trained dermatologists available globally in public health care, accuracy is less than 90% in accurately diagnosing skin cancer [10].

Machine learning involves the application and development of algorithms that enable the conception and management of data. Currently, there is a wide variety of studies related to the detection of skin cancer that include machine learning, in synergy with traditional image classification techniques as tools in the determination of medical diagnoses of this disease. The use of algorithms such as decision trees (DT), support vector machines (SVM), Bayesian classifiers (NB), artificial neural network (ANN) architectures, and a growing number of novel approaches based on artificial intelligence, represent a highly viable option for the assistance of specialist dermatologists and obvious reductions in time and even better performance in skin cancer diagnoses. The main obstacle to the implementation of image classifier models is the obvious need for sufficient knowledge of programming in software and computer systems, seeming at first glance an inconvenience for dermatologists.

This document proposes a block programming model for the development of a skin cancer image classifier that supports its identification. In this way, image classification algorithms are implemented in conjunction with high-performance neural network models for the features extraction from images, thus allowing technical assistance to dermatologists in the diagnosis of this disease.

The proposed approach would be an advantage for users since it does not require high computer of programming knowledge, thus

reducing the time required for a specialized dermatological diagnosis.

In this way, the models used are shown, as well as the different operations carried out in the search to increase the efficiency of the proposed classifier. The best results and the type of operation performed on each type of image are shown. The proposed methodology for the development of the classifier is the evaluation of the types of class sets, which can be combined with each other to determine which is the simplest and most difficult category to classify by the algorithms. This is presented [1], which showed adequate performance classifying biological type images for quality control in wood planks.

1.1 Algorithms and Techniques

The algorithms used in this paper are presented below.

Decision trees. Decision trees are a supervised nonparametric classification method used for classification and regression learning. The goal is to create a model that predicts the value of a target variable by learning the simple decision rules inferred from the data features [13].

k-Nearest neighbors. The principle behind the nearest neighbor methods is to find a predefined number of training samples closest in distance to the new point and predict the label of these. The number of samples can be a user-defined constant (k-nearest neighbor learning) or vary according to the local density of the points (based on the neighbor learning radius). The distance can, in general, be any metric measurement, the standard Euclidean distance being the most common choice [25].

Naive Bayes. The Naive Bayes family of algorithms is based on the Bayes Theorem, which says the following: Let be a set of exclusive and exhaustive events, such that the probability of each of them is different from zero $\{A_1, A_2, \dots, A_n\}$ [18]. Let be any event for which the conditional probabilities are known. Then, the probability is given by the following expression:

$$P(A_i|B) = \frac{P(B|A_i) \times P(A_i)}{P(B)}.$$

Support vector machine. Support vector machine are a group of supervised learning algorithms used for classification, regression, and outlier detection [17, 26]. A hyperplane, called a support vector, separates them as far as possible based on their class. In this way, the vector determines the boundary that serves to classify a new element, so that depending on which part of the space it belongs to, it will be assigned one class or another. This type of algorithm has a series of parameters that allow adjusting its internal configuration and thus optimizing the results during the classification process.

Artificial neural networks. A neural network is a computer model whose layered structure resembles the interconnected structure of neurons in the brain, with layers of connected nodes [27]. A neural network can learn from data, so it can be trained to recognize patterns, classify data, and forecast future events. Neural networks break down inputs into layers of abstraction. Their behavior is defined by how their individual elements are connected, as well as by the importance or weighting of those connections. These weightings are automatically adjusted during training according to a specified learning rule until the neural network performs the desired task correctly. Neural networks that operate on two or three layers of connected neurons are known as shallow neural networks.

2 Related Work

The work described in [11] presents the types of medical images that can be obtained with current technology, which are obtained in analog format, so it is necessary to digitize them for processing. The authors present the best-known ones, such as radiographs, magnetic resonance images, conventional photography, among others. They also present how they are stored and visualized. In medicine, images are used to make diagnoses, plan a form of treatment, and monitor changes over time with respect to some pathology. Image processing is performed to help extract or highlight regions of interest to the physician and to support the classification of a set of images. To perform this processing, it is necessary that the images go through a preprocessing, where they are adapted

according to the processing to be performed. The authors also mention some commonly used preprocessing, such as the improvement and understanding of digital images. These types of preprocessing are useful for the realization of systems, platforms or classifiers developed to support medical specialists, an example of a platform is found in [7] where the authors present an open-source platform that allows to quickly generate prototypes for the processing of medical images with sophisticated visualization mechanisms and user interaction.

The project is built on a flexible cross-platform framework and is developed in C++ which guarantees the integration of heterogeneous C++ modules, and which provides the user with a series of high-level component libraries for the construction of graphical user interfaces, including file management, visualization, interaction and data processing. The authors make use of the platform within research that addresses the automatic detection of coronary artery lesions. In the same sense, systems that allow image processing have been developed, as in the case of [15] the purpose of the work is to find objects of interest in medical images, by detecting edges in images.

Within [16] we found the use of synergy of image processing, pattern recognition, and artificial intelligence to help detect clusters of microcalcifications in digitized images of mammograms. There are other techniques to perform image processing as is the case presented in [14] in which the authors used Raman spectroscopy technique to obtain spectral maps with specific spatial resolution (1 to 5 micrometers) on a selected region of the sample to access and visualize relevant information about the spatial distribution in any sample about its biochemical composition.

Another research found on image processing is [6] which like [16] helps the early diagnosis of breast cancer using image processing, within this research we can find that it uses the technique of segmentation by texture, the images used as a test are from a database, which has images of carcinogenic masses and micro calcifications manually labeled by experts. In [5] they carry out the identification of breast cancer using thermal imaging, performs a digital processing of the

images, using a texture analysis of the images to identify and extract all regions of interest this because thermography is a non-invasive evaluation of physiological functions without involving exposure to radiation and breast compression.

In recent years, there have been numerous works that address the issue of skin cancer detection using computational algorithms and artificial intelligence applications. The analysis of images within these investigations takes a central role because the diagnoses are obtained by classifying the photographs of the dermal incidences of the patients.

The great diversity of this research includes the implementation of simple classification algorithms for machine learning, as is the case of the work in [3], which presents two hybrid techniques for the classification of skin images to predict the existence of cancer. The proposed hybrid techniques consist of three stages, which are: feature extraction, dimensionality reduction and classification. The classifiers used in this research are based on the algorithm of the nearest k-neighbors (k-NN) with a rating performance of 97.5% and an artificial neural network structure of backpropagation and feedback with 95% success in the classification.

On the other hand, in [23] the diagnosis of skin cancer is carried out with a methodology called "intelligent" based on dermoscopy images (direct microscopy of the skin), with a novel variant of the optimization algorithm for evolutionary computation of particle swarm (PSO). This algorithm is used in this work to optimize the differentiation of characteristics between benign cancer and that which is of malignant type.

Despite the different approaches with which the task of classifying images can be approached, specifically images with skin problems, convolution neural networks (CNN) are the artificial neural network structures that are most used obtaining fairly accurate results, in some cases improving the diagnoses of dermatologists specialized in the detection of skin cancer. Derived from the great popularity in the use of deep learning (DL) techniques, convolutional neural networks satisfy in wide spectrum the approach of detection of skin cancer by image classification, [4] and [20] are a clear example of this, since using a deep learning

approach using various CNN, they proved the effectiveness of this type of structures when classifying several types of dermatological diseases. This type of analysis is noticeable since in this work the average results showed that all deep learning models outperformed dermatologists. Above all, it should be noted that the results obtained for various convolutional neural network models exceed 90% effectiveness in their performance.

Similarly, in [24] the authors obtain satisfactory results, emphasizing the high performance that diagnosis through the use of CNN and deep learning, from which activation maps can be obtained from classes with which dermatology professionals can increase the accuracy of diagnoses.

Another related work is presented in [9] in which it is proposed the implementation of a neural network of convolutional type more sophisticated compared to the previous cases, opting for the use of a region-based CNN, which is faster to train without sacrificing its performance to a large extent. This model obtained satisfactory results for different classification variants and types of disease, surpassing the diagnoses of specialists.

From the new computational tools, the deep learning approach begins to take certain aspects in terms of the methodology that is followed. In this sense in [21], the research was directed at the task of classifying skin cancer images using convolutional neural networks in conjunction with deep learning, with the remarkable implementation of a transfer learning architecture, obtaining values greater than 90% in cancer detection and values greater than the approaches of this type previously made.

Moreover, in [10] was implemented deep learning models in the cloud, showing the trend that deep learning techniques are beginning to target to improve their performance computationally speaking. In the aforementioned research, the results obtained in classification performance were higher than 95% showing a robust and efficient model. Beyond the high values in classification performance that the CNN models, the complexity of their architecture compromises the computational cost in which the complete processing of the proposed system is carried out.

The work mentioned in [1] makes it clear that the effectiveness of convolutional neural networks with the deep learning approach can be achieved by implementing traditional machine learning classification techniques and some image processing operations. The extraction of texture characteristics means a suitable approach to the classification of biological images as is the case of the aforementioned work.

In this document, we show the methodology for the classification of three types of skin cancer, comparing the results obtained from classifying the images without any operation against images to which filters and operations have been applied. For fast processing, traditional machine learning classification algorithms are used, as well as CNN models for image feature extraction.

3 Methodology

This section describes the database used and how it was used, as well as the workflow followed and each component that makes it up.

In the same way, the main parameters with which each algorithm used in the classification of the images was implemented are mentioned.

3.1 Database

The images used were extracted from a database called Skin Cancer MNIST: HAM10000, which is a large collection of multi-resource dermoscopy images of pigmented lesions, with 7 different classes (actinic keratoses and intraepithelial carcinoma / Bowen's disease, basal cell carcinoma, benign lesions similar to keratosis, dermatofibroma, melanoma, melanocytic nevi and vascular lesions, of which 3 were taken (melanocytic nevi (Nv2), melanoma (Mel2) and benign lesions similar to keratosis (Bk2)), in total 600 images were counted, 200 for each class. To simplify the workflow, each class was renamed Bk2-(A), Nv2-(B), and Mel2-(C).

The three classes were grouped into a directory for each classification category, thus assigning a label to the image of each subdirectory. Image processing filters were implemented in MATLAB, and each image was treated similarly, the development of the filters was applied over the

general directory of the given extension. Thus, the operating structure of the classifier was as follows:

- Search for the corpus.
- Application of filters.
- Construction of the image recognition model.
- Definition of the classification scenario.
- Application of the classification methods.
- Definition of the evaluation metrics, construction of the confusion matrix and evaluation.

The workflow that was implemented for the realization of the work begins with the search for the corpus that will be used, continuing with the application of filters such as canny, grayscale, among others that are mentioned throughout the work, once the filters were applied to the images, the image recognition models are presented. To continue with the scenarios of classification, cross-validation and training and test set, both scenarios are methods to perform the evaluation of the implemented algorithms, both scenarios have their specific way of separating the set of images into two parts, one of them was taken for training and the other for its part was taken for testing. Once the scenarios have been created, the application of the classification algorithms mentioned in section 1.1 is shown, to finally evaluate the performance of these algorithms with the evaluation metrics and the confusion matrix.

3.2 Operations on Images

A total of 3 processing operations were used on the original images in the following order: grayscale, canny operator and binarization of images, as shown in Figure 1.

In grayscale operation the image should be converted to a more appropriate format to extract certain important features. It was considered that the most relevant of the images to be treated was the outline and the edges, to find the outline as well as the edges, it is not necessary to consider the color of the images, so it is more convenient to convert the image to a grayscale. In the RGB model, an intensity is assigned to each of the three primary colors of light: red, green, and blue, in this way each pixel in a photograph is represented by a value that identifies the intensity of each of these

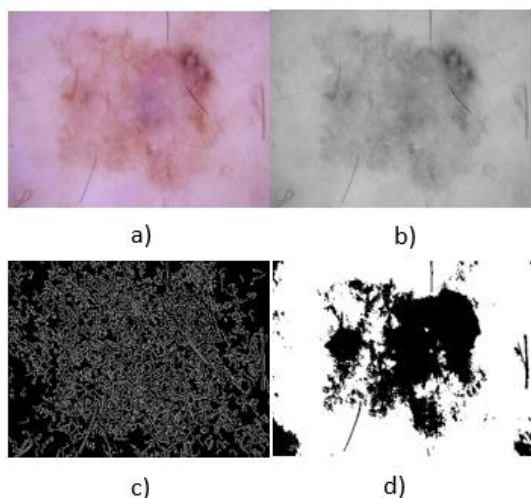


Fig. 1. Operations performed on the images. (a) original image; (b) grayscale image; (c) canny operator; (d) binarized image

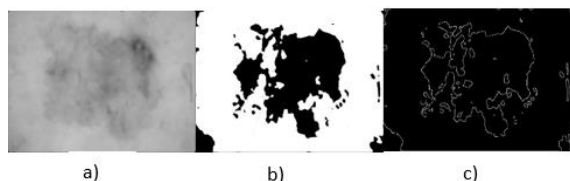


Fig. 2. Additional operations performed on the images. (a) the median filter; (b) filter image of the binarized median; (c) canny operator on the binarized image of the median filter

three colors that, mixed by addition, is closer to the true color of the pixel. In a grayscale each pixel is represented with a single value between 0 and 255, which represents how dark it is. Being a weighted average the sum of the three colors must be equal to 1. The weights used are as follows: $EG = 0.299 R + 0.587 G + 0.114 B$; where R is the channel of the color red, G the color green and B the color blue [2].

A very efficient algorithm for edge detection is the Canny operator, which is a multi-stage algorithm for detecting a wide range of edges in images. Canny's algorithm goes through two stages: reducing noise and finding the intensity of the image gradient.

Once the grayscale image was obtained, a binarization was made. Binarization is a process that transforms an image that is in grayscale to an image in two colors: black and white. This is useful

for recognizing the objects in the image. In this way, converting the images to black and white loses very little relevant information and simplifies the representation of each pixel, going from 255 possible values to only 2.

Additionally, on the gray scale images, a statistical filter type was applied, in this case the median filter. The filter of the median belongs to a special class of filters, the statistics, which in addition to being nonlinear base their action on some operation of the statistical type. It is a nonlinear spatial filter that allows you to remove unwanted artifacts and structures in the image without significantly affecting the edges. It does not produce or calculate a new value to replace a pixel in the image, rather it is an existing value in the region of interest of the filter produced by the data realignment.

On the images filtered with the median, a binarization of the images was reapplied and on these last ones the canny operator was applied to recover the contours present in the image, as shown in Figure 2.

3.3 Classification Categories

As mentioned above, cancers were renamed from benign lesions (Bk2), melanocytic nevi (Nv2) and melanoma (Mel2) to categories A, B and C respectively; thus establishing 4 configurations for each classification model, in this way, considering classes A, B and C, the possible configurations of each classifier were implemented as follows, A-B-C, A-BC, B-AC, and C-AB. This was carried out to observe the performance of each classification algorithm with respect to each class assigned on the types of injuries; in this way, making a comparison of each configuration will yield information about the possible difficulty that may arise when trying to assign a medical diagnosis offering possible points of view for rapid decision making.

The first step was to obtain the reference values (Baseline) for the classifiers and all categories with the original images without any image processing treatment or filtering. Once the images were imported from the directory where the three classes were grouped, the characteristics of the images were extracted and grouped. This was done with orange's Image Embedding application.

3.4 Feature Extractors

Image Embedding is the Orange Data mining application which reads images and uploads them to a remote server or evaluates them locally. The models that this application has are models of deep learning convolutional neural networks (CNN) which are used to calculate a vector of characteristics of each image. The application returns an enhanced data table with additional columns (image descriptors).

Orange has several types of feature extractors also called embedders. **InceptionV3** is Google's deep neural network for image recognition. It is trained on the ImageNet dataset. **SqueezeNet** is a deep model for image recognition that achieves AlexNet-level accuracy on ImageNet with 50 times fewer parameters. The model is trained on the ImageNet dataset. **VGG16** and **VGG19** are deep neural networks for image recognition proposed by the Visual Geometry Group at the University of Oxford.

They are trained in the ImageNet dataset. In Orange, these neural networks are used with the original weights. As an embedder, use activations from the penultimate layer. Image Embedding also includes Painters, an embedding model trained on 79,433 images of paintings by 1,584 painters and won Kaggle's Painter by Numbers contest. Activations of the penultimate layer of the network are used as embeddings. **DeepLoc** is a convolutional network trained on 21,882 images of individual cells that were manually assigned to one of the 15 location compartments.

To verify the influence of the embedders on the performance of each classifier, each feature extractor was evaluated in combination with 5 different learning algorithms in two classification scenarios (cross-validation and training and testing set), and the original images (Baseline). The best performing embedder was chosen for all classifiers and all categories.

3.5 Classification Algorithms

To evaluate each configuration of the proposed classifiers, the machine learning algorithms described in section 1.1 were used. In this way, establishing certain general parameters, the algorithms used were Decision trees, k-Nearest

neighbors, Naive Bayes classifier, Support Vector Machine and Artificial neural networks in a multilayer perceptron structure (MLP).

The decision tree algorithm is a simple algorithm that divides data into nodes in relation to the assigned class. In Orange this algorithm can handle both categorical and numerical data sets. As general parameters, this algorithm used a binary tree induction with which the data is divided into two child nodes; number of instances in sheets of at least 3; with a division for the nodes of at least 5 instances and with a maximum depth of 100 thus specifying the levels of each node. A data division threshold of 95% was maintained

The kNN algorithm that looks for k closest training examples in the feature space and uses its average as a prediction. This algorithm was implemented with a number of closest neighbors of 5 maintaining a Euclidean metric, that is, a straight line, and a uniform configuration for the weights, with this all the points of each neighborhood were weighted equally.

Naive Bayes learns a Naive Bayesian model from the data. Works for sorting tasks. Using the default preprocessing the algorithm was run in the following order: first empty columns are removed from the input data and then numerical values are discretized in 4 containers with the same frequency.

The support vector machine (SVM) is a machine learning technique that separates the attribute space with a hyperplane, thus maximizing the margin between instances of different classes or class values. The technique often produces supreme predictive performance results. Orange incorporates a popular SVM implementation of the LIBSVM package. This widget is a graphical user interface. In this way, this algorithm used a loss penalty value (cost) of 1.4; the distance from the true values was assigned with a value of 0.10 for epsilon; the attribute function was selected as a sigmoid function with a default value of the constant g and o. Numerical tolerance of 0.0010 and a maximum of 100 iterations were used.

The neural network algorithm uses SKLearn's multilayer perceptron algorithm with retro-propagation and can learn linear and nonlinear models. The number of neurons per hidden layer used was 100 with a logistic sigmoid activation function for the hidden layer. The selected solver

Table 1. Results for the different Orange inslators

Learning Technique.	Inslator	Cross Validation	Training Sets
SVM	SqueezeNet	0.615	0.633
	Inception v3	0.658	0.654
	VGG-16	0.639	0.658
	VGG-19	0.649	0.675
	Painters	<u>0.764</u>	<u>0.743</u>
	DeepLoc	0.531	0.566
K-NN	SqueezeNet	0.634	0.621
	Inception v3	0.587	0.603
	VGG-16	0.577	0.594
	VGG-19	0.638	0.632
	Painters	<u>0.68</u>	<u>0.705</u>
	DeepLoc	0.573	0.615
DT	SqueezeNet	0.559	0.55
	Inception v3	0.541	0.512
	VGG-16	0.568	0.514
	VGG-19	0.529	0.532
	Painters	<u>0.594</u>	<u>0.627</u>
	DeepLoc	0.472	0.514
NB	SqueezeNet	0.664	0.675
	Inception v3	0.601	0.6
	VGG-16	0.626	0.623
	VGG-19	0.638	0.646
	Painters	<u>0.691</u>	<u>0.69</u>
	DeepLoc	0.544	0.523
ANN	SqueezeNet	0.69	0.675
	Inception v3	0.639	0.642
	VGG-16	0.685	0.699
	VGG-19	0.664	0.69
	Painters	<u>0.751</u>	<u>0.744</u>
	DeepLoc	0.553	0.545

was the stochastic gradient descent (SGD) with a maximum of 200 iterations.

4 Results

Performance results were analyzed for each embedded using the original images and the A-B-C classification category. The results are presented for the classifiers implemented for all the categories proposed.

Embedders. For the 5 classification techniques, the performance results are shown in Table 1. The images were not treated with filters. Since in 3 of the 5 classification algorithms the Painters embedder showed better performance in the evaluation metrics, it was the feature extractor for the rest of the classifiers. For the testing of the classifiers, during the evaluation stage, 2 classification scenarios were proposed, cross-validation with 10 folds for data training; and

Table 2. Precision for each category A-B-C classification algorithm and scenario

Classification Scenario	Image Type	SVM	K-NN	DT	NB	ANN
Training Sets 80-20	Original	0.743	0.705	0.627	0.69	0.744
	Gray Scale	0.669	0.601	0.538	0.615	0.685
	Canny Operator	0.513	0.561	0.484	0.518	0.611
	Binary	0.572	0.52	0.471	0.497	0.574
	Median Filter	0.629	0.583	0.521	0.578	0.649
	Binary and Median Filter	0.559	0.527	0.472	0.528	0.611
	Canny Operator and Median Filter	0.513	0.561	0.484	0.518	0.611
Cross Validation	Original	0.764	0.68	0.594	0.691	0.751
	Gray Scale	0.669	0.612	0.513	0.617	0.688
	Canny Operator	0.442	0.464	0.447	0.442	0.54
	Binary	0.579	0.555	0.505	0.504	0.601
	Median Filter	0.633	0.611	0.52	0.603	0.687
	Binary and Median Filter	0.562	0.535	0.457	0.522	0.6
	Canny Operator and Median Filter	0.505	0.526	0.48	0.519	0.585

Table 3. Precision for each category A-BC classification algorithm and scenario

Classification Scenario	Image Type	SVM	K-NN	DT	NB	ANN
Training Sets 80-20	Original	0.779	0.788	0.704	0.775	0.791
	Gray Scale	0.741	0.71	0.667	0.711	0.771
	Canny Operator	0.583	0.603	0.594	0.566	0.671
	Binary	0.683	0.717	0.666	0.696	0.756
	Median Filter	0.74	0.71	0.707	0.727	0.771
	Binary and Median Filter	0.665	0.711	0.677	0.704	0.75
	Canny Operator and Median Filter	0.618	0.734	0.653	0.744	0.73
Cross Validation	Original	0.761	0.775	0.714	0.778	0.813
	Gray Scale	0.706	0.739	0.651	0.719	0.779
	Canny Operator	0.603	0.609	0.588	0.623	0.678
	Binary	0.642	0.751	0.669	0.702	0.759
	Median Filter	0.695	0.715	0.677	0.716	0.785
	Binary and Median Filter	0.66	0.734	0.657	0.711	0.75
	Canny Operator and Median Filter	0.652	0.758	0.627	0.71	0.754

training and testing sets with an 80-20 distribution (80% of the data for training and 20% for testing).

Classifiers for category A-B-C. The results of the A-B-C set are presented in Table 2, comparing the results of the original images against the filters used.

It was observed that for both classification scenarios the value with greater precision is obtained from the original images; being the cross-

validation performance together with the SVM algorithm the best value obtained with an accuracy of 76.4% followed by the value of the training set scenario along with the neural network architecture which obtains an accuracy of 74.4%.

Classifiers for category A-BC. As for the previous category, the original image set presents the best ranking performance values, however, this time the highest values were obtained with the

Table 4. Accuracy for each category B-AC classification algorithm and scenario

Classification Scenario	Image Type	SVM	K-NN	DT	NB	ANN
Training Sets 80-20	Original	0.887	0.855	0.793	0.837	0.877
	Gray Scale	0.783	0.796	0.709	0.77	0.819
	Canny Operator	0.654	0.696	0.657	0.633	0.733
	Binary	0.708	0.755	0.656	0.712	0.767
	Median Filter	0.788	0.791	0.72	0.752	0.824
	Binary and Median Filter	0.697	0.72	0.68	0.71	0.754
	Canny Operator and Median Filter	0.687	0.728	0.691	0.7	0.745
Cross Validation	Original	0.894	0.855	0.81	0.839	0.884
	Gray Scale	0.793	0.804	0.737	0.78	0.816
	Canny Operator	0.687	0.701	0.637	0.663	0.766
	Binary	0.727	0.752	0.689	0.713	0.766
	Median Filter	0.779	0.802	0.747	0.772	0.836
	Binary and Median Filter	0.686	0.739	0.696	0.708	0.774
	Canny Operator and Median Filter	0.674	0.728	0.668	0.693	0.759

Table 5. Accuracy for each category B-AC classification algorithm and scenario

Classification Scenario	Image Type	SVM	K-NN	DT	NB	ANN
Training Sets 80-20	Original	0.767	0.739	0.749	0.744	0.786
	Gray Scale	0.686	0.713	0.64	0.748	0.75
	Canny Operator	0.608	0.615	0.601	0.65	0.655
	Binary	0.62	0.633	0.583	0.623	0.641
	Median Filter	0.661	0.696	0.658	0.716	0.721
	Binary and Median Filter	0.621	0.641	0.609	0.636	0.661
	Canny Operator and Median Filter	0.6	0.644	0.605	0.637	0.668
Cross Validation	Original	0.775	0.755	0.697	0.77	0.778
	Gray Scale	0.674	0.706	0.648	0.738	0.739
	Canny Operator	0.613	0.597	0.598	0.662	0.639
	Binary	0.6	0.653	0.61	0.626	0.653
	Median Filter	0.655	0.712	0.65	0.684	0.745
	Binary and Median Filter	0.644	0.65	0.597	0.64	0.662
	Canny Operator and Median Filter	0.579	0.638	0.596	0.642	0.669

classification scenario for cross-validation and neural network (81.3%) and similarly, but for the training set (79.1%). It was evident that grouping the images in this way slightly increases the classification accuracy, as shown in Table 3.

Classifiers for category B-AC. For this category, in Table 4, it was observed that the increase in

accuracy was substantial when the images were grouped according to this category.

Fairly accurate results were obtained for both classification scenarios (cross-validation and training set) but this time for the vector support machine (SVM) algorithm, 89.4% and 88.7% respectively.

Classifiers for category C-AB. For this category, unlike the previous two, the performance of the classifier drops. This time the set of training for the neural network got the best performance (78.6%), while, for the same classification algorithm, but with cross-validation, a similar value was obtained (77.8%), as shown in Table 5.

5 Conclusions

The difference in precision values for the two scenarios posed in the classifiers was not affected for each classification algorithm, resulting in being similar for both cases.

Using filters and processing on images did not significantly improve classification values, however, it comes very close to grading performance for unfiltered images.

It was observed that the best performance in most cases was for the cross-validation scenario. As for the classification algorithms, the artificial neural network architecture and the vector support machine were responsible for presenting a higher performance. Different results were obtained depending on the category proposed and the type of evaluation used. Classification category B-AC got the highest performance values, an indication that the classifier presents problems when classifying classes, A and C separately.

This classifier could be a useful support tool to complement dermatological diagnosis, given its acceptable accuracy.

6 Future work

As future work, it is intended to use texture operations because it was observed that the operations on the pixels do not provide benefits in the performance of precision for the classification of the sets of images.

We plan to create a multiplatform interface easy and intuitive for the user, able to support in the medical diagnosis for skin cancer.

In order to increase the spectrum of use of the classifier, it is advisable to use more data as far as the types of cancer are concerned.

References

1. **Affonso, C., Rossi, A. L. D., Vieira, F. H. A., Carvalho, A. C. P. d. L. (2017).** Deep learning for biological image classification. *Expert Syst. Appl.*, 85(C), 114–122. DOI: 10.1016/j.eswa.2017.05.039.
2. **Cuevas, E., Zaldívar, D. Pérez, M., (2010).** Procesamiento digital de imágenes con MATLAB Simulink. Alfaomega and Ra-Ma. Ciudad de México, México.
3. **Elgamal, M. (2013).** Automatic Skin Cancer Images Classification. *International Journal of Advanced Computer Science and Applications*, 4(3). DOI: 10.14569/ijacsa.2013.040342.
4. **Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., Thruon, S. (2017).** Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118. DOI: 10.1038/nature21056.
5. **Guzmán-Cabrera, R., Guzman-Sepulveda, J. R., Parada, A. G., García, J. J. R., Cisneros, M. T., Baleanu, D., Blanco, P. (2016).** Digital Processing of Thermographic Images for Medical Applications.
6. **Guzmán-Cabrera, R., Guzmán-Sepúlveda, J. R., Torres-Cisneros, M., May-Arrijoja, D. A., Ruiz-Pinales, J., Ibarra-Manzano, O. G., Aviña-Cervantes, G. (2013).** Pattern recognition in photoacoustic dataset. *International Journal of Thermophysics*, 34, 1638–1645. DOI: 10.1007/s10765-013-1452-9.
7. **Hernández Hoyos, M., Dávila Serrano, E. E., Correa Agudelo, E. M., Torres González, J. S., Corredor Jérez, R. A., Gutiérrez Ossa, F. A., Flórez-Valencia, L., Mouton, C., Orkisz, M. (2012).** CreaTools: una plataforma para desarrollo de software de procesamiento de imágenes médicas. Aplicación a la segmentación de arterias coronarias y a la detección automática de lesiones en imágenes de angioTAC. *Revista colombiana de radiología*, 23(3), 3521–3528.
8. **Herrera-Pereda, R., Taboada Crispi, A., Babin, D., Philips, W., Holsbach Costa, M. (2021).** A Review on digital image processing techniques for in-Vivo confocal images of the cornea. *Med Image Anal*, 73, 102188. DOI: 10.1016/j.media.2021.102188.
9. **Jinnai, S., Yamazaki, N., Hirano, Y., Sugawara, Y., Ohe, Y., Hamamoto, R. (2020).** The development of a skin cancer classification system for pigmented skin lesions using deep

- learning. *Biomolecules*, 10(8), 1123. DOI: 10.3390/biom10081123.
10. **Kadampur, M. A., Al Riyae, S. (2020).** Skin cancer detection: Applying a deep learning based model driven architecture in the cloud for classifying dermal cell images. *Informatics in Medicine Unlocked*, 18, 100282. DOI: 10.1016/j.imu.2019.100282.
 11. **Kaur, M., Singh, D. (2021).** Multi-modality medical image fusion technique using multi-objective differential evolution based deep neural networks. *Journal of Ambient Intelligence and Humanized Computing*, 12(2), 2483–2493. DOI: 10.1007/s12652-020-02386-0.
 12. **Li, Y., Zhao, J., Lv, Z., Li, J. (2021).** Medical image fusion method by deep learning. *International Journal of Cognitive Computing in Engineering*, 2, 21–29. DOI: 10.1016/j.ijcce.2020.12.004.
 13. **Liu, H., Cocea, M., Ding, W. (2017).** Decision tree learning based feature evaluation and selection for image classification. *International Conference on Machine Learning and Cybernetics (ICMLC)*, pp. 569–574. IEEE. DOI: 10.1109/ICMLC.2017.8108975.
 14. **Martínez-Espinosa, J. C., Cordova-Fraga, T., Guzmán-Cabrera, R. (2019).** Generation of Raman images through spectral mappings. *Journal of Intelligent and Fuzzy Systems*, 36(5), 4991–4999. DOI: 10.3233/JIFS-179045.
 15. **Pérez-Gutiérrez, J., Magadán-Salazar, A., Pinto-Elías, R., Mejía-Lavalle, M. (2016).** Súper resolución y mejora del algoritmo Canny para la detección de bordes en imágenes médicas. *Pistas Educativas*, Vol. 38, No. 120.
 16. **Quintanilla-Domínguez, J., Barrón-Adame, J. M., Gordillo-Sosa, J. A., Lozano-García, J. M., Estrada-García, H., Guzmán-Cabrera, R. (2016).** Analysis of mammograms using texture segmentation. *Research in Computing Science*, 123(1), 119–126. DOI: 10.13053/rcs-123-1-11.
 17. **R, K. M., Prashantha, D. H. S., Shetty, B. R. (2021).** Comparison of image classification techniques and its applications. *Asian Journal for Convergence in Technology (AJCT)* ISSN - 2350-1146, 7(1), 58–62. DOI: 10.33130/AJCT.2021v07i01.013.
 18. **R, S. (2016).** Melanoma image classification system by NSCT features and Bayes classification. *International journal of advances in signal and image sciences*, 2(2), 27–33. DOI: 10.29284/ijasis.2.2.2016.27-33.
 19. **Ramesh, K. K. D., Guthikonda, K., Swapna, D. K., Datta, D., Rajest, S. (2021).** A review of medical image segmentation algorithms. *EAI Endorsed Transactions on Pervasive Health and Technology*, 7. DOI: 10.4108/eai.12-4-2021.169184.
 20. **Rezvantab, A., Safigholi, H., Karimijeshni, S. (2018).** Dermatologist level dermoscopy skin cancer classification using different deep learning convolutional neural networks algorithms. *arXiv preprint arXiv:1810.10348*.
 21. **Sagar, A., Jacob, D. (2020).** Convolutional neural networks for classifying melanoma images. *Cold Spring Harbor Laboratory*. DOI: 10.1101/2020.05.22.110973.
 22. **Sonka, M., Hlavac, V., Boyle, R. (2014).** *Image Processing, Analysis, and Machine Vision*. Cengage Learning, Stamford, USA.
 23. **Tan, Teck, Zhang, Li, Neoh, Siew Chin, Lim, Chee Peng (2018).** Intelligent skin cancer detection using enhanced particle swarm optimization. *Knowledge-Based Systems*, 158. pp. 118-135. ISSN 0950-7051.
 24. **Tschandl, P., Rinner, C., Apalla, Z., Argenziano, G., Codella, N., Halpern, A., Janda, M., Lallas, A., Longo, C., Malvehy, J., Paoli, J., Puig, S., Rosendahl, C., Soyer, H. P., Zalaudek, I., Kittler, H. (2020).** Human-computer collaboration for skin cancer recognition. *Nature Medicine*, 26(8), 1229–1234. DOI: 10.1038/s41591-020-0942-0.
 25. **Tu, B., Wang, J., Kang, X., Zhang, G., Ou, X., Guo, L. (2018).** KNN-Based representation of superpixels for hyperspectral image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11, 4032–4047. DOI: 10.1109/jstars.2018.2872969.
 26. **Vapnik, V. (2013).** *The nature of statistical learning theory*. Springer Science Business Media, Berlin, Heidelberg, Germany.
 27. **Xie, F., Gao, Q., Jin, C., Zhao, F. (2021).** Hyperspectral image classification based on superpixel pooling convolutional neural network with transfer learning. *Remote sensing*, 13(5), 930. DOI: 10.3390/rs13050930.
 28. **Yu, L. (2014).** *Exploring spatial resolution in image processing*. Rice University.

*Article received on 29/07/2021; accepted on 30/09/2021.
Corresponding author is Rafael Guzmán Cabrera.*