Hybrid Model of Convolutional Neural Network and Support Vector Machine to Classify Basal Cell Carcinoma

Jorge Alexander Ángeles Rojas[1], Hugo D. Calderón Vilca[1], Ernesto N. Tumi Figueroa[2],
Kent Jhunior Cuadros Ramos[1], Steve S. Matos Manguinuri[1], Edwin F. Calderón Vilca[2]

[1] Universidad Nacional Mayor de San Marcos,
Peru

[2] Universidad Nacional del Altiplano,
Peru

{jorge.angeles4, hcalderonv, kent.cuadros, steve.matos }@unmsm.edu.pe,
{nayer.tumi, mcedwin}@unap.edu.pe

Abstract. Skin cancer is one of the most common types of cancer in humans, it covers about one third of all neoplasms. Within skin cancer we find basal cell carcinoma (BCC), this being the most frequent type of cancer worldwide. Solutions with convolutional neural networks generally use the Softmax layer (classic model) to perform a BCC classification, however, in other similar fields such as image classification of microscopic bacteria they have replaced this Softmax layer with a support vector machine (SVM) achieving a better result. Given this, we propose a hybrid model of convolutional neural network and a support vector machine (CNN+SVM) to classify the BCC. Our model is composed of 4 convolution blocks with 32, 64 and 128 filters to carry out the extraction of characteristics and then pass it to the classifier, to which the L1-SVM loss function is implemented. The average results obtained for the CNN+SVM hybrid model were measured with the precision, accuracy, recall and F1-score metrics, obtaining 96.200%, 96.200%, 96.205% and 96.200% respectively compared to the classical model for the metrics of precision, accuracy, recall and F1-score where 95.661%, 95.673%, 95.661%, 95.660% respectively were obtained. The results show that the hybrid model achieves better results than the classic model to classify the BCC.

Keywords. Basal cell carcinoma, convolutional neural network, support vector machine, deep learning.

1 Introduction

Artificial intelligence algorithms are being applied in different solutions: recommendations [32], in semantic search [30], augmented reality [31], questions and answers [28] and with greater emphasis they are applied in the detection of diseases [29, 27].

Skin cancer is one of the most common types of cancer in humans, it covers about one third of all neoplasms; where two large groups are distinguished: malignant melanoma skin cancer; the latter includes basal cell carcinoma (BCC) is a malignant epithelial tumor that originates from pluripotential epithelial cells and is characterized by slow growth, but is locally invasive; although it has a low metastatic potential (cancer reproduction to another part of the body), it has local destructive capacity and compromises extensive areas of tissue, cartilage and even bones, in the most severe clinical forms [1].

Globally, the prevalence of melanoma skin cancer accounts for only 1% of skin cancer cases, so it can be deduced that 99% of skin cancer cases are non-melanoma. In addition, within non-melanoma cancer, eight out of ten cases are from BCC [2, 3].

Continents such as North America and Europe are places where the highest number of new cases of non-melanoma were found in 2018, with the United States, Germany and Australia being the first three countries that report a greater number of new cases of non-melanoma skin cancer (with 436 869, 77 272 and 59 278 new cases respectively); in Latin America, countries such as Brazil, Colombia and Argentina reported a greater

number of new cases of non-melanoma cancer in 2018 (with 32 107, 4 685 and 4 003 new cases respectively). In the case of Peru, this is ranked fifth in new cases of non-melanoma cancer in Latin America with 2 527 new cases [4].

Dermatoscopy has improved the diagnostic accuracy of melanoma by 10% to 27% compared to eye exams. However, the accuracy of dermatoscopy image analysis still depends on the experience of a doctor.

A consensus diagnosis is recommended in which two or more experts participate to obtain the highest possible diagnostic accuracy [5], furthermore it is indicated that a dermatologist not trained in reading dermatoscopy images may be less precise than the analysis at a glance [6].

For the reading of images of non-melanoma cancer, specifically for basal cell carcinoma (BCC), Artificial Intelligence has been used to diagnose BCC, for example [7, 8] used dermotoscopic images, [9, 10] used images such as images of cancerous areas, oral tissue or clinical images. Among the techniques used to classify BCC, it has been reviewed that researchers [7] used ResNet, in research [8] they used Random Forest, [9, 10] used a classic convolutional neural network.

Currently, in the medical field, it is sought to achieve high classification results using artificial intelligence techniques, since an erroneous diagnosis given by a false positive or a false negative can seriously affect the health of patients. According to [11] a series of studies involving deep learning approaches have already performed a considerable number of tasks.

These include, among others, image classification, natural language processing, voice recognition and text classification. The models used in these tasks use the Softmax function (classic model) in the classification layer.

However, studies have been conducted that analyze an alternative to the Softmax function for classification is the use of the support vector machine, in an artificial neural network architecture produces relatively better results than the use of the conventional Softmax function, by what is possible to propose a hybrid model to diagnose BCC.

On the other hand, there are already researchers, who made comparisons and improved their results using the hybrid model of convolutional neural network (CNN) with a support vector machine (SVM) versus the classic model (Softmax) to classify images of microscopic bacteria [12]. Also in [13], they propose a hybrid model and they managed to classify the wear of drills based on images of perforated holes, in another work they present a hybrid model in the classification of a person's gender by the way they walk [14].

Based on the reviews of other research and seeing the improvement using a hybrid model.

In this research, we propose a hybrid model of CNN+SVM that allows classifying basal cell carcinoma (BCC).

The rest of this document is organized as follows. Section 2 addresses works related to this research. In section 3, we include the methodology. Section 4 describes our experimental evaluation and results; the conclusions are presented in the final part.

2 State of the Art

Conducting a review of the scientific literature on the diagnosis of skin cancer in general using Artificial Intelligence, there are investigations that sought new methods and ways to solve this problem, especially in melanoma-like skin cancer. Within melanoma-focused research, there are researchers such as [16, 17, 18, 19, 20, 21, 22, 23, 24], who focused on dermatoscopic images, while [25] used facial images where the treatment of these differ somewhat to dermatological images. Of the used techniques, the vast majority are oriented to Deep Learning, using convolutional networks [17, 22, 23] or adopting specific neural network architectures [16, 17, 18, 20, 25, 21, 24] and several of these investigations focus on classification using the classic neural network with Softmax.

The diagnosis of skin cancer, among which is basal cell carcinoma, has attracted the attention of several researchers, who applied different machine learning techniques to solve this problem, among these we found that [8] presented a segmentation method for Accurate removal of cutaneous blood vessels in dermatoscopic images for BCC classification.

For this, the authors used a dataset composed of 659 images of normal dermatoscopy and cancer lesions, this dataset was made up of different data sources, among which were "Atlas of dermoscopy", the University of Missouri and "Vancouver Skin Care Center".

In addition, for the classification the Random Forest technique was used with 100 trees each constructed considering 4 random characteristics to perform a classification of two kinds of cancerous versus benign lesions.

In the investigation [7], the authors proposed a residual neural network method that seeks to diagnose BCC automatically. For this purpose, they used two datasets, a public dataset called "Skin Lesion Analysis towards Melanoma Detection" of the ISBI 2016 and the second of the International Skin Imaging Collaboration (ISIC). As the method, they used a segmentation technique and a deep residual network (ResNet) to perform the classification where the output was encoded with a range of 0 to 1 that represents a probability of whether an image is BCC or not.

In another study [9], they presented a convolutional neural network architecture to detect BCC using optical coherence tomography (OCT) images. Their data consisted of 40 full-field optical coherence tomography (FFOCT) images. For the classification, they developed a classical architecture of a multilayer convolutional neural network. Although they used a classic neural network they took advantage of VGG ideas, such as: 1) convolutional blocks, consecutive convolutional layers to capture larger inputs with a spare parameter; 2) abandonment layer, a fraction of neurons is removed at random to avoid overfitting; 3) rectified linear unit (ReLU) as an activation function used to accelerate the calculations.

In the research [26], the authors develop a mitotic cell segmentation scheme for automatic recognition of microscopic images of oral squamous cell carcinoma. To achieve this objective, they collected data from the Department of Pathology, Minadpur Medical College and Hospital India. In total, they collected 15 samples of oral tissue from patients from which five images of each sample were taken. As the next step, to identify the possible candidate cells, the images were converted to a gray scale, then segmented and filtered the images to have a series of candidate cells and eliminate other cells.

Then the authors used "the seven moments of Hu's" to discriminate the mitotic cells of the remaining candidate cells that were used as input for the classifier. For classification they used CART, which is a non-parametric decision tree model, they show results with the accuracy, sensitivity and F-score metrics of 83.8%, 73.5% and 78.3% respectively. Comparing with the results of [9] they obtained the best results, which were measured according to the accuracy, sensitivity and specificity with results of 95.35%, 95.2% and 96.54% respectively.

As for skin cancer [16], the authors present a framework for the recognition of dermatoscopic images using Deep learning and a local descriptor coding strategy. For this purpose, they used the ISBI 2016 dataset, which contains 1,279 images of dermatoscopic lesions and they selected images that contained melanoma by taking 173 images and 75 test images. These images were resized (224x224 pixels), normalized and the technique for the data increase with rotations and translations was applied.

Then the images passed through a Deep residual neural network (ResNet) for feature extraction. The authors demonstrated that they can accelerate the convergence of the deep network and maintain accuracy gains. In addition, they used the architectures of AlexNet and VGG16 to evaluate and compare their results, evaluated mainly with the metrics of precision and AUC for the architectures of AlexNet with 84.70% and 82.02%; VGG16 with 84.43% and 81.18%; ResNet50 86.54% and 81.49% respectively.

In the study [10], the authors propose a method based on a stacked scattered autoencoder (SSAE), which is an unsupervised learning method with a fully connected architecture, to detect the translucent areas of BCC clinical image patches using as input images clinics, understood as images of the lesion that can be seen with the naked eye.

For this purpose, 32 clinical images of BCC from 32 patients of the Vancouver Skin Care Center were used. These images had a dimension of 3008x2000 pixels, of these images captured the translucency, which is defined as a jelly-like appearance, which is an important characteristic

feature of BCC at an early stage because the characteristic can be easily observed and for the classification they used a convolutional neural network with Softmax.

In the literature on the diagnosis of skin cancer, there are already several proposals. These investigations [7, 17, 18, 19, 20] were based on dermatoscopic images using public datasets such as PH2, ISBI, MED-NODE and ISIC. On the other hand, studies in [8] were based on dermatoscopic images and the research [10] was based on clinical images, when the authors took the images of the lesion that can be seen with the naked eye. They captured the image without any specialized instruments. In this case, the authors got their data from private institutions such as the University of Missouri or Vancouver Skin Care Center.

As for preprocessing, [7, 17, 18,] performed some standard steps such as resizing images, increasing data or segmentation, [10] used techniques such as ROI mask and stacked scattered autoencoder (SSAE), [19] used the technique of Superpixel based fine-tuning, [20] used the HSV and GLCM (gray level matching matrix) color model and [8] used Independent Component Analysis (ICA).

Among the techniques for classification we can note that [10, 19] used a convolutional neural network, [7, 17] used ResNet, [18] used VGG16, ResNet50 and InceptionV3, [20] used SVM and [8] used a Random Forest classifier. Of these works, the one that obtained the best result was [20] with 95%, compared to [7, 10, 17, 18, 19] with 93%, 83.2%, 85%, 93%, 92.3%, in [8] they reached 82.7% accuracy.

In research [21, 22, 23, 24, 25], the authors propose models to diagnose skin cancer. Among the techniques used for classification we can note that [25] tested using different types of architectures such as ResNet-50, Inception-v3, DenseNet121, Xception and Inception-ResNet-v2. In the investigation [21], the authors used ResNet50, in [22, 23] they used convolutional neural networks, and [24] they used the VGG16 architecture.

As for the preprocessing in [22, 25,] they resized the images in this way they managed to increase data, in [21] they applied a series of techniques: "Median Color Split Algorithm", "Atypical Pigment Network Detection", "Cheng vessel detection", in [23] they added the use of the ROI mask, in [24] they used a convolutional-deconvolutional architecture to obtain a tumor mask.

Regarding the results in [25], the authors show sensitivity and accuracy metrics respectively for Resnet50 74.2% and 60.4%, for InceptionV3 79.2% and 45.5%, for DenseNet121 76.9% and 57.5%, for Xception 83.1% and 65.9%, for Inception-ResnetV2 89.2% and 63.7. In [21], the authors showed higher results with 90% of AUC for the fusion of their two proposed methods. In [22], the authors evaluated with the metric precision for the functions of activation ReLU, tangent and APL and the results were 93.25%, 91.76% and 95.86% respectively. In [23], the obtained results were measured with the metrics of accuracy, sensitivity and specificity with the results of 92%, 94% and 93 % respectively. In [24], the authors used VGG16 being the results 79.7%, 76.3%, 45.4%, 34.1%, 90.7% for the metrics precision, accuracy, sensitivity, specificity and AUC respectively.

Of the investigations focused on basal cell carcinoma, in [7, 8] the authors used dermatoscopic images. [7] used the ISBI 2016 dataset and for the preprocessing they performed some standards steps such as image resizing, data augmentation or segmentation and a ResNet neural network, obtaining the highest results with precision, sensitivity and specificity of 93%, 97% and 96% respectively. In [8], the authors used three different sources of data from "Atlas of dermoscopy", the University of Missouri and "Vancouver Skin Care Center", which formed its dataset. For pre-processing, they used Independent Component Analysis (ICA) and a Random Forest classifier obtaining 82.7% accuracy and 82.4% sensitivity. In [9], the authors used optical coherence tomography images obtained by the same researchers and a convolutional neural network to which they adapted ideas from the VGG architecture, obtaining results that were measured based on accuracy, sensitivity and specificity with results of 95.35%, 95.2% and 96.54% respectively.

On the other hand, in [10], the authors obtained their data from private institutions such as the university of Missouri and Vancouver Skin Care Center, where 32 clinical images of BCC were obtained. For the pre-processing, the ROI mask
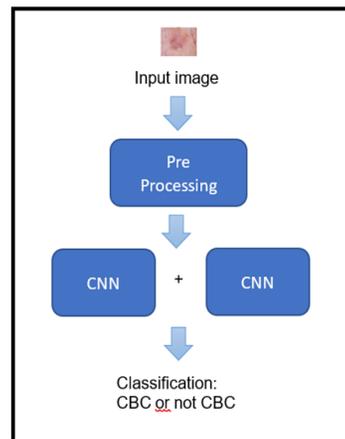
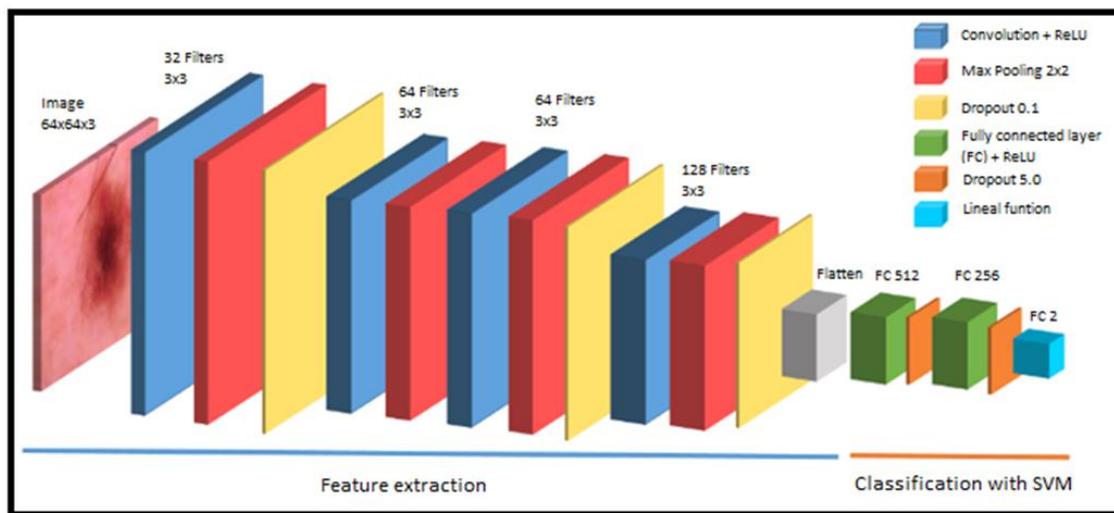**Fig. 1.** Model for classifying BCC at high leve

Fig. 2. CNN+SVM hybrid model to classify the BCC

and stacked scattered autoencoder (SSAE) were used. Further, they used a convolutional neural network for classification, obtaining the results that were measured according to precision, sensitivity and specificity with 93%, 77% and 97.1% respectively.

3 Design of the Hybrid Model for BCC Classification

In this section, we present the methodology to be performed to combine a convolutional neural network plus support vector machine to classify basal cell carcinoma.

The model proposed in this research for the classification of basal cell carcinoma (BCC) is composed of three parts: pre-processing, feature extraction and classification. In pre-processing, the image is prepared to serve as input for the neural network. This is done in order for the network to be trained efficiently.

For the extraction of characteristics, the images of BCC and non-BCC pass through four layers of convolutional with 3x3 filters and then pass through four layers of max pooling with 2x2 filters to obtain

the filters that allow identifying the BCC. Then for classification, the images pass through the fully connected layer composed of two dense layers of 512 neurons and 256 neurons to finally reach a dense layer with 2 neurons, where, in the latter, the L1-SVM and the linear activation functions are used to perform the classification. In figure 1, it is shown the representation of the high-level model of the present investigation.

3.1 Dataset

The dataset used is HAM10000, which is a data set of dermatoscopic images of pigmented skin lesions from two different sites, the Department of Dermatology of the Austrian University of Medicine of Austria and Skin Cancer Practice of Cliff Rosendahlen in Queensland, Australia.

This dataset consists of 10,015 images with different skin diseases among which we can find actinic keratosis (akiec), basal cell carcinoma (bcc), seborrheic keratosis (bkl), dermatofibroma (df), melanocytic nevus (nv), melanoma (mel), vascular lesions (vasc).

3.2 Image Pre-Processing

The images of the HAM10000 are labeled with different abbreviations according to the type of skin cancer. To load the data, BCC images are labeled "bcc" and all non-BCC images have been labeled "nobcc". Then the images are resized to a size of 64 x 64 pixels. As there is a large imbalance of data (514 bcc and 9501 of nobcc), we proceed to perform the data augmentation technique for the bcc group, applying 9 random rotations, 9 random zooms, vertical and horizontal turns for each image. In this way, we practically remove the imbalance between these two classes, finally obtaining 9,766 for bcc and 9,501 for nobcc.

These were then normalized before moving to the convolutional neural network. The normalization consisted of dividing each of the pixels of the image by the value of 225. This is done for obtaining that each pixel has a value between 0 and 1, because entries with large values can slow down the learning process.

3.3 Feature Extraction

After the images are loaded and preprocessed, the model presented in Figure 2 is implemented. The 64x64x3 images (64x64 pixel image with 3 color channels) pass through the convolution layers, where the convolution operation is performed between the input image and the predefined filter or kernel to obtain the feature map:

$$z(t) = (f * g)(t) = \int_{-\infty}^{\infty} f(y)g(t - y)dy.$$

Convolution operation, where:

–   The z function is the feature map of the BCC.
–   The f function is the input image of the HAM10000.
–   The g function is the filter.
–   t is the displacement.

In the first convolution layer, 32 filters with a dimension or kernel size of 3x3 are defined and the ReLU activation function is applied in this layer to increase non-linearity. Since it is normal for BCC images to have non-linear elements, this convolution layer is followed by a max pooling layer with a dimension of 2x2 and a dropout of 0.1. Then two convolution layers with 64 filters are followed with a 3x3 kernel size with the ReLU activation function each with its respective 2x2 max pooling layer and also followed by a dropout of 0.1.

Finally, there is a final convolution layer with 128 filters with a 3x3 kernel size with the ReLU activation function followed by a last 2x2 max pooling layer to which a 0.1 dropout is also applied. At the end of the convolution phase, flattening is applied, obtaining a one-dimensional vector that connects to the fully connected layer with 512 input neurons.

3.4 Classic Classification

After the convolution layers, there are two fully connected layers of 512 and 256 neurons that have the ReLU activation function plus a dropout of 0.5, followed by the last fully connected layer with 2 neurons (since we have a binary classification), but in this case Softmax activation function is used. The Softmax activation function is given by the following equation:

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \; ; \; para \; i = 1, \dots, K; \dots,$$

where:

- z_i is the output of the last layer.
- K is the number of the classes to classify the BCC, in this case it is 2.

Then the cross-entropy loss function is used, which is given by the following formula:

$$CE = -\sum_{i}^{c} t_i \log(s_i) \dots,$$

where:

- t_i y s_i are the expected label and the score for class *i*,
- C is the number of classes to classify.

3.5 Classification with the Hybrid Model Proposed in our CNN+SVM Research

After the features in the convolutional layer are extracted, two fully connected layers of 512 and 256 neurons follow with the ReLU activation function and a dropout of 0.5, followed by a last fully connected with 2 neurons since we have a binary classification. Then the linear activation function is used, which its output is proportional to the input received, and in our case, the input is the output of the last layer.

To perform the classification in the final layer, we have implemented the SVM classifier, and as indicated in [11, 15], it is sufficient to implement the loss function of the L1-SVM standard in the last layer:

$$min \frac{1}{p} w^T w + C \sum_{i=1}^{p} max\left(0, 1 - y_i'(w^T x_i + b)\right).$$

where:

- C is the penalty parameter (it can be an arbitrary value, or a value selected through the hyperparameter setting).
- y' is the real label of the BCC.
- $w^T x + b$ is the predictive function.
- $w^T w$ is the Manhattan Standard (also known as L1 standard).

4 Results and Discussion

The classic model (convolutional neural network using Softmax) has been designed and another CNN+SVM hybrid model has been designed. We evaluate the results of both models in the BCC classification and compare which of them gives better results.

As mentioned in Section 3, the images used for training, validation and testing of the neural network were provided with the HAM10000. The dataset initially had 10,015 dermatoscopic images, but with the data augmentation technique, 19,267 images were available, dividing this data into 10,837 training images, 3,613 validation images that are used for training evaluation and the rest 4,817 were used for the test evaluating the model with metrics of precision, accuracy, recall and f1- score.

To compare the performance of the classic model (Softmax) and the CNN+SVM hybrid model, the configurations for both cases are presented.

The training configuration for the classic model (Softmax) to classify the BCC is:

- Epochs: 150,
- Batch size: 80,
- Optimizer: Adam,
- Activation function in the last layer: Softmax,
- Loss function: Cross entropy.

The training configuration for the CNN+SVM hybrid model to classify BCC is:

- Epochs: 150,
- Batch size: 80,
- Optimizer: Adam,
- Activation function in the last layer: Linear,
- Loss function: L1-SVM.

4.1 Results and Evaluation of the Classic Model to Classify the BCC

For training evaluation, training data were used with 10,837 images and validation with 3,613 images, where precision and losses are evaluated. They were measured on a graph with the number of epochs vs. precision and another graph with the number of epochs vs. losses.

As it can be seen in Figure 3, the training of the validation data tends to increase from epoch 0 to epoch 40 to 94%, but then starting from epoch 40
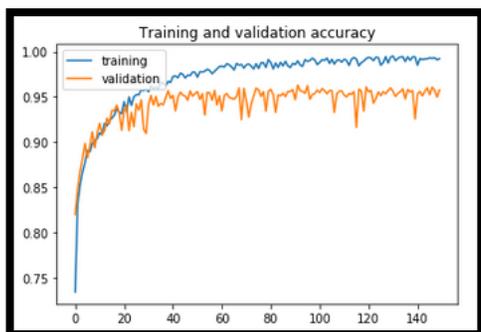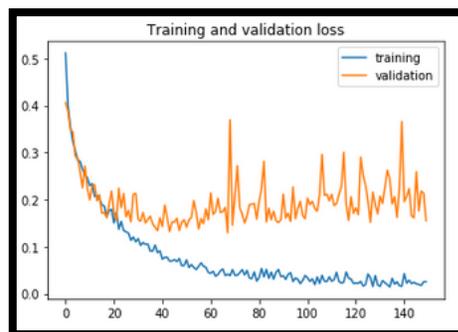
Fig. 3. Precision evaluation for the classic model

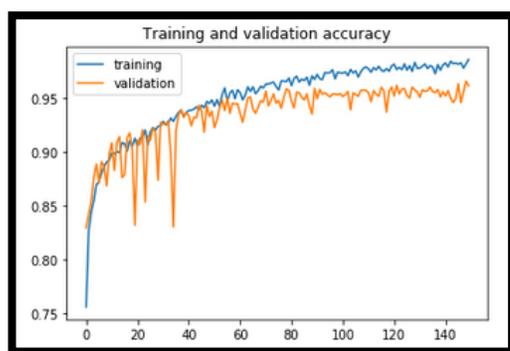**Fig. 4.** Loss evaluation for the classic model

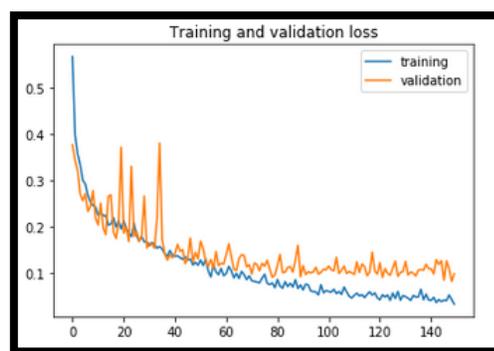**Fig. 5.** Precision evaluation for the CNN+SVM hybrid model

Fig. 6. Loss evaluation for the CNN+SVM hybrid model

training tends to remain constant with certain fluctuations.

In Figure 4, it can be noted that the losses of the validation data tend to decrease until about epoch 20 and then begin to rise gradually with abrupt fluctuations. The classic model has quite acceptable training with approximately 94% precision, but Figure 4 losses tend to rise sharply giving a less stable convergence.

In Figure 3 and in Figure 4, a great overfit is being evidenced in the classic model. With this, we do not ensure that the model when receiving new BCC images can correctly classify whether or not the BCC is presented.

The classic model is then evaluated using the 4,817 test or testing data to calculate the precision, accuracy, recall and f1-score metrics, which are

obtained with the confusion matrix, from which we extract 2328 True Positive (TP), 2280 True Negatives (TN), 85 False Negatives (FN) and 124 False Positives (FP).

From the TP, TN, FN and FP, we can obtain the precision metric, in which 95.66% was obtained, which is understood as the total positive observations correctly predicted among the total of the data and for the following metrics of accuracy, recall and F1-Score. The results are summarized in Table 2, where they are divided into metrics for "BCC", "Non-BCC" or a weighted average of these metrics.

The average accuracy is 95.67%, which is understood as the relationship between the positive observations correctly predicted and the total positive observations correctly predicted and

Table 1. Confusion Matrix of the classic model

BCC Classification		Prediction	
		BCC	Non-BCC
Observation	BCC	2328	85
	Non-BCC	124	2280

Table 2. Metrics of the classic model

	Accuracy	Recall	F1-score
BCC	94.94%	96.48%	95.70%
Non BCC	96.41%	94.84%	95.62%
Average	95.67%	95.66%	95.66%

Table 3. Confusion matrix of the CNN+SVM hybrid model

BCC Classification		Prediction	
		BCC	Non-BCC
Observation	BCC	2334	79
	Non-BCC	104	2300

the total positive observations (true positives and false positives). With the recall metric, 95.66% was obtained, which is the proportion of positive observations correctly predicted to all observations in the expected class. For F1 score, the result of 95.66% was obtained, which is the weighted average of accuracy and sensitivity.

Therefore, this score takes into account both false positives and false negatives. Intuitively, it is not as easy to understand as accuracy, but F1 is usually more useful than accuracy, especially if it has an unequal class distribution. In our data set, there is a small imbalance between the classes "BCC" and "Non-BCC", thus, this metric is taken into consideration.

4.2 Results and Evaluation of the Hybrid Model (CNN+SVM) for BCC Classification

In the case of the CNN+SVM hybrid model, training is also evaluated, where training data with 10837

images and validation with 3613 images are used, to analyze the precision and losses achieved, which we measure with a number graph of epochs vs precision and another graph with the number of epochs vs losses

As can be seen in Figure 5, at the beginning of the training the precision of the validation data fluctuates between 85% and 90% and the precision begins to rise from the epoch 40, then the precision increases along with the number of epochs until reaching the 100th epoch with 95% precision approximately where the precision achieved by the validation data is stabilized.

In the same way for Figure 6, the losses of the validation data tend to fluctuate until epoch 40 and then the losses are gradually stabilized as the number of epochs progresses.

Then the CNN+SVM hybrid model was evaluated using the 4817 test or testing data to calculate the precision, accuracy, recall and f1-score metrics. These metrics are obtained from the confusion matrix, where we obtain 2334 True positive (TP), 2300 True negative (FN), 79 False negative (FN) and 104 False positive (FP).

From the TP, TN, FN and FP we can obtain the average of the precision metric in which 96.2009% was obtained and for the following accuracy, recall and F1-Score metrics Table 4 is presented, where they are divided into metrics for "BCC", "Non-BCC" or a weighted average of these metrics.

Then, from the average results obtained from the CNN+SVM hybrid model with the metrics of precision, accuracy, recall and f1-score, we obtain 96.200%, 96.200995%, 96.205860% and 96.200955% respectively.

4.3 Comparison of the Classic Model Versus the Hybrid Model to Classify the BCC

As shown in Table 5 below, a comparison of the metrics obtained by both models is presented.

It is clearly seen that the CNN+SVM hybrid model achieves better results to classify the BCC in all the proposed metrics, with an improvement of approximately 1% for all the metrics, also if we evaluate the graphs of losses of both models (Figure 4 and Figure 6) we can conclude that the CNN+SVM hybrid model has a much smaller overfitting and does not tend to rise and therefore has a better generalization capacity, so we can conclude that the CNN+SVM hybrid model

Table 4. Metrics of the CNN+SVM hybrid model

	Accuracy	Recall	F1-score
BCC	95.73%	96.73%	96.23%
Non-BCC	96.68%	95.67%	96.17%
Average	96.20%	96.21%	96.20%

Table 5. Comparison of metrics of the classic and CNN+SVM hybrid model

	Accuracy	Recall	F1-score	Precision
Classic	95.67%	95.66%	95.66%	95.66%
Hybrid	96.20%	96.20%	96.20%	96.20%

Table 6. Comparison of results with other investigations that classifies the BCC

Author	Technique	Precision	Accuracy	Recall	F1-score
D. Mandache [9]	CNN	95.93%	-	95.2%	-
H. Huang [10]	CNN	93%	-	77%	-
E. V. Putten [7]	ResNet	93%	-	97%	-
P. Kharazmi el al. [8]	Random Forest	-	82.7%	82.4%	-
Proposed Model	CNN+SVM	96.20%	96.20%	96.20%	96.20 %

exceeds the classic model both in metrics as in model training.

4.4 Discussion

The results obtained from our CNN+SVM hybrid model that classifies the BCC, has been compared against other works considered in the state of art. In Table 6, the comparison with other investigations is shown where the technique used, and the results achieved are described.

The study carried out in [9] presents a convolutional neural network architecture to detect BCC using optical coherence tomography (OCT) images. These types of images are collected using laser scanner, originally their dataset only had 40 images (30 normal and 10 with bcc) which was applied the technique of data increase reaching 59,112 images labeled non BCC and 48,970 images labeled BCC, it should be noted that in their results they obtained 95.35% and 95.2% for

accuracy and recall metrics respectively. Comparing with our research, we use 10,015 dermatoscopic images that are different. They were increased to 19,267 images. Compared with our results, we managed to improve accuracy and recall metrics with 96.20% and 96.21% respectively.

In the study conducted by [7], the authors proposed a residual neural network method that seeks to automatically diagnose BCC, they used two datasets, a public dataset called "Skin Lesion Analysis Towards Melanoma Detection" of the ISBI 2016 and another dataset of the International Skin Imaging Collaboration (ISIC) in total they used 12,160 dermatoscopic images, in their results they reached 93% precision.

In contrast to our research, we used the dataset (HAM10000) with 19,267 dermatoscopic images and we achieved a better result of 96.20%

precision, in addition, we presented training graphs and the losses of the model.

In the study carried out by [10], the authors propose a method based on a scattered stacked autoencoder (SSAE), which is an unsupervised learning method with a fully connected architecture. They manage to detect BCC in the translucent areas of clinical lesion images. They obtained results of precision 93% and recall 77%. Comparing to our research, we use dermatoscopic images, when it comes to results, we manage to overcome them in precision: we reach 96.20% and recall 96.21%.

In the research carried out by [8], the authors present a segmentation method for the precise extraction of cutaneous blood vessels in dermatoscopic images for the classification of BCC. For this, they used a dataset composed of 659 images of normal dermatoscopy and cancer lesions. In their results, they reached 82.7% and 82.4% with the metrics of accuracy and sensitivity respectively, in addition, it is linked to the detection of blood vessels on skin lesions.

In our investigation, we managed to overcome their results with accuracy 96.20% and sensitivity 96.21%. As for the techniques used to classify the BCC as in [9], they used a convolutional neural network, which they adapted to VGG architecture ideas. In another investigation [10], a classical convolutional neural network was used, in [7], ResNet was used, and in [8], Random Forest was used, none of these proposals have proposed using a hybrid model to classify the BCC, comparing with our research, where we propose the CNN+SVM hybrid model and the results exceed the results of the classic model by 1%.

There are proposals for the CNN+SVM hybrid models as can be seen in [12], but they are for classifying images of microscopic bacteria, in [13], with the hybrid model they classified the wear state of drills based on images of perforated holes and in [14] with the hybrid model, they classified a person's gender by the way they walk. Previous research managed to overcome the classic Softmax model in terms of their results. Comparing our research, we classify the existence of the BCC with our CNN+SVM hybrid model, we also surpass the classic model by 1% and, in addition, our hybrid model has better convergence.

5 Conclusions and Future Work

In this investigation, we designed a CNN+SVM hybrid model to classify carcinoma (BCC). The hybrid model developed using the HAM10000 dataset has demonstrated better performance than the classic convolutional neural network to classify the BCC, as could be seen with the results: our hybrid model achieved approximately 1% improvement in all the proposed metrics and a remarkable improvement in the convergence of the model. As future work, different image segmentation techniques can be tested, which allows increasing the accuracy of the CNN+SVM hybrid model. This model can also be adapted to classify other types of carcinoma.

References

1. **Darias-Domínguez, C., Garrido-Celis, J. (2018).** Carcinoma basocelular: Un reto actual para el dermatólogo. Revista Médica Electrónica, Vol. 40, No. 1, pp. 172–182.

2. **American Cancer Society (2019).** Important statistics about melanoma skin cancer. https://www.cancer.org/es/cancer/cancer-de-piel-tipo-melanoma/acerca/estadisticas-clave.html.

3. **American Cancer Society (2019).** What are basal cell and squamous cell skin cancers?

4. **GCO (2019).** Cancer Today: Data visualization tools for exploring the global cancer burden in 2018. http://gco.iarc.fr/today/home.

5. **Kittler, H., Pehamberger, H., Wolff, K., Binder, M. (2002).** Diagnostic accuracy of dermoscopy. The Lancet Oncology, Vol. 3, No. 3, pp. 159–165. DOI: 10.1016/S1470-2045(02)00679-4.

6. **Binder, M., Puespoeck-Schwarz, M., Steiner, A., Kittler, H., Muellner, M., Wolff, K., Pehamberger, H. (1997).** Epiluminescence microscopy of small pigmented skin lesions: short-term formal training improves the diagnostic performance of dermatologists. Journal of the American Academy of Dermatology, Vol. 36, No. 2, pp. 197–202. DOI: 10.1016/S0190-9622(97)70280-9.

7. **Putten, E.V., Kambod, A., Kambod, M. (2018).** Deep residual neural networks for automated Basal Cell Carcinoma detection. IEEE EMBS International Conference on Biomedical & Health Informatics (BHI), pp. 337–340.

8. **Kharazmi, P., Lui, H., Wang, Z.J., Lee,T.K. (2016).** Automatic detection of basal cell carcinoma using

vascular-extracted features from dermoscopy images. IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), pp. 1–4.

9. **Mandache, D., Dalimier, E., Durkin, J.R., Boceara, C., Olivo-Marin, J.C., Meas-Yedid, V. (2018).** Basal cell carcinoma detection in full field OCT images using convolutional neural networks. IEEE 15th International Symposium on Biomedical Imaging, pp. 784–787. DOI: 10.1109/ISBI.2018. 8363689.

10. **Huang, H., Kharazmi, P., McLean, D.I., Lui, H., Wang, Z.J., Lee,T.K. (2018).** Automatic detection of translucency using a deep learning method from patches of clinical basal cell carcinoma images. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference pp. 685–688. DOI: 10.23919/APSIPA.2018.8659685.

11. **Agarap, A.F. (2017).** An Architecture Combining Convolutional Neural Network (CNN) and Support Vector Machine (SVM) for Image Classification.

12. **Ahmed, T., Wahid, M.F., Hasan, M.J. (2019).** Combining deep convolutional neural network with support vector machine to classify microscopic bacteria images. International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–5. DOI: 10.1109/ECACE. 2019.8679397.

13. **Kurek, J. (2018).** Hybrid approach towards the assessment of a drill condition using deep learning and the Support Vector Machine. 22nd International Computer Science and Engineering Conference (ICSEC), pp. 1–5. DOI: 10.1109/ICSEC.2018. 8712622.

14. **Liu, T., Ye, X., Sun, B. (2018).** Combining Convolutional Neural Network and Support Vector Machine for Gait-based Gender Recognition. Chinese Automation Congress (CAC), pp. 3477–3481. DOI: 10.1109/CAC.2018.8623118.

15. **Tang, Y. (2013).** Deep Learning using Linear Support Vector Machines. Contribution to the ICML´13 Challenges in Representation Learning Workshop.

16. **Yu, Z., Jiang, X., Zhou, F., Qin, J., Ni, D., Chen, S., lei, B., Wang, T. (2019).** Melanoma Recognition in Dermoscopy Images via Aggregated Deep Convolutional Features. IEEE Transactions on Biomedical Engineering, Vol. 66, No. 4, pp. 1006–1016. DOI: 10.1109/TBME.2018.2866166.

17. **Sultana, N.N., Mandal, B., Puhan, N.B. (2018).** Deep residual network with regularised fisher framework for detection of melanoma. IET Computer Vision, Vol. 12, No. 8, pp. 1096–1104.

18. **Sachdev, J., Shekhar, S., Indu, S. (2018)** Melanoma Screening Using Deep Neural Networks.

3rd International Conference for Convergence Technology (I2CT), pp. 1–5. DOI: 10.1109/I2CT. 2018.8529724.

19. **Bozorgtabar, B., Sedai, S., Roy, P.K., Garnavi, R. (2017).** Skin lesion segmentation using deep convolution networks guided by local unsupervised learning. IBM Journal of Research and Development, Vol. 61, No. 4/5, pp. 6:1-6:8. DOI: 10.1147/JRD. 2017.2708283.

20. **Waheed, Z., Waheed, A., Zafar, M., Riaz, F. (2017).** An efficient machine learning approach for the detection of melanoma using dermoscopic images. International Conference on Communication, Computing and Digital Systems (C-CODE), pp. 316–319. DOI: 10.1109/C-CODE.2017.7918949.

21. **Hagerty, J.R., Stanley, R.J., Almubarak, H.A., Lama, N., Kasmi, R., Guo, P., Drugge, R.J., Rabinovitz, H.S., Oliviero, M., Stoecker, W.V. (2019).** Deep Learning and Handcrafted Method Fusion: Higher Diagnostic Accuracy for Melanoma Dermoscopy Images. IEEE Journal of Biomedical and Health Informatics, Vol. 23, No. 4, pp. 1385–1391. DOI: 10.1109/JBHI.2019.2891049.

22. **Namozov, A., Cho, Y.I. (2018)** Convolutional Neural Network Algorithm with Parameterized Activation Function for Melanoma Classification. International Conference on Information and Communication Technology Convergence (ICTC), pp. 417–419. DOI: 10.1109/ICTC.2018.8539451.

23. **Ge, Y., Li, B., Zhao, Y., Guan, E., Yan, W. (2018).** Melanoma Segmentation and Classification in Clinical Images Using Deep Learning. ICMLC, pp. 252–256. DOI: 10.1145/3195106.3195164.

24. **Thao, L.T., Quang, N.H. (2017).** Automatic skin lesion analysis towards melanoma detection. (21st Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES), pp. 106–111. DOI: 10.1109/IESYS.2017.8233570.

25. **Wu, Z., Zhao, S., Peng, Y., He, X., Zhao, X., Huang, K., Wu, X., Fan, W., Li, F. (2019).** Studies on Different CNN Algorithms for Face Skin Disease Classification Based on Clinical Images. IEEE Access, Vol. 7, pp. 66505–66511. DOI: 10.1109/ ACCESS.2019.2918221.

26. **Das, D.K., Koley, S., Chakraborty, C., Maiti, A.K. (2014).** Automated segmentation of Mitotic Cells for in vitro histological evaluation of oral squamous cell carcinoma. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 354–357. DOI: 10.1109/ISSPIT.2014. 7300614.

27. **Calderon-Vilca, H.D., Ortega-Melgarejo, L.M., Cardenas-Mariño, F.C. (2020).** Tuberculosis Detection architecture with image processing using the sift and k-means algorithm. Journal Computación

y Sistemas, Vol. 24, No. 3, pp. 989–997. DOI: 10.1109/IESYS.2017.8233570.

28. **Serquen-Llallire, C.E., Calderon-Vilca H.D., Cardenas-Miño, F.C. (2018).** Comparison of two Algorithms for routing questions and answers, applied to group of students software engineering. IEEE Latin American Conference on Computational Intelligence (LA-CCI), pp. 1–6. DOI: 10.1109/LA-CCI.2018.8625262.

29. **Calderon-Vilca, H., Callupe, D.K.E.C., Aliaga, R. J.I., Cuba, J.B., Mariño-Cárdenas, F.C. (2019).** Early Cardiac Disease Detection Using Neural Networks. 7th International Engineering, Sciences and Technology Conference (IESTEC), pp. 562–567. DOI: 10.1109/IESTEC46403.2019.0010.

30. **Silva-Fuentes, M.A., Calderon-Vilca, H.D., Calderon-Vilca E.F., Cárdenas-Mariño, F.C. (2019).** Semantic Search System using Word Embeddings for query expansion. IEEE PES Innovative Smart Grid Technologies Conference - Latin America (ISGT Latin America), pp. 1–6. DOI: 10.1109/ISGT-LA.2019.8894992.

31. **Calderon-Vilca, H.D., Ocrospoma-Callupe, E.G., Cárdenas-Mariño, F.C., Calderon-Vilca, E.F. (2019).** Architecture and mobile application with augmented reality, visualizing videos and 3d objects in museums. IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON), pp. 1–5. DOI: 10.1109/CHILECON47746.2019.8988088.

32. **Calderon-Vilca, H., Chavez, N.M., Guimarey, J.M.R. (2020).** Recommendation of Videogames with Fuzzy Logic. 27th Conference of Open Innovations Association (FRUCT), pp. 27–37. DOI: 10.23919/FRUCT49677.2020.9211082.