IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 5, 2021, accepted February 21, 2021, date of publication February 24, 2021, date of current version March 5, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3061873

Multi-Class Diagnosis of Skin Lesions Using the Fourier Spectral Information of Images on Additive Color Model by Artificial Neural Network

JOSUÉ AARÓN LÓPEZ-LEYVA[1], ESPERANZA GUERRA-ROSAS[1], AND JOSUÉ ÁLVAREZ-BORREGO[2]

[1]CETYS Universidad, Centro de Innovación y Diseño (CEID), Ensenada 22860, Mexico
[2]CICESE, División de FÍsica Aplicada, Departamento de óptica, Ensenada 22860, Mexico

Corresponding author: Josué Aarón López-Leyva (josue.lopez@cetys.mx)

This work was supported by the Institutional Research Coordination, CETyS Universidad, under Project 221.

ABSTRACT This article presents a new methodology to diagnostics ten types of skin lesions based on the image' s Fourier spectral information in an additive color model. All spectral information and correlation coefficients between the skin lesions classes conform the input signals to an Artificial Neural Network. In general, the results show the well-defined classification for all the skin lesions classes based on the high values for Accuracy, Precision, Sensitivity, and Specificity metrics performance and a reduced images misclassification percentage (≈5.9%) for the Testing sub-dataset, and less for Training (≈2.8%) and Validation (≈5.6%) sub-dataset even considering the strange objects, not-clarity, and black sections in some images analyzed. The general achieved classification Accuracy, Precision, Sensitivity, and Specificity percentages of the proposed method are 99.33 %, 94.16 %, 92.9 %, and 99.63 %, respectively. In particular, the skin lesions related to *Basal Cell Carcinoma*, *Seborrhoeic Keratosis*, and *Melanocytic Nevus* present the best performance regarding the Receiver Operating Characteristics, while the *Pyogenic Granuloma* was the worst classified.

INDEX TERMS Artificial neural networks, biomedical computing, image processing, medical diagnostic imaging, Fourier spectral analysis.

I. INTRODUCTION

The skin is considered the largest organ in the human body, and it can present diverse diseases. In particular, some skin diseases can be very common, so they can present very similar symptoms and characteristics, i.e., the skin lesions can be apparently identical for different skin diseases, which can increase the misdiagnoses. For example, skin cancer is a disease caused by unrepaired Deoxyribonucleic acid (DNA) damage that triggers skin mutations which the dermatological evidence is the uncontrolled development of abnormal cells. Also, it is one of the conditions with the highest growth for populations with fair skin around the world in recent years [1].

In general, the skin tumors are classified into malignant melanomas and non-melanoma, and they rank according to the cell in which they originated [2]. The most common non-melanomas are basal cell carcinoma and squamous cell carcinoma. Regarding the malignant skin diseases, Melanoma is the most dangerous skin cancer but less frequently; if it is detected early, it can be curable, having a death rate of 85% of all fatal skin cancer cases [2], [3]. In particular, the main features of the skin lesions for Melanoma are asymmetrical shape, irregular edges, various shades of pigmentation, and its diameter can be greater than 6 mm. Concerning Basal cell carcinoma, which has slow growth and occurs most frequently, can appear on the face and neck as colored or pigmented bumps (e.g., blue, brown, or black), its size varies and can be up to 10 mm in diameter [4], [5]. Another skin disease is Squamous cell carcinoma that is presented as papules, keratotic plaques, or ulcers. Later nodules or ulcerated tumors can develop this carcinoma, which can arise in actinic keratosis [6], [7]. Actinic keratosis manifests as rough papules, patches, or plaques, its size is less than 10 mm, and its edges are irregular; they may itch and burn frequently and appears on the soles of the feet, the palms of

The associate editor coordinating the review of this manuscript and approving it for publication was G. R. Sinha.

the hands, and the face [8]. Also, Intraepithelial carcinoma is another of the malignant neoplasms that produce changes in the cells of the epithelium.

As aforementioned above, there are benign skin tumors among skin diseases such as dermatofibromas, pyogenic granulomas, hemangiomas, seborrhoeic keratosis, and melanocytic nevus. In particular, Dermatofibromas are papules and nodules well defined, usually asymptomatic; their size varies from a few millimeters to more than 2 cm [9]. Pyogenic granuloma is a lesion characterized by a raised nodule that can ulcerate, and presents a slow evolution [10]. Thus, all skin tumors both malign and benign present certain features.

In the dermatology area, visual inspection is the first stage for analyzing spots or skin lesions as part of the complete diagnosis process. However, different skin diseases can present very similar characteristics, such as anatomical distribution, color, and area, among others [11]–[13]. Therefore, it is important to detect skin diseases in the early stages based on particular skin lesions features, especially when they present similar characteristics between dangerous and non-malignant skin diseases. Thus, due to the importance of the early detection of skin diseases, different non-invasive methodologies have been developed for their identification, such as radiography, magnetic resonance, imaging, fluorescence, computed tomography, ultrasound, among others.

Regarding computer techniques, the Artificial Neural Networks (ANN) are digitized models that intent to reproduce the behavior of the human brain for diverse applications as recognizing patterns, classification of input data, among others, based on many learning techniques and algorithms [14]. In particular, the ANN has been used to analyze biological cells, invasive and non-invasive medical processes [15]–[18]. Concerning medical procedures related to the skin, the ANN has been used to study the multi-layered skin model, to analyze the skin in a hyperspectral mode, and quantify the skin damage, among other essential applications [19]–[21]. An important aspect is ANN's uses in medical processes using a limited training dataset, which represents a critical countermeasure related to the conventional ANN uses [22], [23]. Also, as a beginning step toward the ANN design is the definition of the optimal predictors related to the input information analyzed (e.g., signals or images), which can be in spatial, Fourier spectral, temporal domains, or a mixed domain. If the predictors are not well-defined, it is possible that the ANN does not detect patterns and, therefore, cannot perform the correct recognition and classification of the samples [24], [25].

According to the literature review, there are some proposals for multi-class skin lesion classifications using an ANN, signatures via spectral densities, fractional Fourier transform, Hermite transform, statistical information and Asymmetry, Border inconsistency, Color variety, and Diameter metrics (i.e., ABCD metrics). However, most of them consider a reduced class amount and do not consider the spectral information on an additive color model as input to an ANN [26]–[28].

TABLE 1. Database information of skin lesions images.

Class	Abbreviation	Type	Total (L)
1	*ALLMEL*	Malignant Melanoma	76
2	*AK*	Actinic Keratosis	45
3	*ALLIEC*	Intraepithelial Carcinoma	78
4	*ALLDF*	Dermatofibroma	65
5	*ALLSCC*	Squamous Cell Carcinoma	88
6	*PYO*	Pyogenic Granuloma	24
7	*ALVASC*	Hemangioma	97
8	*ALLBCC*	Basal Cell Carcinoma	239
9	*ALLSK*	Seborrhoeic Keratosis	257
10	*ALLML*	Melanocytic Nevus	331

In this work, discrimination from ten skin lesions is performed by calculating the 26 parameters (predictors) related to the Fourier spectral information of images on the RGB additive color model using a Two-Layer neural network. In particular, 16 parameters are associated with the 2-D Fourier Transform, results, i.e., the magnitude and phase for the original and cropped images. The rest parameters (i.e., ten predictors) are related to the correlation coefficient based on the probability density function for each sample image, feature, and class. Thus, all parameters mentioned are the input signals to the artificial neural network proposed in order to recognize patterns and classify each image according to the different skin lesions classes used.

The rest of this article is organized as follows. In Section II, the general methodology is explained. In particular, subsection A describe the public information about the images' database where the diverse skin lesion classes are shown, subsection B describes the features vector's determination process for each sample and type, subsection C explains the process for the determination of general correlation vector for each class and, subsection D presents the background needed in order to understand the neural network used for pattern recognition. Next, Section III presents the results and analysis, and Section IV shows the discussion and comparison with different methods. Finally, conclusions and recommendations are given in Section V.

II. METHODOLOGY

A. GENERAL INFORMATION ABOUT THE IMAGE DATABASE

Firstly, an image database of ten types of skin lesions was used. These types of skin lesions were established as the classes. The images were obtained from Edinburgh Dermofit Library (https://licensing.edinburgh-innovations.ed.ac.uk/i/software/dermofit-image-library.html). Table 1 shows the different categories, the name of the skin lesions, and its abbreviation, as well as the number of images for each class that makes up the database. In particular, the ALLBCC, ALLML, and ALLSK classes have the most significant number of images, 239, 331, and 257, respectively, while the remaining categories have a reduced number of images.

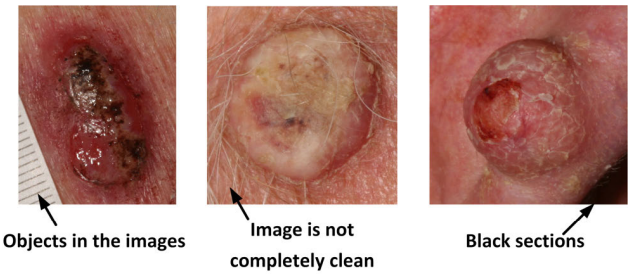

FIGURE 1. Irregularities in the images.

The aforementioned can be an important disadvantage for the traditional use of artificial neural networks; however, the determination of the optimal predictors for each class will be established as a countermeasure to this trade-off.

Because some classes have a large number of images, it is very common that many images present some differences, such as the dimension (related to the pixel per inch), brightness, orientation, among others. Also, Fig. 1 illustrates some image irregularities as stranger objects, black sections, and images not completely clean, among others. Even considering the above, our method can process all the images.

B. DETERMINATION OF FEATURES VECTOR FOR EACH SAMPLE AND CLASS

In this section, we present a detailed description of the process used to obtain each sample's features independently of the classes. Figure 2 shows the first stage of image digital processing, which will be described below. First, each original RGB image (f) with n rows x m columns of dimension for each class (M) is processed by the 2-D Fourier Transform, according to (1):

$$F(p,q) = \sum_{j=0}^{n-1}\sum_{k=0}^{m-1} e^{-i(\omega_p n+\omega_q m)} f(n,m), \tag{1}$$

where ω represents the spatial frequency between consecutive signal samples in each dimension n and m, ω_p and ω_q, respectively. As result, $F(p,q)$ is generated, and consequently, the magnitude ($|F(p,q)| \triangleq \sqrt{F_R(p,q)^2 + F_I(p,q)^2}$) and phase ($\angle|F(p,q)| \triangleq \tan^{-1}\left[F_I(p,q)/F_R(p,q)\right]$) are calculated based on the real and imaginary parts [29]. Next, $f(n,m)$ is decomposed in $f_R(n,m), f_G(n,m)$ and $f_B(n,m)$ in order to perform the 2-D Fourier Transform for each component as follows:

$$F_{R,G,B}(p,q) = \sum_{j=0}^{n-1}\sum_{k=0}^{m-1} e^{-i(\omega_p n+\omega_q m)} f_{R,G,B}(n,m), \tag{2}$$

thus, using $F_{R,G,B}(p,q)$, the magnitude and phase of each component ($|F_R(p,q)|$, $\angle|F_R(p,q)|$, $|F_G(p,q)|$, $\angle|F_G(p,q)|$, $|F_B(p,q)|$ and $\angle|F_B(p,q)|$) have to be calculated and recorded on the features database. Afterward, the original image (f) is cropped only at the center (f_c) to determine particular aspects of the skin lesion spot, considering a reduced image dimension and eliminating the most information about the health skin section. In this way, the 2-D Fourier Transform is performed for the cropped RGB image, which modifies the dimensions n, m, p, and q, as follows [30]:

$$F_c(p/2,q/2) = \times \sum_{j=0}^{(n/2)-1}\sum_{k=0}^{(m/2)-1} e^{-i(\omega_p n/2+\omega_q m/2)} f_c(n/2,m/2). \tag{3}$$

Next, the magnitude ($|F_c(p/2,q/2)|$) and phase ($\angle|F_c(p/2,q/2)|$) are calculated for the $F_c(p/2,q/2)$. Then, the Fourier spectral information for each color component of the cropped image are calculated, i.e., $F_{cR,cG,cB}(p/2,q/2)$, and also the magnitude and phase for each component are determined, i.e., $|F_{cR}(p/2,q/2)|$, $\angle|F_{cR}(p/2,q/2)|$, $|F_{cG}(p/2,q/2)|$, $\angle|F_{cG}(p/2,q/2)|$, $|F_{cB}(p/2,q/2)|$ and $\angle|F_{cB}(p/2,q/2)|$. Considering the digital processing carried out on the original and cropped images, both in RGB format and in each particular component, a Features Vector, $FV_{M,l}$ (16 rows and 1 column for each sample, l, of a specific class, M) for each image and category is created, e.g., 239 feature vectors for class ALLBCC, 331 feature vectors for class ALLML, and so on for all the images and classes.

In particular, until now, the Features Vector is composed of 8 values related to the magnitude and 8 values related to the phase as Fig. 2 shown. Regarding the magnitude values, 4 values are related to the original size image and the rest 4 values to the cropped image. In the same way, concerning the phase values. Besides, the 4 values of both magnitude and phase related to the original size image and cropped image are strictly related to the RGB image, R-component image, G-component image, and B-component image.

C. DETERMINATION OF GENERAL CORRELATION VECTOR FOR EACH CLASS

Figure 3 shows the second stage of the method proposed, where after having calculated the features vector ($FV_{M,l}$) for each one of 1,300 images (N) of the ten different classes (M), we proceed to calculate the features vector for each particular class ($V\mu_{F_M}$). Considering that there are 16 spectral indicators or predictors (W) for each image, therefore also $V\mu_{F_M}$ will have 16 spectral indicators that will define each class based on a specific probability density function (pdf) for each spectral indicator and class, i.e., $pdf(F_M(i))$ [31]–[33], as (4) shows:

$$pdf(F_M(i)) = pdf\left(FV_{M,l}(i), FV_{M,l+1}(i), \cdots, FV_{M,L}(i)\right). \tag{4}$$

Because the probability density function can be different for each indicator and class, the expected value for each spectral indicator and class, $\mu_{F_M(i)}$, have to be formally calculated as follows:

$$\mu_{F_M(i)} = \int_{-\infty}^{\infty} F_M(i)\, pdf(F_M(i) dF_M(i) \tag{5}$$

Next, all the particular values of $\mu_{F_M(i)}$ are concatenated to create a particular features vector for each class, such as

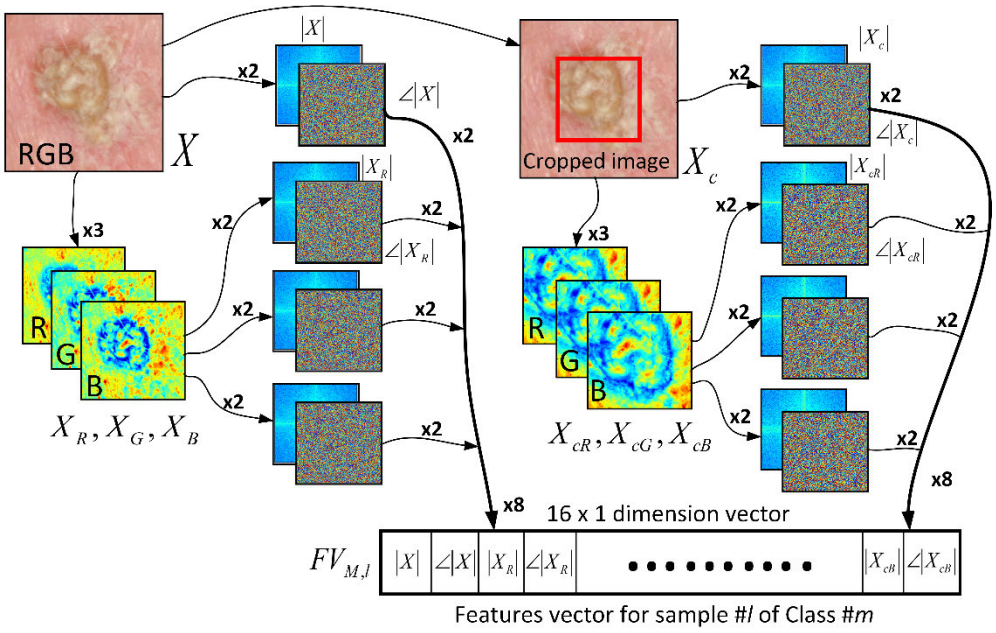

FIGURE 2. Determination of the features vector for each image of each class. The red square represents the cropped image.

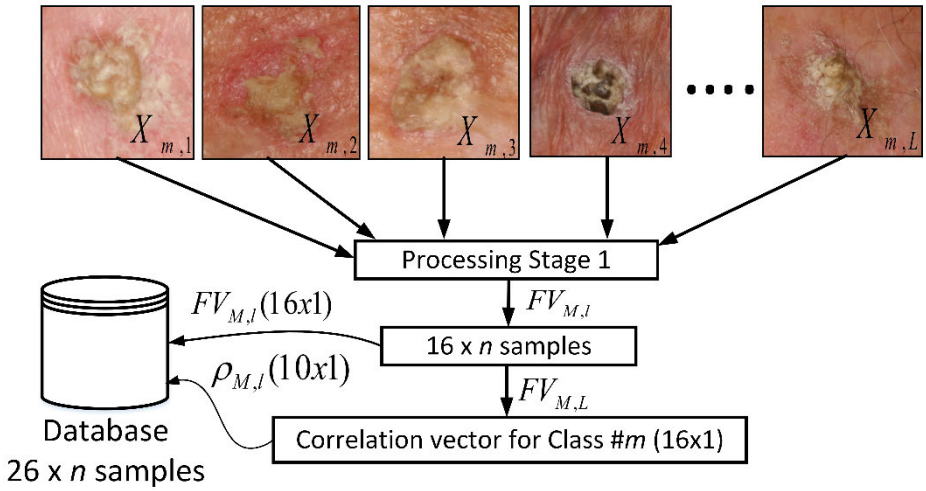

FIGURE 3. Determination of the correlation coefficients vectors.

$V\mu_{F_M} = (\mu_{F_M(1)}, \mu_{F_M(2)}, \cdots, \mu_{F_M(16)})$. Once the $V\mu_{F_M}$ of all the classes have been calculated, each particular sample of each class ($FV_{M,l}$) must be correlated with all the $V\mu_{F_M}$ to calculate the Pearson's correlation coefficient, as shown (6) [34], [35]:

$$\begin{aligned}\rho_{M,l} &= \rho(FV_{M,l}, V\mu_{F_M}) \\ &= \frac{1}{W-1}\sum_{i=1}^{W}\left(\frac{FV_{M,l}(i) - \mu_{FV_{M,l}}}{\sigma_{FV_{M,l}}}\right) \\ &\quad \times \left(\frac{V\mu_{F_M}(i) - \mu_{V\mu_{F_M}}}{\sigma_{V\mu_{F_M}}}\right) \end{aligned} \tag{6}$$

In this way, for each particular sample, there will be ten correlation coefficients. For example, for sample 1, $l = 1$, of class 1, $M = 1$, the complete correlation vector is, $\rho_{1,1} = (\rho\left(FV_{1,1}, V\mu_{F_1}\right), \rho\left(FV_{1,1}, V\mu_{F_2}\right), \cdots, \rho\left(FV_{1,1}, V\mu_{F_{10}}\right))$.

Finally, each image is describing by a 26 rows x 1 column vector, where the first 16 positions represent the Fourier spectral indicators related to the original size image and cropped image as was mentioned in Subsection B, and the ten positions latest the correlation coefficients between all classes. The aforementioned in subsection B and C represents Predictors Rule 1.

D. NEURAL NETWORK DESIGN FOR PATTERN RECOGNITION

Once the database is completed (26 *rows* x 1300 *columns*), a Two-Layer Feed-Forward Neural Network (TLFN) is designed (see Fig. 4) with 26 neurons in the input layer,

TABLE 2. Statistical information related to the classification process for all the classes.

Class	Abbrevi	Classified (%)	Misclassified (%)	Standard Deviation (%)
1	*ALLME*	97.8333	2.1666	± 0.6127
2	*AK*	92.5666	7.4333	± 1.0370
3	*ALLIEC*	93.0666	6.9333	± 4.0185
4	*ALLDF*	91.2666	8.7333	± 4.7590
5	*ALLSC*	96.6	3.4	± 0.8981
6	*PYO*	59.4	40.6	± 4.1012
7	*ALVAS*	94.4666	5.5333	± 0.2222
8	*ALLBC*	99.2333	0.7666	± 0.4713
9	*ALLSK*	98.7333	1.2666	± 0.4108
10	*ALLML*	99	1	± 0.1414

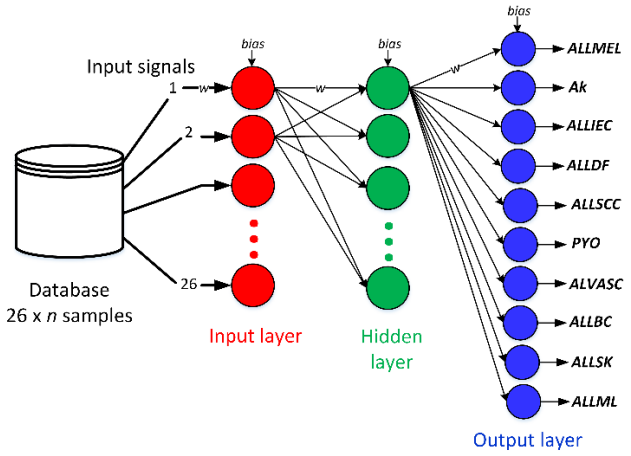

FIGURE 4. Two-layer feed-forward neural network design.

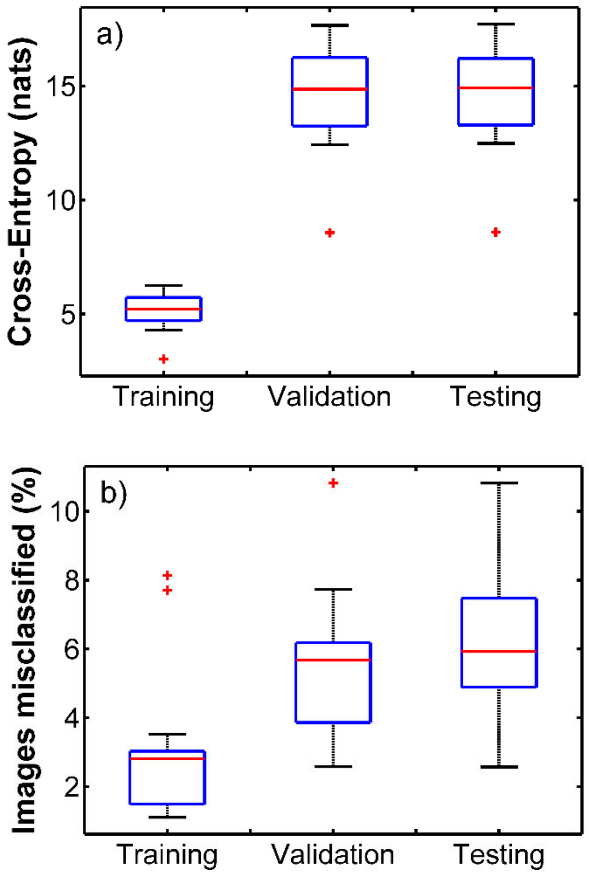

FIGURE 5. Cross-entropy and misclassified results for different sub-datasets.

100 neurons in the hidden layer allowing the neural network to learn more complicated features. In our case, we are using *sigmoid* functions, which is a non-linear activation function widely used in Logistic Regression processes and ANN applications. Also, the ANN proposed uses 10 neurons in the output layer using the *softmax* activation function. It is essential to mention that exists a lot of activation functions able to use on ANN proposals, such as Step, Linear, Hyperbolic Tangent, Rectified Linear Unit (ReLU), Leaky-ReLU, Swish functions. Each of them presents advantages and disadvantages related to the activation range, computational load, ANN convergence, among others [36]–[38].

Besides, it is important to mention that the ANN will be trained. Therefore, the weight (w) and *bias* values will be updated at each iteration independently to minimize the difference between actual output and desired output using the Scaled Conjugate Gradient Backpropagation (SCGB) method [39], [40]. In particular, the sub-dataset used for the Training stage was 70% (910 images) of the entire samples dataset, 15% for both Validation (195 images) and Testing (195 images). However, each stage's portion of images can be modified, maintaining the Training stage's more significant image amount. All sub-datasets considered all the classes, and the image samples used for each sub-dataset were randomly assigned.

III. RESULTS AND ANALYSIS

The complete algorithm runs 40 times to determine an accurate statistical performance of the 10 classes mentioned. Firstly, the results related to the Cross-Entropy (natural unit of information, *nats*) and the percentage of images that were not correctly classified in each stage of the artificial neural network design are shown, i.e., Training, Validation, and Testing stage [41]. In general, Fig. 5 shows the boxplots as a method of descriptive statistics for each Training, Validation, and Testing sub-dataset. Figure 5a) shows that for the Training stage, the cross-entropy mean value was ≈5.21, which is a good result since it should tend towards zero

TABLE 3. Performance metrics for training, validation, and testing sub-dataset.

Class	Abbreviation	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
		Training			
1	*ALLMEL*	99.78 ± 0	97.33 ± 0.90	99.12 ± 0.87	99.82 ± 0.05
2	*AK*	99.78 ± 0	94.94 ± 1.61	98.27 ± 1.72	99.82 ± 0.06
3	*ALLIEC*	99.44 ± 0.11	95.74 ± 0.09	93.82 ± 1.82	99.76 ± 0.05
4	*ALLDF*	99.78 ± 0.11	95.82 ± 1.95	100 ± 0	99.77 ± 0.11
5	*ALLSCC*	99.50 ± 0.05	95.58 ± 1.64	97.81 ± 0.59	99.64 ± 0.12
6	*PYO*	99.51 ± 0.05	100 ± 0	74.34 ± 2.12	100 ± 0
7	*ALVASC*	99.28 ± 0.05	93.21 ± 1.15	97.01 ± 1.29	99.46 ± 0.05
8	*ALLBCC*	100 ± 0	100 ± 0	100 ± 0	100 ± 0
9	*ALLSK*	99.72 ± 0.05	99.71 ± 0.28	98.91 ± 0.51	99.93 ± 0.68
10	*ALLML*	99.78 ± 0	99.78 ± 0.22	99.34 ± 0.21	99.99 ± 0.08
		Validation			
1	*ALLMEL*	99.22 ± 0.25	92.58 ± 0.27	96.15 ± 3.84	99.44 ± 0
2	*AK*	99.48 ± 0.51	100 ± 0	88.89 ± 11.11	100 ± 0
3	*ALLIEC*	97.42 ± 0.52	78.12 ± 3.12	78.12 ± 3.12	98.62 ± 0.30
4	*ALLDF*	97.93 ± 0.51	73.46 ± 3.46	91.66 ± 8.33	98.37 ± 0.02
5	*ALLSCC*	99.74 ± 0.25	100 ± 0	94.44 ± 5.55	100 ± 0
6	*PYO*	99.22 ± 0.26	75 ± 25%	62.5 ± 12.5	99.74 ± 0.26
7	*ALVASC*	97.42 ± 0.15	82.98 ± 10.76	87.5 ± 6.25	98.31 ± 1.12
8	*ALLBCC*	98.71 ± 0.77	97.22 ± 2.77	95.94 ± 1.49	99.36 ± 0.63
9	*ALLSK*	99.48 ± 0	98.61 ± 1.38	98.64 ± 1.35	99.68 ± 0.31
10	*ALLML*	99.49 ± 0	100 ± 0	97.95 ± 0	100 ± 0
		Testing			
1	*ALLMEL*	98.19 ± 0.25	74.60 ± 3.17	80.55 ± 2.77	98.92 ± 0001
2	*AK*	97.42 ± 1.03	65.27 ± 9.72	74.10 ± 11.60	98.39 ± 0.54
3	*ALLIEC*	96.64 ± 0.25	83.75 ± 3.75	76.84 ± 3.15	98.59 ± 0.26
4	*ALLDF*	97.68 ± 0.26	78.02 ± 6.59	75 ± 3.57	98.91 ± 0.024
5	*ALLSCC*	98.71 ± 0.25	79.44 ± 9.44	100 ± 0	98.63 ± 0.021
6	*PYO*	98.96 ± 0	75 ± 25	41.66 ± 8.33	99.74 ± 0.023
7	*ALVASC*	97.93 ± 0.51	86.90 ± 3.57	88.63 ± 2.27	98.8 ± 0.03
8	*ALLBCC*	98.96 ± 0.50	97.14 ± 2.85	96.60 ± 0.45	99.37 ± 0.62
9	*ALLSK*	99.48 ± 0	98.86 ± 1.13	98.52 ± 1.47	99.66 ± 0.33
10	*ALLML*	99.48 ± 0	100 ± 0	98.09 ± 0.09	100 ± 0

for a perfect classification. However, for the Validation and Testing stages, this value was increased, ≈14.8 and ≈14.9, respectively. The minimum and maximum values for Validation and Testing were similar, ≈8.6 and ≈17.7, respectively. Regarding the classification error (see Fig. 5b)), the result for the Training stage was a mean value of ≈2.8% (25 images).

In contrast, for the Validation and Testing stage, the results were ≈5.6% (11 images) and ≈5.9% (12 images), respectively. The Training stage also presents a minimum and maximum value of 1.101% and 8.14%, respectively. In comparison, the Validation stage presents a minimum and maximum value of 2.57% and 10.82%, respectively. These last values were similar to the Testing stage. Table 2 shows the related statistical information for all classes, regarding the percentage of samples classified correctly and incorrectly, and the standard deviation of classification. To clarify and to facilitate the analysis, all the tables presented in this article are based on the information of the 40 confusion matrices generated according to the 40 algorithm runs. For example, the class that presents the best performance regarding recognizing patterns and classification was Class 8, ALLBCC, which gave a mean value of correct classification of 99.23%, a percentage of erroneous classification of 0.76%, as well as a minimum standard deviation, ± 0.4713.

Although the results also show that several classes (e.g., ALLMEL, ALLSCC, ALLBCC, ALLSK, and ALLML) present important results, with a mean value greater than 95% for samples classified correctly. On the other hand, the class that presents the worst classification performance was Class 6 (PYO), with a percentage of correct classification of 59.4% and a standard deviation of ± 4.1012. Besides, according to data science in medicine, evaluating the performance metrics for the classification for all the classes is necessary. For that, the Accuracy (A), Precision (P), Sensitivity (S), and Specificity (SP) metrics were obtained based on all possibilities regarding the classifications such as the True Positive (TP) classification, False Positive (FP) classification, False Negative (FN) classification and the True Negative (TN) classification [42]. These parameters (i.e., TP, FP, FN, and TN) are used to calculate the performance metrics as follow:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%, \qquad (7)$$

TABLE 4. General performance metrics of the ANN.

Class	Abbreviation	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)
1	*ALLMEL*	99.46 ± 0.07	94.23 ± 0.64	96.71 ± 0.65	99.63 ± 0.04
2	*AK*	99.38 ± 0.07	90.34 ± 2.84	92.22 ± 1.11	99.64 ± 0.12
3	*ALLIEC*	98.72 ± 0.03	90.63 ± 1.02	87.66 ± 1.94	99.42 ± 0.08
4	*ALLDF*	99.19 ± 0.11	89.79 ± 1.38	94.61 ± 0.76	99.43 ± 0.08
5	*ALLSCC*	99.42 ± 0.03	93.99 ± 0.51	97.72 ± 0.0001	99.54 ± 0.04
6	*PYO*	99.38 ± 0.07	94.09 ± 0.34	69.56 ± 4.34	99.92 ± 0.0002
7	*ALVASC*	98.80 ± 0.19	90.55 ± 1.44	93.81 ± 1.03	99.20 ± 0.12
8	*ALLBCC*	99.65 ± 0.03	99.15 ± 0.0001	98.95 ± 0.21	99.81 ± 0.0001
9	*ALLSK*	99.66 ± 0.03	99.41 ± 0.19	98.82 ± 0.39	99.85 ± 0.04
10	*ALLML*	99.69 ± 0.01	99.48 ± 0.15	98.94 ± 0.15	99.94 ± 0.05

$$Precision = \frac{TP}{TP + FP} \times 100\%, \tag{8}$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\%, \tag{9}$$

$$Specificity = \frac{TN}{TN + FP} \times 100\%, (10) \tag{10}$$

where $TP + FP$ represents the total number of images with a positive test, $TP + FN$ is the total number of images with a given condition, $TN + FP$ is the total number of images without a given condition, $FN + TN$ represents the total number of images with a negative test, and $TP + TN + FP + FN$ is the total of images in the study. Table 3 shows the performance metrics values obtained for each sub-dataset, stage (i.e., Training, Validation, and Testing), and class. In particular, Class 6 present the worst performance for all ANN stages. Table 4 shows the performance metrics for the ANN considering all the sub-datasets, where can be seen that, concerning the Accuracy and Specificity parameters, all the classes presented adequate results, $A > 98\%$ and $SP > 99\%$, while the Sensitivity parameter presented the lowest value for Class 6 (PYO), which is $69.56 \pm 4.34\%$. In this way, the results show that Class 6 (PYO) has the lowest performance in most of the metrics analyzed. Afterward, the Receiver Operating Characteristics (ROC) analysis is determinate [43]. For that, the True Positive Rate (*TPR*) and False Positive Rate (*FPR*) are calculated in order to determine if the results classifications are relevant or not according to the relationship between both metrics. Thus, it is important to mention that *TPR* is equivalent to the Sensitivity metric mentioned, while *FPR* is calculated as follows [44]:

$$FPR = \frac{FP}{FP + TN}. \tag{11}$$

In particular, the numerical pair (*FPR*, *TPR*) defined as P_{ROC} describes the quality or performance of the classification, e.g., if the numerical pair is (0,1), it means that the classifications were completed successfully or correctly because the Sensitivity and Specificity presented the higher possible value (i.e., 100 %). In general, the ROC space related to *FPR* and *TPR* is given to determine the classification quality [45]. However, in our case, two related parameters will be used as an alternative way to the conventional ROC representation, such as 1) Slope (m) between the

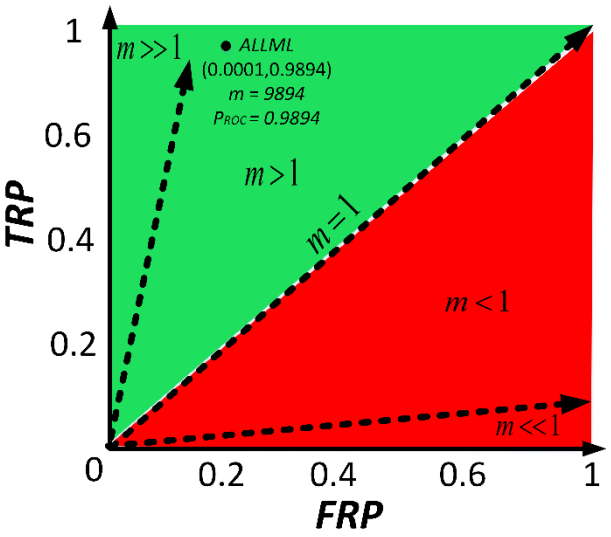

FIGURE 6. ROC space based on FRP and TRP values.

FPR and *TPR* and 2) the magnitude of the numerical pair, $P_{ROC}\sqrt{FPR^2 + TPR^2}$. Thus, to determine if the classifications were relevant or right (i.e., useful diagnostic category), two conditions have to be satisfied: 1) m value for each class must be much higher than 1, i.e., $m = TPR/FPR \gg 1$, and 2) the magnitude value has to be close to unity, $P_{ROC} = 1$. The first condition sometimes is relaxed, considering that the m value has to be greater than 1.

Figure 6 shows the diverse cases related to matters of m. In particular, $m = 1$ means a random classification, i.e., the diagnostic has the same value for both Sensitivity and Specificity metrics. When $m < 1$, the classifications are not considered adequate (i.e., lousy prediction). Besides, if $m \gg 1$, the classifications are considered very good, while for $m \ll 1$, the classifications are regarded as highly non-adequate. It is still necessary to analyze the P_{ROC} value, because the final performance has to consider the two conditions mentioned.

Next, Table 5 shows the m and P_{ROC} values (related to *FPR* and *TPR*) for all the classes. In particular, the m values for all classes are greater than 1, so this can be considered that all classes are adequately classified. The classes best ranked according to m value are those related to ALLSK

TABLE 5. True positive rate, false positive rate, and slope values for each class.

Class	Abbreviation	*FPR*	*TPR*	*m*	P_{ROC}
1	*ALLMEL*	0.0041	0.9671	235.8780	0.9671
2	*AK*	0.0022	0.9222	419.1818	0.9222
3	*ALLIEC*	0.0063	0.8766	139.1428	0.8766
4	*ALLDF*	0.0041	0.9461	230.7560	0.9461
5	*ALLSCC*	0.0052	0.9772	187.9230	0.9772
	PYO	0.0007	0.6956	993.71428	0.6956
7	*ALVASC*	0.0091	0.9381	103.0879	0.9381
8	*ALLBCC*	0.0018	0.9895	549.7222	0.9895
9	*ALLSK*	0.0009	0.9882	1098	0.9882
10	*ALLML*	0.0001	0.9894	9894	0.9894

and ALLML skin lesions. In contrast, the classes with the lowest detection quality were associated with ALVASC and ALLIEC. Even with this, all the classes present a useful classification considering the slope value. Regarding the P_{ROC} value, Classes 8, 9, and 10 (ALLBCC, ALLSK, and ALLML, respectively) present the best performance, 0.9895, 0.9882, and 0.9884, respectively, while Class 6 (PYO) shows the worst performance, 0.6956. In general, and considering the aforementioned related to both parameters, m and P_{ROC}, Class 10 presents the best performance, $m = 9894$ and $P_{ROC} = 0.9894$, while Class 6 presents the worst performance, mainly based on the that $P_{ROC} = 0.6956$. Due to the *FPR* values are so close to zero, the P_{ROC} value is very similar to the *TPR* value for all the classes.

IV. DISCUSSION

Figure 2 shows that the proposed method considers the central cropped image as part of spectral analysis. In this case, the spectral results of a smaller sample image compared with a larger original image presents added-noise due to the spectral windows used by the 2-D Fourier Transform in both images with different dimensions. As another option for not crop an image, segmentation algorithms are usually used to divide an original image into other parts or segments (i.e., segment only the spot representing the skin lesion).

However, the added-noise is still present due to the reduced image dimension. Thus, to clarify our proposal, our method does not compare the Fourier spectral information of images with different dimensions, e.g., the original image and cropped image. Instead, the process compares the spectral information of images with similar dimensions belonging to the same and different classes. Therefore, it is necessary to discuss the impact of the added-noise on the proposed method's overall performance. For comparing the results of the proposed method (based on Predictors Rule 1, PR1), another predictors rule (Predictors Rule 2, PR2) was used, which consist of three intensity profiles for each axis (Horizontal Intensity Profile, HIP, and Vertical Intensity Profile, VIP) for each image. Each intensity profile is decomposed on the real and imaginary parts related to the 2-D Fourier Transform results. In particular, each intensity profile is separated 1/8 of the original image dimension for each axis, as Fig. 7 shows. The analysis was realized in the gray-scale image and RGB channels. Hence, each image sample is described by 48 predictors related to the different intensity profiles. In other words, each image is represented by 12 predictors for each component R, G, B, and RGB, six predictors related to the real part (Re) of three intensity profiles for both horizontal and vertical axes, and the rest predictors are related to the imaginary part (Im) of three intensity profiles on both axes.

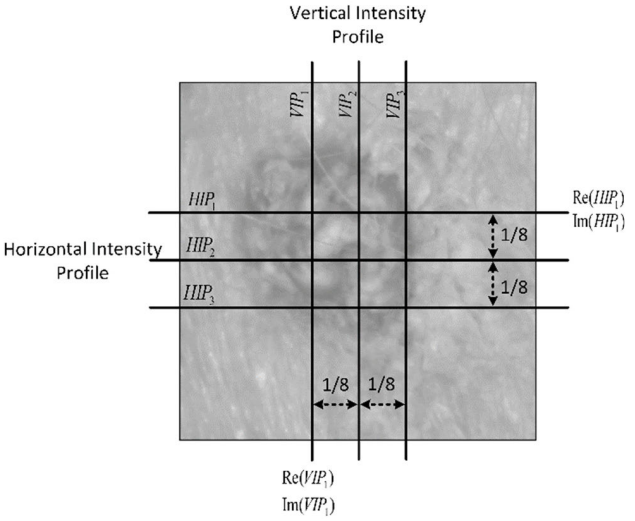

FIGURE 7. Graphical description of the intensity profiles.

TABLE 6. Statistical information related to the classification process for all the classes considering the comparison between different predictors rules.

Class	Abbreviation	Classified (%)		Misclassified (%)		Standard Deviation (%)	
		PR1	PR2	PR1	PR2	PR1	PR2
1	*ALLMEL*	97.8333	21.7105	2.1666	78.2894	± 0.6127	± 6.5010
2	*AK*	92.5666	11.1111	7.4333	88.8888	± 1.0370	± 5.1120
3	*ALLIEC*	93.0666	3.89610	6.9333	96.1038	± 4.0185	± 1.2311
4	*ALLDF*	91.2666	5.3846	8.73333	94.6153	± 4.7590	± 2.5012
5	*ALLSCC*	96.6	15.3409	3.4	84.6590	± 0.8981	± 0.5101
6	*PYO*	59.4	10.8695	40.6	89.1304	± 4.1012	± 0.5021
7	*ALVASC*	94.4666	16.4948	5.5333	83.5051	± 0.2222	± 0.0121
8	*ALLBCC*	99.2333	20.1680	0.7666	79.8319	± 0.4713	± 10.1241
9	*ALLSK*	98.7333	37.1093	1.2666	62.8906	± 0.4108	± 9.0110
10	*ALLML*	99	71.2990	1	28.7009	± 0.1414	± 11.1012

Thus, the input signals to the ANN used are represented by a database of 48 *rows* x 1300 *columns* considering the Predictors Rule 2. Finally, Table 6 shows the performance comparison between the predictors' rules mentioned. As can be seen, the percentage of correct classification for each class is much higher for the first predictors rule. Besides, the standard deviation is much higher for Predictors Rule 2. In general, considering the aforementioned above, the performance of the first predictor rule is better for the classification of various skin lesions, even considering the process of cropping the image and the added-noise. In particular, Class 10 presents the best performance for both predictors rules, 99% and 71.2990 % for images correctly classified for PR1 and PR2, respectively.

Also, the general achieved classification Accuracy, Precision, Sensitivity, and Specificity percentages of the proposed method considering the PR1 are 99.33 %, 94.16 %, 92.9 %, and 99.63 %, respectively, which represents a better performance than the reported in [46]. Also, the method proposed allows to classified a higher number of classes (i.e. ten) in comparison to [46], [47], but our proposal cannot detect images that do not belong to any of the ten classes, which is performed and reported in [46].

V. CONCLUSION

This article presents a new methodology for the discrimination of 10 skin lesions based on Fourier spectral information of images on the RGB additive color model using a Two-Layer neural network. The method used required 16 spectral parameters and ten correlation coefficients to generate a highly representative database for all skin lesions classes. The final classification performance for each category was successful based on the values of the Accuracy, Precision, Sensitivity, Specificity, and Receiver Operating Characteristics metrics. The best performance was for the skin lesions images related to *Basal Cell Carcinoma*, *Seborrhoeic Keratosis*, and *Melanocytic Nevus*, i.e., Class 8, 9, and 10, respectively, while that the worst performance was for *Pyogenic Granuloma* skin lesion, i.e., Class 6. As a recommendation that can improve the performance of the proposed methodology, is the optimization of the predictors, that is, reduce the dimension of the features vector while maintaining or improving performance, to prepare the proposed methodology for the development of realistic fast applications in the healthcare sector based on affordable hardware and software.

ACKNOWLEDGMENT

Esperanza Guerra-Rosas hold a Postdoctoral Position in CETyS Universidad supported by CONACyT scholarship in the framework of the project related to cancer detection based on digital processing techniques.

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest related to this article.

REFERENCES

[1] *Cancer Facts and Figures 2020*, Amer. Cancer Soc., Atlanta, GA, USA, 2020.
[2] J. Jaworek-Korjakowska and P. Kleczek, "eSkin: Study on the smartphone application for early detection of malignant melanoma," *Wireless Commun. Mobile Comput.*, vol. 2018, pp. 1–11, Jan. 2018.
[3] D. S. Rigel, J. Robinson, M. Ross, R. Friedman, C. Cockerell, H. Lim, and J. Kirkwood, *Cancer of the Skin*. Philadelphia, PA, USA: Saunders, 2011.
[4] A. E. Carsin, L. Sharp, and H. Comber, "Geographical, urban/rural and socioeconomic variations in nonmelanoma skin cancer incidence: A population-based study in Ireland," *Brit. J. Dermatol.*, vol. 164, no. 4, pp. 822–829, Apr. 2011.
[5] K. Sellheyer, "Basal cell carcinoma: Cell of origin, cancer stem cell hypothesis and stem cell markers," *Brit. J. Dermatol.*, vol. 164, no. 4, pp. 696–711, Apr. 2011.
[6] O. Warszawik-Hendzel, M. Olszewska, M. Maj, A. Rakowska, J. Czuwara, and L. Rudnicka, "Non-invasive diagnostic techniques in the diagnosis of squamous cell carcinoma," *J. Dermatol. Case Rep.*, vol. 9, no. 4, pp. 89–97, Dec. 2015.
[7] K. Sardana, P. Chakravarty, and K. Goel, "Optimal management of common acquired melanocytic nevi (moles): Current perspectives," *Clin., Cosmetic Invest. Dermatol.*, vol. 7, pp. 89–103, Mar. 2014.
[8] J. B. Travers, C. Poon, D. J. Rohrbach, N. M. Weir, E. Cates, F. Hager, and U. Sunar, "Noninvasive mesoscopic imaging of actinic skin damage using spatial frequency domain imaging," *Biomed. Opt. Exp.*, vol. 8, no. 6, pp. 3045–3052, Jun. 2017.
[9] A. H. Abdelhakim, K. E. Dunbar, K. J. Godfrey, C. A. Ananza, D. N. Silvers, and M. Kazim, "Unique histopathologic features of the eyelid dermatofibroma," *Orbit*, vol. 38, no. 4, pp. 274–278, Jul. 2019.
[10] S. Sharma, S. Chandra, S. Gupta, and S. Srivastava, "Heterogeneous conceptualization of etiopathogenesis: Oral pyogenic granuloma," *Nat. J. Maxillofacial Surg.*, vol. 10, no. 1, pp. 3–7, 2019.
[11] L. Vaienti, L. Lanfranchi, and R. Bosco, "The role of skin expansion: Giant nevi of the hand," *Pediatria Med. Chirurgica*, vol. 24, no. 5, pp. 377–382, 2002.
[12] A. S. Karadag and L. C. Parish, "The status of the seborrheic keratosis," *Clin. Dermatol.*, vol. 36, no. 2, pp. 275–277, Mar. 2018.
[13] M. Durgun, C. Tayyar-Selçuk, B. Özalp, M. Aydınol, and U. Alabalık, "Multiple disseminated pyogenic granuloma after second degree scald burn: A rare two case," *Int. J. Burns Trauma*, vol. 3, no. 2, pp. 125–129, Apr. 2013.
[14] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad, "State-of-the-art in artificial neural network applications: A survey," *Heliyon*, vol. 4, no. 11, Nov. 2018, Art. no. e00938.
[15] G. Dardikman-Yoffe, D. Roitshtain, S. K. Mirsky, N. A. Turko, M. Habaza, and N. T. Shaked, "PhUn-Net: Ready-to-use neural network for unwrapping quantitative phase images of biological cells," *Biomed. Opt. Exp.*, vol. 11, no. 2, pp. 1107–1121, Feb. 2020.
[16] H. Ding, Q. Lu, H. Gao, and Z. Peng, "Non-invasive prediction of hemoglobin levels by principal component and back propagation artificial neural network," *Biomed. Opt. Exp.*, vol. 5, no. 4, pp. 1145–1152, Mar. 2014.
[17] L. Mukherjee, A. Keikhosravi, D. Bui, and K. W. Eliceiri, "Convolutional neural networks for whole slide image superresolution," *Biomed. Opt. Exp.*, vol. 9, no. 11, pp. 5368–5386, Nov. 2018.
[18] Y. Liu, "Artificial intelligence–based neural network for the diagnosis of diabetes: Model development," *JMIR Med. Informat.*, vol. 8, no. 5, May 2020, Art. no. e18682.
[19] S.-Y. Tsui, C.-Y. Wang, T.-H. Huang, and K.-B. Sung, "Modelling spatially-resolved diffuse reflectance spectra of a multi-layered skin model by artificial neural networks trained with Monte Carlo simulations," *Biomed. Opt. Exp.*, vol. 9, no. 4, pp. 1531–1544, Apr. 2018.
[20] J. B. Travers *et al.*, "Quantifying skin photodamage with spatial frequency domain imaging: Statistical results," *Biomed. Opt. Exp.*, vol. 10, no. 9, pp. 4676–4683, Sep. 2019.
[21] E. Zherebtsov, V. Dremin, A. Popov, A. Doronin, D. Kurakina, M. Kirillin, I. Meglinski, and A. Bykov, "Hyperspectral imaging of human skin aided by artificial neural networks," *Biomed. Opt. Exp.*, vol. 10, no. 7, pp. 3545–3559, Jul. 2019.
[22] D. Li, J. Wu, Y. He, X. Yao, W. Yuan, D. Chen, H.-C. Park, S. Yu, J. L. Prince, and X. Li, "Parallel deep neural networks for endoscopic OCT image segmentation," *Biomed. Opt. Exp.*, vol. 10, no. 3, pp. 1126–1135, Mar. 2019.

[23] T. S. Namboodiri and A. Jayachandran, "Multi-class skin lesions classification system using probability map based region growing and DCNN," *Int. J. Comput. Intell. Syst.*, vol. 13, no. 1, pp. 77–84, 2020.
[24] A. A. Ding, "Neural-network prediction with noisy predictors," *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 1196–1203, Sep. 1999.
[25] D. Baptista and F. Morgado-Dias, "A survey of artificial neural network training tools," *Neural Comput. Appl.*, vol. 23, nos. 3–4, pp. 609–615, Sep. 2013.
[26] E. Guerra-Rosas, J. Álvarez-Borrego, and A. Angulo-Molina, "Identification of melanoma cells: A method based in mean variance of signatures via spectral densities," *Biomed. Opt. Exp.*, vol. 8, no. 4, pp. 2185–2194, Mar. 2017.
[27] E. Guerra-Rosas and J. Álvarez-Borrego, "Methodology for diagnosing of skin cancer on images of dermatologic spots by spectral analysis," *Biomed. Opt. Exp.*, vol. 6, no. 10, pp. 3876–3891, Oct. 2015.
[28] E. Garza-Flores, E. Guerra-Rosas, and J. Álvarez-Borrego, "Spectral indexes obtained by implementation of the fractional Fourier and Hermite transform for the diagnosis of malignant melanoma," *Biomed. Opt. Exp.*, vol. 10, no. 12, pp. 6043–6056, Nov. 2019.
[29] A. Averbuch, R. R. Coifman, D. L. Donoho, M. Elad, and M. Israeli, "Fast and accurate polar Fourier transform," *Appl. Comput. Harmon. Anal.*, vol. 21, no. 2, pp. 145–167, Sep. 2006.
[30] X. Li, M. Fang, J.-J. Zhang, and J. Wu, "Learning coupled classifiers with RGB images for RGB-D object recognition," *Pattern Recognit.*, vol. 61, pp. 433–446, Jan. 2017.
[31] C. K. Sirajuddeen, S. Kansal, and R. K. Tripathi, "Adaptive histogram equalization based on modified probability density function and expected value of image intensity," *Signal, Image Video Process.*, vol. 14, no. 1, pp. 9–17, Jun. 2019.
[32] K. Arai, "Image classification considering probability density function based on simplified beta distribution," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 4, pp. 481–486, 2020.
[33] S. Ghafurian, I. Hacihaliloglu, D. N. Metaxas, V. Tan, and K. Li, "A computationally efficient 3D/2D registration method based on image gradient direction probability density function," *Neurocomputing*, vol. 229, pp. 100–108, Mar. 2017.
[34] D. Sasikala and R. Neelaveni, "Correlation coefficient measure of mono and multimodal brain image registration using fast Walsh Hadamard transform," *Int. J. Eng. Technol.*, vol. 3, no. 2, pp. 154–160, Apr. 2011.
[35] S. Kaneko, Y. Satoh, and S. Igarashi, "Using selective correlation coefficient for robust image registration," *Pattern Recognit.*, vol. 36, no. 5, pp. 1165–1173, May 2003.
[36] M. Desai and M. Shah, "An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and convolutional neural network (CNN)," *Clin. eHealth*, vol. 4, pp. 1–11, Jan. 2021.
[37] J. Dheeba, N. A. Singh, and S. T. Selvi, "Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach," *J. Biomed. Informat.*, vol. 49, pp. 45–52, Jun. 2014.
[38] H. Zhang, W. Wu, and D. Wang, "Multi-instance multi-label learning of natural scene images: Via sparse coding and multi-layer neural network," *IET Comput. Vis.*, vol. 12, no. 3, pp. 305–311, Apr. 2018.
[39] M. Kawato, T. Setoyama, and R. Suzuki, "Feedback error learning of movement by multi-layer neural network," *Neural Netw.*, vol. 1, p. 342, Jan. 1988.
[40] A. E. Kostopoulos and T. N. Grapsa, "Self-scaled conjugate gradient training algorithms," *Neurocomputing*, vol. 72, no. 13, pp. 3000–3019, Aug. 2009.
[41] V. Vanchurin, "The world as a neural network," *Entropy*, vol. 22, no. 11, p. 1210, 2020.
[42] R. Deep, *Probability and Statistics*. Cambridge, MA, USA: Academic, 2006, pp. 19–22, 104, and 287.
[43] M. L. Meistrell, "Evaluation of neural network performance by receiver operating characteristic (ROC) analysis: Examples from the biotechnology domain," *Comput. Methods Programs Biomed.*, vol. 32, no. 1, pp. 73–80, May 1990.
[44] A. El Kaid, K. Baïna, and J. Baïna, "Reduce false positive alerts for elderly person fall video-detection algorithm by convolutional neural network model," *Procedia Comput. Sci.*, vol. 148, pp. 2–11, Jan. 2019.
[45] J. Hernández-Orallo, P. Flach, and C. Ferri, "ROC curves in cost space," *Mach. Learn.*, vol. 93, no. 1, pp. 71–91, Feb. 2013.
[46] M. A. Kassem, K. M. Hosny, and M. M. Fouad, "Skin lesions classification into eight classes for ISIC 2019 using deep convolutional neural network and transfer learning," *IEEE Access*, vol. 8, pp. 114822–114832, Jun. 2020.
[47] K. M. Hosny, M. A. Kassem, and M. M. Fouad, "Classification of skin lesions into seven classes using transfer learning with AlexNet," *J. Digit. Imag.*, vol. 33, no. 5, pp. 1325–1334, Jun. 2020.

JOSUÉ AARÓN LÓPEZ-LEYVA was born in Sonora, Mexico. He received the B.S. degree in engineering in electronics from ITESCA and the Ph.D. degree in electronic and telecommunication from CICESE. His research interests include image processing, telecommunication, optical communication, and applications of scientific knowledge to solve industrial problems. He is a member of the Mexican Academy of Optics. He is also a member of the National Research System (SNI) and the Mexican Physic Society.

ESPERANZA GUERRA-ROSAS was born in Durango, Mexico. She received the B.S. degree in computer systems engineering and the M.Sc. degree in electronic engineering from the Instituto Tecnológico de Durango, Mexico, in 2004 and 2010, respectively, and the Ph.D. degree in science (physics) from the Department of Physics Research, Universidad de Sonora, in 2017. She is currently a Postdoctoral Researcher with CETYS University, Mexico. Her research interests include image processing and pattern recognition with applications to biology and medicine. She is a member of the Mexican Academy of Optics and the National Research System.

JOSUÉ ÁLVAREZ-BORREGO received the B.S. degree in physical oceanography from the Facultad de Ciencias Marinas, Ensenada, Mexico, in 1981, and the M.Sc. and Ph.D. degrees in optics from CICESE, Mexico, in 1983 and 1993, respectively. He is currently a Professor with the Applied Physics Division, Department of Optics, CICESE. His research interests include image processing with applications to biogenic particles and image processing of the sea surface. He is a member of the Mexican Academy of Optics. He is also a member of the National Research System (SNI) and the Mexican Sciences Academy.