Abstract—Discovered on 31st December of 2019, the new Coronavirus has a high transmission capacity and was considered pandemic by the World Health Organization. In only six months it was able to spread all over the world and cause more than 600 thousand deaths. Early diagnosis is essential for governments to take public policies, such as social isolation, commerce control, and contact tracking. In order to make these actions, massive tests are required. On the other hand, diagnosis kits are expensive and not accessible to everyone. Medical imaging, such as thoracic x-ray and Computational Tomography (CT) has been used to visualize the lung and to verify at the first moment the presence of viral pneumonia. However, some countries have few radiologists specializing in chest x-ray analysis. The findings in the image are generally not so easy to see and can easily be confused with traditional pneumonia findings. For this reason, studies in Computer Vision are necessary, both to detect anomalies in imaging and to differentiate the other types of pneumonia. This paper addresses the initial results of a research, which developed an image classification methodology to differentiate x-ray images from sick patients, infected with Coronavirus, and healthy patients. The proposed method, based on the extraction and detection of patterns in texture and color features through a Deep Neural Network, obtained an average accuracy of 95% following a k-fold cross-validation experiment.

Index Terms—coronavirus, color-features, deep learning, x-ray

I. INTRODUCTION

SARS-CoV-2, popularly known as Coronavirus, is a virus from Coronaviridae family. It was first identified at the end of 2019 by Wuhan authorities, located in central China. The first reports recognized it as pneumonia without an already known identifiable cause. Soon after, they realized that simple proximity could contribute to the infection, as the virus spreads through droplets from coughing or sneezing of infected people, as well as contact with contaminated areas and surfaces. Then, due to its high spread capability, the virus ended up spreading rapidly to many continents, which resulted in its characterization as a pandemic disease on March, 11th by World Health Organization (WHO) [1]. More than twelve million cases of Coronavirus have been recorded worldwide so far. Also, around six million and 200 thousand patients are still sick. The lethality rate of the disease is approximately 7% and has caused a total of 600 thousand deaths [2].

This virus is responsible for the infectious disease COVID-19, in which the symptoms appear similar to a simple cold, such as dry cough, fever, runny nose, and sore throat, but may quickly evolve to a severe respiratory picture, similar to Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS), which occurred in 2002 and 2012, respectively [3]. In this last case, the patient may have serious breathing difficulties, with the need for intubation for adequate health treatment.

The diagnosis is made in patients with symptoms characteristic of the disease. The diagnosis process is done through the collection of respiratory materials, performing an aspiration of the airways or sputum induction which is subjected to molecular biology exams in order to check for the presence of viral RNA. In addition to these, rapid exams can be an alternative for patients after a certain infection period [4].

Due to the high demand and the lack of testing kits, the use of medical imaging tests can assist doctors in the detection of lung injuries in patients with a certain degree of infection. In this case, the most recommended tests to assist the visualization of the lung are the chest x-ray and Computed Tomography (CT) exams [5].

The diagnosis of x-ray images is performed in order to find signs of abnormalities in contrast to the pattern of a healthy patient. In these cases, images from sick patients show regions with an opaque bilateral irregular effect and with a frosted-glass aspect. However, these clinical findings may be similar to other types of pneumonia, making it difficult to differentiate patients with the new Coronavirus and patients with other viral pneumonia [6].

Therefore, for the search for signs of abnormalities in the image to be effective and to have good efficacy, it must be
performed by a radiologist who specializes in chest x-rays. However, although this is a category of medical images with easy access, since the equipment is easily found in large and medium-sized centers, in some countries there is a low number of specialized professionals to perform this task, which results in the need for more professionals to carry out such classification.

The classification of medical images, between sick and healthy patients, is commonly addressed in the Computer Vision literature applied to Health [7]. Diagnostic aid systems that use Artificial Intelligence have proven to be a powerful ally in the accurate diagnosis of diseases [8].

Computer Vision is a field of study in Computer Science that unites the Analysis and Processing of Images, from Computer Graphics to Artificial Intelligence in order to recognize elements of a scene and extract new knowledge about it. Applied to health sciences, it can assist doctors in the construction of scientific visualization systems and computer-aided diagnosis systems.

With the development of Computer Vision studies, in the context of the new Coronavirus pandemic, it is possible to develop methods that differentiate sick patients from patients with other pulmonary infections by helping general radiologists to interpret the exams indicating possible injuries.

Thus, the objective of this paper is to develop a methodology for classifying x-ray images, using Computer Vision techniques, in order to assist doctors in detecting findings in the image, indicating when a sample comes from a sick, infected patient Coronavirus, or a healthy patient. Despite the researches involving Deep Learning to distinguish healthy patients from those who have had Coronavirus, this paper stands out due to the innovative proposal of using chest x-ray imaging using hybrid features in a classifier based on Deep Neural Networks. Also, the proposed paper brings a study on how effective color features can lead the diagnosis of the new Coronavirus and to improve the results by using combined types of features.

This paper is organized as follows. In Section II we bring related papers for chest x-ray classification. Section III describes the dataset used, the preprocessing stage, and the Neural Network design. Section IV describes the experiments made in order to validate the proposed method. And finally, Section V concludes this work.

II. RELATED PAPERS

Some authors have addressed the problem of binary image classification of the new Coronavirus. In this section, we will consider works that classify x-ray images.

Narin et al. [9] developed an automatic detection system as a diagnostic alternative to COVID-19 using x-ray exams. Using chest X-ray, obtained from an open-source GitHub repository shared by Cohen et al. [10], and another 50 chest x-rays from Kaggle, three different models based on Convolutional Neural Networks (ResNet50, InceptionV3, and InceptionResNetV2) were used to detect patients infected by COVID-19. The authors ran their experiment for over 30 epochs. By the end of the experiment, the ResNet50 model demonstrated a faster training process than the other models and provided the highest classification performance, with 98% accuracy.

Apostolopoulos et al. [11] evaluated the performance of Convolutional Neural Network architectures using Transfer Learning. It collected a database from x-ray images available in public medical repositories for the experiment. The dataset contains 1428 x-ray images, including 224 from patients diagnosed with COVID-19, 700 with common bacterial pneumonia, and 504 in healthy conditions. The x-ray images were resized to 200x266, and in exams with different pixel proportions, a dark background set up to 1:1.5 ratio was added to obtain the 200x266 scale. The proposed method got 96.78%, 98.66%, and 96.46% as its best precision, sensitivity, and specificity.

Also, Ozturk et al. [12] suggested that the use of advanced Artificial Intelligence techniques and x-ray images could help them to detect COVID-19. Their goal was to overcome the problem of the lack of specialized radiologists, mainly in remote villages. The x-ray images were obtained from the image base developed by Cohen et al. [10], using images from various open access sources. The proposed model developed provides accurate diagnoses for binary (COVID-19 vs. Healthy) and multiclass classification (COVID-19 vs. Healthy vs. Pneumonia), getting the results of 98.08% accuracy for binary classes and 87.02% for the multiclass approach.

Hemdan et al. [13] introduced a new Deep Learning framework called COVIDX-Net, to assist radiologists in the diagnosis of COVID-19 through x-ray images. COVIDX-Net includes seven different architectures of Deep Convolutional Neural Network architectures, such as VGG19, DenseNet121, InceptionV3, ResNetV2, Inception-ResNet-V2, Xception, and MobileNetV2. Each Deep Neural Network is capable of analyzing the normalized intensities of the x-ray image to classify the patient’s status in the negative or positive COVID-19 case. Among all tested classifiers, the accuracy of InceptionV3 model was the worst with 50%, while the VGG19 and DenseNet201 models achieved the best values of accuracy (90%).

In Elasnaoui et al. [14] it was performed a comparative study between Deep Learning models to deal with the detection and classification of COVID-19 using x-ray images, Computed Tomography (CT). The experiments found that the use of InceptionResNetV2 and DenseNet201 provides better results compared to other models used, with 92.18% accuracy for InceptionResNetV2 and 88.09% accuracy for DenseNet201.

Wang et al. [15] introduces COVID-Net, a Deep Convolutional Neural Network specialized for detecting COVID-19 cases from chest x-ray images. In the tests, the proposed model performed better than the other analyzed, achieving 93.3% of accuracy, while the VGG-19 and ResNet-50 get 83% and 90.6%. The work also presents a dataset of 13,975 chest x-ray images from 13,870 cases, with the most significant number of positive cases publicly available of COVID-19.

Different from the previous works in the literature, the proposed method brings a traditional Computer Vision classifi-
cation method with a novel Deep Learning approach in which extracted features were used to make up a dataset for pattern recognition using a Fully-Connected Neural Networks. In this work, we analyze the effectiveness of color feature in addition to a hybrid approach mixing color and textural features for pattern recognition.

III. MATERIALS AND METHODS

In this section, we discuss the data acquisition process, the preprocessing performed in order to segment the region of interest and the Neural Network process of design.

All the experiments and development were performed on a computer with Ryzen 3600 processor, 16GB of RAM, SSD M.2 512GB, GPU RTX 2060 Super on the Manjaro 20 operating system. The proposed method were developed using Python 3.7 programming language and the frameworks Keras 2.3.1, Tensorflow 2.1, OpenCV 4, and Scikit-Learn 0.23.

A. Dataset

At the beginning of the disease spread, there was a lack of COVID-19 x-ray and CT images globally, making unfeasible the development of any work using image-based classifying methods. Works like Cohen et al. [10] made it possible by sharing those needed datasets on open data science repositories, allowing the development of several works.

To develop this work, we used datasets from Cohen et al. [10] and Tawsifur et al. [16], totaling 525 chest x-ray images of patients, being them 263 COVID-related and 262 normal.

B. Preprocessing

The datasets’ images dispose of different resolutions, angles, shapes, and color intensities, needing a process to normalize the data. Consequently, every image went through a preprocessing step before the feature extraction. The preprocessing step splits in (i) resizing, (ii) segmentation, and (iii) color-filtering.

In the first step, the OpenCV framework resizes the images to the 512x512 resolution, being this shape the default one through the entire code. Having a default resolution was needed to avoid incompatibilities through the preprocessing steps, and also to have a default model input shape.

The analyzes of the images shows that the patients bones would affect the classifier’s final result, making it necessary to focus on the actual Region of Interest (ROI). There are many approaches for image segmentation, either automatic or manual, in this work we applied the U-Net proposed in Ronneberger et al. [17] since it got excellent results automatically. Their work proposes an automatic lung segmenter based on Neural Networks, which takes an x-ray or CT image as input and creates a segmentation mask separating the ROI from the rest of the image.

The dataset have different color distributions due to its different acquisition processes, cameras, and storage. Said that, at the end of the preprocessing, OpenCV is used to normalize the images through Histogram Equalization, to increase the contrast, and prevent misclassification.

C. Feature Extraction

The feature extraction stage involves recognizing shapes, colors, zones, contours, and others, to identify properties which allow differentiating classes, such as COVID-19 positive and normal. The incorrect choice of descriptive features can lead to misclassification methods.

The x-ray images diagnosis process revealed that COVID-19 manifests itself as pulmonary spots, which influence the Intensity Histogram (IH) of the image. After some tests using this feature as a discriminant, it demonstrated promising results with the classifier. In order to evaluate the effectiveness of a color-based descriptor and to know how much it gains when combined with other types of features on hybrid descriptors, this work extracted both Intensity Histogram and Haralick Textural Features [18].

The proposed method evaluates the classification performance using color-based descriptors and a hybrid descriptor based on textural and color features. The color-based descriptor had 255 values with the referring the intensity of a given histogram level and its belonging class. For every image in the image dataset, a line was inserted on a CSV, making up the feature extracted dataset. The background color (black color) was not considered in this context since it doesn’t bring any information about the ROI.

Also, in order to improve results, a hybrid descriptor vector is proposed mixing color-based features and a hybrid descriptor based on textual and color features. The textural features were extracted based on Haralick’s textural descriptor. This descriptor is composed of 13 values, referring to 13 statistical metrics calculated over the co-occurrence matrix. The 13 values are, respectively, the Entropy of the Sum (ES), the Entropy (En), the Variance of the Difference (VD), the Entropy of the Difference (ED), Haralick’s Correlation (HC), Maximum Correlation Coefficient (MCC), Second Angular Momentum (SAM), Contrast (Con), Correlation (Cor), Sum of Squares - Variance (SSV), Inverse Moment of Difference (IMD), Sum Average (SA), Variance of the Sum (VS). Finally, both descriptors were used during the experiments to evaluate and compare its performance.

D. Deep Neural Network Design

The architecture proposed in this paper is based on a Fully-Connected Neural Network (FCNN), a Deep Neural Network architecture that implements more hidden layers in order to detect complex patterns over the data. FCNN is a traditional Deep Neural Network architecture since it brings elements from an Artificial Neural Network (ANN) but with deep hidden layers.

The modeling of the FCNN proposed begins with the input layer. All the features extracted, as described above, are treated as inputs in the first layer. Also, the data contained in the hybrid vector are subjected to normalization, using the Standard Scaler function from Scikit-Learn, so that there is no pre-defined order of priority among the variables.

The model also had six hidden layers, with 180 neurons each. Each hidden layer used the Rectifier Linear Unit (ReLU) as an activation function. To estimate the error rate between
the predicted class and actual class, the loss function used was the Mean-Squared Error (MSE). Also, the optimizer used was the Stochastic Gradient Descent (SGD).

The output model had one single neuron, representing the binary classification. It’s responsible to say if the given input belongs to a COVID-19 positive exam or if it’s normal. The activation function used in this layer was the Hyperbolic Tangent activation function (Tanh). The whole Neural Network model is represented in Figure 1.

![Image](image.png)

**Fig. 1.** Classification model for the hybrid descriptor. The input layer value 268 corresponds to the number of pixels each value the Intensity Histogram has excluding background information, which \( HI = 256 - 1 \), plus the 13 Haralick’s features.

In order to avoid the occurrence of overfitting in the classification model, six dropout layers were used. They were disposed after each hidden layer and was configured to turn off 20% of the neurons of the following hidden layer. This technique helped to create a generalized classification method.

### E. Hyperparameter Tuning

The Neural Network is composed of several hyperparameters. These hyperparameters, such as activation functions and optimizers are responsible to better detect patterns and create an accurate classification model.

In order to find the best hyperparameters, in this stage, we used the GridSearchCV function from the Scikit-Learn framework to test several hyperparameters combinations to find the best set. In Table I there are all the hyperparameters tested in this stage. The best set found used in the construction of the Neural Network was the Rectifies Linear Unity (ReLU) as a hidden layer’s activation function, the Tanh activation function in the output layer, the SGD optimizer, the mean squared error for loss calculation, and 180 neurons on each hidden layer.

### IV. Experiments and Results

The cross-validation technique has been widely used in classification problems to estimate the performance of models. This protocol defines that the experiment should separate the data into the training set and the test set. There are three different protocols based on cross-validation, they are the k-fold cross-validation, holdout cross-validation, and the leave-one-out cross-validation. The choice for one of the types of validation is defined according to the size of the database and the objective of the experiment.

In order to best evaluate the method, we proposed a three-stage experiment. The first experiment evaluated the efficiency of intensity histogram to classify x-ray images between COVID-19 positive and normal images. This experiment followed the k-Fold Cross-Validation technique. The second experiment evaluated the hybrid approach, which includes the Haralick’s textural feature to the intensity descriptor. The second experiment followed the Holdout Cross-Validation protocol and aimed to verify the existence of overfitting in the model. The third experiment evaluated the hybrid features as well but using K-Fold Cross-Validation. Our goal was to evaluate the method as a whole with different sets of training and validation.

To quantify the efficiency of the proposed method for each experiment, five metrics were calculated. These metrics allow us to verify the effectiveness of the model and also compare it to the other works from the literature.

#### A. Experiment 1: K-Fold Cross-Validation for Color-based Features

Our first concept was to develop a classification method based on Intensity Histogram. During our studies, we observed that the diagnosis was made up by analyzing the presence of certain patterns on the image. These patterns, not going into medical definitions, appear as white spots blurred in the image. This insight leads us to conduct a color-based feature investigation.

For this first test, we conducted an experiment to detect patterns over the intensity histogram. We had a medium-size dataset that allowed us to perform our experiment using the K-Fold Cross Validation protocol. This protocol says that the dataset should be split into k parts. The experiment runs over
k times and on every experiment, part of the data is used for training and part for validation. The results of this experiment can be seen in Table II.

<table>
<thead>
<tr>
<th>#</th>
<th>Accuracy</th>
<th>Precision</th>
<th>F1-Score</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.8571</td>
<td>0.8500</td>
<td>0.8500</td>
<td>0.8500</td>
<td>0.8636</td>
</tr>
<tr>
<td>1</td>
<td>0.8571</td>
<td>0.8421</td>
<td>0.8421</td>
<td>0.8421</td>
<td>0.8696</td>
</tr>
<tr>
<td>2</td>
<td>0.9286</td>
<td>0.9200</td>
<td>0.9388</td>
<td>0.9583</td>
<td>0.8889</td>
</tr>
<tr>
<td>3</td>
<td>0.9286</td>
<td>0.8571</td>
<td>0.9231</td>
<td>1.0000</td>
<td>0.8750</td>
</tr>
<tr>
<td>4</td>
<td>0.8571</td>
<td>0.9231</td>
<td>0.8000</td>
<td>0.7059</td>
<td>0.9600</td>
</tr>
<tr>
<td>5</td>
<td>0.8810</td>
<td>0.8947</td>
<td>0.8718</td>
<td>0.8500</td>
<td>0.9091</td>
</tr>
<tr>
<td>6</td>
<td>0.9286</td>
<td>1.0000</td>
<td>0.9231</td>
<td>0.8571</td>
<td>1.0000</td>
</tr>
<tr>
<td>7</td>
<td>0.8810</td>
<td>0.9091</td>
<td>0.8889</td>
<td>0.8696</td>
<td>0.8947</td>
</tr>
<tr>
<td>8</td>
<td>0.9286</td>
<td>1.0000</td>
<td>0.9388</td>
<td>0.8846</td>
<td>1.0000</td>
</tr>
<tr>
<td>9</td>
<td>0.9512</td>
<td>1.0000</td>
<td>0.9524</td>
<td>0.9091</td>
<td>1.0000</td>
</tr>
<tr>
<td>Mean</td>
<td>0.8999</td>
<td>0.9196</td>
<td>0.8929</td>
<td>0.8727</td>
<td>0.9261</td>
</tr>
</tbody>
</table>

With a mean accuracy of 89%, we noted that this approach was pretty promising. After this evaluation, we tested the dataset over different other Neural Network configurations but the best result with IS stays the same. In order to improve the results, the hybrid descriptor was developed, mixing color and texture features together in order to develop a more accurate model.

B. Experiment 2: Holdout Cross-Validation for Hybrid Features

The holdout cross-validation protocol defines that the entire dataset should be separated into two parts, the training set, and the validation set. The proportion used is 80% for training and 20% for validation. Both subsets have different images in order to train and test with different contents.

Our goal with this experiment was to evaluate the method using a single execution to analyze the learning process over the epochs, get the threshold in which the method converges, and to verify if the method tends to overfit over the time.

This experiment runs for over 5000 epochs. It was observed that the Neural Network proposed quickly converged to its optimum around the epoch 100. It got an accuracy of 94.29%, a sensitivity of 94.24%, a specificity of 94.33%, a precision of 94.34% and 94.29% F-Score. The accuracy and loss were used to plot two graphs (Figure 2) representing their changes over the epochs. Note that the training and test values stay closer and stable.

C. Experiment 3: K-Fold Cross-Validation for Hybrid Features

In order to evaluate the overall performance of the hybrid features, this third experiment was performed splitting the data into 10 parts following the K-Fold Cross-Validation. This experiment used the same architecture of Neural Network but ran for 300 epochs since it converges pretty quickly requiring fewer epochs. The results can be seen in Table III.

The overall accuracy of the proposed method using hybrid features achieved 95%. Note that comparing with the first experiment, which was used only the IH, there was an increase of 6% in the overall accuracy by using Color and Texture Features combined.

D. Comparison with other works

Most of the recent works from literature used Convolutional Neural Networks and its sub-architectures to detect pattern and classify x-ray images from COVID-19 patients. In this work, we investigate the efficiency of color-based and hybrid descriptors, bringing textural features, such as Haralick, alongside Intensity Histogram in order to analyze which one has a better performance with a Fully-Connected Neural Network.

Even with a traditional method and easy to extract features, the proposed method showed itself promising in classifying
x-ray images between sick and healthy patients. The proposed method achieved an overall accuracy of 95%, a competitive value compared to other sophisticated methods in the literature.

In Table IV, the results obtained with the proposed method is compared to the other works in the literature, revised in Section II, which also classify x-ray images. Note that all the results obtained in this work are stable and pretty close to the others but bring an easy to extract and process method.

<table>
<thead>
<tr>
<th>Author</th>
<th>ACC</th>
<th>PREC</th>
<th>F1</th>
<th>SENS</th>
<th>SPEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elasnaoui et al.</td>
<td>0.9218</td>
<td>0.9238</td>
<td>0.9207</td>
<td>0.9211</td>
<td>0.9606</td>
</tr>
<tr>
<td>Hemdan et al.</td>
<td>0.9000</td>
<td>0.9150</td>
<td>0.9000</td>
<td>0.9000</td>
<td>—</td>
</tr>
<tr>
<td>Narin et al.</td>
<td>0.9800</td>
<td>1.0000</td>
<td>0.9800</td>
<td>0.9600</td>
<td>1.0000</td>
</tr>
<tr>
<td>Ozturk et al.</td>
<td>0.9808</td>
<td>0.9803</td>
<td>0.9651</td>
<td>0.9513</td>
<td>0.9530</td>
</tr>
<tr>
<td>Apostolopoulos et al.</td>
<td>0.9875</td>
<td>—</td>
<td>—</td>
<td>0.9285</td>
<td>0.9875</td>
</tr>
<tr>
<td>Proposed Method</td>
<td>0.9548</td>
<td>0.9584</td>
<td>0.9545</td>
<td>0.9534</td>
<td>0.9579</td>
</tr>
</tbody>
</table>

V. CONCLUSION

COVID-19 was discovered recently but it already has become an international viral outbreak. The diagnosis of this new virus is hampered by the absence of biological kits and the delay in obtaining the results. Thus, the use of medical imaging can help doctors diagnose the patient and to better estimate the patient’s contagious level. The difficulty in diagnosing the disease using x-rays is due to the fact that the findings of the Coronavirus and other viral pneumonia are quite similar. The Computer Vision systems can be used to improve diagnosis by helping the radiologist to find abnormal signs in these exams.

The proposed work brings a classification method based on hybrid features. It is shown the effectiveness of the Intensity Histogram alongside textural features, such as the Haralick descriptor. For future works, we plan to work with more complex Deep Learning architectures in order to perform a multiclassification detecting different types of pneumonia.

REFERENCES