Journal of Theoretical and Applied Information Technology
15th June 2024. Vol.102. No. 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FEATURES EXTRACTION USING GEOMETRY, STRUCTURAL, AND TEXTURE ANALYSIS FOR COVID-19 CLASSIFICATION FROM COMPUTED TOMOGRAPHY IMAGES OF THE LUNGS

MOHAMMAD ALFRAHEED[12]

[1]Department of Computer Science, Faculty of Communications and Information Technology, Tafila Technical University, Tafila - Jordan

[2] Department of Computer and Information Science, College of Engineering and Computer Science, University of Michigan – Dearborn, Dearborn – Michigan, USA

E-mail: [1] Alfraheed@ttu.edu.jo, [2] mohalfr@umich.edu

ABSTRACT

Worldwide publishing of the COVID-19 has occurred. One of its impacts has been demonstrated to be a lung infection in the patient. As the virus's virulence increased, the infection's dissemination throughout the lung increased. Utilizing computed tomography images (CT-Images), the severity of COVID-19 has been analyzed and diagnosed. Studies have consequently employed CT-Images to track patients' COVID-19 illness. The contribution here is to use CT-Images to extract the features of the infection. The proposed method uses several filters to eliminate shapes with insufficient grey intensity gradation before extracting the relevant features. The unique aspect of the method lies in the fact that different geometry, structural, and textural characteristics have been retrieved. The variety can be explained by having the ability to characterize the shapes of the infection utilizing those characteristics, in addition to preserving the internal and external appearance of the infection. The CT-Images have been classified using machine learning techniques, demonstrating the efficacy of these extracted features. The classification accuracy has varied between 99% and 100%, and many infection shapes have been identified in the CT-images. By increasing the classification accuracy, the proposed method has outperformed previous methods in terms of performance.

Keywords: *Computed Tomography Images, COVID-1, Feature Extraction, Classification, Machine Learning Algorithm.*

1. INTRODUCTION

Approximately 57 million patients have been diagnosed as Covid-19 virus-infected since 2019. Approximately 24.6% of those individuals passed away because of COVID-19 viral infection [1]. The amount of lung infection caused by the COVID-19 viruses ranges from a high infection level caused by a high dose virus to a low infection level caused by a low dose virus. Furthermore, as noted by the World Health Organization (WHO) and World Meter websites (WMW) [2], the number of affected individuals increases if environmental variables (temperature, humidity, wind speed, and pressure) fluctuate. Consequently, depending on their state of health, certain infected individuals have been experiencing varying degrees of COVID-19 viral infection [3] [4]. Beside of the COVID-19 effects in the world, the educational institutes have been affected by changing the Face-To-Face education to the Online Teaching [5] [6].

To properly analyse the affected area and its effects in the lung, it is necessary to determine the severity of the COVID-19 infection. The doctor's ability to provide vital life-saving treatments is dependent on how severe the lung infections are. It is also necessary to ascertain whether the infections are caused by COVID-19 viruses or other lung ailments [7]. One tool for diagnosing lung infections and assessing the severity of the illness in conjunction with a clinical examination is the use of Computed Tomography Images (CT-Images) [8].

It is not easy to accurately diagnose these CT scans by removing features from them. The features that were collected must also be useful in

the classification process that divides the lung's CT images into infected and non-infected regions [8]. When diagnosing CT-Images, the radiologist must step aside and let the features extraction process take over. He needs to focus his eyes on the regions that are infected and gather these regions from every CT scan of the lung. Many CT-Images of the patient are created throughout each scanning procedure based on various points of view [9]. As a result, driven by this issue, features extraction is an essential phase in both the diagnosis and classification tasks.

These gathered regions need to be accurately identified and converted into statistical data. To uncover the hidden characteristics and numerically depict their underlying subtleties, the latter must be examined. Researchers are encouraged to use the generated results in sophisticated medical diagnostic or advanced classification algorithms by the new statistical forms. Consequently, in terms of the medical applications, the results of this work can be viewed as a crucial first step toward the next round of diagnosis.

The computed tomography scanning procedure has presented another difficulty. Several CT-Images have been produced, beginning with a tiny black spot in the lung and going up to the largest possible black spot where the interior details of the lung are recorded. The biggest area is then reduced in size until it vanishes. The problem lies in the fact that it is unclear which CT-image contains COVID-19 infections. Therefore, it is necessary to exam every CT-Image to check for infections of all sizes and forms.

In the context of deep learning, features extraction is often regarded as the initial stage, particularly for Convolutional Neural Network (CNN) models. In [10], the VGG-19 model was adjusted to identify and diagnose COVID-19. The VGG-19 filters helped to extract the relevant features. The GRADient-weighted Class Activation Mapping (Grad-CAM) approach has been used in [11] to highlight the infection area of the COVID-19 in the CT-Images before pasing them into the CNN. Additionally, the deep learning network ResNet50 has been fine-tuned [12]. Even with their success rate of over 90%, they were still dependent on CNN's internal filters to identify the COVID-19 infection zone. Recently, the white patches in the X-ray images have been considered as a first satege of the COVID-19 detection [13]. Focusing on the infection region instead of the entire CT-image saved more time and memory resources in CNN. However, the infection regions must presicely identified to use them in the classification processing.

As a summary of the problem's outline is that there is a need to extract features for precisely describing the infection region of the COVID-19 in the CT-Images. From the first CT-Image to the last, these features are collected from the CT scan. Rather than extracting features from the entire CT scan, the infection region of the COVID-19 must first be identified, and the lung regions are localized in the CT-Image.

The features—geometry, structure, and texture features— have been here introduced to describe the region of interest. Previous significant works have been done to extract features of interest using a variety of techniques, including wavelet [14], Fourier transform [15], textures [10], and the variation in grey intensities applied by convolution neural network [12]. The related work has addressed these techniques' importance and drawbacks. Nonetheless, the issue at hand is how to extract relevant features while focusing solely on the afflicted regions and preserving all relevant information about them. Even after years of intensive work and substantial advancements, there are still several restrictions, which are as follows:

- ✓ While the CT scans were converted to multiple levels of resolution, some techniques that included Wavelet and Fourier modification caused some significant features to disappear.
- ✓ Because the majority of nowadays CT-Images typically lack appropriate color intensity graduation, analyzing the features based just on the texture's characterization is not precise enough.
- ✓ Intensity-based filters are used by modern deep learning approaches, such as convolution neural networks, to differentiate the object of interest from unrelated objects. These filters' shortcoming is that they only allow the intensity values to describe the objects of interest.
- ✓ None of the existing research examined how to classify COVID-19 cases from normal cases by analyzing different factors and determining the COVID-19 infection locations.

This paper proposes a method to extract characteristics from CT images by considering several notions of region categorization. A few of these ideas have been covered with convolutional neural networks and region-of-interest analysis. Several machine learning methods have evaluated the extracted features for its efficacy in classifying COVID-19 from NORMAL computed tomography images. The novel method here is to provide a

foundation of COVID-19 features that may effectively categorize COVID-19 cases. These features are not restricted and can be enhanced by other features in order to satisfy the machine learning algorithms and boost the classification success rate. The main contributions of this work are summarized as follows:

- ✓ Improving the process of extracting features for the COVID-19 infections observed on the CT images. By focusing the proposed approach on the lung regions rather than evaluating the entire CT-Image, improvements have been made.
- ✓ Extracting the features from the COVID-19 infection regions using the geometry, structure, and texture analyzing. These features have been extracted only from the regions which can be successfully passed the proposed method' filters.
- ✓ Applying several machine learning methods to evaluate the extracted features. The evaluation was carried out to validate the feature's ability to classify CT-Images into COVID-19 and NORMAL instances.

This paper is organized as follows; the first section serves as an introduction. The related work has been covered in the second section. The third section has addressed and discussed the proposed method. The fourth section highlights the findings and results. Lastly, the final section contains the conclusion.

2. RELATED WORK

Extracting features from the computed tomography images (CT-Images) is the key step in COVID-19 classifications. Every research article has proposed its way to carry out this task. In this section, those articles have been discussed.

Using the idea of the Fourier transform is one of the features extraction techniques. In [14], a discrete wavelet model optimized network model has been introduced for the COVID-19 diagnosis and classification. One of the model's stages is to extract features. The input CT-Image has been converted into a frequency domain transformation, highlighting the areas of infection from the surrounding area. The range of their outcomes was 99% to 100%. In [16], robust hybrid features were also extracted as a features matrix using the discrete wavelet transformation. A principal component analysis has been used to further analyse the matrix for producing a second generation of the features matrix. The range of their outcomes was 97.6% to 98.5%. Moreover, one of the frequency domain transformation methods for feature extraction in [15] was the fractional Fourier transform. Their developments have also been taken into consideration as a way to advance the wavelet idea in the extraction of COVID-19 features. The range of their accuracy values was 99.84% to 99.90%. Using wavelet and Fourier processing has the drawback of causing some features to vanish as the CT-Images are converted into several low levels. Even if they have seen success, the vanished features have a significant impact on the COVID infection's severity.

Using texture-based features has been suggested as an additional concept for feature extraction. The Gabor filter has been used in [17] to extract the relevant characteristics. After that, these features were examined using linear discriminative analysis to minimize their dimensionality and identify the robust features. They have a 93.96% classification accuracy rate. In [18], the idea of texture analysis has also been examined. To extract 29 characteristics from the CT images, the Gray Level Co-event Matrix (GLCM) and Gray Level Run Length Matrix (GRLM) have been employed. They achieved about 96.86% accuracy. To extract features from the picture descriptors, researchers in [19] employed the Scale-Invariant Features Transform (SIFT) in their proposed method. Among the algorithms used to extract the stronger point of interest and highlight the stronger region in the input image is the SIFT algorithm. To extract the strong features for their research, the authors in [19] combined the idea of texture analysis. In [20], more texture features have been taken into account in cooperation with an enhanced optimized neural network. The optimized region-growing technique was used to extract and segment about 20 texture features. Their success percentage was 92.37%. Even though the concept of texture analysis has been effectively implemented in earlier research articles, with success rates in the COVID-19 classification, computed tomography scanning typically produces solid images with minimal intensity values. These values provide a unique grey color to each region in the CT image, and there is insufficient color gradation to analyse the features based on textures. This problem has been regarded as one of texture analysis's drawbacks.

Convolution neural networks, or CNNs, have been employed in object classification recently, and numerous studies have been implemented on classification. The object has been identified within CNN by using multiple convolution filters on the input image. A vector of features that will be utilized in the classification procedure was produced by these filters. Here, the

features extraction has been emphasized and a summary of the most current articles has been provided. A multi-task model based on several CNN iterations has been provided to researchers in [21] for the purpose of classifying COVID-19 in chest CT images. By segmenting the regions in CT-Images and testing them using multi-layer perceptron, their hypothesis has been put into practice. They achieved performance levels of about 95.4%. In [22], a CNN framework model for COVID-19 detection in CT images has been presented. VGG16, VGG19, Densenet121, InceptionV3, Xception, and Resnet50 have all been assessed in this work, and the outcomes produced have been contrasted with one another. In order to identify the infection zones, each model has produced its own unique features of interest. The VGG16 model has the highest success rate (98.00% accuracy) in their work. The features extraction procedure has been developed in a distinct phase in [23]. To identify and characterize the infection's zone, the dilated convolutional-based Split-Transform-Merge (STM) blocks have been used in place of the convolution filters. Boundary operations and multi-path region-smoothing have been used to process each convolutional block. They then produced further data regarding the global COVID-19 particular patterns and contrast variation. In addition to this data, each block includes the texture difference between images that are unique to COVID-19 and images that are healthy. The novel CNN has been created using these blocks to analyse and identify the COVID-19 infected zones. The generated results of CNN have been improved as a result of the features extraction procedure adjustment. The analysis of the infected zones was still done using the texture features. They attained a success percentage of 98.21%.

A novel framework called DeepChestNet has been presented by researchers in [24] to solve the difficulty of CNN's inability to capture structural relationships of the infected area. In order to take advantage of the structural relationship for ground glass opacities (GGO)—that is, the radiopacity idea of infection area—the DeepChestNet model has been introduced. Their efforts were focused on developing segmentation procedures for the lung, pulmonary lobe, and GGO. Following segmentation, the structural relationship between these objects has been discovered, allowing for the analysis and COVID-19 detection. They are doing a great job of finding a different way to describe COVID-19 infection, and their success rate is 97.9%. An innovative concept in this work was their introduced description, which was based on the structural properties. Researchers in[25] have been using CNN algorithm's convergence to identify and reduce the number of "False Negatives" features, rather than refining the features extraction method. They have employed the cross-entropy loss function with success, but they have continued to rely on the CNN-style feature extraction mechanism. To extract and categorize COVID-19 features, they employed the DenesNet convolution filters. Rather than enhancing feature extraction, the binary cross entropy loss function was also employed in [26] to enhance CNN's model performance. Given that they have also enhanced the features extraction process, their efforts would be more effective.

In summary, in order to preserve the appearance of the region of interest, the proposed method has been devised to extract features from the CT-Images by first grouping the regions of infections, regardless of their sizes and forms. To maintain the appearance of the region of interest, there has been no consideration given to transforming the CT-Images into Fourier or Wavelet transformation levels. Furthermore, because the CT image's color gradation is insufficient, employing the textural features is not sufficient to identify and analyse the infected regions. However, the proposed method would require the involvement of specific texture features in cooperation with other features. Both characteristics can be useful in identifying the infection area. The proposed method would combine the segmentation and filtering techniques utilized in convolution neural networks to analyse every potential infection region while simultaneously ignoring those insignificant and tiny regions. Reading the structural characteristics, discussed in recent development of the convolution neural network, has been also considered in the proposed method to describe the internal geometry characteristics of the region of interest.

3. METHOD

The algorithm below demonstrates the details of the proposed Features Extraction method. The accompanying code, from Line 1 to Line 12, outlines the 12 phases that make up the proposed method. Figure 1 shows generated images associated with each line of the Algorithm 1. For more details about the proposed method, a description of each line has been given as follows.

Input Step: The x-ray computed tomography for lungs has been produced using several perspectives. The lungs have been represented as a black hole in the first image of the

scanning process. The black hole grew until it completely engulfed the lungs, at which point it shrank till it vanished. Therefore, the proposed method has been developed to search every x-ray computed tomography image for the satisfied features. The input image has been titled CT_Image (x, y).

Line 1: To uniformly distribute the grey intensity level over connected pixel areas, the CT_Image must be converted to a binary image. One of the most modern suitable methods is to use an adaptive threshold [27]. Consequently, a binary image with the name BinaryImage was created and displayed in Figure 1 under the caption Line 1. As shown in the generated binary image, several white connected regions have been formed in place of the grey level intensities.

Line 2: Since they don't connect to the black holes that represent the lungs, some of these white connected regions are useless. The biggest area that contains the lungs' black holes is the important region. Extracting the biggest region enables the proposed method to focus its process on where features of interest are located. Moreover, the radiologist ignores these pointless areas in favor of keeping a watch on the expanding black holes. Each connected region's pixel has been counted and the biggest area has been extracted. The largest region in the binary image has been identified, as shown in Figure 1 under caption Line 2. Compared with the binary image, the useless regions have been eliminated and those small regions located in the black holes have been also deleted. The biggest region has been titled by the MainArea in line 2 of the proposed method, which will be later used.

Algorithm 1: The Proposed Method

procedure EXTRACT_FEATURES(**CT-IMAGES**)
CT_Image (x,y) are x-ray computed tomography images of lunge
$\boldsymbol{Line\ 1}$: BinaryImage $(x, y): Z \times Z \rightarrow \{0,1\}$
$BinaryImageRegions = \{ConnArea_1, ConnArea_2, \dots\dots\dots, ConnArea_N\}$
$\boldsymbol{Line\ 2}$: $MainArea =$ MAX(BinaryImageRegions)
$\boldsymbol{Line\ 3}$: $ForegroundArea = RemoveHoles(ForegroundArea)$
$RemoveHoles = Morph\_Oper\_Shrink(Morph\_Oper\_Fill(B))$
$\boldsymbol{Line\ 4}$: $RegionOfInterest = MainArea \cap ForegroundArea$
$RegionOfInterest(x, y): Z \times Z \rightarrow L_i, i = 1,2,3, \dots\dots\dots M$
$\boldsymbol{Line\ 5}$: $Delete\ RegionOfInterest_i, NumberOfPixel(RegionOfInterest_i) < TValue$
$\boldsymbol{Line\ 6}$: $RegionOfInterest(x', y') = REFINING(RegionOfInterest(x, y))$
$REFINING = Morph\_Dilate(Morph\_Shrink(Morph\_Fill(Morph\_Dilate(B))))$
$\boldsymbol{Line\ 7}$: $ActiveRegion = \prod_{CT_{IMAGE}}(RegionOfInterest(x', y'))$
$ActiveRegion(x, y): Z \times Z \rightarrow \{0, V_1, V_2 \dots\dots\dots V_R)$
$\boldsymbol{Line\ 8}$: $ActiveRegion(x, y) = Value = 0, \quad Value \in \{V_1\}$
$ActiveRegion(x, y): Z \times Z \rightarrow L_J, J = 1,2,3, \dots\dots\dots, T$
$\boldsymbol{Line\ 9}$: $ActiveRegion \in \{ConnArea_1, ConnArea_2, \dots\dots\dots, ConnArea_T\}, R > ThreshValue$
$\boldsymbol{Line\ 10}$: $Delect\ ConnArea_T, \quad StandardDeviation(ConnArea_T) = 0$
$\boldsymbol{Line\ 11}$: $SHAPE(x, y): Z \times Z \rightarrow \{V_2, \dots\dots\dots, V_R\}, ConnArea \neq BorderConnArea$
$ConnArea = \{(x_1, y_1), \dots\dots, (x_n, y_n)\}$
$BorderConnArea = \{(x'_1, y'_1), \dots\dots, (x'_n, y'_n)\}$
$\boldsymbol{Line\ 12}$: $Features_{SHAPE} = \{F_1, F_2, F_3 \dots\dots\dots F_{11}\}$
$\boldsymbol{F_1}: Diameter, \boldsymbol{F_2}: Eccentricity, \boldsymbol{F_3}: ObjectNumber, \boldsymbol{F_4}: PixelRatio$
$\boldsymbol{F_5}: MinorAxis, \boldsymbol{F_6}: Angle, \boldsymbol{F_7}: Perimeter, \boldsymbol{F_8}: Solidity$
$\boldsymbol{F_9}: StandardDeviation, \boldsymbol{F_{10}}: NumLevels, \boldsymbol{F_{11}}: MajorAxis$

Journal of Theoretical and Applied Information Technology
15th June 2024. Vol.102. No. 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

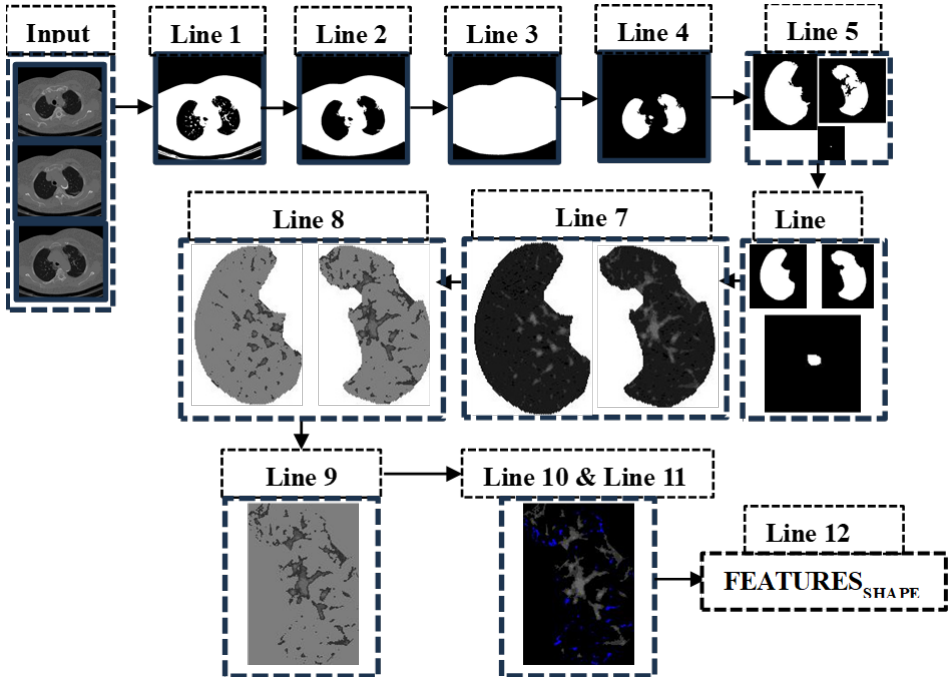


Figure 1. A Photography Flow Chart For The Proposed Method Associated With Code Lines From 1 To 12

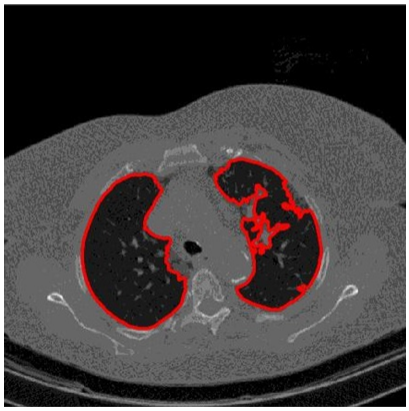
Line 3: The black holes' internal details have been deleted. They do, however, reflect COVID-19 characteristics. It is necessary to first highlight their location in order to retrieve these features. This snap code's concept is to extract the foreground region which represents the entire white region of the MainArea without black holes. Once the Foreground region is generated, these internal features are later isolated by identifying the difference between Foreground and MainArea Region. Furthermore, the foreground needs to encompass the MainArea Region's internal details without introducing any extraneous areas that are situated beyond the MainArea border. To complete the task, two morphological operations — the Fill operation [28] and the Shrink operation [29] — have been introduced. The result has been shown in the Figure 1 under the caption Line 3.

Line 4: Extracting the regions of lungs has been easily identified through the intersection between the MainArea and the Foreground Area. The generated region has been titled by RegionOfInterest and it has been shown in Figure 1 under the caption line 4. A binary image called RegioOfIntereset consists of several connected regions of varying sizes and shapes. Some dispersed regions are meaningless since the primary Regions of Interest encompass both sides of the lungs.

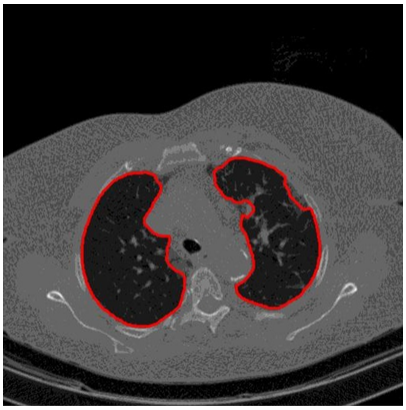
Line 5: These meaningless regions must be eliminated because the proposed method must concentrate its process where the COVID-19 characteristics are located. The proposed method has been developed to be adapted with a parameter named TValue which represents the number of pixels in each connected region of RegionOfInterest. the Region will be deleted and marked as a meaningless region as the RegionOfInterest has pixels less than the TValue. Since these meaningless regions are so small, it is really challenging to display these useless regions as samples in the Figure 1. Under the caption Line 5 of the Figure 1, three regions of interest have been extracted and shown.

Line 6: Due to the extraction process and morphological operation, some regions of the lungs have not included with the RegionOfInterest. Beside of these regions implementing as parts of the lungs, the COVID-19 infection spreads out from margins of lungs while the body is being scanned [19]. It is so necessary to refine the extracted RegionOfInterest for recovering the entire area of the lungs. Refining process consists of Morphological Dilation Operation [29] , Morphological fill operation [28], morphological shrink operation[28] and the Morphological dilation operation [29]. As a result, the internal cracks of the RegionOfInterest have been removed and merged with the lung's region. The

edges of lungs have been refined to smoothly identify where the lungs are located in the CT_Images. For further comparison, the Figure 2 shows the edges of the lung before and after refining process. The edges have been marked by a red color. The edge of the lungs before refining process has been shown in Figure 2 (A). In contrast, the Figure 2 (B) shows the edges after refining process.



(A)



(B)

Figure 2. a comparison between before and after refining process given in line 6 of the proposed method. (A) the lung's edges before refining process coloured by red. (B) the lung's edges after refining process coloured by red

Line 7: To analyse the COVID-19 characterizations and extract the characteristics related to the CT_IMAES, the RegionOfInterest must be projected over the CT_IMAGE after it has been successfully refined. A grey region with varying intensities (i.e., 0, V1, V2, … VR), called ActiveArea, is the name given to the created projected view. Figure 1 shows the ActiveArea under the caption Line 7. The lungs have been taken out and ready for analysis on both sides. At this step, the proposed method replicates the radiologist's duty, which is to concentrate on the area where COVID-19 infections are emerging.

Line 8: It was evident by examining the ActiveArea's internal intensity preview that the intensity values varied, starting at 0 and going through V1, V2, and VR. This primary objective of the line 8 snap code is to create filter that removes superfluous intensities from the ActiveArea, such as the background color and lung scan color. The proposed method has substituted a distinct color for the 0-value in order to eliminate the background color. The inside lung segments are surrounded by the scanning color. A 3D-historgram of the grey intensities [20] has been created and shown Figure 3 (A) in order to preview the scanning color. The Figure 3 demonstrated that the scanning color was represented by the first levels, or V1. The 3D-histogram has been rotated to see intensities values based on X and Z coordinates for displaying the ActiveArea's intensities levels. The new perspective of the 3D-histogram has been shown in Figure 3 (B). The rotated perspective based on Y and Z coordinates has been displayed in Figure 3 (C). It was evident from both pictures that the first level varied from 0 to V1. Figure 3 (D) illustrates how the ActiveArea was converted into a 2D–histogram vertical view in order to observe the first level as a scanning color. In the latter, a solid color encircles the interior lung segments. The new ActiveArea view, with a grey backdrop and scanning color in place of the previous ones, has been displayed in Figure 1 under the caption Line 8.

Line 9: The intensity variation in each lung side has been displayed using the parameter VR, which denotes the number of intensities levels in each ActiveArea. The ActiveRegion has been regarded as a region that includes COVID-19 characterizations after the intensity's variation exceeds ThreshValue. The rationale for employing this filter is based on the theory that, during lung scanning, COVID-19 infections begin to spread from lung edges at high various intensities [19]. As a result, as shown in Figure 3 – B and Figure 3 - C, the left side of the lung has been removed since its intensity levels are insufficient. The 2D-histogram for the right side has been shown in Figure 4. The 2D-histogram based on X and Z coordinates has been shown in Figure 4 (A). Figure 4 (B) shows the 2D-histogram based on Y and Z coordinates. As a

result, the right side of the lung is depicted in Figure 1 under Caption Line 9 as being analysed subsequently for COVID-19 characteristics.

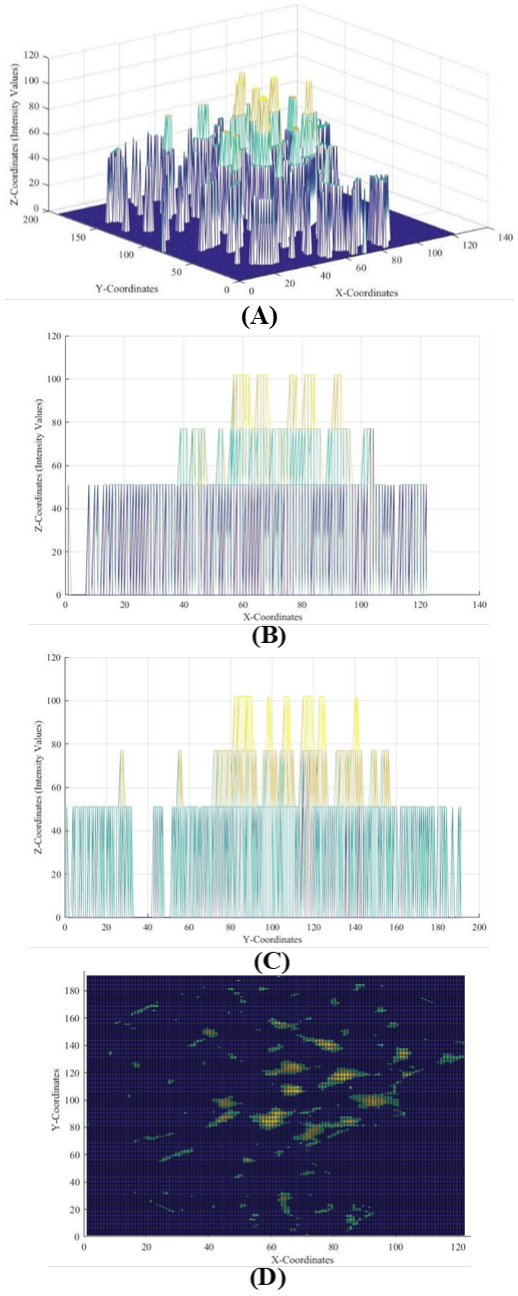


Figure 3. The 3D-Histogram of the ActiveArea in different views. (A) The 3D-historgam for the ActiveArea based X, Y and Z-coordinates (B) The 2D-histogram for the ActiveArea based on X and Z-Coordinates. (C) The 2D-histogram for the ActiveArea based on Y and Z-Coordinates. (D) The vertical view for the 2D-histogram based on X and Y coordinates

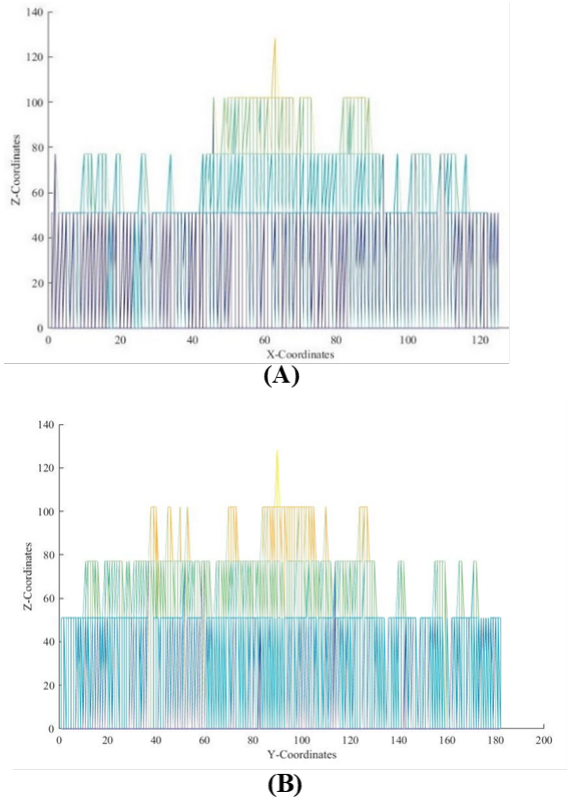


Figure 4. The 2D-histogram for the intensity's levels of lung's right side. (A) The 2D-histogram based on X and Z coordinates. (B) The 2D-histogram based on Y and Z coordinates.

Line 10: The second filter has been applied to the connected regions of the ActiveArea, and each region's standard deviation [18] has been calculated. Due to insufficient intensity variation for COVID-19 infection, the related region associated with a 0-value standard deviation has been eliminated. Eliminating those regions was intended to focus the operations of the proposed method over the region that can be recognized as a COVID-19 characterization. Another goal of the filter that line 10's snap code applies is to preserve the resources of the proposed approach. As shown in the Figure 1 under the caption Line 10, the eliminated regions have been distinguished by color blue.

Line 11: The third filter of the proposed method is to search for regions that can be recognized as shapes. When the border's pixels of the connected region, known as the BorderConnArea, do not match every pixel of the same connected region, the region is identified as a shape. In other words, the region

must have some pixels when border's pixels are removed. The concept underlying this filter is the infection of the COVID-19 has been formed for different shapes while the lungs were scanned. In contrast, regions do not belong to shapes have been formed as tiny regions which have not a sufficient intensity variation. Consequently, compared to the tint regions, the shapes provide more information and features about the COVID-19 characterisation.

Line 12: The last step of the proposed method is to extract features from the regions recognized as shapes. Despite the proposed method has filtered useless regions, some regions are still considered as a target. The capacity to analyse these regions in relation to their appearance, shapes, sizes, intensities values, and other geometry properties has been made possible by the proposed method. The analysis process enables the proposed method to classify in advance these regions into COVID-19 infection regions or normal region localized in the lungs.

For every region in the input image CT_IMAGE that has been recognized as a shape, a vector of features has been initialized. Features in the vector, which range from F1 to F11, each have a value determined by the proper measurement related to the feature itself. A 2D array has been produced as an output of the proposed method. The region's ID is shown in the rows of the array, and the extracted features are shown in the columns. The proposed method is continued until all CT_IMAGE images provided by the scanning operation have been processed. The following is a description of these features.

- ✓ F1: It is a Diameter value of the region. It is defined by the length of the line that touches points located at the region's edge [30].
- ✓ F2: It is the region's eccentricity value. It is determined by measuring the distance between the shape's foci and major axis length [31].
- ✓ F3: It determines differences between the number of objects in the region, and the number of holes in that region. It is also defined in Euler number measurement [32].
- ✓ F4: It shows the rate of change in pixels between the pixels in the region and the other pixels in the rounded rectangle that surrounds it [33].
- ✓ F5: It calculates the length of the area's minor axis. Pixels have been used to compute the length of the minor axis [34].
- ✓ F6: It determines the region's angle which is between the x-axis and the major axis of the region. The angle's value was ranging from -90 to 90 [35].
- ✓ F7: It measures the length of the distance encircling the region. By combining a pair of pixels around the boundaries of the region, the distance is generated[36].
- ✓ F8: It determines the percentage of pixels which are in the region and in the smallest polygon encircling the same region [37].
- ✓ F9: It calculates the standard deviation for the region's intensity values. The feature calculates the distribution of intensity values [38].
- ✓ F10: It shows how many levels can be generated from the region's intensity values [38].
- ✓ F11: It calculates the length of the area's major axis. Pixels have been used to compute the length[34].

These features have generally been differentiated by their variability when it comes to the description of the recognized shapes of the lung from various perspectives. Five elements have been incorporated to illustrate the geometric characteristics: F1, F2, F5, F6, and F11. For every shape, four features have been determined that represent the shape's structure. The available grey intensity gradations, F9 and F10, are still represented by the texture features.

In this regard, the method differs from previous approaches [13] in that it extracts distinct features for the COVID-19 infection regions that may be at risk rather than analyzing the entire computed tomography image for the lungs or chest. The proposed method ignores the white patches that were utilized in [13] as reference points to emphasize the infection locations. Rather, the infection regions are brought to light by the intensity-based filters employed in this work. Additionally, the infection regions can be described by the proposed method using geometric, structural, and texture-based features. The updated COVID-19 infection region description can enhance the mechanism of advanced classification techniques by providing a variety of alternative numerical features.

4. RESULTS AND DISCUSSION

Here, the proposed method has been implemented using Matlab software v2016b. Tools for the computer vision and image

processing had been developed in the proposed method. The experiment and performance evaluation were carried on the dataset published in [39].

The findings of this work have been presented as Research Questions (RQs) to clarify and highlights the study's larger implication based on the generated results. The research questions have been titled as follows.

- ✓ **RQ1**: Do the extracted features satisfy the successful requirements of the machine learning algorithms for classifying the CT-Images into COVID or NORMAL cases?
- ✓ **RQ2**: In terms of the machine learning algorithms' effective classification, can the extracted features be identified by the minimum and maximum values?
- ✓ **RQ3**: Can the proposed method extract a variety of shapes for the infection regions of the COVID-19 offered by CT-Images?
- ✓ **RQ4**: Can the Neural Netwok and Logistic Regression classify the Computed Tomography Images into COVID and NORMAL cases using the extracted features by the proposed method?

There are two categories of data in the dataset. The first relates to COVID-19 instances, where most of the symptoms for COVID-19 infections have been shown in Computed tomography images (CT-Images). The second category includes normal instances, or those who do not have a radiologist's diagnosis of COVID-19. After extracting the required features from those images, the ORANGE Data mining tool [40] has been used to test the features by the advanced machine learning algorithms.

Training was conducted using the first ten cases of each category. There are 1377 CT-Images for COVID cases and 1462 CT-Images for NORMAL cases. There are 5295 shapes created from the NORMAL CT-Images and 7711 shapes generated from the COVID CT-Images. The second five cases of each category were used for testing purposes. There are 739 CT-Images from the COVID case and 694 CT-Images from the NORMAL case.

Referring to RQ1: Showing how much the extracted features satisfy the classification requirements for the COVID-19 computed tomography images is the primary goal of the classification discussed here, which is based on advanced machine learning algorithms. Six advanced machines learning algorithm have been used in COVID-19 classification, which are Logistic Regression [41], Naïve Bayes [42], Random Forest [43], Decision Tree [44], Adaboost [45] and Neural Network [46]. These algorithms have shown in Table 1.

Actual and Predict are the two parameters that are employed in the machine learning algorithm's training process to meet the training requirements. Each algorithm creates training learners of classifications using the instance data linked to the Actual parameter. On the other hand, in order to enhance the effectiveness of the training process, learners have been tested using the Predict instance values. For the purposes of training and local testing, each algorithm separates the provided data into actual and predict parameters. In this case, there are 7711 COVID instances and 5295 NORMAL instances (a total of 13006 instances) from the training dataset. Thirty percent (30%) of the training data instances are allocated to the local testing phase, while seventy percent (70%) are assigned to the training process.

In Table 1, the successful rates for the selected machine learning algorithms have been displayed based on Actual and Predict parameters. The algorithm has been analysed in each pair of rows in Table 1 to display the success rates of both parameters with regard to COVID and NORMAL categorization.

First up is the Logistic Regression technique, which has success rates of 83.4% and 30.3%, respectively, based on actual parameters for the COVID and NORMAL categories. A total of 9105, or 70% of the training data, have been disseminated, 3696 for NORMAL and 5409 for COVID. A total of 9105, or 70% of the training data, have been disseminated, 3696 for NORMAL and 5409 for COVID. With an accuracy rate of 83.4%, the algorithm has successfully identified 4519 instances as COVID instances. With 1073 out of 3696, the NORMAL had a success rate of 30.3%. Regarding the predict parameter, the COVID and NORMAL samples had successful rates of 61.1% and 56.1%, respectively. For the COVID and NORMAL categories, the successful instance values were 4519 and 1073, respectively, out of a total instance value of 7142 and 1963.

The findings of additional algorithms in Table 1 can be interpreted using the previously mentioned discussion. Following a comparison of the COVID category results, the algorithms were ranked from highest to lowest successful rate, with the algorithm's rank indicated on the right side next to the rate. The two greatest

algorithms are the neural network and logistic regression. When it comes to the predict parameter's successful covid rate, the latter has the top position.

Table 1. A list of machine learning algorithms, together with the Training success rates for the COVID-19 and NORMAL categories, based on Actual and Predict parameters

Algorithm's Name	Sucessful Rate for Actual-COVID-19	Sucessful Rate for Actual-NORMAL	Sucessful Rate for Predict- COVID-19	Sucessful Rate for Predict – NORMAL
Logistic Regression	83.40% **1st/6**	30.30%	63.10% **6th/6**	56.10%
NUM.Instance / Total	4519 / 5409	1073 / 3696	4519 / 7142	1073 / 1963
Naïve Bayes	57.60% **6th/6**	54.90%	64.60% **4th/6**	47.60%
NUM.Instance / Total	3083 / 5352	2062 / 3753	3083 / 4774	2062 / 4331
Random Forest	68.60% **4th/6**	48.70%	65.60% **2nd/6**	52.10%
NUM.Instance / Total	3669 / 5352	1828 / 3753	3669 / 5594	1828 / 3511
Tree	60.10% **5th/6**	50.90%	63.60% **5th/6**	47.30%
NUM.Instance / Total	3219 / 5352	1911 / 3753	3219 / 5061	1911 / 4044
AdaBoost	70.10% **3rd/6**	46%	64.90% **3rd/6**	51.90%
NUM.Instance / Total	3752 / 5352	1725 / 3753	3752 / 5780	1725 / 3325
Neural Network	76.60% **2nd/6**	46.30%	67.10% **1st/6**	58.20%
NUM.Instance / Total	4101 / 5352	1739 / 3753	4101 / 6115	1739 / 2990

In terms of testing, the Logistic Regression and Neural Network have been applied to 3901, or 30%, of the training data. The testing data also include the success rates of different machine learning methods. Both Logistic Regression and Neural Network have been chosen to display the generated testing result as indicated in Table 2 due to their best training rates and goal of summarizing.

Based on the COVID category, the logistic regression has an 83% success rate, or 1957 as the instance testing value out of a total of 2359. Table 2 displays the neural network's success rate for the actual parameter, which is 75.3%.

Although the success rate of the Neural Network is higher than that of the Logistic Regression (64.8% vs. 68.6%), Table 2 illustrates that both algorithms produce satisfactory results for classification when employing the Predict parameter. For both algorithms, the success rates at the Actual parameter are less than 50% based on the NORMAL category. Nonetheless, the rate has increased when the Predict parameter is greater than 50%.

The successful rates generated from both training and Testing process gave an indication that the extracted features played an important role to classify CT-images into COVID or NORMAL cases. Generally, the COVID features have enhanced the learners of the machine learning algorithms. This enhancement was coming from the infection areas have been used to distinguish the COVID CT-Images.

Concerning of the RQ1, the success rates obtained from the training and testing process demonstrated that the extracted features were crucial in identifying CT-Images as either COVID-19 or normal cases. In general, the machine learning algorithms' learners have been improved by the COVID features. This improvement originated from the locations of infection that were utilized to differentiate the COVID-19 CT-Images. On the other hand, the NORMAL category lacks any unique characteristics. The regular shapes are only available on the NORMAL category, and the normal shapes that do not fall inside the COVID category have been utilized by the NORMAL learners. The generated classification results varied between the different machine learning algorithms, providing additional evidence that the extracted feature values vary into different satisfied values. These values allow

the machine learning algorithms to be flexible and adaptive, which in turn allows the CT-Images to be successfully classified into both categories.

Table 2. Testing success rates for Logistic Regression and Neural Network algorithms based on Actual and Predict parameters for COVID-19 and NORMAL categories.

Model Name	Sucessful Rate for Actual - COVID-19	Sucessful Rate for Actual - NORMAL	Sucessful Rate for Predict-COVID-19	Sucessful Rate for Predict-NORMAL
Logistic Regression NUM.Instance / Total	83.00% 1957 / 2359	31.10% 479 / 1542	64.80% 1957 / 3020	54.40% 479 / 881
Neural Network NUM.Instance / Total	75.30% 1776 / 2359	47.20% 728 / 1542	68.60% 1776 / 2590	55.50% 728 / 1311

Each machine learning algorithm's successful instance value was evaluated, and the results showed that each instance's success rate ranged from 100% to 50%. The findings of the Neural Network and Logistic Regression algorithms have been used to assess the retrieved features since they were deemed to be the best algorithms. The Neural Network's instance value with a 100% success rate was roughly 60 instance values. Regarding the Logistic Regression, there were roughly 20 instances. Table 3 shows the minimum and maximum values for each extracted features derived from these instance data. The logistic regression has about 100 instance values for the 99% success rate. With a 99% success rate, 223 instance values have been produced for the neural network.

The values of these extracted features, which are displayed in Table 3, are associated with the CT images that the tomography scanning used to detect the COVID infections on. Following the modification of the success rate to 99%, additional shapes and features for COVID infections were added to the value levels.

Concernig of the RQ2, this demonstrates yet evidence where the features extracted using the proposed method are crucial to the COVID classifications and are determined by the minimum and maximum values.

Error! Reference source not found. shows samples of the various shapes that were extracted; the first row shows the shapes that were extracted using the Neural Network classification method, and the second row shows the shapes extracted using the Logistic Regression method. Each entry for the instance value that has a 100% success rate has a shape that was taken from the training datasets. Different shapes have been successfully extracted using the proposed method each time these shapes have been altered.

Furthermore, even if these shapes have altered in the same CT-Image or in other images, the proposed method can still extract them. As shown in **Error! Reference source not found.**, the shapes are identified by the color red.

Concerning of the RQ3, the proposed method's ability for extraction from the surroundings allows for advanced analysis of these shapes using advanced sophisticated machine learning algorithms.

Five COVID-19 and five NORMAL cases make up another Testing Dataset in which the retrieved features have also been assessed. The statistics values and generated findings for the Testing dataset are displayed in Table 4 and Table 5. The CASE ID has been used to identify these cases. In the field TYPE, a case type has been identified for each case. For every case, the number of CT-Images is displayed in the field (named N. Of CT IMAGES). There were 1433 CT-Images in all. For every case, the proposed

method has generated several shapes, denoted by N. of Shapes. There were 4806 shapes in total for all cases. The analysis of the success rate involves presenting the lowest and greatest success rates for each case according to either COVID-19 or NORMAL-19. The COVID success rates, both minimum and maximum, are shown in fields labelled RATE for COVID MIN and RATE for COVID MAX, respectively.

Table 3. The minimum and maximum values of the extracted features for Neural Network and Logistic Regression algorithms based on 100% success rate of COVID category.

Algorithm's Name / Feature's Name	Neural Network		Logistics Regression	
	Maximum	Minimum	Maximum	Minimum
MajorAxisLength	214.698	36.6586	214.698	68.2995
MinorAxisLength	95.6882	17.5887	95.6882	42.3338
Orientation	89.9096	-89.5673	89.9096	-89.5673
EquivDiameter	81.9222	17.9834	81.9222	49.0034
Perimeter	1232.75	189.596	1232.75	359.169
Eccentricity	0.974282	0.478144	0.964218	0.478144
Solidity	0.788449	0.308705	0.718943	0.352064
Extent	0.628684	0.183054	0.52386	0.183054
EulerNumber	-7	-68	-21	-68
StandardDeviation	24.6713	1.75688	24.3838	7.8119
NumLevels	5	2	5	3

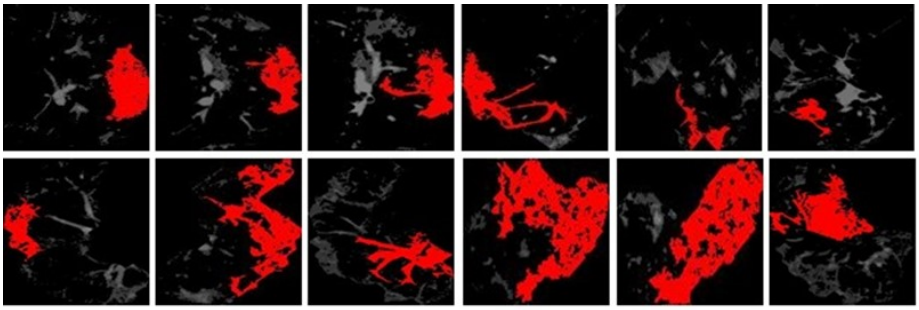


Figure 5. First and second row samples of COVID-19 infections extracted from the 100 success rates for Neural Network and Logistic Regression, respectively

Table 4. The statistics values and the generated results based on the Neural Network algorithms for the 5 cases COVID and 5 cases NORMAL

CASE ID	TYPE	DETECT	N. Of CT IMAGES	N. of Shapes	RATE for COVID MIN	RATE for COVID MAX	RATE for NORMAL MIN	RATE for NORMAL MAX
1	COVID	COVID	139	201	0.501436	0.942056	0.0872676	0.493262
2	NORMAL	NORMAL	132	124	0.0373141	0.490707	0.500742	0.651221
3	COVID	COVID	144	417	0.500135	0.897826	0.0760182	0.499413
4	NORMAL	NORMAL	137	238	0.032017	0.498289	0.507023	0.869985
5	NORMAL	NORMAL	128	330	0.120917	0.49997	0.500507	0.986866
6	COVID	COVID	144	1325	0.500419	1	0.011139	0.499581
7	COVID	COVID	144	739	0.500467	0.999964	0.250651	0.497917
8	NORMAL	NORMAL	143	160	0.0793059	0.49973	0.501446	0.95039

9	COVID	COVID	168	753	0.500173	0.997911	0.209359	0.49921
10	NORMAL	NORMAL	154	519	0.0857266	0.498604	0.500001	0.872109
		SUM	1433	4806				

Table 5. The statistics values and the generated results based on Logistic Regression algorithms for the 5 cases COVID and 5 cases NORMAL.

CASE ID	TYPE	DETECT	N. Of CT IMAGES	N. of Shapes	RATE for COVID MIN	RATE for COVID MAX	RATE for NORMAL MIN	RATE for NORMAL MAX
1	COVID	COVID	139	201	0.500331	0.769391	0.250275	0.499836
2	NORMAL	NORMAL	132	124	0.182279	0.494033	0.517203	0.74528
3	COVID	COVID	144	417	0.500332	0.794131	0.481535	0.499967
4	NORMAL	NORMAL	137	238	0.0506953	0.498817	0.500641	0.642121
5	NORMAL	NORMAL	128	330	0.0737963	0.499848	0.500042	0.936603
6	COVID	COVID	144	1325	0.501306	1	0.0418164	0.499975
7	COVID	COVID	144	739	0.500737	0.999117	0.352228	0.499394
8	NORMAL	NORMAL	143	160	0.209681	0.49805	0.500846	0.980608
9	COVID	COVID	168	753	0.500041	0.991203	0.3085	0.499854
10	NORMAL	NORMAL	154	519	0.0573432	0.499691	0.500265	0.826262
		SUM	1433	4806				

As for the minimum and maximum success rate of NORMAL have been shown in the fields tilted by RATE for NORMAL MIN and RATE for NORMAL MAX respectively. Every shape of the testing instance values has been evaluated by the machine learning algorithm. After evaluating, these testing instance values for each case have been divided into COVID or NORMAL, and the success rate has been then determined. The fields labelled RATE for NORMAL MAX and RATE for NORMAL MIN reflect the minimum and maximum success rates for NORMAL, respectively. The machine learning algorithms have assessed each shape of the testing case values. Following evaluation, the success rate was ascertained by dividing these testing instance values for each case into COVID and NORMAL categories. For the COVID and NORMAL categories, two minimum values as well as two maximum values have been established. The field DETECT is changed to COVID when the case has been identified as COVID and the range of the COVID success rate (i.e., from minimum to maximum) exceeds the NORMAL range. On the other hand, when the reverse comparison is made, the value NORMAL is found. Table 4 and Table 5 both have a darker grey color to further identify the greatest range.

Table 4 displays the Neural Network algorithm's results, which show that the studied cases were correctly divided into COVID and NORMAL categories. Furthermore, as Table 5. illustrates, the Logistic Regression has also been successful in identifying which patients are COVID or NORMAL. Given that the outcomes of both algorithms are essentially the same, the extracted features were essential to making both algorithms flexible.

Based on the Neural Network method, the five COVID examples have, in deep perspective, a success percentage of 94.2%, 89.8%, 100%, 99.9%, and 99.1%, respectively (Table 4). However, for the same COVID scenario, the Logistic Regression produced76.9%,79.4%,100%,99.9, and 99.1%, as indicated in Table 5. In context of the RQ4, these findings demonstrate that, in situations where the performance of the classification tools varies, the extracted features can provide the satisfactory classification values needed to distinguish between the COVID and NORMAL categories. Therefore, the proposed method has 100% an accuracy rate in the COVID-19 classification.

Table 6 presents the accuracy attained by most approaches presented in the article's related work in relation to the concept utilized to extract features of interest, as compared to other

approaches. By utilizing combination features based on geometry, structure, and texture features, the proposed method sets itself apart. The accuracy percentage of the proposed method ranges from 99% to 100%.

Table 6. a comparison of the accuracy rate and feature extraction concept between the proposed method and other approaches.

Article REF	The Used Concept	Accuracy Rate
[16]	Wavelet	From 97.6% to 98.5%.
[15]		From 99.84% to 99.90%.
[17]	Texture	93.96%
[18]		96.86%
[20]		92.37%.
[21]	Convolution	95.4%.
[22]		98.00%
Proposed Method	Geometry, Structural and Texture	99% to 100%

5. CONCLUSION

Diagnosing the COVID-19 patiants is crucial to keeping an eye on the afflicted individual before their condition gets worse. Therefore, in each computed tomography scan, the radiologist needs to monitor the areas affected by the illness. All computed tomography images have different features related to infection. To determine the severity of the infection and diagnose the CT images, it is crucial to extract features that accurately depict the COVID-19 infection. This work's primary contribution is to introduce a proposed method for extracting relevant features from CT-Images while taking into account the texture, geometry, and structural characteristics. The machine learning algorithms have evaluated the proposed method on a Computed Tomography dataset of COVID-19 patients. From then, the CT-Images have been successfully classified into COVID and NORMAL categories using a neural network and logistic regression. There was a range of 99% to 100% in the classification accuracy rate and numerous physical characteristics of the infection region of COVID-19 virus have been identified.

Furthermore, those extracted features will be crucial in the future for refining the advanced Convolution Neural Network (CNN), which will be used to determine the severity of the COVID-19 infection regions and to classify COVID-19 and NORMAL instances. The reason why the retrieved features are significant is because they provide different numerical characteristics for the COVID-19 infection region. CNN can be then adjusted and changed to utilize these numerical values in addition to the infection regions' intensities values.

ACKNOWLEDGEMENT

This study was supported in part by Tafila Technical University grants. The first author was sponsored by a Fulbright scholarship while he was in USA.

REFERENCES:

[1] P. Subramani, S. K, K. R. B, S. R, and P. B D, “Prediction of muscular paralysis disease based on hybrid feature extraction with machine learning technique for COVID-19 and post-COVID-19 patients,” *Pers. Ubiquitous Comput.*, vol. 27, no. 3, pp. 831–844, 2023, doi: 10.1007/s00779-021-01531-6.

[2] A. Neisi *et al.*, “Association of the corona virus (Covid-19) epidemic with environmental risk factors,” *Environ. Sci. Pollut. Res. Int.*, vol. 30, no. 21, pp. 60314–60325, May 2023, doi: 10.1007/s11356-023-26647-5.

[3] H. Musa, T. H. Musa, I. H. Musa, and O. Oderinde, “Letter to Editor COVID-19 outbreak and medical waste: Challenge in hand,” *Int. J. Public Health Sci. IJPHS*, vol. 9, no. 3, Art. no. 3, Sep. 2020, doi: 10.11591/ijphs.v9i3.20523.

[4] L. Amelia and R. A. Syakurah, “Analysis of public search interest towards immune system improvement during the COVID-19 pandemic using Google Trends,” *Int. J. Public Health Sci. IJPHS*, vol. 9, no. 4, Art. no. 4, Dec. 2020, doi: 10.11591/ijphs.v9i4.20518.

[5] K. Syauqi, S. Munadi, and M. B. Triyono, “Students’ Perceptions toward Vocational Education on Online Learning during the COVID-19 Pandemic,” *Int. J. Eval. Res. Educ.*, vol. 9, no. 4, pp. 881–886, Dec. 2020.

[6] D. Sulisworo, N. Fatimah, S. S. Sunaryati, and Sanidi, “A Quick Study on SRL Profiles of Online Learning Participants during the Anticipation of the Spread of COVID-19,” *Int. J. Eval. Res. Educ.*, vol. 9, no. 3, pp. 723–730, Sep. 2020.

[7] A. L. Aswathy, H. S. Anand, and S. S. V. Chandra, “COVID-19 severity detection

using machine learning techniques from CT-images," *Evol. Intell.*, vol. 16, no. 4, pp. 1423–1431, Aug. 2023, doi: 10.1007/s12065-022-00739-6.

[8] S. Sarv Ahrabi, A. Momenzadeh, E. Baccarelli, M. Scarpiniti, and L. Piazzo, "How much BiGAN and CycleGAN-learned hidden features are effective for COVID-19 detection from CT images? A comparative study," *J. Supercomput.*, vol. 79, no. 3, pp. 2850–2881, 2023, doi: 10.1007/s11227-022-04775-y.

[9] D. C. Tran, "An open toolbox for generating map of actively confirmed SARS-CoV-2 or COVID-19 cases in Vietnam," *Bull. Electr. Eng. Inform.*, vol. 9, no. 6, Art. no. 6, Dec. 2020, doi: 10.11591/eei.v9i6.2621.

[10] M. Y. Kamil, "A deep learning framework to detect Covid-19 disease via chest X-ray and CT scan images," *Int. J. Electr. Comput. Eng. IJECE*, vol. 11, no. 1, Art. no. 1, Feb. 2021, doi: 10.11591/ijece.v11i1.pp844-850.

[11] N. Alrefai and O. Ibrahim, "Deep learning for COVID-19 diagnosis based on chest X-ray images," *Int. J. Electr. Comput. Eng. IJECE*, vol. 11, no. 5, Art. no. 5, Oct. 2021, doi: 10.11591/ijece.v11i5.pp4531-4541.

[12] A. A. Abdulmunem, Z. A. Abutiheen, and H. J. Aleqabie, "Recognition of corona virus disease (COVID-19) using deep learning network," *Int. J. Electr. Comput. Eng. IJECE*, vol. 11, no. 1, Art. no. 1, Feb. 2021, doi: 10.11591/ijece.v11i1.pp365-374.

[13] R. Supriyanti, M. R. Kurniawan, Y. Ramadhani, and H. B. Widodo, "Calculating the area of white spots on the lungs of patients with COVID-19 using the Sauvola thresholding method," *Int. J. Electr. Comput. Eng. IJECE*, vol. 13, no. 1, Art. no. 1, Feb. 2023, doi: 10.11591/ijece.v13i1.pp315-324.

[14] V. V. S. Tallapragada, N. A. Manga, and G. V. P. Kumar, "A novel COVID diagnosis and feature extraction based on discrete wavelet model and classification using X-ray and CT images," *Multimed. Tools Appl.*, vol. 82, no. 17, pp. 26183–26224, Jul. 2023, doi: 10.1007/s11042-023-14367-4.

[15] S. Rashidi and A. Nokhostin, "COVID-19 Diagnosis by Extracting New Features from Lung CT Images Using Fractional Fourier Transform." Preprints, May 08, 2023. doi: 10.20944/preprints202305.0489.v1.

[16] H. Elshenbary, E. Ebeid, and D. Baleanu, "COVID-19 classification using hybrid deep learning and standard feature extraction techniques," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 29, pp. 1780–1791, Mar. 2023, doi: 10.11591/ijeecs.v29.i3.pp1780-1791.

[17] R. K. Patel and M. Kashyap, "Automated diagnosis of COVID stages using texture-based Gabor features in variational mode decomposition from CT images," *Int. J. Imaging Syst. Technol.*, 2023, Accessed: Feb. 01, 2024. [Online]. Available: https://doi.org/10.1002/ima.22865

[18] N. Suganthi and K. Sarojini, "Feature Extraction Based on GLCM and GLRM Methods on COVID-19 Dataset," in *Computer Vision and Machine Intelligence Paradigms for SDGs*, R. J. Kannan, S. M. Thampi, and S.-H. Wang, Eds., in Lecture Notes in Electrical Engineering. Singapore: Springer Nature, 2023, pp. 271–279. doi: 10.1007/978-981-19-7169-3_25.

[19] A. A. Akl, K. M. Hosny, M. M. Fouda, and A. Salah, "A hybrid CNN and ensemble model for COVID-19 lung infection detection on chest CT scans," *PloS One*, vol. 18, no. 3, p. e0282608, 2023, doi: 10.1371/journal.pone.0282608.

[20] S. Punitha, T. Stephan, R. Kannan, M. Mahmud, M. S. Kaiser, and S. B. Belhaouari, "Detecting COVID-19 From Lung Computed Tomography Images: A Swarm Optimized Artificial Neural Network Approach," *IEEE Access*, vol. 11, pp. 12378–12393, 2023, doi: 10.1109/ACCESS.2023.3236812.

[21] S. Kordnoori, M. Sabeti, H. Mostafaei, and S. S. A. Banihashemi, "Analysis of lung scan imaging using deep multi-task learning structure for Covid-19 disease," *IET Image Process.*, vol. 17, no. 5, pp. 1534–1545, 2023, doi: 10.1049/ipr2.12736.

[22] N. D. Kathamuthu *et al.*, "A deep transfer learning-based convolution neural network model for COVID-19 detection using computed tomography scan images for medical applications," *Adv. Eng. Softw.*, vol. 175, p. 103317, Jan. 2023, doi: 10.1016/j.advengsoft.2022.103317.

[23] S. H. Khan *et al.*, "COVID-19 detection and analysis from lung CT images using novel channel boosted CNNs," *Expert Syst. Appl.*, vol. 229, p. 120477, Nov. 2023, doi: 10.1016/j.eswa.2023.120477.

[24] M. Ağralı *et al.*, "DeepChestNet : Artificial intelligence approach for COVID -19

detection on computed tomography images," *Int. J. Imaging Syst. Technol.*, vol. 33, Mar. 2023, doi: 10.1002/ima.22876.

[25] A. Motwani *et al.*, "Enhanced framework for COVID-19 prediction with computed tomography scan images using dense convolutional neural network and novel loss function," *Comput. Electr. Eng.*, vol. 105, p. 108479, Jan. 2023, doi: 10.1016/j.compeleceng.2022.108479.

[26] N. Nasir *et al.*, "Multi-modal image classification of COVID-19 cases using computed tomography and X-rays scans," *Intell. Syst. Appl.*, vol. 17, p. 200160, Feb. 2023, doi: 10.1016/j.iswa.2022.200160.

[27] Q. Lei, J. Zhong, and C. Wang, "Joint Optimization of Crack Segmentation With an Adaptive Dynamic Threshold Module," *IEEE Trans. Intell. Transp. Syst.*, pp. 1–15, 2024, doi: 10.1109/TITS.2023.3348812.

[28] Y. Huang *et al.*, "MAST: An Earthquake-Triggered Landslides Extraction Method Combining Morphological Analysis Edge Recognition With Swin-Transformer Deep Learning Model," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 17, pp. 2586–2595, 2024, doi: 10.1109/JSTARS.2023.3342989.

[29] M. Esmaeili, D. Abbasi-Moghadam, A. Sharifi, A. Tariq, and Q. Li, "ResMorCNN Model: Hyperspectral Images Classification Using Residual-Injection Morphological Features and 3DCNN Layers," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 17, pp. 219–243, 2024, doi: 10.1109/JSTARS.2023.3328389.

[30] A. R. Beeravolu *et al.*, "Methods for Detection and Measurement of Potential Broncho-Arterial Pairs in Chest Computed Tomography Scans of Children," *IEEE Access*, vol. 12, pp. 4730–4744, 2024, doi: 10.1109/ACCESS.2023.3347344.

[31] W. Zhang, F. Zhou, Y. Liu, and L. Yang, "Precise Calibration of Binocular Vision System Based on Oblique Cone Projection Model," *IEEE Sens. J.*, vol. 24, no. 1, pp. 437–448, Jan. 2024, doi: 10.1109/JSEN.2023.3330984.

[32] S. Prabu, M. Aravind, J. P. Muthu Aravindh, K. Maruthupandiyan, and M. B. Yokeshwaran, "Detection and Classification of Brain Tumors using Magnetic Resonance Imaging and Artificial Neural Networks: A Novel Approach," in *2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC)*, Jun. 2023, pp. 1–6. doi: 10.1109/ICAISC58445.2023.10199859.

[33] C.-C. Chen, Y.-L. Chen, and J.-D. Lee, "A Model For IC Defect Identification Based on Feature Extraction," in *2023 IEEE 5th Eurasia Conference on IOT, Communication and Engineering (ECICE)*, Oct. 2023, pp. 139–144. doi: 10.1109/ECICE59523.2023.10383129.

[34] S. Sun, J. Wang, and D. Tan, "Ship Detection in SAR Images Based on CFAR and Machine Learning," in *2023 16th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, Oct. 2023, pp. 1–6. doi: 10.1109/CISP-BMEI60920.2023.10373385.

[35] S. Steuernagel, K. Thormann, and M. Baum, "Improved Extended Object Tracking with Efficient Particle-based Orientation Estimation," in *2023 26th International Conference on Information Fusion (FUSION)*, Jun. 2023, pp. 1–8. doi: 10.23919/FUSION52260.2023.10224128.

[36] T. N. S. Yaswanth, K. S. V. Lakshmi, K. G. Vamsi, D. Supraja, and K. Suvarna Vani, "Classification of Medicinal Leaves using SVM," in *2023 World Conference on Communication & Computing (WCONF)*, Jul. 2023, pp. 1–7. doi: 10.1109/WCONF58270.2023.10235087.

[37] T. Kinoshita, K. Yamanaka, K. Konno, and Y. Tokuyama, "A study of geometric shape of polygons for additive manufacturing," in *2018 International Workshop on Advanced Image Technology (IWAIT)*, Jan. 2018, pp. 1–4. doi: 10.1109/IWAIT.2018.8369698.

[38] T. Gao, L. Zheng, X. Wu, H. Ji, B. Yang, and W. Xu, "Research on Modelling the Nonlinear Response Function of TDI CMOS Imaging System," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, pp. 1–12, 2024, doi: 10.1109/JSTARS.2024.3354911.

[39] S. Heidarian *et al.*, "COVID-19 Low-Dose and Ultra-Low-Dose CT Scans." IEEE, May 28, 2021. doi: https://dx.doi.org/10.21227/sed8-6r15.

[40] B. L. Ljubljana University of, "Orange Data Mining," Orange Data Mining. Accessed: Feb. 01, 2024. [Online]. Available: https://orangedatamining.com

[41] X. Lu, H. U. Sami, and B. Güler, "SCALR: Communication-Efficient Secure Multi-Party Logistic Regression," *IEEE Trans.*

Commun., vol. 72, no. 1, pp. 162–178, Jan. 2024, doi: 10.1109/TCOMM.2023.3308954.

[42] A. K. Singh, R. Sunkara, G. R. Kadambi, and V. Palade, "Spectral–Spatial Classification With Naive Bayes and Adaptive FFT for Improved Classification Accuracy of Hyperspectral Images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 17, pp. 1100–1113, 2024, doi: 10.1109/JSTARS.2023.3327346.

[43] D. Saha, Md. E. Hoque, and M. E. H. Chowdhury, "Enhancing Bearing Fault Diagnosis Using Transfer Learning and Random Forest Classification: A Comparative Study on Variable Working Conditions," *IEEE Access*, vol. 12, pp. 5986–6000, 2024, doi: 10.1109/ACCESS.2023.3347345.

[44] I. Krop, Y. Takahashi, T. Sasaoka, H. Shimada, A. Hamanaka, and J. Onyango, "Assessment of Selected Machine Learning Models for Intelligent Classification of Flyrock Hazard in an Open Pit Mine," *IEEE Access*, vol. 12, pp. 8585–8608, 2024, doi: 10.1109/ACCESS.2024.3352733.

[45] R. Zhou, J. Wang, K. Huang, H. Wang, and Z. Chen, "MGRBA: Gas Recognition With Biclustering and Adaboost," *IEEE Access*, vol. 12, pp. 4328–4338, 2024, doi: 10.1109/ACCESS.2023.3342915.

[46] M. S. Pandianchery, V. Sowmya, E. A. Gopalakrishnan, V. Ravi, and K. P. Soman, "Centralized CNN–GRU Model by Federated Learning for COVID-19 Prediction in India," *IEEE Trans. Comput. Soc. Syst.*, vol. 11, no. 1, pp. 1362–1371, Feb. 2024, doi: 10.1109/TCSS.2023.3250656.