Sakarya University Journal of Computer and Information Sciences
http://saucis.sakarya.edu.tr/

DOI: 10.35377/saucis…1309970 RESEARCH ARTICLE

# Rapid and Precise Identification of COVID-19 through Segmentation and Classification of CT and X-ray Images

**Ahmet SAYGILI** [1]

[1]Tekirdağ Namık Kemal University, Çorlu Faculty of Engineering, Department of Computer Engineering, Çorlu/Türkiye

**Corresponding author:**
Ahmet SAYGILI
asaygili@nku.edu.tr

**Submitted:** 05 June 2023
**Revision Requested:** 03 August 2023
**Last Revision Received:** 06 August 2023
**Accepted:** 10 August 2023
**Published Online:** 10 August 2023

**Citation:** Saygılı A. (2023). Rapid and Precise Identification of COVID-19 through Segmentation and Classification of CT and X-ray Images. *Sakarya University Journal of Computer and Information Sciences*. 6 (2) https://doi.org/10.35377/saucis...1309970

## ABSTRACT

The COVID-19 pandemic, caused by a novel coronavirus, has become a global epidemic. Although the reverse transcription-polymerase chain reaction (RT-PCR) test is the current gold standard for detecting the virus, its low reliability has led to the use of CT and X-ray imaging in diagnostics. As limited vaccine availability necessitates rapid and accurate detection, this study applies k-means and fuzzy c-means segmentation to CT and X-ray images to classify COVID-19 cases as either diseased or healthy for CT scans and diseased, healthy, or non-COVID pneumonia for X-rays. Our research employs four open-access, widely used datasets and is conducted in four stages: preprocessing, segmentation, feature extraction, and classification. During feature extraction, we employ the Gray-Level Co-Occurrence Matrix (GLCM), Local Binary Pattern (LBP), and Histogram of Oriented Gradients (HOG). In the classification process, our approach involves utilizing k-Nearest Neighbor (kNN), Support Vector Machines (SVM), and Extreme Learning Machines (ELM) techniques. Our research achieved a sensitivity rate exceeding 99%, which is higher than the 60-70% sensitivity rate of PCR tests. As a result, our study can serve as a decision support system that can help medical professionals make rapid and precise diagnoses with a high level of sensitivity.

**Keywords:** COVID-19, Diagnosis, Imaging techniques, Segmentation methods, Machine learning-based classification

## 1. Introduction

COVID-19, which emerged in Wuhan, China, has rapidly spread globally, affecting a vast number of individuals. As of May 5, 2021, over 153 million people have been afflicted by the disease, and the death toll has surpassed 3.2 million [1]. This pandemic has caused profound impacts on society, manifesting various physical, mental, psychological, and sociological repercussions. A study conducted in early 2020 on 1,210 participants from 194 cities in China revealed that 54% of respondents rated the psychological impact of the COVID-19 outbreak as moderate to severe, with 29% exhibiting moderate to severe anxiety symptoms [2]. COVID-19 can present diverse symptoms in individuals, with fever and cough being the most prevalent [3]. However, the disease's asymptomatic transmission necessitates early diagnosis and isolation to interrupt the transmission chain and contain the epidemic [4]. It is imperative to resolve COVID-19 as soon as possible due to its detrimental effects on humanity.

The gold standard for diagnosing COVID-19 is the reverse transcription-polymerase chain reaction (RT-PCR) test [5], [6]. However, this testing method has limitations, including high false-negative rates and sensitivity rates that can sometimes be as low as 60% [7]. COVID-19 primarily affects the lungs and can be detected through radiological imaging, such as X-rays and computed tomography (CT) scans [8], [9]. Although effective, manual evaluation of these images can be time-consuming, particularly for a widespread disease like COVID-19. To overcome this issue, a computer-aided diagnostic method has been

developed to predict COVID-19 from radiological images, using distinct patterns such as ground-glass opacity and embolism showing linear consolidation [10]. COVID-19 and pneumonia are diseases that have similar symptoms that affect the lungs and cause respiratory problems. Therefore, it is a very difficult process to distinguish between COVID-19 and non-COVID-19 Pneumonia. As a result of all these diagnostic and diagnostic difficulties and findings, our study has been carried out with the idea that a fast and effective automatic diagnosis system can be established. This study builds upon the knowledge and expertise obtained from previous studies [11–13], but it differs in several ways. First, it compares different clustering methods to generate parameters for various classification methods. Second, it uses three different feature selection methods and models with both CT and X-ray images, as well as four different datasets. Another distinguishing feature of this study from [11–13] is the analysis of combined data sets from different image sources, where both healthy and diseased images were merged and classified. Efforts were made to improve the performance of the system in this context.

The fundamental objective of this research is to create an automated system based on segmentation and machine learning techniques, capable of identifying COVID-19 disease on CT and X-ray images. We consolidated four datasets from different COVID-19 patient studies into one study. We performed binary classification to distinguish COVID-19 cases from non-COVID-19 cases, and COVID-19 positive cases from non-COVID-19 pneumonia cases. In the three-class classification process, we classified cases as COVID-19, non-COVID-induced pneumonia, and non-COVID-19. The data sets used were publicly accessible and were chosen for their comprehensiveness. Our study outperforms previous studies in terms of accuracy, speed, and other performance metrics. Furthermore, our study's high-performance rates across different imaging modalities and data sets indicate its adaptability to various data sets.

The subsequent sections of this paper are structured as follows. Section 2 comprises a comprehensive review of pertinent literature regarding COVID-19 diagnosis. Section 3 delineates the clinical aspects of the issue, outlines the utilized methodologies for detecting COVID-19 cases, and describes the datasets used in the experiment. Section 4 presents the experimental design and results of the COVID-19 detection. In Section 5, the outcomes of this study are analyzed and compared with those of other relevant studies in the literature. Finally, Section 6 concludes the article and proposes future research directions.

## 2. Literature Review

Numerous computer-based studies have been conducted in the literature for the detection and diagnosis of the COVID-19 virus, utilizing either CT or X-ray images. Specifically, studies employing X-ray images for the diagnosis of COVID-19 are as follows:

- Abraham et al. [14] conducted a study on the diagnosis of COVID-19 using two different X-ray datasets, where they employed Multi Convolutional Neural Networks (CNN) and BayesNet classifier. The initial dataset in the study was comprised of 453 samples positive for COVID-19 and 497 samples negative for COVID-19. The second dataset included 71 samples positive for COVID-19 and 7 samples negative for COVID-19. The study was able to achieve success rates of 91.16% and 97.44%, respectively, for these datasets. Joshi et al. utilized deep learning techniques for binary and four-class classification processes consisting of COVID-19, non-COVID-19, bacterial pneumonia, and viral pneumonia [15]. The accuracy rate for the binary classification process was determined to be 99.61%, whereas the multiclass classification process obtained an accuracy rate of 94.79%.
- Aslan et al. conducted a study in which they used X-ray images for detecting COVID-19 infection and segmenting lung images [16]. The study used 219 COVID-19 positive, 1341 COVID-19 negative, and 1345 viral pneumonia X-ray images, achieving an accuracy rate of 98.7%.
- Demir suggested an LSTM model-based technique for the automatic detection of COVID-19 cases using X-ray images [17]. The images were subjected to Sobel gradient and watershed segmentation during the preprocessing stage to improve the model's performance.
- Öztürk et al. conducted the DarkCovidNet study, which achieved a success rate of 98.08% using 1125 X-ray images, including 125 COVID-19 positive, 500 No Findings, and 500 Pneumonia images [18].
- Altan and Karasu utilized deep learning techniques in their study and achieved an accuracy rate of 99.69% using positive, negative, and viral pneumonia images [19].
- Nour et al. proposed a system that could classify X-ray images into three categories, COVID-19 positive, COVID-19 negative, and viral pneumonia, using Deep Features and Bayesian Optimization methods [20].
- Gupta et al. obtained a 99% recall and precision value in their study, named Integrated Stacking InstaCovNet-19, using a dataset consisting of 361 COVID-19-positive, 365 healthy, and 362 pneumonia images [21].
- Juel et al. propose a deep learning approach for recognizing Covid-19, Viral Pneumonia, Lung-Opacity, and normal chest from chest X-ray images using a Modified Convolutional Neural Network (M-CNN) and Bidirectional LSTM (BiLSTM) with a Multi-Support Vector Machine (M-SVM) classifier. The proposed method achieves an accuracy of 98.67% [22].
- Kong and Cheng propose a method for classifying chest X-ray images of COVID-19 patients using a fusion of DenseNet and VGG16 networks with an attention mechanism and ResNet segmentation. The proposed model

achieved 98% accuracy in binary and 97.3% three-category classification, making it a useful tool for clinicians and radiologists [23].

The studies utilizing CT images for the detection of COVID-19 in computer-aided studies in the literature are as follows:

- Mishra et al. used VGG16 and ResNet50 transfer learning methods for the detection of COVID-19 in CT images. In addition to a binary classification of COVID-19 and non-COVID-19, the study also performed multiclass classification as COVID-19, normal, and pneumonia. The study achieved 99% success in binary classification with VGG16 and ResNet50. The multiple classification performance results for VGG16 and ResNet50 were 86.74% and 88.52%, respectively [24].
- Chakraborty et al. used meta-heuristic and fuzzy segmentation methods to segment COVID-19 CT images and anomalies [25].
- Ardakani et al. classified COVID-19 and non-COVID-19 CT images using 10 deep-learning models and found that ResNet and Xception models were the most successful ones [26].
- Gilanie et al. classified three different public datasets and achieved accuracy, specificity, and sensitivity values of 96.68%, 95.65%, and 96.24%, respectively [27].
- Kalane et al. used a fully convolutional network method based on U-Net architecture for the automatic detection of COVID-19 in 1000 chest CT images and obtained an overall accuracy value of 94.10% [28].
- In their study, Li et al. utilized the 3D ResNet-18 model and reported a precision rate of 89.6% for detecting COVID-19 cases. The dataset they used consisted of 305 COVID-19-positive cases, 872 Community-Acquired Pneumonia (CAP) cases, and 1498 non-pneumonia CT images [29].
- Xu et al. proposed a new method for the detection of COVID-19 based on segmentation and feature extraction of CT images using the 3D-CNN model [30]. The accuracy rate for the binary classification process was determined to be 99.61%, whereas the multiclass classification process obtained an accuracy rate of 94.79%.
- Jaiswal et al. used the DenseNet201-based deep learning method for the classification of COVID-19 CT images as diseased or healthy and achieved an accuracy of 97% in the study [31].
- Kathamuthu et al. propose a deep transfer learning-based convolution neural network model for COVID-19 detection using computed tomography scan images for medical applications. The model aims to detect the presence of COVID-19 in chest CT images and has shown promising outcomes in detecting and monitoring COVID-19 patients. The VGG16 model in the paper achieved an accuracy of 98.00% [32].

Numerous studies have been conducted in the literature focusing on the detection, segmentation, and classification of COVID-19. In addition to the aforementioned studies, Göreke et al. proposed a system that employs blood data of COVID-19 patients to assist experts [33]. Using deep neural networks, they determined that ethnic and genetic differences have an impact on disease diagnosis. Based on the literature review, it can be observed that chest CT and X-ray imaging play a crucial role in identifying abnormalities for detecting COVID-19, and the implementation of image processing methods with machine learning algorithms has enabled this capability.

## 3. Material and Methods

### 3.1 Dataset

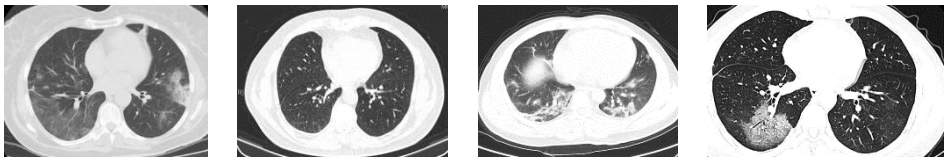
In our study, we performed analysis procedures on four different publicly accessible datasets, consisting of two X-ray and two CT image datasets. We aimed to select datasets frequently used in the literature to ensure easy comparisons of success rates. Table 1 provides details on the datasets, which contain varying numbers of samples from different locations. The study results are expected to be suitable for different COVID-19 variants represented in the datasets from various locations. Notably, the X-ray datasets include three distinct classes, unlike the CT datasets, comprising healthy and diseased lung images and pneumonia class not caused by COVID-19. This differentiation can aid healthcare professionals in distinguishing between pneumonia associated with COVID-19 and pneumonia that is not related to it, which can be difficult due to the severity of COVID-19. Our research aims to make a valuable contribution to the process of expert diagnosis in this regard.

Our study conducted analysis procedures on four different publicly accessible datasets, comprising two X-ray and two CT image datasets. We have selected datasets that are widely used in the literature to ensure that the obtained success rates can be easily compared. The details of the data used in the study are presented in Table 1. As shown in the table, the data sets consist of different numbers of samples obtained from various locations. It can be inferred that the findings of the study apply to different variants of COVID-19 based on the data sets from diverse locations. Moreover, the X-ray data sets consist of three distinct categories, which differ from the CT data sets. Whereas the CT images include healthy and diseased lung images, the X-ray images also incorporate the pneumonia category caused by reasons other than COVID-19. This distinction between COVID-19-induced pneumonia and pneumonia caused by other factors is significant, as medical professionals may sometimes misinterpret this due to the severity of the virus. Hence, this study can potentially help medical professionals to diagnose accurately.

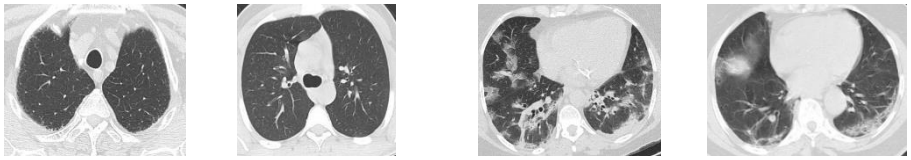
Table 1. Dataset information

| | Type | # of class | # of cases | Location | Citation |
|---|---|---|---|---|---|
| **Dataset 1** | CT | 2 | (+) 349<br>(-) 397 | China | [34] |
| **Dataset 2** | CT | 2 | (+) 1252<br>(-) 1230 | Sao Paulo, Brazil | [35] |
| **Dataset 3** | X-ray | 3 | (+) 125<br>(-) 500<br>Non- COVID Pneumonia 500 | Mixed (Spain, Canada, United Kingdom, etc.) | [18], [36], [37] |
| **Dataset 4** | X-ray | 3 | (+) 1200<br>(-) 1341<br>Non- COVID Pneumonia 1345 | Italy | [38], [39] |

In our study, four distinct data sets containing CT and X-ray images were utilized. Sample images from these data sets are illustrated in Figures 1 and 2. Figure 1 presents the normal and COVID-19 images of Datasets 1 and 2, which include CT images. The contrast between CT images of healthy individuals and those diagnosed with COVID-19 can be observed in the same figure.



Normal Cases (Dataset 1) COVID-19 affected cases (Dataset 1)

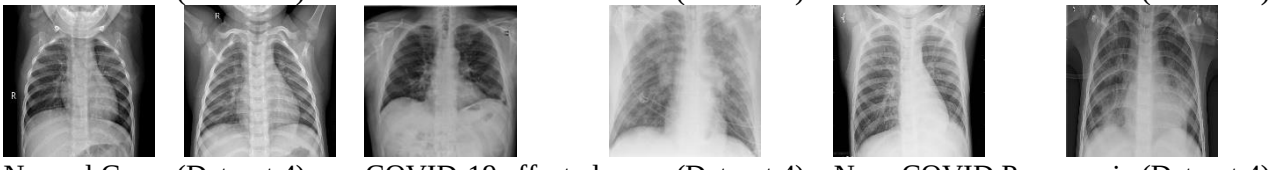


Normal Cases (Dataset 2) COVID-19 affected cases (Dataset 2)

Figure 1. Sample CT images from Dataset 1 [34] and Dataset 2 [35]

Datasets 3 and 4, comprising X-ray images, contain three distinct classes. Figure 2 displays representative images from these datasets that depict normal lung X-rays as well as those depicting COVID-19 and non-COVID-19 pneumonia. Notably, it is essential to develop methods to differentiate between pneumonia caused by COVID-19 and other etiologies.

Normal Cases (Dataset 3) COVID-19 affected cases (Dataset 3) Non- COVID Pneumonia (Dataset 3)



Normal Cases (Dataset 4) COVID-19 affected cases (Dataset 4) Non- COVID Pneumonia (Dataset 4)

Figure 2. Sample X-ray images from Dataset 3 [18], [36], [37] and Dataset 4 [38], [39]

### 3.2 Preprocessing

The data sets utilized in our study contained images with diverse dimensions and sizes. Therefore, we first converted all images to the same format, namely PNG format, and subsequently transformed them into gray levels. This resulted in 2-channel images that are less computationally expensive to process compared to 3-channel images. The sizes of the images in

the four different datasets varied significantly. This guarantee that the study can be used for various data sets, all images were resized and converted to the dimensions of 256 x 256. In the next step, we performed image sharpening by setting the Standard deviation of the Gaussian low pass filter value and sharpening value as '2'. Following that, the image intensity values were normalized from the range of [0, 1] to the range of [0, 0.87], which accentuates the dark regions relatively more. Our final preprocessing step was the application of the Wiener filter [40], which we applied after segmentation because it improved the success of the experimental studies in our research. The Wiener filter is expressed by the formula below:

$$W(u,v) = \frac{H(u,v)}{|H(u,v)|^2 + S_{nx}(u,v)} \quad (1)$$

In the context of image processing, the noise ratio is denoted as $S_{nx}(u,v)$, while the sinc function of the target pixel is represented by H (u, v). A neighborhood dimension of [4 4] is set for the Wiener filter. The parameters of the processes applied in the preprocessing stage were determined experimentally. After conducting experimental studies, the most appropriate values were determined.

### 3.3 Segmentation

The images in the datasets used in our study have homogenous regions due to their characteristics such as gray tones and texture features [41]. These features make them suitable for various clustering or segmentation methods to detect desired regions. In our study, we used k-means (KM) and fuzzy c-means (FCM) segmentation methods to identify anomalies in CT and X-ray images.

The FCM segmentation method is an unsupervised method that has been successfully applied in various fields such as medical imaging, target recognition, and data analysis [42]. This method does not require any class tags since its primary purpose is to group images based on their specific features. The objective of the algorithm is to minimize the J objective function in equation number 2, where X=($x_1$, $x_2$, …, $x_N$) is an image with N pixels that we want to divide into c sets.

$$J = \sum_{j=1}^{N}\sum_{i=1}^{c} \mathrm{u}_{\mathrm{ij}}^{\mathrm{m}} \left\| x_j - v_i \right\|^2 \quad (2)$$

Here $u_{ij}$, denotes the membership value of the $j_{th}$ pixel belonging to the $i_{th}$ cluster, $v_i$ denotes the center point of the $i_{th}$ cluster, and m denotes the blur constantly.

In our study, the blur constant m was chosen as the default value of 2. As can be seen from Formula 2, the FCM method allows a data point to be included in more than one cluster. The decisive factor here is the membership value. The higher the membership value a data point is connected to a cluster, the more similar it is to that cluster. The total membership value of pixels in different clusters is always 1. The membership values and cluster centers are calculated according to formulas 3 and 4;

$$u_{ij} = \frac{1}{\sum_{\mathrm{k=1}}^{\mathrm{c}} \left( \frac{\left\| x_j - v_i \right\|}{\left\| x_j - v_k \right\|} \right)^{2/(m-1)}} \quad (3)$$

$$v_i = \frac{\sum_{j=1}^{N} u_{ij}^m x_j}{\sum_{j=1}^{N} u_{ij}^m} \quad (4)$$

The algorithm starts the first iteration with randomly selected cluster centers, and for the termination process, it decides according to the number of changes in the membership function or cluster centers in two consecutive iterations. Or, it can be terminated after a certain number of iterations.

In our study, we also employ the k-means (KM) clustering algorithm, which was developed by J.B. MacQueen in 1967, making it one of the oldest clustering algorithms [43]. It is a widely used non-parametric learning method. Unlike fuzzy c-means, the KM algorithm allows each data point to belong to only one cluster [44]. The steps of the k-means algorithm are as follows:

- First, the number of cluster centers is selected, and an equal number of samples are randomly chosen to be the centers.
- The remaining samples are assigned to the cluster they are closest to based on their distances from the cluster centers.
- The center points of the clusters are then recalculated based on the new sample assignments, and the distances of the samples to the centers are re-examined.
- The algorithm continues to run until there is no change in the cluster assignments.

The sum squared error (SSE) is used to evaluate the k-means clustering method. The clustering result with the lowest SSE value is considered to be the most successful result. SSE calculation is performed using the following formula [45];

$$SSE = \sum_{i=1}^{k} \sum_{x \in Ci} dist2(\mathrm{mi}, \mathrm{x}) \quad (5)$$

x: An object in $C_i$ cluster, mi: Center point of $C_i$ cluster, k: Number of clusters. Here, dist is the standard Euclidean Distance between two objects, an object whose x value is in the $C_i$ set, and the mi value is the center point of the set $C_i$. The number

of clusters was determined to be 2 in both KM and FCM methods because of the black background and lungs in the anterior part in CT and X-ray images.

### 3.4 Feature extraction

Feature extraction is a process of identifying distinctive parts of an image from a large number of pixels [46]. In our study, we utilized three different feature extraction methods: Gray Level Co-occurrence Matrix (GLCM), Histogram of Oriented Gradients (HOG), and Local Binary Pattern (LBP).

In the first method used in our study, seven textural features based on GLCM were calculated for four different directions (0º, 45º, 90º, and 135º) from each image. Seven Haralick texture descriptors, including Angular Second Moment, Contrast, Inverse Difference Moment/Homogeneity, Dissimilarity, Entropy, Maximum Probability, and Inverse, were extracted from each co-occurrence matrix calculated at each of the four angles [47].

Histogram of Oriented Gradients (HOG) is a feature extraction method that has shown effective results in object and pattern recognition. The primary objective of the HOG method is to display the image as local histograms [48]. The process of extracting HOG features from an image involves applying horizontal and vertical Sobel filters. This is followed by calculating the gradient size and orientation angle during the feature extraction process.

The extraction of features using a Local Binary Pattern (LBP) is a non-parametric approach [49], [50]. LBP has important advantages, such as high tolerance to images with different lighting conditions, sensitivity to small changes in gray-level images, and low calculation cost, making it frequently preferred, especially in image processing studies. The process of calculating the LBP identifier obtained from a 3x3 frame according to the 8-neighborhood is computed using Equation 6.

$$\mathrm{LBP}_{P,R}(x_c) = \sum_{p=0}^{P-1} u(x_p - x_c)2^p, \qquad u(y) = \begin{cases} 0, & \text{if } y < 0 \\ 1, & \text{if } y \geq 0 \end{cases} \tag{6}$$

The notation used in our study defines $x_c$ as the central pixel, $x_p$ as the neighboring pixels surrounding the central pixel, R as the distance between the central pixel and its neighboring pixels, and P as the number of neighbors considered.

### 3.5 Classification

In our study, we employed three different classifiers to classify the features obtained from segmented images: Extreme Learning Machines (ELM), Support Vector Machines (SVM), and the k-nearest neighbor (kNN) method, which is a simple yet effective approach.

ELM is a type of single hidden layer feed-forward neural network (SLFN) [51], [52]. In this method, input weights are randomly determined, while output weights are calculated analytically [53]. In our study, we experimented with many hidden layers set to 4096 and a regularization parameter of 1e-1 for the ELM method. SVM, which stands for Support Vector Machine, is a supervised classification approach created for binary classification tasks but has proven to be effective in handling multi-class problems as well. The primary goal of SVM is to identify the hyperplane that can best differentiate the two classes, as noted in [54]. For multiclass SVM, we followed the one-versus-all approach.

$(x_i, y_i)_{1 \leq i \leq N}$ shows the training examples. Each example shows the size of the $x_i \in R^d$, d feature space. $y_i$ shows class labels. SVM aims to find a hyperplane that separates the data from each other, where the samples with the same label will stay on the same side. For this, a line equation is defined as in 7.

$$y_i(w.x_i + b) > 0, \qquad \text{i=1, … , N} \tag{7}$$

If a hyperplane exists, as shown in Formula 7, it indicates that the data can be separated linearly. However, if the data is distributed nonlinearly, it can still be separated linearly by transforming the feature space into a high-dimensional feature spaceThe Support Vector Machines (SVM) is a type of classifier that is based on kernels. The literature on SVM commonly uses radial, polynomial, and linear kernels as kernel functions [55]. In our study, we employed the polynomial kernel as it performed better than the other kernels for our datasets.

On the other hand, k-Nearest Neighbor (kNN) is a non-parametric classification method that has been widely applied in various classification problems [56]. The classification process is performed by examining the nearest neighbors of the samples and assigning them to the cluster they are most similar. The kNN classifier generates feature vectors and labels of the training samples and applies the same features to the test sample whose class is unknown. Distances between data points are calculated, and the k closest samples are selected. Among these samples, the majority class is assigned to the new instance. In our study, we set the k value to 1 and used Euclidean distance for distance measurement. Despite its simplicity, kNN is a popular and effective method as it produces successful results and is simple and easy to apply.

### 3.6 Modeling

There are various methods for modeling data, with the most popular ones involving partitioning data into separate sets for training and testing. However, in cases where there are imbalanced class distributions, this partitioning approach may not be suitable. Instead, cross-validation (CV) is a widely used method where the entire dataset can be utilized for both training and testing. In our classification studies, we adopt the 10-fold cross-validation method, which involves dividing the dataset into 10 equal parts. During each iteration, one part is used for testing while the other nine parts are used for training. This process is repeated ten times, with the testing part being changed at each iteration. Finally, the average of the values obtained from

each iteration is calculated. This ensures that all data is utilized for both training and testing. A visual representation of the tenfold cross-validation method is illustrated in Figure 3.

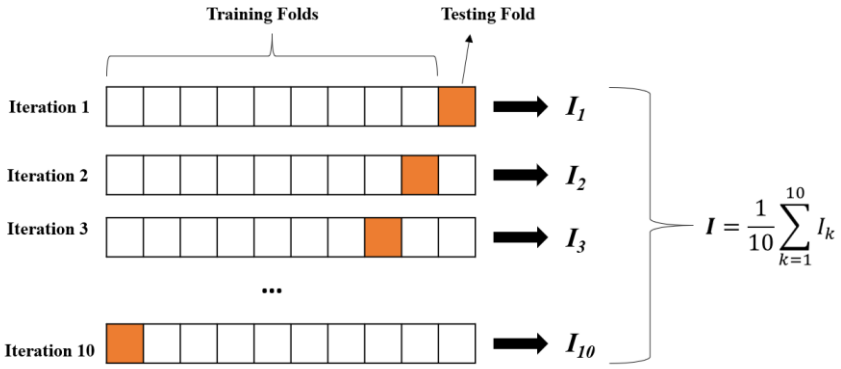



Figure 3. 10-fold cross-validation schema

### 3.7 Performance measurement

The metrics given below will be used to measure the success of the classification processes in our study.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} * 100 \quad (8)$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP+FN})} * 100 \quad (9)$$

$$\text{Specificity} = \frac{\text{TN}}{(\text{TN} + \text{FP})} * 100 \quad (10)$$

$$\text{Precision} = \frac{\text{TP}}{(\text{TP+FP})} * 100 \quad (11)$$

$$\text{NPV} = \frac{\text{TN}}{(\text{TN} + \text{FN})} * 100 \quad (12)$$

TP shows correctly predicted positive samples, TN correctly predicted negative samples, FP incorrectly predicted positive samples, and FN incorrectly predicted negative samples. The choice of which metric to focus on depends on the particular aspect of the test under consideration. For instance, if identifying positive samples is more important, then the recall metric would be a suitable choice. This was the case in our study. On the other hand, the accuracy metric would be more appropriate if we want to assess the correct identification of both positive and negative samples. However, in cases where the class distribution is unbalanced, accuracy may not be a suitable evaluation metric [57].
The flow chart showing the methods and implementation process described in the third section is shown in Figure 4.

## 4. Experimental results

We conducted experimental studies using four different datasets with varying sizes and characteristics. The experiments were performed on a desktop computer equipped with an i5 processor and 4 GB graphic cards. To differentiate between healthy and patient samples, we employed five performance metrics to assess the outcomes. We will report all five metrics in our results; however, Recall and Accuracy metrics are of primary importance to our study.
As depicted in Figure 4, our study included stages of data acquisition, preprocessing, segmentation, feature extraction, classification, and modeling, as well as performance evaluation. The feature extraction process was conducted on segmented images, and the segmented images obtained by employing KM and FCM methods are presented in Figures 5 and 6. The FCM method divided CT images into two clusters, as shown in Figure 5. The segmentation results indicate significant differences between diseased and healthy images. As previously mentioned, COVID-19 anomalies in the lungs become more apparent after segmentation, which is evident in both Figure 5 and Figure 6.

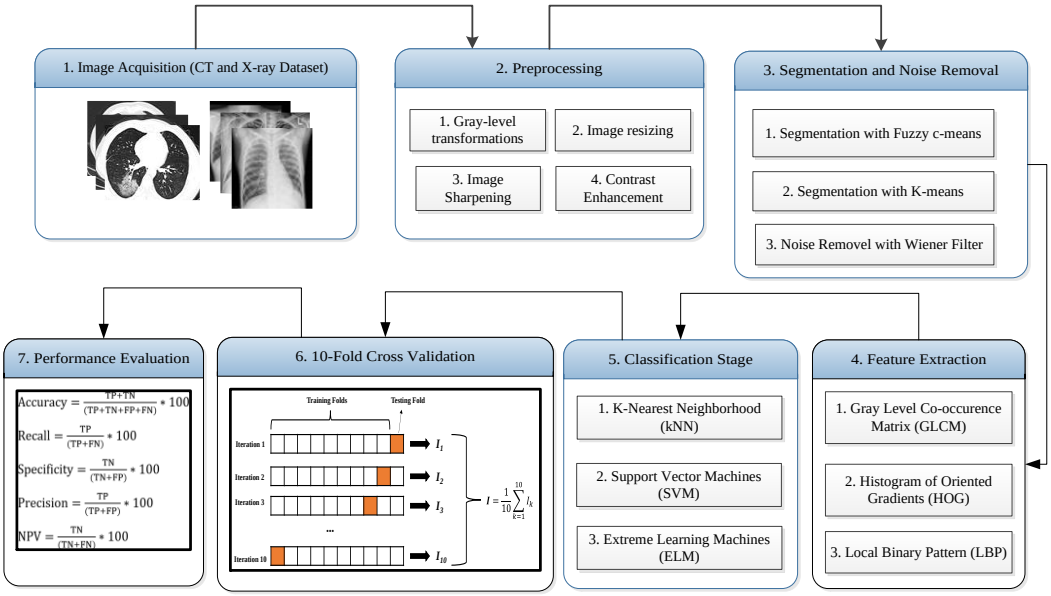



Figure 4. A diagram illustrating the sequence of actions performed by the system.

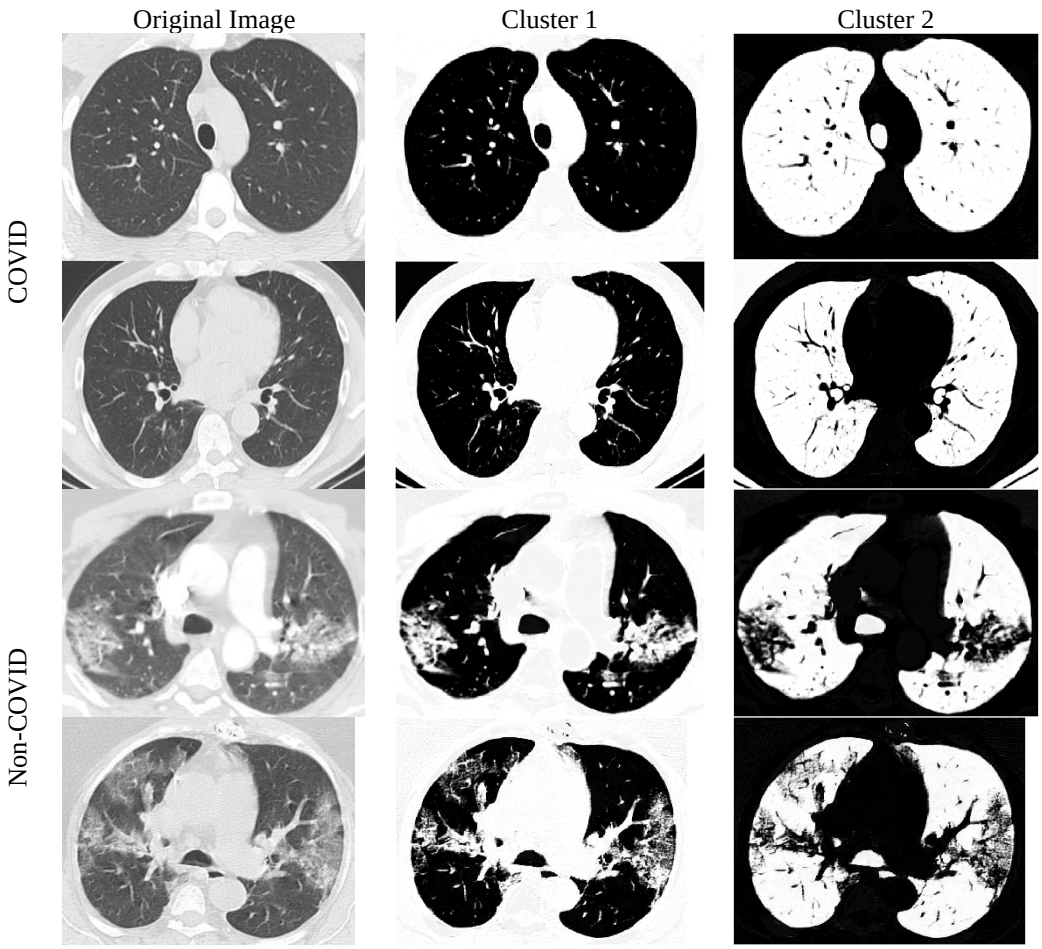



Figure 5. Segmented Sample CT images with FCM method

Figure 6 shows a few examples of diseased and healthy images for obtaining the KM segmentation method on X-ray images. Here, too, the dashed structures on the new images obtained as a result of the segmentation of the images in the first two lines, which are COVID-19, stand out. However, the result sets obtained by segmentation of healthy images seen in the last two lines are seen smoothly and clearly.

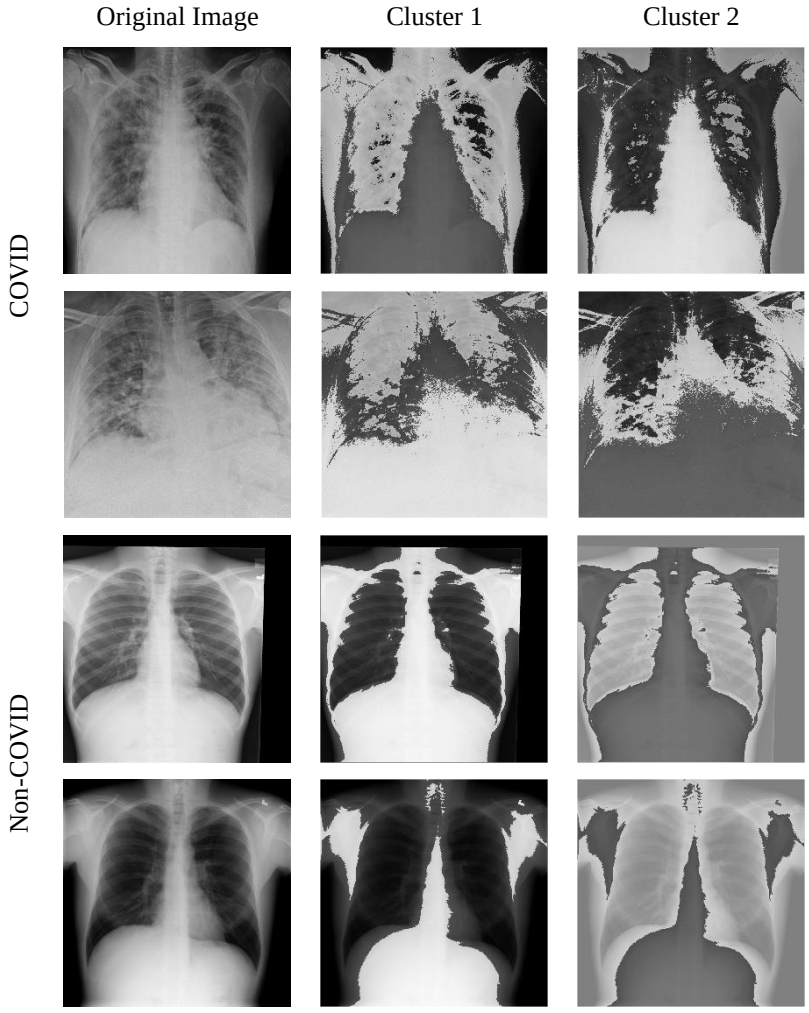



Figure 6. Sample segmented X-ray images with the k-means method

The performance measurement results obtained for four different datasets in the study are presented in Tables 2, 3, 4, and 5. All the detailed results obtained by combining three different classifiers, three different feature extraction methods, and two different segmentation methods are shown in all four tables. Table 2 displays the binary classification results for CT datasets. Upon inspection, it is observed that Dataset 1 yields a Recall value of 90.8% and an Accuracy value of 90.2%. For Dataset 2, these values increase to 99.4% and 98.8%, respectively. The images in Dataset 1 are low resolution and noisy, while Dataset 2 comprises images with higher resolution and less noise. It is believed that this difference accounts for the performance discrepancy. Additionally, Table 2 indicates that the KM and FCM segmentation methods do not offer a significant advantage over one another. Both methods exhibit superior results in different scenarios. FCM segmentation produces better results for Dataset 1, whereas the KM method performs better for Dataset 2. Upon examining the feature extraction methods, it is observed that the GLCM method and its classification results are subpar compared to the HOG and LBP methods, which yield better results across all models. Comparing the HOG and LBP methods, Table 2 demonstrates that the LBP method delivers relatively superior results. Although the HOG method produces comparable results, the LBP method stands out as a feature extraction method in the model that yields the highest values in both CT datasets. Looking at the classification methods in Table 2, it is evident that the kNN method is the most effective, consistent with many prior studies. Additionally, ELM yields relatively better results than SVM.

Table 3 presents the three-class classification results for X-ray data sets, including COVID-19 positive, healthy, and non-COVID-19-derived pneumonia classes. As shown in Table 3, the best results are achieved with the ELM-SVM methods, which differ from Table 2, which presents the results for CT data sets. For Dataset 3, the highest values are obtained with the SVM, LBP, and KM combination, with a Recall value of 84.0% and an accuracy value of 83.9%. However, for Dataset 4, the triplet of ELM, LBP, and KM produced higher results, with a Recall value of 98.7% and an accuracy value of 99.6%. The results for Dataset 4 are considerably better than those for Dataset 3 because the images in Dataset 4 are cleaner, less noisy, and higher in resolution. In particular, Dataset 3 contains many images with various artifacts, which negatively affects the success. Table 3 also shows that the KM segmentation method performs better than the FCM method in X-ray images. Additionally, the LBP method is more successful than the HOG method in both X-ray and CT images, while the GLCM method yields the least successful results as a feature extraction method.

As previously mentioned, our X-ray data sets include three different classes. In contradistinction to the CT image sets, the X-ray datasets comprise images of non-COVID-19 pneumonia, enabling the differentiation between images of COVID-19 and those of healthy individuals, as well as discerning between images of COVID-19 and pneumonia resulting from other causes. The results for these scenarios are presented in Tables 4 and 5. Table 4 displays the classification results of X-ray images according to COVID-19 and pneumonia classes, whereas Table 5 presents the results for COVID-19 and healthy classes. As known, COVID-19 disease causes pneumonia, making it a time-consuming process for an expert to differentiate between pneumonia and not COVID-19, which can lead to errors. As shown in Table 4, while Dataset 3's Recall value was 92.8%, the accuracy value was 97.7%. For Dataset 4, these values were 99.1% and 99.3%, respectively. This finding suggests that the system we have created can accurately differentiate between pneumonia caused by COVID-19 and pneumonia caused by other factors, which is a noteworthy and valuable outcome. When we examine Table 5, which is related to the detection of COVID-19 and healthy individuals from X-ray images, an accuracy rate of 98.6% was achieved for Dataset 3, and an accuracy rate of 99.5% was achieved for Dataset 4. These findings demonstrate that the distinction between healthy and sick individuals can be made with even higher accuracy.
Table 6 presents the average execution time of the classification method utilized in this research. The ELM method, which includes a large number of hidden layers, requires a considerably longer processing time. On the other hand, the kNN classifier provides the shortest processing times. Moreover, the SVM method produces results similar to those of the kNN method. As can be seen from Table 6, the size of the data set also affects the classification process.

Table 7 demonstrates the highest accuracy rates obtained for the four distinct data sets in this study through the confusion matrix. The confusion matrices c and d indicate erroneous classifications between COVID-19 positive and negative, as well as COVID-19 negative and non-COVID-19 ones. Another noteworthy finding is that COVID-19-positive cases and pneumonia can be more successfully distinguished.

Table 7. Confusion Matrices with Performance Metrics for Binary and Multiclass Classification in Four Different Datasets. Yellow values show Accuracy, Bottom-Left blue value Precision, Bottom-right blue value NPV, Right-upper blue recall, and Right-bottom blue specificity values.

**Dataset 1**

| Actual Class | Predicted: COVID (+) | Predicted: COVID (-) | |
|---|---|---|---|
| COVID (+) | 317 | 41 | 88.5 |
| COVID (-) | 32 | 356 | 91.8 |
| | 90.8 | 89.7 | 90.2 |

**Predicted Class**

**Dataset 2**

| Actual Class | Predicted: COVID (+) | Predicted: COVID (-) | |
|---|---|---|---|
| COVID (+) | 1207 | 7 | 99.4 |
| COVID (-) | 22 | 1245 | 98.2 |
| | 98.3 | 99.4 | 98.8 |

**Predicted Class**

**Dataset 3**

| Actual Class | Predicted: COVID (+) | Predicted: Non-COVID | Predicted: Pneumonia | |
|---|---|---|---|---|
| COVID (+) | 105 | 2 | 4 | 94.6 |
| No Findings | 4 | 425 | 82 | 83.2 |
| Pneumonia | 16 | 73 | 414 | 82.3 |
| | 84.0 | 85.0 | 82.8 | 83.9 |

**Predicted Class**

**Dataset 4**

| Actual Class | Predicted: COVID (+) | Predicted: Non-COVID | Predicted: Pneumonia | |
|---|---|---|---|---|
| COVID (+) | 1184 | 5 | 7 | 99.0 |
| No Findings | 9 | 1295 | 79 | 93.8 |
| Pneumonia | 7 | 41 | 1259 | 96.3 |
| | 98.7 | 96.6 | 93.6 | 96.2 |

**Predicted Class**

Table 2**.** Mixed performance measurement results for Dataset 1 and 2 (CT datasets) were obtained with three different classifiers, three different feature extraction, and two different segmentation methods (The highest performance values are shown in bold)

| **Classification Methods** | Feature Extraction Methods | Segmentation Methods | **Dataset 1** | | | | | **Dataset 2** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Accuracy | Recall | Specificity | Precision | NPV | Accuracy | Recall | Specificity | Precision | NPV |
| **ELM** | GLCM | KM | 71.3 | 69.3 | 73.0 | 69.3 | 73.0 | 86.9 | 86.9 | 86.9 | 87.1 | 86.7 |
| | | FCM | 60.3 | 54.7 | 65.2 | 58.1 | 62.1 | 82.0 | 80.4 | 83.6 | 83.3 | 80.7 |
| | HOG | KM | 79.5 | 77.9 | 80.9 | 78.2 | 80.7 | 93.6 | 92.5 | 94.6 | 94.6 | 92.5 |
| | | FCM | 82.0 | 81.1 | 82.9 | 80.6 | 83.3 | 94.6 | 94.4 | 94.8 | 94.8 | 94.3 |
| | LBP | KM | 84.0 | 83.4 | 84.6 | 82.7 | 85.3 | 94.5 | 93.8 | 95.3 | 95.3 | 93.8 |
| | | FCM | 84.2 | 82.5 | 85.6 | 83.5 | 84.8 | 93.3 | 92.8 | 93.8 | 93.9 | 92.8 |
| **kNN** | GLCM | KM | 76.4 | 77.1 | 75.8 | 73.7 | 79.0 | 94.5 | 96.0 | 93.0 | 93.3 | 95.8 |
| | | FCM | 67.0 | 65.9 | 68.0 | 64.4 | 69.4 | 88.7 | 90.2 | 87.2 | 87.8 | 89.7 |
| | HOG | KM | 79.2 | 75.9 | 82.1 | 78.9 | 79.5 | 97.3 | 98.7 | 95.9 | 96.1 | 98.7 |
| | | FCM | 79.8 | 77.1 | 82.1 | 79.1 | 80.3 | 97.1 | 98.6 | 95.5 | 95.7 | 98.5 |
| | LBP | KM | 86.6 | 88.0 | 85.4 | 84.1 | 89.0 | **98.8** | **99.4** | **98.2** | **98.3** | **99.4** |
| | | FCM | **90.2** | **90.8** | **89.7** | **88.5** | **91.8** | 97.9 | 99.0 | 96.9 | 97.0 | 98.9 |
| **SVM** | GLCM | KM | 56.0 | 65.3 | 47.9 | 52.4 | 61.1 | 63.3 | 61.6 | 65.0 | 64.2 | 62.4 |
| | | FCM | 49.5 | 63.3 | 37.2 | 47.0 | 53.6 | 62.2 | 65.6 | 58.8 | 61.9 | 62.7 |
| | HOG | KM | 74.9 | 74.5 | 75.3 | 72.6 | 77.1 | 89.8 | 86.5 | 93.2 | 92.9 | 87.1 |
| | | FCM | 76.7 | 70.8 | 81.9 | 77.4 | 76.1 | 89.0 | 87.1 | 91.0 | 90.8 | 87.4 |
| | LBP | KM | 75.7 | 73.1 | 78.1 | 74.6 | 76.7 | 91.8 | 89.9 | 93.7 | 93.5 | 90.1 |
| | | FCM | 77.9 | 75.4 | 80.1 | 76.9 | 78.7 | 88.7 | 87.9 | 89.5 | 89.5 | 87.9 |

Table 3**.** Multi-class mixed performance measurement results for Datasets 3 and 4 (X-ray datasets) were obtained with three different classifiers, three different feature extraction, and two different segmentation methods (The highest performance values are shown in bold)

| **Classification Methods** | Feature Extraction Methods | Segmentation Methods | **Dataset 3 (Three Class)** | | | | | **Dataset 4 (Three Class)** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Accuracy | Recall | Specificity | Precision | NPV | Accuracy | Recall | Specificity | Precision | NPV |
| **ELM** | GLCM | KM | 72.2 | 54.4 | 97.6 | 73.9 | 94.5 | 74.7 | 78.3 | 91.8 | 81.0 | 90.4 |
| | | FCM | 65.2 | 33.6 | 96.1 | 51.9 | 92.0 | 70.5 | 72.5 | 89.4 | 75.3 | 87.9 |
| | HOG | KM | 80.2 | 76.0 | 99.2 | 92.2 | 97.1 | 96.0 | 98.6 | 99.3 | 98.5 | 99.4 |
| | | FCM | 77.2 | 82.4 | 98.7 | 88.8 | 97.8 | 96.3 | 98.3 | 99.6 | 99.1 | 99.3 |
| | LBP | KM | 83.1 | 85.6 | 99.3 | 93.9 | 98.2 | **98.7** | **99.6** | **99.6** | **99.0** | **99.4** |
| | | FCM | 82.0 | 84.8 | 98.9 | 90.6 | 98.1 | 96.8 | 98.8 | 99.7 | 99.4 | 99.5 |
| **kNN** | GLCM | KM | 71.6 | 48.8 | 97.3 | 69.3 | 93.8 | 80.6 | 86.3 | 95.4 | 89.4 | 94.0 |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | FCM | 67.7 | 60.8 | 95.2 | 61.3 | 95.1 | 79.9 | 86.0 | 97.1 | 92.9 | 93.9 |
| | HOG | KM | 73.9 | 72.8 | 96.7 | 73.4 | 96.6 | 92.1 | 96.0 | 99.3 | 98.4 | 98.2 |
| | | FCM | 72.5 | 77.6 | 96.1 | 71.3 | 97.2 | 90.7 | 94.0 | 99.7 | 99.3 | 97.4 |
| | LBP | KM | 78.8 | 80.8 | 99.2 | 92.7 | 97.6 | 94.4 | 97.3 | 99.7 | 99.2 | 98.8 |
| | | FCM | 77.2 | 83.2 | 99.1 | 92.0 | 97.9 | 94.1 | 97.9 | 99.8 | 99.5 | 99.1 |
| **SVM** | GLCM | KM | 43.1 | 33.6 | 86.9 | 24.3 | 91.3 | 80.6 | 86.3 | 95.4 | 89.4 | 94.0 |
| | | FCM | 43.8 | 48.8 | 91.5 | 41.8 | 93.5 | 79.9 | 86.0 | 97.1 | 92.9 | 93.9 |
| | HOG | KM | 81.2 | 71.2 | 99.1 | 90.8 | 96.5 | 92.1 | 96.0 | 99.3 | 98.4 | 98.2 |
| | | FCM | 79.4 | 81.6 | 99.5 | 95.3 | 97.7 | 90.7 | 94.0 | 99.7 | 99.3 | 97.4 |
| | LBP | KM | **83.9** | **84.0** | **99.4** | **94.6** | **98.0** | 96.1 | 98.3 | 99.4 | 98.6 | 99.3 |
| | | FCM | 83.2 | 83.2 | 99.7 | 97.2 | 97.9 | 96.1 | 98.0 | 99.2 | 98.2 | 99.1 |

Table 4. Binary classification results in which the distinction between COVID-19 and pneumonia is addressed in X-ray data sets (The highest performance values are shown in bold)

| | | | **Dataset 3 (Binary Class)** | | | | | **Dataset 4 (Binary Class)** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Classification Methods** | Feature Extraction Methods | Segmentation Methods | Accuracy | Recall | Specificity | Precision | NPV | Accuracy | Recall | Specificity | Precision | NPV |
| **ELM** | GLCM | KM | 88.6 | 61.6 | 95.4 | 77.0 | 90.9 | 84.4 | 84.1 | 84.8 | 83.1 | 85.6 |
| | | FCM | 82.7 | 36.8 | 94.2 | 61.3 | 85.6 | 78.7 | 75.7 | 81.4 | 78.4 | 78.9 |
| | HOG | KM | 95.4 | 84.0 | 98.2 | 92.1 | 96.1 | 98.9 | 98.8 | 99.1 | 99.0 | 98.9 |
| | | FCM | 97.3 | 92.0 | 98.6 | 94.3 | 98.0 | 99.3 | 98.9 | 99.6 | 99.5 | 99.0 |
| | LBP | KM | 96.8 | 90.4 | 98.4 | 93.4 | 97.6 | **99.3** | **99.1** | 99.4 | 99.3 | **99.2** |
| | | FCM | **97.7** | **92.8** | **99.0** | **95.9** | **98.2** | 99.3 | 98.9 | **99.6** | **99.6** | 99.0 |
| **kNN** | GLCM | KM | 87.8 | 93.2 | 94.0 | 72.5 | 91.1 | 88.4 | 87.2 | 89.5 | 88.1 | 88.7 |
| | | FCM | 86.4 | 64.8 | 91.8 | 66.4 | 91.3 | 91.7 | 89.1 | 94.1 | 93.0 | 90.6 |
| | HOG | KM | 90.2 | 77.6 | 93.4 | 74.6 | 94.3 | 97.8 | 96.7 | 98.9 | 98.7 | 97.1 |
| | | FCM | 90.7 | 81.6 | 93.0 | 74.5 | 95.3 | 98.2 | 96.3 | 99.8 | 99.7 | 96.8 |
| | LBP | KM | 95.4 | 84.8 | 98.0 | 91.4 | 96.3 | 98.8 | 97.9 | 99.6 | 99.5 | 98.2 |
| | | FCM | 94.2 | 82.4 | 97.2 | 88.0 | 95.7 | 98.9 | 98.3 | 99.6 | 99.5 | 98.5 |
| **SVM** | GLCM | KM | 51.8 | 54.4 | 51.2 | 21.8 | 81.8 | 49.9 | 59.8 | 41.1 | 47.5 | 53.4 |
| | | FCM | 70.7 | 62.4 | 72.8 | 36.4 | 88.6 | 57.2 | 65.6 | 49.8 | 53.8 | 61.9 |
| | HOG | KM | 93.3 | 76.0 | 97.6 | 88.8 | 94.2 | 98.9 | 98.5 | 99.2 | 99.1 | 98.7 |
| | | FCM | 95.8 | 85.6 | 98.4 | 93.0 | 96.5 | 99.1 | 98.8 | 99.3 | 99.2 | 98.9 |
| | LBP | KM | 96.6 | 88.0 | 98.8 | 94.8 | 97.1 | 98.7 | 98.5 | 99.0 | 98.8 | 98.7 |
| | | FCM | 96.8 | 89.6 | 98.6 | 94.1 | 97.4 | 98.9 | 98.5 | 99.2 | 99.1 | 98.7 |

Table 5. Binary classification results in which the distinction between COVID-19 and healthy is discussed in the X-ray data sets (The highest performance values are shown in bold)

| | | | **Dataset 3 (Binary Class)** | | | | | **Dataset 4 (Binary Class)** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Classification Methods** | Feature Extraction Methods | Segmentation Methods | Accuracy | Recall | Specificity | Precision | NPV | Accuracy | Recall | Specificity | Precision | NPV |
| **ELM** | GLCM | KM | 92.0 | 72.0 | 97.0 | 85.7 | 93.3 | 93.0 | 91.9 | 94.0 | 93.2 | 92.9 |
| | | FCM | 91.5 | 68.0 | 97.4 | 86.7 | 92.4 | 87.4 | 85.3 | 89.3 | 87.7 | 87.2 |
| | HOG | KM | 97.3 | 89.6 | 99.2 | 96.6 | 97.4 | 99.4 | 99.2 | 99.6 | 99.6 | 99.3 |
| | | FCM | 98.4 | 95.2 | 99.2 | 96.7 | 98.8 | 99.3 | 99.0 | 99.5 | 99.4 | 99.1 |
| | LBP | KM | 98.2 | 92.0 | 99.8 | 99.1 | 98.0 | **99.5** | **99.4** | 99.6 | 99.5 | **99.5** |
| | | FCM | **98.6** | **95.2** | **99.4** | **97.5** | **98.8** | 99.3 | 98.8 | 99.7 | 99.7 | 99.0 |
| **kNN** | GLCM | KM | 90.4 | 64.8 | 96.8 | 83.5 | 91.7 | 95.9 | 93.8 | 97.8 | 97.5 | 94.6 |
| | | FCM | 91.2 | 84.8 | 92.8 | 74.6 | 96.1 | 95.6 | 92.6 | 98.2 | 97.9 | 93.7 |
| | HOG | KM | 95.7 | 88.8 | 97.4 | 89.5 | 97.2 | 98.3 | 97.1 | 99.4 | 99.3 | 97.4 |
| | | FCM | 95.2 | 90.4 | 96.4 | 86.3 | 97.6 | 97.8 | 95.6 | 99.7 | 99.7 | 96.2 |
| | LBP | KM | 97.9 | 92.8 | 99.2 | 96.7 | 98.2 | 99.2 | 98.5 | 99.8 | 99.7 | 98.7 |
| | | FCM | 97.0 | 90.4 | 98.6 | 94.2 | 97.6 | 99.4 | 98.8 | **99.9** | **99.8** | 99.0 |
| **SVM** | GLCM | KM | 57.9 | 52.0 | 59.4 | 24.3 | 83.2 | 55.8 | 68.6 | 44.4 | 52.5 | 61.3 |
| | | FCM | 91.2 | 83.2 | 93.2 | 75.4 | 95.7 | 63.7 | 67.3 | 60.5 | 60.4 | 67.4 |
| | HOG | KM | 97.7 | 91.2 | 99.4 | 97.4 | 97.8 | 99.2 | 98.8 | 99.5 | 99.4 | 99.0 |
| | | FCM | 98.1 | 95.2 | 98.8 | 95.2 | 98.8 | 99.2 | 98.9 | 99.4 | 99.3 | 99.0 |
| | LBP | KM | 98.1 | 92.8 | 99.4 | 97.5 | 98.2 | 99.4 | 99.1 | 99.7 | 99.7 | 99.2 |
| | | FCM | 97.8 | 90.4 | 99.6 | 98.3 | 97.6 | 99.0 | 98.4 | 99.6 | 99.5 | 98.6 |

Table 6. Classification average execution time of the method

| | | **Dataset-1** | **Dataset-2** | **Dataset-3** | **Dataset-4** |
|---|---|---|---|---|---|
| **Time (s)** | ELM | 33.9 | 59.3 | 42.2 | 91.2 |
| | kNN | 29.9 | 3.1 | 0.4 | 2.5 |
| | SVM | 45.6 | 5.3 | 1.0 | 4.7 |

## 4. Discussion

Our study aims to propose an approach that can achieve successful results with various imaging modes and datasets and support early diagnosis of COVID-19. In the modeling, a system has been developed to be applied to datasets that are independent of the dataset, and the results demonstrate success on all four different datasets. Table 8 compares our study with other studies in the literature, focusing on the use of one of the four datasets used in our study. However, it should be noted that comparing different studies using the same dataset may not be entirely accurate since they may use different subsets of the same dataset. Despite this limitation, Table 8 provides a general comparison of our study with other studies. Upon reviewing Table 8, it can be observed that our study achieved a higher success rate than the state-of-the-art studies in nearly all of the four distinct datasets. For instance, in the study of Ozturk et al., while an accuracy of 87.02% was obtained in the multiclass classification of X-ray images, this rate was 83.9% in our study. However, our study produced far more successful results in terms of recall values in the same comparison. Therefore, our study can be considered very successful in detecting COVID-19-positive samples. Similarly, while Chowdhury et al. achieved an accuracy rate of 99.41% for multiclass classification in their study, our study obtained an accuracy rate of 99.3%. Nonetheless, our study achieved a recall value of 99.6%, whereas Chowdhury et al. achieved a recall value of 96.61%.

**Table 8.** Comparison of the performances of the state of the art studies

| Author(s) | Dataset | Accuracy (%) | Recall (%) | Precision (%) |
|---|---|---|---|---|
| Yang et al. [34] | Dataset-1 | 89.0 | NA | NA |
| Soares et al. [35] | Dataset-2 | 97.38 | 95.53 | 99.16 |
| Jaiswal et al. [31] | Dataset-2 | 96.25 | 96.29 | 96.29 |
| Ozturk et al. [18] | Dataset-3 (Binary) | 98.08 | 95.13 | 98.03 |
| Ozturk et al. [18] | Dataset-3 (Multiclass) | 87.02 | 85.35 | 89.96 |
| Rahman et al. [39] | Dataset-4 (Multiclass) | 96.29 | 97.28 | NA |
| Chowdhury et al. [38] | Dataset-4 (Binary) | 99.41 | 99.41 | 99.42 |
| Chowdhury et al. [38] | Dataset-4 (Multiclass) | 97.74 | 96.61 | 96.61 |
| This Study | Dataset-1 | 90.2 | 90.8 | 88.5 |
| This Study | Dataset-2 | 98.8 | 99.4 | 98.3 |
| This Study | Dataset-3 (Binary) | 97.7 | 92.8 | 95.9 |
| This Study | Dataset-3 (Multiclass) | 83.9 | 94.0 | 94.6 |
| This Study | Dataset-4 (Binary) | 99.3 | 99.1 | 99.6 |
| This Study | Dataset-4 (Multiclass) | 98.7 | 99.6 | 99.0 |

In general, we can say that a more effective system has been produced compared to other studies. It is noteworthy that the developed system operates on four distinct datasets acquired through various imaging modes, delivering superior performance when compared to prior studies in the field. Furthermore, the classification times presented in Table 6 indicate that the evaluation duration of the system is less than 1 minute. This is seen as another positive aspect that needs attention.

## 5. Conclusion

Our research yielded a comprehensive system that can diagnose COVID-19 from datasets captured through various imaging modes. This system was created by adopting a holistic approach that integrates image processing, segmentation, feature extraction, and machine learning methods. Our model can be applied to any dataset and has demonstrated its efficacy by achieving successful results. We developed a holistic approach that is not limited to any specific dataset. Our study not only distinguished between COVID-positive and negative cases but also conducted a multi-class classification process, where we identified COVID-19-positive, healthy, and pneumonia classes. Our research achieved a remarkable accuracy rate of 97.7% and 99.3% in two distinct datasets in identifying individuals with COVID-19 and those with pneumonia but without COVID-19. Considering that non-COVID-19 pneumonia and COVID-19 pneumonia exhibit comparable characteristics, making it challenging for healthcare experts to differentiate between them, the achievements of this research are noteworthy.
For more than two years, the entire world has been mobilized to detect and treat COVID-19. Although the contagiousness of the virus has decreased, pneumonia still affects people due to various factors. Therefore, the results obtained from our study are considered important, and they could provide support for medical professionals in their decision-making. A limitation of our study is the lack of information on the period of the disease in which the data were collected. Future studies could create datasets containing this information to provide a more in-depth analysis of the effects of the disease.

**References**


[1] "World Health Organization Coronavirus (COVID-19) Dashboard, Retrieved 05th May 2021. (Web Link: https://covid19.who.int/)." https://covid19.who.int/

[2] C. Wang *et al.*, "Immediate psychological responses and associated factors during the initial stage of the 2019 coronavirus disease (COVID-19) epidemic among the general population in China," *Int J Environ Res Public Health*, vol. 17, no. 5, p. 1729, 2020.

[3] C. M. A. de O. Lima, "Information about the new coronavirus disease (COVID-19)," *Radiol Bras*, vol. 53, no. 2, pp. V–VI, 2020.

[4] M. W. M. Mustafa, "Audiological profile of asymptomatic Covid-19 PCR-positive cases," *Am J Otolaryngol*, vol. 41, no. 3, p. 102483, 2020.

[5] Y. Fang *et al.*, "Sensitivity of chest CT for COVID-19: comparison to RT-PCR," *Radiology*, vol. 296, no. 2, pp. E115–E117, 2020.

[6] L. Lan *et al.*, "Positive RT-PCR test results in patients recovered from COVID-19," *JAMA*, vol. 323, no. 15, pp. 1502–1503, 2020.

[7] T. Ai *et al.*, "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases," *Radiology*, vol. 296, no. 2, pp. E32–E40, 2020.

[8] H. Shi *et al.*, "Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study," *Lancet Infect Dis*, vol. 20, no. 4, pp. 425–434, 2020.

[9] D. Dong *et al.*, "The role of imaging in the detection and management of COVID-19: a review," *IEEE Rev Biomed Eng*, 2020.

[10] C. Hani *et al.*, "COVID-19 pneumonia: a review of typical CT findings and differential diagnosis," *Diagn Interv Imaging*, vol. 101, no. 5, pp. 263–268, 2020.

[11] A. SAYGILI, "Analysis and Segmentation of X-ray Images of COVID-19 Patients using the k-means Algorithm," *Veri Bilimi*, vol. 4, no. 3, pp. 1–6.

[12] A. Saygılı, "A new approach for computer-aided detection of coronavirus (COVID-19) from CT and X-ray images using machine learning methods," *Appl Soft Comput*, vol. 105, p. 107323, 2021.

[13] A. Saygılı, "Computer-aided detection of COVID-19 from CT images based on Gaussian mixture model and kernel support vector machines classifier," *Arab J Sci Eng*, vol. 47, no. 2, pp. 2435–2453, 2022.

[14] B. Abraham and M. S. Nair, "Computer-aided detection of COVID-19 from X-ray images using multi-CNN and Bayesnet classifier," *Biocybern Biomed Eng*, vol. 40, no. 4, pp. 1436–1445, 2020.

[15] R. C. Joshi *et al.*, "A deep learning-based COVID-19 automatic diagnostic framework using chest X-ray images," *Biocybern Biomed Eng*, vol. 41, no. 1, pp. 239–254, 2021.

[16] M. F. Aslan, M. F. Unlersen, K. Sabanci, and A. Durdu, "CNN-based transfer learning–BiLSTM network: A novel approach for COVID-19 infection detection," *Appl Soft Comput*, vol. 98, p. 106912, 2021, doi: https://doi.org/10.1016/j.asoc.2020.106912.

[17] F. Demir, "DeepCoroNet: A deep LSTM approach for automated detection of COVID-19 cases from chest X-ray images," *Appl Soft Comput*, vol. 103, p. 107160, 2021.

[18] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Comput Biol Med*, vol. 121, p. 103792, 2020.

[19] A. Altan and S. Karasu, "Recognition of COVID-19 disease from X-ray images by hybrid model consisting of 2D curvelet transform, chaotic salp swarm algorithm and deep learning technique," *Chaos Solitons Fractals*, vol. 140, p. 110071, 2020.

[20] M. Nour, Z. Cömert, and K. Polat, "A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization," *Appl Soft Comput*, vol. 97, p. 106580, 2020.

[21] A. Gupta, S. Gupta, and R. Katarya, "InstaCovNet-19: A deep learning classification model for the detection of COVID-19 patients using Chest X-ray," *Appl Soft Comput*, vol. 99, p. 106859, 2021.

[22] J. Sikder, N. Datta, and D. Tripura, "A Deep Learning Approach for Recognizing Covid-19 from Chest X-ray using Modified CNN-BiLSTM with M-SVM," in *2022 International Conference on Electrical, Computer and Energy Technologies (ICECET)*, IEEE, Jul. 2022, pp. 1–6. doi: 10.1109/ICECET55527.2022.9872776.

[23] L. Kong and J. Cheng, "Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion," *Biomed Signal Process Control*, vol. 77, p. 103772, Aug. 2022, doi: 10.1016/j.bspc.2022.103772.

[24] N. K. Mishra, P. Singh, and S. D. Joshi, "Automated Detection of COVID-19 from CT scan using Convolutional Neural Network," *Biocybern Biomed Eng*, 2021.
[25] S. Chakraborty and K. Mali, "SuFMoFPA: A superpixel and meta-heuristic based fuzzy image segmentation approach to explicate COVID-19 radiological images," *Expert Syst Appl*, vol. 167, p. 114142, 2021.
[26] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks," *Comput Biol Med*, vol. 121, p. 103795, 2020.
[27] G. Gilanie *et al.*, "Coronavirus (COVID-19) detection from chest radiology images using convolutional neural networks," *Biomed Signal Process Control*, vol. 66, p. 102490, 2021.
[28] P. Kalane, S. Patil, B. P. Patil, and D. P. Sharma, "Automatic detection of COVID-19 disease using U-Net architecture based fully convolutional network," *Biomed Signal Process Control*, vol. 67, p. 102518, 2021.
[29] Y. Li *et al.*, "Efficient and effective training of COVID-19 classification networks with self-supervised dual-track learning to rank," *IEEE J Biomed Health Inform*, vol. 24, no. 10, pp. 2787–2797, 2020.
[30] X. Xu *et al.*, "A deep learning system to screen novel coronavirus disease 2019 pneumonia," *Engineering*, vol. 6, no. 10, pp. 1122–1129, 2020.
[31] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *J Biomol Struct Dyn*, pp. 1–8, 2020.
[32] 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," *Advances in Engineering Software*, vol. 175, p. 103317, Jan. 2023, doi: 10.1016/j.advengsoft.2022.103317.
[33] V. Göreke, V. Sarı, and S. Kockanat, "A novel classifier architecture based on deep neural network for COVID-19 detection using laboratory findings," *Appl Soft Comput*, vol. 106, p. 107329, 2021.
[34] J. Zhao, Y. Zhang, X. He, and P. Xie, "Covid-ct-dataset: a ct scan dataset about covid-19," *arXiv preprint arXiv:2003.13865*, vol. 490, 2020.
[35] P. Angelov and E. Almeida Soares, "SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification," *medRxiv*, 2020.
[36] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "Covid-19 image data collection: Prospective predictions are the future," *arXiv preprint arXiv:2006.11988*, 2020.
[37] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2097–2106.
[38] M. E. H. Chowdhury *et al.*, "Can AI help in screening viral and COVID-19 pneumonia?," *IEEE Access*, vol. 8, pp. 132665–132676, 2020.
[39] T. Rahman *et al.*, "Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images," *Comput Biol Med*, vol. 132, p. 104319, 2021.
[40] J. S. Lim and A. V Oppenheim, "Enhancement and bandwidth compression of noisy speech," *Proceedings of the IEEE*, vol. 67, no. 12, pp. 1586–1604, 1979.
[41] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Comput Vis Graph Image Process*, vol. 29, no. 1, pp. 100–132, 1985.
[42] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Comput Geosci*, vol. 10, no. 2–3, pp. 191–203, 1984.
[43] A. Likas, N. Vlassis, and J. J. Verbeek, "The global k-means clustering algorithm," *Pattern Recognit*, vol. 36, no. 2, pp. 451–461, 2003.
[44] I. Davidson, "Understanding K-means non-hierarchical clustering," *SUNY Albany Technical Report*, vol. 2, pp. 2–14, 2002.
[45] T. Pang-Ning, M. Steinbach, and V. Kumar, "Introduction to data mining Addison-Wesley," 2005.
[46] P. K. Bhagat, P. Choudhary, and K. M. Singh, "Chapter 13 - A comparative study for brain tumor detection in MRI images using texture features," in *Sensors for Health Monitoring*, N. Dey, J. Chaki, and R. Kumar, Eds., Academic Press, 2019, pp. 259–287. doi: https://doi.org/10.1016/B978-0-12-819361-7.00013-0.
[47] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans Syst Man Cybern*, no. 6, pp. 610–621, 1973.
[48] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, Ieee, 2005, pp. 886–893.

[49] T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions," in *Proceedings of 12th international conference on pattern recognition*, IEEE, 1994, pp. 582–585.
[50] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit*, vol. 29, no. 1, pp. 51–59, 1996.
[51] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, pp. 513–529, 2011.
[52] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
[53] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *2004 IEEE international joint conference on neural networks (IEEE Cat. No. 04CH37541)*, Ieee, 2004, pp. 985–990.
[54] V. Vapnik, *The nature of statistical learning theory*. Springer science & business media, 2013.
[55] F. H. Garabaghi, R. Benzer, S. Benzer, and A. Ç. Günal, "Effect of polynomial, radial basis, and Pearson VII function kernels in support vector machine algorithm for classification of crayfish," *Ecol Inform*, vol. 72, p. 101911, Dec. 2022, doi: 10.1016/J.ECOINF.2022.101911.
[56] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," *Journal of the ACM (JACM)*, vol. 45, no. 6, pp. 891–923, 1998.
[57] A.-M. Šimundić, "Measures of diagnostic accuracy: basic definitions," *EJIFCC*, vol. 19, no. 4, p. 203, 2009.

**Acknowledgement**

This study was funded by the Scientific Research Projects Coordination Unit of Tekirdağ Namık Kemal University. Project number: NKUBAP.06.GA.21.317

**Data availability**
Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

**Conflict of Interest Notice**
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Ethical Approval and Informed Consent**
It is declared that during the preparation process of this study, scientific and ethical principles were followed, and all the studies benefited from are stated in the bibliography.

**Plagiarism Statement**
This article has been scanned by iThenticate ™.