

## Classification of Chest X-ray COVID-19 Images Using the Local Binary Pattern Feature Extraction Method

Narin Aslan<sup>1\*</sup>, Sengul Dogan<sup>2</sup>, Gonca Ozmen Koca<sup>1</sup>

<sup>1</sup>Department of Mechatronic Engineering, College of Technology, Firat University, Elazig/ Turkey

<sup>2</sup>Department of Digital Forensics Engineering, College of Technology, Firat University, Elazig/ Turkey

<sup>1\*</sup>[narin.aslan@firat.edu.tr](mailto:narin.aslan@firat.edu.tr), <sup>2</sup>[sdogan@firat.edu.tr](mailto:sdogan@firat.edu.tr), <sup>1</sup>[gozmen@firat.edu.tr](mailto:gozmen@firat.edu.tr)

(Geliş/Received: 24/03/2022;

Kabul/Accepted: 09/08/2022)

**Abstract:** *Background and Purpose:* COVID-19, which started in December 2019, caused significant loss of life and economic losses. Early diagnosis of the COVID-19 is important to reduce the risk of death. Therefore, studies have increased to detect COVID-19 with machine learning methods automatically. *Materials and Methods:* In this study, the dataset consists of 15153 X-ray images for 4961 patient cases in three classes: Viral Pneumonia, Normal and COVID-19. Firstly, the dataset was preprocessed. And then, the dataset was given to the Cubic Support Vector Machine (Cubic SVM), Linear Discriminant (LD), Quadratic Discriminant (QD), Ensemble, Kernel Naive Bayes (KNB), K-Nearest Neighbor Weighted (KNN Weighted) classification methods as input data. Then, the Local Binary Model (LBP) texture operator was applied for feature extraction. *Results:* These values were increased from 94.1% (without LBP) to 98.05% using the LBP method. The Cubic SVM method's highest accuracy was observed in these two applications. *Conclusions:* This study demonstrates that the performance of the presented methods with LBP feature extraction is improved.

**Keywords:** Covid-19, local binary pattern, feature extraction, machine learning, classification.

## Chest X-ray COVID-19 Görüntülerinin Yerel İkili Model Özellik Çıkarımı Yöntemi Kullanılarak Sınıflandırılması

**Öz:** *Arka Plan ve Amaç:* Aralık 2019'da başlayan COVID-19, önemli can ve ekonomik kayıplara neden oldu. Ölüm riskini azaltmak için COVID-19'un erken teşhisi çok önemlidir. Bu nedenle, COVID-19'u makine öğrenmesi yöntemleriyle otomatik olarak tespit etmeye yönelik çalışmalar artmaktadır. *Materyal ve Metot:* Bu çalışmada veri seti, Viral Pnömoni, Normal ve COVID-19 olmak üzere üç sınıftaki 4961 hasta vakası için 15153 X-ray görüntüsünden oluşmaktadır. Öncelikle, veri seti ön işlemden geçirildi. Daha sonra, Cubic Support Vector Machine (Cubic SVM), Linear Discriminant (LD), Quadratic Discriminant (QD), Ensemble, Kernel Naive Bayes (KNB), K-Nearest Neighbor Weighted (KNN Weighted) sınıflandırma metodlarına girdi datası olarak verildi. Daha sonra özellik çıkarımı için Yerel İkili Model (LBP) doku operatörü uygulandı. *Bulgular:* Bu değerler LBP yöntemi kullanılarak %94,1'den (LBP kullanılmadan) %98,05'e yükseltildi. Bu iki farklı uygulamada en yüksek doğruluk Cubic SVM yönteminde gözlemlendi. *Sonuçlar:* Bu çalışma, LBP öznelik çıkarımı ile sunulan yöntemlerin performansının arttığını göstermektedir.

**Anahtar kelimeler:** Covid-19, yerel ikili model, özellik çıkarma, makine öğrenimi, sınıflandırma.

### 1. Introduction

Coronavirus 2019 (COVID-19), a pandemic virus type, causes respiratory tract infections in humans. This epidemic, which emerged in Wuhan city of China in December 2019, caused the death of millions of people. COVID-19 poses a danger to global health and should be detected early. Therefore, it is important to facilitate the early diagnosis of this disease, predict the recovery day of the patient, and help the specialists who diagnose the disease [1, 2]. Due to these necessary reasons, studies on the machine learning method for the specification of COVID-19 have gained momentum. In addition, studies on lung computed tomography (CT) and X-ray images were analyzed [3]. Some related studies in the literature are shortly mentioned below:

Hasoon et al. [4] used feature extractors with Local Binary Pattern (LBP), Gradient Histogram, and Haralick texture features after preprocessing the COVID-19 data. KNN and SVM were applied in classification methods. An average of 98.66% accuracy performance was observed in the LBP-KNN model. Jawahar et al. [5] suggested a Local Binary Model technique to predict COVID-19 disease using X-ray images and extract the images distinctive features. The features were given as input data to various classifiers. As a result of the study, 77.7%

<sup>1\*</sup>Corresponding author: [narin.aslan@firat.edu.tr](mailto:narin.aslan@firat.edu.tr). ORCID Number of authors: <sup>1\*</sup> 0000-0002-7609-1557, <sup>2</sup> 0000-0001-9677-5684, <sup>1</sup> 0000-0003-1750-8479

accuracy was obtained in the Random Forest classifier. Tuncer et al. [6] proposed a sample Local Binary Model (ResExLBP) feature generation method to detect COVID-19. The work consists of preprocessing, feature selection, and feature extraction. Grayscale conversion and image resizing were applied in the preprocessing stage. Iterative ReliefF (IRF) was used in the feature selection phase. In the classification phase, they worked on decision tree (DT), SVM, subspace discriminant (SD), and K-Nearest Neighbor (KNN) methods. 100% accuracy was obtained in the SVM classifier. Lakshmi et al. [7] classified 2815 COVID CT images using two different datasets, COVID and non-COVID. The logarithmic transformation of the LBP (LT-LBP) was applied in the feature extraction. KNN, SVM, Random Forest (RF), and Logistic Regression (LR) methods were used for classification. The accuracy value for the LD-LBP method combined with SVM was calculated as 95.7%. Alquran et al. [8] applied LBP, Gabor Filter, and Gray Level Co-Occurrence Matrix (GLCM) texture features extraction methods to machine learning methods to detect COVID-19. The 1929 X-ray image of the lung was analyzed. SVM, RF, KNN, ANN, and Ensemble were used for classification. The best accuracy obtained using the Ensemble classifier was observed as 93.1%. Abed et al. [9] detected COVID-19 disease using SVM, radial basis function (RBF), linear kernel, DT, KNN, CN 2 rule induction techniques, and deep learning models. 800 X-ray images were used. The best accuracy of 95% was observed in the SVM method. Barstugan et al. [10] analyzed 150 CT images for COVID-19 classification. Gray Level Working Length Matrix (GRLLM), GLCM, Discrete Wavelet Transform (DWT), Local Directional Model (LDP), and Gray Level Dimension Region Matrix (GLSZM) algorithms were applied as feature selection methods. These features extracted by SVM method are classified. With the GLSZM feature extraction method, 99.68% classification accuracy was obtained. Rohman and Bustamam [11] analyzed the classification of COVID-19 disease based on tissue feature selection using X-ray and CT scan images. GLCM, Histogram of Oriented Pattern (HOG), and LBP were implemented to select texture features. 1100 X-ray and 1100 CT images were used for analysis. SVM method was applied in classification. Predictive accuracy of 97% and 99% was observed in the classification of CT and X-ray images, respectively. Amini and Shalbah [12] determined the severe, moderate, and mild severity of COVID-19 from 956 CT images. They used second-order statistical and a number of quantitative first texture features. Variance, kurtosis, and skewness are the first-order tissue features extracted from the histogram. GLCM, GLRLM, and GLSZM constitute the quadratic texture feature extraction methods. It was detected with 90.95% accuracy in the classification made using random forest (RF).

In this study, an open-access dataset containing X-ray images was used to detect COVID-19. Local Binary Model (LBP) feature extraction method was applied to this data set. The most important features obtained in LBP feature extraction were given as input data to classification methods such as support vector machine (SVM), Linear Discriminant (LD), Quadratic Discriminant (QD), Ensemble, Naive Bayes (NB), and K-Nearest Neighbor. In the diagnosis of COVID-19, classification and performance criteria were measured, and the results were compared. We can summarize the contribution of our work as follows:

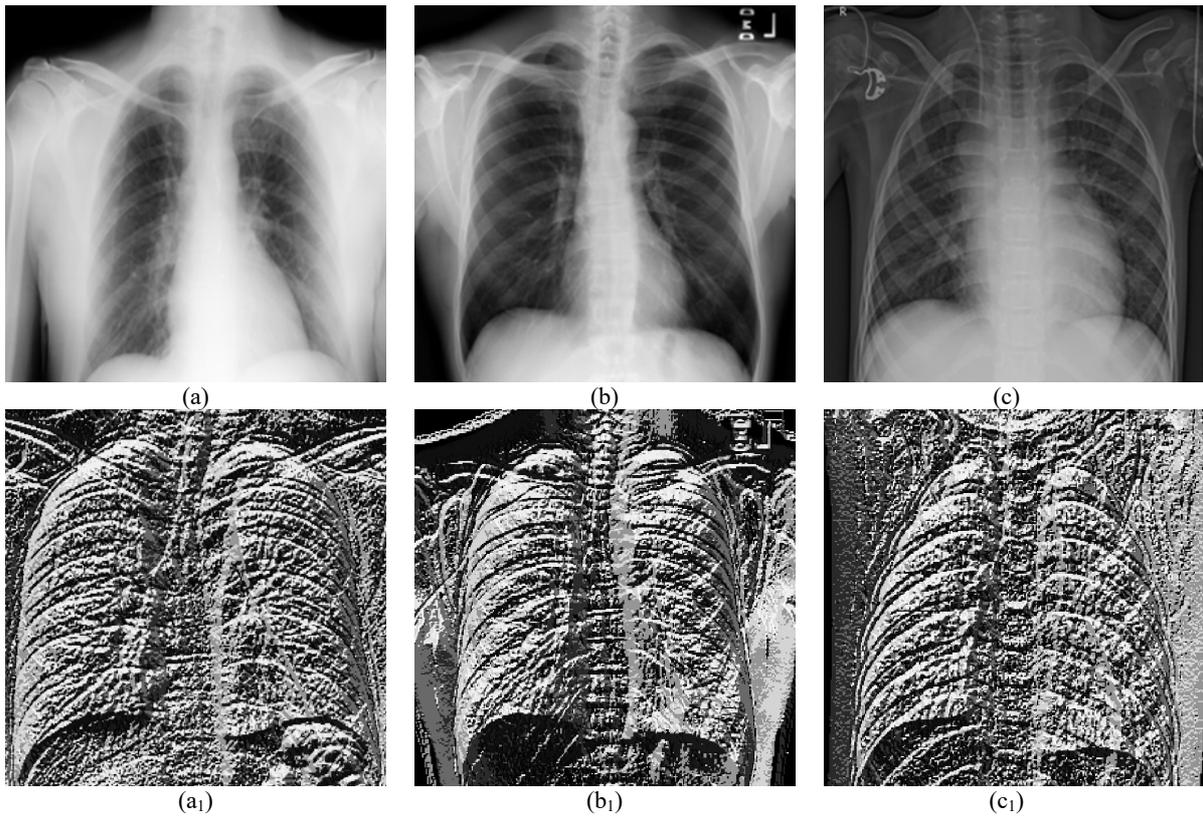
- Six different classification methods were combined Cubic SVM, LD [13], QD [14], Ensemble [15], KNN, and KNN with the LBP feature extraction operator.
- A large dataset of 15153 X-ray images obtained very high predictive values in performance criteria.
- LBP feature extraction was used to extract high-level features from images.
- Among the other five classification models, the best results were seen in the LBP-Cubic SVM model.
- We achieve the highest performance criteria with 98.05% accuracy, 90.99% sensitivity, 95.39% specificity, and 91.71% F score.

The remainder of the paper study is structured as follows: First, dataset preparation and preprocessing, feature selection, and evaluation metrics are discussed in **Section 2**. In **Section 3**, the experimental results are described. In **Section 4** provides a comprehensive discussion of similar literature studies. Finally, the conclusion and future work are discussed in **Section 5**.

## 2. Material and Method

### 2.1 Dataset

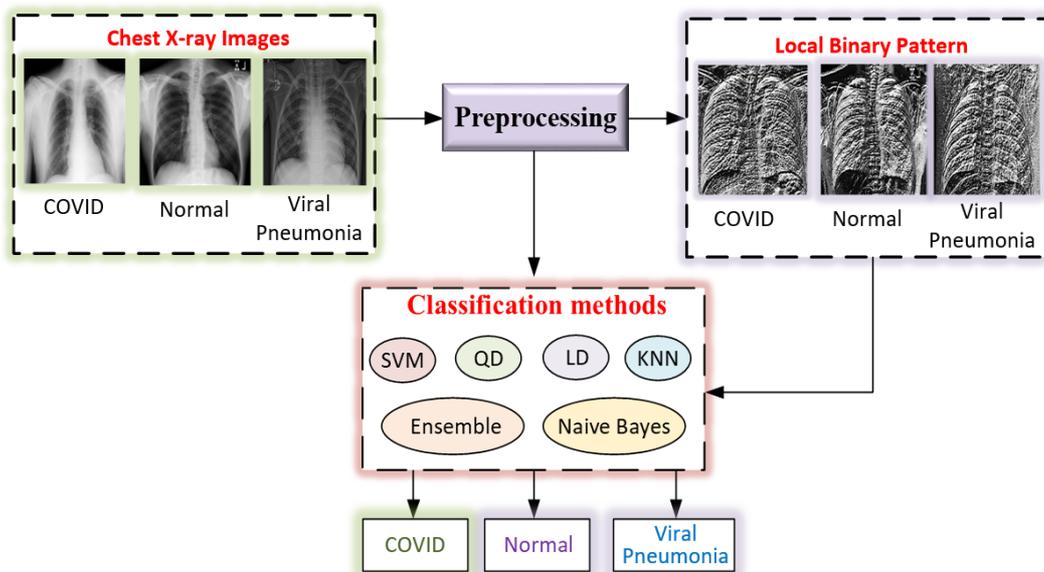
Chest X-ray images are a data set containing three classes: 10192 Normal, 3616 COVID-19, and 1345 Viral Pneumonia [16]. The resulting dataset has a total of 15153 images. These images consist of 299×299 Portable Network Graphics (PNG) image files. The original X-ray images of Normal, COVID-19, and Viral Pneumonia are presented in three classes, and textural images of the same images obtained by applying LBP are shown in Figure 1.



**Figure 1.** The original images of COVID-19 (a), Normal (b), and Viral Pneumonia (c) X-ray images and the feature images of these images obtained by applying LBP are (a<sub>1</sub>), (b<sub>1</sub>), and (c<sub>1</sub>), respectively.

## 2.2 Method

The flow chart covering the content of the study is shown in Figure 2.



**Figure 2.** Workflow of proposed study framework for classifying the COVID-19 status in X-Ray images.

The steps followed in the presented study are mentioned below:

**Step 1:** The chest X-ray images are tagged in three classes: COVID, Normal, and Viral Pneumonia.

**Step 2:** Since LBP is applied to a one-dimensional image, these images containing Portable Network Graphics (PNG) of  $299 \times 299$  in the preprocessing part were converted to  $224 \times 224$  data size.

**Step 3:** These images are given as input data to SVM, QD, LD, KNN, Ensemble, and Naive Bayes classification methods.

**Step 4:** The performance criteria of the classification results in three classes (COVID, Normal and Viral Pneumonia) are calculated.

**Step 5:** The input data is reduced to one dimension before applying LBP feature extraction.

**Step 6:** These input data, which applied LBP feature extraction, are given as input data again to SVM, QD, LD, KNN, Ensemble, and Naive Bayes classification methods.

**Step 7:** After applying LBP feature extraction, performance criteria for classification results in three classes (COVID, Normal and Viral Pneumonia) are calculated.

### 2.2.1. Local Binary Pattern

Local Binary Pattern (LBP) is a very effective tissue operator that labels the pixels of the image and evaluates the result as a binary number. It is a texture measurement method independent of gray level. It is a common approach in various applications due to its computational simplicity and distinctive power. Computational simplicity is its most important feature, making it possible to analyze images in real-time adjustment. The original LBP operator constrains every pixel  $3 \times 3$  neighborhood of every pixel to its center value. It evaluates the result as a binary number and creates labels for the image pixels [17, 18]. The following equations are used in the LBP method to obtain the feature set.

$$LBP = \sum_{n=0}^7 S(I_n - I_c) \times 2^n \quad (1)$$

$$A_{i,j}^1 = S(I_1, I_c) \times 2^7 + S(I_2, I_c) \times 2^6 \quad (2)$$

$$A_{i,j}^2 = S(I_3, I_c) \times 2^5 + S(I_4, I_c) \times 2^4 \quad (3)$$

$$A_{i,j}^3 = S(I_5, I_c) \times 2^3 + S(I_6, I_c) \times 2^2 \quad (4)$$

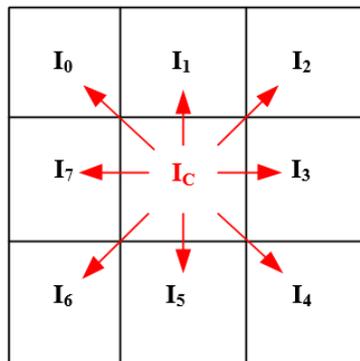
$$A_{i,j}^4 = S(I_7, I_c) \times 2^1 + S(I_8, I_c) \times 2^0 \quad (5)$$

$$A = A_{i,j}^1 + A_{i,j}^2 + A_{i,j}^3 + A_{i,j}^4 \quad (6)$$

$$LBP = Histogram_A \quad (7)$$

$$S(z) = \begin{cases} 0, & z < 0 \\ 1, & z \geq 0 \end{cases} \quad (8)$$

Here,  $I_c$  and  $I_n$  represent the center pixel value and the neighboring pixel value, respectively.  $z$  represents the difference between the neighboring pixel, and the center pixel  $S(z)$  represents the bits produced as a result of the LBP operator. The pixel labeling scheme in the LBP operator is given in Figure 3.



**Figure 3.** Labeling pixels in the LBP operator.

The performance criteria used in this study are given below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$Sensitivity = \frac{TP}{Positive} \quad (10)$$

$$Sensitivity = \frac{TN}{Negative} \quad (11)$$

$$F\_score = \frac{((1 + \beta^2)) * (sensitivity * precision)}{((\beta^2) * (precision + sensitivity))} \quad (12)$$

where, the value of beta is 1. Also, True Negative, True Positive, False Positive, and False Negative are TN, TP, FP, and FN, respectively [19-21].

### 3. Experimental Results

The dataset used in this study includes 15153 Viral Pneumonia, Normal, and COVID-19 Chest X-ray images. In the first stage of this study, the dataset size consisting of 299×299 Portable Network Graphics (PNG) image files is converted to 224×224 data size. Then, Cubic SVM, LD, QD, Ensemble, Kernel Naive Bayes, and KNN Weighted classification methods are given separately as the input dataset. The classification learner of MATLAB R2020a provides this study with a personal computer with 16 GB RAM and a 3.30 GHz processor. For each classification method, 10-fold cross validation is used. In the second step, important features are selected from the dataset with LBP feature extraction, and these features are applied again as input data to the specified classification methods. Finally, the Accuracy, Sensitivity, Specificity, and F\_Score performance criteria of the experimental results obtained without feature extraction are calculated, and these results are shown in Table 1.

**Table 1.** Classification results without applying LBP feature extraction.

|                        | Accuracy | Sensitivity | Specificity | F_score |
|------------------------|----------|-------------|-------------|---------|
| Cubic SVM              | 94.1     | 91.25       | 95.61       | 91.86   |
| Linear Discriminant    | 87.0     | 80.87       | 89.91       | 81.52   |
| Quadratic Discriminant | 86.7     | 86.54       | 91.64       | 82.29   |
| Ensemble               | 88.4     | 79.93       | 90.43       | 83.12   |
| Kernel Naive Bayes     | 71.2     | 72.66       | 83.89       | 64.48   |
| KNN Weighted           | 90.0     | 82.82       | 91.37       | 85.51   |

As seen in Table 1, the highest estimation accuracy of the Cubic SVM method is 94.1%. Sensitivity, Specificity, and F\_score performance criteria of Cubic SVM are 91.25%, 95.61%, and 91.86%, respectively. In addition, these performance criteria are higher than the performance criteria of other classification methods. Although the lowest accuracy is seen in the Kernel Naive Bayes method with a rate of 71.2%, the Linear Discriminant and Quadratic Discriminant methods showed close values of 87.0% and 86.7%, respectively.

The confusion matrix is used to compare target feature estimates and actual values to measure the performance of classification methods [22]. The confusion matrix results obtained without LBP feature extraction are shown in Figure 4.

|            |   | Predicted Class |          |      |                     |   | Predicted Class |      |      |                        |   | Predicted Class |      |      |
|------------|---|-----------------|----------|------|---------------------|---|-----------------|------|------|------------------------|---|-----------------|------|------|
|            |   | 1               | 2        | 3    |                     |   | 1               | 2    | 3    |                        |   | 1               | 2    | 3    |
| True Class | 1 | 3221            | 360      | 35   | True Class          | 1 | 2506            | 974  | 136  | True Class             | 1 | 2843            | 689  | 84   |
|            | 2 | 240             | 9856     | 96   |                     | 2 | 338             | 9608 | 246  |                        | 2 | 497             | 9060 | 635  |
|            | 3 | 33              | 129      | 1183 |                     | 3 | 65              | 217  | 1063 |                        | 3 | 41              | 65   | 1239 |
| Cubic SVM  |   |                 |          |      | Linear Discriminant |   |                 |      |      | Quadratic Discriminant |   |                 |      |      |
|            |   | Predicted Class |          |      |                     |   | Predicted Class |      |      |                        |   | Predicted Class |      |      |
|            |   | 1               | 2        | 3    |                     |   | 1               | 2    | 3    |                        |   | 1               | 2    | 3    |
| True Class | 1 | 2651            | 895      | 70   | True Class          | 1 | 2096            | 1033 | 487  | True Class             | 1 | 2614            | 890  | 112  |
|            | 2 | 329             | 9790     | 73   |                     | 2 | 1559            | 7539 | 1100 |                        | 2 | 152             | 9968 | 72   |
|            | 3 | 109             | 289      | 947  |                     | 3 | 108             | 79   | 1158 |                        | 3 | 49              | 242  | 1054 |
| Ensemble   |   |                 |          |      | Kernel Naive Bayes  |   |                 |      |      | KNN Weighted           |   |                 |      |      |
|            |   | 1               | COVID-19 | 2    | Normal              | 3 | Viral Pneumonia |      |      |                        |   |                 |      |      |

**Figure 4.** Confusion matrix of classification results without applying LBP feature extraction.

Confusion matrices by class Covid-19 (1), Normal (2), and Viral Pneumonia (3) are used to visualize their predictive accuracy. Pink cells correct prediction values and green cells indicate incorrect prediction values are shown. It is also understood from the confusion matrix that the best result is seen in the Cubic SVM method. Here, the Normal class achieved a classification accuracy of 96.70%, while the COVID-19 class has an accuracy of 89.08% on 3616 test samples, classifying as 360 samples Normal and 35 samples as Viral Pneumonia. In the normal class 10192 test samples, 240 samples are included COVID-19 and 96 samples of Viral Pneumonia. While the Viral Pneumonia class distribution reached an accuracy of 87.96% in 1345 test samples, 33 samples are in the COVID-19, and 129 samples are in the Normal class.

LBP feature extraction is applied to increase the predictive accuracy of disease diagnosis in Chest X-ray images and improve performance criteria. The results obtained after applying LBP feature extraction are presented in Table 2.

**Table 2.** Classification results applying LBP feature extraction.

|                        | Accuracy | Sensitivity | Specificity | F_score |
|------------------------|----------|-------------|-------------|---------|
| Cubic SVM              | 98.05    | 90.99       | 95.39       | 91.71   |
| Linear Discriminant    | 95.36    | 80.56       | 89.78       | 81.24   |
| Quadratic Discriminant | 95.71    | 86.41       | 91.62       | 82.10   |
| Ensemble               | 95.80    | 79.38       | 90.06       | 82.61   |
| Kernel Naive Bayes     | 89.03    | 72.46       | 83.78       | 64.25   |
| KNN Weighted           | 96.31    | 82.27       | 91.12       | 84.94   |

As seen in Table 2, the Cubic SVM method's highest prediction accuracy is 98.05%. The accuracy value is 94.1% in the study without LBP feature extraction. When the other methods in the table examined, an increase in the accuracy performance criterion observed. The Linear Discriminant, Quadratic Discriminant, and Ensemble classification methods have almost the same predictive accuracy, with 95.36%, 95.71%, and 95.80%, respectively. The KNN method has 96.31% predictive accuracy. After applying LBP feature extraction to the input data, the best prediction accuracy is presented in Figure 5, again with the Cubic SVM method.

|            |                 |          |      |        |                     |                 |      |      |                        |      |      |      |
|------------|-----------------|----------|------|--------|---------------------|-----------------|------|------|------------------------|------|------|------|
|            | Predicted Class |          |      |        | Predicted Class     |                 |      |      | Predicted Class        |      |      |      |
|            |                 | 1        | 2    | 3      |                     | 1               | 2    | 3    |                        | 1    | 2    | 3    |
| True Class | 1               | 3195     | 392  | 29     | 1                   | 2496            | 984  | 136  | 1                      | 2844 | 683  | 89   |
|            | 2               | 243      | 9852 | 97     | 2                   | 348             | 9596 | 248  | 2                      | 504  | 9047 | 641  |
|            | 3               | 31       | 131  | 1183   | 3                   | 69              | 220  | 1056 | 3                      | 44   | 66   | 1235 |
|            | Cubic SVM       |          |      |        | Linear Discriminant |                 |      |      | Quadratic Discriminant |      |      |      |
|            | Predicted Class |          |      |        | Predicted Class     |                 |      |      | Predicted Class        |      |      |      |
|            |                 | 1        | 2    | 3      |                     | 1               | 2    | 3    |                        | 1    | 2    | 3    |
| True Class | 1               | 2604     | 938  | 74     | 1                   | 2092            | 1037 | 487  | 1                      | 2588 | 909  | 119  |
|            | 2               | 352      | 9770 | 70     | 2                   | 1572            | 7507 | 1113 | 2                      | 166  | 9950 | 76   |
|            | 3               | 108      | 292  | 945    | 3                   | 112             | 78   | 1155 | 3                      | 52   | 249  | 1044 |
|            | Ensemble        |          |      |        | Kernel Naive Bayes  |                 |      |      | KNN Weighted           |      |      |      |
|            | 1               | COVID-19 | 2    | Normal | 3                   | Viral Pneumonia |      |      |                        |      |      |      |

Figure 5. Confusion matrix of classification results applying LBP feature extraction.

While classifying 392 samples as Normal and 29 samples as Viral Pneumonia out of 3616 test samples, a classification accuracy of 98.05% is achieved. 10192 test data of the normal class are contained 243 data COVID-19 and 97 Viral Pneumonia data. The Viral Pneumonia class distribution in 1345 test data are included 31 data in the COVID-19 class and 131 data in the Normal class.

#### 4. Discussion

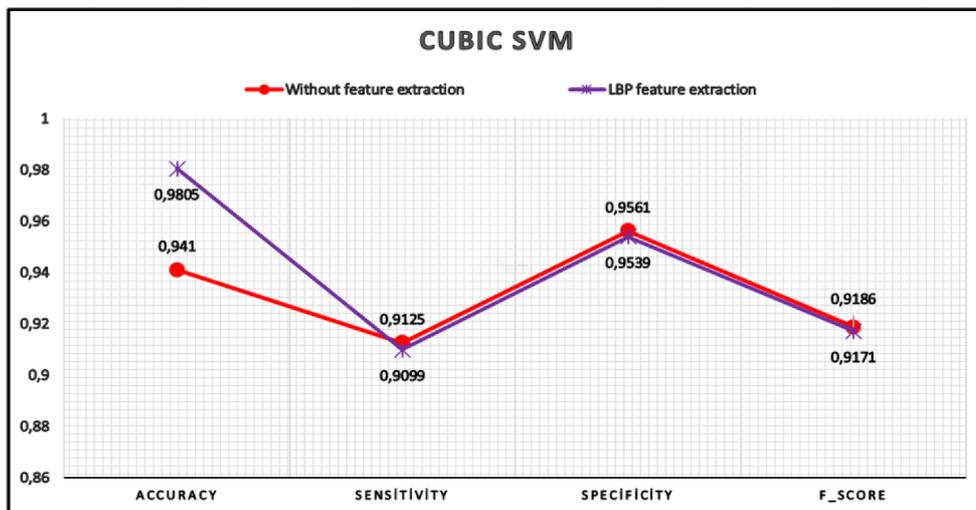
In this section, the classification studies of COVID-19 disease with the machine learning method in the literature are compared in Table 3.

Table 3. Comparison of machine learning methods and chest X-ray models.

| Ref. | Number of Sample | Feature extraction | Method | Accuracy | Sensitivity | Specificity | F_score |
|------|------------------|--------------------|--------|----------|-------------|-------------|---------|
| [4]  | 5000             | LBP                | KNN    | 98.66    | 97.76       | 100         | -       |
|      |                  |                    | SVM    | 94.25    | 99.94       | 88.10       | -       |
|      |                  | HOG                | KNN    | 94.26    | 79.95       | 70.19       | -       |
|      |                  |                    | SVM    | 89.20    | 78.61       | 65.30       | -       |
|      |                  | Haralick           | KNN    | 95.51    | 93.98       | 97.88       | -       |
|      |                  |                    | SVM    | 94.88    | 99.96       | 89.23       | -       |

|           |       |       |                    |       |       |       |       |
|-----------|-------|-------|--------------------|-------|-------|-------|-------|
| [10]      | 150   | GLCM  | SVM                | 98.91 | 98.52 | 99.23 | 98.81 |
|           |       | LDP   |                    | 50.70 | 42.47 | 57.71 | 44.14 |
|           |       | GLRLM |                    | 96.41 | 98.78 | 94.38 | 96.20 |
|           |       | GLSZM |                    | 98.77 | 97.72 | 99.67 | 98.65 |
|           |       | DWT   |                    | 97.81 | 96.8  | 98.66 | 97.60 |
| [11]      | 2200  | GLCM  | SVM                | 95.4  | 96.8  | -     | -     |
|           |       | LBP   |                    | 97.5  | 98.0  | -     | -     |
|           |       | HOG   |                    | 97.8  | 100   | -     | -     |
| Our study | 15153 | LBP   | Cubic SVM          | 98.05 | 90.99 | 95.39 | 91.71 |
|           |       |       | LD                 | 95.36 | 80.56 | 89.78 | 81.24 |
|           |       |       | QD                 | 95.71 | 86.41 | 91.62 | 82.10 |
|           |       |       | Ensemble           | 95.80 | 79.38 | 90.06 | 82.61 |
|           |       |       | Kernel Naive Bayes | 89.03 | 72.46 | 83.78 | 64.25 |
|           |       |       | KNN Weighted       | 96.31 | 82.27 | 91.12 | 84.94 |

In the study, Hasoon et al. [4] used 5000 datasets, which applied LBP, HOG, and Haralick feature extraction methods. KNN and SVM methods were used as classification methods. The performance criteria values were calculated by taking the average of the 5-fold cross-validation predictive values. The highest accuracy value was seen at 98.66% in the KNN classification made with the LBP feature extraction method. The lowest accuracy estimate was observed at 89.20% in the SVM classification made with the HOG feature extraction method. In addition, in the HOG-SVM method the lowest sensitivity and specificity values were found to be 78.61% and 65.30%, respectively. When this study is compared with ours, the number of the dataset used in classification is less than ours. Although 15153 data are used in our study, in the LBP-SVM method is 98.05%, the accuracy value and the value in the LBP-KNN method in [4] is close to the accuracy rate of 98.66%. Barstugan et al. [10] classified with the different number of patches. 150 data were used in the study, which was carried out by applying 10-fold cross-validation. The highest predictive accuracy was found to be 98.91% in the GLCM-SVM method. The lowest accuracy, sensitivity, specificity, and F\_score performance criteria were observed in the LDP-SVM method as 50.70%, 42.47%, 57.71%, and 44.14%, respectively. Although 2200 data were used in the study by Rohmah and Bustamam [11], an accuracy rate of 97.5% was observed in the analysis performed using the LBP-SVM method. Our study calculated the predictive accuracy value as 98.05% in the LBP-SVM method. When our study is compared with the studies in the literature, although the number of data in the study is high, the performance criteria values calculated with the Cubic SVM, LD, QD, Ensemble, Kernel Naive Bayes, and KNN Weighted methods used in classification vary between 98.05% and 64.25%. In the Cubic SVM classification method, the performance criterion result graph in the analysis performed LBP feature extraction, and without LBP feature extraction is shown in Figure 6.



**Figure 6.** Performance criteria graph after feature selection in Cubic SVM classification method and LBP feature selection.

Although sensitivity, specificity, and F\_score values are close, a noticeable increase in accuracy values is observed. Although the accuracy value in the analysis performed without applying feature extraction to the input data is 94.10%, the accuracy value in the analysis performed after applying LBP feature extraction is 98.05%.

## 5. Conclusion

In this study, classification is performed using 15153 Chest X-ray images belonging to the Normal, COVID-19, and Viral Pneumonia classes. First, the input data with the size of 229×229 is converted to the size of 224×224, and these data are applied as input data to the Cubic SVM, LD, QD, Ensemble, Kernel Naive Bayes, and KNN Weighted classification methods. Then, LBP feature extraction is applied to the input data, and these features are given as inputs to the respective classification methods. The performance criteria calculated in both applications are compared. The analysis performed after LBP feature extraction observed that the accuracy value increased from 94.1% to 98.05%. In the future, we intend to use and test variants of other feature extraction and classification operators.

## References

- [1] L. Wynants *et al.*, "Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal," *bmj*, vol. 369, 2020.
- [2] D. Uphade and A. Muley, "Identification of parameters for classification of COVID-19 patient's recovery days using machine learning techniques," *J. Math. Comput. Sci.*, vol. 12, no. 3, p. Article ID 56, 2022.
- [3] S. H. Kassania, P. H. Kassanib, M. J. Wesolowskic, K. A. Schneidera, and R. Detersa, "Automatic detection of coronavirus disease (COVID-19) in X-ray and CT images: a machine learning based approach," *Biocybernetics and Biomedical Engineering*, vol. 41, no. 3, pp. 867-879, 2021.
- [4] J. N. Hasoon *et al.*, "COVID-19 anomaly detection and classification method based on supervised machine learning of chest X-ray images," *Results in Physics*, vol. 31, p. 105045, 2021.
- [5] M. Jawahar *et al.*, "Diagnosis of covid-19 using optimized pca based local binary pattern features," *International Journal of Current Research and Review*, pp. 37-41, 2021.
- [6] T. Tuncer, S. Dogan, and F. Ozyurt, "An automated Residual Exemplar Local Binary Pattern and iterative ReliefF based COVID-19 detection method using chest X-ray image," *Chemometrics and Intelligent Laboratory Systems*, vol. 203, p. 104054, 2020.
- [7] P. P. Lakshmi, M. Sivagami, and V. Balaji, "A novel LT-LBP based prediction model for COVID-CT images with Machine Learning," in *2021 International Conference on Information Systems and Advanced Technologies (ICISAT)*, 2021: IEEE, pp. 1-5.
- [8] H. Alquran, M. Alsleti, R. Alsharif, I. A. Qasmieh, A. M. Alqudah, and N. H. B. Harun, "Employing texture features of chest x-ray images and machine learning in covid-19 detection and classification," in *Mendel*, 2021, vol. 27, no. 1, pp. 9-17.
- [9] M. Abed *et al.*, "A comprehensive investigation of machine learning feature extraction and classification methods for automated diagnosis of COVID-19 based on X-ray images," *Computers, Materials, & Continua*, pp. 3289-3310, 2021.
- [10] M. Barstugan, U. Ozkaya, and S. Ozturk, "Coronavirus (covid-19) classification using ct images by machine learning methods," *arXiv preprint arXiv:2003.09424*, 2020.
- [11] L. N. Rohmah and A. Bustamam, "Improved classification of coronavirus disease (covid-19) based on combination of texture features using ct scan and x-ray images," in *2020 3rd International Conference on Information and Communications Technology (ICOIACT)*, 2020: IEEE, pp. 105-109.

- [12] N. Amini and A. Shalbaf, "Automatic classification of severity of COVID-19 patients using texture feature and random forest based on computed tomography images," *International Journal of Imaging Systems and Technology*, vol. 32, no. 1, pp. 102-110, 2022.
- [13] S. Balakrishnama and A. Ganapathiraju, "Linear discriminant analysis-a brief tutorial," *Institute for Signal and information Processing*, vol. 18, no. 1998, pp. 1-8, 1998.
- [14] S. Srivastava, M. R. Gupta, and B. A. Frigyik, "Bayesian quadratic discriminant analysis," *Journal of Machine Learning Research*, vol. 8, no. 6, 2007.
- [15] Y. Ren, L. Zhang, and P. N. Suganthan, "Ensemble classification and regression-recent developments, applications and future directions," *IEEE Computational intelligence magazine*, vol. 11, no. 1, pp. 41-53, 2016.
- [16] T. Rahman, M. Chowdhury, and A. Khandakar. "COVID-19 Radiography Database, COVID-19 Chest X-ray Database, <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>." (accessed).
- [17] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern recognition*, vol. 29, no. 1, pp. 51-59, 1996.
- [18] T. Ojala and M. Pietikäinen, "Unsupervised texture segmentation using feature distributions," *Pattern recognition*, vol. 32, no. 3, pp. 477-486, 1999.
- [19] X. Li, W. Tan, P. Liu, Q. Zhou, and J. Yang, "Classification of COVID-19 chest CT images based on ensemble deep learning," *Journal of Healthcare Engineering*, vol. 2021, 2021.
- [20] D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.
- [21] M. J. Warrens, "On the equivalence of Cohen's kappa and the Hubert-Arabie adjusted Rand index," *Journal of classification*, vol. 25, no. 2, pp. 177-183, 2008.
- [22] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, "An improved method to construct basic probability assignment based on the confusion matrix for classification problem," *Information Sciences*, vol. 340, pp. 250-261, 2016.