



Automatic Computer Aided Diagnostic for COVID-19 Based on Chest X-Ray Image and Particle Swarm Intelligence

Suhaila N. Mohammed¹ Fatin Sadiq Alkinani^{1*} Yasmin A. Hassan¹

¹*Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq*

* Corresponding author's Email: fatin.alkinani@sc.uobaghdad.edu.iq

Abstract: COVID-19 is a vital zoonotic illness caused by Severe Acute Respiratory Syndrome Corona Virus 2 (SARS-CoV-2). COVID-19 is a very wide-spread among humans thus the early detection and curing of the disease offers a high opportunity of survival for patients. Computed Tomography (CT) plays an important role in the diagnosis of COVID-19. As chest radiography can give an indicator of coronavirus. Though, an automated Computer Aided Diagnostic (CAD) system for COVID-19 based on chest X-Ray image analysis is presented in this article. It is designed for COVID-19 recognition from other MERS, SARS, and ARDS viral pneumonia. The optimal threshold value for the segmentation of a chest image is deduced by exploiting Li s' method and particle swarm intelligence. Laws' masks are then applied to the segmented chest image for secondary characteristics highlighting. After that, nine different vectors of attributes are extracted from the Grey Level Co-occurrence Matrix (GLCM) representation of each Law's mask result. Support vector machine ensemble models are then built based on the extracted feature vectors. Finally, a weighted voting method is utilized to combine the decisions of ensemble classifiers. Experimental findings show an accuracy of 98.04 %. It indicates that the suggested CAD scheme can be a promising supplementary COVID-19 diagnostic tool for clinical doctors.

Keywords: COVID-19, Chest X-ray, Li s' method, Particle swarm intelligence, Laws' masks, Gray level co-occurrence matrix (GLCM), Weighted voting, Support vector machine (SVM).

1. Introduction

In December 2019, a large outbreak of a novel infection with coronavirus occurred in Wuhan, Hubei province, China. In humans, coronaviruses are among the spectrum of viruses that cause the common cold as well as more severe respiratory diseases, specifically, Severe Acute Respiratory Syndrome (SARS), Acute Respiratory Distress Syndrome (ARDS) and Middle East Respiratory Syndrome (MERS). The disease caused by the virus, named Corona Virus Disease (COVID-19) by the World Health Organization (WHO), can spread through human-to-human contact. Since then, this highly contagious COVID-19 has spread worldwide, with a rapid rise in the number of deaths. This rapid spreading encourages the need for accurate identification and detection methods that can be

used in hospitals and clinics responsible for the diagnosis of COVID-19 [1] [2].

Currently, there is no effective cure for this virus. Also, there is an urgent need to increase global information of its infection mechanisms, distribution of lung parenchyma damage, and related trends, which leads to the necessary prompt identification or enhancement method of the diagnosis to ease the design of curative therapy. Computer Tomography (CT) is an effective imaging tool in the diagnosis and treatment of patients with coronavirus disease pneumonia. Many studies are emerging on the radiological presence of COVID-19 pneumonia. As the prevalent trend seen in COVID-19 pneumonia is ground-glass opacification, it is likely to be challenging to detect COVID-19 with chest X-Ray, in which this form of abnormality is often imperceptible, particularly in patients with few symptoms or low severity. By contrast, chest x-rays were commonly used in the diagnosis of SARS,

since opacification and consolidation were present early [3].

Artificial Intelligence (AI) techniques can be used with X-Ray image of the chest region to detect and follow-up of the disease. AI algorithms and discriminative features derived from the X-Ray image would be of big support to undertake Computer Aided Diagnostic (CAD) program that could be used in any country with access to X-Ray equipment to help in the diagnosis of COVID-19 [4]. Farid et al. (2020) presented a technique for recognizing the COVID-19 in CT images by proposing a Composite Hybrid Feature Extraction (CHFS). The selected features were classified by the Stack Hybrid Classification system (SHC). A total of 51 CT-images collected from Kaggle database website were used for model evaluation. An accuracy of about 96.07% has been achieved when using a Naïve Bayes as a meta-classifier in a hybrid classification [5]. Xu et al. (2020) established an early screening model using deep learning techniques to distinguish COVID-19 pneumonia from just Influenza pneumonia using CT images. A 3-dimensional deep learning model has been used to segment the CT image set. The infection type and total confidence score of this CT case were calculated with Noisy-or Bayesian function. The experiments result on a dataset consists of 1,710 CT samples, including 357 COVID-19, 390 Influenza-A-viral-pneumonia, and 963 irrelevant-to-infection (ground truth) showed that the overall accuracy of the proposed system was 86.7 % [6].

COVID-19 diagnoses have confined practices on chest X-Ray studies. Though, the main goal of this paper is to design a diagnosis system of identifying COVID-19 from other MERS, SARS, and ARDS viral pneumonia using chest X-Ray.

To enhance the diagnostic process of COVID-19, an automated CAD system for COVID-19 is presented in this article with the following two minor goals:

- Overcoming the uneven grey level distribution effect of the chest image during the segmentation process by combining Li s' method and particle swarm intelligence.
- Improving the detection ability of the proposed CAD system by using an ensemble method of classification. This will in effect help the doctors to accurately determine the state of chest X-Ray image for the better level of treatment.

The remaining of the paper is structured as follows: Section 2 shows the detailed description of the proposed CAD system for COVID-19. Section 3

illustrates the experimental results that are achieved when applying the proposed CAD system on real chest X-Ray images. Finally, work conclusions and ideas for future work are presented in section 4.

2. The proposed CAD method

As illustrated in Fig. 1, the proposed CAD system consists of five main steps which are: (1) contrast enhancement using contrast normalization, (2) automatic chest X-Ray image segmentation based on Li s' method and particle swarm intelligence (3) Laws' filter masks to transform the segmented image into another representation that highlights the important details of the chest, (4) GLCM feature extraction to extract texture features from the GLCM representation of each of Law s' masks' result, and (5) ensemble classification by building a set of SVMs models which then combined by using voting fusion method to classify the chest image into either positive (i.e., the patient suffers from COVID-19) or negative type (i.e. non-COVID-19).

2.1 Contrast enhancement

The essential improvement required to be performed on the chest X-Ray image is to elevate the contrast between the entire chest and the background region. To achieve this goal, contrast normalization is utilized. In contrast normalization, the range of intensities in the digitized X-Ray image of the chest is stretched to make full use of all possible intensities' range (i.e., [0-255]). The following function is used to stretch the contrast of the X-Ray image [7]:

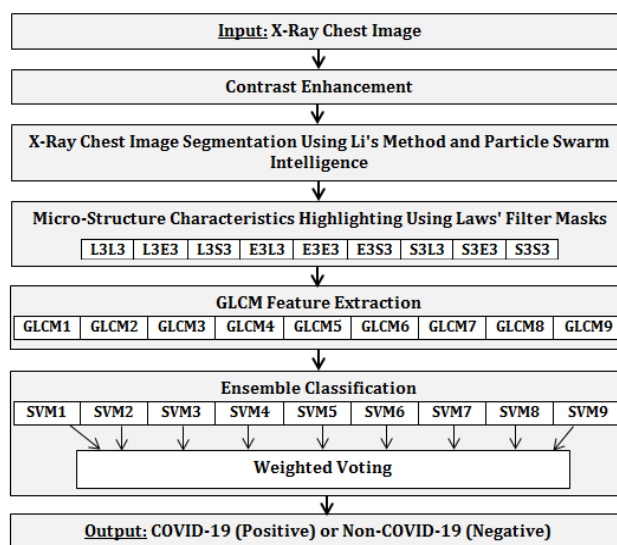


Figure. 1 The general design of the proposed CAD for COVID-19

$$I_s(x, y) = 255 \times \left(\frac{I(x, y) - \text{Min}}{\text{Max} - \text{Min}} \right) \quad (1)$$

Where, *Min* and *Max* are the minimum intensity value and maximum intensity value in the original image *I*, respectively, *I_s* represents the enhanced image and *(x, y)* is the coordinates of image pixel under processing.

2.2 Chest X-ray image segmentation

Chest image segmentation is a critical step in the proposed CAD framework because the segmented objects will then be transformed into high-level meaningful attributes. The thresholding technique is used to conduct image segmentation. It consists of choosing a value that separates the object from its context; this chosen value is called the threshold. However, finding the optimal (or near-optimal) threshold value is a challenging task due to the variability among different X-Ray image conditions. To fully automate the selection process of the ideal threshold value, a new method is proposed in this paper. The proposed method is based on the approach suggested by Li [8] to find the optimal threshold value. Li takes into consideration the variance between the foreground and background when finding the threshold. This approach can be defined as [9]:

$$J(\beta, t) = \beta \left(\sigma_1^2(t) + \sigma_2^2(t) \right) + (1 - \beta) \sigma_D^2(t) \quad (2)$$

and

$$\sigma_D^2(t) = \sigma_1(t) \sigma_2(t) \quad (3)$$

Where, $\sigma_D^2(t)$ refers to the measure of the variance degree between the object and the background, $\sigma_2(t)$ and $\sigma_1(t)$ are their standard deviation, respectively. In Li's equation, β is a weight balance parameter that can be tuned to determine the contributions of variance discrepancy and variance sum in threshold calculation. To obtain the optimal threshold t^* , Li proposed to minimize the following criterion:

$$J(a, t^*) = \text{Min}_{0 \leq t \leq L-1} J(\beta, t) \quad (4)$$

To find the optimal value of $t(t^*)$ that minimizes Eq. (4) automatically, the behavior of particle swarm during flying has been exploited in this work. The proposed segmentation method can be summarized with the following points:

- The population of the swarm consists of *N* particles. The initial position of each particle (*P_i*) is calculated as follows:

$$P_i = P_{i-1} + \left\lfloor \frac{255}{N} \right\rfloor \quad (5)$$

$$\text{where } P_1 = \left\lfloor \frac{255}{N} \right\rfloor$$

- Particles' velocities (*V_i*) are updated as in the following equation:

$$V_i = wV_i + \alpha_1(P_i - \text{Lbest}) + \alpha_2(P_i - \text{Gbest}) \quad (6)$$

Where α_1 and α_2 are learning factors that represent learning speed, *w* is inertia factor used to prevent the particle from changing its direction suddenly, *Lbest* is the best particle position in the current iteration and *Gbest* is the best particle position for all iterations.

- Particles' positions are updated using the following equation:

$$P_i = P_i + V_i \quad (7)$$

- To prevent particles from exceeding the acceptable threshold range [0,255], an additional condition is added after updating each particle position to check whether the particle position becomes more than 255 or less than 0. If this condition occurs, then the particle position must be returned to its acceptable range.
- At the end of the optimization process, the global best particle (*Gbest*) will be considered as the optimal threshold value for the X-Ray image under segmentation.

Algorithm 1 shows the main steps taken during the search for the optimal threshold value for automatic chest image segmentation.

Algorithm (1): Automatic Chest X-Ray Image Segmentation	
Input	Img: X-Ray chest image, N: population size, α_1 and α_2 : Learning rate factors, <i>w</i> : Inertia factor, Maxv: Maximum velocity, β : Li s' weight balance, Epoch: Number of iterations
Output	OpT: Optimal Threshold value
Steps	Initialize particles' positions using Eq. (5) For each particle (<i>P_i</i>) Do Set <i>V_i</i> ← 0 End For Set <i>Gbest</i> ← 0, Set <i>itr</i> ← 0 While (<i>itr</i> < Epoch) OR (fitness (<i>Gbest</i>) <> 0) Do For each particle (<i>P_i</i>) Do Compute <i>P_i</i> fitness value using Eq. (2) End For Find local best particle (<i>Lbest</i>) that minimizes

```

Eq. (4)
Update particles' velocities using Eq. (6)
Update particles' positions using Eq. (7)
If fitness(Lbest) is better than fitness(Gbest)
Then Set Gbest←Lbest
End While
Return Gbest

```

As a post-processing operation, grey scale erosion morphological operation is applied to the image resulted from the segmentation process. Erosion used to clean the conducted image from the segmentation by removing pixels which made small undesirable regions with intensity less than the value of predefined intensity threshold (I_{th}). The erosion of image I by structuring element SE is given by g in which SE is positioned with its origin at (x, y) and the new pixel value $g(x, y)$ is defined using the following rule [10]:

$$g(x, y) = \begin{cases} 1 & \text{if } SE \text{ hits } I(x, y) \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

2.3 Micro-structure characteristics highlighting

Law's masks are used to improve the texture information of the segmented chest image and highlight its micro-structure characteristics. The I - D Laws 'vectors of length three and five are [11]:

$$\begin{aligned}
 L3 &= [1 \ 2 \ 1], \\
 E3 &= [1 \ 0 \ -1], \\
 S3 &= [1 \ -2 \ 1], \\
 L5 &= [1 \ 4 \ 6 \ 4 \ 1], \\
 E5 &= [1 \ -2 \ 0 \ 2 \ 1], \\
 S5 &= [-1 \ 0 \ 2 \ 0 \ -1], \\
 W5 &= [-1 \ 2 \ 0 \ -2 \ 1], \\
 R5 &= [1 \ -4 \ 6 \ -4 \ 1]
 \end{aligned}$$

where L refers to Level, E to Edge, S to Spot and R to Ripple. The level vector (L) defines the average gray level or local weighted average, the edge (E) vector is equal to the gradient operator, the spot vector (S) denotes the spot extraction, the ripple vector (R) detects ripples from the image while the wave vector reacts to any image pixel modifications.

However, to apply these filters on 2 - D image, 2 - D filters of size 3×3 or 5×5 are computed by convolution any horizontal I - D vertical vector with another one. Therefore, nine 3×3 and twenty-five 5×5 possible combinations of 2 - D masks are generated [12]. The following nine 3×3 Law's texture masks are applied on the segmented chest image to produce a new texture image for every convolution mask [13]:

$$\begin{array}{lll}
 \text{L3L3} & & \text{L3E3} & & \text{L3S3} \\
 = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} & = & \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} & = & \begin{bmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{bmatrix} \\
 \text{E3L3} & & \text{E3E3} & & \text{E3S3} \\
 = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} & = & \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix} & = & \begin{bmatrix} -1 & -2 & 1 \\ 0 & 0 & 0 \\ -1 & 2 & -1 \end{bmatrix} \\
 \text{S3L3} & & \text{S3E3} & & \text{S3S3} \\
 = \begin{bmatrix} -1 & -2 & -1 \\ 2 & 4 & 2 \\ -1 & -2 & -1 \end{bmatrix} & = & \begin{bmatrix} -1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} & = & \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}
 \end{array}$$

2.4 GLCM feature extraction

GLCM-based attributes are extracted to discernment between the X-Ray chest image with positive COVID-19 and negative cases. The GLCM utilizes pixel pairs of a joint probability distribution. The joint probability distribution between pixels' pairs calculated by using angle " θ " and distance " d ". The $(i, j)^{th}$ entry in the GLCM matrix refers to the frequency with which the gray level i is followed by the gray level j with distance " d " and angle " θ ".

Since the size of the co-occurrence matrix depends on the number of gray levels existence in X-Ray image; the number of gray levels is reduced to M discrete levels to save the computational time. Nine different $M \times M$ GLCM matrixes are built to work as a base for nine different texture feature vectors extraction.

Consequently, the normalization process is performed by dividing each entry into GLCM by the total number of pixel pairs. The purpose of the normalization is to make the extracted features independent of the size of the image. Normalized $GLCM(N_{GLCM}(i, j))$ defined by:

$$N_{GLCM}(i, j) = \frac{GLCM(i, j)}{\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} GLCM(i, j)} \quad (9)$$

The following texture features are extracted from each normalized $GLCM(N_{GLCM})$ [14, 15]:

1) *Contrast*. This feature measures the local contrast of the chest image. Higher contrast is achieved when the difference between intensities of neighboring pixels increases. The contrast feature can be calculated using Eq. (10).

$$Contrast = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} N_{GLCM_{i,j}}(i - j)^2 \quad (10)$$

2) *Correlation*. It provides the degree of correlation between pixel pairs. A higher correlation is obtained when the intensities of adjacent pixels vary together. Pair's correlation can be obtained using Eq. (11).

$$\text{Correlation} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} N_{GLCM_{i,j}} \left[\frac{(i - \mu_i)(j - \mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right] \quad (11)$$

where μ_i and σ_i^2 refer to the mean and variance of $\sum_{i=0}^{M-1} N_{GLCM_{i,j}}$, μ_j and σ_j^2 refer to the mean and variance of $\sum_{j=0}^{M-1} N_{GLCM_{i,j}}$.

3) *Dissimilarity*. Dissimilarity measures the difference between any two adjacent pixels' intensities. The higher of this feature, the more differences among adjacent pixels' intensities. It can be found using Eq. (12).

$$\text{Dissimilarity} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} N_{GLCM_{i,j}} |i - j| \quad (12)$$

4) *Energy*. It measures the number of repeated pairs. An image containing a small number of repeated pairs has more energy than an image containing approximately the same number of pairs. Energy can be calculated using Eq. (13).

$$\text{Energy} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} (N_{GLCM_{i,j}})^2 \quad (13)$$

5) *Entropy*. The entropy feature measures the randomness or disorder of the image area. The maximum entropy is achieved if the probabilities of adjacent intensities are equal. This feature is defined with Eq. (14).

$$\begin{aligned} \text{Entropy} \\ = - \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} N_{GLCM_{i,j}} (\text{Log}_2 N_{GLCM_{i,j}}) \end{aligned} \quad (14)$$

6) *Homogeneity*. It measures the smoothness of the image area. A higher value of homogeneity is produced for a smoother image. Homogeneity can be obtained using Eq. (15).

$$\text{Homogeneity} = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \frac{N_{GLCM_{i,j}}}{1 + (i - j)^2} \quad (15)$$

7) *Rows Mean* (μ_i). μ_i is the average value in the *GLCM* concerning the rows. It can be calculated using the following equation:

$$\mu_i = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} i (N_{GLCM_{i,j}}) \quad (16)$$

8) *Columns Mean* (μ_j). μ_j is the average value in the *GLCM* concerning the columns as shown in Eq. (17).

$$\mu_j = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} j (N_{GLCM_{i,j}}) \quad (17)$$

9) *Rows Variance* (σ_i^2). It measures the spread of the *GLCM* frequency values concerning the rows and can be computed using Eq. (18).

$$\sigma_i^2 = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} N_{GLCM_{i,j}} (i - \mu_i)^2 \quad (18)$$

10) *Columns Variance* (σ_j^2). σ_j^2 measures the spread of the *GLCM* frequency values concerning the columns. It can be obtained using Eq. (19).

$$\sigma_j^2 = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} N_{GLCM_{i,j}} (j - \mu_j)^2 \quad (19)$$

2.5 Ensemble classification

The primary concept behind the ensemble model is that a group of weak learners form a powerful learner and thus increase the precision of the model. Different methods can be used for combining the decisions of the weak learners and producing the final prediction such as min, max, voting...etc. The weighted voting method has been utilized in this article in which the voting takes a convex combination of the base learners (c_1, c_2, \dots, c_L) to produce the final prediction y as illustrated in the following equation [16]:

$$y = f(c_1, \dots, c_L | \varphi) = \sum_{j=1}^L w_j c_j \quad (20)$$

where w_j and c_j are the weight and prediction output of learner j with $w_j \geq 0$ and $\sum_{j=1}^L w_j = 1$, $\varphi = (w_1, \dots, w_L)^T$ are the parameters that represent the weight of each classifier. w_j can be obtained based

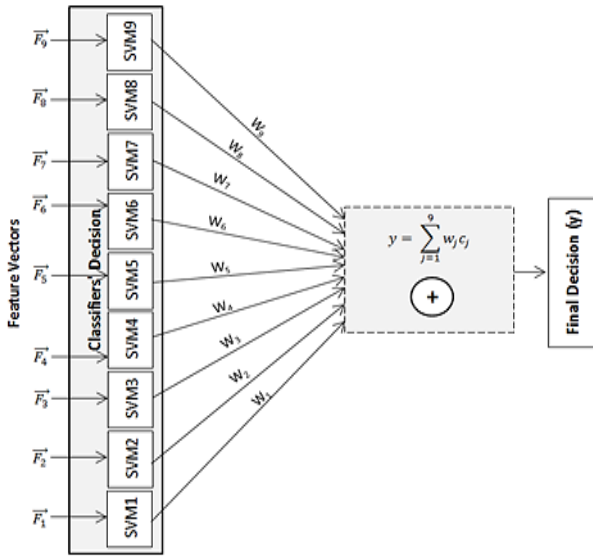


Figure. 2 Ensemble learning method used by the proposed CAD system

on the accuracy of each classifier (A_1, \dots, A_L) as shown in Eq. (21).

$$w_j = \frac{A_j}{\sum_{i=1}^L A_i} \quad (21)$$

Nine different binary SVMs models are constructed in this article as weak classifiers using the nine different feature vectors extracted from *GLCM* matrixes. The kernel function used in the proposed CAD system is the spherical kernel because it is an anisotropic stationary kernel and has positive definite in R^3 . Spherical kernel function can be defined as shown in Eq. (22) [17, 18]:

$$k(x, y) = \begin{cases} 1 - \frac{3}{2} \frac{\|x-y\|}{\sigma} + \frac{1}{2} \left(\frac{\|x-y\|}{\sigma}\right)^3 & \text{if } \|x-y\| < \sigma \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

Fig. 2 illustrates the ensemble classification method used by the proposed CAD scheme. As shown in the figure, the extracted feature vectors ($\overrightarrow{Fv1} \dots \overrightarrow{Fv9}$) are fed to the SVM classifiers and then the decisions (c_1, \dots, c_9) are fused using the weighted voting function. The final prediction result (y) is obtained as a final decision result

3. Experimental results and analysis

A dataset contains 51 chest X-Ray images obtained from Kaggle website is used for system evaluation purposes. The dataset contains two classes named positive (with COVID-19) and

negative (non-COVID-19). The positive class includes 39 X-Ray images collected from real cases in China, Korea, the US, Canada, and Taiwan. On the other hand, the negative class includes 12 X-Ray images for patients suffering from MERS, SARS, and ARDS [19]

The proposed CAD scheme has nine tunable parameters: population size N , learning factors $\alpha 1$ and $\alpha 2$, inertia factor w , maximum velocity $Maxv, Lis'$ weight balance β , number of iterations N , erosion structure element size SE , SVM kernel function sigma σ . Based on the extensive experiments, N was obtained as 10, $\alpha 1$ and $\alpha 2$ are determined as 2, the w value was chosen as 1, $Maxv$ was determined as 10, β was tuned as 0.5, N was chosen as 10, the best SE was 3×3 and the σ value was chosen as 1.5.

Table 1 shows obtained results with the proposed automatic segmentation methods when applied on ten different X-Ray chest images (five images for COVID-19 case and five images for non-COVID-19 case). The table also demonstrates the optimal threshold value selected by the proposed segmentation method below each image in the segmentation result column.

As shown in Table 1, the optimal threshold value is varied from case to case based on the distribution of grey level values and degree of transparency between the object (chest region) and the background. GLMC features are then extracted from the results of Laws' masks of the segmented chest image. Fig. 3 illustrates the application results of the nine different Laws' masks on an example chest image.

GLCM features are extracted concerning different d and ϑ values and the ensemble SVMs classifiers are then trained on the extracted features. Each classification model's accuracy is calculated using Eq. (23) [21]. Tables 2- 4 show the accuracy acc., which is ca achieved for the nine different weak SVMs models in addition to ensemble fusion results. As seen in the tables, the best classification accuracy (98.04%) is obtained with the proposed method when $d=2$ and $\vartheta = 45^\circ$. The second-best performance is (96.08%) obtained when $d=2$ and $\vartheta=135^\circ$.

$$Acc. = \frac{Corr.}{N} \times 100 \quad (23)$$

Where N is the total samples and cor. is the corrected classified samples.

Table 1. Segmentation step results for different chest X-ray cases

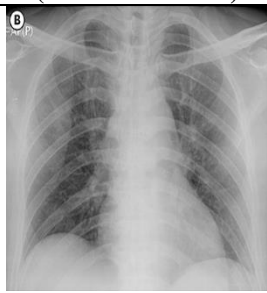

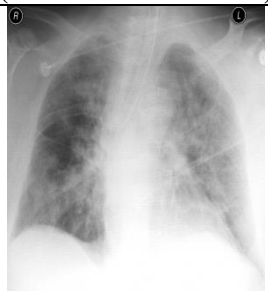
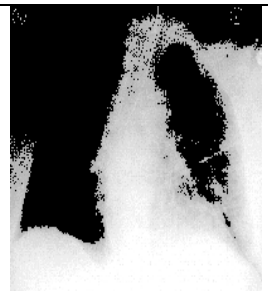
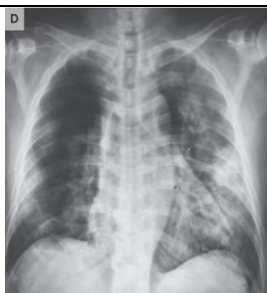
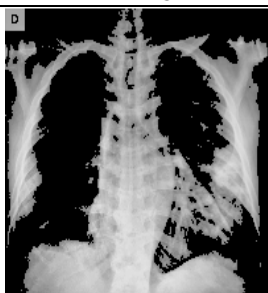

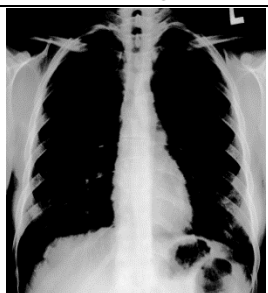

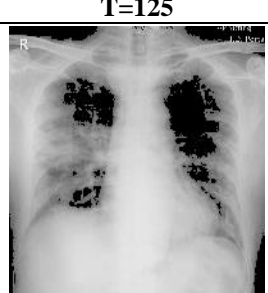


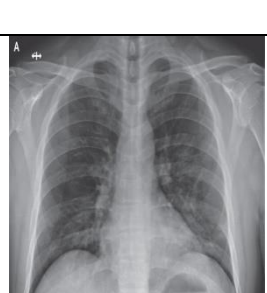
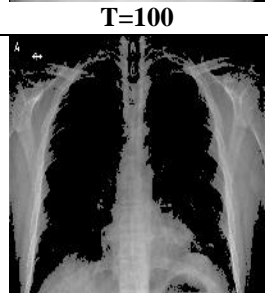

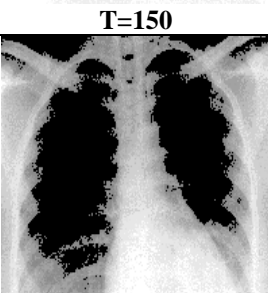
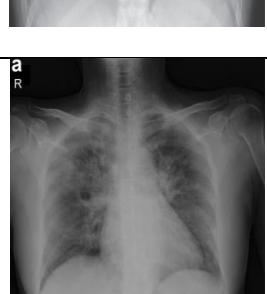
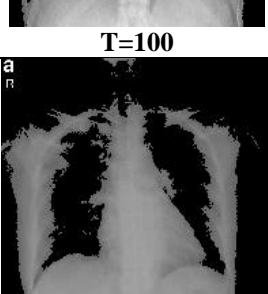


Chest X-Ray Image (COVID-19 Case)	Segmentation Result	Chest X-Ray Image (Non- COVID-19 Case)	Segmentation Result
	 T=128		 T=175
	 T=125		 T=128
	 T=100		 T=150
	 T=100		 T=128
	 T=99.11		 T=150

Table 2. Ensemble classification results for GLCM features with d=1

Classifier	$\vartheta = 135^\circ$	$\vartheta = 90^\circ$	$\vartheta = 45^\circ$	$\vartheta = 0^\circ$
SVM1	92.16	96.08	94.12	94.12
SVM2	92.16	88.24	94.12	90.20
SVM3	92.16	90.20	96.08	92.16
SVM4	92.16	88.24	94.12	90.20
SVM5	92.16	94.12	94.12	90.20
SVM6	92.16	94.12	90.20	92.16
SVM7	90.20	90.20	94.12	90.20
SVM8	92.16	90.20	96.08	92.16
SVM9	90.20	94.12	94.12	90.20
Weighted voting Result	92.16	94.12	98.04	96.08

Table 3. Ensemble classification results for GLCM features with d=2

Classifier	$\vartheta = 135^\circ$	$\vartheta = 90^\circ$	$\vartheta = 45^\circ$	$\vartheta = 0^\circ$
SVM1	94.12	94.12	90.20	94.12
SVM2	90.20	94.12	88.24	94.12
SVM3	94.12	92.16	88.24	94.12
SVM4	92.16	94.12	88.24	94.12
SVM5	88.24	92.16	90.20	94.12
SVM6	94.12	90.20	94.12	94.12
SVM7	92.16	90.20	94.12	94.12
SVM8	94.12	90.20	94.12	94.12
SVM9	92.16	88.24	94.12	94.12
Weighted voting Result	96.08	94.12	90.20	94.12

Table 4. Ensemble classification results for GLCM features with d=3

Classifier	$\vartheta = 135^\circ$	$\vartheta = 90^\circ$	$\vartheta = 45^\circ$	$\vartheta = 0^\circ$
SVM1	82.54	87.16	94.11	90.20
SVM2	87.40	88.46	94.11	88.24
SVM3	84.13	90.33	92.15	82.94
SVM4	83.14	85.84	94.11	86.27
SVM5	85.32	85.54	92.15	88.24
SVM6	88.90	85.54	94.11	88.24
SVM7	87.78	85.54	92.15	94.12
SVM8	85.32	85.54	94.11	94.12
SVM9	88.54	85.54	92.15	94.12
Weighted voting Result	91.67	89.68	95.11	90.20

To illustrate the effectiveness of using PSO for optimal threshold selection during the segmentation process, a comparison has been made between the

accuracy achieved using PSO and without using PSO as shown in Table 5. The optimal settings for d and ϑ with ensemble classification of the generated GLCM features (i.e., Tables 2, 3, and 4) are obtained and the experiments are conducted again with threshold value equals to 128 for all image cases. As clearly shown in the table, the accuracy is degraded with an about 4% when a single threshold value is used for all X-Ray image cases.

The efficiency of the proposed CAD scheme can be further evaluated by measuring sensitivity and specificity values. Sensitivity is the percentage of positive samples that are correctly labeled, while specificity is the percentage of negative data samples that are labeled correctly. These measures can be defined as in the following equations [20]:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100 \tag{24}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100 \tag{25}$$

where TP refers to the number of true positive samples, TN refers to the number of samples which are true negative, FP refers to the number of false positives samples, and FN is the number of false negatives samples. The calculated specificity and sensitivity values for the cases that give the best accuracy within different d and ϑ combinations are given in Table 6.

Table 5. Achieved accuracy when using PSO for threshold determination and without using PSO (with threshold =128)

Case	Using PSO	Without using PSO
d=1, $\vartheta = 45^\circ$	98.04	94.12
d=2, $\vartheta = 135^\circ$	96.08	92.16
d=3, $\vartheta = 45^\circ$	95.11	88.24

Table 6. The calculated sensitivity and specificity values

Case	Sensitivity	Specificity
d=1, $\vartheta = 45^\circ$	100.00%	91.67%
d=2, $\vartheta = 135^\circ$	100.00%	83.33%
d=3, $\vartheta = 45^\circ$	100.00%	75.00%

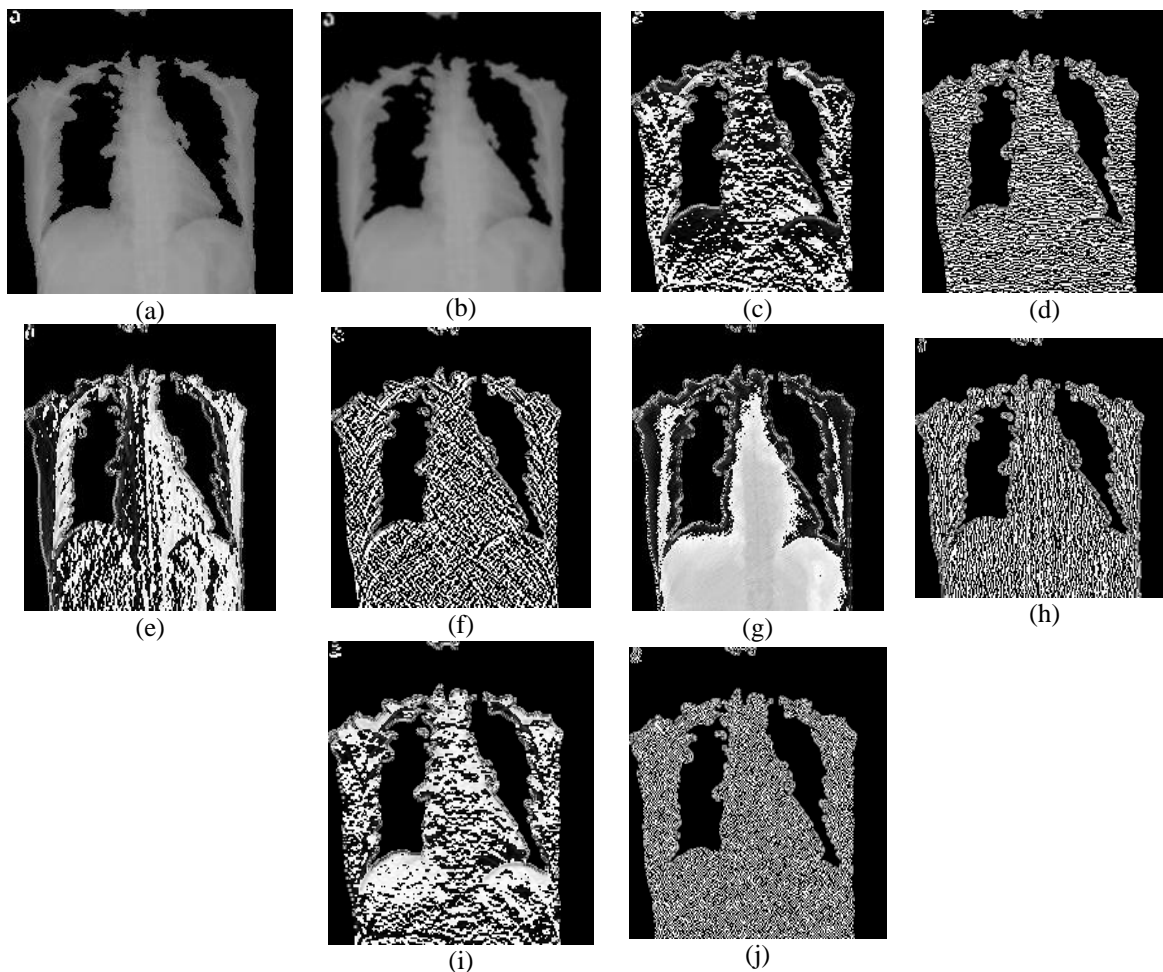


Figure. 3 Results of application Laws' masks: (a) Original segmented chest image, (b) L3L3, (c) L3E3, (d)L3S3, (e) E3L3, (f) E3E3, (g) E3S3, (h) S3L3, (i) S3E3, and (j) S3S3

Since the work on this research problem is started in a time not far from now and the used dataset is newly collected; There are few works to compare with. Table 7 shows a comparison between the proposed method and another study that used the same dataset.

Table 7. Comparison with other studies

Study	Aim	Accuracy
[5]	The study has two parts. First, it detects the only true COVID-19 from COVID-19 samples. Second, it detects the true SARS from the true SARS samples.	Both Avg. = 96%
Proposed	identifying COVID-19 from other MERS, SARS, and ARDS viral pneumonia.	98.04%

4. Conclusions and future work

CAD system for COVID-19 has been presented in this paper based on the CT scan of the chest region. Particle swarm intelligence helps in achieving the automatic segmentation regardless of different grey-scale distribution among different chest X-Ray images. Laws' masks and GLCM based features showed an effective role in the extraction of the regular patterns in the segmented chest image. According to the experimental outcomes, the suggested CAD tool is effective in classifying the chest X-Ray image into either positive or negative cases with classification accuracy more than 98% when $d=1$ and $\vartheta=45^\circ$. As future work, different transforms such as radon, wavelets, discrete cosine transform can be used to extract additional features from the texture of the chest image. Also, other fusion methods can be utilized such as bagging and boosting which in turn may result in enhancing the diagnostic accuracy.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, Suhaila N. Mohammed, Fatin Sadiq Alkinani, and Yasmin A. Hassan; methodology and implementation, Suhaila N. Mohammed; writing—original draft preparation, Suhaila N. Mohammed and Fatin S. Alkinani; writing—review and editing, Fatin Sadiq Alkinani; supervision and funding acquisition, Yasmin A. Hassan.

References

- [1] Y. Li and L. Xia, “Coronavirus Disease 2019 (COVID-19): Role of Chest CT in Diagnosis and Management”, *American Journal of Roentgenology*, pp. 1-7, 2020. doi: doi.org/10.2214/AJR.20.22954
- [2] S. Kooraki, M. Hosseiny, L. Myers, and A. Gholamrezanezhad, “Coronavirus (COVID-19) Outbreak: What the Department of Radiology Should Know”, *Journal of the American College of Radiology*, Vol. 17, No. 4, pp. 447-451, 2020.
- [3] E. Lee, M. Ng, and P. Khong, “COVID-19 pneumonia: what has CT taught us?”, *The Lancet Infectious Diseases*, Vol. 20, No. 4, pp. 384-385, 2020.
- [4] F. Song, N. Shi, F. Shan, Z. Zhang, J. Shen, H. Lu, Y. Ling, Y. Jiang, and Y. Shi, “Emerging 2019 Novel Coronavirus (2019-nCoV) Pneumonia”, *Radiology*, Vol. 295, No. 1, pp. 210-217, 2020.
- [5] A. Farid, G. Selim, and H. Khater, “A Novel Approach of CT Images Feature Analysis and Prediction to Screen for Corona Virus Disease (COVID-19)”, *International Journal of Scientific and Engineering Research*, Vol. 3, No. 1, pp. 1141-1148, 2020.
- [6] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Y. Chen, J. Su, G. Lang, Y. Li, H. Zhao, K. Xu, L. Ruan, and W. Wu, “Deep Learning System to Screen Coronavirus Disease 2019 Pneumonia”, *arXiv:2002.09334 [physics.med-ph]*, 2020.
- [7] S. Mohammed and L. George, “Illumination - Invariant Facial Components Extraction Using Adaptive Contrast Enhancement Methods”, *British Journal of Applied Science & Technology*, Vol. 12, No. 3, pp. 1-13, 2016.
- [8] Z. Li, C. Liu, G. Liu, Y. Cheng, X. Yang, and C. Zhao, “A Novel Statistical Image Thresholding Method”, *International Journal of Electronic Communication (AEU)*, Vol. 64, No. 12, pp. 1137-1147, 2010.
- [9] E. Chehade, R. Abdel-Kader, and A. El-Zaart, “Segmentation of MRI Images for Brain Cancer Detection”, In: *Proc. of International Conference on Information and Communications Technology (ICOIACT)*, Yogyakarta, Indonesia, pp. 929-934, 2018.
- [10] K. Sreedhar and B. Panlal, “Enhancement of Images Using Morphological Transformations”, *International Journal of Computer Science & Information Technology (IJCSIT)*, Vol. 4, No. 1, pp. 33-50, 2012.
- [11] R. Shenbagavalli and K. Ramar, “Classification of Soil Textures Based on Law’s Features Extracted from Preprocessing Images on Sequential and Random Windows”, *Bonfring International Journal of Advances in Image Processing*, Vol. 1, Special Issue, pp. 15-18, 2011.
- [12] S. Dash and U. Jena., “Texture Classification Using Laws’ Filter in Various Color Spaces”. *International Journal of Engineering and Technology (IJET)*, Vol. 9, No. 2, pp. 454-460, 2017.
- [13] S. Dash and U. Jena, “Gaussian pyramid based laws' mask descriptor for texture classification”, In: *Proc. of International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, Chennai, India, pp. 654-658, 2017.
- [14] P. Kumar and S. Chatterjee, “Computer Aided Diagnostic for Cancer Detection Using MRI Images of Brain”, In: *Proc. of IEEE Annual India Conference (INDICON)*, Bangalore, India, pp.1-6, 2017.
- [15] A. Goshtasby, *Image Registration Principles, Tools and Methods*, Springer-Verlag, London, 2012.
- [16] T. Hastie, R. Tibshirani, and J. Friedman, *Data Mining, Inference, and Prediction*, 2nd edition, Springer, Stanford, California, 2001.
- [17] M. Somvanshi and P. Chavan, “A Review of Machine Learning Techniques Using Decision Tree and Support Vector Machine”, In: *Proc. of 2016 International Conference on Computing Communication Control and Automation (ICCUBEA)*, Pune, pp. 1-7, 2016.
- [18] S. Boughorbel, P. Tarel, F. Fleuret, and N. Boujemaa, “The GCS Kernel for SVM-Based Image Recognition”, In: *Proc. of Duch W., Kacprzyk J., Oja E., Zadrozny S. (eds) Artificial Neural Networks: Formal Models and Their Applications – ICANN 2005*, Lecture Notes in

Computer Science, Vol. 3697. Springer, Berlin, Heidelberg, 2005.

- [19] Chest X-Ray database Website, <https://www.kaggle.com/bachrr/covid-chest-xray>
- [20] H. Jiawei and P. Jian, *Data Mining Concepts and Techniques*, 3rd edition, Elsevier, USA, 2012.
- [21] A. Hassan and S. Mohammed, "A novel facial emotion recognition scheme based on graph mining", *Defence Technology*, 2019. doi: 10.1016/j.dt.2019.12.006