ORIGINAL ARTICLE

Artificial Intelligence Assisted Detection of Respiratory Infectious Diseases Signs from Computed Tomography Images

Faezeh Shalbafzadeh [1, 2], Fatemeh Taherpour-Dizaji [1, 2], Mohammad Reza Fouladi [1, 2], Sharareh Baradaran [3], Matin Ghadiri [4], Hossein Ghadiri [1, 2*]

[1] Department of Medical Physics and Biomedical Engineering, School of Medicine, Tehran University of Medical Sciences, Tehran, Iran

[2] Research Center for Molecular and Cellular Imaging, Advanced Medical Technologies and Equipment Institute, Tehran University of Medical Sciences, Tehran, Iran

[3] Department of Radiology, West Nikan Hospital, Tehran, Iran

[4] School of Electrical and Computer Engineering, University of Tehran, Tehran, Iran

*Corresponding Author: Hossein Ghadiri

Received: 14 March 2025 / Accepted: 15 July 2025

Email: h-ghadiri@sina.tums.ac.ir

Abstract

Purpose: Respiratory infectious diseases often manifest as Ground-Glass Opacity (GGO) or consolidation signs in the lungs. Artificial Intelligence (AI)-assisted systems utilizing data mining algorithms such as Waikato Environment for Knowledge Analysis (Weka) can be used for the detection and segmentation of these signs. In this study, we propose using Weka as a comprehensive data mining and machine learning tool to develop the most accurate models for detecting lung signs in chest CT images of patients with respiratory infectious diseases.

Materials and Methods: First, we manually selected specific signs from chest Computed Tomography (CT) images from 600 cases using the Graphical User Interface (GUI) Weka plugin. We then trained the random forest algorithm based on different features and presented the best combined model obtained for the automatic detection of the aforementioned signs. Lastly, the model's performance was evaluated with different metrics.

Results: Our findings indicate that the hybrid texture description features, including “Structure”, “Entropy”, “Maximum”, “Anisotropic”, and “Laplacian” available in Weka, demonstrated the lowest Out-of-Bag (OOB) error rate, highest Area Under the ROC Curve (AUC) value of 0.992, and accuracy of 98.1%.

Conclusion: By leveraging the combination of Weka features, we have successfully developed models for the detection and segmentation of lung signs associated with infectious diseases, from chest CT images. These findings contribute to the field of medical image analysis and hold promise for improving the diagnosis and treatment outcomes of patients with respiratory infectious disorders.

Keywords: Respiratory Infectious Diseases; Artificial Intelligence; Random Forest Algorithm; Weka; Chest Computed Tomography Images.

DOI: https://doi.org/10.18502/fbt.v13i1.20336

1. Introduction

Lung disease encompasses a variety of disorders affecting the lungs, including asthma, Chronic Obstructive Pulmonary Disease (COPD), infectious diseases such as influenza, pneumonia, and tuberculosis, as well as lung cancer and other respiratory problems. Lung infections are the number one cause of transmissible deaths worldwide and rank as the third leading cause of death globally. These infections affect all age groups, particularly the young, elderly, and chronically ill. Globally, in 2019, there were 488.9 million incident cases and 2.4 million deaths due to lower respiratory infections [1].

In addition to known clinical presentations, the imaging or radiological features of infectious respiratory diseases include Ground-Glass Opacity (GGO) and consolidation signs. GGO is a radiological term indicating an area of hazy, increased lung opacity through which vessels and bronchial structures may still be seen. It is less opaque than consolidation, in which such structures are obscured [2]. If the signs caused by respiratory diseases in the lungs are detected, and the severity of the disease is determined, it is possible to determine how much the disease has progressed or in what phase it is in which phase it is.

Given the pandemic nature of infectious respiratory diseases and their global impact in recent years, along with the abundance of available data, this study focuses on analyzing chest CT images of COVID-19 patients. In terms of imaging characteristics, COVID-19 is one of the most infectious diseases of the lung, and in its advanced stages, GGO and consolidation appear [3]. However, the results and insights gained from this research can potentially be applied to all infectious lung diseases.

Artificial Intelligence (AI) techniques have gained significant attention for the diagnosis of contagious respiratory diseases, particularly in analyzing chest CT images. The majority of studies in this area investigate two main aspects: Machine Learning (ML) and Deep Learning (DL). DL methods, such as Convolutional Neural Networks (CNNs), have been widely explored for feature extraction from images. For instance, a study conducted in 2021 proposed an ensemble ResNet_NRC model for COVID-19 detection on CT images, utilizing the relative majority voting method to merge individual classifiers [4]. Moreover, several ML methods have been proposed for sign detection and classification using chest CT images. For example, researchers used the random forest model in the machine learning method to realize automatic severity assessment of COVID-19 based on chest CT images [5]. However, these existing algorithms have certain limitations. DL algorithms often lack interpretability for features that may not contribute significantly to performance, require substantial amounts of data for effective training, and demand high computational resources and preprocessing steps. Machine learning algorithms, on the other hand, rely on accurate ground truth images for optimization and often have limited availability [6].

To address these limitations and develop a more robust and accessible AI-assisted diagnosis system, this study uses the Weka tool. Weka provides a comprehensive set of visualization tools and algorithms for data analysis and forecasting, which, with the help of a user-friendly graphical user interface (GUI), facilitates easy access to these functions. With its proficiency in data mining, including image processing, clustering, segmentation, and regression, Weka offers a promising framework for developing accurate models for detecting lung signs in chest CT images of people with respiratory infectious diseases [7].

For example, Shankarlal et al implemented Weka to classify COVID-19 chest CT images using a multilayer perceptron with various imaging filters (edge histogram, histogram equalization, etc.) and demonstrated an accuracy of approximately 89.6%. This confirms Weka's value in medical image classification tasks and supports its suitability for our approach [8].

The contributions of this study can be summarized as follows: (I) manual selection of specific signs or pneumonia lesions associated with contagious respiratory disorders from chest CT images of patients, (II) training a random forest algorithm using various combinations of features available in the Weka plugin and exploring different combinations, and (III) selecting the best model for detecting infectious respiratory diseases' signs and evaluating its performance against ground truth images.

In this study, we proposed an AI-assisted approach for detecting radiological signs of infectious diseases in the lungs, with a particular focus on ground-glass opacity (GGO) and consolidation patterns observed in CT images. Although the initial dataset involves COVID-19 cases, the methodology was not limited to this disease and can be generalized to other respiratory infections such as influenza and pneumonia, which remain prevalent and clinically significant worldwide. The novelty of our work lies in the automated identification of these key radiological signs (GGO and consolidation), which are critical yet underutilized features in many current AI-based diagnostic models. Furthermore, we explored different combinations of feature selection techniques and classification algorithms using the Weka platform, providing a comparative evaluation that contributed new insights to the field. We believe this approach has the potential to lay the groundwork for future diagnostic systems focused on radiological sign recognition.

2. Materials and Methods

2.1. Data Collection

A total of 1440 CT images of 600 COVID-19 patients were retrospectively collected from the Picture Archiving and Communication System (PACS) of the radiology wards during August and September 2020. Initially, the collected images were divided into two groups based on sex: male and female. The average age for male and female patients was 50 and 48 years, respectively. The study was approved by the ethics committee of our institute, and individual consent for this retrospective analysis was waived. Each CT image collected was anonymized before its use in this study (Approval number: IR.TUMS.MEDICINE.REC.1399.471).

The selected images were categorized based on specific signs associated with GGO, consolidations, and both signs. We considered 400 images for each group, with an equal distribution of male and female subjects (200 females and 200 males). Each category was then divided into two subsets: a training set comprising 60% of the images, and a test set comprising the remaining 40%. (240 images for the training and 160 images for the testing in each category). Therefore, a total of 1200 images belong to the GGO, consolidation, and both sign groups. Additionally, a separate category was created for images without any signs (referred to as "Normal" images). A total of 240 images of normal images are considered for testing our models. The patient grouping flowchart used in this study is shown in Figure 1.

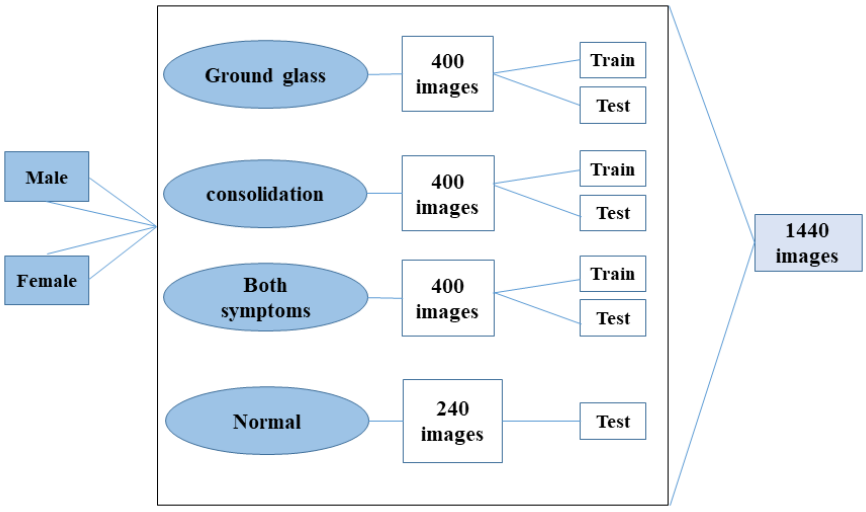

Figure 1. Flowchart categorizing the images used in this study, a total of 1440 images are used in this study. 400 images for each category of GGO, consolidation, and both symptoms which are divided into train and test groups. Also, 240 images of the Normal category are used for testing

CT images were acquired using a 16-slice CT scanner. The patients were placed in the Head-First Supine (HFS) position, and the gantry rotation mode was set to helical. The CT protocol employed for image acquisition included the following parameters: 120 kV, spacing of 6 mm between slices, and slice thickness of 5 mm. Images were obtained with the patients in a breath-hold state at full inspiration, and no intravenous contrast injection was administered. The acquired images had a matrix size of 512 × 512 pixels. The scanning range covered the entire span, from the apex to the base of the lungs. For image interpretation, two different window settings were utilized: a mediastinal window with a window width of 400 Hounsfield Units (HU) and a window level of 40 HU, and a lung window with a window width of 1500 HU and a window level of -700 HU.

For this study, CT images of the lungs were reviewed by two radiologist with over 10 years of experience in radiology. The radiologists were blinded to the test being conducted. Their role was to identify and delineate the presence or absence of pneumonia lesions associated with signs of respiratory infectious diseases. The radiological features considered in this study were GGO and consolidations, which are commonly observed in CT images of patients with respiratory infectious diseases. In the final stage of this study, these labeled images were used as the ground-truth images. They were used as a reference to compare with the images obtained by applying models to the test images.

2.2. Training Procedure

2.2.1. Random Forest Algorithm Training

In this section, we describe the training process of the random forest algorithm. The random forest algorithm is a machine learning technique known for its simplicity and robustness, enabling automatic and efficient classification of samples from large datasets. This algorithm mitigates overfitting by randomly selecting and inputting data into training trees, and then combining the outputs of numerous random trees into the best classifier [9]. For each gender group (male and female) and within each group, specifically for all three categories of GGO, consolidation, and both signs, 400 chest CT images were selected. Among these images, 240 were designated as training images, while the remaining 160 were set aside as test images.

Subsequently, all the training images were traced using the TWS plugin available in the ImageJ (Fiji) software and the Weka graphical environment [10]. The radiologists manually selected the boundaries of signs in the training images, and each distinct part was assigned to a specific categories as follows:

1. Class 1: Regions corresponding to the signs of the disease, GGO, or consolidations.
2. Class 2: The normal lung tissue.
3. Class 3: Anatomical structures outside the lung tissue.

For images exhibiting GGO or consolidation signs, three classes were considered, while the images showing both signs were assigned to four classes due to the presence of both signs simultaneously and is as follows:

1. Class 1: Regions corresponding to the sign GGO.
2. Class 2: Regions corresponding to the sign consolidation.
3. Class 3: The normal lung tissue.
4. Class 4: Anatomical structures outside the lung tissue.

2.2.2. Feature Selection and Extraction

In the second stage, after tracing all training images, the random forest algorithm implemented in the Weka plugin was trained using various features. The 19 features investigated in this study using the Weka plugin include the following: Mean, Maximum, Minimum, Variance, Kuwahara linear, Bilateral, Gaussian Blur, Anisotropic diffusion (Perona-Malik diffusion), Sobel operator, Lipschitz continuity, Second order Derivatives, Discrete Laplacian, Structure, Difference of Gaussian, Membrane projections, Hessian matrix, Local Entropy, Neighbors (Local Binary Patterns). These Weka features can be categorized into three groups: edge detection, texture description, and noise removal features [11]. During the algorithm training, all features were utilized, forming different combinations of one, two, three, four, and five features. Features that

did not demonstrate satisfactory performance were removed from consideration. All satisfactory combinations of features were saved as a model or a classifier. Figure 2 schematically illustrates all steps involved in feature selection and algorithm training.

The selection of the best model involved comparing the Out-of-Bag (OOB) error rates associated with different feature combinations. The feature combinations with the lowest error rates were considered the most effective. After evaluating various combinations and comparing their OOB error rates, the best model for all groups was determined based on the five-feature combination with the lowest OOB error rate.

2.2.3. Detection of Signs

The third step involved applying the models obtained from the training stage to the test images. The best models obtained for each group were applied to the test images within that respective group, generating maps. When applying the models to the test images, the Weka plugin offers two types of images: segmentation images and map images [4].

In the context of this study, the selected map images and the trained algorithm demonstrated the probability of signs in different lung regions, represented as grayscale images with values ranging from 0 to 1. For each of the classes regarding the boundaries of signs, one map image was created. A sample result is illustrated in Figure 3. The selected map images, along with the trained algorithm, provide a visual representation of the likelihood of signs in specific lung regions.

2.3. Positive (Diseased) Criteria

In this stage, a post-processing algorithm in the ImageJ software was employed to apply the minimum threshold method for binarization and evaluate the best models. The binarization method served three purposes in this study:

2.3.1. Diseased or Normal Classification

Ten map images that exhibited errors in identifying GGO and consolidation signs were selected. This process generated mask images that were represented in black and white (false=0 and true=1). Dilation was then applied to the mask images to enhance the boundary separation of signs, particularly for consolidation signs located in the peripheral lung areas. The number of true pixels (N_T) in the affected areas was calculated using a histogram, and the corresponding mean (M) and Standard Deviation (SD) were calculated as well. Finally, for the remaining test images, diseased or normal classes were determined using the condition shown in Equation 1 [12]:

$$if\ \left(N_T - (M + SD)\right) > 0\ \ Class = Positive \qquad (1)$$

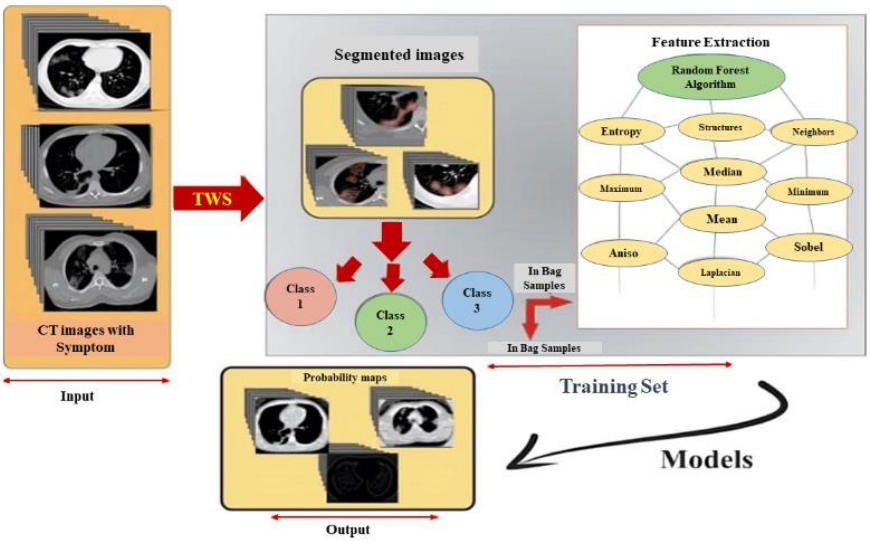

Figure 2. Proposed framework for the selecting features and training images steps. 120 training images were traced, and the random forest algorithm implemented in the Weka plugin was trained using various features. A total of 19 Weka features, including edge detection, texture description, and noise removal features, were investigated and utilized in different combinations to form satisfactory models or classifiers

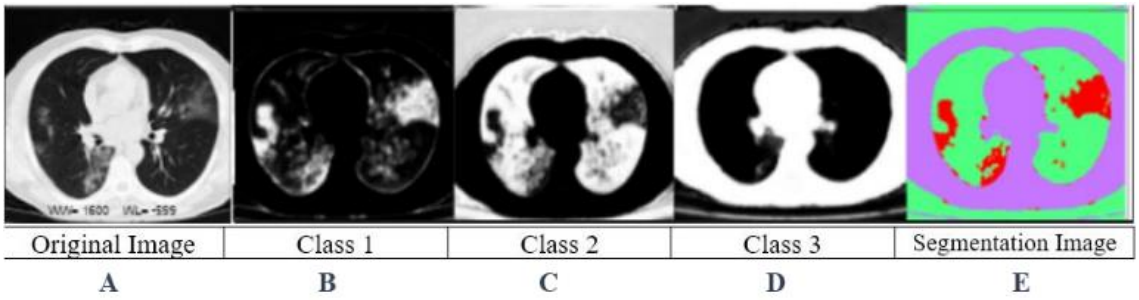

Figure 3. Images obtained using the best GGO models. (A) Original chest CT images with GGO signs. (B) Probability maps for Class 1. (C) Probability maps for Class 2. (D) Probability maps for Class 3. (E) A colored segmentation image

2.3.2. Calculation of CT Severity Score (CT-SS)

For each group, the mask image obtained using the method described previously was used. In Class 1 and 2 (images with one sign) and Class 3 (images with both signs), the number of true pixels in the remaining lung tissue was calculated using the histogram and list options. The percentage of lung tissue involvement, representing the severity score, was determined by dividing the number of pixels involved by the total number of pixels in the lung tissue in each slice, according to Equation 2:

$$\begin{aligned} &\textit{everity Score}\ (\%) \\ &= \frac{N_T\ \textit{in sign classes}}{\textit{Total number of pixels (including signs \& healthy)}} \\ &\times 100 \end{aligned} \qquad (2)$$

Indexing this parameter in the clinic can vary depending on its definition [13] [14].

2.3.3. Calculation of the Confusion Matrix

The third application of binarization was to calculate the Confusion Matrix. The following steps were performed:

- Ground truth images were converted from 16-bit to 8-bit using the Image/Type menu in ImageJ, ensuring the gray levels ranged from 0 to 255.
- A subtraction of 10 pixels was applied to the original image using the Process/Math/Subtract menu in ImageJ, ensuring that the highest pixel value in the images is 255.

Using the area selection tools, the desired signs of the disease were determined. Then, through the Analyze/Tool/ROI manager menu, all the selected areas were saved and filled with red color.

- Using the Flatten option in ImageJ, the image with red areas representing disease signs turned into an RGB image. This RGB image had three channels, each with pixel values ranging from 0 to 255. Due to the 10-pixel subtraction, the maximum pixel value was limited to 245. The flatten option completely fixes all the selected areas in the image layers, allowing the layer corresponding to the red color to be separated from the rest of the image layers.
- Using the Image/Color/Split Channel menu, different color channels in the image were separated, and color channels corresponding to the blue and green were eliminated, retaining only the red color channel for further steps.
- Using the Image/Adjust menu, the image related to the red color channel was converted into a binary image with the help of the default thresholding, which ensured that only the areas related to the sign were present, while all other areas were removed.
- Map images regarding the related ground truth image were chosen, and the classes associated with the sign were also converted into masks and binary images.
- Using the Process/Image calculator menu and the Add option, the mask images were added to the binary ground truth images, which led to the

segmented areas of both images overlapping each other. It should also be noted that, at this step, the final image is in 32-bit format.

- The histogram of the final image was generated using the Analyze/Histogram menu. By selecting the List option, the number of pixels for each value was calculated. The probability metrics were then determined based on these pixel counts. The two images have pixel values ranging from 0 to 255 and 0 to 245, respectively. When these images are added, the areas correctly representing the disease signs have a combined value of 500.

Figure 4 illustrates the method for calculating the Confusion Matrix. It uses binarized mask images and defines true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values.

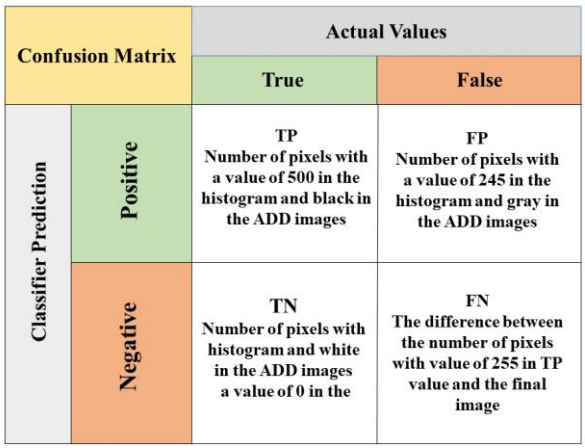

Figure 4. The method of calculating the Confusion Matrix using the binarization of mask images. The definition of TP, TN, FP, and FN is described in this figure

2.4. Performance Parameters

The predictive model in this study is based on the AI random forest algorithm. The model's performance was evaluated using various metrics. The primary metric for assessing feature performance was the OOB error, which is an unbiased estimator of misclassification errors and serves as an alternative to cross-validation. It approximates the number of samples excluded during the creation of each decision tree [15].

Additional evaluation metrics used in this study for supervised learning include the following:

- **TP:** Number of pixels with a value of 500 in the histogram and the color black in the ADD images
- **FP:** Number of pixels with a value of 245 in the histogram and the color gray in the ADD images
- **TN:** Number of pixels with a value of 0 in the histogram and the color white in the ADD images
- **FN:** The difference between the number of pixels with a value of 255 in the TP value and the final image
- **Precision:** Precision measures the proportion of correctly predicted positive instances by the random forest classifier out of all instances predicted as positive.
- **Recall:** Also known as sensitivity, it indicates the percentage of positive patients correctly identified by the random forest classifier.
- **F-Measure:** A metric that combines the metrics precision and recall to provide an overall measure of a classifier's performance.
- **Accuracy**: A metric that shows how often the Random Forest model correctly predicts the outcome.
- **Receiver-Operating Characteristic (ROC) area:** ROC curves are used to evaluate the performance of classification algorithms. The ROC space is formed by two indicators: False Positive Rate (FPR) on the horizontal axis and True Positive Rate (TPR) on the vertical axis. The area under the ROC curve, known as the AUC, is considered a crucial validation method in this study.

Classifier performance was evaluated using 10-fold cross-validation. The data were divided into ten equal parts. The model was trained on nine parts and tested on the remaining part. This process was repeated ten times, so that each part served as the test set once.

3. Results

In this study, we evaluated the performance of the random forest algorithm using various features in the

Weka plugin. By combining these features, we aimed to identify the best combination that would yield the lowest OOB error rate and the highest AUC.

3.1. Effect of the Number of Combinations of Features

Tables 1 and 2 demonstrate the OOB error and AUC values for different combinations of features, ranging from one to five, in both male and female groups. The AUC values were calculated using the Weka plugin. It's important to note that the AUC value was calculated for all classes of the models, but only the average values for the classes are provided in these tables (Table 1 and Table 2). Within each group of positive (diseased) patients, which includes GGO, consolidation, and both signs, we selected and compared five combinations of features. The combinations that are determined in Tables 1 and 2 represent the best models obtained for each group. All the best models consisted of five features. According to the results from Table 1, the best model for male patients with both signs consists of a combination of features found in the GGO and consolidation models. The features "Structures, Median, Maximum, and Laplacian" are the same in both the GGO and consolidation models, with the only difference being the inclusion of the "Entropy and Anisotropic" features. The model with GGO alone has the lowest OOB error rate (2.65%), while the model with both signs has the highest error rate (5.69%). Similarly, Table 2 shows that the best models for GGO and consolidation have the same features, including "Structures, Entropy, Maximum, and Laplacian," with the only difference being the "Minimum and Median" features. Out of all the models, the combination of features from both groups in the best model for female patients with both signs, considering the removal of the minimum feature, results in the lowest OOB error rate (8.11%). ¬

During the final stage of algorithm training, we used a combination of male and female images, a total of 240 images (twice the size of the initial training data), to reduce the detection error rate of the models. This training was specifically conducted on images exhibiting consolidation and both signs (GGO and consolidation). The results obtained from this stage are presented in Table 3. However, due to the time-consuming nature of the training algorithm and to expedite the process, the models obtained at this stage only include combinations of the five features that were selected earlier. Additionally, based on Tables 1 to 3, it can be concluded that increasing the number of features initially leads to an increase in AUC up to a certain plateau value. Beyond this maximum value, further increasing the number of features has minimal impact on the OOB error rate and does not change the AUC value significantly. The choice of five features for the best models was made to maintain a constant AUC value. All AUC values, as shown in Figure 5, range from 0.9 to 1, indicating an excellent model performance [16]. In all of the training instances, increasing the number of features from one to five

Table 1. OOB error and AUC parameters in combination of different features in Male images Str: Structures, Ent: Entropy, Aniso: Anisotropic diffusion, Max: Maximum, Lap: Laplacian, Med: Median, Min: Minimum, ROC: Receiver operating characteristic, OOB: Out Of Bag error

Sign	Features	OOB error %	AUC
GGO	Str	4.34	0.966
	Str & Med	2.96	0.998
	Str & Med & Max	2.74	0.999
	Str & Med & Max & Aniso	2.66	0.999
	Str & Med & Max & Aniso & Lap	2.65	0.999
consolidation	Str	5.85	0.993
	Str & Ent	3.77	0.997
	Str & Ent & Med	3.71	0.997
	Str & Ent & Med & Lap	3.53	0.998
	Str & Ent & Med & Lap & Max	3.14	0.998
Both Signs	Str	8.65	0.998
	Str & Aniso	6.5	0.994
	Str & Aniso & Max	5.83	0.995
	Str & Aniso & Max & Ent	5.72	0.996
	Str & Aniso & Max & Ent & Lap	5.69	0.996

resulted in a decrease in the OOB error rate. However, it should be noted that the training time of the algorithm also increased. This relationship between the number of features and the OOB error rate is depicted in Figure 6.

Table 2. OOB error and AUC parameters in combination of different features in Female images**.** Str: Structures, Ent: Entropy, Aniso: Anisotropic diffusion, Max: Maximum, Lap: Laplacian, Med: Median, Min: Minimum, ROC: Receiver operating characteristic, OOB: Out Of Bag error

Sign	Features	OOB error %	AUC
GGO	Str	7.12	0.965
	Str & Med	6.51	0.993
	Str & Med & Max	5.21	0.999
	Str & Med & Max & Ent	5.11	0.994
	Str & Med & Max & Ent & Lap	5.04	0.994
consolidation	Str	5.85	0.993
	Str & Ent	4.64	0.996
	Str & Ent & Med	3.48	0.997
	Str & Ent & Max & Lap	3.19	0.997
	Str & Ent & Max & Lap & Min	3.16	0.997
Both Signs	Str	12.63	0.985
	Str & Med	9.89	0.987
	Str & Med & Max	8.83	0.989
	Str & Med & Max & Ent	8.39	0.992
	Str & Med & Max & Ent & Lap	8.11	0.992

Table 3. Combination of different features in Male & Female images. Str: Structures, Ent: Entropy, Aniso: Anisotropic diffusion, Max: Maximum, Lap: Laplacian, Med: Median, Min: Minimum, ROC: Receiver operating characteristic, OOB: Out Of Bag error

Sign	Features	OOB error %	AUC
consolidation	Str & Ent & Med & Max & Min	4.52	0.98
	Str & Ent & Lap & Max & Min	4.47	0.98
	Str & Lap & Med & Max & Min	4.4	0.981
	Str & Ent & Med & Max & Lap	11.41	0.923
	Str & Med & Max & Aniso & Lap	11.36	0.919
Both signs	Str & Ent & Max & Lap & Aniso	11.18	0.924
	Str & Ent & Med & Max & Min	4.52	0.98
	Str & Ent & Max & Min & Lap	4.47	0.98
	Str & Med & Max & Min & Lap	4.4	0.981
	Str & Ent & Med & Max & Lap	11.41	0.923

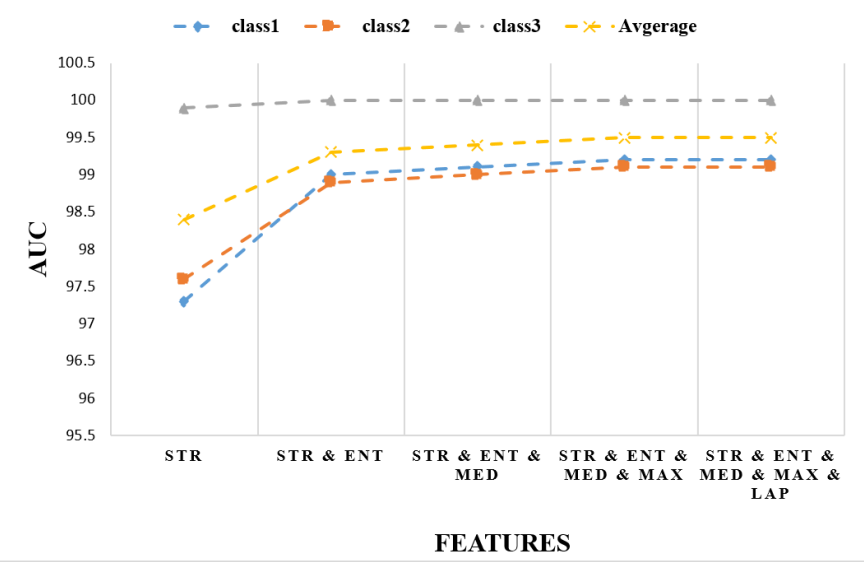

Figure 5. Increasing AUC values with the number of features in the models derived from images of men with GGO symptoms
Str: Structures, Ent: Entropy, Aniso: Anisotropic diffusion, Max: Maximum, Lap: Laplacian, Med: Median, Lap: Laplacian

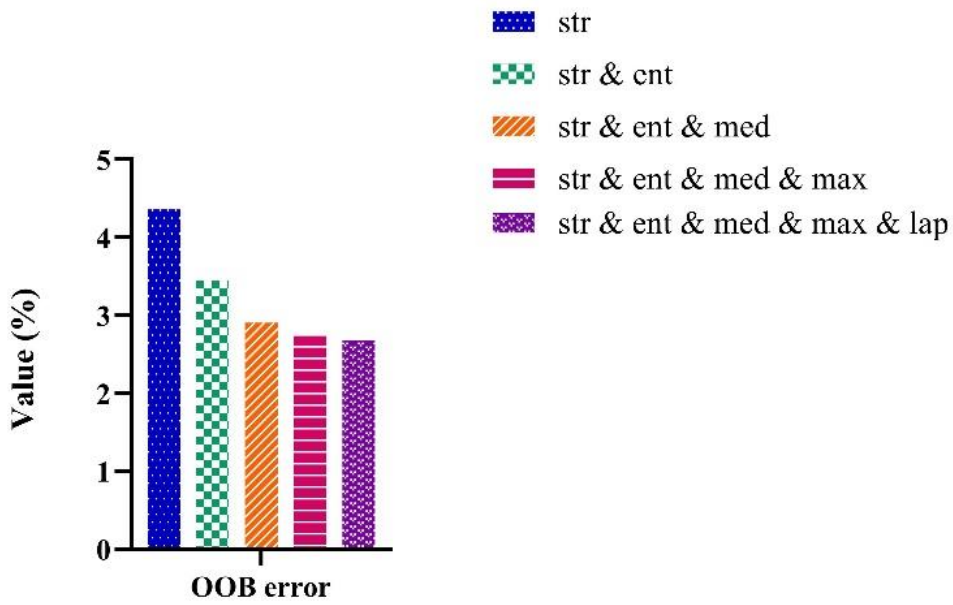

Figure 6. Reducing the OOB error rate while increasing the number of features. Comparing OOB rate of different features. As the number of features increases from 1 to 5, the OOB error rate decreases. Str: Structures, Ent: Entropy, Max: Maximum, Lap: Laplacian, Med: Median

3.2. Balanced Classes

The selection of balanced classes in the Weka plugin produced another key outcome. Table 4 presents the results obtained from this selection. The Class Balancer is a filter that assigns weights to samples, ensuring that each group or class has the same weight while maintaining the total weights of the samples in the dataset. Ideally, each class should have a similar or equal detection rate. Having a balanced dataset enhances the accuracy of the model [17]**.** As Table 4 indicates, selecting the Weka balancing option led to a decrease in the OOB error rate for all models by approximately 43% or one-third of the previous value. It should be pointed out that in the best models trained with twice the training data, the balanced class option was utilized to reduce the error rate.

A consistent finding across all models, regardless of the number of features or gender groups, was that the structure feature consistently had the lowest OOB error rate. This feature is one of the texture features in Weka, which differentiates pixels based on variations in texture content [18].

3.3. Cross-validation Parameters

In the validation phase of the study described in Figure 4, the confusion matrix was used to calculate performance parameters such as accuracy, recall, specificity, precision, and F1-score. These parameters are presented in Table 5.

Table 4. The effect of balanced classes on algorithm training results

Sign	Features	OOB error	
		Class unbalanced	Class balanced
Both signs (female)	**Str & Ent & Med & Max & Lap**	**12.95**	**5.77**
	Str & Ent & Aniso & Max & Lap	**13.21**	**5.76**
GGO (Female)	**Str & Ent & Med & Max & Lap**	5.95	2.54

Based on the results shown in Table 5, all the best models achieved an accuracy of over 94%. The highest accuracy (98.1%) was obtained by the model trained on both signs using twice the training data, i.e., a combination of male and female images. The lowest accuracies were observed in the consolidation sign models trained with 120 images (94.6%) and 240

images (95.1%), as well as in class 2 models with both disease signs (95.4%, 95.1%, 95.7%). Notably, despite having a higher OOB error rate (11.18%), the model trained on both signs with twice the training data demonstrated superior performance in terms of validation accuracy (98.1%). The OOB score is calculated using data excluded from the model analysis and is based on a subset of decision trees without the OOB samples in the bootstrap training dataset. In contrast, the validation score is computed using all the decision trees of the random forest algorithm [19].

Table 5. Classifier performance using 10-fold cross-validation

Best Models		FPR	Specificity	Accuracy
GGO	Male	0.016	0.983	0.97
	Female	0.016	0.983	0.971
Consolidation	Male	0.016	0.977	0.967
	Female	0.031	0.968	0.946
Consolidation	Male & female	0.032	0.967	0.951
Both signs (male)	Class 1	0.008	0.991	0.98
	Class 2	0.012	0.987	0.973
Both signs (female)	Class 1	0.012	0.987	0.973
	Class 2	0.022	0.977	0.951
Both signs (male & female)	Class 1	0.022	0.997	0.981
	Class 2	0.013	0.986	0.957

3.4. Evaluating Models and Obtaining Probability Map Images

In the final stage, the obtained models were tested on 80 test images, which included three groups of disease signs and both male and female groups. Different map images corresponding to the three or four classes were generated. Additionally, models trained on one group of signs were tested on images with the same sign, and vice versa. The models were also applied to normal images without signs to assess the performance of the algorithm. The TWS plugin generated colored segmented images using the same color scheme as the classes during training. In images with both disease signs (Figures 7 to 10), four colors represent the classes (a color for each class), while images with only one sign have three colors in the segmentation images.

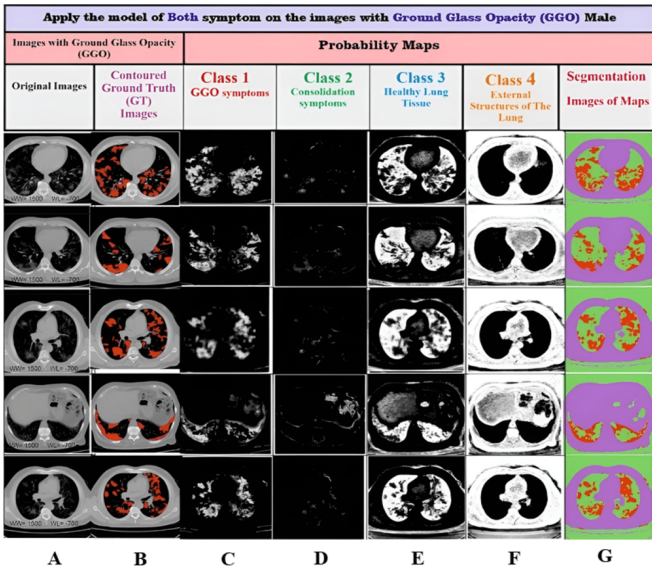

Figure 7. Probability maps and segmentation images generated by applying the best model for both symptoms to GGO images in the male group. (A) Original Images, (B) Contoured Ground Truth (GT) Images, (C) Class 1: GGO Symptoms, (D) Class 2: Consolidation Symptoms, (E) Class 3: Healthy Lung Tissue, (F) Class 4: External Structures of the Lung, (G) Images of maps

In Figure 7, the model accounts for both signs of the disease, resulting in four distinct classes. Class 1 represents the GGO sign, and class 2 represents the consolidation sign. Since the test images only have the GGO sign, areas related to this sign are depicted in color white (pixel values of 1) in class 1. In class 2, which corresponds to the consolidation sign, all areas -except for a few erroneous points- are displayed in white, while the rest are shown in black. Class 3 represents the healthy structure of the lung, where all healthy areas without the GGO sign are shown in white. Finally, class 4 represents structures outside the lung, with all external lung areas displayed in white and the lung tissue itself appears in black.

Similarly, in Figure 8, the model for both signs is applied to test images with both signs. Here, four classes are present, including class 1 (GGO sign), class 2 (consolidation sign), class 3 (healthy lung structure), and class 4 (structures outside the lung). In Figure 9, the GGO model is applied to test images with both disease signs. As the best model focuses on one sign,

the resulting maps have three classes: class 1 corresponds to the GGO sign, class 2 represents the healthy structure of the lung, and class 3 represents the structure outside the lung. In all three classes, the corresponding structures are depicted in white pixels, while the remaining areas appear as black. Figure 10 demonstrates the application of the model for both signs to normal images. Since these images are without disease, both classes 1 and 2 have black pixels, indicating the model's correct performance. In class 3, which represents healthy lung tissue, all lung tissue is depicted in white pixels, accurately identified by the model. All the figures from 6 to 9 include original images as well as ground truth images contoured by the radiologist in the areas related to the disease.

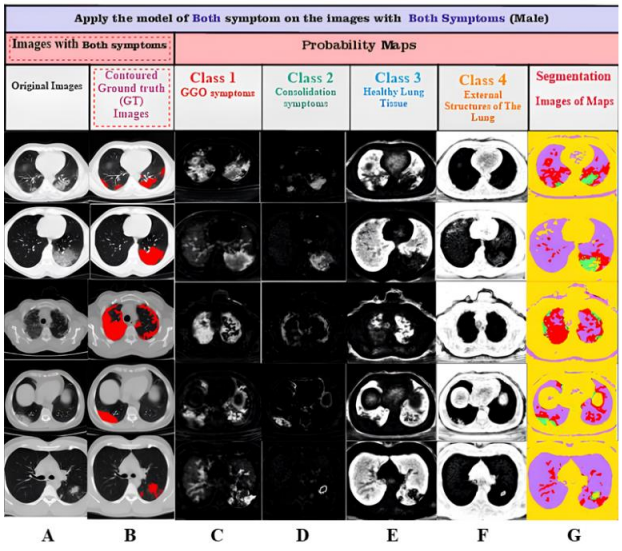

Figure 8. Probability maps and segmentation images generated by applying the best model for both symptoms to images with both disease symptoms in the male group. (A) Original Images, (B) Contoured Ground Truth (GT) Images, (C) Class 1: GGO Symptoms, (D) Class 2: Consolidation Symptoms, (E) Class 3: Healthy Lung Tissue, (F) Class 4: External Structures of the Lung, (G) Images of maps

3.5. Result of Thresholding

The results of the thresholding stage showed that by applying certain GGO models to both male and female groups (a total of 180 images, each model containing 60 images), positive and negative values were obtained. Positive values corresponded to images with the GGO sign, indicating that the number of pixels correctly detected as having the sign was higher than the average number of error pixels. Negative values were mainly related to images that only contained the consolidation sign, where the number of pixels recognized as having a value of one (indicative of the sign) was lower than the average. In the consolidation models, positive values indicated images with the consolidation sign or both signs, while on the other hand, negative values corresponded to images with the GGO sign. This distinction arises because the count of pixels with a value of one was lower than the sum of the mean and standard deviation. In models with both signs, positive values for COVID-19 disease indicated images with both signs, while negative values were for images with only one kind of sign, which were either GGO signs or consolidation ones. The results of this stage revealed that the presence or absence of the disease, as well as the type of sign present in the lung, depended on the specific model applied to the test images.

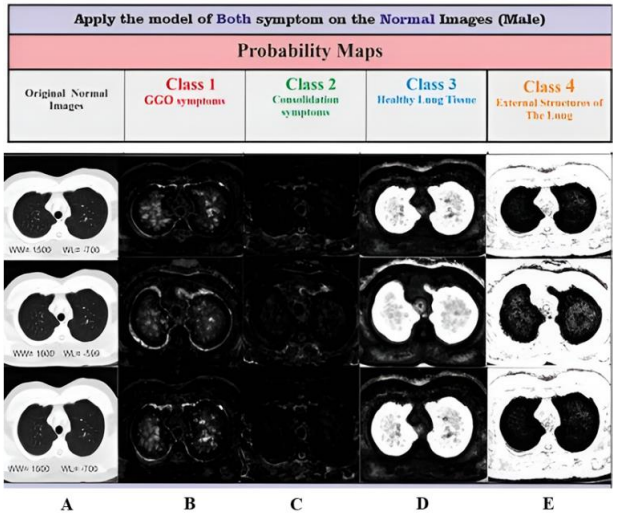

Figure 9. Probability maps and segmentation images generated by applying the best GGO model to images marked with both symptoms in the male group. (A) Original Images, (B) Contoured Ground Truth (GT) Images, (C) Class 1: GGO Symptoms, (D) Class 2: Healthy Lung Tissue, (E) Class 3: External Structures of the Lung, (F) Images of maps

3.6. Result for CT Severity Score (CT-SS)

In terms of the CT Severity Score (CT-SS), the image processing method using Fiji ImageJ software was used to determine the percentage of lung involvement in patients. In ground truth images related to the GGO and both signs, the calculated severity percentage was lower than the percentage obtained

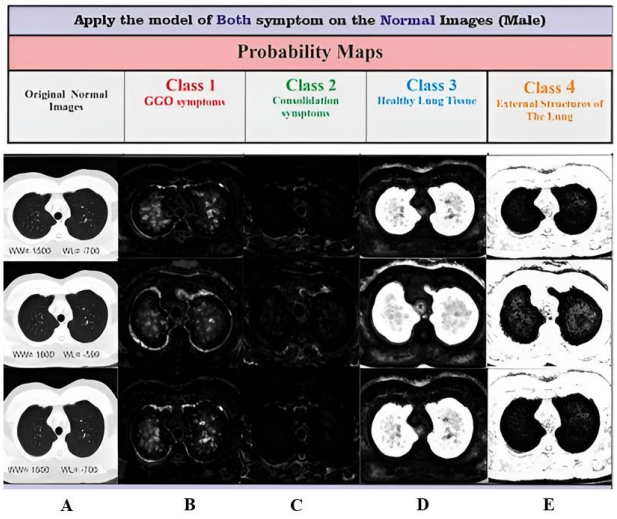

Figure 10. Probability maps generated by applying models trained with both symptoms on normal male images without disease symptoms. (A) Original Images, (B) Class 1: GGO Symptoms, (C) Class 2: Consolidation Symptoms, (D) Class 3: Healthy Lung Tissue, (E) Class 4: External Structures of the Lung

from the model's map images. This discrepancy can be attributed to the high false-positive rate of GGO signs. In other words, some lung tissue areas, including bronchus branches and airways, were mistakenly considered as GGO signs. On the other hand, all values for calculating the severity of COVID-19 involvement in ground truth images were higher than the values obtained from the consolidation model maps. This is due to the large, local, and peripheral nature of the consolidation sign. Consolidation models had a higher rate of false negatives compared to other models. This is because some consolidation signs, falling within the CT number range of -300 to +50, closely resemble extra-pulmonary structures and are depicted as white in the images. Another factor contributing to diagnostic errors is that consolidation signs are predominantly located in the peripheral and posterior regions of the lung, close to external structures. In some patients, these areas may merge with regions outside the lung, making it challenging to differentiate between them. Tables 6-8 present the results for CT-SS of patients with different signs.

4. Discussion

AI-assisted diagnosis systems have shown promise in the early detection of lung infections such as the

Table 6. Severity score of patient images with GGO sign

Image number	Number of GGO pixels		Number of Normal lung tissue pixels		Severity score obtained by the model	Severity score of the GT image
	GT image	Class 1 map image	GT image	Class 2 map image		
1	2666	5106	39668	35482	12.58	5.4
2	5657	5188	47538	41448	11.12	10.63
3	7572	11991	45521	19462	38.1	14.26
4	2845	3456	53888	58624	5.56	5.01
5	9555	7380	44348	31088	19.18	17.72
6	1661	2241	57846	62856	3.44	2.79
7	16173	25692	36209	15010	63.12	30.87
8	4169	4711	37223	37632	11.12	10.07
9	10910	14319	39751	11579	55.28	21.53
10	7663	7882	17326	6898	53.32	30.66

Table 7. Severity score of patient images with consolidation sign

Image number	Number of consolidation pixels		Number of Normal lung tissue pixels		Severity score obtained by the model	Severity score of GT image
	GT image	Class 1 map image	GT image	Class 2 map image		
1	10521	9754	38356	44533	17.96	21.52
2	4106	3764	30566	38970	8.8	11.84
3	3823	3747	56775	64778	5.46	6.3
4	9460	8952	74378	78342	10.25	11.28
5	3608	3667	42974	52855	6.48	7.74
6	4008	3594	33781	43431	7.64	10.6
7	4707	3871	41272	51997	6.92	10.23
8	19236	5627	30705	34645	13.97	38.51
9	5734	3946	15900	16831	18.99	26.50
10	2275	1921	20966	29695	6.07	9.78

Table 8. Severity score of patient images with both signs

Image number	Number of pixels with both signs		Number of Normal lung tissue pixels		Severity score obtained by the model	Severity score of GT image
	GT image	Class 1 & 2 map image	GT image	Class 2 map image		
1	16099	18009	16812	7337	71.05	48.91
2	13059	24938	18785	6666	78.9	41
3	3326	6301	70696	71400	8.1	4.49
4	2631	3354	38172	71569	4.47	3.71
5	2117	4120	30897	24233	14.53	6.41
6	15530	16444	16294	12070	57.66	48.79
7	21734	29587	11643	13444	68.75	65.11
8	15230	34420	50815	41396	45.39	23.06
9	5839	12118	70284	70801	14.61	7.67
10	5846	11709	17465	13946	45.64	25.07

COVID-19 virus and influenza H1N1 (A). These systems are able to significantly improve decision-making by creating proficient algorithms and facilitating research through data analysis. In this study, we proposed a novel approach for AI-assisted diagnosis systems using the Weka software and the random forest algorithm. The aim of our method was to simultaneously detect both the GGO and consolidation signs, and only lung involvement was considered. Our approach successfully addressed both signs using Weka and the random forest algorithm.

The combination of seven features, including Structure, Anisotropic, Laplacian, Minimum, Entropy, Maximum, and Median, yielded the best performance among the 19 features available in Weka. While Anisotropic and Laplacian features served for edge detection and noise removal, respectively, the remaining features were able to distinguish between the structural and textural content of pixels in the image [20]. These findings helped identify distinct structural patterns in lung tissue affected by the COVID-19 disease compared to healthy lung tissue. Moreover, based on the results from Tables 1 and 2, it can be observed that the five best models for GGO and consolidation in both gender groups differ only in one feature, while the other four features remain the same.

Throughout this study, a total of six models were developed to detect COVID-19 signs. Among them, the model trained using data with a combination of signs from both male and female patients achieved the highest accuracy of 98.1%. These models can also be applied for diagnosing other infectious lung diseases, including influenza A(H1N1).

To directly compare our obtained method with other existing ones in this field is a challenging task due to the novelty of our proposed method. However, a brief comparison is presented in Table 9, which highlights respiratory infectious disease detection using the random forest algorithm. Our proposed models proved to have superior accuracy (98.1%) compared to the results reported by studies [15] (87.9%) and [21] (91.79%). Researchers in study [5] utilized the random forest algorithm with 63 quantitative features and lung lobe segmentation to determine disease severity, but their validation parameters had lower accuracy (87.5%) compared to our proposed models. In another study [22] in early 2021, the Otsu thresholding method was used for lung area segmentation, and multiple algorithms, including random forest, were employed. While their models achieved higher accuracy (99.6%) and AUC (1.000) than ours, it should be noted that the differences in methodology and the specific focus on the RF algorithm for respiratory infectious disease sign detection in our study. Building on this, our approach aligns with recent research [23], which utilized deep learning methods (U-Net and DenseNet121) for segmentation and classification of lung CT images for detecting COVID-19 abnormalities. This study yielded lower accuracy than the present study. Similarly, study [24] contourlet transform and convolutional neural network have been employed to extract features individually from the segmented CT scan images and to fuse them in one feature vector. Then, an ML/DL-based ensemble classifier considering the bagging technique has been employed to detect COVID-19 from the CT images. methods used in this article were not explored in our study, but could enhance our methodology. These comparisons highlight both the strengths of our current method and potential areas for future enhancement by integrating broader features analysis techniques.

Furthermore, the lower recall and precision parameters in our study were due to the high values of FP and FN in the confusion matrix. This means that the model has difficulty distinguishing between signs of infectious respiratory diseases and anatomical structures of the lung such as bronchi with CT numbers similar to the CT numbers of the signs. It is worth noting that the problem of sign boundary detection is also a difficulty that many physicians face and our model is very successful in classifying healthy and diseased cases, which is a major task in clinics.

Table 9. Comparison of validation parameters between the proposed models of the study with previous studies in the field of detection of respiratory infectious disease symptoms

Studies	FPR	Specifity	AUC	Accuracy
[13]	0.302	0.907	0.942	0.879
[5]	0.380	0.933	0.91	0.875
[17]	0.310	0.930	0.963	0.9179
[18]	0.008	0.989	0.990	0.9885
[21]	0.187	0.813	0.8865	0.886
[22]	0.0433	0.9567	0.9707	0.9766
Our study	0.002	0.997	0.937	0.981

Table 3 also demonstrates the OOB error rates for the best models obtained using twice the training data. It is noted that these models have higher OOB error rates in comparison with the models trained with 120 images. The error rate for the best model with consolidation is higher than that of GGO in both male and female groups (4.4%), and it reaches 11.18% in the model with both signs.

Overall, what our study demonstrates is the potential of the random forest algorithm in AI-assisted diagnosis systems for the accurate detection of respiratory infectious disease signs. The combination of different features in Weka allowed for robust models that can be applied not only to COVID-19 but also to other infectious lung diseases. Further research and validation are needed to refine and expand the application of AI-assisted diagnosis systems in the field of lung disease diagnosis and management.

The limitation of our study was finding images that exclusively exhibited consolidation signs, as most images typically had both GGO and consolidation signs together. This limited our ability to focus solely on the consolidation sign. Despite this limitation, this study offers valuable insights into the use of AI-assisted diagnosis systems and the random forest algorithm for detecting lung infection signs. It paves the way for future research on the refinement of the methodology and expanding the application of AI in the diagnosis and management of infectious lung diseases.

5. Conclusion

In this study, our objective was to achieve accurate lung image segmentation by training the random forest algorithm in Weka. The results obtained from this research have significant implications not only for the diagnosis of infectious lung diseases, including influenza(A) and COVID-19, but also for their treatment. The best models developed through the training of the random forest algorithm can be utilized to assist radiologists in clinical settings.

Currently, radiologists often examine chest CT scan slices of patients sequentially, since there exists a considerably large number of patients that they need to assess. This approach may not allow for such detailed evaluation of each slice; However, our trained models can analyze each lung image in a separate manner and provide a comprehensive identification of disease signs. By employing our method, radiologists can benefit from a more accurate assessment of all lung slices, improving the diagnosis and, on top of it, the treatment process.

The outcomes of this study have the potential to enhance the speed and efficiency of disease detection and management. By automating the process of sign identification in lung images, our proposed method can contribute to faster and more accurate diagnoses of infectious lung diseases. This can ultimately lead to highly improved patient outcomes and more effective treatment strategies.

In conclusion, the utilization of the random forest algorithm trained in Weka for lung image segmentation holds promise in the field of diagnosis. By providing enough detailed information on disease signs, this method can support radiologists in making informed decisions.

Acknowledgment

We extend our heartfelt thanks to Alghadir Hospital for providing the dataset essential for our study. The images were meticulously verified by an experienced radiologist to confirm the presence or absence of disease, ensuring the integrity of our clinical data. We recognize that this dataset is not publicly available due to privacy and ethical considerations related to patient confidentiality. Nonetheless, we appreciate the hospital's cooperation in this research, which underscores the balance between data utility and ethical responsibility.

References

1- Saeid Safiri *et al.*, "Global burden of lower respiratory infections during the last three decades." *Frontiers in Public Health,* Vol. 10p. 1028525, (2023).

2- Farid Rahimi and Amin Talebi Bezmin Abadi, "Implications of the Emergence of a New Variant of SARS-CoV-2, VUI-202012/01." *Archives of Medical Research,* Vol. 52 (No. 5), pp. 569-71, (2021).

3- Summer E Galloway, "Emergence of SARS-CoV-2 b. 1.1. 7 lineage—united states, december 29, 2020–january 12, 2021." *MMWR. Morbidity and mortality weekly report,* Vol. 70(2021).

4- Tao Zhou, Huiling Lu, Zaoli Yang, Shi Qiu, Bingqiang Huo, and Yali Dong, "The ensemble deep learning model for novel COVID-19 on CT images." *Applied soft computing,* Vol. 98p. 106885, (2021).

5- Zhenyu Tang *et al.*, "Severity assessment of COVID-19 using CT image features and laboratory indices." *Physics in Medicine & Biology,* Vol. 66 (No. 3), p. 035015, (2021).

6- Huijun Chen *et al.*, "Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: a retrospective review of medical records." *The lancet,* Vol. 395 (No. 10226), pp. 809-15, (2020).

7- Mohammed Mahmoud Dawoud, Tamer Mahmoud Dawoud, Noha Yousef Amin Ali, and Hanan Ahmad Nagy, "Chest CT in COVID-19 pneumonia: a correlation of lung abnormalities with duration and severity of symptoms." *Egyptian Journal of Radiology and Nuclear Medicine,* Vol. 51pp. 1-12, (2020).

8- S Sreejith, J Ajayan, NV Uma Reddy, and Shashank Rebelli, "Analysis of COVID-19 CT Chest Image Classification using Dl4jMlp Classifier and Multilayer Perceptron in WEKA Environment." *Current Medical Imaging,* Vol. 20 (No. 1), p. e170423215872, (2024).

9- Bo Wang *et al.*, "AI-assisted CT imaging analysis for COVID-19 screening: Building and deploying a medical AI system." *Applied soft computing,* Vol. 98p. 106897, (2021).

10- Michael Chung *et al.*, "CT imaging features of 2019 novel coronavirus (2019-nCoV)." *Radiology,* Vol. 295 (No. 1), pp. 202-07, (2020).

11- Lin Li *et al.*, "Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT." *Radiology,* (2020).

12- Michael F McNitt-Gray *et al.*, "The Lung Image Database Consortium (LIDC) data collection process for nodule detection and annotation." *Academic radiology,* Vol. 14 (No. 12), pp. 1464-74, (2007).

13- Yasufumi Oi *et al.*, "Prediction of prognosis in patients with severe COVID-19 pneumonia using CT score by emergency physicians: a single-center retrospective study." *Scientific Reports,* Vol. 13 (No. 1), p. 4045, (2023).

14- Ran Yang *et al.*, "Chest CT severity score: an imaging tool for assessing severe COVID-19." *Radiology: Cardiothoracic Imaging,* Vol. 2 (No. 2), p. e200047, (2020).

15- Chuangsheng Zheng *et al.*, "Deep learning-based detection for COVID-19 from chest CT using weak label." *MedRxiv,* p. 2020.03. 12.20027185, (2020).

16- Huseyin Yasar and Murat Ceylan, "A novel comparative study for detection of Covid-19 on CT lung images using texture analysis, machine learning, and deep learning methods." *Multimedia Tools and Applications,* Vol. 80 (No. 4), pp. 5423-47, (2021).

17- Feng Shi *et al.*, "Large-scale screening to distinguish between COVID-19 and community-acquired pneumonia using infection size-aware classification." *Physics in Medicine & Biology,* Vol. 66 (No. 6), p. 065031, (2021).

18- Feng Pan *et al.*, "Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia." *Radiology,* (2020).

19- Wajid Arshad Abbasi *et al.*, "COVIDC: An expert system to diagnose COVID-19 and predict its severity using chest CT scans: Application in radiology." *Informatics in Medicine Unlocked,* Vol. 23p. 100540, (2021).

20- Ali Abbasian Ardakani, U Rajendra Acharya, Sina Habibollahi, and Afshin Mohammadi, "COVIDiag: a clinical CAD system to diagnose COVID-19 pneumonia based on CT findings." *European radiology,* Vol. 31pp. 121-30, (2021).

21- Liang Sun *et al.*, "Adaptive feature selection guided deep forest for covid-19 classification with chest ct." *IEEE Journal of Biomedical and Health Informatics,* Vol. 24 (No. 10), pp. 2798-805, (2020).

22- Amira Hasan, Hala Abd El Kader, and Aya Hossam, "An intelligent detection system for Covid-19 diagnosis using CT-images." *JES. Journal of Engineering Sciences,* Vol. 49 (No. 4), pp. 476-508, (2021).

23- Abdoulreza S Moosavi, Ashraf Mahboobi, Farzin Arabzadeh, Nazanin Ramezani, Helia S Moosavi, and Golbarg Mehrpoor, "Segmentation and classification of lungs CT-scan for detecting COVID-19 abnormalities by deep learning technique: U-Net model." *Journal of Family Medicine and Primary Care,* Vol. 13 (No. 2), pp. 691-98, (2024).

24- Md Nur-A-Alam, Mostofa Kamal Nasir, Mominul Ahsan, Md Abdul Based, Julfikar Haider, and Marcin Kowalski, "Ensemble classification of integrated CT scan datasets in detecting COVID-19 using feature fusion from contourlet transform and CNN." *Scientific Reports,* Vol. 13 (No. 1), p. 20063, (2023).