ORIGINAL ARTICLE

Expert Systems   WILEY

Effective hybrid deep learning model for COVID-19 patterns identification using CT images

Dheyaa Ahmed Ibrahim[1] | **Dilovan Asaad Zebari**[2] | **Hussam J. Mohammed**[3] |
Mazin Abed Mohammed[4]

[1]Communications Engineering Techniques Department, Information Technology Collage, Imam Ja'afar Al-Sadiq University, Baghdad, Iraq

[2]Department of Computer Science, College of Science, Nawroz University, Duhok Kurdistan Region, Iraq

[3]Computer Center, University of Anbar, Al Anbar, Iraq

[4]Information systems Department, College of Computer Science and Information Technology, University of Anbar, Al Anbar, Iraq

Correspondence
Dheyaa Ahmed Ibrahim, Communications Engineering Techniques Department, Information Technology Collage, Imam Ja'afar Al-Sadiq University, Baghdad, Iraq.
Email: dheyaa.ibrahim@sadiq.edu.iq

Abstract

Coronavirus disease 2019 (COVID-19) has attracted significant attention of researchers from various disciplines since the end of 2019. Although the global epidemic situation is stabilizing due to vaccination, new COVID-19 cases are constantly being discovered around the world. As a result, lung computed tomography (CT) examination, an aggregated identification technique, has been used to ameliorate diagnosis. It helps reveal missed diagnoses due to the ambiguity of nucleic acid polymerase chain reaction. Therefore, this study investigated how quickly and accurately hybrid deep learning (DL) methods can identify infected individuals with COVID-19 on the basis of their lung CT images. In addition, this study proposed a developed system to create a reliable COVID-19 prediction network using various layers starting with the segmentation of the lung CT scan image and ending with disease prediction. The initial step of the system starts with a proposed technique for lung segmentation that relies on a no-threshold histogram-based image segmentation method. Afterward, the GrabCut method was used as a post-segmentation method to enhance segmentation outcomes and avoid over-and under-segmentation problems. Then, three pre-trained models of standard DL methods, including Visual Geometry Group Network, convolutional deep belief network, and high-resolution network, were utilized to extract the most affective features from the segmented images that can help to identify COVID-19. These three described pre-trained models were combined as a new mechanism to increase the system's overall prediction capabilities. A publicly available dataset, namely, COVID-19 CT, was used to test the performance of the proposed model, which obtained a 95% accuracy rate. On the basis of comparison, the proposed model outperformed several state-of-the-art studies. Because of its effectiveness in accurately screening COVID-19 CT images, the developed model will potentially be valuable as an additional diagnostic tool for leading clinical professionals.

KEYWORDS
COVID-19 identification, CT scan images, deep learning models, feature fusion

1 | INTRODUCTION

Coronavirus disease 2019 (COVID-19) has posed a global threat to human life throughout the world (Alyasseri et al., 2021). Individuals infected with SARS-CoV-2 will experience fever, coughing, muscle aches, migraine, and other flu-like symptoms (Hasoon et al., 2021). COVID-19, a previously unknown human pathogen, is able to move from different animal species to individual populations, where it can spread very rapidly. Once the coronavirus starts to spread throughout the population, it can overwhelm the medical system of countries within 4 weeks (Allioui et al., 2022). The World Health Organization named illnesses that are caused by viruses as COVID-19. COVID-19 can be contained, and the mortality rate can be reduced through the early discovery and treatment of suspected patients (Alyasseri et al., 2021; World Health Organization, 2020).

In this context, reverse transcription-polymerase chain reaction (RT-PCR) is typically utilized to determine the presence of COVID-19 (Ai et al., 2020). However, the RT-PCR test has a number of limitations, including limited RT-PCR kits, lengthy procedure, and high rate of false negatives; thus, patients cannot be diagnosed and treated on time (Fang et al., 2020; Xie et al., 2020). Therefore, chest computed tomography (CT) is utilized to identify suspected diseases. In recent years, CT has used by many experts, owing to the possibility that an initial chest CT may reveal abnormal signs of COVID-19 (Pan et al., 2020). Furthermore, CT has a quick turnaround time, high positive rate, and better diagnostic accuracy due to having access to pathology-specific information (Xu et al., 2020; Zebari et al., 2020). In the literature, several works based on computerized tomography (CT) images were suggested as secondary examination for suspicious persons who hold COVID-19 and shows symptoms even their RT-PCR results were negative. For example, in China, Wuhan the test for COVID-19 for 1014 persons has been performed, during the test based on RT-PCR it shows that 59% of patients are holding COVID-19 whereas when the test has been perfumed with CT scan it shows that 88% of patients are holding COVID-19 with difference of 29% between both tests. Moreover, among the results of COVID-19 of RT-PCT the CT scans have obtained higher sensitivity of 97%. Therefore, CT scans are able to identify COVID-19 with higher accuracy compared to RT-PCR. Furthermore, early detection of lesions in the lung-based CT scan images can be performed, thus they can be utilized by radiologists for COVID-19 diagnosis (Qiblawey et al., 2021). Although CT scan has demonstrated immense possibility in COVID-19, the diagnosis of pneumonia, radiographic features detection performs manually. Peripheral ground-glass opacities have frequently offered lesser capability to distinguish COVID-19 among other types of pneumonitis including viral and bacterial pneumonias (Li et al., 2020). In addition, the number of people with COVID-19 rapidly increases every day; therefore, multiple CT scans are performed with an average 300 slices for each patient. This leads to an increasing number of CT images. Thus, radiologists face a significant challenge to deal with these images particularly in epidemic areas. A possible solution for a quick and an accurate identification of COVID-19 cases from a large number of CT slices is the development of a computer-aided detection system by using the data mining techniques and convolutional neural networks (CNNs), which is a new artificial intelligence (AI) technique that used widely in the medical fields.

Data mining can be a highly useful method for predicting medical issues and assisting caregivers in making precise medical decisions. The ability to execute sophisticated computing processes, such as finding patterns in a vast number of datasets-based data mining techniques that can be able to extract a valuable information and patterns in the medical cases. Furthermore, COVID-19 infections can be identified with high accuracy and sensitivity by using medical datasets. The classification of COVID-19 patients is in fact one of the data analyses processes that is used to assign them to their corresponding classes. There are numerous classification methods supporting vector machines such as the Bayesian method, decision trees, KNNs, and artificial neural networks (Shariaty et al., 2019; Thamilselvan & Sathiaseelan, 2015). Classification techniques have the ability to diagnose COVID-19 from CT images on the basis of common features that are extracted. Therefore, it is necessary to perform the feature extraction process on CT images in order to transform them into a set of relevant features before performing the detection process. This process will aid the classification method in making accurate decisions (Lingayat & Tarambale, 2013). Recently, various feature extraction methods have been developed, such as texture features, morphological features, co-occurrence matrices, Gabor features, deep-based CNN features, and wavelet transform-based features.

This work provides the following contributions:

1. A multi-level segmentation method was proposed to extract the lung from the CT images. In this method, we enhanced the performance of GrabCut by combining it with the threshold method. A no-threshold histogram-based image segmentation method was used as an initial segmentation stage to binarize the image, keep the whole region of interest (ROI), and avoid the over-segmentation problem. Then, the GrabCut method was utilized as a post-segmentation method to identify the border of the RIO and avoid the under-segmentation problem.
2. An affective training strategy was used to build a powerful trainable model by using the augmentation method to increase the dataset size, which can help avoid overfitting issues.
3. Three pre-trained models (Visual Geometry Group Network [VGGNet], convolutional deep belief network [CDBN], and high-resolution network [HRNet]) were analysed and investigated in this study. Each model presents a good accuracy. However, their accuracy is not enough for the sensitive COVID-19 issues. Therefore, more investigations have been conducted and showed that the VGGNet, CDBN, and HRNet can work together to create a powerful decision. For this reason, a novel effective hybrid deep learning (DL) model for COVID-19 detection was developed and presented in this study.

4. A fusion mechanism was proposed to combine the previously described pre-trained models to increase the system's overall prediction capabilities.

The structure of this paper is as follows: Section 2 describes the most recent state-of-the-art studies. Section 3 presents the proposed method including COVID-19 CT scan dataset, proposed segmentation model, three pre-trained DL models, and proposed fusion model. Section 4 depicts the evaluation of the proposed approach. Finally, Section 5 presents the conclusion and future work.

2 | LITERATURE REVIEW

This section discusses several notable proposed studies that have addressed the issue of COVID-19 identification on CT scan images and have had a direct impact on the development of this work utilizing AI techniques. These studies will highlight the most important facts about COVID-19 identification utilizing AI, such as the feature extraction phase, which will be automated for feature learning; classification model; and the database and image type that were utilized in the experiments.

In this regard, the study of Barstugan et al. (Barstugan et al., 2020) proposed a technique for classifying COVID-19 patients that relied on machine learning methods. The classification step was carried out after classifying data into four different sets of data from 150 CT images with 16×16, 32×32, 48×48, and 64×64 taking patches based. In Barstugan et al. (Barstugan et al., 2020) the feature extraction phase was implemented on the basis of five different feature extraction techniques to extract the most relevant features from the CT scan images. This phase was implemented to separate the infected patches more accurately. Afterward, the patients were classified on the basis of the retrieved features, which were fed into a support vector machine (SVM) as an input. As presented in Barstugan et al. (Barstugan et al., 2020) the grey-level size zone matrix with support vector machines (GLSZM-SVM) worked successfully. In Wang et al. (Wang et al., 2021) COVID-19 could be analysed by using a probabilistic method as a classification method. As a result, the CT scans with COVID-19 might be classified into groups on the basis of the most relevant features through the feature extraction phase. Afterward, the feature selection phase was executed on the extracted features. Finally, the classification phase was performed using the stack hybrid classification (SHC) method. To ameliorate the prediction performance, SHC relied on ensemble methods, which combine several models. On the basis of the experimental findings in Wang et al. (Wang et al., 2021) the new strategy was found to outperform the traditional categorization methods.

In Farid et al. (2020) DL methods could extract specific graphical features from the COVID-19 image. Similarly, a clinical diagnosis can be introduced before pathogenic testing is performed. This technique attempted to save valuable time for the diagnosis of the disease. On the basis of the experimental results in Farid et al. (2020) the efficiency of DL techniques to extract graphical features to diagnose patients with COVID-19 was proven. Machine learning techniques could be used to diagnose 150 images from COVID-19 cases and non-COVID-19 images (Barstugan et al., 2020). When it came to feature extraction, different techniques were considered, including GLSZM and discrete wavelet transform. The extracted features were then classified utilizing SVM, which was trained on the data. K-fold cross-validations were carried out in the experiments with 2-, 5-, and 10-folds of data. According to the findings of this study, SVM using the GLSZM features was able to obtain an accuracy rate of 99.68%. In Gozes et al. (2020) a complete system for identifying COVID-19 cases was proposed for comparison with other types of cases. COVID-19 identification in CT scan was part of the proposed system, and the case was marked as COVID-19 when the number of COVID-19-positive slices exceeded a predetermined threshold. When it came to the training and testing phases, several datasets were considered, and the network that has been prepared in advance COVID-19 was found using ResNet50. This system had a sensitivity of 94%, specificity of 98%, and AUC score of 0.9940.

When it came to diagnosing COVID-19 patients on the basis of the features extracted from chest CT scan images, classification methods can be utilized effectively. Before applying the detection model to CT images, feature extraction needs to be conducted. Features extracted from a CT image are used to aid the classification approach in making correct decisions about the image's contents. The CT images we have examined revealed that texture is likely the most prominent visual characteristic contained within them. Consequently, we investigated some texture descriptors that have been used in literature, considering both handcrafted and non-handcrafted methods for texture description. Table 1 summarizes the research covered in this section, highlighting the most significant findings from each study. One of the primary goals of this table is to provide a convenient way to quickly locate some critical information about those studies.

As discussed above, machine learning and DL technologies are important in tackling the problem of automated COVID-19 identification. However, despite the positive outcomes from the previous research, some improvements can be made in the diagnosis of COVID-19 patients on the basis of chest CT. It is demonstrated that there is still a difficult effort and an important area of research to examine how to use procedures that have acceptable diagnostic accuracy. Because machine learning and DL algorithms produce good outcomes for COVID-19 detection, identifying good diagnostic tools has become more difficult. It was possible to discover, via the thorough research of these works, that certain procedures do not perform well with variability in the samples (the results of the work), because it uses mechanisms of inflexible extraction with little adaptability. Global features (e.g., edges) are removed from an image using these methods independent of local information. By looking at the extraction of key features, we try to better diagnose the problem by identifying subtle textural variations within points that are considered to be a

TABLE 1 Presents most recent state-of-art works

Ref.	Method	Dataset	Feature	Problem	Class	Results	Limitation
(Shi et al., 2020)	Infection size aware Random forest method	2685 private	Volume, number, histogram, and surface FS: LASSO	Classification of COVID-19 from CT images	2	Ac = 87.9, Sn = 90.7, Sp = 83.3	Follow-up CT scan results were excluded, and clinical features associated with pneumonia were not taken into account.
Tang et al. (2020)	Random Forest model	176 private	30 quanti- tative	Severity assessment of COVID-19 patients		Ac = 0.875, TP = 0.933, TN = 0.74	It was decided to use only two of the COVID-19 severity categories (i.e., with binary categorization) rather than all four (i.e, mild, common, severe and critical)
Barstugan et al. (2020)	SVM	150	GLCM, LDP, GLRLM, GLSZM, DWT	For the classification of COVID-19		Ac = 99.68, Sn = 97.56, Sp = 99.68, Pre = 99.62, F1-score = 98.58	SVM cannot deal properly with huge datasets, and it struggled mightily when confronted with data that had more noise.
Al-Karawi et al. (2020)	SVM	470	Gabor	To differentiate between positive from negative cases	2	Ac = 95.37, Sn = 95.99, Sp = 94.76	–
Özkaya et al. (2020)	CNN with Fusion and Ranking method	150 private	VGG-16, GoogleNet and ResNet-50	Classification of COVID-19 images	---	Ac = 98.27, Sn = 98.93, SP = 97.6, Pre = 97.63, F1-score = 98.28, MCC = 96.54	–
Alom et al. (2020)	Improved Inception Recurrent Residual Neural Network (IRRCNN) and NABLA-3 network models	420 CT publicly	–	For identification of COVID-19 patients from X-ray and CT images	2	X-ray Ac = 84.67 CT Ac = 98.78	because of the dearth of labelled data for COVID-19 lung segmentation in CT, the COVID-Seg CT generates results that contain a small number of true positives
Wang et al. (2020)	DenseNet121-FPN and COVID-19Net	5372	DL feature	For COVID-19 Diagnostic and Prognostic Analysis	2	Ac = 85, Sn = 79.35, Sp = 71.43	diagnostic performance of the DL is low

more representative sample. According to the researchers, the study intends to improve the way CT scan images may be used to better understand texture or texture information. Several studies have focused on the detection of COVID-19, which can assist clinicians in accurately diagnosing COVID-19 and evaluating treatment response. The offered models' performance is still poor of studies that used the same as our dataset as us. COVID-19 patients will be identified and classified into two severity levels as part of this study. In addition, various DL strategies for recognizing COVID-19 are investigated in the research.

3 | PROPOSED MODEL

COVID-19 detection based on the chest X-ray and CT scan images has become the main subject of numerous studies to date. Comparing with X-rays, the CT scans have less false-positive rates, and this is the foundation of this research. The most difficult part was obtaining a model that has the desired accuracy. This research has employed a variety of DL techniques to obtain a model with an effective and a desired accuracy. Figure 1 shows the complete proposed system model that addresses COVID-19 detection on CT scan images. The proposed model consists of three main stages: firstly, is the proposed segmentation method to extract the ROI from the CT scan image and use the segmented region as input to the DL method to identify the risk of COVID-19. In this stage, two segmentation steps were proposed to identify the ROI including the initial segmentation method by applying the threshold-based binarization method to binarize the CT image and capture the whole or part of the ROI, as well as unwanted objects. The post-segmentation method (GrabCut method) was then used to get the whole ROI (avoiding the under-segmentation problem) and ignore unwanted objects. The main contribution of this stage is enhancing the achievement of GrabCut by combining it with the threshold to obtain fully automatic segmentation methods. Secondly, three DL models (CDBN, HRNet, and VGGNet model) were trained and tested to evaluate and determine the accuracy of each model. Finally, the outcome of the three models was fused to benefit from the advantages of each model and obtain the best result.

3.1 | Data acquisition

We utilized a publicly available open-source dataset called the COVID-19 CT dataset to examine our proposed model. Many people who were afflicted with COVID-19 between January and April of 2020 found the therapy and diagnostic to be really valuable which has been proven by radiologists at Tongji Hospital, Wuhan, China. This dataset consists of 349 and 397 CT images of positive and negative COVID-19, respectively, which were gathered from 216 patients. Figure 2 shows the positive and negative samples of COVID-19. Preprints on medRxiv and bioRxiv regarding COVID-19 yielded positive CT images, which show diverse kinds of COVID-19. The CT images have various sizes because they were obtained from various sources (Zebari et al., 2020).

3.2 | Segmentation

Two segmentation methods were built in this stage to achieve the segmentation objective. First, the threshold-based method was used to generate the initial mask that can be used as input to the GrabCut method. Then, the GrabCut method was used as a post-segmentation method to extract the whole ROI, avoid under-segmentation, and ignore unwanted objects. In this process, the GrabCut method was enhanced by the threshold technique to achieve a fully automatic segmentation method by generating the initial mask automatically. In the following section, the main segmentation stages will be described in detail.

3.2.1 | Threshold-based method

In this study, a binary image containing the ROI was produced using the segmentation-based threshold approach. Determining the ROI within an image is the main obstacle in the field of image processing to be analysed. Despite the fact that grey-level thresholding frequently yields unhelpful results, this method continues to be the focus of numerous studies that propose new approaches for automatically determining the correct grey-level threshold. Thus, the threshold-based approach was utilized as an initial segmentation step to extract all or part of the ROI and utilize it as a mask for the subsequent segmentation step in this research. A no-threshold histogram-based image segmentation method was applied as an initial segmentation step.

To illustrate the process of the method, two concepts should be explained: 1) for the sake of simplicity, only a first-order (one-dimensional) histogram will be discussed in this note. As a result, histograms in two and three dimensions may be simply generated by applying the process to pixels with multiple characteristics. 2) Given that the classes are blended here rather than separated as in supervised processes, the probability
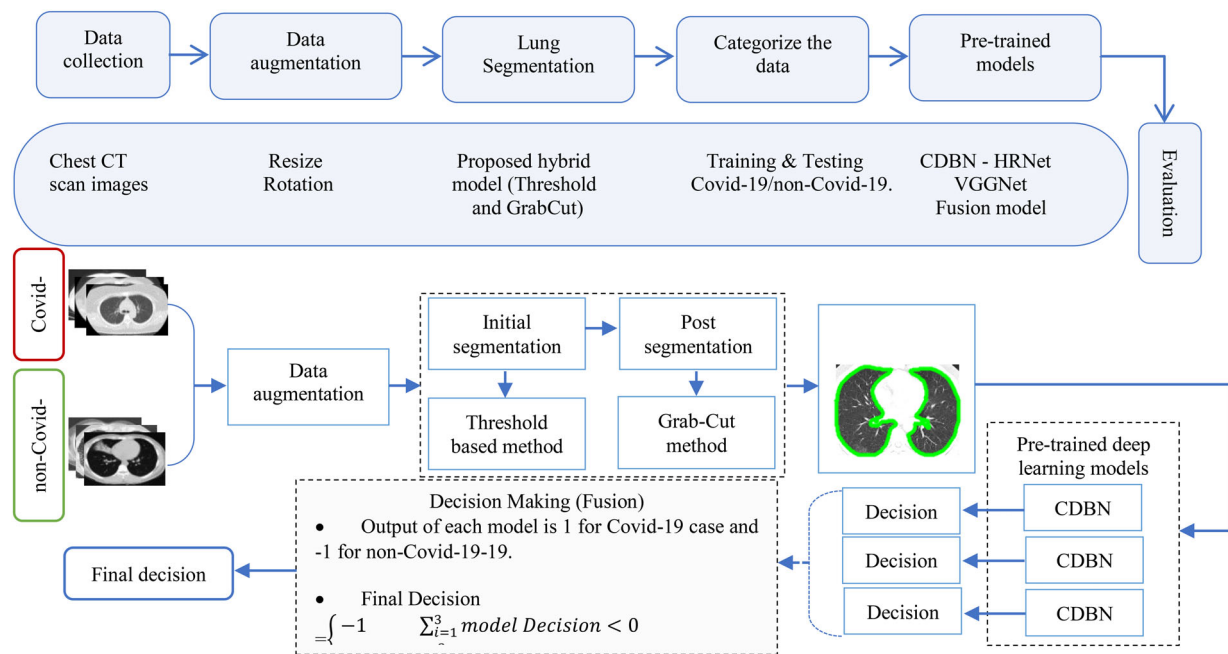
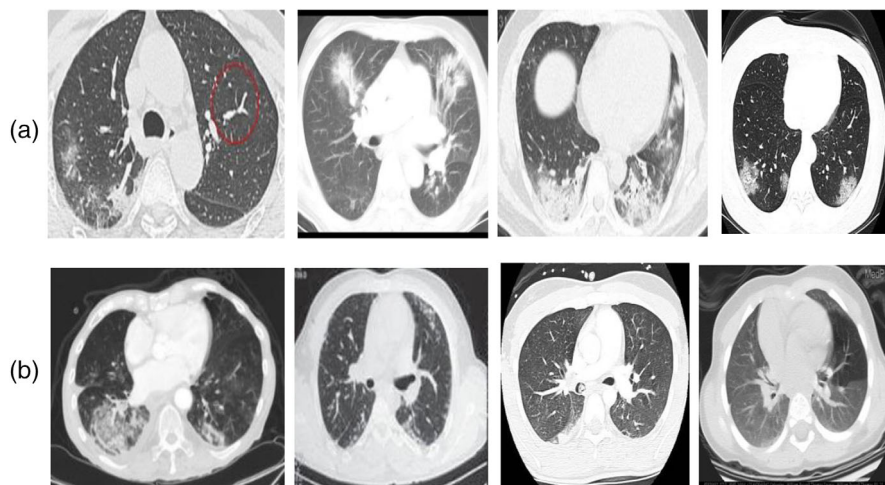**FIGURE 1** The proposed architecture for COVID-19 identification

FIGURE 2 Chest CT scan samples from COVID-19 datasets, first row (a) presents positive cases, while second row (b) presents negative cases

density function (PDF) cannot be estimated individually. For the purpose of calculating the global PDF for each pixel class in the image, the image's histogram is generated in the first order. To do this, the histogram should be more regular. The standard Parzen–Rosenblatt approach can be used with a Gaussian or other-shaped kernel to accomplish this. Once the predicted PDF has undergone regularization, we may consider that each mode represents a different pixel class. One of the approaches already proposed for difficult partitioning of the one-dimensional histogram such as the selection of a threshold (Ren et al., 2019).

The method goes one step further because it performs image segmentation without the need to look for borders between categories in the parameter (grey level) space. We can achieve this by simply characterizing each grey level gl (or, more generally, each sampling point in the space of the parameter), for instance, by specifying the grades at which it belongs in each class c of grey levels, as follows: $(gl,c), gl = 1,...,GL; c = 1,...,C$. The following is an example of how the findings can be displayed in the image space: $I(x,y) = gl \rightarrow \mu_{x,y}(c) = \mu(gl,c)$. Now, rather than just carrying out the most basic of tasks, defuzzyfunction $\mu_{x,y}(c) \rightarrow label = argmax_c(\mu(gl,c))$, it is preferable to attempt to ameliorate the outcomes achieved while ignoring the pixel coordinates. This can be accomplished using probabilistic relaxing, in which the membership grades are iteratively modified to take into account the membership grades of neighbouring pixels. The final point that has to be discussed is how we can extract the

membership grades of $\mu(gl,c)$ from the global predestinated PDF. Distinctly, there are different possibilities that can be envisaged, each of which relies on the fuzziness definition that we use. In this brief note, we simply want to demonstrate the concept of the technique by using a straight-forward approach to illustrate the idea of the method. Overall, we want to account for the image statistics while also taking into consideration the form and height of the various PDF modes. Equation (1) defines a possibility.

$$\mu(gl,c) = \frac{\exp\left(-x[\text{cost}(gl,m_c)]^2\right)}{\sum\limits_{c'=1}^{c} \exp\left(-x[\text{cost}(gl,m_{c'})]^2\right)'} \tag{1}$$

A path's cost is determined by the relative amplitudes of the modes' heights. There are one-dimensional histograms where only the path from a parameter space point to the histogram mode is fully defined, as indicated above. Multi-dimensional histograms require a more detailed description that incorporates concepts such as the least expensive path being the optimum.

3.2.2 | Grab-cut method

In this study, the GrabCut method was employed in the post-segmentation step to efficiently isolate the ROI from the background and other unwanted information. This method mostly depends on the graph cuts method. In order to use the GrabCut algorithm, the user must draw a sur-round box on the objects of the input image to be segmented. The initial phase of segmentation was necessary to generate an initial mask before feeding it into the GrabCut function as input. Afterward, the Gaussian mixture model (GMM) is employed to determine the scene's foreground object's colour distribution. GMM analyzes an image and assigns labels to all of the pixels that are unknown. This information is gleaned from the image's data. According to GMM's colour statistics, every pixel is classified as foreground or background (Basavaprasad & Hegadi, 2014; Jaisakthi et al., 2018).

For example, as illustrated in Figure 3, the GrabCut method treats the image as a graph, whereas the individual pixels represent the vertices. In addition, the feature link between pixels will represent the edges between the vertices. In each iteration of the image pixels, the GrabCut method eliminates the weak connections among pixels and then labels them either background or foreground. As a result, the accuracy of the method might be considerably influenced by the bounding box that has been defined because it is outside of the box that contains the majority of the background information. The pixels within the image are deemed to be a member of the background when the features of the pixels within and outside the bounding box are similar. However, the selected pixel is presumed to be a part of the intended foreground item. The bounding box is generated and passed into the GrabCut method utilizing (X, Y, H, and W) as fixed coordinates, which are specified to be (50, 50, 400, 400), respectively. A rectangle containing the intended foreground object is formed by performing these coordinates over all the images in the bean col-lection of datasets (bean leaves). It is necessary to manually review and correct a number of the segmented images that are obtained from the GrabCut method.

Figure 4 shows the outcome of the whole segmentation process. In this figure, three CT images have been used as examples to show the effect of the proposed model. As we can see, the first column shows the original CT images. The initial segmentation process has been highlighted in the second column, and the under-segmentation problem is clear at this stage. This limitation encouraged us to use the post-segmentation method to extract the whole ROI as shown in the third column. Finally, the border of the ROI has been highlighted in the fourth column to show the power of the proposed model.

3.3 | DL models

DL networks are automatically utilized to extract deep features from the fully connected (FC) layer of any network. In order to distinguish between distinct input classes, only representative characteristics are used from the FC layer. This study aims to distinguish between positive and negative COVID-19 CT images. To identify COVID-19 from the CT scan, a new fusion model-based DL network has been presented in this study. Three effective architectures are used to extract FC layer features (CDBN, HRNet, and VGGNet), which is a popular strategy because the FC layer precedes a softmax classifier. Depending on the architectures selected, the images will be processed to extract various deep features. CNN's default classifier is softmax, a powerful and widely utilized form of discriminant classifier. The nonlinear conversion of the distance between both sets of training and testing images is used by the softmax discriminant function to allocate a new input of the testing images to the output class. The learning rule for binary units in a softmax classifier is comparable to the standard binary unit law. The softmax function model is identical to the function of logistic sigmoid aside from resolving classification concerns when there are more than two potential values. Because the primary goal of this research is to accurately categorize COVID-19 instances, we used a fusion of three different architectures to create a model-based voting system that is powerful and efficient.
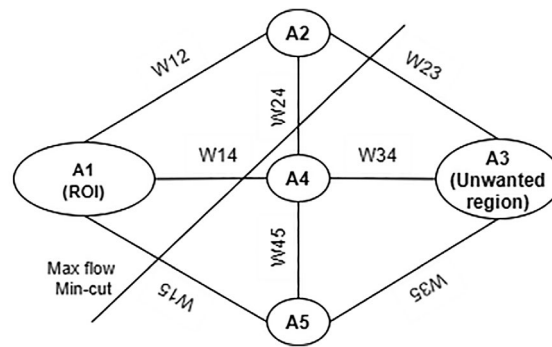
FIGURE 3 Grab-cut method to segment objects into ROI and unwanted objects (Basavaprasad & Hegadi, 2014)
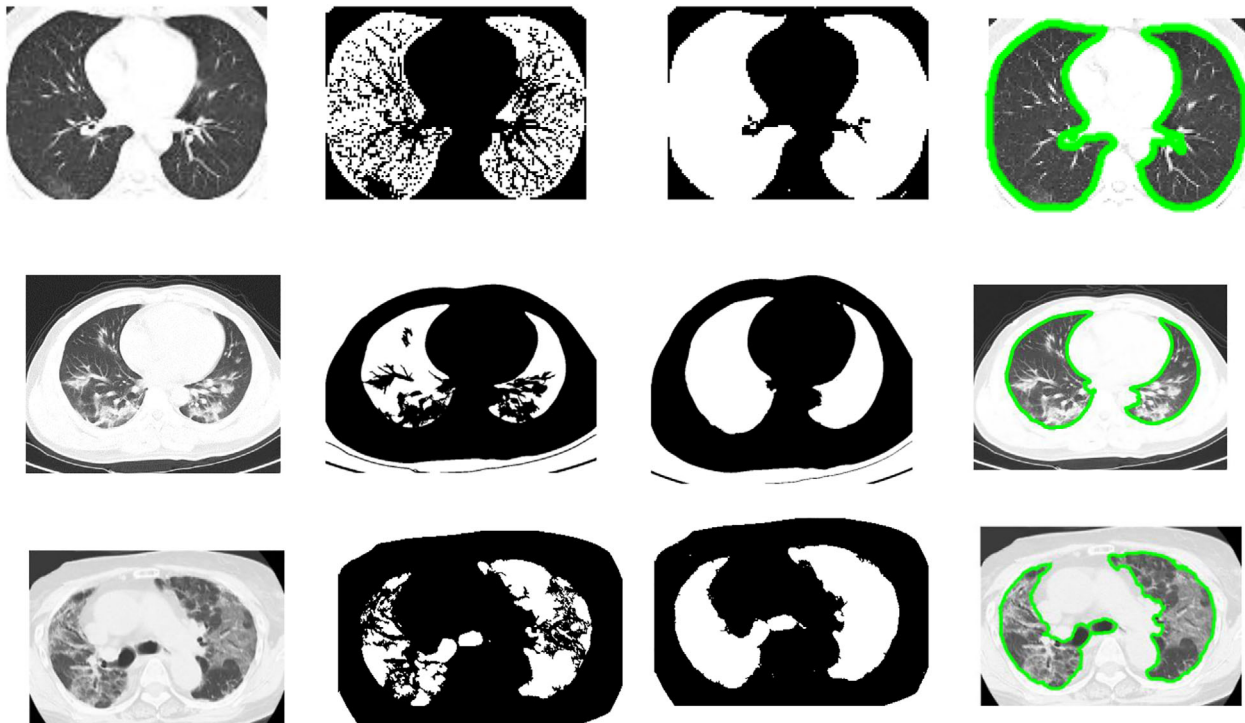

FIGURE 4 COVID-19 CT scan images segmented using the proposed model

3.3.1 | CDBN model

In the DBN, a multilayer of RBM is superimposed on top of each other to extract deep features from the image. The joint potential allocation between input data v and l-layer hidden layer h^k in the visible layer is illustrated in Equation (2). The unsupervised greedy method is utilized to calculate the weight of the data. In the beginning, the first layer of the RBM is trained in order to fix the first layer training parameters. The output of the first RBM layer's hidden layer is then used as the input for the second layer's RBM, and the first layer's parameters are progressively trained. The softmax regression classifier is connected to the final hidden layer, and the supervised gradient descent method completes the fine-tuning (Almanaseer et al., 2021; Dai et al., 2020).

$$P\left(v,h^1,h^2,...,h^l\right) = \left(\sum_{k=0}^{l-2} P\left(h^k|h^{k+1}\right)\right)P\left(h^{l-1},h^l\right) \tag{2}$$

$P\left(h^{l-1},h^l\right)$ has been defined as the probability allocation between the visible and hidden layers of the topmost RBM.

Figure 5 illustrates the procedure of RBM training. Set X of training samples is given after initialization. The RBM network structure has k visible layers (v), with the visible layers only being influenced by the jth hidden layers after start-up. After each parameter has been initialized, the

period of training and learning rate are given at the same time. The comparative distribution method is used to modify the training parameters. Assuming the process is successful, the output will continue; otherwise, the parameter training will continue on the basis of Equations (3 to 5). Figure 6 shows the training procedure of DBN. The first layer of RBM is trained using Figure 5 after initialization. In addition, the parameters and hidden layers are shown in Figure 6. In this manner, the hidden layer is used to train the first layer of RBM and used as input to the second layer, and the training is repeated until the last layer of RBM is trained. Finally, the output of this layer is connected to the softmax regression as classification classifier, which has been fine-tuned.

$$W = W + \in \left(h_1 x_1' - Q(h_2 = 1|x_2)x_2'\right) \quad (3)$$

$$b = b + \in (x_1 - x_2) \quad (4)$$

$$c = c + \in (h_1 - Q(h_2 = 1|x_2)) \quad (5)$$

where ε indicates the learning rate, $Q(h_2 = 1|x_2)$ it is a vector consisting of $Q(h_{2i} = 1|x_2.)$

Convolutional restricted Boltzmann machines (CRBMs) are commonly used in CDBNs. Instead of using filters to explore an item, CNNs employ them to establish connections between the layers. Neurons of the CNN are not fully connected to each other, whereas the DBN architecture, each neuron of visible layer are connected to each neuron of hidden layer. The nodes in the visible layer are not connected to any other visible nodes. Similarly, no hidden nodes are not connected to any other hidden nodes in the graphical topology of a simple RBM that have undirected connections with one another. As a result, the feature extractors such as CDBN has widely utilized in pattern recognition in the resent years due to their ability to create hierarchical feature structures. Using a CDBN model, it is possible to make efficient probabilistic inferences both from the bottom up and from the top down. Several layers of max-pooling CRBMs are located on the top of this structure, and training procedure is achieved using the greedy layer-wise method, such as in a traditional DBN. The system learns high-level features including stroke groups or object parts by constructing a CDBN. In context of system's tests, two layers of CRBM were used to train the CDBN, and feed-forward approximation was utilized for inference. The CDBN is built on the top of CRBM. It is possible to train the CDBN approach by performing a series of CRBMs, each CRBM feeds into another one. Figure 7 shows the structure of CDBN. The visible and hidden layers are connected by sets of local and shared parameters of CRBM's architecture. Binary-valued or real-valued units can be shown; however, binary-valued units are hidden (Elleuch et al., 2015; Jaisakthi et al., 2018).

This study uses three convolutional layers as well as three different max-pooling layers. The size of the kernel window is determined as 2×2 also in each layer the filter number is increased in order to contain a more complex pattern of the image in the training. For testing, this study used the image with size 128×128 with batch size 200. In Table 2 the parameters of the used CDBM model have been summarized.
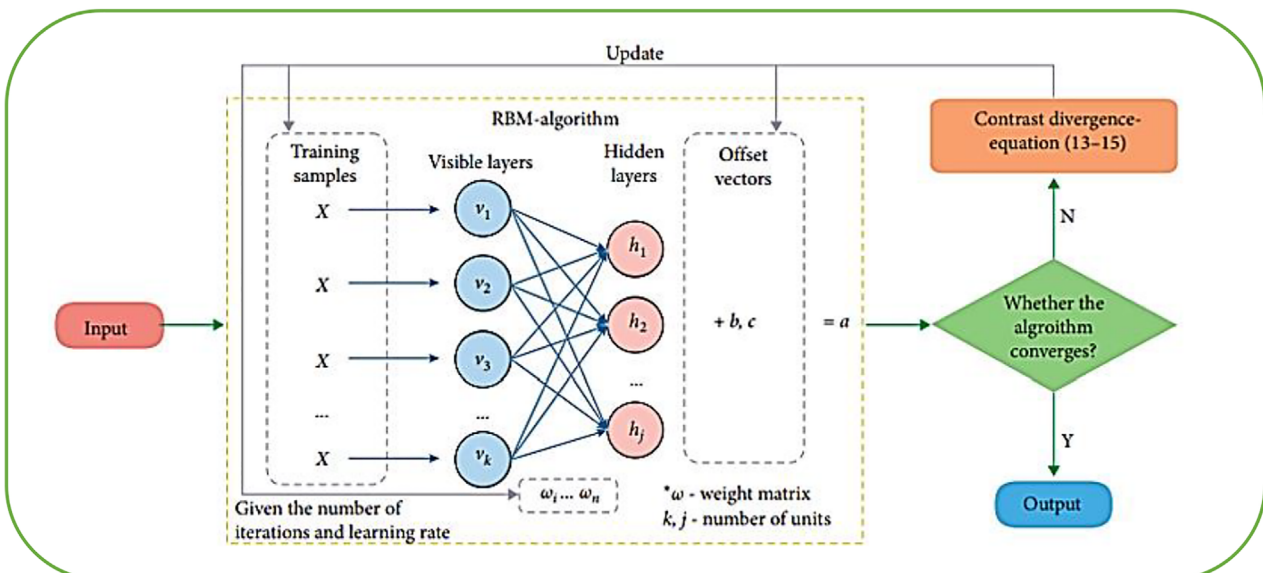

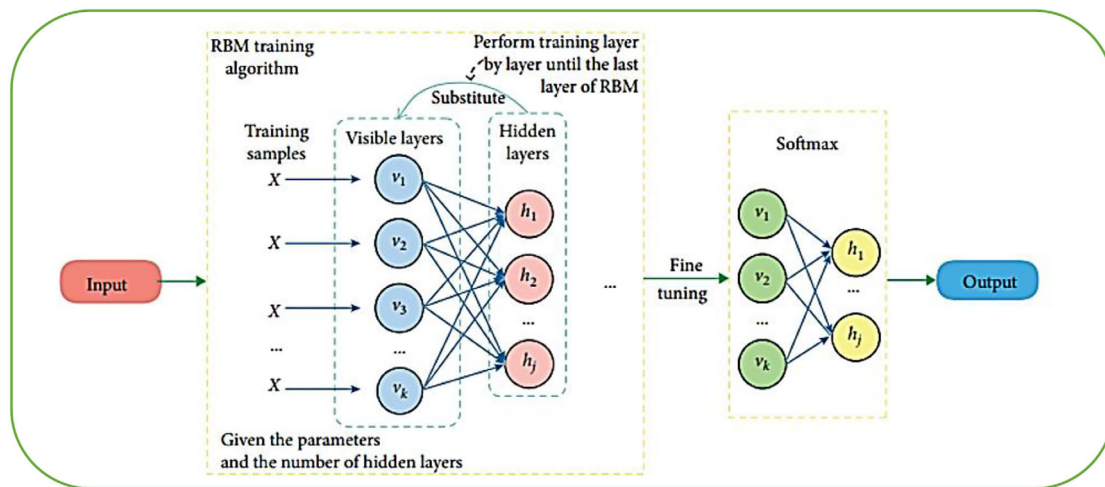**FIGURE 5** The training of RBM method (Almanaseer et al., 2021)

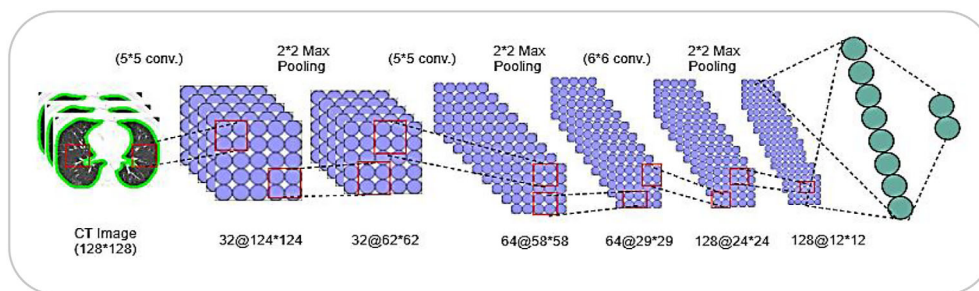**FIGURE 6**   The training of DBN model (Almanaseer et al., 2021)

**FIGURE 7**   Structure of CDBN model

3.3.2 | HRNet model

A HRNet comprises several phases, the first of which is the establishment of a high-resolution subnetwork. Following this, the addition of individually high-to-low-resolution subnetworks is carried out in order to build additional phases, accompanied by the connection of the multiresolution subnetworks in tandem (Seong & Choi, 2021). The information obtained from other parallel representations is fed back into each of the high-to-low-resolution representations on an ongoing basis as it performs repeated multi-scale fusions, resulting in very rich high-resolution representations. Because of this, the low-to-high process, as the name suggests, strives to produce high-resolution representations, whereas the high-to-low process attempts low-resolution and low-level representations. It is possible that the two processes will be repeated multiple times in order to increase the overall performance, but this is not guaranteed. Network design patterns that are representative of this type involve (i) symmetric processes from high to low and from low to high. High-to-low processes are designed to be a mirror image of one other, as is the case with the hourglass and its offspring. (ii) High-to-low weighting with heavy low-to-high weighting is used. By contrast, the low-to-high process is essentially a few bilinear upsampling or transposition convolution layers, whereas the high-to-low process is based on the ImageNet classification network (e.g., ResNet). (iii) Dilated convolutions are used in conjunction with the combination.

As shown in Figure 8, the fundamental design of HRNet comprises four stages, each of which contains four subnetworks that are interconnected in parallel. The resolution of subnetworks is reduced to half of its original value, whereas the set of feature mappings (connections) is gradually increased to twice of its original value. It consists of four residual blocks that were generated with 64 channels of a bottleneck at the beginning (width). This is continued by the use of one convolution operation to limit the width of feature maps to C = 32 by employing a kernel with a pixel size of (3 9 3). The remaining three stages are made up of (1, 4, and 3) multi-resolution blocks (exchange blocks), which are arranged in the following order: Each exchange block is made up of four residual blocks, each of which has two convolutional layers of (3 9 3) pixels, with each block consisting of four residual blocks.There were four resolution feature representations used in this study, and they were fed into a bottleneck to increase their width (the number of channels) to 128, 256, 512, and 1024 in order to implement the classification task on the pretrained HRNet model on the COVID-19 CT scan dataset. These down-sampled high-resolution maps were then added to the second high-resolution representation by executing two stride convolutional processes on the (3 9 3) pixel widths, which resulted in 256 widths. To obtain 1024 channels at

TABLE 2　Specific parameters of CDBN model

Parameters	Input size
Input size	128×128
Number of layers	2
1st layer of conv kernel	7×7 [$32 \times 124 \times 124$]
Max pooling	2×2 [$32 \times 62 \times 62$]
2nd layer of conv kernel	5×5 [$64 \times 58 \times 58$]
Max pooling	2×2 [$64 \times 29 \times 29$]
3rd layer of conv kernel	6×6 [$128 \times 24 \times 24$]
Max pooling	2×2 [$128 \times 12 \times 12$]
Batch size	200
Epoch	20
Learning rate	0.001
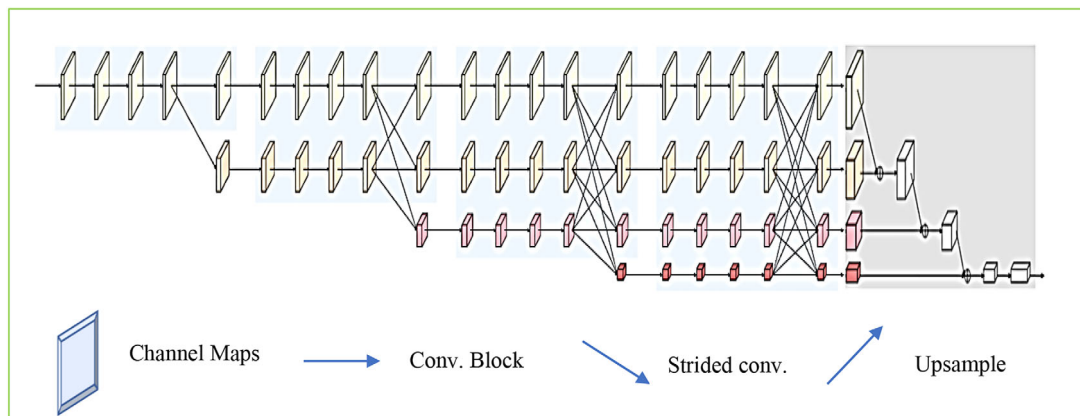
FIGURE 8　The HRNet model

a low resolution, the same technique was executed twice. Eventually, a 2048-dimensional vector was created by applying a convolutional operation to 1 9 1 pixels on top of 1024 channels and then applying an average pooling layer on top of that. In order to make a final choice and allocate the image representation to one of the anticipated classes, this extracted feature vector was passed into the softmax classifier.

3.3.3 | VGGNet

The VGGNet is a DL model-based multi-layered process that is used to recognize geometric shapes. Using the COVID-19 CT image dataset, the VGGNet model, which is based on the CNN model, is implemented in (Haque & Abdelgawad, 2020; Mateen et al., 2019). VGG-19 is beneficial because of its simplicity as 3×3 convolutional layers which are placed on top of each other to increase in depth as the depth level increases. Layers of max pooling were utilized as a handler in VGGNet in order to lower the volume size. With 4096 neurons, two FC layers were utilized. Figure 9 shows that the segmented COVID-19 images were utilized as input data for the VGGNet deep neural network. This study utilized convolutional layers as feature extractors in the training stage. The dimensionality reduction of features was performed on the basis of connecting max-pooling layers with different convolutional layers. For extracting deep features from images, 64 kernels with a size filter of 3×3 were utilized by the first convolutional layer. The feature vector was created by combining layers that were fully connected. Eventually, during the testing stage, 10-fold cross-validation was carried out to identify the positive and negative cases of COVID-19 on the basis of the softmax activation classifier. The VGGNet model was compared with existing feature extraction architectures, namely, CDBN, HRNet, and the proposed fusion model, to determine their performance.

　　Moreover, the parameter values that are utilized in VGGNet model are shown in Table 3. The VGGNet network structure described in Figure 9 is a model to extract deep significant features from CT scan images of Covid-19. This network has used several layers during the features extraction namely, convolutional layer, max-pooling layer, and fully connected layer. The image input size of 128×128 was used, the number of
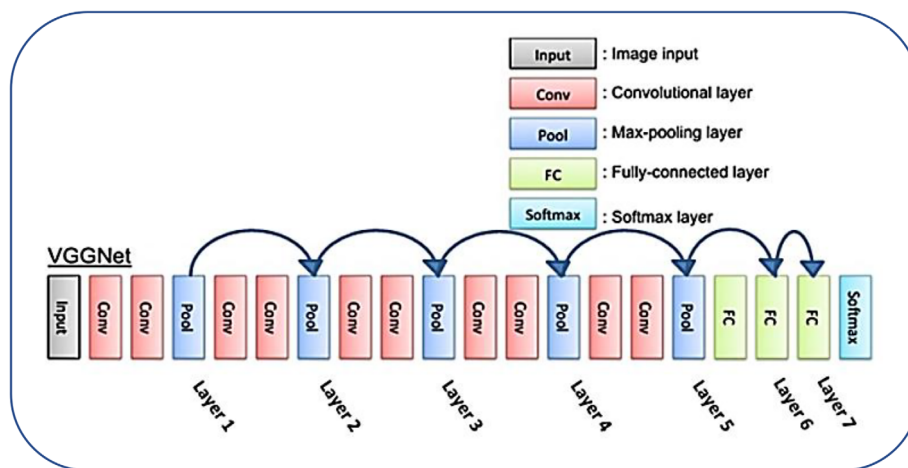
FIGURE 9   The VGGNet model

TABLE 3   Specific parameters of CDBN model

Parameter	Input value
Size of input image	128 × 128
Number of conv layers	10
Maxpooling layers	5
FC layers	3
Batch size	150
Number of epochs	200
Hidden layer size	8–96 neurons
dropout	0.1
Learning rate	0.0001
Activation function	ReLu softmax
Loss function	Cross entr.
Optimizer	SGDM
Kernel size	2 × 2

batch size has identified as 150 with 200 epochs. Finally, the ReLu softmax is used activation function and SGDM optimizer is used in the last convolution layer.

3.4 | Decision making (fusion)

Several studies have successfully implemented fusion in a variety of applications over the past several years, including fingerprint recognition, irises, facial expression, and hand geometry. Overall, four different levels of fusion have been examined, which are pixel-level fusion, feature-level fusion, score-level fusion, and decision-level fusion. Pixels can be fused at the pixel level by applying some form of processing to the source images (e.g., edge extraction or texture analysis) to create an abstract meta-image from which features for image classification can be extracted. This meta-image can then be used to classify the original images. However, the feature level of fusion brings together different feature vectors that have been produced through the use of numerous feature extraction methods. Frequently, the feature vectors are combined in an individual feature vector, or some type of feature combination is used to combine the feature vectors. In score-level fusion, the classification scores are aggregated into one final score that decides the class even further. Basically, there are three types of score-level fusion: level fusion using transformation, classifier, and density methods. Furthermore, it is possible to integrate classification decision outputs from several classifiers or feature vectors into a single final decision by utilizing decision-level fusion, which can be used to make a final conclusion. In this work, decision-level

TABLE 4 The decision level fusion for all possible cases

Cases	CDBN	HRNet	VGGNet	Decision-level fusion
Case 1	Positive	Positive	Positive	Positive
Case 2	Positive	Positive	Negative	Positive
Case 3	Positive	Negative	Positive	Positive
Case 4	Negative	Positive	Positive	Positive
Case 5	Negative	Negative	Negative	Negative
Case 6	Positive	Negative	Negative	Negative
Case 7	Negative	Positive	Negative	Negative
Case 8	Negative	Negative	Positive	Negative

fusion was applied by combining the decision outcomes of the DL models (CDBN, HRNet, and VGGNet model). The decision of these three models will be calculated as explained in Table 4:

4 | EXPERIMENTS RESULTS

This section outlines our experimental results and demonstrates the effectiveness of our fine-tuned networks through a series of trials. In the beginning, we described the used CT imaging databases. In the next section, we determine the experimental states and criteria that were used in performance evaluation. As a third step, we use the results of each DL model applied and our proposed model to our dataset. It is in fact a difficult issue to train DL models with limited data without overfitting. Three DL networks were tested to see how well the ensemble architecture is used and performs. The proposed DL model resolves the problem with considerable evaluation performance by fusing VGG-16, CDBN, and HRNet models. Therefore, a new model fusion procedure is then created using these models to significantly improve detection performance. The model utilized in this experiment can be considered as an effective model that incorporates all of the aforementioned foundational learners. A single layer of input replicates and transmits the input data into three levels of base learners in each of the system's three networks. Every base learner receives these input images individually, and then an extracted feature is constructed from each base learner that predicts the input labels. All three basic learners produce feature maps, which are then input to a classifier. Afterward, the softmax attempts to efficiently categorize the input data into the relevant classes in an effective way. Only the softmax is used in the developed framework, and it is trained solely on the training set's data. When training the softmax, we used an initial learning rate of 0.0003, 100 epochs per batch size of 32, and 100 epochs of training data on each base model. Numerous experiments were conducted to verify the validity of the suggested research project. For the networks, we used MATLAB (2020b) and a computer with a Core i7 processor, 32 GB of RAM, and Windows 10 operating system.

Five metrics were utilized in the evaluation process to assess the performance of the model. For every classification, the performance measures including accuracy, sensitivity (recall), precision, F1-score, and specificity were calculated. The prediction results were qualified as true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Accuracy measures how well a prediction has performed in comparison to all other predictions. TP predictions are represented by the recall, which is the ratio of TPs to FPs. In statistics, precision is defined as the percentage of correctly predicted events versus all predicted events. An F1-score is derived by taking the harmonic mean of their precision and recall scores together. Accurate positive predictions versus FPs are known as the specificity ratio. Formula (6 to 10) gives the criteria used in this study. Figure 10 shows the relationship between the sensitivity and specificity using the FP, FN, TP, and TN.

$$Recall = Sensitivity\,(SN) = \frac{TP}{TP + FN} \tag{6}$$

$$Specificity\,(SP) = \frac{TN}{FP + TN} \tag{7}$$

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

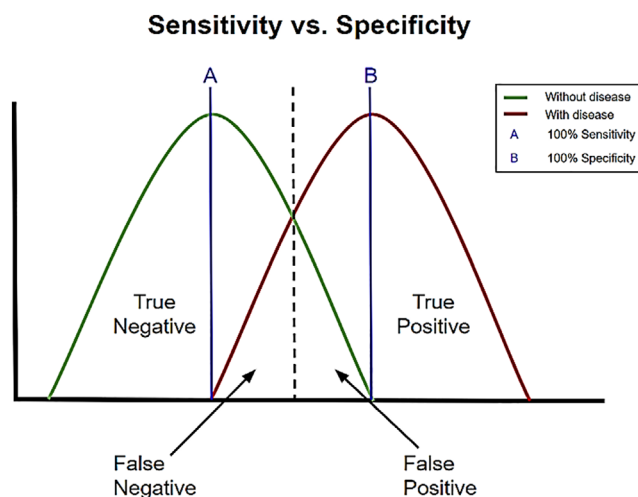$$F1 - score = 2X\frac{Precision + Recall}{Precision + Recall} \tag{9}$$

Sensitivity vs. Specificity

FIGURE 10 Comparison the data of the sensitivity and specificity (Serrao et al., 2018)

$$\text{Accuracy (AC)} = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

TP represents the subjects accurately identified in a present (positive) class, FN reflects the individuals who were incorrectly assigned to the opposite (negative) category, FP indicates the misclassified individuals who were incorrectly assigned to the predefined (positive) group, and TN refers to subjects that were classified correctly in the other (negative) class.

4.1 | Data augmentation

We trained and tested our proposed model on COVID-19 CT scan images. This dataset consists of 397 and 349 CT images as COVID-19 and non-COVID-19 cases, respectively. These CT scan images were taken from 216 patients with various height and width (minimum 153 and 124, average 491 and 383, and maximum 1853 and 1485, respectively). Patients with a positive test result had age and gender information of 169 and 137, respectively. There are 86 more male patients than female patients in the group. Because our proposed model uses DL networks, the number of images in training and testing has a substantial impact on its performance. Thus, the size of the dataset is a crucial factor; data augmentation is required for neural network training in order to attain high generalizability.

To efficiently map a specific input to output, large amounts of data are required in computer vision. The volume of data has a significant impact on the learning process, particularly for DNN training. The amount of data in the COVID-19 CT dataset is too small for DNN training. It is necessary to increase the size of our dataset because we only have a small amount of training data. In this study, data augmentation operations are used on the original images to obtain more data. It makes no difference whether a crack is oriented in a positive or a negative direction for classification purposes in this example. Cracks can form on either the upper or bottom edge; it makes no difference. As a result, many strategies for data augmentation can be successfully used. In this study, data augmentation is achieved through the use of parameterized transformations. Prior to CNN training, we performed offline data augmentation on the training set. To increase the variety of defect patterns, we rotated all images to 90°, 180°, and 270° before adding them to the original dataset. By rotating by the same number of degrees, we can simply create the corresponding mask. Because scaling has some relationship to rotation, we simply rotate the image by 0.5 and 1.5 times to generate an additional 200 images and masks. Table 5 displays the distribution of the datasets that were utilized.

4.2 | Classification result

In this paper, we present and analyse the findings obtained by employing our fine-tuned deep networks to detect COVID-19 on the CT image datasets. We present the quantitative results, as well as the confusion matrices, for each and every architecture of the networks that were used in this study. First, we applied our proposed segmentation method. The main objective of the segmentation stage is to crop the ROI and extract the texture inside the ROI. Figure 11 shows the steps of the segmentation method. Table 4 shows the average values of performance metrics obtained by our diverse networks on each CT image dataset, as calculated by our algorithms. When applicable, we also compare our findings to

TABLE 5   Total number of dataset in training and testing after data augmentation

COVID-CT dataset	Original data	Original data and augmentation	Training	Testing
Positive cases	349	1396	1116	280
Negative CASES	397	1588	1268	320
Total	746	2984	2384	600
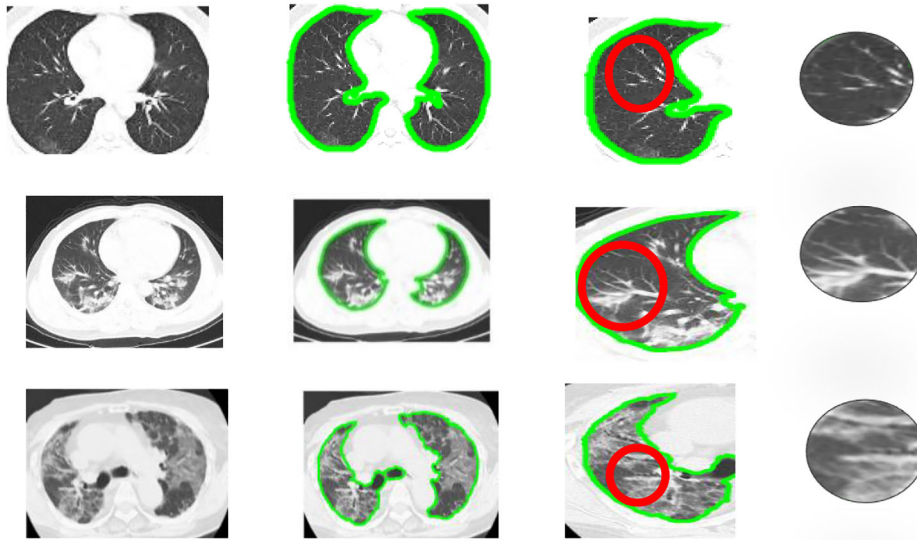
FIGURE 11   Segmentation steps of the proposed method: a) original image, b) segmented image, c) extracted ROI and d) the texture sign inside the ROI that can help to identify the positive COVID-19 cases

those that have been previously proposed in the literature. When we compared our acquired models to equivalent models from a recently published paper, we found that our models were significantly superior.

According to our dataset, the overall performance of a network in terms of evaluation criteria varies from one network to the next in terms of assessment metrics. This can be attributed to the fact that the CT scans in the dataset are heterogeneous across different sources. Because of the difficulty in distinguishing between COVID-19 and other findings associated with lung diseases, the non-COVID-19 CT images were obtained from a variety of sources and revealed a wide range of findings. In addition, there are significant differences between CT images in the COVID19-CT dataset when it comes to contrast, spatial resolution, and other visual properties, all of which could impair the model's potential to extract more discriminative and generalizable features. Our proposed models obtained fairly good evaluation results compared with the recent state-of-the-art works utilizing the exact dataset. To summarize the results of the three algorithms for the prediction of COVID-19 CT images, the accuracy, sensitivity, specificity, precision, and F1-score measures obtained from the three models and our proposed model are shown in Table 6. With respect to all evaluation metrics, the CDBN and HRNet models obtained nearly identical outcomes (Table 6). We can observe that the CDBN model outperforms the HRNet pretrained network in terms of accuracy, specificity, and precision. However, HRNet obtained better sensitivity, and both networks had similar F1-score. On the other hand, the VGGNet model obtained lower results compared with the CDBN and HRNet models.

To improve the prediction model for COVID-19, we stacked all three models based on fusion into one. In accordance with the results shown in Table 6, a proposed fusion model of the previously employed models was developed. Table 6 shows that the aggregate of models has a higher recognition performance than each model independently. Using all three models together resulted in a recognition accuracy of 95%. Moreover, our proposed deep model achieved a sensitivity, specificity, precision, and F1-score of 95%, 96%, 96%, and 95%, respectively. Table 6 shows that compared with CDBN and HRNet, VGGNet does not perform well when taking into account all of the criteria. To increase the stacking model's effectiveness and reduce prediction error, we used a FC neural network as a meta-learner. Because the neural network incorporates all of the previous predictions, the total model's performance is improved. The neural network is fine-tuned to disregard incorrect predictions given by the underlying models and to use only predictions that increase the classification rate. Perpetrate errors could be made by various DL techniques on various samples. Therefore, combining various DL techniques based on fusion can assist in producing better performance. As shown in Table 6,

TABLE 6 The deep learning model compared to the proposed fusion model

	Whole data	COVID	Non-COVID	Training-COVID	Training non-COVID	Testing-COVID	Testing non-COVID		AC	SN	SP	Precision	F-score
								CDBN					
Original Data	746	349	397	279	317	70	80		0.793	0.814	0.775	0.76	0.78
Augmented Data	2984	1396	1588	1116	1268	280	320		0.917	0.904	0.92	0.91	0.91
	Whole data	COVID	Non-COVID	Training-COVID	Training Non-COVID	Testing-COVID	Testing Non-COVID	**HRNet**					
									AC	SN	SP	Precision	F-Score
Original Data	746	349	397	279	317	70	80		0.78	0.78	0.78	0.76	0.77
Augmented Data	2984	1396	1588	1116	1268	280	320		0.915	0.92	0.90	0.89	0.91
	Whole data	COVID	Non-COVID	Training-COVID	Training Non-COVID	Testing -COVID	Testing Non-COVID	**VGGNet**					
									AC	SN	SP	Precision	F-Score
Original Data	746	349	397	279	317	70	80		0.69	0.72	0.66	0.65	0.68
Augmented Data	2984	1396	1588	1116	1268	280	320		0.88	0.88	0.88	0.86	0.87
	Whole data	COVID	Non-COVID	Training-COVID	Training Non-COVID	Testing -COVID	Testing Non-COVID	**fusion (CDBN, HRNet and VGGNet)**					
									AC	SN	SP	Precision	F-Score
Original Data	746	349	397	279	317	70	80		0.82	0.82	0.82	0.80	0.81
Augmented Data	2984	1396	1588	1116	1268	280	320		**0.95**	**0.95**	**0.96**	**0.96**	**0.95**

the proposed fusion model correctly classifies CT images that were originally categorized as positive COVID-19 but were incorrectly classified by one model, namely, CDBN, HRNet, and VGGNet, as negative COVID-19.

In medical research, particularly for life-threatening disorders (noncommunicable diseases) such as COVID-19, it is crucial to decrease the FP and FN as much as possible when developing a prediction system. FN should be kept as low as possible because when positive cases of COVID-19 are incorrectly classified as negative cases, it may result in the reoccurrence of otherwise preventable deaths. Furthermore, the misclassification of negative COVID-19 cases as positive COVID-19 (FP) may result in the reoccurrence of otherwise preventable deaths. As a result, it is also necessary to reduce the number of FP cases that occur. As demonstrated in Table 7, the confusion matrices clearly illustrate the significant difference in the performance of these models when tested against one another on a given test set. As shown in the table, the likelihood of inaccurate classification is substantial when only one model is used. Although only 14 positive COVID-19 samples are wrongly identified as non-COVID-19 in the final confusion matrix of the proposed DL model based on the fusion process, 11 non-COVID-19 cases are incorrectly labelled as COVID-19 cases in the model's final confusion matrix. Compared with other models, this can be regarded as a major improvement.

Receiver operating characteristic (ROC) is a curve measurement that measures the problem of classification performance based on the number of threshold values. ROC is known as a curve probability that is able to measure the separability degree between two different categories such as positive and negative cases. The scoring higher ROC result is considered as a better performance classification. This study uses ROC to measure the classification performance between COVID-19 and non-COVID-19 using different deep learning models as well as the proposed fusion model. This study scored the better performance for any classification networks at classifying COVID-19 as COVID-19 and non-COVID-19 as non-COVID-19. Figure 12 illustrates the achieved ROC of CDBN, HRNet, VGGNet, and proposed models. A competitive analysis among well-known deep learning models as well as proposed model based on measuring true positive rate and false-positive rates. It is clearly illustrating that the proposed model outperformed in ROC values as compared to other deep learning models.

This study presents an efficient comparison of the most recent state-of-the-art works on COVID-19 CT (Table 8). The study of (Pathak et al., 2020) achieved an accuracy of 93.01% using ResNet-32. By contrast, all other studies obtained an accuracy ranging from 80.3% to 88.3%, with the study of (Wang et al., 2020) achieving the lowest accuracy (78.6%). As a result, the baseline study illustrates that even pre-existing models can produce remarkable performance. As depicted in Table 8, we achieved 95% performance accuracy with our proposed fusion model after incorporating data augmentation. According to our results, our models performed reasonably well compared with a recently published study that utilized the exact dataset. However, developing a classification-based model is still considered as open research in this field to improve the performance of accuracy.

The whole proposed framework includes two main stages to achieve the identification objective: segmentation and identification. The segmentation stage is implemented by enhancing the GrabCut mechanism, where the entail mask is generated by a no-threshold histogram-based image segmentation method and uses it as a binary image for the GrabCut method. The limitation of this stage are as follows: First, the ROI in some cases have irregular objects. In this case, we cannot obtain the whole ROI as one object, which confuses the model into detecting the ROI from the FP objects as shown in Figure 13. Second, the texture of the ROI in some cases overlaps with the background, resulting in an over- and under-segmentation problem.

Despite achieving interesting performance results for the identification stage, this work still has substantial limitations. First, the similarity between the negative and positive cases confuses the proposed framework into identifying the positive cases as shown in Figure 14. Therefore, we should conduct more investigations to obtain a more effective model that can identify the positive cases from the negative.

Second, despite the fact that various alternative networks obtained outstanding results in two networks, the best performing solution for the COVID-19 class relies on CDBN, HRNet, and VGGNet architectures. Improving the networks' features, which take their properties into account, could ameliorate results, especially in terms of improving the ability to differentiate between various classes. Third, no pre-processing phase was applied under any of the experimental conditions. To provide a comprehensive model in the future, pre-processing techniques such as contrast enhancement and denoising in images should be employed. A final round of experiments focused on patients verified CDBN and HRNet positive outcomes despite some patients being misclassified when compared to the findings of VGGNet. It is vital to put out effort in this area in order to better understand the CT scan that is critical in this critical circumstance. Finally, because of the way they are arranged, the COVID-19 class is

TABLE 7 Obtained results-based confusion matrix

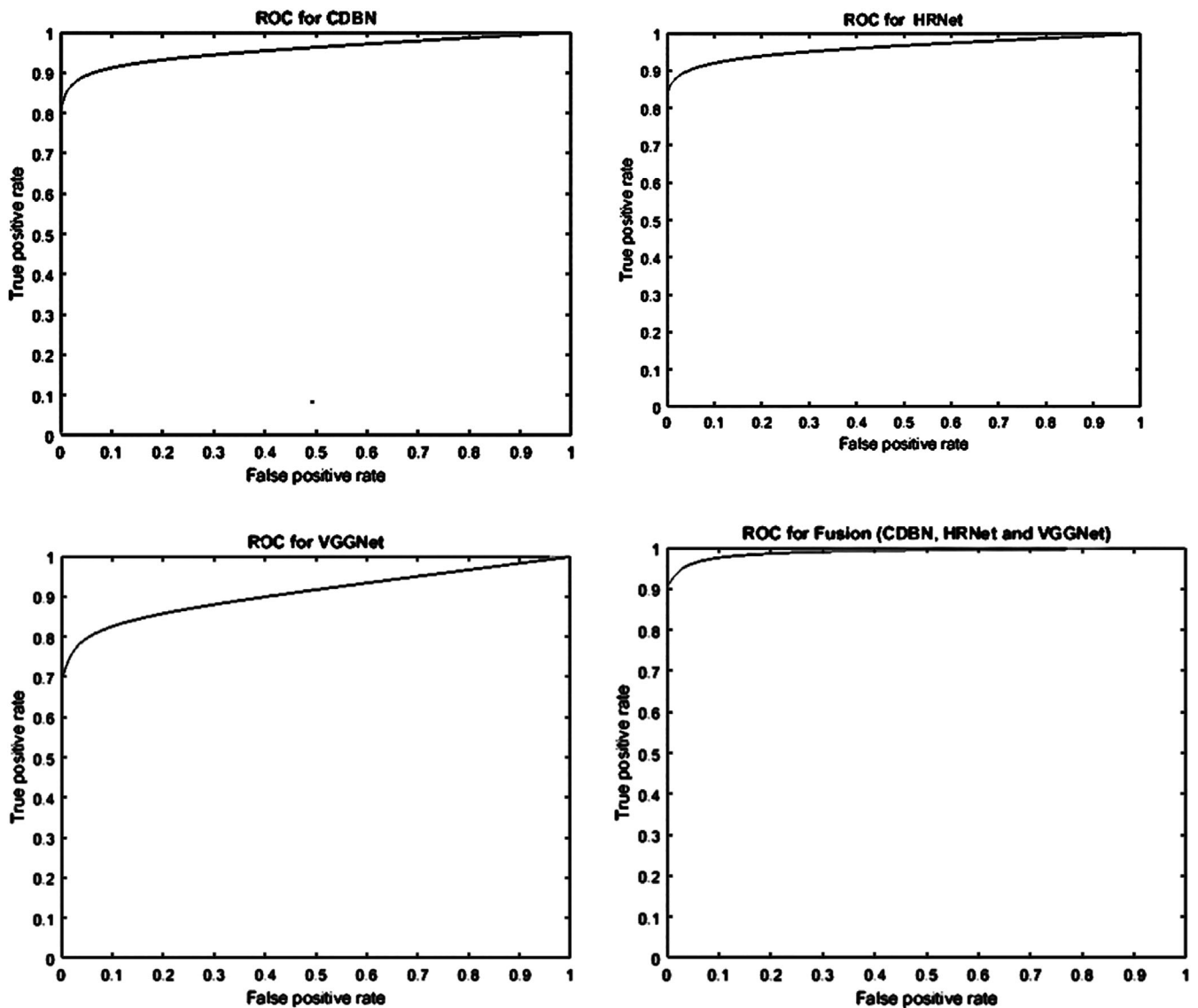| Model	Whole data	COVID	Non-COVID	Training-COVID	Training non-COVID	Testing -COVID	Testing non-COVID	TP	FN	TN	FP
CDBN	2984	1396	1588	1116	1268	280	320	253	27	297	23
HRNet	2984	1396	1588	1116	1268	280	320	259	21	290	30
VGGNet	2984	1396	1588	1116	1268	280	320	249	31	282	38
Proposed model	2984	1396	1588	1116	1268	280	320	266	14	309	11

FIGURE 12 The analysis of different deep learning and proposed fusion model using receiver operating characteristic (ROC)

TABLE 8 Comparison between proposed model and recent state-of-art studies

Study	Network	AC	SN	SP	Precision	F1-score
(Mobiny et al., 2020)	DECAPS+Peekaboo	87.6	91.5	85.2	84.3	87.1
Pathak et al. (2020)	ResNet-32	93.01	91.4	94.7	95.1	-
Dey et al. (2020)	Feature fusion +KNN	87.75	-	-	-	-
He et al. (2020)	DenseNet169	83	-	-	-	81
Mishra et al. (2020)	Decision function	88.3	-	-	-	86.7
Saqib et al. (2020)	ResNet101	80.3	85.7		78.2	81.8
Shamsi et al. (2021)	DenseNet121 + SVM	85.9	84.9	86.8	-	-
Martinez (2020)	DenseNet169	87.7	85.6		90.2	87.8
Wang et al. (2020)	Contrastive Learning	78.6	79.7	-	78	78.8
Proposed Model	VGGNet + CDBN + HRNet	95	95	96	96	95

more difficult to distinguish from the others as indicated by the class-wise performance. Consequently, previous research has shown that hand-crafted descriptors and potentially the aggregation of diverse descriptors aid in the recognition of the most difficult cases.
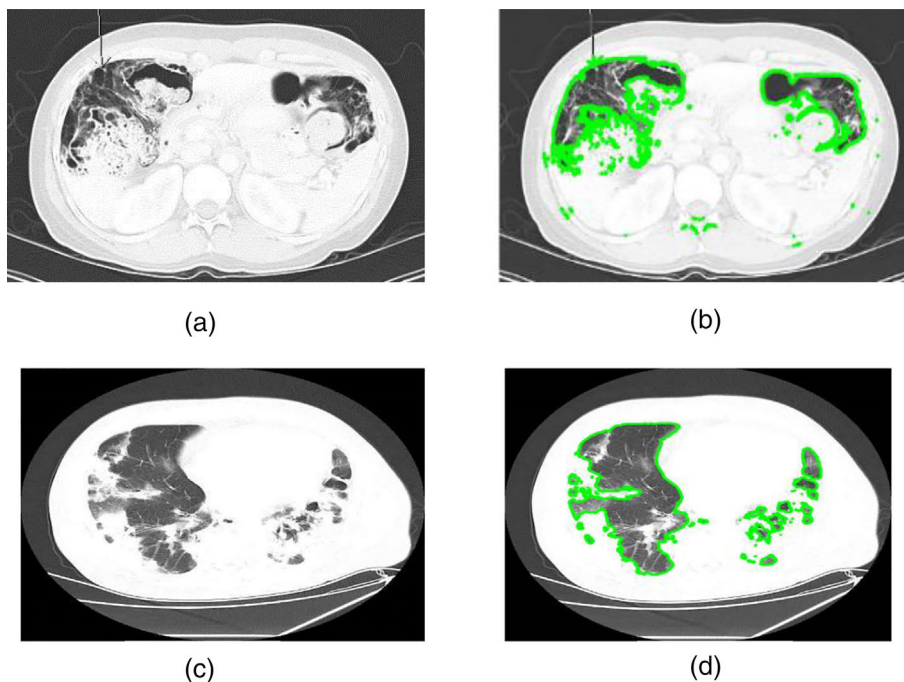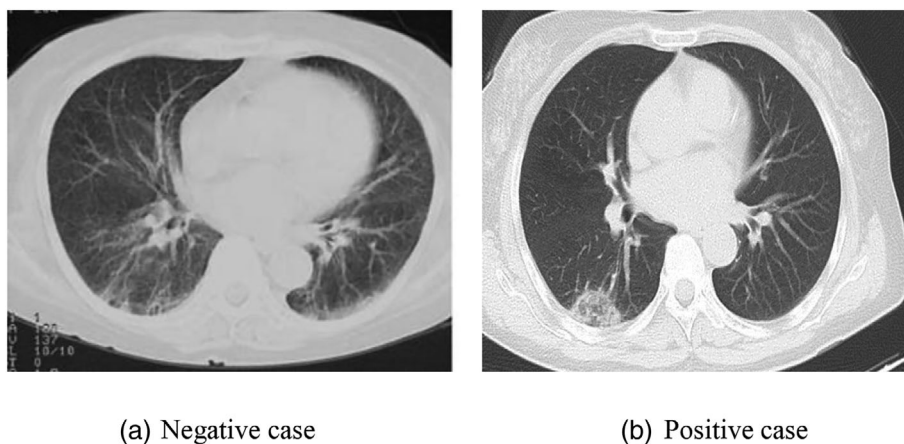
FIGURE 13 The outcome of the proposed segmentation method for the difficult cases, a and c shows the orginal Ct images, b and d shows the segmented image

(a) Negative case (b) Positive case

FIGURE 14 The similarity between the positive and negative cases for the COVID-19 CT image is shown. (a) Negative case and (b) positive case

5 | CONCLUSION AND FUTURE WORKS

COVID-19 examination can be done with the help of a chest CT scan, which is a quick and painless procedure. This study investigated the ability of implement a model that can help the expert to identify the positive COVID-19 cases. To achieve this objective, new powerful model has been utilized, namely, effective hybrid DL model, for identifying COVID-19 patterns using CT images. The proposed model includes two main stages: segmentation and identification stage. In the segmentation stage, the GrabCut is enhanced by combining it with a no-threshold histogram-based image segmentation method to extract the ROI and use it as a segmented image for the identification stage. In the identification stage, the VGGNet, CDBN, and HRNet models are used to diagnose COVID-19 cases, obtaining identification accuracies of 88%, 91.7%, and 91.5%, respectively. To increase the diagnosis capabilities, an ensemble-based learning or fusion strategy was suggested by combining the strength of all three models and attained an accuracy of 95%. Compared with numerous current literature approaches, the identification performance of the proposed model for COVID-19 was superior. However, we found a few instances where the results were incorrect because the overlapping or the similarity between the texture of the negative and positive cases. In the future, various datasets will be employed to train and evaluate the proposed model to ensure it is efficient and robust in order to ameliorate the model's diagnosis ability. Moreover, to improve the diagnosis quality of COVID-19,

more affective DL models will be used to extract more powerful deep features and produce various decisions on the basis of different fusion processes to increase the performance of the proposed model.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

We have trained as well as tested our proposed model on Covid-19 CT scan images. This dataset consists of 397 and 349 CT images as Covid-19 and non-Covid-19 cases, respectively. The dataset available online on: https://github.com/UCSD-AI4H/COVID-CT/tree/master/Images-processed

ORCID

Dheyaa Ahmed Ibrahim https://orcid.org/0000-0001-7357-0045

Dilovan Asaad Zebari https://orcid.org/0000-0002-7643-6359

Hussam J. Mohammed https://orcid.org/0000-0001-9925-3945

Mazin Abed Mohammed https://orcid.org/0000-0001-9030-8102

REFERENCES

Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., Tao, Q., Sun, Z., & Xia, L. (2020). Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases. *Radiology*, *296*(2), E32–E40.

Al-Karawi, D., Al-Zaidi, S., Polus, N., & Jassim, S. (2020). Machine learning analysis of chest CT scan images as a complementary digital test of coronavirus (COVID-19) patients. *MedRxiv*.

Allioui, H., Mohammed, M. A., Benameur, N., al-Khateeb, B., Abdulkareem, K. H., Garcia-Zapirain, B., Damaševičius, R., & Maskeliūnas, R. (2022). A multi-agent deep reinforcement learning approach for enhancement of COVID-19 CT image segmentation. *Journal of Personalized Medicine*, *12*(2), 309.

Almanaseer, W., Alshraideh, M., & Alkadi, O. (2021). A deep belief network classification approach for automatic Diacritization of Arabic text. *Applied Sciences*, *11*(11), 5228.

Alom, M. Z., Rahman, M. M., Nasrin, M. S., Taha, T. M., & Asari, V. K. (2020). COVID_MTNet: COVID-19 detection with multi-task deep learning approaches. arXiv preprint arXiv:2004.03747.

Alyasseri, Z. A. A., Al-Betar, M. A., Doush, I. A., Awadallah, M. A., Abasi, A. K., Makhadmeh, S. N., Alomari, O. A., Abdulkareem, K. H., Adam, A., Damasevicius, R., Mohammed, M. A., & Zitar, R. A. (2021). Review on COVID-19 diagnosis models based on machine learning and deep learning approaches. *Expert Systems*, *39*(3), e12759.

Barstugan, M., Ozkaya, U., & Ozturk, S. (2020). Coronavirus (covid-19) classification using ct images by machine learning methods. arXiv preprint arXiv: 2003.09424.

Basavaprasad, B., & Hegadi, R. S. (2014). Improved grabcut technique for segmentation of color image. *International Journal of Computers and Applications*, *975*, 8887.

Dai, X., Cheng, J., Gao, Y., Guo, S., Yang, X., Xu, X., & Cen, Y. (2020). Deep belief network for feature extraction of urban artificial targets. *Mathematical Problems in Engineering*, *2020*, 1–13.

Dey, N., Rajinikant, V., Fong, S. J., Kaiser, M. S., & Mahmud, M. (2020). Social-group-optimization assisted Kapur's entropy and morphological segmentation for automated detection of COVID-19 infection from computed tomography images. *Cognitive Computation*, *12*, 1011–1023.

Elleuch, M., Tagougui, N., & Kherallah, M. (2015). Deep learning for feature extraction of Arabic handwritten script. In *International conference on computer analysis of images and patterns* (pp. 371–382). Springer.

Fang, Y., Zhang, H., Xie, J., Lin, M., Ying, L., Pang, P., & Ji, W. (2020). Sensitivity of chest CT for COVID-19: Comparison to RT-PCR. *Radiology*, *296*(2) 200432. https://doi.org/10.1148/radiol.2020200432

Farid, A. A., Selim, G. I., & Khater, H. A. A. (2020). A novel approach of CT images feature analysis and prediction to screen for Corona virus disease (COVID-19).

Gozes, O., Frid-Adar, M., Greenspan, H., Browning, P. D., Zhang, H., Ji, W., A & Siegel, E. (2020). Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis. arXiv preprint arXiv:2003.05037.

Haque, K. F., & Abdelgawad, A. (2020). A deep learning approach to detect COVID-19 patients from chest X-ray images. *AI*, *1*(3), 418–435.

Hasoon, J. N., Fadel, A. H., Hameed, R. S., Mostafa, S. A., Khalaf, B. A., Mohammed, M. A., & Nedoma, J. (2021). COVID-19 anomaly detection and classification method based on supervised machine learning of chest X-ray images. *Results in Physics*, *31*, 105045.

He, X., Yang, X., Zhang, S., Zhao, J., Zhang, Y., Xing, E., & Xie, P. (2020). Sample-efficient deep learning for COVID-19 diagnosis based on CT scans. *IEEE Journal of Biomedical and Health Informatics*, *24*, 2806–2813.

Jaisakthi, S. M., Mirunalini, P., & Aravindan, C. (2018). Automated skin lesion segmentation of dermoscopic images using GrabCut and k-means algorithms. *IET Computer Vision*, *12*(8), 1088–1095.

Li, X., Fang, X., Bian, Y., & Lu, J. (2020). Comparison of chest CT findings between COVID- 19 pneumonia and other types of viral pneumonia: A two-center retrospective study. *European Radiology*, 1–9. https://doi.org/10.1007/s00330-020-06925-3

Lingayat, N. S., & Tarambale, M. R. (2013). A computer based feature extraction of lung nodule in chest x-ray image. *International Journal of Bioscience, Biochemistry and Bioinformatics*, *3*(6), 624–629.

Martinez, A. R. (2020). Classification of COVID-19 in CT scans using multi-source transfer learning. *arXiv*, arXiv:2009.10474.

Mateen, M., Wen, J., Song, S., & Huang, Z. (2019). Fundus image classification using VGG-19 architecture with PCA and SVD. *Symmetry*, *11*(1), 1.

Mishra, A. K., Das, S. K., Roy, P., & Bandyopadhyay, S. (2020). Identifying COVID19 from chest CT images: A deep convolutional neural networks based approach. *Journal of Healthcare Engineering*, *2020*, 1–7.

Mobiny, A., Cicalese, P. A., Zare, S., Yuan, P., Abavisani, M., Wu, C. C., Ahuja Jitesh, Groot de Patricia M., Van Nguyen, H. (2020). Radiologist-level covid-19 detection using CT scans with detail-oriented capsule networks. arXiv preprint arXiv:2004.07407.

Özkaya, U., Öztürk, Ş., & Barstugan, M. (2020). Coronavirus (COVID-19) classification using deep features fusion and ranking technique. In *Big data analytics and artificial intelligence against COVID-19: Innovation vision and approach* (pp. 281–295). Springer.

Pan, Y., Guan, H., Zhou, S., Wang, Y., Li, Q., Zhu, T., Hu, Q., & Xia, L. (2020). Initial CT findings and temporal changes in patients with the novel coronavirus pneumonia (2019-nCoV): A study of 63 patients in Wuhan, China. *European Radiology*, 1–4. https://doi.org/10.1007/s00330–020-06731-x

Pathak, Y., Shukla, P. K., Tiwari, A., Stalin, S., & Singh, S. (2020). Deep transfer learning based classification model for COVID-19 disease. *IRBM*, *43*(2), 87–92.

Qiblawey, Y., Tahir, A., Chowdhury, M. E., Khandakar, A., Kiranyaz, S., Rahman, T., … Ayari, M. A. (2021). Detection and severity classification of COVID-19 in CT images using deep learning. *Diagnostics*, *11*(5), 893.

Ren, M., Zhang, Q., & Zhang, J. (2019). An introductory survey of probability density function control. *Systems Science & Control Engineering*, *7*(1), 158–170.

Saqib, M., Anwar, S., Anwar, A., & Blumenstein, M. (2020). COVID19 detection from radiographs: Is deep learning able to handle the crisis? *TechRxiv*, 1–14.

Seong, S., & Choi, J. (2021). Semantic segmentation of urban buildings using a high-resolution network (HRNet) with channel and spatial attention gates. *Remote Sensing*, *13*(16), 3087.

Serrao, N. R., Reid, S. M., & Wilson, C. C. (2018). Establishing detection thresholds for environmental DNA using receiver operator characteristic (ROC) curves. *Conservation Genetics Resources*, *10*(3), 555–562.

Shariaty, F., Hosseinlou, S., & Rud, V. Y. (2019). Automatic lung segmentation method in computed tomography scans. In *Journal of Physics: Conference Series*(Vol. 1236, No. 1, p. 012028). IOP Publishing.

Shamsi, A., Asgharnezhad, H., Jokandan, S. S., Khosravi, A., Kebria, P. M., Nahavandi, D., … Srinivasan, D. (2021). An uncertainty-aware transfer learning-based framework for covid-19 diagnosis. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(4), 1408–1417.

Shi, F., Xia, L., Shan, F., Wu, D., Wei, Y., Yuan, H., Jiang H, He Y, Gao Y, Sui H, Shen, D. Large-scale screening of covid-19 from community acquired pneumonia using infection size-aware classification (2020). arXiv preprint arXiv:2003.09860.

Tang, Z., Zhao, W., Xie, X., Zhong, Z., Shi, F., Liu, J., & Shen, D. (2020). Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest CT images. arXiv preprint arXiv:2003.11988.

Thamilselvan, P., & Sathiaseelan, J. (2015). A comparative study of data mining algorithms for image classification. *International Journal of Education and Management Engineering*, *5*, 1–9.

Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., Cai, M., Yang, J., Li, Y., Meng, X., & Xu, B. (2021). A deep learning algorithm using CT images to screen for Corona virus disease (COVID-19). *European Radiology*, *31*(8), 6096–6104.

Wang, S., Zha, Y., Li, W., Wu, Q., Li, X., Niu, M., Wang, M., Qiu, X., Li, H., Yu, H., Gong, W., Bai, Y., Li, L., Zhu, Y., Wang, L., & Tian, J. (2020). A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. *European Respiratory Journal*, *56*(2), 2000775.

Wang, Z., Liu, Q., Dou, Q. Contrastive cross-site learning with redesigned net for COVID-19 CT classification.

World Health Organization. (2020). *Coronavirus disease (COVID-19) pandemic*. World Health Organization; Retrieved May 25, 2020, from https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Xie, X., Zhong, Z., Zhao, W., Zheng, C., Wang, F., & Liu, J. (2020). Chest CT for typical 2019-nCoV pneumonia: Relationship to negative RT-PCR testing. *Radiology*, *296*(2), 200343. https://doi.org/10.1148/radiol.2020200343

Xu, B., Xing, Y., Peng, J., Zheng, Z., Tang, W., Sun, Y., … & Peng, F. (2020). Chest CT for detecting COVID-19: A systematic review and meta-analysis of diagnostic accuracy. *European Radiology*, *30*(10), 5720-5727.

Zebari, D. A., Abdulazeez, A. M., Zeebaree, D. Q., & Salih, M. S. (2020, December). A fusion scheme of texture features for COVID-19 detection of CT scan images. In *2020 international conference on advanced science and engineering (ICOASE)* (pp. 1–6). IEEE.

AUTHOR BIOGRAPHIES

Dheyaa Ahmed Ibrahim is a scientific researcher with more than 8 years' experience in the field of Computing and awarded PhD in Computing from the University of Buckingham. He received his BSc and MSc in computer science from University of Anbar, Iraq. He has experience in coding, designing, and carrying out experiments, testing different types of cancer using AI solutions, and writing technical reports. His research interest is in the field of Computing/Artificial intelligence, Machine & Deep Learning. The project that he has been working on it since 2013, is to investigate, develop effective and novel computer-based solutions in automatically classifying medical Images based on different image processing and machine learning techniques.

Dilovan Asaad Zebari received the BSc degree in computer science from the University of Duhok (UoD), Kurdistan Region, Iraq, in 2011, the MSc degree in computer information systems from Near East University, North of Cyprus, Turkey, in 2013, and the PhD degree from the Faculty of Engineering, School of Computing, Universiti Teknologi Malaysia (UTM), Johor Bahru, Malaysia. His graduate work was focused on breast cancer identification. Currently, he is working as a specialist lecturer in computer science at Nawroz University in Duhok, Kurdistan Region, Iraq. His current research interests include artificial intelligence, machine learning, deep learning, medical image analysis, image encryption, and steganography.

Hussam J. Mohammed is a doctor of philosophy in computer science. He received his BSc and MSc in computer science from University of Anbar, Iraq, and his PhD in digital forensic and cybersecurity from University of Plymouth, UK. He currently works at the Computer Center, University of Anbar. He does research in the digital forensics, machine learning and AI. His current research interest in cybersecurity and digital evidence.

Mazin Abed Mohammed completed his BSc degree in computer science from College of Computer, University of Anbar, Iraq, Master of IT from Universiti Tenaga Nasional Malaysia and PhD, from Universiti Teknikal Malaysia Melaka, Malaysia. He is currently working as an associated professor in the Department of Information Systems in College of Computer Science and Information Technology, University of Anbar, Ramadi, Iraq. He teaches a variety of university courses in Computer Science, such as, Operating Systems, Database design, Mobile systems Programing, Software Project Management, Web technologies and Software Requirements and design. His areas of research interest include Artificial Intelligence, Medical image processing, machine learning, Computer Vision, computational intelligence, IoT, Biomedical Computing, bio-informatics, and fog Computing. He has published more than 120 papers International Journals and conferences. His outstanding scientific production spans over120+ contributions published in high standard ISI journals, such as IEEE Access, Future Generation Computer Systems, Medical Informatics, Computers & Electrical Engineering, Computational Science, and Medical Systems.