IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 30 September 2023, accepted 28 October 2023, date of publication 6 November 2023,
date of current version 9 November 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3330238

RESEARCH ARTICLE

Identification and Severity Assessment of COVID-19 Using Lung CT Scans

ANAND THYAGACHANDRAN[1], (Student Member, IEEE), AATHIRA BALACHANDRAN[2], AND HEMA A. MURTHY[1], (Senior Member, IEEE)

[1]Computer Science and Engineering Department, Indian Institute of Technology Madras, Chennai 600036, India
[2]Radio Diagnosis Department, Government Medical College Kollam, Kollam 691574, India

Corresponding authors: Anand Thyagachandran (tanand@cse.iitm.ac.in) and Hema A. Murthy (hema@cse.iitm.ac.in)

ABSTRACT The COVID-19 pandemic, caused by the SARS-CoV-2 virus, continues to have a significant impact on the global population. To effectively triage patients and understand the progression of the disease, a metric-based analysis of diagnostic techniques is necessary. The objective of the present study is to identify COVID-19 from chest CT scans and determine the extent of severity, defined by a severity score that indicates the volume of infection. An unsupervised preprocessing pipeline is proposed to extract relevant clinical features and utilize this information to employ a pretrained ImageNet EfficientNetB5 model to extract discriminative features. Subsequently, a shallow feed-forward neural network is trained to classify the CT scans into three classes, namely COVID-19, Community-Acquired Pneumonia, and Normal. Through various ablation studies, we find that a domain-specific preprocessing pipeline has a significant positive impact on classification accuracy. The infection segmentation mask generated from the preprocessed pipeline performs better than state-of-the-art supervised semantic segmentation models. Further, the estimated infection severity score is observed to be well correlated with radiologists' assessments. The results confirm the importance of domain-specific preprocessing for training machine learning algorithms.

INDEX TERMS COVID-19, CT scans, infection segmentation, semi-supervised augmentation, severity assessment.

I. INTRODUCTION

The COVID-19 pandemic, a highly contagious and primarily respiratory illness, has been of significant concern with devastating effects on public health, the world economy, and the social fabric of society. The unabated spread of infection, contrasted with other respiratory illnesses like SARS, has been ascribed to the ability of the virus to infect other people when the infected carrier is clinically asymptomatic [1]. This poses a significant challenge in terms of early detection and containment. In severe cases, the progression of the disease often leads to respiratory problems, which can be identified by noticeable changes in chest X-rays or CT scans, such as lung fibrosis and opaqueness [2].

The diagnosis of COVID-19 is crucial in identifying the pathogenicity of the virus and the severity of the disease. The gold standard for this purpose in a community setting is Real-Time Polymerase Chain Reaction (RT-PCR), which uses nasal and nasopharyngeal swab samples [3]. While RT-PCR is a more reliable method for detecting infection, it suffers from two limitations – 1. Although RT-PCR estimates the viral load, there seems to be an ambiguous correlation between viral load and the severity of the disease. This is owing to the nature of the swab taken, whereby nasopharyngeal swabs provide proof of the viral load in the upper respiratory tract, while the severe disease is usually associated with the lower respiratory tract. 2. RT-PCR does not quantify the clinical features of the patient under study, i.e., it is not a test for the response of the human immune system to the pathogen. Concerning the scope of the present work with an emphasis on COVID-19 being a respiratory disease primarily, the clinical diagnosis of COVID-19 severity is analyzed through radiological techniques like chest X-ray, which can be in one angle (traditional X-ray) or along several planes to provide for tomographic reconstruction of the chest CT Scans [4].

The associate editor coordinating the review of this manuscript and approving it for publication was Zhan-Li Sun.

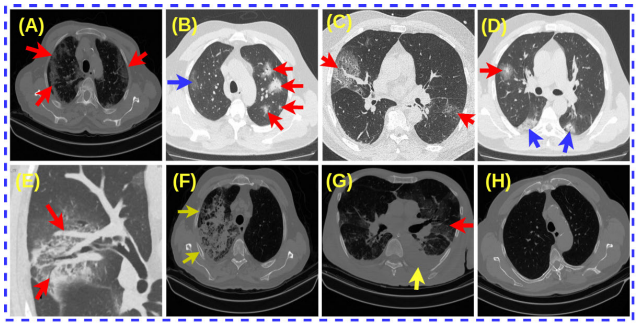

FIGURE 1. The clinical imaging observations of COVID-19 and CAP in chest CT scans. Images (A)-(E) show the COVID-19 patient's chest CT scan slices, (F) and (G) show the CAP patient's chest CT scan slices. (A) Peripheral distribution of GGO in both lungs (red arrows). (B) Patchy areas of consolidations in the left lung (red arrow) and GGOs in the right lung (blue arrow). (C) Crazy paving pattern: GGOs with superimposed septal thickening (red arrows). (D) Halo sign: consolidations surrounded by GGOs (red arrow), consolidations at the lower lobes of lungs (blue arrow). (E) Subsegmental vessel enlargement near the lesion (red arrows). (F) An extensive area of consolidations with GGO in the right lung – severe disease. (G) Pleural effusion (blue arrow) and GGOs (red arrows) in the left lung. (H) Normal chest CT scan slice.

Chest X-ray images have emerged as one of the most cost-effective and time-efficient tools for assessing the severity of infection. AI-based solutions are effectively utilized for conducting severity analysis using chest X-ray images [5], [6]. However, the limitation of traditional X-ray imaging lies in its line-of-sight integration, leading to lower resolution and a less accurate prognosis of COVID-19 when compared to CT scans. In the context of image processing, CT scans represent higher dimensional input data, with the ability of the processing algorithms being tested on the ability to recognize essential features. Although algorithms can encode high-dimensional images into a set of low-dimensional features, the overlap between features of various diseases (respiratory in the present context) results in erroneous classification as shown by [7], [8]. This can lead to a failure in managing the complications of COVID-19, such as cytokine storms, a significant cause of fatalities in COVID-19 cases [9]. Motivated by these issues, the primary purpose of the present work is to perform better classification between a widely occurring respiratory disease – Community-Acquired Pneumonia (CAP), COVID-19 and also against Normal CT scans through a preprocessing routine. Chest CT scans involve the projection of the X-ray bursts at different planes (termed as slices) to obtain high-resolution images of the chest region. The superior performance of CT scans in diagnosing respiratory illness lies in their ability to localize areas of abnormal opacity, which is usually a result of inflammation. These are often characterized in the form of ground-glass opacities (GGO), consolidation, a combination of GGOs and consolidation, halo sign (central consolidations surrounded by ground-glass opacities), reverse halo sign (central ground-glass lucent area with peripheral consolidation), and crazy paving patterns [10], [11], [12]. Common features of GGO include opaque foggy regions that do not obstruct the pulmonary vessels, while consolidation is marked by high opacity, thereby rendering the visualization of the pulmonary vascular structures impossible. Crazy-paving patterns are linear patterns superimposed on the background of GGOs resembling irregularly shaped paving stones. The bilateral distributions of GGO with or without consolidation in the posterior or peripheral lung regions are the primary indicators for COVID-19. As the disease severity progresses, consolidations, crazy-paving patterns, and vascular enlargement [13], [14] are the hallmark features. Pleural effusion and significant mediastinal lymphadenopathy are less commonly observed findings in COVID-19 infection. Pleural effusion is accumulating excessive fluid in the pleural space surrounding each lung. In the case of pneumonia, the features are more localized with the observation of a unilateral distribution of GGOs and consolidation and associated with pleural effusion and significant mediastinal lymphadenopathy compared to COVID-19 [7]. An experienced radiologist was consulted to obtain the Hounsfield Unit (HU – a measure of the range of radiation attenuation values) corresponding to different regions in the CT scan. The clinical features observed for COVID-19 and CAP are shown in FIGURE 1. The HU values of pixels in the GGO region lie between −650 HU and −50 HU, the consolidation region is between 10 HU and 45 HU, and the pleural effusion region is between 0 HU and 35 HU.

The proposed research encompasses the development of a preprocessing pipeline that enhances the features of interest from a CT scan of a COVID-19 patient, leading to a better assessment of the severity of the infection.

The contributions of the paper are as follows:

- An unsupervised preprocessing module to segment the regions of infection from chest CT scans of COVID-19 patients.
- An ablation study to evaluate the contribution of each stage in the preprocessing pipeline to classification accuracy.
- COVID-19 CT scan severity analysis based on the infections developed in the lungs.

The paper is organized as follows– Segmentation and identification of COVID-19 from chest CT scans is presented in Section II. Section III briefly discusses the datasets used in the proposed work. Section IV presents approaches to extract relevant lung features for the identification of infection. We also briefly explain the training strategy to classify CT scans into COVID-19, CAP, and Normal categories. Section V details the set of experiments and discussion on results and inferences. Section VI concludes the work.

II. RELATED WORK

Most existing research on the "classification of COVID-19 CT scans" uses raw images for the classification tasks. Image preprocessing techniques such as image enhancement and segmentation have been extensively used in the literature for enhancing the quality of the image, removing unwanted

regions, and extracting the region of interest. In the proposed work, we amalgamate generic image preprocessing techniques to aid image classification algorithms in identifying the prevalence of clinical features from chest CT scans. Although several algorithms for preprocessing do exist, we propose a domain-dependent preprocessing approach before embarking on classification. The severity of the disease is best estimated by volumetric analysis, and it requires slice-level information. Most datasets do not include slice-level information. Available slice-level information is used to attach labels to all slices of CT scans. This data is then used to augment training data leading to an improvement in accuracy.

A. CT SCAN SEGMENTATION

Segmentation using sophisticated image processing techniques is required to address the issue of varying contrasts observed in CT scans procured through different sources. Unsupervised image segmentation algorithms are broadly categorized into a) threshold-based, b) region-based, c) boundary-based, d) machine learning-based, and e) deep learning-based models. Threshold-based methods, as utilized in [15], such as Huang [16], Kapur [17], and Otsu [18], are employed to binarize the CT scans and generated a region adjacency graph (RAG) [19] to demarcate COVID-19 lesions from CT scans. The absence of sufficient contrast in threshold-based approaches can lead to a significant loss of information and result in poor performance. Region-based methods segment an area by assessing the homogeneity of the neighboring pixels. Widely practised region-based algorithms include region-growing [20], [21], [22], watershed [23], graph cuts [24], [25], and fuzzy connectedness [26]. Boundary-based methods are computationally intensive but provide highly accurate segmentation when the initial iteration is in the vicinity of the actual boundary. Boundary-based methods include snakes [27], active contours [28], and level sets [29], [30]. The boundary-based and region-based methods capture variations in attenuation but fail to segment regions of infection (such as consolidations and pleural effusion) near the lung boundary owing to a similar range of HU values. Hu et al. [31] proposed a threshold-based approach to segment lung regions from CT scans by processing the left and right lungs separately. Further modifications are performed using morphological operators for fine-tuning the identification of irregular boundaries of the GGOs. Zhu et al. [32] proposed a method to extract GGOs by modeling their intensity distributions and using the Markov random field model to improve boundary identification. Ye et al. [33] utilized textural features from CT image intensity parameters viz. entropy, contrast, roughness, and coarseness for segmentation of infection region. These morphological features are then used to demarcate GGOs from the image but were found to be insufficient in distinguishing the consolidations from pulmonary vessels. Wang et al. [34] proposed a novel approach for the detection of COVID-19 features using a 3D deep convolutional neural network (CNN) called "DeCoVNet" on CT volumes. The authors employed a combination of the activation maps generated by the DeCoVNet with a 3D connected component (3DCC) algorithm to identify lesions from the CT scans. Although the model demonstrated high recall, it suffered from a high incidence of false positives. This is seen to be the result of the formulation based on a black-box approach, thereby rendering the attention of the gradient to the infection region. The attention of activation maps cannot be guaranteed to accurately focus on the region of interest. This may lead to less accurate predictions and a higher rate of false positives.

Inspired by the intensity distribution adaptive model using MAP as proposed by Zhu et al. [32], we propose to use a three-mixture Gaussian mixture model using adaptive thresholding instead of a single Gaussian to extract the clinical features in the proposed work.

B. CT SCAN CLASSIFICATION

Prior studies have made significant progress in COVID-19 identification through CT scan images, broadly categorized into two approaches: 3D CT scan-based classification and 2D CT scan-based classification. In 3D CT scan-based classification, a 3D CNN is trained on volumetric CT scans, and a probability score is evaluated for each scan [34], [35], [36], [37], [38], [39]. Among these techniques, segmentation of the lung region using image preprocessing methods is applied before performing classification. Owing to the varying dimensionality of 3D CT scans, interpolation or truncation of the slices is applied to convert them to fixed dimensionality [40].

In 2D CT scan-based classification, a 2D CNN is trained on individual slices, generating slice-level probability scores. Further, threshold-based [40], majority voting [41], [42], [43], weighted average methods [44], and sequence models (such as recurrent neural network (RNN) [45] and bidirectional long short term memory (BiLSTM) [46]) are used to obtain patient-level COVID-19 classification. Threshold-based and majority voting methods create higher false negatives at regions where traces of infections are not visible, as in most CT scan slices. Considering the difficulty in obtaining annotated CT scans, transfer learning methods have been extensively employed in COVID-19 classification. Some transfer learning works [42], [44], [46], [47], [48] have explored different CNN models trained on ImageNet dataset [49] for classification tasks. Transfer learning methods reduce the training requirement for every dataset and provide discriminative features for classifying COVID-19, even while being applied to raw CT scan images.

Chaudhary et al. [44] employs transfer learning in generating features from raw CT images using the EfficientNet-B5 model [50]. The slice-level scores are obtained using a shallow feed-forward neural network (FFNN). Further, patient-level classification was performed using a weighted average method on the slice-level scores. We use this model

TABLE 1. Different datasets used in the experiment. "*"" – slice-level labels are available; "" – slice-level labels are unavailable, but the CT scan patient-level label is available. "C" and "S" denote the classification and segmentation datasets respectively.**

Dataset	Class	COVID-19	CAP	Normal	C/S	Format
SPGC	Train	55*+116**	25*+35**	76	C	DICOM
	Test	28**	51**	51		
LDCT	Test	104**	0	56		
LDCT-PCR	Test	100*	0	0		
Mosmed	Test	854**	0	254	C-S	NIfTI
Mosmed	Test	50*	0	0		
MedSeg	Test	100* slices	0	0	S	NIfTI
MedSeg_1	Test	9* (638 slices)	0	0		
Mehta	Test	14**	0	0		DICOM

as the baseline for comparison and use a similar pipeline for classification in the proposed approach. Inspired by this technique [44], the present work attempts to enhance the performance in classifying COVID-19 in a CT scan by using a preprocessing pipeline rather than raw CT scan images. We further provide an objective understanding of the severity of a COVID-19 patient by introducing a severity analysis module in the proposed architecture.

III. DATASETS

The research work utilizes six publicly available datasets (in TABLE 1). These datasets consist of CT scans, which are volumetric scans composed of axial slices. Each slice has dimensions of 512 × 512 pixels. We have chosen datasets from diverse imaging devices that are available in either Digital Imaging and Communications in Medicine (DICOM) or Neuroimaging Informatics Technology Initiative (NIfTI). One of the primary challenges lies in the availability and quality of labeled data. Creating accurate and comprehensive annotations for COVID-19 CT scans, especially at the slice level, can be labor-intensive and time-consuming, often requiring expert radiologists. We briefly describe the characteristics of each of the datasets.

For infection region segmentation analysis in COVID-19 patients, three additional available datasets (Mosmed [51], MedSeg [52], and MedSeg_1 [53]) are employed. Expert radiologists delineated the lesion regions in the CT scans. The MedSeg dataset contains 100 CT slices collected from over 40 COVID-19 patients, featuring infections spanning from minimal to severe cases. However, this dataset is inadequate for lesion segmentation assessment, as the severity analysis relies on the CT scan volume, which may or may not contain infections across all slices. The Mosmed and MedSeg_1 datasets comprise CT scan volumes from 50 and 9 patients, respectively, collected from different geographic locations. The Mosmed dataset demonstrates infection rates below 25%, while the MedSeg_1 dataset exhibits infection severity variations from minimal to severe.

The SPGC dataset [54] encompasses CT scan slices belonging to COVID-19, CAP, and Normal classes, presented in DICOM format is used for training the classification model. Expert radiologists assign slice-level labels for COVID-19 and CAP cases. The SPGC training dataset also contains unlabeled CT scan slices, but the patient-level labels for the entire CT scan are provided. To evaluate the classification model's robustness, three additional publicly accessible datasets (LDCT [55], LDCT-PCR [55], and Mosmed [51]) are employed. The LDCT and LDCT-PCR datasets are available in DICOM format, whereas the Mosmed dataset is provided in the NIfTI format.

COVID-19 patient severity is measured using the CT severity score (CTSS), ranging from 0 to 25. The Mehta dataset encompasses CT scans from fourteen COVID-19 patients collected at Mehta Multi-speciality Hospital (a nearby hospital) is employed for the severity analysis. It comprises comprehensive diagnostic reports for each CT scan, containing CTSS, COVID-19 Reporting and Data System (CO-RADS) score, symptoms, and lung infection volume. Patient information within the CT scan metadata and diagnosis reports is anonymized. A subset of the SPGC dataset (36 patients) is assigned to an experienced radiologist to predict the CTSS. These datasets aid in determining the correlation of CTSS predicted by the proposed model.

IV. PROPOSED SYSTEM

The proposed system is designed to achieve three primary objectives: 1) Propose an innovative preprocessing pipeline for extracting pertinent clinical features used by radiologists, 2) Generally, a single classification score is given for each CT scan, we have tried to classify each of the slices in a CT scan, and 3) Predict CTSS for COVID-19 patients through the utilization of the preprocessing pipeline. We first preprocess the CT scans using image processing and machine learning algorithms. The pre-processed CT scan slices are then fed to a pretrained ImageNet model to extract high-dimensional discriminative features. These features are then used to train an FFNN that predicts scores at the CT scan slice level. Subsequently, a weighted average approach is employed to compute the definitive score, enabling the categorization of CT scans into three targeted classes: COVID-19, CAP, and Normal. Moreover, an additional module for severity analysis is introduced, aimed at determining the CTSS for COVID-19 patients. Overall, the proposed model is devised to effectively classify COVID-19, CAP, and Normal CT scans while offering predictions for the CT severity score in COVID-19 patients.

A. PREPROCESSING PIPELINE

Radiologists commonly rely on typical clinical features, such as GGOs, consolidation, crazy paving pattern, halo sign, reverse halo sign, and pleural effusion, to differentiate between COVID-19 and CAP in comparison to healthy individuals. The abnormalities in the CT scan exhibit attenuation variations with respect to the severity of the infection. Utilizing insights from FIGURE 1, an image processing pipeline is proposed, focusing primarily on distinctions among the three classes: COVID-19, CAP, and Normal. As illustrated in FIGURE 2, this unsupervised

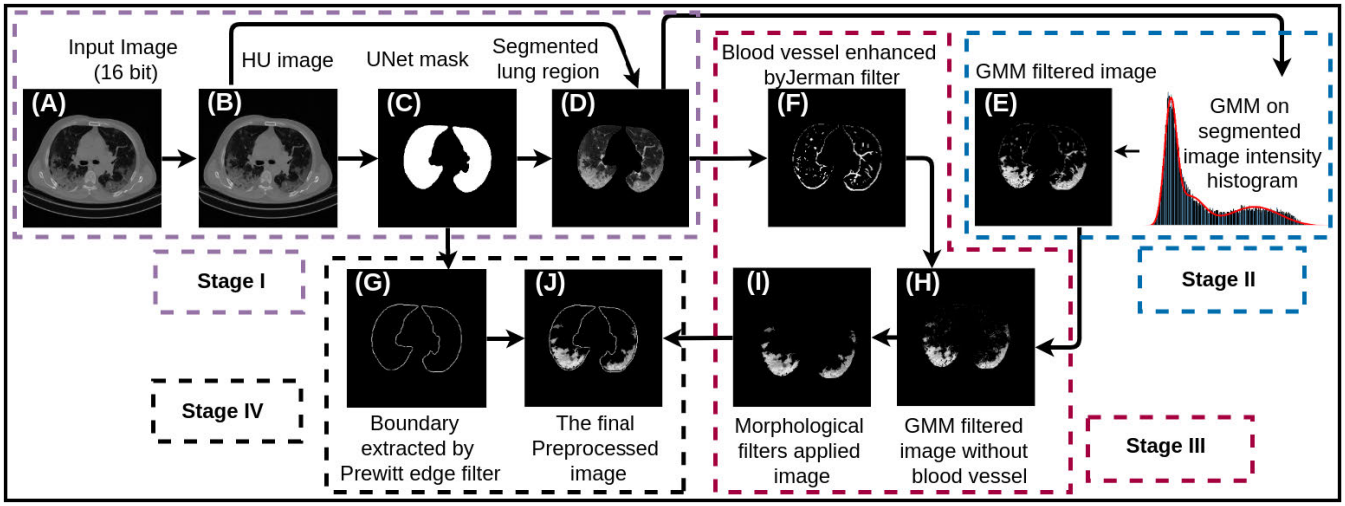

FIGURE 2. Proposed preprocessing pipeline for CT scan slices.

preprocessing pipeline amalgamates a pretrained deep learning model with classical image processing methodologies and machine learning techniques. The preprocessing pipeline aligns with domain-knowledge-based protocols employed by radiologists during CT scan analysis. The pipeline extracts appropriate clinical features for infection segmentation and classification tasks through four stages.

1) STAGE-I: LUNG MASK GENERATION

A CT scan is a 3D scan composed of numerous slices, each with a 512 × 512 pixels resolution. Within a chest CT scan, tissues, heart, stomach, blood vessels, and bones exhibit higher attenuation (measured in Hounsfield Units - HU) values than air and lung regions. The preprocessing pipeline operates on the HU scale of CT scan slices. CT scans are primarily available in two formats: NIfTI and DICOM. The NIfTI-formatted CT scan slices provide pixel intensities in HU units. Conversely, CT scan slices in DICOM format are transformed into HU scale slices using a linear transformation.

The initial preprocessing step involves eliminating unnecessary regions in the chest CT scan and identifying the area of interest, specifically, the lung region. The HU-scaled image is fed to a pretrained UNet model [56] to extract the region of interest. This model, introduced by Hofmanninger et al., is tailored to detect lung regions amidst severe pathological conditions present in chest CT scans. The backbone of this model is based on UNet, a state-of-the-art model in medical image segmentation [57]. The pretrained UNet [56] constitutes an end-to-end, fully convolutional neural network incorporating an encoder module that compresses input CT images through convolution and max-pooling operations into a fixed-length feature map. Subsequently, the decoder module (enhancing spatial representation) upsamples these feature maps to create the lung mask. The integration of skip connections between the encoder and decoder modules enhances semantic features for segmentation. This pretrained UNet model generates a lung region mask, segmenting the chest CT scan into three classes: the left lung region, the right lung region, and the background.

Commonly, the initial and final slices of the CT scan comprise structures like bones, the trachea, the diaphragm, the heart, and the stomach, which are irrelevant to subsequent analysis. A fixed threshold for the number of slices is impractical due to the varying slice count among patients. Instead, the area of the lung mask generated by the UNet is employed as a threshold to exclude initial and final slices. A criterion based on this threshold determines the number of slices for subsequent analysis in each patient. This approach ensures that slices most relevant to further analysis are considered.

2) STAGE-II: GMM-BASED ADAPTIVE FILTER

The range of attenuation values (HU) characterizing clinical features such as GGO, consolidation, and pleural effusion is discussed in Section I. The pixel intensity values in GGO, consolidation, and pleural effusion can vary based on the severity of the infection. Regions with more severe infections tend to have higher pixel intensities than milder ones. This attenuation variance is captured by modeling pixel intensity histograms using a three-mixture Gaussian Mixture Model (GMM). Each mixture within the GMM seeks to capture patterns corresponding to GGO, consolidation, pleural effusion, and the background. This unsupervised clustering technique is applied to the Stage-I output image (lung region segmented CT scan slice). A consultation with radiologists suggests that an adaptive threshold is appropriate. Each GMM mixture is characterized by a mean μ, a standard deviation σ, and a threshold ($\mu \pm 1.5 \times \sigma$) is employed for pixel selection. Pixel intensities falling within the specified range are retained for subsequent analysis. While the GMM adaptive filter effectively segments infection regions, the

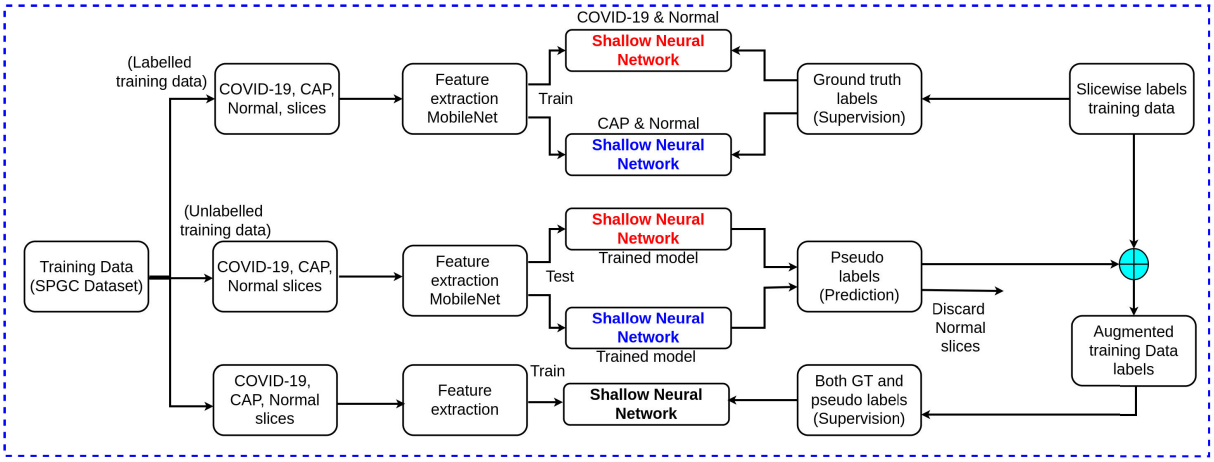

FIGURE 3. The proposed semi-supervised training method for the classifier.

presence of blood vessels introduces complexities to the severity analysis. The images undergo further vascular analysis modules to mitigate this concern.

3) STAGE-III: VASCULAR ENHANCEMENT AND MORPHOLOGICAL FILTERS

After Stage-II of preprocessing, the CT scan slice contains essential features of COVID-19, CAP, and primary pulmonary vessels. The pulmonary vessels share similar HU attenuation values as consolidations and pleural effusions. A *shape-based* filter is needed to retain the consolidation and pleural effusion regions and remove the blood vessels. The Jerman blood vessel enhancement filter [58] is widely used in angiographic images to enhance blood vessels. This filter can identify the local structures in the images based on the shape (elongated or circular) by evaluating the sign and magnitude of the Eigenvalues of the image's Hessian matrix. The Jerman filter examines the largest to smallest eigenvalue ratio and assigns a probability score for each pixel to be a part of elongated local structures. The output from stage-I is subjected to this filter with a threshold of 0.75 to generate a binary mask for marking and removing the blood vessels. The blood vessels underneath the GGOs are also enhanced and removed from the slice, creating holes in the GGO region. The flood fill algorithm [59] is then applied to homogenize these holes using the intensities of the pixels in the neighborhood. Following this, a morphological operation (dilation) is applied to enhance the infection boundary in the resulting image. The small white regions, which are generated due to the removal of the blood vessels, are removed using the area opening morphological method [60].

4) STAGE-IV: GENERATE LUNG BOUNDARY AND GRAYSCALE IMAGE

The final preprocessing stage consists of generating the lung boundary from the UNet mask generated in Stage-I and rendering the HU scale preprocessed CT scan slice compatible with ImageNet pretrained models. The Prewitt edge detector [61] is employed to extract the lung boundary from the lung mask produced during Stage-I. This extracted lung boundary is then superimposed onto the output CT scan slice of the preprocessing pipeline, aiding in localizing and assessing the distribution of clinical features within the CT scan. The resulting HU scale slice is then normalized into an 8-bit grayscale image. Next, the grayscale image is converted to a three-channel image with appropriate dimensions as required by a pretrained ImageNet model.

B. TRAINING

The SPGC training dataset [54] doesn't suffice to train a deep CNN model from scratch independently. This dataset offers a larger number of patients with CT scan slices without slice-level labels. Nevertheless, patient-level labels are available. A semi-supervised approach is therefore implemented to label unannotated CT scan slices within the training dataset. The training process consists of two phases: i) Data augmentation and ii) Training the classifier. Transfer learning is employed from the state-of-the-art computer vision models trained on the ImageNet dataset to extract high-dimensional features of the preprocessed CT scan slices. We used a pretrained MobileNet [62] in the data augmentation and EfficientNetB5[1] [50] for training the final classifier.

In the data augmentation phase, the labeled slices undergo preprocessing and are fed to the pretrained MobileNet to extract slice-wise feature maps. Each feature map is subsequently input to a global-average-pooling layer, generating a 1024-dimensional feature vector. Later, two shallow FFNNs, featuring 1024 neurons and 2 neurons in their initial and final layers, are trained using these extracted 1024-dimensional feature vectors. To address the class imbalance problem, we train two separate FFNNs.

[1]A number of feature extraction techniques are available, namely ResNet50, ResNet101, EfficientNetB1. EfficientNetB5 works best for the proposed features identified by radiologists.

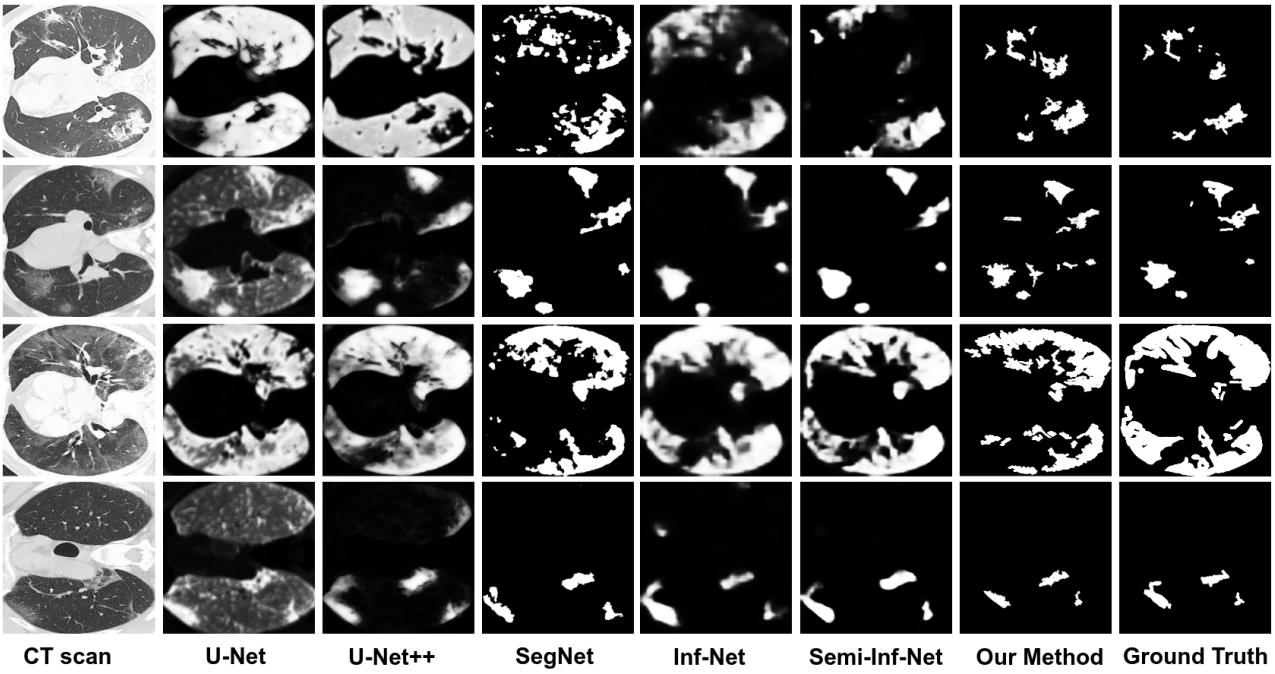

FIGURE 4. The region of infection extracted from the MedSeg dataset using different baseline models and the proposed model.

One that distinguishes between Normal and COVID-19 and another that distinguishes between Normal and CAP. While training these networks, it is ensured that the same number of examples are chosen from each class. These models are respectively termed MobileNet_COVID and MobileNet_CAP for future reference. The unlabeled CT scan slices undergo the preprocessing pipeline described in Section IV-A. The preprocessed CT scan slices corresponding to COVID-19 classes are provided to the MobileNet_COVID model, and similarly, unlabeled CAP slices are fed to the MobileNet_CAP model. The unlabeled slices in the training dataset are labeled using these networks, which give slice-level labels. The slice-level labels corresponding to infection are then used to augment the training dataset. The final classifier comprises EfficientNetB5 to extract the features and shallow FFNN to learn these extracted features. The FFNN consists of three layers: two dense layers (with 2048 and 1024 neurons) and a final layer (comprising three neurons). Each neuron in the last layer generates a probability score for a corresponding class: CAP, COVID-19, and Normal. The final classifier model is trained with approximately an equal number of preprocessed CT scan slices from each class. Random sampling is employed for classes with large amounts of data. The training pipeline is shown in FIGURE 3.

C. CT SCAN CLASSIFICATION

The three-class classifier produces probabilistic scores for each CT scan slice. However, due to the volumetric nature of CT scans, which have dimensions of (n * 512 × 512), CT scan annotations are preferred over slice-level classification. Thus, a weighted averaging technique is employed to aggregate the classifier's probability scores.

The CT scan volume, consisting of 'n' slices, is divided into three equal segments across the axial plane, each associated with distinct weights (W_1, W_2, W_3), where W_1, W_2, and W_3 are assigned values of 0.7, 1, and 0.7, respectively. This weightage is primarily based on the observation that the middle region contributes to a larger lung area and thus contributes most to information regarding infections. In our experimentation with the training data, the average lung area of the upper extreme region is approximately 0.7, while the lower region is roughly 0.8 times the middle region. This was further confirmed by the radiologist that slices in the extreme regions account for smaller lung area due to organ occlusion and that slices from the lower axial plane are susceptible to artifacts from holding the breath during the scan. As a result, greater weights are allocated to slices from the middle region, while a relatively smaller weight is attributed to the other two regions.

For classification, if the weighted sum of the probabilities for the COVID-19 and CAP categories exceeds that of the Normal class, the CT scan is considered abnormal. Subsequently, it is classified either as COVID-19 or CAP based on the scores. The LDCT, Mosmed, and LDCT-PCR datasets lack CAP CT scans. For these datasets, CT scan slices predicted are considered as belonging to the COVID-19 class. Consider the predicted score for the i^{th} slice as P_i. The CT scan score is calculated as

follows:

$$FS = \max\left(\sum_{i=1}^{n/3} P_i * W_1 + \sum_{i=n/3}^{2n/3} P_i * W_2 + \sum_{i=2n/3}^{n} P_i * W_3\right)$$

V. RESULTS AND DISCUSSIONS

This section begins by outlining the diverse evaluation metrics employed in this study. Subsequently, it delves into the performance of the proposed unsupervised preprocessing pipeline's infection segmentation with supervised deep-learning models. Ablation studies are conducted to assess the preprocessing pipeline's impact on classification performance. The effectiveness of the three-class classifier is examined across various test datasets collected from diverse geographic locations and utilizing different scanning devices, aiming to assess the model's robustness when applied to varied CT scan slices. The CTSS prediction for both the Mehta dataset and a subset of the SPGC dataset is discussed.

A. EVALUATION METRICS

The evaluation of the proposed pipeline involves a comprehensive set of widely used metrics, addressing both segmentation and classification tasks. For the assessment of segmentation performance, the following metrics are employed: Dice score (Dice), sensitivity (Sen.), specificity (Spec.), precision (Prec.), and mean absolute error (MAE). The Dice score quantifies segmentation errors by measuring the degree of overlap between predicted and annotated areas. Precision assesses the proportion of accurately predicted white pixels within the total white pixel predictions made by the model. Conversely, specificity indicates the ratio of correctly predicted black pixels relative to the total black pixels in the ground truth image. MAE computes the average absolute difference between predicted and annotated binary masks, thereby quantifying the quality of the predicted Region of Interest (ROI). A lower MAE value signifies superior segmentation performance. For classification evaluation, metrics such as sensitivity, precision, and F1 score are employed. Sensitivity measures the count of accurately predicted CT scans with the ground truth. Precision captures the count of CT scans correctly predicted from the overall predictions by the model. The F1 score, analogous to the Dice score in the segmentation context, represents the harmonic mean of sensitivity and precision. The proposed pipeline predicts the CTSS and assesses these predictions using the Pearson correlation coefficient and cosine similarity. The Pearson correlation coefficient gauges the correlation between the covariance of the predicted and ground truth scores, along with their respective standard deviations. This metric provides insights into the trend between the proposed and ground truth CTSS predictions. On the other hand, cosine similarity computes the angle between the vectors of the proposed and ground truth CTSS predictions. Both the Pearson correlation coefficient and cosine similarity range from -1 to 1. Values greater than zero indicate positive correlations and higher levels of similarity between the predictions and the ground truth.

TABLE 2. The infection segmentation results with the MedSeg Dataset (48 CT scan slices).

Model	Dice	Sen.	Spec.	MAE
UNet [57]	0.439	0.534	0.858	0.186
Attention UNet [63]	0.583	0.637	0.921	0.112
Gated UNet [64]	0.623	0.658	0.926	0.102
Dense UNet [65]	0.515	0.594	0.840	0.184
UNet++ [66]	0.581	0.672	0.902	0.120
SegNet [67]	0.67	0.748	0.925	0.089
InfNet [68]	0.682	0.692	0.943	0.082
SemiInfNet [68]	0.739	0.725	0.960	0.064
Our Method	**0.673**	**0.678**	**0.9852**	**0.0356**

B. INFECTION SEGMENTATION RESULTS

The output from Stage-III of the preprocessing pipeline (FIGURE 2-I) is used as the infection mask for the CT scan slices. This preprocessed image (I) is generated by adaptive filtering of the clinical features using a three-mixture GMM; then, blood vessels are removed by the Jerman filter, and the resultant image is further processed using morphological operators to fill the holes and enhance the infection regions. The primary contribution of this work is on infection segmentation. We compare our approach with 5 baseline models given in Fan et al. [68], based on different variants of the UNet [57] architecture, and SegNet model [67] given in TABLE 2. The MedSeg dataset [52] contains a diverse range of infections, but non-infectious slices are absent. The infection segmentation of the proposed model is assessed using the same dataset and CT scan slices employed for evaluating the baseline models. The infection segmentation of sample CT scan slices from the MedSeg dataset is shown in FIGURE 4. Compared to the baseline models, the proposed model demonstrates superior performance regarding dice score, specificity, and MAE. The proposed model works equally with the InfNet and SemiInfNet models.

Given that the MedSeg dataset is a collection of slices from various patients, the authors of the study [68] also conducted infection segmentation using another dataset named MedSeg_1. This dataset comprises nine patients with 638 slices, of which 285 are non-infected, and 353 are infected. The infection segmentation results of the baseline, InfNet, SemiInfNet, and Proposed model are summarized in TABLE 2. The proposed model's Dice score, precision, and MAE scores are better than the baseline models. The baseline, InfNet, and SemiInfNet models generate a more oversized infection segmentation mask (higher sensitivity and lower precision) than the ground truth mask. The proposed model's precision to sensitivity score ratio is higher than other models in TABLE 2 and ensures that the proposed model infection mask is more consistent with the infections in the

TABLE 3. The infection segmentation results with the MedSeg_1 Dataset (9 patient's real CT scan volumes).

Model	Dice	Sen.	Spec.	Prec.	MAE
UNet [57]	0.308	0.678	0.836	0.265	0.214
SegNet [67]	0.374	0.435	0.949	0.339	0.053
Attention UNet [63]	0.466	0.723	0.930	0.390	0.095
Gated UNet [64]	0.447	0.674	0.956	0.375	0.066
Dense UNet [65]	0.410	0.607	0.977	0.415	0.167
UNet++ [66]	0.444	0.877	0.929	0.369	0.106
InfNet [68]	0.579	0.870	0.974	0.500	0.047
SemiInfNet [68]	0.597	0.865	0.977	0.515	0.033
Our Method	**0.589**	**0.657**	**0.994**	**0.572**	**0.0095**

TABLE 4. The infection segmentation results with the Mosmed Dataset (50 patients CT scan volume).

	Task 1			Task 2		Proposed method
	MSD	Struct seg	NSCLC	Separate	Union	Infection Mask
Dice	0.392	0.443	0.301	0.588	0.482	**0.575**
Spec.	1.00	1.00	1.00	0.999	0.999	**0.998**
Sen.	0.364	0.422	0.249	0.575	0.601	**0.559**
Prec.	0.614	0.607	0.614	0.679	0.577	**0.681**

CT scan slice. The preprocessing pipeline's enhancements in performance are particularly in identifying uninfected CT scan slices. A similar performance is obtained with the InfNet and SemiInfNet models. Moreover, the proposed pipeline showcases the ability to detect even minor infections based on the statistical properties of attenuation within infection regions.

Ma et al. [69] used a 'no new UNet' (nnUNet) [70] based baseline model for infection segmentation. The baseline model is mainly trained with two datasets, out-of-domain and in-domain. The out-of-domain datasets include Medical Segmentation Decathlon (MSD) lung tumor segmentation (MICCAI 2018 challenge dataset), StructSeg lung cancer volume segmentation (MICCAI 2019 challenge dataset), and NSCLC pleural effusion segmentation [71]. The in-domain datasets [53] contain twenty COVID-19 CT scans with an infection range of 0.01% - 59%. Three baseline results are provided by training the nnUNet on each out-of-domain dataset (Task 1). Two baseline results are provided by training the nnUNet with in-domain and out-of-domain datasets (Task 2). In Task 2, one model is trained for lung and infection segmentation (Union), and the other is designed only for infection segmentation (separate). The Mosmed dataset [51] is used as a blind test data for the baseline models. The baseline results, along with our proposed model results, are summarized in TABLE 4. The proposed model provides better dice scores and precision results than the Task 1 baseline and Task 2 (Union) models. Since the infection region is less than 25% in the CT scan, it is observed that the proposed model can identify small infection regions from the CT scan slices and outperform the baseline models.

C. CLASSIFICATION RESULTS

One of our goals is to classify CT scans into three categories: COVID-19, CAP, and Normal. The baseline model by Chaudhary et al. [44] uses raw CT scan slices with EfficientNet-B5 [50] for feature extraction, and these features are learned by a shallow FFNN for classification. The pretrained weights for the baseline model are publicly available, and the baseline results cited in the paper are based on these weights. The proposed model follows a similar architecture but uses preprocessed images for training. Ablation studies are conducted to assess the impact of each preprocessing pipeline stage on the classification task. The best-performing model from the ablation study is adopted as the proposed model for subsequent analysis. The ablation results are shown in TABLE 5. The baseline and proposed models are assessed by observing accuracy, loss, and F1 scores for each epoch across training and validation datasets. These outcomes are shown in FIGURE 5. A Gradient weighted Class Activation Map (GradCAM) analysis [72] on raw and preprocessed images is employed, and results are visualized in FIGURE 6. Next, the baseline and proposed model are evaluated using unseen datasets such as LDCT [55], LDCT-PCR [55], and Mosmed [51]. The results are presented in TABLE 6.

1) ABLATION STUDY

The primary novelty of the proposed work is in its preprocessing pipeline. Ablation studies are performed to evaluate the effect of each of the stages in the preprocessing pipeline. Six different ablation studies are performed:

1) Raw Image (RI)
2) Preprocessed image with GMM filter (GMM)
3) Same as 2 with Morphological operations (GMM+MO)
4) Same as 2 using only features from the lung region (GMM + B)
5) Same as 4 with Morphological operations (GMM + MO + B)
6) Same as 5 with blood vessels removed from the image using Jerman Filter (GMM + MO + B + JF)

Each ablation study leads to a particular system. The preprocessed image in each case is subjected to the EfficientNetB5 [50] feature extractor. Nevertheless, to ensure that the preprocessing pipeline is not biased toward the EfficientNetB5 network, we have experimented with other feature extraction models; namely, ResNet50 [73], ResNet101 [73], and EfficientNetB1 [50]. Similar results are obtained, where system 5 consistently gives the best results. The blood vessel removal causes a loss of information in the preprocessed images, impacting the classification performance. Notably, the literature suggests that blood vessel enlargement can occur in the later stages of COVID-19 infection [13], [14]. The ablation study shows the significant contribution of the preprocessed image (GMM + MO + B) to the classification accuracy with the SPGC dataset. Based on the results in TABLE 5, System 5 is used as the proposed model for further analysis. The baseline and proposed model are compared with metrics such as accuracy, loss, and F1 score during the training and validation phase. The results are shown in

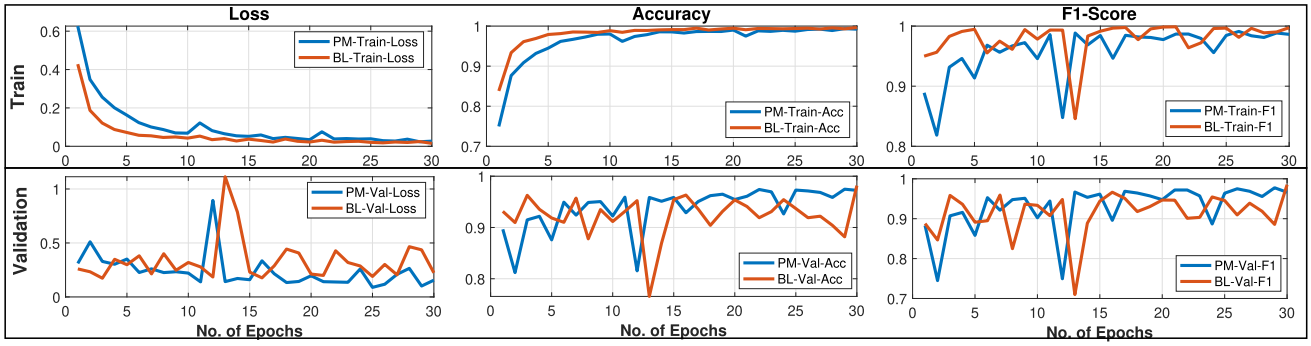

FIGURE 5. The training and validation performance of both the baseline and proposed model is assessed in terms of loss, accuracy, and F1 score in every epoch.

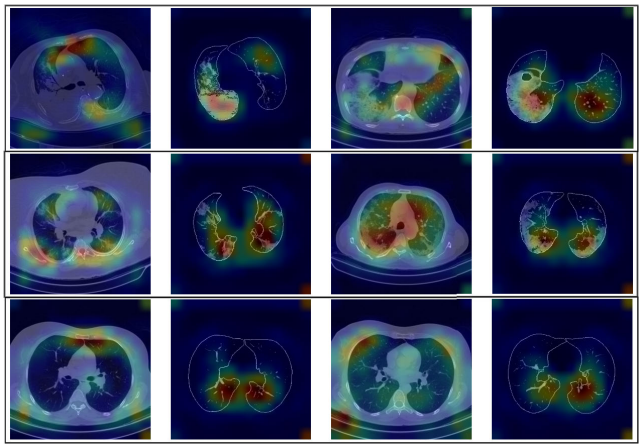

FIGURE 6. Gradient weighted class activation map for the CT scan images. The first row represents the original and preprocessed CAP CT scan slices with GradCAM representation. The second row depicts the GradCAM for the original and preprocessed COVID-19 CT scan slices. The third row shows the GradCAM for the raw and preprocessed Normal CT scan slices.

TABLE 5. An ablation study on the different preprocessing stages of the SPGC test dataset with the EfficientNetB5 model. The baseline model with given weights is used to evaluate the original image, and the results are shown in red color. The best result is shown in blue color.

RI	GMM	MO	B	JF	EfficientNet-B5		
					COVID-19	CAP	Normal
✓					**0.77**	**0.77**	**0.84**
	✓				0.77	0.57	0.79
	✓	✓			0.73	0.73	0.8
	✓		✓		0.80	0.78	0.87
	✓	✓	✓		**0.83**	**0.78**	**0.93**
	✓	✓	✓	✓	0.74	0.73	0.83

FIGURE 5. We can observe from the validation scores that the proposed model performs consistently better compared to the baseline model. GradCAM [72] is a technique that highlights the regions of an image that contribute to the classification accuracy. GradCAM images are obtained for the raw and pre-processed images from each class: COVID-19, CAP, and Normal, and are given in FIGURE 6. The preprocessing enables the network to focus on the lung regions, whereas the raw image is also activated by regions corresponding to that of the bones.

2) CT SCAN CLASSIFICATION

Since the CT scan is volumetric data, a weighted average method is utilized on the probabilistic score obtained from the slice-level classifier to estimate the final score. The baseline model [44] selects only the middle region slices based on a threshold applied to the number of images relative to the total number of images in the CT scan. A weighted average method is applied to the selected CT scan slices to evaluate the final score for the CT scan classification. The baseline model requires a minimum of 40 CT scan slices to make the final score using the weighted average method. In contrast, the proposed method employs a threshold based on the UNet lung mask's area generated during the preprocessing pipeline (FIGURE 2-C). The chosen images are categorized into three equivalent regions, and a weighted average is computed across these three groups (as elaborated in Section IV-C).

The baseline and proposed model are evaluated with the SPGC test dataset. The proposed model, which utilizes EfficientNet-B5, exhibits an average 6% improvement in the F1 score for the classification task compared to the baseline model. The model with preprocessed slices (GMM + MO + B) outperforms the baseline model. The confusion matrices for the baseline and proposed models are presented in FIGURE 7. The baseline model is more biased toward the COVID-19 class since ten patients from the CAP and Normal classes are misclassified into the COVID-19 class. While the proposed model performance is similar for COVID-19 CT scans, it has an edge over the baseline for CAP identification.

The SPGC [54] dataset is alone used for training. The scalability of the preprocessing pipeline is illustrated using publicly available datasets, namely, LDCT [55], LDCT-PCR [55], and Mosmed [51] datasets. The LDCT, LDCT-PCR, and Mosmed datasets consist of two classes of CT scans: COVID-19 and Normal. Since COVID-19 and CAP share similar clinical features in CT scans, CAP scores are augmented to COVID-19 scores for the CT scan classification. The proposed model achieves better results

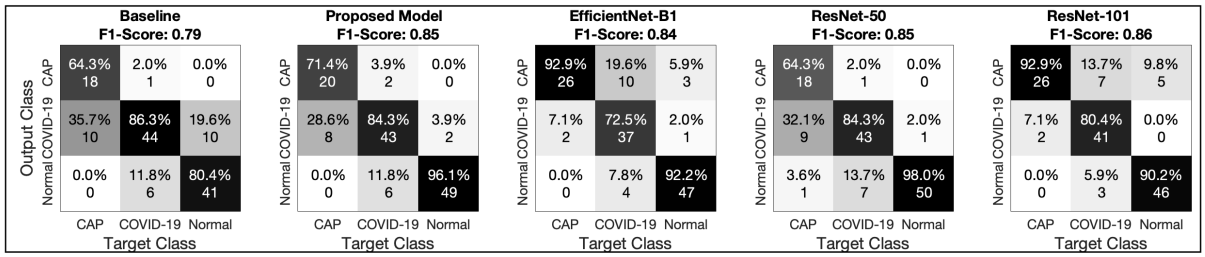

FIGURE 7. CT scan classification results for the baseline and proposed models with SPGC dataset.

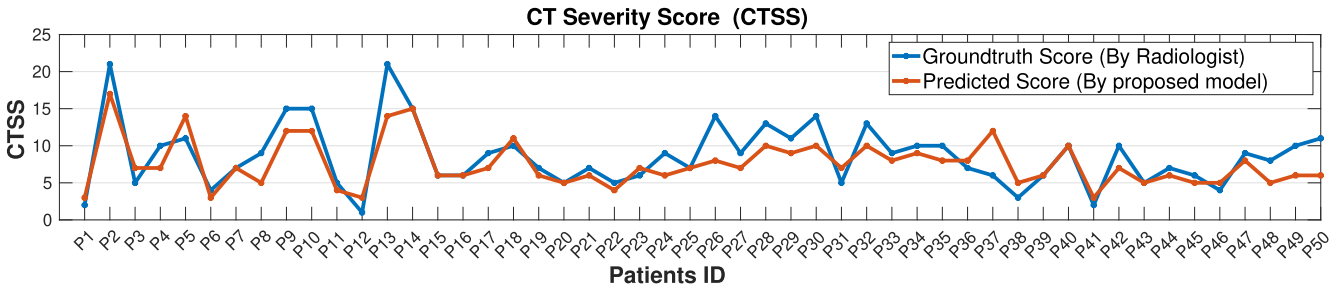

FIGURE 8. CTSS for 14 patients (Mehta) and 36 patients (a subset of SPGC datasets).

TABLE 6. A comparative study of baseline and proposed model with different test datasets used in the experiments.

Dataset	Baseline Model				Proposed Model			
	COVID-19		Normal		COVID-19		Normal	
	Prec.	Sen.	Prec.	Sen.	Prec.	Sen.	Prec.	Sen.
LDCT	0.98	0.86	0.81	0.97	0.96	0.83	0.77	0.95
LDCT-PCR	0.98	0.81	0.70	0.97	0.98	0.85	0.76	0.97
Mosmed	—	—	—	—	0.98	0.60	0.42	0.96

with the LDCT-PCR dataset, while the baseline model achieves better results with the LDCT dataset. The baseline model works only for the DICOM image format, but the Mosmed dataset is available in NifTI format. So, the proposed model is evaluated with the Mosmed dataset. The results are shown in TABLE 6.

The Mosmed dataset divided the CT scans into four subsets: CT-0, CT-1, CT-2, and CT-3. The CT-0 consists of CT scans having less than 5% of infections in CT scans. Since the ground truth was not given for each CT scan, CT-0 refers to the Normal class, and other subsets CT-1, CT-2, and CT-3 to the COVID-19 class. The CT-1 subset contains 684 CT scans with less than 25% of infections, but the number of infected slices is much less compared to the total number of slices. The CT-2 and CT-3 categories have significant lung infections in the CT scans and are classified as COVID-19 with minor misclassification. Our proposed model can identify the infected slices, but the number of infected slices is very small compared to the overall slices in the case of CT-0 and CT-1.

D. SEVERITY ANALYSIS

The CT severity score (CTSS) is the standard metric used for severity analysis, and it ranges from 0 (no involvement) to 25 (maximum involvement). This score is evaluated by segmenting the right lung into three lobes and the left lung into two, respectively. The infection rate in each lobe is computed, and a severity score, ranging from 0 to 5, is assigned to each lobe. The infection rate and severity score mapping is shown in FIGURE 9. The cumulative score, ranging from 0 to 25, across these five lobes constitutes the CTSS for the CT scan. An analogous algorithm is devised for the axial slices of the CT scan. The lung mask (FIGURE 2-C) and infection mask (FIGURE 2-I) are used to estimate the CTSS for the CT scan. The pretrained UNet [56] categorizes each CT scan image into right, left, and background masks. The right lung mask is divided into three regions, and the left lung mask into two regions, as shown in FIGURE 9. The infection rate in the region is estimated by superimposing the infection mask onto the lung mask. Each region's score is obtained using the mapping table shown in FIGURE 9. The sum of the average scores across regions in each slice yields the proposed CTSS for the CT scan.

The Mehta dataset (14 patients) and a subset from the SPGC dataset (36 patients) are used for the CTSS analysis. An expert radiologist provides the CTSS for the subset of the SPGC dataset. The CTSS predicted by the proposed method and radiologist-predicted CTSS are compared in FIGURE 8. A strong correlation of 0.82 (Pearson correlation) between the CTSS predicted by the proposed method and the ground truth is observed, as well as a high cosine similarity of 0.97 between the proposed and ground truth CTSS score vectors of 50 patients. These results indicate that the CTSS predicted by the proposed method follows the same trend as the ground truth provided by the radiologist.

CTSS	
% of infection	Score
0	0
<= 5	1
5 - 25	2
25 - 50	3
50 - 75	4
>= 75	5

1) Divide the right lung into three lobes and left lung into two lobes
2) Find the percentage of infection in each lobes and assign a score based on the Table (CTSS)

FIGURE 9. A lung mask generated by UNet and an infection mask are used for estimating the severity score for a CT scan slice.

E. CHALLENGES & FUTURE WORK

One of the primary challenges lies in the availability and quality of labeled data. Creating accurate and comprehensive annotations for COVID-19 CT scans, especially at the slice-level, can be labor-intensive and time-consuming, often requiring expert radiologists. Apart from the SPGC dataset, all the other datasets do not contain slice-level information regarding COVID-19, CAP, and Normal classes. This makes the classification of CT scans difficult. We have employed a semi-supervised training method with the available labeled CT scan slices to augment the unlabelled slices in the SPGC dataset to train the classification model.

The proposed model is a multitasking system for COVID-19 detection, lung infection segmentation, and infection severity quantification. While classification models are effective with 2D CT scan slices, they may not capture dependencies between neighboring slices in volumetric CT scans. To address this, 3D CNN or temporal models could be explored to better capture slice-to-slice dependencies. Additionally, a single-stage preprocessing pipeline for infection region extraction can minimize error propagation compared to a sequential preprocessing pipeline. In terms of severity assessment, exploring the non-linear relationship between infection rate and severity score could enhance the accuracy of the assessment process.

VI. CONCLUSION

The proposed model is an automatic diagnosis system integrating COVID-19 detection, lung infection segmentation, and infection severity quantification into a unified framework. The intermediate stage output of the proposed unsupervised preprocessing pipeline is employed for infection segmentation, classification, and severity prediction. The domain expertise of radiologists helps to identify the lung infection regions during the preprocessing stages. The proposed unsupervised infection segmentation method performs better than state-of-the-art supervised semantic segmentation models. The preprocessed images (GMM+MO+B) are used to train the COVID-19 classifier in a semi-supervised way. The semi-supervised approach is adopted along with transfer learning to overcome the limitation of small amounts of training data. The proposed preprocessing pipeline has a positive impact on classification accuracy. The CTSS achieves a high correlation with the manual assessment of radiologists. The results confirm the importance of domain-specific expertise for feature extraction for COVID-19 detection and severity assessment.

REFERENCES

[1] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, and W. Tan, "A novel coronavirus from patients with pneumonia in China, 2019," *New England J. Med.*, vol. 382, no. 8, pp. 727–733, Feb. 2020.

[2] C. Long, H. Xu, Q. Shen, X. Zhang, B. Fan, C. Wang, B. Zeng, Z. Li, X. Li, and H. Li, "Diagnosis of the coronavirus disease (COVID-19): RRT-PCR or CT?" *Eur. J. Radiol.*, vol. 126, May 2020, Art. no. 108961.

[3] M. Dramé, M. T. Teguo, E. Proye, F. Hequet, M. Hentzien, L. Kanagaratnam, and L. Godaert, "Should RT-PCR be considered a gold standard in the diagnosis of COVID-19?" *J. Med. Virol.*, vol. 92, no. 11, pp. 2312–2313, Nov. 2020.

[4] E. Benmalek, J. Elmhamdi, and A. Jilbab, "Comparing CT scan and chest X-ray imaging for COVID-19 diagnosis," *Biomed. Eng. Adv.*, vol. 1, Jun. 2021, Art. no. 100003.

[5] S. Tabik, A. Gómez-Ríos, J. L. Martín-Rodríguez, I. Sevillano-García, M. Rey-Area, D. Charte, E. Guirado, J. L. Suárez, J. Luengo, M. A. Valero-González, P. García-Villanova, E. Olmedo-Sánchez, and F. Herrera, "COVIDGR dataset and COVID-SDNet methodology for predicting COVID-19 based on chest X-ray images," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 12, pp. 3595–3605, Dec. 2020.

[6] C. Öksüz, O. Urhan, and M. K. Güllü, "COVID-19 detection with severity level analysis using the deep features, and wrapper-based selection of ranked features," *Concurrency Comput., Pract. Exper.*, vol. 34, no. 20, Sep. 2022, Art. no. e6802.

[7] A. Nambu, "Imaging of community-acquired pneumonia: Roles of imaging examinations, imaging diagnosis of specific pathogens and discrimination from noninfectious diseases," *World J. Radiol.*, vol. 6, no. 10, p. 779, 2014.

[8] C. Hani, N. H. Trieu, I. Saab, S. Dangeard, G. Chassagnon, and M.-P. Revel, "COVID-19 pneumonia: A review of typical CT findings and differential diagnosis," *Diagnostic Interventional Imag.*, vol. 101, no. 5, pp. 263–268, May 2020.

[9] M. I. Mustafa, A. H. Abdelmoneim, E. M. Mahmoud, and A. M. Makhawi, "Cytokine storm in COVID-19 patients, its impact on organs and potential treatment by QTY code-designed detergent-free chemokine receptors," *Mediators Inflammation*, vol. 2020, Sep. 2020, Art. no. 8198963.

[10] T. C. Kwee and R. M. Kwee, "Chest CT in COVID-19: What the radiologist needs to know," *RadioGraphics*, vol. 40, no. 7, pp. 1848–1865, Oct. 2020.

[11] M. Chung, A. Bernheim, X. Mei, N. Zhang, M. Huang, X. Zeng, J. Cui, W. Xu, Y. Yang, Z. A. Fayad, A. Jacobi, K. Li, S. Li, and H. Shan, "CT imaging features of 2019 novel coronavirus (2019-nCoV)," *Radiology*, vol. 295, no. 1, pp. 202–207, Apr. 2020.

[12] L. Lin, G. Fu, S. Chen, J. Tao, A. Qian, Y. Yang, and M. Wang, "CT manifestations of coronavirus disease (COVID-19) pneumonia and influenza virus pneumonia: A comparative study," *Amer. J. Roentgenol.*, vol. 216, no. 1, pp. 71–79, Jan. 2021.

[13] A. Esposito et al., "Chest CT-derived pulmonary artery enlargement at the admission predicts overall survival in COVID-19 patients: Insight from 1461 consecutive patients in Italy," *Eur. Radiol.*, vol. 31, no. 6, pp. 4031–4041, Jun. 2021.

[14] R. Aoki, T. Iwasawa, E. Hagiwara, S. Komatsu, D. Utsunomiya, and T. Ogura, "Pulmonary vascular enlargement and lesion extent on computed tomography are correlated with COVID-19 disease severity," *Jpn. J. Radiol.*, vol. 39, no. 5, pp. 451–458, May 2021.

[15] A. Khan, R. Garner, M. L. Rocca, S. Salehi, and D. Duncan, "A novel threshold-based segmentation method for quantification of COVID-19 lung abnormalities," *Signal, Image Video Process.*, vol. 17, no. 4, pp. 907–914, Jun. 2023.

[16] L.-K. Huang and M.-J.-J. Wang, "Image thresholding by minimizing the measures of fuzziness," *Pattern Recognit.*, vol. 28, no. 1, pp. 41–51, Jan. 1995.

[17] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Comput. Vis., Graph., Image Process.*, vol. 29, no. 1, p. 140, Jan. 1985.
[18] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
[19] A. Tremeau and P. Colantoni, "Regions adjacency graph applied to color image segmentation," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 735–744, Apr. 2000.
[20] M. S. Brown, M. F. McNitt-Gray, N. J. Mankovich, J. G. Goldin, J. Hiller, L. S. Wilson, and D. R. Aberie, "Method for segmenting chest CT image data using an anatomical model: Preliminary results," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 828–839, Dec. 1997.
[21] R. Adams and L. Bischof, "Seeded region growing," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 6, pp. 641–647, Jun. 1994.
[22] J. Dehmeshki, H. Amin, M. Valdivieso, and X. Ye, "Segmentation of pulmonary nodules in thoracic CT scans: A region growing approach," *IEEE Trans. Med. Imag.*, vol. 27, no. 4, pp. 467–480, Apr. 2008.
[23] A. P. Mangan and R. T. Whitaker, "Partitioning 3D surface meshes using watershed segmentation," *IEEE Trans. Vis. Comput. Graph.*, vol. 5, no. 4, pp. 308–321, Oct./Dec. 1999.
[24] Y. Boykov and M.-P. Jolly, "Interactive organ segmentation using graph cuts," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Berlin, Germany: Springer, 2000, pp. 276–286.
[25] G. Li, X. Chen, F. Shi, W. Zhu, J. Tian, and D. Xiang, "Automatic liver segmentation based on shape constraints and deformable graph cut in CT images," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5315–5329, Dec. 2015.
[26] J. K. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation," *Graph. Models Image Process.*, vol. 58, no. 3, pp. 246–261, May 1996.
[27] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 359–369, Mar. 1998.
[28] S. C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 9, pp. 884–900, Sep. 1996.
[29] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *Int. J. Comput. Vis.*, vol. 50, no. 3, pp. 271–293, Dec. 2002.
[30] J. S. Schildkraut, S. Chen, M. Heath, W. G. O'Dell, P. Okunieff, M. C. Schell, and N. Paul, "Level-set segmentation of pulmonary nodules in radiographs using a CT prior," *Proc. SPIE*, vol. 7259, Mar. 2009, Art. no. 72593B.
[31] S. Hu, E. A. Hoffman, and J. M. Reinhardt, "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images," *IEEE Trans. Med. Imag.*, vol. 20, no. 6, pp. 490–498, Jun. 2001.
[32] Y. Zhu, Y. Tan, Y. Hua, G. Zhang, and J. Zhang, "Automatic segmentation of ground-glass opacities in lung CT images by using Markov random field-based algorithms," *J. Digit. Imag.*, vol. 25, no. 3, pp. 409–422, Jun. 2012.
[33] B. Ye, X. Yuan, Z. Cai, and T. Lan, "Severity assessment of COVID-19 based on feature extraction and V-descriptors," *IEEE Trans. Ind. Informat.*, vol. 17, no. 11, pp. 7456–7467, Nov. 2021.
[34] X. Wang, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, and C. Zheng, "A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2615–2625, Aug. 2020.
[35] X. Wu, C. Chen, M. Zhong, J. Wang, and J. Shi, "COVID-AL: The diagnosis of COVID-19 with deep active learning," *Med. Image Anal.*, vol. 68, Feb. 2021, Art. no. 101913.
[36] S. Xue and C. Abhayaratne, "COVID-19 diagnostic using 3D deep transfer learning for classification of volumetric Computerised tomography chest scans," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 8573–8577.
[37] J. Wang, Y. Bao, Y. Wen, H. Lu, H. Luo, Y. Xiang, X. Li, C. Liu, and D. Qian, "Prior-attention residual learning for more discriminative COVID-19 screening in CT images," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2572–2583, Aug. 2020.
[38] T. Javaheri et al., "CovidCTNet: An open-source deep learning approach to diagnose COVID-19 using small cohort of CT images," *npj Digit. Med.*, vol. 4, no. 1, p. 29, Feb. 2021.
[39] Z. Han, B. Wei, Y. Hong, T. Li, J. Cong, X. Zhu, H. Wei, and W. Zhang, "Accurate screening of COVID-19 using attention-based deep 3D multiple instance learning," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2584–2594, Aug. 2020.
[40] T. Anwar, "COVID19 diagnosis using AutoML from 3D CT scans," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 503–507.
[41] M. N. Teli, "TeliNet: Classifying CT scan images for COVID-19 diagnosis," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 496–502.
[42] P. Afshar, S. Heidarian, N. Enshaei, F. Naderkhani, M. J. Rafiee, A. Oikonomou, F. B. Fard, K. Samimi, K. N. Plataniotis, and A. Mohammadi, "COVID-CT-MD, COVID-19 computed tomography scan dataset applicable in machine learning and deep learning," *Sci. Data*, vol. 8, no. 1, p. 121, Apr. 2021.
[43] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, K. Cao, D. Liu, G. Wang, Q. Xu, X. Fang, S. Zhang, J. Xia, and J. Xia, "Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the diagnostic accuracy," *Radiology*, vol. 296, no. 2, pp. 65–71, Aug. 2020.
[44] S. Chaudhary, S. Sadbhawna, V. Jakhetiya, B. N. Subudhi, U. Baid, and S. C. Guntuku, "Detecting COVID-19 and community acquired pneumonia using chest CT scan images with deep learning," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 8583–8587.
[45] D. Kollias, A. Arsenos, L. Soukissian, and S. Kollias, "MIA-COV19D: COVID-19 detection through 3-D chest CT image analysis," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshops (ICCVW)*, Oct. 2021, pp. 537–544.
[46] P. Garg, R. Ranjan, K. Upadhyay, M. Agrawal, and D. Deepak, "Multi-scale residual network for COVID-19 diagnosis using CT-scans," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 8558–8562.
[47] T. D. Pham, "A comprehensive study on classification of COVID-19 on computed tomography with pretrained convolutional neural networks," *Sci. Rep.*, vol. 10, no. 1, p. 16942, Oct. 2020.
[48] M. Singh, S. Bansal, S. Ahuja, R. K. Dubey, B. K. Panigrahi, and N. Dey, "Transfer learning-based ensemble support vector machine model for automated COVID-19 detection using lung computerized tomography scan data," *Med. Biol. Eng. Comput.*, vol. 59, no. 4, pp. 825–839, Apr. 2021.
[49] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
[50] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. 36th Int. Conf. Mach. Learn.*, 2019, pp. 6105–6114.
[51] S. P. Morozov, A. E. Andreychenko, I. A. Blokhin, P. B. Gelezhe, A. P. Gonchar, A. E. Nikolaev, N. A. Pavlov, V. Y. Chernina, and V. A. Gombolevskiy, "MosMedData: Data set of 1110 chest CT scans performed during the COVID-19 epidemic," *Digit. Diag.*, vol. 1, no. 1, pp. 49–59, Dec. 2020.
[52] H. B. Jenssen, 2021, "MedSeg COVID dataset," *MedSeg*, doi: 10.6084/m9.figshare.13521488.
[53] M. Jun, G. Cheng, W. Yixin, A. Xingle, G. Jiantao, Y. Ziqi, Z. Minqing, L. Xin, D. Xueyuan, C. Shucheng, W. Hao, M. Sen, Y. Xiaoyu, N. Ziwei, L. Chen, T. Lu, Z. Yuntao, Z. Qiongjie, D. Guoqiang, and H. Jian, Apr. 2020, "COVID-19 CT lung and infection segmentation dataset," *Zenodo*, doi: 10.5281/zenodo.3757476.
[54] S. Heidarian. (2021). *SPGC-COVID Dataset*. [Online]. Available: https://figshare.com/articles/dataset/SPGC-COVID_Dataset/166-32397/1
[55] P. Afshar, M. J. Rafiee, F. Naderkhani, S. Heidarian, N. Enshaei, A. Oikonomou, F. B. Fard, R. Anconina, K. N. Plataniotis, and A. Mohammadi, 2021, "COVID-19 low-dose and ultra-low-dose CT scans," *IEEE Dataport*, doi: 10.21227/sed8-6r15.
[56] J. Hofmanninger, F. Prayer, J. Pan, S. Röhrich, H. Prosch, and G. Langs, "Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem," *Eur. Radiol. Experim.*, vol. 4, no. 1, p. 50, Dec. 2020.
[57] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.

[58] T. Jerman, F. Pernuš, B. Likar, and Ž. Špiclin, "Enhancement of vascular structures in 3D and 2D angiographic images," *IEEE Trans. Med. Imag.*, vol. 35, no. 9, pp. 2107–2118, Sep. 2016.
[59] P. Soille, *Morphological Image Analysis: Principles and Applications*. Cham, Switzerland: Springer, 2013.
[60] L. Vincent, "Morphological area openings and closings for grey-scale images," in *Shape in Picture*. Berlin, Germany: Springer, 1994, pp. 197–208.
[61] J. M. Prewitt, "Object enhancement and extraction," *Picture Process. Psychopictorics*, vol. 10, no. 1, pp. 15–19, 1970.
[62] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*.
[63] O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz, B. Glocker, and D. Rueckert, "Attention U-Net: Learning where to look for the pancreas," 2018, *arXiv:1804.03999*.
[64] J. Schlemper, O. Oktay, M. Schaap, M. Heinrich, B. Kainz, B. Glocker, and D. Rueckert, "Attention gated networks: Learning to leverage salient regions in medical images," *Med. Image Anal.*, vol. 53, pp. 197–207, Apr. 2019.
[65] X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes," *IEEE Trans. Med. Imag.*, vol. 37, no. 12, pp. 2663–2674, Dec. 2018.
[66] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham, Switzerland: Springer, 2018, pp. 3–11.
[67] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder–decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
[68] D.-P. Fan, T. Zhou, G.-P. Ji, Y. Zhou, G. Chen, H. Fu, J. Shen, and L. Shao, "Inf-Net: Automatic COVID-19 lung infection segmentation from CT images," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2626–2637, Aug. 2020.
[69] J. Ma, Y. Wang, X. An, C. Ge, Z. Yu, J. Chen, Q. Zhu, G. Dong, J. He, Z. He, T. Cao, Y. Zhu, Z. Nie, and X. Yang, "Toward data-efficient learning: A benchmark for COVID-19 CT lung and infection segmentation," *Med. Phys.*, vol. 48, no. 3, pp. 1197–1210, Mar. 2021.
[70] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "NnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021.
[71] K. J. Kiser, S. Ahmed, S. Stieb, A. S. R. Mohamed, H. Elhalawani, P. Y. S. Park, N. S. Doyle, B. J. Wang, A. Barman, Z. Li, W. J. Zheng, C. D. Fuller, and L. Giancardo, "PleThora: Pleural effusion and thoracic cavity segmentations in diseased lungs for benchmarking chest CT processing pipelines," *Med. Phys.*, vol. 47, no. 11, pp. 5941–5952, Nov. 2020.
[72] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626.
[73] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

ANAND THYAGACHANDRAN (Student Member, IEEE) received the bachelor's degree in computer science and engineering from the Cochin University of Science and Technology, Kerala, India, in 2011, and the master's degree in information technology from the University of Hyderabad, Telangana, India, in 2016. He is currently pursuing the Ph.D. degree with the Computer Science Department, IIT Madras, Chennai, India. Since 2019, he has been a Teaching Assistant with IIT Madras. His research interests include medical image processing, computational brain research, and other areas of machine learning and signal processing.

AATHIRA BALACHANDRAN received the Bachelor of Medicine and Bachelor of Surgery degrees from the Government Medical College Trivandrum, Kerala, India, in 2011, and the master's degree in radiodiagnosis from the Government Thirumala Devaswom Medical College, Alappuzha, Kerala, India, in 2016. From 2017 to 2020, she was a Consultant Radiologist with the Government Medical College Kollam and Metro Scans & Laboratory, Kollam, Kerala, India. She is currently an Assistant Professor with the Radio Diagnosis Department, Government Medical College Kollam. Her research interests include diagnostic values of ultrasound in resource-limited settings in the areas of head and neck malignancies, gynecological malignancies, and antenatal ultrasound scans.

HEMA A. MURTHY (Senior Member, IEEE) received the bachelor's degree from Osmania University, Hyderabad, India, in 1980, the master's degree from McMaster University, Hamilton, ON, Canada, in 1986, and the Ph.D. degree from IIT Madras, Chennai, India, in 1992. From 1980 to 1983, she was a Scientific Officer (SO/SC) with the Speech and Digital Systems Group, TIFR, Mumbai. She is currently a Professor with the Department of Computer Science and Engineering, IIT Madras. She has over 39 journal publications, two book publications, and over 220 articles. Her research interests include speech processing, computer networks, music information retrieval, computational brain research, and other areas of machine learning and signal processing.