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

Anand Thyagachandran, Aathira Balachandran, and Hema A Murthy

***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 pre-trained ImageNet model to extract discriminative features. Subsequently, a shallow feed-forward neural network is trained to classify the available CT scans into three classes namely COVID-19, Community-Acquired Pneumonia, and Normal. Through various ablation studies, we find that a domain-specific pre-processing pipeline improves classification accuracy significantly. In terms of classification accuracy, our approach when evaluated on publicly available datasets is seen to have an absolute improvement of 6% F1 score over the baseline model. Further, the estimated infection severity score is observed to be well correlated with radiologists' assessments. The results support the necessity of data-driven pre-processing before implementing learning algorithms.**

***Index Terms*— COVID-19, CT scans, infection segmentation, severity assessment, semi-supervised augmentation.**

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. 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.

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 [1]. 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. Also, 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 Scan) [2]. The limitation of traditional X-ray imaging in being a line-of-sight integration results in poorer resolution and hence a poorer prognosis of COVID-19 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 [3], [4]. This can lead to a failure in managing the complications of COVID-19, such as cytokine storms, which are a major cause of fatalities in COVID-19 cases [5]. 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) and COVID-19 and also against normal CT-scan through a pre-processing 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 regions 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 [6]–[8]. Common features of GGO include opaque foggy areas which do not obstruct the pulmonary vessels while consolidation is marked

Manuscript submitted on 5/2/23.

Anand Thygachandran and Hema A Murthy are with Computer Science and Engineering Department, Indian Institute of Technology Madras, India (e-mail: tanand@cse.iitm.ac.in, hema@cse.iitm.ac.in).

Aathira Balachandran is with Radio Diagnosis Department, Government Medical College Kollam, India (e-mail: draathirab@gmail.com).

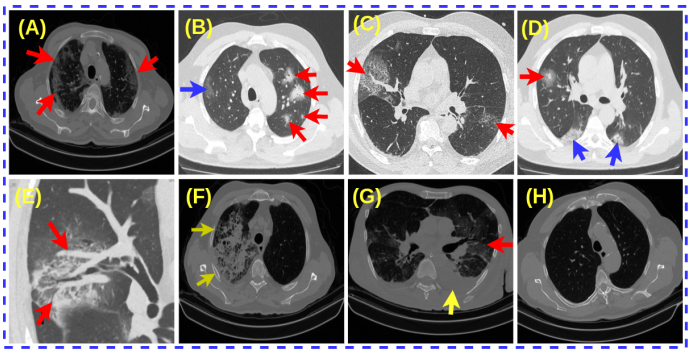

Fig. 1: Clinical image findings: The imaging findings of COVID-19 and CAP in chest CT scans. Images (A)-(E) show the COVID-19 patient's chest CT scan slices, (F) and (G) show the CAP patient's chest CT scan slices. (A) Peripheral distribution of GGO in both lungs (red arrows). (B) Patchy areas of consolidations 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 (yellow arrow) and GGOs (red arrows) in the left lung. (H) Normal chest CT scan slice.

by higher opacity thereby rendering the visualization of the pulmonary vascular structures to be 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 regarded as the primary indicators for COVID-19. As the disease severity progresses, consolidations, crazy-paving patterns, and vascular enlargement [9], [10] are seen to be the hallmark features. Pleural effusion and significant mediastinal lymphadenopathy are less commonly observed findings in COVID-19 infection. Pleural effusion refers to the accumulation of 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 [3]. An experienced radiologist was consulted to obtain the Hounsfield Unit (a measure of the range of radiation attenuation values – HU) corresponding to different regions in the CT scan. The clinical features observed for COVID-19 and CAP are shown in Fig. 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 contributions of the paper are as follows:

- A pre-processing module to segment regions of infection derived from chest CT scans of COVID-19 patients.
- Prediction of CT severity score (CTSS) for severity assessment.

The paper is organized as follows- Discussion on segmentation and identification of COVID-19 from chest CT scans as presented in Section II. Section III describes the segmentation of relevant lung features in the CT scan contributing to infection. We also briefly explain the training strategy to classify CT scans into COVID-19, CAP, and Normal categories. Section IV details the set of experiments and discussion on results and inferences. Section V concludes the work.

II. Related Work

Prior to the classification/identification of COVID-19, CT scans have primarily relied on raw images, i.e. the data as obtained from a diagnostic procedure. In related fields, however, pre-processing techniques such as image enhancement and segmentation have been employed extensively for enhancing image quality through the focus on dominant features and noise removal. In our work, we amalgamate generic image pre-processing techniques to aid image classification algorithms to achieve the objective of identifying the prevalence of COVID-19 from chest CT scans. Although several algorithms for pre-processing do exist, we utilize knowledge-based machine learning approaches for the said purpose before embarking on classification. Considering the limitation in obtaining slice-level labels, which however is required for better classification of the CT scans, a classifier is trained using the available data, with the training data being further augmented with pseudo-labeled data to improve 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. An unsupervised way of segmentation based on domain knowledge is extensively used in the proposed work. 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. Azrin *et. al.* [16] utilized the threshold-based segmentation algorithms such as Huang [17], Kapur [18], and Otsu [19], to binarize the CT scans and generated a region adjacency graph (RAG) [20] to demarcate COVID-19 lesions from CT scans. Threshold-based methods are very simple in approach and implementation if provided with images bearing sufficient contrast. The absence of sufficient contrast leads to a downgrade in performance while considering differences in image attenuation. Region-based methods segment an area by assessing the homogeneity of the neighboring pixels. Widely practised region-based algorithms include region-growing [21]–[23], watershed [24], graph cuts [25], [26], and fuzzy connectedness [27]. 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 [28], active contours [29], and level sets [30], [31]. 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

TABLE I: Different datasets used in the experiment. "*" – slice-level labels are available; "**" – slice-level labels are not available but the CT scan patient-level labels are available.

Dataset	Class	COVID-19	CAP	Normal	Classification(C)/ Segmentation(S)	Source	Data Format
SPGC [11]	Train	55*+116**	25*+35**	76	C	ICASSP COVID-19 SignalProcessing Grand Challenge (SPGC) 2021	DICOM
	Test	28**	51**	51	C		DICOM
LDCT [12]	Test	104**	0	56	C	IEEE Dataport, 2021	DICOM
LDCT-PCR [12]	Test	100*	0	0	C		DICOM
Mosmed [13]	Test	854**	0	254	C	Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Department of Health Care of Moscow, Russia	NIfTI
Mosmed [13]	Test	50*	0	0	S		NIfTI
MedSeg [14]	Test	100* slices	0	0	S	Italian Society of Medical and Interventional Radiology	NIfTI
MedSeg_1 [15]	Test	9* (638 slices)	0	0	S	Italian Society of Medical and Interventional Radiology	NIfTI
Mehta	Test	14**	0	0	S	Mehta multispeciality hospital, Chennai, India	DICOM

et al. [32] 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.* [33] proposed a method to extract GGOs by modeling their intensity distributions and using the Markov random field model to improve the identification of the boundary. Ben *et al.* [34] 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. [35] 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 be accurately focused on the region of interest. This may lead to less accurate predictions and a higher rate of false positives.

To avoid the fore-described shortcomings, we propose a three-mixture Gaussian mixture model inspired by the intensity distribution adaptive model using maximum-a-posteriori (MAP) adaptation as proposed by Zhu *et al.* [33]. We formulate, as the name suggests, three Gaussian distributions using adaptive thresholding to extract the clinical features (defined by the radiologists in Section I)) in the proposed work.

B. CT scan classification

Previous studies have made significant progress in identifying COVID-19 using CT scan images, which can be broadly divided into two categories: 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 [35]–[40]. Among these techniques, segmentation of the lung region using image pre-processing methods is applied before performing classification. Owing to the varying dimensionality of 3D CT scans, interpolation or truncation of the slices are applied to convert them to fixed dimensionality which might lead to information loss [41]. In 2D CT scan-based classification, a 2D CNN is trained on individual slices and slice-level probability scores are generated. Further, threshold-based [41], majority voting [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 the traces of infections are not visible as in the majority of CT scan slices. Considering the difficulty in obtaining annotated CT scans, transfer learning methods have been extensively employed in COVID-19 classification. Some of the transfer learning works [43], [44], [46]–[48] have explored different CNN models trained on ImageNet [49] dataset for classification tasks. Transfer learning methods reduce the requirement for training for every dataset and also provide discriminative features for classifying COVID-19, even while being applied to raw CT scan images.

Chaudary *et al.* [44] employed transfer learning in generating features from raw CT images using the EfficientNet-B5 model [50]. The slice-level scores were obtained using a shallow FFNN. Further, patient-level classification was performed using a weighted average method on the slice-level scores. This model with higher accuracy in classifying COVID-19 serves as the baseline for comparison with the present work. Inspired by this technique [44], the present work attempts to enhance the performance in classifying COVID-19 in a CT scan by using 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. Proposed System

The proposed system is designed to accomplish three key tasks: 1) A novel pre-processing pipeline to extract relevant clinical features, 2) A Semi-supervised method for the classification of CT scans, and 3) A CT severity score is generated for patients by utilizing the pre-processing pipeline. The proposed system pre-processes the CT scan images initially, which are then fed to a model trained on the ImageNet dataset for extracting high-dimensional features. These features are subsequently used to train a shallow feed-forward neural network (FFNN) to predict slice-level scores. A weighted average method is then applied to calculate the final score for the CT scans. Additionally, a severity analysis module is

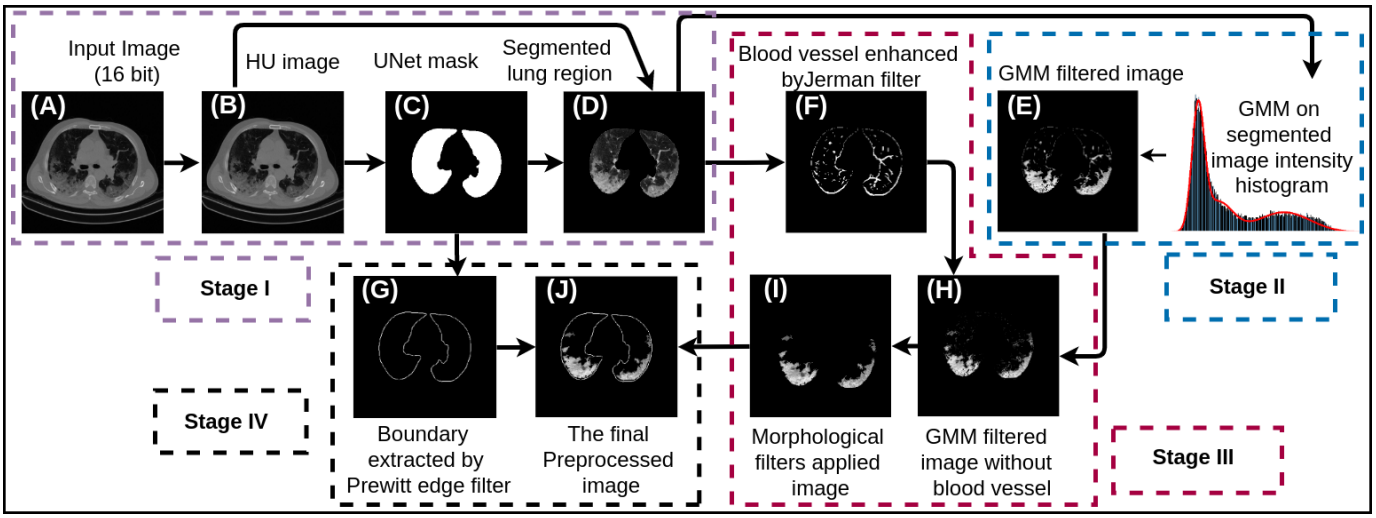

Fig. 2: Proposed model architecture.

implemented to determine the CT severity score (CTSS). We evaluate the system's performance on six publicly available online datasets, the details of which are summarized in Table I.

A. Pre-processing pipeline

Radiologists often use common clinical features such as GGO, consolidation, crazy paving pattern, halo sign, reverse halo sign, and pleural effusion to distinguish between COVID-19 and CAP from healthy individuals. The abnormalities present in the CT scan exhibit attenuation variations with respect to the severity of the infection. Based on the observations in Fig. 1, an image processing pipeline is proposed with primary emphasis on the differences between the three classes of interest — COVID-19, CAP, and Normal. The proposed pre-processing pipeline is shown in Fig. 2. The pre-processing pipeline is developed in an unsupervised manner, by integrating traditional image processing techniques in tandem with machine learning and deep learning algorithms.

1) Stage-I: Lung mask generation

A CT scan is a volumetric scan consisting of multiple slices of resolution of 512x512 pixels. In a chest CT scan, tissues, heart, stomach, blood vessels, and bones possess higher attenuation (HU) values than air and lung areas. The pre-processing pipeline is applied on the HU scale of CT scan images. CT scans are mainly available in two formats namely, Neuroimaging Informatics Technology Initiative (NIfTI) and Digital Imaging and Communications in Medicine (DICOM). The pixel intensity values of the CT scan slices are HU units in NIfTI format, whereas DICOM format CT scan slices are converted to HU scale images using a linear transformation. The preliminary task is to remove the unnecessary areas in the CT scan and identify the region of interest (lung region). For this, the HU scale image is fed to the pre-trained UNet model to extract the region of interest. The pre-trained UNet [51] is an end-to-end, fully convolution neural network containing an encoder module that compresses input CT images using convolution and max-pooling operations into a fixed-length feature map. The decoder module (which provides a better spatial representation) upsamples these feature maps to the lung mask. The skip connections between the encoder and decoder modules enhance the semantic features for segmentation. The pre-trained UNet model segments the lung CT scan into three classes: the left lung region, the right lung region, and the background. The UNet model removes unnecessary objects such as bones, trachea, and organs and retains only the lung region. The lung region provides a clear segmented image that makes it easier to identify any potential issues.

The initial and final slices of the CT scan typically contain structures such as bones, the trachea, the diaphragm, the heart, and the stomach which are not relevant to further analysis and hence not considered. A fixed threshold on the number of slices is not possible as the number of slices can vary between patients. Instead, a lung mask is used as a threshold to remove the initial and final slices. A criterion based on the degree of lung involvement is used to decide the number of slices for further analysis for every individual patient. This approach ensures that only the slices that are most relevant to COVID-19 analysis are included, improving the accuracy of the diagnosis.

2) Stage-II: GMM-based adaptive filter

The range of attenuation values (HU) of features like GGO, consolidation, and pleural effusion are discussed in (Section I). The pixel intensity values of GGO, consolidation, and pleural effusion regions can vary depending on the severity of the infection. Severe infection regions have higher pixel intensities than mild infection regions. This variability in the attenuation is captured by modeling the pixel's intensity histogram with a three-mixture GMM. Each mixture attempts to capture the GGO, consolidation, pleural effusion, and background patterns. This unsupervised clustering technique is applied to the pre-processing (Stage I) images. HU ranges for the clinical features were defined by the radiologist. Each mixture in GMM is represented by a mean μ, a standard deviation σ, and a threshold of $(\mu \pm 1.5 * \sigma)$ is considered for the pixel selection. The pixel intensities lie between the range as defined above and the radiologist proposed are considered for further analysis. While the GMM adaptive filter accurately segments the infection regions, the presence of blood vessels affects the severity analysis. To address this issue, the images are further subjected to vascular analysis.

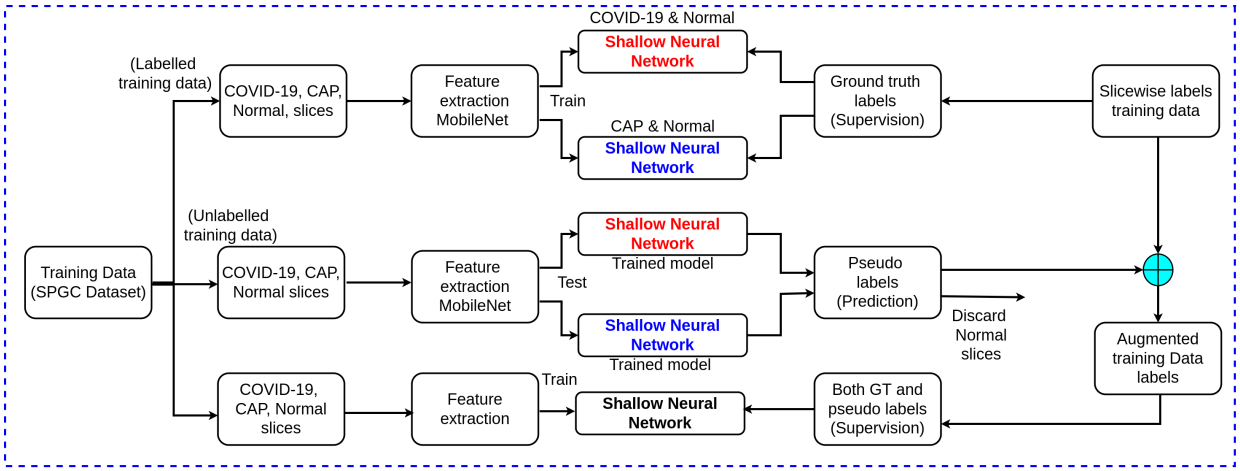

Fig. 3: The proposed semi-supervised training method for the classifier

3) Stage-III: Vascular enhancement and morphological filters

After Stage II of pre-processing, the image contains essential features of COVID-19 and CAP but also details of bronchi and primary pulmonary vessels, which can resemble tubular structures. The vessels share similar HU values as that of 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 [52] is a widely used filter 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 ratio of the largest to the smallest eigenvalue 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 marked and removed from the image thereby creating holes in the GGO region. The flood fill algorithm [53] 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 [54].

4) Stage-IV: Generate lung boundary and grayscale image

The Prewitt edge detector, as proposed in [55], is employed to extract the lung boundary from the lung mask generated from Stage-I. This lung boundary aids to identify the localization and distribution of the clinical features in the CT scan. Furthermore, the resultant HU scale image is converted to an 8-bit grayscale image using min-max normalization. This grayscale image is then resized to a three-channel image to match the input dimension of the model pre-trained on the ImageNet dataset.

B. Feature extraction by models pre-trained on the ImageNet dataset

Considering that the training dataset does not contain sufficient labeled 2D CT scan slices to train a deep CNN, state-of-the-art computer vision models trained on the ImageNet dataset are used to extract high-dimensional features of the pre-processed CT scan images. In this work, five deep CNN models namely, MobileNet [56], ResNet-101 [57], ResNet-50 [57], EfficientNet-B5 [50], and EfficientNet-B1 [50] are used to extract discriminative features from the pre-processed CT scan images. The MobileNet [56] is a lightweight model that generates a 1024-dimensional feature vector for each 2D image. The novel depth-wise convolution in each layer reduces the number of parameters compared to a network with regular convolutions with the same depth. The ResNet models [57] involve the idea of skip connections which allow for building deep CNN models. These models perform better than models without skip connections. The deep CNN model with skip connections can handle the problem of vanishing gradient by following an alternate method for backpropagation. Tan *et al.* [50] proposed a simple and effective way of scaling up CNN models to obtain better accuracy using compound coefficients. The technique allows the EfficientNet models to achieve higher accuracy with fewer parameters by focusing on the design choice of the model, thereby performing better than random scaling up of the model's width, depth, or the resolution of the feature map. Compared to other convolutional neural networks, the compound scaling method scales the input dimensions at a constant ratio to enhance classification accuracy. The ResNet and EfficientNet models generate a 2048-dimensional feature vector for each pre-processed CT scan image.

C. Training

In the present work, a semi-supervised approach is proposed to label the unannotated CT scan images in the SPGC training dataset [11] to augment the final training dataset with more COVID-19 and CAP slices, as the dataset is not sufficient to train a deep CNN model from scratch. The labeled images are first pre-processed and then fed to the pre-trained MobilNet to extract slice-level feature maps. Each feature map is then

given to the global average pooling layer to make it a 1024-dimensional feature vector. Two shallow feed-forward neural networks (FFNN), with 1024 neurons and two neurons in the first and last layers, respectively, are trained on the 1024-dimensional extracted feature vectors. Among the two models, one FFNN is trained with an equal amount of features of COVID-19 and Normal CT scans, and the other FFNN is trained with an equal amount of features of CAP and Normal CT scans, named Mobilenet_COVID and MobileNet_CAP for future references, respectively. Although the subset of the SPGC training dataset does not contain labeled CT scan images at the slice-level, it provides patient-level labels for the entire set of CT scans. Unlabeled slice-level CT scan images are pre-processed using the pre-processing pipeline developed in Section III-A and then fed to the pre-trained ImageNet models for feature extraction. The features corresponding to the COVID-19 patients are given to the MobileNet_COVID model; similarly, the features of unlabelled CAP slices are given to the MobileNet_CAP model. CT scan slices with predicted labels as COVID-19 and CAP are further used to augment the final training dataset, and slices with the label as Normal are discarded. The pseudo-labeled COVID-19 and CAP slices, along with the labeled training dataset, are used to train the final classifier. The final classifier consists of three layers: two dense layers (2048 and 1024 neurons) and a final layer (three neurons). Each neuron in the last layer gives a probability score for each class: CAP, COVID-19, and Normal. The final classifier is trained with an equal number of CT scans slices from COVID-19, CAP, and Normal categories to avoid class imbalance. The complete pipeline of the training process is shown in Fig. 3.

D. Patient-level classification

The three-class classifier generates a probabilistic score for each CT scan slice. Since the CT scan is volumetric and has a dimension of (n * 512x512), patient-level annotation is preferable to slice-level classification. Thus, a weighted averaging method is applied to the probability score generated by the classifier. The CT scan volume (with 'n' images) is divided into three equal regions, and each part is associated with different weights (W_1, W_2, W_3), where W_1, W_2, W_3 are 0.7, 1, and 0.7 respectively. The middle region CT scan slices have a large lung region and contain more information. Hence, more weightage is assigned to the slices from this region. If the weighted sum of the probabilities of COVID-19 and CAP categories is greater than that of the Normal class, the CT scan is considered abnormal and classified as either COVID-19 or CAP based on the scores. The LDCT, Mosmed, and LDCT-PCR datasets do not contain any CAP CT scans. The CT scan slices predicted as CAP class are considered to belong to the COVID-19 category for these datasets. Consider the predicted score for the i^{th} slice is P_i. The patient label is calculated as:

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

IV. Results and Discussions

The proposed pipeline is thoroughly evaluated on several aspects including segmentation and classification accuracy.

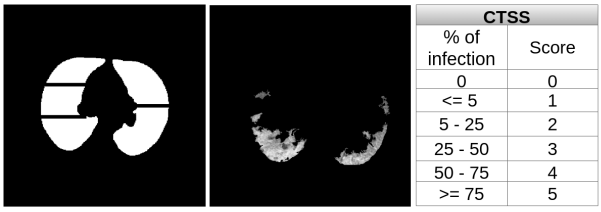

Fig. 4: A lung mask generated by UNet and an infection mask is used for estimating the severity score for a CT scan slice.

Additionally, ablation studies are conducted to evaluate the significance of the image pre-processing pipeline in the overall performance of the proposed method.

A. Evaluation metrics

The proposed pipeline is evaluated using several widely used metrics for segmentation and classification. For segmentation, the pipeline is evaluated using metrics such as Dice score (Dice), sensitivity (Sen.), specificity (Spec.), precision (Prec.), and mean absolute error (MAE). Dice score computes the error in segmentation by computing the overlap between annotated and predicted areas. Precision denotes the number of accurate white pixel predictions out of the overall white pixel predictions by the model. At the same time, specificity represents the number of correct predictions of black pixels from the total black pixels in the ground truth image. The MAE finds the average absolute difference between the predicted and annotated binary masks and quantifies the quality of the ROI predicted by the model. A lower MAE value indicates a better segmentation by the model. For classification, the pipeline is evaluated using metrics such as sensitivity, precision, and F1 score. Sensitivity denotes the number of CT scan volumes correctly predicted with respect to the ground truth. It determines how well the model can discriminate patients' input CT scans with respect to the ground truth. Precision denotes the number of CT scans correctly predicted by the model's overall prediction. F1 score can be quantified as the harmonic mean of sensitivity and precision. The F1 score is similar to the Dice score in the segmentation task. The proposed pipeline also predicts the CT severity score (CTSS) and evaluates the prediction using the Pearson correlation coefficient and cosine similarity. The Pearson correlation measures the ratio between the covariance of the predicted and ground truth scores and their standard deviations, showing the trend between the proposed and ground truth CTSS. Cosine similarity calculates the angle between the proposed and ground truth CTSS vectors, providing a measure of the prediction accuracy relative to the ground truth. The Pearson correlation coefficient and cosine similarity range from -1 to 1, with positive correlations and higher similarities indicated by values greater than zero.

B. Segmentation results

The output image from Stage-III of the pre-processing pipeline (image (I) in Fig. 2) is used as the infection mask for the CT scan slices. This pre-processed image (I) is generated by adaptively filtering the clinical features using a three-mixture

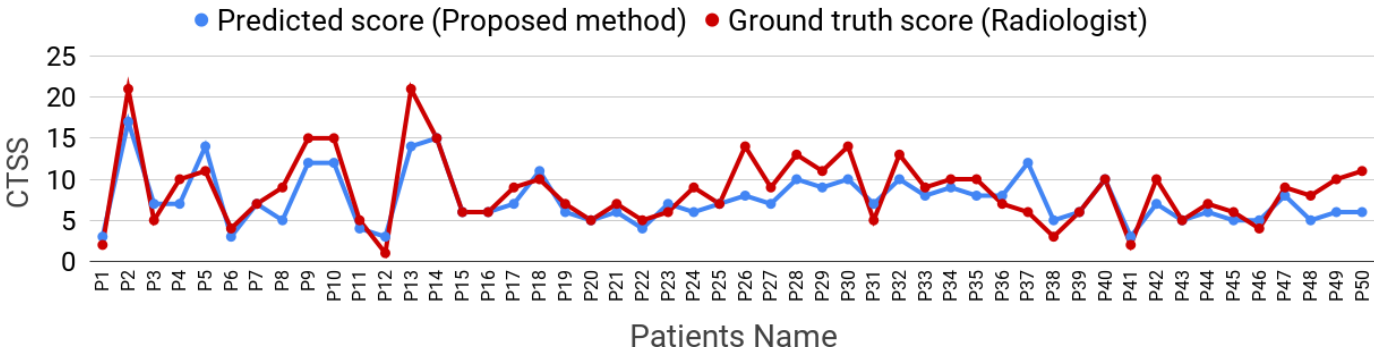

Fig. 5: CTSS for 14 patients (Mehta) and 36 patients (a subset of SPGC datasets).

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

Model	Dice	Sen.	Spec.	MAE
UNet [59]	0.439	0.534	0.858	0.186
Attention UNet [60]	0.583	0.637	0.921	0.112
Gated UNet [61]	0.623	0.658	0.926	0.102
Dense UNet [62]	0.515	0.594	0.840	0.184
UNet++ [63]	0.581	0.672	0.902	0.120
InfNet [58]	0.682	0.692	0.943	0.082
SemiInfNet [58]	0.739	0.725	0.960	0.064
Our Method	**0.673**	**0.678**	**0.9852**	**0.0356**

TABLE III: The infection segmentation results with the MedSeg_1 Dataset (nine patients' real CT scan volumes).

Model	Dice	Sen.	Spec.	Prec.	MAE
UNet [59]	0.308	0.678	0.836	0.265	0.214
Attention UNet [60]	0.466	0.723	0.930	0.390	0.095
Gated UNet [61]	0.447	0.674	0.956	0.375	0.066
Dense UNet [62]	0.410	0.607	0.977	0.415	0.167
UNet++ [63]	0.444	0.877	0.929	0.369	0.106
InfNet [58]	0.579	0.870	0.974	0.500	0.047
SemiInfNet [58]	0.597	0.865	0.977	0.515	0.033
Our Method	**0.726**	**0.775**	**0.996**	**0.711**	**0.0057**

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

	Task 1			Task 2		Our method
	MSD	Struct seg	NSCLC	Separate	Union	Infection Mask
Dice score	0.392	0.443	0.301	0.588	0.482	**0.752**
Specificity	1.00	1.00	1.00	0.999	0.999	**0.998**
Sensitivity	0.364	0.422	0.249	0.575	0.601	**0.757**
Precision	0.614	0.607	0.614	0.679	0.577	**0.812**

GMM; then, blood vessels are removed by the Jerman filter and the extracted features are fine-tuned with the morphological filters. In a previous study, Fan *et al.* [58] evaluated five baseline models based on the different variants of the UNet [59] architecture with the MedSeg dataset (48 slices). The slices in the MedSeg dataset contain a wide range of infections, but non-infectious slices are absent. The infection segmentation results of the proposed model are evaluated with the same dataset with the same CT scan slices, and the results are summarized in Table II. The proposed model achieves better results in dice score, specificity, and MAE with respect to the baseline, and also the models proposed by Fan *et al.*, namely, InfNet and SemiInfNet. Since the MedSeg dataset is a random collection of slices from different patients, Fan *et al.* conducted infection segmentation with another dataset (MedSeg_1) comprising nine patients with 638 slices (285 non-infected and 353 infected slices). The results are summarized in Table III. It is again observed that the preprocessing pipeline improves performance, especially when the non-infected slices are identified more accurately than the baseline, InfNet, ad SemiInfNet models. Moreover, the proposed pipeline is able to detect even minor infections based on the statistical properties of the attenuation in the infection regions.

Ma *et al.* [64] used a nnUNet (no new UNet) 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 [65]. The in-domain datasets contain 20 COVID-19 CT scans with an infection range of 0.01% - 59% [66]. 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). Out of the two benchmark results, one model is trained for both lung segmentation, and infection segmentation (Union), and the other is prepared only for infection segmentation (separate). The Mosmed dataset [13] is used as blind test data for the baseline models, and the baseline results along with our proposed model results are summarized in Table IV. The proposed model provides better results regarding dice score, sensitivity, and precision. 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. Severity Analysis

Determining the severity of infection in a COVID-19 patient is essential in determining the appropriate course of treatment. 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 percentage of infection in each lobe is calculated based on the infection rate as shown in

The Mehta dataset (14 patients) and a subset from the

TABLE V: An ablation study on the different pre-processing stages of the SPGC test dataset with four different features.

Category	Feature Extractor	OI	GMM	GMM + MO	GMM + B	GMM + MO + B	GMM+ MO + JF	GMM + MO + JF + B
COVID-19	**ResNet-50**	0.71	0.78	**0.83**	0.73	0.81	0.68	0.77
	ResNet-101	0.46	0.83	0.78	0.80	**0.87**	0.81	0.71
	EfficientNet-B1	0.72	**0.81**	0.77	**0.81**	**0.81**	0.79	0.76
	EfficientNet-B5	0.77	0.71	0.73	0.80	**0.83**	0.75	0.74
CAP	**ResNet-50**	0.64	0.78	**0.89**	0.51	0.81	0.29	0.70
	ResNet-101	0.65	**0.81**	**0.81**	**0.81**	0.78	0.80	0.50
	EfficientNet-B1	0.72	0.74	0.74	**0.85**	0.78	0.76	0.76
	EfficientNet-B5	**0.85**	0.57	0.73	0.79	0.78	0.68	0.73
Normal	**ResNet-50**	0.64	0.90	0.86	0.85	**0.93**	0.76	0.84
	ResNet-101	0.68	0.87	0.83	0.85	**0.92**	0.89	0.79
	EfficientNet-B1	0.77	0.88	0.84	0.85	**0.92**	0.87	0.82
	EfficientNet-B5	0.83	0.79	0.82	0.87	**0.93**	0.88	0.83

TABLE VI: The slice-level classification of the LDCT dataset with different features. C-19 and NL represent COVID-19 and Normal categories.

Feature Extractor	Sensitivity		Precision		F1 Score	
	C-19	NL	C-19	NL	C-19	NL
ResNet50	0.90	0.86	0.88	0.89	0.89	0.88
ResNet101	0.82	0.89	0.89	0.82	0.86	0.86
EfficientnetB5	0.88	0.88	0.89	0.87	0.88	0.87
EfficientNetB1	0.80	0.93	0.92	0.81	0.86	0.86

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 is compared in Fig. 5. A strong correlation of 0.82 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.

D. Classification results

Features are extracted using pre-trained networks, as stated earlier. A shallow feed-forward neural network is trained using a semi-supervised learning method. The classification results for the ablation study, as well as the slice-level and patient-level, are summarised in Table V, Table VI, Table VII, and Fig. 6 respectively.

1) Ablation Study

The proposed work differs significantly from previous work in its image processing pipeline. To evaluate the contribution of each stage in the pre-processing pipeline to classification accuracy, a series of ablation studies are performed. These studies include additional pre-processing steps in the pipeline. The different intermediate pre-processed images used in the ablation study are original image (OI), Stage-II with GMM (GMM), Stage-II with morphological operations (GMM + MO), Stage-II with lung region boundary (GMM + B), Stage-II with morphological operations and lung boundary (GMM + MO + B), Stage-III output image (GMM + MO + JF) and Stage-IV output image (GMM + MO + JF + B). Here again, we use the pre-trained ImageNet model for feature extraction and train a shallow FFNN to classify the images. Ablation studies are performed on the SPGC dataset, and the results are summarized in Table V.

As seen in Table V that GMM-filtered images with morphological operations and lung region boundary (GMM + MO + B) have consistent and uniform F1 scores across all the models pre-trained on the ImageNet dataset. GMM + MO images performed well with the ResNet-50 model, but the other models with the same image produced less classification accuracy. It is observed from the ablation study that blood vessel removal from the CT scan images degrades the classification performance. The blood vessel removal algorithm introduces an information loss in the CT scan images for the classification task. It is observed from the literature that blood vessel enlargement can happen in the later stages of COVID-19 infection [9], [10]. From the ablation study on the SPGC dataset, the (GMM + MO + B) image in pre-processing pipeline appears to be contributing most to accuracy. The F1 score is used as the evaluation metric for the ablation study since the test dataset has a class imbalance with the different categories of CT scans.

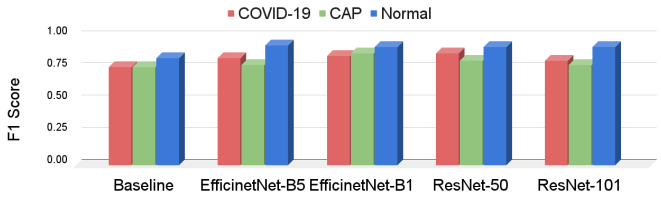

Fig. 6: Patient-wise classification results for the baseline and proposed models with SPGC dataset.

2) Slice-level classification

An experiment is conducted with the LDCT dataset, which provides slice-level labels for 160 patients. The (GMM + MO + B) images from the pre-processing pipeline are used to train and test the models. Four different models pre-trained on the ImageNet dataset are used to extract the features. Each feature is trained with different shallow FFNN models. It is observed from Table VI that all the features are suitable for discriminating between the classes. Since the CT scan is volumetric data, patient-level classification is preferable to slice-level identification.

3) Patient-level classification

The proposed three-class classifier model is evaluated with the SPGC test dataset and summarised the results along with the baseline model in Fig. 6. The baseline model [44] used the EfficientNet-B5 as the feature extractor from the raw CT scan images, and a shallow FFNN is trained on the extracted features. The proposed model with the EfficientNet-B5 as a feature extractor shows an average improvement of 6% F1 score for the classification task. Similarly, other features are also performing better than the baseline model. Fig. 6

TABLE VII: A comparative study of different features with different test datasets used in the experiments. C-19 and NL represent COVID-19 and Normal categories.

Feature Extractor	LDCT		RT-PCR		Mosmed	
	C-19	NL	C-19	NL	C-19	NL
ResNet-50	0.88	0.82	0.89	0.82	0.72	0.54
ResNet-101	0.89	0.82	0.90	0.83	0.77	0.58
EfficientNet-B1	0.89	0.82	0.94	0.89	0.79	0.59
EfficientNet-B5	0.88	0.82	0.89	0.82	0.75	0.59

TABLE VIII: Different severity classes and respective classification results.

Category	Prediction		Number of patients	Percentage of infection
	COVID-19	Normal		
CT-0	30	224	254	<= 5
CT-1	412	272	685	5 - 25
CT-2	110	15	125	25 - 50
CT-3	37	8	44	50 - 75

shows that the pre-processing pipeline introduces a significant improvement in the classification task.

The LDCT, LDCT-PCR, and Mosmed datasets consist of two classes of CT scans: COVID-19 and Normal. The proposed model achieves better results with both LDCT and LDCT-PCR datasets. However, the Mosmed dataset provides comparatively less accuracy than the LDCT dataset. The Mosmed dataset divided the patients into four categories: CT-0, CT-1, CT-2, and CT-3, based on the volume and severity of infection in the CT scans. The CT-0 category consists of COVID-19 cases having less than 5% of infections in CT scans. Since the ground truth was not given for each CT scan, CT-0 corresponds to the Normal class, and other categories CT-1, CT-2, and CT-3 to the COVID-19 class. A category-wise experiment is conducted with ResNet-101 generated features, and the classification results are given in Table VII. From Table VIII, it can be observed that patients in CT-0 are classified as Normal with less misclassification. 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 CT scan volume in the case of CT-0 and CT-1. The classifier is able to perform well for patients with greater than 25% infections developed in the lungs.

V. Conclusion

The present work improves upon the ability to identify COVID-19 infection and predict CTSS to assess the severity of infections developed in the lungs. The primary scope for improvisation is related to the scarcity of labeled images and the lack of models specifically trained on COVID-19 characteristics. An image pre-processing module is introduced to clearly mark the infected lung regions using the domain expertise of radiologists. The preprocessing module contains image processing algorithms along with an adaptive GMM filter. The pre-processing module is unsupervised and is capable of extracting the clinical features from the CT scans which are consistent with the boundaries of the infected region as identified by the domain experts. The pre-processed images are subsequently used to train the COVID-19-CAP-Normal classifier in a semi-supervised manner. The semi-supervised approach is adopted in conjunction with transfer learning to overcome the challenge of limited training data. As a result of the listed approach, specifically the pre-processing module, an absolute improvement of 6% is observed in the F1 score compared to the baseline model. Further, CTSS achieves a correlation of 82% with manual assessment by radiologists. The improvement in the identification of infected regions and evaluation of chest infection severity would serve as a reliable and quick diagnostic for the underlying COVID-19 infection, which would complement human effort and enhance the treatment outcome.

References

[1] M. Dramé *et al.*, "Should rt-pcr be considered a gold standard in the diagnosis of covid-19?," *Journal of medical virology*, 2020.

[2] E. Benmalek *et al.*, "Comparing ct scan and chest x-ray imaging for covid-19 diagnosis," *Biomedical Engineering Advances*, vol. 1, p. 100003, 2021.

[3] A. Nambu *et al.*, "Imaging of community-acquired pneumonia: roles of imaging examinations, imaging diagnosis of specific pathogens and discrimination from noninfectious diseases," *World Journal of Radiology*, vol. 6, no. 10, p. 779, 2014.

[4] C. Hani *et al.*, "Covid-19 pneumonia: a review of typical ct findings and differential diagnosis," *Diagnostic and interventional imaging*, vol. 101, no. 5, pp. 263–268, 2020.

[5] M. I. Mustafa *et al.*, "Cytokine storm in covid-19 patients, its impact on organs and potential treatment by qty code-designed detergent-free chemokine receptors," *Mediators of Inflammation*, vol. 2020, 2020.

[6] 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, 2020.

[7] M. Chung, A. Bernheim, X. Mei, N. Zhang, M. Huang, X. Zeng, J. Cui, W. Xu, Y. Yang, Z. A. Fayad, *et al.*, "Ct imaging features of 2019 novel coronavirus (2019-ncov)," *Radiology*, vol. 295, no. 1, pp. 202–207, 2020.

[8] L. Lin *et al.*, "Ct manifestations of coronavirus disease (covid-19) pneumonia and influenza virus pneumonia: A comparative study," *American Journal of Roentgenology*, vol. 216, no. 1, pp. 71–79, 2021.

[9] 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," *European radiology*, vol. 31, no. 6, pp. 4031–4041, 2021.

[10] R. Aoki *et al.*, "Pulmonary vascular enlargement and lesion extent on computed tomography are correlated with covid-19 disease severity," *Japanese journal of radiology*, vol. 39, no. 5, pp. 451–458, 2021.

[11] S. Heidarian *et al.*, "SPGC-COVID Dataset." {https://figshare.com/articles/dataset/SPGC-COVID_Dataset/166-32397/1},DOI={10.6084/m9.figshare.16632397.v1}. 2021.

[12] P. Afshar *et al.*, "COVID-19 Low-Dose and Ultra-Low-Dose CT Scans." https://dx.doi.org/10.21227/sed8-6r15. IEEE Dataport, 2021.

[13] S. P. Morozov *et al.*, "Mosmeddata: data set of 1110 chest ct scans performed during the covid-19 epidemic," *Digital Diagnostics*, vol. 1, no. 1, pp. 49–59, 2020.

[14] S. T. MedSeg, Jenssen HB, "MedSeg Covid Dataset." https://doi.org/10.6084/m9.figshare.13521488.v2. figshare Dataset, January 2021.

[15] M. Jun *et al.*, "COVID-19 CT Lung and Infection Segmentation Dataset." {https://doi.org/10.5281/zenodo.3757476}, DOI={10.5281/zenodo.3757476}. April, 2020.

[16] 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 and Video Processing*, pp. 1–8, 2022.

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

[18] J. N. Kapur, P. K. Sahoo, and A. K. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Computer vision, graphics, and image processing*, vol. 29, no. 3, pp. 273–285, 1985.

[19] N. Otsu, "A threshold selection method from gray-level histograms,"

IEEE transactions on systems, man, and cybernetics, vol. 9, no. 1, pp. 62–66, 1979.

[20] A. Trémeau and P. Colantoni, "Regions adjacency graph applied to color image segmentation," *IEEE Transactions on image processing*, vol. 9, no. 4, pp. 735–744, 2000.

[21] M. S. Brown *et al.*, "Method for segmenting chest ct image data using an anatomical model: preliminary results," *IEEE transactions on medical imaging*, vol. 16, no. 6, pp. 828–839, 1997.

[22] R. Adams and L. Bischof, "Seeded region growing," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 16, no. 6, pp. 641–647, 1994.

[23] J. Dehmeshki, H. Amin, M. Valdivieso, and X. Ye, "Segmentation of pulmonary nodules in thoracic ct scans: A region growing approach," *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 467–480, 2008.

[24] A. P. Mangan and R. T. Whitaker, "Partitioning 3d surface meshes using watershed segmentation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 5, no. 4, pp. 308–321, 1999.

[25] Y. Boykov and M.-P. Jolly, "Interactive organ segmentation using graph cuts," in *International conference on medical image computing and computer-assisted intervention*, pp. 276–286, Springer, 2000.

[26] 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 Transactions on Image Processing*, vol. 24, no. 12, pp. 5315–5329, 2015.

[27] J. K. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation," *Graphical models and image processing*, vol. 58, no. 3, pp. 246–261, 1996.

[28] C. Xu and J. L. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Transactions on image processing*, vol. 7, no. 3, pp. 359–369, 1998.

[29] S. C. Zhu and A. Yuille, "Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 18, no. 9, pp. 884–900, 1996.

[30] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *International journal of computer vision*, vol. 50, no. 3, pp. 271–293, 2002.

[31] J. S. Schildkraut, S. Chen, M. Heath, W. G. O'Dell, P. Okunieff, M. Schell, and N. Paul, "Level-set segmentation of pulmonary nodules in radiographs using a ct prior," in *Medical Imaging 2009: Image Processing*, vol. 7259, pp. 1103–1116, SPIE, 2009.

[32] S. Hu *et al.*, "Automatic lung segmentation for accurate quantitation of volumetric x-ray ct images," *IEEE transactions on medical imaging*, vol. 20, no. 6, pp. 490–498, 2001.

[33] Y. Zhu *et al.*, "Automatic segmentation of ground-glass opacities in lung ct images by using markov random field-based algorithms," *Journal of digital imaging*, vol. 25, no. 3, pp. 409–422, 2012.

[34] B. Ye *et al.*, "Severity assessment of covid-19 based on feature extraction and v-descriptors," *IEEE Transactions on Industrial Informatics*, 2021.

[35] X. Wang *et al.*, "A weakly-supervised framework for covid-19 classification and lesion localization from chest ct," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2615–2625, 2020.

[36] X. Wu *et al.*, "Covid-al: The diagnosis of covid-19 with deep active learning," *Medical Image Analysis*, vol. 68, p. 101913, 2021.

[37] S. Xue and C. Abhayaratne, "Covid-19 diagnostic using 3d deep transfer learning for classification of volumetric computerised tomography chest scans," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8573–8577, 2021.

[38] J. Wang *et al.*, "Prior-attention residual learning for more discriminative covid-19 screening in ct images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2572–2583, 2020.

[39] T. Javaheri *et al.*, "Covidctnet: an open-source deep learning approach to diagnose covid-19 using small cohort of ct images," *NPJ digital medicine*, vol. 4, no. 1, pp. 1–10, 2021.

[40] Z. Han *et al.*, "Accurate screening of covid-19 using attention-based deep 3d multiple instance learning," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2584–2594, 2020.

[41] T. Anwar, "Covid19 diagnosis using automl from 3d ct scans," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 503–507, 2021.

[42] M. N. Teli, "Telinet: Classifying ct scan images for covid-19 diagnosis," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 496–502, 2021.

[43] P. Afshar *et al.*, "Covid-ct-md, covid-19 computed tomography scan dataset applicable in machine learning and deep learning," *Scientific Data*, vol. 8, no. 1, pp. 1–8, 2021.

[44] S. Chaudhary *et al.*, "Detecting covid-19 and community acquired pneumonia using chest ct scan images with deep learning," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8583–8587, IEEE, 2021.

[45] D. Kollias *et al.*, "Mia-cov19d: Covid-19 detection through 3-d chest ct image analysis," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 537–544, 2021.

[46] P. Garg *et al.*, "Multi-scale residual network for covid-19 diagnosis using ct-scans," in *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8558–8562, IEEE, 2021.

[47] T. D. Pham, "A comprehensive study on classification of covid-19 on computed tomography with pretrained convolutional neural networks," *Scientific reports*, vol. 10, no. 1, pp. 1–8, 2020.

[48] M. Singh *et al.*, "Transfer learning–based ensemble support vector machine model for automated covid-19 detection using lung computerized tomography scan data," *Medical & biological engineering & computing*, vol. 59, no. 4, pp. 825–839, 2021.

[49] J. Deng *et al.*, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, pp. 248–255, Ieee, 2009.

[50] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, pp. 6105–6114, PMLR, 2019.

[51] J. Hofmanninger *et al.*, "Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem," *European Radiology Experimental*, vol. 4, no. 1, pp. 1–13, 2020.

[52] T. Jerman *et al.*, "Enhancement of vascular structures in 3d and 2d angiographic images," *IEEE transactions on medical imaging*, vol. 35, no. 9, pp. 2107–2118, 2016.

[53] P. Soille, *Morphological image analysis: principles and applications*. Springer Science & Business Media, 2013.

[54] L. Vincent, "Morphological area openings and closings for grey-scale images," in *Shape in Picture*, pp. 197–208, Springer, 1994.

[55] J. M. Prewitt *et al.*, "Object enhancement and extraction," *Picture processing and Psychopictorics*, vol. 10, no. 1, pp. 15–19, 1970.

[56] A. G. Howard *et al.*, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[57] K. He *et al.*, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[58] D.-P. Fan *et al.*, "Inf-net: Automatic covid-19 lung infection segmentation from ct images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2626–2637, 2020.

[59] O. Ronneberger *et al.*, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.

[60] O. Oktay *et al.*, "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.

[61] J. Schlemper *et al.*, "Attention gated networks: Learning to leverage salient regions in medical images," *Medical image analysis*, vol. 53, pp. 197–207, 2019.

[62] X. Li *et al.*, "H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes," *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.

[63] Z. Zhou *et al.*, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pp. 3–11, Springer, 2018.

[64] J. Ma *et al.*, "Toward data-efficient learning: A benchmark for covid-19 ct lung and infection segmentation," *Medical physics*, vol. 48, no. 3, pp. 1197–1210, 2021.

[65] K. J. Kiser *et al.*, "Plethora: Pleural effusion and thoracic cavity segmentations in diseased lungs for benchmarking chest ct processing pipelines," *Medical physics*, vol. 47, no. 11, pp. 5941–5952, 2020.

[66] "Covid-19 ct segmentation dataset. accessed: Apr. 11, 2020. [online]. available: https://medicalsegmentation.com/covid19/," 2020.