COVID-19 identification from volumetric chest CT scans using a progressively resized 3D-CNN incorporating segmentation, augmentation, and class-rebalancing

Md. Kamrul Hasan[a,1,*], Md. Tasnim Jawad[a], Kazi Nasim Imtiaz Hasan[b], Sajal Basak Partha[b], Md. Masum Al Masba[b], Shumit Saha[c,d]

[a]*Department of Electrical and Electronic Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh*
[b]*Department of Computer Science and Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh*
[c]*University Health Network, Toronto, ON, Canada*
[d]*Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada*

Abstract

The novel COVID-19 is a global pandemic disease overgrowing worldwide. Computer-aided screening tools with greater sensitivity is imperative for disease diagnosis and prognosis as early as possible. It also can be a helpful tool in triage for testing and clinical supervision of COVID-19 patients. However, designing such an automated tool from non-invasive radiographic images is challenging as many manually annotated datasets are not publicly available yet, which is the essential core requirement of supervised learning schemes. This article proposes a 3D Convolutional Neural Network (CNN)-based classification approach considering both the inter- and intra-slice spatial voxel information. The proposed system is trained in an end-to-end manner on the 3D patches from the whole volumetric CT images to enlarge the number of training samples, performing the ablation studies on patch size determination. We integrate progressive resizing, segmentation, augmentations, and class-rebalancing to our 3D network. The segmentation is a critical prerequisite step for COVID-19 diagnosis enabling the classifier to learn prominent lung features while excluding the outer lung regions of the CT scans. We evaluate all the extensive experiments on a publicly available dataset, named MosMed, having binary- and multi-class chest CT image partitions. Our experimental results are very encouraging, yielding areas under the ROC curve of 0.914 ± 0.049 and 0.893 ± 0.035 for the binary- and multi-class tasks, respectively,

1

applying 5-fold cross-validations. Our method's promising results delegate it as a favorable aiding tool for clinical practitioners and radiologists to assess COVID-19.

Keywords: COVID-19, 3D convolutional neural network, Volumetric chest CT scans, 3D patches, Progressive resizing.

1. Introduction

Pneumonia of unknown cause discovered in Wuhan, China, was published to the World Health Organization (WHO) office in China on 31st December 2019. It was consequently assigned to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) because of having similar genetic properties to the SARS outbreak of 2003. On 11th February 2020, WHO termed that new disease as COVID-19 (Coronavirus disease), which displays an upper respiratory tract and lung infection [106]. The clinical characteristics of critical COVID-19 pandemic are bronchopneumonia that affects cough, fever, dyspnea, and detailed respiratory anxiety ailment [14, 60, 100]. According to the WHO reports, COVID-19's general indications are equivalent to that of ordinary flu, including fever, tiredness, dry cough, shortness of breath, aches, pains, and sore throat [48]. Those shared signs turn it challenging to recognize the virus at an ancient step. The aforementioned is a virus, which works on bacterial or fungal infections [48, 108] with no possibility that antibiotics can restrict it. Besides, people suffering from medical complications, like diabetes, chronic respiratory and cardiovascular diseases, are prone to undergo COVID-19. An explanatory statement of the Imperial College advised that the affection rate will be more than 90.0% of the world's people, killing 40.6 million people if no reduction actions are grasped to combat the virus [99].

---

*I am corresponding author

Email addresses: m.k.hasan@eee.kuet.ac.bd (Md. Kamrul Hasan), jawad1703006@stud.kuet.ac.bd (Md. Tasnim Jawad), hasan1607097@stud.kuet.ac.bd (Kazi Nasim Imtiaz Hasan), partha1607101@stud.kuet.ac.bd (Sajal Basak Partha), masba@cse.kuet.ac.bd (Md. Masum Al Masba), shumit.saha@mail.utoronto.ca (Shumit Saha)

[1] Department of EEE, KUET, Khulna-9203, Bangladesh.

Advanced presumed discovery of COVID-19 is also a challenge for public health security and control of pandemic. The COVID-19 detection failure increases the mortality rate exponentially. The incubation period, which is a time between catching the virus and causing to have indications of the illness, is $1 \sim 14$ days, making it remarkably challenging to identify COVID-19 infection at a preliminary stage of an individual's symptoms [48]. The clinical screening test for the COVID-19 is Reverse Transcription Polymerase Chain Reaction (RT-PCR), practicing respiratory exemplars. However, it is a manual, complicated, tiresome, and time-consuming fashion with an estimated true-positive rate of 63.0% [103]. There is also a significant lack of RT-PCR kit inventory, leading to a delay in preventing and curing coronavirus disease [112]. Furthermore, the RT-PCR kit is estimated to cost around $120 \sim 130$ USD. It also requires a specially designed biosafety laboratory to house the PCR unit, each of which can cost $15,000 \sim 90,000$ USD [1]. Nevertheless, the utilization of a costly screening device with a delayed test results makes it more challenging to restrict the disease's spread. Inadequate availability of screening workstations and measurement kits constitute an enormous hardship to identify COVID-19 in this pandemic circumstance. In such a situation, speedy and trustworthy presumed COVID-19 cases are an enormous difficulty for related personals.

However, it is observed that most of the COVID-19 incidents have typical properties on radiographic CT and X-ray images, including bilateral, multi-focal, ground-glass opacities with a peripheral or posterior distribution, chiefly in the lower lobes and early- and late-stage pulmonary concentration [18, 42, 88, 110]. Those features can be utilized to build a sensitive Computer-aided Diagnosis (CAD) tool to identify COVID-19 pneumonia, which is deemed an automated screening tool [59]. Currently, deep Convolutional Neural Networks (CNNs) allow for building an end-to-end model without requiring manual and time-consuming feature extraction and engineering [57, 58], demonstrating tremendous success in many domains of medical imaging, such as arrhythmia detection [4, 28, 113], skin lesion segmentation and classification [17, 23, 24, 35], breast cancer detection [13, 19, 31], brain disease segmentation and classification [93, 97], pneumonia detection from chest X-ray images [79], fundus image segmentation [34, 94], and lung segmentation [26]. Most recently, various deep CNN-based

methods have been published for identifying COVID-19 from X-rays and CT images, summarizing and bestowing in Table 1. Though the results obtained in the current articles are promising, they exhibit limited scope for use as a CAD tool, as most of the works, especially on x-ray images, have been based on data coming from different sources for two distinct classes (Covid Vs. Normal) [8, 38, 53, 67, 74, 84, 88, 102]. This brings inherent bias on the algorithms as the model tends to learn the distribution of the data source for binary classification problem [32]. Therefore, these models perform very low when used in practical settings, where the models have to adapt to data from different domains [32]. Recently, Morozov et al. [66] launched a public chest volumetric CT scan dataset with 1110 COVID-19 related studies (see details in subsection 2.1). However, the published articles [65, 114] on this dataset consider only intra-slice spatial voxel information to isolate COVID-19 and regular healthy patients.

This article aims to evaluate the proposed 3D-CNN classifier's performance for identifying COVID-19 utilizing volumetric chest images, where the volumes have come from the same source (details in subsection 2.1). However, the core contributions in this article are enlisted as follows:

- Designing a 3D-CNN-based classification network for volumetric CT images as the 3D networks account for the inter- and intra-slice spatial voxel information while the 2D networks consider only the intra-slice spatial voxel information [37, 44, 52, 89, 114, 118]

- Conducting 3D patch-based classification as it increases the sample numbers in the smaller datasets, where we perform ablation studies to determine a proper patch size

- Progressively increasing the input patch size of our network up to the original CT size of $R \times C \times S$, where the trained network with the patch size of $(R/2^{n+1}) \times (C/2^{n+1}) \times (S/2^{n+1})$ is a pre-trained model of a network with the patch size of $(R/2^n) \times (C/2^n) \times (S/2^n)$

- Developing an unsupervised lung segmentation pipeline for allowing the classifier to learn salient lung features while omitting the outer lung areas of the CT scans

Table 1: Numerous published articles for the COVID-19 identification with their respective utilized datasets and performances exhibiting different metrics such as mSn, mSp, and mF1 respectively for mean sensitivity, specificity, and F1-score. The mixed datasets indicate that data have come from different open-sources.

Different methods	Datasets	Results
A pre-trained 2D MobileNet-v2 [82] architecture on ImageNet [20] was used to extract massive high-dimensional features to classify six different diseases using the fully-connected layers [7]	Mixed	mSn: 0.974 mSp: 0.994
DeTraC [2, 3], where the network was trained first using a gradient descent optimization [81], and then, the class-composition layer of DeTraC was used to refine the final detection results [2]	Mixed	mSn: 0.979 mSp: 0.919
A multi-objective differential evolution–based CNN method fine-tuning iteratively using mutation, crossover, and selection operations to discover the best possible results [88]	Mixed	mSn: 0.907 mSp: 0.906
An ensemble of VGG-16 [87], Inception [92], Xception [16], Inception-ResNet [91], MobileNet [41], DenseNet [43], and NasNet [77] optimizing the hyperparameters using a greedy search algorithm [10, 78]	Mixed	mSn: 0.990 mSp: 0.990
Support vector machine [31, 107]-based method to classify the in-depth features from the pre-trained MobileNet and SqueezeNet [46] from the restructured the data using a fuzzy color technique [96]	Mixed	mSn: 0.983 mSp: 0.997
An ensemble of three lightweight pre-trained SqueezeNet, ShuffleNet [116], and EfficientNet-B0 [95] at various depths and consolidates feature maps in diverse abstraction levels [70]	Mixed	mSn: 0.978 mSp: 0.985
A fusing and ranking of in-depth features for classifying using a support vector machine, where the pre-trained CNN models on ImageNet were used to extract the COVID-19 features [72]	Mixed	mSn: 0.989 mSp: 0.976
A DenseNet-201 [43]-based transfer learning to extract features using its learned weights on the ImageNet was used to classify the patients as COVID infected or not [49]	SARS-COV-2 [6]	mSn: 0.960 mSp: 0.960
A transfer learning-based approach using one of the VGG, ResNet [36], Inception, or Xception pre-trained deep learning model on ImageNet as a backbone [55]	Zhao et al. [117]	mSn: 0.996 mSp: 0.100
A weakly-supervised learning schema, where the lung region was segmented using a pre-trained UNet [80]; then, a 3D network was used to predict the probability of COVID-19 infectious [104]	Wang et al. [104]	mSn: 0.911 mSp: 0.881
A multi-scale-multi-encoder ensemble of CNN model aggregating the outputs from two different encoders and their different scales to obtain the final prediction probability [30]	Mixed	mSn: 0.997 mSp: 0.997
Advanced deep network architectures proposing a transfer learning strategy on ImageNet using a custom-sized input tailored for each architecture to achieve the best possible results [5]	Mixed	mSn: 0.996 mSp: 0.998
A pre-trained CNN-based schema leveraging the strength of multiple texture descriptors and base classifiers at once, where data was re-balanced using resampling algorithms [76]	Mixed	mSn: − mF1: 0.889
A deep ResNet-based transfer learning technique with a top-2 smooth loss function and a cost-sensitive attribute to handle noisy and imbalanced COVID-19 datasets [75]	Mixed	mSn: 0.915 mSp: 0.948
An auxiliary classifier generative adversarial network-based design to generate synthetic images, where the synthetic images produced CNN's enhanced results for the prediction [98]	Mixed	mSn: 0.900 mSp: 0.970
A framework consisting of a CNN-based feature extractor and k-nearest neighbor [29, 33], support vector machine, and decision tree [33]-based classifiers using the Bayesian algorithm [69]	Mixed	mSn: 0.894 mSp: 0.998
An architecture based on the deep residual neural network using two parallel levels with different kernel sizes for capturing both local and global features of the inputs images [71]	Mixed	mSn: − mF1: 0.967
A classification architecture combining ResNet and Xception to investigate the challenges and limitations of deep CNN and different datasets for building generic COVID-19 classifiers [32]	Mixed	mSn: 0.976 mSp: −
An average rank pooling, multiple-way augmentation, and deep feature fusion-based CNN and graph CNN was developed to fuse individual image-level features and relation-aware features [102]	Wang et al. [102]	mSn: 0.963 mSp: 0.970
An end-to-end DarkCovidNet architecture [74] based on DarkNet [74] gradually increasing the number of filters, where each convolutional layers were followed by BatchNorm [47] and LeakyReLU [109]	Mixed	mSn: 0.951 mSp: 0.953
A CoroNet model based on pre-trained Xception architecture on ImageNet for automated detection of COVID-19 infection and trained in end-to-end manners [53]	Mixed	mSn: 0.993 mSp: 0.986
Comparative analyses of different pre-trained models considering several important factors such as batch size, learning rate, epoch numbers, and type of optimizers to find the best-suited model [68]	Mixed	mSn: 0.100 mSp: 0.967
A comparative analysis of different CNN models, such as VGG, Resnet, Inception, Xception, Inception-ResNet, DenseNet, and NASNet-Large [119], to decide a proper one for multi-modal image classification minimizing the image quality imbalances in the image samples as a preprocessing [40]	Mixed	mSn: 0.820 mSp: − mF1: 0.820
A pipeline consisting of segmentation and subsequent classification employing both 3D and 2D CNNs, where the promising results for detecting were obtained in the 3D-CNNs than the 2D CNNs [37]	He et al. [37]	mSn: 0.891 mSp: 0.911

- Class rebalancing and augmentations, such as intensity- and geometry-based, are employed to develop a general network, although a small dataset is being utilized

The remainder of the article is prepared as follows. Section 2 details the materials and methods practiced in the study, including a brief introduction to the methodology and end-to-end 3D-CNN training. Section 3 describes the experimental operations and their corresponding obtained results. Lastly, section 4 concludes the article.

2. Materials and Methods

In this section, we describe the utilized materials and methods to conduct the widespread experiments. We summarize the adopted dataset in the first subsection 2.1. The essential integral preprocessing, such as segmentation, augmentation, and class-rebalancing, are reported in the second subsection 2.2. The design of the proposed 3D-CNN-based COVID-19 classifier, along with its training protocol, is explained in the third subsection 2.3. Finally, in the fourth subsection 2.4, we represent used hardware to execute the aimed method and evaluation criterion.

2.1. Dataset

This article's experimentations utilize a publicly usable MosMedData dataset administered by municipal hospitals in Moscow, Russia, from March to April 2020 [66]. This dataset includes anonymized human chest lung CT scans with and without COVID-19 related findings of 1110 studies. The population of MosMedData is distributed as 42% male, 56% female, and 2% others, where the median age of the subjects is 47 years ($18 \sim 97$ years). All the studies ($n = 1110$) are distributed into five following categories, as presented in Table 2. We design two experimental protocols using the MosMedData dataset, such as binary- and multi-class identification, to evaluate our proposed workflow. In binary-class evaluation, we use NOR vs. NCP (Novel COVID-19 Positive), where NCP includes MiNCP-, MoNCP-, SeNCP-, and CrNCP-classes, while in multi-class evaluation, we use NOR vs. MiNCP vs. MoNCP vs. SeNCP. In multi-class protocols, we merge SeNCP- and CrNCP-classes, naming

Table 2: Distribution of utilized MosMedData dataset for COVID-19 identification with a short class description.

Class acronym	Description	PPI*	Samples (%)
NOR	Not consistent with pneumonia, including COVID-19, and refer to a specialist	–	254 (22.8 %)
MiNCP	Mild novel COVID-19 positive with ground-glass opacities and follow-up at home using mandatory telemonitoring	$=< 25\%$	684 (61.6 %)
MoNCP	Moderate novel COVID-19 positive with ground-glass opacities and follow-up at home by a primary care physician	$25-50\%$	125 (11.3 %)
SeNCP	Severe novel COVID-19 positive with ground-glass opacities and immediate admission to a COVID specialized hospital	$50-75\%$	45 (4.1 %)
CrNCP	Critical novel COVID-19 positive with diffuse ground-glass opacities and emergency medical care	$>= 75\%$	2 (0.2 %)
Total Samples (%)			1110 (100 %)

PPI*: Pulmonary parenchymal involvement

them as SeNCP, as CrNCP has only two samples in the MosMedData dataset. We have applied a cross-validation technique to choose training, validation, and testing images as those are not explicitly given by the data provider. The class-wise distribution of MosMedData dataset in Table 2 illustrates that the class distribution is imbalanced. Such an imbalanced class distribution produces a biased image classifier towards the class having more training samples. We apply various rebalancing schemes to develop a generic classifier for COVID-19 identification, even though the dataset is imbalanced.

2.2. Preprocessing

The recommended integral preprocessing consists of segmentation, augmentations (both geometry- and intensity-based), and class-rebalancing, which are concisely explained as follows:

Segmentation. The segmentation, to separate an image into regions with similar properties such as gray level, color, texture, brightness, and contrast, is the significant element for automated detection pipeline [39]. It is also a fundamental prerequisite for the COVID-

19 identification as it extracts the lung region and delivers explanatory information about the shapes, structures, and textures. However, this article proposes an unsupervised Lung Segmentation (LS) technique applying different image processing algorithms, as a massive number of annotated COVID-19 images are not available yet in this pandemic situation. Fig. 1 depicts the pipeline of the proposed LS method. The proposed threshold-based LS's
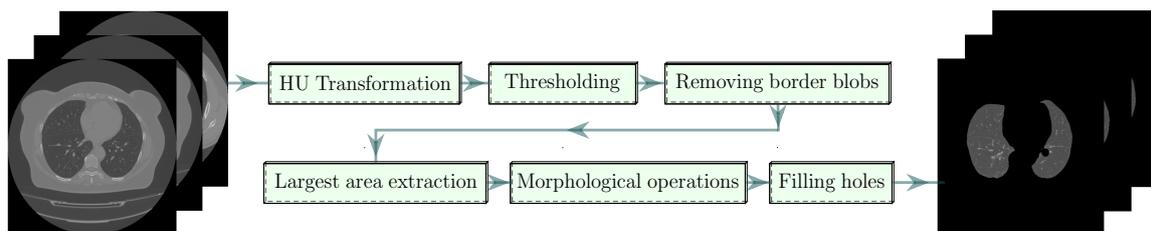

Figure 1: The proposed block diagram of an unsupervised lung segmentation pipeline, without requiring a manually annotated lung region.

primary step is transforming all the CT volumes to Hounsfield units (HU), as it is a quantitative measure of radiodensity for CT scans. We set the HU unit as -1000 to -400 as the study shows that lung regions are within that range, which was also used in many articles [56, 85, 101]. The thresholded binary lung masks are then refined to exclude different false-positive regions, such as the connected blobs with the image border and other small false-positive areas, and false-negative regions, such as small holes in the lung regions. Firstly, the border connected regions are eradicated. Secondly, the two largest areas are picked using the region properties algorithm. Thirdly, morphological erosion to separate the lung nodules attached to the blood vessels and morphological closing to keep nodules attached to the lung wall. Finally, the false-negative regions are removed using binary hole fill algorithms. Such an unsupervised thresholding-based segmentation method is better in terms of efficiency, taking only a few seconds, and yields utterly reproducible LS.

Augmentation. The CNN-based classifiers are profoundly dependent on large data samples to evade the overfitting. Lamentably, various medical imaging fields, especially the current COVID-19 pandemic, suffer from an inadequate dataset size as manually annotated

8

massive training samples are still not available. In such a scenario, the augmentations are very dormant preprocessing for increasing the training samples as they are incredibly discriminative [45]. Data augmentation incorporates a method that magnifies training datasets' size and property to develop a better-CNN classifier [86]. The geometric-based augmentation, including a rotation (around $row/2$ and $col/2$) of $-25°$, $-15°$, $10°$, $30°$ and height & width shifting by 20%, the intensity-based augmentation, including gamma correction & adding Gaussian random noise, and Elastic deformation[2] are applied in this article as a part of the recommended preprocessing. Two values of gamma (γ), such as 0.7 and 1.7, have used in gamma correction to adjust the luminance of the CT volumes by $V_{out} = V_{in}^{\gamma}$, where V_{out} and V_{in} individually denote the output and input values of the luminance.

Rebalancing. The utilized dataset in Table 2 is imbalanced. This situation is pretty obvious in the medical diagnosis field due to the scarcity of massive manually annotated training samples, especially in COVID-19 datasets. The undesired class-biasing occurs in the supervised learning systems towards the class with majority samples. However, we apply two techniques to rebalance the imbalanced class distribution, such as adding extra CT volumes from the publicly available CC-CCII dataset [115] and weighting the loss function for penalizing the overrepresented class. The latter approach rewards more extra consideration to the class with minority samples. Here, we estimate the class weight using a portion of $W_n = N_n/N$, where W_n, N, and N_n separately denote the nth-class weight, the total sample numbers, and the samples in n^{th}-class. We employ both the class-rebalancing strategies in the binary-class protocol, whereas the only class weighting method is adopted in the multi-class protocol.

2.3. Methodologies

2.3.1. Architecture

The deep neural network is a machine learning framework with a wide range of applications, from natural language processing [21] to medical image classification [12], segmentation [12], and registration [25]. In special, CNNs have become a prevalent technique in the

---

[2]`https://pypi.org/project/elasticdeform/`

computer vision community. They are practiced in diverse tasks, including object detection [50], classification [22], and localization [63]. The CNN-based deep neural systems are also popularly adopted in recent pandemic for COVID-19 identification [11, 73] (see in Table 1). CNN is an excellent discriminant feature extractor at various abstraction levels, which is translation-invariant. Consequently, utilizing it to classify medical images evades complicated and expensive feature engineering [83]. The early few CNN layers learn low-level image features and later layers learn high-level image features particular to the application types [52]. However, the 2D-CNNs are frequently employed in natural RGB and grayscale images to extract the spatial features only in two dimensions [92]. The 2D-CNN also can be applied to the volumetric medical image datasets taking cross-sectional 2D slices of the CT, MRI, or similar scans. However, the recent experimental results have revealed the advantages of 3D-CNN over 2D-CNN, where the 3D-CNN accepts the volumetric spatial information as an input [62]. Conventional 2D-CNNs' effectiveness is degraded due to loss of spatial voxel information for volumetric 3D medical imaging tasks. A 3D-CNN, a 3D space implementation of convolution and pooling operation, is practiced to overcome spatial voxel information loss as in the 2D-CNNs. The image becomes scalable in the spatial direction using a 3D-CNN, allowing accurate image detection with different frame sizes [64]. Therefore, we propose a classifier based on 3D-CNN to identify COVID-19 from the volumetric CT scans.

Fig. 2 represents the constructional structure of our proposed COVID-19 base classifier. The proposed base network in Fig. 2 essentially consists of two modules, such as feature extractor and feature classifier. The former module is a stack of convolutional, pooling, and batch normalization layers, whereas the latter module is a stack of fully-connected layers followed by a softmax layer. We involve 3D layers for all the feature extractor module components to operate on volumetric medical images for extracting the most discriminating features, accounting for both the intra- and inter-slice spatial voxel information. In our network, each 3D convolutional layer with ReLU activation is followed by a 3D max-pooling layer, where the pooling layer increases translational invariances of the network. The pooled feature maps are then used as an input to the successive layers, which may dynamically change during training at each training epoch [89]. The more enormous changes prone
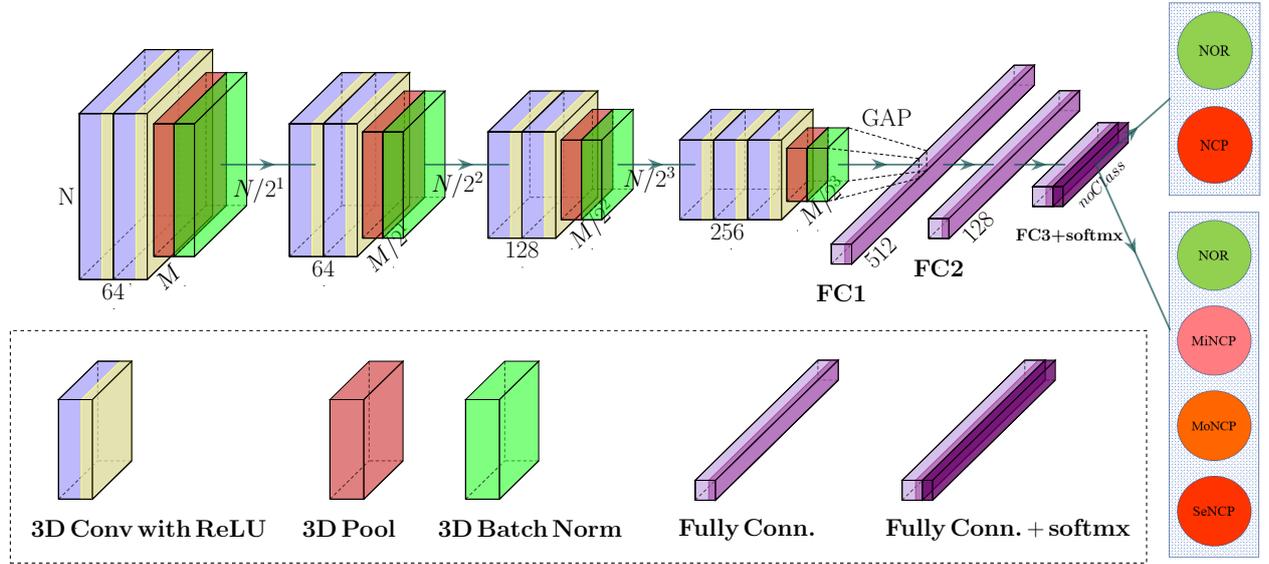
Figure 2: The architectural construction of the proposed base network, training with the most smaller 3D patches. This trained base network is applied as a pre-trained model for the next bigger patches. Best view in the color figure.

to bring difficulties for searching an optimal parameter or hyperparameter; often become computationally expensive to reach an optimal value [47]. Such a problem is mitigated by integrating batch normalization layers in our network [47]. It also facilitates the smooth training of the network architectures in less time [89]. The Global Average Pooling (GAP) [61] is used as a bridge layer between the feature extractor and feature classifier modules, converting the feature tensor into a single long continuous linear vector. In GAP, only one feature map is produced for each corresponding category, achieving a more extreme dimensionality compression to evade overfitting [61]. A dropout layer [90] is also employed as a regulariser, which randomly sets half of the activation of the fully-connected layers to zero through the training of our network.

Again, as mentioned earlier, the CNNs are heavily reliant on the massive dataset to bypass overfitting and build a generic network. The acquisition of annotated medical images is arduous to accumulate, as the medical data collection and labeling are confronted with data privacy, requiring time-consuming expert explanations [111]. There are two general

resolving directions: accumulating more data, such as crowdsourcing [51] or digging into the present clinical reports [105]. Another technique is investigating how to enhance the achievement of the CNNs with small datasets, which is exceptionally significant because the understanding achieved from the research can migrate the data insufficiency in the medical imaging fields [111]. Transfer learning is a widely adopted method for advancing the performance of CNNs with inadequate datasets [15]. To our most trustworthy knowledge, there is no public pre-trained 3D-CNN model for the COVID-19 identification from the volumetric chest images with limited samples. Therefore, we create a pre-trained model by training our base model (see in Fig. 2) on the extracted 3D patches from whole chest CT scans (see details in subsection 2.3.2). Then, we double the patches' size and use them for training the modified base network, where we also double the base model's input size applying a stack of convolutional, pooling, and batch normalization layers (see details in Fig. 3). At the same time, we keep the base model's trained weights for the smaller patches. We repeat to enlarge (n^{th}-times) the patch and network sizes until we arrive at the provided CT scans' size, as pictured in Fig. 3. Such training is called progressive resizing [9], where the training begins with smaller image sizes followed by a progressive expansion of size. This training process is continued until the last patch and network sizes are as same as the initial image dimension.

2.3.2. Training protocol

We first extract five different patches with different sizes (see in Fig. 4) to begin the experimentations. We perform ablation studies in subsection 3.1 looking for the best patch size. The weights of the base network in Fig. 2 is initialized with Xavier normal distribution. The weights of the first progressively resized network are initialized with the weights of the base network. In general, the weights of the network with the patch size of $(R/2^n) \times (C/2^n) \times (S/2^n)$ are initialized with the weights of the network with the patch size of $(R/2^{n+1}) \times (C/2^{n+1}) \times (S/2^{n+1})$ for the original CT volume size of $R \times C \times S$.

Categorical cross-entropy and accuracy are utilized as a loss function and metric, respectively, for training all the networks in this article. We use Adam [54] optimizer with
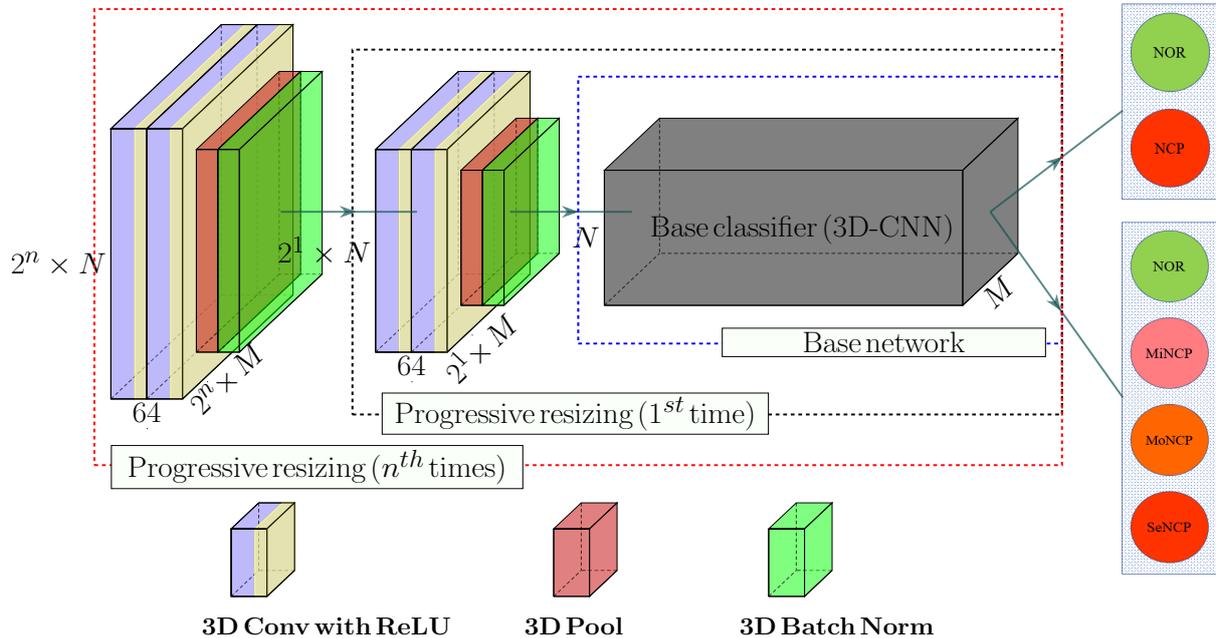
Figure 3: The proposed progressively resized network's architectural structure, where the base model (see in Fig. 2) is trained with the smaller 3D patches and sequentially doubles the base network's size from smaller to larger sizes. The network trained with the smaller patches is the pre-trained model for the next bigger patches. Best view in the color figure.

initial learning rate (LR), exponential decay rates (β_1, β_2) as $LR = 0.0001$, $\beta_1 = 0.9$, and $\beta_2 = 0.999$, respectively, without AMSGrad variant. The exponential decaying LR schedule is also employed for the networks' optimization. Initial epochs are set as 200, and training is terminated if validation performance stops growing after 15 epochs.

2.4. Hardware and evaluation criterion

We execute all the comprehensive experiments on a *Windows*-10 machine utilizing the Python, with various Keras [27] and image processing APIs, and MATLAB programming languages. The device configurations of the used machine are: Intel® Core™ i7-7700 HQ CPU @ $3.60\,GHz$ processor with a install memory (RAM) of $32.0\,GB$, and GeForce GTX 1080 GPU with a memory of $8.0\,GB$ (GDDR5).

We evaluate all the experimental outcomes by employing numerous metrics, such as recall, precision, and F1-score, for evaluating them from diverse perspectives. The recall mea-

sures the type-II error (the patient having positive COVID-19 characteristics, erroneously abandons to be repealed), whereas the precision estimates the positive predictive values (a portion of absolutely positive-identification amid all the positive-identification). The harmonic mean of recall and precision is manifested using the F1-score, conferring the tradeoff between these two metrics. Furthermore, we also quantify the prognostication probability of an anonymously selected CT sample using a Receiver Operating Characteristics (ROC) with its Area Under the ROC Curve (AUC) value.

3. Results and Discussion

In this section, the achieved results from different experiments are reported with comprehensive discussion. In subsection 3.1, we confer the results of COVID-19 identification utilizing various 3D patches and compare them with original CT image utilization on the same experimental conditions and network. We discuss the results of progressive resizing over a single fixed size in subsection 3.2. We demonstrate the effects of different proposed preprocessing on COVID-19 identification in subsection 3.3. Finally, in subsection 3.4, we show the results for binary- and multi-class COVID-19 identification applying our proposed network and preprocessing.

3.1. Patch Selection

We extract five different 3D patches, named P_1, P_2, P_3, P_4, and P_5, having respective size of $16 \times 16 \times 9$, $32 \times 32 \times 12$, $64 \times 64 \times 15$, $128 \times 128 \times 20$, and $256 \times 256 \times 27$. The original CT scans having size of $512 \times 512 \times 36$ is named as P_6. The height and width of the patch P_5 is half of the P_6, whereas these dimensions of the patch P_4 is one-fourth of the P_6, and so on. We extract 2^n number of patches for a nth-time reduction of the height and width. Therefore, we train and test our network with 71040, 35520, 17760, 8880, 4440, and 1110 samples for the 3D volumes P_1 to P_6, respectively. The examples of the extracted patches are shown in Fig. 4, where we select the middle slices of the extracted patches of the same CT scan. Different patches in Fig. 4 shows their respective resolutions, where it is seen that
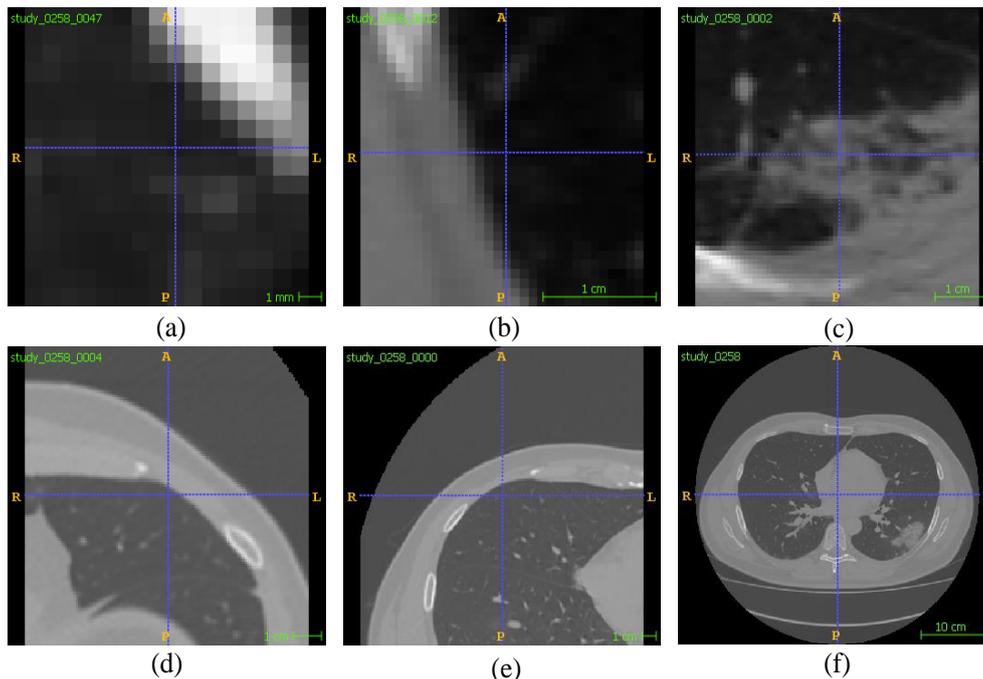
14

Figure 4: Example of various extracted patches having different sizes, as mentioned earlier, where patches P_1 to P_6 are displayed in a) to f), respectively. The middle slices of each 3D patches are illustrated for the same sample (*study_0258.nii.gz*) in the MosMedData dataset. Slices are captured using a ITK-Snap windows version[a].

---

[a]http://www.itksnap.org/pmwiki/pmwiki.php?n=Downloads.SNAP3

the patches P_1 and P_2 demonstrate very low resolutions. However, the effects of those patch resolutions are judged by classifying the NOR vs. NCP classes (see in subsection 2.1).

The classification results are presented in Fig. 5 for all the patches (P_1 to P_5) and original CT scans (P_6) employing our 3D network without any type of preprocessing. The results show that the network inputting with P_1 patch outputs COVID-19 identification with type-II errors as 69.0 % and 25.0 % for NOR- and NCP-classes, respectively. Such results confirm that NCP-class has been identified more accurately (44.0 % more in NCP-class), pointing that classifier is biased towards the NCP-class. On the other hand, the utilization of patch P_2 produces identification results with type-II errors as 56.0 % and 39.0 % for NOR- and NCP-classes, which reduce the differences between two classes (only 17.0 % more in NCP-class). Although the P_1 patch has double samples, it fails to provide a class-balanced performance
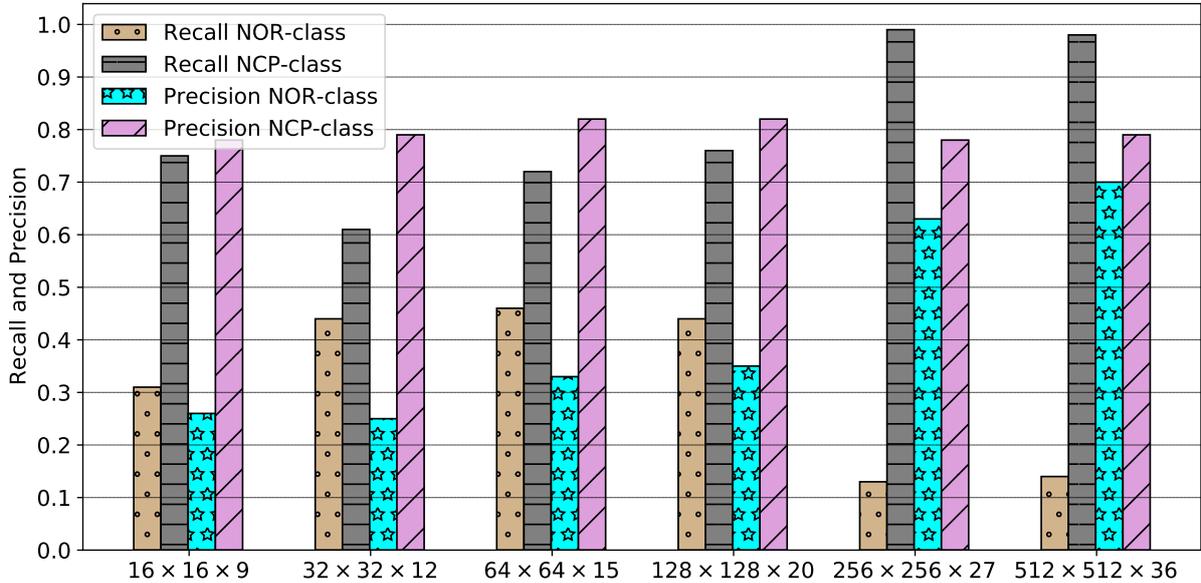
Figure 5: The binary classification results from our 3D-CNN utilizing different 3D-patch sizes, where the bars with dots, horizontal lines, stars, and diagonal hatching respectively denote recall and precision of NOR- and NCP-classes. Best view in the color figure.

than the P_2 patch. This is because of having a better-resolution in the P_2 patch than the P_1 patch (see in Fig. 4), as other experimental settings are constant. Again, the patch P_3 further improves the identification results with type-II errors as 54.0 % and 28.0 % for NOR- and NCP-classes. Approximately, the patch P_4 also provides similar results to the P_3 patch. It is noteworthy from those experimentations that P_3 or P_4 patches have much fewer samples than P_1 (4-times and 8-times, respectively); still, they outperform the identification results of P_1 and P_2 patches with the same experimental settings.

Furthermore, the utilization of patch P_5 further reduces the performances (type-II errors as 6.0 % and 99.0 % for NOR- and NCP-classes) than all the previous patches discussed above. Such a result shows that it produces a more biased model towards the NCP-class. From Fig. 4 shows that the patch P_4 and P_5 are visually looking similar but P_4 has two-times samples as of P_5. This experiment exposes that having fewer samples also generates class-biased classifiers if input images are similar in resolution.

Finally, the network with the original images also provides less COVID-19 identification

16

performance as in the patch P_5 (see in Fig. 5). All the experiments show that our network with P_3 or P_4 patches has outputted better-identification results. Such experimental results undoubtedly prove that both the input resolution and the number of samples play an important role in CNN-based classifiers. We can not increase the number of samples taking the smaller patch sizes, as it has a shallow resolution, which adversely affects the classifiers.

3.2. Progressive Resizing

The aforementioned results reveal that the utilization of better-resolution with more sample numbers increases the performance of CNN. Therefore, we propose to employ progressive resizing of our proposed 3D-CNN (see details in subsection 2.3). Firstly, we begin training our network with a suitable 3D patch with more training samples from the previous experiments, acting as a base model. Then, we add some CNN layers to the input of the base model with the higher resolution (2-times more in this article), where the base model is adopted as a pre-trained model (see details in subsection 2.3). We repeat this network resizing until we reach to original given CT size (P_6).

The results for such a progressive resizing are presented in the confusion metrics in Table 3 and ROC curves (with respective AUC values) in Fig. 6. The confusion matrix in Table 3,

Table 3: Normalized confusion matrix employing our network with progressive resizing, where we progressively increase the input resolution from P_4 to P_5 then to P_6 (original resolution). The first table (left) for the resolution of P_4, the second table (middle) for resolution of $P_4 \longmapsto P_5$, and the last (right) for resolution of $P_4 \longmapsto P_5 \longmapsto P_6$.

P_4		Actual	
		NOR	NCP
Predict	NOR	24.26 %	13.52 %
	NCP	75.74 %	86.48 %

$P_4 \longmapsto P_5$		Actual	
		NOR	NCP
Predict	NOR	21.08 %	5.38 %
	NCP	78.92 %	94.62 %

$P_4 \longmapsto P_5$ $P_5 \longmapsto P_6$		Actual	
		NOR	NCP
Predict	NOR	39.22 %	9.30 %
	NCP	60.78 %	90.70 %

for more detailed analysis of the identification results, points that 24.26 %-NOR samples are accurately classified as NOR, whereas 86.48 %-NCP samples are correctly classified as NCP while utilizing the 3D patch P_4 with 8880 samples. This training is set as a base model. Now, employing the base model as a pre-trained model, the utilization of P_5 patch, with
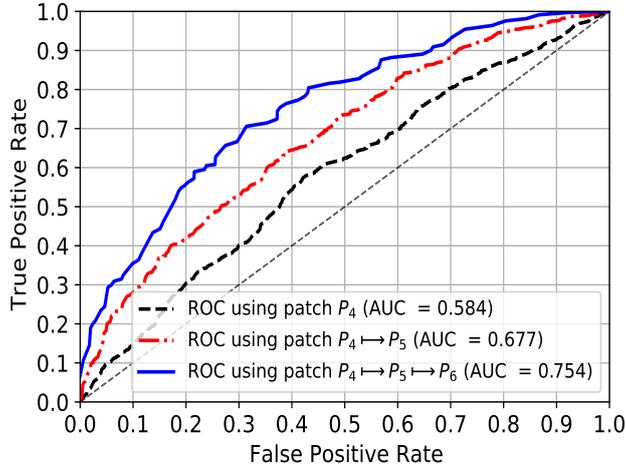
Figure 6: The ROC curves for the progressive resizing of our 3D network. Best view in the color figure.

4440 samples, decreases the false-negative rate of NCP by 8.14% although false-positive rate increases by 3.18% (see in Table 3 (left and middle)). This training is a first-time progressive resizing $(P_4 \longmapsto P_5)$. Again, employing $P_4 \longmapsto P_5$ as a pre-trained model, the utilization of P_6 (original CT scans), with 1110 samples, increases the false-negative rate of NCP by 3.92%, still less than baseline false-negative rate of 13.52%. It also decreases the false-positive rate by a margin of 18.14%, which is less than the former two false-positive rates (see all tables in Table 3). Furthermore, the proposed final progressively resized network $(P_4 \longmapsto P_5 \longmapsto P_6)$ obtains an AUC of 0.754, which indicates that the probability of correct COVID-19 identification is as high as 75.4% for any given random CT samples (see in Fig. 6). It has beaten the baseline P_4 and $P_4 \longmapsto P_5$ respectively by 17.0% and 7.70% in terms of AUC, as presented in Fig. 6. Although the final progressively resized network $(P_4 \longmapsto P_5 \longmapsto P_6)$ has an input of the original CT scans, its performance is far better than the network training with P_6 alone (see in Fig. 5). All the above discussions in this subsection experimentally validate the progressive resizing supremacy for the COVID-19 identification instead of training using single size input CT scans.

18

3.3. Prepossessing Employment

This subsection presents the COVID-19 identification results from our progressively resized 3D network employing different preprocessing, such as augmentation, segmentation, and class-rebalancing.

Table 4 bestows different experimental results, where we explicitly explicate the outcomes of each preprocessing for the COVID-19 identification from volumetric CT scans. The base-

Table 4: The COVID-19 identification results on the MosMedData dataset from our 3D-CNN network utilizing different preprocessing.

| Different experiments | Class-wise and weighted average metrics | | | | | |
| | Recall | | | Precision | | |
	NOR	NCP	Avg.	NOR	NCP	Avg.
Baseline model	0.137	0.983	0.789	0.700	0.793	0.772
Progressively Resized Network (PRN)	0.392	0.907	0.789	0.556	0.834	0.770
PRN with Augmentation (PRNA)	0.529	0.884	0.803	0.574	0.864	0.798
PRN with Lung Segmentation (PRNS)	0.333	0.971	0.825	0.773	0.831	0.818
PRNA and PRNS with Class-rebalancing (PRNASCR)	0.706	0.919	0.870	0.720	0.913	0.869

line model, without progressive resizing and inputting with original CT scans (P_6), produces low identification consequences resulting in type-II errors of 86.3 % and 1.7 % respectively for NOR- and NCP-classes, showing high class-imbalanced results. The weighted average type-II error is also only 21.1 % with respective average positive predictive value as 77.2 %. Highly imbalanced training samples ($NOR : NCP = 1 : 3.37$) with less intra-class heterogeneity and high inter-class similarity are the probable causes for providing such a poor result. However, the utilization of different 3D patches improves intra-class heterogeneity and inter-class similarity and appliance of progressive resizing, where the base model acts as a pre-trained model, can mitigate those aforementioned difficulties, which reflects in the PRN results (see in the second row of Table 4). The appliance of PRN successfully reduces the class-imbalanced results improving the type-II error of NOR-class by a margin of 25.5 %, while the weighted average type-II error is identical (21.1 %).

19

Augmentation. The employment of different image augmentations, such as random rotation, height & width shifting, gamma correction, adding Gaussian noise, and Elastic deformation (see details in subsection 2.2) with PRN further improves the COVID-19 identification results, showing far better class-balance (type-II error of NOR-class improved by a margin of 13.7 % with significantly less reduction as 2.3 % in NCP-class). The weighted average type-II error is increased by 1.4 % with respective increases in average positive predictive value by 2.8 % for the appliance of augmentations with the PRN.

Segmentation. The well-defined segmentation, with less-coarseness, is an essential requirement for further identification. The incorporation of segmentation with the PRN further promotes the identification results than the PRN alone, as exposed in Table 4. Several examples of the segmented lung from our proposed unsupervised pipeline (as described in subsection 2.2) are depicted in Fig. 7 for qualitative evaluation. However, the COVID-19
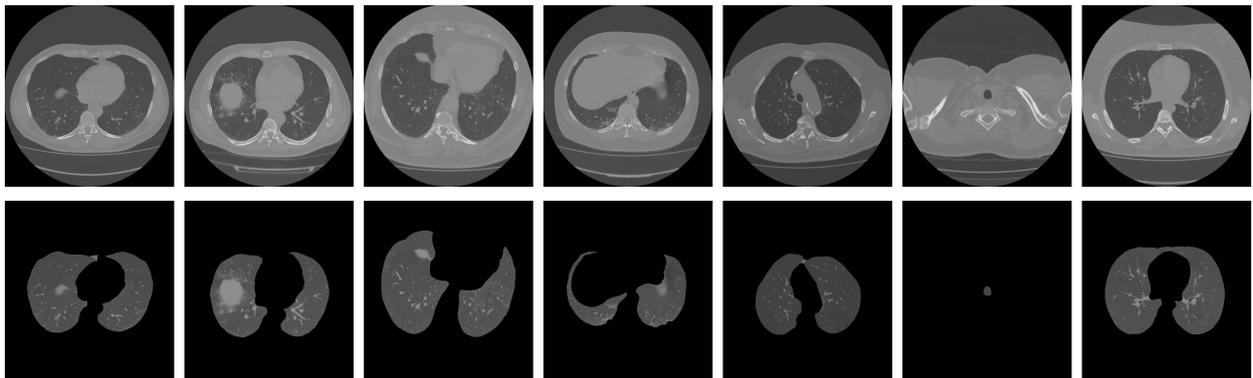

Figure 7: Examples of lung segmentation results applying our unsupervised pipeline, as described in subsection 2.2.

identification results incorporating lung segmentation with the PRN reflects its supremacy over the PRN alone, extending the weighted average type-II error by 1.4 % with respective improvements in average positive predictive value by 2.8 % (see in Table 4). The class-imbalanced identification is also dwindled due to segmented lung area utilization over the full CT volumes. The reasonable ground for those enhanced performances due to the segmentation is that it extracts an abstract region, enabling the classifier to learn only the

precise lung areas' features while avoiding the surrounding healthy tissues of the chest CT scans.

Augmentation, Segmentation, and Class-rebalancing*.* The combination of augmentations, segmentation, and class-rebalancing with the PRN provides the best COVID-19 identification results of this article. This experiment identifies the COVID-19 from the chest CT scans with relatively less class-imbalance with the weighted average type-II error of 13.0 % with respective average positive predictive value as 13.1 %. All the preprocessing employment heightens the former metric by a margins of 8.1 % and the latter metric by 9.7 % from the baseline model (see in Table 4) with less class-imbalance performance. Besides, Fig. 8 displays the ROC curves of our PRN with/without all the preprocessing and a baseline model with their corresponding AUC values. The proposed PRNASCR achieves
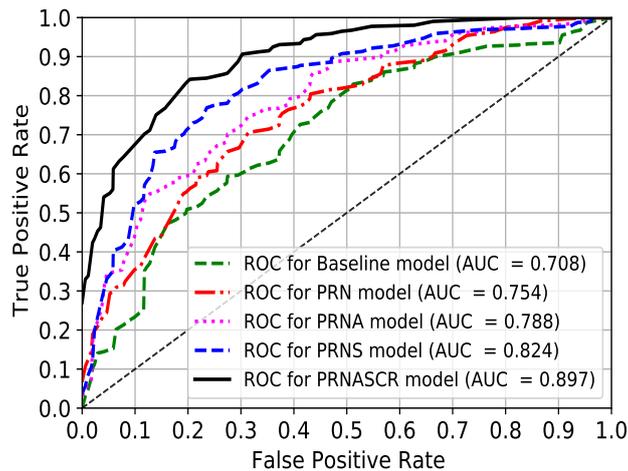

Figure 8: The ROC curves for the employment of various preprocessing to our 3D network. Best view in the color figure.

an AUC of 0.897, showing the probability of accurate COVID-19 recognition is as large as 89.7 % for any yielded random CT sample. For AUC, the proposed PRNASCR betters the baseline model, PRN, PRNA, and PRNS respectively by 18.9 %, 14.3 %, 10.9 %, and 7.3 %. From Fig. 8 and given 10.0 % false-positive rates, the true-positive rates of COVID-19 identification from the baseline model, PRN, PRNA, PRNS, and PRNASCR are approximately

21

22.0 %, 35.0 %, 46.0 %, 52.0 %, and 67.0 %, respectively, showing the improvements of 45.0 % from the baseline 22.0 %.

3.4. Binary- Vs. Multi-class Evaluation

This subsection displays the COVID-19 identification results using our proposed PRNASCR for binary- and multi-class (see in subsection 2.1) utilizing the 5-fold cross-validation. The

Table 5: The confusion matrix for the COVID-19 identification on the MosMedData dataset from our proposed 3D-CNN network and preprocessing for both the binary- (left) and multi- (right) class problems.

	4-classes	Actual			
		NOR	MiNCP	MoNCP	SeNCP
	NOR	188	67	3	2
		74.02 %	9.80 %	2.40 %	4.26 %
Predict	MiNCP	62	580	29	13
		24.41 %	84.80 %	23.2 %	27.66 %
	MoNCP	3	22	86	1
		1.18 %	3.22 %	68.80 %	2.13 %
	SeNCP	1	15	7	31
		0.39 %	2.18 %	5.60 %	65.95 %

	2-classes	Actual	
		NOR	NCP
Predict	NOR	167	8
		65.75 %	0.94 %
	NCP	87	848
		34.25 %	99.06 %

detailed class-wise performance of our PRNASCR for both the binary- and multi-class is exhibited in the confusion metrics in Table 5 (left) and Table 5 (right), correspondingly.

The binary-classification results in Table 5 (left) show that among 254-NOR CT samples, correctly classified samples are 167 (67.75 %), whereas only 87 (34.25 %) samples are erroneously classified as NCP (false positive). It is also noteworthy that among 856-NCP samples, rightly classified samples are 848 (99.06 %), whereas only 8 (0.94 %) samples are wrongly classified as NOR (false negative). Again, the matrix in Table 5 (right) for multi-class recognition reveals the FN and FP for the COVID-19 identification, where number of wrongly classified CT images (type-I or type-II errors) are 66/256 (25.78 %), 104/684 (15.20 %), 39/125 (31.20 %), and 16/47 (34.04 %) respectively for the NOR-, MiNCP-, MoNCP-, and SeNCP-classes. Those binary- and multi-class results expose that the NOR-class performance has been improved by 8.27 % margin with other constant experimental

settings. The identification results for the severity prediction (MoNCP vs. SeNCP) confer tremendous success in our pipeline, where barely 5.60 %-MoNCP and 2.13 %-SeNCP samples are prognosticated as SeNCP- and MoNCP-classes, respectively (see in Table 5). Although overall macro-average AUC of the binary classification defeats the multi-class recognition (see in Fig. 9) by a margin of 2.1 %, the later protocol has better class-balance results. The

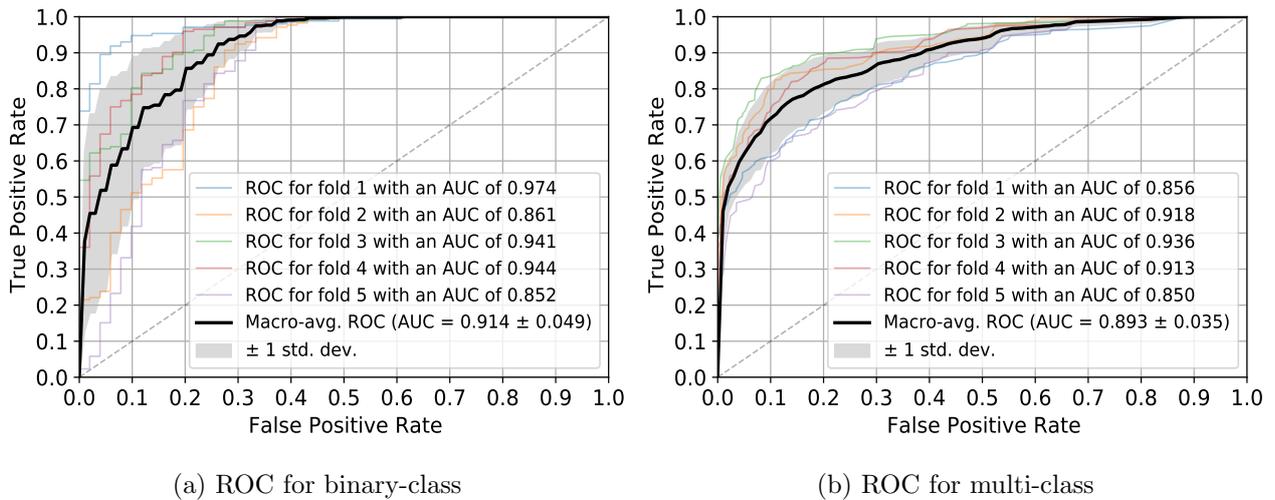(a) ROC for binary-class   (b) ROC for multi-class

Figure 9: The ROC curves for the binary- and multi-class identification of COVID-19, applying 5-fold cross-validations. Best view in the color figure.

multi-class protocol also provides less inter-fold variation than the binary-class, as depicted in Fig. 9. However, our approach for the COVID-19 identification exhibits praiseworthy achievement with high AUC values with less inter-fold variation in both of the class protocols.

4. Conclusion

During the current COVID-19 pandemic emergency, to mitigate the permanent lung damage due to coronavirus, precise recognition with negligible false negative is highly essential. This article aimed to design an artificial screening system for automated COVID-19 identification. A progressively resized 3D-CNN classifier is recommended in this study, incorporating lung segmentation, image augmentations, and class-rebalancing. The experimental

analysis confirms that the CNN classifier's training with the suitable smaller patches and progressively increasing the network size enhance the identification results. Furthermore, incorporating the lung segmentation empowers the classifier to learn salient and characteristic COVID-19 features than utilizing whole chest CT images, driving to improved COVID-19 classification performance. Again, the augmentations and class-rebalancing result in improved COVID-19 identification with high class-balanced recognition, shielding the network from being biased to a particular overrepresented class. In the future, the proposed pipeline will be employed in other volumetric medical imaging domain to validate its efficacy, versatility, and robustness. We also aim to deploy our trained model to a user-friendly web application for clinical utilization. The proposed system can be an excellent tool for clinicians to fight this deadly epidemic by the quicker and automated screening of the COVID-19.

CRediT authorship contribution statement

M. K. Hasan: Conceptualization, Methodology, Software, Formal analysis, Investigation, Visualization, Writing- Review & Editing, Supervision; **M. T. Jawad:** Software, Validation, Data Curation, Writing- Original Draft; **K. N. I. Hasan:** Data Curation, Writing- Original Draft; **S. B. Partha:** Writing- Original Draft; **M. M. A. Masba:** Writing- Original Draft;

Acknowledgements

Conflict of Interest

All authors have no conflict of interest to publish this research.

References

[1] A. J. NEWS, 2020. *Bangladesh scientists create $3 kit. Can it help detect COVID-19?* . `https://bit.ly/aj2020corona` [Accessed: 14 July 2020].

[2] Abbas, A., Abdelsamea, M.M., Gaber, M.M., 2020a. Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network. arXiv:2003.13815 .

[3] Abbas, A., Abdelsamea, M.M., Gaber, M.M., 2020b. Detrac: Transfer learning of class decomposed medical images in convolutional neural networks. IEEE Access 8, 74901–74913.

[4] Acharya, U.R., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M., Gertych, A., San Tan, R., 2017. A deep convolutional neural network model to classify heartbeats. Computers in biology and medicine 89, 389–396.

[5] Alshazly, H., Linse, C., Barth, E., Martinetz, T., 2020. Explainable covid-19 detection using chest ct scans and deep learning. arXiv:2011.05317 .

[6] Angelov, P., Almeida Soares, E., 2020. Explainable-by-design approach for covid-19 classification via ct-scan. medRxiv .

[7] Apostolopoulos, I.D., Aznaouridis, S.I., Tzani, M.A., 2020. Extracting possibly representative covid-19 biomarkers from x-ray images with deep learning approach and image data related to pulmonary diseases. Journal of Medical and Biological Engineering , 1.

[8] Apostolopoulos, I.D., Mpesiana, T.A., 2020. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Physical and Engineering Sciences in Medicine , 1.

[9] Arani, E., Marzban, S., Pata, A., Zonooz, B., 2021. Rgpnet: A real-time general purpose semantic segmentation, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 3009–3018.

[10] Bergstra, J., Bengio, Y., 2012. Random search for hyper-parameter optimization. The Journal of Machine Learning Research 13, 281–305.

[11] Bhattacharya, S., Maddikunta, P.K.R., Pham, Q.V., Gadekallu, T.R., Chowdhary, C.L., Alazab, M., Piran, M.J., et al., 2021. Deep learning and medical image processing for coronavirus (covid-19) pandemic: A survey. Sustainable cities and society 65, 102589.

[12] Cai, L., Gao, J., Zhao, D., 2020. A review of the application of deep learning in medical image classification and segmentation. Annals of translational medicine 8.

[13] Celik, Y., Talo, M., Yildirim, O., Karabatak, M., Acharya, U.R., 2020. Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images. Pattern Recognition Letters .

[14] Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., et al., 2020. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in wuhan, china: a descriptive study. The Lancet 395, 507–513.

[15] Cheplygina, V., de Bruijne, M., Pluim, J.P., 2019. Not-so-supervised: a survey of semi-supervised,

multi-instance, and transfer learning in medical image analysis. Medical image analysis 54, 280–296.

[16] Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1251–1258.

[17] Codella, N.C., Nguyen, Q.B., Pankanti, S., Gutman, D.A., Helba, B., Halpern, A.C., Smith, J.R., 2017. Deep learning ensembles for melanoma recognition in dermoscopy images. IBM Journal of Research and Development 61, 5–1.

[18] Corman, V.M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D.K., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M.L., et al., 2020. Detection of 2019 novel coronavirus (2019-ncov) by real-time rt-pcr. Eurosurveillance 25, 2000045.

[19] Cruz-Roa, A., Basavanhally, A., González, F., Gilmore, H., Feldman, M., Ganesan, S., Shih, N., Tomaszewski, J., Madabhushi, A., 2014. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks, in: Medical Imaging 2014: Digital Pathology, International Society for Optics and Photonics. p. 904103.

[20] Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L., 2009. Imagenet: A large-scale hierarchical image database, in: 2009 IEEE conference on computer vision and pattern recognition, Ieee. pp. 248–255.

[21] Deng, L., Liu, Y., 2018. Deep learning in natural language processing. Springer.

[22] Dhruv, P., Naskar, S., 2020. Image classification using convolutional neural network (cnn) and recurrent neural network (rnn): A review. Machine Learning and Information Processing , 367–381.

[23] Dutta, A., Hasan, M.K., Ahmad, M., 2020. Skin lesion classification using convolutional neural network for melanoma recognition. medRxiv .

[24] Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S., 2017. Dermatologist-level classification of skin cancer with deep neural networks. nature 542, 115–118.

[25] Fu, Y., Lei, Y., Wang, T., Curran, W.J., Liu, T., Yang, X., 2020. Deep learning in medical image registration: a review. Physics in Medicine & Biology 65, 20TR01.

[26] Gaál, G., Maga, B., Lukács, A., 2020. Attention u-net based adversarial architectures for chest x-ray lung segmentation. arXiv:2003.10304 .

[27] Géron, A., 2019. Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media.

[28] Hannun, A.Y., Rajpurkar, P., Haghpanahi, M., Tison, G.H., Bourn, C., Turakhia, M.P., Ng, A.Y., 2019. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. Nature medicine 25, 65.

[29] Hasan, M., Ahamed, M., Ahmad, M., Rashid, M., et al., 2017. Prediction of epileptic seizure by analysing time series eeg signal using-nn classifier. Applied bionics and biomechanics 2017.

[30] Hasan, M., Alam, M., Elahi, M., Toufick, E., Roy, S., Wahid, S.R., et al., 2020a. Cvr-net: A deep convolutional neural network for coronavirus recognition from chest radiography images. arXiv:2007.11993 .

[31] Hasan, M., Aleef, T.A., et al., 2019. Automatic mass detection in breast using deep convolutional neural network and svm classifier. arXiv:1907.04424 .

[32] Hasan, M.K., Alam, M.A., Dahal, L., Elahi, M.T.E., Roy, S., Wahid, S.R., Marti, R., Khanal, B., 2020b. Challenges of deep learning methods for covid-19 detection using public datasets. medRxiv .

[33] Hasan, M.K., Alam, M.A., Das, D., Hossain, E., Hasan, M., 2020c. Diabetes prediction using ensembling of different machine learning classifiers. IEEE Access 8, 76516–76531.

[34] Hasan, M.K., Alam, M.A., Elahi, M.T.E., Roy, S., Martí, R., 2020d. Drnet: Segmentation and localization of optic disc and fovea from diabetic retinopathy image. Artificial Intelligence in Medicine , 102001.

[35] Hasan, M.K., Dahal, L., Samarakoon, P.N., Tushar, F.I., Martí, R., 2020e. DSNet: Automatic dermoscopic skin lesion segmentation. Computers in Biology and Medicine 120, 103738.

[36] He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778.

[37] He, X., Wang, S., Shi, S., Chu, X., Tang, J., Liu, X., Yan, C., Zhang, J., Ding, G., 2020. Benchmarking deep learning models and automated model design for covid-19 detection with chest ct scans. medRxiv .

[38] Hemdan, E.E.D., Shouman, M.A., Karar, M.E., 2020. Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images. arXiv:2003.11055 .

[39] Hesamian, M.H., Jia, W., He, X., Kennedy, P., 2019. Deep learning techniques for medical image segmentation: Achievements and challenges. Journal of digital imaging 32, 582–596.

[40] Horry, M.J., Chakraborty, S., Paul, M., Ulhaq, A., Pradhan, B., Saha, M., Shukla, N., 2020. Covid-19 detection through transfer learning using multimodal imaging data. IEEE Access 8, 149808–149824.

[41] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861 .

[42] Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., et al., 2020. Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. The lancet 395, 497–506.

[43] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017a. Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700–4708.

[44] Huang, X., Shan, J., Vaidya, V., 2017b. Lung nodule detection in ct using 3d convolutional neural networks, in: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), IEEE. pp. 379–383.

[45] Hussain, Z., Gimenez, F., Yi, D., Rubin, D., 2017. Differential data augmentation techniques for medical imaging classification tasks, in: AMIA Annual Symposium Proceedings, American Medical Informatics Association. p. 979.

[46] Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., Keutzer, K., 2016. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size. arXiv:1602.07360 .

[47] Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167 .

[48] Jain, G., Mittal, D., Thakur, D., Mittal, M.K., 2020. A deep learning approach to detect covid-19 coronavirus with x-ray images. Biocybernetics and biomedical engineering 40, 1391–1405.

[49] Jaiswal, A., Gianchandani, N., Singh, D., Kumar, V., Kaur, M., 2020. Classification of the covid-19 infected patients using densenet201 based deep transfer learning. Journal of Biomolecular Structure and Dynamics , 1–8.

[50] Ji, Y., Zhang, H., Zhang, Z., Liu, M., 2021. Cnn-based encoder-decoder networks for salient object detection: A comprehensive review and recent advances. Information Sciences 546, 835–857.

[51] Jiménez-Sánchez, A., Albarqouni, S., Mateus, D., 2018. Capsule networks against medical imaging data challenges, in: Intravascular Imaging and Computer Assisted Stenting and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis. Springer, pp. 150–160.

[52] Jnawali, K., Arbabshirani, M.R., Rao, N., Patel, A.A., 2018. Deep 3d convolution neural network for ct brain hemorrhage classification, in: Medical Imaging 2018: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 105751C.

[53] Khan, A.I., Shah, J.L., Bhat, M.M., 2020. Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images. Computer Methods and Programs in Biomedicine , 105581.

[54] Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv:1412.6980 .

[55] Ko, H., Chung, H., Kang, W.S., Kim, K.W., Shin, Y., Kang, S.J., Lee, J.H., Kim, Y.J., Kim, N.Y., Jung, H., et al., 2020. Covid-19 pneumonia diagnosis using a simple 2d deep learning framework with a single chest ct image: model development and validation. Journal of medical Internet research 22, e19569.

[56] Ko, J.P., Betke, M., 2001. Chest ct: automated nodule detection and assessment of change over time—preliminary experience. Radiology 218, 267–273.

[57] Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, pp. 1097–1105.

[58] LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. nature 521, 436–444.

[59] Lee, E.Y., Ng, M.Y., Khong, P.L., 2020. Covid-19 pneumonia: what has ct taught us? The Lancet Infectious Diseases 20, 384–385.

[60] Li, Q., Guan, X., Wu, P., Wang, X., Zhou, L., Tong, Y., Ren, R., Leung, K.S., Lau, E.H., Wong, J.Y., et al., 2020. Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia. New England Journal of Medicine .

[61] Lin, M., Chen, Q., Yan, S., 2013. Network in network. arXiv:1312.4400 .

[62] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Medical image analysis 42, 60–88.

[63] Long, Y., Gong, Y., Xiao, Z., Liu, Q., 2017. Accurate object localization in remote sensing images based on convolutional neural networks. IEEE Transactions on Geoscience and Remote Sensing 55, 2486–2498.

[64] Lu, H., Wang, H., Zhang, Q., Yoon, S.W., Won, D., 2019. A 3d convolutional neural network for volumetric image semantic segmentation. Procedia Manufacturing 39, 422–428.

[65] Mahmud, T., Alam, M., Chowdhury, S., Ali, S.N., Rahman, M.M., Fattah, S.A., Saquib, M., et al., 2021. Covtanet: A hybrid tri-level attention based network for lesion segmentation, diagnosis, and severity prediction of covid-19 chest ct scans. arXiv:2101.00691 .

[66] Morozov, S., Andreychenko, A., Pavlov, N., Vladzymyrskyy, A., Ledikhova, N., Gombolevskiy, V., Blokhin, I., Gelezhe, P., Gonchar, A., Chernina, V., et al., 2020. Mosmeddata: Chest ct scans with covid-19 related findings. medRxiv .

[67] Narin, A., Kaya, C., Pamuk, Z., 2020. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. arXiv:2003.10849 .

[68] Nayak, S.R., Nayak, D.R., Sinha, U., Arora, V., Pachori, R.B., 2020. Application of deep learning techniques for detection of covid-19 cases using chest x-ray images: A comprehensive study. Biomedical Signal Processing and Control 64, 102365.

[69] Nour, M., Cömert, Z., Polat, K., 2020. A novel medical diagnosis model for covid-19 infection detection based on deep features and bayesian optimization. Applied Soft Computing 97, 106580.

[70] Öksüz, C., Urhan, O., Güllü, M.K., 2020. Ensemble-cvdnet: A deep learning based end-to-end classification framework for covid-19 detection using ensembles of networks. arXiv:2012.09132 .

[71] Ouchicha, C., Ammor, O., Meknassi, M., 2020. Cvdnet: A novel deep learning architecture for detection of coronavirus (covid-19) from chest x-ray images. Chaos, Solitons & Fractals 140, 110245.

[72] Ozkaya, U., Ozturk, S., Barstugan, M., 2020. Coronavirus (covid-19) classification using deep features fusion and ranking technique. arXiv:2004.03698 .

[73] Ozsahin, I., Sekeroglu, B., Musa, M.S., Mustapha, M.T., Uzun Ozsahin, D., 2020. Review on diagnosis of covid-19 from chest ct images using artificial intelligence. Computational and Mathematical Methods in Medicine 2020.

[74] Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O., Acharya, U.R., 2020. Automated detection of covid-19 cases using deep neural networks with x-ray images. Computers in Biology and Medicine , 103792.

[75] Pathak, Y., Shukla, P.K., Tiwari, A., Stalin, S., Singh, S., Shukla, P.K., 2020. Deep transfer learning based classification model for covid-19 disease. IRBM .

[76] Pereira, R.M., Bertolini, D., Teixeira, L.O., Silla Jr, C.N., Costa, Y.M., 2020. Covid-19 identification in chest x-ray images on flat and hierarchical classification scenarios. Computer Methods and Programs in Biomedicine , 105532.

[77] Pham, H., Guan, M.Y., Zoph, B., Le, Q.V., Dean, J., 2018. Efficient neural architecture search via parameter sharing. arXiv:1802.03268 .

[78] Rajaraman, S., Siegelman, J., Alderson, P.O., Folio, L.S., Folio, L.R., Antani, S.K., 2020. Iteratively pruned deep learning ensembles for covid-19 detection in chest x-rays. arXiv:2004.08379 .

[79] Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., et al., 2017. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv:1711.05225 .

[80] Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer. pp. 234–241.

[81] Ruder, S., 2016. An overview of gradient descent optimization algorithms. arXiv:1609.04747 .

[82] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4510–4520.

[83] Sarvamangala, D., Kulkarni, R.V., 2021. Convolutional neural networks in medical image understanding: a survey. Evolutionary Intelligence , 1–22.

[84] Sethy, P.K., Behera, S.K., 2020. Detection of coronavirus disease (covid-19) based on deep features. Preprints 2020030300, 2020.

[85] Shojaii, R., Alirezaie, J., Babyn, P., 2005. Automatic lung segmentation in ct images using watershed transform, in: IEEE International Conference on Image Processing 2005, IEEE. pp. II–1270.

[86] Shorten, C., Khoshgoftaar, T.M., 2019. A survey on image data augmentation for deep learning. Journal of Big Data 6, 60.

[87] Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition.

arXiv:1409.1556 .

[88] Singh, D., Kumar, V., Kaur, M., 2020a. Classification of covid-19 patients from chest ct images using multi-objective differential evolution–based convolutional neural networks. European Journal of Clinical Microbiology & Infectious Diseases , 1–11.

[89] Singh, S.P., Wang, L., Gupta, S., Goli, H., Padmanabhan, P., Gulyás, B., 2020b. 3d deep learning on medical images: a review. Sensors 20, 5097.

[90] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research 15, 1929–1958.

[91] Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A., 2016. Inception-v4, inception-resnet and the impact of residual connections on learning. arXiv:1602.07261 .

[92] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., 2015. Going deeper with convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1–9.

[93] Talo, M., Yildirim, O., Baloglu, U.B., Aydin, G., Acharya, U.R., 2019. Convolutional neural networks for multi-class brain disease detection using mri images. Computerized Medical Imaging and Graphics 78, 101673.

[94] Tan, J.H., Fujita, H., Sivaprasad, S., Bhandary, S.V., Rao, A.K., Chua, K.C., Acharya, U.R., 2017. Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. Information sciences 420, 66–76.

[95] Tan, M., Le, Q.V., 2019. Efficientnet: Rethinking model scaling for convolutional neural networks. arXiv:1905.11946 .

[96] Toğaçar, M., Ergen, B., Cömert, Z., 2020. Covid-19 detection using deep learning models to exploit social mimic optimization and structured chest x-ray images using fuzzy color and stacking approaches. Computers in Biology and Medicine , 103805.

[97] Tushar, F.I., Alyafi, B., Hasan, M.K., Dahal, L., 2019. Brain tissue segmentation using neuronet with different pre-processing techniques, in: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), IEEE. pp. 223–227.

[98] Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., Pinheiro, P.R., 2020. Covidgan: Data augmentation using auxiliary classifier gan for improved covid-19 detection. IEEE Access 8, 91916–91923.

[99] Walker, P.G., Whittaker, C., Watson, O.J., Baguelin, M., Winskill, P., Hamlet, A., Djafaara, B.A., Cucunubá, Z., Mesa, D.O., Green, W., et al., 2020. The impact of covid-19 and strategies for mitigation

and suppression in low-and middle-income countries. Science .

[100] Wang, D., Hu, B., Hu, C., Zhu, F., Liu, X., Zhang, J., Wang, B., Xiang, H., Cheng, Z., Xiong, Y., et al., 2020a. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus–infected pneumonia in wuhan, china. Jama 323, 1061–1069.

[101] Wang, J., Li, F., Li, Q., 2009. Automated segmentation of lungs with severe interstitial lung disease in ct. Medical physics 36, 4592–4599.

[102] Wang, S.H., Govindaraj, V.V., Górriz, J.M., Zhang, X., Zhang, Y.D., 2020b. Covid-19 classification by fgcnet with deep feature fusion from graph convolutional network and convolutional neural network. Information Fusion 67, 208–229.

[103] Wang, W., Xu, Y., Gao, R., Lu, R., Han, K., Wu, G., Tan, W., 2020c. Detection of sars-cov-2 in different types of clinical specimens. Jama 323, 1843–1844.

[104] Wang, X., Deng, X., Fu, Q., Zhou, Q., Feng, J., Ma, H., Liu, W., Zheng, C., 2020d. A weakly-supervised framework for covid-19 classification and lesion localization from chest ct. IEEE Transactions on Medical Imaging .

[105] Wang, X., Peng, Y., Lu, L., Lu, Z., Summers, R.M., 2018. Tienet: Text-image embedding network for common thorax disease classification and reporting in chest x-rays, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 9049–9058.

[106] World Health Organization, 2020. *Naming the coronavirus disease (COVID-19).* `https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it` [Accessed: 16 July 2020].

[107] Wu, W.J., Lin, S.W., Moon, W.K., 2012. Combining support vector machine with genetic algorithm to classify ultrasound breast tumor images. Computerized Medical Imaging and Graphics 36, 627–633.

[108] Wu, Y.C., Chen, C.S., Chan, Y.J., 2020. The outbreak of covid-19: An overview. Journal of the Chinese Medical Association 83, 217.

[109] Xu, B., Wang, N., Chen, T., Li, M., 2015. Empirical evaluation of rectified activations in convolutional network. arXiv:1505.00853 .

[110] Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., Yu, L., Ni, Q., Chen, Y., Su, J., et al., 2020. A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering .

[111] Yadav, S.S., Jadhav, S.M., 2019. Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big Data 6, 1–18.

[112] Yang, T., Wang, Y.C., Shen, C.F., Cheng, C.M., 2020. Point-of-care rna-based diagnostic device for covid-19.

[113] Yıldırım, Ö., Pławiak, P., Tan, R.S., Acharya, U.R., 2018. Arrhythmia detection using deep con-

volutional neural network with long duration ecg signals. Computers in biology and medicine 102, 411–420.

[114] Yip, S.S., Klanecek, Z., Naganawa, S., Kim, J., Studen, A., Rivetti, L., Jeraj, R., 2020. Performance and robustness of machine learning-based radiomic covid-19 severity prediction. medRxiv .

[115] Zhang, K., Liu, X., Shen, J., Li, Z., Sang, Y., Wu, X., Zha, Y., Liang, W., Wang, C., Wang, K., et al., 2020. Clinically applicable ai system for accurate diagnosis, quantitative measurements, and prognosis of covid-19 pneumonia using computed tomography. Cell .

[116] Zhang, X., Zhou, X., Lin, M., Sun, J., 2018. Shufflenet: An extremely efficient convolutional neural network for mobile devices, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6848–6856.

[117] Zhao, J., Zhang, Y., He, X., Xie, P., 2020. Covid-ct-dataset: a ct scan dataset about covid-19. arXiv:2003.13865 .

[118] Zhou, X., Yamada, K., Kojima, T., Takayama, R., Wang, S., Zhou, X., Hara, T., Fujita, H., 2018. Performance evaluation of 2d and 3d deep learning approaches for automatic segmentation of multiple organs on ct images, in: Medical Imaging 2018: Computer-Aided Diagnosis, International Society for Optics and Photonics. p. 105752C.

[119] Zoph, B., Vasudevan, V., Shlens, J., Le, Q.V., 2018. Learning transferable architectures for scalable image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 8697–8710.