Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
COVID-19 identification from volumetric chest CT scans using a progressively resized 3D-CNN incorporating segmentation, augmentation, and class-rebalancing

Md. Kamrul Hasan a,∗, Md. Tasnim Jawad a, Kazi Nasim Imtiaz Hasan b, Sajal Basak Partha b, Md. Masum Al Masba c, Shumit Saha d, Mohammad Ali Moni d,e

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 Department of Electronics and Communication Engineering, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh
d Department of Computer Science & Engineering, Pabna University of Science and Technology, Pabna 6600, Bangladesh
e School of Health and Rehabilitation Sciences, Faculty of Health and Behavioural Sciences, The University of Queensland, St Lucia, QLD 4072, Australia

A R T I C L E I N F O

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

A B S T R A C T

The novel COVID-19 is a global pandemic disease overgrowing worldwide. Computer-aided screening tools with greater sensitivity are 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 end-to-end on the 3D patches from the whole volumetric Computer Tomography (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 into 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 Receiver Operating Characteristics (ROC) curve of 0.914 ± 0.049 and 0.893 ± 0.035 for the binary- and multi-class tasks, respectively, 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.

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. Therefore, on 11th February 2020, WHO termed that new disease as COVID-19 Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) because of having similar genetic properties to the SARS outbreak of 2003. Therefore, on 11th February 2020, WHO termed that new disease as COVID-19 (Coronavirus disease), which displays an upper respiratory tract and lung infection [1]. The clinical characteristics of critical COVID-19 pandemic are bronchopneumonia that affects cough, fever, dyspnea, and detailed respiratory anxiety ailment [2–4]. 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 [5]. Those shared signs turn it challenging to recognize the virus at an ancient step. Moreover, the aforementioned is a virus, which works on bacterial or fungal infections [5,6] with no possibility that antibiotics can restrict it. Besides, people suffering from medical complications, like diabetes and chronic respiratory and cardiovascular diseases, can undergo COVID-19. An explanatory statement of the Imperial College advised that the affection rate will be more than

∗ Correspondence to: Department of EEE, KUET, Khulna-9203, Bangladesh.
E-mail addresses: m.k.hasan@eee.kuet.ac.bd (M.K. Hasan), jawad1703006@stud.kuet.ac.bd (M.T. Jawad), hasan1607097@stud.kuet.ac.bd (K.N.I. Hasan), m.asba1607101@stud.kuet.ac.bd (M.A. Masba), shumit.saha@mail.utoronto.ca (S. Saha), m.moni@uq.edu.au (M.A. Moni).

https://doi.org/10.1016/j.imu.2021.100709
Received 27 June 2021; Received in revised form 11 August 2021; Accepted 18 August 2021
Available online 28 August 2021
2352-9148/© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license.
90.0% of the world’s people, killing 40.6 million people if no reduction actions are grasped to combat the virus [7].

Advanced presumed discovery of COVID-19 is also a challenge for public health security and control of pandemics. 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 ~ 14 days, making it remarkably challenging to identify COVID-19 infection at a preliminary stage of an individual’s symptoms [5]. 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% [8]. There is also a significant lack of RT-PCR kit inventory, leading to a delay in preventing and curing coronavirus disease [9]. Furthermore, the RT-PCR kit is estimated to cost around 120 ~ 130 USD. It also requires a specially designed biosafety laboratory to house the PCR unit, each of which can cost 15,000 ~ 90,000 USD [10]. Nevertheless, using a costly screening device with delayed test results makes it more challenging to restrict the disease’s spread. Inadequate availability of screening workstations and measurement kits constitutes an enormous hardship to identify COVID-19 in this pandemic circumstance. In such a situation, speedy and trustworthy presumed COVID-19 cases are a massive difficulty for related personals.

However, it is observed that most of the COVID-19 incidents have typical properties on radiographic Computed Tomography (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 [11–14]. 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 [15]. Currently, deep Convolutional Neural Networks (CNNs) allow for building an end-to-end model without requiring manual and time-consuming feature extraction and engineering [16–19], demonstrating tremendous success in many domains of medical imaging, such as arrhythmia detection [20–22], skin lesion segmentation and classification [23–26], breast cancer detection [27–29], brain disease segmentation and classification [30,31], pneumonia detection from chest X-ray images [32], fundus image segmentation [33,34], and lung segmentation [35]. 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, where the mixed datasets indicate that data have come from different open-sources. 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) [14,75,76,79,84–87]. This brings inherent bias on the algorithms as the model tends to learn the distribution of the data source for binary classification problems [74]. Therefore, these models perform very low when used in practical settings, where the models have to adapt to data from different domains [74]. Recently, the authors in [88] launched a public chest volumetric CT scan dataset with 1110 COVID-19 related studies (see details in Section 2.1). However, the published articles [89,90] 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 Section 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 [83,90–94]
- 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
- 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 materials utilized and methods to conduct the widespread experiments. We summarize the adopted dataset in the first Section 2.1. The essential integral preprocessing, such as segmentation, augmentation, and class-rebalancing, are reported in the second Section 2.2. The design of the proposed 3D-CNN-based COVID-19 classifier, along with its training protocol, is explained in the third Section 2.3. Finally, in the fourth Section 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 [88]. This dataset includes anonymized human chest CT scans in Neuroimaging Informatics Technology Initiative (NIFTI) format 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 ~ 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 them as SeNCP, as CrNCP has only two samples in the MosMedData dataset. We have applied a 5-fold 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 the 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. Therefore, 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 in the proposed framework (see in Fig. 1) consists of segmentation, augmentations (both geometry- and intensity-based), and class-rebalancing, which are concisely explained as follows:
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 (recall), specificity, and F1-score (see metric details in Section 2.4).

| Different methods | Datasets | Results |
|-------------------|----------|---------|
| A pre-trained 2D MobileNet-v2 [36] architecture on ImageNet [37] was used to extract massive high-dimensional features to classify six different diseases using the fully-connected layers [38] | Mixed | mF1: 0.974 mSp: 0.944 |
| DeTrAc [39,40], where the network was trained first using a gradient descent optimization [41], and then, the class-composition layer of DeTrAc was used to refine the final detection results [40] | Mixed | mF1: 0.970 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 [14] | Mixed | mF1: 0.907 mSp: 0.906 |
| An ensemble of VGG-16 [42], Inception [43], Xception [44], Inception-ResNet [45], MobileNet [46], DenseNet [47], and NasNet [48] optimizing the hyperparameters using a greedy search algorithm [49,50] | Mixed | mF1: 0.990 mSp: 0.990 |
| Support vector machine [29,51]-based method to classify the in-depth features from the pre-trained MobileNet and SqueezeNet [52] from the restructured the data using a fuzzy color technique [53] | Mixed | mF1: 0.983 mSp: 0.997 |
| An ensemble of three lightweight pre-trained SqueezeNet, ShuffleNet [54], and EfficientNet-B0 [55] at various depths and consolidates feature maps in diverse abstraction levels [56] | Mixed | mF1: 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 [57] | Mixed | mF1: 0.989 mSp: 0.976 |
| A DenseNet-201 [47]-based transfer learning to extract features using its learned weights on the ImageNet was used to classify the patients as COVID infected or not [58] | [59] | mF1: 0.960 mSp: 0.960 |
| A transfer learning-based approach using one of the VGG, ResNet [60], Inception, or Xception pre-trained deep learning model on ImageNet as a backbone [61] | [62] | mF1: 0.996 mSp: 0.100 |
| A weakly-supervised learning schema, where the lung region was segmented using a pre-trained UNet [63]; then, a 3D network was used to predict the probability of COVID-19 infectious [64] | [64] | mF1: 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 [65] | Mixed | mF1: 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 [66] | Mixed | mF1: 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 [67] | Mixed | mF1: -- |
| 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 [68] | Mixed | mF1: 0.915 mSp: 0.948 |
| An auxiliary classifier generative adversarial network-based to generate synthetic images, where the synthetic images produced CNN’s enhanced results for the prediction [69] | Mixed | mF1: 0.900 mSp: 0.970 |
| A framework consisting of a CNN-based feature extractor and k-nearest neighbor [70,71], support vector machine, and decision tree [71]-based classifiers using the Bayesian algorithm [72] | Mixed | mF1: 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 [73] | Mixed | mF1: -- mSp: 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 [74] | Mixed | mF1: 0.976 |
| 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 [75] | [75] | mF1: 0.963 mSp: 0.970 |
| An end-to-end DarkCovidNet architecture [76] based on DarkNet [76] gradually increasing the number of filters, where each convolutional layers were followed by BatchNorm [77] and LeakyReLU [78] | Mixed | mF1: 0.951 mSp: 0.953 |
| A GorNet model based on pre-trained Xception architecture on ImageNet for automated detection of COVID-19 infection and trained in end-to-end manners [79] | Mixed | mF1: 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 [80] | Mixed | mF1: 0.100 mSp: 0.967 |
| A comparative analysis of different CNN models, such as VGG, Resnet, Inception, Xception, Inception-ResNet, DenseNet, and NASNet-Large [81], to decide a proper one for multi-modal image classification minimizing the image quality imbalances in the image samples as a preprocessing [82] | Mixed | mF1: 0.820 mSp: -- |
| A pipeline consisting of segmentation and subsequent classification employing both 2D and 2D CNNs, where the promising results for detecting were obtained in the 3D-CNNs than the 2D CNNs [83] | [83] | mF1: 0.891 mSp: 0.911 |

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 [95]. 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. 2 depicts the pipeline of the proposed LS method. The proposed threshold-based LS’s 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 [96–98]. 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 areas, 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 [99]. Thirdly, morphological erosion separates 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 [100]. Such an unsupervised thresholding-based segmentation method is better for efficiency, taking only a few seconds, and yields utterly reproducible LS.
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.

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 massive training samples are still unavailable. 

Rebalancing. The utilized dataset in Table 2 is imbalanced. This situation is 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 (90 samples) from the publicly available CC-CCII dataset (see details in [103]) 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 nth-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.

Fig. 1. The intended framework for the automatic COVID-19 identification from the volumetric 3D CT scans, consisting of essential integral preprocessing.

Fig. 2. The proposed block diagram of an unsupervised lung segmentation pipeline, without requiring a manually annotated lung region.

1 https://pypi.org/project/elasticdeform/.
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 [104] to medical image classification [105], segmentation [105], and registration [106]. In special, CNNs have become a prevalent technique in the computer vision community. They are practiced in diverse tasks, including object detection [107], classification [108], and localization [109]. The CNN-based deep neural systems are also popularly adopted in recent pandemic for COVID-19 identification [110, 111] (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 [112]. The early few CNN layers learn low-level image features and later layers learn high-level image features particular to the application types [91]. However, the 2D-CNNs are frequently employed in natural RGB and grayscale images to extract the spatial features only in two dimensions [43]. 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 [113]. Conventional 2D-CNNs' effectiveness is degraded due to the 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 [114]. Therefore, we propose a classifier based on 3D-CNN to identify COVID-19 from the volumetric CT scans.

Fig. 3. The architectural construction of the proposed base network, inputted with a size of \( M \times N \), training with the most smaller 3D patches. This trained base network is applied as a pre-trained model for the following bigger patches. Best view in the color figure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 3 represents the constructional structure of our proposed COVID-19 base classifier of the proposed framework in Fig. 1. The proposed base network in Fig. 3 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. In addition, 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 Rectified Linear Unit (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 [92]. The more enormous changes prone to bring difficulties for searching an optimal parameter or hyperparameter; often become computationally expensive to reach an optimal value [77]. Such a problem is mitigated by integrating batch normalization layers in our network [77]. It also facilitates the smooth training of the network architectures in fewer times [92]. The Global Average Pooling (GAP) [115] 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 [115]. A dropout layer [116] is also employed as a regularizer, 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 [117]. There are two general resolving directions: accumulating more data, such as crowdsourcing [118] or digging into the present clinical reports [119]. 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 because the understanding achieved from the research can migrate to the medical data collection and labeling are confronted with data privacy, requiring time-consuming expert explanations [117].

Transfer learning is a widely adopted method for advancing the performance of CNNs with inadequate datasets [120]. There is no public pre-trained 3D-CNN model for the COVID-19 identification from the volumetric chest images with limited samples to our most trustworthy knowledge. Therefore, we create a pre-trained model by training our base model (see in Fig. 3) on the extracted 3D patches from whole chest CT scans (see details in Section 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. 4). At the same
time, we keep the base model’s trained weights for the smaller patches. We repeat to enlarge (nth-times) the patch and network sizes until we arrive at the provided CT scans’ size, as pictured in Fig. 4. Such training is called progressive resizing [121], where the training begins with smaller image sizes followed by a progressive expansion of size. This training process continues until the last patch, and network sizes are the same as the initial image dimension.

2.3.2. Training protocol
We first extract five different patches with different sizes (see in Fig. 5) to begin the experiments. We perform ablation studies in Section 3.1 looking for the best patch size. The weights of the base network in Fig. 3 are initialized with Xavier’s normal distribution. The filters in each layer of the network are initialized with the “he normal” distribution. 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 [122] optimizer with initial learning rate (LR), exponential decay rates (β₁, β₂) as LR = 0.0001, β₁ = 0.9, and β₂ = 0.999, respectively, without AMSGrad [123] 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 [124] and image processing APIs, and MATLAB programming languages. The former Python language is utilized to implement a deep learning framework, and the latter MATLAB language is used to access the NIfTI Tool² for volumetric medical images. 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 mathematical formulations of those metrics are provided in Eq. (1).

\[
\text{Recall} = \frac{TP}{TP + FN}, \\
\text{Precision} = \frac{TP}{TP + FP}, \\
F1\text{-score} = \frac{2 \times TP}{2 \times TP + FP + FN},
\]

where TP, FP, and FN indicate true positive, false positive, and false negative of the detection, respectively. The recall measures the type-II error (the patient having positive COVID-19 characteristics erroneously abandons to be repealed). In contrast, 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 trade-off between these two metrics. Furthermore, we also quantify the prognostication probability of an anonymously picked 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 a comprehensive discussion. In Section 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 Section 3.2. We demonstrate the effects of different proposed preprocessing on COVID-19 identification in Section 3.3. In this subsection, we also investigate a statistical ANOVA test to validate the proposed preprocessing for the aimed task. Finally, in Section 3.4, we dispense the results for binary- and multi-class COVID-19 classification applying our proposed network and preprocessing.

---

² https://in.mathworks.com/matlabcentral/profile/authors/757722.
3.1. Patch selection

We extract five different 3D patches, named \( P_1, P_2, P_3, P_4, \) and \( P_5 \), having respective sizes 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 \( P_0 \). The height and width of the patch \( P_5 \) is half of the \( P_0 \), whereas these dimensions of the patch \( P_4 \) is one-fourth of the \( P_0 \), and so on. We extract 2\( n \) number of patches for a \( n \)-th-time reduction of the height and width. Therefore, we train and test our network with \( 71040 (= 2^6 \times 1110), 35520 (= 2^5 \times 1110), 17760 (= 2^4 \times 1110), 8880 (= 2^3 \times 1110), 4440 (= 2^2 \times 1110), \) and 1110 samples for the 3D volumes \( P_1 \) to \( P_6 \), respectively. The examples of the extracted patches are shown in Fig. 5, where we select the middle slices of the extracted patches of the same CT scan.

Various patches in Fig. 5 confers their respective resolutions, where it is seen that the patches \( P_1 \) and \( P_2 \) demonstrate very low resolutions, which may degrade the network’s learning capability. However, the effects of those patch resolutions are judged by classifying the NOR vs. NCP classes (see class details in Section 2.1).

The classification results are presented in Fig. 6 for all the patches (\( P_1 \) to \( P_6 \)) and original CT scans (\( P_0 \)), employing our 3D network without any type of preprocessing. The results determine 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 \( P_2 \) has better identification results. Such experimental results unquestionably prove that the input resolution and the number of samples play an important role in CNN-based classifiers’ learning. Unfortunately, we cannot increase the number of specimens taking the smaller patch sizes of the given images in the patch-based strategies, 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 for image identification. Therefore, we propose to employ progressive resizing of our proposed 3D-CNN (see details in Section 2.3). Firstly, we begin training the recommended 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 Section 2.3). Finally, we repeat this network resizing until we reach to original given CT size (\( P_0 \)).

The results for such a progressive resizing are presented in the confusion metrics in Table 3 and ROC curves (with respective AUC.
which is less than the former two false-positive rates (see all tables in Table 3). Furthermore, the proposed final progressively resized network \( (P_4 \rightarrow P_3 \rightarrow P_2) \) 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. 7). It has beaten the baseline \( P_4 \) and \( P_3 \rightarrow P_2 \) respectively by 17.0% and 7.70% in terms of AUC, as presented in Fig. 7. Those ROC curves in Fig. 7 also depict that for a given false-positive rate, the possibility of getting a true-positive rate is significantly higher for the \( P_4 \rightarrow P_3 \rightarrow P_2 \) than the other two approaches. Although the final progressively resized network \( (P_4 \rightarrow P_3 \rightarrow P_2) \) has an input of the original CT scans, its performance is considerably better than the network training with \( P_4 \) alone (see in Fig. 6). Transferring the knowledge from the prior smaller patch inputted network is responsible for achieving much better results than the random initialization of the filter kernels of the proposed network. All the above discussions of this subsection experimentally certify the progressive resizing supremacy for the COVID-19 identification instead of training using single-size input CT scans.

### 3.3. Prepossessing employment

This subsection exhibits the COVID-19 identification results from our progressively resized 3D network, employing different preprocessings for the ablation studies, such as augmentation, segmentation, and class-rebalancing. At the end of this subsection, the effectiveness of the introduced preprocessing has been validated by applying a statistical ANOVA test.

Table 4 bestows different experimental results, where we explicitly illustrate the outcomes of each preprocessing for the COVID-19 identification from volumetric chest CT scans. Without progressive resizing and inputting with original CT scans \( (P_4) \), the baseline model produces low identification consequences resulting in type-II errors of 86.3% and 1.7% respectively for NOR- and NCP-classes, conferring high class-imbalance results. The weighted average type-II error is only 21.1% with respective average positive predictive value as 77.2%.

| P₁ | NOR (24.26%) | NCP (75.74%) | Predict | NOR (21.00%) | NCP (78.92%) | Actual | NOR (39.22%) | NCP (60.78%) |
|----|--------------|--------------|---------|--------------|--------------|--------|--------------|--------------|

![Normalized confusion matrix employing our network with progressive resizing, where we progressively increase the input resolution from \( P_4 \) to \( P_3 \) (original resolution). The first table (left) for the resolution of \( P_4 \), the second table (middle) for resolution of \( P_3 \rightarrow P_2 \), and the last (right) for resolution of \( P_2 \rightarrow P_1 \).](image)

![ROC curves for the progressive resizing of our 3D network. Best view in the color figure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)](image)
The experimental results determine that, on average, 77.2% samples are correctly recognized as positive out of all positive identifications, which are 70.0% and 79.3% for the NOR- and NCP-classes, respectively. 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 of the intended task. However, the utilization of different 3D patches in the following experiment of progressive resizing (PRN) improves the classification results, where the base model on the smaller patches acts as a pre-trained model. This PRN also mitigates the class-imbalanced outcomes, as reflected in the PRN results (see in the second row of Table 4).

The appliance of PRN successfully improves the type-II error of NOR-class by a margin of 25.5%, while the weighted average type-II error is identical to the baseline experiment (21.1%). In the PRN, the network is trained with more sample numbers and better resolution (discussed in Section 3.1 and Section 3.2), which improves intra-class heterogeneity and inter-class similarity, leading to enhanced outcomes in the PRN experiment. However, the following three paragraphs explain the results of our progressively resized 3D network, combining Augmentation, Segmentation, and Class-rebalancing in the manner of ablation studies.

**PRN with Augmentation (PRNA).** The employment of different image augmentations, such as random rotation, height & width flipping, gamma correction, adding Gaussian noise, and Elastic deformation (see details in Section 2.2) with PRN further improves the COVID-19 identification results (see in Table 4). An example of such augmentations is displayed in Fig. 8. It gives far better class-balanced (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 augmentation with the PRN. This experiment reveals that the augmentations bring variability in the training sample with high intra-class heterogeneity and less inter-class similarity, which empowers the CNN model to learn from many examples, driving to enhanced classification results.

**PRN with Segmentation (PRNS).** The well-defined segmentation, with less coarseness, is an essential requirement for further classification. 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 Section 2.2) are depicted in Fig. 9 for qualitative evaluation. Those segmentation results quantitatively demonstrate that the segmented lung regions are pretty much accurate visually. However, the COVID-19 identification results incorporating lung segmentation with the PRN reflects its supremacy over the PRN alone, extending the weighted average type-II error by 3.6% with respective improvements in average positive predictive value by 4.8% (see in Table 4). Those results indicate that 4.8% additional specimens are perfectly identified as positive out of all positive classifications. The class-imbalanced identification is also lesserened due to segmented lung area utilization over the complete 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.

**PRN with Augmentation, Segmentation, and Class-rebalancing (PRNASCR).** Combining augmentations, segmentation, and class-rebalancing with the PRN provides this article’s best COVID-19 identification results. This experiment identifies the COVID-19 from the chest CT scans with relatively more minor 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. 10 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 an AUC of 0.897, determining the probability of accurate COVID-19 recognition is as large as 89.7% for any yielded random chest CT sample. For AUC, the proposed PRNASCR better the baseline model, PRN, PRNA, and PRNS respectively by 18.9%, 14.3%, 10.9%, and 7.3%. From Fig. 10 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 22.0%, 35.0%, 46.0%, 52.0%, and 67.0%, respectively, showing the improvements of 45.0% from the baseline 22.0%.

Moreover, the experimental results from different experiments, such as baseline model, PRN, PRNA, PRNS, and PRNASCR, have been confirmed employing a statistical test called ANOVA and 5-fold cross-validation. Fig. 11 dispenses the Box and Whisker plot of the AUC values of this validation test. For ANOVA testing, a = 0.05 is applied as a threshold to reject the Null hypothesis (all experiments’ mean AUC values are equal) if p-value ≤ 0.05, which outcomes significant results. The ANOVA test demonstrates a p-value of 0.000837 (≤ 0.05), which reveals that an alternative hypothesis is accepted, strongly pointing that none of the mean AUCs are equal (also displayed in Fig. 11). Again, a post hoc T-test (Bonferroni correction) is combined with the ANOVA test for determining the better experiment for the recommended COVID-19 detection system, which confirms the superiority of the offered PRNASCR scheme for the aimed task.

### 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 Section 2.1), utilizing the 5-fold cross-validation. The 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 (90.06%), 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-,...
Table 5
The confusion matrix for the COVID-19 identification on the MosMedData dataset from our proposed 3D-CNN network and preprocessing for the binary- (left) and multi- (right) class problems.

| Actual | 2-classes | Actual | 4-classes |
|--------|-----------|--------|-----------|
|        | NOR | NCP | NOR | MINCP | MoNCP | SeNCP |
| Predict | NOR | 167 | 8 | 188 | 67 | 3 | 2 |
|         | 65.75% | 0.94% | 74.02% | 9.80% | 2.80% | 4.26% |
| NCP     | 87  | 848 | 24.25% | 99.06% |

Fig. 8. Qualitative examples of augmentation results applied for automatic COVID-19 identification from the proposed framework.

Fig. 9. Examples of lung segmentation results applying our unsupervised pipeline, as described in Section 2.2.

Fig. 10. The ROC curves for the employment of various preprocessing to our 3D network. Best view in the color figure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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. 10) by a margin of 2.1%, the later protocol has better class-balance results. The multi-class protocol also provides more minor inter-fold variation than the binary-class, as depicted in Fig. 12. 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.
To compare our proposed framework with the existing works [125–130], we consider the articles, which are tested on the same MosMedData dataset. Fig. 13 exhibits a quantitative comparison of different methods for the same task and dataset, dispensing recall and accuracy. The comparative results in Fig. 13 point that the recommended framework achieves the highest recall of 0.991, delivering significantly less type-II error of COVID-19 identification. Although the accuracy of our system is less, it is still better than other methods of Ibrahim et al. [125], Qiblawey et al. [126], Demir et al. [127], Mittal and Oh [128], Lacerda et al. [129], Mahmud et al. [130] for medical diagnostic procedures because our system has significantly fewer false-negative rates with slightly higher false-positive rates. Additionally, most of the authors in [125–130] split data for training, validation, and testing without cross-validation. However, selecting test images in such a splitting cannot guarantee not choosing the most straightforward samples, which perhaps exhibited better accuracy from those articles.

4. Conclusion

During the current COVID-19 pandemic emergency, precise recognition with negligible false negative is essential to mitigate the permanent lung damage due to coronavirus. 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 rather than utilizing whole chest CT images, driving 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 domains 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**

Md. Kamrul Hasan: Conceptualization, Methodology, Software, Formal analysis, Investigation, Visualization, Writing – review & editing, Supervision. Md. Tasnim Jawad: Software, Validation, Data curation, Writing – original draft. Kazi Nasim Imtiaz Hasan: Data curation, Writing – original draft. Sa jal Basak Partha: Writing – original draft. Md. Masum Al Masba: Writing – original draft. Shumit Saha: Supervision. Mohammad Ali Moni: Writing – review & editing.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgments**

The authors acknowledge the contributors of the publicly available MosMed dataset in this COVID-19 pandemic situation. The authors also acknowledge the Keras developer team for building the deep learning API.

**References**

[1] World Health Organization. Naming the coronavirus disease (COVID-19). 2020. https://bit.ly/3qE6q5s. [Accessed 16 July 2020].

[2] Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. JAMA 2020;323(11):1061-9.

[3] Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet 2020;395(10223):507–13.

[4] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of coronavirus-infected pneumonia. N Engl J Med 2020.

[5] Jain G, Mittal D, Thakur D, Mittal MK. A deep learning approach to detect Covid-19 coronavirus with X-Ray images. Biocybern Biomed Eng 2020;40(4):1391–405.

[6] Wu YC, Chen C-S, Chan YJ. The outbreak of COVID-19: An overview. J Chin Med Assoc 2020;83(2):1217.

[7] Walker PG, Whitkater C, Watson OJ, Baguelin M, Winskll P, Hamlet A, et al. The impact of COVID-19 and strategies for mitigation and suppression in low and middle-income countries. Science 2020.

[8] Wang W, Xu Y, Gao R, Li J, Wang X, Gao R, et al. Detection of SARS-CoV-2 in different types of clinical specimens. JAMA 2020;323(18):1843–4.

[9] Yang T, Wang YC, Shen CF, Cheng CM. Point-of-care RNA-based diagnostic device for COVID-19. 2020.

[10] NEWS AJ. Bangladesh scientists create $3 device for COVID-19. 2020. https://bit.ly/3zQEqn5. [Accessed 16 July 2020].

[11] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China. J Natl Med Assoc 2020;112(1):51–4.

[12] Edalatifar M, Tavakoli MB, Ghalambaz M, Setoudeh F. Using deep learning to learn physics of conduction heat transfer. J Therm Anal Calorim 2020;1–18.

[13] Edalatifar M, Tavakoli MB, Ghalambaz M, Setoudeh F. A dataset for conduction heat transfer and deep learning. MedInd Data 2020.

[14] Vildrum O, Plawiak P, Tan RS, Acharya UR. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. Comput Biol Med 2018;102:41–20.

[15] Hannun AY, Rajpurkar P, Haghpanah M, Tison F, Bourn C, Turakhia MP, et al. Cardio-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep convolutional neural network. Nat Med 2019;25(1):1365.

[16] Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adam M, Gertych A, et al. A deep convolutional neural network model to classify heartbeats. Comput Biol Med 2018;99:389–96.

[17] Kazi Nasim Imtiaz Hasan: Writing – original draft. Mohammad Ali Moni: Writing – review & editing.

**Declaration of competing interest**

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

**Acknowledgments**

The authors acknowledge the contributors of the publicly available MosMed dataset in this COVID-19 pandemic situation. The authors also acknowledge the Keras developer team for building the deep learning API.
Wang SH, Govindaraj VV, Górriz JM, Zhang X, Zhang YD. Covid-19 classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network. Inf Fusion 2020;2017:2017.

Jaiswal A, Gianchandani N, Singh D, Kumar V, Kaur M. Classification of the Covid-19 infection detection based on deep features and Bayesian optimization. Appl Bionics Biomech 2020;8:76516–31.

Zhang X, Zhou S, Shi S, Chu X, Tang J, Liu X, et al. Benchmarking deep learning models and automated model design for COVID-19 detection with chest CT scans. 2020, medRxiv.

He X, Wang Y, Song C, Niu Y. Data augmentation using auxiliary classifier gan for improved pneumonia diagnosis using a simple 2D deep learning framework with a single chest CT image: model development and validation. J Med Internet Res 2020;12:e19569.

Zhang X, Zhao X, Sun J, Ren X, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016. p. 770–8.

Wang J, Zhang Y, He X, Xie P. COVID-CT-dataset: a CT scan dataset about COVID-19. 2020, arXiv:2003.13865.

Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention. Springer; 2015, p. 234–41.

Wang X, Deng X, Fu Q, Zhou Q, Feng J, Ma H, et al. A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT. IEEE Trans Med Imaging 2020;39:374–84.

Hasan M, Alam M, Elahi M, Toutkif E, Roy S, Wahid SR, et al. Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. Comput Methods Programs Biomed 2020;2017:2017.

Hasan MK, Alam MA, Das D, Hossain E, Hasan M. Diabetes prediction using ensembling of different machine learning classifiers. IEEE Access 2020;8:76516–31.

Nour M, Comert Z, Polat K. A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization. Appl Soft Comput 2020;97:106580.

Hasan MK, Alam MA, Dhahal L, Elahi MTE, Roy S, Wahid SR, et al. Challenges of deep learning methods for COVID-19 detection using public datasets. 2020, medRxiv.

Wang SH, Govindaraj VV, Górriz JM, Zhang X, Zhang YD. Covid-19 classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network. Inf Fusion 2020;2017:2017.

Hasan MK, Alam MA, Deo MV, Deo J. Efficient neural architecture search via parameter sharing. 2018, arXiv:1802.03268.

Nguyen DV, Tran BD, Le QV. Efficientnet: rethinking model scaling for convolutional neural networks. 2019, arXiv:1905.11946.

Öksüz C, Urban O, Gülük MK. Ensemble-CVDNet: A deep learning based end-to-end classification framework for COVID-19 detection using ensembles of networks. 2020, arXiv:2012.09132.

Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. ICML 2015;2015.

Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K. SqueezeNet: AlexNet-level accuracy with <0.5 MB model size. 2016, arXiv:1602.07360.

Toğaçar M, Ergen B, Cömert Z. COVID-19 detection using deep learning models exploiting Social Mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches. Comput Biol Med 2020;103805.

Wang X, Deng X, Fu Q, Zhou Q, Feng J, Ma H, et al. A weakly-supervised classification by FGCNet with deep feature fusion from graph convolutional network and convolutional neural network. Inf Fusion 2020;2017:2017.

Wang X, Shi S, Chu X, Tang J, Liu X, et al. Benchmarking deep learning models and automated model design for COVID-19 detection with chest CT scans. 2020, medRxiv.

Apollonioulous ID, Mpemiana TA. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 2020;1–8.

Sethy PK, Behera SK. Detection of coronavirus disease (covid-19) based on deep features. 2020, p. 2020, Preprints 2020030300.

Hemdan EED, Shouman MA, Karar ME. CovidNet: a framework of deep learning classifiers to diagnose covid-19 in x-ray images. 2020, arXiv:2003.11055.

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

Jnawali K, Arbabshirani MR, Rao N, Patel AA. Deep 3D convolution neural network for CT brain hemorrhage classification. In: Medical imaging 2018: computer-aided diagnosis. vol. 10575, International Society for Optics and Photonics; 2018, p. 105751C.

Singh SP, Wang L, Gupta S, Golli H, Padmanabhan P, Gulyas B. 3D deep learning approaches and applications. John Wiley & Sons; 2019, p. 6848–56.

Hemdan EED, Shouman MA, Karar ME. CovidNet: a framework of deep learning classifiers to diagnose covid-19 in x-ray images. 2020, arXiv:2003.11055.

Zhou X, Yamada K, Kojima T, Takayama R, Wang S, Zhou X, Hara T, Fujita H. Performance evaluation of FGCNet and 3D deep learning approaches for automatic segmentation of multiple organs on CT images. In: Medical imaging 2018: computer-aided diagnosis. vol. 10575, International Society for Optics and Photonics; 2018, p. 105752C.

Hasman MH, Jia W, He X, Kennedy P. Deep learning techniques for medical image segmentation: Achievements and challenges. J Dig It Imaging 2019;32(4):582–96.

Shoji Ali, Aliaraieez J, Babyn P. Automatic lung segmentation in CT images using watershed transform. In: IEEE international conference on image processing. vol. 2005, 2005, pp. II–1270.

Wang L, Li F, Li Q. Lung nodule detection and severity prediction of COVID-19 chest CT scans. 2021, arXiv:2101.00691.

Yip SS, Klannecz K, Naganawa S, Kim J, Studen A, Rivetti L, et al. Performance and robustness of machine learning-based radiomic COVID-19 severity prediction. 2020, medRxiv.

Mahmad T, Alum M, Chowdhury S, Ali SN, Rahman MM, Fattha SA, et al. CovidNet: A hybrid tri-level attention based network for lesion segmentation, disease diagnosis, and severity prediction of COVID-19 chest CT scans. 2021, arXiv:2101.00691.

Usmani AM, Jawad A, Al-Dhaheri A, Al-Riyami A, Al-Azouni K, Al-Belushi H, et al. COVID-19: a survey of machine learning approaches and applications. John Wiley & Sons; 2020.
[105] Cai L, Gao J, Zhao D. A review of the application of deep learning in medical image classification and segmentation. Ann Transl Med 2020;8(11).

[106] Fu Y, Lei Y, Wang T, Curran WJ, Liu T, Yang X. Deep learning in medical image registration: a review. Phys Med Biol 2020;65(20):20TR01.

[107] Ji Y, Zhang H, Zhang Z, Liu M. CNN-based encoder-decoder networks for salient object detection: A comprehensive review and recent advances. Inform Sci 2021;546:835–57.

[108] Dhruv P, Naskar S. Image classification using convolutional neural network (CNN) and recurrent neural network (RNN): A review. Mach Learn Inf Process 2020;367–81.

[109] Long Y, Gong Y, Xiao Z, Liu Q. Accurate object localization in remote sensing images based on convolutional neural networks. IEEE Trans Geosci Remote Sens 2017;55(5):2486–98. http://dx.doi.org/10.1109/TGRS.2016.2645610.

[110] Ozsalehin I, Sekeroglu B, Musa MS, Mustapha MT, Uzun Ozsalehin D. Review on diagnosis of COVID-19 from chest CT images using artificial intelligence. Comput Math Methods Med 2020;2020.

[111] Bhattacharya S, Madikunta PKR, Pham QV, Gadekallu TR, Chowdhary CL, Alzab M, et al. Deep learning and medical image processing for coronavirus (COVID-19) pandemic: A survey. Sustainable Cities Soc 2021;65:102589.

[112] Sarvamangala D, Kulkarni RV. Convolutional neural networks in medical image understanding: a survey. Evol Intell 2021;1:1–22.

[113] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60–88.

[114] Lu H, Wang H, Zhang Q, Yoon SW, Won D. A 3D convolutional neural network for volumetric image semantic segmentation. Procedia Manuf 2019;39:422–8.

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

[116] Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res 2014;15(1):1929–58.

[117] Yadav SS, Jadhav SM. Deep convolutional neural network based medical image classification for disease diagnosis. J Big Data 2019;6(1):1–18.

[118] Jiménez-Sánchez A, Albarquini S, Mateu D. 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; 2018, p. 150–60.

[119] Wang X, Peng Y, Lu L, Lu Z, Summers RM. 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, 2018. p. 9049–58.

[120] Cheplygina V, de Bruinie M, Pluim JP. Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Med Image Anal 2019;54:280–96.

[121] Arani E, Marzban S, Pata A, Zonooz B. Rgnet: A real-time general purpose semantic segmentation. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision, 2021. p. 3009–18.

[122] Kingma DP, Ba J. Adam: A method for stochastic optimization. 2014, arXiv: 1412.6980.

[123] Reddi SJ, Kale S, Kumar S. On the convergence of adam and beyond. 2019, arXiv:1904.09237.

[124] Géron A. Hands-on machine learning with scikit-learn, keras, and TensorFlow: concepts, tools, and techniques to build intelligent systems. O’Reilly Media; 2019.

[125] Ibrahim MR, Yousef SM, Fathalla KM. Abnormality detection and intelligent severity assessment of human chest computed tomography scans using deep learning: a case study on SARS-COV-2 assessment. J Ambient Intell Humaniz Comput 2021;1:1–24.

[126] Qiblawey Y, Tahir A, Chowdhury ME, Khandakar A, Kiranyaz S, Rahman T, et al. Detection and severity classification of COVID-19 in CT images using deep learning. Diagnostics 2021;11(5):1893.

[127] Demir U, Irmakli I, Keles E, Topcu A, Xu Z, Spampinato C, et al. Information bottleneck attribution for visual explanations of diagnosis and prognosis. 2021, arXiv:2104.02869.

[128] Mittal B, Oh J. CovNet: Covid-19 diagnosis using machine learning analyses for computerized tomography images. In: Thirteenth international conference on digital image processing. vol. 11878, International Society for Optics and Photonics; 2021, 1187816.

[129] Lacerda P, Barros B, Albuquerque C, Conci A. Hyperparameter optimization for COVID-19 pneumonia diagnosis based on chest CT. Sensors 2021;21(10):2174.

[130] Mahmud T, Alam MJ, Chowdhury S, Ali SN, Rahman MM, Fattah SA, et al. CovTANet: A hybrid tri-level attention-based network for lesion segmentation, diagnosis, and severity prediction of COVID-19 chest CT scans. IEEE Trans Ind Inf 2020;17(9):6489–98.