Internet of Things 14 (2021) 100377

Contents lists available at ScienceDirect

Internet of Things

journal homepage: www.elsevier.com/locate/iot

Deep learning assisted COVID-19 detection using full CT-scans

Varan Singh Rohila[a], Nitin Gupta[a,*], Amit Kaul[a], Deepak Kumar Sharma[b]

[a] National Institute of Technology Hamirpur, India
[b] Netaji Subhas University of Technology, Delhi, India

ARTICLE INFO

Article history:
Received 17 November 2020
Revised 1 February 2021
Accepted 14 February 2021
Available online 24 February 2021

Keywords:
COVID-19
Internet of Things
Medical imaging
Deep learning
CT-scan
Convolutional neural networks
Supervised learning

ABSTRACT

The ongoing pandemic of COVID-19 has shown the limitations of our current medical institutions. There is a need for research in automated diagnosis for speeding up the process while maintaining accuracy and reducing the computational requirements. This work proposes an automated diagnosis of COVID-19 infection from CT scans of the patients using deep learning technique. The proposed model, *ReCOV-101*, uses full chest CT scans to detect varying degrees of COVID-19 infection. To improve the detection accuracy, the CT-scans were preprocessed by employing segmentation and interpolation. The proposed scheme is based on the residual network that takes advantage of skip connection, allowing the model to go deeper. The model was trained on a single enterprise-level GPU. It can easily be provided on a network's edge, reducing communication with the cloud, often required for larger neural networks. This work aims to demonstrate a less hardware-intensive approach for COVID-19 detection with excellent performance that can be combined with medical equipment and help ease the examination procedure. With the proposed model, an accuracy of 94.9% was achieved.

1. Introduction

An outbreak of pneumonia caused by a novel coronavirus was first believed to be originated in December 2019 in Wuhan, China [1]. The virus was then officially named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the World Health Organisation (WHO). The WHO further declared this outbreak a pandemic on March 11, 2020. The transmission capacity of the SARS-CoV-2 is much stronger [2] as compared with the SARS-CoV outbreak in 2003. COVID-19 is a complicated disease that can happen in multiple modes and levels of severity varying from moderate to critical with the chance of organ malfunction and demise. The rapid increase in confirmed cases makes the prevention and control of COVID-19 extremely serious. As of 31st January 2021, more than 101.9 million cases have been reported across 218 countries and territories with more than 2.20 million deaths; more than 56.9 million people have also recovered [3].

The increasing number of COVID-19 positive cases and records of patients experiencing severe respiratory failure, heart complications, and other symptoms are substantial reasons to be extremely concerned about this viral infection's outcomes. COVID-19 spreads mainly through the medium of air and contact, transmitted in the form of small droplets produced by an infected person while sneezing and coughing. These droplets can land on the hands/cloths of nearby people, from there to eyes, nose, or mouth, and eventually reach the lungs. Touching the nose, mouth, or eyes after making contact with the surface exposed to the virus can result in an infection [4]. Although the virus's risk of spreading from an animal to a human

* Corresponding author.
E-mail address: nitin3041@gmail.com (N. Gupta).

https://doi.org/10.1016/j.iot.2021.100377
2542-6605/

is quite low, it can still spread in some situations, as reported by a few pet owners [5]. The people infected with the COVID-19 display symptoms similar to influenza-like illness and other respiratory ailments. The reported symptoms range from mild to severe symptoms. Fever, cough, shortness of breath, headache, diarrhea are a few that occur in people having the COVID-19 virus. These symptoms may take up to 2–14 days to appear after exposure to the virus [6].

Presently, there are several methods for the diagnosis of COVID-19. The standard procedure for testing is the real-time reverse transcription-polymerase chain reaction (rRT-PCR). Along with this laboratory testing, scans of the chest area in the form of X-rays and CT scans are proven to help its diagnosis [7]. Chest imaging in the patients have the multilobar involvements and rounded as well as peripheral airspace opacities. The chest CT scans are more sensitive to the ground-glass opacities, which could be missed in the X-ray scans [8]. Ground-glass opacity refers to an area of increased attenuation in the lung on CT or X-ray scan with preserved bronchial and vascular markings.

COVID-19's impact on sensitive populations, including the elderly, infants, and peoples with other comorbidities, is noticeable. The scale and the destructive potential of this virus has made it into a multidisciplinary issue. Experts from various fields such as epidemiology, the pharmaceutical industry, and diagnosis system modeling are working to combat the virus. Since there is no known cure for the virus, the number of people infected is rising day by day. Among the spread of the virus, a significant challenge of research and development has been witnessed in numerous science fields globally to decrease the developing trends of the disease's spread and fasten the diagnosis procedure. The increasing number of infected people poses a problem for the current screening procedures as they are overwhelmed by the number of tested people. Huge queues are formed to get a CT scan. Evaluation of the scan is time-consuming and puts an extra load on the doctors involved in quarantining and treating the infected patients. To overcome this problem, the process of diagnosis through CT scans can be fastened through state-of-the-art (SOTA) computer vision and deep learning techniques. These algorithms can be used to evaluate COVID-19 as they can identify the lesions and the varied opacities at a faster rate than the existing methods.

Recently, deep learning-based algorithms have bestowed enormous progress in medical data interpretation and automation due to their capability to elicit valuable features from the available medical datasets. The proposed work considers CNN, a specific class of neural networks having a grid-like structure for processing the data. Data could be a one-dimensional time series or a two-dimensional array of pixels, as found in the images. This type of neural net employs the convolution operation, which is a kind of linear operation. CNN uses convolution instead of standard matrix multiplication for calculation in at least one layer of the network [9]. CNN forms the basic building block of most of the deep and complex image-based deep learning algorithms. For training and maintaining such large algorithms, cloud-based architectures provide an efficient solution as they possess greater horsepower than average desktops. Still, it increases the communication overhead and computational latency considerably.

Recently Internet of Medical Things (IoMT) has made a big impact on the health care system. Generally, IoMT consists of interconnected medical devices sharing the data like pacemakers, CT-scan machines, and MRI [10]. Usually, a smart connected CT scanner sends scanned images of the lungs of the COVID-19 patients to the cloud server for further analysis and processing because these images require high-end GPUs to get the automatic results. However, transmitting data to the cloud and receiving the processed results is time-consuming, increasing latency. One of the objectives of this work is to design such a CNN that the models can be trained on a single enterprise-level GPU, which can easily be provided on the edge of the network, on the path between the data sources and the cloud data centers. This will help in reducing communication overhead and latency.

The contributions of the proposed work are as follows:

1. The proposed work provides the state-of-the-art solution to evaluate varying degrees of COVID-19 infection from CT-scans fuelled by techniques such as transfer learning, segmentation, and a robust preprocessing pipeline, which reduce the overall bias that may occur in datasets that are comparatively smaller in size.
2. We build the model's architecture to utilize comparatively less hardware and achieve good performance. Since the model can work efficiently on a single GPU, we can provide it on the network's edge rather than on the cloud.

The rest of the paper is organized as follows. The Section 2 discusses the recent works in the literature related to the COVID-19 detection using CT scan images. The Section 3 presents the proposed scheme to evaluate CT scan images. In Section 4, various methods used along with the proposed scheme are discussed, followed by the proposed scheme's performance evaluation and conclusions.

2. Related work

Recently, after the COVID-19 outbreak, few researchers have worked on COVID-19 detection using CT-scan images. Zhao et al. [11] prepared a data set of the CT-scan images of the COVID-19 patients. To address overfitting due to the small dataset size, the authors used data augmentation and transfer learning methods, along with their deep learning model. Later, Morozov et al. [12] published an extensive dataset of the full chest CT-scans of the patients with varying degrees of the COVID-19 infection aimed at felicitating research in this field. Further, this dataset is also used in our proposed approach. Wang et al. in [13] proposed a deep CNN named as COVID-Net, for detection of the COVID-19 using the chest X-ray images. The authors constructed an initial network prototype based on human-driven design principles. However, for the ease of training, the authors only considered four densely connected 1×1 convolution layers, which slightly reduces the accuracy.

The authors in [14] considered a simple CNN and applied a pre-trained AlexNet model on the X-rays and CT scan images dataset where the learned weights, bias, and features are transferred to the proposed approach. The authors considered a CNN model consisting of only one convolutional layer and 16 filters where each filter was constructed based on 5×5 filter size, batch normalization, and rectified linear unit (ReLU), and few other connected layers. Further, preprocessing like cropping and resizing is also performed on the images. Asmaa et al. in [15] adapted a deep CNN called Decompose, Transfer, and Compose (DeTraC) for the classification of the COVID-19 chest X-ray images. The authors proposed to add a class decomposition layer to the already trained models. This added layer partitioned each class into various sub-classes within the image dataset. Then, newly created sets were assigned new labels, and each subset was treated as an independent class. At last, these subsets were assembled back to produce the final predictions. However, it also suffered from the unavailability of the large dataset. The authors in [16–18], also used the CNN to detect the COVID-19 cases using the X-ray or CT-Scan images.

The majority of the works in computer-aided diagnosis of the COVID-19 infection focused on using the X-ray images of the infected patients as a primary dataset. Islam et al. in [19], presented a combination of CNN with Long Short Term Memory network (LSTM) for the detection of COVID-19. Their novel approach made it possible for them to achieve an AUC score of 99.9%. Similarly, Nour et al. [20] were able to achieve an accuracy of 98.97% using a deep CNN model combined with a Bayesian optimization algorithm for hyperparameter tuning. These works primarily focused on the differentiation between pneumonia and COVID-19, but not the infected patients. On the other hand, this proposed work focuses on detecting varying degrees of the COVID-19 infection for easy patient segregation and monitoring. The above works served as inspiration for our approach in terms of the model architecture and parameter tuning. Polsinelli et al. in [21] proposed a light CNN model based on SqueezeNet, a SOTA model, for the efficient discrimination of the COVID-19 CT scan images with the other CT scan images. Their light and efficient method helped transform the presented model's initial concept into a more robust yet memory-efficient model. The common shortage of large datasets among the works directed us to incorporate techniques like transfer learning and sampling techniques to compensate for datasets' size.

Unlike other works, this proposed work analyzed a full chest CT scan in one go for a much better analysis, emulating how a physician would perform the diagnosis. Most of the above-mentioned works require high computational resources, which may only be available at the cloud servers. The proposed work uses a single SOTA model as its backbone to increase the overall robustness without increasing hardware needs. Further, transfer learning, segmentation, efficient preprocessing, and hyperparameter tuning are combined with a unique architecture for scans classification. The framework used in this work, along with the CNN architecture, is explained in the following section.

3. Proposed approach

This section discusses the proposed CNN approaches used to identify COVID-19 cases through CT-scan images and details of proposed CNN model implementations. Traditionally, 3D CNN networks are incorporated in problems where the image dataset contains 3D images. These networks, however small, require a fair amount of data to churn out desirable performance. The dataset is not large enough to support a medium to high-depth 3D CNN model in the presented case. This problem is tackled by introducing a 2D CNN backbone that receives data in high volumes. It breaks down a sample into its 2D components, provides an output for each part, which is further aggregated and processed to arrive at an outcome for that sample.

The proposed work presents *ReCOV-101*, a deep CNN model with the ResNet-101 as its backbone. ResNet or residual network [22] solves the problem of the vanishing/exploding gradient using a technique called skip connections. The skip connection skips training from a few layers and connects directly to the output. The advantage of adding this type of skip connection is that if any layer effects the performance of the architecture, then it is skipped by regularization. In the proposed work, the classification layer of the ResNet-101 backbone is removed. The backbone model can be understood by an analogy of a physician analyzing full CT scan at once, providing importance to slices with relevant information. This backbone model is presented with slices of a single scan and it outputs a 3D tensor with 2048 channels for each slice which are then passed through a 1×1 convolution network with 512 filters for down sampling the number of channels. For merging these tensors, an addition layer is used. Before flattening this layer, a residual block of 3×3 convolution and 128 filters is added for more robustness. Finally, a 2D max pooling with 2×2 kernel size is added, and its output is flattened. For output, a softmax layer of four nodes is attached to the flattened nodes. The architecture of *ReCOV-101* is shown in Fig. 1. The dimensions at each step of the model are discussed in the next paragraph.

The model's input is a 4D tensor of shape (20, 224, 224, 3). The input tensor is split into 20 3D tensors, each of the shapes (224, 224, 3), which are input to the backbone model. After processing the tensors, the backbone model outputs 20 3D tensors of shape (7, 7, 2048), which are each passed through a 1×1 convolution network with 512 filters and then combined using addition operation. The tensor shape at this step is (7, 7, 512). This tensor is passed through a Residual Block, which outputs a tensor of shape (7, 7, 128), then passed through the max-pooling layer. The max-pooling layer outputs a (3, 3, 128) tensor. This tensor is flattened and fed into a densely connected layer of 1152 nodes. Finally, a four-node dense layer outputs the class probabilities.

For comparing and evaluating the above-mentioned approach, we choose a simple 3D CNN model as the baseline (Fig. 4). This baseline model consists of $3 \times 3 \times 3$ 3D convolution layer with 32 filters, followed by a 3D max pooling layer with $2 \times 2 \times 2$ kernel size. Another $3 \times 3 \times 3$ convolution layer with 64 filters is added followed by a max-pooling layer with $2 \times 2 \times 2$ kernel size. The output is then flattened and connected to a 256 nodes dense layer, which finally connects to a

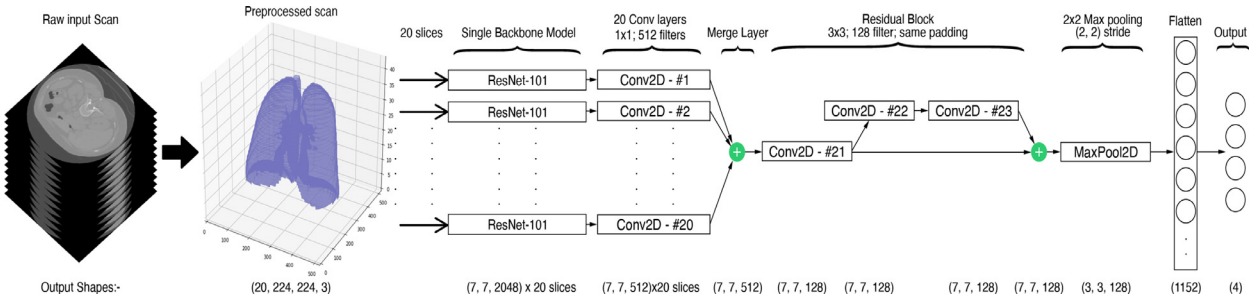

Fig. 1. ReCOV-101 Architecture.

Table 1
Data distribution between splits and classes.

Class	CT-0	CT-1	CT-2	CT-3	Total
Train	163	438	80	29	710
Validation	40	109	20	7	176
Test	50	136	25	9	220

softmax layer of four nodes. ReLU is used as the activation for all layers except the last layer. The next section describes the data and the preprocessing pipeline used to prepare the data.

4. Methodology

The partitioning of the dataset used in the proposed work based on the degree of severity of the infection is explained in this section. The preprocessing methods used to highlight the region of interest in the CT-scan images before feeding them to the network are also explained here.

4.1. Dataset

The dataset considered in the proposed work is *MosMedData: Chest CT Scans with COVID-19 Related Findings*. The dataset contains full chest CT scans of the 1110 patients, annotated into the five classes. The classes are as follows [12]:

a) **CT-0:** Normal lung tissue, no CT-signs of viral pneumonia.
b) **CT-1:** Several ground-glass opacifications, the involvement of the lung parenchyma is less than 25%.
c) **CT-2:** Ground-glass opacifications, the involvement of the lung parenchyma is between 25 and 50%.
d) **CT-3:** Ground-glass opacifications and regions of consolidation, the involvement of the lung parenchyma is between 50 and 75%.
e) **CT-4:** Diffuse ground-glass opacifications and consolidation as well as reticular changes in the lungs. Involvement of lung parenchyma exceeds 75%.

Some example scan slices are shown in Fig. 3. Each scan is an array of three-dimensional pixels called voxels describing the patient's chest radiodensity in the Hounsfield unit (HU). The attenuation of the X-ray beam is directly proportionate to the material density of the tissue. The Hounsfield unit is then calculated based on a linear transformation of the X-ray beam's baseline linear attenuation coefficient. Water is arbitrarily defined to be zero Hounsfield Units and air defined as -1000 HU[23]. Different scans contain a different number of slices, with 31 being the lowest and 71 being the highest number of slices. The width and length of the scans are consistent, both being 512 units. The varying number of slices poses an issue for the proposed model since the input shape is fixed. This problem is resolved in the next sub-section.

The classes CT-0 to CT-4 contain 253, 683, 125, 45, and 2 patient scans. The classes or categories are imbalanced, with CT-4 having the least number of occurrences. To tackle this issue, a combination of random oversampling and random undersampling is used in this work[24]. Since oversampling will not be fruitful for the category CT-4, it contains only two instances, which are 0.002% of the CT-1 class; therefore, this category is dropped for this research work. However, in the future, the *ReCOV-101* can be modified to accommodate this class using transfer learning when this class has enough instances.

Before sampling the data, the data is distributed into training, validation, and testing splits. Testing splits being 20% of the data from each class, validation being 20% of the remaining data, and training is the rest. The sampling is only applied to the training split to maintain the data's imbalance and authenticity during testing and validation. The detailed distribution of the data among classes and split is described in Table 1.

Firstly, the dominant class is randomly undersampled for sampling the training data, i.e., CT-1 to 70%. The remaining classes CT-0, CT-2, and CT-3 are randomly oversampled to make them 75%, 65%, and 55% of the dominant class, respectively.

Table 2
Sampling of training data.

Class	Before Sampling	After Sampling	% change
CT-0	163	229	+40%
CT-1	438	307	-30%
CT-2	80	199	+148%
CT-3	29	168	+479%

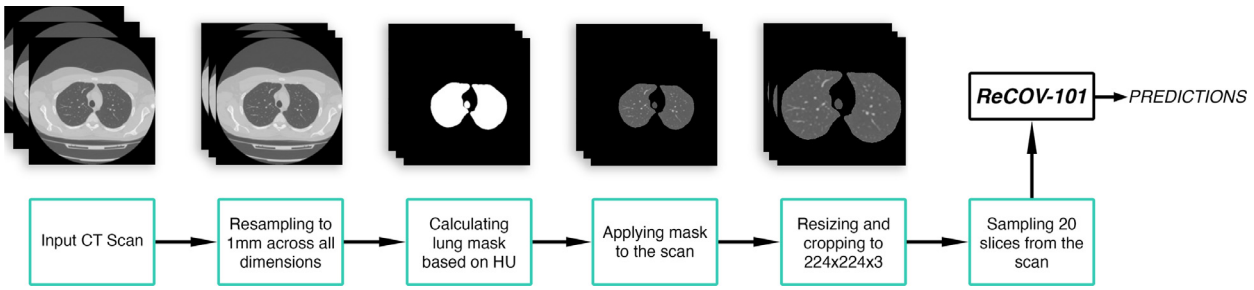

Fig. 2. Process Flowchart.

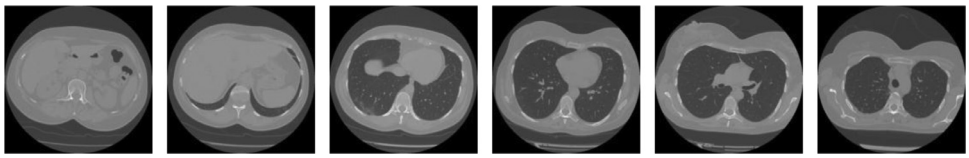

Fig. 3. Slices of a Patient Scan from the Dataset.

Finally, after sampling, the training data consists of a total of 903 scans. The detailed distribution among the classes after sampling is given in Table 2. The oversampling percentages of classes CT-2 and CT-3 might seem like an issue that could lead to poor results due to duplicate copies. Since the validation and the testing splits only contain unseen data, the accuracy achieved is justified. In the next section, the accuracy of the *ReCOV-101* between different classes is also studied to ensure the authenticity of the results.

4.2. Preprocessing pipeline

Each scan in the dataset was preprocessed before feeding into the model pipeline using various steps as shown in Fig. 2. The scans are in 3D NIfTI array format, with a varying number of slices and have different pixel spacing. The presented procedure handles all these discrepancies and outputs a 4D tensor with (slices, length, width, channels) as dimensions. A scan may have a pixel spacing of [2.5, 0.5, 0.5], which means that the distance between slices is 2.5 mm. For a different scan this may be [1.5, 0.725, 0.725]. A method of dealing with this is resampling the full dataset to a certain isotropic resolution. For the proposed approach, the scans are resampled to [1, 1, 1] pixels. This helps in making the scans invariant to slice thickness.

After intensive research and analysis, it was found that the portion of the scans excluding the lung area was not a determining factor of the degree of COVID-19 infection. If present, the areas excluding the lungs caused the model to have bias and affected the model's overall reliability. To overcome this issue, segmentation is used. Lung masks are calculated for each scan using the Hounsfield unit (HU) range of the lung tissue [25] as shown in Fig. 5a. Detailed segmentation process is discussed in the next sub-section. The masks are then applied to the scans, removing the unnecessary portions. However, dataset segmenting raises another issue. The segmented scan contains considerably large portions of no information or null pixel value. To reduce these portions, the scans are cropped. First, bounding box dimensions are obtained from the slice having the largest lung portion. Then, all the slices are cropped using those dimensions. This helps in the center positioning of the scan while also reducing the areas with no information.

ReCOV-101 requires each scan to be of a consistent dimension. The backbone of the model, ResNet-101, requires each slice to be of the shape 224×224 pixels with 3 channels. To achieve this, each slice of the scan is resized and repeated three times since they only contain a single channel. The scan values (in HU) are also mapped to a signed floating-point range of 256 for temporary storage. These are normalized on the fly during training. Further, the number of slices in each scan is limited to 20. These slices are randomly sampled from the full scan while maintaining their order, excluding a few slices from top and bottom which do not contain any significant lung information. The final shape of each scan is (20, 224, 224,

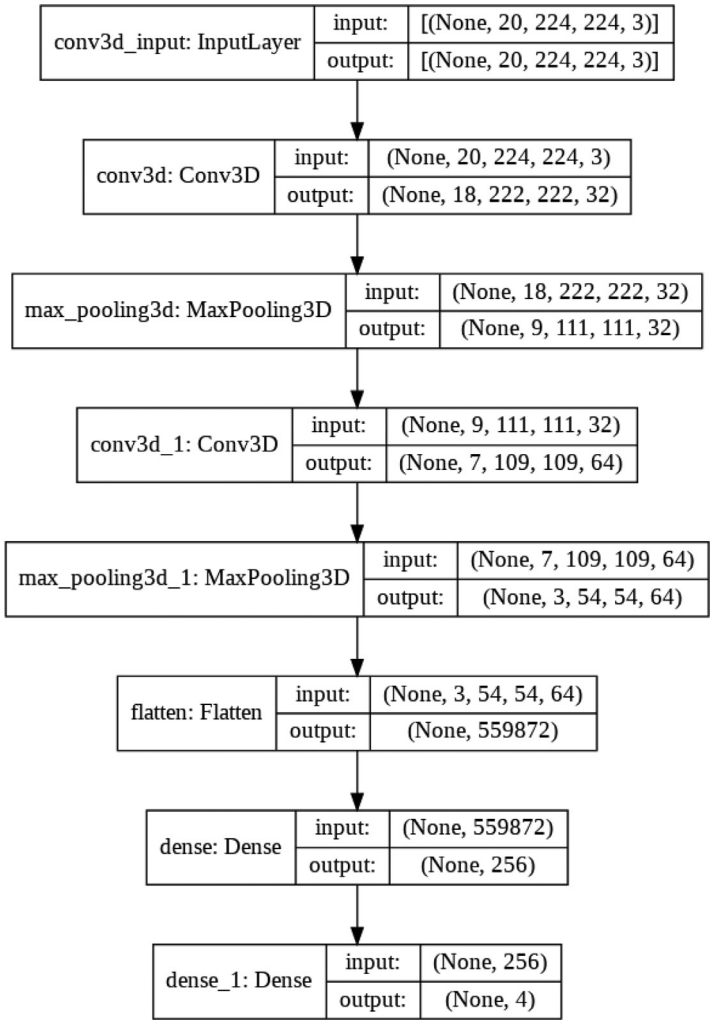

Fig. 4. Baseline 3D CNN model.

3). Since there are 903 scans for training, each having 20 slices, the total number of images that the ResNet-101 backbone receives amounts to be 18,060.

4.3. Segmentation

Segmentation is the process of dividing a digital image into multiple segments, which simplifies and makes the image's representation straightforward to analyze. Typically, segmentation refers to the process of location of objects and boundaries in digital photos. Additionally, segmentation helps the model to focus on a particular area of interest. Many datasets may contain various images with tiny lung portion. Maguolo et al. in [26] critically evaluate existing methods of COVID-19 detection. They discussed the possibility of achieving high accuracy scores with images that do not contain most of the lungs. Although the paper only considered X-ray images, a similar pattern can creep into CT scans also. To avoid such behavior, the lung portions are segmented from the CT scans using the HU range of lung tissue, region growing, and morphological operations. The typical HU value for lung tissue is about 700, and the air is 1000. Only separating lung tissue does not help since air is contained between the lungs. To differentiate between the air inside and outside the lungs, the lung tissue provides a boundary. Then, using the region growing algorithm, the air pocket inside the lungs is labeled. During this process, it is possible to label multiple air pockets since a few present inside the body. The final step in segmentation is to choose the largest region since the lung is the largest air pocket. The results of this process are shown in Fig. 5.

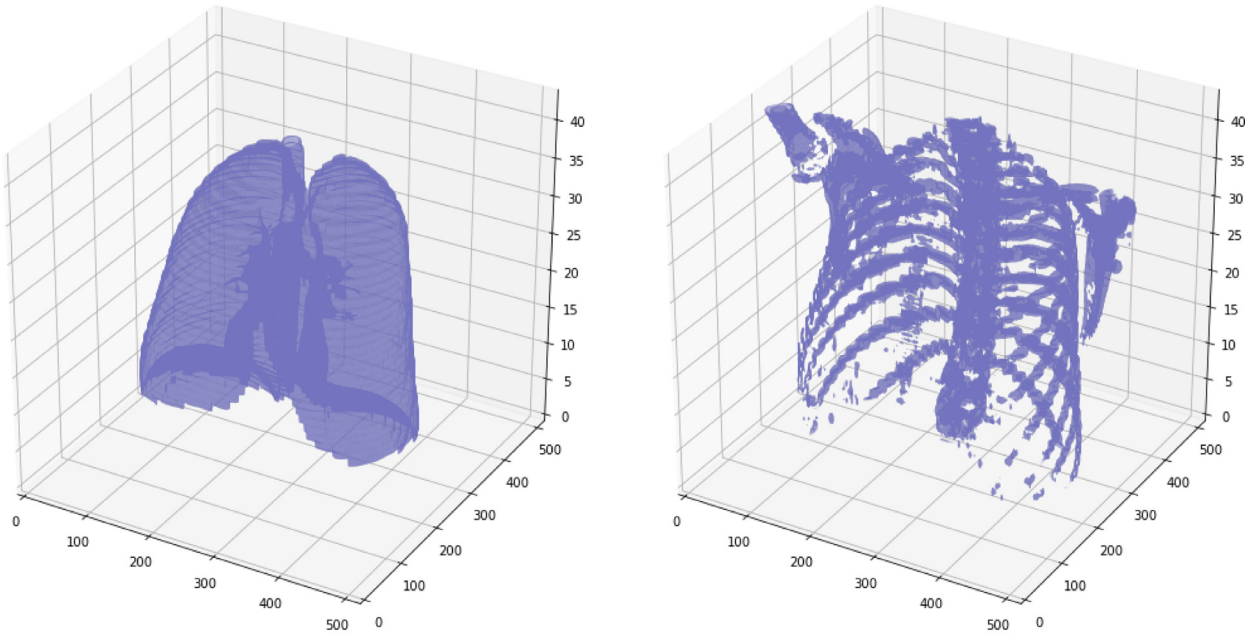

Fig. 5. Segmentation using HU Range of Tissues.

4.4. Transfer learning

Owing to the depth of *ReCOV-101*, the convergence of the gradient posed a problem. Convergence took a very long time or would plateau at a local minimum. This issue is removed by using transfer learning on the backbone models. They were pre-trained on the same dataset with the slices as raw input before inserting it in *ReCOV-101*. The slices were sampled from the middle portion of the scans to ensure maximum lung presence. The preprocessing procedure was kept the same to ensure consistency between training. In comparison, *ReCOV-101* performed much better with transfer weights than with randomly initialized weights. The pre-trained ResNet-101 achieved an accuracy of 92.4% on its testing.

4.5. Hyperparameter tuning

ReCOV-101 on itself performed better than many approaches aimed at COVID-19 detection from CT scans. To further enhance the model, various hyperparameters were tuned, such as the learning rate, optimizer, backbone model, merge layer, and the number of additional residual layers. An exponential learning rate scheduler [27] is used for learning rate, which keeps the initial learning rate constant for 10 epochs and then decreases it exponentially. The initial learning rate is chosen to be 0.0001. Pre-trained ResNet-50, ResNet-101, DenseNet-169 and DenseNet-201 were tested as the backbone models in which ResNet-101 performed the best. The optimum number of additional residual blocks, optimizer, and batch size for training were selected using grid search and cross-validation, which came out to be 1, Adam [28], and 64, respectively. Stochastic gradient descent and Adam are chosen as possible optimizers. The number of residual blocks can be one or two. Training for all settings was most stable for 64 as the batch size. The performance of each setting is evaluated in the next section.

For the merge layer, addition, average, and concatenation were considered. The addition layer and average layer are similar in implementation since they do not change the tensor dimensions. However, the concatenation layer altered the dimensions that needed slight modification in the architecture as considered in DenseNet. Out of the three, the addition layer outperformed others in accuracy and efficiency in every setting. All models compared in the next section contain addition layer as the merge layer.

5. Performance evaluation

The dataset considered for measuring the proposed approach's performance with accuracy, as the metric, and categorical cross-entropy, as the loss, is *MosMedData* [12]. To ensure that *ReCOV-101* does not overfit, regularisation and early stopping are used during training. First, we evaluate the baseline model and the models with different hyperparameters to arrive at the best performing combination.

Table 3
Hyperparameter search results.

Setting	ResNet-50	ResNet-101	DenseNet-169	DenseNet-201
adam-1	78.6	**94.9**	83.2	90.5
adam-2	82.4	88.8	82.4	84.5
sgd-1	77.5	91.4	86.5	89.1
sgd-2	79.2	86.5	87.3	85.4

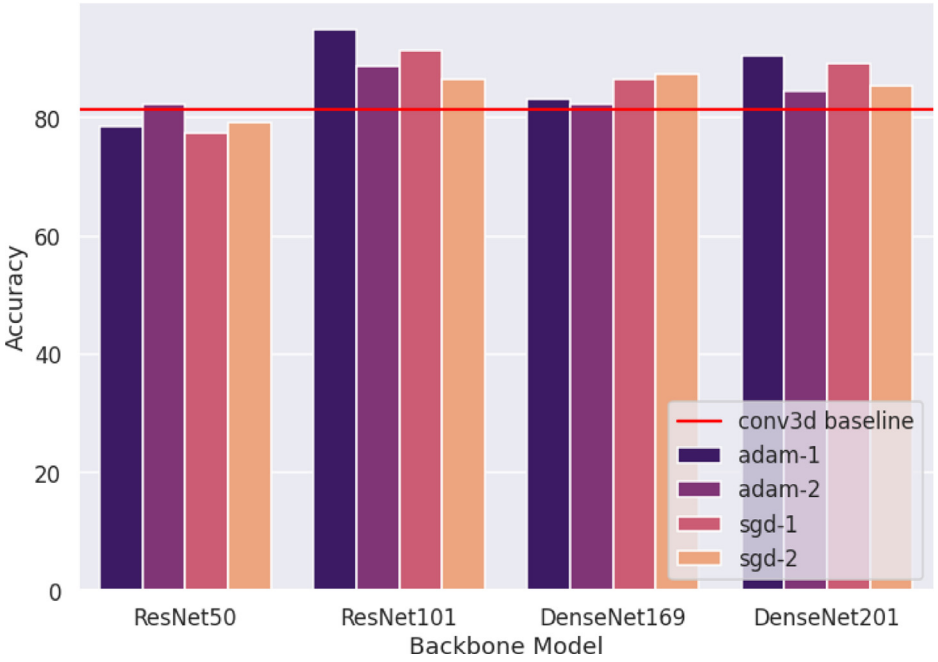

Fig. 6. Accuracies of Different Settings.

When trained with stochastic gradient descent, the baseline model did not achieve a decent accuracy even during training. The model's gradient was stuck in local minima and did not recover from it. When trained with Adam optimizer, the gradient converged, and training was successful. The baseline model hence was able to achieve an accuracy of 81.4%. The results of the hyperparameter search are presented in Table 3. The columns represent the chosen backbone model, and rows represent the optimizer hyphen the number of residual blocks. The models with two residual blocks generally performed poorly compared to models with one residual block. A departure from this pattern is observed when ResNet-50 is taken as the backbone model, where two residual blocks achieve better accuracy. Overall, both optimizers achieve similar accuracies, with adam performing slightly better in almost all cases. A more visual form of the search is presented in Fig. 6. The best performing model setting is with ResNet-101 as the backbone and one residual block, trained with adam optimizer. When referencing *ReCOV-101*, we refer to these settings. For a deeper understanding, *ReCOV-101*'s performance between different classes is also analyzed. To measure the impact of segmentation, training of *ReCOV-101* is also carried out separately for segmented and non-segmented data.

As shown in Fig. 7, for the segmented data, after about 50 epochs, the model achieved an accuracy of 88.7% on the testing data. Continuing the training, the testing accuracy kept on increasing up to about the 80th epoch. Further training increased the training accuracy to about 100%. The training was complete in about 100 epochs, which is quite early for such a large network. This anomaly could be understood by analyzing the dataset. A single example in the dataset consists of 20 images, which are all used to train the backbone model. Moreover, the backbone model was pre-trained on the same dataset separately. Hence, the overall convergence took a small number of epochs. The final accuracy achieved on the training, validation, and testing data were 99.8%, 95.2%, and 94.9%, respectively. The performance of the model in different classes is shown in Table 4 (rounded to nearest decimal). Comparatively, the model performs poorly in the CT-3 class with about 91.7% accuracy, which could be due to the lesser number of examples present in this class. The model was only shown about 30 scans from this class, although they were oversampled to about 168. Oversampling ensured that the model did not bias towards a particular class, but the lack of distinguishable examples hampered its performance. A similar pattern can be seen in the class CT-2, which also had comparatively fewer examples. This pattern is non-existent during training since class imbalance was compensated using sampling.

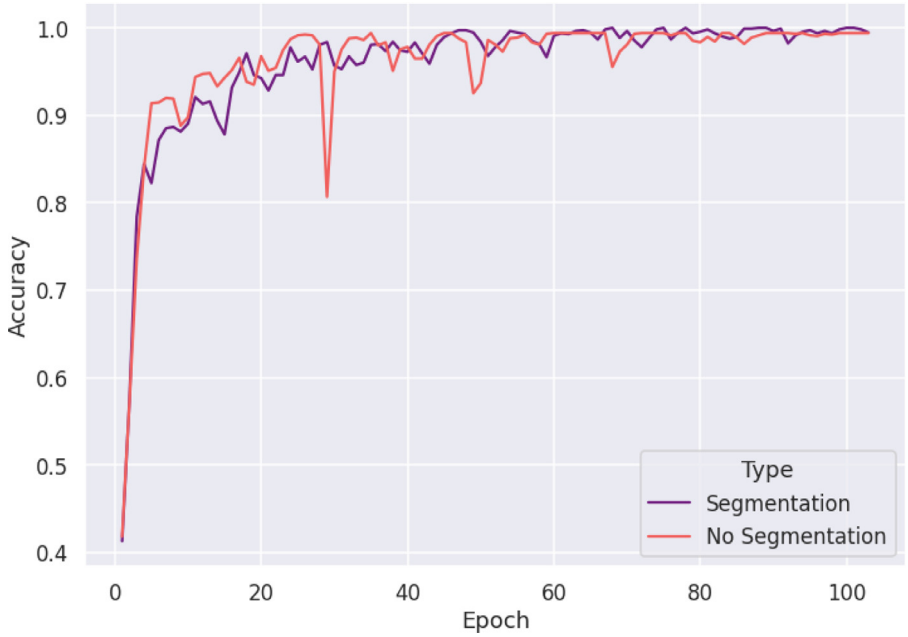

Fig. 7. Comparison of training process.

Table 4
ReCOV-101 Experimental Results (Accuracy in %).

	With Segmentation			Without Segmentation		
Class	Training	Validation	Testing	Training	Validation	Testing
CT-0	99.8	97.5	96.2	100	96.2	94.6
CT-1	100	96.9	98.6	100	94.8	92.4
CT-2	99.8	94.2	93.1	99.9	92.3	88.3
CT-3	99.6	92.2	91.7	100	90.7	85.1
Overall	99.8	95.2	**94.9**	99.9	93.5	90.1

On the other hand, *ReCOV-101* with non-segmented data was less stable during the training. It achieved an accuracy of 90.1% and similar accuracy distribution among classes. In this particular use case, detection without using segmentation can produce destructive results even if it seems to achieve decent results because the model could be fitting to patterns outside the lung area that do not indicate COVID-19 presence (small but considerable probability) [8]. Still, for the sake of research, the results from both approaches are included, and the training process is shown in Fig. 7.

6. Conclusions

In this work, *ReCOV-101*, a convolution-based model, has been presented and analyzed to detect COVID-19 from CT scan images. The work also proposes methods to preprocess full CT scans using a combination of techniques and later use them to develop a robust model. *ReCOV-101* was trained on a single enterprise-level GPU and is not hardware intensive as compared to other approaches. Further, other techniques such as data augmentation can also be incorporated to artificially increase the dataset's size with consultancy from a radiology specialist. Unsupervised Visual Representation, such as *Momentum Contrast* [29], can be used to further increase the models' performance. However, it is hardware intensive and requires two or more GPUs to run, which inhibits its usage to a few individuals and enterprises. For any deep learning model, the more data, the better the performance. In the future, when the database for COVID-19 research grows big enough, *ReCOV-101* has potential and is expected to outperform its currently presented form.

Declaration of Competing Interest

The authors declare that there is no conflict of interest.

References

[1] A. Deslandes, V. Berti, Y. Tandjaoui-Lambotte, C. Alloui, E. Carbonnelle, J. Zahar, S. Brichler, Y. Cohen, SARS-COV-2 was already spreading in France in late December 2019, Int. J. Antimicrob. Agents (2020) 106006.

[2] Y.-Y. Zheng, Y.-T. Ma, J.-Y. Zhang, X. Xie, Covid-19 and the cardiovascular system, Nat. Rev. Cardiol. 17 (5) (2020) 259–260.
[3] Who coronavirus disease (covid-19) dashboard, (https://covid19.who.int/). Accessed: 3rd Jan, 2020.
[4] Z.Y. Zu, M.D. Jiang, P.P. Xu, W. Chen, Q.Q. Ni, G.M. Lu, L.J. Zhang, Coronavirus disease 2019 (covid-19): a perspective from china, Radiology (2020) 200490.
[5] M.A. Shereen, S. Khan, A. Kazmi, N. Bashir, R. Siddique, Covid-19 infection: origin, transmission, and characteristics of human coronaviruses, J. Adv. Res. (2020).
[6] H.A. Rothan, S.N. Byrareddy, The epidemiology and pathogenesis of coronavirus disease (covid-19) outbreak, J. Autoimmun. (2020) 102433.
[7] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi, D. Shen, Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19, IEEE Rev. Biomed. Eng. (2020).
[8] W. Kong, P.P. Agarwal, Chest imaging appearance of covid-19 infection, Radiology 2 (1) (2020) e200028, doi:10.1148/ryct.2020200028.
[9] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[10] A. Limaye, T. Adegbija, Hermit: a benchmark suite for the internet of medical things, IEEE Internet Things J. 5 (5) (2018) 4212–4222.
[11] J. Zhao, Y. Zhang, X. He, P. Xie, Covid-ct-dataset: a ct scan dataset about covid-19, arXiv:2003.13865(2020).
[12] S. Morozov, A. Andreychenko, N. Pavlov, A. Vladzymyrskyy, N. Ledikhova, V. Gombolevskiy, I.A. Blokhin, P. Gelezhe, A. Gonchar, V.Y. Chernina, Mosmeddata: chest ct scans with covid-19 related findings dataset, arXiv:2005.06465(2020).
[13] L. Wang, A. Wong, Covid-net: a tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images, arXiv:2003.09871(2020).
[14] H.S. Maghdid, A.T. Asaad, K.Z. Ghafoor, A.S. Sadiq, M.K. Khan, Diagnosing covid-19 pneumonia from x-ray and ct images using deep learning and transfer learning algorithms, arXiv:2004.00038(2020).
[15] A. Abbas, M.M. Abdelsamea, M.M. Gaber, Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network, arXiv:2003.13815(2020).
[16] P. Afshar, S. Heidarian, F. Naderkhani, A. Oikonomou, K.N. Plataniotis, A. Mohammadi, Covid-caps: a capsule network-based framework for identification of covid-19 cases from x-ray images, arXiv:2004.02696(2020).
[17] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, et al., A deep learning algorithm using ct images to screen for corona virus disease (covid-19), MedRxiv (2020).
[18] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, arXiv:2003.10849(2020).
[19] M.Z. Islam, M.M. Islam, A. Asraf, A combined deep CNN-LSTM network for the detection of novel coronavirus (covid-19) using x-ray images, Inform. Med. Unlocked (2020) 100412.
[20] M. Nour, Z. Cömert, K. Polat, A novel medical diagnosis model for covid-19 infection detection based on deep features and Bayesian optimization, Appl. Soft Comput. (2020) 106580.
[21] M. Polsinelli, L. Cinque, G. Placidi, A light CNN for detecting covid-19 from ct scans of the chest, arXiv:2004.12837(2020).
[22] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, CoRR (2015). abs/1512.03385.
[23] D. TD, S. J., Hounsfield unit, (https://www.ncbi.nlm.nih.gov/books/NBK547721/). Accessed: 3rd Jan, 2020.
[24] N.V. Chawla, K.W. Bowyer, L.O. Hall, W.P. Kegelmeyer, Smote: synthetic minority over-sampling technique, J. Artif. Intell. Res. 16 (2002) 321–357.
[25] R. Wiemker, A. Zwartkruis, Optimal thresholding for 3d segmentation of pulmonary nodules in high resolution CT, in: International Congress Series, 1230, Elsevier, 2001, pp. 653–658.
[26] G. Maguolo, L. Nanni, A critic evaluation of methods for covid-19 automatic detection from x-ray images, arXiv:2004.12823(2020).
[27] Z. Li, S. Arora, An exponential learning rate schedule for deep learning, arXiv:1910.07454(2019).
[28] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv:1412.6980(2014).
[29] K. He, H. Fan, Y. Wu, S. Xie, R. Girshick, Momentum contrast for unsupervised visual representation learning, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 9729–9738.