 doi:10.20944/preprints202105.0605.v1

*Article*

# COVID-19 diagnosis from chest CT Scans: a weakly supervised CNN-LSTM approach

**Mustafa Kara[1], Zeynep Öztürk[1], Sergin Akpek and Ayşegül Turupcu [1,*]**

[1] Artificial Intelligence and Data Analytics, KoçDigital with BCG, Istanbul, Turkey; mustafa.kara@kocdigital.com (M.K); zeynep.ozturk@kocdigital.com (Z.Ö.);

[2] Department of Radiology, VKF American Hospital, Istanbul, Turkey; sakpek@amerikanhastanesi.org (S.A);

* Correspondence: aysegul.turupcu@kocdigital.com;

**Abstract:** Advancements in deep learning and availability of medical imaging data have led to use of CNN based architectures in disease diagnostic assisted systems. In spite of the abundant use of reverse transcription-polymerase chain reaction (RT-PCR) based tests in COVID-19 diagnosis, CT images offer an applicable supplement with its high sensitivity rates. Here, we study classification of COVID-19 pneumonia (CP) and non-COVID-19 pneumonia (NCP) in chest CT scans using efficient deep learning methods to be readily implemented by any hospital. We report our deep network framework design that encompasses Convolutional Neural Networks (CNNs) and bidirectional Long Short Term Memory (biLSTM) architectures. Our study achieved high specificity (CP: 98.3%, NCP: 96.2% Healthy: 89.3%) and high sensitivity (CP: 84.0%, NCP: 93.9% Healthy: 94.9%) in classifying COVID-19 pneumonia, non-COVID-19 pneumonia and healthy patients. Next, we provide visual explanations for the CNN predictions with gradient-weighted class activation mapping (Grad-CAM). The results provided a model explainability by showing that Ground Glass Opacities (GGO), indicators of COVID-19 pneumonia disease, were captured by our CNN network. Finally, we have implemented our approach in three hospitals proving its compatibility and efficiency.

**Keywords:** deep learning; computed tomography; image classification; COVID-19; medical image analysis; pneumonia; CNN, LSTM, medical diagnosis

 doi:10.20944/preprints202105.0605.v1

## 1. Introduction

The outbreak of a novel coronavirus, severe acute respiratory syndrome coronavirus 2, (SARS-CoV-2) started in Wuhan, China in December 2019 [1]. Afterwards, the aggressive human-to-human spread of the virus infected the entire world now widely referred to as COVID-19 (coronavirus disease 2019). The total number of reported cases is more than 110 million at the time of writing (Feb 22, 2021). [2] Although governments take drastic measures to fight the transmission of the virus, emerging variants are threatening progress [3].

Currently, there are two types of diagnostic tests used in the identification of COVID-19. These are molecular tests such as real-time reverse transcriptase polymerase chain reaction (RT-PCR), detecting virus's genetic material and antigen tests, detecting specific proteins from the virus. Although, RT-PCR test is done to patients showing clinical symptoms, the current tests have very high false negative rates resulting in the lack of treatment of COVID-19 patients and consequently failing to prevent transmission of the virus to others by these patients [4]. Therefore, as an alternative, medical doctors are utilizing chest imaging technologies for diagnosing COVID-19. Computed tomography (CT) of lungs is found to be a sensitive and accurate test for COVID-19 diagnosis [5,6]. Chest CT scans have further advantage of identifying patients who are at the very start of the symptoms and even identifying asymptomatic patients [7].

In addition to diagnosis challenges, the prevalence of the virus and ever-growing number of confirmed cases created a high-level demand of healthcare workers. The healthcare providers are facing intense workload due to the pandemic [8]. To relieve the overwhelming workload, AI systems are being used to detect and identify COVID-19 using medical imaging technologies [9-17]. Recent studies on radiology demonstrate promising results on COVID-19 pneumonia classification using chest CTs with the help of deep learning methodologies.

Although there are works proposing deep Convolutional Neural Networks (CNNs) without performing segmentation on CT slices [9,13], performing segmentation of CT slices is usually preferred as a preprocessing step before classification networks. The reason why the segmentation before classification is commonly preferred is that lung regions in CT images can easily be extracted with segmentation networks which are efficacious in removing redundant information and eliminating noise. In these related works, pre-trained well-known architectures such as U-Net [18], U-Net++ [19], and Deeplabv3 [20] are used for the segmentation task. As for the classification, deep CNNs, ResNext [21] based classifiers and LSTMs are utilized. There are different class definition approaches in the COVID-19 related studies like binary or multiclass. Binary classification is usually applied as Covid pneumonia and non-Covid pneumonia [12,17] or Covid positive and Covid negative[14,16]. Multiclass classification separates Covid negative further into Covid pneumonia, non-Covid pneumonia and no pneumonia [9-11,13,15]. For radiologists diagnosing hundreds of patients with COVID pneumonia and distinguishing Covid pneumonia and non-Covid pneumonia in a short time is challenging. Therefore, a multiclass classifier is beneficial for radiologists for making fast diagnosis.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 May 2021 doi:10.20944/preprints202105.0605.v1

Another important difference in the methods applied among the related studies is the use of slice-based or patient-level labeling. While most of the COVID-19 classification studies exploit slice-based CT image labels [10,12,13,15-17] because of its convenience, it has a big labour cost requiring manual labelling of slices by radiologists. Therefore, patient-level labelling has also been used [9,11,14] instead of slice-based labelling where slice-based features can be obtained with the help of deep learning methods. Notwithstanding, the studies making patient-level predictions apply restrictions on the input slice size of a single CT Volume causing information loss. These are done via random selection or interpolation.

At the beginning of the outbreak, most of the infected patients were diagnosed with pneumonia with unknown causes [22]. Resemblance of the chest CT scans of patients with COVID-19 and non-Covid pneumonia, made identification of the disease difficult. Therefore, it is crucial to distinguish between non-Covid pneumonia and pneumonia caused by COVID-19. In this study, we developed a multiclass classification model on chest CTs of the patients with COVID-19 symptoms. The dataset used in this study was provided by a Turkish hospital.

## 2. Materials and Methods

### *2.1. Dataset*

The dataset used in this study consists of 26935 CT images of 348 patients with varying numbers from 13 to 261 for each patient. For each individual CT slice image, the labels of the patients were used.The dataset was supplied by American Hospital and labelled by their expert radiologists. CT images were exported onto a local server from the CT machines via the picture archiving and communication systems (PACS) in DICOM format. All slices were converted to PNG format with adaptive histogram equalization.

In this study three subclasses were used as Covid Pneumonia (CP), Non-Covid Pneumonia (NCP) and healthy. Out of 348 patients, 270 were assigned to the training set and 78 to the test set with these three subclasses. The distribution of the dataset is given in Table 1. For each three subclasses CT slice images are represented in Figure 1 with original, manually segmented, AI-based segmented and final merged format.

**Table 1.** Distribution of CT image dataset in three studied cases.

| Image Dataset | Training | Testing |
|---|---|---|
| Covid19 pneumonia cases | 95 | 25 |
| Other pneumonia cases | 96 | 33 |
| Healthy cases | 79 | 20 |

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 May 2021 doi:10.20944/preprints202105.0605.v1

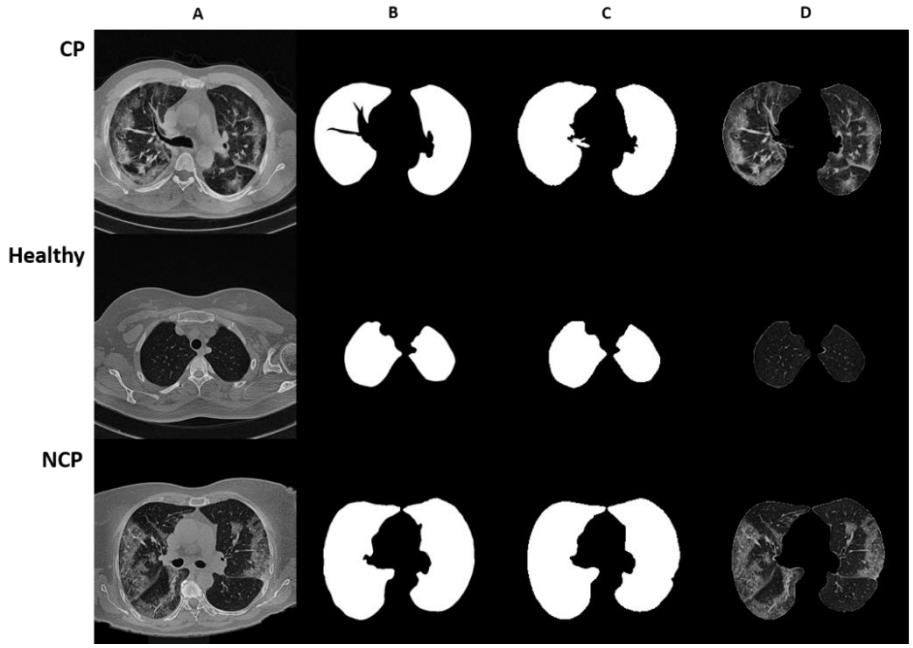



**Figure 1.** Representative images of lung CT segmentation. Rows are depicting CT slices of patients with Covid pneumonia, no pneumonia and non-Covid pneumonia, respectively. A) Original CT slice images; B) manually segmented CT slice images; C) AI-based segmentation of CT slice images; D) merged image of the segmentation mask and original CT slice images

*2.2. Preprocessing*

Spatial resolution of the CT images is 512 pixels and slice thicknesses are between 0.75 and 5.0 mm. In the image preprocessing step, CT images were scaled to 256x256 pixels. To eliminate redundant information in lung CT slices, CT slices were segmented and obtained masks were used for elimination of small lungs. The slices of each patient were grouped to have 50 slices in each group to feed the network.

2.2.1. Lung Segmentation

U-net [18] medical image segmentation architecture was used for CT slices segmentation. In this study U-Net architecture was trained from scratch since it is not designed for CT images. The segmentation helped to reduce the background information and assisted classifier model by focusing on the areas with pneumonia characteristics.

The segmentation model was trained with 301 manually annotated CT slice images. The validation set was assigned randomly from 10% of the annotated dataset. The model was trained for 30 epochs achieving 0.9681 Dice-Coefficient Index proving an accurate segmentation. Learning rate of the segmentation model was set to 2e-5. For data augmentation, rotating, flipping, blurring and brightening were used. The generalization of the model was increased by applying augmentation in random range with randomly selected augmentation type through a data generator.

2.2.2. Segmentation Based Elimination

The areas containing no lung information were eliminated. CT slices having small lung regions were also eliminated since those were not supplying valuable information on pneumonia characteristics. The original CT slices were masked by the ratio obtained

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 May 2021 doi:10.20944/preprints202105.0605.v1

by the maximum CT lung area. The threshold values of 0.4 for up-side of lung and 0.9 for down-side of lung were used for masking. The segmentation mask was applied to those having a scaled CT lung area smaller than the defined thresholds.

#### 2.2.3. Slice Grouping

In this study, we adopted a patient-based labelling without imposing any restrictions on input size to not have an information loss. Patient-based labelled dataset requires the number of slices in a group which has to be defined considering the number of occurrences of pneumonia characteristics. We grouped CT images since each slice can contribute to pneumonia characteristics, We set the group size to be 50 CT slices since an exact number of slices needs to be fed to the network. In the studied dataset, the number of slices ranges from 13 to 261 as mentioned in Section 2.1. Therefore, in case the number of slices for a patient was not divisible by 50; when there were more than 25 slices remaining, the last group of slices were augmented with padding of black images, when there were less than 25 slices remaining then the amount of slices in the last two groups were downsampled to 50 using interpolation.

### *2.3. Deep Learning Model*

In this study, CNNs were used to obtain feature maps of CT slices and bidirectional LSTMs were used for classification. This framework was applied to CT Volumes consisting of multiple CT slices for each patient. CT groups generated from CT slices for each patient were fed to the bidirectional LSTM. Then, segmented and eliminated CT groups of each patient were processed by the framework.

Two separate steps of training were performed in an end-to-end pipeline. Initially, the slices were fed to the proposed Convolutional Neural Network block for extracting spatial features. Once the spatial features were extracted from each slice, the feature maps were given to the bidirectional LSTM. The bidirectional LSTM makes the classification for the CT group by exploiting the axial dependency in the input slices and transforming the spatial features to axial features. Finally, label-based majority voting was performed for classification outputs of multiple CT groups for a particular patient. The proposed framework was motivated by detection of violence in video frames by Hanson et al. [23].

The proposed framework was trained on a machine with a 4 NVIDIA Tesla V100 GPU and implemented on the Python programming language with Keras library with Tensorflow backend.

#### 2.3.1. Convolutional Neural Network

The proposed architecture was formed of 4 convolutional layers with ReLU activation function whose outputs were concatenated after applying Global Max Pooling operations. The architecture of the network is shown in Figure 2. This concatenation ensures that features from all layers are used equally.

Most of the related studies use pre-trained classification models [9,10,14,24]; however, these architectures are not particularly designed for classification of CT Volumes. The architectures such as ResNet [25] and Inception [26], are designed for ImageNet dataset [27] where the images contain a wide variety of objects. For the ImageNet classification task, a large set of features must be learned for inference, hence

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 May 2021 doi:10.20944/preprints202105.0605.v1

deeper networks need to be utilized. On the other hand, CT Volumes can be inferred with smaller feature maps in comparison to ImageNet. The CNN encoder architecture is designed to be more efficient in terms of both memory and runtime inference.

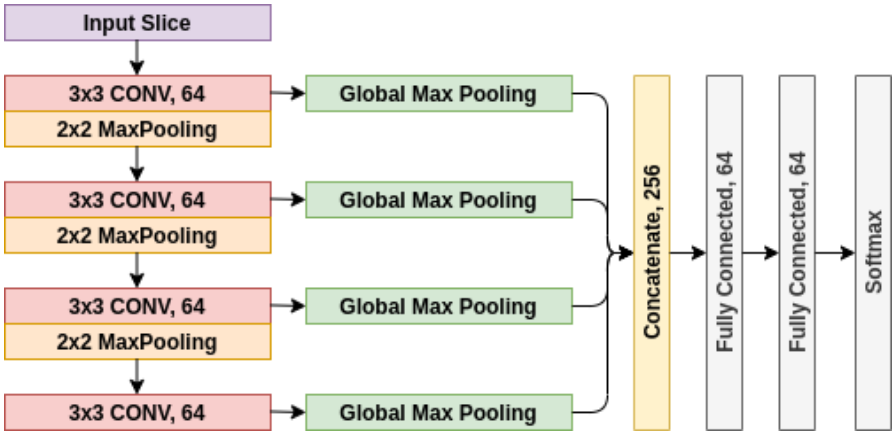



**Figure 2.** CNN block architecture for learning feature maps.

The main purpose of training these CNN blocks is to obtain feature maps of ground glass opacities (GGO) which are prevalent in Covid pneumonia [5,28]. After training, the infection feature maps belonging to each slice were fed to the bidirectional LSTM to maintain the relationship between slices.

All layer weights were initialized randomly. The CNNs were trained for 100 epochs with an early stopping set to 25 epochs to avoid overfitting. Learning rate is set to 1e-4. The categorical cross-entropy loss function was used to calculate the loss between predictions and ground-truth labels. Rotating, flipping, blurring and brightening were used as data augmentation techniques. Data generator at each step augments data in random range with randomly selected augmentation type to increase generalization of model.

2.3.2. Bidirectional LSTM based classifier

The concatenation of the output of CNN blocks for each CT slice was fed to bidirectional LSTM. Sequence model was used to emphasize the relationship between spatial feature maps of multiple and ordered CT slices. Since this relationship is both forward and backward, a bidirectional sequence model was used. The output of each bidirectional LSTM cell was preserved and flattened. It is assumed that each cell output has equal importance given the fact that spatial feature maps are agnostic in terms of the amount of infection in the lungs. The aggregated features obtained from bidirectional LSTM were then transferred to a dense layer with dropout. The obtained output of the final dense layer is a multiclass classifier of three classes for each CT group of a single patient. The architecture of the network is shown in Figure 3.

CNN blocks (explained in detail in section 2.2.2) were initialized with the weights of feature extracting layers. CNN weights were frozen and the rest were initialized randomly. The classifier was trained for 200 epochs with an early stopping set to 25 epochs

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 May 2021 doi:10.20944/preprints202105.0605.v1

to avoid overfitting. Same parameters in the CNN were used for biLSTM (see section 2.3.2).

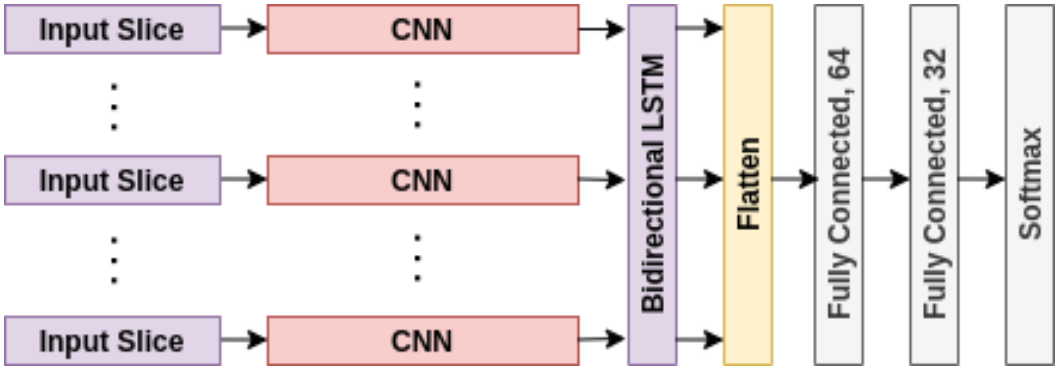



**Figure 3.** Illustration of the CNN-biLSTM classifier architecture. End-to-end pipeline for classification where segmented slices are fed to CNN blocks and then CNN outputs are processed by bidirectional LSTM. Each state output is preserved, and every output is combined prior to dense layers and softmax layer.

*2.4. Performance and Assessment*

In the case of patients having more than one CT group, label-based majority voting was applied. First, the label with the highest classification probability for each group was selected. Next, the label having the majority vote was assigned to the label of the patient. When the output was healthy and the group was classified as NCP or CP, the probability of non-healthy class was checked further. When the probability was greater than the predefined threshold (0.9), then the label changed in favor of the non-healthy class.

Accuracy, specificity, sensitivity and f1-score were calculated from the label-based majority voting method outputs since the infections of pneumonia might not be apparent in all slices of a patient. The flowchart for the algorithm is illustrated in Figure 4.

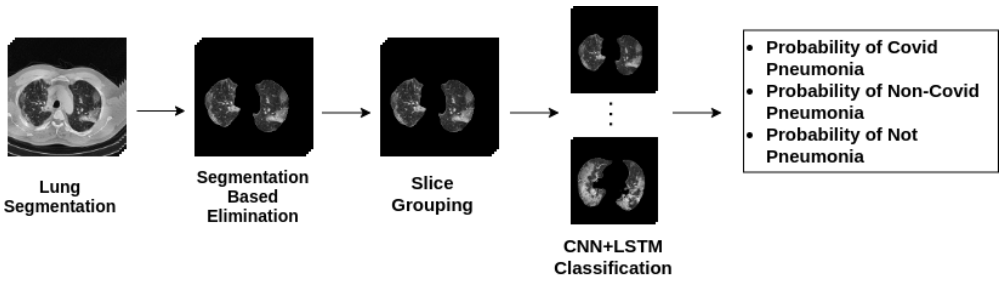



**Figure 4.** Flowchart diagram for the classification algorithm.

**3. Results**

CT images, collected from PACS server, were preprocessed then fed to the CNN + biLSTM network. CNN blocks were used for obtaining feature maps of ground glass opacities (GGO), then feature maps were fed to the biLSTM network which maintained the relationship between spatial feature maps of multiple and ordered CT slices. The framework was designed to be efficient in runtime inference. As a result of this approach,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 May 2021 doi:10.20944/preprints202105.0605.v1

average inference runtime for each CT group in our test set is 0.106 second with a single NVIDIA RTX 2060 GPU and 0.409 second with a Intel i7-9750H CPU. The proposed network performance was evaluated on an independent test set. Accuracy, specificity, sensitivity and f1-score for each class were calculated and summarized in Table 2.

On an independent test dataset, we achieved high specificity (98.3%) and high sensitivity (84%) rates in detection of COVID-19. Moreover, we obtained high specificity (96.2%) and high sensitivity (93.9%) rates in detection of non-COVID-19 pneumonia as well (Table 2). The areas under receiver operating characteristics for COVID-19 and non-COVID-19 pneumonia were 0.90 and 0.91, respectively (Figure 6a).

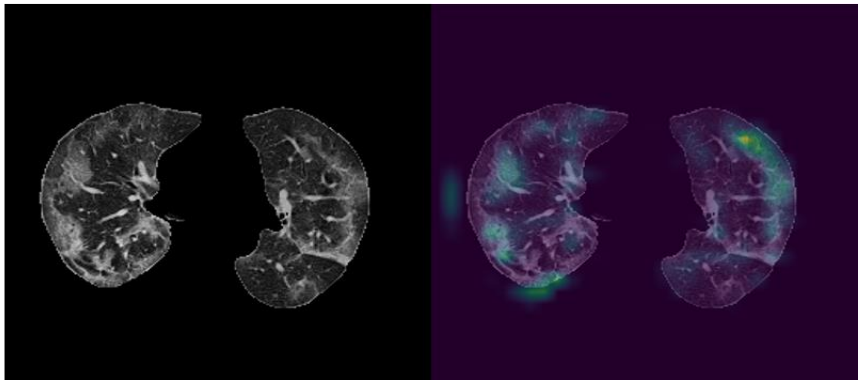
CNN blocks of the proposed network were utilized to obtain feature maps of ground glass opacities (GGO). To illustrate the effectiveness of the approach used in this study, the last class activation map (CAM) was visualized with the help of Grad-CAM (Figure 5). As it can be seen from Figure 5, features were extracted mostly from ground glass opacities.

The same algorithm (Figure 4) was applied for each patient in the test set and represented. By comparing the patients' ground-truth labels, ROC curves and confusion matrix are plotted as shown in Figure 6. In the ROC curve we obtained a macro average AUC value as 0.92. Table 2 shows the classification results of the patients. Our network classifies NCP patients more accurately (93.9%) than other labels and its overall accuracy is 89.7%.

Specificity and sensitivity are mostly used for evaluating diagnosis applications. Our model's specificity value on CP patients is 98.3% and NCP patients is 96.2%. These values show that our model can classify pneumonia affected patients efficiently.

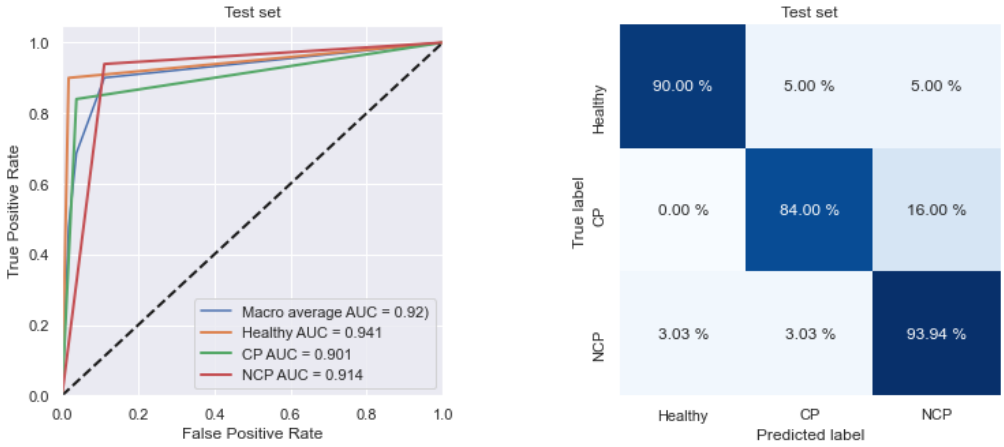
**Table 2.** Performance metrics (accuracy, specificity, sensitivity, f1-score) for each class of test set.

| | **accuracy** | **specificity** | **sensitivity** | **f1-score** |
|---|---|---|---|---|
| CP | 0.840 | 0.983 | 0.840 | 0.875 |
| NCP | 0.939 | 0.962 | 0.939 | 0.899 |
| Healthy | 0.900 | 0.889 | 0.900 | 0.923 |
| Average | 0.897 | 0.949 | 0.893 | 0.899 |



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 May 2021 doi:10.20944/preprints202105.0605.v1

**Figure 5.** Visualization of CNN outputs. Left: Cropped CT image, right: visualized feature map of CNN output. The right image shows CNN blocks are capable of capturing Ground Glass Opacities in CT images.




(**a**) (**b**)

**Figure 6.** Classification performance represented with (a) Receiver operating characteristics (ROC) curves. (b) Normalized confusion matrix. In the ROC curve, the blue line denotes the macro-average area under the curve (AUC).

## 4. Discussion

In this work, we designed and implemented a two-staged deep learning model for the diagnosis of COVID-19 disease from chest CT volumes. Our model was able to assist the radiologists with its performance of high specificity (CP: 98.3%, NCP: 96.2%) and high sensitivity (CP: 84%, NCP: 93.9%) rates in detection of COVID-19 pneumonia and non-COVID-19 pneumonia. This study offers a reduced time and effort in diagnosing CP and NCP as well as distinguishing healthy patients. This work was especially motivated to be easily implemented and used in COVID-19 pandemic to manage overwhelming workload.

We were able to collect a large number of chest CT scans from American Hospital in Turkey. The dataset included a total of 26935 CT scans belonging to 348 distinct patients. We used a patient-based labelling due to its effectiveness both in cost and time. Although slice-based labelling is advantageous for deep learning methods, these slice-based features can be obtained via CNNs. The use of slice based, or patient-based labeling was among the differences in the model architectures used in the related studies (see Table 3 for all compared features). We implemented a grouping method for the varying numbers of slices in the CT Volumes dataset. There are other methods like performing interpolations on CT Volumes or random selection of slices of a CT Volume; however, these cause information loss. Another advantage of the grouping method was exploiting each single slice of every patient. Since we grouped the CT slices of the patients, the evaluation of the class probabilities of groups was done by utilizing the majority-voting method.

 doi:10.20944/preprints202105.0605.v1

In Table 3, we report a summary of the previous works that are focusing on multiclass classification on chest CT scans distinguishing COVID-19 pneumonia from community-acquired pneumonia and healthy patients. Only exception is the study from El-bana et al. [24] where they are compartmentalizing the community-acquired pneumonia by viral and bacterial. In this table we compared the related studies with our work in terms of number and types of classes, sample size as well as the architectures used for segmentation and feature extraction tasks. Furthermore, the use of pre-trained models was examined for both segmentation and feature extraction methods. The studies using sequence models to relate CT scans in a CT volume was also noted. Unlike other studies using prominent segmentation architectures, in this study, we trained on our dataset of chest CT scans since U-Net is not trained on chest CT scans.

We designed an efficient architecture to be readily used in every hospital. Since deeper architectures require high memory and high computing power, not every hospital has a suited infrastructure to fulfill those requirements. Our model utilizes inference task with moderate memory and computational power (0.409 seconds average runtime for each CT group on CPU). This work has several limitations. Foremost, we used a single dataset from a single hospital since this study was an immediate response to an urgent need. Correspondingly, this dataset is collected from a particular type of CT machine. It would be desirable to work with other datasets from other hospitals. Hence, we are in the process of collaborating with other hospitals. Nonetheless, we believe our study is a compatible and fast response to the hospitals in need. Our proposed framework can be used to help relieve the burden on radiologists as well as reduce the time to diagnose COVID-19 pneumonia and non-COVID-19 pneumonia. In addition, our framework can decrease misdiagnosis of CP and NCP thanks to its high specificity and sensitivity rates.

Although explainability and interpretability are major concerns for deep learning models, there have been significant efforts in explaining the decisions of deep learning models used for medical diagnosis. Since the doctors are going to make decisions for patients relying on the AI system, it is useful to understand why the AI system makes one choice over another. There are many methods developed to explain deep learning methods trained on medical images [29], such as Grad-CAM [30]. With this in mind, we visualized the feature maps generated via Grad-CAM to provide insights on our framework.

## 5. Conclusions

Fast and accurate diagnosis of COVID-19 forms a challenge for all countries. The available tests for fast diagnosis either have sensitivity or accuracy problems. In this regard, chest CT scans found to be very effective in identifying COVID-19 patients. This method showed a superior practice over other methods by its power in identifying COVID-19 patients even who are at the very early stages of the disease or patients with no symptoms. In addition to this, there is a pronounced level of confirmed cases causing an additional demand of healthcare workers.

We built an AI system that can accurately differentiate COVID-19 pneumonia patients and non-COVID-19 pneumonia patients. The framework is implemented and

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 25 May 2021 doi:10.20944/preprints202105.0605.v1

ready for use to assist radiologists. Our two-staged deep learning classifier enables us to identify CP, NCP and healthy patients from chest CT volumes. First, this framework is capable of learning pneumonic features in CT scans. Second, it can classify COVID-19 pneumonia and non-COVID-19 pneumonia with 84% and 93.9% accuracy. Our model achieves high specificity for CP: 98.3% and NCP: 96.2% and high sensitivity for CP: 84% and NCP: 93.9%. We have further analyzed the area under receiver operating characteristics curve (ROC-AUC) to demonstrate the effectiveness and efficiency of our proposed framework. We provide an AI system for automated, fast and accurate diagnosis of COVID-19 using medical imaging technologies. Along with recent efforts in radiology we demonstrate a promising COVID-19 pneumonia classification using chest CTs augmented with the help of deep learning methodologies. We showed that our system is assisting the radiologists. It is implemented in real-time in selected hospitals through a user interface.

**Table 3.** Comparison of model architectures used in the related studies and in this study.

| Study | Target Specifications | Patient-level labelling | Total # of images | Segmentation model | Feature Extraction Model | Sequence Model | Segmentation-Feature Extraction Pretrained |
|---|---|---|---|---|---|---|---|
| Amyar 2020 | CP [1], Healthy, Other | no | 1,369 | - | Multitask U-net | - | N/A-no |
| Li 2020 | CP, NCP [2], Healthy | no | 4,352 | U-net | Resnet50 | - | yes-yes |
| Zhang 2020 | | yes | 617,775 | Deeplabv3 | 3d Resnet-18 | 3d Resnet-18 | no-no |
| Wang 2020 | | no | 104,009 | - | Conv2d | - | N/A-yes |
| Hasan 2020 | | no | N/A | Image processing | CNN & QDE | LSTM | N/A-no |
| Lee 2021 | | yes | 436,265 | - | Inception 3d | Inception 3d | N/A-yes |
| This work | | yes | 26,935 | U-net | Conv2d | biLSTM | no-no |
| El-bana 2020 | CP, PV [3], PB [4], Healthy | no | 100 | - | Inception v3 | - | N/A-yes |

[1] Covid Pneumonia, [2] Non-Covid Pneumonia, [3] Pneumonia Virus, [4] Pneumonia Bacteria

As a next step, we showed that our framework is able to detect Ground Glass Opacities (GGO) in patients suffering from COVID-19 pneumonia via visualizations using Grad-CAM algorithm [30]. Image pixel attributions can be visualized upon calculating the gradient from the output to a given deeper layer. Grad-CAM reconstructs maps as a weighted combination of forward neuron activation, with weights based on global average pooling and backpropagation outputs to a target layer. This way COVID-19 patients can be distinguished from the others by providing demonstration of the proposed

 doi:10.20944/preprints202105.0605.v1

models with pneumonia indicating top features. Note that this framework not only identifies CP but also NCP owing to its capability to generate two-dimensional feature maps. This can be easily further compartmentalized for additional pneumonia types according to needs of the hospitals and healthcare workers. Overall, our work offers an easily adaptable framework which has been already implemented and being used. As such, it offers implementation in numerous hospitals with low computing powers

**Author Contributions:** M.K. conceptualized the framework, developed the methodology, implemented and validated the framework. S.A. provided the curated dataset. Z.O. made the literature review and wrote the manuscript and A.T. reviewed and edited it. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** This work was supported by the KoçDigital. The authors thank Dr. Baturalp Güner from American Hospital for providing the labelled data for this project. We thank Huawei Turkey for providing the servers. The authors gratefully acknowledge the support and guidance of Acar Erdinç on development of the architecture. Also the authors thank Kerem Ozkarakas for supporting the framework development.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

**Abbreviations**

The following abbreviations are used in this manuscript:

AI: Artificial Intelligence
AUC: Area Under Curve
biLSTM: bidirectional Long Short Term Memory
COVID-19: Coronavirus Disease 2019
CP: COVID-19 pneumonia
CNN: Convolutional Neural Network
CT: Computed Tomography
DICOM: Digital Imaging and Communications in Medicine
DL: Deep Learning
Grad-CAM: gradient-weighted class activation mapping
GGO: Ground Glass Opacities
PACS: Picture Archive and Communication System.
ROC: Receiver Operating Characteristic
RT-PCR: Reverse Transcription Polymerase Chain Reaction
NCP: non-COVID-19 pneumonia

## References


1. Hui, D.S.; E, I.A.; Madani, T.A.; Ntoumi, F.; Kock, R.; Dar, O.; Ippolito, G.; McHugh, T.D.; Memish, Z.A.; Drosten, C., et al. The continuing 2019-nCoV epidemic threat of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. *Int J Infect Dis* **2020**, *91*, 264-266, doi:10.1016/j.ijid.2020.01.009.
2. WHO Covid-19. Available online: https://covid19.who.int/ (accessed on 22 February).
3. CDC Covid Data Tracker Weekly Review. Available online: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html (accessed on February 23).
4. Xiao, A.T.; Tong, Y.X.; Zhang, S. False negative of RT-PCR and prolonged nucleic acid conversion in COVID-19: Rather than recurrence. *J Med Virol* **2020**, *92*, 1755-1756, doi:10.1002/jmv.25855.
5. Ai, T.; Yang, Z.; Hou, H.; Zhan, C.; Chen, C.; Lv, W.; Tao, Q.; Sun, Z.; Xia, L. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* **2020**, *296*, E32-E40, doi:10.1148/radiol.2020200642.
6. Panwar, H.; Gupta, P.K.; Siddiqui, M.K.; Morales-Menendez, R.; Bhardwaj, P.; Singh, V. A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. *Chaos Solitons Fractals* **2020**, *140*, 110190, doi:10.1016/j.chaos.2020.110190.
7. Shi, H.; Han, X.; Jiang, N.; Cao, Y.; Alwalid, O.; Gu, J.; Fan, Y.; Zheng, C. Radiological findings from 81 patients with COVID-19 pneumonia in Wuhan, China: a descriptive study. *Lancet Infect Dis* **2020**, *20*, 425-434, doi:10.1016/S1473-3099(20)30086-4.
8. Moazzami, B.; Razavi-Khorasani, N.; Dooghaie Moghadam, A.; Farokhi, E.; Rezaei, N. COVID-19 and telemedicine: Immediate action required for maintaining healthcare providers well-being. *J Clin Virol* **2020**, *126*, 104345, doi:10.1016/j.jcv.2020.104345.
9. Lee, E.H.; Zheng, J.; Colak, E.; Mohammadzadeh, M.; Houshmand, G.; Bevins, N.; Kitamura, F.; Altinmakas, E.; Reis, E.P.; Kim, J.K., et al. Deep COVID DeteCT: an international experience on COVID-19 lung detection and prognosis using chest CT. *NPJ Digit Med* **2021**, *4*, 11, doi:10.1038/s41746-020-00369-1.
10. Li, L.; Qin, L.; Xu, Z.; Yin, Y.; Wang, X.; Kong, B.; Bai, J.; Lu, Y.; Fang, Z.; Song, Q., et al. Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy. *Radiology* **2020**, *296*, E65-E71, doi:10.1148/radiol.2020200905.
11. Zhang, K.; Liu, X.; Shen, J.; Li, Z.; Sang, Y.; Wu, X.; Zha, Y.; Liang, W.; Wang, C.; Wang, K., et al. Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography. *Cell* **2020**, *181*, 1423-1433 e1411, doi:10.1016/j.cell.2020.04.045.
12. Chen, J.; Wu, L.; Zhang, J.; Zhang, L.; Gong, D.; Zhao, Y.; Chen, Q.; Huang, S.; Yang, M.; Yang, X., et al. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. *Sci Rep* **2020**, *10*, 19196, doi:10.1038/s41598-020-76282-0.
13. Wang, X.; Deng, X.; Fu, Q.; Zhou, Q.; Feng, J.; Ma, H.; Liu, W.; Zheng, C. A Weakly-Supervised Framework for COVID-19 Classification and Lesion Localization From Chest CT. *IEEE Trans Med Imaging* **2020**, *39*, 2615-2625, doi:10.1109/TMI.2020.2995965.
14. Mohammed, A.; Wang, C.; Zhao, M.; Ullah, M.; Naseem, R.; Wang, H.; Pedersen, M.; Cheikh, F.A. Weakly-Supervised Network for Detection of COVID-19 in Chest CT Scans. *IEEE Access* **2020**, *8*, 155987-156000, doi:10.1109/ACCESS.2020.3018498.
15. Hasan, A.M.; Al-Jawad, M.M.; Jalab, H.A.; Shaiba, H.; Ibrahim, R.W.; Al-Shamasneh, A.R. Classification of Covid-19 Coronavirus, Pneumonia and Healthy Lungs in CT Scans Using Q-Deformed Entropy and Deep Learning Features. *Entropy (Basel)* **2020**, *22*, doi:10.3390/e22050517.

16. Ahsan, M.M.; Gupta, K.D.; Islam, M.M.; Sen, S.; Rahman, M.L.; Shakhawat Hossain, M. COVID-19 Symptoms Detection Based on NasNetMobile with Explainable AI Using Various Imaging Modalities. *Machine Learning and Knowledge Extraction* **2020**, *2*, 490-504.
17. Alshazly, H.; Linse, C.; Barth, E.; Martinetz, T. Explainable COVID-19 Detection Using Chest CT Scans and Deep Learning. *Sensors (Basel)* **2021**, *21*, doi:10.3390/s21020455.
18. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Cham; pp. 234-241.
19. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support (2018)* **2018**, *11045*, 3-11, doi:10.1007/978-3-030-00889-5_1.
20. Chen, L.-C.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587* **2017**.
21. Xie, S.; Girshick, R.; Dollár, P.; Tu, Z.; He, K. Aggregated Residual Transformations for Deep Neural Networks. In Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 21-26 July 2017; pp. 5987-5995.
22. WHO Novel Coronavirus - China. Available online: http://www.who.int/csr/don/12-january-2020-novel-coronavirus-china/en/ (accessed on Jan 19).
23. Hanson, A.; PNVR, K.; Krishnagopal, S.; Davis, L. Bidirectional Convolutional LSTM for the Detection of Violence in Videos. Cham; pp. 280-295.
24. El-bana, S.; Al-Kabbany, A.; Sharkas, M. A multi-task pipeline with specialized streams for classification and segmentation of infection manifestations in covid-19 scans. *PeerJ Computer Science* **2020**, *6*, e303.
25. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of Proceedings of the IEEE conference on computer vision and pattern recognition; pp. 770-778.
26. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of Proceedings of the IEEE conference on computer vision and pattern recognition; pp. 1-9.
27. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M., et al. ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision* **2015**, *115*, 211-252, doi:10.1007/s11263-015-0816-y.
28. Chung, M.; Bernheim, A.; Mei, X.; Zhang, N.; Huang, M.; Zeng, X.; Cui, J.; Xu, W.; Yang, Y.; Fayad, Z.A., et al. CT Imaging Features of 2019 Novel Coronavirus (2019-nCoV). *Radiology* **2020**, *295*, 202-207, doi:10.1148/radiol.2020200230.
29. Singh, A.; Sengupta, S.; Lakshminarayanan, V. Explainable Deep Learning Models in Medical Image Analysis. *Journal of Imaging* **2020**, *6*, 52.
30. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International Journal of Computer Vision* **2020**, *128*, 336-359, doi:10.1007/s11263-019-01228-7.