CoRSAI: A System for Robust Interpretation of CT Scans of COVID-19 Patients Using Deep Learning

28

MANVEL AVETISIAN, ILYA BURENKO, KONSTANTIN EGOROV, VLADIMIR KOKH, and ALEKSANDR NESTEROV, Sberbank AI Laboratory
ALEKSANDR NIKOLAEV, Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies, Russia
ALEXANDER PONOMARCHUK and ELENA SOKOLOVA, Sberbank AI Laboratory
ALEX TUZHILIN, Sberbank AI Laboratory and New York University
DMITRY UMERENKOV, Sberbank AI Laboratory

Analysis of chest CT scans can be used in detecting parts of lungs that are affected by infectious diseases such as COVID-19. Determining the volume of lungs affected by lesions is essential for formulating treatment recommendations and prioritizing patients by severity of the disease. In this article we adopted an approach based on using an ensemble of deep convolutional neural networks for segmentation of slices of lung CT scans. Using our models, we are able to segment the lesions, evaluate patients' dynamics, estimate relative volume of lungs affected by lesions, and evaluate the lung damage stage. Our models were trained on data from different medical centers. We compared predictions of our models with those of six experienced radiologists, and our segmentation model outperformed most of them. On the task of classification of disease severity, our model outperformed all the radiologists.

CCS Concepts: • **Applied computing** → **Health care information systems**; • **Computing methodologies** → **Computer vision**; **Ensemble methods;**

Additional Key Words and Phrases: Convolutional neural network, deep learning, ensembling, COVID-19, segmentation, lesion detection

ACM Reference format:
Manvel Avetisian, Ilya Burenko, Konstantin Egorov, Vladimir Kokh, Aleksandr Nesterov, Aleksandr Nikolaev, Alexander Ponomarchuk, Elena Sokolova, Alex Tuzhilin, and Dmitry Umerenkov. 2021. CoRSAI: A System for Robust Interpretation of CT Scans of COVID-19 Patients Using Deep Learning. *ACM Trans. Manag. Inf. Syst.* 12, 4, Article 28 (September 2021), 16 pages.
https://doi.org/10.1145/3467471

Authors listed in alphabetical order.
Authors' addresses: M. Avetisian (corresponding author), Sberbank AI Laboratory, Oruzheynyy Pereulok, 41, Moscow, Russia; email: avetisian.m.s@sberbank.ru; I. Burenko (corresponding author), K. Egorov, V. Kokh, A. Nesterov, A. Ponomarchuk, E. Sokolova, and D. Umerenkov, Sberbank AI Laboratory, Moscow, Russia; emails: {burenko.i.m, egorov.k.ser, kokh.v.n, AINesterov, ponomarchuk.a.v, Sokolova.El.Vladimirov}@sberbank.ru, D.Umerenkov@gmail.com; A. Nikolaev, email: a.e.nikolaev@yandex.ru; A. Tuzhilin, Sberbank AI Laboratory, Russia and New York University, New York, NY 10003, United States, New York, USA; email: atuzhili@stern.nyu.edu.

2158-656X/2021/09-ART28 $15.00
https://doi.org/10.1145/3467471

1 INTRODUCTION

Coronavirus (COVID-19) has spread widely around the world since the beginning of 2020, and an extensive effort to combat the pandemic was launched that year. As a result of this effort, there have been several diagnostic tests developed in the medical community to detect COVID cases. One of the most prominent methods to confirm a COVID-19 infection is by conducting the **reverse transcriptional polymerase chain reaction (RT-PCR)** test, which has a lower sensitivity of 65% to 95%. Although useful and popular, the RT-PCR test has the problems of producing negative results even if the patient is infected and having to wait for the test results. Therefore, in some countries a chest **computed tomography (CT)** scan is widely used in clinical practice to detect typical changes in the pulmonary parenchyma associated with COVID-19 [6, 8, 24, 28] as a complement to the RT-PCR test, especially since CT is effective for early detection and diagnosis of COVID-19 [11, 18] and the results of CT scans can be analyzed immediately [1, 15]. Multifocal **ground-glass opacifications (GGOs)** are the most common finding of the CT scan, usually localized peripherally in both lungs, while a single ground-glass lesion can be common at an early stage of the disease [44]. Clinical manifestations of COVID-19 pneumonia and their severity correlate with the volume of lung damage, which can be assessed using visual or quantitative scale.

Although it is easy to assess the severity of lung damage using a visual scale, this is a subjective assessment that can vary substantially among radiologists [11]. Therefore, there exists a more objective classification of lung damage widely used in some countries, including Russia, consisting of the following five stages (referred in the article as CT classes): CT-0: absence of damage; CT-1: **pulmonary parenchymal involvement (PPI)** being $\leq$ 25%; CT-2: PPI being in the range of 25% to 50%; CT-3: PPI in the range of 50% to 75%; and CT-4: PPI $\geq$ 75% [28]. In the context of the current COVID-19 pandemic, radiologists in specialized departments need to process a very large number of CT images of subjects with suspected COVID-19, sometimes up to several hundred patients per day, which puts an incredible burden on them and also delays the COVID-19 detection event. Therefore, an automated system that can accurately detect the presence of COVID-19 and calculate the pathology of lung volume will significantly reduce the burden on the radiologist, help objectively assess the severity of the disease, make it possible to prioritize the radiologist work schedule, and provide better insights into the follow-up studies to assess the dynamics of the disease.

In this article, we present the CoRSAI system[1] that takes CT scans of COVID-19 patients and does the image classification and segmentation tasks using **Deep Learning (DL)**-based methods to find the affected areas, to determine the severity of the disease, and to track disease progression. The proposed system uses a novel ensemble of previously developed DL-based models that was architected specifically with the goal of detecting lung damage caused by COVID-19.

To test our system, we compared its performance with two existing DL-based baselines on two open datasets. As a result, our system outperformed these baselines. In addition, we also conducted a study in which we compared CoRSAI's performance with that of six radiologists having at least 3 years of practical experience across the following typical tasks:

- *Segmentation:* detection of the affected areas of the lungs
- *Patient's dynamics:* detection of positive response to the therapy or disease progression
- *Lesion share estimation:* assessment of the lung damage share (ratio of lesion volume to lung volume)
- *Classification:* identification of lung damage stage according to the CT class

[1]CoRSAI stands for *Ru*S*sian* *CO*ronovirus *AI*-based detection system.

We performed these four experiments using 58 CT scans on 49 patients at a large Russian hospital and used the services of six experienced radiologists.

The results of this study show that our system outperformed the experienced radiologists for the segmentation and classification tasks on average. In all the cases, our system correctly determined the patients' dynamics. The results of the lesion share estimation are not directly usable due to a high degree of radiologists' subjectivity on this task. Correcting for this subjectivity bias allows our system to outperform all six radiologists on the classification task, three of them with statistical significance.

These results imply that our system can be used as a second-opinion tool that would help radiologists to deal with the coronavirus pandemic. In fact, our system has been favorably received by the medical community in Russia and has been successfully deployed in several hospitals in the country.

In this article, we make the following contributions:

- First, we propose an ensemble method specifically designed for the COVID detection problem for the CT scan data that we implemented as a part of the CoRSAI system.
- Second, we empirically compare CoRSAI with two existing baselines and demonstrate that our method outperforms these baselines on the public and on our proprietary CT scan data.
- Third, we conducted a study in which the CoRSAI system outperformed six experienced radiologists across various COVID detection tasks.

We give an overview of existing approaches to classification and segmentation of CT scans and chest X-ray studies in Section 2; in Section 3 we give a detailed description of datasets that we used for classification and segmentation tasks as well as for experiments with doctors; in Section 4 we describe models that we utilized, how we preprocess data, and how we combine individual models into an ensemble; Section 5 is devoted to experiments that we conducted and results we obtained; we give a conclusion and some final thoughts in Section 6.

2 RELATED WORK

Using **Convolutional Neural Networks (CNNs)** is a common practice for the task of image segmentation. Since its appearance in 2015, the U-Net architecture [35] and its modifications have been widely used for the medical segmentation tasks during the analysis of X-rays, CT scans, MRIs, and ultrasound signals for detecting pneumonia [33], breast cancer [40], stroke [5], liver tumor segmentation [23], prostate cancer [27], and many other medical problems [26].

Furthermore, there is a large body of work on applying CNNs to the task of nodule detection in the chest CT images [20], segmentation of the interstitial lung disease [2], chest organ segmentation [7, 36], and other related tasks [22].

There is a large body of recent work dedicated to the task of detecting COVID-19 lesions in lungs based on X-ray studies and CT scans. In particular, [43] and [29] focus on differentiating coronavirus-induced pneumonia from other pneumonia types and healthy controls. Both papers describe the experiments conducted on large samples of cases and produce comparable results with high levels of differentiation between these two types of pneumonias. In [39], a model was developed that classified whether a CT scan contains COVID-19 lesions or not, achieving ROC AUC of 0.959 based on the CT-level annotations. In [14], the authors describe a supervised and a semi-supervised approach to segmentation of lesions and lungs. In [41] a joint classification and segmentation model is built in order to achieve higher quality by extending an expensive segmentation dataset with a classification dataset. In [9] the same problem was solved by using contrastive learning to train a neural network that can later be adopted for the classification task.