

Received May 26, 2020, accepted June 25, 2020, date of publication June 29, 2020, date of current version July 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3005510

Weakly Supervised Deep Learning for COVID-19 Infection Detection and Classification From CT Images

SHAOPING HU¹, YUAN GAO^{1b,2,3}, (Member, IEEE), ZHANGMING NIU^{3,4}, YINGHUI JIANG^{4,5}, LAO LI^{4,5}, XIANGLU XIAO^{3,5}, MINHAO WANG^{4,5}, EVANDRO FEI FANG⁶, WADE MENPES-SMITH³, JUN XIA⁷, HUI YE⁸, AND GUANG YANG^{1b,9,10}, (Member, IEEE)

¹Radiology Department, Hospital of Wuhan Red Cross Society, Wuhan 430015, China

²Institute of Biomedical Engineering, University of Oxford, Oxford OX3 7DQ, U.K.

³Aladdin Healthcare Technologies Ltd., London EC1Y 0UH, U.K.

⁴Hangzhou Ocean's Smart Boya Company, Ltd., Hangzhou 310016, China

⁵Mind Rank Ltd., Admiralty, Hong Kong

⁶Department of Clinical Molecular Biology, University of Oslo, 0315 Oslo, Norway

⁷Radiology Department, Shenzhen Second People's Hospital, Shenzhen 518035, China

⁸PET-CT Center, Hunan Cancer Hospital, Changsha 410013, China

⁹NHLL, Imperial College London, London SW3 6LY, U.K.

¹⁰Royal Brompton Hospital, London SW3 6NP, U.K.

Corresponding authors: Hui Ye (yuxin75831@163.com) and Guang Yang (g.yang@imperial.ac.uk)

This work was supported in part by the European Research Council Innovative Medicines Initiative on Development of Therapeutics and Diagnostics Combatting Coronavirus Infections Award 'DRAGON: rapiD and secuRe AI imaging based diaGnosis, stratification, fOllow-up, and preparedness for coronavirus paNdemics' under Grant H2020-JTI-IMI2 101005122, and in part by IIAT Hangzhou.

ABSTRACT An outbreak of a novel coronavirus disease (i.e., COVID-19) has been recorded in Wuhan, China since late December 2019, which subsequently became pandemic around the world. Although COVID-19 is an acutely treated disease, it can also be fatal with a risk of fatality of 4.03% in China and the highest of 13.04% in Algeria and 12.67% Italy (as of 8th April 2020). The onset of serious illness may result in death as a consequence of substantial alveolar damage and progressive respiratory failure. Although laboratory testing, e.g., using reverse transcription polymerase chain reaction (RT-PCR), is the golden standard for clinical diagnosis, the tests may produce false negatives. Moreover, under the pandemic situation, shortage of RT-PCR testing resources may also delay the following clinical decision and treatment. Under such circumstances, chest CT imaging has become a valuable tool for both diagnosis and prognosis of COVID-19 patients. In this study, we propose a weakly supervised deep learning strategy for detecting and classifying COVID-19 infection from CT images. The proposed method can minimise the requirements of manual labelling of CT images but still be able to obtain accurate infection detection and distinguish COVID-19 from non-COVID-19 cases. Based on the promising results obtained qualitatively and quantitatively, we can envisage a wide deployment of our developed technique in large-scale clinical studies.

INDEX TERMS COVID-19, deep learning, weakly supervision, CT images, classification, convolutional neural network.

I. INTRODUCTION

Coronavirus disease 2019 (COVID-19) has been widespread worldwide since December 2019 [1], [2]. It is highly contagious, and severe cases can lead to acute respiratory distress or multiple organ failure [3]. On 11 March 2020, the WHO has made the assessment that COVID-19 can be characterised as a pandemic. As of 8th April 2020, in total, 1,391,890 cases

The associate editor coordinating the review of this manuscript and approving it for publication was Shuihua Wang^{1b}.

of COVID-19 have been recorded, and the death toll has reached 81,478 with a rapid increase of cases in Europe and North America.

The disease can be confirmed by using the reverse-transcription polymerase chain reaction (RT-PCR) test [4]. While being the gold standard for diagnosis, confirming COVID-19 patients using RT-PCR is time-consuming, and both high false-negative rates and low sensitivities may put hurdles for the presumptive patients to be identified and treated early [3], [5], [6].