doi • 10.5578/tt.20219606

Tuberk Toraks 2021;69(4):486-491

Received/Geliş Tarihi: 02.08.2021 • Accepted/Kabul Ediliş Tarihi: 07.09.2021

Contribution of artificial intelligence applications developed with the deep learning method to the diagnosis of COVID-19 pneumonia on computed tomography

Nevin AYDIN¹(ID)
Özer ÇELİK²(ID)

¹ Department of Radiology, Eskişehir Osmangazi University
Faculty of Medicine, Eskişehir, Turkey

¹ Eskişehir Osmangazi Üniversitesi Tıp Fakültesi, Radyoloji Anabilim Dalı,
Eskişehir, Türkiye

² Department of Mathematics and Computer, Eskişehir Osmangazi University
Faculty of Science and Letters, Eskişehir, Turkey

² Eskişehir Osmangazi Üniversitesi Fen Edebiyat Fakültesi, Matematik ve
Bilgisayar Bilimleri Bölümü, Eskişehir, Türkiye

Cite this article as: Aydın N, Çelik Ö. Contribution of artificial intelligence applications developed with the deep learning method to the diagnosis of COVID-19 pneumonia on computed tomography. Tuberk Toraks 2021;69(4):486-491.

Yazışma Adresi (Address for Correspondence)

Dr. Nevin AYDIN
Department of Radiology,
Eskişehir Osmangazi University
Faculty of Medicine,
ESKİŞEHİR - TURKEY
e-mail: drnevinaydin@gmail.com

ABSTRACT

Contribution of artificial intelligence applications developed with the deep learning method to the diagnosis of COVID-19 pneumonia on computed tomography

Introduction: Computed tomography (CT) is an auxiliary modality in the diagnosis of the novel Coronavirus (COVID-19) disease and can guide physicians in the presence of lung involvement. In this study, we aimed to investigate the contribution of deep learning to diagnosis in patients with typical COVID-19 pneumonia findings on CT.

Materials and Methods: This study retrospectively evaluated 690 lesions obtained from 35 patients diagnosed with COVID-19 pneumonia based on typical findings on non-contrast high-resolution CT (HRCT) in our hospital. The diagnoses of the patients were also confirmed by other necessary tests. HRCT images were assessed in the parenchymal window. In the images obtained, COVID-19 lesions were detected. For the deep Convolutional Neural Network (CNN) algorithm, the Confusion matrix was used based on a Tensorflow Framework in Python.

Results: A total of 596 labeled lesions obtained from 224 sections of the images were used for the training of the algorithm, 89 labeled lesions from 27 sections were used in validation, and 67 labeled lesions from 25 images in testing. Fifty-six of the 67 lesions used in the testing stage were accurately detected by the algorithm while the remaining 11 were not recognized. There was no false positive. The Recall, Precision and F1 score values in the test group were 83.58, 1, and 91.06, respectively.

Conclusion: We successfully detected the COVID-19 pneumonia lesions on CT images using the algorithms created with artificial intelligence. The integration of deep learning into the diagnostic stage in medicine is an important step for the diagnosis of diseases that can cause lung involvement in possible future pandemics.

Key words: COVID-19 pneumonia; CT; deep learning; neural networks; computed tomography

ÖZ

Derin öğrenme yöntemi kullanılarak geliştirilen yapay zeka uygulamalarının bilgisayarlı tomografide COVID-19 pnömonisi tanısına katkısı

Giriş: Bilgisayarlı tomografi (BT) novel Coronavirüs hastalığının tanısında yardımcı bir tetkik olup akciğer tutulumu varlığında yönlendirici olmaktadır. Biz çalışmamızda BT’de tipik COVID-19 pnömonisi tanısı almış hastalarda derin öğrenmenin tanıya katkısını araştırmayı hedefledik.

Materyal ve Metod: Hastanemizde kontrastsız yüksek rezolüsyonlu bilgisayarlı tomografiye göre tipik COVID-19 pnömonisi tanısı almış ve diğer gerekli testlerle tanısı doğrulanmış 35 hastadan elde edilen 690 lezyon retrospektif olarak değerlendirildi. Yüksek rezolüsyonlu BT’de görüntüler parankim penceresinde değerlendirildi. Elde edilen görüntülerde COVID-19 lezyonlarına yönelik saptama yapıldı. Bizim Convolutional Neural Network (CNN) algoritmamızda Python’un Tensorflow Framework tabanlı Confusion matrisi kullanıldı.

Bulgular: Hasta görüntülerinde 224 kesitten elde edilen 596 etiketlenmiş lezyon eğitim basamağında kullanıldı. Yirmi yedi kesitten elde edilen 89 etiketlenmiş lezyon validasyonda, 25 görüntüden elde edilen 67 etiketlenmiş lezyon test aşamasında kullanıldı. Test aşamasında kullanılan lezyonların 67’sinin 56’sı doğru tespit edildi. 11’i bulunamadı. Yanlış bulunan olmadı. Test grubunda Recall, Precision ve F1 skor değerleri sırasıyla 83.58, 1, 91.06 olarak saptandı.

Sonuç: Çalışmamızda yapay zeka uygulamalarıyla BT görüntüleri üzerinden yapılan programlarla COVID-19 hastalığı lezyonlarını saptamada başarılı olduk. Derin öğrenmenin tıp alanında tanı aşamasına entegrasyonu bundan sonraki olabilecek salgınlarda akciğer tutulumu yapabilen hastalıkların tanısı için önemli bir adımdır.

Anahtar kelimeler: COVID-19 pnömonisi; BT; derin öğrenme; nöral network; bilgisayarlı tomografi

INTRODUCTION

The novel Coronavirus (COVID-19) disease caused by the SARS-CoV-2 virus, which emerged in Wuhan, China at the end of 2019, has affected the whole world (1). In early 2020, this disease was declared a public health emergency by the World Health Organization. To detect the disease, the reverse-transcription polymerase chain reaction (RT-PCR) method is applied to nasopharyngeal swab samples. However, since the sensitivity of this method cannot exceed 60-70% under optimal conditions and the result cannot be obtained in a short time, other auxiliary diagnostic techniques have been evaluated (2,3). Computed tomography (CT) is an auxiliary modality in the diagnosis of the COVID-19 pneumonia and guides this process in the presence of lung involvement. The typical lung involvement of the disease is the presence of ground-glass opacities involving the lower lobes with peripheral location in either or both lungs (4). CT is important for the early diagnosis and treatment of the disease. Deep learning using convolutional neural network (CNNs) is an important component of artificial intelligence (5). The deep learning component of artificial intelligence has been used in the diagnosis of COVID-19 pneumonia and has been successful in this regard (6). Recognition and correct labeling of lesions by artificial intelligence in COVID-

19 pneumonia may help the patient to be diagnosed early and start supportive treatments early. In this study, we aimed to investigate the contribution of deep learning to diagnosis in patients with typical findings of COVID-19 pneumonia on CT.

MATERIALS and METHODS

The study included the images of 35 patients that presented to the Medical Faculty between March and May 2020 and were diagnosed with COVID-19 pneumonia based on typical findings on non-contrast high-resolution CT (HRCT). All patients had RT-PCR positivity. A total of 690 lesions obtained from 276 cross-sectional images of the 35 patients were retrospectively evaluated. Prior to the study, the approval of the ethics committee was obtained (No. 25403353-050.99-E.46807, Date: 08.05.2020). CT scans were performed using the 64-slice CT scanner (Siemens, SOMATOM 2015) at a section thickness of 0.6 mm. HRCT images were evaluated in the parenchymal window. CT examinations were performed with a window width of 1,400, windows height of -400 at 100 kV and 243 mA. In the parenchymal window, the images with a matrix size of 672 x 672 pixels were used, and the images taken in the axial plane were converted to the png format. In machine learning, data labeling is the process of identifying raw data

(images, text files, videos, etc.) and adding one or more meaningful and informative labels to provide context so that a machine learning model can learn from it. For example, labels might indicate whether a photo contains a bird or car, which words were uttered in an audio recording, or if an x-ray contains a tumor. Data labeling is required for a variety of use cases including computer vision, natural language processing, and speech recognition (7).

The lesions located in the sections were labeled. Eighty percent of the data set was allocated for the training of the algorithm, 10% for testing, and 10% for validation. The algorithm was used to detect the COVID-19 lesions on the images. In mathematics and computer science, an algorithm is a set of operations that run from a defined initial state to a clearly defined final state to do a work (8). Artificial intelligence algorithms are methods that perform model creation using various statistical approaches from the training data set (9).

HRCT images of patients with bilateral ground glass opacities, predominantly in the lower lobe, were included in the study. Some patients underwent more than one HRCT scan. The findings obtained from a single examination were included in the evaluation for each patient. HRCTs were evaluated by a thoracic radiologist with 3 years of experience. To calculate the success of the model, the confusion matrix was used.

The test dataset was used to calculate the diagnostic and predictive accuracy, sensitivity, specificity, nega-

tive predictive value, negative predictive value, precision, and confusion matrix, with the deep CNN algorithm based on a Tensorflow framework in Python. 16 GB RAM and an NVIDIA GeForce GTX 1060Ti graphic card were used as personal computer equipment. Data collection, labeling of all lesions, and evaluation of the results were performed in approximately one month.

RESULTS

Mean age of the patients was calculated as 38.9 ± 12.04 , and the mean number of lesions in the HRCT scans was calculated as 17 (minimum 5, maximum 27).

All patients had more than five lesions, and the lesions were predominantly bilateral and located in the lower lobe. 581 of the lesions were in the form of ground glass density and 15 lesions were in the appearance of crazy paving pattern. 407 of the lesions were located in the lower lobes. A total of 596 labeled lesions obtained from 224 sections of the CT images were used in training, 89 labeled lesions from 27 sections in validation, and 67 labeled lesions from 25 images in testing. Of the 67 lesions used in the testing stage, 56 were accurately detected by the constructed model while 11 were not recognized (Figure 1 and Figure 2). There was no false negative. The recall, precision and F1 score values of the test group were found to be 83.58, 1 and 91.06, respectively (Table 1). The receiver operating characteristic (ROC) curve and area under the curve (AUC) values of the test group of our model

Figure 1. The image obtained from the axial plane in the parenchymal window. The three lesions labeled as having ground glass density in the right lower lobe were accurately identified by the deep learning method.

Figure 2. Another image obtained from the axial plane in the parenchymal window. Of the four lesions labeled in the upper and lower lobe of the left lung, two with a peripheral location were accurately identified using the deep learning method while the remaining two that were centrally located could not be detected.

Table 1. Recall, precision and F1 score values of the COVID-19 pneumonia test group

	Recall	Precision	F1 Score
COVID-19 Pneumonia	83.58	1	91.06

Figure 3. ROC-AUC Graph.

are given in Figure 3. Of the 11 undetected lesions in the test group, five were peripherally and six were centrally located. All undetectable lesions in the test

group were of ground glass density. For 10 of 11 undetected lesions in the test group, the longest diameter was less than 2 cm in the axial plane.

DISCUSSION

The cross-sectional images of 35 patients that presented to our hospital over a period of approximately three months and were diagnosed with COVID-19 pneumonia based on typical HRCT findings and RT-PCR positivity were included in this study. In the diagnostic stage of the patients, the criteria of bilateral lesions, nodular ground glass density, predominant distribution in the lower lobe, and multiple lesions were taken into consideration. The sensitivity of RT-PCR test varies between 56 and 83% (10,11). In our study, we performed the evaluation based on lesions and included a total of 690 lesions in this evaluation. We obtained the recall, precision and F1 score values as 83.58, 1 and 91.06, respectively, indicating that our deep learning method was successful in detecting COVID-19 lesions. We used two-dimensional images in our study. In the literature, Li et al. developed a deep learning method using three-dimensional images and similarly showed high sensitivity and specificity in the diagnosis of COVID-19 pneumonia (12).

In many studies, lung involvement of the COVID-19 related lung disease has been defined as multiple peripheral ground glass opacities predominantly located in the lower lobe (4). In addition, other findings, such as consolidation, crazy paving pattern, and reverse halo sign can be observed (13,14). Five hundred and eight-one of the lesions were in the form of ground glass density and 15 lesions were in the appearance of crazy paving pattern in our study. This lack of lesion variety can be considered as a limitation of our study. However, the lesions having similar characteristics can also be one of the factors that led to the success of our model. Studies to be conducted with a large number of cases, including lesions with other patterns, such as consolidation, crazy paving, and reverse halo sign will shed light on this issue.

Similar to our study, Brunese et al. have used the deep learning method in X-ray examinations to differentiate the COVID-19 pneumonia from other pulmonary diseases. Despite the use of X-ray images, the authors have obtained high accuracy values (15). In another deep learning study investigating the detection of COVID-19 on lung radiography images, Vaid et al. have reported >96.3% accuracy (16). Ardakani et al. have used 10 different CNNs to differentiate COVID-19 and non-COVID diseases and compared

the results of the deep learning methods (17). Zhang et al. have performed the detection and quantification of lesions on CT images and reported that the software they used was successful in this process (18). In our study, we used the deep learning algorithm only for detection.

In this study, the lesions that could not be detected by the deep learning method were mostly centrally located. This may be due to the proximity of these lesions to the bronchovascular structures in the central area. Furthermore, of the 11 lesions that were not detected by the model, 10 were under 2 cm in size. From these findings, it can be concluded that the detection success rate of the deep learning method may be lower in small lesions. In our study, we did not include other viral pneumonias that usually accompany ground-glass lesions. Wang et al. have compared the findings of the COVID-19 disease and influenza pneumonia that accompany ground-glass lesions (19). Our patients were those that presented to our hospital during the duration of the pandemic. These patients had a clinical manifestation consistent with COVID-19 pneumonia, confirmed by the typical thorax CT findings. However, the possibility of confusing COVID-19 pneumonia with other viral pneumonia is another limitation of our study. This is also a limitation, as lesions from the same patient were included in the training, validation, and test group sets. Last, we did not use a control group consisting of healthy patients with normal thorax CT findings.

CONCLUSION

As in many studies, in our work, we successfully detected COVID-19 lesions on CT images using the program developed with artificial intelligence. Radiological and artificial intelligence applications will continue to be complementary areas in the coming years. In order to protect medical personnel in highly infectious diseases, such as the COVID-19 disease, which has affected the whole world, there is a need for further artificial intelligence applications to be used in diagnosis and treatment. Every study conducted on this subject will be valuable in terms of contributing to future diagnostic algorithms.

Ethical Committee Approval: This study approval was obtained from Eskişehir Osmangazi University Non-Invasive Researches Ethical Committee (Decision No: 25403353-050.99-E.46807, Date: 08.05.2020).

CONFLICT of INTEREST

The authors declare that they have no conflict of interest.

AUTHORSHIP CONTRIBUTIONS

Concept/Design: All of authors

Analysis/Interpretation: All of authors

Data acquisition: All of authors

Writing: NA

Clinical Revision: All of authors

Final Approval: All of authors

REFERENCES

- Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020; 395(10223): 497-506.
- Ai T, Yang Z, Hou H, Zhan C, Chen C, Lv W, et al. Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* 2020; 296(2): E32-E40.
- Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, et al. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* 2020; 296(2): E115-E117.
- Wang Y, Dong C, Hu Y, Li C, Ren Q, Zhang X, et al. Temporal changes of CT findings in 90 patients with COVID-19 pneumonia: a longitudinal study. *Radiology* 2020; 296: E55-E64.
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: In Pereira F, Burges CJC, Bottou L, Weinberger KQ (eds). *NIPS'12. Proceedings of the 25th International Conference on Neural Information Processing Systems*; 2012 Dec 1. Red Hook, NY, USA. Curran Associates Inc.; 2012 p. 1097-1105.
- Alshazly H, Linse C, Barth E, Martinetz T. Explainable COVID-19 Detection Using Chest CT Scans and Deep Learning. *Sensors (Basel)*. 2021;21(2):455.
- Chen HL, Chuang KT, Chen MS. On data labeling for clustering categorical data. *IEEE Transactions on knowledge and Data Engineering*. 2008; 20(11): 1458-1472.
- Merriam-Webster. (n.d.). Algorithm. In Merriam-Webster.com dictionary. Retrieved August 16, 2021, from <https://www.merriam-webster.com/dictionary/algorithm>.
- Çelik Ö, Aslan AF, Osmanoğlu, UÖ, Çetin N, Tokar B. Estimation of renal scarring in children with lower urinary tract dysfunction by utilizing resampling technique and machine learning algorithms. *Journal of Surgery and Medicine* 2020; 4(7): 573-7.
- Yang Y, Yang M, Shen C, Wang F, Yuan J, Li J et al. Evaluating the accuracy of different respiratory specimens in the laboratory diagnosis and monitoring the viral shedding of 2019-nCoV infections. *The Innovation* 2020; 1(3): 100061.
- Wang W, Xu Y, Gao R, Lu R, Han K, Wu G et al. Detection of SARS-CoV-2 in different types of clinical specimens. *JAMA* 2020; 323(18): 1843-1844.
- Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, 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(2): E65-E71.
- Bernheim A, Mei X, Huang M, Yang Y, Fayad ZA, Zhang N, et al. Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection. *Radiology* 2020; 295(3): 200463.
- Kanne JP. Chest CT Findings in 2019 Novel Coronavirus (2019-nCoV) Infections from Wuhan, China: Key Points for the Radiologist. *Radiology* 2020; 295(1): 16-17.
- Brunese L, Mercaldo F, Reginelli A, Santone A. Explainable Deep Learning for Pulmonary Disease and Coronavirus COVID-19 Detection from X-rays. *Comput Methods Programs Biomed* 2020; 196: 105608.
- Vaid S, Kalantar R, Bhandari M. Deep learning COVID-19 detection bias: accuracy through artificial intelligence. *Int Orthop* 2020; 44(8): 1539-42.
- Ardakani AA, Kanafi AR, Acharya UR, Khadem N, Mohammadi A. Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. *Comput Biol Med* 2020; 121: 103795.
- Zhang HT, Zhang JS, Zhang HH, Nan YD, Zhao Y, Fu EQ, et al. Automated detection and quantification of COVID-19 pneumonia: CT imaging analysis by a deep learning-based software. *Eur J Nucl Med Mol Imaging* 2020; 47(11): 2525-32.
- Wang H, Wei R, Rao G, Zhu J, Song B. Characteristic CT findings distinguishing 2019 novel coronavirus disease (COVID-19) from influenza pneumonia. *Eur Radiol* 2020; 30(9): 4910-4917.