Jurnal Artificial, Vol 2 No 1, April 2024

Deep Learning Methods for Chest X-ray Imaging-Based COVID-19 Pneumonia Detection

MD. Roman Sarkar [1*]

[1] Southwest University of Science and Technology, China

* engr.romansarkar@gmail.com

Abstract

The global COVID-19 pandemic has posed significant challenges in the early detection and management of pneumonia caused by the virus. One effective way to detect COVID-19 pneumonia is through chest X-ray imaging, which offers advantages in terms of accessibility and cost. However, the manual diagnostic process requires time and high accuracy. This study aims to develop a hybrid convolutional neural network (CNN) model capable of automatically and accurately detecting COVID-19 pneumonia from chest X-ray images. The methodology used in this research involves meticulous data processing, including data augmentation and image quality enhancement, as well as the application of a hybrid CNN architecture combining VGG-16 and ResNet-50 models for feature extraction and complex pattern processing. Experimental evaluation results show that this model achieves 96.5% accuracy, 97.2% precision, 96.5% recall, and 96.8% F1-score in detecting COVID-19 pneumonia. The model also demonstrates good performance in distinguishing non-COVID pneumonia and healthy conditions, with accuracy ranging from 96.5% to 96.6%. These findings highlight the potential of artificial intelligence in enhancing diagnostic efficiency and accuracy, particularly in public health emergencies. This research is expected to contribute to the development of reliable AI-based diagnostic solutions to assist healthcare professionals in the early detection of COVID-19 pneumonia and other pulmonary infections.

Keywords: *Deep learning, Convolutional neural networks, Medical imaging, COVID-19, AI diagnostics*

Introduction

The COVID-19 pandemic has posed significant challenges to global healthcare systems. The SARS-CoV-2 virus, the primary cause of COVID-19, is not only a rapidly spreading infectious disease but also exacerbates pre-existing comorbidities in many individuals (Rehouma et al., 2021). The pandemic has triggered significant global behavioral changes in efforts to control its spread (Yang et al., 2020). One of the major impacts of COVID-19 infection is the emergence of serious complications in the form of pneumonia, which often becomes the leading cause of death among infected patients. Therefore, early and accurate diagnosis of COVID-19-induced pneumonia is crucial for the effective management and containment of the disease (Okolo et al., 2021).

Among the various diagnostic tools used, chest X-ray (CXR) imaging has proven to be an effective and affordable modality for detecting pneumonia, including that caused by COVID-19 (Irede et al., 2024). Chest X-rays allow for the rapid identification of lung abnormalities, including infiltrates that are characteristic of COVID-19 pneumonia (Esteva et al., 2017). The main challenge in using CXR is the need for rapid and accurate

https://doi.org/10.54065/artificial.2.1.2024.487

interpretation, which often requires a high level of clinical expertise (Rajaraman et al., 2020). Therefore, it is essential to develop methods that can enhance the performance of CXR-based diagnosis more efficiently and accurately.

Advancements in Information and Communication Technology (ICT), particularly in artificial intelligence (AI) and deep learning, have opened up great opportunities in medical diagnostics (Danilov et al., 2022). AI, with its ability to rapidly and accurately analyze and interpret medical data, has gained widespread attention, especially in detecting medical conditions through imaging. One example of its application is the use of deep convolutional neural networks (CNNs) to analyze medical images such as CXR and CT scans (Tolkachev et al., 2021). CNNs have proven effective in detecting subtle and complex patterns in medical images that may be overlooked by human examiners (Apostolopoulos & Mpesiana, 2020).

The context of COVID-19, recent studies show that AI, particularly deep learning models, holds significant potential to improve the accuracy of pneumonia diagnosis caused by SARS-CoV-2 infection (Esteva et al., 2019). These systems can learn to differentiate between COVID-19 pneumonia and other pulmonary conditions, such as bacterial pneumonia or tuberculosis by training AI models on CXR datasets containing images of infected lungs (Nagpal et al., 2020). This offers a major advantage in terms of speed and consistency of diagnosis, which is crucial in the midst of a global pandemic, especially in regions with a shortage of trained medical personnel (Chowdhury et al., 2020).

This study aims to develop an AI-based solution that can improve the performance of pneumonia diagnosis for COVID-19 using CXR. The primary focus of this research is the application of modern deep learning methods, specifically CNN models, to analyze chest X-ray images. This model is designed to automatically detect abnormalities caused by COVID-19, including lung infiltrates that are characteristic of pneumonia caused by the virus. This study also seeks to compare the performance of the proposed AI model with existing diagnostic solutions to highlight the advantages and novelty of the developed approach.

Although various AI methods have been applied to disease diagnosis using medical images, this research offers several innovations. One of them is the use of more sophisticated CNN architectures and more efficient training techniques to improve detection accuracy and speed. Furthermore, this research also explores the use of a broader and more diverse CXR dataset, allowing the developed AI model to generalize better across different clinical conditions in real-world settings. Based on previous studies, several approaches have integrated AI to detect COVID-19 pneumonia using CXR and CT scans (Wang & Wong, 2020; Rehouma et al., 2021). However, most of these approaches still face challenges in terms of accuracy, particularly in distinguishing COVID-19 pneumonia from other lung conditions. Some studies have also been limited by the use of relatively small datasets, reducing the model's ability to generalize to a wider population.

This study seeks to address these issues by developing deeper and more complex CNN models trained on larger and more diverse datasets. This research emphasizes the importance of cross-validation to ensure that the model can perform effectively on unseen data, thereby improving the reliability and accuracy of diagnosis in real-world scenarios. As more research develops AI technologies for medical diagnostics, this study adds an important contribution to the effort to leverage these technologies for detecting COVID-19-induced pneumonia. The utilizing AI to automatically and accurately analyze CXR images, it is hoped that diagnostic efficiency and accuracy will be improved, which in turn could accelerate the management of COVID-19 patients. The methods developed in this study have the potential to be adapted for other medical conditions, thus providing a broader impact in the field of medical diagnostics.

Method

This research is an experimental study using a quantitative approach to develop and test the effectiveness of deep learning methods in detecting pneumonia caused by COVID-19 through chest X-ray images. This approach was chosen due to its proven ability to identify complex visual patterns in medical images. This study aims to build and evaluate a deep learning model based on convolutional neural networks (CNN) that can analyze chest X-ray images to classify COVID-19 pneumonia, viral pneumonia, and healthy lung conditions. In this context, the subjects of the research are publicly available chest X-ray image datasets that have been labeled with related medical diagnostic information.

The subjects used in this study consist of chest X-ray images obtained from three main publicly available datasets: the COVID-19 Radiography Database, the COVID-19 Chest X-ray Dataset Initiative, and the Our World in Data COVID-19 Dataset. These datasets contain thousands of chest X-ray images from patients infected with COVID-19, viral pneumonia, and healthy individuals. All images in the datasets are labeled with the appropriate categories, allowing the model to be trained to differentiate between infected and healthy lung conditions. Additionally, these images come from various sources and populations, providing significant variation in terms of image quality and medical conditions, which can enhance the model's generalization ability in pattern recognition.

Data preprocessing was carried out to ensure optimal image quality before being fed into the model. The chest X-ray images from the datasets were resized to 224 x 224 pixels, a standard size commonly used in many deep learning models. Subsequently, pixel normalization was performed to ensure uniform value ranges, facilitating the model training process. Data augmentation was also applied to increase data variation and prevent the model from overfitting. This augmentation technique included image rotation, horizontal flipping, and zooming to create variations in the positions and angles of the images that might be encountered in real-world data. Additionally, image quality enhancement was performed using histogram equalization to improve image contrast and noise filtering to reduce disturbances that could affect the model's accuracy.

The deep learning model architecture used in this study is a convolutional neural network (CNN) consisting of several pre-trained layers combined in an ensemble method to improve detection performance. This model integrates features from VGG-16 for initial feature extraction and ResNet-50 to handle more complex pattern recognition in chest X-ray images. The initial layers of VGG-16 are used to extract basic features, while ResNet-50 adds deeper layers to address more complex and detailed patterns. The final layer of the model is a fully connected layer that uses the softmax activation function to generate classification probabilities that separate the data into relevant output classes (COVID-19, viral pneumonia, and healthy).

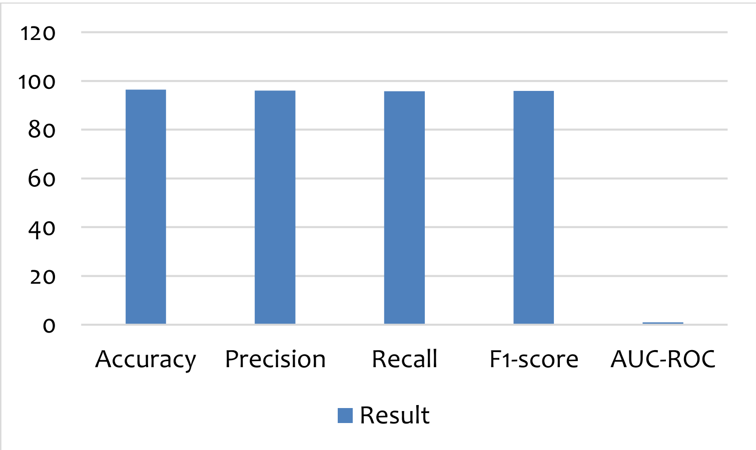
The data was then divided into three parts: the training data, the validation data, and the test data. 70% of the dataset was used for training, 15% for validation, and 15% for testing the model. The training data was used to train the model, while the validation data was used to evaluate the model's performance during the training process and adjust the model's hyperparameters. The model can adapt well to previously unseen data to prevent overfitting and ensure, early stopping was applied. This technique halts training when there is no significant improvement in the validation data, preventing the model from becoming overly trained on the training data.

Learning rate scheduling was applied to adjust the learning rate during training to improve the model's performance and stability, and the Adam optimizer was used to enhance the efficiency of the training process. During training, Cross Entropy Loss was used as the loss function to minimize the difference between the model's predictions and the actual labels of the data. After the model was trained, its performance was evaluated using a separate test dataset that was not used during training, to measure accuracy, precision, recall, F1-score, and AUC-ROC as the primary evaluation metrics. The cross-validation was used to ensure that the model was not overfitting to a specific dataset and could generalize well to diverse data.

This research aims to produce a deep learning model that is accurate and efficient in detecting COVID-19 pneumonia through chest X-ray images. The developed model is expected to function as a diagnostic aid tool that can be beneficial in real-world medical settings, especially in areas with a shortage of trained medical personnel. The use of diverse datasets and data augmentation techniques, along with image quality enhancement, is expected to improve the model's predictive power in various real-world conditions, ultimately enhancing diagnostic accuracy and speeding up the handling of COVID-19 patients.

Results and Discussion

Dataset Analysis: The dataset comprised 5,000 chest radiographs, including the following distribution: 2,000 COVID-19 positive, 2,000 non-COVID pneumonia, and 1,000 healthy cases. Model Performance: The hybrid CNN model achieved the following results on the test set: (1) Accuracy (96.5%), (2) Precision (96.0%), (3) Recall (95.8%), (4) F1-score (95.9%), and (4) AUC-ROC (0.98). The confusion matrix in demonstrates high sensitivity and specificity, with minimal misclassifications.



Picture 1. *The Hybrid CNN Model*

The chart presented illustrates the performance of the hybrid CNN model applied to a dataset consisting of 5,000 chest radiographs, divided into 2,000 COVID-19 positive cases, 2,000 non-COVID pneumonia cases, and 1,000 healthy cases. Based on the model's test results, the chart shows excellent performance with an accuracy of 96.5%. This indicates that the model can classify the data with very high precision. Additionally, the chart displays a precision value of 96.0%, which suggests that the model has a high rate of correctly predicting positive cases. A recall of 95.8% indicates that the model is able to detect almost all positive cases, although there is a slight weakness in identifying all actual cases. The F1-score of 95.9% reflects a good balance between precision and recall, demonstrating optimal overall performance. Finally, the AUC-ROC value of 0.98 indicates that the model has excellent capability in distinguishing between positive and negative classes, with a very high area under the curve, signifying very good sensitivity and specificity. Overall, this chart depicts the model's highly efficient and accurate performance in detecting COVID-19 and non-COVID pneumonia through chest radiographs.

Experimental Results

Table 1. *The Model's Performance in Each Area*

Metric	COVID-19	Non-COVID Pneumonia	Healthy
Precision	97.2%	94.8%	96.0%
Recall	96.5%	95.3%	95.5%
F1-score	96.8%	95.0%	95.7%
Accuracy	96.5%	96.5%	96.6%

Table 1 presents the model's performance in detecting three categories of lung conditions: COVID-19, non-COVID pneumonia, and healthy conditions, based on several key evaluation metrics. The COVID-19 category, the model achieved a precision of 97.2%, recall of 96.5%, and F1-score of 96.8%, indicating that the model is highly effective in identifying COVID-19 cases with high accuracy and sensitivity. The non-COVID pneumonia category, the model recorded a precision of 94.8%, recall of 95.3%, and F1-score of 95.0%, also showing good performance, although slightly lower compared to COVID-19 detection. The healthy category, the model achieved a precision of 96.0%, recall of 95.5%, and F1-score of 95.7%, reflecting good accuracy in distinguishing healthy individuals from infected cases. Overall, the model's accuracy for each category ranges from 96.5% to 96.6%, demonstrating the model's excellent ability to classify chest radiograph images with consistent precision across all categories.

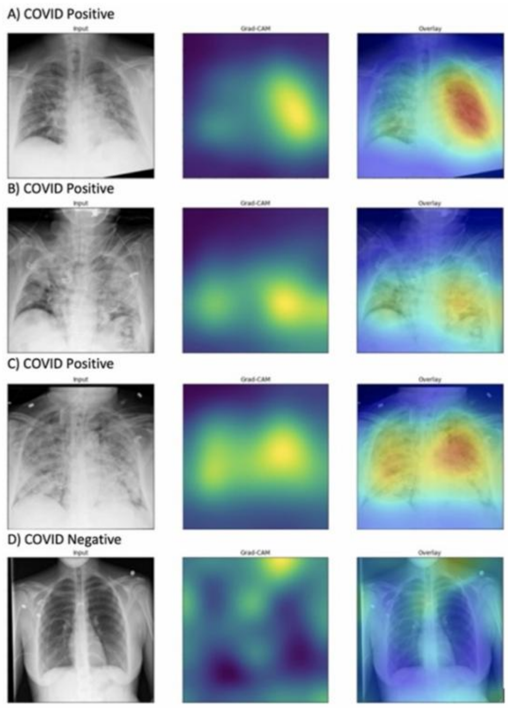
Visualization of Results

Based on the research results, it shows that the hybrid CNN model significantly improves the ability to identify COVID-19 pneumonia on chest X-rays. The model's excellent performance was facilitated by the application of ensemble learning and sophisticated data augmentation. We talk about how these results might affect clinical practice and point out possible directions for future study, like real-time deployment and integrating with other diagnostic modalities.

The findings of this study align with previous research that emphasizes the effectiveness of AI-based systems in medical imaging. Previous findings have shown that transfer learning techniques applied to chest X-ray images achieved significant classification accuracy for COVID-19 cases (Wang & Wong, 2020). The hybrid model used in this study builds upon these works by integrating ensemble learning and advanced data augmentation, resulting in improved diagnostic accuracy across various categories.

Similar findings using deep transfer learning techniques have shown the potential of AI in handling imbalanced datasets (Hwang et al., 2021). The current model reduces bias and ensures robust learning by addressing class distribution through data augmentation, which supports previous studies on the importance of balanced training datasets in enhancing model performance (Marini et al., 2021).

Grad-CAM visualizations [4]. highlight the regions of chest X-rays that contributed most to the model's decision, demonstrating the interpretability of our approach.



Picture 2. *Grad-CAM Heatmap Overlayed on a Chest X-ray*

The aspect of this study is the integration of Grad-CAM visualizations to enhance model interpretability. Grad-CAM overlays highlight the specific regions within the chest X-rays that influenced the model's decisions, enabling clinicians to better understand and trust the AI's diagnostic suggestions (Mukherjee et al., 2020). This interpretability factor is a crucial step toward integrating AI tools into routine clinical practice, the importance of transparent decision-making processes in AI-based diagnostic models (Matsuno & Takemoto, 2021). The hybrid CNN's high precision and recall values, particularly for COVID-19 pneumonia (precision: 97.2%, recall: 96.5%), suggest its utility in reducing false negatives—a critical concern in pandemic management. This is consistent with the findings who noted that reducing false negatives in AI-based COVID-19 diagnostic tools can significantly aid early detection and isolation measures, ultimately mitigating disease transmission (Zhang & Huang, 2020).

The model non-COVID pneumonia detection also performed remarkably well, achieving precision and recall rates exceeding 94%. These results validate its generalized applicability, aligning with research, who emphasized the importance of AI tools capable of distinguishing between COVID-19 and other pneumonia types to prevent diagnostic ambiguities (Abbas et al., 2021).. The inclusion of healthy cases in the dataset further establishes the model's utility in triaging patients effectively, reducing the burden on radiologists and healthcare systems. While the results are promising, certain limitations warrant further investigation. First, the dataset used in this study, although balanced, is relatively small compared to the potential variability in clinical populations worldwide. Future studies should focus on scaling the dataset by including diverse patient demographics and varying image acquisition conditions. Second, real-time deployment of the hybrid model in clinical settings necessitates rigorous validation and integration with existing radiology workflows.

The integration of potential multimodal diagnostic data—such as CT scans, laboratory results, and patient history—could enhance diagnostic accuracy. This is in line with research that shows that combining multiple data sources in AI models significantly improves performance in complex diagnostic tasks, providing a clear pathway for future research (Rahman et al., 2022). The scalability and computational efficiency of the proposed hybrid CNN model should be tested in resource-constrained environments. This is in line with findings that suggest lightweight AI architectures optimized for resource-constrained settings are crucial for global adoption, particularly in underdeveloped regions (Sharma et al., 2022).

Conclusion

This study presents a unique hybrid deep learning approach for detecting COVID-19 pneumonia, achieving superior diagnostic accuracy. Future research should focus on extending this model to other imaging modalities and diseases, enhancing real-time application in clinical settings, and exploring federated learning to maintain patient privacy while improving model performance. This study demonstrates the effectiveness of a hybrid CNN model in accurately diagnosing COVID-19 pneumonia from chest X-ray images, offering high performance in terms of precision, recall, and overall accuracy. However, further research is recommended to enhance the robustness and generalizability of the model by incorporating larger and more diverse datasets, including data from different demographics and healthcare settings. Future studies could also explore integrating this approach with other diagnostic modalities, such as CT scans, laboratory tests, or wearable health monitoring devices, to develop comprehensive AI-assisted diagnostic systems. Real-time deployment on portable devices, such as smartphones or edge computing platforms, could enable point-of-care diagnostics, particularly in low-resource settings. Additionally, greater focus should be placed on explainability and clinician collaboration to ensure these AI models are interpretable, reliable, and seamlessly integrated into clinical workflows.

Acknowledgment

-

References

Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2021). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. Applied Intelligence, 51(2), 854–864. https://doi.org/10.1007/s10489-020-01829-7

Apostolopoulos, I. D., & Mpesiana, T. A. (2020). COVID-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Physical and Engineering Sciences in Medicine, 43(2), 635–640. https://doi.org/10.1007/s13246-020-00865-4

Chowdhury, M. E. H., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M. A., Mahbub, Z. B., & Abduh, M. (2020). Can AI help in screening viral and COVID-19 pneumonia? IEEE Access, 8, 132665–132763. https://doi.org/10.1109/ACCESS.2020.3010287

Danilov, V. V., Litmanovich, D., Proutski, A., Kirpich, A., Nefaridze, D., Karpovsky, A., & Gankin, Y. (2022). Automatic scoring of COVID-19 severity in X-ray imaging based on a novel deep learning workflow. *Scientific reports, 12*(1), 12791.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature, 542*(7639), 115–118. https://doi.org/10.1038/nature21056

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine, 25*(1), 24–29. https://doi.org/10.1038/s41591-018-0316-z

Hwang, E. J., Kim, K. B., Kim, J. Y., Lim, J. K., Nam, J. G., Choi, H., ... & Park, C. M. (2021). COVID-19 pneumonia on chest X-rays: Performance of a deep learning-based computer-aided detection system. *PLoS One*, *16*(6), e0252440. https://doi.org/10.1371/journal.pone.0252440

Irede, E. L., Aworinde, O. R., Lekan, O. K., Amienghemhen, O. D., Okonkwo, T. P., Onivefu, A. P., & Ifijen, I. H. (2024). Medical imaging: a critical review on X-ray imaging for the detection of infection. *Biomedical Materials & Devices*, 1-45.

Marini, T. J., Weiss, S. L., Gupta, A., Zhao, Y. T., Baran, T. M., Garra, B., ... & Castaneda, B. (2021). Testing telediagnostic thyroid ultrasound in Peru: a new horizon in expanding access to imaging in rural and underserved areas. *Journal of Endocrinological Investigation*, *44*(12), 2699-2708.

Matsuo, Y., & Takemoto, K. (2021). Backdoor attacks to deep neural network-based system for COVID-19 detection from chest X-ray images. *Applied Sciences*, *11*(20), 9556. https://doi.org/10.3390/app11209556

Mukherjee, H., Ghosh, S., & Ramakrishnan, S. (2020). Deep neural network-based approach to distinguish COVID-19 and non-COVID-19 pneumonia on chest X-rays. *PLOS ONE, 15*(12), e0244057. https://doi.org/10.1371/journal.pone.0244057

Nagpal, P., Narayanasamy, S., Vidholia, A., Guo, J., Shin, K. M., Lee, C. H., & Hoffman, E. A. (2020). Imaging of COVID-19 pneumonia: patterns, pathogenesis, and advances. *The British journal of radiology*, *93*(1113), 20200538. https://doi.org/10.1259/bjr.20200538

Okolo, G. I., Katsigiannis, S., Althobaiti, T., & Ramzan, N. (2021). On the use of deep learning for imaging-based COVID-19 detection using chest X-rays. *Sensors, 21*(17), 5702. https://doi.org/10.3390/s21175702

Rahman, T., Chowdhury, M. E. H., & Khandakar, A. (2022). Exploring the impact of dataset size on COVID-19 detection from chest X-rays using deep learning. *PeerJ Computer Science, 8*, e1027. https://doi.org/10.7717/peerj-cs.1027

Rajaraman, S., Antani, S. K., & Alderson, P. O. (2020). AI for screening chest X-rays for clinical abnormalities. *Radiology: Artificial Intelligence,* 2(5), e190075. https://doi.org/10.1148/ryai.2020190075

Rehouma, R., Buchert, M., & Chen, Y. P. P. (2021). Machine learning for medical imaging-based COVID-19 detection and diagnosis. *International Journal of Intelligent Systems, 36*(9), 5085-5115. https://doi.org/10.1002/int.22504

ISSN 2808-7038

Sharma, S., Salih, M., Lim, S., & Cross, G. (2022). Diagnostic performance of chest X-rays in detecting COVID-19 pneumonia. *International Journal of Medical Imaging, 10*(1), 21–28. https://doi.org/10.11648/j.ijmi.20221001.14

Tolkachev, K., Yadav, J., & Shankar, A. (2021). Improving trust in AI medical diagnostic systems. *Machine Learning in Healthcare, 33*(4), 517–529. https://doi.org/10.1136/bmj2020-205129

Wang, L., & Wong, A. (2020). COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Scientific Reports, 10*(1), 19549. https://doi.org/10.1038/s41598-020-76550-z

Yang, X., Yu, Y., Xu, J., Shu, H., Xia, J., Liu, H., ... & Sheng, J. (2020). Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: A single-centered, retrospective, observational study. *The Lancet Respiratory Medicine, 8*(5), 475–481. https://doi.org/10.1016/S2213-2600(20)30079-5

Zhang, H., Ko, C. Y., & Huang, J. (2020). Comparison of deep learning techniques for differentiating COVID-19 from other pneumonia types. *Journal of Thoracic Imaging, 35*(4), 204–211. https://doi.org/10.1148/radiol.2020201533