IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# Covid-19 Prediction With CT Scan Image Of Lungs Using Convolutional Neural Network

Dr. S. Sahaya Tamil Selvi

Associate Professor and Head,
Department of Computer Science
St. Joseph's College for Women, Tirupur, Tamilnadu, India

***Abstract:*** The virus causing corona disease is transmitted through droplets that are released by an infected person when he/she coughs, sneezes, or exhales. People get infected by either breathing in the droplets through atmosphere if they are released by an infected person nearby or by touching a surface which has been contaminated by the virus. Most of the people who get infected suffer mild to moderate symptoms. However, people with severe symptoms develop Acute Respiratory Distress Syndrome (ARSD) characterized by rapid onset of inflammation in the lungs. It can cause serious conditions such as blood clots, multi-organ failures etc. to happen suddenly sometimes even leading to death of the patient. In other words, COVID-19 affects our lungs with adverse infection due to which a person is unable to breathe with decorum. The broad aim of this research paper is to study the CT scan images of patients infected from coronavirus and compare them to the CT scan images of patients which may be not be infected or may have normal pneumonia (i.e. not due to coronavirus) using the classifier as Convolutional Neural Network.

***Index Terms*** **- COVID-19, Convolutional Neural Network, CT scan images, Pooling Layer, Fully connected layer**

## I. INTRODUCTION

In recent times COVID-19 has had a tremendous effect on the lives of individuals globally. In the past, there were many such viruses but nothing as strong as this Coronavirus. People sure did not anticipate facing a virus like this in the present age of technology. Evaluating the time required for the diagnostic process and the expense of the kits needed to test cases, AI and deep learning-based application was required to help support doctors working on constraining the spread of the virus and treating patients combating the illness.

The Covid 19 pandemic has been spreading rapidly throughout the world since last year. The rapid spread of Covid-19 and the intensity of the symptoms have claimed the lives of numerous humans. Individuals have lost employment, family members and friends. It has put society in a quarantine. One of the issues we are facing while fighting this pandemic is the lack of quick and affordable diagnosis. COVID19 diagnosis can take a lot of time and can be expensive for some people. In this system, we intend to use Lung X-Ray scans to identify and classify them into COVID and non-COVID images. This reduces both the cost and time taken for COVID19 diagnosis.

## II. RELATED WORK

Sharma et al. [3] created an effective model that is trained on chest X-ray images using deep learning for quick screening of COVID-19 patients. Wu et al. [6] suggested explainable COVID-19 diagnosis, in which real-time diagnosis is carried out using the Joint Classification and Segmentation (JCS) system. Deep Pneumonia, a deep learning-based CT diagnosis algorithm to find patients with COVID-19, was created by Song et al. [4]. The model is able to distinguish between COVID-19-infected patients and patients with bacterial pneumonia. A modified Inception transfer-learning model was proposed by Wang et al. [5], and both internal and external validation were completed. To develop the algorithm, the author gathered CT pictures from patients with pathogen-confirmed COVID-19 instances as well as images from those with normal viral pneumonia.

## III. PROPOSED SYSTEM

The scope of this paper is to classify X-ray images and distinguish them as Covid positive or Normal using machine learning algorithms based on image datasets. The dataset of this work has been collected from the Kaggle repository, which has been collected from multiple sources and contains Chest X-Ray scans of Covid-19 affected and normal lungs. This collected dataset is not meant to claim the diagnostic ability of any Deep Learning model but to research various possible ways of efficiently detecting Coronavirus infections using computer vision techniques. The features of the system are:

- The application will detect covid automatically.
- The application is user friendly and easily accessible.
- It will save time, reduce the effort of the user.
- Provide cheaper and accurate systems to the user which can be easily accessible and give accurate recognition of Covid disease.

It can be applicable to every economic level.

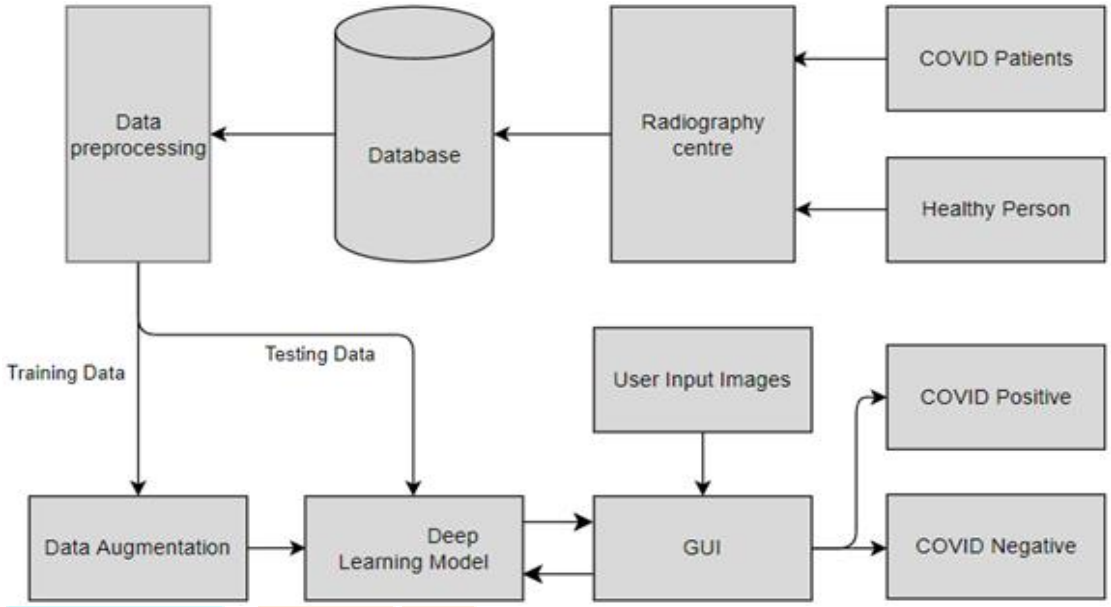


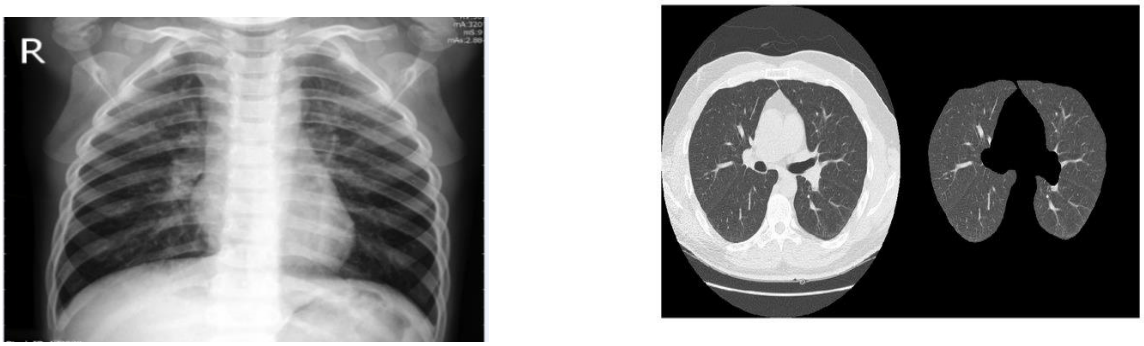

Fig.1 Block Diagram of the proposed system.

### 3.1 Data and Sources of Data

The COVID-19 radiography dataset collected from kaggle.The dataset is open source and can be used by anyone. It contains chest x-ray images of patients. The data is formatted such that it contains folders for each of the three classes and metadata files giving attributes of the data such as name, file type(PNG), file path and image size in pixels. The dataset does not have any CSV file with image data along with their labels. The dataset contains 219 images of COVID-19 patient's lungs, 1341 images of healthy patient's images and 1345 images of viral pneumonia patients. These images are 1024 X 1024 pixels with all three channels that are RGB.

### 3.2 Image Preprocessing

It includes resizing, cropping and removal of noise from the given image. The preprocessing step aims to enhance some image features that are required for further processing. This step includes the removal of some undesired features from the given image. For example, the background and an irrelevant portion of the image is discarded to reduce image processing time.

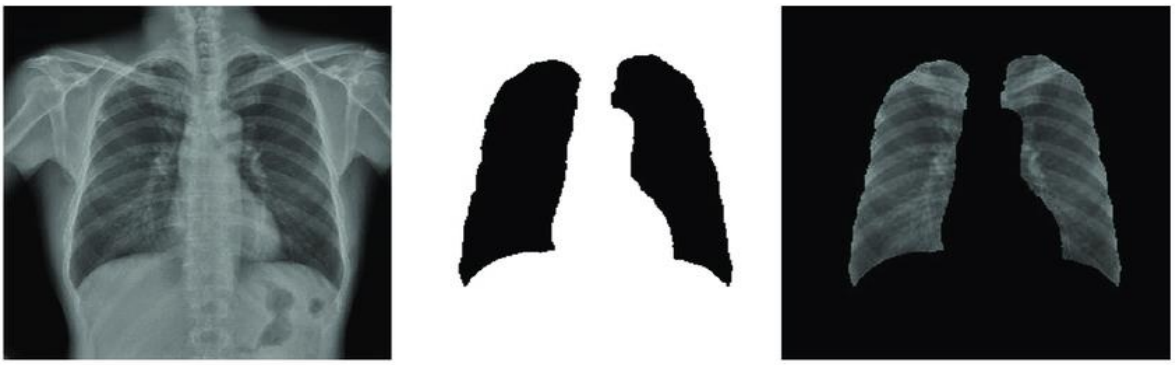



Chest X-ray image Preprocessed image

Fig.2 Preprocessing process.

### 3.3 Segmentation

Segmentation is an important step in object recognition tasks. It transforms images into a form that is more meaningful and less complex to analyze. Here, an image can be divided into some regions based on the desired feature. A watershed segmentation technique was applied to segment the fracture regions of each image owing to its relatively less computational complexity and capability of providing high accuracy in segmentation. This method separated touching objects in an X-ray image and provided a complete division. Fig.10 shows the process of segmentation.



a. Chest X-ray image b.Area to be Segmented c.Segmented image

Fig.3 Segmentation process.

**3.4 Feature Extraction**

Feature Extraction is the process extorts the features from the segmented based on shapes, colors, and textures. Some shape-based features are area, axis, and angle. Feature extraction in image processing is a method of transforming large redundant data into a reduced data representation. Transforming the input data into the set of features is called feature extraction. Data fusion was applied in several machine learning and computer vision applications. Particularly, feature fusion can combine more than one feature vector. Two feature extractors provide a feature vector of 1 × 4096 and 1 × 3780. Fig.11 represents the features in the lung scan.

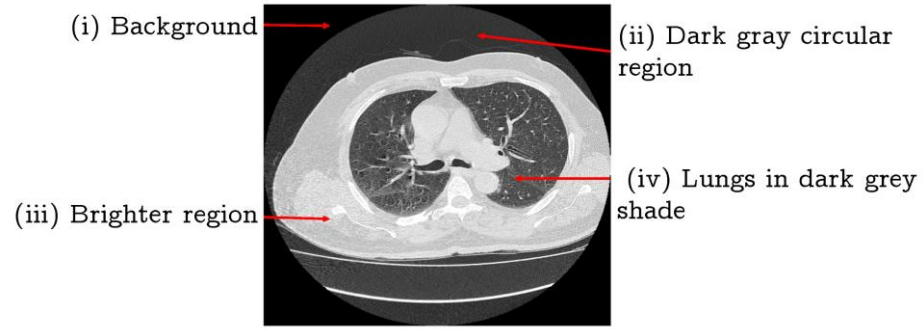



Fig.4 Feature Extraction process.

**3.5. Classification**

Classification is an important module in lung disease detection systems. It is defined as a process of categorizing Lung images based on identified diseases. In this system, Keras- Convolution Neural Network (CNN) is used. Fig.12 shows the classification of three classes they are Covid-19, Normal and Pneumonia.

**3.5.1 Convolutional Neural Networks (CNN)**

Convolutional Neural Networks (CNN's) are a class of Deep Neural Networks that can recognize and classify particular features from images and are widely used for analysing visual images. Their applications range from image and video recognition, image classification, medical image analysis, computer vision and natural language processing.

The term 'Convolution" in CNN denotes the mathematical function of convolution which is a special kind of linear operation wherein two functions are multiplied to produce a third function which expresses how the shape of one function is modified by the other. There are two main parts to CNN architecture:

- A convolution tool that separates and identifies the various features of the image for analysis in a process called as Feature Extraction
- A fully connected layer that utilizes the output from the convolution process and predicts the class of the image based on the features extracted in previous stages.Fig.13 Shows the two main parts of CNN architecture.

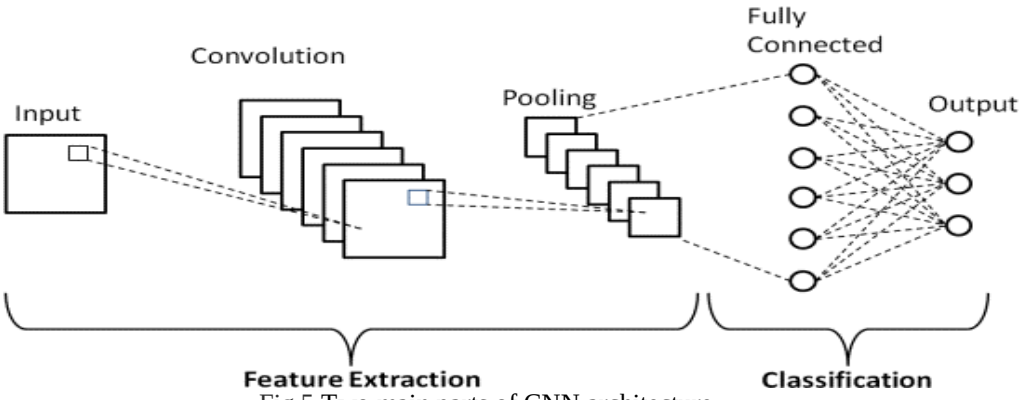



Fig.5 Two main parts of CNN architecture

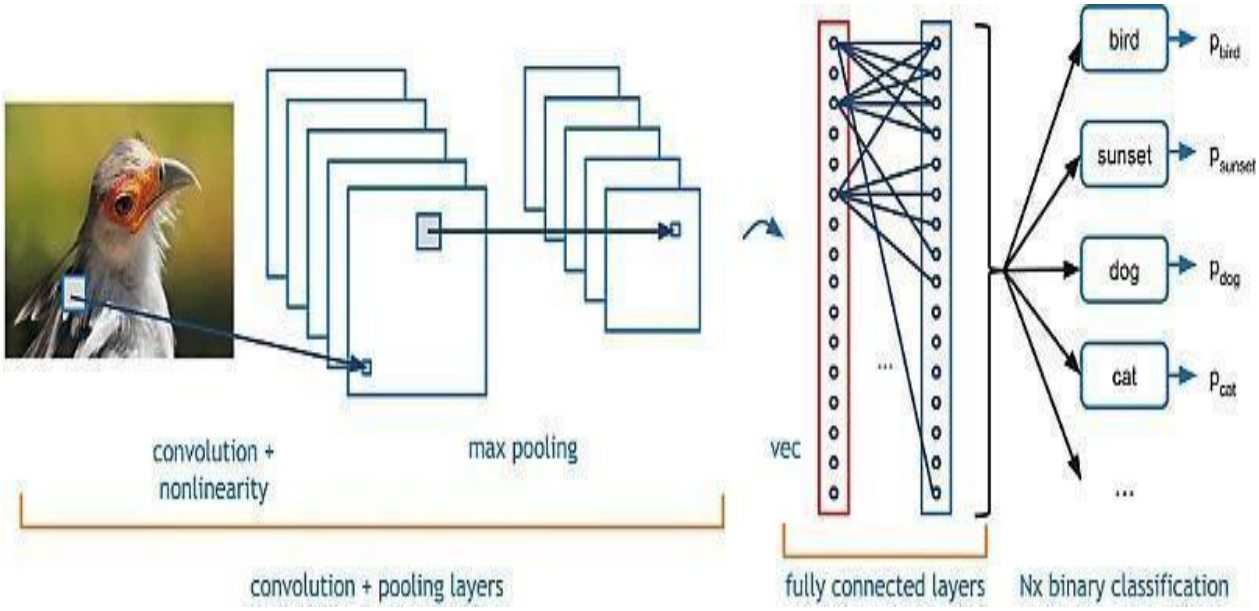



Fig.6 Feature Extraction process.

Convolutional Neural Networks have several types of layers:

- **Convolutional layer**: A "filter" passes over the image, scanning a few pixels at a time and creating a feature map that predicts the class to which each feature belongs.
- **Pooling layer (down sampling)**: Reduces the amount of information in each feature obtained in the convolutional layer while maintaining the most important information.
- **Fully connected input layer (flatten)**: Takes the output of the previous layers, "flattens" them and turns them into a single vector that can be an input for the next stage.
- **The first fully connected layer**: Takes the inputs from the feature analysis and applies weights to predict the correct label.
- **Fully connected output layer**: Gives the final probabilities for each label

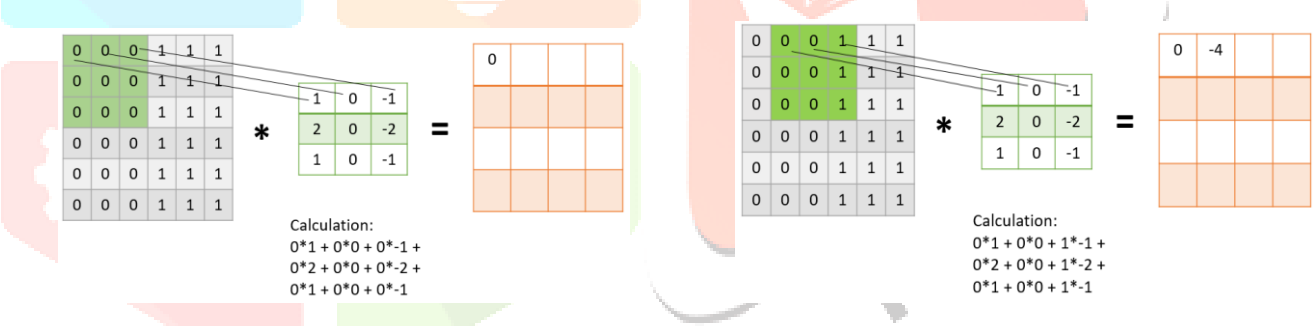



Fig.7 Convolution process (a) Multiplication of pixel (b) Addition part of pixel

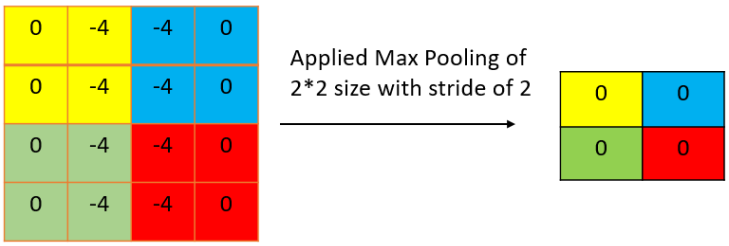



Fig.8 Max pooling process.

### 3.5.2 Train and Test set

For training and testing purposes of our model, we should have our data broken down into three distinct dataset splits.

#### 3.5.2.1 The Training Set

It is the set of data that is used to train and make the model learn the hidden features/patterns in the data.
In each epoch, the same training data is fed to the neural network architecture repeatedly, and the model continues to learn the features of the data. The training set should have a diversified set of inputs so that the model is trained in all scenarios and can predict any unseen data sample that may appear in the future.

**3.5.2.2 The Test set**

The test set is a separate set of data used to test the model *after* completing the training.It provides an unbiased final model performance metric in terms of accuracy, precision, etc.

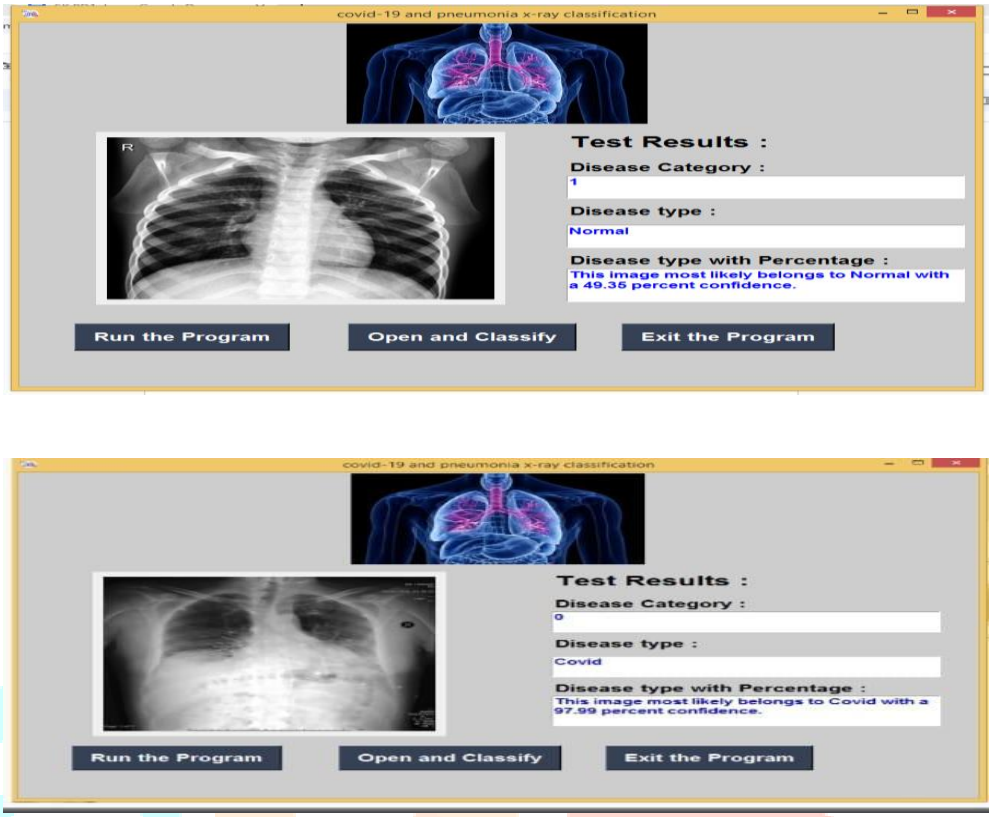



Fig.9 Results of the system.

## IV. RESULTS AND DISCUSSION

After preprocessing of the dataset, the final dataset consisted of a total of 285 X-ray images. For training and testing the proposed CNN, the dataset was partitioned into two subsets. The training dataset contained 95 COVID-19 X-ray images , 95 normal X-ray images, and 95 Pneumonia X-ray images, making a total of 285 X-ray images. The testing dataset similarly contained 210 X-ray images, in which 70 X-ray images were from each class COVID-19 , Normal and Pneumonia. Then, the training subset containing 210 X-ray images has been passed to the model with 25% validation size. So, out of 285 X-ray images, with each epoch, 1208 X-ray images train the model, and 75 X-ray Images validate the model. As mentioned in the proposed architecture of the CNN model, it consisted of 38 layers in which 6 are convolutional, 6 max pooling layers, 6 dropout layers, 8 activation function layers, 8 batch normalization layers, 1 flattening layer, and 3 fully connected layers. The CNN model thus achieved an extraordinary performance with an accuracy of 93 % with the test data subset used from the processed dataset of this system.

## V. CONCLUSION

This model can be used as an efficient tool for early detection and diagnosis of covid-19 which is boon for the medical field. Using the deep learning models could open new insights for detecting and improving testing and treatment for covid-19 patients. Our active experimentation is the effective implementation of the system, evaluation parameters, and GUI execution. COVID-19 was detected utilizing chest X-ray pictures using a diagnostic algorithm based on CNN. With the use of an augmented dataset, the model was able to detect COVID-19 quickly and reliably.

## VI. ACKNOWLEDGMENT

Dr. S. Sahaya Tamil Selvi is an Associate Professor and Head in department of Computer Science, St. Joseph's College for Women, Tirupur with over 22 years of experience. With area of interest in Artificial Intelligence and Digital Image Processing, she has published 10 research articles in various reputed national and international journals and has presented 10 research papers in national and international Conferences and Seminars. She has published a book named "A Comprehensive Handbook of C++, Java and Python". She is an Editorial Board Member for various national /international journals.

## REFERENCES


**[1]** Zhu N., Zhang D., Wang W., Li X., Yang B., Song J. A novel coronavirus from patients with pneumonia in China, 2019. N Engl J Med. 2020;382(8):727–733.

**[2]** Nardelli P., Jimenez-Carretero D., Bermejo-Pelaez D., Washko G.R., Rahaghi F.N., Ledesma-Carbayo M.J. Pulmonary artery-vein classification in CT images using deep learning. IEEE Trans Med Imaging. 2018;37(11):2428–2440.

**[3]** Sharma A, Rani S, Gupta D. Artificial intelligence-based classification of chest X-ray images into COVID-19 and other infectious diseases. Int J Biomed Imaging 2020;2020:10. https://doi.org/10.1155/2020/8889023. Article ID 8889023pages.

**[4]** Song Y, Zheng S, Li L, Zhang X, Zhang X, Huang Z, et al. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. medRxiv 2020.

**[5]** Wang S, Kang B, Ma J, Zeng X, Xiao M, Guo J, et al. A deep learning algorithm using CT images to screen for corona virus disease (COVID-19). medRxiv 2020.

**[6]** Y.H. Wu, S.H. Gao, J. Mei, J. Xu, D.P. Fan, C.W. Zhao, M.-M. Cheng. JCS: an explainable COVID-19 diagnosis system by joint classification and segmentation. 2020. arXiv preprint arXiv:2004.07054

**[7]** Cohen J., Normile D. New SARS-like virus in China triggers alarm. Science. 2020;367(6475):234–235.

**[8]** Corman V.M., Landt O., Kaiser M., Molenkamp R., Meijer A., Chu D.K.W. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Euro Surveill. 2020;25(3):23–30.

**[9]** Huang C., Wang Y., Li X., Ren L., Zhao J., Hu Y. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497–506.