International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

# Deep Learning for COVID-19 Prediction and Lung Infection Segmentation from CT Images

**Yogita Kajabe[1], Dr. Brijendra Gupta[2]**

[1]Student, Information Technology, Siddhant College of Engineering
[2] Professor, Information Technology, Siddhant College of Engineering

**Abstract**

Lung disorders like pneumonia, cancer, and COVID-19, a common and potentially deadly respiratory infection, are very challenging to identify accurately and promptly. Using advances in machine learning (ML) techniques, this research aims to produce a reliable and efficient pneumonia detection system. Chest CT images are used to build a convolutional neural network (CNN) model to differentiate between lung pictures with pneumonia and those that are healthy. Training and validation datasets are a diverse collection of chest CT images collected from multiple sources. With strong sensitivity and accuracy, the proposed approach detects pneumonia with promising results. COVID-19 can also cause long-term harm to the lungs and other organs. When an infected individual coughs, sneezes, or exhales, droplets are released that are the main way the COVID-19 virus is transmitted. Being too heavy to float in midair, these drops quickly land on floors or other objects. The corona virus disease 2019 (COVID-19) started to spread globally in early 2020, as everyone is aware. causing a global health crisis associated with existential issues. Thus, automating the diagnosis of lung infections from computed tomography (CT) images presents a significant possibility to improve the traditional healthcare approach to COVID-19. Nonetheless, there are several challenges in distinguishing infected regions from CT slices, including the broad variety of infection characteristics and the poor contrast between healthy tissues and infections. Moreover, collecting large amounts of data in a short period of time is impracticable, which hinders the training of a deep model. Our proposed technique would analyze the CT scan of the lung to determine the affected lung part and the infected region. The technology will assess the intensity of the inflection and help patients take the appropriate response.
.

**Keywords:** CT images, lung infections, pneumonia, COVID-19, convolutional neural networks, and machine learning

## 1. Introduction

Currently, medical imaging analysis is growing more and more prevalent in the sector, especially in clinical evaluation and non-invasive treatment. These restorative images, including ultrasonic imaging, x-rays, computer tomography (CT), and so forth, are employed for a comprehensive diagnosis. Among medical imaging methods, CT is an essential filtering tool that utilizes the fascinating domains for image collection. A study found that one of the main causes of the 1.61 million deaths that take place there annually is lung cancer. A different study found that the chance of survival is increased when cancer is detected early. Lung cancer ranks second among males and tenth among women globally, making it the

International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

second most frequent disease. One of the most dangerous disease in world is Lung cancer. However, early diagnosis and treatment can save a life.

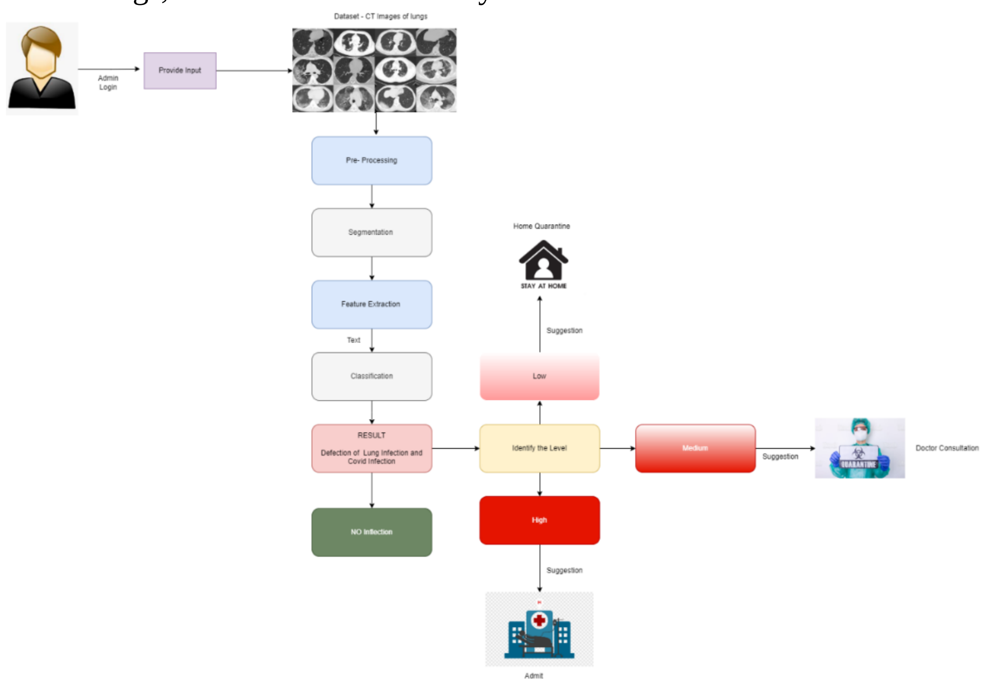
It is difficult for clinicians to interpret and identify cancer from CT scan data, even though it is the most efficient imaging technique in the medical field. Computer-aided diagnostics can therefore help doctors accurately detect cancerous cells. The number of deaths from lung cancer is greater than the number of deaths from breast, colon, and prostate cancer combined. taking into account the current situation, in which we are fighting a disease like corona, which significantly affects the lungs.

COVID-19 has had a significant impact on people and healthcare systems around the world. Computed tomography scans can be a helpful tool in addition to reverse transcription-polymerase chain reaction testing. This study employed a convolutional neural network to screen for COVID-19. We examined the performance of different pre-trained models on CT testing and discovered that utilizing larger, out-of-field datasets increases the models' testing capability. This suggests that a priori knowledge of the models acquired from out-of-field training can also be advantageous for CT images.

The recommended approach to transfer learning performs better than the current approaches covered in the literature. We believe that the identification performance that our approach has so far provided is at the forefront of the field. Our model performs satisfactorily in experiments using randomly chosen training datasets. The relevant visual characteristics of the CT scans used by the model were examined; they could aid medical practitioners in manual screening.

## 2. System Implementation

In our system's admin module allows for the submission of a CT image. To give us the best outcome, the system will further examine the uploaded image and perform several operations, including pre-processing, segmentation, feature extraction, classification, and more. He will be informed by the system, which will also make recommendations, including the necessity of emergency care or a straightforward home quarantine. There isn't a system in place right now that can analyze CT scans, generate an alert based on the findings, and enable the necessary action to be taken.

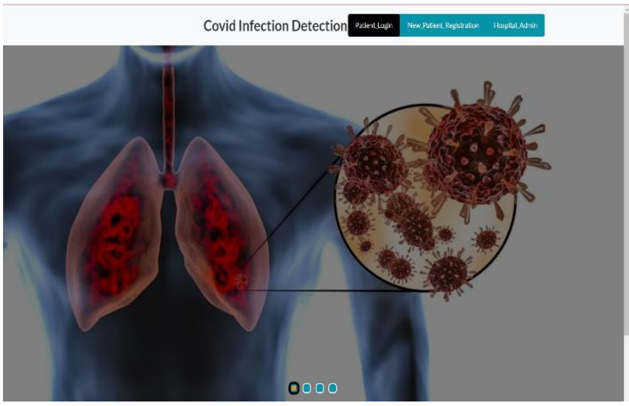


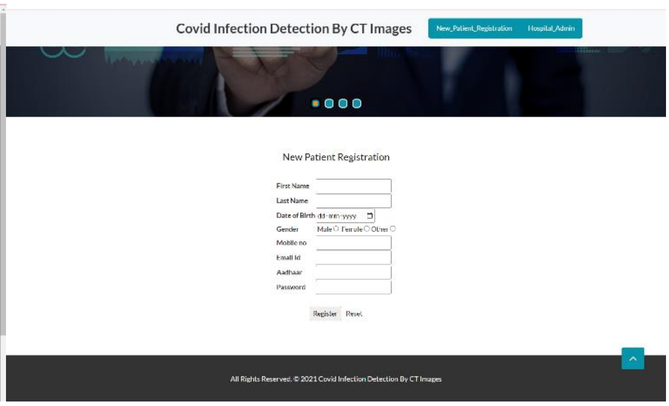
## 3. Module Description

- The lung CT scan can be uploaded by the administrator.
- The system will do pre-processing, segmentation, feature extraction, and classification on those CT images.
- The output will be shown by the system at last. For example, whether the lung is infected.
- **Output:** Depending on the severity of the infection, the system will show the user whether they have COVID-19 or lung cancer and, if so, what they need to do.

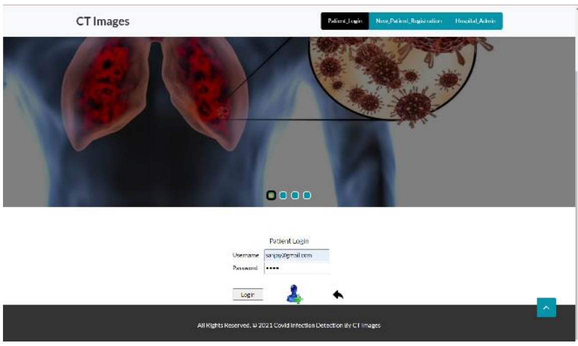
## 4. Details of Modules Implementation

Developing a system with three primary activities is the aim of the current study. In order to enable radiologists to input CT scan images of patients with lung infections brought on by cancer, COVID-19, etc., a web-based application will first be created. Radiologists will also be able to determine the illness stage by analyzing the patient's symptoms. The photos will then be pre-processed and segmented by the system. Utilizing the watershed method, the CT scan provides the generated pictures to a convolutional neural network in order to detect lung infections. This method provides a faster and more accurate method of measuring nodules, which ultimately leads to improved staging of lung infection by considering the patient's symptoms.




Above are screenshots of Home page and Registration Form. User can fill the form and create an account in our System.



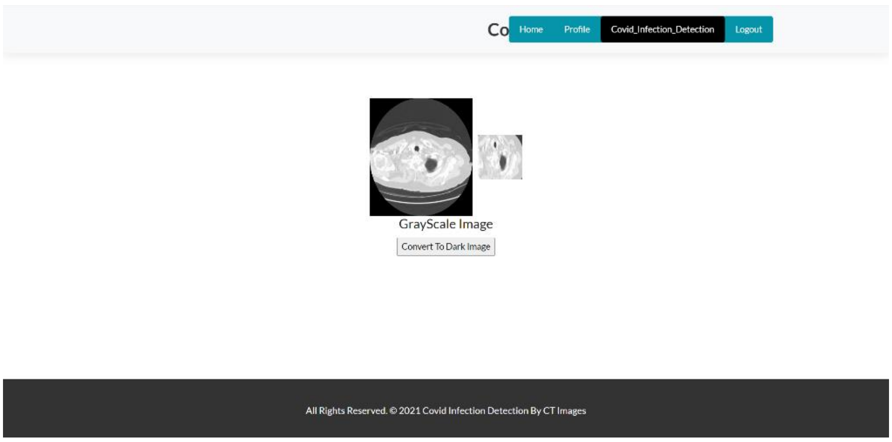


Once Registration is successful, User can login and browse the CT image for the checking infection in the lungs.

Then our System will start below different processing techniques to identify the infection.

**Image Preprocessing Phase:**

Computed tomography (CT) scans are an essential diagnostic and treatment planning tool in modern medicine. However, they often include ionizing radiation exposure, which can be detrimental to the lungs or other organs. Pre-processing methods are therefore used to provide excellent images with the least amount of radiation exposure. One common pre-processing technique for CT images is filtering, which is removing unwanted noise and features from the images. One common filtering technique used in CT image pre-processing is the Gaussian filter. Noise and other high-frequency information are removed from an image using a smoothing filter known as the Gaussian filter. The image is convolution ally transformed using a Gaussian kernel, a mathematical function, to produce a bell-shaped curve. Larger curves result in more smoothing; the width of the curve determines how much smoothing is applied to the image. For CT scans in particular, Gaussian filtering is useful since it preserves edges while reducing noise. The Gaussian filter is a low-pass filter, which means that high-frequency information is lowered but low-frequency information is maintained. In CT scans, low-frequency information depicts the underlying architecture of the lung tissue, whereas high-frequency information displays noise and other undesirable features. For effective diagnosis and treatment planning, pre-processing creates a clearer, more accurate image of the lung tissue. By eliminating noise and higher frequency data, filtering procedures improve the clarity of the CT scan image and help doctors make more accurate diagnosis and treatment decisions.




**Segmentation:**

The second stage of image segmentation is pixel or voxel separation, which is achieved by extracting the pixel array and associated metadata from the DICOM file. The Hausfield units generated from the recovered pixel array are then used to calculate the radiodensity of the scanned tissue. Because they help differentiate lung tissue from other thoracic cavity components, Hounsfield units are crucial for picture segmentation, especially when it comes to lung segmentation. The internal, external, and watershed markers required for marker-controlled segmentation must be prepared by calculating Hounsfield units. The marker-controlled segmentation technique uses these markers along with the Sobel gradient to derive lung segments from pictures. Two essential processes in the image segmentation process are resizing and resampling. While scaling involves altering the image's dimensions to match the desired output size, resampling involves changing the voxel size and spacing to accommodate for differences in acquisition settings across different CT scans. The watershed segmentation approach is widely used to refine the segmentation results. This technique uses Hounsfield units and gradient information to first

International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

identify regions of interest, and then it uses a marker-based watershed transformation to split these regions. Pixel or voxel separation, Hounsfield unit computation, marker-controlled segmentation using Sobel gradient and watershed segmentation, as well as resizing and resampling, are all steps in the image segmentation process that are necessary to generate accurate and trustworthy results.


**Feature Extraction Phase:**

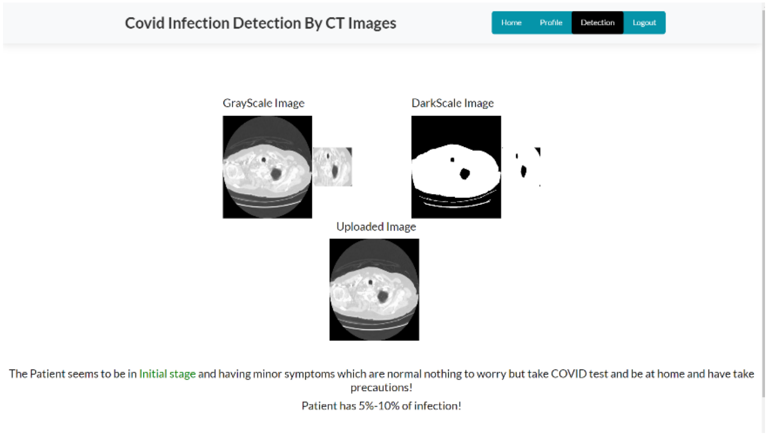
Even though many types of pre-trained CNN models were used to extract the features, the statistical results indicated that DenseNet-169 was the best model for the feature extraction stage. This stage therefore focuses on explaining the DenseNet-169 model architecture and its function in feature extraction.

**Classification Phase:**

For the classification challenge, numerous classifiers, such as Random Forest and Support Vector Machine, were used after feature extraction. But the best results were found when Support Vector Machine was used as the problem's classifier. Better results were obtained by combining DenseNet-169 features with an SVM classifier in the best proposed model. An explanation of the SVM kernel and parameters is provided below:

Assume that a set of training data, such (x1,y1), (x2,y2),..(xn,yn), needs to be split into two sets of classes. The label class is represented by yi ε (0,1), and the feature vector is represented by xi ε Fd. By determining the ideal hyperplane for the training data displayed above—that is, the one with the greatest margin between the classes—a support vector machine used for binary classification can distinguish between the data points of one class and the other. The performance of SVM is greatly influenced by the kernel and parameter selections.

Significant effects on SVM performance are caused by the gamma and C parameters of the RBF kernel. It makes intuitive sense to utilize the gamma parameter, whose smaller value denotes "far" and whose greater value denotes "close," to calculate the appropriate level of influence for single training examples. The gamma parameter thus shows the inverse of the radius of effect of the samples selected by the model as support vectors. On the other hand, the C parameter compensates for the training sample misclassification. A low C provides a smooth surface, whereas a high C strives to accurately classify every training sample by giving the model permission to select extra samples as support vectors.




## 5. Result and Discussion

When a user or administrator uploads a CT picture to our system, a number of pre-processing procedures are applied to the submitted image. In addition, procedures including segmentation, feature extraction, and classification will be used to determine whether or not the CT image contains a COVID infection. If the user tests positive for COVID, the system will determine the severity and recommend the appropriate course of action, such as hospitalization or home quarantine. Furthermore, no additional action is required if the outcome is unfavorable.

One hundred patients' CT scans were used in this study.

| No of Test Images (Slices) | Correctly detected | Accuracy (%) | Computation Time |
|---|---|---|---|
| 20 × 20 = 400 | 320 | 80% | 120 Seconds |

Eighty patients are utilized for training, while the remaining patients are used for testing. Twenty slices per patient, or 20 × 20 = 400 slices, are employed for testing, and these numbers are higher than those used in the other earlier studies.
Above table shows the precision and processing time of our suggested detection system. We used a computer running 64-bit Windows 10 OS with an Intel Core i5-7200U CPU running at 2.50 GHz, Intel HD Graphics 4000, and 16 GB of RAM to complete the calculation. We have attained a detection accuracy of almost 80%.

## References

1. H.J. Aerts, T. Coroller, I. Franco, C. Parmar, R. Zeleznik, A. Hosny, and Y. Xu Using serial medical imaging, deep learning forecasts the response to treatment for lung cancer.
2. F.H. Khan, Z.S. Khan, S. Bashir, and M.I. Faisal An assessment of machine learning ensembles and classifiers for lung cancer early stage prediction
3. Using deep learning for classification and mutation prediction from non-small cell lung cancer histopathology images, N. Coudray, P.S. Ocampo, T. Sakellaropoulos, N. Narula, M. Snuderl, D. Fenyö, …, A. Tsirigos
4. "The role of chest imaging in patient management during the COVID-19 pandemic: A multinational consensus statement from the Fleischner society," by G. D. Rubin, L. B. Haramati, and colleagues, Radiology, p. 201365.

5. The Lancet, vol. 395, no. 10223, pp. 497–506; Huang, Y. Wang, et al., "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China."
6. V. Vanitha, R Srinidhi, R Darshana, and S K M Nivetha, 022 IEEE 3rd Global Conference, COVID-19 Detection using CT-Scan.
7. Wenzhi Lv 1, Qian Tao, Ziyong Sun, Liming Xia, and Tao Ai, "Correlation of chest CT and rt-pcr testing in coronavirus disease 2019(COVID-19) in China: A report of 1014 cases," Radiology, vol. 2019, p. 200642.