# Deep Convolutional Neural Networks and Transfer Learning Based Approach for Lung Cancer Detection from CT Scan Images

*Md. Ahasanul Islam, Samreen Sohail, Shahnewaz Siddique*
Department of Electrical and Computer Engineering, North South University, Dhaka-1229, Bangladesh
Email: islamahsanul79@gmail.com, samreensohail03@gmail.com, shahnewaz.siddique@northsouth.edu

***Abstract*—A recent WHO survey ranked lung cancer as the second most frequent reason for death worldwide. As a result, early detection of lung cancer has become essential, as early detection can save many lives. Modern approaches to deep learning and image processing have made the detection and treatment of lung cancer much more straightforward. The ability of convolutional neural networks (CNNs) and transfer learning-based models to identify Cancer from CT scan pictures are investigated in this research. Deep CNN models already trained, like ResNet50, MobileNetV2, and VGG19, were employed to extract the deep features. Two datasets were collected and merged to increase training and validation datasets. MobileNetV2 had a 97% accuracy rate. VGG19 had an almost 100% accuracy rate, but the Model suffered from overfitting. ResNet50 had a 78% accuracy rate, and CNN had 99% accuracy with overfitting issues. Thus among these networks, MobileNetV2 provided the highest level of accuracy. These accuracies assist in detecting malignancies early before they cause physical side effects such as paralysis and other complications.**

***Index Terms*—Lung Cancer, CT Scan Images, Custom CNN model, Transfer Learning, MobileNetV2, ResNet50, VGG-19**

## I. Introduction

One of the harshest medical disorders in the world is considered to be lung cancer. WHO estimated that Cancer caused approximately 10 million deaths in 2020. Lung cancer caused 2.21 million deaths or nearly 22.1%. [1] Lung cancer causes tumor growth when lung cells multiply out of control. These may interfere with breathing and spread to various body areas. Because Cancer is most effectively treated when discovered in its early stages, screening for the disease is crucial to preventive healthcare. Computed tomography (CT) is one of the most essential instruments for diagnosing lung cancer. However, this screening technique necessitates the expertise of a qualified specialist, which is not always available. [2] Since its introduction in the early 1960s, digital processing in medicine has reduced processing times and improved specificity and sensitivity. [3] Another issue for the health community is the steadily rising size and variety of medical data modalities. Computer algorithms are the best options for offering the ability to consume and analyze these data. These algorithms' use in medicine has increased steadily during the past few years. [4] The preprocessing, segmentation, feature extraction, and classification steps in computerized medical data analysis are typically included. Lung cancer results from uncontrolled lung cell proliferation, which promotes tumor growth. These can spread to numerous bodily regions and impede respiration. Alcohol use and smoking are two main risk factors for lung cancer. Nevertheless, the reality is that lung cancer affects both smokers and nonsmokers equally. Early detection can help reduce the aggressiveness of Cancer and increase the likelihood of survival. When Cancer is detected early, the chance of survival in an advanced stage is lower than survival during cancer therapy. Several methods are used to find lung disease early. Among the most typical in recent years is CT scan image detection. To detect them, tissues are separated into non-cancerous and malignant groups. Artificial intelligence (AI) can be used to achieve the utmost efficiency and precision. This study examines how CT scans can detect lung cancer via transfer learning. In addition to using three transfer learning models, we constructed a customized CNN model and conducted a comparison study between them.

## II. Literature Review

Many research investigations have already suggested systems to increase the precision of lung cancer detection. So many different architectures have been proposed and compared to other architectures.

Tekade et al. [5] proposed two convenient architectures called CN, achieving an accuracy of about 95%. They have used 3Dmultipath architecture, which uses VGG16. They have chosen VGG because it is faster and lightweight. They have also segmented the lung images after getting predictions from this architecture. They have used uNET architecture for segmentation. This is useful for predicting whether the patient will have lung cancer.

Abdul et al. [6] used only one architecture, CNN, and has an accuracy of 96%. They suggested building CNN with the first convolutional layer with ReLU activation and 32 filters. Sixteen filters are included in the second convolutional layer, followed by ReLU activation.

Fang et al. [7] used transfer learning. They have used google net. The proposed network also inherits GoogLeNet's powerful feature detection functionality to detect lung canc including three softmax layers and nine inception modules.

979-8-3503-2294-1/23/$31.00 ©2023 IEEE

In particular, inception modules are integrated to broaden and deepen networks by processing data simultaneously through multiple convolutions. The second method they used was median intensity projections. The three-axis median projected images were combined into three distinct channels of a single composite RGB image. By combining the two architectures, they got 81% relatively poor accuracy.

Jakimovski et al. [8] trained a dual-convolutional Deep Neural Network (CDNN) along with standard CDNN using Computed Tomography (CT) scans. Using CDNN, they get 87% accuracy, whereas using double CDNN, the accuracy increases to about 96%. Their study shows that In contrast to the regular DNN, which only discovered Cancer in the final step, stage 4, the double CDNN detected Cancer at stage 3.

Sajja et al. [9] used three architectures, AlexNet, GoogleNet and ResNet, with 89, 95, and 96% accuracy. Their suggested network is constructed using a sparse network. There are 27 layers in the entire light network. A GoogleNet-based deep neural network with a maximum percentage of dropouts was created to decrease processing time. By utilizing the dropout layer, this network lessens overfitting during the learning phase. One of the three pertained ResNet50 produced the highest accuracy among all architectures. From this literature, this can be concluded that CNN is always better to have good accuracy along with Resent. Combining different approaches always gives different ways of handling data and improving results. So we have decided to combine CNN with ReseNt, MobileNet, and VGG19 for better accuracy.

Nobrega et al. [10] investigated the deep learning models extractor as a feature to address the problem of lung nodule malignancy classification. VGG-16, VGG-19, Mobile-Net, Xception, Inception V3, ResNet50, DenseNet 169, NASNet-Mobile, NASNetLarge, and DenseNet201 were used as feature extractors. Later these features were classified using different classifiers. The CNN-ResNet50 and SVM-RBF combo produced the most outstanding results, with an accuracy of roughly 88%.

## III. METHODOLOGY

### A. Experimental Methodology

Python is a high-level programming language with numerous uses, including developing deep learning algorithms and data analysis. Applications for image and data manipulation, visualization, machine learning, data science, and many other fields are developed using Python libraries. Due to its vast library access, the Python program- ming is especially effective for deep learning-based issues. Anaconda Navigator and Jupyter Notebook were employed for the preprocessing of the dataset. As well as for handling large datasets and training the Model, Google Colab was used. All the codes, data, graphs, and works were saved to the local machine and online GitHub to retrieve it anytime using any GPU. We will explain how we can solve or can do the project or, in other words, how we can go through each step explained in this chapter. One is the traditional deep learning method, and the other is transfer learning. We have used a custom CNN model and three transfer learning models, which are MobileNetV2, ResNet50, and VGG19, have been trained by our dataset. This chapter will discuss details of our used block diagram and system architecture.

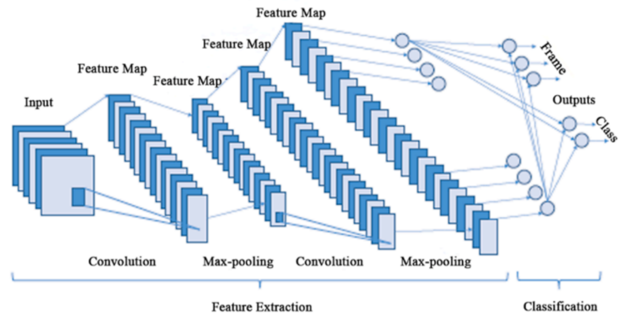


Fig. 1. The CNN's general layout.

### B. Datasets

We have used two datasets. These two datasets are accessible to the general public. The first is the popular LIDC IDRI (The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans (LIDC-IDRI)), [11] and the other is collected from Mendeley. [12] We combined them initially. Even after merging the whole dataset, we saw that the dataset size was too small to train our models. Then we augmented the entire dataset and got a final dataset with a length of around 15 thousand data. Then we did preprocessing and other stuff.

### C. Deep Learning

Using patterns and feature extraction identification, deep learning involves a specific machine learning algorithm, whether supervised or unsupervised. Several non-linear information processing layers are used. [13] More hidden layers of artificial neural networks are used, and at each layer, a feature pertinent to the task at hand is learned. In this design, features found in each layer serve as input data for the layers below. The result is a framework in which information is collected from the most basic functions to the highest layers. [14] In contrast to prior machine learning algorithms, deep learning techniques demand a significant quantity of hardware and data lot of computing capacity to analyze that data. Convolutional neural networks (CNNs), the name for most of these methods, are utilized in applications like image categorization, as shown in Figure 1.

### D. Convolutional Neural Network Models

Among the most diverse domains of deep learning applications, such as image and video processing, image classification, object recognition, and segmentation, CNNs have been extensively exploited by researchers in recent years. In addition to the idea, they are applied to other signals that contain sequential and connected data. [15] The three fundamental layers that comprise CNNs are the layer of Convolution, the ultimately linked layer and the pooling layer. [15] The following layers make up CNN:

*1) The Input Layer:* The top layer of a convolutional neural network is called the input layer. The CNN input layer should contain image data. The data size is crucial for the Model to work effectively in this phase. If the supplied image's specified size is sizable, the strong memory required, the test duration per image, and the training time may increase. Furthermore, their chances of success might increase.

*2) Convolutional Layer:* The feature extraction filters on a photo can be compared to this layer (kernels). [16] Each filter illustrates a specific matrix that convolutes the input image.

*3) The Pooling Layer:* Convolution layers can be sandwiched between pooling layers, one of the most popular methods in deep learning systems. By shrinking how big the feature maps are, this layer hopes to manage the complexity of the image. The Model proposed in our research is the Keras sequential model with max pooling layers.

*4) Activation Layer:* The objective of the activation layer will introduce nonlinearity into a convolutional level-based system that mostly computes linear operations. The system's input is analyzed by how the activation layer works, which then selects the value the system will output.

*5) Whole Layer Connectivity:* The CNN models feature completely interconnected layers connecting every neuron in one layer. Theoretically, an entire layer connected acts as a multi-layer perceptron (MLP). The only distinction is the input format that feeds the convolutional and thoroughly combined layers. There are two classes (Cancer and regular) in this study. The output value of our Model's ultimately linked layer two is provided.

*6) Dropout:* The dropout layer is used to stop the network from memorizing. The Model has a high learning rate and can remember training data. If the network undergoes excessive learning, its learning capacity is gone. During the dropout phase, some network nodes are randomly disabled. To avoid overlearning, two dropout layers are used in this study; the first layer's area is computed to be 108x108x128 and the other dropout layer's location to be 128, respectively.

### *E. Block Diagram*

Figure 2, is the block diagram depicting the overall workflow of our lung cancer detection system. We commence by acquiring two publicly accessible datasets, the LIDC-IDRI dataset and a Mendeley dataset. These datasets include CT scan images of lung tissue samples. For data quality assurance, we preprocess the entire dataset. This preprocessing phase involves applying a cropping technique based on the height of the images to remove noise from the images. We improve the accuracy of our ensuing analysis by removing noise. After preprocessing, the two datasets are merged to produce a new combined dataset. However, it was determined that the magnitude of the final dataset was insufficient for accurate analysis. We employ an augmentation technique to enhance its volume in order to circumvent this limitation. Through augmentation, we generate a larger dataset consisting of 14,486 images, including 7,593 images of malignant samples and 6,904 images of healthy samples. Next, we divide the dataset into training, validation, and testing groups. 70% of the data is allocated to the training set, 20% to the validation set, and the remaining 10% is used for testing. This partitioning enables us to accurately evaluate the efficacy of our models. Having prepared the dataset, we proceed to train our models. We employ custom CNN architectures and transfer learning utilising pre-trained models including MobileNetV2, VGG19, and ResNet50. This combination of custom CNN and transfer learning improves our model's ability to detect lung cancer. Finally, we evaluate the trained models on the test dataset to determine if they can classify CT scan images into two categories: cancerous or healthy. The output of the models provides valuable insights for lung cancer diagnosis and plays a crucial role in enhancing the outcomes of early detection and treatment

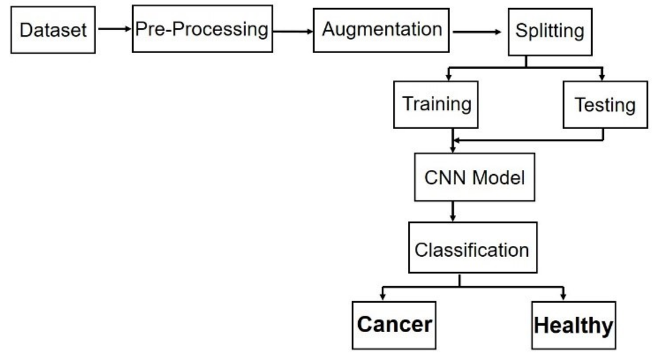


Fig. 2. System Block Diagram

### *F. Proposed Architecture*

At figure 3, the proposed architecture of our system has been drawn. This is our suggested system architecture. The convolutional layers are the foundational layers in deep transfer learning. These layers only filter the original images; weight ranges are multiplied by the input

## IV. RESULT AND DISCUSSION

The results of how we solved the project or, in other words, how we went through each step are explained in this chapter. One is the traditional deep learning method, and the other is transfer learning. We have used a custom CNN model, and three transfer learning models, which are MobileNetV2, ResNet50, and VGG19, have been trained by our dataset. Table 1 shows the summary of all the Model's accuracy and loss.

TABLE I
COMPARISON OF DIFFERENT DEEP LEARNING MODELS

| Model | Accuracy (%) | Validation Accuracy (%) | Loss (%) | Validation Loss (%) |
|---|---|---|---|---|
| Custom CNN | 99 | 88 | 02 | 78 |
| MobileNetV2 (Best one with no OverFitting Issue) | 97 | 94 | 08 | 15 |
| VGG19 | 100 | 100 | 02 | 24 |
| ResNet50 | 78 | 73 | 45 | 52 |

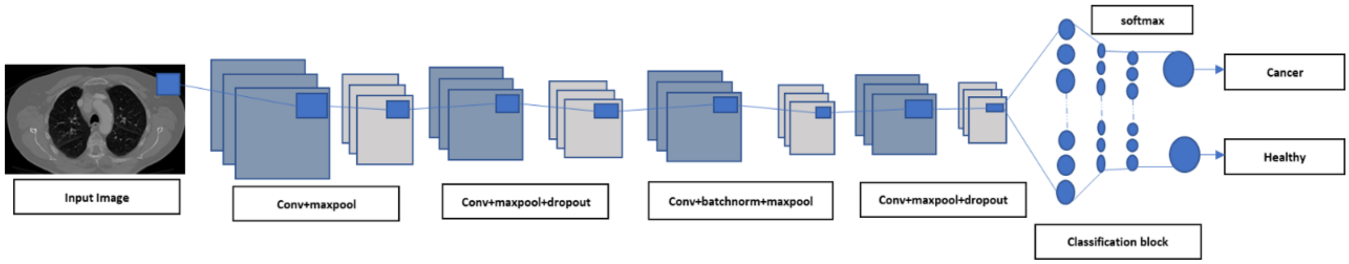


Fig. 3. Proposed System Architecture.

The final epoch accuracy of the trained custom CNN model was 99%, and the validation accuracy was 88%. There were 200 epochs in total. Right away, the Model showed good accuracy. It achieved 98% and 84% validation accuracy in the first epoch. Respectively, 86% and 92% of the pre-trained Model VGG19 had a 100% training accuracy and a 100% validation accuracy after ten epochs. Then we got 97% and 94% training and validation accuracy from MobileNetV2. Furthermore, we got 78% and 73% training and validation accuracy from our final Model, ResNet50. Among these models, we got the highest accuracy from MobileNetV2, and the Model has no over-fitting issues.

### *A. Factors Influencing Variation in Model Accuracy*

Although the same dataset was used to train multiple models, the resulting accuracy of each model was distinct. Numerous factors contribute to this variation in precision. Each model's architecture and design are distinctive, including the number of layers, activation functions, and parameters. These distinctions can affect the models' ability to generalise to new data and distinguish between cancerous and healthy images. In addition, overfitting in specific models suggests they may have learned the training data excessively, resulting in poor performance on unobserved data. This can occur when a model is overly complex or when the dataset is limited, as is the case here. Overfitting can increase accuracy during training but decrease generalisation on validation or test data. In conclusion, the differences in accuracy between models can be attributed to disparities in their architecture and overfitting. By comprehending these factors, additional research can be conducted to optimise model performance and improve the precision of lung cancer detection from CT scan images.

### *B. Model Accuracy and Loss*

After plotting the accuracy, It is evident that the prototype was getting good accuracy from the start. The accuracy in the first epoch was 86%. The accuracy improved in the following epoch and persisted into the final epoch. The model had a 99% absolute accuracy. The model's validation accuracy began at 62%. After ups and downs, it finally got 78%. Figure 4 demonstrates the accuracy and model loss of the Custom CNN model.

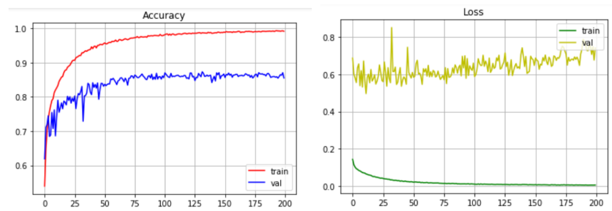


Fig. 4. Model Accuracy and Loss (Custom CNN).

The model loss plot shows that the line of the training loss has decreased, and the line of validation loss has increased gradually. In the first epoch, training loss was 18%; after 200 epochs, it decreased to 2%. The validation loss started at 60% and increased to 78%.

### *C. Pretrained Model Accuracy and Loss*

*1) VGG-19:* Four pre-trained models were employed in this research, along with transfer learning. These trained models produced predictions that were substantially smoother. VGG-Net has a homogeneous design and 16 convolutional layers. More filters were used; however, the filter size was only 33. The architectural consistency of VGGNet was its best quality. VGGNet is hence the best choice for feature extraction from photos. For this reason, pre-trained neural networks are widely used as feature extractors in various applications and problems. However, this network struggles with the 138 million factors, making it tough and challenging to handle. The accuracy and validation accuracy in the VGG19 model were both 100The model's first epoch yielded close to 86% and 91% validation accuracy. In the tenth epoch, it reached 100% accuracy after progressively rising. In the first epoch, it achieved 91% validation accuracy. It dropped to 80% in the third epoch before gradually rising to 100% in the tenth. The model provided a 25% validation loss and a 2% training loss. It offered a 30% training loss in the first epoch and subsequently fell to 2% in the next nine. It offered 65% in the first and 25 in the tenth epoch as compensation for the validation loss. It slowly fell off in the middle epochs. The VGG-19 model's accuracy and loss are depicted in Figure 5.

*2) MobileNetV2:* A neural network design called MobileNetV2 performs exceptionally well when balancing resource limitations and identification accuracy. A neural net-

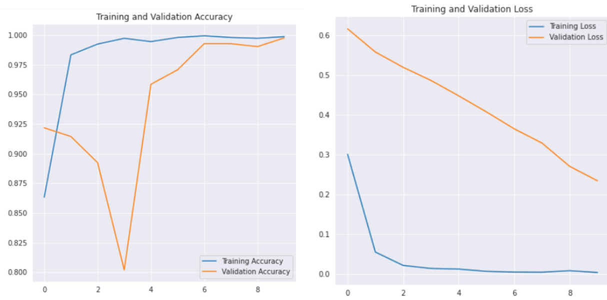


Fig. 5. Model Accuracy and Loss (VGG-19).

work design called MobileNetV2 excels at balancing source limits with recognition accuracy. Being usable on embedded systems and mobile devices is also one of the most significant advantages. There are several difficulties with deep CNN designs, such as network optimization, vanishing gradient issues, and distortion problems. MobileNetV2 had the highest accuracy among these pre-trained models with no overfitting problems. It achieved 96% validation accuracy and 99% accuracy. Its accuracy in the first epoch, which was 72%, was lower than that of the VGG19 model. Additionally, the initial validation accuracy was lower than VGG-19, which starts at 86% and rises to 94% over time. The training loss for MobileNetV2 is 8%, and the validation loss is 15%. The precision and loss of the MobileNetV2 model are depicted in Figure 6.

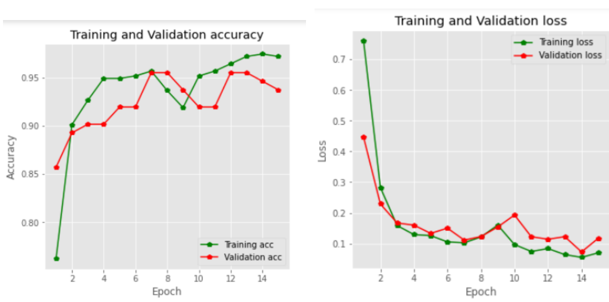


Fig. 6. Model Accuracy and Loss (MobileNetV2).

*3) ResNet50:* All of these problems, including saturation and accuracy loss, are addressed in the training process of the ResNet50 design using residual blocks. These models were selected for our investigation because deep learning applications regularly use them. [17] The least accurate of these pre-trained models was ResNet50. It achieved an accuracy of 78% and a validation accuracy of 73%. Its accuracy is lower in the first epoch, at 61% than that of the MobileNetV2 and VGG19 models, and then gradually rises to 78%. Additionally, the validation accuracy for the initial epoch was lower than that of the VGG-19 and MpbileNetV2, which is 64% and progressively rises to 73%. The validation loss for ResNet50 is 45%, and the training loss is 52%. ResNet50 did not offer a good validation loss in this instance. The model with the most significant validation loss out of the four is this one. Figure 7 shows how the ResNet50 model performed in terms of accuracy and loss.

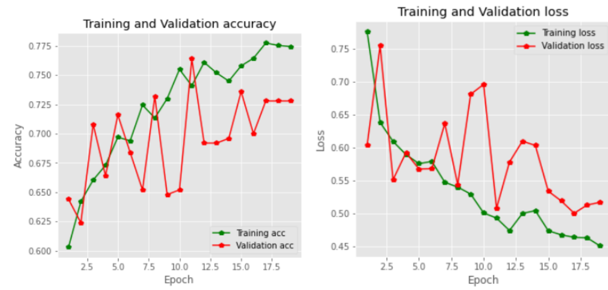


Fig. 7. Model Accuracy and Loss (ResNet50).

## D. Model Test

Here we have two images as our input; our model can recognize healthy and Cancer lungs. Moreover, it will be followed by some procedures. At first, the model extracts the features from image pre-processing and lung segmentation. Then the features can be used as a lung cancer cell identification tool, and we get the final result. The results of Model Testing are shown in Figure 8.


Fig. 8. Model Testing.

## V. Conclusion

Accurate diagnosis of lung cancer is crucial for early diagnosis and effective treatment. In this study, the transfer learning models MobileNetV2, VGG19, and ResNet50 are combined with a customised Convolutional Neural Network (CNN) model to increase the accuracy of lung cancer detection from CT scan images. Using the publicly accessible LIDC-IDRI (The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans (LIDC-IDRI) dataset and a dataset from Mendeley, we aim to surpass the 90% accuracy benchmark established by prior research. We train and evaluate the performance of our models by combining and augmenting the datasets and using diligent preprocessing techniques. Notably, the MobileNetV2 model obtains a remarkable level of precision of approximately 97% without exhibiting any signs of overfitting. Our results demonstrate the effectiveness of transfer learning models in improving the accuracy of lung cancer detection and emphasise the potential for further improvement with larger and more diverse datasets. This study contributes to the advancement of lung cancer detection methods, offering optimistic avenues for early diagnosis and enhanced patient outcomes.

## REFERENCES


[1] "Cancer," World Health Organization. [Online]. Available: https://www.who.int/en/news-room/fact-sheets/detail/cancer. [Accessed: 07-Dec-2022].

[2] H. F. Kareem, M. S. AL-Huseiny, F. Y. Mohsen, E. A. Khalil, and Z. S. Hassan, "Evaluation of SVM performance in the detection of lung cancer in marked CT Scan Dataset," Indonesian Journal of Electrical Engineering and Computer Science. [Online]. Available: https://ijeecs.iaescore.com/index.php/IJEECS/article/view/22853. [Accessed: 07-Dec-2022].

[3] Gomathi, "A computer aided diagnosis system for lung cancer detection using support vector machine," American Journal of Applied Sciences, vol. 7, no. 12, pp. 1532–1538, 2010.

[4] M. S. AL-Huseiny, N. K. Abbas, and A. S. Sajit, "Diagnosis of arrhythmia based on ECG analysis using CNN," Bulletin of Electrical Engineering and Informatics, vol. 9, no. 3, pp. 988– 995, 2020.

[5] R. Tekade and K. Rajeswari," Lung Cancer Detection and Classification Using Deep Learning," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697352.

[6] W. Abdul," An Automatic Lung Cancer Detection and Classification (ALCDC) System Using Convolutional Neural Network," 2020 13th International Conference on Developments in eSystems Engineering (DeSE), 2020, pp. 443- 446, doi: 10.1109/DeSE51703.2020.9450778., 2018, pp. 1-5, doi: 10.1109/ICCUBEA.2018.8697352.

[7] T. Fang," A Novel Computer-Aided Lung Cancer Detection Method Based on Transfer Learning from GoogLeNet and Median Intensity Projections," 2018 IEEE International Conference on Computer and Communication Engineering Technology (CCET), 2018, pp. 286-290, doi: 10.1109/CCET.2018.8542189.

[8] G. Jakimovski and D. Davcev, "Using double convolution neural network for lung cancer stage detection," MDPI, 28-Jan-2019. [On- line]. Available: https://www.mdpi.com/2076- 3417/9/3/427. [Accessed: 24-Nov-2022].

[9] Sajja, T.K., Devarapalli, R.M., Kalluri, H.K. (2019). Lung cancer detection based on CT scan images by using deep transfer learning. Traitement du Signal, Vol. 36, No. 4, pp. 339-344. https://doi.org/10.18280/ts.360406

[10] R. V. Nobrega, S. A. Peixoto, S. P. da Silva, and P. P. Filho, "Lung nodule classification via deep transfer learning in CT lung images," 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), 2018.

[11] "Data from LIDC-IDRI.," The Cancer Imaging Archive (TCIA) Public Access - Cancer Imaging Archive Wiki. [Online]. Available: https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=1966254. [Accessed: 16-Dec-2022].

[12] H. Alyasriy, "The IQ-OTHNCCD lung cancer dataset," Mendeley Data, 04-Jan2021. [Online]. Available: https://data.mendeley.com/datasets/bhmdr45bh2/2. [Accessed: 16-Dec-2022].

[13] L. Deng, "Deep learning: Methods and applications," Foundations and Trends® in Signal Processing, vol. 7, no. 3-4, pp. 197–387, 2014.

[14] W. Rawat and Z. Wang, "Deep convolutional neural networks for Image Classification: A Comprehensive Review," MIT Press, 01-Sep-2017. [Online]. Available: https://direct.mit.edu/neco/article-abstract/29/9/2352/8292/Deep-Convolutional-Neural-Networks-for-Image. [Accessed: 30-Mar-2023].

[15] M. Emin Sahin, "Deep learning-based approach for detecting COVID-19 in chest X-rays," Biomedical Signal Processing and Control, vol. 78, p. 103977, 2022.

[16] Q. Li, E. Al., Author AffiliationsFrom the Chinese Center for Disease Control and Prevention, A. S. Fauci, D. W. and Others, A. C. and Others, S. V. G. and Others, P. B. Gilbert and Others, M. I. and Others, and J. K. Olynyk and G. A. Ramm, "Early Transmission Dynamics in Wuhan, China, of novel coronavirus–infected pneumonia: Nejm," New England Journal of Medicine, 26-Mar-2020. [Online]. Avail- able: https://www.nejm.org/doi/10.1056/NEJMoa2001316. [Accessed: 16- Dec-2022].

[17] A. U. Ibrahim, M. Ozsoz, S. Serte, F. Al-Turjman, and P. S. Yakoi, "Pneumonia classification using deep learning from chest X-ray images during COVID-19," Cognitive Computation, 2021.