

HUMAN DISEASE DETECTION USING ANN

(Covid-19 Detection)

Hamza Iftikhar (21-CS-45) (21-CS-44 @students.uettaxila.edu.pk)

Zunysa Naveed (21-CS-75) (21-CS-75 @students.uettaxila.edu.pk)

University of Engineering and Technology, Taxila
Computer Science department

Keywords: COVID-19 detection, Artificial Neural Networks (ANN), Convolutional Neural Network (CNN), X-ray images, Data preprocessing, Data augmentation, Health care, Cost-effective Health care solution, Deep learning, Model Accuracy, Future Improvements.

Abstract

This study seeks to enhance Covid-19 detection by implementing the concept of Artificial Neural Networks inspired by functioning of human brain. Our project embeds the concept of Convolutional Neural Network for efficient and accurate detection of Covid-19 by taking x-rayed images of lungs of patients as input. The model development involves systematic steps including data acquisition, data preprocessing and data augmentation followed by applying Convolutional Neural Network to the prepared data. The dataset utilized for this project of Covid-19 detection using Artificial Neural Network (CNN) is obtained from the source of website of Kaggle. The dataset consists of three classes namely normal, viral and Covid-19 effected. The visual data of these three classes is utilized for the purpose of training and testing of the model. PCR testing is the most used technique used for the Covid-19 detection, but this technique is pricey for people who belongs to the middle- or lower-class families so, our project overcomes this finance barrier by using X-ray images of the patient to detect whether patient is infected with Covid-19 or not. Accurate identification of COVID-19 cases is vital for controlling its transmission. Minimizing false negatives ensures timely care for infected individuals, reducing spread. Our project achieves accuracy of 95% by using multiple layer Convolutional Neural Network.

Introduction

The project on Covid-19 detection is a superlative intersection of Artificial Intelligence and health care. The project implements the concepts of neural networks. Neural networks are the networks inspired by the Human Brain in terms of functioning. The project utilizes the concept of neural network to analyze the data and then make accurate prediction of Covid-19.

The Covid-19 detection project sights to contribute to the Healthcare by contributing efficient and reasonable model of covid-19 detection with the help of X-rays images which helps further in addressing critical challenges in health care field. Nowadays healthcare is a big challenge faced by every country in various ways and it's very crucial to take appropriate steps regarding the healthcare sector to improve the overall health conditions of any country. Our project aims to contribute to the healthcare sector by providing a project or tool which can help to detect the Covid-19 disease in an easy and efficient manner. Our project aims to provide a tool which is cost effective as well. Nowadays it's very difficult for every individual to afford good healthcare treatment, so our project promises to provide an easy and cost-effective method. It allows the individuals to detect whether they are infected with covid or not and tells them if there is any other virus their lungs are affected with the help of taking x-ray images of their lungs. Our project is GUI based which makes sure that we provide a user-friendly interface for health care individuals to use the tool. Health care accessibility and affordability is a significant concern, and our model aims to focus on this concern and solve this affordability barrier and provide a tool that is not costly and efficient as well.

The project aims on the early detection of Covid-19 with reliable predictions of presence of disease with high accuracy. Neural Networks are based on the concept of Deep Learning, and deep learning models involve continuous learning which results in high accuracies and prediction.

[1] In recent study CNN model called CoroDet for automatic selection of COVID-19 using Chest X-Ray and CT-Scan images have been proposed in recent study. CoroDet is developed to serve develop to deal with 2 class (accuracy 99%), 3 class (accuracy 94%) and 4 class classification (accuracy 91%). [2] Recently a model C-CovidNet is developed that is a lightweight model CNN based model which can compete with heavy transfer learning-based approach. [3] In past COVID-19 Radiography Database, lung segmentation, data augmentation and finally the architectures mAlexNet and Hybrid (that includes BiLSTM Layer) used for classification.[4] Alazab, M. utilized 1000 images dataset for Covid-19 detection and implemented ANN and achieved accuracy of 94.80% in his work. He utilized LSTM to predict the number of Covid-19 detections, recoveries in next 7 upcoming days. He also made predictions by relating the age of a patient and chances of that person being affected with Covid-19 in his studies.[5] Mohammed Qasim, H. and O. Ata. Have done work on real time Covid-19 detection. His study utilized the deep learning optimization system that can work with imbalanced dataset and then he applied various techniques for balancing it and further processing it. He implemented recursive feature elimination algorithm to extract features in his studies and he achieved an accuracy of 94.98 %. [6] Walter, J.R., et al. uses the artificial intelligence for the cough detection and early detection of Covid-19. He utilized different

sensors for the purpose of detection of cough and Covid-19 in his studies. His study highlights the potential value of wearable devices in early disease detection and monitoring. The final performance of the algorithm achieved an F-1 score of 83.3%. [7] Kuvvetli, Y., et al. have made a predictive analytical model for the pandemic of Covid-19 by using ANN. His model future number of daily patients and deaths with covid-19. He achieved an accuracy of 91% in his working by applying model of ANN. [8] Saha, P., et al. They utilized CT scans and CXRs for the detection of Covid-19. They used GIN based model to detect Covid-19 from CT scans in their workings. [9] Shoeibi, A., et al. have worked on Covid-19 detection using real-time reverse transcription–polymerase chain reaction (RT-PCR) tests with a turnaround time of 2–3 days. [10] Zhao, W., W. Jiang, and X. Qiu their study analyzed the pretrained models of Covid-19 detection using CT images and they concluded that training the model on large number of datasets can tremendously enhanced the performance of the model and their study applies this analysis of them. They enhanced the pretrained models by enhancing the datasets utilized and they achieved an accuracy of 93 %. [11] Hariri, W. and A. Narin have utilized the images and acoustic based techniques for the detection of Covid-19. In their work they applied the technique of transfer learning and also utilized auto encoder-based models for the detection of Covid-19, and they achieved an accuracy of 92%.

Artificial Neural Network is an efficient approach for the COVID-19 detection but there is a notable gap in developing the ANN models that are precise, highly accurate and are generalized as well. Existing research lacks the standardized methodologies for model development and validation which will in return create many problems including bias and disparities. Healthcare areas are very sensitive because they can cost anybody's life. So, the models that should be built or designed for health care areas should be highly efficient and precise and works with highest accuracy as much as possible but there is an issue which is that understanding how the ANN works is sometimes difficult which makes it challenging to use it in health-care areas where the cases are sensitive and can cost patients life if a wrong prediction is made by the model, therefore, to overcome this problem we must create large amount of data to train models and get the accurate and precise predictions using ANN.

PCR testing is the most used technique used for the Covid-19 detection, but this technique is pricey for people who belongs to the middle- or lower-class families so, our project overcomes this finance barrier by using X-ray images of the patient to detect whether patient is infected with Covid-19 or not.

This research will employ the Convolutional Neural Network (CNN) which is a type of Artificial Neural Network for Covid-19 detection. Convolutional Neural Network is a type of Artificial Neural Network which is used for the various purposes. CNN consists of single or multiple convolutional and pooling layers to first analyze and then detect the useful features from the visual data and then conclude efficient results from them. Our project aims to enhance the sensitivity and specify for our detection model, ultimately contributing to more reliable and efficient Covid-19 detection. The CNN model also possesses the capability to learn relevant features from the images, making it the right choice for our problem of COVID-19 detection. Further details regarding the

number of convolutional and pooling layers used in our model, as well as the sizes of kernels, are elaborated below in the Model Architecture section of this document.

Accurate identification of COVID-19 cases is vital for controlling its transmission. Minimizing false negatives ensures timely care for infected individuals, reducing spread. However, high false positives can lead to unnecessary quarantines, straining healthcare systems. Balancing false negatives and positives is crucial for effective disease control and resource management.

The dataset under scrutiny consists of three main classes namely 'Normal', 'Virus', 'COVID-19 affected'. The visual data of images in the Train folder of dataset is used for the training of Neural Network while the data of test folder is for testing out model that how well it performs.

Section 2 of our research paper delves into literature review giving important insights in the current state of research in the field of Covid-19 detection. In Section 3, We meticulously highlight the methodology employed in our study, elucidating the various steps undertaken to conduct our investigation effectively. Moving forward to Section 4 and 5, we present the results obtained from our experimentation and classification analyses, offering comprehensive perception into the performance and results of the different models.

Literature Review

The project on Covid-19 detection is a superlative intersection of Artificial Intelligence and health care. The project implements the concepts of neural networks. Neural networks are the networks inspired by the Human Brain in terms of functioning. The project utilizes the concept of neural network to analyze the data and then make accurate prediction of Covid-19. The Covid-19 detection project sights to contribute to the Healthcare by contributing efficient and reasonable model of covid-19 detection with the help of X-rays images which helps further in addressing critical challenges in health care field. Our project utilizes X-ray images of the patient to detect whether the patient is infected with Covid-19 or not. Accurate identification of COVID-19 cases is vital for controlling its transmission. Minimizing false negatives ensures timely care for infected individuals, reducing spread [12] Ibrahim M. Nasser, Samy S. Abu-Naser (2019) have applied the concept of feed forward neural network for the lung cancer detection and achieved the accuracy of almost 94%. Lungs Cancer is a death causing disease and his model is very beneficial in way that it helps in early detection of such a death causing disease which is rewarding for the health care industry in various perspectives.[13] U. Rajendra Acharya et al. (2002): have used the 3-layered feed forward neural network for cardiac monitoring. their model is used for the classification of heart rate based on ECG sampling. they achieved the accuracy in between 85-95%. [7] Kuvvetli, Y., et al. have made a predictive analytical model for the pandemic of Covid-19 by using ANN. His model future number of daily patients and deaths with covid-19. He achieved an accuracy of 91% in his working by applying model of ANN. [8] Saha, P., et al. They utilized CT scans and CXRs for the detection of Covid-19. They used GIN based model to detect Covid-19 from CT scans in their workings. [14] Muhammad Azeem et al. (2023): They have used the Convolutional Neural Network for the purpose of the COVID-19 detection. Their model uses

the multi-layer convolutional Neural Network and has achieved an accuracy of 93%. Further their model can be improved by using more complex datasets and by using additional layers of CNN to enhance the efficiency and accuracy.[15] D. K. Ravish et al. (2014): They worked on the monitoring predicting and preventing heart attacks. They utilized the back propagation Neural Network alongside the ECG heartbeat categorization. Their study is beneficial for healthcare industry to monitor, predict and prevent such a life taking disease with the help of their model which has accuracy of almost 95%.[10] Zhao, W., W. Jiang, and X. Qiu their study analyzed the pretrained models of Covid-19 detection using CT images and they concluded that training the model on large number of datasets can tremendously enhanced the performance of the model and their study applies this analysis of them. They enhanced the pretrained models by enhancing the datasets utilized and they achieved an accuracy of 93 %. [4] Alazab, M. utilized 1000 images dataset for Covid-19 detection and implemented ANN and achieved accuracy of 94.80% in his work. He utilized LSTM to predict the number of Covid-19 detections, recoveries in next 7 upcoming days. He also made predictions by relating the age of a patient and chances of that person being affected with Covid-19 in his studies.

[16] Nasreen Sameer et al. (2018): Has designed a model with Artificial Neural Network to predict diabetes. Their model works in a way that predicts whether a patient is diabetic or not. They have made a binary classification model with an accuracy of 87.3%. His work has had a very good impact on the health care industry as now days its becoming a common disease and needs to be detected early, easily and efficiently. [11] Hariri, W. and A. Narin have utilized the images and acoustic based techniques for the detection of Covid-19. In their work they applied the technique of transfer learning and also utilized auto encoder-based models for the detection of Covid-19 and they achieved an accuracy of 92%.

[17] Dipali M. Joshi et al. (2010): used Artificial Neural Network for brain cancer classification. She utilizes the neuro-fuzzy classifier for the detection of tumors. In addition to overcoming the problems that usually arise from the quality and quantity of dataset her model serves as a foundational step utilizing advanced technologies for the detection of brain cancer. [9] Shoeibi, A., et al. have worked on Covid-19 detection using real-time reverse transcription–polymerase chain reaction (RT-PCR) tests with a turnaround time of 2–3 days. [18] Rahila Parveen et al. (2017): has utilized the Artificial Neural Netwok for the detection of malaria. She achieved accuracy of 85% which showcase that either the dataset utilized is not sufficient for training the data or either the training of the model is not done properly which results in low accuracy. %. [6] Walter, J.R., et al. uses the artificial intelligence for the cough detection and early detection of Covid-19. He utilized different sensors for the purpose of detection of cough and Covid-19 in his studies. His study highlights the potential value of wearable devices in early disease detection and monitoring. The final performance of the algorithm achieved an F-1 score of 83.3%. [19] Behrouz Alizadeh et al. (2015): worked on asthma diagnosis using Artificial Neural Networks. He utilizes the concept of back propagation in his project for asthma diagnosis and achieved an accuracy of 91.5%. His model facilitates the early and accurate diagnosis of asthma, which nowadays as well is a common disease.[20] Victor Gotliv et al. (2013): has done the diagnosis of thalassemia minor. He applied the concept of Feed forward neural network in his model for the diagnosis of thalassemia. Their

study offers new insight into the disease mechanism as well. Their study is helpful for the healthcare industry for the early detection of thalassemia as well.

S#	AUTHOR	PROBLEM	DATASET	TECHNIQUE	ACCURACY	LIMITATIONS
1	Ibrahim M. Nasser, Samy S. Abu-Naser (2019)	Lungs Cancer Detection	Lungs Cancer Detection	Feedforward Neural Network	94.67%	Imbalanced dataset
2	U. Rajendra Acharya...	Classification of Heart Rate	Sampling Frequency of ECG	3-Layered Feedforward Neural Network	85-95%	Small dataset
3	Muhammad Azeem, ... (2023)	Detection of Covid-19	Covid-19	Convolutional Neural Network	93.73%	Complexity Privacy and Security Deficient training data
4	D. K. Ravish, ... (2014)	Monitoring, Prediction and Prevention of Heart Attack	ECG Heartbeat Categorization	Feature selection Back Propagation Neural Network	95%	Cardiologist suggestion more clinical trials treadmill ECG test (TET)
5	Nasreen Sameer, .(2018)	Diabetes prediction using ANN	Association of Diabetic city of Urmia	Just Neural Network (JNN) Environment	87.3%	Small dataset
6	Dipali M. Joshi, .. (2010)	Classification of Brain cancer using ANN	Brain Tumor Detection	Feature Extraction Neuro Fuzzy Classifier	75%	Low quantity dataset Limited detection
7	Angona Biswa, . (2021)	Brain Tumor Types Classification	MRI Brain Tumor Dataset	K-mean clustering Feedforward Neural Network	95.4%	Deficient training data

8	Rahila Parveen, ... (2017)	Prediction of Malaria using ANN	Self-collected Data	Feedforward Neural Network	85%	Low training testing data
9	Behrouz Alizadeh, ... (2015)	Diagnosis of Asthma using ANN	Clinical dataset	Back Propagation Neural Network	91.5%	Imbalanced Dataset
10	Victor Gotliv, ... (2013)	Diagnostic of Thalassemia Minor by ANN	Self-Collected Data	Feedforward Neural Network	93.7%	Deficient training data

Materials and Methods

Covid-19 detection model includes a systematic approach. It consists of a series of various steps which includes data acquisition, data preprocessing, training, testing, evaluation, and output. Dataset obtained from the source of 'Kaggle' website. The dataset obtained consists of three classes folders namely 'Normal', 'Viral' and 'COVID-19 affected'. After the accession of the data, data is preprocessed by using different techniques to make it ready for use. The preprocessed data is then fed into the either Convolutional Neural Network Model or VGG 16 depending on which model is chosen, Both of these models are consist of multiple convolutional and pooling layers to extract important features by analyzing the X-ray images to make efficient identification of Covid-19 and then these layers are followed by the flattening of data and then full connected network is created and results are passed to the output layer and then the model is compiled and evaluation is done to validate models performance.

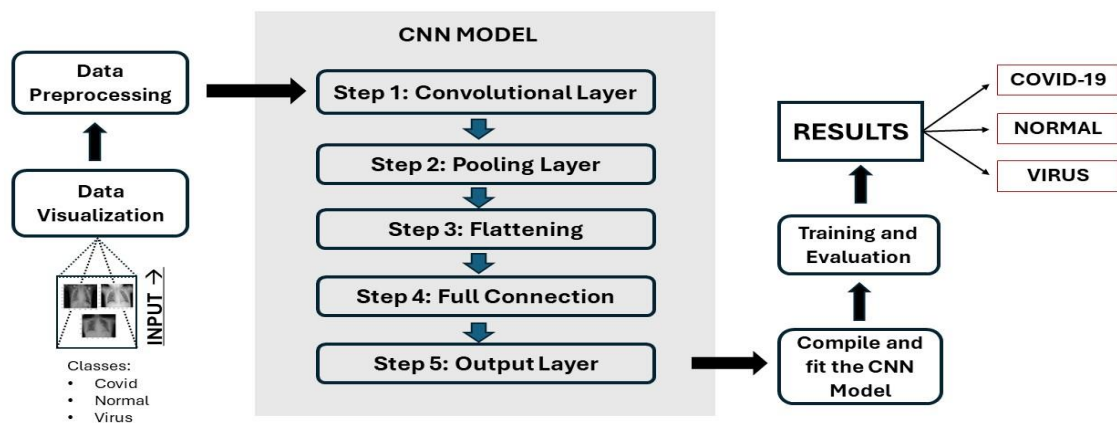


Figure 1: Methodology Diagram

Data Acquisition

The dataset utilized for this project of Covid-19 detection using Artificial Neural Network (CNN) is obtained from the source of website of Kaggle. The dataset is consisting of three classes:

- Normal.
- Viral.
- COVID-19 affected patients.

In total the dataset consists of approximately 1500 images divided into the three classes mentioned above. These three classes are used for the training and testing of our CNN model. Normal class consists of the normal healthy images of chests while viral class consists of images of chests effected with different viruses and class name COVID-19 affected patients contains images of chests affected with disease of COVID-19. All these three classes are combinedly used for the purpose of training and testing of our model.

Data preprocessing

Data preprocessing is a crucial step in machine learning and deep learning tasks. It involves various types of techniques to transform the data into a suitable format for further processing. Our model of covid-19 detection uses the 'Rescaling' technique of data preprocessing. In this technique every pixel is divided by the maximum pixel value to scale down all pixel values to the range $[0,1]$. In addition to this during preprocessing stage the data is split into training and testing for the purpose of first training the model by using train data and then testing the model with the rest of the data in test. The color mode is set to 'rgb', so the images are read in RGB format.

Data Augmentation

Data augmentation incorporates various procedures or techniques to grow dataset by making varieties of existing information to upgrade performance of model. Our model of Covid-19 detection uses various data augmentation techniques including shear range, zoom range, horizontal flip, flow from directory. Shearing distorts the image along a specified axis. Our model uses a range of 0.2 which means images are sheered by maximum angle of 0.2. Horizontal flip is used to flip the images horizontally and add those flipped images to the dataset results in growing the size of dataset. Zoom range is used to zoom in or out the images, our model uses zoom range of 0.2 which means images are zoomed in or out by a factor of 20%.

Data Visualization

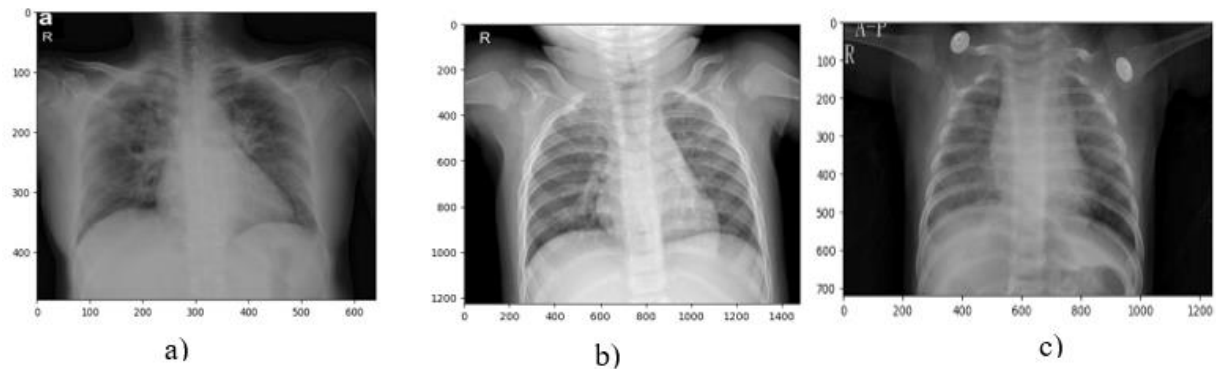


Figure 2 a) Covid, b) Normal, c) Virus

Deep Learning Model

Selecting the appropriate model is a very essential part in problem solving. While choosing between different models to use for solving a particular problem make sure to take into consideration very precisely what the actual problem is and what are the expected outcomes. Furthermore, before finalizing the model, it's also very essential to evaluate performance by examining the F1 score, accuracy, Precision and recall of the model output for the desired problem. The problem we are solving is the COVID-19 detection and we aim to classify images as normal and COVID-19 affected, making it a binary classification task.

We aim at classifying the inputs as positive or negative (positive states that the input image is an image of a healthy lungs while negative states that the input image of chest or lungs is affected with the virus COVID-19).

The model which we selected considering our problem and the outputs we expect are **Convolutional Neural Network (CNN) and VGG16**.

Convolutional Neural Network

Convolutional Neural Network consists of different layers including different convolutional layers and pooling layers. Convolutional layers include different kernels that are used to capture features. Kernel slides over the whole image to detect features. Multiple kernels are used to detect multiple features which then lead to the formation of feature maps.

Proceeding to the convolutional layer there is an activation layer which contains an activation function. This layer is used to learn complex patterns and features. The most common activation function used in CNN is Rectified Linear Unit (ReLU) activation function.

Moving forward there are pooling layers which are used to extract only useful features. It reduces the spatial dimension of feature maps which results in improving efficiency and accuracy as well.

The CNN model contains fully connected layer after convolutional and pooling layers. In a fully connected layer, each neuron is connected to every neuron in the previous layer. The output generated by the fully connected layer is fed to the softmax function which then performs further classification. CNN model contains the flattening layer which is used to convert the 3-D feature maps to 1-D which after conversion can be fed into the fully connected layer.

The CNN model also possesses the capability to learn relevant features from the images, making it the right choice for our problem of COVID-19 detection. Further details regarding the number of convolutional and pooling layers used in our model, as well as the sizes of kernels, are elaborated below in the Model Architecture section of this document.

Model Architecture

The model we utilized in our project of COVID-19 detection is Convolutional Neural Network (CNN) which uses the x-ray images as input and then predict whether the input image of lungs is affected with COVID-19, or the lungs are affected with any other virus, or the input image of lungs are healthy.

Our model processes in such a way that firstly the dataset is loaded, and necessary libraries are imported which will be used further in the model. Then the data is preprocessed to make it ready for use and then it's fed into the CNN.,

Our Convolutional Network consists of 4 layers and each layer contains two layers that are 'Convolutional Layer' to extract features and 'Pooling layer' to spatial down the dimensions of the features maps. Below are the details of these four layers:

- In the initial layer the image is given as input for the processing. The kernel which is used in convolutional layer is of size (3,3) to extract features, and the activation function which is used by the convolutional layer is ReLu activation function. Afterwards the convolutional layer there is a max pooling layer in which pooling is done and pool size used for this purpose is (2,2).
- Moving towards the second layer, it consists of the convolutional layer which has the kernel of size (3,3) to learn features and then convolutional layer is followed by the max pooling layer which has the pool size of (2,2) and the activation function for this layer was also ReLu like the initial layer.
- The third layer also contains the convolutional and pooling layers like the previous two layers and the kernel size and pool size utilized in this layer are also (3,3) and (2,2) with an activation function of ReLu.
- Just like the previous three layers the fourth layer performs the same functionality with the same activation function and kernel and pool size as utilized in previously three addressed layers.

Using multiple layers has many advantages like learning the complex features or generalizing well to the new examples or many more.

After the model is trained by using CNN then the model is visualized and examined with the help of confusion matrix, accuracy measures and classification report and in many other ways.

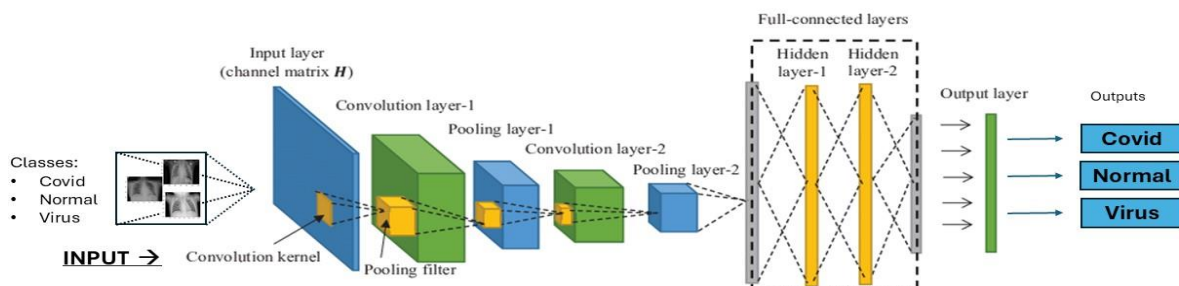


Figure 3 CNN Model Architecture

VGG16

VGG16 stands for Visual Geometric Group. 16 refers to 16 weighted layers, in which there are 13 convolutional layers and 3 fully connected layers. Convolutional layers include different kernels that are used to capture features. Kernel slides over the whole image to detect features. Multiple kernels are used to detect multiple features which then lead to the formation of feature maps.

Proceeding to the convolutional layer there is an activation layer which contains an activation function. This layer is used to learn complex patterns and features. The most common activation function used in VGG16 is Rectified Linear Unit (ReLU) activation function. There are 5 max-pooling layers in the model as well in which pooling is done. VGG16 consists of 3 fully connected layers. Fully connected layers are used to combine all features detected by all previous layers to make the final decision of classification. The output generated by the fully connected layer is fed to the softmax function which then performs further classification.

Model Architecture

VGG16 consists of 13 weighted layers, 3 fully connected layers and 5 max-pooling layers. 13 weighted layers are used to detect features, 3 fully connected layers are used to combine all the features detected by the previous layers and max pooling layers are used to perform pooling to spatial down the dimensions of the features maps.

In VGG16 convolutional layers contain kernels of size 3x3 to extract the features efficiently. The max pooling layers in VGG16 contain filter of size 2x2 with a stride of 2.

Fully connected layers contain 4096 neurons which is equal to the number of classes in classification.

Using multiple layers has many advantages like learning the complex features or generalizing well to the new examples or many more.

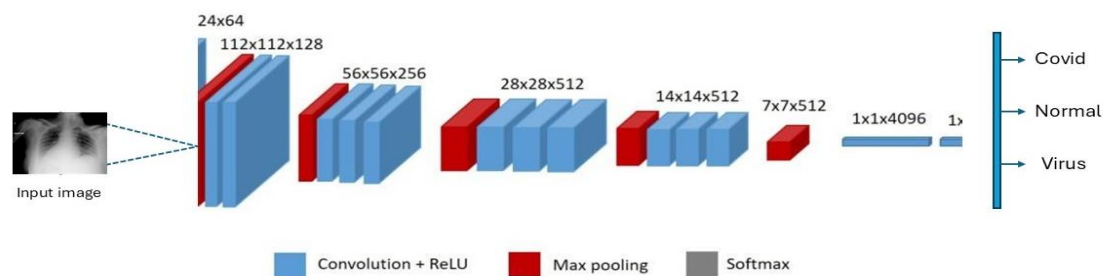


Figure 4 VGG16 Model Architecture

Results

Our project of Covid-19 detection achieved accuracy of 95% in detection of Covid-19 with the help of X-ray images by embedding the concepts of **convolutional neural network** which is a type of Artificial Neural Network and achieved accuracy 98 % by embedding the concepts of **VGG16**. of the model underwent excessive training and testing using the dataset obtained from Kaggle.

During the training phase the CNN model and VGG16 model showed an excessive accuracy of 97.68 % and 99.7% respectively by learning complex features and patterns from the visual data given to it as input and learning complex features and patterns. As model implements the concept of Convolutional Neural Network so it holds the potential of generalizability as well which makes it more efficient.

Furthermore, the testing accuracy of our both models matches close to the training accuracy which states that the model has done learning efficiently and can generalize very efficiently and gives precise and accurate outputs.

Given below is the comparison of our model with the working of others in terms of accuracies.

<i>Authors</i>	<i>Accuracies</i>
<i>Aneesha Manke</i>	92% (Back propagation Neural Network)
<i>Kamana Simkhada</i>	95% (Neuro Fuzzy Classifier)
<i>Vamshi Pallapu</i>	94% (Back propagation Neural Network)
<i>Fatih Mehmet koc.</i>	91% (CNN)
<i>Kuvvetli, Y., et al</i>	91% (Feed Forward Neural Network)
<i>Zhao, W., W. Jiang, and X. Qiu</i>	93% (JNN)
<i>Hariri, W., and A. Narin</i>	92% (Transfer learning and auto encoder modles)
OUR MODEL	95% (CNN) 98.33% (VGG16)

Aneesha Manke, Kamana Simkhada, Vamshi Pallapu and Faith Mehmet Koc. Achieved accuracies of 92%, 95%, 94% and 91% respectively. Most of the used double layer architecture of CNN, while our model of CNN stood out with accuracy of 95% by adding additional convolutional and pooling layers which improved the accuracy and out model of VGG16 stood out with the accuracy of 98% and makes the model more accurate for the COVID-19 detection.

Here is the visual representation of our model's performance as well:

CNN

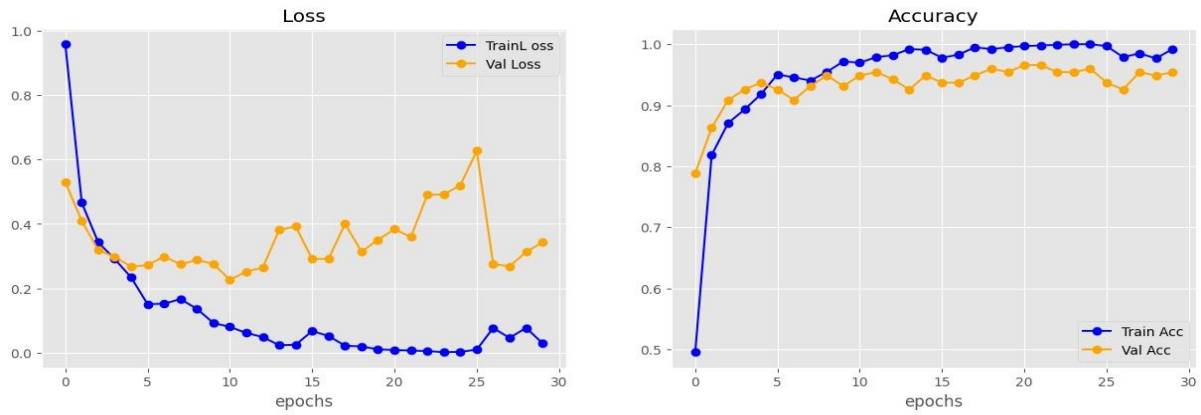


Figure 5 CNN Loss and Accuracy graphs

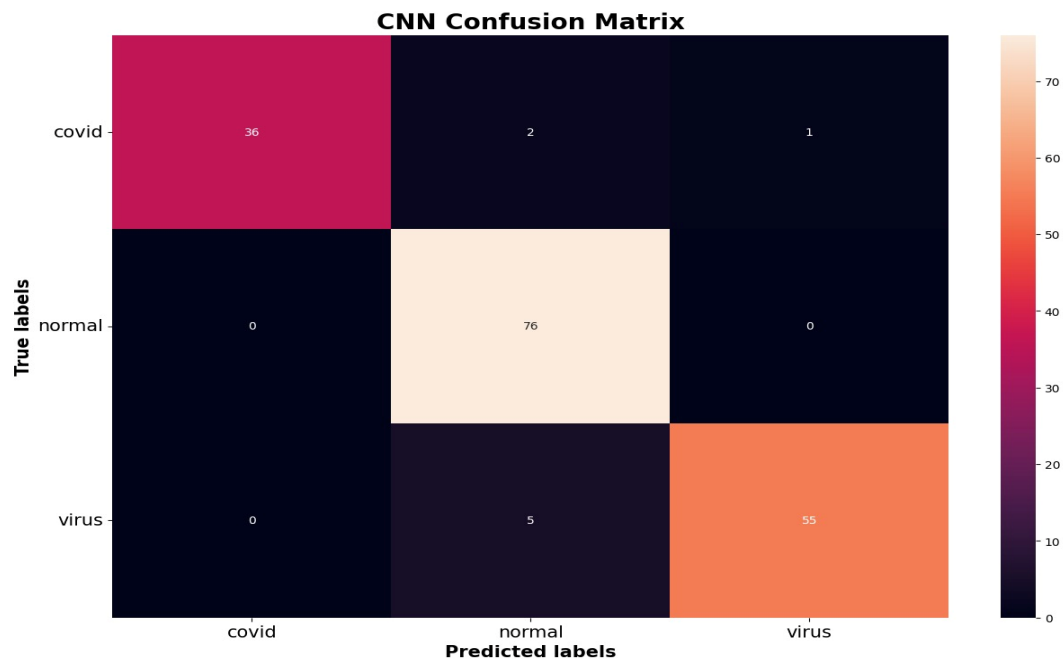


Figure 6 CNN Confusion Matrix

The confusion metrics of model shows that:

- 36 instances of covid class are predicted correct by the model.
- 2 instances of covid class are predicted wrongly as normal.
- Single instance of covid class is predicted wrong as virus.
- The model hasn't predicted any instance of normal class wrong.
- In total 76 instances of normal class are predicted right as normal.
- No instance of normal class is predicted as virus.
- No instance of normal class is predicted as covid-19 as well.
- 5 instances of virus class are wrongly predicted as normal.
- 55 instances of virus class are predicted correctly as virus by our model.

Classification report summarized the overall performance of the model. It contains information regarding precision, recall, F1-Score. Our model gives an overall 95% testing accuracy.

By using the information generated by the confusion metrics different calculations can be performed to get various values including values of F1 Score, Precision, Recall etc.

Here are the calculations of the values:

- **Accuracy= (Total InstancesTPcovid+TPnormal+TPvirus) / Total Instances**

$$\begin{aligned}\text{Accuracy} &= (36+76+55) / (36+2+1+76+0+0+0+5+55) \\ &= 167/175 \\ &=0.954\end{aligned}$$

- **Precision**

- Covid
 $P=36 / (36+0+0) = 1.0$
- Normal
 $76 / (2+76+5) = 0.92$
- Virus
 $55 / (1+0+55) = 0.982$

- **Recall**

- Covid
 $36 / (36+2+1) = 0.92$
- Normal
 $76 / (0+76+0) = 1.0$
- Virus
 $55 / (0+5+55) = 0.92$

- **F1 Score**

F1 score = $2 * (\text{Precision} * \text{Recall}) / (\text{precision} + \text{Recall})$

- Covid
 $F1 = 2 * (1.0 * 0.92) / (1.0 + 0.92) = 0.96$
- Normal
 $F1 = 2 * (0.92 * 1.0) / (0.92 + 1.0) = 0.95$
- Virus
 $F1 = 2 * (0.98 * 0.92) / (0.98 + 0.92) = 0.94$

The above calculations show the overall performance of our model. Accuracy tells how well the model performs and how often it makes right predictions. The accuracy of our model tells how often it predicts right weather the lungs image given as input are affected with Covid-19 or not.

Precision tells us about that in all positive predictions how much the model has predicted positive right.

Recall tells that in all correct items how much model predicted right. It tells how much model is good at doing what it is supposed to do.

F1 score is the harmonic means of precision and recall. It combines precision and recall into a single number.

Calculating all these values helps to have a comprehensive look at the overall performance of a model. These factors won't only help at looking at the overall model's performance but in addition these factors can help to detect where the model is lacking and where it's performing good.

Here is the classification report of our model:

Classification Report is :				
	precision	recall	f1-score	support
covid	1.00	0.92	0.96	39
normal	0.92	1.00	0.96	76
virus	0.98	0.92	0.95	60
accuracy			0.95	175
macro avg	0.97	0.95	0.95	175
weighted avg	0.96	0.95	0.95	175

Figure 7 CNN Classification Report

VGG16

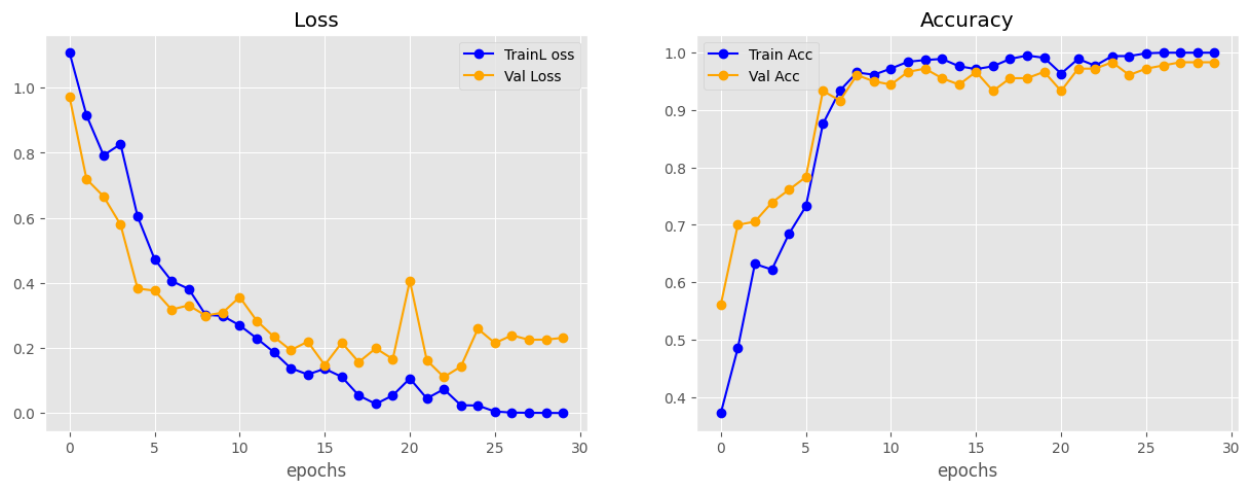


Figure 8 VGG16 Loss and Accuracy graphs

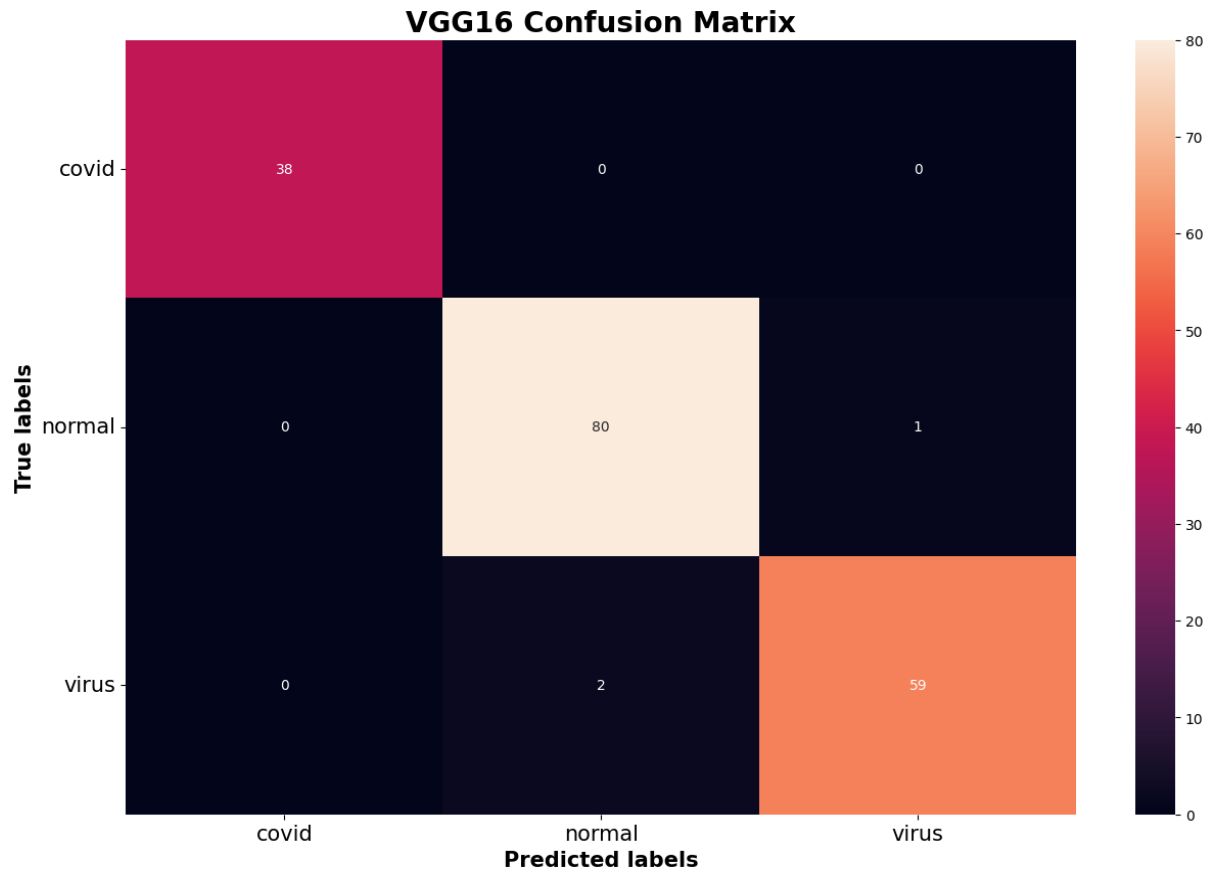


Figure 9 VGG16 Confusion Matrix

The confusion metrics of model shows that:

- 38 instances of covid class are correctly predicted by the model as covid.
- 80 instances of normal class are correctly predicted by the model as normal.
- 59 instances of virus class are correctly predicted by the model as virus.
- Single instance of normal class is predicted wrong as virus by the model.
- 2 instances of viral class are wrongly predicted by the model as normal.

By using the information generated by the confusion metrics different calculations can be performed to get various values including values of F1 Score, Precision, Recall etc.

Here are the calculations of the values:

- **Accuracy** = $(\text{Total Instances} \text{TP}_{\text{covid}} + \text{TP}_{\text{normal}} + \text{TP}_{\text{virus}}) / \text{Total Instances}$

$$\text{Accuracy} = (38 + 80 + 59) / (38 + 0 + 0 + 0 + 80 + 1 + 0 + 2 + 59)$$

$$= 177 / 180$$

$$= 0.9833$$

- **Precision**

- Covid

$$P = 38 / (38 + 0 + 0) = 1.0$$

- Normal

$$80 / (0 + 80 + 2) = 0.97$$

- Virus

$$59 / (0 + 1 + 59) = 0.983$$

- **Recall**

- Covid

$$38 / (38 + 0 + 0) = 1.0$$

- Normal

$$80 / (0 + 80 + 1) = 0.98$$

- Virus

$$59 / (0 + 2 + 59) = 0.96$$

- **F1 Score**

$$\text{F1 score} = 2 * (\text{Precision} * \text{Recall}) / (\text{precision} + \text{Recall})$$

- Covid

$$F1 = 2 * (1.0 * 1.0) / (1.0 + 1.0) = 1.0$$

- Normal

$$F1 = 2 * (0.97 * 0.98) / (0.97 + 0.98) = 0.98$$
- Virus

$$F1 = 2 * (0.98 * 0.96) / (0.98 + 0.96) = 0.98$$

The above calculations show the overall performance of our model. Accuracy tells how well the model performs and how often it makes right predictions. The accuracy of our model tells how often it predicts right weather the lungs image given as input are affected with Covid-19 or not.

Precision tells us about that in all positive predictions how much the model has predicted positive right.

Recall tells that in all correct items how much model predicted right. It tells how much model is good at doing what it is supposed to do.

F1 score is the harmonic means of precision and recall. It combines precision and recall into a single number.

Calculating all these values helps to have a comprehensive look at the overall performance of a model. These factors won't only help at looking at the overall model's performance but in addition these factors can help to detect where the model is lacking and where it's performing good.

Here is the classification report of our model:

Classification Report is :

	precision	recall	f1-score	support
covid	1.00	1.00	1.00	38
normal	0.98	0.99	0.98	81
virus	0.98	0.97	0.98	61
accuracy			0.98	180
macro avg	0.99	0.98	0.99	180
weighted avg	0.98	0.98	0.98	180

Figure 10 VGG16 Classification Report

Conclusion and Future Work

The ANN project on COVID-19 detection demonstrates the application of Artificial Intelligence in healthcare. With the help of advanced methods, the project exhibited markable accuracy by introducing additional convolutional and pooling layers. The project tends to help healthcare to detect COVID-19 affected patients very efficiently in a cost-effective way. Utilizing technology in this manner allows the healthcare to analyze the patients in an easy manner by using their X-ray images to predict the possibility of being affected by COVID-19 or not. Accurate identification of COVID-19 cases is vital for controlling its transmission. Minimizing false negatives ensures timely care for infected individuals, reducing spread. Our project achieves accuracy of 95% by using multiple layer Convolutional Neural Network and 98.33 % by using VGG16.

The factors which limit our model are complexity and data quality. Understanding how ANN works and makes predictions is a difficult task. Incorrect prediction by the model can lead to many consequences like wrong treatment, it can endanger patients' life etc. Secondly data quality is also one of the limitations of our project, the precise detection of covid-19 depends on the quality of input image as well. Poor quality input image may lead to the wrong prediction.

Furthermore, the future improvements which can be made to this model is to incorporate more diverse and extensive for more precise detection. This can be done by obtaining data from different sources. Another improvement which can be incorporated in this project in future is to utilize CT scan images as well along with the Xray images for more efficient detection of Covid-19. Real-time processing can also be imbedded in to provide immediate feedback which will help us to make quick decisions. Furthermore, the project work can be enhanced by incorporating additional data along with X-ray images. Additional data can include patient symptoms, blood tests and various other elements which can help for the detection of Covid-19. By incorporating the additional data, it will help to make more accurate predictions of Covid-19.

References:

1. Hussain, E., et al., *CoroDet: A deep learning based classification for COVID-19 detection using chest X-ray images*. Chaos, Solitons & Fractals, 2021. **142**: p. 110495.
2. Rajawat, N., et al., *C-COVIDNet: A CNN model for COVID-19 detection using image processing*. Arabian Journal for Science and Engineering, 2022. **47**(8): p. 10811-10822.
3. Aslan, M.F., et al., *CNN-based transfer learning-BiLSTM network: A novel approach for COVID-19 infection detection*. Applied Soft Computing, 2021. **98**: p. 106912.
4. Alazab, M., et al., *COVID-19 prediction and detection using deep learning*. International Journal of Computer Information Systems and Industrial Management Applications, 2020. **12**: p. 14-14.

5. Mohammedqasim, H. and O. Ata, *Real-time data of COVID-19 detection with IoT sensor tracking using artificial neural network*. Computers and Electrical Engineering, 2022. **100**: p. 107971.
6. Walter, J.R., et al., *Use of artificial intelligence to develop predictive algorithms of cough and PCR-confirmed COVID-19 infections based on inputs from clinical-grade wearable sensors*. Scientific reports, 2024. **14**(1): p. 8072.
7. Kuvvetli, Y., et al., *A predictive analytics model for COVID-19 pandemic using artificial neural networks*. Decision Analytics Journal, 2021. **1**: p. 100007.
8. Saha, P., et al., *Retracted article: Graphcovidnet: A graph neural network based model for detecting COVID-19 from ct scans and x-rays of chest*. Scientific reports, 2021. **11**(1): p. 8304.
9. Shoeibi, A., et al., *Automated detection and forecasting of covid-19 using deep learning techniques: A review*. Neurocomputing, 2024: p. 127317.
10. Zhao, W., W. Jiang, and X. Qiu, *Deep learning for COVID-19 detection based on CT images*. Scientific Reports, 2021. **11**(1): p. 14353.
11. Hariri, W. and A. Narin, *Deep neural networks for COVID-19 detection and diagnosis using images and acoustic-based techniques: a recent review*. Soft computing, 2021. **25**(24): p. 15345-15362.
12. Nasser, I.M. and S.S. Abu-Naser, *Lung cancer detection using artificial neural network*. International Journal of Engineering and Information Systems (IJEAIS), 2019. **3**(3): p. 17-23.
13. Acharya, U.R., et al., *Classification of heart rate data using artificial neural network and fuzzy equivalence relation*. Pattern recognition, 2003. **36**(1): p. 61-68.
14. Azeem, M., et al., *Neural Networks for the Detection of COVID-19 and Other Diseases: Prospects and Challenges*. Bioengineering, 2023. **10**(7): p. 850.
15. Ravish, D., et al. *Heart function monitoring, prediction and prevention of heart attacks: Using artificial neural networks*. in *2014 International conference on contemporary computing and informatics (IC3I)*. 2014. IEEE.
16. El_Jerjawi, N.S. and S.S. Abu-Naser, *Diabetes prediction using artificial neural network*. 2018.
17. Joshi, D.M., N. Rana, and V. Misra. *Classification of brain cancer using artificial neural network*. in *2010 2nd international conference on electronic computer technology*. 2010. IEEE.
18. Parveen, R., et al., *Prediction of malaria using artificial neural network*. Int J Comput Sci Netw Secur, 2017. **17**(12): p. 79-86.
19. Alizadeh, B., et al., *Developing an intelligent system for diagnosis of asthma based on artificial neural network*. Acta Informatica Medica, 2015. **23**(4): p. 220.
20. Barnhart-Magen, G., et al., *Differential diagnostics of thalassemia minor by artificial neural networks model*. Journal of clinical laboratory analysis, 2013. **27**(6): p. 481-486.