Covid-19 CT Scan Image Classification Using Deep Learning Models

Abstract—**Covid-19 has become a major global health crisis, with many people infected and dying due to late detection and ineffective treatment processes. The commonly used method for diagnosing Covid-19 is the RT-PCR test, but due to its limitations, particularly in accessibility and speed, Chest CT scans have proven to be more effective for early detection, diagnostic confirmation, and assessing complications. Despite its benefits, there are cases where Chest CT scans may show negative results even if a person is positive. In recent years, researchers have explored the use of Deep Learning models to detect and classify Covid-19 through CT scan images. This paper presents the development of a Covid-19 detection model based on deep learning techniques using Chest CT scan images.**

Index Terms—**Covid-19, CT Scan, Deep Learning, Image Classification, VGG16, ResNet50, Transfer Learning.**

I. Introduction

The global phenomenon Covid 19 has occurred in 2019 and this disease spread rapidly to the entire world and a lot of people are affected and a lot of people are died and faced the side effects . According to the reports of 2022 there were around 476 million covid cases and around 6.1 million people have died . With the lack of facilities firstly for detecting these diseases the common approach across was RT PCR test but it has a certain time limit and it has limitations to and when it comes to the accessibility and most important after the test the results will take a lot of time and lack of availability of labs too .

For this kind of issues, the Chest CT imaging methods for example the X ray and the ultrasound imaging has entered the field at the time when these methods had proved vital for detecting COVID 19.These kinds of deep advance methods not only predict the diseases but it has an ability to predict how severe that disease was.Then the medical imaging with the IoT framework came into the field for COVID-19 detection and prevention, and it also used to check the spread of the virus.t the disease how severity that disease was.After that medical imaging with the IoT framework was came to the field for COVID 19 detection and prevention and it also used to check the spread of virus

However, their current evaluation is normal practice and fairly qualitative; hence, the call for computer-based decision support systems.Consequently, with technology advancing, medical imaging possesses a diversified and diverse capacity in the diagnosis of diseases through the employment of knowledge ML and deep technology. become popular for detecting the COVID 19 . This kid of deep advance methods not only predict the disease but it has an capability to predict the disease how severity that disease was.After that medical imaging with the IoT framework was came to the field for COVID 19 detection and prevention and it also used to check the spread of virus

Nevertheless, their current assessment is considered standard practice and rather qualitative; thus, there is the demand for computer-based decision support systems. Medical imaging has a rich and variant potential for the knowledge-based diagnosis of diseases using machine learning (ML) and deep learning technologies. Further, deep learning techniques have been discussed to improve the diagnostic system effectiveness and prove more effective particularly in combination with medical imaging techniques.

In this case, images of the chest computed tomography for COVID-19 patients have been subjected to deep learning.CT scarcity studies have shown that this technique is much sensitive than the RT-PCR assays Deep Neural networks such as Convolution Neural Networks are using often in the image classification as well as the image processing, image detection methods are also using in the Covid 19 detection. disease how severity that disease was.After that medical imaging with the IoT framework was came to the field for COVID 19 detection and prevention and it also used to check the spread of virus .

Although COVID-19 classification has been used with many established deep learning architectures, such as VGG16, ResNet-50, and InceptionV3, the application of deformable CNNs is still little explored. Two deformable CNN models based on the conventional CNN architecture and ResNet-50 are proposed in this study for the detection of COVID-19 from chest CT images. These models are deformable models intended to improve performance by creating what is otherwise unfeasible for the network to learn or adapt to: complex and varied patterns in the input data.

II. Dataset

For this project, we have obtained a dataset from Kaggle, containing both Covid-19 and non-Covid-19 Chest CT scan images. The dataset consists of 80 Covid images and 203 non-Covid images, with each category labeled in separate folders for classification purposes.

III. Data Visualization

The dataset was visualized using various techniques. In Figure 1, we display the first 15 Covid images arranged in a 3x5 grid. A similar visualization was performed for non-Covid images. To understand the pixel intensity distribution of the images, histograms were plotted for randomly selected images. This was done using the `histogram()` function, which converts the images into 1D arrays and plots the intensity distribution for both Covid and non-Covid images.

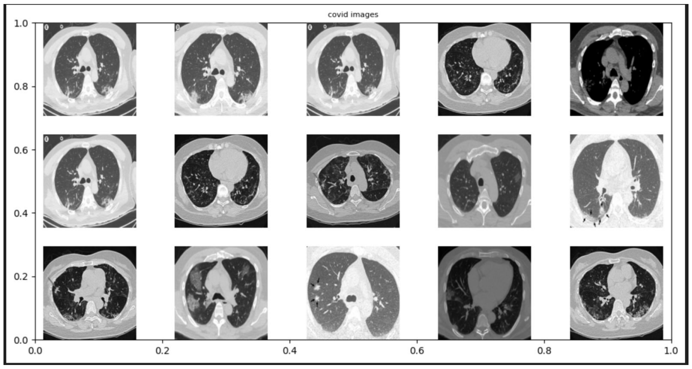

Fig. 1: covid images

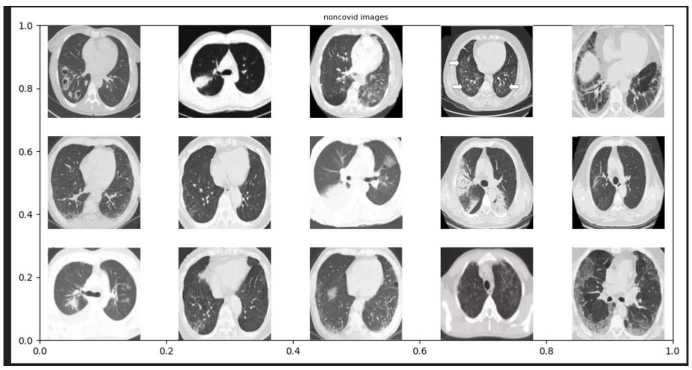

Fig. 2: noncovid images

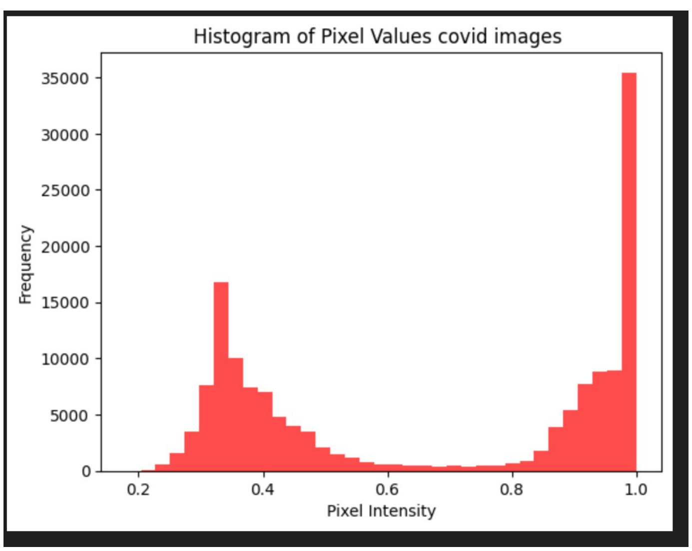

Fig. 3: Histogram of pixes covid images

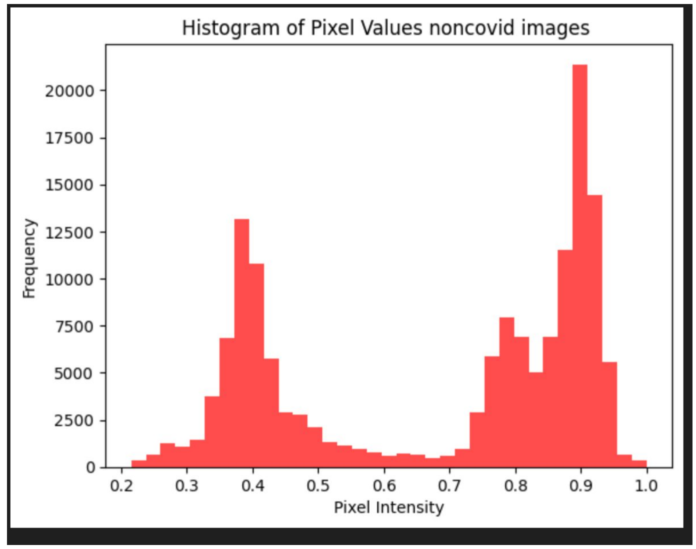

Fig. 4: Histogram of pixes noncovid images

IV. Existing Methods

Machine learning (ML) and deep learning (DL) models have been extensively researched to diagnose COVID-19 from CT scans or X-rays. A nine-layer CNN model was created in the experiment to have 96.28% accuracy, while a 23-layer CNN had 95.99% accuracy. Additionally, transfer learning was applied on well-known architectures such as ResNet-50 and achieved strong results such as 82.91%. Other studies relied on multitask learning and 3D U-Net and MVP Net for segmentation and detection, achieving accuracies in the range of 94%.

Precision has allowed for the popularity of ensemble learning methods in COVID-19 diagnostics. Studies using ensemble models such as VGG, Xception, and ResNet achieved F1 scores ranging from 0.94 to 0.95. Compared to single-model approaches, these ensemble methods exceed single-model accuracy, with some implementations achieving up to 85.4%. The robustness of ensemble approaches was demonstrated by an ensemble method using VGG16, ResNet-50, and Xception.

Indeed, results are encouraging; however, the problems of unbalanced datasets and high false prediction rates remain challenges. The presented limitations highlight the opportunity to improve prediction accuracy and reduce false positive and negative rates. Optimizing models for robustness has been an area of active research. While ensemble models and transfer learning have advanced significantly, optimization remains an ongoing challenge.

V. Algorithms

A. *VGG16*

The VGG16 architecture is a popular deep learning model used for image classification tasks. It consists of 16 layers, including 13 convolution layers and pooling layers, with small kernel filters (3x3) to maintain the spatial resolution. The architecture uses ReLU activation functions and includes fully connected layers for classification. VGG16 is known for its deep architecture, which helps learn complex features while avoiding excessive parameters.

Architecture of VGG16:

- Convolution layers and with the use of small kernel filters which has the size of (3,3) and using the stride 2 and padding 2 to maintain and not lose the spatial resolution and then we are going to increase the depth of the network by increasing the filters this will help model to learn the complex features from the image and it will also abstract the features from the edges.after every convolution we are using the batch normalization to improve the stability and speed of the neural networks .
- In the VGG16 architecture after building the series of convolution we have the pooling layers to reduce the spatial resolution by hals and main aim of the using pooling layers was to reduce the power of the computation and not loose and need to maintain the key features from the images and the major problem we are avoiding this was the model will not go under Overfitting and this also allows the models to mainly focus on the spatial hierarchies.
- In the VGG16 architecture we have three fully connected layers and mainly the first two connected layers with the 4096 neurons and the final output layers has a neurons which was equal to the number of classes and sometimes we can use the softmax activation for the image classification tasks .
- In the model for very convolution layers the model has the ReLu activation function to enhance the non linearning and it will enable the model to learn the complex mappings
- VGG16 takes in input images and performs series of convolutional operations and when merged, gives us hierarchical features moving from edges to patterns. Since pooling layers limit the spatial dimensions, a feature rich compressed representation of the input is obtained. This information is then used downstream by fully connected layers in order to classify the input by learned features. VGG16 has a systematic structure with uniform 3x3 filters, and this computational efficient and effective deep feature extraction content makes VGG16 fit for many applications.

B. ResNet50

ResNet50 is a deep learning architecture with 50 convolution layers, designed to address the vanishing gradient problem through residual connections. These residual connections allow gradients to flow more effectively through deeper networks by skipping intermediate layers.

Architecture of ResNet50:

- Building Blocks : Building Blocks: Residual blocks known as "bottleneck" residual blocks comprise the majority of ResNet-50. Every block has three layers: A 1x1 convolution reduces the dimensionality, a 3x3 convolution processes it, and a final 1x1 convolution restores it to dimensionality. These blocks are replicated up to 50 levels deep across the network.
- When we are working on the Neural Network model while doing we faces the Vanishing Gradient problem where in the deeper network the gradient becomes very small to update the weights efficiently. To tackle this problem in this residual networks we have the method called skip connections . With use of skip connections it will pass the input directly to the deeper networks and if it is necessary it will skip the intermediate layers to avoid the vanishing gradient problem .In this way residual connects allows the gradient to flow the network very easily .
- This model will contain the 50 layers of network or stages each will be doubling the number of filters at every stage like 64,128,256,512 and each stage will begin with the convolution layer and it will be followed by the multiple residual blocks . This residual network ends with an fully connected layer and an average pooling and it has the softmax activation for the classification
- While coming to the activation function each convolution layer has the ReLU activation function that will introduce the non linearity. The batch normalization will reduce the spatial dimensions in max pooling and the training
- Residual connections help train this extremely deep network from degrading, and from flowing gradients. Instead of ruining the output classification accuracy and robustness by learning the last 50 layers, we show that even as these learned characteristics have to be aggregated to the final layers, we can get good output classification accuracy and robustness even with 50 layers. A big leap forward, skip connections in ResNet-50 give ResNet-50 a significant boost in effectiveness and potency for picture assignments

The ResNet50 model benefits from pre-trained weights, which are fine-tuned for this task to improve performance.

VI. Methodology

The methodology for this project involves the following steps:

- The methodology in this project focuses on creating a deep learning model to classify CT scan images into COVID or Non COVID images. It is a dataset with 224x224 image dimension and standardize images in order to have more effect on the model. Firstly, a custom VGG16 like CNN architecture was developed in a scratch way, i.e., each block consisted of convolution, batch normalization, and max pooling, demonstrating VGG16's layered approach. The depth of these blocks is progressively increased and the features thus extracted are very rich, which are necessary for image classification.
- To build the VGG16 network, we built it block by block – each block consists of a number of convolutional layers and max pooling to extract deep hierarchical feature – which is the strong suit of VGG16 when it comes to

the classification tasks. Using these techniques and architectural choices, the model differentiates well between COVID and Non-COVID cases, proving the applicability of both custom built and pre trained architectures in medical image classification.

- An additional step was then taken to put together a ResNet-50 model to boost performance further. ResNet-50, a feature of which is the use of residual connections which well tackle vanishing gradient problem in deep learning, was known to be effective to overcome vanishing gradient problem. In particular, here only the final layers of ResNet-50 were changed in order to take advantage of the transfer learning that the pre-trained weights on ImageNet already have.
- Data augmentation techniques (random rotation, shifts, and flips) were applied to augment data and generalize model to get a diverse training set that reduces overfitting. We performed model evaluation by confusion matrix, accuracy metrics and ROC curves to visualize the performance and predict the suitable hyper parameters like dropout and learning rate.

VII. VGG16 Results

After evaluating the VGG16 model on a set of CT scans, we analyzed predictions for each scan, indicating whether each image was classified as COVID-19 or Non-COVID. Each sample is labeled with its predicted class, class confidence level (percentage), and actual class. In the visual representation, green indicates correct predictions, while red indicates incorrect predictions. This type of visualization helps assess model performance, providing a direct comparison between predictions and actual diagnoses. Such insight is particularly valuable in medical image classification, where understanding the strengths and weaknesses of the model is crucial.

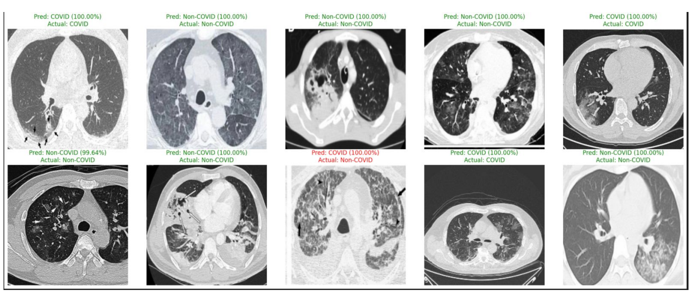

Fig. 5: vgg16 predicted outputs

A. Confusion Matrix

The figure below displays the confusion matrix for the VGG16 model after evaluation, showing the performance of COVID classification on the test dataset. Here's a breakdown of the matrix:

- **True Negatives (TN)**: The top-left cell shows 38 TN, where Non-COVID images were correctly classified as Non-COVID.
- **False Positives (FP)**: There is 1 FP, indicating that a Non-COVID image was incorrectly classified as COVID.
- **False Negatives (FN)**: We have 0 FN, meaning no COVID images were missed.
- **True Positives (TP)**: There are 18 TP, where COVID images were correctly classified as COVID.

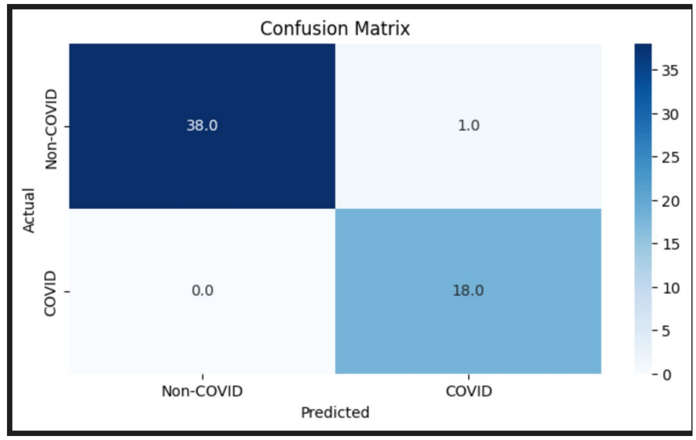

Fig. 6: Confusion$_{m}atrix$

B. Receiver Operating Characteristic (ROC) Curve

The Receiver Operating Characteristic (ROC) curve shows the performance of the binary classifier. The curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold settings. For this model, we achieved perfect classifier performance, with the area under the ROC curve (AUC) being 1.00. The ROC curve is shown as an orange line. As a baseline comparison, a random classifier is represented by a dashed blue line, where TPR equals FPR.

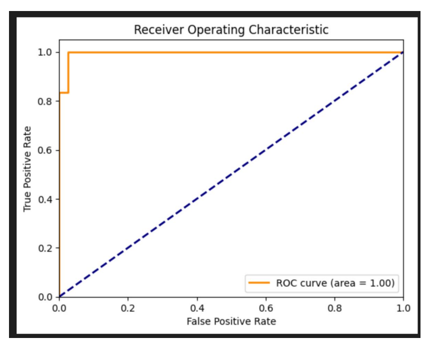

Fig. 7: roc curve

C. Model Accuracy

- **Training Accuracy**: The model's training accuracy improves steadily, reaching approximately 0.95 by epoch 20.
- **Validation Accuracy**: Starting at 0.7, validation accuracy dips significantly between epochs 10 and 30 but recovers to around 0.9 by epoch 50. However, it appears unstable.

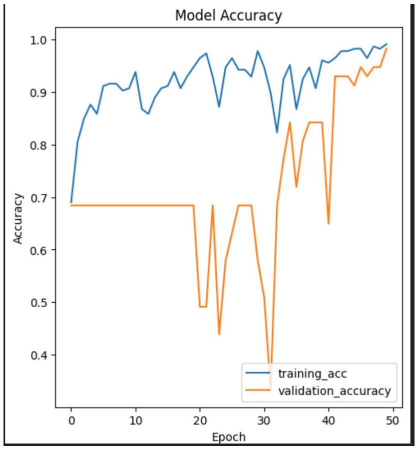

Fig. 8: Accuarcys graph

D. Model Loss

- **Training Loss**: Training loss remains low and stable, indicating a good fit to the training data.
- **Validation Loss**: The validation loss shows high spikes, with values exceeding 50, indicating significant fluctuations and suggesting poor generalization and potential overfitting.

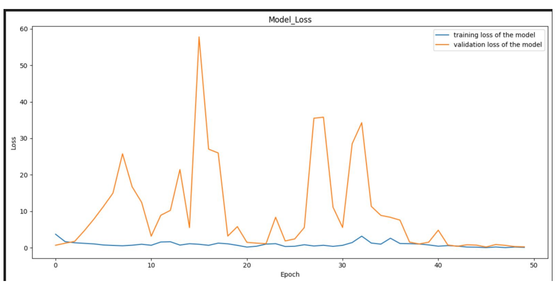

Fig. 9: loss graph

	precision	recall	f1-score	support
Non_COVID	1.00	0.97	0.99	39
COVID	0.95	1.00	0.97	18
accuracy			0.98	57
macro avg	0.97	0.99	0.98	57
weighted avg	0.98	0.98	0.98	57

Fig. 10: classification report of vgg16

VIII. ResNet-50 Results

The first image shows 10 predictions of lung CT scans by the model, displaying both predicted labels (`COVID`, `Non-COVID`) and ground truth labels. Each prediction is accompanied by a confidence percentage, with some predictions made with high confidence (up to 100%) and others with lower confidence (e.g., 86.75%). The model correctly identifies several cases, but some misclassifications occur, such as when a `Non-COVID` label is predicted as `COVID`. The varying confidence levels indicate that the model is more or less certain about some predictions, depending on the individual case.

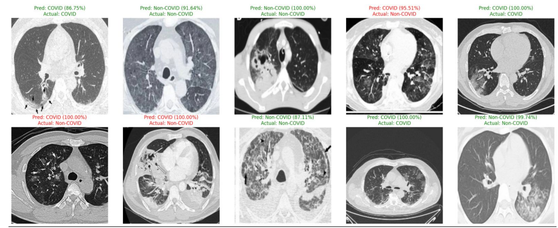

Fig. 11: predicted output of resent

A. Confusion Matrix

The confusion matrix provides insight into the model's classification performance for `COVID` and `Non-COVID` cases:

- **True Positives (TP)**: The model correctly predicted 17 cases as `COVID`.
- **True Negatives (TN)**: The model correctly predicted 27 cases as `Non-COVID`.
- **False Positives (FP)**: 12 `Non-COVID` cases were incorrectly classified as `COVID`.
- **False Negatives (FN)**: 1 `COVID` case was incorrectly classified as `Non-COVID`.

This confusion matrix highlights the model's strengths and weaknesses, with an emphasis on minimizing false positives and false negatives for reliable performance in medical image classification.

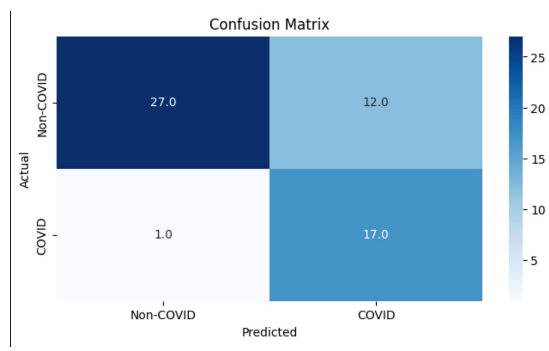

Fig. 12: confusion matrix of resent50

B. Receiver Operating Characteristic (ROC) Curve

The ROC curve measures the performance of the binary classifier, plotting the True Positive Rate (Sensitivity) on the Y-axis and the False Positive Rate on the X-axis. For both `Class 0` (`Non-COVID`) and `Class 1` (`COVID`), an Area Under the Curve (AUC) of 0.92 is achieved, indicating strong classification capability for both classes. The dashed diagonal line (AUC = 0.5) represents random guessing, and the fact that both curves lie significantly above this line shows that the model outperforms random guessing by a large margin.

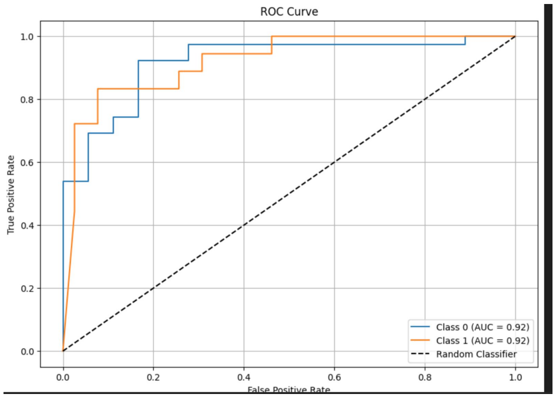

Fig. 13: Resnet50 roc curve

C. Model Accuracy

The first image illustrates the model's accuracy over 100 epochs, with the training accuracy represented by the blue line and validation accuracy by the orange line. Model performance is observed to be unstable across epochs, with considerable fluctuations in both training and validation accuracy. Although the overall trend is similar for both, substantial increases and decreases are seen at certain points, indicating inconsistent learning behavior.

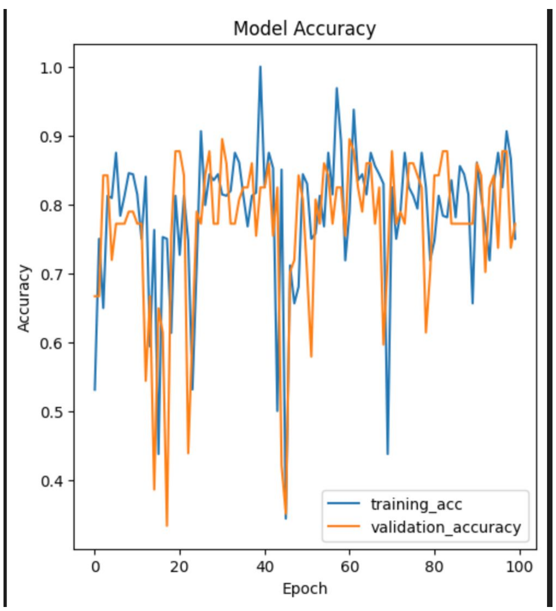

Fig. 14: model$_a ccuracyofresent50$

D. Model Loss

The second image displays the model's loss over the same 100 epochs, with blue representing training loss and orange representing validation loss. Both losses show significant sharp increases and decreases, which could be attributed to over-fitting, as the model fits the training data well but performs poorly on validation data. The sharp fluctuations and spikes in both training and validation losses indicate inconsistencies in the model's learning process and suggest a potential for overfitting.

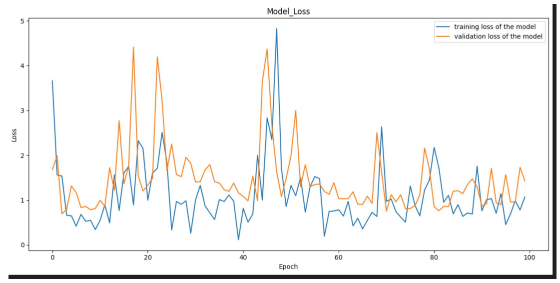

Fig. 15: model loss of resnet50

IX. Comparison between the Models

The accuracy of the VGG16 and ResNet-50 models was compared, revealing significant differences in performance. The VGG16 model achieved a high accuracy of 98.25%, demonstrating that this simpler architecture with deep convolutional layers is particularly effective for this specific task. The VGG16 model's deep convolutional layers are well-suited to capture a variety of features in large datasets, making it a strong performer for image classification tasks like COVID-19 detection in CT scans.

In contrast, the ResNet-50 model, which employs residual learning and skip connections to facilitate a deeper network, achieved an accuracy of 77.19%. While ResNet-50's architecture is generally beneficial in addressing issues such as vanishing gradients and in supporting very deep networks, it may require additional fine-tuning to reach optimal performance for this dataset. The model's lower accuracy here might be attributed to its more complex architecture, which, although effective in preventing gradient disappearance and making computation efficient, may not be perfectly optimized for this specific dataset.

Overall, while ResNet-50 is known for its usefulness in very deep networks and when computational efficiency is a priority, VGG16's performance in this study suggests that, for certain tasks, a simpler yet deep convolutional approach may yield better results.

Model	Architecture	Accuracy
VGG16	Deep convolutional layers	98.25%
ResNet-50	Residual learning with skip connections	77.19%

TABLE I: Comparison of VGG16 and ResNet-50 Model Performance

X. Conclusion

Overall, VGG16 performs better than ResNet-50 in terms of accuracy for this specific task, achieving a higher accuracy and showing strong feature extraction capabilities due to its deep convolutional layers. VGG16's simpler architecture proves effective for COVID-19 detection in CT scans, with minimal tuning required to capture relevant features in the data. Although VGG16 is more prone to overfitting, this can be mitigated through techniques such as data augmentation and regularization.

ResNet-50, on the other hand, consumes less computational resources and is better suited for handling very deep networks due to its residual connections, which help prevent vanishing gradient issues. However, in this application, ResNet-50 was slightly less effective, possibly due to the need for additional fine-tuning to optimize its performance on this dataset.

Each model has its strengths depending on the specific requirements, whether that's maximizing accuracy, improving computational efficiency, or accommodating deeper networks. With further fine-tuning, ResNet-50's performance could potentially be improved. However, for this task, VGG16 stands out as the better choice.

References

[1] Mukherjee H., Ghosh S., Dhar A., Obaidullah S. M., Santosh K. C., Roy K. Deep neural network to detect COVID-19: one architecture for both CT Scans and Chest X-rays. *Applied Intelligence*, 2020; 51(5):2777–2789. doi: https://doi.org/10.1007/s10489-020-01943-610.1007/s10489-020-01943-6.

[2] Xie C., Jiang L., Huang G., et al. Comparison of different samples for 2019 novel coronavirus detection by nucleic acid amplification tests. *International Journal of Infectious Diseases*, 2020; 93:264–267. doi: https://doi.org/10.1016/j.ijid.2020.02.05010.1016/j.ijid.2020.02.050.

[3] Loey M., Manogaran G., Khalifa N. E. M. A deep transfer learning model with classical data augmentation and CGAN to detect COVID-19 from chest CT radiography digital images. *Neural Computing & Applications*, 2020:1–13. doi: https://doi.org/10.1007/s00521-020-05437-x10.1007/s00521-020-05437-x.

[4] Singh D., Kumar V., Kaur M., Kaur M. Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks. *European Journal of Clinical Microbiology & Infectious Diseases*, 2020; 39(7):1379–1389. doi: https://doi.org/10.1007/s10096-020-03901-z10.1007/s10096-020-03901-z.

[5] Wang J., Bao Y., Wen Y., et al. Prior-Attention residual learning for more discriminative COVID-19 screening in CT images. *IEEE Transactions on Medical Imaging*, 2020; 39(8):2572–2583. doi: https://doi.org/10.1109/tmi.2020.299490810.1109/tmi.2020.2994908.

[6] Ifani P., Shalbaf A., Vafaeezadeh M. Automated detection of COVID-19 using ensemble of transfer learning with deep convolutional neural network based on CT scans. *International Journal of Computer Assisted Radiology and Surgery*, 2020; 16(1):115–123. doi: https://doi.org/10.1007/s11548-020-02286-w10.1007/s11548-020-02286-w.

[7] Sun L., Mo Z., Yan F., et al. Adaptive feature selection guided deep forest for COVID-19 classification with chest CT. *IEEE Journal of Biomedical and Health Informatics*, 2020; 24(10):2798–2805. doi: https://doi.org/10.1109/jbhi.2020.301950510.1109/jbhi.2020.3019505.