Detection of COVID-19 using Deep Neural Networks with Chest X-Ray Images-Comparison of DarkNet and ResNet

CS7313 Machine Learning and Pattern Recognition

Punitha Valli Devaraj
Department of Computer Science
Texas State University
p_d120@txstate.edu

Aishwarya Baskaran
Department of Computer Science
Texas State University
a_b1011@txstate.edu

***Abstract*—The catastrophic outbreak of Severe Acute Respiratory Syndrome - Coronavirus (SARS-CoV-2) also known as COVID-2019 has brought the worldwide threat to the living society. It first appeared in Wuhan City of China in December 2019, spread rapidly around the world and evolved into a pandemic. The artificial intelligence researchers are focusing their expertise knowledge to develop mathematical models for analyzing this epidemic situation using nationwide shared data. Application of Artificial Intelligence coupled with radiology imaging can be helpful for the accurate detection of the disease and also to invaluable to overcome the problem of lack of physicians in remote areas. In this study, we propose a deep learning model to identify COVID-19 using the patient's chest X-ray images. We used transfer learning to tune the pre-trained ResNet-50 model, to maximize the overall accuracy to 91.95% whereas the existing DarkCovidNet-17 multiclassification model has 87.02%. This shows the potential for transfer learning and deep neural networks to distinguish amongst normal, pneumonia, and COVID-19 using chest X-rays. This helps to ease the pressure off traditional PCR diagnostic tests, while the further implementation could use deep learning for diagnostic radiology. Our model is available in (https://github.com/punithadevaraj/Machine-Learning-Project)github repository.**

***Keywords*— Covid-19, deep learning, transfer learning, ResNet, chest X-ray images.**

I. Introduction

A newly identified coronavirus, SARS-CoV-2, has caused a worldwide pandemic of respiratory illness, called COVID-19. COVID-19 appeared in Wuhan, a city in China, in December 2019. Although health officials are still tracing the exact source of this new coronavirus, early hypotheses thought it may be linked to a seafood market in Wuhan, China. Some people who visited the market developed viral pneumonia caused by the new coronavirus. A study that came out on Jan. 25, 2020, notes that the individual with the first reported case became ill on Dec. 1, 2019, and had no link to the seafood market. Investigations are ongoing as to how this virus originated and spread(https://coronavirus.jhu.edu/). As of now, researchers know that the new coronavirus is spread through droplets released into the air when an infected person coughs or sneezes. The droplets generally do not travel more than a few feet, and they fall to the ground (or onto surfaces) in a few seconds — this is why physical distancing is effective in preventing the spread. From the current numbers, it has affected 64.2 million cases worldwide, with 1.49 million death cases and 41.2 million recovered cases(https://www.worldometers.info/coronavirus/).

Current diagnostic testing relies on reverse transcription polymerase chain reaction (RT-PCR) which can be time consuming and lead to delays in test results (Ai et al., 2020). Chest radiological imaging such as computed tomography (CT) and X-ray have vital roles in early diagnosis and treatment of this disease. Due to the low RT-PCR sensitivity of 60%-70%, even if negative results are obtained, symptoms can be detected by examining radiological images of patients. It is stated that CT is a sensitive method to detect COVID-19 pneumonia and can be considered as a screening tool with RT-PRC. CT findings are observed over a long interval after the onset of symptoms, and patients usually have a normal CT in the first 0-2 days. In a study on lung CT of patients who survived COVID-19 pneumonia, the most significant lung disease is observed ten days after the onset of symptoms.

The Centers for Disease Control (CDC) prioritizes testing to hospitalized patients, healthcare workers, elderly, immuno-compromised individuals, and those with mild symptoms (CDC, 2020). Despite the knowledge that people without symptoms may still be contagious, the CDC states that asymptomatic individuals are a non-priority (CDC, 2020). Making COVID-19 testing more accessible, available, and fast could help alleviate these problems.

In this study, we propose a solution to the shortage of tests through the development of a convolutional neural network (CNN) for the automatic diagnosis of COVID-19 to identify COVID-19 amongst normal and pneumonia cases using chest X-rays. X-ray imaging is readily available in almost all hospitals and faster than running PCR, allowing more widespread availability and faster results. Thus, the proposed model has an end-to-end architecture without using any feature extraction methods, and it requires raw chest X-ray images to return the diagnosis.

II. BACKGROUND

At the beginning of the pandemic, Chinese clinical centers had insufficient test kits, which are also producing a high rate of false-negative results, so doctors are encouraged to make a diagnosis only based on clinical and chest X-ray results. Researchers state that combining clinical image features with laboratory results may help in early detection of COVID-19. Radiologic images obtained from COVID-19 cases contain useful information for diagnostics. Some studies have encountered changes in chest X-ray and CT images before the beginning of COVID-19 symptoms. Significant discoveries have been realized by investigators in imaging studies of COVID-19(Kong et al) observed right airspace opacities in a COVID-19 patient. Peripheral focal or multifocal GGO affecting both lungs in 50% -75% of patients is observed. Similarly, Zu et al. and Chung et al. discovered that 33% of chest CTs can have rounded lung opacities. In Fig.1, chest X-ray images taken at days 1, 4, 5 and 7 for a 50-year-old COVID-19 patient with pneumonia are given.

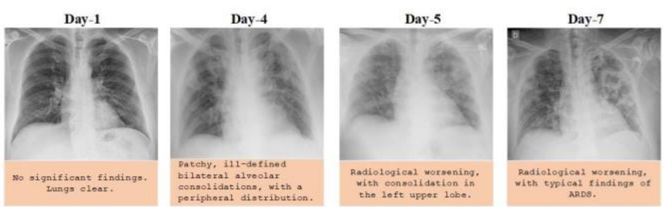

Figure 1. Chest X-ray images of a 50-year-old COVID-19 patient with pneumonia over a week.

Application of machine learning methods for automatic diagnosis in the medical field have recently gained popularity by becoming an adjunct tool for clinicians. Deep learning, which is a popular research area of artificial intelligence (AI), enables the creation of end-to-end models to achieve promised results using input data, without the need for manual feature extraction. Deep learning techniques have been successfully applied in many problems such as arrhythmia detection, skin cancer classification, breast cancer detection, brain disease classification, pneumonia detection from chest X-ray images, fundus image segmentation and lung segmentation. The COVID-19 epidemic's rapid rise has necessitated the need for expertise in this field. This has increased interest in developing the automated detection systems based on AI techniques. It is a challenging task to provide expert clinicians to every hospital due to the limited number of radiologists. Therefore, simple, accurate and fast AI models may be helpful to overcome this problem and provide timely assistance to patients. Although radiologists play a key role due to their vast experience in this field, the AI technologies in radiology can be assistive to obtain accurate diagnosis. Additionally, AI approaches can be useful in eliminating disadvantages such as insufficient number of available RT-PCR test kits, test costs and waiting time of test results.

Our work with X-ray image detection of COVID-19 proposes one possible solution based on the fact that COVID-19 often presents as pneumonia and acute respiratory distress syndrome (CDC, 2020). Past research shows high accuracy in using CNNs to classify X-ray images for thorax diseases (L. Wang & Wong, 2020; X. Wang et al., 2017). Recently, X-ray imaging has also been shown to help detect respiratory distress in COVID-19 patients (Xu et al., 2020). Because COVID-19 acutely attacks the respiratory system, it is likely that X-rays can be used to help identify COVID-19. (Hall et al., 2020; Narin et al., 2020) both used deep learning models to classify normal, pneumonia, and COVID-19 chest X-rays. However, both groups fail to recognize that their normal and pneumonia images are from pediatric patients aged 1 to 5, making their models inapplicable to adults. Wang and Wong proposed a deep model for COVID19 detection (COVID-Net), which obtained 92.4% accuracy in classifying normal, non-COVID pneumonia, and COVID-19 classes. Ioannis et al. developed the deep learning model using 224 confirmed COVID-19 images.

In this paper, we aim to combat this issue by using a more representative dataset of normal, pneumonia, and COVID-19 chest X-rays from adults and develop a machine learning model using transfer learning to accurately distinguish amongst these three classes. So far, only one published paper has done similar work, building a deep CNN to classify chest X-rays into the same three classes (L. Wang & Wong, 2020). Overall, we aim to use transfer learning to tailor a CNN to distinguish amongst normal, pneumonia and COVID-19 adult chest X-rays to propose a potential future solution to the shortage of RT-PCR diagnostic tests for COVID-19.

III. DATASET EXPLORATION

In this study, dataset consisting of X-ray images obtained from two different sources were used for the diagnosis of COVID-19. First is from (Cohen JP (2020) COVID-19 image data collection. A COVID-19 X-ray image database was developed by Cohen JP using images from various open access sources. This database is constantly updated with images shared by researchers from different regions. At present, there are 127 X-ray images diagnosed with COVID-19 in the database. Fig. 2 shows a few COVID-19 cases obtained from the database and the findings of the experts.

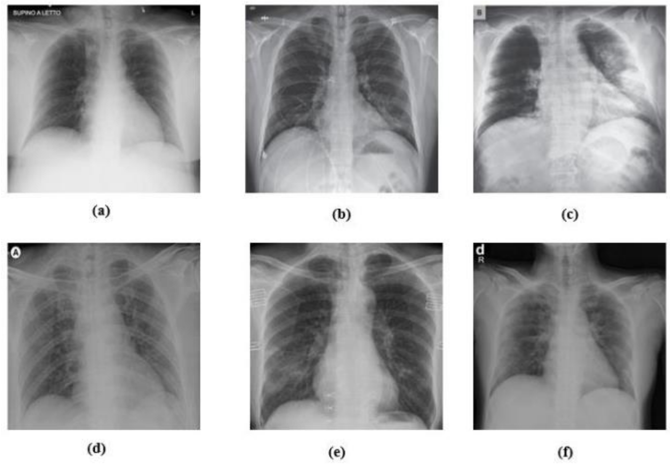

Figure 2. A few COVID-19 cases and findings by dataset: (a) Cardio-vasal shadow within the limits, (b) Increasing left basilar opacity is visible, arousing concern about pneumonia, (c) Progressive infiltrate and consolidation, (d) Small consolidation in right upper lobe and ground-glass opacities in both lower lobes, (e) Infection demonstrates right airspace opacities, and (f) Progression of prominent bilateral perihilar infiltration and ill-defined patchy opacities at bilateral lungs.

Although this repository is an ongoing data collection project, at the time of our data collection stage, we had access to 115 total COVID-19 images, among which 43 are female and 72 are male cases that were found to be positive. The age information of 26 COVID-19 positive subjects is given, and the average age of these subjects is approximately 55 years.

Second dataset for pneumonia and normal chest X-rays were gathered from the NIH Clinical Center (X. Wang et al., 2017). We focused on pneumonia X-rays from the NIH dataset based on prior research that suggests that many COVID-19 patients develop pneumonia (CDC, 2020). The NIH dataset consists of 16,756 images, but due to computational limitations when training our model, we have randomly chosen 1360 normal and 1360 pneumonia images. The existing reference paper of Dark Net has only 500 normal and 500 pneumonia cases. We preferred choosing more images to train the model efficiently as well as with more data.

The final dataset consists of 2,835 X-rays images from the posteroanterior view, sized 224 by 224 pixels (Figure 3). This dataset was split into train and test datasets with 2,625 and 210 images in each, ensuring that patients in the training and test sets are mutually exclusive to prevent multiple X-rays from one patient appearing in both the training and test sets. To ensure generalization, the training data was split into 60% training and 40% validation sets. In comparison to other splitting patterns (Figure 4), a 60-40 split showed the best generalization performance in test accuracy.

Figure 3. Images from the final dataset into 3 classes – Covid-19, Pneumonia, Normal.

Train – Test Split of data	Training Accuracy(%)	Testing Accuracy(%)
80% - 20%	95.01	86.55
60% - 40%	95.08	91.95
50% - 50%	96.51	85.43

Figure 4. Training and Testing split of data

IV. CHOOSING THE MODEL

In this study, before starting to choose the model, we are aware from the available dataset that, one of the major challenges to be faced is the lack of COVID-19 X-rays, consisting of only 4.4% of the training and 0.47% of the test datasets. Second, every patient's body can react differently to infections, which may be apparent in chest infection localization. Furthermore, X-rays are best suited for detecting dense objects, such as bones, and may not provide as high-quality imaging on soft tissue as other techniques such as CT scans or MRIs. This is potentially problematic if X-rays are unable to detect smaller or less aggressive infections.

A. Transfer Learning

The advent of deep learning technology has revolutionized artificial intelligence. The word deep refers to the increase in the size of this network with the number of layers. The structure is named after convolution, a mathematical operator. A typical CNN structure has a convolution layer that extracts features from the input with the filters it applies, a pooling layer to reduce the size for computational performance, and a fully connected layer, which is a neural network. By combining one or more such layers, a CNN model is created, and its internal parameters are adjusted to accomplish a particular task, such as classification or object recognition.

Instead of initiating a deep model development from scratch, a more rational approach is to construct a model using already proven models. In computer vision, transfer learning is usually expressed through the use of pre-trained models. A pre-trained model is a model that was trained on a large benchmark dataset to solve a problem similar to the one that we want to solve. Accordingly, due to the computational cost of training such models, it is common practice to import and use models from published literature (e.g. VGG, Inception, MobileNet). A comprehensive review of pre-trained models' performance on computer vision problems using data from the ImageNet (Deng et al. 2009) challenge is presented by Canziani et al. (2016).

From the practical perspective, the entire transfer learning process can be summarized as

- Select the pre – trained model.
- Classify the problem according to size-similarity matrix (see Figure 5).
- Fine-tune the model

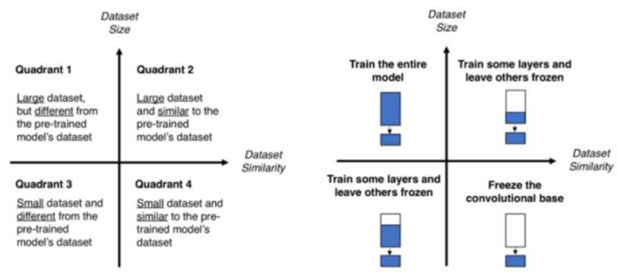

Figure 5. Size-Similarity matrix (left) and decision map for fine-tuning pre-trained models (right).

Therefore, we have selected Quadrant 1 since ours is large dataset(>1000 images) and different from pre-trained models dataset (cats and dogs).

B. SVM Classifier

Support vector machine is highly preferred by many as it produces significant accuracy with less computation power. Support Vector Machine, abbreviated as SVM can be used for both regression and classification tasks. But, it is widely used in classification objectives. The objective of the support vector machine algorithm is to find a hyperplane in an N-dimensional space(N - the number of features) that distinctly classifies the data points. Support vectors are data points that are closer to the hyperplane and influence the position and orientation of the hyperplane. Using these support vectors, we maximize the margin of the classifier. Deleting the support vectors will change the position of the hyperplane. These are the points that help us build our SVM(Figure 6).

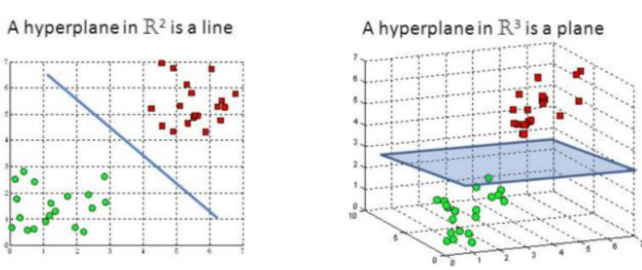

Figure 6. Hyperplanes in 2D and 3D feature space.

Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyperplane can be attributed to different classes. Also, the dimension of the hyperplane depends upon the number of features. If the number of input features is 2, then the hyperplane is just a line. If the number of input features is 3, then the hyperplane becomes a two-dimensional plane. In the existing reference paper, DarkNet-19 is used as a classifier, but our model uses SVM with linear kernel. The epsilon value of the linear kernel is set as 0.04.A kernel trick is employed to transfer the input data to another hyperplane where the data is more convenient for linear separation. The best hyperplane can be determined by using Eq.1

$$\left\{\begin{matrix} \min \frac{\|w\|^2}{2} \\ y_i\left(w^T x_i + b\right) \geq 1 i = 1, 2, \cdots, M \end{matrix}\right)$$

where w indicates the weight vector and b is bias value used to determine the position of the hyperplane. The linear separation of the positive and negative samples can be handled using Eq.2

$$f(x) = w^T x + b = 0$$

C. Selecting the Pre-trained Model – ResNet 50

After developing a baseline SVM, we used transfer learning to leverage architecture and pre-trained ImageNet weights, then tuned hyper parameters, added layers, and trained on our X-images to better fit our project. We selected the final model based on highest overall accuracy on the test data, and analyzed performance using a variety of metrics including class precision and recall. The decision tree for choosing the model is shown below (Figure 7).

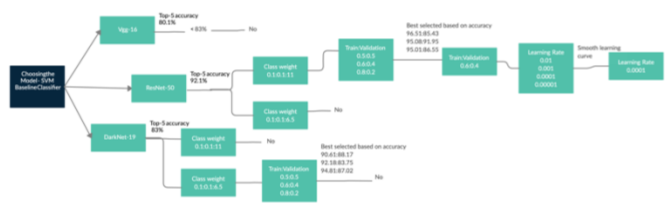

Figure 7. Decision Tree for Model Selection.

Decision tree is one of the predictive modelling approaches that identifies ways to split data set based on different condition and widely used in data mining, statistics and machine learning. Decision trees are non-parametric supervised learning method used for both classification and regression tasks. They depend on important factors such as follows:

- Information gain - measures how much information a feature gives us about the class. The split with the highest information gain will be taken as the first split and the process will continue until all children nodes are pure, or until the information gain is 0.
- Gini Impurity – Pure generally means, in a selected sample of dataset all data belong to same class. Impure means it is mixture of different classes. It is the measurement of the likelihood of an incorrect classification of a new instance of a random variable, if that new instance were randomly classified according to the distribution of class labels from the data set.

For the test accuracy data split See (Figure 4). Learning rate at 0.0001 produced smoother learning curve (Figure 8) than higher learning rates and enabled faster convergence than lower learning rates.

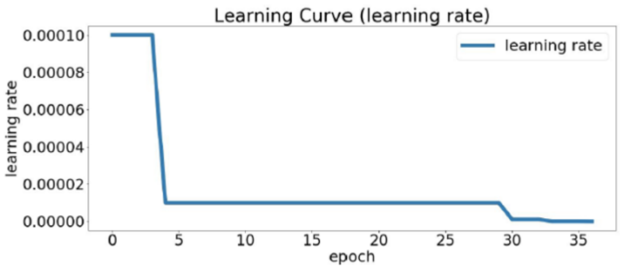

Figure 8. Smooth learning curve at 0.0001 learning rate

At the beginning of training, a higher learning rate can help model find different possibilities and avoid being trapped into local minimums too early. After we trained for several epochs, we saved the best result from these epochs and proceeded the training on that result with lower learning rates which helps the model converge smoothly.

V. RESNET-50 ARCHITECTURE

According to the universal approximation theorem, given enough capacity, we know that a feedforward network with a single layer is sufficient to represent any function. However, the layer might be massive, and the network is prone to overfitting the data. Therefore, there is a common trend in the research community that our network architecture needs to go deeper. However, increasing network depth does not work by simply stacking layers together. Deep networks are hard to train because of the notorious vanishing gradient problem, as the gradient is back propagated to earlier layers, repeated multiplication may make the gradient infinitively small. As a result, as the network goes deeper, its performance gets saturated or even starts degrading rapidly.

The core idea of ResNet – Residual Network is introducing a so-called "identity shortcut connection" that skips one or more layers. The diagram below(Figure 9) illustrates skip connection. It stacks convolution layers but also adds the original input to the output of the convolution block. This is called skip connection.

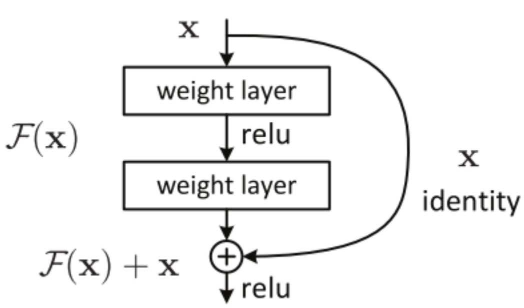

Figure 9. Residual block

The Resnet-50 model (Figure 10) consists of 5 stages each with a convolution and Identity block. Each convolution block has 3 convolution layers and each identity block also has 3 convolution layers. The ResNet-50 has over 23 million trainable parameters.

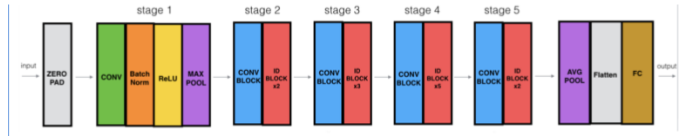

Figure 10. ResNet-50 Model Architecture

Winning the ImageNet challenge in 2015, ResNet uses residual learning, which decreases the computational complexity and expense of training deep neural networks while solving the issue of degrading training accuracy. In addition, based on the research results done by X. Wang et al., ResNet-50 is the most efficient pre-trained model for thorax disease identification (X. Wang et al., 2017). After testing various available pre-trained models through Keras, we found that ResNet-50 had the best performance in overall accuracy.

VI. PROPOSED RESNET-50 ARCHITECTURE

The proposed architecture for this project has some slight variations to the existing pre trained ResNet model. The architecture of our model is shown in Figure 11.

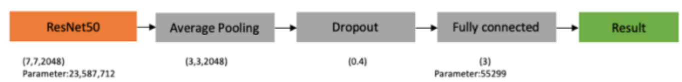

Figure 11. Proposed ResNet-50 Model Architecture

We implemented an average pooling layer between ResNet-50 output and fully connected layer to reduce the dimensionality of the ResNet50 output and retain high feature variation. Since X-ray images include many high value pixels representing white colors, applying a Max Pooling layer may cause loss of variation in the image, especially for pneumonia and COVID-19 patients, leading to poor performance. Hence, the Average Pooling layer is a better choice for this project. But the existing DarkNet model uses Max pooling layer.

After reducing the dimensionality, we decreased the output shape from (7, 7, 2048) to (3, 3, 2048). This was still high dimensional data compared to our output of 3 classes and can increase the chance of overfitting. Therefore, we implemented a Dropout layer as a regularization method to avoid overfitting and reduce computation time. Our final model architecture has almost 24 million parameters, so we augmented our data by randomly rotating, shifting, flipping, cropping, and shearing the training data.

This augmentation can gauge model accuracy and prevent overfitting to a relatively small number of images in our dataset without sacrificing computation ability or choosing a smaller network (Scott et al., 2017). These operations help our model converge within 30 epochs without overfitting. Early Stopping feature is also implemented in this project, where this feature is widely used to combat the overfitting issue. The way it does is to stop training as soon as the validation error reaches a minimum. As the epochs go by, the algorithm leans and its error on the training set naturally goes down, (Figure 12) and so does its error on the validation set. However, after a while, the validation error stops decreasing and actually starts to go back up. This indicates that the model has started to overfit the training data. With Early Stopping, you just stop training as soon as the validation error reaches the minimum

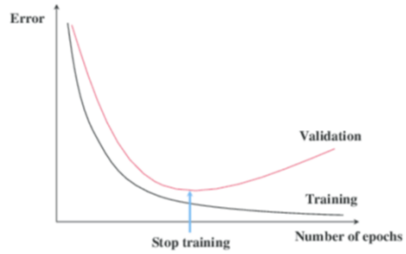

Figure 12. Early Stopping

In our project, we are monitoring validation loss and patience=10, mode=min. In this mode, training will stop when the measured quantity has stopped decreasing.

We have used the Adam optimizer same as of DarkNet for weight updates, categorical cross entropy loss function instead of cross entropy because they are best in multiple class classification problem and selected learning rate as 0.0001.

VII. Experimental results

A. Precision Recall(PR) and Receiver operating characteristics(ROC) curves

We performed experiments to detect and classify COVID-19 using X-ray images in two different scenarios. First, we have trained the Resnet-50 with certain tuning features as mentioned above with early stopping and without early stopping. The data model to classify X-ray images into three categories: Normal, Pneumonia, COVID-19. They are given label categorization as 0,1 and 2 respectively. We are training the model to 100 epochs.

Findings/Accuracy	Train		Test	
	Precision	Recall	Precision	Recall
Normal	0.95	0.96	0.94	0.89
Pneumonia	0.94	0.95	0.90	0.92
Covid-19	0.96	0.95	0.83	0.88
Accuracy	95.08		91.95	

Figure 13. Class Precision, Recall, overall accuracy for training and test data.

The final model was selected based on highest overall accuracy with train and test dataset accuracies at 95.08% and 91.95% respectively (Figure 13). We also gauged model performance on class precision and recall. On the test data, these results showed that the final model performed well in identifying all three classes, with both precision and recall for normal, pneumonia and COVID-19 classes at 0.92, 0.91 and 0.93 respectively.

Two diagnostic tools that help in the interpretation of probabilistic forecast for binary (two-class) classification predictive modeling problems are ROC Curves and Precision-Recall curves.

- ROC curves summarize the trade-off between the true positive rate and false positive rate for a predictive model using different probability thresholds.
- Precision-Recall curves summarize the trade-off between the true positive rate and the positive predictive value for a predictive model using different probability thresholds.
- ROC curves are appropriate when the observations are balanced between each class, whereas precision-recall curves are appropriate for imbalanced datasets.

The main difference between ROC curves and precision-recall curves is that the number of true-negative results is not used for making a PRC.

We evaluated model performance by plotting the PR and ROC curves (Figure 14). In the PR curves (Figures 14a, 14b), the curves for all three classes (0: normal, 1: pneumonia, 2: COVID-19) hug the upper right corner, confirming that the final model performs well in terms of precision and recall for all three classes. In the ROC curves (Figures 14c, 14d), the curves for all three classes hug the upper left corner, indicating that the model performs well in terms of true and false positive rates. Overall, these curves by class confirm that the model performs well in distinguishing amongst the three classes.

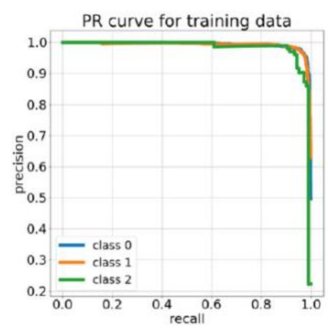

Figure 14a

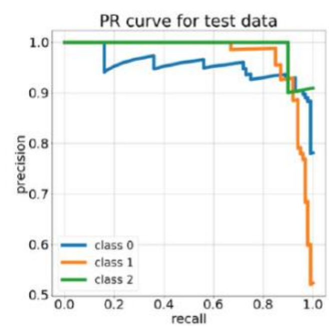

Figure 14b

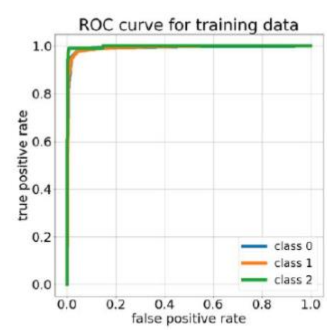

Figure 14c

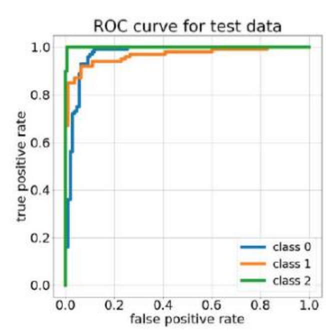

Figure 14d

Figure 14. PR(a,b) and ROC(c,d) curves on the training (a,c) and test(b,d) datasets where class 0:Normal 1:Pneumonia 2:Covid-19.

B. Confusion Matrix

It is a performance measurement for machine learning classification problem where output can be two or more classes. It is a table with 4 different combinations of predicted and actual values, in case of 2 classes. In our project, there are 9 values since it is 3 classes.

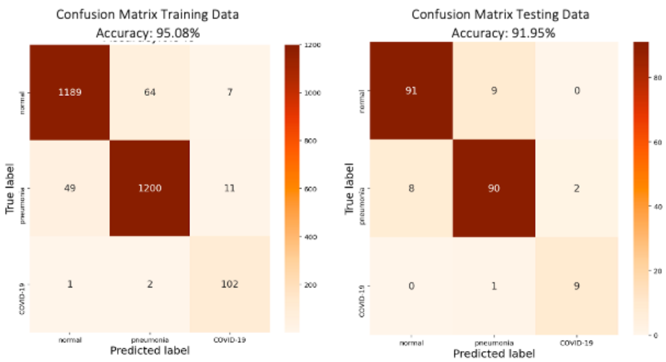

Figure 15. Confusion Matrix for Training and Testing Data

Based on the confusion matrices (Figure 15), our model performs well, with low counts on the off- diagonal indicating few incorrect predictions. On the test data, only 1 image were falsely classified as having pneumonia or CO VID-19 (Figure 15), which is a relatively low false positive rate. The overlapped confusion matrix of DarkCovidNet (referenced paper) is shown in Figure 16. It achieved classification accuracy of 87.02% to classify: no-findings(normal), covid-19 and pneumonia.

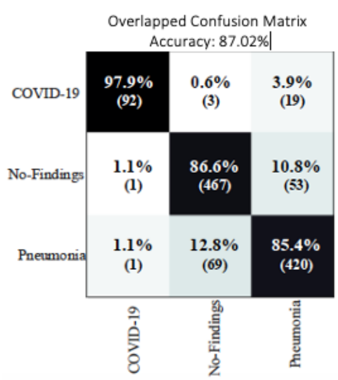

Figure 16. Overlapped Confusion Matrix – DarkCovidNet Model

We compared our final ResNet model to a support vector machine (SVM) trained on the same training data (Figure 17).SVMs are advantageous in image classification because they can handle nonlinear class boundaries and high-dimensional data while being flexible enough to yield high-performing models that can often generalize well in comparison to other more inflexible classifiers like logistic regression. The accuracy of the SVM is 99.0% on the training data and 83.81% on the test data, indicating that the SVM is overfitting to the training data. Based on overall accuracy, our ResNet model performs the best. ResNet performs slightly better than the SVM for normal and pneumonia classes, but the SVM clearly struggles with COVID-19. By analyzing class precision and recall, we can clearly see the strength of the CNN in identifying COVID-19.

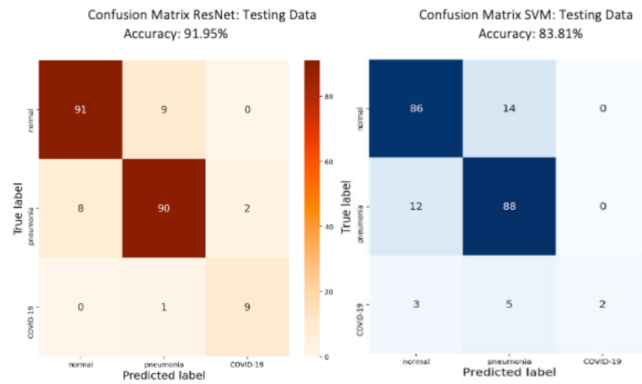

Figure 17. ResNet and SVM Confusion Matrix on Testing data

C. Classification Report

Classification report is generated for referenced DarkCovidNet model, our Resnet model and SVM (Figure.18)

	PERFORMANCE METRICS (%)			
Model	Precision	Recall	F1 score	Accuracy
DarkCovidNet	89.96	85.35	87.37	87.02
ResNet-50	91.67	90.38	90.1	91.95
SVM	87.79	89.42	85.81	83.81

Figure 18. Classification Report – Comparison.

From this report, it is clearly evident that ResNet-50 model performs better than the other two.

D. Proposed Model's battle and defeat

We have clearly examined classification patterns to identify areas where the model succeeds and defeats the misclassifications, while on the other hand battles hard and struggles. Specifically, we selected two misclassified images in the test data:
Figure 19a - Predicted pneumonia but actually COVID-19
Figure 19b - Predicted COVID-19 but actually pneumonia.

First, Figure 19a shows white areas of inflammation similar to those seen in pneumonia X-rays (Simpson et al., 2020), indicated by the red arrows. Second, it is possible that the model struggles when images contain a large amount of white matter, as in Figure 19b, leading to misclassification as COVID-19. Overall, in looking at these two specific misclassified images from the test data, we hypothesize that the model struggles and battles hard when images contain a lot of white matter in less clearly defined regions. Further research and advice from radiologists could clarify areas where our model struggles.

We can also identify areas where the model succeeds by selecting images with high probabilities of correct predictions (Figure 20). We hypothesize that when the X-ray images are clearer and have less “haze,” the model can more accurately classify. It is also possible that the model succeeds when images are already well-centered rather than shifted or not filling the entire space.

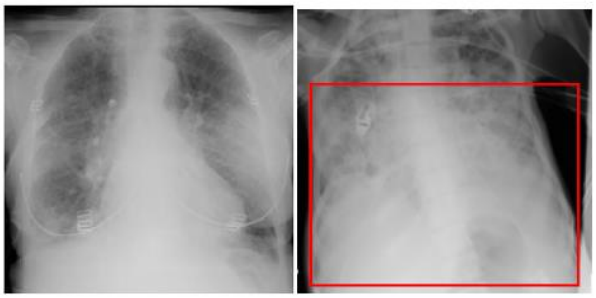

Figure 19a and Figure 19b. X-rays Misclassification – Model battles

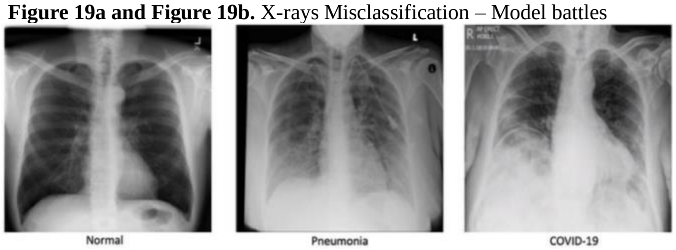

Figure 20. Correct classification – Model defeats and succeeds.

VIII. CONCLUSION

In this study, we have proposed a transfer learning-based model to detect and classify COVID-19 cases from X-ray images. Our model is fully automated with an end-to-end structure without the need for manual feature extraction. This system can be used in remote places in countries affected by COVID-19 to overcome a shortage of radiologists. Also, such models can be used to diagnose other chest-related diseases including tuberculosis and pneumonia. Limitation of the study is the use of a limited number of COVID-19 X-ray images. We intend to make our model more robust and accurate by using more such images from our local hospitals.

Future work as follows:

- Should seek advice from a radiologist to better understand these X-rays to improve model performance.
- We have to get more images for the dataset and try training with all the existing Pre-trained models with different learning rates
- A real-time webpage can be created for the doctors to use and test this X-rays which updates regularly
- An App can be created.

IX. CONTRIBUTION

Team Members:

- Punitha Valli Devaraj
- Aishwarya Baskaran

The contributions are mentioned as below:

Punitha Valli Devaraj

- Set up the environment with all the necessary libraries installed.
- Trained with 60:40,80:20 data split to select the best train: test split accuracy
- Code implementation on validation and train splits, assisted in data preprocessing, defining the model, training and testing with confusion matrix.
- Trained and tested ResNet50-100 epochs, with and without early stopping.
- Performance metrics was developed and analyzed for the same
- Made the decision tree and tried various optimization techniques.
- Prepared the Project presentation ppt
- Prepared the Report (Since my teammate had to travel to India due to personal emergency reasons).

Aishwarya Baskaran

- Dataset exploration – from the 2 sources
- Trained with 50-50 data split to select the best train: test split accuracy
- Coded with path set up, data loading, data preprocessing
- Trained and tested the model with SVM classifier, Vgg-16
- Performance metrics was developed and analyzed for the same
- Assisted when training ResNet-50 with 100 epochs and analyzing the results.
- Classification report generation of SVM, ResNet and DarkNet
- Helped in writing the Introduction, background, data split and conclusion part in report

As a team

- Researched through many reference papers, read those and then had a clear understanding of what project we are going to implement
- Read many articles and papers regarding the existing Pre-Trained models and fixed on the model of ResNet in comparison with existing DarkNet paper.
- We used shared Google Colab Notebook, so both can edit and run the code.
- As a team, we have discussed and made conclusion to go with 60-40% data split.
- Adding the layers, deciding on learning rate, what all data augmentations has to be done is decided as a team

- We trained and tested multiple times all the above-mentioned models
- Code and Report is tested multiple times in terms of precision, accuracy both quantitatively and qualitatively.

X. ACKNOWLEDGEMENT

We would like to take this moment to offer our sincere thanks to our Professor Dr. Yan Yan for teaching us various techniques and methodologies in Machine learning class. Also the referenced paper[1] really helped a lot for our basic understanding of this project and move forward with the help of that score. I would like to thank all the authors in that paper.

REFERENCES

[1] Tulin Ozturk,a Muhammed Talo,b Eylul Azra Yildirim,c Ulas Baran Baloglu,d Ozal Yildirim,e,∗ and U. Rajendra Achary af,g,h,"Automated detection of COVID-19 cases using deep neural networks with X-ray images",April28,2020

[2] Wu, F., Zhao, S., Yu, B., et al. (2020). A new coronavirus associated with human respiratory disease in China. Nature, 579(7798), 265-269.

[3] Huang, C., Wang, Y., et al. (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. The Lancet, 395(10223), 497-506.

[4] World Health Organization. (2020). Pneumonia of unknown cause–China. Emergencies preparedness, response, Disease outbreak news, World Health Organization (WHO).

[5] Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., ... & Xu, B. (2020). A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). medRxiv.

[6] Zheng, C., Deng, X., Fu, Q., Zhou, Q., Feng, J., Ma, H., ... & Wang, X. (2020). Deep Learning-based Detection for COVID-19 from Chest CT using Weak Label. medRxiv.

[7] Shuo Wang, Yunfei Zha, Weimin Li, Qingxia Wu, Xiaohu Li, Meng Niu, Meiyun Wang, Xiaoming Qiu, Hongjun Li, He Yu, Wei Gong, Yan Bai, Li Li, Yongbei Zhu, Liusu Wang, JiTian. A Fully Automatic Deep Learning System for COVID-19 Diagnostic and Prognostic Analysis.ERS Publications

[8] Moutaz Alazab , Albara Awajan , Abdelwadood Mesleh, Ajith Abraham, Vansh Jatana , Salah Alhyari. COVID-19 Prediction and Detection Using Deep Learning. Research Gate

[9] Rajpurkar, P., Irvin, J., et al. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. arXiv preprint arXiv:1711.05225.

[10] Hemdan, E. E. D., Shouman, M. A., & Karar, M. E. (2020). COVIDX-Net: A framework of deep learning classifiers to diagnose COVID-19 in x-ray images. arXiv preprint arXiv:2003.11

[11] Ioannis D. Apostolopoulos1, Tzani Bessiana, COVID-19: Automatic detection from X-ray images utilizing Transfer Learning with Convolutional Neural Networks, arXiv:2003.11617

[12] Narin, A., Kaya, C., & Pamuk, Z. (2020). Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. arXiv preprint arXiv:2003.10849.

[13] Song, Y., Zheng, S., Li, L., Zhang, X., Zhang, X., Huang, Z., ... & Chong, Y. (2020). Deep learning Enables Accurate Diagnosis of Novel Coronavirus (COVID-19) with CT images. medRxiv.

[14] https://openaccess.thecvf.com/content_cvpr_2017/papers/Wang_ChestX-ray8_Hospital-Scale_Chest_CVPR_2017_paper.pdf

[15] https://pjreddie.com/darknet/imagenet/

[16] https://keras.io/api/applications/

[17] https://acutecaretesting.org/en/articles/precision-recall-curves-what-are-they-and-how-are-they-used#:~:text=A%20precision%2Drecall%20curve%20shows,for%20every%20possible%20cut%2Doff

[18] https://towardsdatascience.com/transfer-learning-from-pre-trained-models-f2393f124751

[19] https://towardsai.net/p/programming/decision-trees-explained-with-a-practical-example-fe47872d3b53

[20] https://towardsdatascience.com/decision-tree-in-machine-learning-e380942a4c96