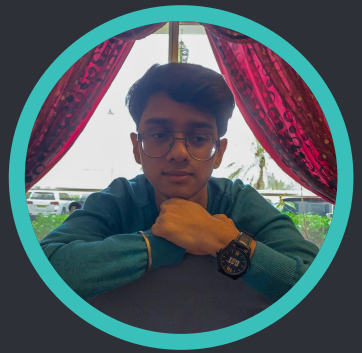
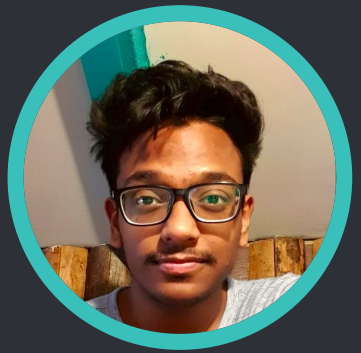
Image Super-Resolution based Classification of COVID-19 patients using CNN-SVM.

Group Members

Milan Ashvinbhai Bhuva
IIT2018176
"Why is you leaving?"



Manav Kamlesh Agrawal
IIT2018178
Yeh kaha phas gaya bich mein??



Mohammed Aadil
IIT2018179
"Why are you running?"

"Soo many question soo few time"

Problem Statement and Objective

1

The first case of the Coronavirus Disease (COVID-19) was registered in December 2019 in Wuhan, China. This extremely contagious disease spread worldwide in a very short time, thus resulting in a global pandemic. To curb the spread of COVID-19, it is important to detect the virus in an infected person at the earliest. To do we are implementing a AI system that can identify covid 19 patients using X-ray images. For developing the AI system we will use VGG16 as our CNN model and SVM as our classifier.

2

Literature Survey

A brief summary of the papers we surveyed.

RP 01: Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based CNN

Year: 2020

Method: They proposed multi-objective differential evolution (MODE) based on CNN for classification of COVID-19-infected patients from chest CT images. Initially they used CNN, ANN, and ANFIS models and these provided good performance for classification of COVID-19 patients, however CNN suffered from hyperparameters tuning issue.

They divided the process into two classes to classify COVID-19-infected patients using CNN model.
Feature Extraction is done by CNN.

They use a multi objective fitness function: $f(t) = S_n + S_p$, S_n and S_p define the sensitivity and specificity parameters, respectively.

Dataset: Not mentioned.

Results:

Model	ACC	SEN	SPE
MODE based CNN	98	98.2	92.2

RP 02: Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep CNN

Year: 2020

Method: Chest X-ray images of 341 COVID-19 patients have been obtained from the open source GitHub repository shared by Dr. Joseph Cohen et al. 80% for training and 20% for testing. All images were resized to 224x224 pixel size in the datasets.

They implemented three different binary classifications with four classes (COVID-19, normal (healthy), viral pneumonia and bacterial pneumonia) by using 5-fold cross validation, and used ResNet50, ResNet101, ResNet152, InceptionV3 and Inception-ResNetV2 as CNN models.

Dataset: COVID Chest X-rays: https://github.com/ieee8023/covid-chestxray-dataset
Hospital-scale X-ray Dataset: https://doi.org/10.1109/CVPR.2017.369
Kaggle X-ray images: https://www.kaggle.com/paultimothymooney/chest-xray-pnemonia

Results:

Model	ACC	REC	SPE	F1
ResNet50	96.1	91.8	96.6	83.5
ResNet101	96.1	78.3	98.2	81.2
ResNet152	93.9	65.4	97.3	69.8
InceptionV3	95.4	90.6	96.0	81.1
Inception-ResNet	94.2	83.5	95.4	74.8

RP 03: Automated Systems for Detection of COVID-19 Using Chest X-ray Images and Lightweight CNN **[Base 1]**

Year: 2020

Method: Dataset: Cohen dataset (GitHub), Italian Society of Medical and Interventional Radiology (SIRM) website, Radiopaedia and Radiological Society of North America (RSNA), a collection of normal and pneumonia images are taken from Kaggle publicly available dataset. Used augmentation to prevent CNN from overfitting and memorizing the training set. Images were preprocessed and the following classifiers were used :

- CNN used Softmax classifiers and methods used were AOCTNet, shufflenet, mobilenet, transfer learning.
- Support vector machine (SVM) used RBF
- K nearest neighbor (KNN) used hamming distance function with 3 neighbours
- Random Forest used 100 bags for bootstrapping

Dataset: COVID Chest X-rays: https://github.com/ieee8023/covid-chestxray-dataset

Results:

Model	ACC	SEN	SPE
AOTCNeT-Softmax	99.24	99.24	99.62
SVM	98.60	98.60	99.30
RF	99.46	99.46	99.73
KNN	99.46	99.46	99.73
MobileNet-KNN	99.46	99.46	99.73
ShuffleNet-KNN	99.35	99.35	99.68

RP 04: Deep MLP-CNN Model Using Mixed-Data to Distinguish between COVID-19 and Non-COVID-19 Patients

Year: 2020

Method: They considered both X-ray images and numerical/categorical dataset which contains data from 184 different patients containing information such as age, gender, temperature, survival, intubation, partial pressure of oxygen dissolved in the blood and classification like COVID-19, SARS, Pneumocystis, E. coli, Streptococcus, or “no findings” patients
They tested the CNN model on this dataset using three different optimization algorithms:

- Adaptive learning rate optimization algorithm (Adam)
- Stochastic gradient descent (Sgd)
- Root mean square propagation (Rmsprop)

Furthermore, they considered Relu as their activation function.

Dataset: COVID Chest X-rays: https://github.com/ieee8023/covid-chestxray-dataset

Results:

Model	ACC	REC	SPE	F1
CNN-Adam	94.6	93.5	94.5	90.7
CNN-RMSprop	88.9	85.9	88.9	86.6
CNN-SGD	88.4	88.7	88.5	83.0

RP 5: CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images

Year: 2020

Method: This paper proposes using CoroNet, a DCNN model to classify COVID-19 patients by merely grazing chest X-ray images. This proposed model is pre-trained on ImageNet dataset and trained end-to-end on a dataset of COVID-19 and other chest pneumonia X-ray images.

Dataset: COVID Chest X-rays: https://github.com/ieee8023/covid-chestxray-dataset
Hospital-scale X-ray Dataset: https://doi.org/10.1109/CVPR.2017.369

Results:

Model	ACC	REC	SPE	F1
CORO-Net	90.21	90	95	91

RP 6: FIND NEW PAPER**

Year: 2020

Method: This paper suggests 5 step methodology consisting of :

- Database formulation
- Lung Segmentation : is an important as it resolves noises or irrelevant patterns in the image which can lead to false predictions.
- Data Augmentation : As the dataset is really small, and DCNN relies heavily on a large dataset, we create one using this trick.
- Feature Extraction : They propose to use the mAlexNet CNN architecture to extract features, and while they are at it use it to test the mAlexNet CNN prediction.
- Hybrid Model : They use BiLSTM as the new hybrid model.

Dataset: COVID-19 Radiography Database: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database

Results:

Model	ACC	SEN	SPE	F1
mAlexNet	98.14	98.26	99.06	98.2
mAlexNet + BiLSTM	98.7	98.76	99.33	98.76

RP 7: A Review on Deep Learning Techniques for the Diagnosis of Novel Coronavirus

Year: 2020

Results: The paper reviewed a total of 45 COVID-19 diagnosis systems.

- Among these, 23 systems used pre-trained models for diagnosis purposes and 22 models used custom deep learning techniques for COVID-19 detection.
- The two most popular datasets were CT scans and X-Ray images, while some used the mix of these.
- It was also noticed that pre-trained models were faster than custom models and also had a better average prediction accuracy. The highest accuracy noted was 99.51%.
- The major challenge in implementing these systems in real time lies in the fact that the image datasets provided are incomplete, noisy, ambiguous, and inaccurate labels in some cases,
- A small ratio of positive to negative covid samples is plainly one-sided.
- However, LSTM (shallow long short- term memory) and synthetic generation of data can be used for these shortcomings.

RP 8: Review on Machine and Deep Learning Models for the Detection and Prediction of Coronavirus

Year: 2020

Results: Various DL based algorithms have conceived that Dl can make remarkable impacts in detecting COVID-19. Their accuracy is beyond imagination and with growing data they might be able to replace existing COVID-19 checks in the near future. Most of the algorithms have produced an accuracy of over 90% or an f score of over 0.89.

Conclusion: Due to the rapid spread of coronavirus, these DL algorithms can prove to be instrumental in the future, however we cannot currently implement these due to the current lack of data and limitation of related studies.

Future research could incorporate distinguishing mild and harsh symptoms and could also work on improving the accuracy of detection of covid 19 in their early stages.

RP 9: GoogleNet CNN Neural Network towards Chest CT- Coronavirus Medical Image Classification

Year: 2020

Method: The paper has divided the proposed methodology in three main stages :

- Preprocessing : which includes conversion of image from grayscale to RGB format, followed by binarization.
- Retraining GoogleNet CNN : The CNN algorithms was retrained over the new dataset named COVID-CT-DATASET
- Classification : Softmax classifier along with Stochastic gradient descent algorithm was used.

Dataset: X-ray and CT snapshots of CONVID-19 patients: https://www.kaggle.com/andrewmvd/convid19-x-rays

Results:

Model	ACC
GoogleNet CNN	82.14

RP 10: A Deep Learning Based Hybrid Approach for COVID-19 Disease Detections

Year: 2020

Method: The freely available kaggle chest X-Ray dataset was used for this study. The hybrid model developed used Resnet 50 architecture. The input layer of resnet 50 was modified to 240*240*1 and 5 layers of the Resnet architecture were replaced by 10 new layers The algorithm utilised along with this CNN model was noted to be SGD (Stochastic gradient descent) and mini batch size was kept to be 10.

Dataset: Kaggle Chest X-rays (Pneumonia): https://www.kaggle.com/paultimothymooney/chest-xray-pnemonia

X-ray and CT snapshots of CONVID-19 patients: https://www.kaggle.com/andrewmvd/convid19-x-rays

Results:

Model	ACC	REC	SPE	F1
Improved ResNet	96.30	96	98	96
ResNet 50	92.59	92.5	94	92
AlexNet	88.89	88.2	95	86.5
GoogleNet	90.74	92.5	96.1	90.12

RP 11: Transfer learning to detect COVID-19 automatically from X-ray images, using CNN. **[Base 2]**

Year: 2020

Method: Very Similar to the standard CAT v DOG classifier using CNN + DNN, where we first pass the data through a CNN(Pre-Trained Model) to extract the features then, passed it into a DNN after flattening the features

They used (ReLU) in all CNN layers. A dropout layer was added to prevent the occurrence of overlapping for NN using 2 hidden layers.
The CNNs were compiled using the RMSprop(lr=1e-5) optimization method.
The training lasted 15 epochs, with the batch size set to 32 and time taken was at max 2 mins.

Dataset: COVID Chest X-rays: https://github.com/ieee8023/covid-chestxray-dataset
COVID-19 Radiography Database: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database

Results:

Model	ACC	F1
COVID-Net	92.4	90
Bayes-SqueezeNet	98.3	98.3
VGG-19	93.48	-
MobileNet-V2	94.72	-
ResNet 50+SVM	-	95.52
DarkCOVIDNet	87.02	87.37
CORO NET	95	95.6
VGG-16	99.57	99.36

RP 12: Covid-19: Automatic detection from X-ray images utilizing transfer learning with CNN

Year: 2020

Method: They used (ReLU) in all CNN layers. A dropout layer was added to prevent the occurrence of overlapping for NN using 2 hidden layers.
The CNNs were compiled using the RMSprop(lr=1e-5) optimization method.
The training lasted 15 epochs, with the batch size set to 32 and time taken was at max 2 mins.

Dataset: COVID Chest X-rays: https://github.com/ieee8023/covid-chestxray-dataset
COVID-19 Radiography Database: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database

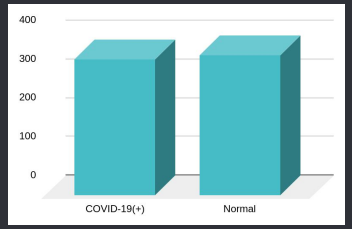
Results:

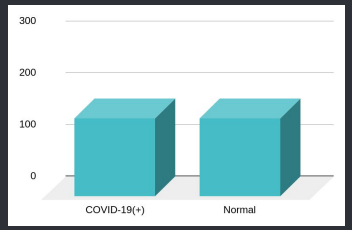
Model	ACC	F1
COVID-Net	92.4	90
Bayes-SqueezeNet	98.3	98.3
VGG-19	93.48	-
MobileNet-V2	94.72	-
ResNet 50+SVM	-	95.52
DarkCOVIDNet	87.02	87.37
CORO NET	95	95.6
VGG-16	99.57	99.36

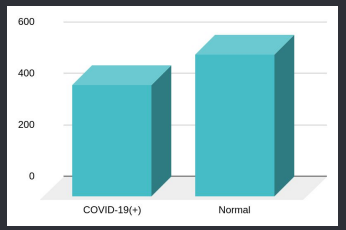
RP 13: COVID-19 DATASETS: A SURVEY AND FUTURE CHALLENGES

Datasets :

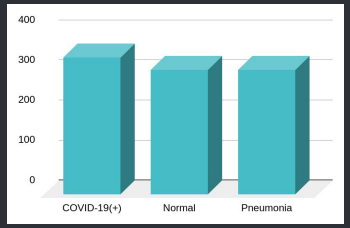
2 Classes

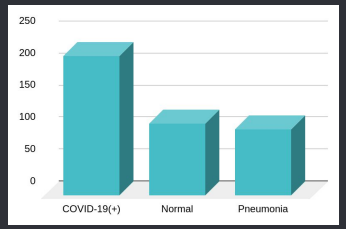


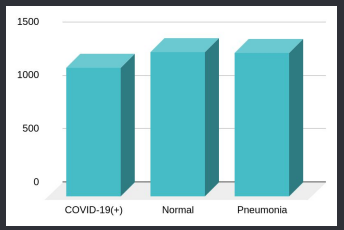




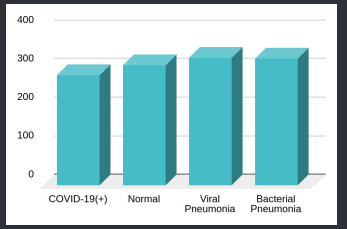
3 Classes

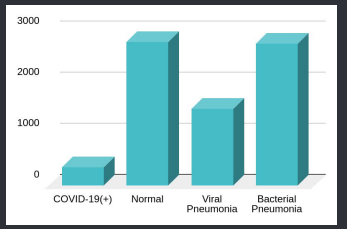






4 Classes





RP 14: Rapid COVID-19 Diagnosis Using Deep Learning of the Computerized Tomography Scans

Year: 2020

Method: This paper aims to review and implement AI image-based diagnosis methods in order to detect coronavirus infection with much better results. Besides the already existing AI image-based medical diagnosis method for the other well-known disease, this study aims on finding the most accurate COVID-19 detection method among AI methods such as ML and ANN, ensemble learning (EL) methods.
They have not experimented with too many mixed models, but they tried ensemble models but the accuracy was too low. So they ended up simply using a single model for their classifications.

Dataset: Kaggle Chest X-rays (Pneumonia): https://www.kaggle.com/paultimothymooney/chest-xray-pnemonia

Results:

Model	ACC	REC	F1
MLP	0.9400	0.8715	0.9268
CNN	0.9760	0.9724	0.9724
SVM	0.9920	1.000	0.990
NB	0.9400	0.9541	0.9327
AdaBoost	0.9600	0.9633	0.9545
GBDT	0.9520	0.9816	0.9469

RP 15: Covid-2019 Detection Using X-Ray Images And Artificial Intelligence Hybrid Systems

Year: 2020

Method: Images are classified using different ML algorithms and their performance was tested to recognize the best of them, these algorithms include CNN, SVM, and Random Forest. The CNN is exploited here in two scenarios, the first one in classifying the X-ray images, and the second one in extracting the graphical features for hybrid system implantation.
Same CNN with multiple layers and ReLu activation function. Then he extracts features using the CNN model and feeds to other classifiers like SVM and RF.

Dataset: COVID Chest X-rays: https://github.com/ieee8023/covid-chestxray-dataset
Mendeley Data: https://data.mendeley.com/datasets/2fxz4px6d8/4

Results:

Model	ACC	SEN	SPE
CNN-Softmax	95.2	93.3	100
CNN-SVM	90.5	86.7	100
CNN-RF	81	76.5	100

RP 16: COVID-19 Detection using SVM Classifier

Year: 2020

Method: In order to classify the normal and COVID-19 cases, a CAD system is developed using Support Vector Machine (SVM) Classifier with a Graphical User Interface. The pre-processed X-ray images are ROI selected of 25x25 pixel size window and GLCM features are extracted.
The extracted contrast, correlation, energy and homogeneity features from 30 images are used to train the SVM model and the best curve fit with 5-fold cross validation is found to classify the COVID-19 against normal images.

Dataset: No access link.

Results:

Model	ACC
GLCN -> SVM	57.1

RP 17: An Artificial Intelligence System for Classification of COVID-19 Suspicious Person using Support Vector Machine (SVM) Classifier

Year: 2020

Method: The data is collected from the people of two corona-free villages in the form of questionnaires. The dataset contains integer and categorical types of 20 attributes. After the initial pre-processing, this pre-processed dataset is provided as an input to the various supervised machine learning algorithms.
K-fold Cross Validation technique is been used with k = 10 is selected in the implementation since the accuracy obtained using these values were the best.

Dataset: No access link.

Results:

Model	Mean	STD
LR	96.667	10
SVM	97.78	6.67
DT	93.33	13.333
KNN	78.33	23.629

RP 18: Detection of coronavirus disease (COVID-19) from X-ray images using deep convolutional neural networks
[Base 3]

Year: 2021

Method: The database consists of three type CXR images for COVID-19 positive cases, Viral Pneumonia images and Normal CXR images. There are 1200
COVID-19 positive case, 1345 viral pneumonia and 1341 normal images in COVID-19 radiography database. Adam optimization algorithm was used for training deep neural networks. The Trained CNN model and the pre-trained MobileNet model were used to examine their performance using different classifiers.

Dataset: COVID-19 Radiography Database: https://www.kaggle.com/tawsifurrahman/covid19-radiography-database

Results:

Model	Class	ACC	REC	F1
CNN	COVID-19		99.79	99.69
	Pneumonia	96.90	99.44	95.53
	Normal		90.94	94.89
CNN-SVM	COVID-19		99.75	99.85
	Pneumonia	99.53	99.48	99.40
	Normal		99.41	99.39

RP 19: Convolutional Support Vector Models: Prediction of Coronavirus Disease Using Chest X-rays

Year: 2020

Method: Images used in this analysis came from anteroposterior radiographs (X-rays). After selecting the images, a hashing technique was used to detect duplicates and 30 duplicated images were removed, resulting in 217 COVID-19 detected X-rays. After removing duplicates, a grayscale/color conversion was applied to each of them through the OpenCV library in Python. This was the only pre-processing technique applied. These 217 X-ray were passed into the CNN. The features extracted were then passed into the SVM models.

Dataset: COVID Chest X-rays: https://github.com/ieee8023/covid-chestxray-dataset

Open Access Biomedical Image Search Engine: https://openi.nlm.nih.gov/faq#collection

Results:

Model	ACC	F1	MCC
CNN_1	96.67	96.63	93.48
CNN_2	96.73	96.67	93.74
SVM_{Lin}	80.79	80.21	61.98
SVM_{Pol}	77.90	77.24	56.30
SVM_{RBF}	83.45	83.86	67.39

RP 20: Detection of coronavirus Disease (COVID-19) based on Deep Features and Support Vector Machine

Year: 2020

Method: The dataset consists of 127 confirmed COVID-19, 127 confirmed pneumonia and 127 healthy X-ray images. i.e. in total 381. The deep features of 13 deep CNN models are extracted and used by SVM classifier for detection of COVID-19.

Finally, a comparative analysis of deep feature plus SVM and traditional image classification method (LBP+SVM, HOG+SVM and GLCM+SVM) is carried out.

Dataset: COVID Chest X-rays: https://github.com/ieee8023/covid-chestxray-dataset

X-ray and CT snapshots of CONVID-19 patients: https://www.kaggle.com/andrewmvd/convid19-x-rays

Results:

Model	SEN	FPR	F1
AlexNet	94.86	2.56	94.85
GoogleNet	91.73	4.13	91.74
InceptionV3	90.26	4.86	90.28
MobileNetV2	94.46	2.76	94.46
ResNet18	94.26	2.86	94.25
ShuffleNet	65.26	17.36	58.79
VGG16	94.20	2.90	94.20

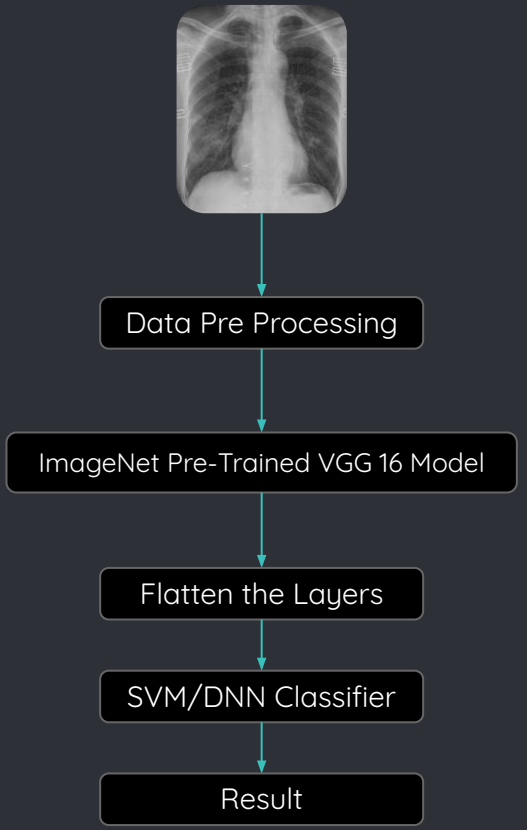
3

Proposed methodology

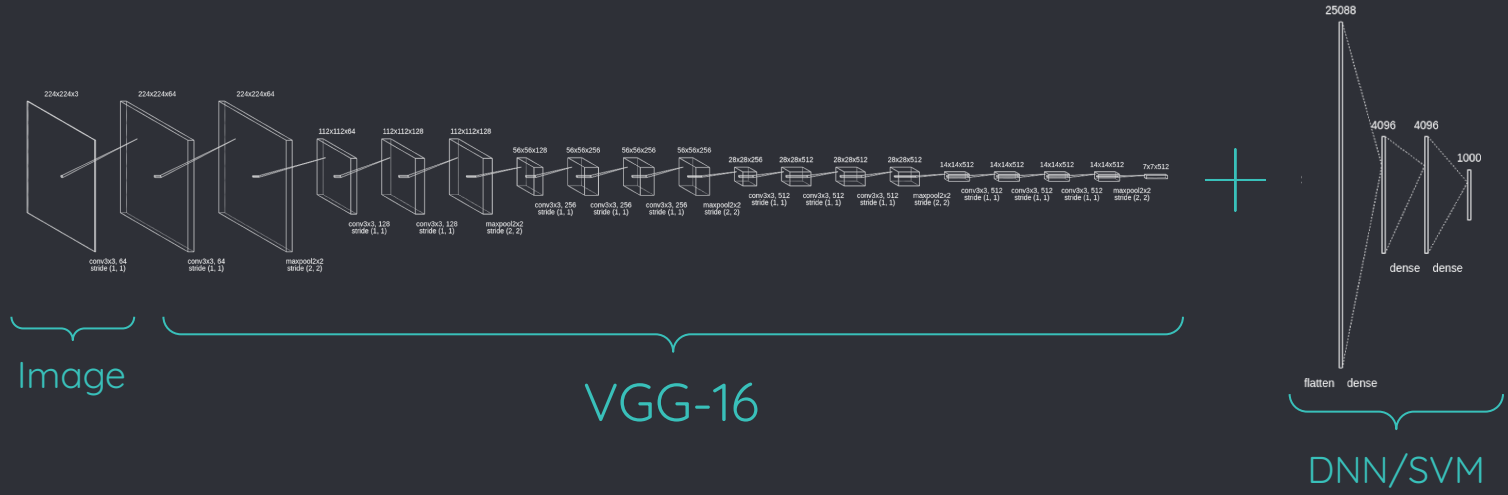
The Inner workings of our model

Flow Chart

- Image Preprocessing
- Dataset Split
- Feature Extraction
- Classifier



Visual Representation



Image

VGG-16

DNN/SVM

What is IMAGENET?

1. It is an image database organized according to the WordNet hierarchy, in which each node of the hierarchy is depicted by hundreds and thousands of images.
2. It consists of millions of images containing over 1000 classes and many thousand groups.
3. It is used to train various models over long periods of time, to attain superior accuracy.

What is VGG-16 ?

VGG16 is a convolutional neural network model proposed by K. Simonyan and A. Zisserman from the University of Oxford in the paper “Very Deep Convolutional Networks for Large-Scale Image Recognition”.

Why VGG-16 ?

It performs outstandingly on the ImageNet dataset. It achieves a 92.7% accuracy which is a top-5 test accuracy in ImageNet

4

Dataset

COVID - Data

Our primary dataset is obtained from a GitHub repository.

Class	No. of Images
Healthy	500
Pneumonia	500
COVID-19(+)	125

The second dataset is the Mendeley Dataset.

Class	No. of Images
Covid	878
Non Covid	912

5 Software & Technology

Software	Hardware
Numpy	Intel i7
Pandas	8 GB RAM
Pytorch	NVIDIA 1080-Ti
sklearn	
Keras	
Tensorflow	

Thank You.