Data Collection And Preprocessing Phase

Date	6 JULY 2024
Team ID	SWTID1720110768
Project Name	Covidvision: Advanced Covid-19 Detection From Lung X-rays With Deep Learning
Maximum Marks	6 Marks

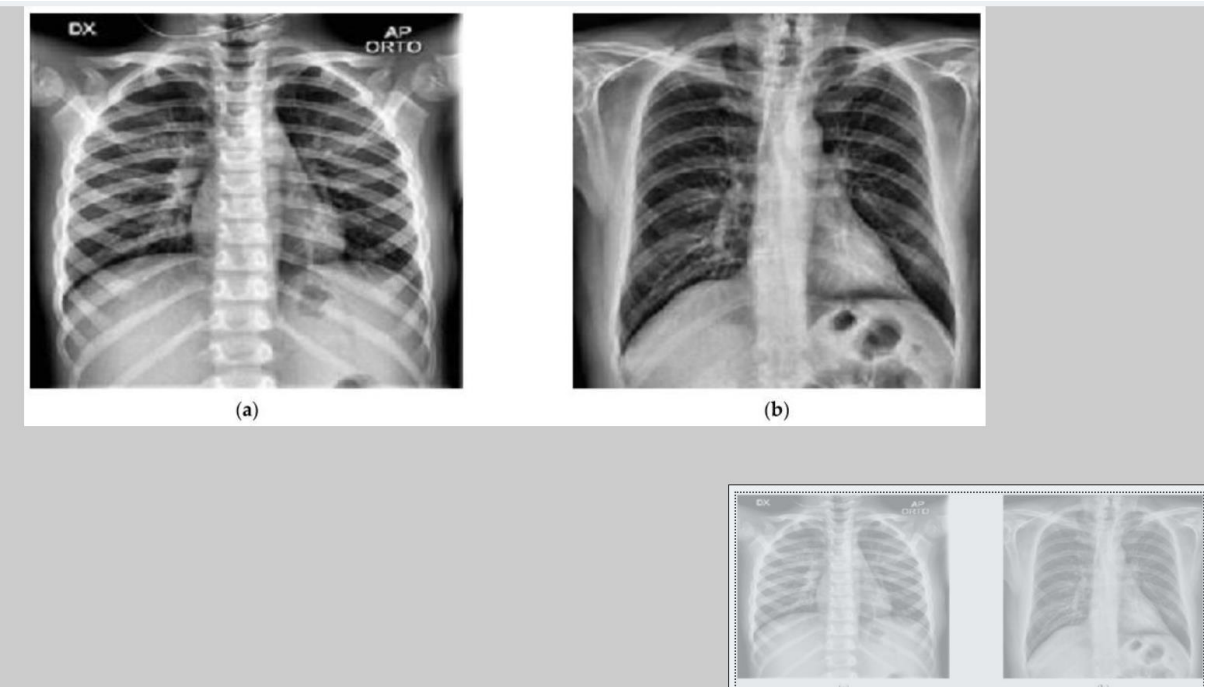
Data Exploration And Preprocessing Template:

Data Exploration and Preprocessing for CovidVision involves collecting and understanding X-ray images, cleaning and balancing the dataset, applying data augmentation, normalizing and resizing images, and finally splitting the data into training, validation, and test sets for effective deep learning model development.

Reference	Number of Image Samples	Classes	Architectures	Best Performing Architecture	Performance/Accuracy
[35]	1428	3	VGG19, MobileNetV2, Inception, Xception, InceptionResNetV2	MobileNetV2	Acc = 96.78%
[36]	204	2	VGG16 + Resnet50	VGG16 + Resnet50 + custom CNN	Acc = 89.2%
[37]	100	2	ResNet50, InceptionV3 and InceptionRes-NetV2	ResNet50	Acc = 98%
[38]	21,152	2	CNN	CNN	Acc = 94.64%
[39]	5184	2	ResNet18, ResNet50, SqueezeNet, DenseNet-121	SqueezeNet	Sensitivity = 98%, Specificity = 92.9%
[40]	13,975	2	COVID-CAPS	COVID-CAPS	Acc = 95.7%,
[41]	400	2	VGG16, InceptionResNetV2, ResNet50, DenseNet201, VGG19, MobilenetV2, NasNetMobile, and ResNet15V2	NasNetMobile	Acc = 93.94%
[42]	75	2	VGG19, Xception, ResNetV2, DenseNet201, InceptionV3, MobileNetV2, InceptionResNetV2	VGG19, DenseNet	F1 scores = 0.91
[43]	1127	2	Modified Darknet	Modified Darknet	Acc = 98%
[44]	1257	3	Xception	Xception	Acc = 94%
[45]	2356	3	ACoS system	ACoS	Acc = 91.33%
[46]	6100	3	SVM, LR, nB, DT, and kNN + VGG16, ResNet50, MobileNetV2, DenseNet121	Mean result	Acc = 98.5%

[47]	1428	2	VGG16	VGG16	Acc = 96%
[48]	79,500	3	Grad-CAM	Grad-CAM	Acc = 91.5%
[49]	6200	4	CSEN-based classifier	CSEN-based Classifier	Sensitivity = 98% Specificity = 95%
[50]	13,975	19	COVID-Net	COVID-Net	Acc = 93.3%
[51]	196	3	DeTrac	Detrac	Acc = 93.1%
[52]	3150	3	CapsNet	CapsNet	Acc = 97%
[53]	1127	3	Xception	Xception	Acc = 97%
[54]	7470	2	MD-Conv	MD-Conv	Acc = 93.4%
[55]	380	2	Novel CNN Model	Novel CNN Model	Acc = 91.6%
[56]	247	2	BMO-CRNN	BMO-CRNN	Sensivity = 97.01% Acc = 97.31% F-value = 97.53%

Dataset preparation:



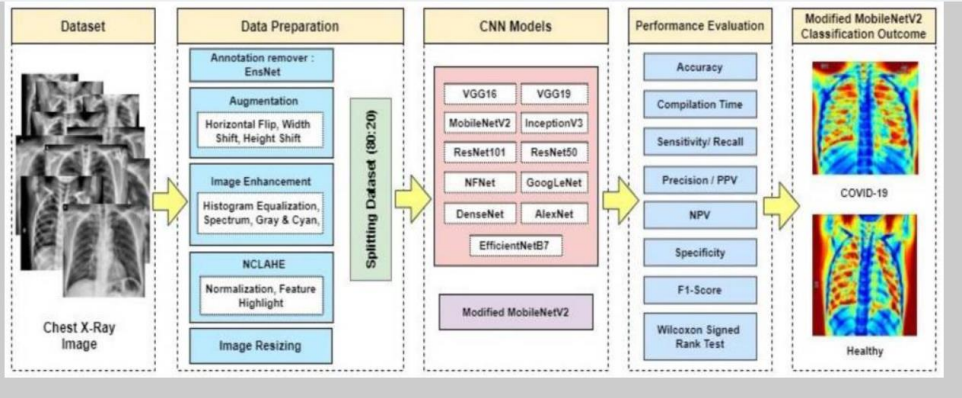
Chest X-ray image data samples. (**a**) Healthy; (**b**) COVID-19.

Parameters of data augmentation.

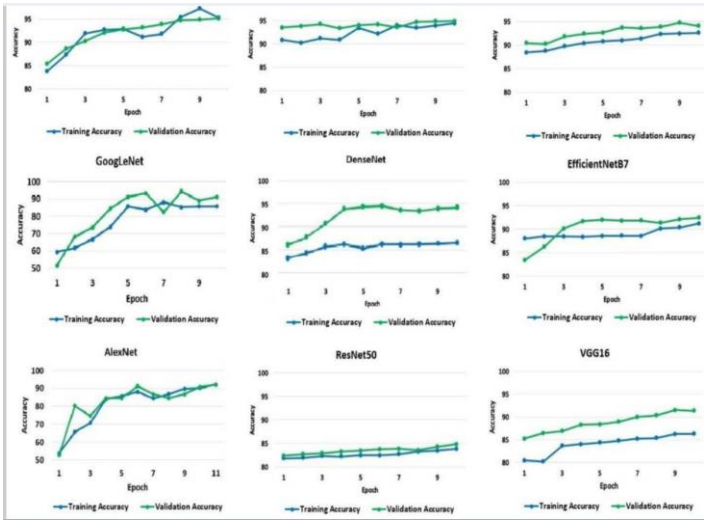
Augmentation Technique	Range
Horizontal flip	True
Rotation range	10
Width shift range	0.1
Height shift range	0.1
Vertical flip	False

Dataset Description.

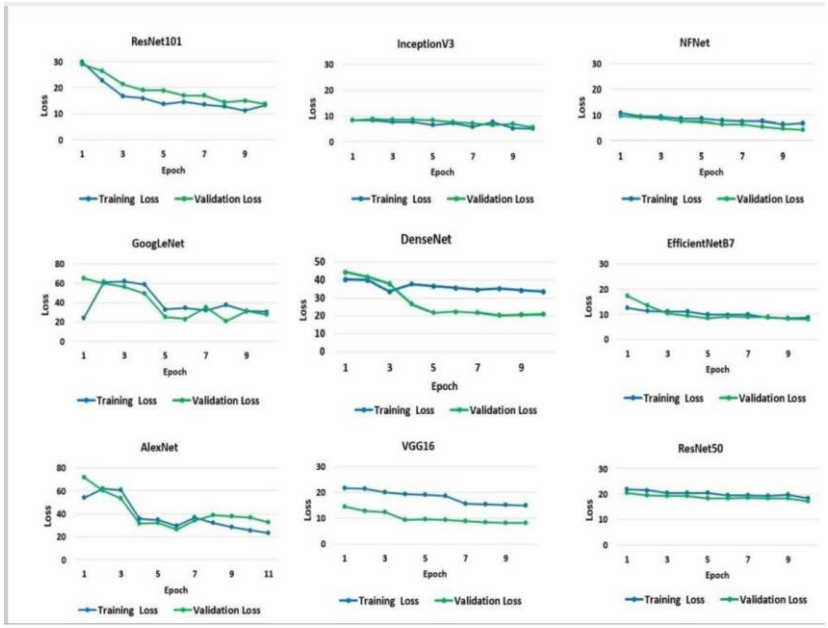
Features	Values
Total Number of Images	52,000
Disease Types	2
Dimension (Size in Pixel)	Classifier's Resolution (i.e., AlexNet is 256 × 256 pixels)
Color Grading	Grays, Cyan, Spectrum
COVID-19 Images	26,000 (After Augmentation)
Healthy Images	26,000 (After Augmentation)
Training Images	41,600
Testing Images	10,400



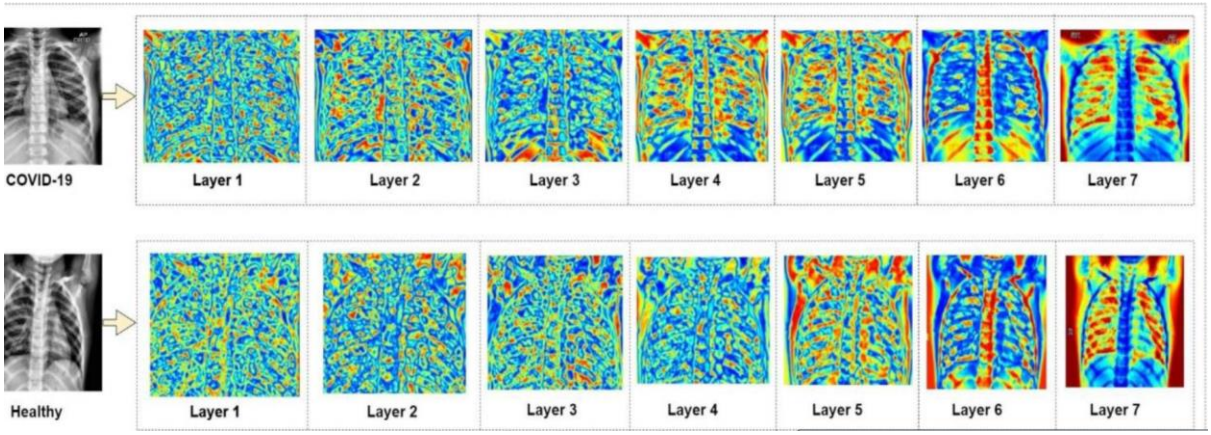
Proposed system diagram.



Training and validation accuracy of the models.



Loss in training and validation of the models.



Heat map of modified MobileNetV2 classification.