COVID-19 Detection using Chest X-RAY

[1]Jai Shankar K N, [1]Poornima G R, [2]Narayanappa C K
[1]Department of ECE, SVCE, Bengaluru, India
[2]Department of Medical Electronics Engineering, M S Ramaiah Institute of Technology
Bengaluru, India

Received: April 20, 2021. Revised: February 11, 2022. Accepted: March 4, 2022. Published: March 28, 2022.

Abstract— **In view of the COVID-19 pandemic, the exponential increase in the COVID-19 patients is leading to the enormous demand on the healthcare systems across the world. The allocation of resources towards the detection of the people affected by the virus plays a key role in curbing the pandemic and slowing down the spread of the virus to a greater extent. While traditional procedures are utilized to discover COVID-19 individuals, testing each individual with a limited number of testing kits is a massive undertaking. Most healthcare systems include X-ray equipment, and most of them being digitized, can be utilized as a way of screening for COVID-19 patients. This paper proposes AI model that can analyze and predict a possible COVID-19 patient, which can be used to prioritize the people for further testing. Further we propose the automation of this process where the models can be deployed in a remote server or an edge computing device where the X-ray images can be screened by the deep learning model to give predictions with very less turnaround time.**

Keywords—* *Chest X-ray, Convolutional Neural Networks, Covid-19, Deep Learning

I. Introduction

The rapid surge of COVID-19 patients has put a tremendous strain on healthcare systems throughout the world. COVID19 patients have already overburdened healthcare systems in several nations [1]. The limited number of doctors, hospital beds, personal protective equipment (PPE) and testing kits makes it a nearly impossible task to use only the conventional methods for the detection of a COVID-19 patient which need to be addressed and used to change the processes[2][3]. The detailed study of various factors and symptoms in the detection of COVID-19 patients helped in exploring various alternatives to the traditional detection methods[4][5]. These alternative resources are critical in improving detection processes and maximizing the utilization of today's healthcare infrastructure. In this paper we propose the use of X-ray machines as one such alternative resource for the detection of COVID-19 patients using Anterior to Posterior (AP) or Posterior to Anterior (PA) chest X-ray. Portable X-ray equipment are also available, allowing testing to take place wherever it is needed, minimizing the need for PPE kits, which is a valuable resource in this situation. The deep learning models proposed in this paper use the AP or PA chest x-ray of a person to classify the person as Normal, Covid or Viral Pneumonia patient. This method is much more cost effective, consumes less time when compared the conventional testing methods (RT-PCR), does not require any transportation facility and can be tested in bulk. Further automation of this process through deployment of the model in a remote server and accessing it through internet makes it easily available to the healthcare systems. The following is a breakdown of the structure of this paper: The overview of comparable studies in this field is discussed in section II. The dataset analysis and models proposed in our work are described in part III, followed by the model analysis and findings in section IV. Part V discusses the conclusions, while section VI discusses future study.

II. LITERATURE REVIEW

A multitude of methods for detecting COVID-19 patients using chest x-ray have been proposed. Rahul et al.[6] shows one such approach which uses the modified ResNet152 to perform the feature extraction followed by the synthetic minority oversampling technique (SMOTE) to balance the data points. The output features are then given as input to the ML classifiers to predict a COVID-19 patient.

Arpan et al.[7] propose a deep learning model which uses the CheXNet architecture as a backbone with 121-layer Dense Convolutional Network and modify the final classifier of 14 classes to 4 classes classifier, classifying the input as Normal, Bacterial Pneumonia, Viral Pneumonia and COVID-19 respectively. The deep convolutional layers are used in the CheXNet architecture to train.

Tulin et al.[8] use the DarkCovidNet architecture inspired by the DarkNet-16 model. This model consists of 17 convolutional layers and gives an accuracy of 87.02% on dataset developed by Cohen JP which contains about 127 x-ray images of COVID-19 patient.

III. METHODOLOGY

Finding a well-balanced dataset with a sufficient number of images for each of the dataset's categories is the most important stage in developing an effective deep learning model. The dataset is evaluated through dataset analysis. This process is followed by various image processing steps which involve normalizing the images and augmenting them. The second step is to create a deep learning model architecture and train it on the dataset which is a iterative process. Once the deep learning model is trained, the learning curves of the model are analyzed and the model is refined and re-trained until the best model is obtained. Experimental analysis is a part of this iterative process and helps in fine tuning the deep learning model to increase its performance. These steps are explained in detail below.

A. *DATASET ANALYSIS*

The dataset used for training the deep learning models is a combination of two datasets obtained from COVID19, Pneumonia and Normal Chest X-ray PA Dataset (Dataset A) [9] and a Kaggle dataset created by Chowdhury et al (Dataset B)[10][11]. The Dataset A and Dataset B consists of a total of 4,575 and 15,153 images respectively corresponding to 3 categories namely, Covid-19, Normal and Viral Pneumonia. The composition of number of images in Dataset A, Dataset B and the combined dataset is shown in TABLE I, TABLE II and TABLE III respectively

Dataset A	
Categories	***Number of images***
Covid-19	1525
Normal	1525
Viral Pneumonia	1525
Total	**4575**

TABLE I. Composition of the number of images in Dataset A

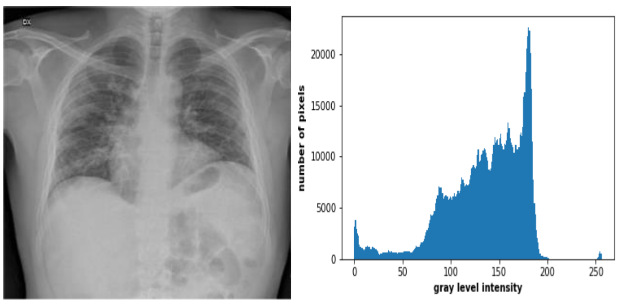
Dataset B	
Categories	***Number of images***
Covid-19	3616
Normal	10192
Viral Pneumonia	1345
Total	**15153**

TABLE II. Composition of the number of images in Dataset B

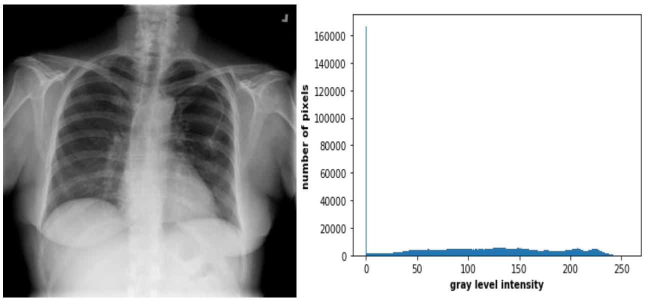
Combined Dataset				
Dataset	***Covid-19***	***Normal***	***Viral Pneumonia***	***Total***
Training	2009	2009	2009	**6027**
Validation	574	574	574	**1722**
Test	287	287	287	**861**
Total	**2870**	**2870**	**2870**	**8610**

TABLE III. Composition of the number of images in Combined Dataset

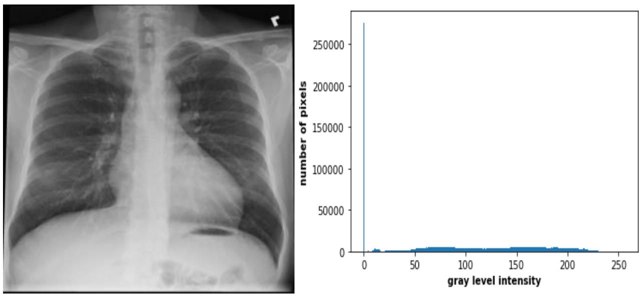
The combined dataset consists of randomly selected images from Dataset A and Dataset B. Each category contained equal number of images to prevent class imbalance.



(a)



(b)



(c)

Figure 1. Sample images of the chest x-ray and histograms of (a) Covid-19 person, (b) Normal person and (c) viral pneumonia person.

The whole combined dataset was split into 70% of training data, 20% of validation data and 10% of test data.

The histograms of few sample images belonging to the 3 categories were analyzed and found to be different for images belonging to different categories which is shown in Figure 1.

The covid-19 pneumonia condition affects the lungs by filling the air sacs with fluid. This hinders the ability of the lungs to take oxygen, thus causing serious breathing problems to the patient. As shown in Figure 1 (a), the lungs of the patient are filled with fluid which is denser than air, hence the gray level intensity of the x-ray image is high in covid-19 patient. Whereas in the case of Figure 1 (b) and (c), the gray level intensities of the x-ray images are considerably low. Further, the analysis of the images and their histograms shows that it is possible to classify the person as normal, covid-19 positive or affected by viral pneumonia using the AP or PA chest x-ray images of the person.

B. *MODEL TRAINING*

The Combined Dataset was used to train 8 different Convolutional Neural Networks (CNN) models. Seven of these models were trained along with transfer learning from the pre-trained models. Transfer learning is a machine learning approach in which pre-trained models that have been trained on one task are re-trained on another task by freezing the weights of some layers and re-training the other layers according to the task's requirements. The pre-trained models used in this paper are VGG16, VGG19[12] , Densenet121[13], Mobilenet[14], Resnet50[15], Inceptionv3[16], Xception[17] models. These models were trained on the ImageNet dataset which is a benchmark dataset for image classification and object detection. The numbers in the names of the models signify the number of layers in the models. For example, VGG16 model contains 16 layers.

layers in the models. For example, VGG16 model contains 16 layers.

The images were loaded after pre-processing the images as per the pre-processing method proposed by each of the pre-trained models. The input size of the image was 224x224x3 for VGG16, VGG19, Densenet121, Mobilenet and Resnet50 and 299x299x3 for InceptionV3 and Xception models. The output layer of each model was replaced with 3 neurons with softmax activation function to classify the image into 3 categories of Covid-19, Normal and Viral Pneumonia with class indices 0, 1 and 2 respectively.

Each pre-trained model was optimized to train well on the dataset by re-training few layers of the model which was hyperparameter and was tuned by trial-and-error basis. Re-training the last 10 layers was found to be optimum for VGG16, VGG19, Densenet121 and Mobilenet models while the last 20 layers were re-trained for ResNet50, InceptionV3 and Xception models.

A CNN model architecture called Cowinet was also designed and trained on the dataset. The image pre-processing proposed for Densenet121 was used to pre-process the images before training the model on the dataset. All the images were resized to resolution of 256x256 containing 3 channels. The model consists of an input layer which receives input of size 256x256x3 followed by 3 CBCP layers. Each CBCP layer starts with a Convolutional layer with a 3x3 kernel size, then a Batchnormalization layer, another Convolutional layer with a 3x3 kernel size, an activation function called ReLU (Rectified Linear Units), and finally a Maxpooling layer with a 2x2 kernel size. The number of neurons for each of the CBCP layers were 32,64 and 128 respectively. The outputs from these layers were passed to the flatten layer. The input from the flatten layer is passed to the output layer through the two fully connected layers. The fully connected layer consists of 1024 neurons and activation function of ReLU and the output consisted of 3 neurons with the activation function of softmax.

C. *EXPERIMENTAL ANALYSIS*

All the Deep Learning models were trained on the hardware containing intel core i7 8th Gen processor with the hardware accelerator NVIDIA GeForce MX150 Graphics card. A virtual environment was created for training the Deep Learning models which was accelerated with the Computer Unified Device Architecture (CUDA) and NVIDIA CUDA Deep Neural Network (cuDNN) libraries.

The Deep Learning models were compiled using the Adaptive Moment Estimation (Adam) optimizer with a learning rate of 0.0001. The loss function considered for training the deep learning models is Categorical Crossentropy which is most suitable for training classification models when the output of the models are categories. Certain generalization methods like EarlyStopping and Modelcheckpoint from Keras were used to prevent the models from overfitting on the training dataset EarlyStopping functions stops the training of the model once model stops improving for a certain number of iterations. Modelcheckpoint function saves the model which performs the best on the validation dataset with loss function as the parameter for evaluation. Learning rate is one the hyperparameters of the optimizer which cannot be tuned easily. ReduceLROnPlateau function was used to reduce the learning rate of the model when the model did not improve for 4 epochs. The dynamic learning rate played a vital role in optimizing the model's performance. The Adam optimizer was found to be better at finding solution that was closer to the global minima of the model in comparison with the Stochastic Gradient Descent (SGD) optimizer.

IV. RESULTS

The trained models were tested on the test dataset for evaluation. The models' performance was evaluated using accuracy, precision, recall (Sensitivity), F1 score as the metrics.

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Precision = TP / (TP + FP)

Recall = TP / (TP + FN)

F1 score = 2 * (Precision * Recall) / (Precision + Recall)

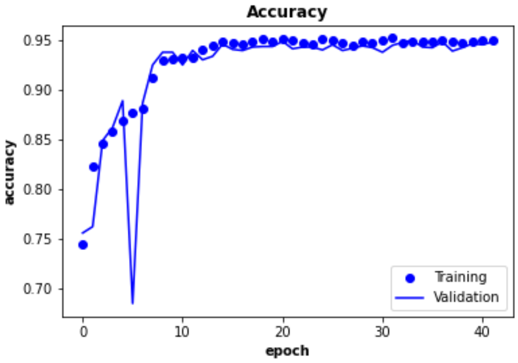
Where TP, TN, FP and FN refer to the True Positive, True Negative, False Positive and False Negative respectively.

The model performance of the respective models on the test dataset is shown in TABLE IV.

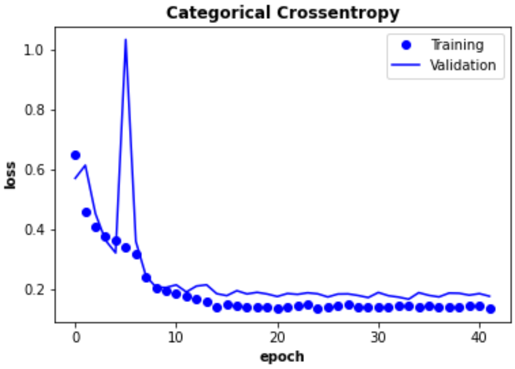
TABLE IV. Model's Performance on the test dataset

Deep Learning Models	**Metrics (Percentage)**			
	Accuracy	***Precision***	***Recall (Sensitivity)***	***F1-score***
VGG16	95.35	95.67	95.33	95.67
VGG19	94.66	95.00	94.67	95.00
Densenet121	96.16	96.33	96.33	96.33
Mobilenet	95.81	95.67	95.67	95.67
Resnet50	96.16	96.33	96.33	96.33
InceptionV3	96.05	96.33	96.00	96.33
Xception	87.10	88.67	87.00	86.67
Cowinet	92.92	93.00	93.00	92.67

Densenet121 and Resnet50 outperforms the rest of the models. Even while pre-trained models provide substantial results on the test dataset, they tend to overfit on the training dataset, implying that the model does not work consistently and requires a lot of computing resources due to the large number of layers. On the other hand, the Cowinet model shows consistency in the performance which is its performance on the training, validation and test dataset.



(a)

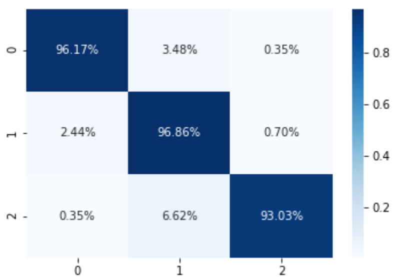


(b)

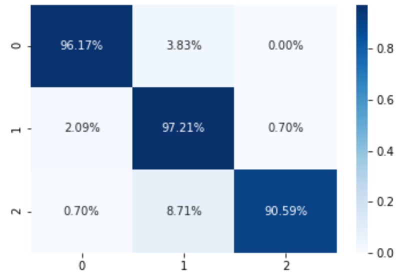
Figure 2. (a) accuracy and (b) loss of Cowinet model on Training and Validation dataset

The consistency in the performance of the Cowinet model can be found by analyzing the learning curves of the Cowinet model shown in Figure 2. Since the training versus validation loss are very close to each other, we can conclude that the Cowinet model is a good fit model which makes it more reliable for predictions.

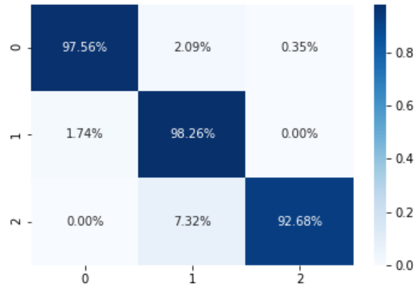
Further the confusion matrix of the individual models is shown in Figure 3. The analysis of confusion matrices of all the models shows that the covid-19 patients are classified with very good accuracy from the normal or viral pneumonia patients. Comparatively, the model shows lesser accuracy in classifying normal person versus viral pneumonia patients when compared to classifying covid-19 patients versus normal patients or covid-19 patients versus viral pneumonia. Overall, the CNN models successfully classify a covid-19 patient from a normal or viral pneumonia patient with very good accuracy.



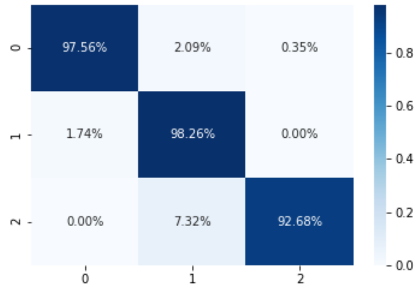
(a)



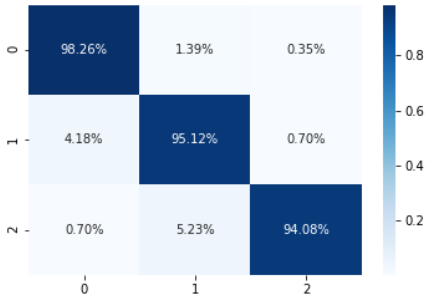
(b)



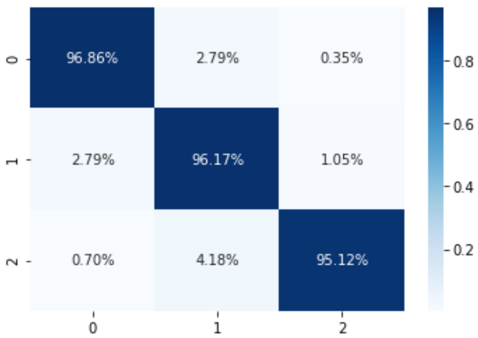
(c)



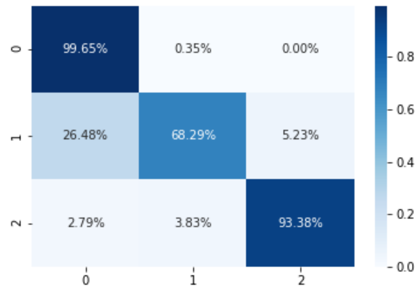
(d)



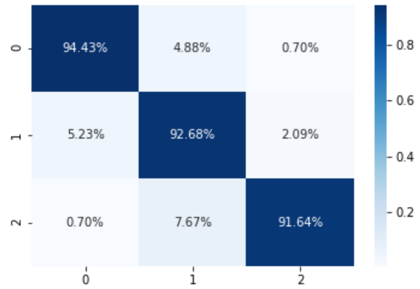
(e)



(f)



(g)



(h)

Figure 3. Confusion matrix of (a) VGG16, (b) VGG19, (c) Densenet121, (d) Mobilenet, (e) Resnet50, (f) InceptionV3, (g) Xception and (h) Cowinet models

The Figure 3 (a), (b), (c), (d), (e), (f), (g) and (h) shows the confusion matrix obtained for the VGG16 algorithm, VGG19, Densenet121, Mobilenet, Resnet50, InceptionV3, Xception and Cowinet models respectively. The class in the confusion matrix are the normal patient, viral pneumonia patient and the Covid Patient. The average of the diagonal elements of the matrix results the accuracy of the algorithm. Similarly the other parameters are obtained using the elements of the confusion matrix. The horizontal values are the actual values, the vertical parameters are the predicted values. the prediction is the ratio of the correct predictions to total predictions. The average of each class will give the average prediction of the algorithm. Similarly, the recall is calculated for the considered algorithm. Recall is defined as the ratio of the correct classification to the actual classification. From the matrix, it is clear that the Figure 3 (h) shows that accuracy of the Cowinet Models has the best accuracy amongst all the algorithm. The Cowinet deep learning algorithms will give an accuracy of 92.92%. 93% come recall of 93% and F1 score of 92.67%. When compared to other models, the Cowinet algorithm gives a consistent result in all the 4 parameters. Among all the confusion matrix the figure 3 (g) shows Xception algorithm which gives the least performance. As the blueness of the colour increases, the correlation increases as the blueness of the colour decreases, the correlation decreases.3

V. CONCLUSION

The detection of Covid-19 patients is critical to controlling the Covid-19 pandemic. The recent advancements in the technology have enabled us to find a faster and more effective way of detecting the Covid-19 patients at a larger scale and in a more streamlined manner. This paper proposes one such solution of using the x-ray machines as an alternate resource to maximize the utilization of the healthcare systems and prioritize the people for the conventional testing process. The Deep Learning models presented in this paper classify the images with very less turnaround time. Further we propose the automation of this process through the deployment of the Deep Learning models in a cloud-based system which can be accessed by any system around the world or in an edge computing device which is portable to increase the efficiency of testing, while reducing the turnaround time to a greater extent.

VI. FUTURE WORKS

The models' performance is limited by its architecture and the dataset. We want to further train the models on a much larger and more diverse dataset which addresses a larger variety of categories. The models will then be trained on this dataset with more fine tuning to its architecture along with usage of various hyper parameter tuning methods to obtain the optimal performance from each model.

REFERENCES.

[1] IBISWorld. Effects of COVID-19 on Global Healthcare Systems. Apr 16 2020. Available: https://www.ibisworld.com/industry-insider/coronavirus-insights/effects-of-covid-19-on-global-healthcare-systems

[2] Maia Dorsett. Point of no return: COVID-19 and the U.S. healthcare system: An emergency physician's perspective, 26 June 2020. Available: https://advances.sciencemag.org/content/6/26/eabc5354

[3] Kariman Ramzy El Helow, Abdel-Badeeh M. Salem, "Are Artificial Intelligence (AI) and Machine Learning (ML) Having an Effective Role in Helping Humanity Address the New Coronavirus Pandemic?", WSEAS Transactions on Biology and Biomedicine, ISSN / E-ISSN: 1109-9518 / 2224-2902, Volume 17, 2020, Art. #14, pp. 119-124

[4] Huang C., Wang Y. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497–506.

[5] Kong W., Agarwal P.P. Chest imaging appearance of COVID-19 infection. Radiology: Cardiothoracic Imaging. 2020;2(1)

[6] Rahul Kumar, Ridhi Arora, Vipul Bansal, Vinodh J Sahayasheela, Himanshu Buckchash , Javed Imran , Narayanan Narayanan, Ganesh N Pandian, and Balasubramanian Raman. Accurate Prediction of COVID-19 using Chest X-Ray Images through Deep Feature Learning model with SMOTE and Machine Learning Classifiers, 17 April 2020

[7] Arpan Mangal, Surya Kalia, Harish Rajgopal, Krithika Rangarajan, Vinay Namboodiri, Subhashis Banerjee, and Chetan Arora. CovidAID: COVID-19 Detection Using Chest X-Ray, 29 May 2020

[8] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Balgolu, Ozal Yildirim, and U. Rajendra Acharya. Automated detection of COVID-19 cases using deep neural networks with X-ray images, 28 April 2020

[9] Asraf, Amanullah; Islam, Zabirul (2021), "COVID19, Pneumonia and Normal Chest X-ray PA Dataset", Mendeley Data, V1, doi: 10.17632/jctsfj2sfn.1

[10] M.E.H. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M.A. Kadir, Z.B. Mahbub, K.R. Islam, M.S. Khan, A. Iqbal, N. Al-Emadi, M.B.I. Reaz, M. T. Islam, "Can AI help in screening Viral and COVID-19 pneumonia?" IEEE Access, Vol. 8, 2020, pp. 132665 - 132676.

[11] Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S.B.A., Islam, M.T., Maadeed, S.A., Zughaier, S.M., Khan, M.S. and Chowdhury, M.E., 2020.

Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images.

[12] Sanjay H S, Basavraj V Hiremath, Prithvi B S, P A Dinesh, “Machine Learning Based Assessment of Auditory Threshold Perception in Human Beings”, S N Applied sciences, Vol 2(142), 1-10, (2020) (DOI: 10.1007/s42452-019-1929-7)

[13] Gao Huang and Zhuang Liu and Laurens van der Maaten and Kilian Q. Weinberger, “Densely Connected Convolutional Networks”, arXiv:1608.06993v5 [cs.CV], 28 January 2018

[14] Andrew G. Howard and Menglong Zhu and Bo Chen and Dmitry Kalenichenko and Weijun Wang and Tobias Weyand and Marco Andreetto and Hartwig Adam, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications”, arXiv:1704.04861v1 [cs.CV], 17 Apr 2017.

[15] Kaiming He, Xiangyu Zhang, Shaoqing Ren and Jian Sun, “Deep Residual Learning for Image Recognition”, arXiv:1512.03385v1 [cs.CV], 10 Dec 2015

[16] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens and Zbigniew Wojna, “Rethinking the Inception Architecture for Computer Vision”, arXiv:1512.00567v3 [cs.CV], 11 Dec 2015

[17] François Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions”, arXiv:1610.02357v3 [cs.CV], 4 Apr 2017