An Ensemble Deep Learning System for the Automatic Detection of COVID-19 in X-Ray Images.

Beltus Nkwawir Wiysobunri
Instanbul Technical University
nkwawir18@itu.edu.tr

Hamza Salih Erden
Istanbul Technical University
erdenh@itu.edu.tr

Behcet Ugur Toreyin
Instanbul Technical University
toreyin@itu.edu.tr

Abstract

At the dusk of the year 2019 and the dawn of the new year, 2020, the world was awakened with the birth of a novel virus called the coronavirus nicknamed COVID-19. Originating from a small city in China, this virus swiftly sprawled across the four corners of the world wreaking havoc on a global scale and costing millions of human lives. The early detection, quarantine, and treatment of positive cases in order to prevent further spread of the virus, and possible eradication is the world's number one objective.

Modern research based on radiology imaging suggests that chest X-ray images contain some important information about the COVID-19 virus. With the existence of sophisticated artificial intelligence(AI) algorithms especially in deep learning, this presents an opportunity for the development of advanced AI-based applications in association with radiological imaging for the detection of COVID-19.

In this study, we present an ensemble deep learning system based on Max (Majority) voting scheme of 5 classical deep Convolutional Neural Network(CNN) architectures: ResNet50, ResNet34, VGG-19, MobileNet-V2, and DenseNet201 for the automatic detection of the COVID-19 disease using chest X-Ray images. Our ensemble system produced a remarkable 99% performance accuracy. This system is highly robust with a significant decrease in misclassification errors and an increase in both reliability and confidence. The ensemble deep learning system can be used as an automatic diagnostic tool to provide assistance not only to expert radiologists but can be deployed in remote areas where the availability of such experts is minimal. The system can play a significant role in the world especially in areas that are plagued by the deadly COVID-19 virus and have no expert radiologists. This system is highly versatile as it can also be adapted to detect other chest-related diseases such as pneumonia and tuberculosis. The code used for the project implementation can be found using the link below https://github.com/Beltus/COVID-19-ENSEMBLE-MODEL-

1. Introduction

Between the dusk of the year 2019 and the dawn of the new year, 2020, the world was awakened with the birth of a new virus, COVID-19. The first traces of this virus appeared in Wuhan, Hubei Province in China and rapidly sprawled across the five continents of the world[1]. Its devastating blow has inflicted great harm and left the world with an everlasting scar. Globally, as of September, 29th 2020, there has been a total of 33.206.004 confirmed cases of COVID-19, including 999.239 deaths reported by the World Health Organization(WHO)[2].

The virus that causes the COVID-19 pandemic disease was officially designated as severe acute respiratory syndrome coronavirus 2, SARS-CoV-2[3]. Coronaviruses have been identified in both avian hosts and various mammals, including bats, camels, dogs, and masked palm civets. However, out of the seven types of coronaviruses, only SARS-CoV(20022003), Middle East respiratory syndrome (2012), and SARS-CoV-2(2019) can inflict severe respiratory disease and possible death in humans [4]. COVID-19 rapidly proved to have human-to-human transmission spreading through the world at a fast pace, jeopardizing the modern world[5] The typical clinical features of COVID-19 include fever, cough, sore throat, headache, fatigue, muscle pain, and shortness of breath [6].

1.1. Literature Review on Existing Detection Methods

Up to date, the main screening method used as a reference standard for detecting positive COVID-19 cases is reverse transcriptase-polymerase chain reaction (RT-PCR) testing that can detect SARS-CoV-2 RNA from respiratory specimens[7]. The RT-PCR test for COVID-19 is known to have very high specificity and low sensitivity in the range of 60% to 70%. This means that excluding a diagnosis of COVID-19 requires multiple negative tests, with test kits running in short supply[8].

Chest radiographic techniques such as Chest Tomography(CT) and X-Ray imaging have been used for early clinical diagnosis and detection of COVID-19 Pneumonia, hence playing a crucial role in the early prevention and control of COVID-19 [9]. In [10], the authors suggest the combination of CT imaging with clinical and laboratory findings as a possible way of facilitating the diagnosis of COVID-19 pneumonia. This is because the virus manifests by showing abnormalities in chest CT images, even in asymptomatic patients.

With the rapid parallel development of sophisticated artificial intelligence algorithms and cheap efficient computational systems, machine learning has gain popularity in the field of clinical research. Clinicians are now leveraging these tools to foster research in medicine. Due to the availability of data, deep learning; a sub-field of machine learning has found favor in the eyes of medical specialists[11][12] [13]. This is because it has the capabilities of hierarchical learning, automatic feature extraction, multi-tasking, and weight sharing [14]. In addition, deep learning models can learn representations from grid-like data and research has shown a significant performance boost in diverse Machine Learning and Computer vision tasks [15].

Currently, some of the deep learning applications in the medical science field include the classification of skin cancer [16][17], detection of breast cancer[18], brain disease detection using MRI images[19], Chest X-Ray lung segmentation[20], Tuberculosis screening[21] and many others. The rampant spread of the Covid-19 disease to every angle of the world has let to an increasing shortage of not only test kits but also expert radiologists. Hence, it's super important to come up with AI-based advanced tools to assist radiologists to automatically diagnose COVID-19 in X-ray images. These AI models will go a long way to facilitate the limited number of radiologists in the fast and accurate diagnosis of patients with the virus. Also, the time between testing and results confirmation can be tremendously cut down. In addition to this, more time can be spent by these experts in treating the patients than critically analyzing X-Ray images.

Since the outbreak of the deadly virus, scientists all over the world have been developing deep learning models using the available X-Ray images from multiple sources to identify patients with COVD-19. In [22], the authors implemented 7 classical deep learning models and then proposed a COVIDX-Net model architecture that consists of these CNN models in order to diagnose COVID-19 in X-ray images. In [23]. the authors developed a network architecture called COVID-Net tailored for the detection of COVID-19 cases from CXR images which gave a decent 91.0% detection accuracy of COVID-19 cases. The authors of [24] implemented several standard deep learning network architectures using 224 confirmed COVID-19 images. Their results show that VGG19 and MobileNet models achieve the best accuracy. In [25], the authors implemented three deep learning models using chest X-ray images and the performance results showed that the ResNet50 pre-trained model yielded the highest accuracy of 98%. Also, in [26] X-rays images were used by the authors to detect COVID-19 cases by extracting features using a convolutional neural network (CNN) model with Support Vector Machine (SVM) classifier. The research paper concludes that the ResNet50 model with the SVM classifier provided the best performance. The above outlined deep learning approaches give a glimpse of how deep learning is actively being used in the field of medicine for the fight against the killer virus, however, it's not an exclusive list. There is research pouring out everyday from various research institutes than can be kept track of. We encourage the reader to do further research if need be.

1.2. Hypothesis and Paper Contributions

In the literature of this study, the deep learning approaches that were implemented focused on 2 main techniques. The authors either decided to implement a series of CNN models, train the models and then select the best based on some performance metric as is the case in [24], or they modified the already existing classical deep CNNs to spring to life a new variant which was then used for training, as the authors in [27] did, in order to detect COVID-19.

The performance accuracies of these models are quite impressive. However, the dependence on a single model for the prediction of positive cases can be limiting. This is because each model extract features from training examples differently. That is, an image that is misclassified by one model can be correctly classified by another model and vice-versa. In addition to this, given the fact that we are dealing with human lives and a global pandemic, sensitivity is of critical importance. An error in such a diagnosis can be catastrophic beyond measure. This is especially true in the case of a false negative.

Hence we propose an Ensemble deep learning approach that leverages 5 popular state-of-the-art classical deep CNN models in order to automatically detect COVID-19 in chest X-Ray images. Ensemble learning algorithms are based on the construction of a set of classifiers that classify new data points by taking a (weighted) vote of their predictions. Our objective is to improve the prediction accuracy of COVID-19, decrease error-rate of misclassification and increase robustness of such a system by harnessing the strengths of the already existing sophisticated deep learning models. Therefore, an Ensemble Deep Learning system for the detection and classification of COVID-19 using X-Ray images became the ultimate choice.

2. Proposed Method

In machine learning, ensemble learning methods are meta-algorithms that combine several classifiers into one predictive model and then classify new data points by taking a (weighted) vote of their predictions. Normally, ensemble learning methods are used to decrease variance(bagging), bias(boosting), or improve predictions(stacking).

In our approach, we propose an ensemble deep learning method that will help to improve deep learning prediction accuracies of COVID-19 and decrease the error-rate of misclassification by combining 5 different models. These models include: VGG19, ResNet50, ResNet34, MobilNetV2, and DenseNet201.This approach allows the production of a better predictive performance model compared to a single model.

2.1. Theoretical Background of Proposed Method

At the core of our approach are two main concepts: Transfer Learning, also known as fine-tuning and Ensemble learning. By combining these powerful techniques we were able to build a robust model for the detection of COVID-19 in X-Ray images.

2.1.1. Ensemble Learning

Ensemble learning is a machine learning paradigm where multiple learners are trained to solve the same problem. In contrast to ordinary machine learning approaches that try to learn one hypothesis from training data, ensemble methods try to construct a set of hypotheses and combine them to use[28]. An ensemble is made up of a number of learners that are usually called base learners. In our approach, these base learners are VGG19, ResNet50, ResNet34, DenseNet201, and MobileNetV2. These base learners are usually based on a combination scheme. The most popular of these schemes include Majority or Max voting, Averaging, and Weighted averaging.

Max Voting: In this approach, multiple models are used to make predictions for each data point. The predictions by each model are considered as a 'vote'. The predictions which we get from the majority of the models are used as the final prediction. Max voting is generally used for classification tasks.

Averaging: In this technique, multiple predictions are made for each data point. An average of the predictions from all the models is then computed and used to make the final prediction. Averaging is typically used in regression tasks.

Weighted Average: This is an extension of the averaging scheme. In this technique, each model is assigned different weights defining its importance for prediction. The fact that our problem is a typical classification task, we decided to implement the max voting ensemble learning approach in this paper.

Figure 1. shows the schematic of the ensemble deep learning approach implemented in this study.

2.1.2. Transfer Learning

Transfer learning is the improvement of learning a new task through the transfer of knowledge from a related task that has already been learned[29]. In this project, transfer learning was achieved by leveraging the pretrained weights of 5 deep CNNs: VGG19, ResNet50, ResNet34, MobileNetV2, and DenseNet201. These are standard model architectures that have been trained with millions of data points (images) and hence act as a perfect base to layer up the X-Rays data set images for better performance accuracy.

a) VGG-19 Model:

VGG19 is a variant of the original VGG-16 convolutional neural network architecture proposed the Zisserman et al in [30]. In VGG-16 architecture we have a CNN that is 16 layers deep. However, in the VGG-19 architecture that's used in this paper, the model is made up of 19 layers deep. We loaded the pretrained version of the network trained on more than a million images from the ImageNet database[31]. The pretrained network can classify images into 1000 object categories, including keyboard, mouse, pencil, and many animals. Hence, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224 by 224. This network is a pretty large network that has about 143 million trainable parameters.

b) ResNet Model: Residual Networks(ResNets) is a classic neural network introduced by Kaiming He et al (2015) in [32]. ResNet-50 and ResNet34 which are later versions are deep CNNs that are 50 and 34 layers deep respectively. The ResNet-50 model 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 ResNet50 and ResNet34 have over 20 million trainable parameters. In this paper, we used the pretrained version of the network trained on more than a million images from the ImageNet database. The pretrained network can classify images into 1000 object categories, such as a keyboard, mouse, pencil, and many animals. However, it should be noted that, in our case, we're only interested in classifying COVID and non-COVID cases.

c) MobilNetV2: The MobileNet is a CNN that was proposed in [33] by Andrew G. Howard et al of Google Research team. This model was designed to effectively maximize accuracy while being mindful of the restricted resources for an on-device or embedded application. Mo-

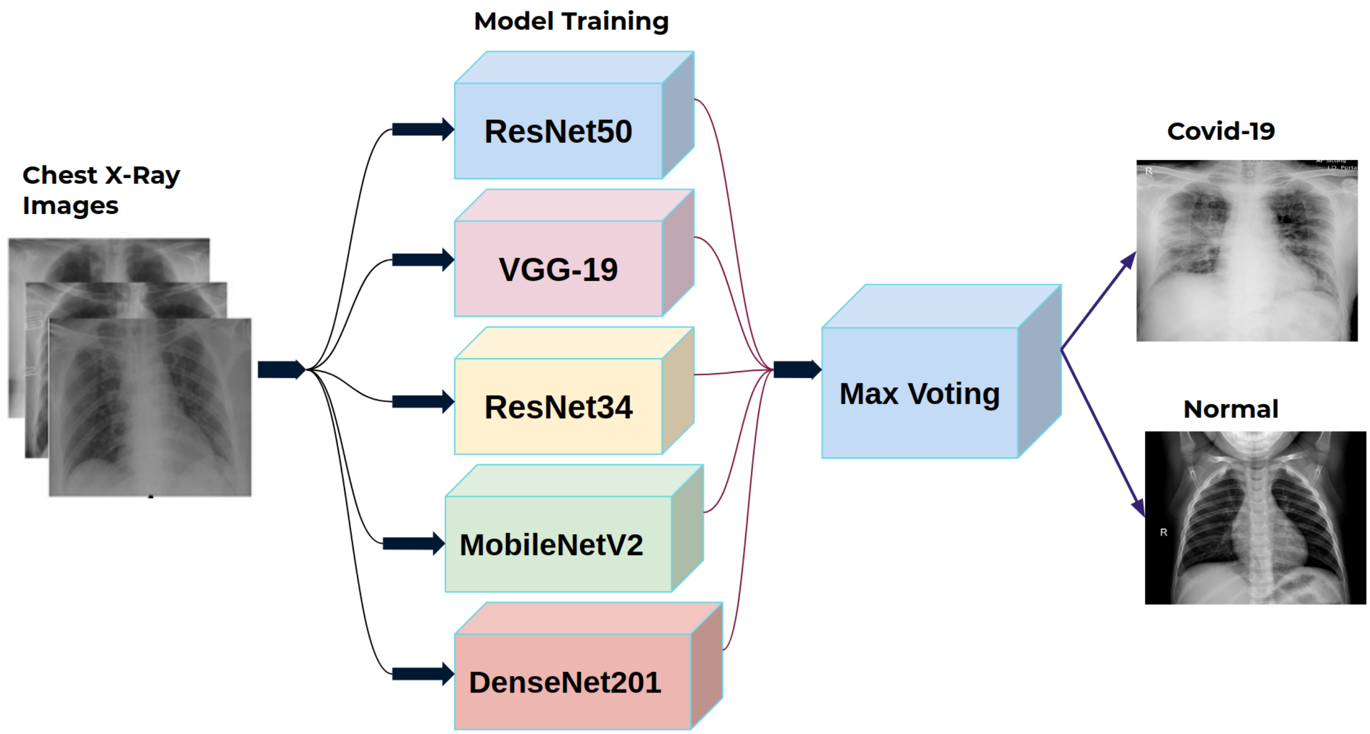

Figure 1. Ensemble Deep Learning Model Schematic with 5 Models

bileNets are small, low-latency, low-power models parameterized to meet the resource constraints of a variety of use cases. In MobileNetV2, there are two types of blocks. One is a residual block with a stride of 1. Another one is a block with a stride of 2 for downsizing. There are 3 layers for both types of blocks. The first layer is 11 convolution with ReLU6. The second layer is the depthwise convolution. The third layer is another 11 convolution but without any non-linearity. It is claimed that if ReLU is used again, the deep networks only have the power of a linear classifier on the non-zero volume part of the output domain.

d) DenseNet201: DenseNet(Densely Connected Convolutional Networks) is a deep CNN that was proposed by Gao Huang et al-2017[34]. DenseNet201 is a CNN that is 201 layers deep. In DenseNet each layer connects to every other layer in a feed-forward fashion, meaning that DenseNet has n(n+1)/2 connections in total. For each layer, the feature maps of all preceding layers are used as inputs, and its own feature-maps are used as inputs to all subsequent layers. DenseNet201 has 20 million trainable parameters.

2.2. Dataset of Chest X-Ray Images

In this research study, we used X-ray images obtained from two different sources for the diagnosis of COVID-19. The first source is from the COVID-19 Radiography Database created by a team of researchers from Qatar University, Doha, Qatar, and the University of Dhaka, Bangladesh along with their collaborators from Pakistan and Malaysia in collaboration with medical doctors. This database is constantly updated. In the latest release used in this study, there are 219 COVID-19 positive chest X-ray images, 1341 normal X-ray images, and 1345 viral pneumonia images. We discarded the pneumonia class images. This database is hosted on the Kaggle website[35]

The second source of X-ray images was obtained from the Github repository[36]. It was made up of 100 chest X-ray images. 50 images were COVID-19 positive chest X-ray images and 50 were normal X-ray images.

Figure 2 shows sample chest X-rays images for both Covid-19 patient(upper) and non-covid patient(lower)

By combining these 2 datasets into a single dataset of a total of 269 positive Covid-19 X-ray images and 1391 normal X-ray images. 80% of the entire dataset was allocated for training and 20% for testing. From both sources, information regarding age or gender of the subjects was not provided in the meta-data.

3. Experimental Results

In this project, we implemented an ensemble deep learning system to detect and classify COVID-19 using X-ray im-

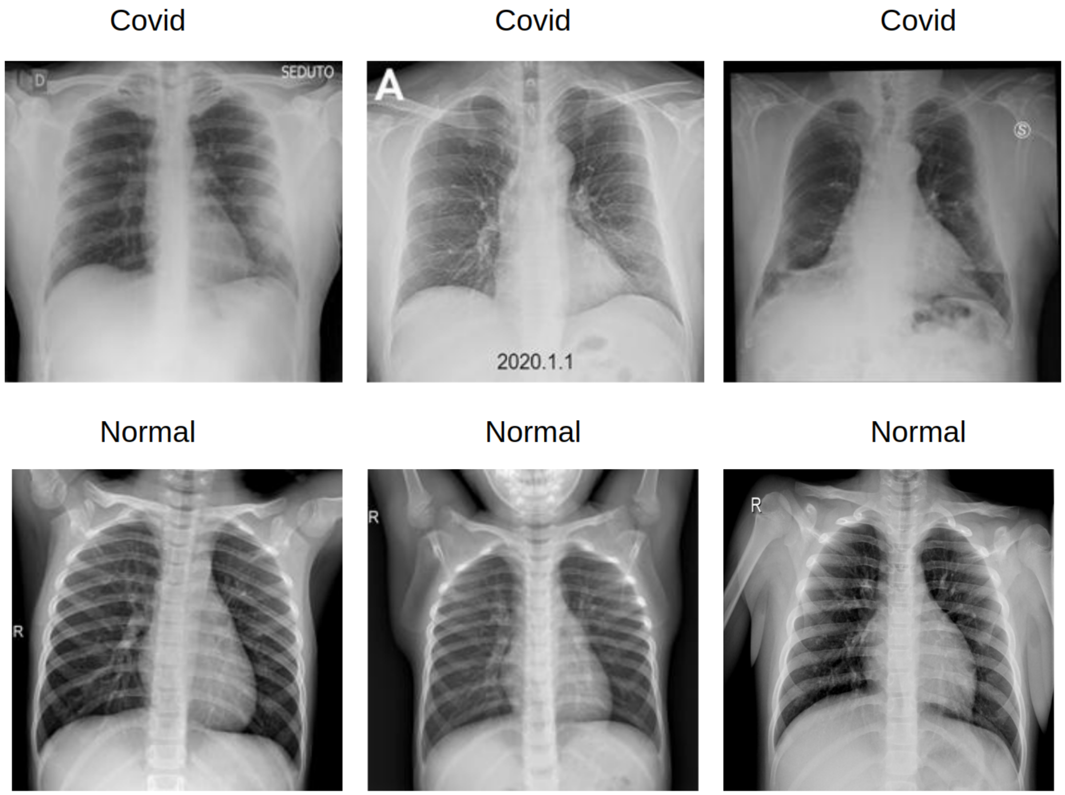

Figure 2. Sample Images of Chest X-Ray Images of Covid and Non-Covid Patients

ages by training 5 different classical state-of-the-art models: ResNet50, VGG-19, ResNet34, MobileNetV2, and DenseNet201. 80% of chest X-ray images were used for training and 20% for validation. This resulted in 1328 images for the training split and 332 images for the test(validation)split. Each of the models was trained for 10 epochs using a single 12GB NVIDIA Tesla K80 Graphics Processing Unit, GPU provided by Google Colabs. After training, each model outputs a prediction using the validation examples which is then put to a vote by our system. For each image, the system chooses the class label with the maximum number of votes as the correct prediction. Figure 3. shows the graphs of the training loss, accuracy, and error rate of ResNet50, ResNet34, VGG-19, MobileNetV2, DenseNet201 respectively.

The low ratio of positive COVID-19 X-ray images of approximately 16% of the total dataset required the use of other metrics of evaluation apart from accuracy. So, each model performance was also evaluated based on precision, sensitivity, and F1-score as can be seen in Table 1.

In addition to this, the result of confusion matrices for each of the 5 models was computed as can be seen in Figure 4.

By applying the max voting(majority) scheme to the 5 models, our ensemble deep learning system results in a performance accuracy of 99%.

4. Discussions and Performance Comparisons

The results shown in the previous section are all based on the dataset obtained from 2 independent sources(269 positive Covid-19 X-ray images and 1391 normal X-ray images). Previous studies done by researchers to combat this same problem have involved using a lesser number of training examples or getting training data from different sources. These factors have influenced the results achieved by these studies.

In [22], the authors proposed COVIDX-Net to diagnose COVID-19 in X-ray images and obtained a 90% accuracy rate using only 25 COVID-19 positive and 25 normal images.A deep CNN model(COVIDNet) designed by the authors in [23] for the detection of COVID-19 had a performance accuracy of 92.4% using 13,975 radiography images.By using the concept of transfer learning, the authors

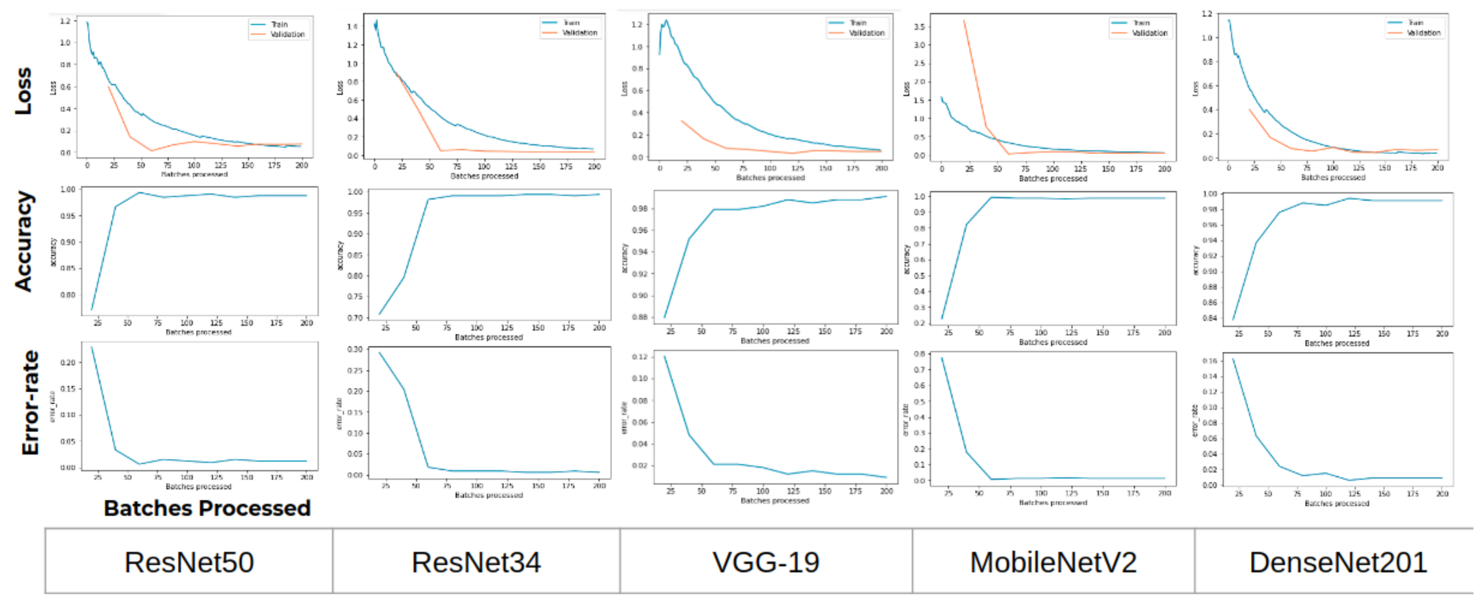

Figure 3. Graph of the Training loss, Accuracy, and Error Rate of ResNet50, ResNet34, VGG-19, MobileNetV2, DenseNet201

Table 1. Precision, Sensitivity, F1-Score, and Training Accuracy of 5 base models.

Model	Precision	Sensitivity(Recall)	F1-Score	Training Accuracy	Validation Accuracy
ResNet50	0.98	0.92	0.95	0.98	0.98
ResNet34	1.00	0.97	0.98	0.99	0.99
VGG-19	0.98	0.95	0.97	0.99	0.99
MobileNetV2	1.00	0.93	0.97	0.99	0.98
DenseNet201	1.00	0.95	0.97	0.99	0.99

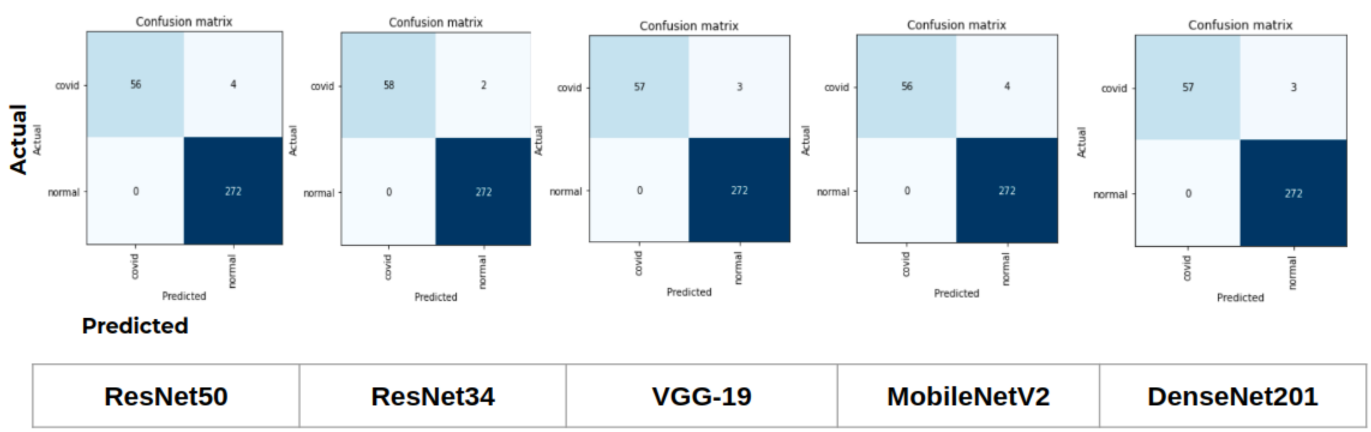

Figure 4. Confusion Matrix Plot of ResNet50, ResNet34, VGG-19, MobileNetV2, DenseNet201

in [24] were able to train a deep CNN using 1431 radiology images to obtain a stunning 98.75% performance accuracy in the case of COVID and non-COVID. However, they obtained a 93.48% performance for the 3-class problem; pneumonia included. In another related study, the authors of [25] implemented ResNet50, InceptionV3, and InceptionResNetV2 models using only 50 positive COVID-19 chest X-ray images and 50 normal chest X-ray images. The ResNet model performed the best with a 98% accuracy. In [37], the authors modified the Inception model and trained with CT images, achieving a classification accuracy of 82.9%. Also, in [38], the authors proposed a 3-

dimensional deep CNN model to detect COVID-19 from CT imagery and reported a 90.8% accuracy. Last but not the least, in [27] the authors developed a deep learning model called DarkCovidNet for the detection of COVID-19 using a total of 1125 images (125 COVID-19, 500 Pneumonia and 500 No-Findings) and obtained an accuracy of 98.08% and 87.02% for binary and three classes, respectively.

In this research study, an ensemble deep learning system of 5 deep CNN models was developed for the detection of COVID-19. Each of the models was trained using a total of 1660 chest X-ray images(269 positive COVID-19 images and 1391 normal images). From Table 1, we can clearly see that three models: ResNet34, VGG19, and DenseNet201 gave outstanding validation performance accuracies of 99%. However, when all the other metrics are taken into consideration, ResNet34 outperforms the remaining four models with a recall(sensitivity) of 97%, precision of 100%, and F1-Score of 98%. The lowest performance was achieved with the ResNet50 model having a validation accuracy of 98%, recall of 92%, precision value of 98%, and F1-score of 95%.

The performance accuracy obtained from the ensemble of these 5 models using the max voting scheme is 99%.

Comparing our system with other previous systems from the literature, it is evident that, our system outperforms all the existing implementations. An interesting observation is that the worse performance single model of the ensemble network(ResNet50) resulted in an accuracy equal to that of the best models in the literature. The greatest strength of our ensemble deep learning system is that the prediction of a sample is done collectively by 5 models in a voting format. This makes our model robust and less susceptible to misclassification errors. This is contrary to the other approaches that base their prediction on a single trained model. The diagnosis of positive and negative COVID-19 cases can be done with a greater degree of confidence.

This proposed system can be used for the diagnosis of COVID-19 using X-ray radiographs. One important point to note here is that most of the earlier studies conducted to combat this same issue were performed with limited data. Based on the fact that deep learning models are data-hungry models, we can conclude that this negatively affected the performance accuracy of these studies.

4. Conclusion

In this study, we have proposed an ensemble deep learning system to detect and classify COVID-19 cases from 1660 chest X-ray images. Our system performs this binary classification task by employing the ensemble max voting scheme based on the prediction of 5 classical deep learning models: ResNet50, ResNet34, VGG-19, MobileNetV2, and DenseNet201.

The performance accuracy achieved by our system is 99%. The system is highly robust and decreases misclassification error by ensuring that the prediction of every new unseen example doesn't rely solely on a single trained model but on 5 models. With this remarkably high-performance accuracy coupled with impressive system robustness, we are confident that such a system can play a significant role in the world especially in areas that are plagued by the deadly COVID-19 virus and have no expert radiologists. This system is versatile as it can also be adapted to detect other chest-related diseases such as pneumonia and tuberculosis. As more data is made available over time, we'll continue to re-train the base models making the system more robust and boosting the performance accuracy.

5. Future Work

In the future, I plan on working hand-in-hand with expert radiologists to have a broader perspective on what they have to say with respect to the results obtained by our proposed system. This will lay the foundation for deeper research and development of a more reliable and robust system that can be deployed in hospitals to circumvent the insufficient number of experts in the field of radiology and also act as a support system in the case of a pandemic where test kits become limited in supply. Such a system can be deployed in remote locations that are inaccessible to help diagnose respiratory diseases and save lives.

6. Acknowledgement

I'm grateful to Turkiye Burslari for organising such a challenging competition. I learned so much during the implementation of this project. I'm also grateful to Prof. Behcet Ugur Toreyin and Prof. Hamza Erden Salih for their unceasing support, guidance and knowledge.

7. References

[1] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, J. Y. Wong *et al.*, "Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia," *New England Journal of Medicine*, 2020.

[2] "WHO Coronavirus Disease (COVID-19) Dashboard kernel description," https://covid19.who.int/, accessed: 2020-10-01.

[3] "Naming the coronavirus disease (COVID-19) and the virus that causes it kernel description," https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/

naming-the-coronavirus-disease-(covid-2019)
-and-the-virus-that-causes-it, accessed: 2020-10-01.

[4] J. Zheng, "Sars-cov-2: an emerging coronavirus that causes a global threat," *International journal of biological sciences*, vol. 16, no. 10, p. 1678, 2020.

[5] S. El Hamichi, A. Gold, T. G. Murray, and V. K. Graversen, "Pandemics, climate change, and the eye," 2020.

[6] Z. Wu and J. M. McGoogan, "Characteristics of and important lessons from the coronavirus disease 2019 (covid-19) outbreak in china: summary of a report of 72 314 cases from the chinese center for disease control and prevention," *Jama*, vol. 323, no. 13, pp. 1239–1242, 2020.

[7] W. Wang, Y. Xu, R. Gao, R. Lu, K. Han, G. Wu, and W. Tan, "Detection of sars-cov-2 in different types of clinical specimens," *Jama*, vol. 323, no. 18, pp. 1843–1844, 2020.

[8] J. P. Kanne, B. P. Little, J. H. Chung, B. M. Elicker, and L. H. Ketai, "Essentials for radiologists on covid-19: an updateradiology scientific expert panel," 2020.

[9] Z. Y. Zu, M. D. Jiang, P. P. Xu, W. Chen, Q. Q. Ni, G. M. Lu, and L. J. Zhang, "Coronavirus disease 2019 (covid-19): a perspective from china," *Radiology*, p. 200490, 2020.

[10] E. Y. Lee, M.-Y. Ng, and P.-L. Khong, "Covid-19 pneumonia: what has ct taught us?" *The Lancet Infectious Diseases*, vol. 20, no. 4, pp. 384–385, 2020.

[11] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[12] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annual review of biomedical engineering*, vol. 19, pp. 221–248, 2017.

[13] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Computer methods and programs in biomedicine*, vol. 161, pp. 1–13, 2018.

[14] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A review," *Neurocomputing*, vol. 187, pp. 27–48, 2016.

[15] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[16] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[17] N. C. Codella, Q.-B. Nguyen, S. Pankanti, D. A. Gutman, B. Helba, A. C. Halpern, and J. R. Smith, "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM Journal of Research and Development*, vol. 61, no. 4/5, pp. 5–1, 2017.

[18] Y. Celik, M. Talo, O. Yildirim, M. Karabatak, and U. R. Acharya, "Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images," *Pattern Recognition Letters*, 2020.

[19] M. Talo, O. Yildirim, U. B. Baloglu, G. Aydin, and U. R. Acharya, "Convolutional neural networks for multi-class brain disease detection using mri images," *Computerized Medical Imaging and Graphics*, vol. 78, p. 101673, 2019.

[20] G. Gaál, B. Maga, and A. Lukács, "Attention u-net based adversarial architectures for chest x-ray lung segmentation," *arXiv preprint arXiv:2003.10304*, 2020.

[21] F. Pasa, V. Golkov, F. Pfeiffer, D. Cremers, and D. Pfeiffer, "Efficient deep network architectures for fast chest x-ray tuberculosis screening and visualization," *Scientific reports*, vol. 9, no. 1, pp. 1–9, 2019.

[22] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images," *arXiv preprint arXiv:2003.11055*, 2020.

[23] L. Wang and A. Wong, "Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images," *arXiv preprint arXiv:2003.09871*, 2020.

[24] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, p. 1, 2020.

[25] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *arXiv preprint arXiv:2003.10849*, 2020.

[26] P. K. Sethy and S. K. Behera, "Detection of coronavirus disease (covid-19) based on deep features," *Preprints*, vol. 2020030300, p. 2020, 2020.

[27] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of covid-19 cases using deep neural networks with x-ray images," *Computers in Biology and Medicine*, p. 103792, 2020.

[28] C. Zhang and Y. Ma, *Ensemble machine learning: methods and applications*. Springer, 2012.

[29] L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*. IGI global, 2010, pp. 242–264.

[30] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

[32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[33] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.

[34] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[35] "COVID-19 Radiography Database," https://www.kaggle.com/tawsifurrahman/covid19-radiography-database, accessed: 2020-10-01.

[36] "COVID-19 Chest X-Ray Images kernel description," https://github.com/drcerenkaya/COVID-19-Detection, accessed: 2020-10-01.

[37] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng *et al.*, "A deep learning algorithm using ct images to screen for corona virus disease (covid-19)," *MedRxiv*, 2020.

[38] C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, and X. Wang, "Deep learning-based detection for covid-19 from chest ct using weak label," *medRxiv*, 2020.