Using neural network to classify pneumonia in chest radiographs

Mary Chris Go, Zhanbota Bissaliyeva

***Abstract*—Early diagnosis of pneumonia is critical in preventing the disease. Doing it remains a challenging task that relies on skilled radiologists. Utilizing the current technology, medical diagnosis can be more efficient through artificial intelligence. With this motivation, the project aims to accurately classify frontal chest radiographs using a convolutional neural network (CNN) and depthwise separable CNN. It was run with a batch size of 10 for 15 epochs. Two models were created for each architecture: with and without data augmentation. After training both models and both architectures with an open-source dataset containing chest radiographs, the results show a significant difference between the recall of standard CNN versus depthwise separable CNN. The latter showed a reliable recall making it a robust architecture in medical applications. For both cases, data augmentation helped in improving the results.**

I. Introduction

Pneumonia is inflammation of the lung tissue, usually of infectious or bacteriological origin, with a predominant lesion of the alveoli. The alveoli of a person with pneumonia fill with pus and fluid, which affects breathing making it painful and reduces the oxygen intake. It is one of the most crucial diseases which affects more than 450 million people globally and takes the lives of 4 million of them each year [1]. Moreover, pneumonia remains the main cause of death in underdeveloped and poor countries, especially among children younger than 5 years and elderly people. However, the problem to solve here is not to cure pneumonia, but to detect the disease in chest X-rays. This challenging task relies on the abilities of an expert radiologist. Since the help of a trained specialist isn't always available, specifically in developing countries, the diagnosis of this disease can be improved by resolving the challenge of improving the accuracy and efficiency of pneumonia classification with the use of machine learning.

There have been various attempts of image classification of medical images. One famous study is a CheXNet [2]. It is a model which was proposed by a group of Stanford students and showed good results comparing to specialists with different years of experience. This is a 121-layer convolutional neural network that inputs a chest X-ray image and outputs the probability of pneumonia along with a heatmap localizing the areas of the image most indicative of pneumonia. To train the model they used the ChestX-ray14 dataset, released by Wang et al. in 2017 [3], which contains 112,120 frontal-view chest X-ray images of 30,805 unique patients. Each image is individually labeled with up to 14 thoracic diseases, including pneumonia . The model outperformed radiologists with an F1 score of 0.435 while the radiologist average was 0.387, thus showing that the performance of their model is statistically higher than performance of the radiologists.

mary.go@estudiante.uam.es, zhanbota.bissaliyeva@estudiante.uam.es

Convolutional neural network (CNN) made its way to various applications in the field of computer vision [4]. The development of CNN gave its ability in classification and detection. It includes many processing layers that analyze the input image by convolving multiple filters with the image. It produces a feature map that will serve as an input to the next layer. The input image goes through all the layers. There are various attempts of using neural networks for medical diagnosis [6]. This project is motivated by the idea of rapidly reviewing an enormous amount of medical images accurately in a short span of time. As artificial intelligence (AI) paves its way to revolutionize medical diagnosis, there are various proposed methods for improving AI such as CNN with depthwise separable convolution [7]. This technique deals with lesser computations and lesser parameters hence computationally cheaper and less prone to overfitting. It is assumed to be a spatial convolution to each channel of the input. It is proceeded by a pointwise convolution that manipulates the dimensions of the final image and combines the channels [7].

Despite the potential of AI, it is still a challenge to produce a robust model. This project aims to introduce another convolutional neural network method called depthwise separable convolution and explore its performance in classifying pneumonia from chest radiographs. The main goal is to implement and explore the results of implenting the standard CNN and depthwise separable CNN with varying models. As well as to understand the impact of using data augmentation strategies using a dataset obtained in [21] which we refer as Ped-ChestXray. Second objective is to study the state-of-the-art to give us enough foundation on what are the current researches in this given area. The project is organized as follows: Section II, we present all the review of papers we have studied related to the topic. Section III expounds on the proposed approach which is the depthwise separable convolutional neural network and a brief description of the standard convolutional neural network. Section IV details the implementation of the algorithms. In Sections V and VI, we analyze and discuss experimental results leading to a set of conclusions and recommendations in section VII.

II. Pneumonia Classification - State of the Art

The use of deep learning models in medical imaging tasks lately has experienced increasing interest, conditioned by the accessibility of huge datasets and the usage of new techniques. Recent research works demonstrate that deep learning appears to be a widely used powerful tool in numerous applications in the medical imaging field, such as medical image segmentation [8], tumors [9], diabetic retinopathy [10], and skin cancer [11] detection. There took place, among others, various relatively successful approaches for pneumonia detection in chest X-rays images. Reviews of published papers, with the dataset used for this project and other datasets, are presented.

A. Pneumonia Classification trained using Ped-ChestXray

In one study, dedicated to pneumonia detection from frontal chest radiographs, the authors made an emphasis on implementation a CNN model from scratch instead of using transfer learning [12]. The algorithm consists of two main elements: the feature extraction layer and a classifier. Proposed extractor composed of convolutional and max-pooling layers with a RELU activation function connecting them. The feature maps, obtained from it, are represented as 2D planes and need to be transformed into feature vectors for the classifier. In order to make a small dataset more suitable for deep convolutional neural network architecture several data augmentation methods were deployed, and the learning rate experienced variations. To guarantee substantial results authors reiterated training of the algorithm several times without results getting worse. For obtained results refer to Table 1.

Another research paper [13] uses the method of transfer learning. It is a method in which knowledge from a fully-trained network is utilized for the related target task. Furthermore, the authors of the paper decided to use MobileNets, instead of a traditional CNN. The main difference of mentioned model from the ordinary one lies in architecture, in which a single convolution layer is replaced by a depthwise convolution combined with a pointwise convolution. Although the authors have not used any preprocessing techniques for the dataset, they decided to consider three different proportions for dataset splitting. According to the comparison of these proportions the best ratio to train the model is 80%, 10%, and 10% of the train, valid, and test sets respectively. To give more accurate results authors implemented 16 versions of MobileNets, from which MobileNets 1.0-244 demonstrated the best performance and accuracy. This model was trained in different number of epochs $[1500, 2000, 2500]$.

One more research [14] has demonstrated and proved a better performance of depthwise separable CNN in comparison with traditional CNN in terms of computational cost. As for the dataset, it was split into training, validation, and test. In order to increase accuracy, data analysis and subsequent preprocessing such as resizing was employed. The authors of the research have compared several mainstream models for pneumonia classification, according to which they concentrated attention on the MobileNet's algorithm. Although, it is essential to mention that they tried ResNet-18, ResNet-50, VGG19, and a manually architectured CNN. Out of all the 5 mentioned models, MobileNet's outperformed other networks. However, the emphasis was made not only on the performance outcome, but the amount of computations needed for algorithm to obtain the prediction. Computational resources were prioritized, as the main objective of the work is to develop a model that can be used in embedded devices. It is to provide quick and relatively accurate pneumonia detection in countries experiencing a lack of equipment or radiologists. This paper uses the same architecture with the former discussed paper. Although, this model was trained for 20 epochs only.

B. Pneumonia Classification trained using ChexNet dataset

In another study [2], the authors implemented an algorithm called CheXNet, that has 121 layers composing the whole convolutional neural network. It was trained on ChestX-ray14 benchmark, which is an open-source dataset provided by NIH Clinical Center. Part of the evaluation is comparing the performance of the algorithm versus the individual annotation of four radiologists. The model outperformed three radiologists. The paper, therefore, concludes that CheXNet exceeds average radiologist performance. In addition, the authors compared the performance of their model to other published papers and saw better results from their algorithm. The algorithm is extended to detect 14 thoracic diseases. With this innovation, the authors hope that this step can improve healthcare delivery and efficient medical diagnosis in places where skilled radiologists are limited. In a research paper written by Paras Lakhani and Baskaran Sundaram [15] a set of four HIPAA-compliant datasets was used. These datasets are composed of more than a thousand posteroanterior chest radiographs. The authors decided to implement an ensemble of two Deep CNNs, which are AlexNet and GoogLeNet, to classify the radiographs as having pulmonary tuberculosis or as healthy. Moreover, performances of both untrained and pre-trained networks were compared. In order to improve the performance of the model in case of disagreement of DCNNs, images were interpreted by a certified cardiothoracic radiologist. To evaluate the performance of the algorithm AUCs were used. As a result radiologist-augmented approach is able to classify images with an AUC of 0.99, the sensitivity of 97.3%, and 100% specificity.

C. COVID-19 Related work

In terms of the current pandemic situation caused by coronavirus (COVID-19), a number of various approaches aimed at the identification of this disease were conducted. For instance, one study [16] set a goal to classify chest radiographs into COVID-19 and a number of other lungs related infectious diseases. One interesting features of the work is the age-based filtering of the chest images. Thus, selection criteria such as age more than 18 years lead to higher probability in classification between two groups: adults and pediatric age group. The trade-off of this dataset is the biased learning of the model.

One COVID-19 related work [17] based their DarkCovidNet classifier on Darknet-19 model. The model was designed to assist radiologists in two scenarios: in case of binary classification, to determine whether a patient has COVID infection or none, and three class classification, for COVID, pneumonia, and healthy images. The fact that COVID-19 is a subset of pneumonia resulted in an algorithm evaluating patients with COVID-19 as pneumonia. Therefore accuracy for binary classes appeared to be higher (98.08%) in comparison with the one for multi-class experiments (87.02%).

TABLE I: State-of-the-art Results

Reference	Model	Preprocessing and Augmentation	train_acc	train_loss	val_acc	val_loss
[12]	CNN from scratch	rescale rotation translation flipping	95,31	12,88	93,73	18,35
[13]	D-CNN with transfer learning	—	98	Not provided		
[14]	D-CNN	resizing	94,45	15,3	87,12	25.51

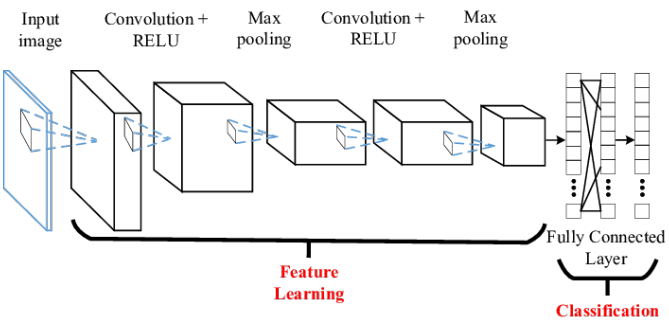


Fig. 1: The visual representation of how the implemented CNN works. Image adapted from [19].

III. Standard CNN vs Depthwise Separable CNN

A. Convolutional neural network (CNN)

This deep learning algorithm has been widely used for its many applications specifically in classification tasks. The goal of the whole architecture, for surpervised learning, is to solve an objective function. The algorithm are composed of neurons with learnable weights and biases. The visual representation of the flow can be seen in Figure 1. As shown, it consists of several layers in which each layer inputs a multi-dimensional array of numbers as input and outputs another multi-dimensional array of numbers as output. This output will become the input of the next layer.

Basically, there is an input image that goes through different layers. Each layer has its own task. The kernel convolves with the 3-channel input image and the result of this convolution will then be processed by the remaining layers. The goal of this convolutional operation is to solve a problem [20]. Convolution has advantages as they provide descriptions that are invariant to translations, rotations, and so on.

It still has not perfected the task of classification. With the complexity of its method, this architecture has been continuously developing with proposed variants to be more reliable. Also, it requires large computational resources. Therefore, we introduce a variant of standard CNN on the next part.

B. Depthwise separable CNN

CNN has been doing good work in many image classification tasks. Computer scientists use this as a foundation in innovating various techniques for a more efficient model with less parameters, but with similar results compared to the standard CNN. A standard CNN involves the convolution of input, output, width, and height on getting the parameters. This formula means that it can have more parameters which we try to avoid since it increases the chance of overfitting. With this problem, people try to solve it by modifying the standard CNN.

The basis of this convolution is from the idea that spatial dimensions can be separated. This method involved two methods: depthwise convolution and pointwise convolution. There are many ways to to do this. Depthwise separable CNN is not limited to dividing the image on the first layer. It can be used anywhere in the architecture, at any point, for all layers, or just for some of them. For simplicity of grasping the theory behind it, think of an input image with channels that are divided into R, G, and B. Each channel corresponds as one input channel [7]. From Figure 2, we have a three-channel image and a three-channel filter. These channels are then separated. The convolution happens between one image channel with the corresponding channel. Then, the resulting images are stacked thereafter. The second method involves replicating the normal convolution to produce the same effect. It uses a kernel that iterates every point. It has depth depending on the number of channels the input image has. This method is used for the output image to have a three-channel image convert to a 1 image channel [7]. With its less computations, the main goal of depthwise separable CNN is to be an architecture that yields similar results but with less parameters. Because of less multiplications, the computational complexity decreases as well making it more computationally efficient.

Despite of it being a stellar architecture theoretically, there are disadvantages on using this architecture. As it reduces the number of parameters in a convolution, it also means the there are few parameters. With lesser parameters, the network might fail to properly learn during the training. But with proper usage, it can manage to enhance efficiency without significantly reducing the effectiveness which makes it a robust architecture.

IV. Implementation of Standard CNN and Depthwise Separable CNN

All models were run locally with a GPU of GeForce GTX 1050ti and 8GB RAM. Certain parameters such as epoch and batch size were dependent on these factors to cater its capability and limitations.

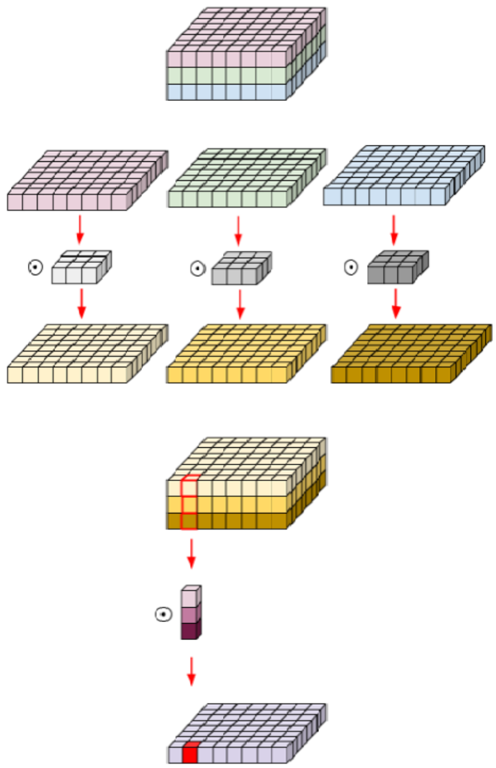


Fig. 2: The visual representation of how depthwise separable CNN treats each channel. Image adapted from [5].

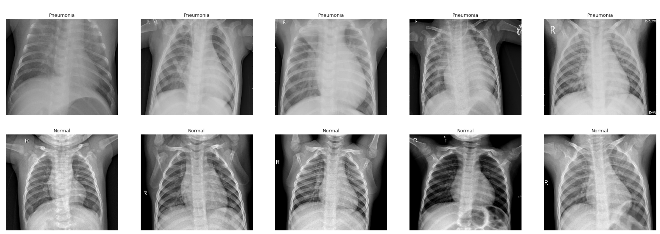


Fig. 3: These images were randomly selected from the dataset. The images on the top row are the pneumonia positive while the bottom row shows the normal chest radiographs. Images adapted from [3].

A. Dataset

The chosen dataset, Ped-ChestXray, was obtained from an openly available dataset [21]. It contains three folders: train, test, and validation. Each folder has subfolders for a normal and with pneumonia chest radiographs. It has 5, 863 grayscale images in JPEG format with 3883 images labeled as pneumonia positive. It contains various dimensions of the images ranging form 384×127 to 2772×2098. Samples of the randomly picked images can be seen in Figure 3. The images were from randomly selected patients of pediatric patients from one to five years old residing in Guangzhou Women and Children's Medical Center, China. It was performed as part of their clinical care. The images passed through quality control by two expert physicians. They simply removed low-quality scans and made sure that the labels were correct. Further analysis was done by a third expert for assurance.

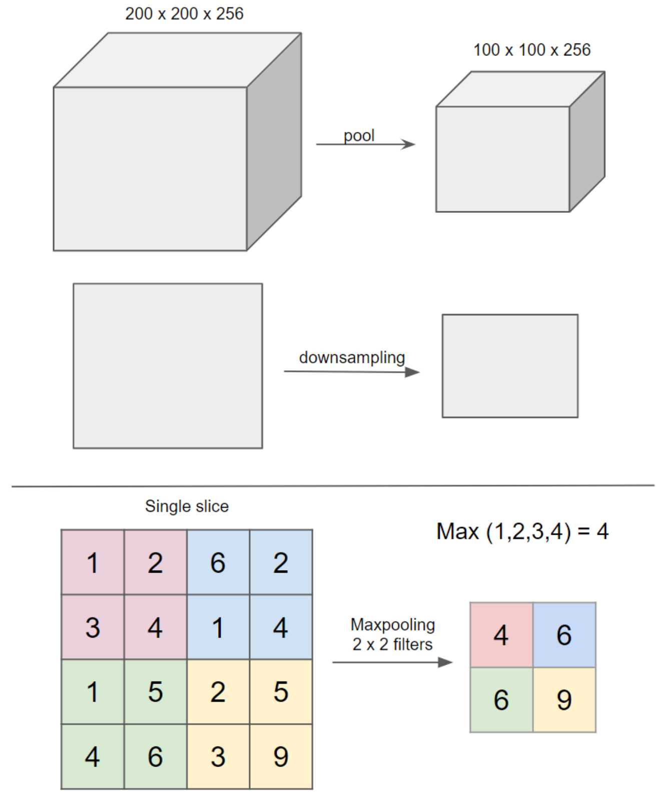


Fig. 4: Visual representation of the pooling operation.

B. Convolutional neural network

This method was implemented by using Xingyu Bian's Kaggle notebook [18]. From the previous section, we see the great disproportionality of the positive and negative images. The solution to this issue is data augmentation. The image generator provides various approaches in augmenting the images.

For this classification task, we use the manually implemented architecture used in [19] for face recognition using CNN. The input to the first layer is the input image. A simple CNN was implemented, meaning there is a sequence of layers which each layer transforms activations to another using a differentiable function. Three main types of layers were used to build the desired CNN architecture: convolution layer, pooling layer, and fully-connected layer. The mentioned layers were then stacked to form the whole architecture.

- **INPUT** : it holds the raw pixel of the input image. For uniformity, all images were resized to $[200 \times 200]$
- **CONV** : it computes the output of neurons. These neurons are connected to the local regions in the input. The dot product of the weights and the small region they are connected to happens on this layer. The volume output in this case is $[200 \times 200 \times 256]$ stating that we decided to use 256 filters.

TABLE II: The CNN layers that was implemented to classify chest radiographs. Colors were used to show the grouping of each block.

Layers of Standard CNN (in order)
Input layer
Conv2D
Activation
MaxPooling2D
Batch Normalization
Conv2D
Activation
MaxPooling2D
Batch Normalization
Conv2D
Activation
MaxPooling2D
Batch Normalization
Flatten
Dropout
Dense
Activation
Dropout
Dense
Activation

- **RELU** : it is responsible for applying an element-wise activation function. It outputs an unchanged size of $[200 \times 200 \times 256]$.
- **POOL** : the down-sampling operation along the width and height of the input is done on this layer. This layer is essential for the model to require less computational resources. It also acts as a noise suppressant but performs de-noising for some dimensionality reduction [20]. This layer results to a $[100 \times 100 \times 256]$ as the filter size is 2 as shown in Figure 4. We preserved here the volume depth. The used variant of pooling here is max pooling as it is the most commonly used downsampling operation.
- **FULLY-CONNECTED** : this layer is responsible for computing the class scores. As the name suggests, neurons in this layer are connected to the neurons of the previous volume.

By the use of the layers mentioned above, the architecture transforms the image layer by layer. It is important to remember that some layers do not have parameters. To be specific, the convolution layers and fully-connected layers perform transformations that are functions of the parameters such as weights and biases, and the activations of the input volume. The weights and biases of these layers were trained with gradient descent so the computed class scores are consistent with the labels in the train set. Figure II shows the detailed structure of the layers in which the details of the input and output sizes are indicated.

C. Depthwise convolutional neural network

This kind of convolution was done with the same set of parameters with the standard CNN such as batch size, epochs, and learning rate. The architecture is very similar to the standard CNN except for the first layers of each block. As seen in Figure II, depthwise separable CNN was also implemented in this way except the *conv_2d* layers were converted to *SeparableConv2D* which means that there has

TABLE III: The depthwise separable CNN layers that was implemented to classify chest radiographs. Colors were used to show the grouping of each block.

Layers of Depthwise Separable CNN (in order)
Input Layer
Conv2D
Max Pooling 2D
Separable Convolution 2D (input)
Separable Convolution 2D
Batch Normalization
Max Pooling 2D
Separable Convolution 2D (input)
Separable Convolution 2D
Batch Normalization
Max Pooling 2D
Separable Convolution 2D (input)
Separable Convolution 2D
Batch Normalization
Max Pooling 2D
Flatten
Dense
Dropout
Dense
Dropout
Dense
Dropout

been a division of channels before passing through batch normalization and max pooling. This is the main difference between the two architectures. They are very much alike but in depthwise, the channels of the image are treated as a one input. These were easily implemented through Keras package. *tf.keras.layers.SeparableConv2D* is responsible for performing a depthwise spatial convolution that is followed by pointwise convolution that compiles the output channels. The following layers can be seen in Table III.

V. Experimental Results

A. Setup

The architecture shown in Figure 1 was implemented for standard CNN and Figure 2 for depthwise CNN. The hyper-parameter, learning rate, was set to 0.0001. With this slow learning rate, the model was able to learn an optimal set of weights. Although, this led to a slower training time. It uses an Adam optimizer and a binary cross-entropy loss method. The latter was chosen as this loss as it minimizes the distance between two probability distributions. It is tailored for binary classification.

The training was done in 15 epochs and a batch size of 10. It was trained for approximately 2 hours for both models (CNN with and without data augmentation). These hyperparameters were based from the capacity of the local machine the model was trained in.

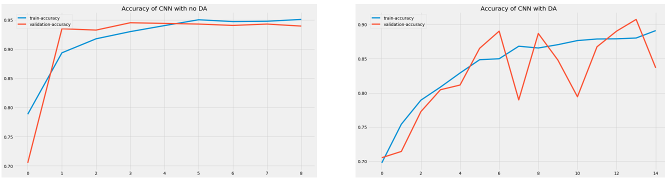
As mentioned earlier, data augmentation was explored to solve the disproportionality of the dataset. Techniques such as $90°$ for random rotations, zoom, width shift, height shift, horizontal flip, and vertical flip were used. These methods were the simple ones and was randomly applied to the dataset.

B. Performance metrics

Before reading this section, it is important to note that there are two metrics that are considered: precision and recall.

TABLE IV: Obtained results from the standard CNN model using mean per class precision and recall method.

	Precision	Recall
with DA	93%	84%
without DA	97%	79%



(a) CNN with no data augmentation (b) CNN with data augmentation

Fig. 5: The behavior of the accuracy of train and validation sets.

These metrics are not greatly affected by imbalanced number of radiographs with or without pneumonia. For unbalanced datasets, accuracy is avoided as it might give results with bias. Therefore, these two metrics will be used as it exhibits the ability of the model to label pneumonia right. In the terms of pneumonia detection, recall refers to how many chest radiographs were labeled correctly. On the other hand, precision refers to the number of chest radiographs that was labeled.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{1}$$

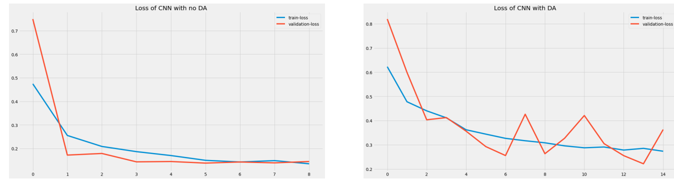
$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{2}$$

Furthermore, it is also essential to take note of the running time as one of the goals of the depthwise separable CNN is the reduce computational expenses. First, we analyze the results of CNN.

C. Results of CNN

As the main goal of the paper is to compare both kinds of the neural network, we focused on exploring and implementing the best possible setting of standard CNN for the first part. Two settings were explored: with data augmentation and without data augmentation. The obtained values were computed using mean per class recall and precision. The model was tested using 628 randomly selected chest radiographs. Both precision results yielded a good result of 93% for CNN with data augmentation and 97% for CNN without data augmentation. Table IV summarizes the result of the standard network.

The model was trained for approximately 2 hours for the model with data augmentation and 1.6 hours for the model without data augmentation. This is expected as the former has an additional step of generating more images to balance the dataset. To check if the training is working well, we graphed the loss and accuracy for both train and validation sets as shown in Figures 5a, 5b, 6a, and 6b.

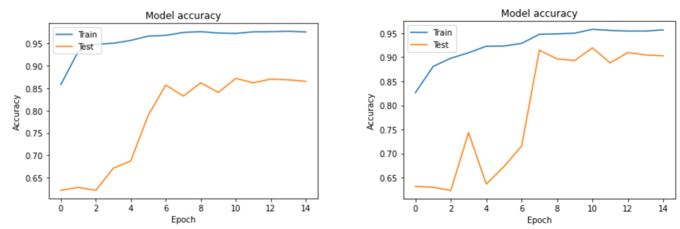


(a) CNN with no data augmentation (b) CNN with data augmentation

Fig. 6: The behavior of the loss of train and validation sets.

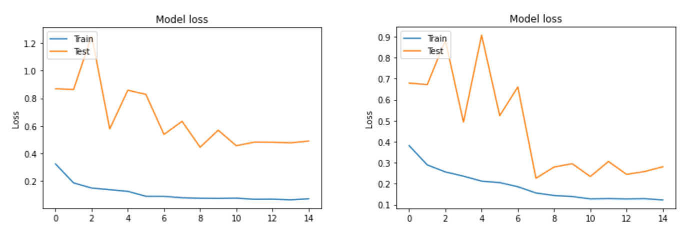
TABLE V: Obtained results from the depthwise separable CNN model using mean per class precision and recall method.

	Precision	Recall
with DA	87%	98%
without DA	82%	99%



(a) Depthwise separable CNN with no data augmentation (b) Depthwise separable CNN with data augmentation

Fig. 7: The behavior of the accuracy of train and validation sets.



(a) Depthwise separable CNN with no data augmentation (b) Depthwise separable CNN with data augmentation

Fig. 8: The behavior of the loss of train and validation sets.

D. Results of depthwise separable CNN

We explore now the results for the depthwise separable convolutional neural network. To recall, our hypothesis is to this algorithm to have better results in terms of recall and with more efficiency in terms of usage of computational resources.

As observed in Table V, the recall improved a lot although there is a trade-off in precision. The model was trained for approximately an hour for the model with data augmentation and for the model without data augmentation. It agrees with the initial hypothesis that there will be a less computing time as compared to the standard CNN. To check if the training is working well, we graphed the loss and accuracy for both train and validation sets as shown in Figures 7a, 7b, 8a, and 8b.

VI. Discussion

A. Standard CNN

The data presented in Table IV represents the highest results we were able to achieve. We observed a good performance in precision but not in the recall. We did many attempts of implementing this right with our desired parameters. These models showed a well-trained model. It neither underfit nor overfits the training process. Figures 5a and 6a showed the best performing scenario. Given that these graphs are the ones without data augmentation, we can think of the dataset to be an easy one. For Figures 5b and 6a, we see that there is still a room for improvement. The model has not yet stopped learning as the training accuracy has not yet stopped increasing and the loss has not yet stopped decreasing. One solution we can do here is to train it with more epochs. This can be done for the next semester. More work is needed to achieve consistency reported in the literature. However, it is clear that the introduction of data augmentation helped to improve the reliability of the model.

B. Depthwise Separable CNN

The data presented in Table V shows the highest obtained results from training the algorithm. One thing to take note deduct from the table is the good recall performance. As mentioned before, this is the most important metric in terms of medical applications. This metric means that we do not have large number of false negative which in the medical field means diagnosing the patient all clear when in fact he/she has pneumonia. Another observation if the improvement of having data augmentation. For both models, using data augmentation technique gives an improvement consistently. Figures 7a and 8a showed the best performing scenario. Although the test was not close to the train, it followed the same pattern of increasing values for accuracy and slowing down for the loss. For Figures 7b and Figure 8b, these graphs were the best training we obtained although there are still spikes that shows the instability of the training of the algorithm. The model are converging as the loss shows a decrese in value. For the loss of depthwise separable CNN with data augmentation, it can be seen that it hasn't stopped learning. A solution for these problems that we see is to increase the number of epochs as well. But as stated, computational resources are limited and we try to limit it to 15 epochs to have a proper baseline for comparison.

VII. Conclusions

After studying several papers, we have discovered that there are several datasets to explore but the networks that they use are usually demanding in terms of computational resources. We have explored some implementations such as ChexNet [2] that has 121 layers with their dataset larger than what we are using. This implementation is a recommendation for future work.

Basically, the fundamentals of implementing these two networks are explored. From the obtained values of CNN, we observed the significant difference of embedding data augmentation to the whole architecture. We can conclude the data augmentation really helped both algorithm in improving the result. Although, there is a trade-off. Greater precision results were seen for the standard CNN without data augmentation. On the other hand, greater recall results were observed for the depthwise separable CNN for both cases of with and without data augmentation.

Another observation are the problems caused by the unbalanced datasets. Although it was handled through data augmentation, future work requires more techniques such as focal loss [22], class-balanced loss [23], and modeling tail [24]. These techniques can be explored, although changing to a more curated dataset is also an option as larger datasets are now available [2]. This can enhance the robustness of the model as it provides more ground truth training data.

These findings agree with the initial hypothesis of depthwise separable CNN producing better results and with a higher computational efficiency. The main goal of the paper was reached through implementation of both algorithms with the same parameters for fair comparison. For future work, it is also recommendeded to explore other parameters such as different learning rate, higher number of epochs and batch size, and different arrangement of layers. These mentioned ideas can be further explored to maximize the capacity of both algorithms.

References

[1] Ruuskanen, Olli et al. "Viral pneumonia." Lancet (London, England) vol. 377,9773 (2011): 1264-75. doi:10.1016/S0140-6736(10)61459-6

[2] Pranav Rajpurkar, Jeremy Irvin, Kaylie Zhu, Brandon Yang, Hershel Mehta, Tony Duan, Daisy Ding, Aarti Bagul,Robyn L. Ball, Curtis Langlotz, Katie Shpanskaya, Matthew P. Lungren, Andrew Y. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning

[3] Wang, Xiaosong, Peng, Yifan, Lu, Le, Lu, Zhiyong,Bagheri,Mohammadhadi, and Summers,Ronald M. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. arXiv preprint arXiv:1705.02315, 2017.

[4] Krizhevsky, Alex and Sutskever, Ilya and Hinton, Geoffrey. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Neural Information Processing Systems. 25. 10.1145/3065386.

[5] Pandey, A. (2018, September 09). Depth-wise convolution and depth-wise separable convolution. Retrieved May 05, 2021, from https://medium.com/@zurister/depth-wise-convolution-and-depth-wise-separable-convolution-37346565d4ec

[6] D. S. Kermany, M. Goldbaum, W. Cai, C. C.S. Valentim, H.-Y. Liang, S. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, Cell 172, 1122 (2018).

[7] F. Chollet, in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Hawaii, USA, 2017), CVPR 2017.

[8] Hesamian, M.H., Jia, W., He, X. et al. Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. J Digit Imaging 32, 582–596 (2019). https://doi.org/10.1007/s10278-019-00227-x

[9] Zilong Hu, Jinshan Tang, Ziming Wang, Kai Zhang, Ling Zhang, Qingling Sun, Deep learning for image-based cancer detection and diagnosis. A survey, Pattern Recognition, Volume 83, 2018, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2018.05.01

[10] Arcadu, F., Benmansour, F., Maunz, A. et al. Deep learning algorithm predicts diabetic retinopathy progression in individual patients. npj Digit. Med. 2, 92 (2019). https://doi.org/10.1038/s41746-019-0172-3

[11] E. Andre, K. Brett, A Roberto et al., Dermatologist-level classification of skin cancer with deep neural networks, Nature, vol. 542, no. 7639, pp. 115–118, 2017.

[12] Okeke Stephen, Mangal Sain, Uchenna Joseph Maduh, Do-Un Jeong, An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare, Journal of Healthcare Engineering, vol. 2019, Article ID 4180949, 7 pages, 2019. https://doi.org/10.1155/2019/4180949

[13] Nghia Duong-Trung, Tuyen Tran Ngoc, Hiep Xuan Huynh, Automated Pneumonia Detection in X-Ray Images via Depthwise Separable Convolution Based Learning, Conference Paper, 2019

[14] Zhenjia Yue, Liangping Ma, Runfeng Zhang, Comparison and Validation of Deep Learning Models for the Diagnosis of Pneumonia, Computational Intelligence and Neuroscience, vol. 2020, Article ID 8876798, 8 pages, 2020. https://doi.org/10.1155/2020/8876798

[15] Lakhani P, Sundaram B, Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks, Radiology, 2017 Aug;284(2):574-582. doi: 10.1148/radiol.2017162326. Epub 2017 Apr 24. PMID: 28436741.

[16] Sharma, Arun Rani, Sheeba Gupta, Dinesh, Artificial Intelligence-Based Classification of Chest X-Ray Images into COVID-19 and Other Infectious Diseases, International Journal of Biomedical Imaging, 2020, 1-10. 10.1155/2020/8889023.

[17] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, U. Rajendra Achary, Automated detection of COVID-19 cases using deep neural networks with X-ray images,Computers in Biology and Medicine, Volume 121, 2020, 103792, ISSN 0010-4825

[18] Bian Xingyu. (2020, June 8). Pneumonia Detection Using Deep Learning. Kaggle. https://www.kaggle.com/therealcyberlord/pneumonia-detection-using-deep-learning.

[19] Kamencay, Patrik and Benco, Miroslav and Mizdos, Tomas Radil, Roman. (2017). A New Method for Face Recognition Using Convolutional Neural Network. Advances in Electrical and Electronic Engineering. 15. 10.15598/aeee.v15i4.2389.

[20] Jiuxiang Gu, Zhenhua Wang, Jason Kuen, Lianyang Ma, Amir Shahroudy, Bing Shuai, Ting Liu, Xingxing Wang, Gang Wang, Jianfei Cai, and Tsuhan Chen. 2018. Recent advances in convolutional neural networks. Pattern Recogn. 77, C (May 2018), 354–377. DOI:https://doi.org/10.1016/j.patcog.2017.10.013

[21] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H. Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, F. Yan, et al., Identifying medical diagnoses and treatable diseases by image-based deep learning, Cell 172, 1122 (2018).

[22] T. Lin, P. Goyal, R. Girshick, K. He and P. Dollar, "Focal Loss for Dense Object Detection" in IEEE Transactions on Pattern Analysis Machine Intelligence, vol. 42, no. 02, pp. 318-327, 2020. doi: 10.1109/TPAMI.2018.2858826

[23] Y. Cui, M. Jia, T. Lin, Y. Song and S. Belongie, "Class-Balanced Loss Based on Effective Number of Samples," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp.9260-9269, doi: 10.1109/CVPR.2019.00949.

[24] Wang, Yu-Xiong et al. "Learning to Model the Tail." NIPS (2017).