MobileNetV2 for Identification of Pneumonia in Pediatric Chest X-Rays

Ari Barnett
University of Central Florida – Institute for Modeling, Simulation, and Training
3039 Technology Pkwy, Orlando, FL 32826
aribarnett.roldan@knights.ucf.edu

Introduction

Respiratory illness in pediatric patients can be especially devastating in the first few years of life. Pneumonia is particularly troublesome for those infected. Methods of infection include viral, bacterial, and fungi – therefore for vulnerable populations quick diagnosis of ailment is important to beginning appropriate treatments. For diagnosis of Pneumonia, chest x-ray is commonly utilized and presents as an opaque whitening within the lung [1]. The severity of presentation and pathology varies the opaqueness present (figure 1).

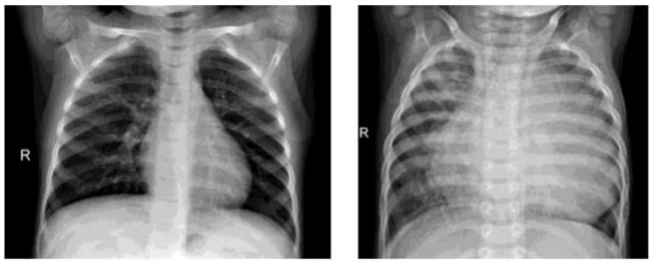
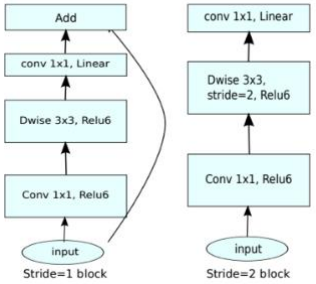


Figure 1 Example of Normal Chest X-Ray (Left) and Bacterial Pneumonia (Right) [2]

For this reason, we propose utilizing a CNN (MobileNetV2) for performing binary-classification of healthy and pneumonia images among pediatric patient examples. First, the CNN will be trained from scratch and evaluated for accuracy. Secondly, we will leverage the same CNN trained utilizing the ImageNet dataset to perform transfer learning and hyper tuning to attempt to increase accuracy of the task. For both tasks we provide loss-plots as well as confusion plots and failure examples.

Model & Training

For this project the choice to utilize MobileNetV2 is due to limitation in computational power. Therefore, we aimed to select a smaller CNN. Within MATLAB this is performed by first initializing the model from the deep learning toolbox. The model is then augmented to perform binary classification tasks by replacing both the final fully connected layer as well as the classification layer. The layers are replaced to perform binary classification by altering the output size of each layer to 2 instead of 1000 that is present for use on the ImageNet dataset. For task 1, the network is initialized without weights, and for task 2 with ImageNet weights. The general architecture is presented below (figure 2). The framework is the same for both iterations (with/without ImageNet weights). Training options used are as follows (figure 2).



Training Option	Selection
Solver	**SGDM**
Max Epochs	**2**
Mini Batch Size	**128**
Execution Environment	**Parallel CPU**
Validation Freq.	**10**
Validation Patience	**3**
L2 Regularization	**0.0001**
Validation Shuffle	**Every Epoch**

Figure 2 General Architecture for MobileNetV2 and Training Options Utilized for Project [3]

Data & Preprocessing

The dataset utilized is comprised of 5,863 chest x-rays from pediatrics patients of Guangzhou Women and Children's Medical Center. Ages of patients range from 1 to 5 years. Images were all collected retrospectively and were not taking outside of routine provided care. The dataset used is publicly available for use at [4]. Examples are provided below (figure 3). Upon initial examination of the dataset a noticeable discrepancy is the imbalance present among normal and pneumonia examples (figure 3). Due to this, we employ data augmentation methods to try and mitigate the effect on

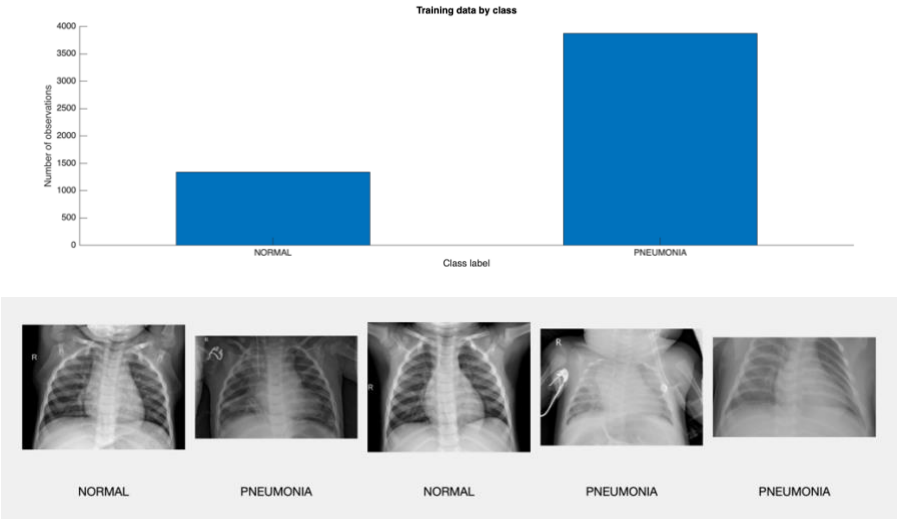


Figure 3 Data Split Prior to augmentation

accuracy. To perform this the normal images in the training set are first divided into four groupings: images 1-600 (2x), images 601-1200, images 1201-END. These groupings then underwent the following augmentations: 30% increase in size, flip image from top-bottom, rotation by 20°, and flip image horizontally. These augmentations then create a more equal training set for use (figure 4). Alternatively, the generation of synthetic data can be utilized to aid in resolving data imbalance – methods such as SMOTE, and or GAN could prove useful.

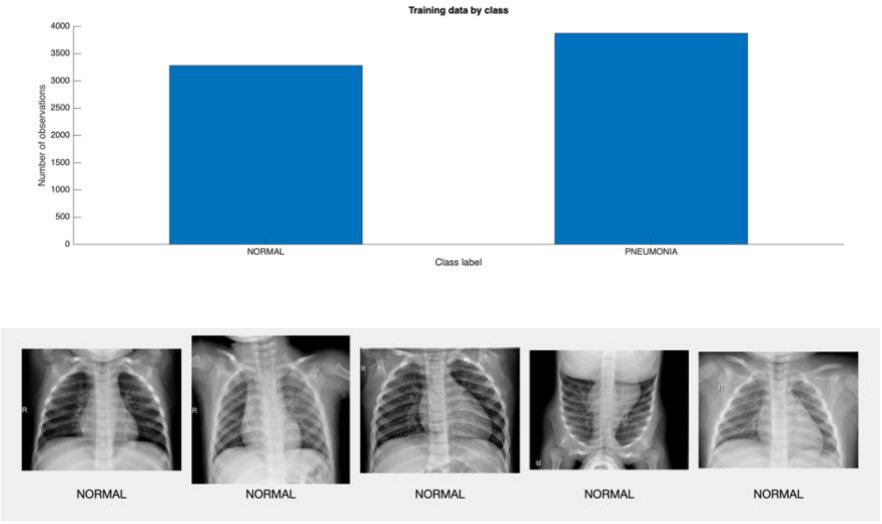


Figure 4 Data Split after augmentation

Results

For both tasks, resulting accuracy is reported as 78.37% and 79.49% respectively. It is noted that even with data augmentation and L2 Regularization, models consistently demonstrate issue with classifying healthy vs infected images. This can furthermore be seen in failure case examples (figure 5). This is further evaluated within the discussion section.

Task 1 Loss and Confusion Matrix

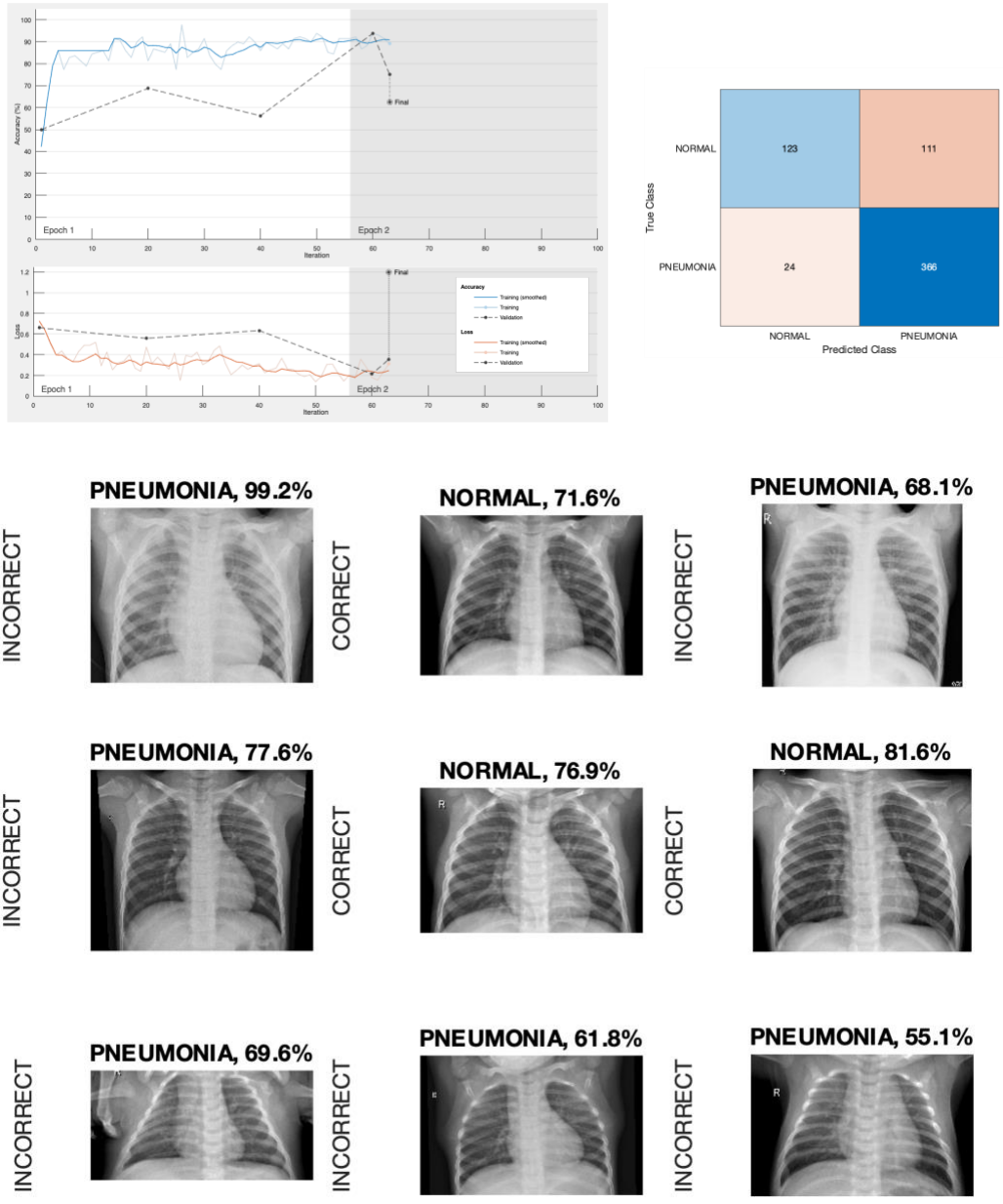
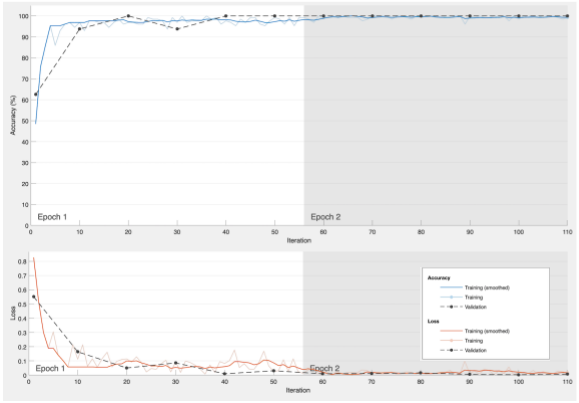
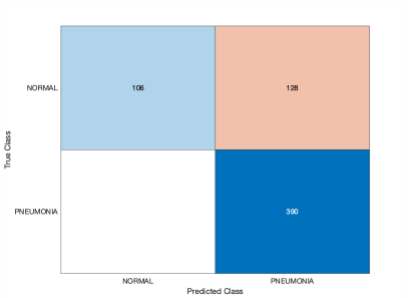


Figure 5 Loss/Accuracy Plots and Confusion Matrix for Task 1 and 2

Task 2 Loss and Confusion Matrix





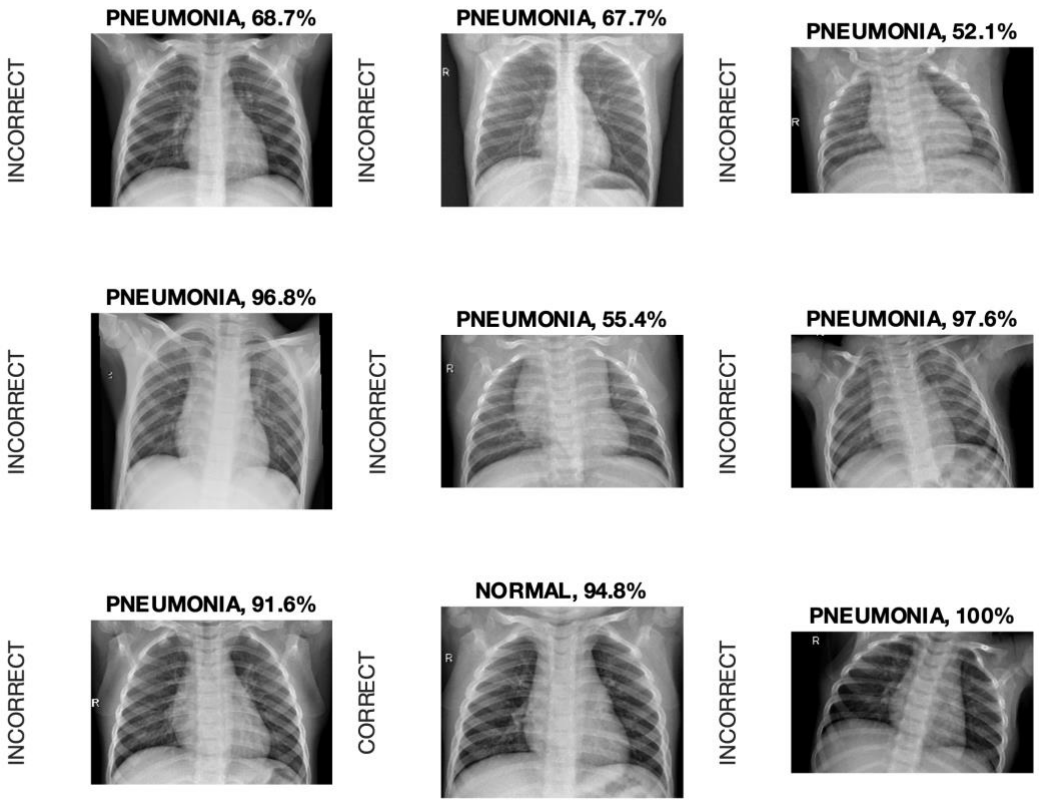


Figure 5 Loss/Accuracy Plots and Confusion Matrix for Task 1 and 2

Discussion

While both models provide an additional tool for clinical decision support it is important to note the overwhelming bias caused by the data. As seen in figure 5 both models have little to no issue classifying imaging where pneumonia is truly present. However, the issue arises in instances of false positives when presented with normal healthy images. To evaluate this further we utilize gradient-weighted class activation mapping to visualize target areas for the classification process. Upon examination (figure 6) we can see that in instances of false positives there is heavy focus on potential overexposure within the image as opposed to important clinical features.

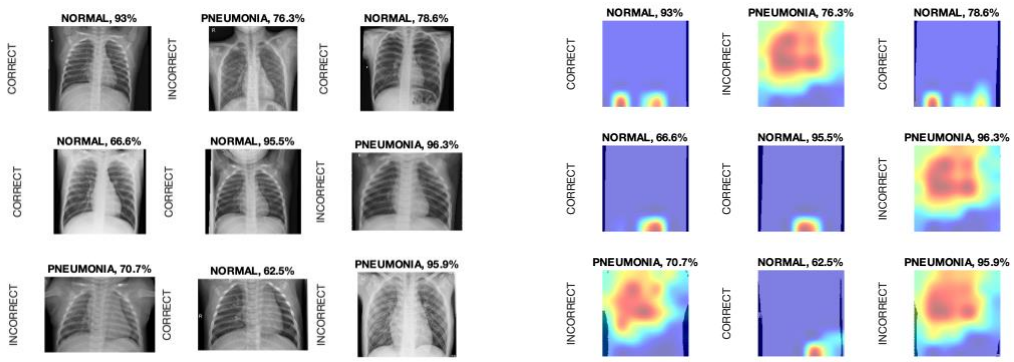


Figure 6 gradient-weighted class activation mapping of normal imaging examples.

Overexposure can commonly be caused by improper settings on the digital imaging device, and or improper patient positioning/stature. The latter could be a viable option due to the inability for younger patients to properly follow clinical procedures. During an imaging session, patients are required to hold their breath when the x-ray is captured – failure to perform this step results in a blurred image which is reminiscent of overexposure. Another approach previously mentioned is approach the data disparity via a synthetic minority oversample technique (SMOTE) to create in-silico patients. While this can address the imbalance within the dataset – it cannot directly denoise the images. Therefore, a more appropriate approach could be to utilize a DCGAN to produce synthetic clarified images to aid in reduction of false-positive instances.

Further directions for a work like this would be to provide comparisons on other CNN models (Resnet18, MobileNetV1, SqueezeNet, etc) and evaluate the potential of off-the-shelf approaches for quick and easy clinical deployment as decision-making aids. Technologies like DL approaches can provide a useful tool to aid in the training of resident physicians on complex cases where imaging has difficult interpretation.

References

[1] C. Ebeledike and T. Ahmad, “Pediatric Pneumonia,” in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2022. Accessed: Mar. 20, 2023. [Online]. Available: http://www.ncbi.nlm.nih.gov/books/NBK536940/

[2] “Transfer learning exploits chest-Xray to diagnose COVID-19 pneumonia,” *News-Medical.net*, Dec. 18, 2020. https://www.news-medical.net/news/20201218/Transfer-learning-exploits-chest-Xray-to-diagnose-COVID-19-pneumonia.aspx (accessed Mar. 20, 2023).

[3] “Papers with Code - MobileNetV2 Explained.” https://paperswithcode.com/method/mobilenetv2 (accessed Mar. 20, 2023).

[4] “Chest X-Ray Images (Pneumonia).” https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia (accessed Mar. 20, 2023).