Diagnosing Pneumonia from X-Ray Images Using Machine Learning

Augusto José Peterlevitz[1] , José Roberto C. S. Filho[2]
[1]Institute of Physics 'GlebWataghin' (UNICAMP), contact: augusto.peterlevitz@gmail.com
[2]Institute of Computing (UNICAMP), contact: jrcdsf@gmail.com

Abstract

Medical diagnosis based on observation of images is not always an easy task. It's the case of detecting pneumonia in lungs using X-Ray images, where physicians try to locate the inflammations looking at the image. In this project, an implementation of fine tune and transfer learning associated with Support Vector Machines [1] and Random Forests [2] using the Inception-V3 [3] CNN is proposed as solution of X-Ray images classification from normal and with pneumonia children with the age between 1 and 5 years old.

1. Introduction

Pneumonia is a infection in the lungs caused by bacterias, viruses and fungi. This infection causes inflammations on the alveoli, filling it with fluid or pus and making difficult to breath. The most common symptoms are: coughing, fever, shortness of breath and chest pain. The pneumonia diagnoses by observing X-Ray images is not always an easy task, as can be observed at figures 1 and 2. In this project, fine tune and transfer learning associated with Support Vector Machines (SVM) and Random Forests using the Inception CNN are used in an attempt to predict the state of the patient: normal or with pneumonia.

2. The Dataset

An dataset containing 5863 chest X-Ray images (JPEG) organized into train, validation and test with two categories (Pneumonia/Normal) was used to train the models. The images were obtained from children of one to five years old from Guangzhou Women and Childen's Medical Center [4]. All low quality or unreadable scans were removed and then the diagnoses for the images were graded by two experts physicians before being cleared to be available for training usage. Figure 1 and 2 shows the X-Ray images from children without and with pneumonia.

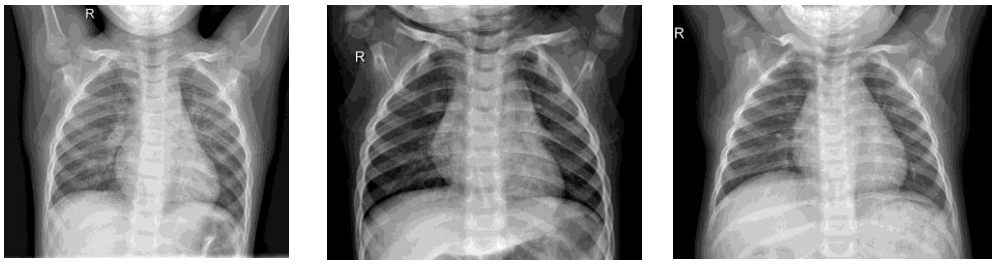

Figure 1: Chest X-Ray images from chidren without pneumonia.

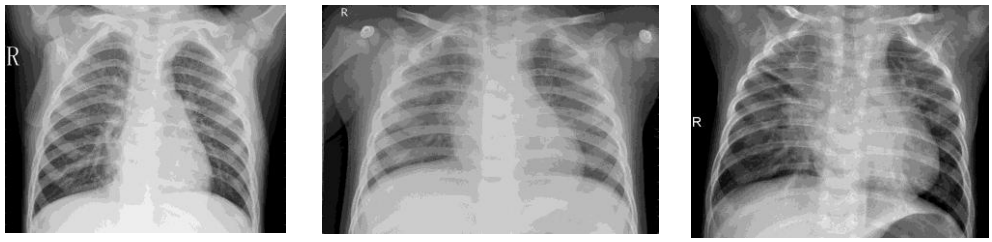

Figure 2: Chest X-Ray images from chidrem with pneumonia.

3. Preprocessing Step

First it was noted that some images on the dataset has only one channel (grayscale) and were transformed into three channels (RGB). Then, a histogram was plotted to each part of the dataset (train/validation/test) to check the balancing between the two classes (figure 3). These histograms show a high unbalanced data for the training and test and a low amount of data for validation. To solve these problems, the train and validation passed through a data augmentation process to raise the number of samples to 5000 to each class. This process was held doing random rotations with angle between -45º and 45º, random contrast adjustment (with lower limit between 2% and 10% and higher limit between 90% and 98%), histogram equalization and contrast limited adaptive histogram equalization (an algorithm for local contrast contrast enhancement, that uses histograms computed over different tile regions of the image , it was used a random clipping limit between 0 and 1). For a more variation, after the first augmentation iteration on the original data, the process was applied on the data that was generated previously until the number of samples reached 5000 for each class.

4. Experiments and Discussion

4.1 Fine Tune using Inception-V3

In a first approach, a fine tune model using the convolutional neural network (CNN) Inception-V3 was applied on the data. The CNN was loaded without its tops layers and a global spatial average pooling and a logistic layers were added. Then, a pre-training process was used for adjustment of weights of the new layers added for 50 epochs with Gradient Descent (SGD) as optimizer using a initial learning rate of 0.001 that was reduced by a factor of 0.1 if no improvement on the validation loss was observed in 2 epochs. The accuracy and the loss reached in this step can be seen at figure 4.

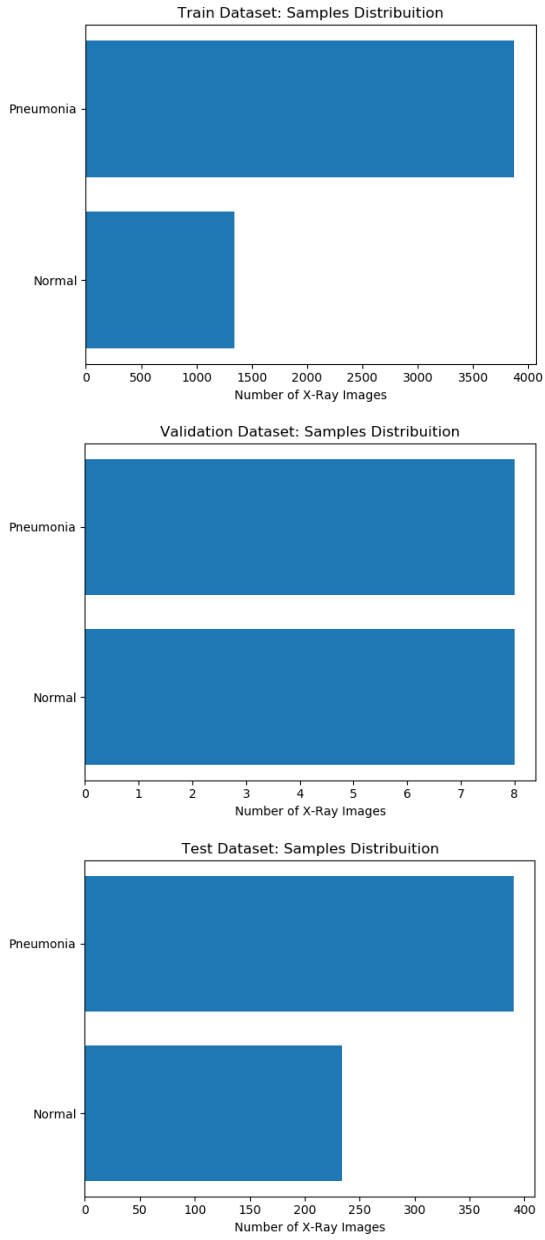

Figure 3: Data distribuition between the classes of each dataset.

After the pre-training step, the first 172 first layers were freezed and the rest were trained for more 50 epochs with the same conditions of the SGD used before. The results are pointed at figure 5. The better accuracy reached at the validation data was 90.31% at the forth epoch. The weights calculated at this epoch were used to predict the

outcomes of the test data, hitting an accuracy of 88.30%. A normalized confusion matrix was also plotted (figure 6) to analyze the performance of the model to predict the classes of the test dataset. As can be observed, although the model presents a good accuracy, it failed to predict the normal class.

4.2 Transfer Learning using Inception-V3 and Support Vector Machines

Transfer Learning was applied to the data to try to predict the outcome of the X-Ray images: 2048 features were extracted using the Inception-V3 from the "pool_3" layer and then used as input of a Support Vector Machines (SVM) model. For this task, a radial basis function (RBF) was used as kernel, kernel coefficient $\gamma = 1/2048$, decision function as one-vs-one (OVO) and penalty parameter $C = 1$. The accuracy obtained was 88.18% at the validation and 87.18% at the test data, figure 7 shows its normalized confusion matrix.

This model exhibited a better performance than the previous one, especially at predicting the pneumonia class.

4.3 Transfer Learning using Inception-V3 and Random Forests (RF)

Also, a transfer learning associated with Random Forests was tested. To decide the number of trees, a plot of the out-of-bag error in function of the number of trees was generated for the case of the maximum number of features used to calculated the trees being the square root of the total number of features available (2048) and for the case of being the $\log_2$ of the same number. The plot is showed at figure 8.

With a number of trees of 250, it seemed that the out-of-bag error has stabilized, also it was chosen the squared root function as maximum number of features of each tree. The accuracy obtained was 84.40% at the validation data and 86.70% at the test.

The figure 9 shows the normalized confusion matrix. The model presented almost the same performance of the transfer learning using Support Vector Machines.

Figure 4: Accuracies and losses reached in the pre-training step (fine tune).

Figure 5: Accuracies and losses reached in the training (fine tune).

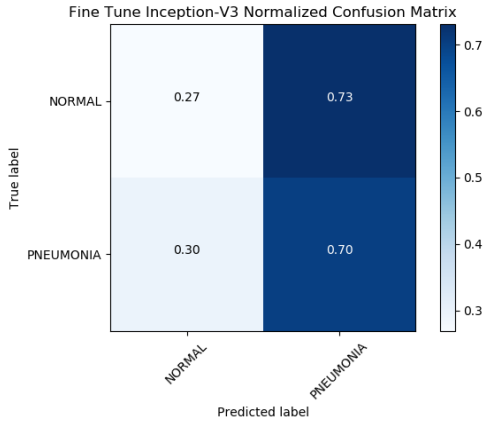

Figure 6: Normalized Confusion Matrix of the Fine Tune using Inception-V3.

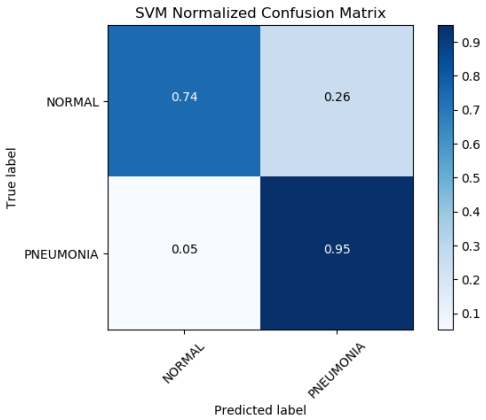

Figure 7: Normalized Confusion Matrix of the Transfer Learning with Inception-V3 and Support Vector Machines.

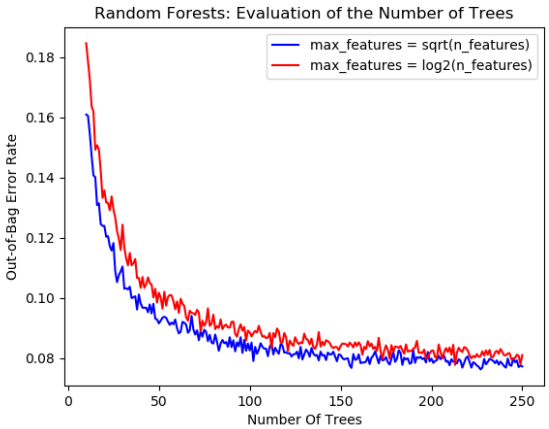

Figure 8: Evaluation of the number of trees.

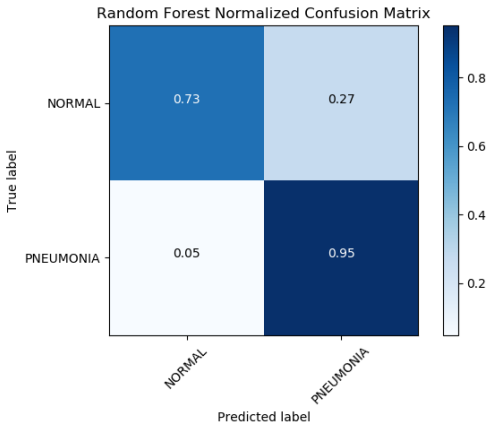

Figure 7: Normalized Confusion Matrix of the Transfer Learning with Inception-V3 and Random Forests.

5. Conclusion and Future Work

Dealing with unbalanced data is a common problem with medical data. In this project it was noted the importance of data augmentation to help to solve this problem. The firsts approaches,

when no augmented data was being used, the models were given an accuracy of about 50% on the validation data. After the augmentation, it was observed a great rise on the accuracy. Unfortunately, looking only at the accuracy of the model isn't a good manner to analyze the performance of a model, especially when the data is unbalanced. The models presented a much better prediction on the pneumonia class than the normal, suggesting that the augmentation on the normal class was not enough. If the model predicts that a person is with pneumonia, it has a good chance to be right, but if it predicts it as normal, it hasn't a very good accuracy at this class to tell if the prediction is right.

For future works, we pretend to work only with histogram equalized pictures and not use the original ones. The first ones have a better quality to detect the inflammations of the disease and working only with them can direct to better results. Besides that, a better augmentation is needed, not only equalizing the number of samples of each class like was done at this work, but introducing more variation on the data generated.

References

[1]CORTES, C.; VAPNIK, V. Support-Vector Networks. *Machine Learning*, v. 20, p. 273-297, 1995.

[2]HO, T. K. Random Decision Forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition*,Montreal, QC, p. 278-282, August 1995.

[3]SZEGEDY, C. et al. Rethinking the Inception Architecture for Computer Vision, 2015. arXiv:1512.00567v1

[4]Kermany, Daniel; Zhang, Kang; Goldbaum, Michael (2018), "Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification", Mendeley Data, v2.