Identifying Pneumonia from Chest X-Ray Images using Deep Learning

Faber Silitonga	Fikhri Masri	Dimas Anom Priyayi
faber.yosua@gmail.com	masrifikhri@gmail.com	priyayidimas@gmail.com
abrosua	fikhrimasri	priyayidimas

This project is available on github and the dataset is available on Kaggle.

Dataset Metadata

Citation	http://www.cell.com/cell/fulltext/S0092-8674(18)30154-5	
Usage Information	License	CC BY 4.0
	Visibility	Public
Maintainer	Dataset Owner	Paul Mooney
Updates	Last Updated	2018-03-25 (Version 2)
	Created Date	2018-03-22
Dataset	Train	Normal: 1341 images; Pneumonia: 3875 images
	Validation	Normal: 8 images; Pneumonia: 8 images
	Test	Normal: 234 images; Pneumonia: 390 images

Problem Framing → Binary Classification

By using image classification to determine whether someone has pneumonia or not from the x-ray image of the lungs, we can probably speed the diagnosis process, thus helping the medical center works efficiently.

The output labels would be “normal” or “pneumonia”.

The success rate of this model depends on precision and recall rate of the model. The model is presumed success if precision rate is over 90% and recall rate is over 95%.

Preparation

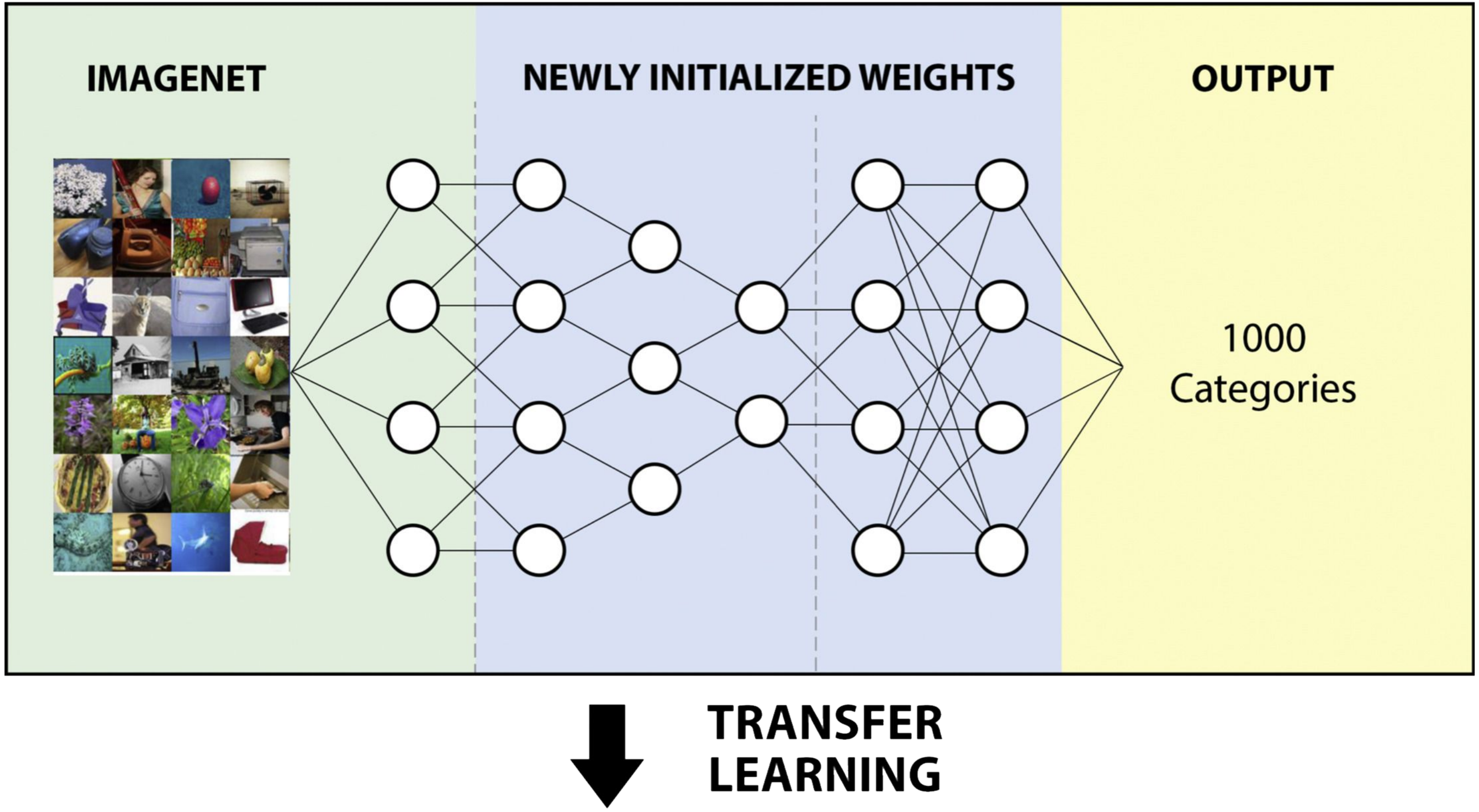
Hyperparameters to be tuned:

1. Network architecture
2. Optimizer method
   - Adam or RMSProp
3. Initial learning rate value
   - 1e-4, 5e-5 or 1e-5
4. LR Scheduler (ReduceLROnPlateau) factor value
   - 0.2 or 0.5

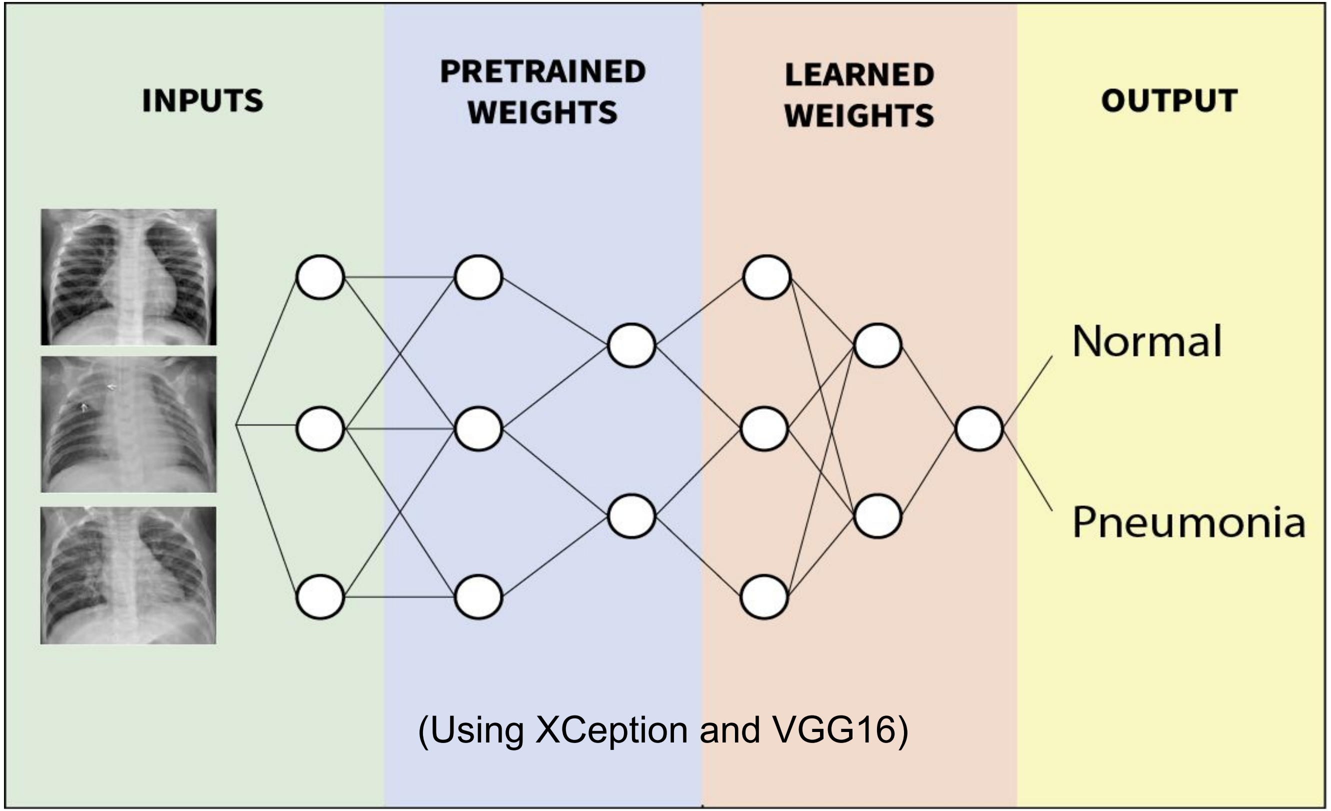
Data Preprocessing (Image augmentation):

1. Pixel value rescaling
2. Rotation : -5 to 5 degree
3. Width shift : -0.1 to 0.1
4. Height shift : -0.05 to 0.05
5. Shear : -0.1 to 0.1
6. Zoom : -0.15 to 0.15
7. Horizontal flipping

Technique → Neural Network, Deep Transfer learning



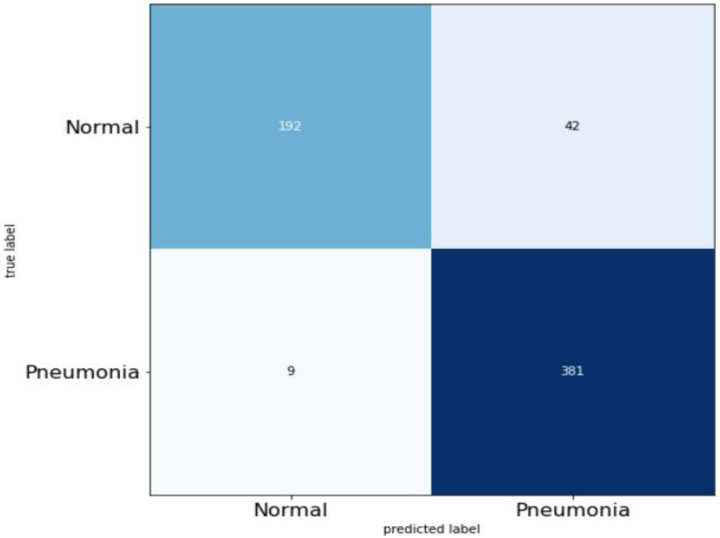
Technique → Neural Network, Deep Transfer learning



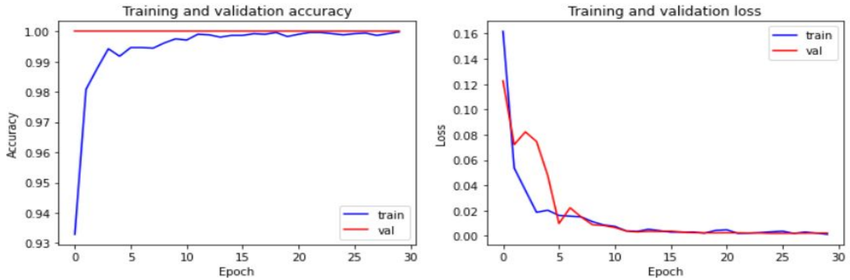
Result

Xception Model Result

Confusion Matrix Data Test



Training & Validation Loss



In exception model we get confusion matrix data test

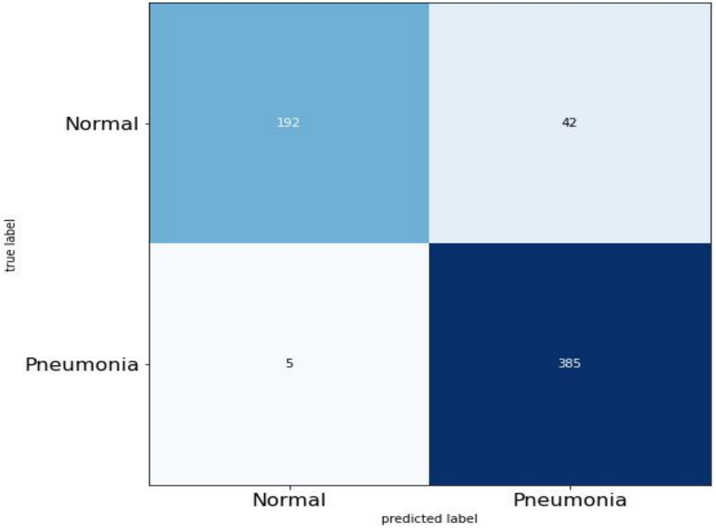
Precision : 90.07%

Recall : 97.69%

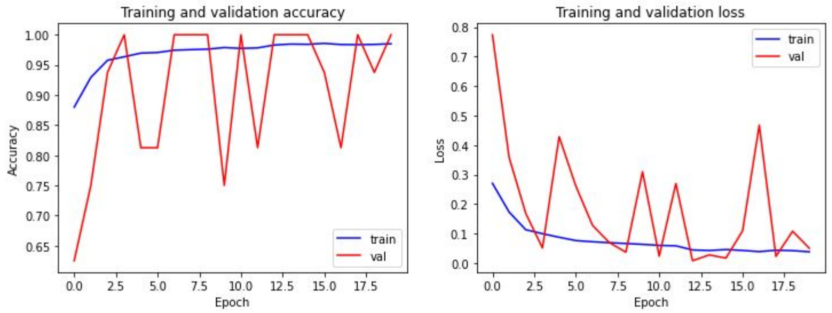
F1 Score : 93.73%

VGG16 Model Result

Confusion Matrix Data Test



Training & Validation Loss



In exception model we get confusion matrix data test

Precision : 90.16%

Recall : 98.72%

F1 Score : 94.25%

Conclusion

Based on the confusion matrix, VGG16 gives better result than Xception (e.g., VGG f1-score is 94.25% while Xception only gets 93.73%). Despite having a better result, this model might be quite unstable (based on the fluctuated loss graph) which may lead to overfitting when applied to real world cases. Because, deep learning results may vary due to several aspects (e.g., weights initialization, etc) thus robust design is preferable.

*Note: the network configuration is shown in the repository.