Ensemble of LeNet-5, AlexNet and VGG-16 for CoronaHack Chest X-ray dataset

Lazar Jelic

November 14, 2020

1 Dataset

Preparation. Dataset preparation phase involves basic formatting of pandas dataframes which hold paths to input images and corresponding output classes (*Normal*, *Pneumonia*). Dataset comes with validation split of 89/11 for 5910 images in total.

It can be observed that the data is imbalanced, meaning that there is a 73% of images classified as *Pneumonia* and 27% of images classified as *Normal*. To create a model that generalizes well, despite feeding it with imbalanced data, usage of a weighted cost function is advised.

Another solution to this problem is a runtime data augmentation process which consists of applying random rotation and zoom effects on training images, while validation images need to stay the same. Normalization of images' RGB values is also done during this process by multiplying every pixel value with $\frac{1}{255}$.

2 Model

Architecture. Ensemble consists of LeNet-5 [3], AlexNet [2] and VGG-16 [4] convolutional neural networks with their standard architectures. Only changes done to these architectures include the following:

- Addition of a single-unit fully connected output layer
- Usage of sigmoid activation function in the output layer
- Usage of binary cross-entropy loss function in the output layer
- Addition of a dropout layer at the end of pre-trained VGG-16
- Usage of SELU function at the end of AlexNet

An odd number of models is composing this ensemble because a mode is calculated to determine a final classification outcome.

Hyperparameters. Every model in the ensemble is trained on 3-channel 64×64 images from the whole dataset in batches of size 32 for 16 epochs. Using these hyperparameters' values, *2 hours and 50 minutes* is needed for the ensemble training phase.

Optimizer. Adaptive Moment Estimation (ADAM) [1] is the optimization method of choice because it incorporates a learning rate decay for regularization purposes.

Weighted loss function. As the above-mentioned imbalanced data, i.e., the skewed distribution of the classes, can prevent a model to generalize, although it could have high accuracy, loss function can be changed to account for such apriori biases. The idea is to assign different weights to both the majority and minority classes while calculating a loss function.

For example, instances of a minority class will cause larger weights to be assigned to losses, which will ultimately cause an optimizer to focus on minorities.

Metrics. Accuracy, precision, recall and F1-score are used to measure models' performances, while the performance of the ensemble is measured only by accuracy which is 92%. Classification reports, as well as training charts and confusion matrices, are shown below.

	precision	recall	f1-score	support
Normal	0.89	0.89	0.89	234
Pneumonia	0.94	0.93	0.93	390
accuracy			0.92	624
macro avg	0.91	0.91	0.91	624
weighted avg	0.92	0.92	0.92	624

Table 1: LeNet-5 classification report

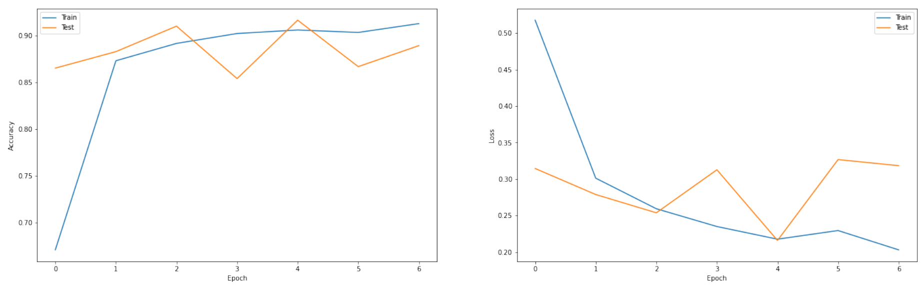

Figure 1: LeNet-5 training charts

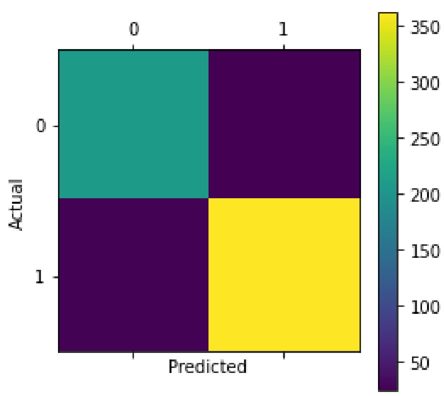

Figure 2: LeNet-5 confusion matrix

	precision	recall	f1-score	support
Normal	0.87	0.88	0.87	234
Pneumonia	0.93	0.92	0.92	390
accuracy			0.90	624
macro avg	0.90	0.90	0.90	624
weighted avg	0.90	0.90	0.90	624

Table 2: AlexNet classification report

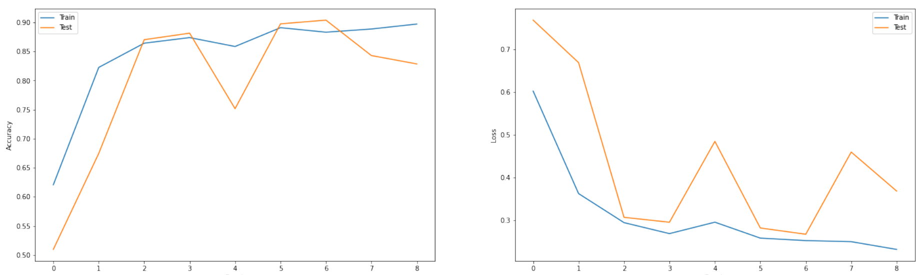

Figure 3: AlexNet training charts

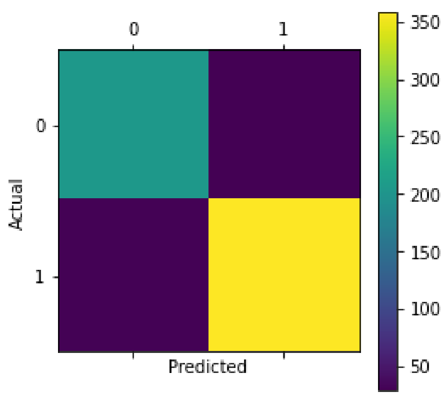

Figure 4: AlexNet confusion matrix

	precision	recall	f1-score	support
Normal	0.84	0.86	0.85	234
Pneumonia	0.91	0.90	0.91	390
accuracy			0.88	624
macro avg	0.88	0.88	0.88	624
weighted avg	0.89	0.88	0.88	624

Table 3: VGG-16 classification report

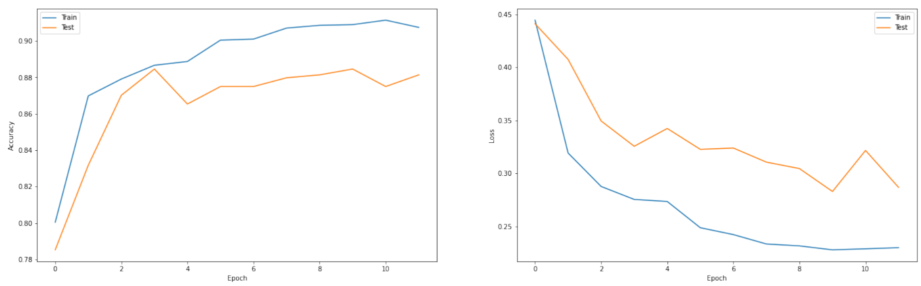

Figure 5: VGG-16 training charts

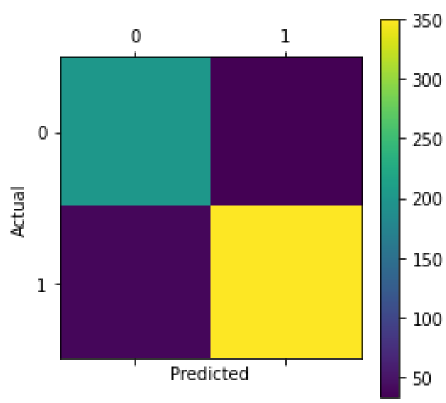

Figure 6: VGG-16 confusion matrix

3 Resources

Dataset: `https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset`

Code: `https://github.com/jelic98/raf_du/tree/main/homework_1`

References

[1] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

[2] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, 2012.

[3] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998.

[4] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2015.