

Computer Vision, Assignment - 2

Tulsi Jain

1 Introduction

German Traffic Sign Data consists of 34500 training images, 12631 test images and 3870 images for validation data (dev set). Images are to be classified in 43 classes.

2 Class Imbalance

Class imbalance is major problem in deep learning model as model tends to learn biases in dataset instead of right learning representation. German Traffic Sign dataset is highly imbalance. Class with highest number of sample has 2160 training sample and class with lowest number of sample has 120 training sample. We resolve this problem by producing more examples of under represented classes by apply rotation and translation in both direction. This proved helpful in increasing the accuracy[3].

3 Data Pre-processing

Size of training sample is not uniform so data pre-processing was required. Every images is convert to 32x32x3(RGB) size. And data is also normalized across pixels values. Local contrast normalization is also applied as mention in 7. Affine transformation is also applied a mentioned in 7 but does not seem to improve may be because of different architecture.

4 Data Augmentation

After resolving class imbalance we have 87597 training sample but this was sufficient for the given problem. But we can't apply horizontal/vertical flip and rotation as it can change the sample class. To make sure that transformation does not change the sample class rotation degree, translation and shear are kept small. Shear, rotation and translation is applied three times making the data four fold. Finally, we have 350388 number of training samples.

5 Model Architecture

Couple of model architectures starting from three convolution layers to 20 convolution layers with different value of hyper parameters[3]. DenseNet introduced in [2] seems to work provide better result[3]. Model has 100 layers, with batch normalization, Relu as activation function, dropout is added to avoid overfitting and at the end softmax is added to classify. With growth rate 32, Depth 100, Reduction=0.5 and Bottleneck set to True.

6 Fine tuning Hyper Parameters

6.0.1 Optimizer

Adam is used with amsgrad set to true[4][6].

6.1 Early Stopping

Validation set proved to be very useful to know when to stop the training.

6.2 Batch Size

Various batch is fine-tuned and 64 is observed to be worked best.

6.3 Weight Initialization

Weight initialization is done using Xavier's initialization[1].

6.4 Learning Rate

Initial learning rate is 0.01 and reduced exponentially per epoch.

7 Result

Final accuracy reported on public leader board is 98.590 percent. This is link to the trained model.

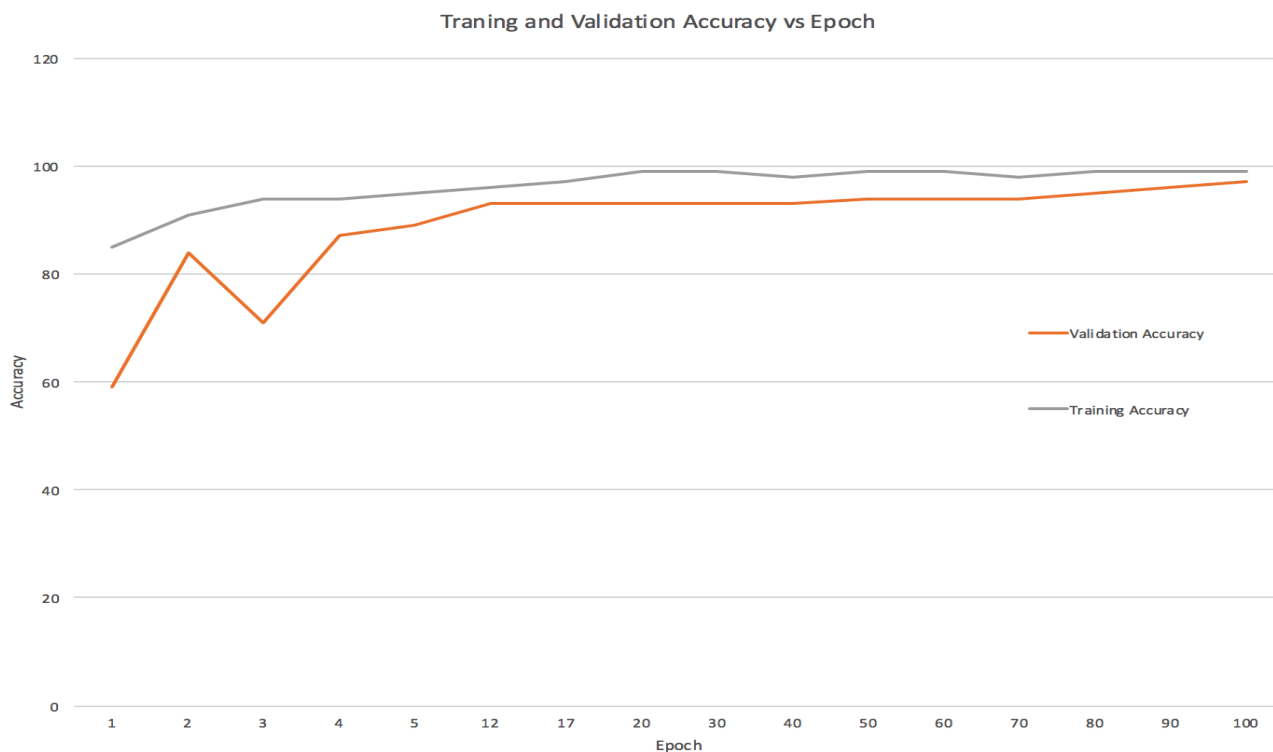


Figure 1: Accuracy Vs Epoch

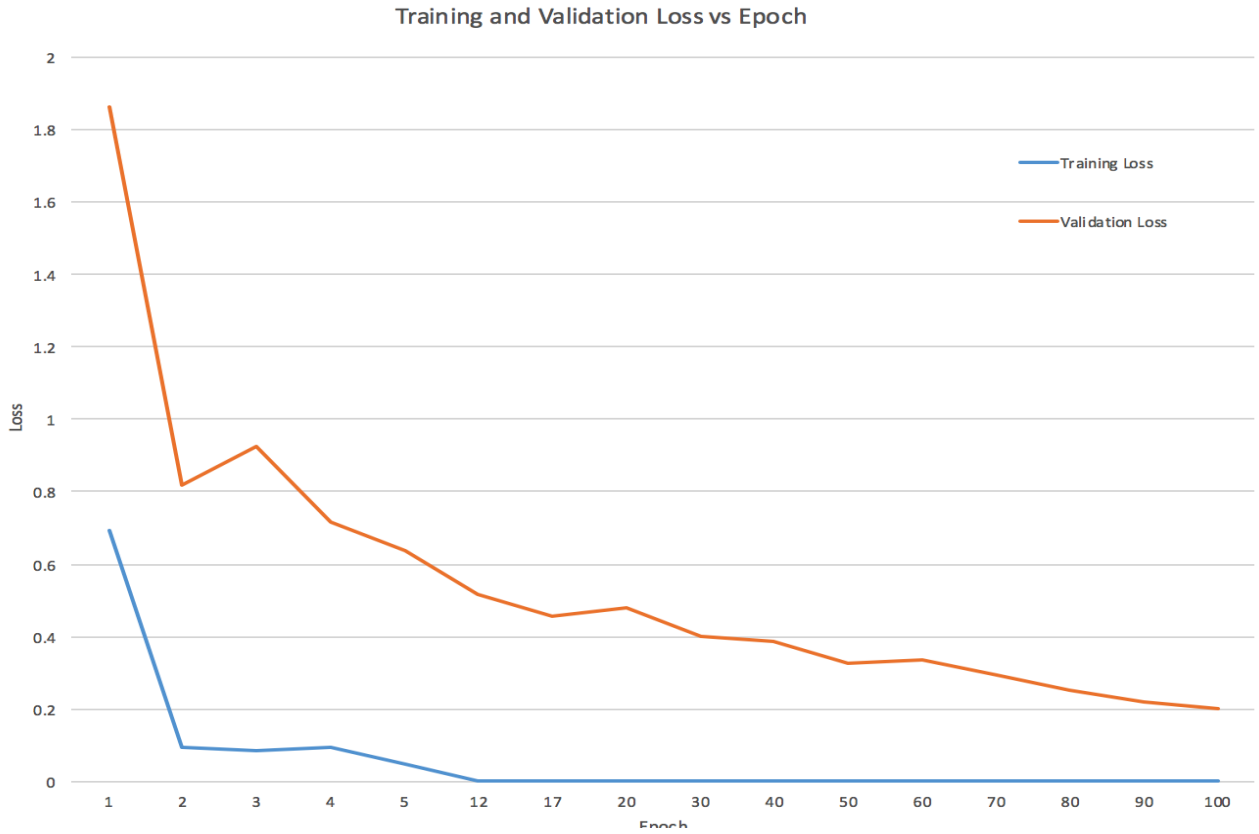


Figure 2: Loss Vs Epoch

References

- [1] Glorot, X. and Bengio, Y. [*Understanding the difficulty of training deep feedforward neural networks*]. In Proceedings of AISTATS 2010, volume 9, pp. 249256, May 2010.
- [2] Gao Huang, Zhuang Liu, Kilian Q. Weinberger, and Laurens van der Maaten. [*Densely connected convolutional networks*]. arXiv preprint arXiv:1608.06993, 2016b.
- [3] Yadav Vivek. [*Dealing with unbalanced data: Generating additional data by jittering the original image*]. <https://medium.com/@vivek.yadav/dealing-with-unbalanced-data-generating-additional-data-by-jittering-the-original-image-7497fe2119c3>. Nov 3, 2016.
- [4] Diederik P. Kingma, Jimmy Ba [*Adam: A Method for Stochastic Optimization*]. arXiv preprint arXiv:1412.6980 [cs.LG], 2017.
- [5] Amos, Brandon and Kolter, J. Zico [*A PyTorch Implementation of DenseNet*]. <https://github.com/bamos/densenet.pytorch>, 10-11-2018.
- [6] S. Ruder [*An overview of gradient descent optimization algorithms*]. <http://sebastianruder.com/optimizing-gradient-descent/index.html#momentum>, 10-15-2018.
- [7] Dan Ciresan, Ueli Meier, Jonathan Masci and Jurgen Schmidhuber [*Multi-Column Deep Neural Network for Traffic Sign Classification*]. 2012.