Development Of Image Processing Models Using NVIDIA DIGITS For Classification Of CARS Dataset

Pratyaksh Rao
MSc Data Science
University Of Surrey

1. Introduction

The image classification using Convolutional neural networks was carried out through NVIDIA DIGITS. CARS dataset was set as the base classification data from which three Data sets with varying augments were created:

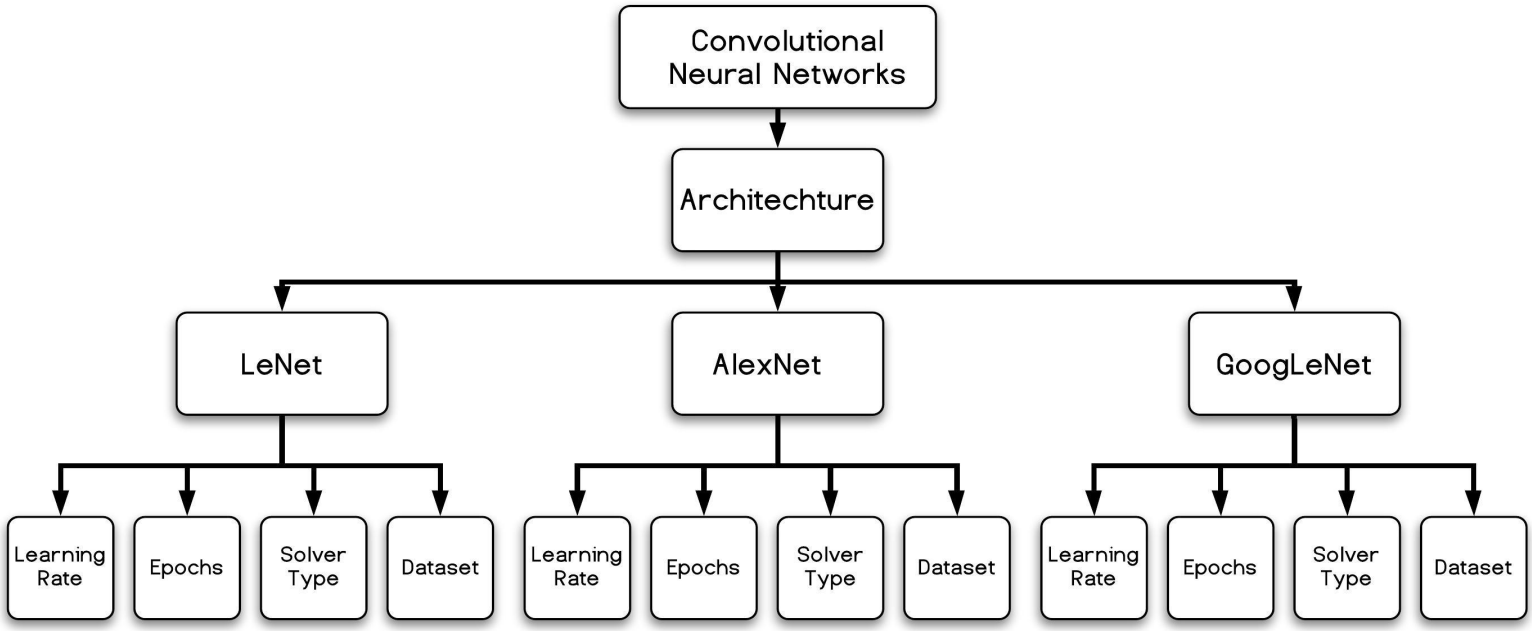
1. Base set
2. Greyscale set
3. Rotated set
4. Split Change set

which were tested using three CNN architectures:

1. GoogLeNet
2. AlexNet
3. LeNet

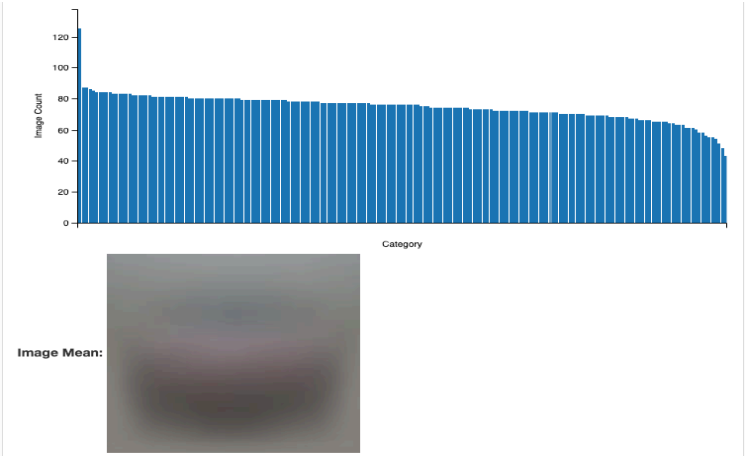
With varying Parameters values of namely three metaparameters:

1. Learning Rate with values 0.1, 0.01, 0.001
2. Epochs with values 30, 60, 90
3. Solver Type being Stochastic Gradient Descent (SGD), Nesterov's Accelerated Gradient (NAG)

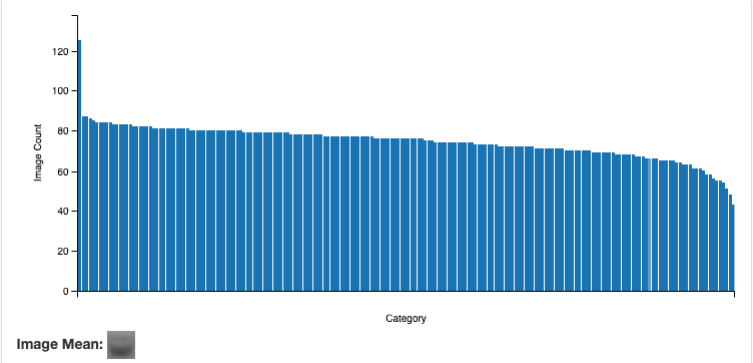


2. Datasets

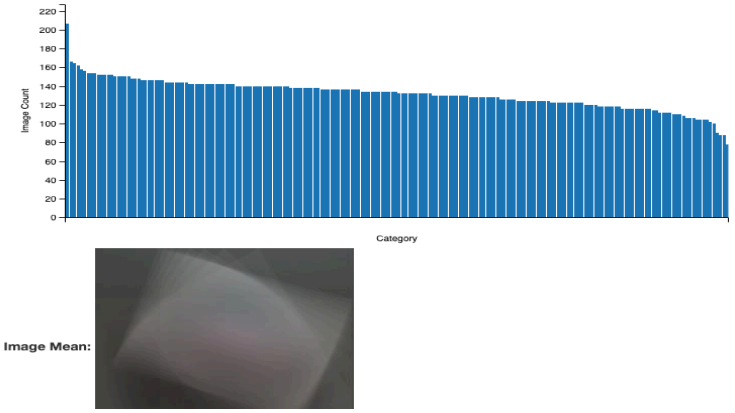
1.Base Set
- a. 90:9:1 Split
- b. Normal Image Setting
- c. 256 x 256 Size



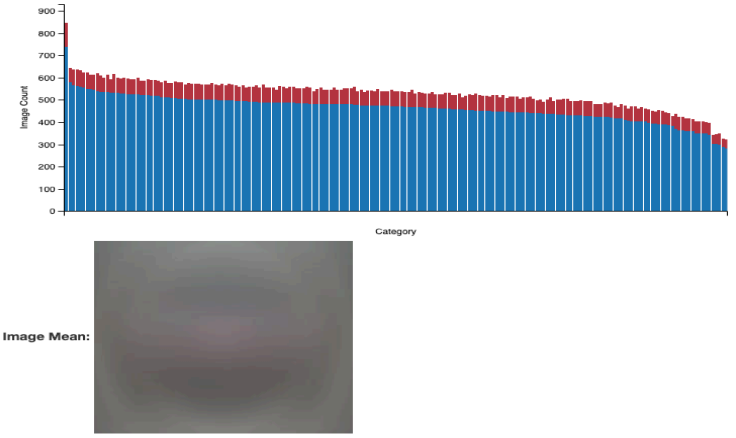
2.Greyscale Set
- a. 90:9:1 Split
- b. Greyscale Image Setting
- c. 28 x 28 Size



3.Rotated Set
- a. 80:10:10 Split
- b. Rotated Image Setting
- c. 256 x 256 Size

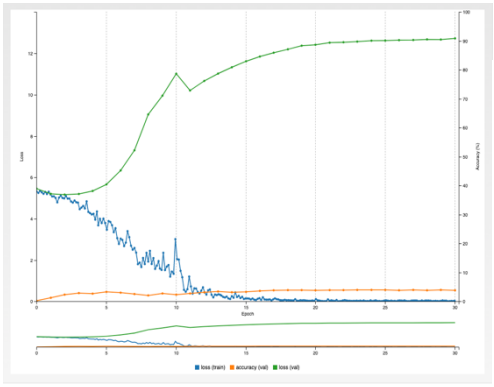


4.Split Change Set
- a. 60:20:20 Split
- b. Normal Image Setting
- c. 256 x 256 Size



3. LeNet Architecture Models

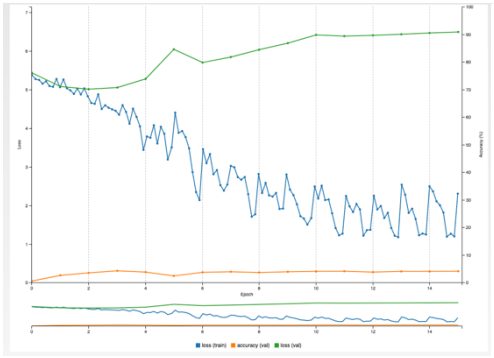
1. Model 1
    a. 30 Epochs
    b. 0.01 Learning Rate
    c. Greyscale set
    d. SGD Solver Type
    e. Accuracy (val) = 3.87%
    f. Loss (val) = ~12



Predictions

Aston Martin Virage Coupe 2012	43.61%
Dodge Challenger SRT8 2011	24.96%
Volkswagen Golf Hatchback 2012	13.68%
Ferrari 458 Italia Coupe 2012	8.1%
Lamborghini Aventador Coupe 2012	6.11%

2. Model 2
    a. 15 Epochs
    b. 0.01 Learning Rate
    c. Greyscale set
    d. SGD Solver Type
    e. Accuracy (val) = 4.14%
    f. Loss (val) = ~6

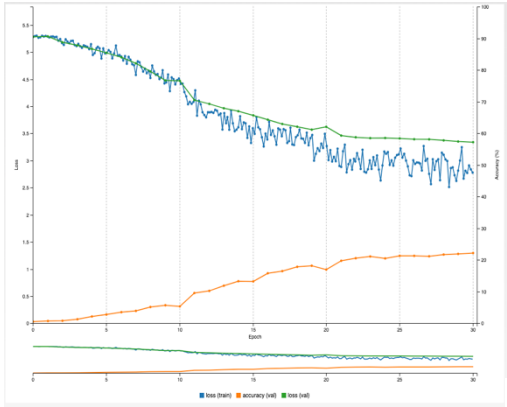


Predictions

Chevrolet Camaro Convertible 2012	7.24%
BMW 3 Series Wagon 2012	6.51%
BMW X5 SUV 2007	5.42%
Chevrolet Traverse SUV 2012	4.66%
Lamborghini Aventador Coupe 2012	4.56%

4. AlexNet Architecture Models

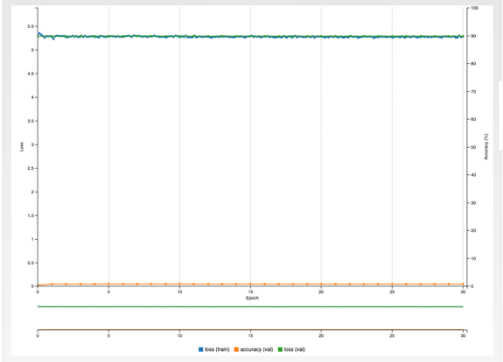
1. Model 1
    a. 30 Epochs
    b. 0.01 Learning Rate
    c. Base set
    d. SGD Solver Type
    e. Accuracy (val) = 24.66%
    f. Loss (val) = ~3.5



Audi TT RS Coupe 2012	32.18%

Bentley Continental Supersports Conv. Convertible 2012	10.78%

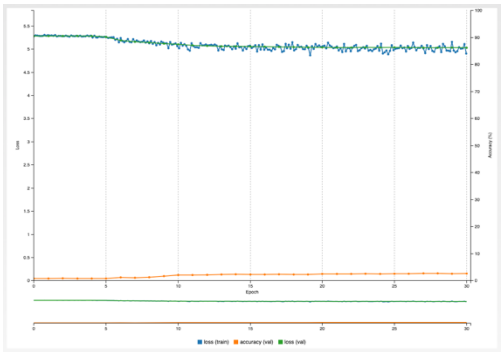
2. Model 2
    a. 30 Epochs
    b. 0.1 Learning Rate
    c. Base set
    d. SGD Solver Type
    e. Accuracy (val) = 0.74%
    f. Loss (val) = ~5.5



GMC Savana Van 2012	0.86%

GMC Savana Van 2012	0.86%

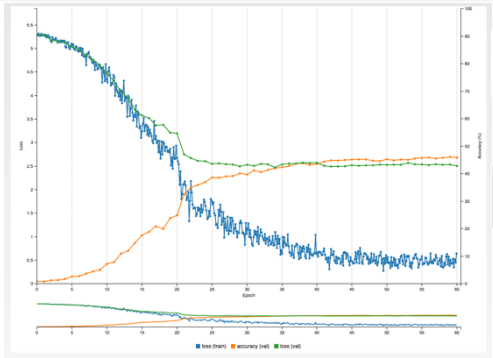
3. Model 3
    a. 30 Epochs
    b. 0.001 Learning Rate
    c. Base set
    d. SGD Solver Type
    e. Accuracy (val) = 2.58%
    f. Loss (val) = ~5.5



McLaren MP4-12C Coupe 2012 2.36%

GMC Savana Van 2012 1.48%

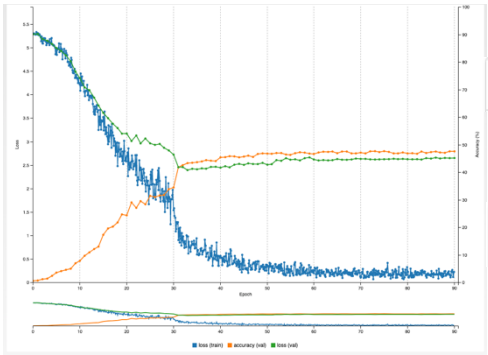
4. Model 4
    a. 60 Epochs
    b. 0.01 Learning Rate
    c. Base set
    d. SGD Solver Type
    e. Accuracy (val) = 48.77%
    f. Loss (val) = ~2.5



Audi TT RS Coupe 2012 99.96%

Dodge Challenger SRT8 2011 81.52%

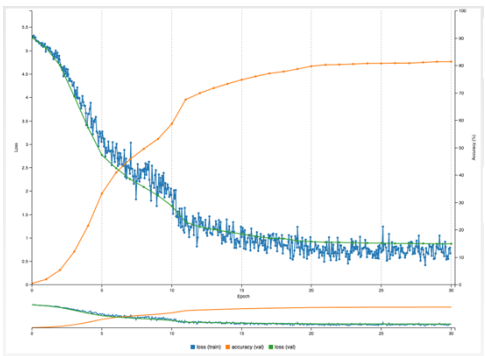
5. Model 5
    a. 90 Epochs
    b. 0.01 Learning Rate
    c. Base set
    d. SGD Solver Type
    e. Accuracy (val) = 47.62%
    f. Loss (val) = ~2.5



Audi TT RS Coupe 2012 87.28%

Dodge Challenger SRT8 2011 99.66%

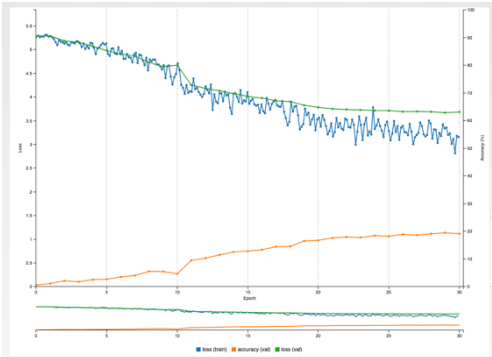
6. Model 6
    a. 30 Epochs
    b. 0.01 Learning Rate
    c. Split Change set
    d. SGD Solver Type
    e. Accuracy (val) = 81.43%
    f. Loss (val) = ~1



Audi TT RS Coupe 2012 90.67%

Dodge Challenger SRT8 2011 100.0%

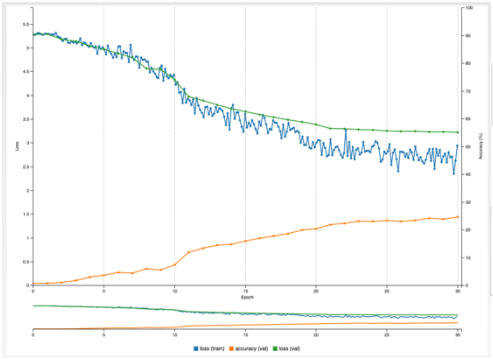
7. Model 7
    a. 30 Epochs
    b. 0.01 Learning Rate
    c Rotated set
    d. SGD Solver Type
    e. Accuracy (val) = 19.05%
    f. Loss (val) = ~4



Ford GT Coupe 2006 22.82%

Bugatti Veyron 16.4 Convertible 2009 7.41%

8. Model 8
    a. 30 Epochs
    b. 0.01 Learning Rate
    c. Base set
    d. NAG Solver Type
    e. Accuracy (val) = 24.66%
    f. Loss (val) = ~3.5

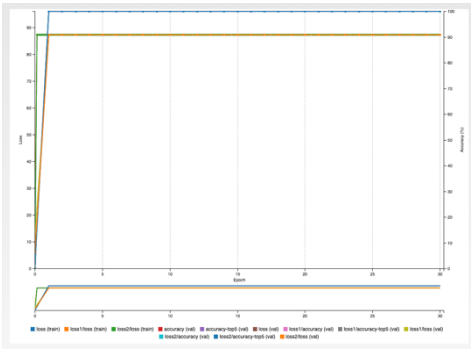


Audi TT RS Coupe 2012 82.04%

Bentley Continental Supersports Conv. Convertible 2012 11.07%

5. GoogLeNet Architecture Models

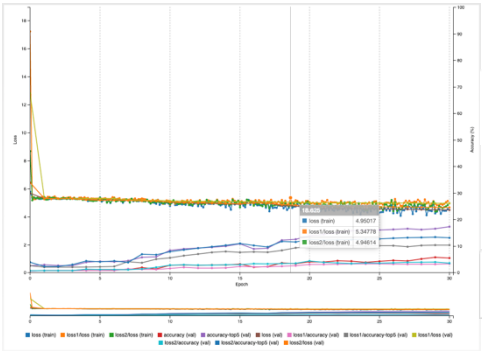
1. Model 1
    a. 30 Epochs
    b. 0.1 Learning Rate
    c. Base set
    d. SGD Solver Type
    e. Accuracy (val) = 0.50%
    f. Loss (val) = ~90



AM General Hummer SUV 2000 nan%

AM General Hummer SUV 2000 nan%

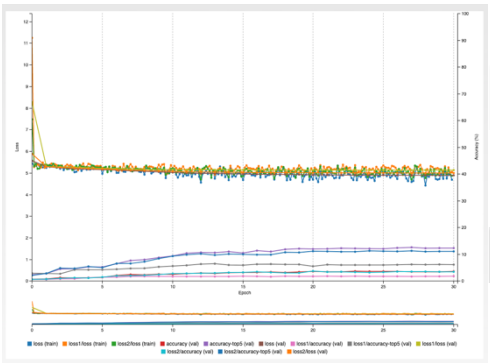
2. Model 2
    a. 30 Epochs
    b. 0.01 Learning Rate
    c. Base set
    d. SGD Solver Type
    e. Accuracy (val) = 5.08%
    f. Loss (val) = ~6



Lamborghini Aventador Coupe 2012 12.92%

Acura TL Sedan 2012 1.34%

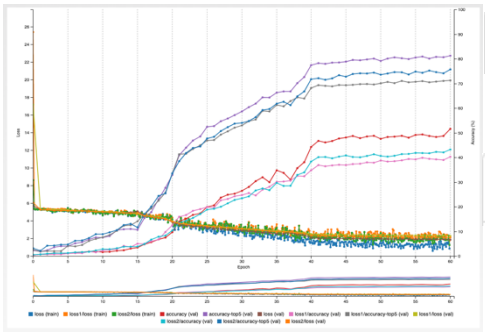
3. Model 3
    a. 30 Epochs
    b. 0.001 Learning Rate
    c. Base set
    d. SGD Solver Type
    e. Accuracy (val) = 3.57%
    f. Loss (val) = ~5



Ferrari California Convertible 2012 6.46%

Chevrolet Malibu Sedan 2007 0.98%

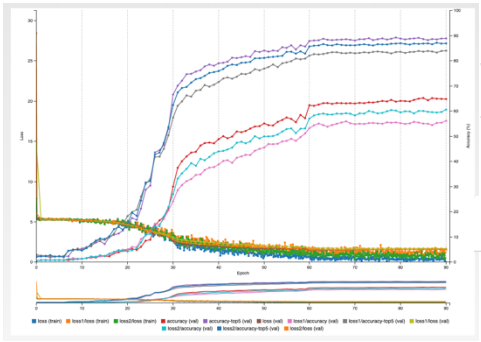
4. Model 4
    a. 60 Epochs
    b. 0.01 Learning Rate
    c. Base set
    d. SGD Solver Type
    e. Accuracy (val) = 44.15%
    f. Loss (val) = ~3



Audi TT RS Coupe 2012 98.76%

Dodge Challenger SRT8 2011 97.09%

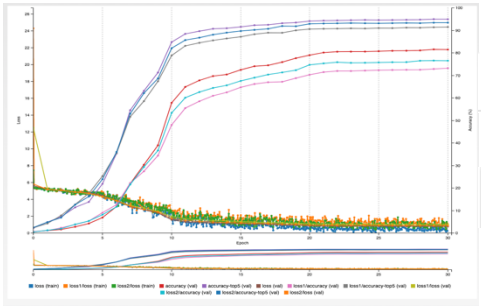
5. Model 5
    a. 90 Epochs
    b. 0.01 Learning Rate
    c. Base set
    d. SGD Solver Type
    e. Accuracy (val) = 64.74%
    f. Loss (val) = ~3



Audi TT RS Coupe 2012 100.0%

Dodge Challenger SRT8 2011 100.0%

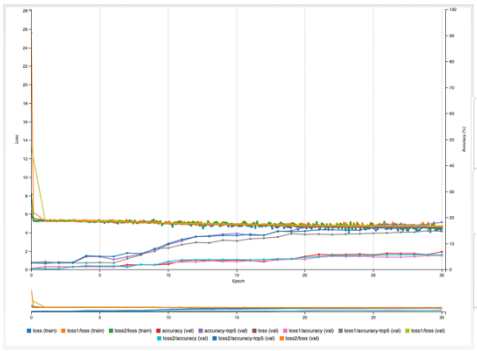
6. Model 6
    a. 30 Epochs
    b. 0.01 Learning Rate
    c. Split Change set
    d. SGD Solver Type
    e. Accuracy (val) = 81.45%
    f. Loss (val) = ~1



Audi TT RS Coupe 2012 76.28%

Dodge Challenger SRT8 2011 100.0%

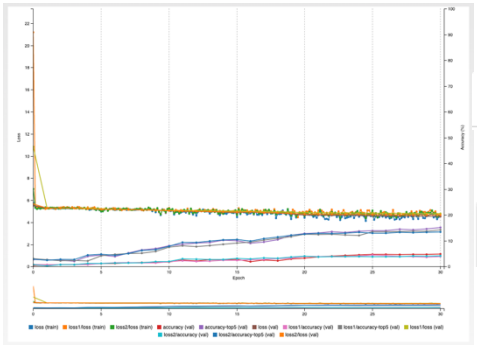
7. Model 7
    a. 30 Epochs
    b. 0.01 Learning Rate
    c. Rotated Set
    d. SGD Solver Type
    e. Accuracy (val) = 6.87%
    f. Loss (val) = ~6



Ferrari California Convertible 2012 12.71%

Jaguar XK XKR 2012 2.81%

8. Model 8
    a. 30 Epochs
    b. 0.01 Learning Rate
    c. Base set
    d. NAG Solver Type
    e. Accuracy (val) = 4.87%
    f. Loss (val) = ~6



Ferrari California Convertible 2012 16.37%

Acura TL Sedan 2012 1.78%

LeNet

Model Number	Epochs	Learning Rate	Solver Type	Dataset	Accuracy
model 1	30	0.0100000	SGD	Greyscale set	0.03870000
model 2	15	0.0100000	SGD	Greyscale set	0.04140000

AlexNet

Model Number	Epochs	Learning Rate	Solver Type	Dataset	Accuracy
model 1	30	0.1000000	SGD	Base set	0.007400
model 2	30	0.0100000	SGD	Base set	0.246600
model 3	30	0.0010000	SGD	Base set	0.025800
model 4	60	0.0100000	SGD	Base set	0.487700
model 5	90	0.0100000	SGD	Base set	0.476200
model 6	30	0.0100000	SGD	Split Change set	0.814300
model 7	30	0.0100000	SGD	Rotated set	0.190500
model 8	30	0.0100000	NAG	Base set	0.246600

GoogLeNet

Model Number	Epochs	Learning Rate	Solver Type	Dataset	Accuracy
model 1	30	0.1000000	SGD	Base set	0.005000
model 2	30	0.0100000	SGD	Base set	0.050800
model 3	30	0.0010000	SGD	Base set	0.035700
model 4	60	0.0100000	SGD	Base set	0.441500
model 5	90	0.0100000	SGD	Base set	0.647600
model 6	30	0.0100000	SGD	Split Change set	0.814500
model 7	30	0.0100000	SGD	Rotated set	0.068500
model 8	30	0.0100000	NAG	Base set	0.048700

6. Result Interpretation

a) LeNet :

-> A special dataset had to be produced for LeNet as it only accepts greyscale images with a size of 28 x 28
->The first base model was done with 30 epochs and 0.01 learning rate and according to the result graph it was seen that the training loss converges after 15 epochs indicating the rest 15 epochs were wasted and only consuming GPU power.
->The next model was then created with 15 epochs giving a similar accuracy to that with 30 epochs (4.14 ~ 3.87) saving 15 epochs worth of GPU computation

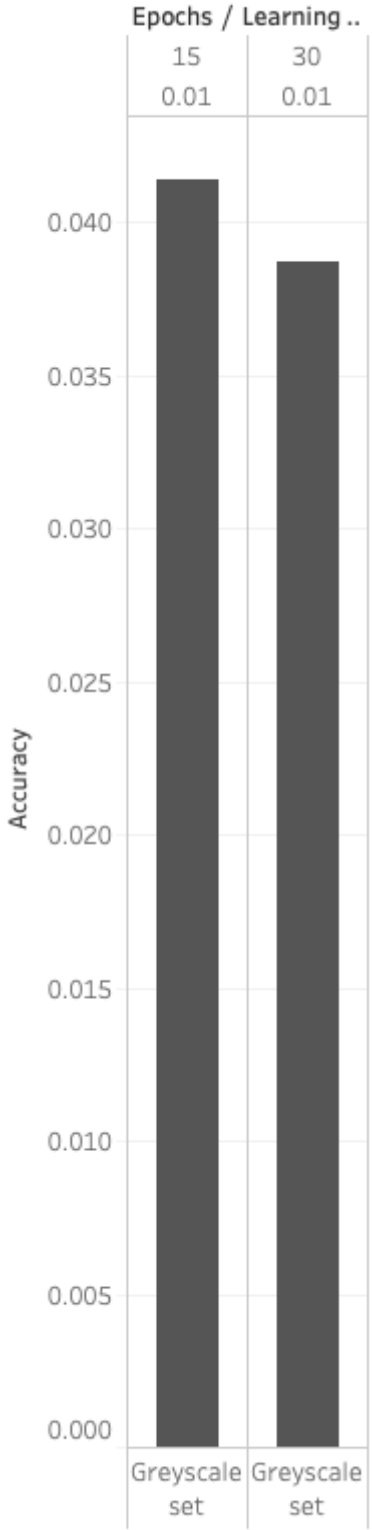
b) AlexNet :

->The first base model was done with 30 epochs and 0.01 learning rate which was then further modified based on learning rate. For High and low learning rate (0.1 ~ 0.001) similar to AlexNet, the accuracy was considerably low hence 0.01~0.05 was considered the best learning rate range for the model.
->Alterations were then made based on Epochs showing that higher the epoch the better the accuracy as the model will train on the data with higher iterations. Around 60~50 epochs was considered the best Epoch range for the model when under AlexNet architecture
->Changing the solver type to NAG gave a similar accuracy with that of SGD, but since SGD gave minutely higher accuracy it was considered the better choice.
->Training the model on a rotated dataset (where the images were rotated by an angle) gave low accuracy while training it on the base dataset but with 60:20:20 data split ratio have the highest accuracy recorded for AlexNet Model (81.43%) hence it was considered that the Split Change Set was the best set with the best split ratio.

c) GoogLeNet :

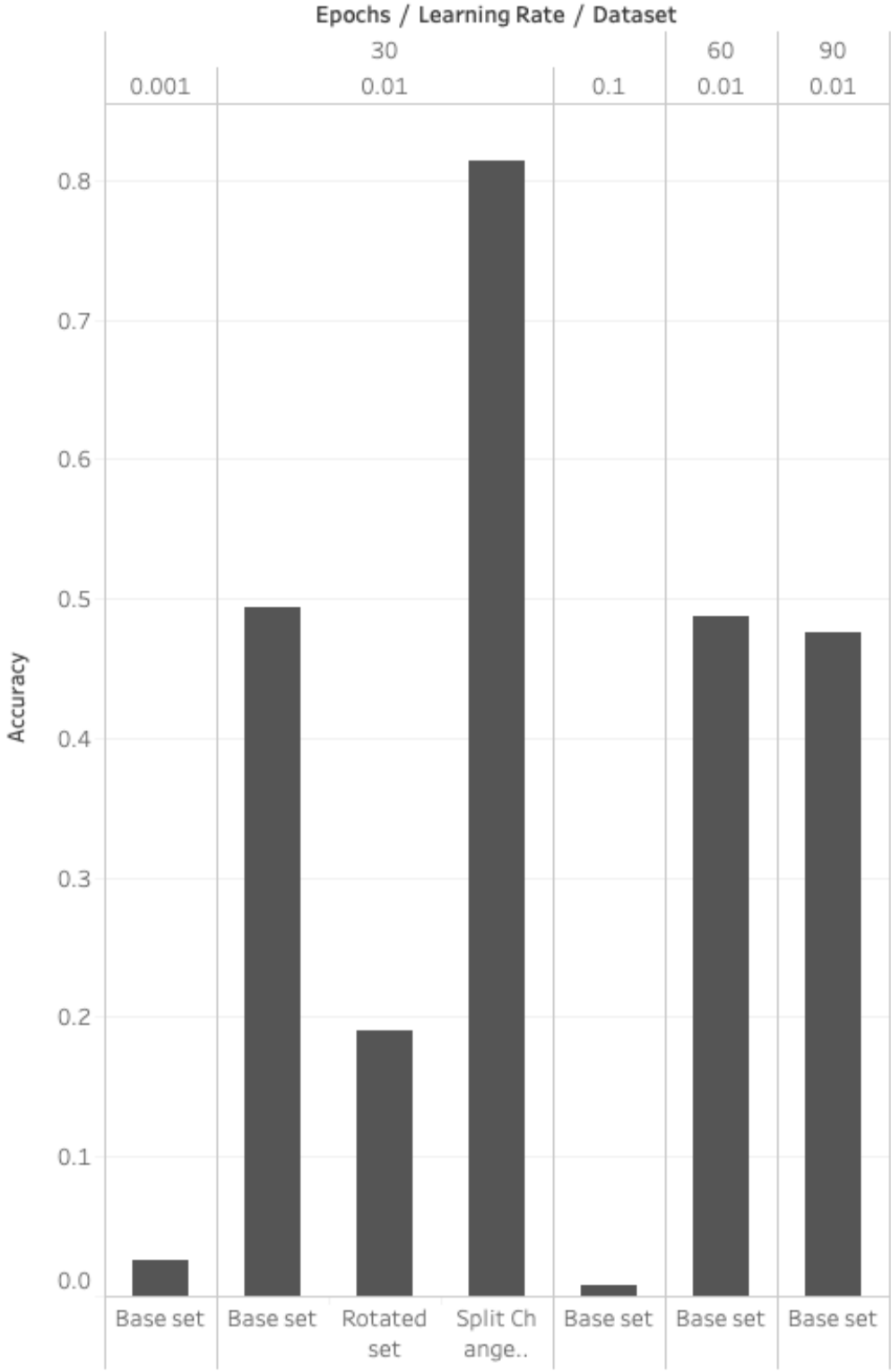
-> The first base model was done with 30 epochs and 0.01 learning rate which was then further modified based on learning rate. For High and low learning rate (0.1 ~ 0.001) the accuracy was considerably low hence 0.01~0.05 was considered the best learning rate range for the model.
-> Alterations were then made based on Epochs showing that higher the epoch the better the accuracy as the model will train on the data with higher iterations. The 90 Epoch Model gave a ~64% accuracy which was relatively high compared to its counterpart in AlexNet with same 90 epochs giving a accuracy of ~47%.Hence around 80~90 epochs was considered the best Epoch range for the model when under the GoogLeNet architecture.
-> Changing the solver type to NAG gave a similar accuracy with that of SGD, but was a little higher than the NAG Alexnet Model(4.8% of GoogLeNet to 2.4% of AlexNet) .Regardless since SGD gave minutely higher accuracy it was considered the better choice.
-> Training the model on a rotated dataset (where the images were rotated by an angle) gave low accuracy while training it on the base dataset but with 60:20:20 data split ratio have the highest accuracy recorded for GoogLeNet (81.45%) hence it was considered that the Split Change Set was the best set with the best split ratio. Also the highest Accuracies for both AlexNet and GoogLeNet was achieved when it was using the Split Change Set, while both accuracies had similar scores.

LeNet

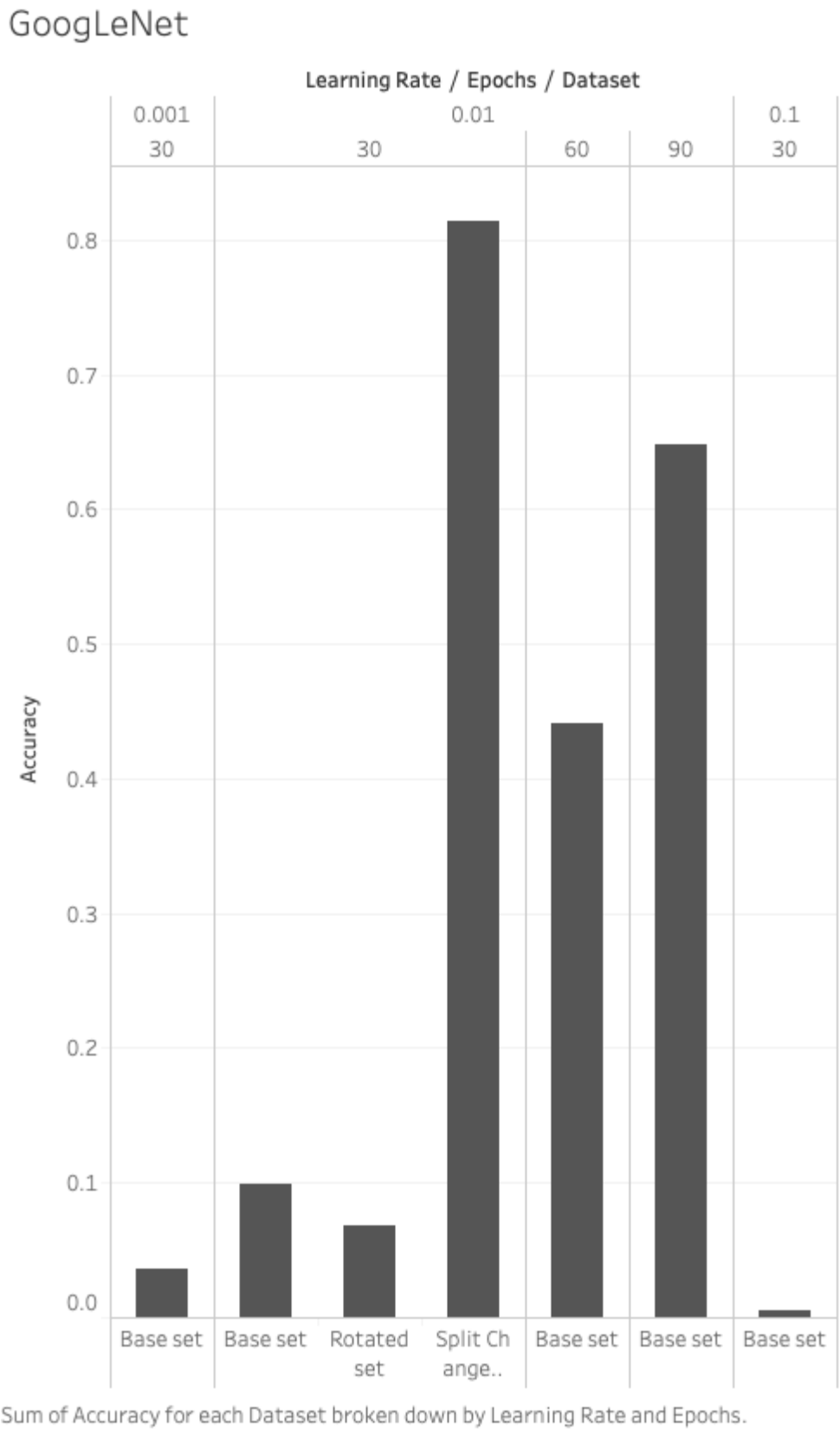


Sum of Accuracy for each Dataset broken down by Epochs and Learning Rate.

AlexNet



Sum of Accuracy for each Dataset broken down by Epochs and Learning Rate.



Sum of Accuracy for each Dataset broken down by Learning Rate and Epochs.

7. Future Scope

This model can be further expanded by customizing various factors and values such as adding/removing CNN layers for AlexNet and GoogLeNet. Custom values for Step size, Batch size And random seed might help with accuracy and better model design.