Comparison of Pre-Trained Models for DeepSteer

Mark Kocherovsky

16 December 2021

Introduction to DeepSteer

DeepSteer (DS) aims to enable a car to drive based on what it sees using Neural Networks.

Given a picture, can DS return the correct angle of the steering wheel based on the picture's content?

Based on work by Dr. CJ Chung, Dr. Giuseppe DeRose, Ian Timmis, and Nicholas Paul.

Dr. G. DeRose

Pretrained Models

Pretrained models are used for **Fine-Tuning**.

I tested eight pretrained networks with the same regression model on top.

There are several models for each family provided in TensorFlow, but I tested one of each.

DenseNet	Xception
EfficientNet	VGG16
MobileNet	VGG19
ResNet50	InceptionV3

Abadi, Agarwal, et. al. arXiv: 1603.04467
Huang, Liu, et. al. arXiv: 1608.06993
Howard, Zhu, et. al. arXiv: 1704.04861
He, Zhang, et. al. arXiv:1512.03385
Chollet. arXiv:1610.02357
Simonyan, Zisserman. arXiv:1409.1556
Szegedy, Vanhoucke, et. al. arXiv:1512.00567

Each pre-trained model was fine-tuned with the top 20% of layers unfrozen.

Modified version of
NEW_Steering_Regession_v3_Tamarack_Only.ipynb
by Dr. G. DeRose

```
if model_name == "VGG16":
  base_model = VGG16(weights="imagenet", include_top=False, input_shape=data_dict["x_train"][1].shape)
elif model_name == "VGG19":
  base_model = VGG19(weights="imagenet", include_top=False, input_shape=data_dict["x_train"][1].shape)
elif model_name == "DenseNet":
  base_model = DenseNet201(weights="imagenet", include_top = False, input_shape=data_dict["x_train"][1].shape)
elif model_name == "EfficientNet":
  base_model = EfficientNetB7(weights="imagenet", include_top = False, input_shape=data_dict["x_train"][1].shape)
elif model_name == "MobileNet":
  base_model = MobileNetV3Large(weights="imagenet", include_top = False, input_shape=data_dict["x_train"][1].shape)
elif model_name == "ResNet":
  base_model = ResNet50(weights="imagenet", include_top = False, input_shape=data_dict["x_train"][1].shape)
elif model_name == "Xception":
  base_model = Xception(weights="imagenet", include_top = False, input_shape=data_dict["x_train"][1].shape)
else:
  base_model = InceptionV3(weights="imagenet", include_top=False, input_shape=data_dict["x_train"][1].shape)
```

Data

5487 training, validation, and preliminary testing images from the LTU ACTor driving down a road.

1544 Test Images from the ACTor.

Data taken either by Dr. DeRose or with his assistance.

Each image is saved with the steering wheel angle in Radians as part of the filename.

Each image is also resized (By G. DeRose) to 320x240 Pixels.

Dr. G. DeRose

219_-0.741765.jpg

Indicates -0.74 radians (turning right)

Dr. CJ Chung

Comparison of Test Evaluation

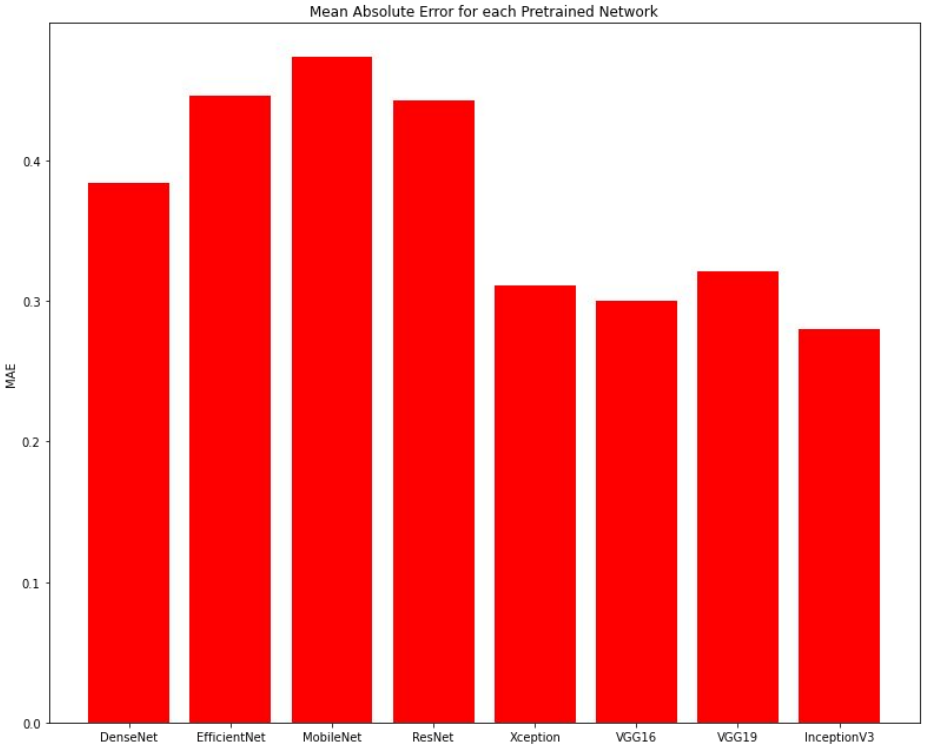
Each model was tested on the same test dataset.

DeepSteer_PT_Comparison.ipynb

InceptionV3 had the lowest MAE (0.2799 rad)

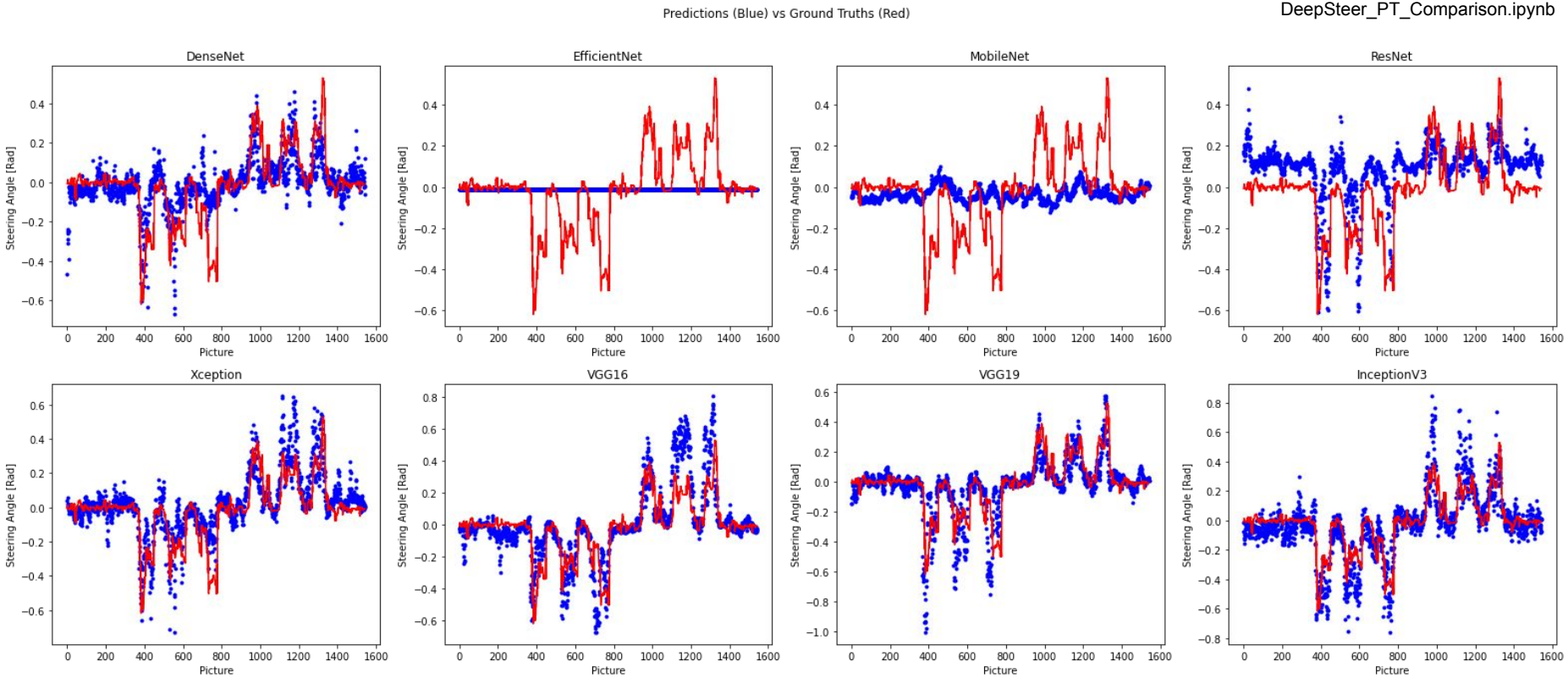
MobileNet had the highest MAE (0.4455 rad)

MobileNet's high error is likely because of its architecture: it is designed for mobile devices and applications.



Predictions vs Ground Truths

DeepSteer_PT_Comparison.ipynb



Prediction Analysis

EfficientNet, MobileNet, and ResNet perform poorly.

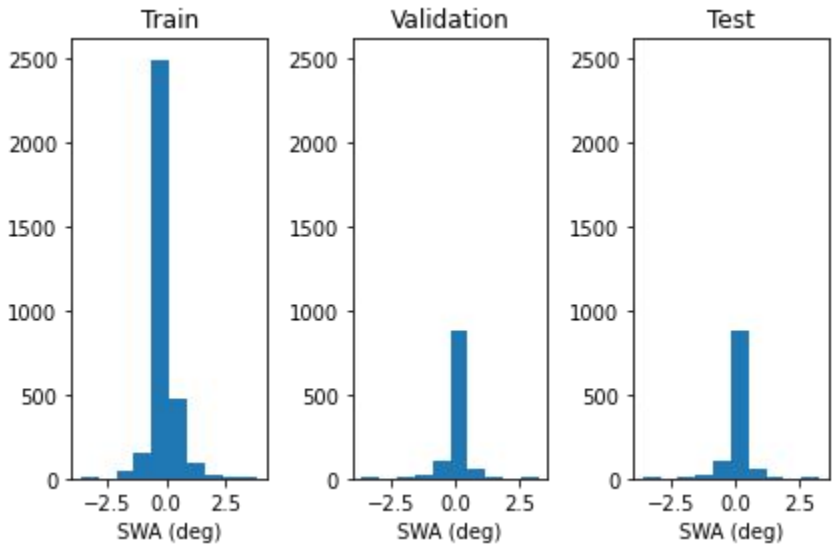
Their architectures might not be suitable for this kind of data or fine-tuning.

InceptionV3, VGG16, and VGG19 show the best performance out of all the models.

The high performing models seem to perform better at low-angle turns, and worse at higher-angle turns. This indicates a biased dataset.

This bias is confirmed by creating a histogram of the steering angles in the original set.

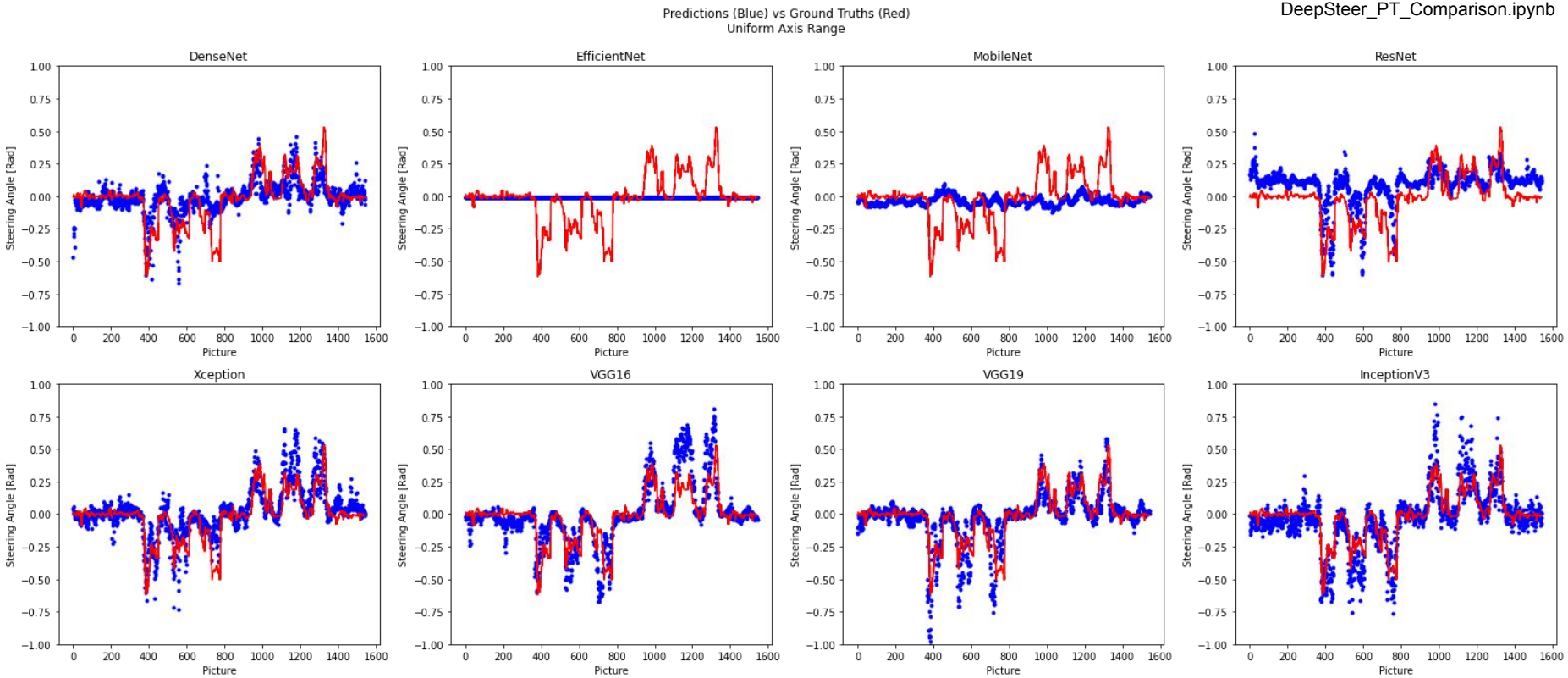
Modified version of
NEW_Steering_Regression_v3_Tamarack_Only.ipynb
by Dr. G. DeRose



Dr. G. DeRose

Predictions vs Ground Truths (Uniform Y-Axis)

DeepSteer_PT_Comparison.ipynb



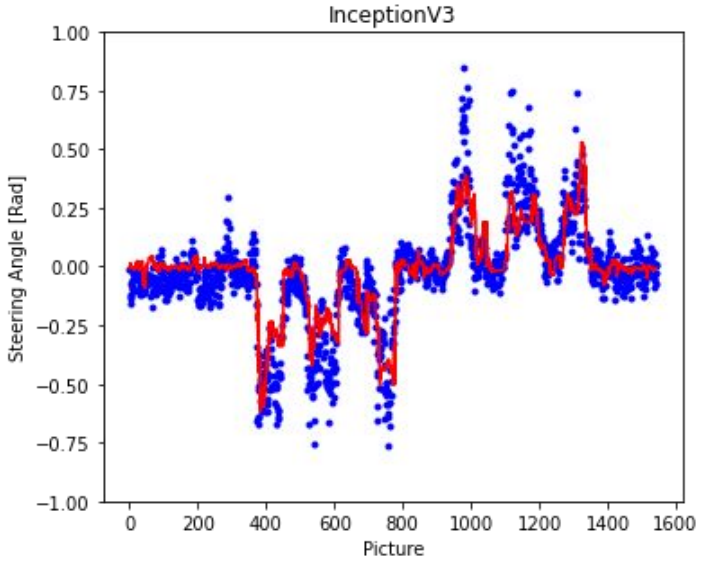
Prediction Analysis - Uniform Axis

These graphs allow us to better compare each model.

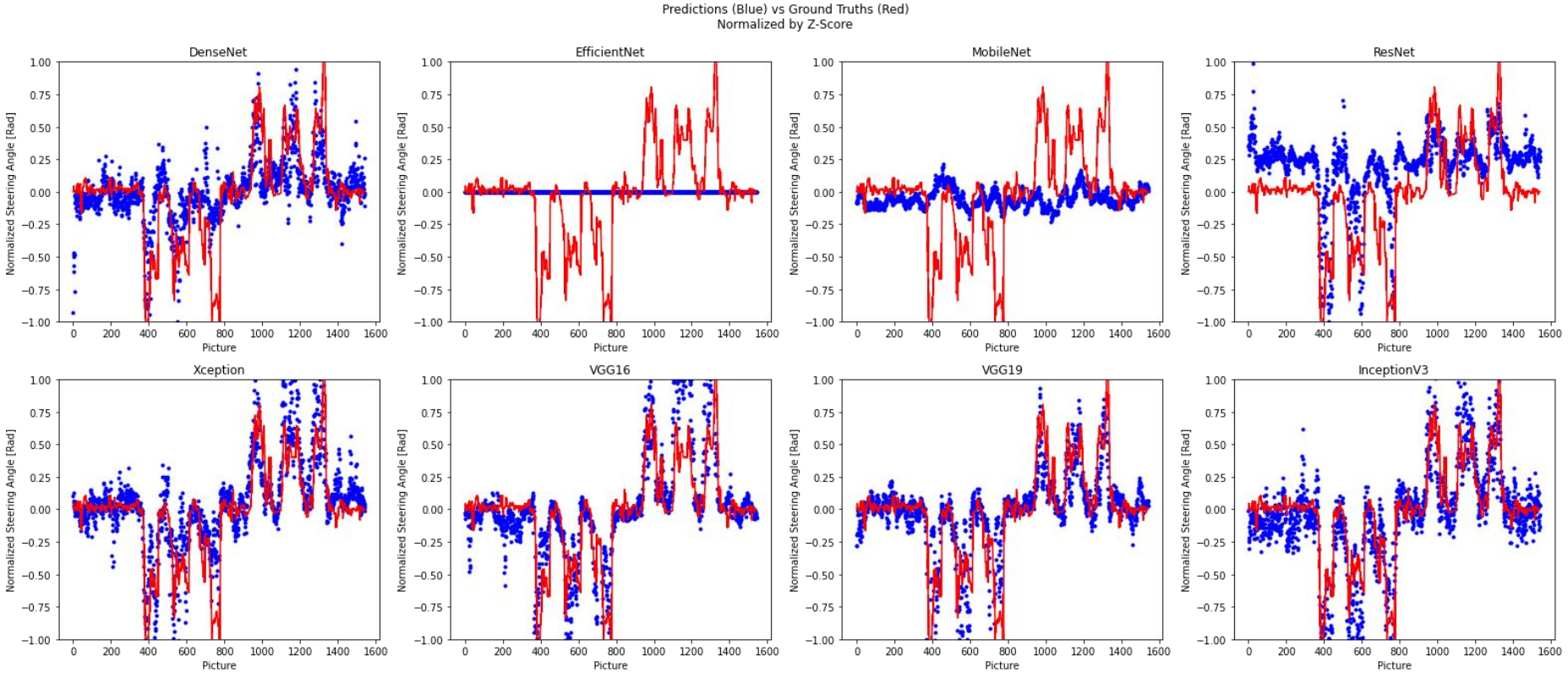
VGG16 seems to have slightly better agreement to the ground truth than InceptionV3.

InceptionV3 has better agreement during left turns.

InceptionV3 performs worse at low turning angles compared to VGG16 and VGG19.



Predictions vs Ground Truths (Normalized)



Prediction Analysis - Normalized

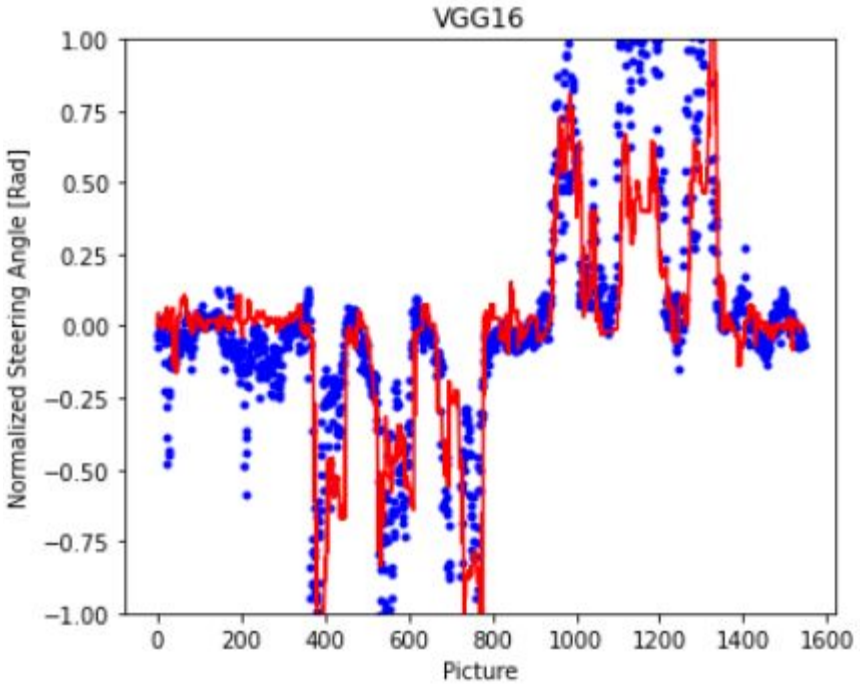
Each data point is normalized by z-score.

These plots confirm what we see in the previous two sets of graphs.

VGG16, VGG19, and InceptionV3 have the highest performances.

The higher variance of predictions at higher turning angles indicates a biased dataset.

VGG16 has error but that error has less variance than InceptionV2.



Summary

Out of the eight tested pretrained models, it is shown that DenseNet, EfficientNet, MobileNet, ResNet, and Xception do not perform sufficiently well for DeepSteer.

InceptionV3, VGG16, and VGG19 perform far better.

Higher errors at higher turning angles indicates a dataset biased towards low-angle data.

Recommendations for DeepSteer improvement and future study:

- Balance the training dataset.
- Continue to use InceptionV3 for the time being.
- Consider switching to VGG16 based on continued testing.
- Test different degrees of fine-tuning.

I would like to acknowledge Dr. Giuseppe DeRose, Dr. Chan-Jin Chung, Ian Timmis and Nicholas Paul for their work on DeepSteer and J. Schulte and A. Houck for linguistic assistance.

Links

Data, weights, and program:
https://drive.google.com/drive/folders/1x21UTSontbM28K2pPLOadk7t1jV0kIez?usp=sharing

YouTube Video: https://youtu.be/P8Vcy7ws1do