

Computer Vision Features

1 Find the Best Models

All the raw program output can be found in the *Resut-ComputerVision.txt* file. Here is the Best Models we found for the different features.

1.1 Y-Gradient

With the grid search of the following parameters of the Random Forests Model:

- Number of Decision Tree Model
 - Range 25 to 150, step 25
- Min samples to split of node
 - Range 25 to 150, step 25
- Restrict Features number for each Decision Tree Model
 - 0, 10, 20
 - 0 means use all features in every Decision Tree Model
- Use Bagging when training

We can get the accuracy of **85.5611%** on Test Set, when using the following parameters:

- Number of Decision Tree Model 150
- Min samples to split of node 25
- Restrict Features number for each Decision Tree Model 20

1.2 X-Gradient

With the same grid search of the parameters of the Random Forests Model for Y-Gradient, we can get the accuracy of **85.9736%** on Test Set, when using the following parameters:

- Number of Decision Tree Model 75
- Min samples to split of node 25
- Restrict Features number for each Decision Tree Model 10

1.3 Y-Gradient Histogram

With the grid search of the following parameters of the Random Forests Model:

- Number of Decision Tree Model
 - Range 25 to 150, step 25
- Min samples to split of node
 - Range 25 to 150, step 25
- Restrict Features number for each Decision Tree Model
 - 0, 2, 4
 - 0 means use all features in every Decision Tree Model
- Use Bagging when training

We can get the accuracy of **57.0132%** on Test Set, when using the following parameters:

- Number of Decision Tree Model 25
- Min samples to split of node 125
- Restrict Features number for each Decision Tree Model 0

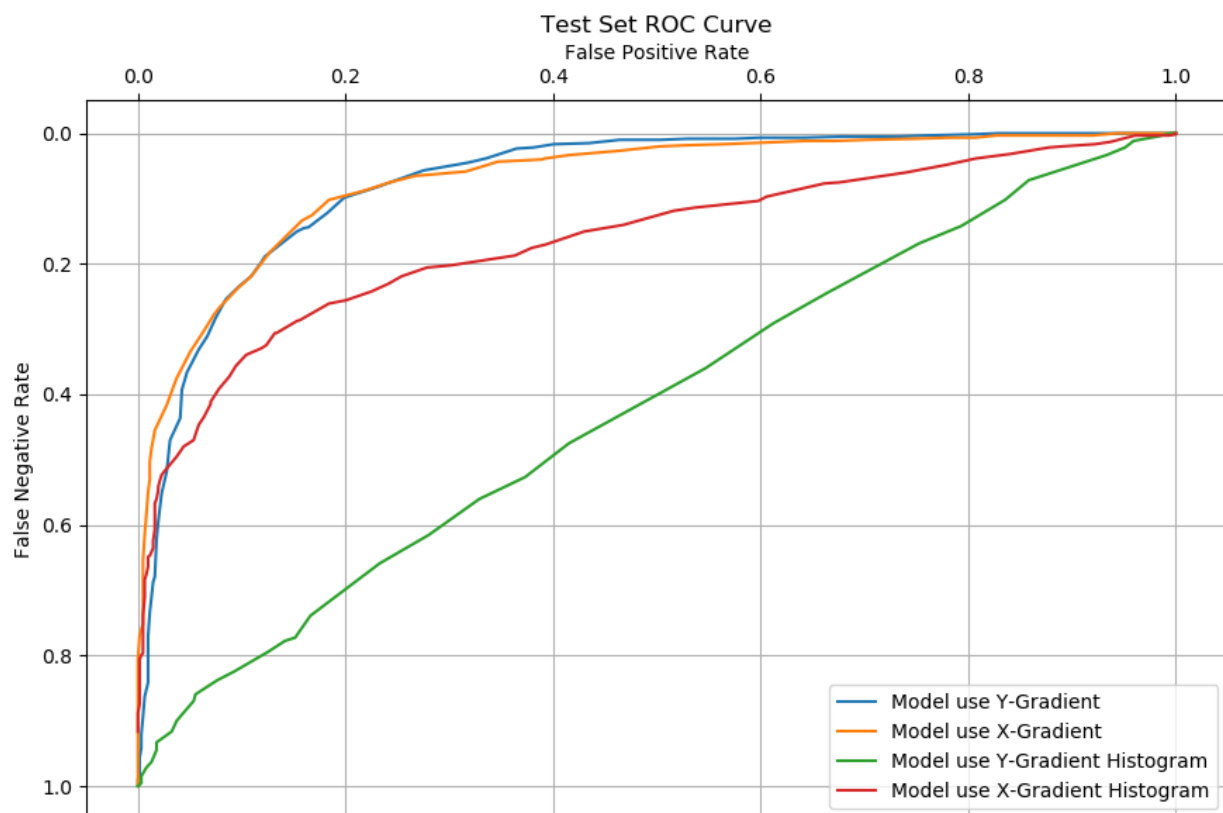
1.4 X-Gradient Histogram

With the grid search of the following parameters of the Random Forests Model for Y-Gradient Histogram, we can get the accuracy of **78.5479%** on Test Set, when using the following parameters:

- Number of Decision Tree Model 100
- Min samples to split of node 25
- Restrict Features number for each Decision Tree Model 0

2 Compare the Models

By comparing the 4 Best Models, we can generate the ROC Curves for them.



From the ROC Curves we can see that

- Directly use Gradient will performance better than use Histogram of the gradient
- When need less False Negative, Y-Gradient as the feature will be better
- When need less False Positive, X-Gradient as the feature will be better