**Methods**

To make comparisons available, Histogram of Gradients were extracted for each frame other than the black-to-white ratio. The implementation of the extraction of HOG was made up of four major steps:

1. Load the video and the ground truth annotations (.mat) file.
2. Create a video data matrix of size num\_frames x (hog\_length+1). The last column is reserved for class labels (0, 1, 2, 3 cars and so forth).
3. Extract HOG features for each frame with a cell size of 128x128 and a block size of 4x4. The sizes are ambiguous. The reason for the cell size and block size to be the specific values above is that it would not generate a result that is either too large or too small.
4. Iterate over the annotations matrix, looking for the corresponding frame numbers both in its cells and in the video data matrix. For every same frame encountered, one is added to the class label.

After obtaining HOG features of the video data from the video and the class labels from the annotations matrix, the data were trained with an SVM classifier. LIBSVM 3.2.2 was used for the implementation of both training and testing.

To do the training and testing, the video data was randomly split into a training dataset and a testing dataset with a ratio of 80%-to-20%.

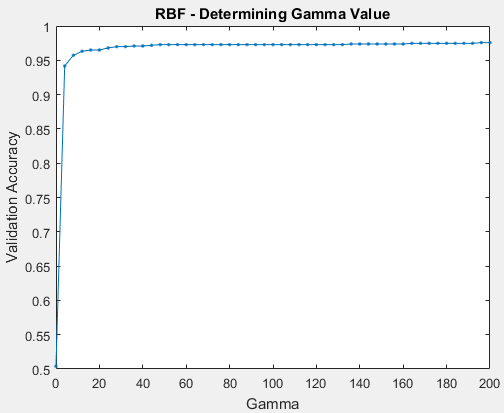
In the process of processing the video ‘PS12\_3\_3’ and getting video data frame by frame, I found the last frame (#5148) has all 0’s for HOG features. It may be since the last frame is empty. Therefore, I removed the last row of the data before saving it to file.

**Results**

PS12\_3\_3:

RBF Kernel:

* Best gamma = 196



* Validation
  + Confusion matrix:

517 2 0 0

22 319 1 0

0 0 128 0

0 0 0 41

* + Accuracy = 97.57% (1005/1030)
* Testing
  + Confusion matrix:

532 0 0 0

11 328 1 0

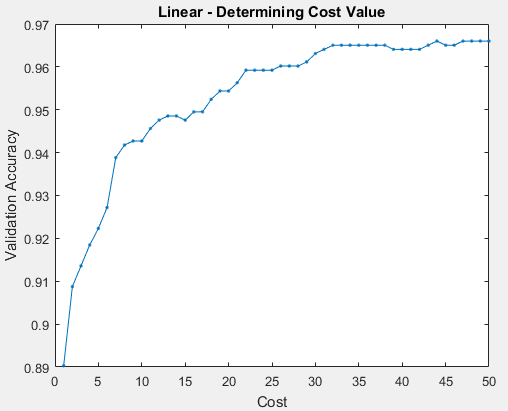
0 0 115 0

0 0 0 42

* + Accuracy = 98.83% (1017/1029)

Linear Kernel:

* Best cost = 44



* Validation
  + Confusion matrix:

519 0 0 0

34 307 1 0

0 0 128 0

0 0 0 41

* + Accuracy = 96.60% (995/1030)
* Testing
  + Confusion matrix:

532 0 0 0

29 308 3 0

0 1 114 0

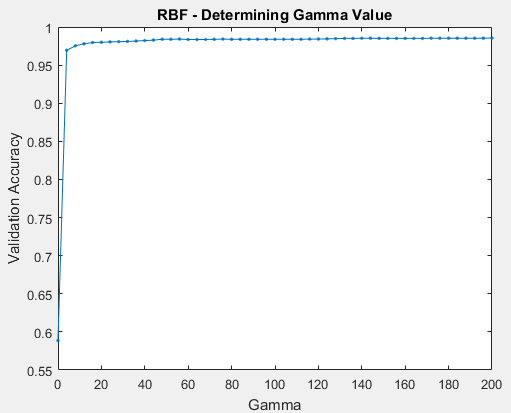
0 0 0 42

* + Accuracy = 96.79% (996/1029)

PS12\_1\_7

RBF Kernel

* Best gamma = 200



* Validation
  + Confusion matrix:

2075 5 0 0

34 1238 3 0

1 8 131 0

0 0 0 39

* + Accuracy = 98.56% (3483/3534)
* Testing
  + Confusion matrix:

2130 3 0 0

35 1171 2 0

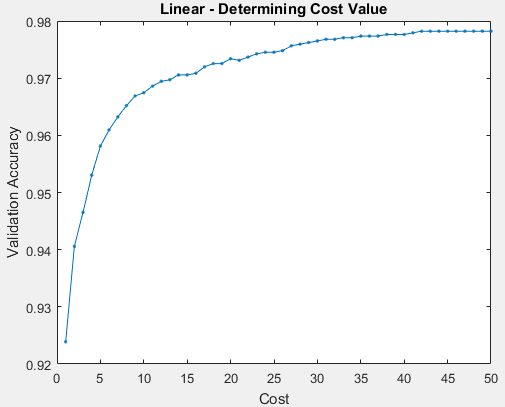
1 7 133 0

0 0 0 51

* + Accuracy = 98.64% (3485/3533)

Linear Kernel

* Best cost = 42



* Validation
  + Confusion matrix:

2074 6 0 0

42 1232 1 0

1 27 112 0

0 0 0 39

* + Accuracy = 97.82% (3457/3534)
* Testing
  + Confusion matrix:

2127 6 0 0

47 1159 2 0

2 26 113 0

0 0 0 51

* + Accuracy = 97.65% (3450/3533)