OBJECT CLASSIFICATION FOR SURVEILLANCE
CS 771A COURSE PROJECT

Instructor: Dr. Harish Karnick

Group 45

Vicki Anand

Parikshit Khanna

Atique Firoz

Akash Gupta

Proposal

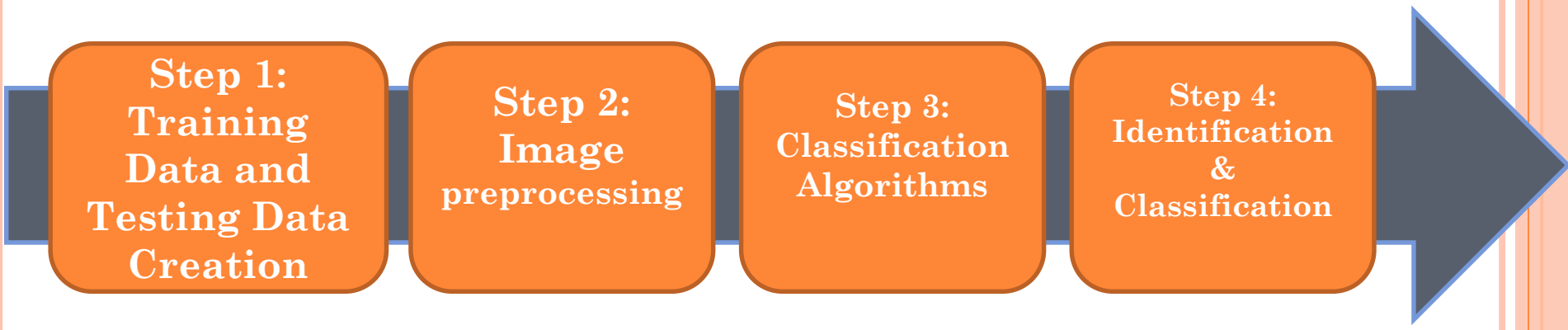
We classified the dynamic objects in the feed into the following categories :

1. Pedestrians
2. Slow moving two-wheelers (Bicycles)
3. Fast moving two wheelers (Motor-cycles)
4. Slow moving three-wheelers(Manual Rickshaw)
5. Fast Moving three wheelers (Auto-Rickshaws)
6. Fast Moving four wheelers (Cars)

Implementation

- Phase 1: Extracted images and refined the crowd-sourced labeled videos.
- Phase 2: Experimented with various classifiers and tuned them for best results
- Phase 3: Foreground and background subtraction using frame differencing and localization. Used for porting our learnt classifier with new videos

Classification & Identification Overview



Data Description

Train Data-Set :

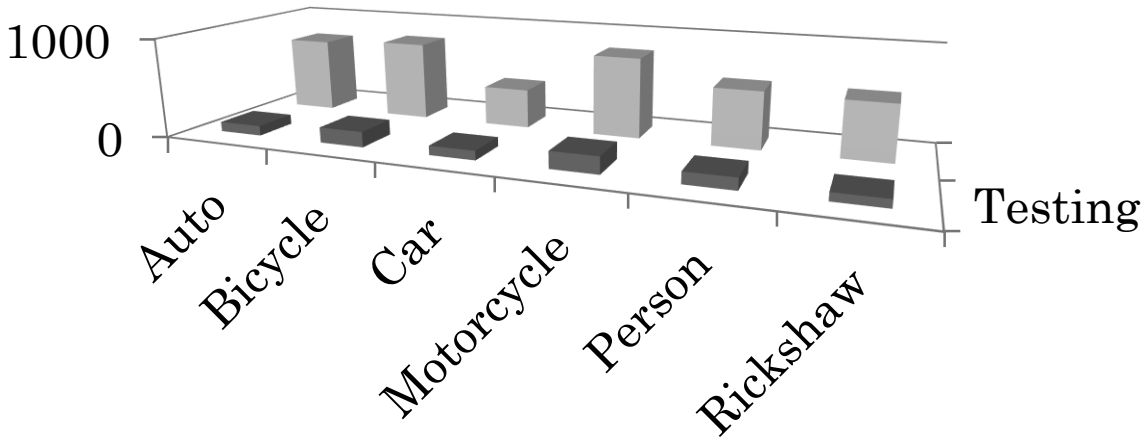
1. Datasample1
2. input_video_sample1
3. input_video_sample2
4. input_video_sample3
5. videosample5

Test Data-Set :

1. dec21h1330
2. nov92015-1
3. nov92015-2

Distribution Of Labels in our Sets

Object	Training Set	Testing Set
Auto	756	98
Bicycle	798	154
Car	396	92
Motorcycle	812	169
Person	580	116
Rickshaw	564	80
Total	3906	709



Preprocessing of Cropped Images

- **Rescaling**: Images rescaled to a common size of 30 x 30 and 50 x 50.
- **Padding**: Extra pixels set to zero for rectangular images to maintain their aspect ratio.
- Data converted to greyscale (900) & RGB (2700) arrays.
- No significant gain by increased sizes or switching from greyscale to RGB.

Feature Extraction

Tuned parameters for following algorithms:

- HOG (Histogram of Oriented Gradients)
- PCA (Principal Component Analysis)

HOG (Histogram of Oriented Gradients)

- Parameter tuning done using linear SVC

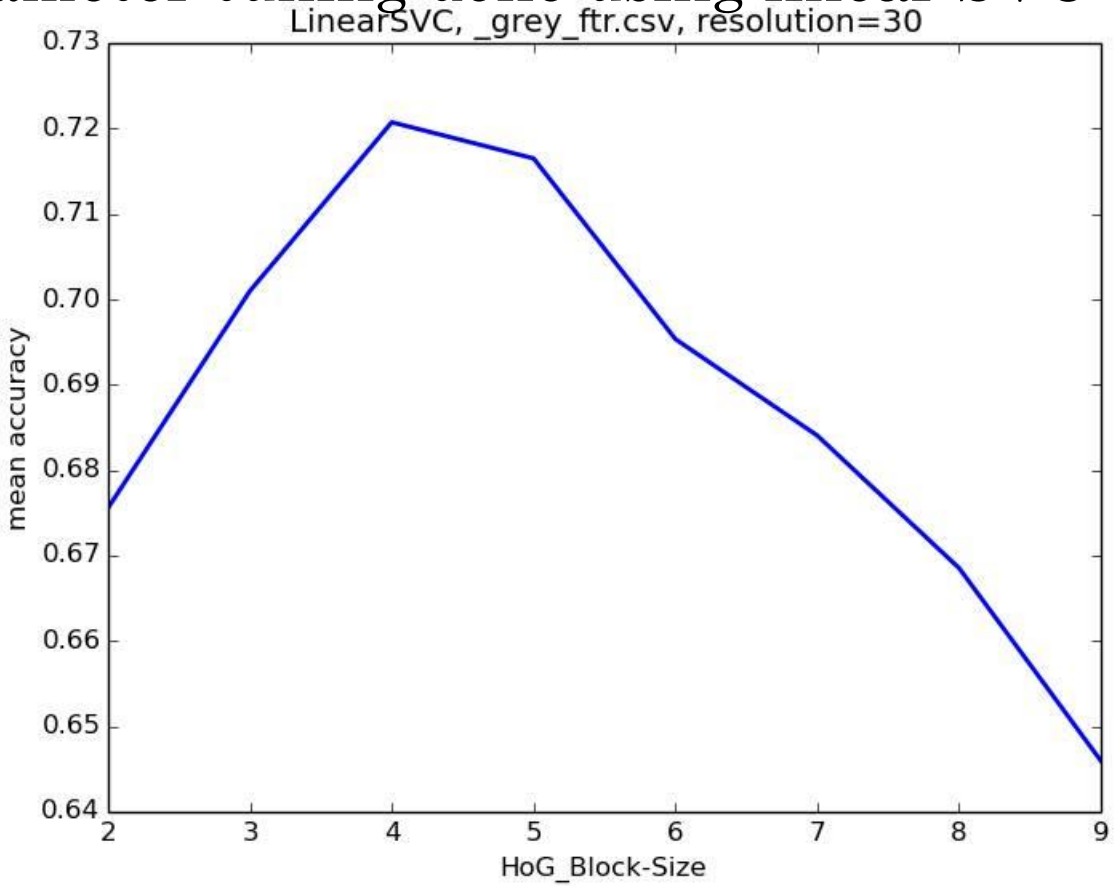


Fig : Accuracy vs Hog Block Sizec

HoG : Accuracy vs Cell Size

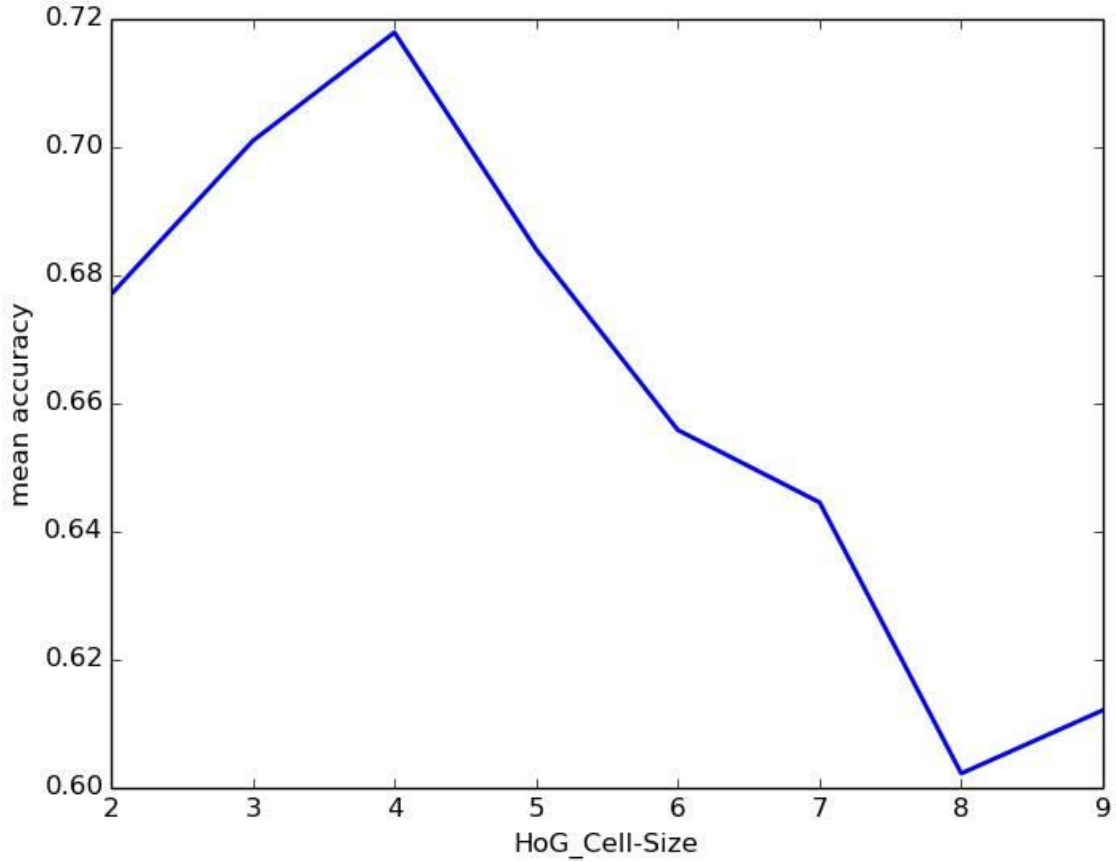


Fig : Accuracy vs Cell-Size

Accuracy vs HoG Orientation

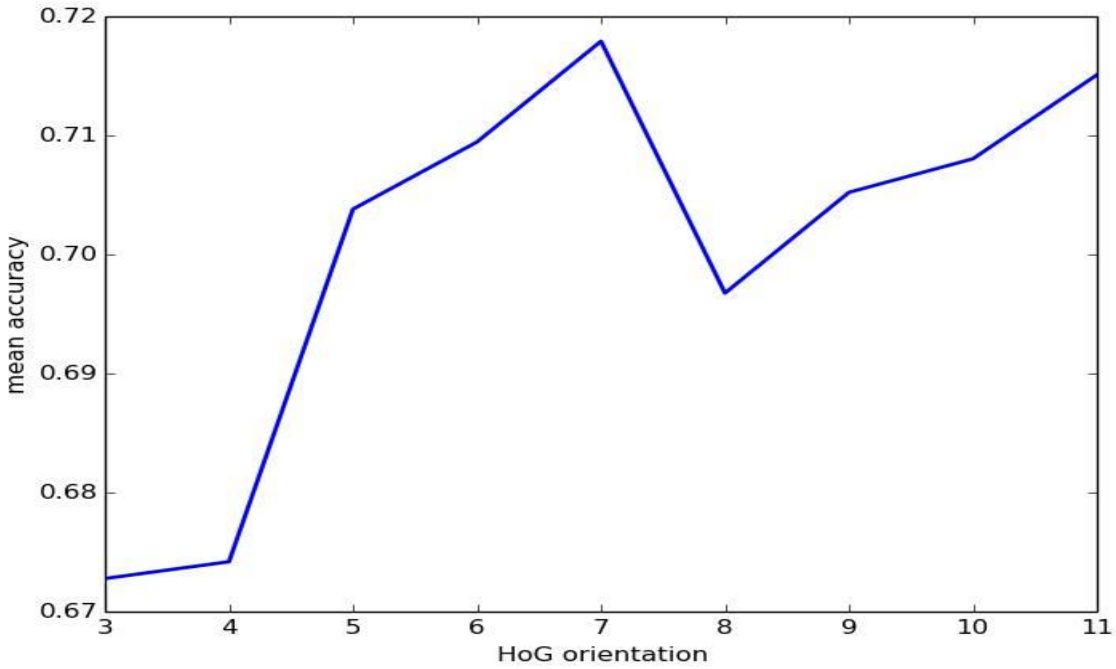


Fig : Accuracy vs HoG Orientation

Best accuracy for

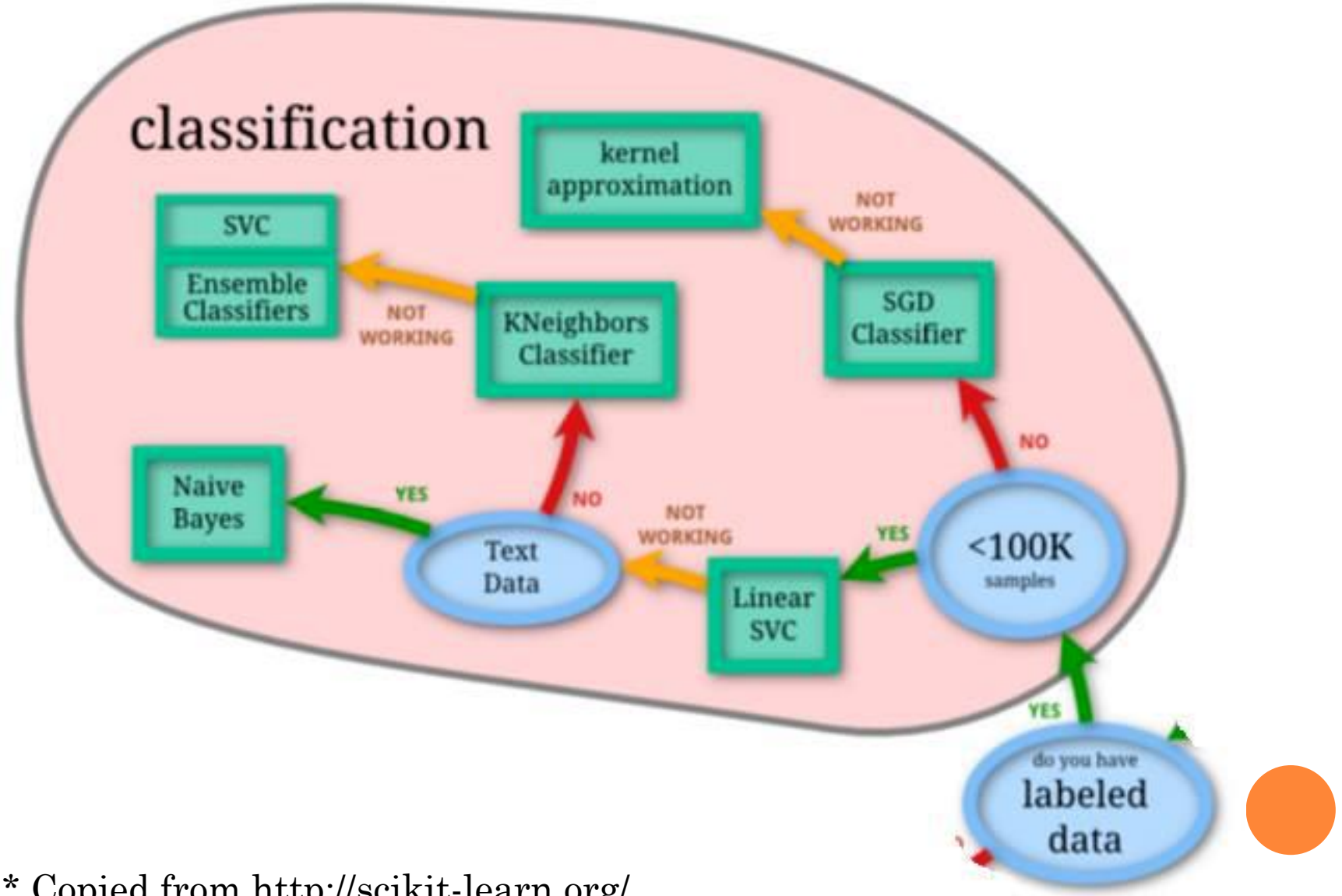
- Orientation = 7
- Pixels per cell = 4
- Cells per block = 4

Principal Component Analysis

Employing PCA led to significant reduction in accuracy . We believe that the reasons for such reductions are :

1. Proportion of variance explained is low for higher components, i.e variance is uniformly distributed across all the components.
2. Due to different dimensions, a significant degree of padding was involved which reduced the rank leading to fewer eigen values.

Classifier Road Map *



* Copied from http://scikit-learn.org/

Classifiers Used

1. K-NN Classifier
2. SVC Classifier
3. Decision Trees Classifier
4. Random Forest
5. Adaptive Boosting Algorithm

KNN Classifier (HoG)

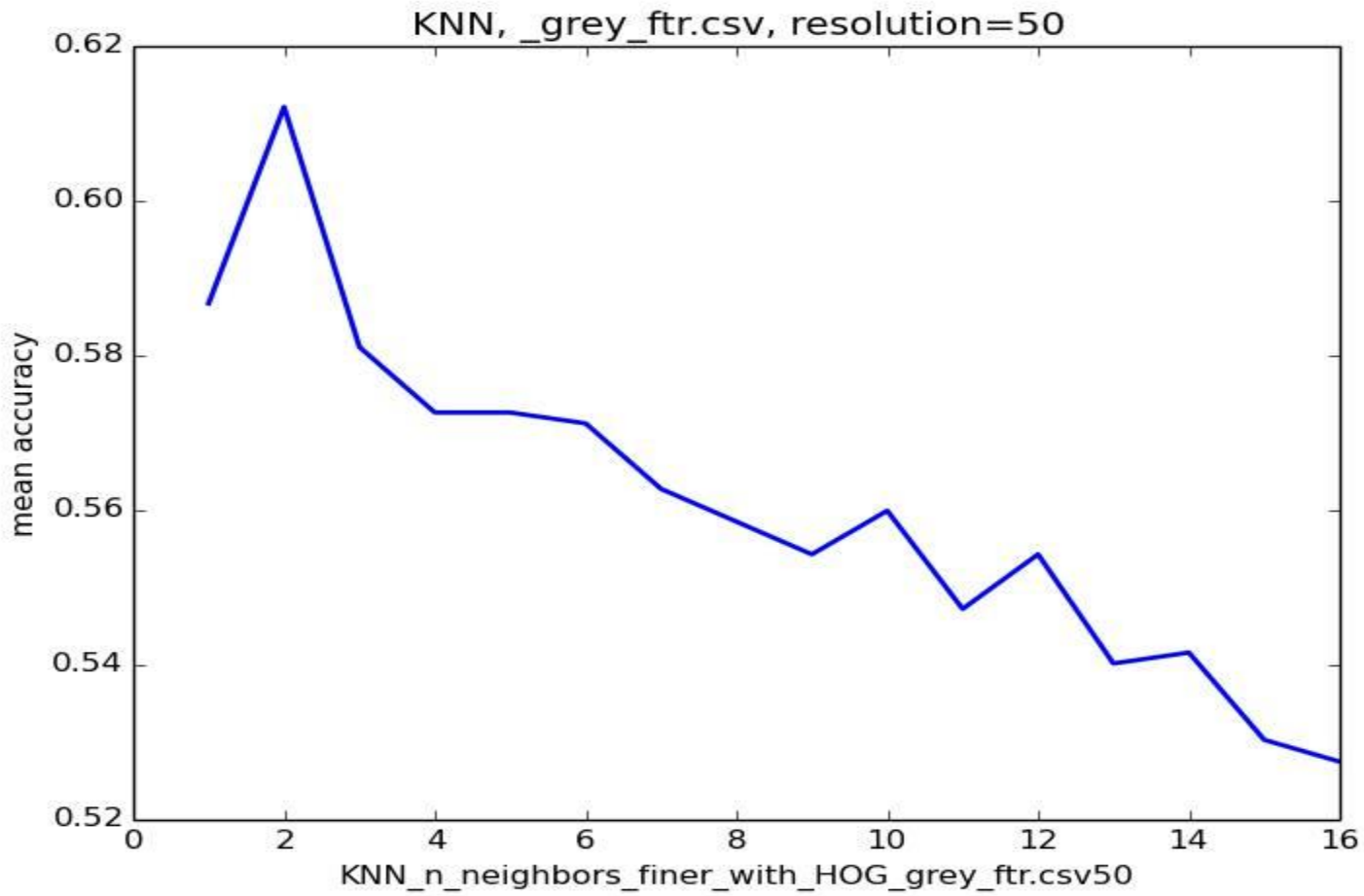


Fig : Accuracy vs K (HoG)

KNN Classifier (Without HoG)

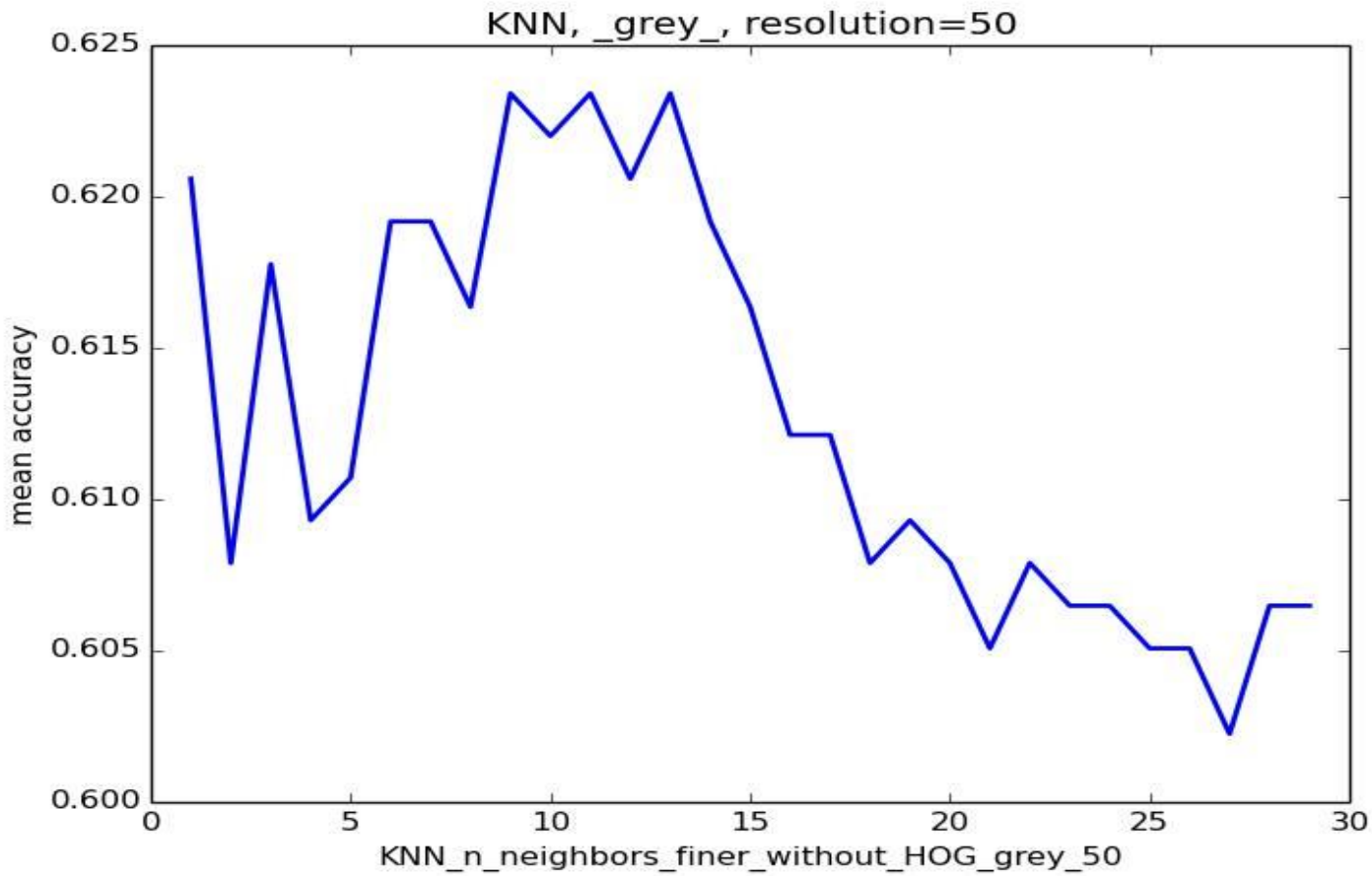
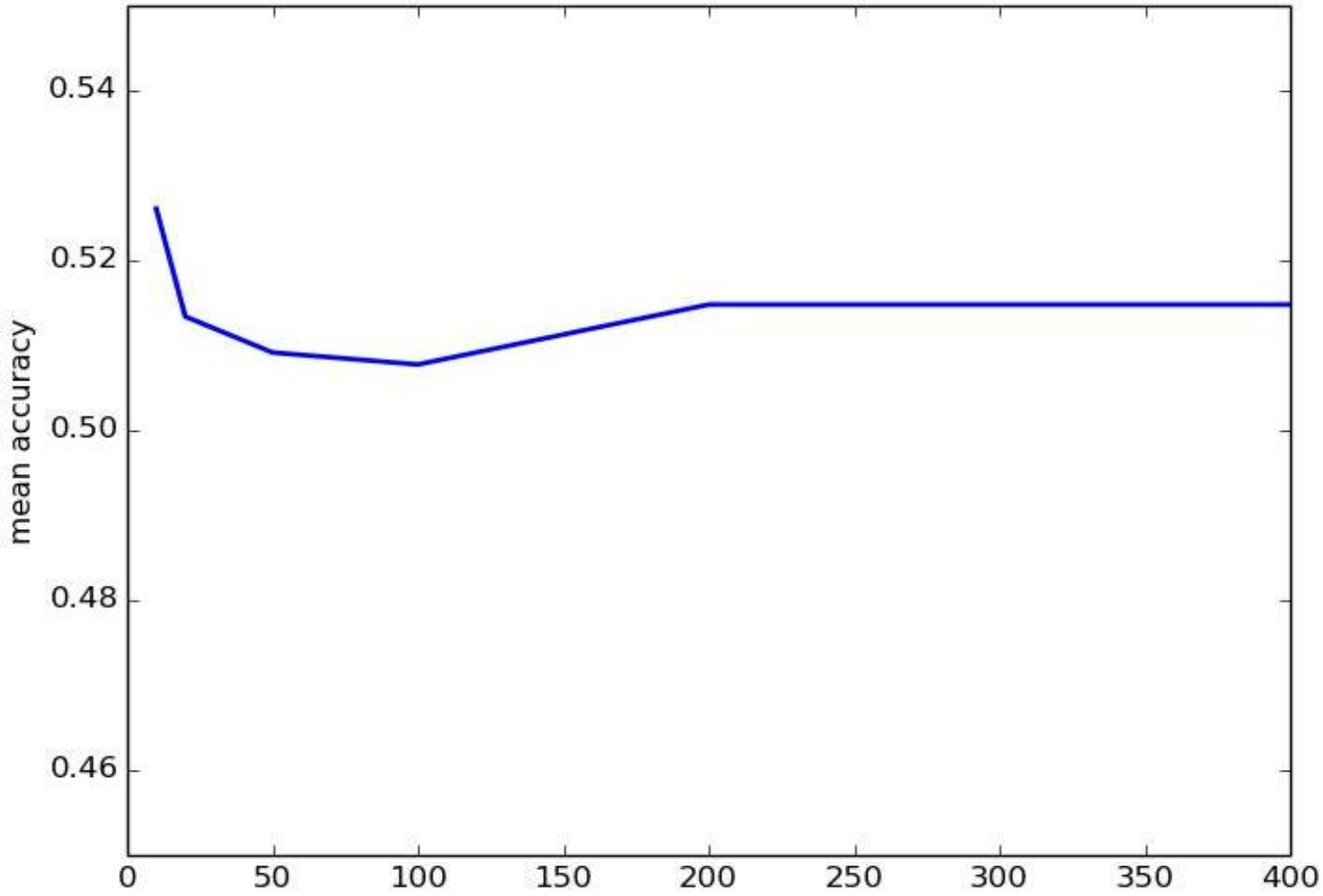


Fig : Accuracy vs K

Decision Tree (HoG)



Accuracy vs Max Depth (HoG)

Decision Tree (Without Hog)

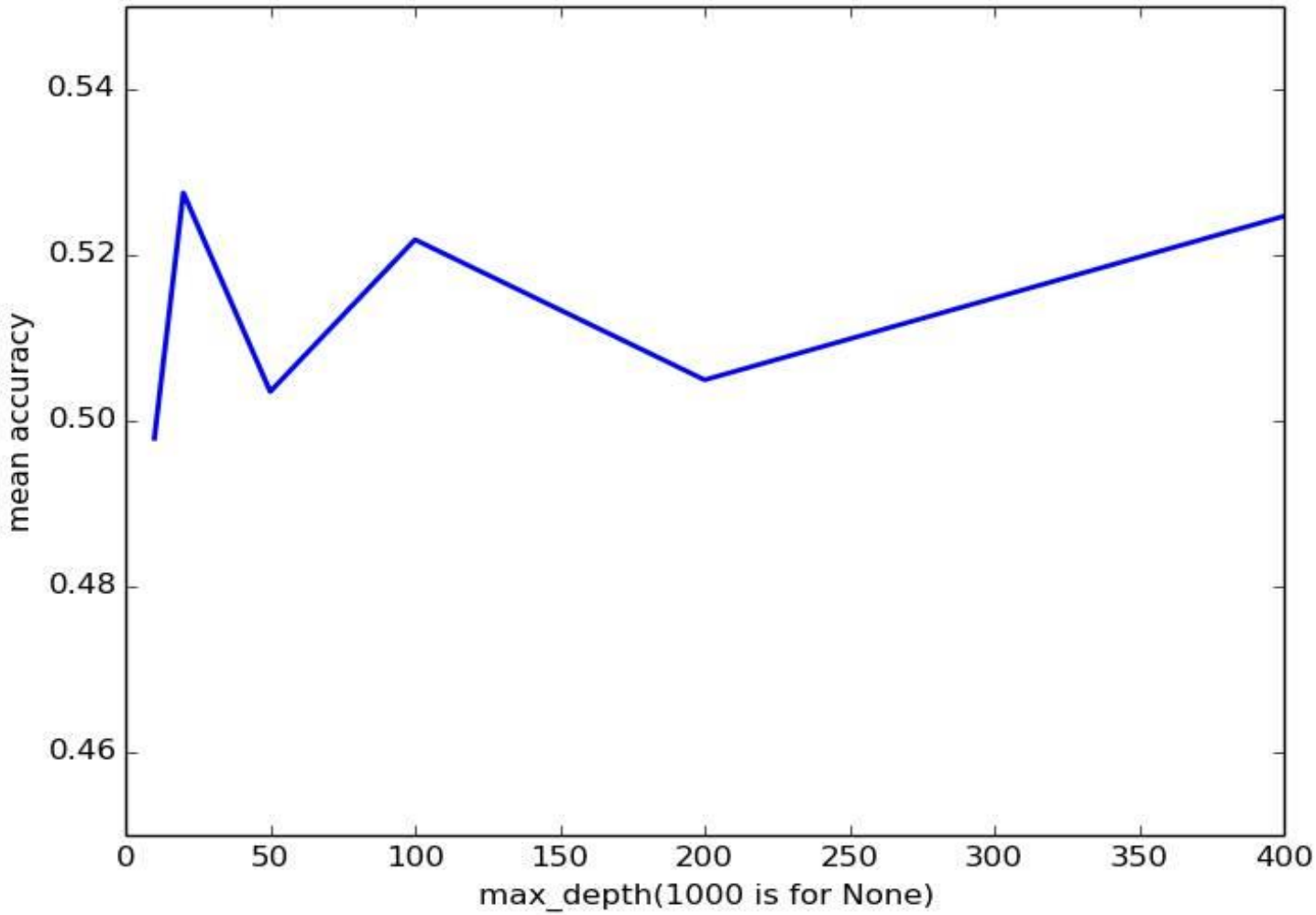
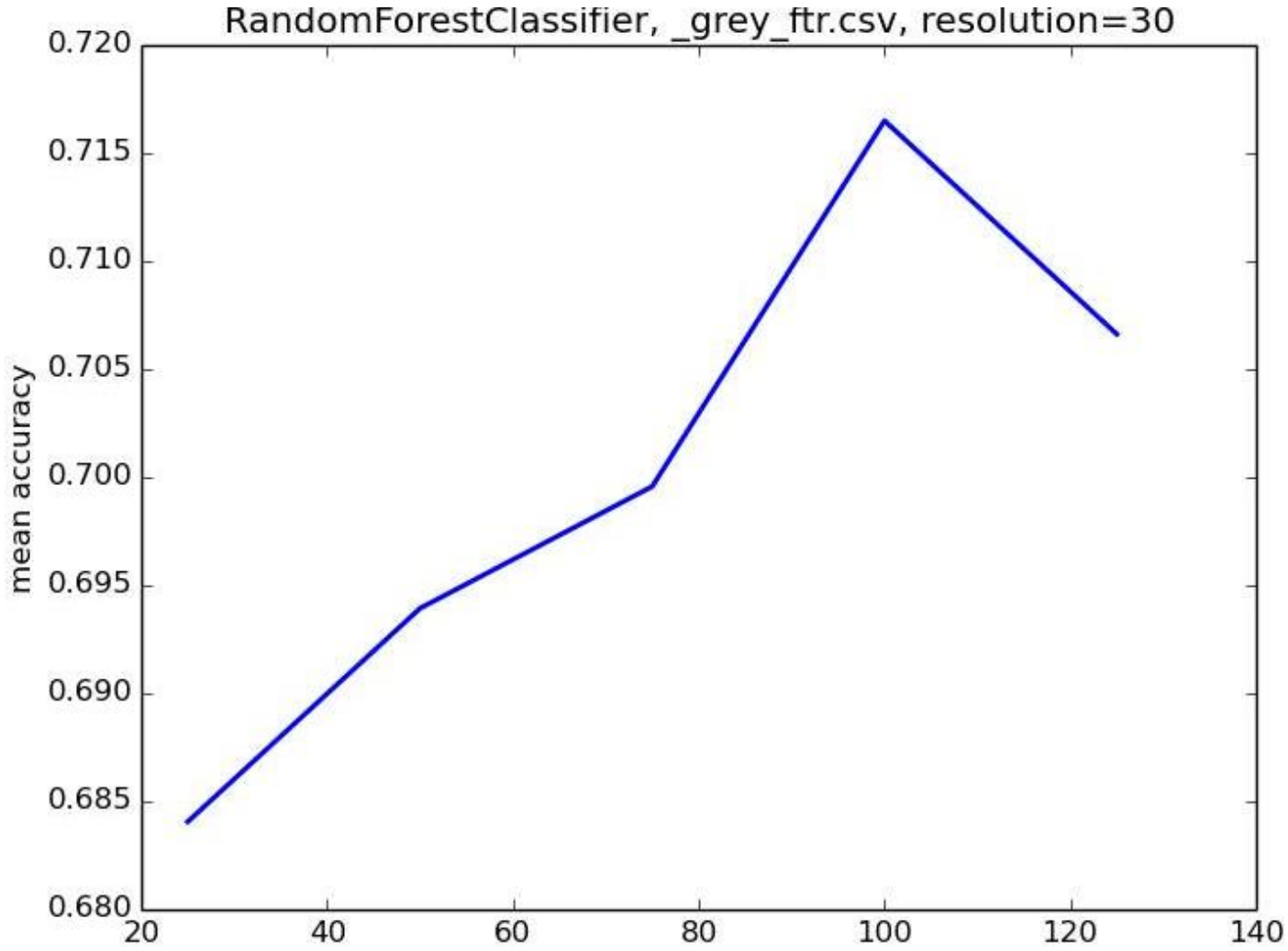


Fig : Accuracy vs Max Depth (HoG)

Random Forest Classifier (HoG)



Accuracy vs Number of Estimators (HoG)

Random Forest Classifier (Without HoG)

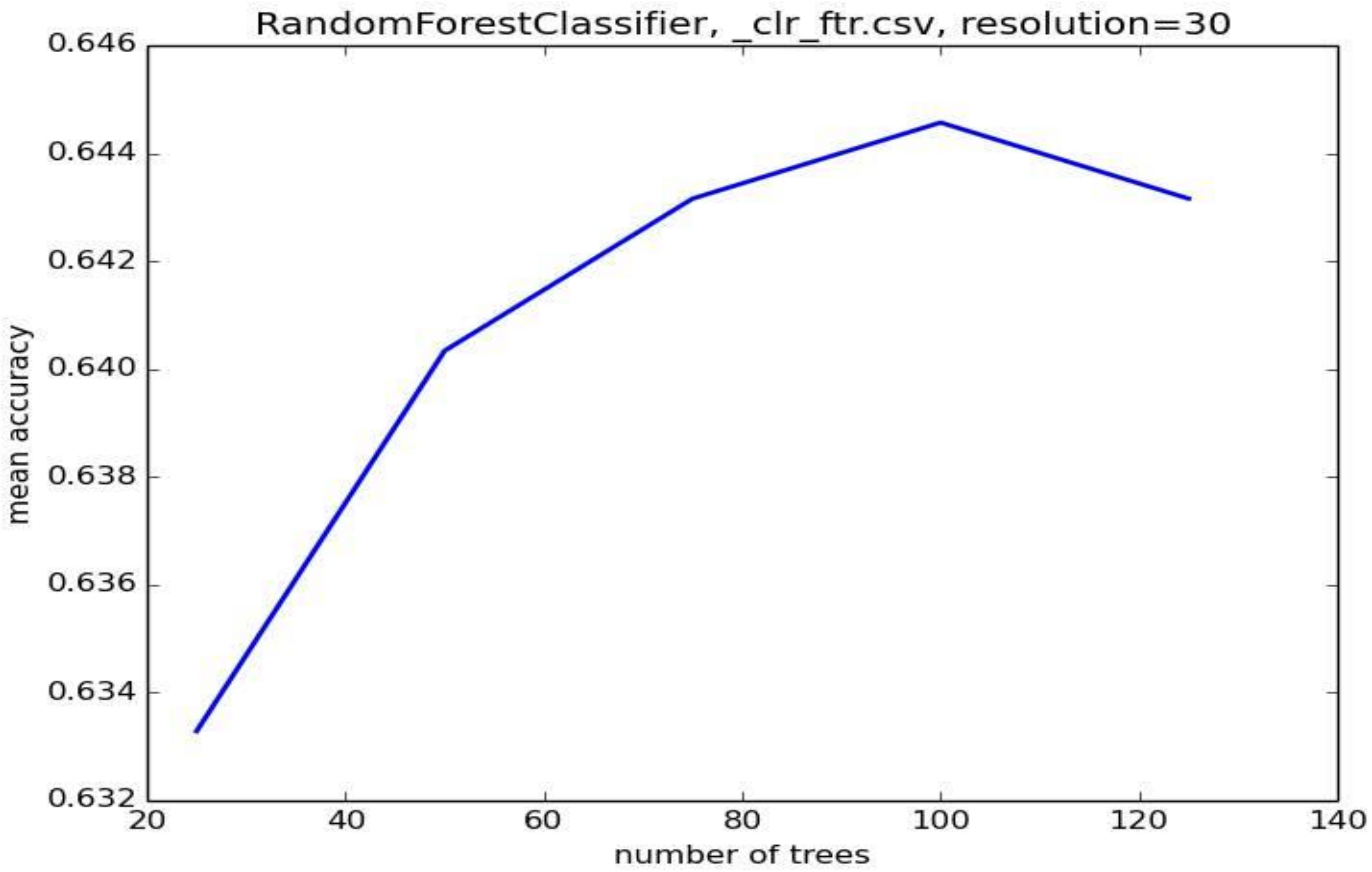


Fig : Accuracy vs Number of Estimators

Adaboost Algorithm

IDENTIFICATION (BACKGROUND SEPARATION)

- Underlying Assumputions : The background of our video stream is largely static and unchanging over consecutive frames of a video. While modeling the background, we monitored it for substantial changes. If there was a substantial change, we detected it. This change normally corresponds to motion on our video.
- Used the last local frame as reference for the detection of motion and applied the frame differencing between the current and last frame. (experimented with different frame gaps).

Motion Detection

- Because of open environment, even background of consecutive frames of a video stream are effected by camera sensors and lighting conditions etc.
- We need to account for this and reduce this effect by applying Gaussian blurring to average pixel intensities across the region.

CONTINUED..

Fig : Video Frame vs Gaussian Blur

CONTINUED..

- Converted image to binary using threshold value.

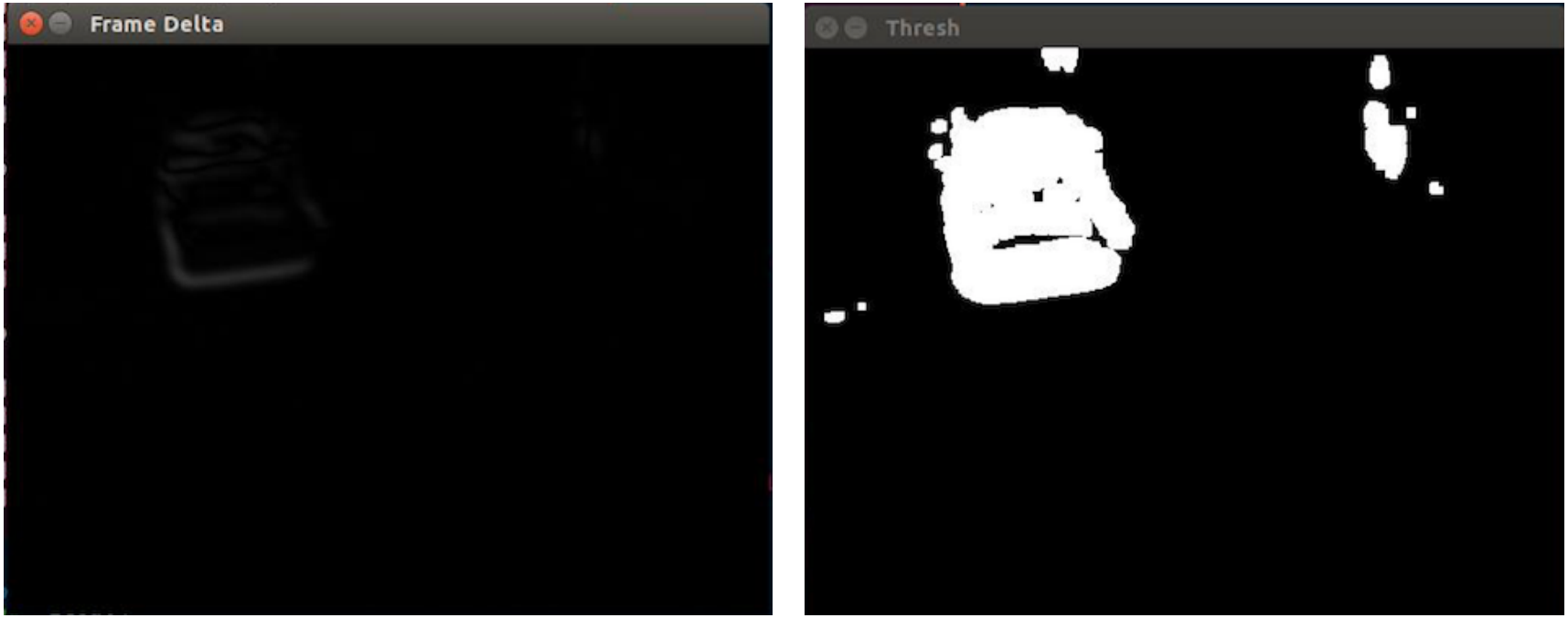


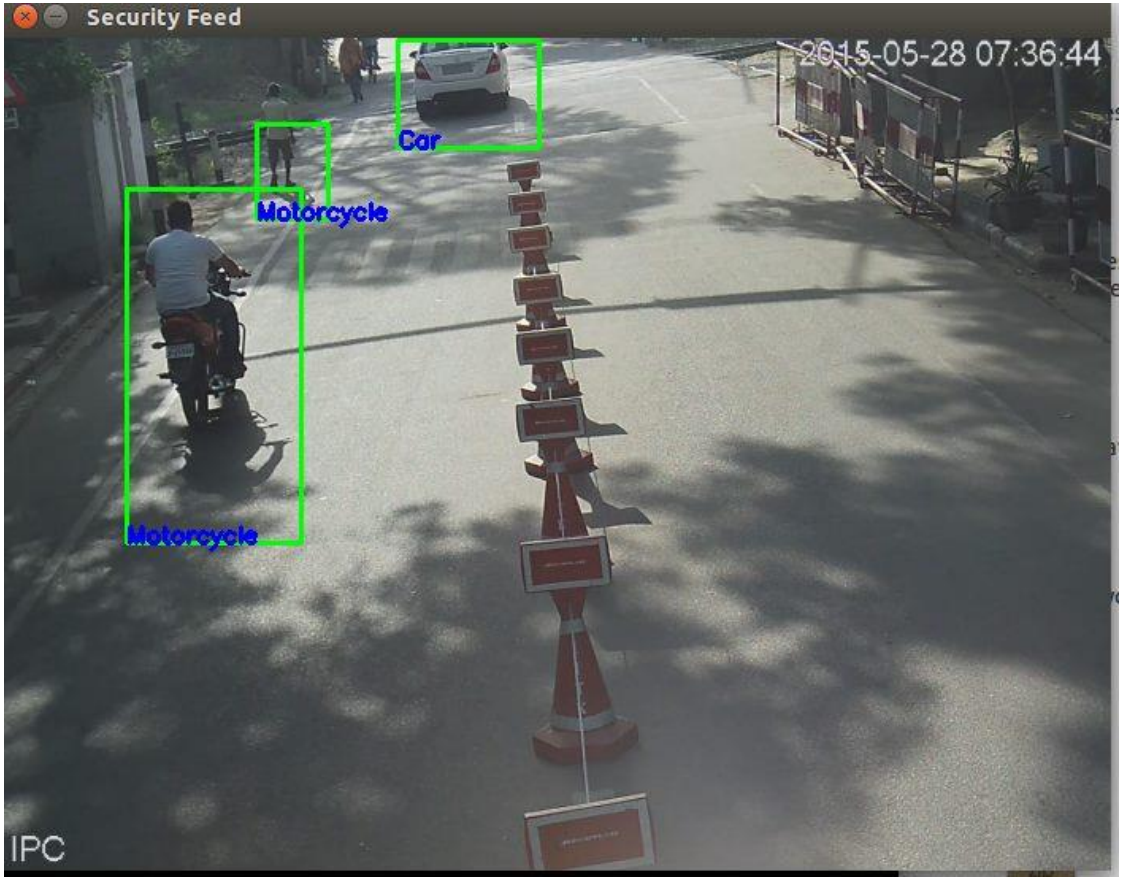
Fig : Frame differencing and Binary image

CONTINUED..

- Contour detection to find the outlines of the white regions.

CONTINUED..

- Classification on the cropped objects detected and labeling.



Results

Model	Hog	Without Hog
K-NN	61.9%	62.5%
SVM	**71.79 %**	61.8%
Decision Tree	52.60%	52.75%
Random Forest	71.6%	64.5%
AdaBoost	50.02%	44.87%

Confusion Matrix for SVM (Best Result)

Person : Bicycle : Motorcycle : Ricksaw : Auto : Car

[79	32	4	1	0	0]
[2	84	40	17	6	5]
[0	15	106	19	22	7]
[0	4	6	65	1	4]
[0	2	13	4	74	5]
[0	1	5	0	0	86]

Confusion Matrix