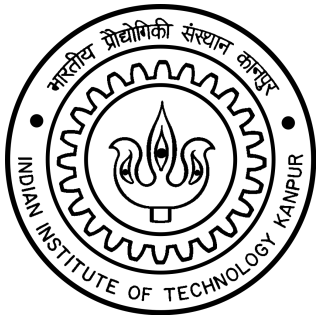
CS771: Machine Learning: Tools, techniques, applications

Project Title: Background Foreground Separation and Object Classification

Instructor: Professor Harish Karnick

Group Number : 6

- Avani Samdariya
- Rahul Kumar Wadbude
- Ritika Mulagalapalli
- Shubham Agrawal



The Problem Statement

Object Classification

- Classification of objects into : pedestrian, 2/3/4 wheelers(coarse classification)
- Classification of objects into 6 classes: person, bicycle, motorcycle, rickshaw, auto-rickshaw and cars(fine classification)

Background Foreground

- Separation of Background and Foreground in the videos to get images of objects for Label Prediction

Problem break down

- Pre trained deep convolutional neural network

- Decision Trees
- Random Forests
- One-vs-rest SVM

- Background and foreground separation to detect moving objects
- Draw contours around them
- Use classifier to predict labels on contours

Feature Extraction

Feature Extraction

- Used Caffe framework with gpu hardware acceleration for feature extraction
- Pretrained **BVLC GoogLeNet** model is used :
  - It uses deep convolutional neural network architecture to predict labels on images
  - It was responsible for setting the new state of the art for classification and detection in the ImageNet Large-Scale Visual Recognition Challenge 2014 (ILSVRC 2014)
- Also considered SURF features and compared them with deepnet features on **Caltech dataset** with same classification algorithm
  - Deepnet features outperformed SURF features

Feature Extraction - SURF vs Deepnet

- Used one-vs-rest svm for comparison
- **SURF** :
  - Tuned parameters : Kernel = Sigmoid, C = 10000
  - Accuracy = **0.5994**
- **Deepnet :**
  - Tuned parameters : Kernel = Polynomial, C = 1
  - Accuracy = **0.944297**

Classification Model

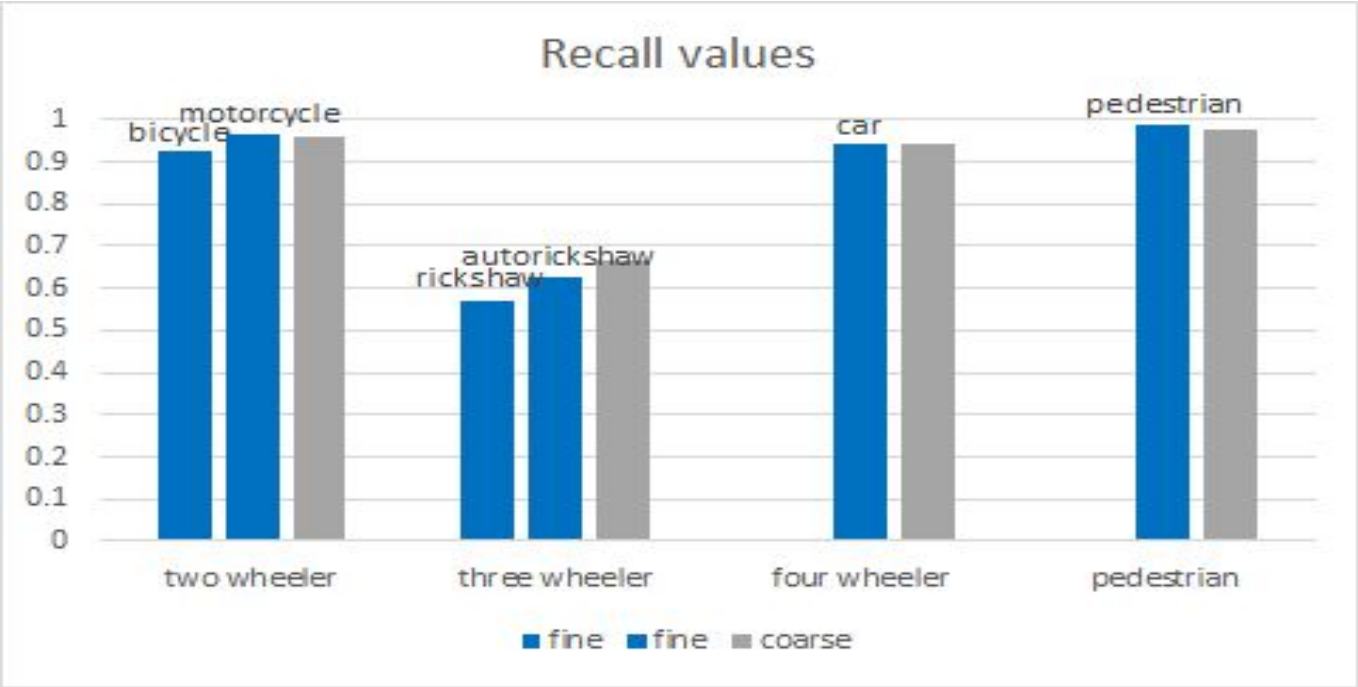
Steps for Classification:

- Coarse vs Fine: choice of classes to be used
  - Coarse classes: pedestrian, 2/3/4 wheelers
  - Fine classes: person, bicycle, motorcycle, rickshaw, autorickshaw, car

Since Caffe gave good results, we decided to go for fine classification also

- Classifiers used: SVM, Decision Trees, Random Forests
  - SVM: tuned parameters- C(penalty error term) and kernel
  - Decision trees: tuned parameters- minimum samples at leaf nodes and impurity criteria among gini and entropy
  - Random forest: varied number of trees

Best Classifier: SVM for C = 10

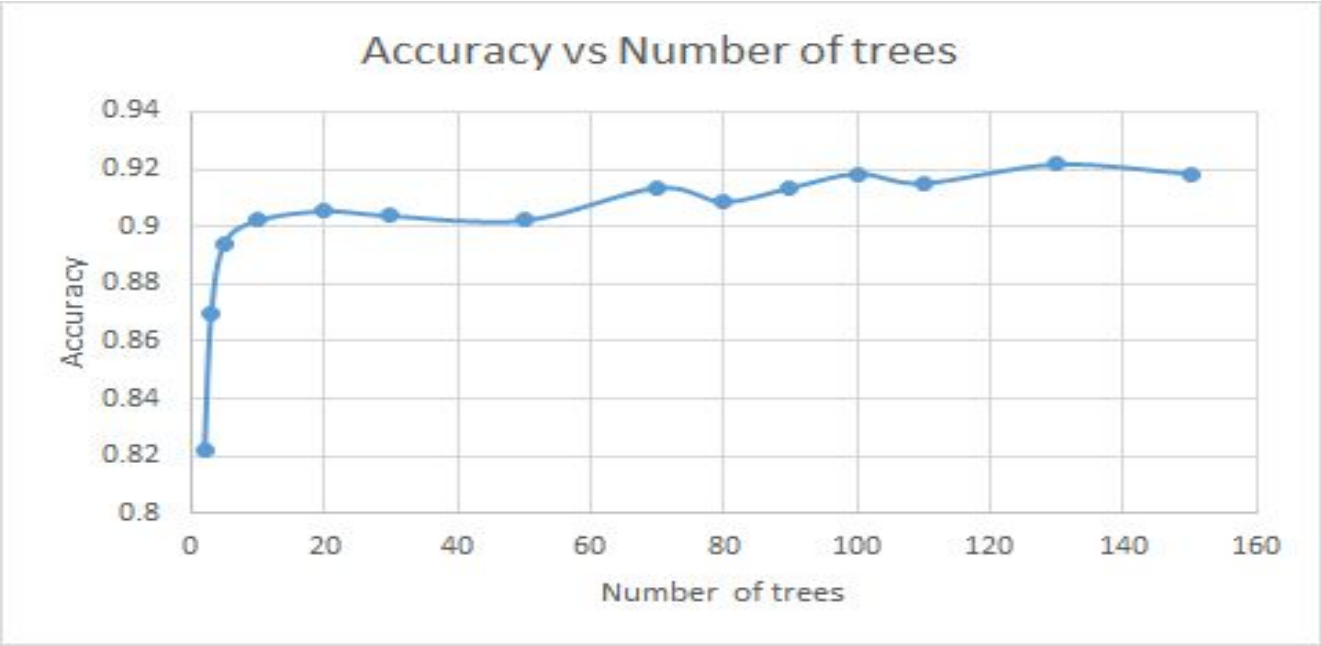


Classifier used: *SVM with C = 10*
Kernel function: *rbf (for fine), poly (for coarse)*
Results are better for coarse. The following data on classifiers is for coarse classification.

Decision Trees

S No.	Impurity Index	Min no of vectors at leaf node	Accuracy
1	‘gini’	1	84.9918
2	‘entropy’	1	85.155
3	‘gini’	3	84.9918
4	‘entropy’	3	85.155
5	‘gini’	5	86.7863
6	‘entropy’	5	86.2969

Random Forests



Converges to 0.92
Threshold no of trees = 30

SVM

Kernal	C=0.1	C=1	c=10	C=100	C= 1000	C=10000
linear	95.758	95.758	95.758	95.758	95.758	95.758
poly	94.943	96.737	97.063	97.063	97.063	97.063
rbf	92.659	96.085	96.737	96.737	96.737	96.737
sigmoid	90.701	91.354	87.928	85.318	85.154	84.829

Clearly, for coarse classification, SVM gives the highest accuracy with C=10 and polynomial kernel which we used in our final prediction.

FINAL PREDICTION

1. Background and foreground separation :
    a. Tried MOG and MOG2 Algorithms:
    b. MOG2 has better performance.
    c. Applied various filters like blurring , hole-filling, noise removal.
2. Draw bounding boxes over contours found in foreground mask obtained in step 1.
3. Predict the label of objects in bounding box.

Challenges faced:

1. Incorrect labels in labelled data.
2. Labelled data of two wheeler contain only vehicle but on bgs, we get vehicle with the rider. Classifier faces difficulty there.
3. Shadow of Trees and objects makes the task difficult.
4. Naive implementation of BGS(mog 2 without filters) clubs objects that are close in space.
5. Occluded objects were harder to classify.

Future Extensions:

- In case of bicycles and motorcycles, data was labelled by cropping person sitting on it. But background foreground separator draw contours around entire bicycle + person. Since our classifier was trained only for bicycle, it was confusing object with person and bicycle.
- Try object localization in video for finding contours
- We noticed that when object was big in size, our classifier predicted label correctly. We can use this in following way :
    - Trace path of object in video
    - Predict all labels along the path
    - For predicting label along the path, Increase weightage of that part of path where object size was bigger