Car Detecting Method using high Resolution images

Swapnil R. Dhawad
Department of Electronics and Telecommunication Engineering
JSPM's Rajarshi Shahu College of Engineering, Savitribai Phule Pune University, Pune, India
swapnildhawad555@gmail.com

Mrs. Rajashri R. Itkarkar
Department of Electronics and Telecommunication Engineering
JSPM's Rajarshi Shahu College of Engineering, Savitribai Phule Pune University, Pune, India
itkarkarrajashri@yahoo.com

Abstract— A car detection method is implemented using high resolution images, because these images gives high level of object details in the image as compare with satellite images. There are two feature extraction algorithms are used for implementation such as SIFT (Scale Invariant Feature Transform) and HOG (Histogram of Oriented Gradient). SIFT keypoints of objects are first extracted from a set of reference images and stored in a database. HOG descriptors are feature descriptors used in image processing for the purpose of object detection. The HOG technique counts occurrences of gradient orientation in localized portions of an image. The HOG algorithm used for extracting HOG features. These HOG features will be used for classification and object recognition. The classification process is performed using SVM (Support Vector Machine) classifier. The SVM builds a model with a training set that is presented to it and assigns test samples based on the model. Finally get the SIFT results and HOG results, then compare both results to check better accuracy performance. The proposed method detects the number of cars more accurately

Keywords - Scale Invariant Feature Transform (SIFT), Histogram of Oriented Gradient (HOG), Support Vector Machine (SVM)

*****

I. INTRODUCTION

In the last few years, there are number of cars increasing on the roads and also in parking lots. So that some problems are generated, such as traffic jams and parking problems which are mainly found in big city areas. Nowadays, one of the classes of objects to which the research community is giving particular attention to the cars. The determination of the number of cars on roads or in parking lots represents one of the most discussed and interesting issues in the field of object detection. It can be solve urban problems that could be encountered almost every day. Especially in big cities, knowing the concentration of cars in roads or in parking lots will help us to optimize the urban traffic planning and management. Car detecting also useful for security purposes and many other applications.

The high resolution images are used for car detecting method, because these images describe the objects present in the analyzed areas with an extremely high level of detail. So it is helpful for detection process. These high resolution images are taken from very low altitude so objects in images shown clearly as compare with satellite images.

Scale invariant feature transform (SIFT) is an algorithm in computer vision to detect and describe local features in images. The SIFT algorithm was introduced by David Lowe in 1999. The algorithm is patented in the US; the owner is the University of British Columbia. Also Navneet Dalal and Bill Triggs, researchers first described HOG descriptors at the 2005 Conference on Computer Vision and Pattern Recognition (CVPR). They focused on pedestrian detection in static images The HOG algorithm used for extracting HOG features.

II. LITERATURE SURVEY

T. Moranduzzo and F. Melgani [1] presents a solution to solve the car detection and counting problem in images acquired by means of unmanned aerial vehicles (UAVs). They used screening step, SIFT keypoints and SVM classfier. F. Melgani and T. Moranduzzo [2] presents a new method for the automatic detection of cars in unmanned aerial vehicle (UAV) images acquired over urban contexts. They used screening step and filtering operations. J. Leitloff, S. Hinz and U. Stilla [3] presents an approach for automatic vehicle detection from optical satellite images. J. Gleason, A. Nefian and G. Bebis [4] introduced a real-time system for vehicle detection in rural environments from aerial images. Their approach consists of a cascade detection algorithm with the first stage serving as a fast detection solution that rejects most of the background and selects patterns corresponding to manmade objects.

T. Zhao and R. Nevatia [5] present a system to detect passenger cars in aerial images, where cars appear as small objects. They pose this as a 3D object recognition problem to account for the variation in viewpoint and the shadow. S. Hinz [6] introduces a new approach to automatic car detection in monocular large scale aerial images. The extraction is performed on a hierarchical 3D-model. W. Shao, W. Yang and G. Liu [7] propose a robust and effective framework for car detection from high-resolution aerial imagery. S. Wang [8] proposed an efficient vehicle detection algorithm for aerial images based on an improved shape matching algorithm. Q. Tan, J. Wang [9] In the paper, we adopted an object-oriented image analysis method to detect and classify road vehicles from airborne color digital orthoimagery at a ground pixel resolution of 20cm.

III. METHODOLOGY

A. Block Diagram:

- Training Phase:

Fig.1 Training Phase of Methodology

- Testing Phase:

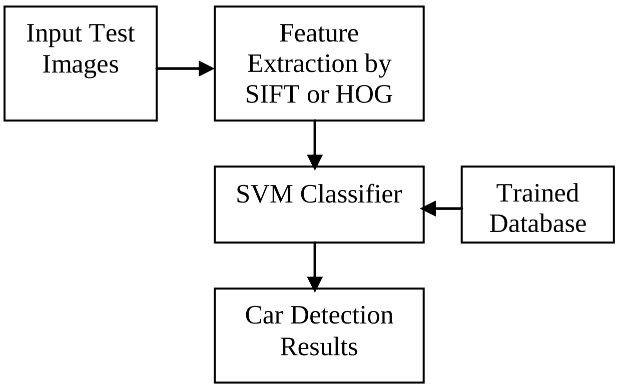

Fig.2 Testing Phase of Methodology

a) Training Phase:

The fig.1 shows the training phase of methodology, in this training phase collect different training samples such as positive samples means different car models and negative samples means background.

Now from given training samples the features are extracted using SIFT (Scale Invariant Feature Transform) algorithm. The SIFT feature extraction algorithm has four main steps such as Scale Space Detection, Keypoint Localization, Orientation Assignment and Keypoint Descriptor. These extracted SIFT features are stored in trained database which is useful in classification process. So using SIFT algorithm all the features of training samples are extracted successfully and SIFT features stored in trained database.

Similarly instead of SIFT algorithm use HOG algorithm for feature extraction process. The HOG feature extraction algorithm has also four main steps such as Gradient computation, Orientation binning, Blocks Descriptor and Blocks Normalization. So using HOG algorithm all the features of training samples are extracted successfully and HOG features stored in trained database.

b) Testing Phase:

The fig.2 shows the testing phase of methodology, In this testing phase give different input test images with high resolution. Now extract the features of test images using SIFT algorithm. Then these extracted SIFT features are apply to SVM classifier for classification process. Also the trained database from training phase is applied to SVM classifier. The SVM will compare the extracted SIFT feature of test image and the trained database with different classes. After classification process get the car detection results of SIFT algorithm i.e. SIFT results.

Similarly instead of SIFT algorithm use HOG algorithm for feature extraction process. These extracted HOG features are apply to SVM classifier for classification process. SVM will compare the extracted HOG feature of test image and the trained database with different classes. After classification process get the car detection results of HOG algorithm i.e. HOG results.

Finally after getting results of SIFT and HOG algorithm, compare the SIFT results with HOG results to check better accuracy performance between them.

B. Feature Extraction:

The identification of the features is very useful for the detection of cars. These features have to be invariant to image scale, rotation and translation also there is no effect of the illumination changes. In this method use SIFT algorithm as well as HOG algorithm.

a) SIFT algorithm:

SIFT algorithm can be transforms image data into the scale-invariant coordinates which relative to the local features. SIFT algorithm firstly extract the SIFT key points, these key points are highly distinctive. Then every identified key point will be characterized by a feature vector which describes the area surrounding such key point. In SIFT, the key points are invariant to image scaling, rotation and also partially invariant to illumination changes

SIFT algorithm consists four main steps such as [10]

1) Scale Space Detection
2) Keypoint Localization
3) Orientation Assignment
4) Keypoint Descriptor

1) Scale Space Detection:

In this first step, identify the possible locations with invariant to scale changes. Detecting possible locations by searching stable features across all possible scales using scale space. The scale space of an image is defined as a function, L(x, y, σ), it is produced by convolution between input image I (x, y) and a variable-scale Gaussian filter

$$L(x, y, \sigma) = I(x, y) * G(x, y, \sigma)$$

Where "*" is the convolution operator, σ is a scale factor.

$$\text{and } G(x, y, \sigma) = \frac{1}{2\Pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

Now to detect stable keypoint locations in scale space, scale space extrema in the difference of Gaussian (DoG) function convolved with the original image

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$$

Here k is a constant multiplicative factor which separates the new image scale from the original image.

For identification of which points become possible key points, each pixel in the DoG is compared with the 8 neighbors at the same scale and the 18 neighbors of the two other scales. If pixel is larger or smaller than all the other 26 neighbors then it is known as keypoint.

2) Keypoint Localization:
Once the possible keypoint locations are found by comparing a pixel with its neighbors. The next step is to get more accurate results. In keypoint localization, the keypoints will be rejected which have low contrast also eliminates edge keypoints and what remains is strong interest keypoints.

3) Orientation Assignment:
There are also need invariant locations to the rotation point of view, so orientation is assigned to each keypoint to achieve invariance to image rotation. To select the Gaussian smoothed image L, the scale of the keypoint will be used. For the each Gaussian smoothed image sample L(x, y) at this scale, the magnitude m(x, y) and the orientation θ(x, y) are evaluated using pixel differences

$$m(x, y) = \sqrt{(L(x+1,y)-L(x-1,y))^2 + (L(x,y+1)-L(x,y-1))^2}$$

And $$\theta(x, y) = \tan^{-1} \frac{(L(x,y+1)-L(x,y-1))}{(L(x+1,y)-L(x-1,y))}$$

4) Keypoint Descriptor:
In the last step, a vector is assigned at each key point which contains image gradients at the selected locations to give further invariance. To create key point descriptor, the gradient magnitude and the orientation at each location are computed in a region around the key point location.

The computed values are weighted by the Gaussian window. Then they accumulated into orientation histograms over 4 × 4 sub regions, with length of each arrow are to the sum of gradient magnitude near that direction within the region. The descriptor will be formed as a vector with consist of values of all the entries of orientation histogram.

b) HOG algorithm:

HOG is a feature descriptor which is used in computer vision and also in image processing for the object detection. The HOG algorithm can be implemented for extraction of HOG features. Then these features will be used for the classification process as well as object recognition.

The HOG algorithm consists four main steps as [11]-[12]
1) Gradient Computation
2) Orientation Binning
3) Blocks Descriptor
4) Blocks Normalization

1) Gradient Computation
In this first step, calculate the computation of the gradient values. The 1D centered point discrete derivative mask is a most common method used. This method requires filtering the grayscale image with the help of following filter kernels

$$D_X = [-1 \quad 0 \quad 1] \text{ and } D_Y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Consider given an image I, so using a convolution operation we obtain the x and y derivatives

$$I_X = I * D_X \quad \text{and} \quad I_Y = I * D_Y$$

The magnitude of the gradient $|G|$ and the orientation of the gradient θ are as follows

$$|G| = \sqrt{I_X^2 + I_Y^2} \text{ and } \theta = \tan \frac{I_Y}{I_X}$$

2) Orientation Binning
In the second step cell histograms will be created. A weighted vote can be casts within the cell in each pixel for an orientation based histogram channel which is based on the values found in the gradient computation. The cells are rectangular and the histogram channels are evenly spread over 0 to 180 degrees or 0 to 360 degrees, it depends on the unsigned or signed gradient.

3) Blocks Descriptor
In illumination and contrast changes, the gradient strengths can be locally normalized, it requires larger and connected blocks by grouping the cells together. Then from all of the block regions, the HOG descriptor is the vector components of the normalized cell histograms. These blocks are typically overlap, that means each cell contributes more than once to the final descriptor. There are two main block geometries available, one is rectangular R-HOG block and other is circular C-HOG block

4) Blocks Normalization
There are various different methods for block normalization. Consider v is the non-normalized vector which contains all histograms in a given block. $\|v_k\|$ is its k-norm for k = 1, 2 and e is the some small constant. Then the normalization factor will be one of the following

$$\text{L2 norm: } f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}}$$

$$\text{L1 norm: } f = \frac{v}{\|v\|_1 + e}$$

$$\text{L1 sqrt: } f = \sqrt{\frac{v}{\|v\|_1 + e}}$$

C. Classification:

- SVM Classifier:

Once the feature extraction is done by using SIFT and HOG algorithm then classification process is performed by using SVM classifier.

Support Vector Machine (SVM) is a supervised learning algorithm that is used for classification and regression analysis. The SVM builds a model with a training set that is presented to it and assigns test samples based on the model. An SVM model represents points of samples in space, mapped in a way that the samples of the separate categories are divided by a clear gap that is as wide as possible. The challenge is to find the optimal hyperplane that maximizes the gap. New samples are then mapped into that same space and class predictions are made as belonging to either category based on which side of the gap they fall on. The performance of the SVM is greatly dependent on its kernel functions.

IV. EXPERIMENTAL RESULTS

A. Dataset Description:

a) Training Database:

Fig.3 Positive Training Samples (20 samples)

Fig.4 Negative Training Samples (20 samples)

In this paper, we use set of high resolution images which are captured by camera. The camera name is Fujifilm Fine Pix HS35EXR with 16 mp. The places in the pune city where the images are taken are SNDT region, karve road area, pune university region, COEP area.

MATLAB14 software is used for method implementation. In the training database we use 50 positive training samples such as different car models and 80 negative training samples as background. The more informative the samples are, the better the trained model will be. The fig.3 & fig.4 shows some 20 positive and 20 negative samples respectively.

b) Testing Database:

In the testing database, we use 5 test image samples. The fig.5 shows 5 testing image samples.

Test Image 1

Test Image 2

Test Image 3

ISSN: 2321-8169
197 - 203

Test Image 4

Test Image 5

Fig.5 Testing Image Samples (5 samples)

B. Final Results:

We take Test Image 1 for example and see the SIFT result & HOG result as shown in fig.6 and fig.7

Fig.6 SIFT result of Test Image 1

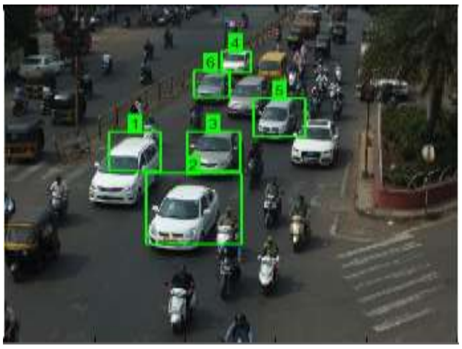

Fig.7 HOG result of Test Image 1

Fig.8 Final result of Test Image 1

Fig.8 shows final result of Test Image 1 that is comparison between SIFT & HOG results. In this final result there are 9 cars are present in the image where SIFT results gives 5 TP (True Positive = correctly detected cars) & 3 FP (False Positive = incorrectly detected cars)) also HOG results gives 6 TP and 0 FP. So we can say that HOG results give maximum TP and minimum FP as compare to SIFT results.

Similarly get the final result of other Test Images such as Image 2, Image 3, Image 4 and Image 5 which shown in fig. 9, fig.10, fig.11, and fig. 12 respectively.

Fig.9 Final result of Test Image 2

Fig.10 Final result of Test Image 3

Fig.11 Final result of Test Image 4

Fig.12 Final result of Test Image 5

C. Accuracy Percentage:

To assess the capability of our methodology to correctly detect the number of cars in the test images, we adopted a measure of accuracy based on results of two methods such as SIFT and HOG algorithm.

The producer's accuracy (Pacc) is compute by dividing the number of correctly detected cars for a class by the total number of cars, N (number of cars present in the image).

$$\text{Pacc} = \frac{\text{TP}}{\text{N}}$$

The user's accuracy (Uacc) is calculated by dividing the number of correct samples for a class by the total number of samples assigned to that class.

$$\text{Uacc} = \frac{\text{TP}}{\text{TP}+\text{FP}}$$

The final or total accuracy, which we adopted to quantify our results, is the average of the two previous accuracies, as

$$\text{Tacc} = \frac{\text{Pacc}+\text{Uacc}}{2}$$

Such accuracy has the advantage of taking into account both the number of car correctly classified and the number of false alarms.

D. Observation Table for Accuracy:

a) SIFT Accuracy Results:

Test Image	TP	FP	N	Pacc (%)	Uacc (%)	Tacc (%)
Image 1	5	3	9	55.55	62.5	59.02
Image 2	3	3	6	50	50	50
Image 3	3	6	7	42.85	33.33	38.09
Image 4	4	5	5	80	44.44	60.2
Image 5	3	2	7	42.85	60	51.42
Total/Avg.	18	19	34	54.25	50.05	51.74

Table 1 SIFT Accuracy Results

b) HOG Accuracy Results:

Test Image	TP	FP	N	Pacc (%)	Uacc (%)	Tacc (%)
Image 1	6	0	9	66.66	100	83.33
Image 2	6	0	6	100	100	100
Image 3	5	1	7	71.42	83.33	77.37
Image 4	4	1	5	80	80	80
Image 5	6	1	7	85.71	85.71	85.71
Total/Avg.	27	3	34	80.75	89.80	85.28

Table 2 HOG Accuracy Results

c) Accuracy for SIFT & HOG with SVM classifier:

Feature Set	Classifier	% Accuracy
SIFT	SVM	51.74
HOG	SVM	85.28

Table 3 Accuracy for SIFT & HOG with SVM classifier

d) Average Time Required for SIFT & HOG:

Test Image	Execution time for SIFT (sec.)	Execution time for HOG (sec.)
Image 1	13.63	4.15
Image 2	12.27	4.03
Image 3	11.78	3.91
Image 4	12.33	4.29
Image 5	12.77	3.48
Average Time	12.55	3.97

Table 4 Average Time Required for SIFT & HOG

Table 1 & 2 shows SIFT & HOG Accuracy Results respectively. In this we use 5 test images for testing & calculate TP, FP, Pacc, Uacc & Tacc for all 5 test image then finally we get average accuracy result. So average accuracy of SIFT results is 51.74% & average accuracy of HOG result is 85.28%. Table 3 shows Accuracy for SIFT & HOG with SVM classifier. So HOG algorithm gives better accuracy results as compare to SIFT algorithm.

Table 4 shows average time required for SIFT & HOG. We calculate execution time of all 10 test images, then finally we get average time. So average time required for SIFT is 12.55 sec & average time required for HOG is 3.97 sec. We can say that HOG algorithm is less time consuming as compare to SIFT algorithm.

V. CONCLUSION

A car detection method is implemented using high resolution images, because these images give high level of object details in the image. There are 2 feature extraction algorithms are used for implementation such as SIFT and HOG algorithm. Also classification process performed using SVM classifier. Finally get the SIFT results and HOG results, then compare both results to check better accuracy performance. We use 5 test images for testing & calculate TP, FP, Pacc, Uacc & Tacc for all 5 test images then finally we get average accuracy results. So average accuracy for SIFT is 51.74% & average accuracy for HOG is 85.28%. So HOG algorithm gives better accuracy results as compare to SIFT algorithm. Also average time required for SIFT is 12.55 sec & average time required for HOG is 3.97 sec. So HOG algorithm is less time consuming as compare to SIFT algorithm.

ACKNOWLEDGEMENTS

The author would like to thank Mrs. R.R. Itkarkar for valuable support and guidance for completing this work.

REFERENCES

[1] T. Moranduzzo and F. Melgani, "An automatic car counting method for unmanned aerial vehicle images,"

IEEE Trans. on Geoscience and Remote Sensing ,vol. 52, no. 3, pp. 1635–1647, Mar. 2014

[2] T. Moranduzzo, and F. Melgani, "Detecting Cars in UAV Images With a Catalog Based Approach", *IEEE Trans. on Geoscience and Remote Sensing,* vol. 52, no. 10, Oct. 2014

[3] J. Leitloff, S.Hinz and U. Stilla, "Vehicle Detection in Very High Resolution Satellite Images of City Areas", *IEEE Trans. on Geoscience and Remote Sensing,* vol. 48, no. 7, July. 2010

[4] J. Gleason, A. V. Nefian, and G. Bebis, "Vehicle detection from aerial imagery," inProc. IEEE Int. Conf. Robot. Autom., May 2011, pp. 2065–2070

[5] T. Zhao and R. Nevatia, "Car detection in low resolution aerial images," *inProc. IEEE Int. Conf. Comput. Vis*., 2001, pp. 710–717.

[6] S. Hinz, "Detection and counting of cars in aerial images," *inProc. IEEE Int. Conf. Image Process*, 2003, pp. 997–1000

[7] W. Shao, W. Yang, G. Liu, and L. J. , "Car detection from high resolution aerial imagery using multiple features," *inProc. IGARSS*, 2012, pp. 4379–4382.

[8] S. Wang, "Vehicle detection on aerial images by extracting corner features for rotational invariant shape matching," *in Proc. IEEE Int. Conf. Comput. Inf. Technol*., Aug.–Sep. 2011, pp. 171–175.

[9] Q. Tan, J. Wang, and D. A. Aldred, "Road vehicle detection and classification from very-high-resolution color digital orthoimagery based on object-oriented method," inProc. IEEE Int. Geosci. Remote Sens. Symp., vol. 4. Jul. 2008, pp. 475–478

[10] D. Lowe, "Distinctive image features form scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004

[11] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 886–893, 2005

[12] Histogram of Oriented Gradient, Pattern Recognition Systems – Lab 5