NEURAL CLASSIFIER FOR OBJECT CLASSIFICATION WITH CLUTTERED BACKGROUND USING STATISTICAL CENTRAL MOMENT BASED FEATURES

Nagarajan .B[1] and Balasubramanie .P[2]
[1]Department of Computer Applications, Bannari Amman Institute of Technology, Sathyamangalam, Tamil Nadu, India
[2]Department of Computer Science and Engineering, Kongu Engineering College, Perundurai, Tamil Nadu, India
E-Mail: [1]bnr_mca@rediffmail.com

Abstract

Object classification in static images is a difficult task since motion information in no longer usable. The challenging task in object classification problem is the removal of cluttered background containing trees, road views, buildings and occlusions. The goal of this paper is to build a system that detects and classifies the car objects amidst background clutter and mild occlusion. This paper addresses the issues to classify objects of real-world images containing side views of cars with cluttered background with that of non-car images with natural scenes. The threshold technique with background subtraction is used to segment the background region to extract the object of interest. The background segmented image with region of interest is divided into equal sized blocks of sub-images. The statistical central moment based features are extracted from each sub-block. The features of the objects are fed to the back-propagation neural classifier. Thus the performance of the neural classifier is compared with various categories of block size. Quantitative evaluation shows improved results of 93.8%. A critical evaluation of our approach under the proposed standards is presented.

Key words: Object Classification, Background Segmentation, Statistical Central Moments, Neural Classifier

I. INTRODUCTION

Object detection and classification are necessary components in an artificially intelligent autonomous system. Especially, object classification plays a major role in applications such as security systems, traffic surveillance system, target identification, etc. We expect these artificially intelligent autonomous systems to venture onto the street of our world, thus requiring detection and classification of car objects commonly found on the street. In reality, these classification systems face two types of problem. (i) Objects of same category with large variation in appearance. (ii) The objects with different viewing conditions like occlusion, complex background containing buildings, people, trees, road views, etc. This paper tries to bring out the importance of the background elimination with statistical based feature extraction method of varying sub-block size for object classification. Since dynamic motion information is no longer usable for static images, background elimination becomes a more difficult task. Thus background removed and statistical features of squared sub-blocks of the images are fed to the neural classifier. The objects of interest being a car and non-car images are classified.

Image understanding is a major area where researchers design computational systems that can identify and classify objects automatically. A new approach to object detection that makes use of a sparse, part-based representation is proposed by Agarwal et al. [1]. This study gives very promising results in the detection of vehicles from a group of non-vehicle category of natural scenes. Nagarajan and Balasubramanie [2], [3] have proposed their work based on wavelet and moment invariant features towards object classification with cluttered background. Nagarajan and Balasubramanie [4] have also made a study on object classification based on statistical texture features with cluttered background. Xu and Li [5] have presented their work on geometric moment invariants to describe and represent the shape of solid objects. A comprehensive study for classification of texture and object categories using local features has been presented by Zhang and Marszalek [6]. Identification and classification of vehicles has been a focus of investigation over last decades [8], [11], [12].

II. BACKGROUND REMOVAL & MAPPING FUNCTION

The overall complexity increases for the natural images as the object of interest is lying on the background region. In object classification problem, it is essential to distinguish the object of interest and the background. Segmentation of object is done through background subtraction technique. This method is more suitable when the intensity levels of the objects fall outside the range of levels in the background.

An object with natural background is shown in Fig. 1. Initially morphological operations are applied to suppress the residual errors with help of open and close pair statements [7], [10]. The small regions are removed by filling the holes.

Then the image subtraction is applied with the previous result. Thus the object is segmented from the background. The gray level intensity is restored for the region of interest through the mapping function (1).

$$f(x,y) = \begin{Bmatrix} 0, \ if \ \ d(x,y)=0 \\ I(x,y), \ Otherwise \end{Bmatrix} \tag{1}$$

Where, *f(x,y)* is the transformed image, *d(x,y)* is image difference after fill operation and *I(x,y)* is the original image.

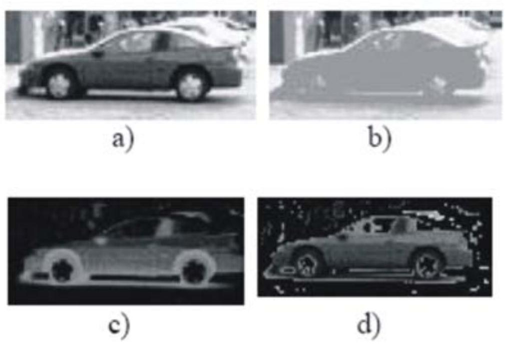

Fig. 1: a) A sample image with complex background denoted as *I(x,y).* b) The small regions are removed by filling the holes. c) Image difference obtained by subtraction (a) by (b) denoted as *d(x,y).* d) Gray level intensity restored for the region of interest through mapping function *f(x,y).*

III. STATISTICAL FEATURES

Statistical functions such as mean, median, standard deviation and moments are most common to characterize data set, which have been used as pattern features in many applications [3-6]. One of the principal approaches for describing the shape of a histogram is via its central moments. This is also called moments about the mean.

Let Z_i be a discrete random variable that denotes intensity levels in an image, and let ($p(z_i)$, I = 0,1,2,......... L - I, be the corresponding normalized histogram, where *L* is the number of possible intensity values. A histogram component $p\ (z_j)$ is an estimate of the probability of occurrence of intensity value z_j and the histogram may be viewed as an approximation of the intensity probability density function. Thus the central moments are defined in (2).

$$\mu_n = \sum_{i=0}^{L-1} (z_i - m)^n \ p(z_i) \tag{2}$$

Where, *n* is the moment order, and *m* is the mean (3).

$$m = \sum_{i=0}^{L-1} z_i \ p(z_i) \tag{3}$$

Since the histogram is assumed to be normalized, the sum of all its components is 1, so, from the preceding equations (2) and (3), 10=μ_0=1 and .01=μ_1=0 The second moment (4) is the variance.

$$\mu_2 = \sum_{i=0}^{L-1} (z_i - m)^2 \ p(z_i) \tag{4}$$

Features extraction is done by computing the mean (3) and statistical central moments up to order *n,* of a histogram whose components are in vector P. The length of P must equal 256 (2^8) or 65536 (2^{16}). The moment of order 0 is always 1, and the moment of order 1 is always 0. So, these two moments are ignored and output feature vector V is defined as,

$$\begin{aligned} V(1) &= \text{mean} \\ V(2) &= \text{variance} \\ V(3) &= 3^{rd} \text{ central moment} \\ V(4) &= 4^{th} \text{ central moment} \\ V(5) &= 5^{th} \text{ central moment} \\ \text{and } V(6) &= 6^{th} \text{ central moment} \end{aligned} \tag{5}$$

Features extracted values (5) are normalized to the range [0, 1], so all moments are in this range. Thus six statistical measures of central moments are calculated for every sub-block of an image. The feature vector is populated with multiples of six with that of number of sub-blocks in an image.

IV. BUILDING A NEURAL CLASSIFIER

A binary Artificial Neural Network (ANN) classifier is built with back-propagation algorithm that learns to classify an image as a member or nonmember of a class. The number of input layer nodes is equal to the dimension of the feature space obtained from the statistical central moment features. The number of output nodes [9] is usually determined by the application, which is 1 (either "Yes/No") where, a threshold value nearer to 1 represents "Yes" and a value nearer to 0 represents "No". The neural classifier is trained with different choices for the number of hidden layer. The final architecture is chosen with single hidden layer shown in Fig. 2 that results with better performance.

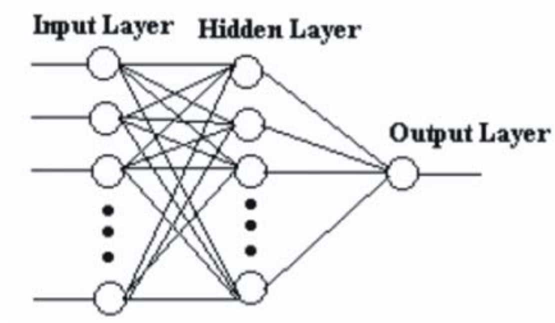

Fig. 2: The Three Layer Neural Architecture

The connections carry the outputs of a layer to the input of the next layer have a weight associated with them. The node outputs are multiplied by these weights before reaching the inputs of the next layer. The output neuron (6) will be representing the existence of a particular class of object.

$$O_j^l(k) = f\left(\sum_{m=0}^{N_{l-1}} w_{jm}^l O_m^{l-1}\right) \tag{6}$$

V. PROPOSED WORK

This paper addresses the issues to classify objects of real-world images containing side views of cars amidst background clutter and mild occlusion. The objects of interest to be classified are car (positive) and non-car (negative) images taken from University of Illinois at Urbana-Champaign (UIUC) standard database. The image data set consists of 1000 real images for training and testing having 500 in each class. The sizes of the images are uniform with the dimension 40X100 pixels.

The proposed framework consists of three methods followed by background removal as given in section II. Method-I: 10 Blocks of size 20x20 each, Method-II: 40 Blocks of size 10x10 each and Method-III: 160 Blocks of size 5x5 each. Six statistical features are calculated from each single block of sub-image using equations mentioned in section-III. Data normalization is applied for the statistical features, which are the deviated from its mean by standard deviation. This process improves the performance of the neural classifier. The overall flow of the framework is shown in Fig. 3.

VI. IMPLEMENTATION

We trained our methods with different kinds of cars against a variety of background, partially occluded cars of positive class. The negative training examples include images of natural scenes, buildings, and road views. The training is done with 400 images (200 positive and 200 negative) against all the methods. The testing of images are done with 1000 images (500 positive and 500 negative) taken from the same image database.

The feed-forward network for learning is done for 10 blocks of size 20x20 namely method-I, 40 blocks of size 10x10 namely method-II and 160 blocks of size 5x5 namely method-III respectively. The input nodes for method-I is 60 (10 blocks x 6 features), method-II is 240 (40 blocks x 6 features) and method-III is 960 (160 blocks x 6 features) respectively. Optimal structure validation is done and the structure given in Fig.2 performs well and leads to better results. Thus the optimal structure (Fig. 2) of the neural classifier for method-I is 60-20-1, method-II is 240-10-1 and method-III is 960-10-1 respectively.

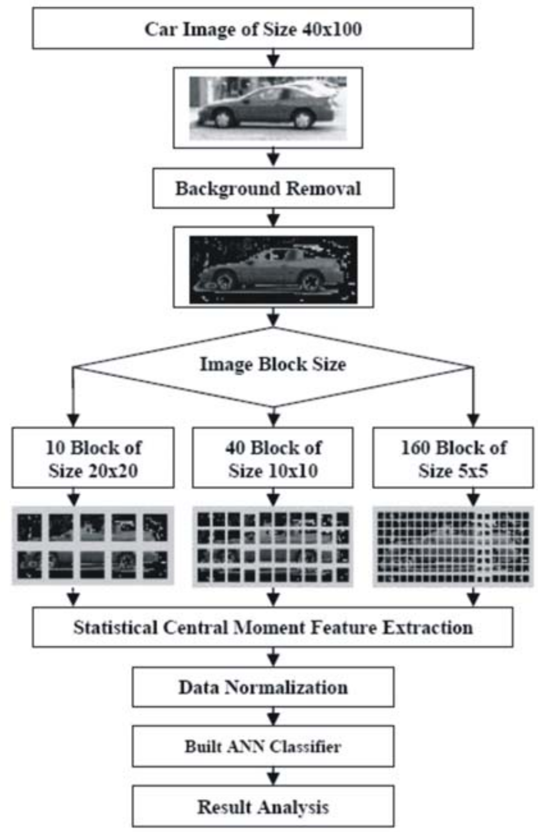

Fig. 3 : The Description of The Proposed Work.

The various parameters for the Neural classifier training for all the methods are given in Table 1. The Performance graph of the neural classifier for method-I method-II and method-III are shown in Fig.4, Fig.5 and Fig.6 respectively.

Table 1 : Parameters for training of the neural classifier

Parameters	Method-I	Method-II	Method-III
Learning Rate	0.5	0.5	0.5
Performance Goal	0.01	0.01	0.01
No. of Epochs taken to meet the performance goal.	10501	321	233
Time taken to learn	105.32 Secs	7.312 Secs	12.375 Secs

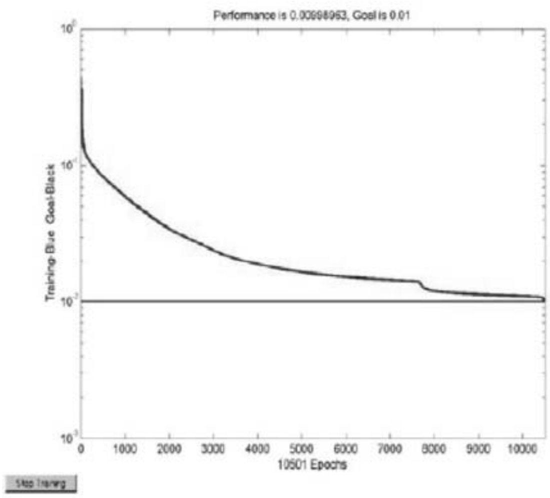

Fig. 4. The Performance Graph of Neural Network Training For Method-I: 10 Blocks of size 20x20.

VII. DISCUSSION

In object classification problem, the four quantities of results category are given below.

(I) True Positive (TP) : Classify a car image into class of cars.

(ii) True Negative (TN) : Misclassify a car image into class of Non-cars.

(iii) False Positive (FP) : Classify a non-car image into class of non-cars.

(iv) False Negative (FN) : Misclassify a non-car image into class of cars.

The objective of any classification is to maximize the number of correct classification denoted by True Positive Rate (TPR) and False Positive Rate (FPR) where by minimizing the wrong classification denoted by True Negative Rate (TNR) and False Negative Rate (FNR).

$$TPR = \frac{Number\ \ of\ true\ \ positive\ \ (TP)}{Total\ \ number\ \ of\ \ positive\ \ in\ data\ \ set\ \ (nP)}$$

$$TNR = \frac{Number\ \ of\ true\ negative\ \ (TN)}{Total\ number\ \ of\ negative\ \ in\ data\ set\ (nN)}$$

$$FPR = \frac{Number\ \ of\ \ false\ \ positive\ \ (FP)}{Total\ number\ \ of\ \ positive\ \ in\ data\ set\ (nP)}$$

$$FNR = \frac{Number\ \ of\ \ false\ negative\ \ (FN)}{Total\ number\ \ of\ negative\ \ in\ data\ set\ (nF)}$$

The values of *nP* and *nN* used as testing samples are 500 and 500 respectively. Most classification algorithm includes a threshold parameter for classification accuracy which can be varied to lie at different trade-off points between correct and false classification. The comparison of results of the proposed methods is shown in Table 2 which is obtained with an activation threshold value of 0.7. Classified images of category car and non-car as resultant sample images are shown below in the Fig. 7 and Fig. 8 respectively.

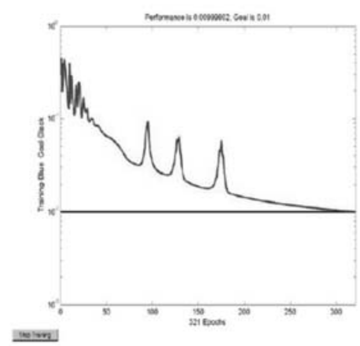

Fig. 5. The Performance Graph of Neural Network Training For Method-II: 40 Blocks of size 10x10.

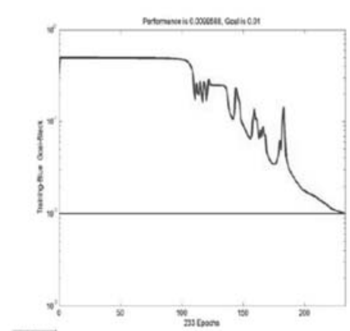

Fig. 6. The Performance Graph of Neural Network Training For Method-III: 160 Blocks of size 5x5.

It is evident that the classifier with 160 blocks of size 5x5 (Method-III) is showing improved overall results of 93.8% of classification accuracy comparatively with that of 40 blocks of size 10x10 (Method-II) and 10 blocks of size 20x20 (Method-I). Method-II is also comparatively good in classification with the accuracy of 93.2%.

Table 2. Comparison of experimental methods

Threshold for classification : 0.7	Classifying Positive Images (Car Images)			Classifying Negative Images (Non-Car Images)
	TPR	TNR	**FPR**	FNR
Method-I : 10 Blocks of size 20x20	**86.8 %**	13.2 %	**84.6 %**	15.4 %
	Method-I : Overall Classification Accuracy (TPR+FPR)/2 is **85.7%**			
Method-II : 40 Blocks of size 10x10	**90.4 %**	9.6 %	**96.0 %**	4.0 %
	Method-II : Overall Classification Accuracy (TPR+FPR)/2 is **93.2%**			
Method-III : 160 Blocks of size 5x5	**89.8 %**	10.2 %	**97.8 %**	2.2 %
	Method-III :Overall Classification Accuracy (TPR+FPR)/2 is **93.8%**			

Fig. 7. Sample results of the neural classifier of the category car images with cluttered background and mild occlusion.

Fig. 8. Sample results of the neural classifier of the category non-car images containing trees, road view, bike, wall, buildings and persons.

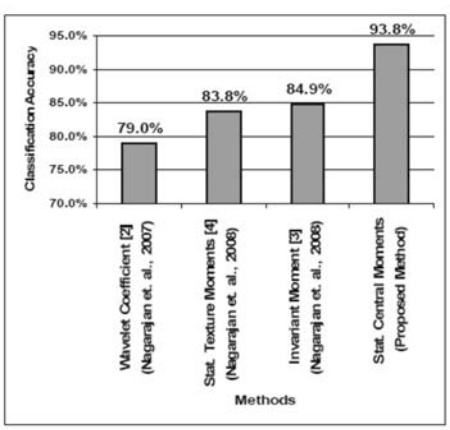

Fig. 9 Classification Accuracy of
The Proposed Method With Previous Works

VIII. CONCLUSION

Thus an attempt is made to build a system that classifies the objects amidst background clutter and mild occlusion is achieved to certain extent. Thus the goal to classify objects of real-world images containing side views of cars with cluttered background with that of non-car images with natural scenes is presented. Comparing the results in Table 2, the performance of the proposed method with 160 blocks of size 5x5 with statistical central moment based features after background removal gives a satisfactory classification rate of 93.8%. The classification accuracy has improved substantially in comparison with previous works [2-4] as depicted in Fig. 9. The limitation of this method is the object with a high degree of occlusion for classification. Further work extension can be made to improve the performance of the classifier system with the inclusion of feature selection process. This complete work is implemented using neural network and image processing toolbox of Matlab 6.5.

ACKNOWLEDGMENT

The authors would like to thank the software MATLAB from Mathworks and UIUC standard image database. They are also grateful to the anonymous reviewers of our works for their insightful comments that helped to improve the quality of the work.

REFERENCES

[1] Agarwal, S., A. Awan, and D. Roth, 2004, Learning to Detect Objects in Images via a Sparse, Part-Based Representation, IEEE Trans. on Pattern Anal. and Machine Intell., 26 (11), 1475-1490.

[2] Nagarajan, B., and P. Balasubramanie, 2007, Wavelet Feature based Neural Classifier System for Object Classification with Complex Background, iccima07, IEEE Computer Society Press, 1, 302-307.

[3] Nagarajan, B., and P. Balasubramanie, 2008, Neural Classifier System for Object Classification with Cluttered Background Using Invariant Moment Features, International Journal of Soft Computing, 3(4), 302-307.

[4] Nagarajan, B., and P. Balasubramanie, 2008, Object Classification in Static Images with Cluttered Background Using Statistical Feature Based Neural Classifier, Asian Journal of Information Technology, 7(4), 162-167.

[5] Dong Xu and Hua Li, 2008, Geometric Moment Invariants, Pattern Recognition, 41, 240-249.

[6] Zhang, J and M. Marszalek, 2006, Local Features and Kernels for Classification of Texture and Objects Categories: A Comprehensive Study, International Journal of Computer Vision, 10, 1-26.

[7] Richord J. R., *et al*., 2005, Image Change Detection Algorithms :A Systematic Survey, IEEE Trans. Image Proc., 14(3), 294–306.

[8] Hsieh, J.W. *et al.,* 2006, Automatic traffic surveillance system for vehicle tracking and classification, IEEE Trans. Intell. Trans. Sys., 7(2), 175-187.

[9] Khotanzand, A. and C. Chung, 1998, Application of Multi-Layer Perceptron Neural Networks to Vision Problem. Neural Computing & Applications, Springer-Verlag London Limited, 249-259.

[10] Li, L. *et al*., 2004, Statistical modeling of complex backgrounds for foreground object detection, IEEE Trans. Image Proc., 13(11), 1459–1472.

[11] Shan, Y. *et al.,* 2005, Vehicle Identification between Non-Overlapping Cameras without Direct Feature Matching, 10th IEEE International Conf. on Computer Vision, pp: 1550-1558.

[12] Sun, Z., *et al.,* 2006, Monocular Precrash Vehicle Detection: Features and Classifiers, IEEE Trans. Image Proc., 15, 2019-2034.

B. Nagarajan received MCA degree from Madras University, India in 1997, and M.Phil. degree in Computer Science from Manonmaniam Sundaranar University, India in 2002. Currently he is pursuing the Ph.D. His area of interest in research includes Image Processing and Neural Networks. He has published 11 papers in National/International Conferences of repute and 3 papers in International Journals. He has worked as a Co-Investigator for the DRDO funded research project in the area of Image Processing. He is a Life member of Indian Society of Technical Education (ISTE).

P. Balasubramanie post graduated from Bharathiar University, India in 1990. He obtained his M.Phil. Degree in Mathematics and Ph.D. Degree in Discrete Mathematics from Anna University in 1992 and 1996 respectively. He was awarded Research fellowship by Council of Scientific and Industrial Research (CSIR) in 1990. He has published more than 20 papers in National and International Journals. He is the author of three books, One on Operational Research and other two are on Theory of Computation. His area of interest includes Discrete Mathematics, Theoretical Computer Science and Image Processing. Presently he is working as Professor of Computer Science and Engineering, Kongu Engineering College, Tamil Nadu, India.