IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 11, November 2015

Detect Pedestrian Orientation by Integrating Multiclass SVM Utilizing Binary Decision Tree

G. Santoshi[1], G Gowri Pushpa[2]

Asst. Prof, CSE, ANITS Engineering College, Visakhapatnam, India[1,2]

Abstract: In Driver assistance systems the pedestrian protection is an essential component which should be able to predict the probability of collision after detecting the pedestrian and it is also important to consider all the cues available in order to make the prediction. One such cue is the direction in which the pedestrian is facing which could be used in predicting where the pedestrian may move in future. The proposed method describes a novel approach for solving the problem of Pedestrian orientation classification by SVM based Binary Decision Tree (SVMBDT) architecture for solving multiclass problem. The hierarchy of binary decision tree subtasks the use of SVM for decision making. The proposed algorithm takes advantage of both the efficient computation of decision tree and high classification accuracy of SVM. Experiment shows the performance of estimating orientation and describes to show the promise of the approach.

Keywords: Orientation, Multiclass problem, SVMBDT, Binary decision tree.

I. INTRODUCTION

Pedestrian detection is an important and necessary task in Driving Assistance system and an obvious extension to automotive applications to improving safety systems. An environment with a large number of pedestrian moving particularly in the busy streets, still it is challenging to estimate the future states required for the decision-making and path planning. To detect pedestrian with direction will involve large number of categories or class types (Back, Front, Right ,Left and Non-Pedestrian).

Several surveys are made to solve the problem of multiclass approach .One of the popular methods is applying SVMs to multi-class classification problems by decomposing the multi-class problems into several two-class problems and assemble several binary classifiers whose outputs are then composed to decide the direction class. Some of the approaches are One against all and One against one[10] .First method is One against all in which N 2-class SVM classifiers are constructed[2].The i^{th} training set is trained while labelling all the samples in the i^{th} class as positive examples and the rest as negative examples. During recognition, a test example is presented to all N SVMs and labelled according to the maximum output among the N classifiers. The demerit of this method is the number of training samples is too large and difficult to train. Second method is [2] One against one, constructs N (N-1)/2 2-class classifiers using all the binary pair-wise combinations of the N classes. Every classifier is trained using the samples of the first class as positive examples and the samples of the second class as negative examples. This results in faster training but slower testing, especially when the number of the classes in the problem is big.

We will focus on pedestrian orientation classification performance, by using Multi-class SVM with Decision tree technique as our pedestrian detection method for solving the multiclass problem. The processes of learning splits of binary decisions are done using SVM and the dataset for training of SVM is grouped using mean shift clustering. The proposed algorithm takes advantage of both of the decision tree architecture and SVMs. As, decision tree is good at efficient computation and SVM for it's the high classification accuracy.

Additionally, we will show that pose classification minimizes the use of computing resources. Numerous works showed the advantage of using tracking for improving the detection. However, we chose to work with the still images as predicting the direction of the pedestrian stands on the sidewalk near to a zebra crossing and do not move. Tracking at this stage of detection would not be beneficial in multiple inner-city scenarios. The paper is organized as follows 2) Explain the proposed work.3) The experimental result shows accuracy and trace the pedestrian along with the direction. Final the paper contains conclusions with future works.

II. PROPOSED WORK

We propose an architecture which uses binary decision tree SVMs for learning the splits at the nodes. To deploy this architecture N-1 SVMs need to be trained for an N-class [9] problem (like in one-against-all), but on the average only [$\log_2$ N] SVMs are required to be consulted to classify a sample, which can lead to a pronounced improvement in pedestrian detection along with direction. Our orientation estimate involves approximating the density function of pedestrian body orientation. This is quite unlike [16], where pedestrian heading is only recovered in terms of predefined direction classes, e.g. front, back, right etc., work on multi-class classification techniques. Such classes are completely contained in our approach by integrating the density function [7]. The proposed algorithm detects the direction in two phases Learning and Detection. During learning phase the Feature Extraction of the pedestrian is done

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 11, November 2015

using Histogram Orientation of Gradients(HOG) ,the different combination of classes with different directions are clustered into one group depending on the similarities using Mean shift clustering algorithm. Once the classes are defined, training of multiclass is done using SVM classifiers.

During detection the given samples direction is detected using Binary decisions where the processes of learning splits of binary decisions is done using SVM. Feature extraction [12] using HOG have been extensively studied for object recognition. In this work we demonstrate how a mean shift clustering algorithm applied to identify the orientation classes of positive and negative datasets and its effectiveness by using binary decisions for splitting the categorization tasks.

2.1 Mean Shift Clustering Algorithm

The pedestrian direction is classified into multiple classes. There prevail many ways to divide N classes into 2 groups, and it is necessary to have proper grouping for the good performance of SVM-BDT. The hierarchy of binary decision independent tasks should be carefully designed before the training of each binary SVM classifier. Mean shift is a procedure for locating the maxima of a density function from a given sampled discrete data [7]. It is useful for detecting the modes of this density. We start with an initial estimate x, and the iteration is applied. Let a centroid function $K(x_i - x)$ be given. This is used to determine the weight of nearby points for re-estimation of the mean. The weighted mean of the density in the window determined by K is

$$m(x) = \frac{\sum_{x_i \in N(x)} K(x_i - x) x_i}{\sum_{x_i \in N(x)} K(x_i - x)} \quad (1)$$

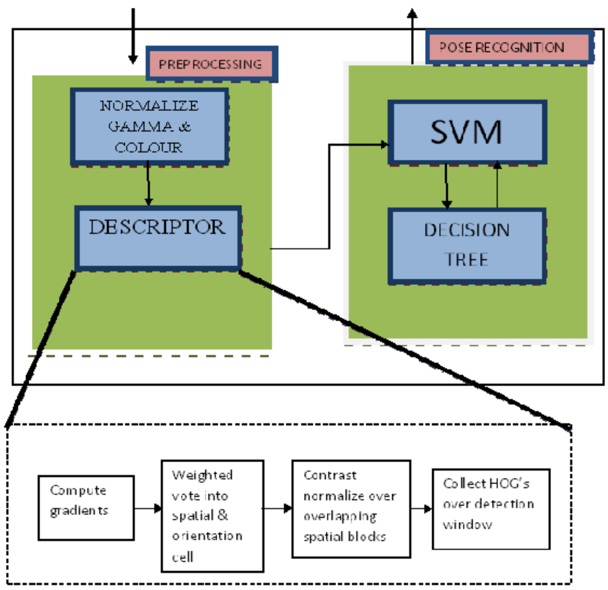

Fig.1 Flowchart for Pedestrian detection using SVM with Decision tree

where $N(x)$ is the neighborhood of x, a set of points for which $K(x) \neq 0$.

$K(x_i - x) = e^{-c\|x_i - x\|^2}$, Gaussian kernel on the distance to the current estimate. The mean-shift algorithm now sets $x \leftarrow m(x)$, and repeats the approximation until $m(x)$ converges.

The mean shift clustering method divides the classes' recursively into two disjoint groups depending on the similarity of the features. SVM is trained using these two disjoint groups, instead of following one-against-all and One against one method for training. The steps for the algorithm are

1)The number of clusters are defined as k=2 to obtain a binary decisions. The input to the Mean shift is the predictions values of the pedestrian feature values dataset generated using HOG. From which the centroids are considered in a cunning way as different locations causes different results.
2)Next step takes each points belonging to a given dataset which are associated to the nearest centroid are grouped together.
3)At this point we need to re-calculate k new centroids as braycenters of the clusters resulting from the previous step
4) After we have these k new centroids, a new binding of the nearest new centroid and the same data set points is made.
5)Repeat Steps 2 and 3 until the centroids no longer move. These results a separation of the objects into groups from which the metric to be minimized can be calculated.

Suppose that we have n sample test samples the vectors $\mathbf{x}_1$, $\mathbf{x}_2$... $\mathbf{x}_n$ are the predictions of same images, and we know that they fall into 2 different clusters, the clusters are well separated using Euclidean-distance classifier to separate them. We can say that $\mathbf{x}$ is in cluster i if $\| \mathbf{x} - \mathbf{m}_i \|$ is the minimum of all the k distances.

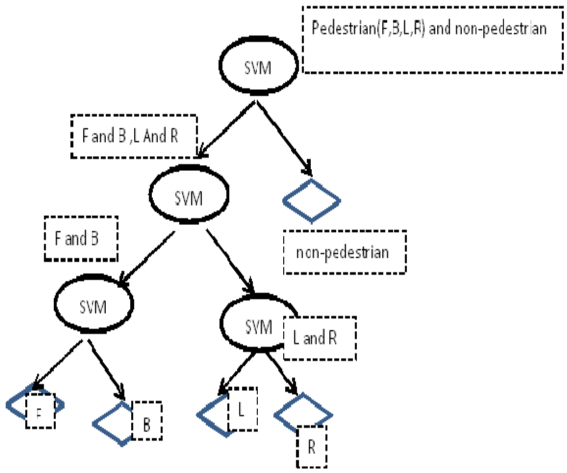

Fig.2 Learning Splits using SVM

2.2 SVM utilizing Binary Decision Tree (SVM-BDT)

The SVM-BDT solves an N-class pedestrian direction detection problem utilizing a binary decision tree, which recursively dividing the classes in two disjoint groups at every node of the decision tree and training a SVM that

DOI 10.17148/IJARCCE.2015.411120

will decide in which group the incoming unknown sample should be assigned. The recognition of each pedestrian starts at the root of the decision tree. At each node of the binary tree a decision is being made about the assignment of the input pedestrian into one of the two possible groups represented by transferring the image to the left or right sub-tree ,which will decide in which of the direction the incoming unknown sample matches appropriately.

Assume that we have N classes of images at a given node. We uniformly sample N binary variables (b) and assign all examples of a particular class c_i a class label of b_i. This method improves the scalability of our method to a large number of classes and results in well-balanced trees as the binary splits are performed at each node using SVM. Using the feature representation f of an image region as described in Section 2, we find a binary split of the data:

$$\begin{cases} w^T f \leq 0, \text{ go to left child} \\ \text{otherwise, go to right child} \end{cases}$$

where w is the set of weights learned from a linear SVM.

We evaluate each binary split that corresponds to an image region or pairs of regions with the information gain criteria, which is computed from the complete training images that fall at the current tree node. The splits that maximize the information gain are selected and the splitting process is repeated with the new splits until the tree reaches maximum tree depth. For example, on Fig.2 which illustrates clustering of 5 classes, the SVM classifier in the root is trained by considering samples from the classes {Pedestrian(Front, back, right, Left} as positives examples and samples from the classes {*non-pedestrian*} as negative examples. The SVM classifier in the left child of the root is then trained by considering samples from the classes {*Back*, *Front*} as positives and samples from the classes {*right*, *left*} as negative examples. This way, we will train (k−1) SVMs for k-class problem.

III. EXPERIMENTAL RESULTS

This section aims to prove that the proposed algorithm of Multiclass SVM Utilizing Binary Decision Tree is efficient and reduce the computability time. The algorithm uses the INRIA dataset which consist of 2000 images. To make a balanced decision splits. The trained dataset for Left facing are 1164, Right facing are 1603, front facing are 874 and the back facing are 1003 and Non-pedestrian are 453 all are of size [64,128] taken into consideration .The Training of multi class SVM is takes approximately 51.5 sec as multiclass are grouped according to the mean shift algorithm, which reduces the training time of the multiclass SVM.

The accuracy result for the test dataset (1000 images) is predicted using the matrix below. The confidence matrix explains that the prediction of direction is better than previously defined by system. The number of inputs of each case (non-ped, front, back, left and right) is 200. From the Table .1, shows that the Non-Pedestrian is detected 100% accurately. Front, Left is predicted 99% accurately. Right is predicted 95% accurately and back is

Table 1 Confusion Matrix explaining Number of accurately predicted result

	Front	*Left*	*Back*	*Right*
Front	*0.99*	*0.01*	*0.08*	*0.01*
Left	*0.00*	*0.99*	*0.02*	*0.05*
Back	0.00	0.00	0.90	0.00
Right	0.01	0.02	0.00	0.95

predicted 90% accurately. Most of the back facing images is predicted as front. The table with lighter (grey colour) boxes shows that how the pedestrian is predicted, which indicates 95.75% the algorithm can detect it correctly. The false ratio of the matrix is 4.25%. Most of the images are detected as front even if they are in other directions.

Processes speed: When run on 2GB RAM. We obtained the results, for test dataset of 1000 images of size [65,129] contains Left facing are 200, Right facing are 200, front facing are 200 and the back facing are 200 and Non-pedestrian are 20. For all these images, the taken to predict the direction is 336.978259 sec. Therefore to compute the average time for executing each image is 336.9/1000=0.3369 sec approximately. To detect up to 1000 images direction with an accuracy of 95.75%.

IV. CONCLUSION

In this contribution, we presented a pedestrian detection and orientation classification system implemented on still images by using multi class SVM utilizing binary Decision tree. The detection and classification are done using Histograms of Oriented Gradients (HOG) features fed to a cascade of linear Support Vector Machines trained to make decisions. We ensure our system reduces the computational power needed, as the system takes the advantages of decision tree for faster classification.

The work exhibit that by selecting the right amount of classes, we can tremendously improve the quality of the classification and increase the accuracy. By grouping together classes that are visually similar, and applying the decision to classify we can estimate the pose of a pedestrian based solely on their visual appearance. Of course, these results are not final. Reason explaining why the results can be improved is the fact that we do not perform non-maxima suppression after the detection stage, so even detections with low confidence are used for training and testing the pose classification. The accuracy can be improved by including scaling factor in the system. The system does not perfectly distinguish the front and back classes yet.

REFERENCES

[1] R. Cucchiara, C. Grana, A. Prati, and R. Vezzani, "Probabilistic posture classification for human-behavior analysis," IEEE Transactions on Systems, Man, And Cybernetics, Part A: Systems and Humans, vol. 35, no. 1, pp. 42–54, 2005.

[2] T. Gandhi and M. M. Trivedi, "Pedestrian protection systems: Issues, survey, and challenges," IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 3, pp. 413–430, September 2007.

[3] D. M. Gavrila, "Sensor-based pedestrian protection," IEEE Intelligent Systems, vol. 6, no. 5, pp. 77–81, 2001.
[4] G. Antonini, S. Venegas, J. P. Thiran, and M. Bierlaire, "A discrete choice pedestrian behavior model for pedestrian detection in visual tracking systems," in Proc. Advanced Concepts for Intelligent Vision Systems, September 2004.
[5] G. Antonini, S. V. Martinez, M. Bierlaire, and J. P. Thiran, "Behavioral priors for detection and tracking of pedestrians in video sequences," International Journal of Computer Vision, vol. 69, no. 2, pp. 159–180, 2006.
[6] C. Wakim, S. Capperon, and J. Oksman, "A markovian model of pedestrian behavior," in Proc. IEEE Int. Conf. on Systems, Man, and Cybernetics, October 2004, pp. 4028–4033.
[7] M. Enzweiler and D. Gavrila." Integrated pedestrian classification and orientation estimation." In Computer Vision and Pattern Recognition (CVPR), pages 982–989. IEEE, 2010.
[8] C.-C. Chang and C.-J. Lin, LIBSVM : A Library for Support Vector Machines, Last updated June 2007, http://www.csie.ntu.edu.tw/ cjlin/papers/libsvm.pdf
[9] Bottou, L., et al.: Comparison of classifier methods: A case study in handwriting digit recognition. In: Proc. Int. Conf. Pattern Recognition., pp. 77–87 (1994)
[10] Platt, Cristianini, N., Shawe-Taylor, J.: Large margin DAG's for multiclass classification. Advances in Neural Information Processing Systems 12, 547–553 (2000).
[11] C. Papageorgiou and T. Poggio, "A Trainable Pedestrian Detection system", International Journal of Computer Vision (IJCV) in the year 2000
[12] N. Dalal, B. Triggs, "Histograms of oriented gradients for human detection", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), page 1:886-893, 2005.
[13] A. Agarwal and B. Triggs, "Recovering 3d human pose from monocular images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 1, pp. 44–58, January 2006.
[14] Bo Wu and Ram Nevatia, "Detection of Multiple, Partially Occluded Humans in a Single Image by Bayesian Combination of Edgelet Part Detectors", IEEE International Conference on Computer Vision(ICCV), pages 1:90-97, 2005
[15] B.Leibe, E. Seemann, and B. Schiele. "Pedestrian detection in crowded scenes" IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pages 1:878-885, 2005
[16] Integrated Pedestrian and Direction Classification using a Random Decision Forest. (Junli Tao and Reinhard Klette University of Auckland, Auckland, New Zealand 2013 IEEE)

 DOI 10.17148/IJARCCE.2015.411120