International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

AUTOMATIC HELMET DETECTION

A.Rajesh[1], V. Sai[2], P. Janardhana rao[2], K. Sai Sandeep[2], K. Aravindh[2]

[1]*Assistant Professor, Department of Electronics and Communication Engineering*
[2, 3,4,5] *Student, Department of Electronics and Communication Engineering Aditya Institute of Technology and Management, Tekkali, Srikakulam-AP*

ABSTRACT

We develop an helmet detection method combining classification and cluster. Helmet detection is an important, yet challenging vision task. It is a critical part in many applications such as traffic surveillance. Our proposed method work is as follows, Pre-processing, Feature Extraction and classification. We demonstrate our proposed work by using surveillance traffic videos. Finally, our method will classify whether the person is wearing helmet or not. As far as the robustness and effectiveness are concerned, our method is better than the existing algorithms. The person is not wearing helmet and recognized.

1. INTRODUCTION

1.1 DIGITALIMAGEPROCESSING:

The identification of objects in an image and this process would probably start with image processing techniques such as noise removal, followed by (low-level) feature extraction to locate lines, regions and possibly are as with certain textures

The clever bit is to interpret collections of these shapes as single objects, e.g. cars on a road, boxes on a conveyor belt or cancerous cells on a microscope slide. One reason this is an AI problem is that an object can appear very different when viewed from different angles or under different lighting. Another problem is deciding what features belong to what object and which are background or shadows etc. The human visual system performs these tasks mostly unconsciously but a computer requires skillful programming and lots of processing power to approach human performance. Manipulation of data in the form of an image through several possible techniques. An image is usually interpreted as a two-dimensional array of brightness values, and is most familiarly represented by such patterns as those of a photographic print, slide, television screen; or movie screen. An image can be processed optically or digitally with a computer.

2. LITERATURE SURVEY

This study proposed a computational vision methodology for the detection of helmet use by motorcyclists on public roads. The study is divided into two stages: vehicle segmentation and classification, and the detection of helmet use. The stage of vehicle segmentation and classification has the following objectives: determining which objects are moving in the scene and classifying these objects. Similar to the majority of computational vision systems, the proposed system requires a calibration stage. In the calibration stage, parameters that are required for the operation of the system are adjusted. In the calibration stage of the proposed system, a cross line (CL) is defined. This line will be marked by the system operator and should cross the public road wherethesystemwillberesponsibleforcapturingthevehicles.ThemovingobjectsthatcrosstheCLaresubsequently extracted from the video frame. The next step involves extracting the features of the segmented objects. The wave let transform (WT) was employed. The vectors are the input parameters of the classifier. The random forest classifier was employed to classify the vehicles. The images are grouped into two classes: motorcycle or no motorcycle. This classification is adopted as it is sufficient for assessing whether an object is a motorcycle in the proposed system. The second stage consists of the detection of helmet use. To reduce the computational cost and to increase the precision, a region of interest (RoI) was defined. The HOG descriptor was employed in this stage. The descriptor obtains different vectors for an image of a motorcyclist with a helmet and without a helmet. The extraction of features for the detection of helmet use is considered as a critical step in this study, as helmet detection is the main objective of the proposed system. The MLP classification algorithm was used to classify the images into two classes: with helmet or without helmet. The diagram of the proposed system shows all of the stages and sub stages of the problem.

3. METHODOLOGY

3.1 INTRODUCTION

A convolutional neural network (CNN) is a multi layer neural network. It is a deep learning method designed for image recognition and classification tasks. It can solve the problems of too many parameters and difficult training of the deep neural networks and can get better classification effects. The structure of most CNNs consists of input layer convolutional layer (Conv layer)-activation function-pooling layer-fully connected layer (FClayer).

3.2 CHARACTERISTICS

The main characteristics of CNNs are local connectivity and parameter sharing in order to reduce the number of parameters and increase the efficiency of detection. The Conv layer and the pooling layer are the core parts, and they can extract the object features. Often, the convolutional layer and the pooling layer may occur alternately. The Conv layers can extract and reinforce the object features. The pooling layers can filter multiple features; remove the unimportant features, and compress the features. The activation layers use nonlinear activation functions to enhance the expression ability of the neural network models and can solve the nonlinear problems effectively. The FC layers combine the data features of objects and output the feature values. By this means the CNNs can transfer the original input images from the original pixel values to the final classification confidence layer by layer. In order to better extract the object features and classify the objects more precisely, Hinton proposed the concept of deep learning which is to learn object features from vast amounts of data using deep neural networks and then classify new objects according to the learned features. Deep learning algorithm based on convolutional neural networks has achieved great results in object detection, image recognition, and image segmentation. Girshick proposed R-CNN detection framework (region with CNN features) in 2014. Many models based on R-CNN were proposed after that including SPP-net (spatial pyramid pooling network), FastR-CNN (fast region with CNN features), and FasterR-CNN (faster region with CNN features).

3.3 CLASSIFICATION

Classification-based CNN object detection algorithms such as Faster R-CNN are widely used methods. However, the detection speed is slow and cannot detect in real time. Regression-based detection algorithms are becoming increasingly important. Redmon proposed YOLO (You Only Look Once) algorithmin2016.At the end of 2016, Liu combined the anchor box of FasterR-CNN with the bounding box regression of YOLO and proposed a new algorithm SSD (Single Shot Multi Box Detector) with higher detection accuracy and faster speed. Although the SSD algorithm is not capable of the highest accuracy, the detection speed of the SSD algorithm is much faster and comparable to the YOLO algorithm and the precision can be higher than that of the YOLO algorithm when the sizes of the input images are smaller. While the FasterR-CNN algorithm tends to lead to more accurate models, it is much slower and requires at least 100 ms per image .Therefore, considering the real-time detection requirements, the SSD algorithm is chosen in the research. In order to reduce greatly the calculation amount and model thickness, the Mobile Net model is added. Therefore, in the paper, the SSD-Mobile Net model is selected to detect safetyhelmetswornbytheworkers.TheSSDalgorithmisbasedonafeed-forwardconvolutionalnetworkto produce bounding boxes of fixed sizes and generate scores for the object class examples in the boxes. A non maximum suppression method is used to predict the final results. The early network layers of the SSD model are called the base network, based on a standard frame work to classify the image. The base network is truncated before the classification layers, and the convolutional layers are added at the end of the truncated base network. The sizes of the convolutional feature maps decrease progressively to predict the detections at multiple scales. The SSD algorithm sets a series of fixed and different size default boxes on the cell of each feature map as shown in Figure1. Each default box predicts two kinds of detections. One is the location of bounding boxes including 4offsets (cx, cy, w, h), which represent, respectively, x and y coordinates of the center of the bounding box and the width and height of the bound ing box; the other is the score of each class. If there are C classes of the objects, the SSD algorithm predicts a total of C+1 score including the score of the background. The setting of default boxes can be divided into two aspects: size and aspect ratio

3.4 IMAGE FUSION

Multi sensor data fusion has become a discipline which demands more general formal solutions to a number of application cases. Several situations in image processing require both high spatial and high spectral information in a single image. This is important in remote sensing. However, the instruments are not capable of providing such information either by design or because of observational constraints. One possible solution for this is data fusion.

3.5 STANDARD IMAGE FUSION METHODS

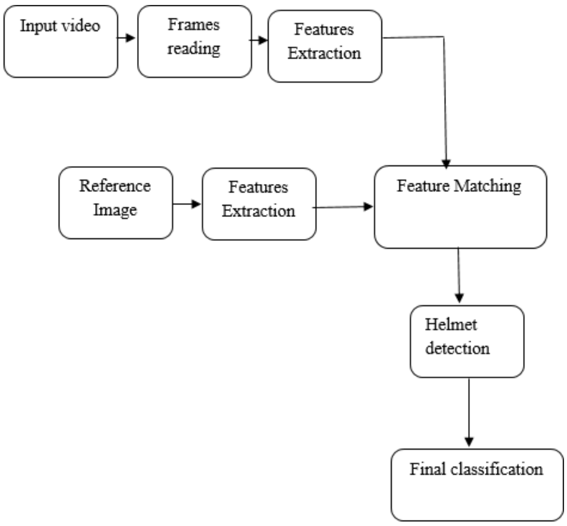
Image fusion methods can be broadly classified into two groups - spatial domain fusion and transform domain fusion. The fusion methods such as averaging, Bovey method, principal component analysis (PCA) and IHS based methods fall under spatial domain approaches. Another important spatial domain fusion method is the high pass filtering based technique. Here the high frequency details are injected into up sampled version of MS images. The disadvantage of spatial domain approaches is that they produce spatial distortion in the fused image. Spectral distortion becomes a negative factor while we go for further processing, such as classification problem. Spatial distortion can be very well handled by frequency domain approaches on image fusion. The multi resolution analysis has become a very useful tool for analyzing remote sensing images. The discrete wavelet transform has become a very useful tool for fusion. Some other fusion methods are also there, such as Laplacian pyramid based, curve let transform based etc. These methods show a better performance in spatial and spectral quality of the fused image compared to other spatial methods of fusion.

The images used in image fusion should already be registered. Misregistration is a major source of error in image fusion. Some well-known image fusion methods are:

3.6 IMAGEFUSIONMETRICS:

Comparative analysis of image fusion methods demonstrates that different metrics support different user needs, sensitive to different image fusion methods, and need to be tailored to the application. Categories of image fusion metrics are based on information theory, features, structural similarity, or human perception.

4. WORKING MODEL



Block Diagram

5. RESULTS

In this section, the results are presented and discussed. In addition, a comparative analysis is performed with other algorithms to describe and classify the images. The results are divided into two groups: – results of the vehicle classifier, and–results of the helmet detector. Information about the image databases, generated from the segmentation of vehicles and the methodology employed for the classification of the results, a real so presented in this section.

5.1 IMAGE DATABASE

The image databases used for the tests of the algorithms were obtained from the application of vehicle segmentation in videos captured on public roads. Two image databases were obtained: database1 anddatabase2. 1 The first database was employed in the classification of the vehicles, and the second database was employed in the detection of the helmets. The videos were obtained from a charge-coupled device (CCD) video camera that was installed on public roads. The videos were recorded during the day and at night, and under different lighting conditions (cloudy and sunny days).

Fig. 10 Example of images from *database2*

5.2 METHODS OF EVOLUTION OF RESULTS

Some known metrics from the literature were used to evaluate the performance of the classification algorithms. This section shows the results of these evaluation metrics. The confusion matrix is a table that shows the classification results. The matrix is composed of four values: true positive (TP), false positive(FP), falsenegative (FN) and true negative (TN). From these rates, the values of specificity (S), negative predictive value (NPV), precision (P), recall (R), accuracy (A), F-Measure(FM) and Kappa coefficient (K) are calculated. As shown in Table1, the level of accuracy of the Kappa coefficient was classified according to the level of accuracy established by Land is and Kock. Another result evaluation method involves Receiver Operating Characteristic (ROC) curves. The ROC curves show the true positive rates versus the false positive rates. The curve is constructed by varying the threshold of the classifier and observing the resultsthat are generated from the modification. The ROC curve can be used to evaluate the efficiency in terms ofhitrateofthemachinelearningalgorithms.ThemaininformationthatisextractedfromaROCcurveistheareaunder thecurve(AUC);thelargertheareais,thebetter the classifier.

5.3 VEHICLE CLASSIFICATION

The statistical method-fold cross-validation (k= 10)was employed to generate the resultsintheclassificationstage.Inthe10-foldcross-validation,thesetoforiginaldataisrandomlypartitionedin10datasubsetswiththe same size. From the ten subsets, nine subsets are used as training data, and only one subset is selected as the validation set to test the model. The cross-validation process is repeated ten times. Each of the ten subsets is only used once to validate the data. The mean of the ten generated results is computed to produce a single estimate. The advantage of this method is that all subsets are employed for testing and training, and each subset is only used once for testing. The proposed system used the Wavelet transform as feature descriptor and the random forest as classifier. An accuracy of 0.9778 was obtained, that is, of the3,245 vehicles, only 72 were misclassified. The Kappa coefficient classified the results as “Excellent”accordingtoTable1,whichshowsthelevelofaccuracy.TheFMalsoreturnedasatisfactoryresult(0.9754),which reflects the P and R rates, as the FM is calculated based on these rates. The values of S (0.9930) and NPV(0.9793)reflect that the proportion of true negatives was satisfactory.

Table 1 Level of accuracy of a classification according to the Kappa coefficient value

Kappa Coefficient (K)	Quality
K < 0.2	Poor
0.2 ≥ K < 0.4	Reasonable
0.4 ≥ K < 0.6	Good
0.6 ≥ K < 0.8	Very Good
K ≥ 0.8	Excellent

5.4 DETECTION OF HELMET USE

The results obtained for the detection of helmet use are discussed in this section. In this stage, the images from database2 were employed. The same metrics used for vehicle classification were also used here to evaluate the obtained results. The proposed method calculates a sub-window of the RoI, and the HOG descriptorisusedtoextractthefeatures.Oncethefeaturesareextracted, theclassificationisperformedwiththe MLP network. The obtained result was an accuracy of 0.9137. From a total of 255 images, 22 images were incorrectly classified. Similar to the vehicle classification stage, the Kappa coefficient classified the result as “Excellent “according to the table that lists the accuracy levels(K≥0.8).TheFMvaluewas0.9281,which reflects the P (0.9161) and R (0.9404) rates. Figure 12 shows the ROC curve that was generated for the classifier. The AUC was 0.9556.

5.5 ANALYSIS

In addition to the results of the proposed system, a comparative analysis was performed using other descriptors and classifiers. The descriptors tested here were applied to the sub-window in the pre-processing stage. The following descriptors and combinations were tested: WT, HOG, LBP, WT+LBP, WT+HOG, HOG+LBP and WT+HOG+LBP. The hybrid forms of the descriptors were assembled by combining the feature vectors. The descriptors were applied with the same parameters of the vehicle classification stage.

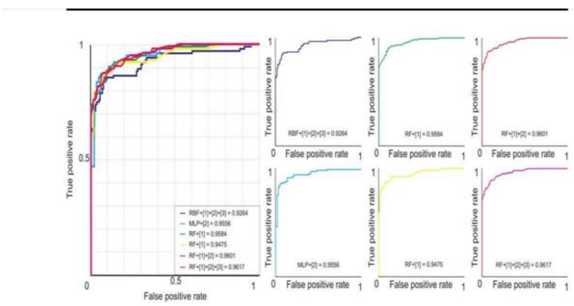


Fig. 12 Bests ROC curves in detection of helmet use, left aggregate and right individual, respectively

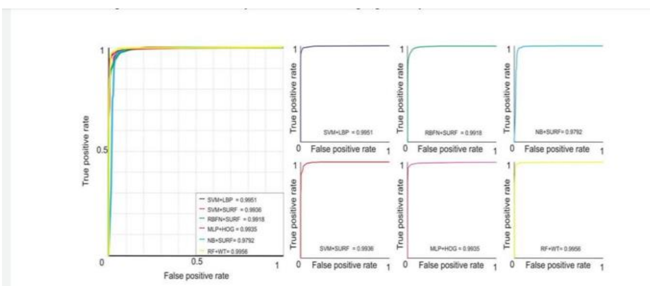
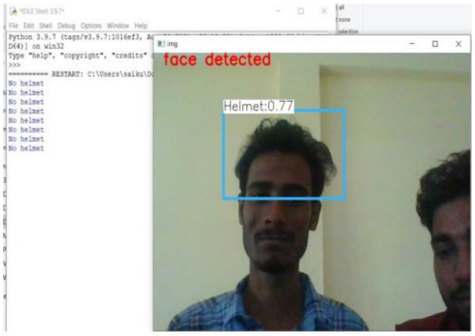


Fig. 11 Bests ROC curve for the vehicle classification, left aggregate and right individual, respectively



6. CONCLUSION

This study addresses the detection of motor cyclists without helmets on public roads. A computational vision system was proposed and classified as follows: – vehicle segmentation andclassification,and–detectionofhelmetuse.Inthestageofvehiclesegmentationandclassification,algorithms for the background calculation and tracking of objects, descriptors and classifiers that exhibited reasonable hit rates and low processing times were selected from the literature achieving an accuracy of 0.9778. In the stage of detection of helmet use, algorithms for the extraction of features in images and classification algorithms were employed. The proposed system also obtained satisfactory hit rates. The MLP classifier that incorporated the HOG descriptor obtained the best results, with an accuracy of 0.9137

6.1 FUTURE STUDEIS

The results are promising but can be improved. An important step for improving the results is the stage of image capturing, which should produce better quality images. The images of database1were not employed in the stage of helmet detection due to their low quality. Future studies should focus on the detection and recognition of the registration plate of the vehicle. A better quality image is necessary to recognize the characters on the plate. Hybrid descriptors were not employed in the vehicle segmentation stage. They will be employed to improve the results. The use of algorithms for feature selection can be evaluated to increase the obtained hit rates. The descriptors returned a large number of features, which frequently hinders the classification of objects. In these feature vectors, the possibility of features that are insignificant or duplicates is high. Therefore, a thorough analysis of the literature in search of features election algorithms is necessary. Another future study is the detection of passengers on motorcycles. A motorcycle has the capacity to carry a driver and a passenger. This functionality can be extended to the detection of more than one passenger, as two or more passengers constitute a traffic violation.

REFERENCES

[1] AdelsonEH,BergenJR(1985)Spatiotemporalenergymodelsfortheperceptionofmotion.JOptSocAmA2(2):284–299

[2] Agurto C, Murray V, Barriga E, Murillo S, Pattichis M, Davis H, Russel S, Abramoff M,SolizP(2010)Multiscaleam-fmmethodsfordiabeticretinopathylesiondetection.IEEETrans Med Imag 29:502–512

[3] BouguetJ-Y(2000)Pyramidalimplementationofthelucaskanadefeaturetracker.IntelCorporation,Microprocessor Research Labs

[4] BreimanL(2001)Randomforests.MachLearn45(1):5–32

[5] ChenZ, EllisT, VelastinSA (2012) Vehicledetection, trackingandclassification inurbantraffic.In:15thInternationalIEEEconferenceonintelligenttransportationsystems,pp 951–956 .

[6] ChiuC-C,Min-YuK,ChenH-T(2007)Motorcycle detectionandtrackingsystemwithocclusion segmentation. In: Eighth International workshop on image analysis formultimediainteractiveservices. USA

[7] ChivertonJ(2012)Helmetpresenceclassificationwithmotorcycledetectionandtracking.IntellTransp Syst6(3):259–269

[8] CortesC,VapnikV(1995)Support-vectornetworks.MachLearn20:273–297

[9] DalalN,TriggsB(2005)Histogramsoforientedgradients forhumandetection.In:IEEEComputersocietyconferenceoncomputervisionandpatternrecognition,pp886–893

[10] DaubechiesI(1992)Tenlecturesonwavelets.Societyforindustrialandappliedmathematics.