International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 05 | May - 2021 ISSN: 2582-3930

Hel Must - System to Report Helmet Violation in Traffic using Deep Learning

Aravindhan G[1],Mani kandan NP[2], Jayasurya S[3], Shanmughapriya M[4]
[1,2,3] IV Year Student,
Department of Information Technology, Sri Sai Ram Institute of Technology,
[4]Assistant Professor,
Department of Information Technology, Sri Sai Ram Institute of Technology

Abstract

Helmet is an amazing discovery which saves lives of two wheeler users during accidents. Usage of Helmets is encouraged in all two wheeler driving situations regardless of duration of drive and traffic conditions. Though government imposes many rules to regulate and mandate the use of Helmets in Roads, it is not achieved with optimal usage. Continuous monitoring and constant alerts need to be given to make Helmets to be used predominantly by two wheeler drivers. The proposed system tries to achieve the possibility of remote monitoring of traffic for helmet violators. The traffic videos are analyzed and for two wheeler identification and helmet detection. The deep learning algorithm helps in combining the process of vehicle localization along with helmet detection. The proposed system uses the IITH dataset. The IITH Dataset [1] obtained is then preprocessed tagged and used for this project. The proposed system uses CNN for Helmet Detection and a back ground subtraction algorithm to get moving objects without overlay from video frames. The system deals with Illumination, Occlusion of surveillance videos and challenges associated with processing and working with poor quality surveillance videos.

Keywords – Helmet Detection, Traffic Surveillance, Convolutional Neural Networks, Deep Learning,

1. Introduction

Two wheelers are affordable and used as a common mode of transportation in India. Nearly 17.42 million two wheelers sold in the year 2020 alone. Two wheeler accidents also increase along with its usage, the main problem faced during two wheeler accidents is non usage of helmets while driving. The usage of helmet to be mandated and governed to decrease the causalities of head associated injuries during accidents. Remote monitoring of traffic videos for helmet detection in moving vehicles will help in this regard, K. Dahiya et al [2] discusses about the challenges faced in human assisted video monitoring, they have also proposed a system to detect helmets using Object Detection in Grey Scale images using Template matching techniques. The proposed system focuses on Object Detection using back ground subtraction and deep learning CNN models to achieve the helmet detection from surveillance video.

The architecture of the proposed model Figure 1 illustrates the various steps involved in video processing and the system developed as a front end to enable access to the application.

- The application takes any traffic related surveillance input video and detects helmet in that video,
- The major scope of this proposed system is to enhance the detection accuracy irrespective of the illumination and ambient occlusion which may lead to poor quality video.

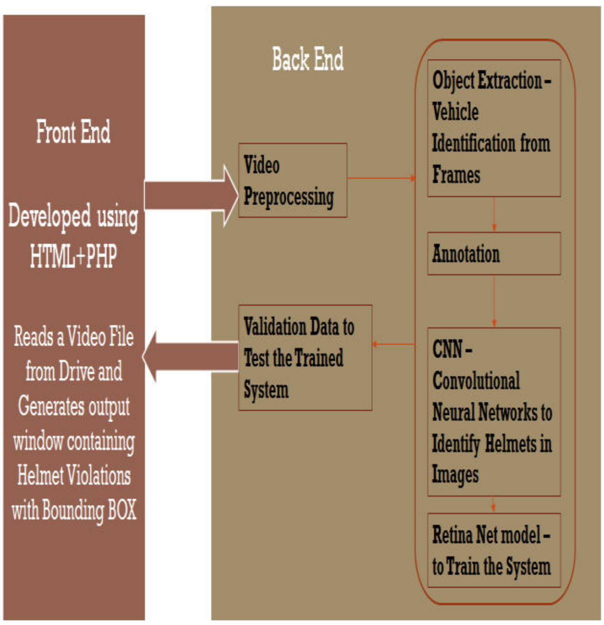

Figure 1: Architecture of Hel Must System

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 05 | May - 2021 ISSN: 2582-3930

- The use case driven approach in this proposed system can be found in figure 2. Many researchers [3],[4],[5] have proposed system to detect helmets in videos, but major limitations of those systems includes less deterministic methods for object identification and classification.
- The proposed model focuses on the use of convolution neural network over features that are hand crafted. The proposed model also has a way to train the system with new data sets.

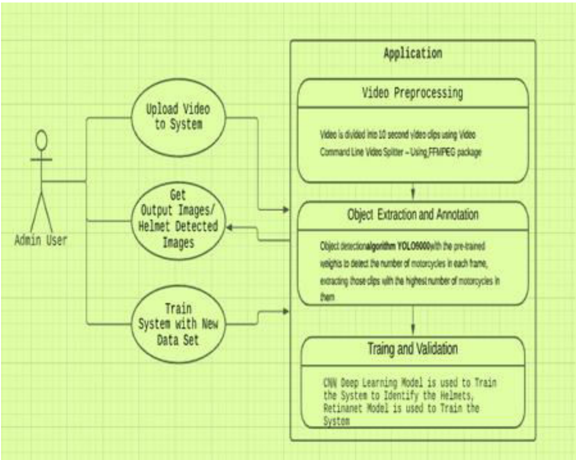

Figure 2: Use Case Model of the proposed system

2. Literature Review

This section of the paper focuses on various methods and data sets used for helmet detection from videos.

The existing work that solves the problem by image processing solutions use technologies like HOG, LBP, WT [5][6][7] . The system proposed by isolates the bikes from images and by approximation crops the most probable area where helmet might be present and then feeds it to the feature extraction and matching system. Chiverton[5] proposed the use of circular arc to identify helmet in a video feed, it has very low accuracy. On the other hand, given the number of vehicles on the speed at a given instant, the computation that required is very heavy and consumes lots of resources. These methods will determine any circular object around the bike rider as helmet. In [6] two phases were used for helmet detection. In the first phase, moving objects were determined where cross line was specified. It is then checked whether it is a motorbike or not. In the second phase, a region of interest was used to improve efficiency. A SVM classifier was used to classify moving object into two classes. Three classification families were used viz. geometric, periodic, and tree based. Videos were captured at 25 fps and image size was of 1280x720 [8].

Talks about a system very similar to the one proposed in this paper which identifies bike riders without helmet and captures the number plate of all the offenders on a COCO database. It classifies motor bike and helmet using YOLO and the technology used for license place recognition is Open ALPR. Both of these technologies charge monthly fees and hence are not economically feasible. A detection method for circular arcs was proposed by Wen et al. [9] based on the modified circular Hough transform (CHT) [10] .The edges of the image are calculated using defined threshold value. The circle Hough transforms is subsequently applied. Circular regions such as a helmet are searched by the transform. In automated teller machines (ATM), this method was used for surveillance systems. To identify a helmet in the image with use of geometric resources is the main limitation of this study. For locating the helmet, Geometric characteristics are not sufficient; the helmet can be confused with a human head, as their shapes are similar. Chiu et al. [11] proposed a vehicle counting system based on computational vision. The objective is to detect and track motorcycles that are partially occluded by another vehicle. Motorcycle is detected using Helmet detection system. The system assumes that the helmet region has a shape of a circle. The edges of the image are calculated over its possible region where the motorcycle is located to detect the helmet. The numbers of edge points that resemble a circle are subsequently counted. If this number is greater or equal to a predefined value during the calibration of the system, the region will correspond to a helmet. A motorcycle is assumed to exist in the same location if helmet is detected by the system. Some parameters such as helmet radius, camera angle and height are required to be given as input by the system operator at the calibration stage. If any condition, such as camera height or the road on which the system is in operation, changes, all parameters should be altered. Waranusast et al. [12] proposed most recent study of the detection of helmet use. Moving objects from videos is extracted using the AGMM algorithm. The system classifies extracted objects as motorcycles or other objects. Three features are employed for this purpose: the area of the rectangle that contains the image, the ratio between the width and the height of the rectangle and the standard deviation of the H band in the hue-saturationvalue (HSV) colour space around a rectangle at the Centre of the object. The next step uses k-nearest neighbours (KNN) classifier with the calculated features. The counting of passengers, which is performed by the number of heads that appear on the image, is the primary advantage of this study. The final step performs classification using geometric information of the head region and colour information. KNN classifier reapplied these features to classify the images of motorcyclists with helmets and without helmets. Hit rate of 95 % is obtained at the motorcycle detection stage. The passenger counting stage obtained a total of 83.82 % hits. In the helmet detection stage, the hit rate was 89 %. The images of the head region were manually cut in the latter stage. A flaw of the system was the images were perpendicularly captured by the camera, that is, the images show the side view of motorcycles, as the vehicle registration plate is difficult to capture in that position. More than one person on the motorcycle is identified in the images using this method. Another angle images had been taken from other side, one of the persons on the motorcycle would most likely be superimposed on another image, which would generate an occlusion

3. Helmet use detection algorithm

3.1. Method After the creation of the dataset was finished, we applied a state-of-the-art object detection algorithm to the annotated data, to facilitate motorcycle helmet use detection on a frame-level. In this process, data from the training set is used to train the object detection algorithm. In the process of training, the validation set is used to find the best generalizing model, before the algorithm's accuracy in predicting helmet use is tested on data that the algorithm has not seen before, the so-called test set. Generally, the state-of-the-art object detection algorithms can be divided into two types: two-stage and single-stage approaches. The two-stage approaches first identify a number of potential locations within an image, where objects could be located. In a second step, an object classifier (using a convolutional neural network) is used to identify objects a these locations. While two-stage approaches such as Fast R-CNN [12], achieve a higher accuracy than single-stage approaches, they are very time-consuming. In contrast, single-stage approaches simultaneously conduct object location and object identification. Single stage approaches like YOLO [12] and RetinaNet [13] therefore are much faster than two-stage approaches, although there is a small trade-off in accuracy. In this paper, we used RetinaNet [13] for our helmet use detection task. While it is a single-stage approach, it uses a multi-scale feature pyramid and focal loss to address the general limitation of one-stage detectors in accuracy. Figure 4 illustrates the framework of RetinaNet. 3.2. Training Since the task of detecting motorcycle riders' helmet use is a classic object detection task, we fine-tuned RetinaNet instead of training it from scratch. I.e. we use a RetinaNet model1 which is already trained for general object detection and fine tune it to specifically detect motorcycles, riders, and helmets. In our experiments, we used ResNet50 [13] as the backbone net, initialized with pre-trained weights from ImageNet [12]. The backbone net provides the specific architecture for the convolutional neural network. In the learning process, we used the Adam optimizer [13] with a learning rate of $\alpha = 0.00001$ and a batch size of 4 and stopped training when the weighted mean Average Precision (weighted mAP, explained in the following) on the validation set stopped improving with a patience of 2. To assess the accuracy of our algorithm, we use the Average Precision (AP) value [13]. The AP integrates multiple variables to produce a measure for the accuracy of an algorithm in an object detection task, including intersection over union, precision, and recall. The intersection over union describes the positional relation between algorithm generated and human annotated bounding boxes. Algorithm generated bounding boxes need to overlap with human annotated bounding boxes by at least 50%, otherwise they are registered as an incorrect detection. The precision presents the number of correct detections of all detections made by the algorithm (precision = true positives true positives + false positives). The recall variable measures how many of the available correct instances were detected by the algorithm. (recall = true positives true positives + false negatives). For a more in-depth explanation of AP please see [13] and [13]. Since the number of frames per class was very imbalanced in our dataset (Table 3), the final performance for all classes is computed as weighted average of AP for each class, defined as: weighted mAP = X C i=1 wiAPi , (1) where weights wi across all C classes will sum to one, and set to be proportional to the number of instances. It can be observed that training loss is constantly decreasing, i.e. the prediction error is getting smaller, while the deep model learns useful knowledge for the helmet use detection from the training set. Consequently, the weighted mAP of the training set is constantly increasing. At the same time, the validation loss, i.e. the prediction error on the validation set is getting smaller in the first 9 epochs. Correspondingly, the mAP on the validation set is increasing in the first few epochs before it stops to improve after 9 epochs, which means the algorithm starts to overfit on the training set. Therefore, we stopped training and selected the optimal model after 9 epochs, obtaining 72.8% weighted mAP on the validation set.

Our models were implemented using the Python Keras library with Tensorflow as a backend [13] and ran on two NVIDIA Titan Xp GPUs.

4. Results

In the following, we report the helmet use detection results of the algorithm on the test set, using the optimal model developed on the validation set (where it obtained 72.8% weighted mAP). We achieved 72.3% weighted mAP on the test set, with a processing speed of 14 frames/second. The AP for each class is shown in Table 3. It can be observed that RetinaNet worked well on common classes but not on underrepresented classes due to the small number of training instances. Considering only common classes (up to two riders), our trained RetinaNet achieved 76.4% weighted mAP. This is a very good performance considering a lot of factors such as occlusion, camera angle, and diverse observation sites. Detection results on some sample frames are displayed in Fig. 6. Due to the imbalanced classes, there are some missing detections, e.g., Fig. 6 (a), (g) and (h). Example videos, consisting of algorithm annotated frames of the test set can be found in the supplementary material.

4. Conclusion and future work

The lack of representative motorcycle helmet use data is a serious global concern for governments and road safety actors. Automated helmet use detection for motorcycle riders is a promising approach to efficiently collect large, up-to-date data on this crucial measure. When trained, the algorithm presented in this paper can be directly implemented in existing road traffic surveillance infrastructure to produce real-time helmet use data. Our evaluation of the algorithm confirms a high accuracy of helmet use data, that only deviates by a small margin from comparable data collected by human observers. Observation site specific training of the algorithm does not involve extensive data annotation, as already the annotation of 270 s of video data is enough to produce accurate results for e.g. the Yangon II observation site. While the sole collection of data does not increase road safety by itself [13], it is a prerequisite for targeted enforcement and education campaigns, which can lower the rate of injuries and fatalities [13].

References

[1]. Vishnu, Chalavadi & Singh, Dinesh & Chalavadi, Krishna Mohan & Babu, Sobhan. (2017). Detection of Motorcyclists without Helmet in Videos using Convolutional Neural Network. 10.1109/IJCNN.2017.7966233. Data Set Obtained from IITH https://docs.google.com/forms/d/e/1FAIpQLSclwXdJXYofwiGlvhLESGREzvYgkbVWhy-7Y3iUFzJ5ZNN4Fg/viewform

[2]. K. Dahiya, D. Singh, and C. K. Mohan, "Automatic detection of bikeriders without helmet using surveillance videos in real-time," in Proc.Int. Joint Conf. Neural Networks (IJCNN), Vancouver, Canada, 24–29 July 2016, pp. 3046–3051.

[3]. C.-C. Chiu, M.-Y. Ku, and H.-T. Chen, "Motorcycle detection and tracking system with occlusion segmentation," in Proc. Int. Workshop on Image Analysis for Multimedia Interactive Services, Santorini, Greece, 6–8 June 2007, pp. 32–32

[4]. J. Chiverton, "Helmet presence classification with motorcycle detection and tracking," IET Intelligent Transport Systems (ITS), vol. 6, no. 3, pp. 259–269, 2012. [7] R. Silva, K. Aires, T. Santos, K. Abdala, R. Veras, and A. Soares, "Automatic detection of motorcyclists without helmet," in Proc. Latin American Computing Conf. (CLEI), Puerto Azul, Venezuela, 4–6 October 2013, pp. 1–7.

[5]. W. Rattapoom, B. Nannaphat, T. Vasan, T. Chainarong, and P. Pattanawadee, "Machine vision techniques for motorcycle safety helmet detection," in Proc. Int. Conf. Image and Vision Computing New Zealand (IVCNZ), Wellington, New Zealand, 27–29 November 2013, pp. 35–40.

[6]. Jie Li, Huanming Liu, Tianzheng Wang, Min Jiang, Kang Li and Xiaoguang Zhao, "Safety Helmet Wearing Detection Based on Image Processing and Machine Learning" in 2017 International Conference on Advanced Computing Intelligence, ICACI, Feb 4-6, 2017

[7]. The New Indian Express report on "Traffic violation: Cops collects Rs 20 lakh from 55 in just 24 hours" https://www.newindianexpress.com/cities/bengaluru/2019/sep/13/traffic-violation-cops-collect-rs-20-lakh-from55-in-just-24-hours-2032987.html

[8]. Chiverton J (2012) Helmet presence classification with motorcycle detection and tracking. Intell Transp Syst 6(3):259–269

[9]. 'Automatic detection of motorcyclists without helmet' by Romuere Silva, Kelson Aires, Thiago Santos, Kalyf Abdala, Rodrigo Veras, André Soares, Departamento de Computação, 2013 Latin America Cmputiong Conference.

[10]. 'Detection of motorcyclist without helmet using video convolution neural network' by C. Vishnu,Dinesh Singh,C. Krishna Mohan,Sohan Babu; 2017 International Joint conference on Neural Networks(IJCNN).

[11]. An Automatic Detection of Helmeted and NonHelmeted Motorcyclist with License Plate Extraction using Convolutional Neural Network' by Jimit Mistry, Aashish K. Misraa, Meenu Agarwal, Ayushi Vyas, Vishal M. Chudasama, Kishor P. Upla, Sardar Vallabhbhai National Institute of Technology (SVNIT),

[12]. Wen C-Y, Chiu S-H, Liaw J-J, Chuan-Pin L (2003) The safety helmet detection for atm's surveillance system via the modified hough transform. In: IEEE 37th Annual international carnahan conference on security technology, pp 364–369

[13]. D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980.

[14]. M. Everingham, L. Van Gool, C. K. Williams, J. Winn, A. Zisserman, The pascal visual object classes (voc) challenge, International Journal of Computer Vision 88 (2) (2010) 303–338.

[15]. G. Salton, M. J. McGill, Introduction to modern information retrieval, McGraw Hill Book Co., 1983. [33] F. Chollet, et al., Keras, https://keras.io (2015).

[16]. A. A. Hyder, Measurement is not enough for global road safety: implementation is key, The Lancet Public Health 4 (1) (2019) e12–e13.

[17]. World Health Organization, Helmets: a road safety manual for decisionmakers and practitioners, Geneva: World Health Organization, 2006.