Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 13083 - 13091
Received 25 April 2021; Accepted 08 May 2021.

Traffic Rule Violation Detection System using Surveillance Videos in real-time

Darshana A Naik[1*], Pramod Sunagar[2] , Preetham Naik[3], Sai Ootej Reddy[4]
[1,2,3,4] M S Ramaiah Institute of Technology, Bangalore India (Affiliated to VTU)
*darshananaik@msrit.edu

ABSTRACT

In many countries, especially in India, two-wheeler vehicles are the most common mode of transportation. However, because of the lack of security, the danger is extremely high. Governments are working on imposing strict rules and law-abiding behavior in traffic to minimize the risk and injuries. There is a lack of detailed data on the safety-critical behavioural metric of motorbike helmet usage, especially in developing countries like India, where motorcycles are the primary mode of transportation. Due to a lack of evidence, tailored compliance and outreach campaigns, which are critical for accident prevention, are not possible. Despite this, there is a growing need for a secure and timely intelligent device that does not rely on a human observer to detect helmet use by motorcycle riders. As a result, we used a deep learning technique to build a model that can automatically detect the rider's use of a motorcycle and a helmet from video data. Since wearing a motorcycle helmet reduces the risk of fatal injury to motorcycle riders in road traffic accidents by 42 percent, the government has enacted legislation making helmet use mandatory for bike riders. The key goal of our work is to reduce the likelihood of motorcycle riders being killed in traffic accidents. The algorithm's solution is Convolutional Neural Networks (CNN). It's a kind of deep neural network that's often used for visual imagery. The suggested method first determines whether the car is a motorcycle or not. It then decides whether the cycle rider is wearing a helmet or not and whether or not using optical features. On real-world surveillance evidence, the experimental results show that helmet detection accuracy is about 85%.

Keywords: Convolutional Neural Networks (CNN), vehicle detection, vehicle categorization, helmet recognition

Introduction

Motorcycles are commonly utilized as a mode of commuting in numerous countries. Their low costs and low activity costs in testing of various vehicles are important preferences. The number of bicycle accidents has increased during the previous decade. Motorbike is an extremely well known method of transportation in pretty much every nation. Nonetheless, there is a high hazard included due to less security. To lessen the included hazard, it is exceptionally alluring for bicycle riders to utilize head protector. Watching the convenience of protective cap, Governments have declared riding a bicycle without a helmet a criminal offence and have resorted to manual methods to apprehend offenders. Nonetheless, new video observation-based approaches are ineffective and require a great deal of human assistance. Overall, such systems are impractical due to the interaction of individuals, whose productivity deteriorates over time. Computerization of this process is extremely appealing for thorough and thorough inspection of this violation while also dramatically reducing the amount of human resource required. Likewise, numerous nations are embracing frameworks including reconnaissance cameras at open spots. Along these lines, the answer for recognizing violators utilizing the current foundation is likewise financially savvy. Using background subtraction and object segmentation, the proposed method first determines if the vehicle is a motorcycle or not. It then decides whether the cycle rider is wearing a helmet or not and whether or not optical features are being used.

Feed: By and large, CCTV cameras catch low goals video. Additionally, conditions, for example, low light, terrible climate entangle it further. Because of such confinements, assignments, for example, division, classification and following become much more difficult.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 13083 - 13091
Received 25 April 2021; Accepted 08 May 2021.

Literature Review

Over the previous years, numerous works were completed in traffic examination on open roads, including vehicle location and classification, and head protector discovery. Background and foreground image computing techniques are required to segregate and categorize moving objects. Following that, various helmet detection-related works are displayed. In [2], authors have suggested a method for detecting cycle riders without helmets in real-time using surveillance images. Using context subtraction and target segmentation, the suggested approach first identifies motorbike riders in surveillance recordings. The system then combines visual cues and a boolean classifier to determine whether or not the motorist is using a helmet. We also provide a convergence method for compliance reporting, which seeks to increase the dependability of the suggested technique. In [3] authors presented a method to detect motorcyclists without helmet on public roads. The system was divided in three steps:

1) Moving Object Segmentation
2) Moving Object Classification
3) Helmet Detection

In [6], a computer vision technology was proposed with the aim of detecting and segmenting motorcycles that were partially obscured by another car. A helmet recognition device is used, and the appearance of a helmet indicates the presence of a motorcycle. To identify the existence of a helmet, the edges of the probable helmet zone are calculated. The edge detector Canny is used. A helmet area is described by the number of edge points that resemble a circle. The method needs a lot of data from the user (helmet radius, camera angle, camera height, and so on) [7]. Chiverton [5] identified and tested a method for classifying and monitoring motorcycle riders wearing and not wearing helmets. The SVM algorithm, trained on histograms, is used in the scheme. The histograms were created via both static images and individual picture frames from surveillance video taken from the heads of motorcycle riders. While the accuracy rate was good, the quantity of test photos was insufficient.

Proposed Work

This part illustrates our methodology of detecting bike riders without helmets in real-time.

A. First stage: In the primary stage, the task is to recognize a person riding the bicycle in the video outline. For the subsequent stage, the task is to find the top of the bicycle rider to identify if the rider is making use of a helmet or not. To reduce false expectations, we integrate the results of back-to-back cases for a convincing projection. The proposed technique is depicted in Figure 1. The approach begins with the identification of moving objects and then moves on to the detection of bike riders. In the following stage, we subtract those who are wearing helmets and identify those who are riding without helmets. The helmet is highly significant only if there are moving bicycle riders; therefore creating whole case becomes a computational burden that does not improve detection rate. To proceed, we use background deduction on dim scale outlines with the purpose of recognizing moving and static things.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 13083 - 13091
Received 25 April 2021; Accepted 08 May 2021.

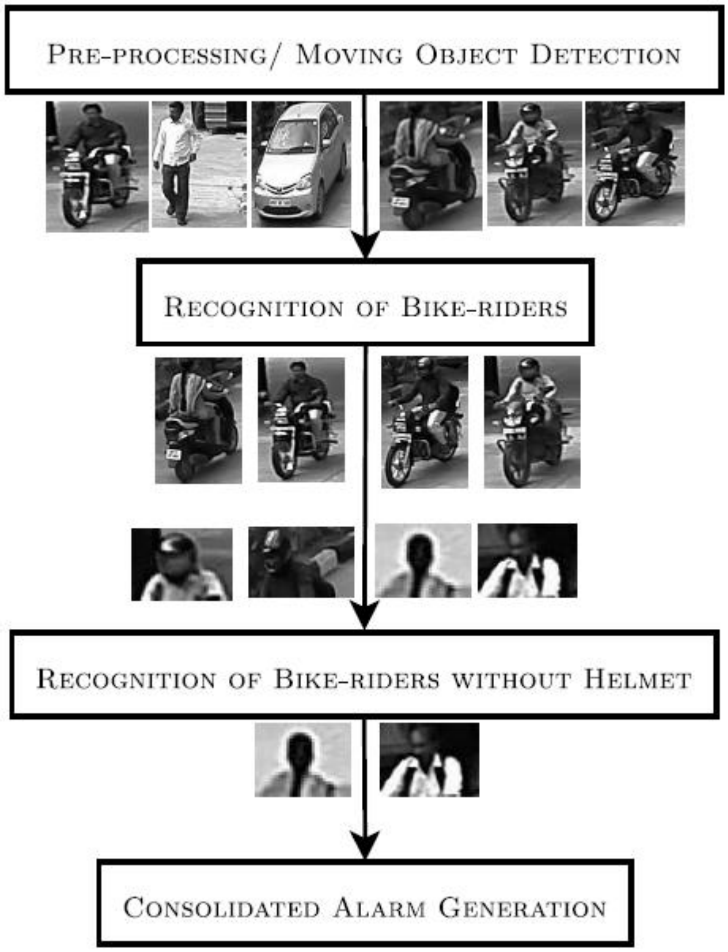

Figure 1. Proposed methodology for automatically detecting helmet-less bike riders [13]

B. Foundation Modeling: At first, the background deduction technique is utilized to isolate the articles moving, for example, bicycle, people, vehicles from static items, for example, trees, streets and structures. Be that as it may, there are sure difficulties when managing information from a solitary predetermined camera. Light differences during the day, shadows, trembling tree limbs, and other rapid changes make it difficult to recover and refresh backdrop from a continuous stream of frames. Due to the complexity and variability of the circumstances, a single Gaussian is insufficient to fully represent these variations. As a result of this reasoning, it is critical to employ a variable number of Gaussian models for each pixel. Here, K, the number of Gaussian segments for each pixel, is fixed between 3 and 5, which is estimated experimentally. The variable number of Gaussian segments enables the backdrop model to easily change its borders depending on the situation. In any event, a few errors may occur due to the nature of deeply obstructed 12t objects and consolidated shadows. Let us consider this, I... I shall be a pixel's power for the past t, continual casings. At that point at time t likelihood of watching force an incentive for a pixel is given by:

$$P(I^t) = \sum_{j=1}^{K} w_j^t \times \eta(I^t, \mu_j^t, \sigma_j^t) \qquad (1)$$

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 13083 - 13091
Received 25 April 2021; Accepted 08 May 2021.

In the above formula, the w^t_j is the weight, $\eta(\cdot, ,\cdot)$ is j^{th} Gaussian likelihood thickness, mean μ^t_j and σ^t_j as fluctuation at t^{th} time. Gaussian parts with low variance and high weight keep up a correspondence to the foundation class with every pixel, while others with big shift match up to the closer view class; at that point, the j^{th} component is regarded as a match. Similarly, the present pixel is referred to as the foundation or the foreground depending on the category of the j^{th} Gaussian model. The weight update rule is given by:

$$\left|\mu_j^t - I^t\right| < e_j \sigma_j^t, \quad (2)$$

$$w_j^t = (1-\alpha) w_j^{t-1} + \alpha(M_j^t), \quad (3)$$

$$M_j^t = \begin{cases} 0, & \text{for matched model} \\ 1, & \text{otherwise}, \end{cases} \quad (4)$$

where α is the learning rate, which decides how much of the time boundaries are balanced. Here, EJ is a limit which has a noteworthy effect when various locales have distinctive lighting. For the most part, t, the estimation of EJ is set at 3, as μt ± 3σjt records for 99% of information. Likewise, different boundaries of coordinated models are refreshed as:

$$\mu^t = (1-\rho)\mu^{t-1} + \rho I^t, \quad (5)$$

$$(\sigma^2)^{(t)} = (1-\rho)(\sigma^2)^{(t-1)} + \rho(I^t - \mu^t)^2. \quad (6)$$

Here, ρ = η(It|μj,σj). Since there is no associated part, a second Gaussian model with the existing pixel value as the mean, a low weight, and a high variance is generated.

Phase-I: Detecting Bike Riders

This stage includes the location of motorcycle riders in an edge. This progression utilizes identification of objects to segregate the ‘motorcycle rider' and 'others', in light of their visual highlights. This stage includes two stages: highlight extraction and grouping.

A. Feature Extraction

Item grouping requires some reasonable portrayal of diagram highlights. In writing, “HOG,” “SIFT,” and “LBP” are shown to be capable of object discovery and recognition. As a result, we analyze the accompanying highlights:

- **Histogram of Gradients**: The “HOG” descriptions have been shown to be quite effective in object recognition. Through inclinations, these descriptions capture neighborhood forms. In this nine containers are utilized, 8×8 pixels/cell and 2×2 cells per square. The subsequent element vector is h, where h ∈ Rn and n=3780.
- **Scale Invariant Feature Transform**: This approach seeks to capture critical focal points in the image. It splits highlight vectors for each key-point. These descriptors' scale, pivot, and enlightenment invariance provide heartiness in changing settings. A bag of words technique was utilized to craft a jargon V with a size of 6000. Planning SIFT descriptors to V at that moment results in element vector s, where s ∈ Rn and n is 6000. Highlight vectors are used to determine the proximity of images.

- **Nearly Binary Patterns**: These highlights capture surface data near the border of the screen. The pixels in the round neighborhood allow a double number threshold for each pixel, yielding the vector l ∈ Rn, where n is 26.

In the below pictures, the examples of stage I grouping inside 2-D space utilizing t-SNE is displayed. Scattering of the HOG includes the vectors, which demonstrates that the 2 classes, namely ‘bike riders' (blue color) and 'others' (red color), are a part of obvious regions consisting of not many exceptions as shown in the figure 2. This shows the element vectors proficiently speak to the movement and contains discriminative data, which further gives trust in great grouping precision.

B. Classification

After component extraction, the subsequent stage will divide the riders into 'bike riders' and 'other'. This is done using a 2-fold classifier. The SVM classifier is used because of its strength in arrangement execution in any event when prepared from less number of highlight vectors. Likewise, we utilize various bits, for example, direct, sigmoid (“MLP”), spiral premise work (“RBF”) to show up, best case scenario hyper-plane.

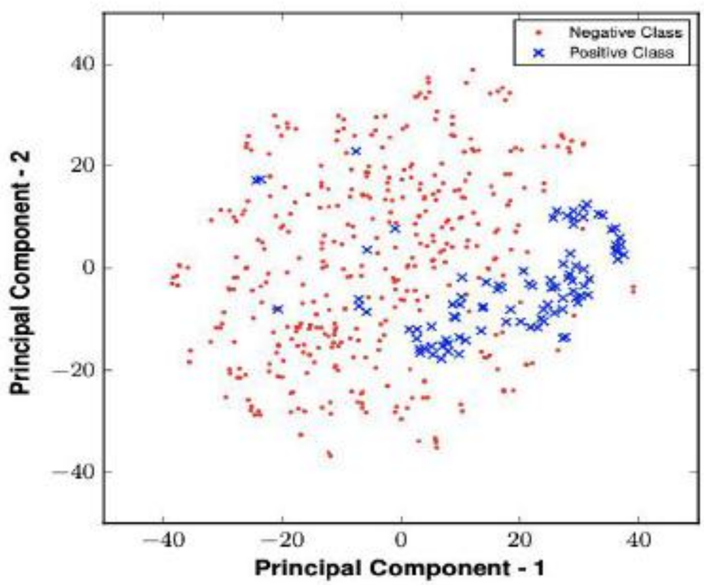

Figure 2. HOG feature vectors visualization for categorization of bike riders vs. others using t-SNE [11]

Phase-II: Recognition of Bike riders who do not wear Helmet

In the previous stage the clear identification was done for bike riders and others. In this stage, the task is to identify if the motorbike rider is wearing a helmet or not. Common face location calculations may not be adequate for phase-2 because of the accompanying reasons: i) Low goals represents an extraordinary test to catch facial subtleties, for example, eyes, nose, and mouth. ii) The point of development of the bicycle might be at coldhearted edges.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 13083 - 13091
Received 25 April 2021; Accepted 08 May 2021.

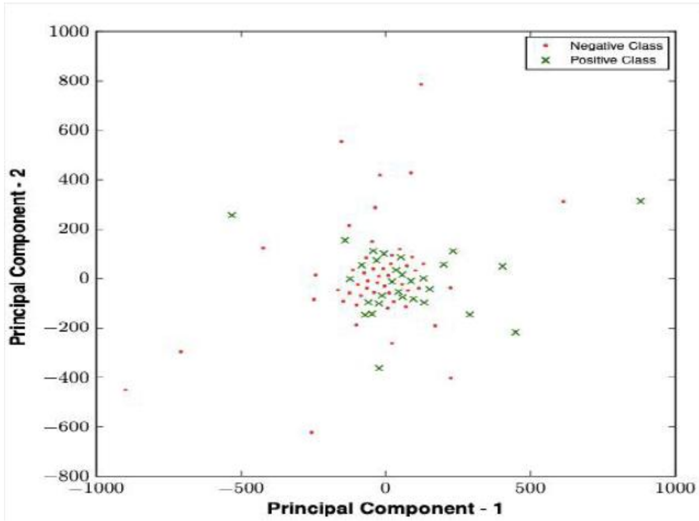

Figure 3. HOG feature vectors visualization for categorization of riders with helmet vs. non-helmet categorization using t-SNE [11]

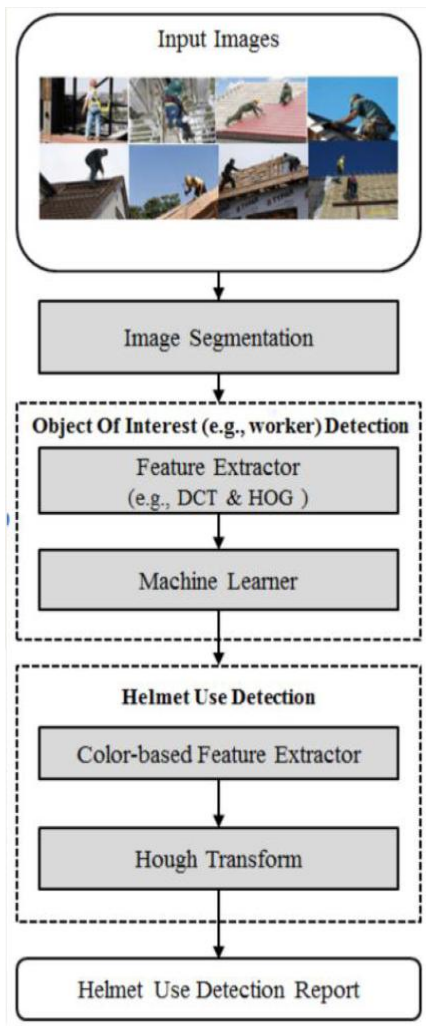

Figure 4. Complete Workflow

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 13083 - 13091
Received 25 April 2021; Accepted 08 May 2021.

In such cases, the face may not be obvious by any stretch of the imagination. So the proposed system recognizes the area around the head and afterward continue to decide if the bike rider is utilizing a protective cap or not. So as to find the head of the bike rider, the proposed structure utilizes the way that the fitting area of the cap will likely be in the upper regions of the bike rider. The bike riders wearing helmet can be detected by considering the top 1/4th of the image from the surveillance video. From this image, the part of the head is detected. Then construct a CNN model to distinguish the riders who are wearing the helmets from those who are not wearing the helmets. This progression is effectively reflected in our characterization results for stage II.

Results

To see the accuracy, the model is run for 15 epochs such that in each and every epoch the validation loss decreases and error loss increases. This confirms that the model gains the increasing of accuracy for each epoch and the same is plotted in the figure 5.

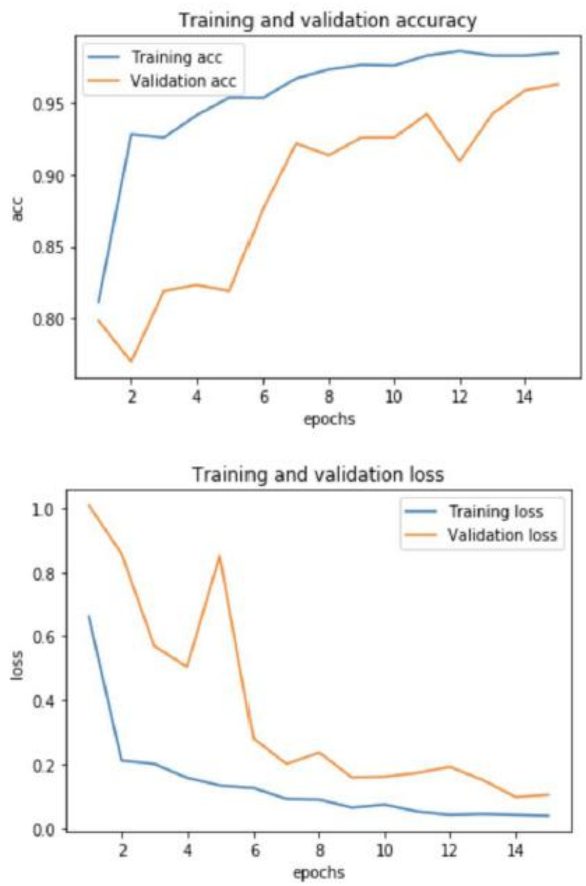

Figure 5. Training and Validation Accuracy and Loss

As the above figures clearly mentioned the validation loss and validation accuracy and the accuracy rate is measured from range 0 to 1, if the accuracy rate is more than 0.6, then it falls under good accuracy if it is less than 0.5 or less, it falls under low accuracy model.

Conclusion

This work is a study of the system which helps in detecting the bike riders without helmet from the surveillance videos. This will assist the law-enforcing authorities to take any action on such violators and also aid in maintaining the proper law and order in the city. Test results show the precision of 85% and 87% for recognition of bike riders. This work gives the result in 10.5ms with low error rate. Using the hybrid deep learning techniques, there is a scope to improve the accuracy of the model.

Acknowledgement

This work was supported by M S Ramaiah Institute of Technology, Bangalore-560054, and Visvesvaraya Technological University, Jnana Sangama, Belagavi - 590018.

References

[1] Dahiya, Kunal, Dinesh Singh, and C. Krishna Mohan. "Automatic detection of bike-riders without helmet using surveillance videos in real-time." *2016 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2016.

[2] Duan, Bobo, et al. "Real-time on-road vehicle and motorcycle detection using a single camera." *2009 IEEE International Conference on Industrial Technology*. IEEE, 2009.

[3] Chiu, Chung-Cheng, Min-Yu Ku, and Hung-Tsung Chen. "Motorcycle detection and tracking system with occlusion segmentation." *Eighth International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS'07)*. IEEE, 2007.

[4] Silva, Romuere, et al. "Automatic detection of motorcyclists without helmet." *2013 XXXIX Latin American Computing Conference (CLEI)*. IEEE, 2013.

[5] Zivkovic, Zoran. "Improved adaptive Gaussian mixture model for background subtraction." *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*. Vol. 2. IEEE, 2004.

[6] Otsu, Nobuyuki. "A threshold selection method from gray-level histograms." *IEEE transactions on systems, man, and cybernetics* 9.1 (1979): 62-66.

[7] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*. Vol. 1. Ieee, 2005.

[8] e Silva, Romuere Rodrigues Veloso, Kelson Rômulo Teixeira Aires, and Rodrigo de Melo Souza Veras. "Helmet detection on motorcyclists using image descriptors and classifiers." *2014 27th SIBGRAPI Conference on Graphics, Patterns and Images*. IEEE, 2014.

[9] Stauffer, Chris, and W. Eric L. Grimson. "Adaptive background mixture models for real-time tracking." *Proceedings. 1999 IEEE computer society conference on computer vision and pattern recognition (Cat. No PR00149)*. Vol. 2. IEEE, 1999.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 6, 2021, Pages. 13083 - 13091
Received 25 April 2021; Accepted 08 May 2021.

[10] Guo, Zhenhua, Lei Zhang, and David Zhang. "A completed modeling of local binary pattern operator for texture classification." *IEEE transactions on image processing* 19.6 (2010): 1657-1663.

[11] Van der Maaten, Laurens, and Geoffrey Hinton. "Visualizing data using t-SNE." *Journal of machine learning research* 9.11 (2008).

[12] Vishnu, C., Singh, D., Mohan, C. K., & Babu, S. (2017, May). Detection of motorcyclists without helmet in videos using convolutional neural network. In 2017 International Joint Conference on Neural Networks (IJCNN) (pp. 3036-3041). IEEE.

[13] Singh, D., Vishnu, C., & Mohan, C. K. (2016, December). Visual big data analytics for traffic monitoring in smart city. In 2016 15th IEEE international conference on machine learning and applications (ICMLA) (pp. 886-891). IEEE.