Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. . 4727 - 4733
Received 25 April 2021; Accepted 08 May 2021.

Video Analysis of Vehicle and Pedestrian Using Neural Network

Varsha P[1] , Priyadharshini D K[2], Swetha S[3], A.Tamizharasi[4]
Department of CSE, R.M.D. Engineering College,Kavaraipettai,Tamil Nadu 601206,India.
vijaya.varsha032@gmail.com[1], dkpriya1999@gmail.com[2], swetha10899@gmail.com[3], tamizh.cse@rmd.ac.in[4]

ABSTRACT

In this paper, we suggest an algorithm for detection and tracking analysis of vehicles, pedestrians in video frames. The extracted features function for visual perception tasks and are used to address a tracking problem. The algorithm can be split into 2 sections the first part is classification algorithm, which is predicted on convolutional neural networks. The CNN model is implemented using the You Only Look Once (YOLO) algorithm in deep learning. And the second part is algorithm for tracking regions. Using this algorithm, we suggest an idea of detecting, counting of vehicles and pedestrian.

Keywords
Classification, Tracking, Convolutional Neural Network, You Only Look Once, CNN model.

1. INTRODUCTION

Video monitoring of vehicles and pedestrians is now in use and has business and academic applications. Several works have been represented in the prior art using various methods, some are now considered standard, with numerous examples from business and academic implementations. Classification of images and object detection is one of the best applications of convolutional neural networks. It can be commonly used in video surveillance, and also in crowd monitoring to detect persons for counting for general statistics and many other uses. In this paper, we outline a video-tracking algorithm for scanning vehicles and pedestrians that employs a deep convolutional neural network (CNN) to extract features, detect objects, monitor them, and classify them from each frame of video. The standard performance of CNN object detection can be seen as boxes surrounding the frame of the object. We suggest an algorithm for laying out a video detecting result in video frames for vehicular and pedestrian monitoring. For several years, vehicle and pedestrian tracking algorithms separated the scene into foreground and background (FG-BG) before examining a binary threshold with contours or blob detection. There are some drawbacks to this FG-BG technique. For example, object shadows are tough to separate from the object, so this approach cannot provide an easy solution for organizing the targeted and detected object. And it is primarily influenced by stage lighting and noise in the picture.Deep convolutional neural network object detection has emerged as a viable alternative to FG GB in road video analysis, owing to the increased interest in deep learning among the general public due to its superior performance in image classification and detection. As a result of these works in deep learning neural networks,we comewith idea of providing trained sets and networks that make the tracking algorithm easier and also works either for FG-BG and image tracking, object evaluations.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4727 - 4733
Received 25 April 2021; Accepted 08 May 2021.

2. SYSTEM ANALYSIS

2.1 EXISTING SYSTEM

Here we outline a video monitoring algorithm for scanning vehicles and pedestrians that employs a deep convolutional neural network (CNN) to extract features, detect objects, monitor them, and classify them from each frame of video. The standard performance of CNN object detection can be seen asboxes surrounding the frame of the object in the picture. Here we use YOLO algorithm of version 2 in the convolutional neural network because of its higher speedwhen compared to other versions of YOLO. There are about 20 classes of objects that YOLO can identify and used for training and testing.

DISADVANTAGES OF EXISTING SYSTEM

The object shadows are tough to split from the object, thus this approach cannot give easy solution to organize the targeted and detected object. And it is primarily influenced by stage lighting and noise in the picture.

2.1 PROPOSED SYSTEM

We suggest an algorithm for laying out a video detecting result in video frames for vehicular and pedestrian monitoring. The vehicle and pedestrian tracking algorithms separated the scene into foreground and background (FG-BG) before examining a binary threshold with contours or blob detection.

ADVANTAGES OF PROPOSED SYSTEM

We come up with an idea of target tracking algorithm that has trained sets and networks of YOLO, and works either for FG-BG and image tracking, object evaluations.

3. MODULES

The entire project mainly consists of 4 modules,
which are:

GRAYING AND DE NOISING OF IMAGE

Graying of an image necessitates near instantaneous in vehicle recognition and object monitoring. On the other hand, the complex enumeration of images in various colors reduces the processing efficiency and results in inadequate real time output. The transition of various colored images to gray needs enumeration, enhanced processing speed, and refined features. The weighted arithmetic mean method is employed for gray scaling the images. Noise reduction in images is carried out when devices and the surroundings yield high-frequency sound whilst capturing videos. Such grainy images affect the general efficiency of auto recognition and locating. As a result, noise reduction in images is carried out in order to increase the precision of successive enumerations.

IMAGE SEGEMENTATION

To distinguish the object from complicated and inoperable backdrops, image sectionalization is carried out to help the variance between the objective and background characteristics. Currently, image sectionalization algorithms, including optical flow and frame variance techniques, are used for vehicle detection and locating applications.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4727 - 4733
Received 25 April 2021; Accepted 08 May 2021.

On the other hand, the stable background prototype will be unable to form in the course of the experiment due to the complex surroundings. A case in the point of this setting is the state of automobiles in various states, such as when they are driving at high pace rates, travelling slowly, or stopping at the side of the lane. Frontal vehicle states should be recognized and monitored.

VEHICLE DETECTION

Vehicle recognition mainly focuses on to determine the proportions, path, and additional details of the objective vehicle utilizing the geometric and entity features of image in prior processing. The current version is the most widely used system for frontal automobile recognition. Most detection and recognition techniques, like optical flow and frame variance, are not appropriate to our videos and images because of the complicated backdrop, lighting changes, and thus the different states of vehicles.

TARGET TRACKING

The presence of the objective, as well as all other details about the detected position, scale, and pace of target vehicles from prior segments, is resolved by video-based target monitoring. Currently, there are three types of video-based target tracking algorithms: 3-D model-based tracking, feature-based tracking and region-based tracking methods. Feature based tracking method indulges simple enumeration. On the other hand, complicated movements cannot be designated using this technique. This approach is additionally vulnerable to noise in the course of initiation.

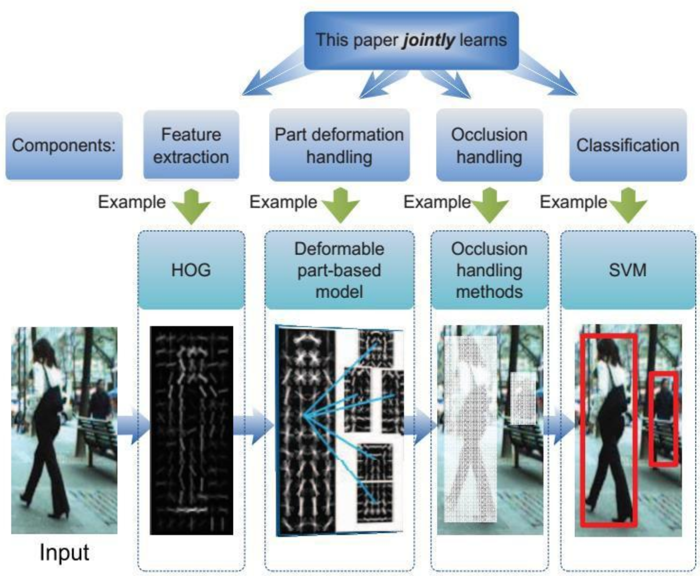
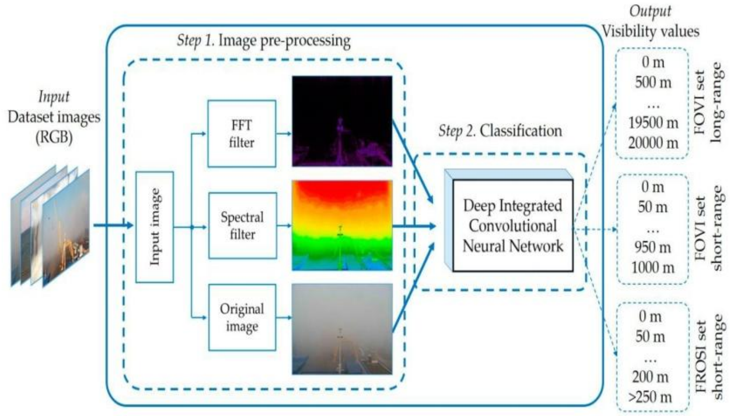
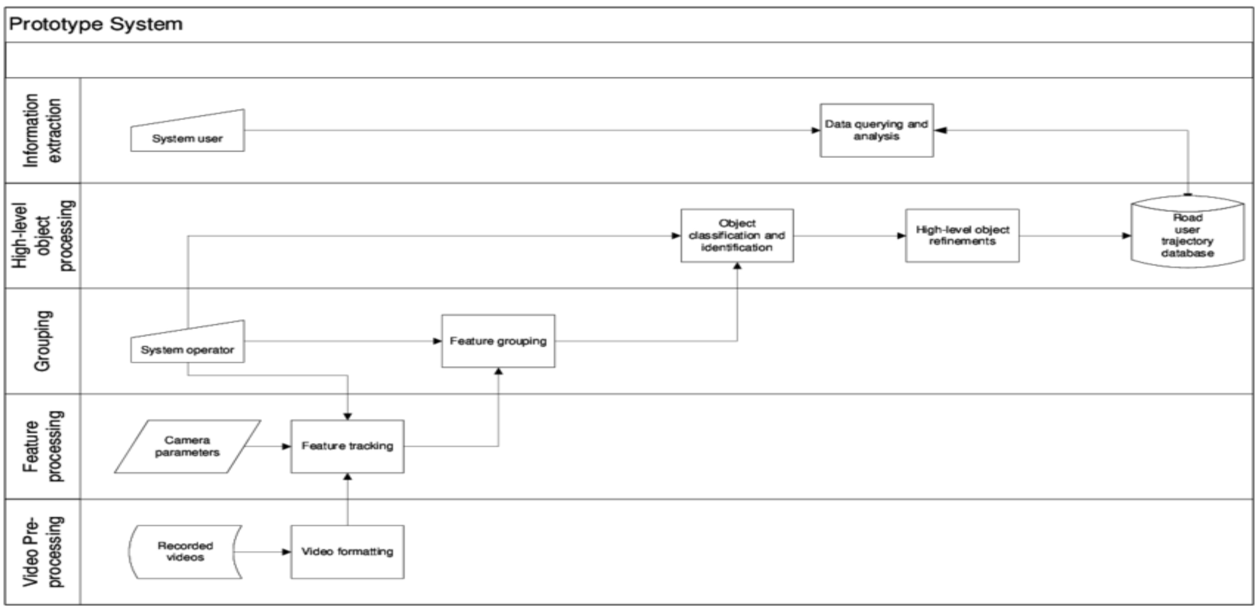


Figure 1. Motivation of this paper to jointly learn the four key components in pedestrian detection: feature extraction, deformation handling models, occlusion handling models, and classifiers.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4727 - 4733
Received 25 April 2021; Accepted 08 May 2021.





4. SYSTEM FRAMEWORK

Object segmentation:

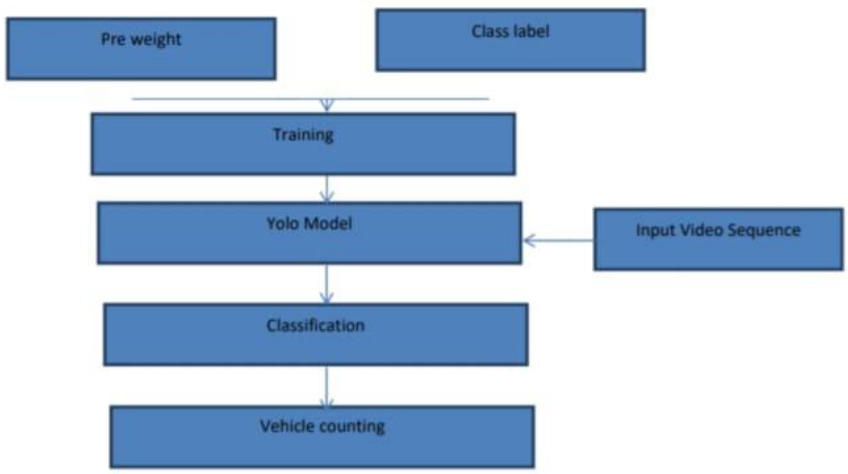
Most tracking algorithms need a contiguous perception of the object in a video stills for initial input. The thing is separated using the performance of the YOLO algorithm by making use of Rectangular region Of Interest(ROI). In our implementation, the number of frames used was 20 FPS. The object recognition algorithm is utilized in each frame.

Clustering:

Since the identification of an object is determined in numerous frames, the algorithms supported by the neural networks issue numerous ROIs for the equivalent object, these rectangles are divided into two sectionsinforeground-background algorithms, usuallythe object is segmentedas manyoverlapping rectangles.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4727 - 4733
Received 25 April 2021; Accepted 08 May 2021.

The output of YOLO oftenproduce several rectangles for the same object, yet these rectangles are extremely overlapped.



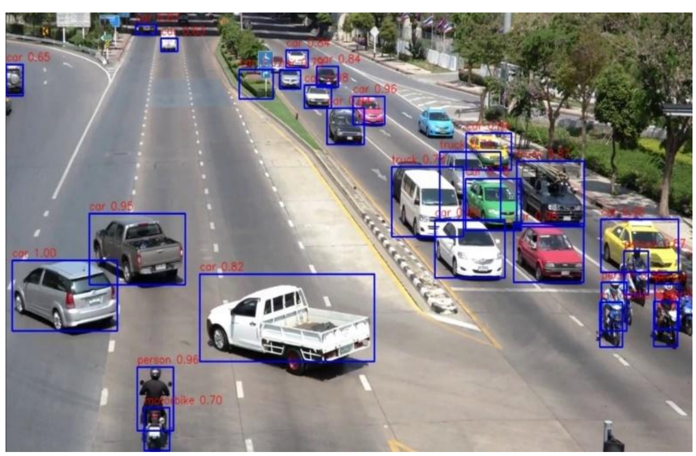
Feature extraction:

The tracking algorithm gathers details to combine the objects in a singular frame using the ROI – centroid pixel coordinates, the rectangle dimension and the YOLO estimated mark.

5. RESULT AND DISCUSSION

The expected result provides the identity of the vehicle and the pedestrian.

Finds the actual speed of theobject in motion and differentiates the object and its shadow.



The FG-BG algorithm produces a number of issues. In vehicle video processing, deep convolutional neural network object recognition has emerged as a viable replacement to FG-BG.

Deep learning has good image recognition and object perception efficiency as well as a lot of estimating ability in contemporary graphics card and is simple to implement.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4727 - 4733
Received 25 April 2021; Accepted 08 May 2021.

The training sets and trained networks will be provided to differentiate foreground background and object detection predictions.

6. CONCLUSION

For more than 60 years, the invention of vehicle measuring instruments has been a constant function used to document the application automobiles on the railroad tracks. Elongating this skill to additional forms of transportation like walkers and cyclists helps in gaining a better grasping of citizen space allotment and recognize requirements and possibilities in civic areas. The most common roadblocks in processing an algorithm creation is the inadequacy of data sets to test and evaluate various algorithms which is likely to be established. This problem is partially solved with the data provided; however, future work would be required to reconsider the need for additional new groups to be categorized, in addition to the nine already identified. In the absence of closures amidst objects, identification is nearly ideal, under this presumption, vehicle and pedestrian calculating is solved by properly positioning the camera. In the state of non-occlusion, the identification and counting of bicycles exceeds a detection threshold of 97.6% and is only impaired when objects are close together and it is impossible to differentiate between them. To estimate an actual value under congested situations, the adequacy of vehicles or bicycles per square meter have to be calculated and the total area occupied equally by clustered objects have to be predicted. This work will guide the people to develop new identification and tracking algorithms making it easier to extract civic traffic criteria.

7. REFERENCES

[1]. A. M. Ocana and F. Calderon, “Preliminary studies on the taxonomy of object’s tracking algorithms in video sequences,” in 2012 XVII Symposium of Image, Signal Processing, and Artificial Vision (STSIVA). IEEE, sep 2012, pp.153–157.

[2]. A. Sobral and A. Vacavant, “A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos,” Computer Vision and Image Understanding, vol. 122, pp. 4 – 21,2014.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4727 - 4733
Received 25 April 2021; Accepted 08 May 2021.

[3]. J. H. B. S. Rodrigo Benenson, Mohamed Omran, "Ten years of pedestrian detection, what have we learned?" ECCV 2014 Workshops. Lecture Notes in Computer Science, vol. 8926, no. 1,2015.

[4]. Y. Xu, J. Dong, B. Zhang, and D. Xu, "Background modeling methods in video analysis: A review and comparative evaluation," CAAI Transactions on Intelligence Technology, vol. 1, no. 1, pp. 43 – 60, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2468232216000068

[5]. G. Urrego, F. Calderon, A. Forero, and J. Quiroga, "Adquisici´on de variables de tr´afico veh´ıcular usando visi´on por computador," Revista de ingenier´ıa Universidad de los Andes, vol. 1, no. 30, pp. 7–15,2009.

[6]. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105. [Online]. Available:http://papers.nips.cc/paper/4824-imagenet-classification-with-deep- convolutional-neural-networks.pdf

[7]. D. Tom`e, F. Monti, L. Baroffio, L. Bondi,

M. Tagliasacchi, and S. Tubaro, "Deep convolutional neural networks for pedestrian detection," Signal Processing: Image Communication, vol. 47, pp. 482 – 489, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0923596516300637

[8]. Y. Hsu, S. Chang, and J. Guo, "A multiple vehicle tracking and counting method and its realization on an embedded system with a surveillance camera," in 2018 IEEE International Conference on Consumer Electronics-Taiwan(ICCE-TW), May 2018, pp. 1–2

[9]. P. Kuang, T. Ma, F. Li, and Z. Chen, "Real-time pedestrian detection using convolutional neural networks," International Journal of Pattern Recognition and Artificial Intelligence, vol. 32, no. 11, p. 1856014, 2018. [Online]. Available: https://doi.org/10.1142/S0218001418560141

[10]. "Urban dataset, making the visual information of urban scenes available for all," http://urban-dataset.com/, 12 14, (Accessed on14/12/2018).

[11]. J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," CoRR, vol. abs/1612.08242, 2016. [Online]. Available:http://arxiv.org/abs/1612.08242

[12]. "Yolov3: An incremental improvement," CoRR, vol. abs/1804.02767, 2018. [Online]. Available:http://arxiv.org/abs/1804.02767

[13]. R. Real and J. M. Vargas, "The probabilistic basis of jaccard's index of similarity," Systematic Biology, vol. 45, no. 3, pp. 380–385,1996.

[14]. D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two Frame Stereo Correspondence Algorithms," International Journal of Computer Vision, vol. 47, no. 1-3, pp. 7–42, 2002. [Online]. Available: http://link.springer.com/article/10.1023/A:1014573219977

[15]. P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," PAMI, vol. 34,2012.