ESKİŞEHİR TECHNICAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY
A- APPLIED SCIENCES AND ENGINEERING

2020, 21(3), pp. 446-453, DOI: 10.18038/estubtda.629701

REAL TIME PEDESTRIAN ALERT SYSTEM FOR VEHICLES

Şahin IŞIK [1,*], Göksu VATANSEVER [1], Yıldıray ANAGÜN [1], Mehmet ÇELİKHAN [1], Kemal ÖZKAN [1]

[1] Department of Computer Engineering, Faculty of Engineering and Architecture, Eskisehir Osmangazi University, Eskisehir, Turkey

ABSTRACT

In this study, we have developed a pre-collision alert system for vehicles in terms of detection pedestrians in road. The system is consisting from deep learning models and transfer learning methodologies. For this purpose, pre-trained convolutional models were considered to detect pedestrian and road. Finally, the segmented road mask and pedestrian mask were utilized to reveal the intersection of these two masks. The system generates an alert if the number of pixels is higher than predefined threshold value. By considering the test results, the proposed system gives 83.33% accuracy rate when it comes to use it to avoid collision.

Keywords: Deep learning models, Pedestrian alert system, Road detection, Pedestrian detection

1. INTRODUCTION

Depending on the increasing population and the number of vehicles, careless driving raises the risks of traffic injuries for pedestrians. The main causes of inevitable accidents are usually reasoned by dangerous driving and pedestrians crossing a road in a reckless manner. To minimize these risks in traffic, efforts have been geared up to upgrade the systems in vehicles, i.e., enabling communication between vehicle to vehicle (v2v) and vehicle to pedestrian (v2p). With the help of integrated smart car systems, the breaking system can instantly activate in case of an unexpected collision occurred.

There are many awesome automotive technologies in smart vehicle systems. One of them is Pre-Collision assist system, which is developed by FORD company [1]. Based on a camera, the system detects a potential collision, when a vehicle or pedestrian stated in front of the car in case of both nighttime or day driving. Pre-Collision assist with automatic breaking avoids frontal collision situations with three ways either slowing down or stopping immediately. First, it provides an audible chime, which presents a warning message. Second, if the risk of collision increases, then Automatic Emergency Braking (AEB) will pre-charge and facilitate the control over the brake in terms of rapid breaking. Finally, if the collision is eminent, automatic breaking system activated without driver input help. Again, VOLVO uses pedestrian detection using cameras and radars to prevent any collision to the pedestrians with automatically brake system.

On the other hand, HONDA provides v2p technology by communicating with pedestrians' mobile phones to prevent possible collisions [2]. With the DSRC (Dedicated Short Range Communication) messages, the vehicle communicates with pedestrians' phones, and if a possible collision has been forecasted between that vehicle and the pedestrian, an audible and visual warning is sent to the pedestrians' phone and the vehicle. In this way, both careless pedestrians and careless drivers would be warned to avoid instant collisions. If again the critical value is exceeded, the vehicle automatically activates the braking system and ensures the safety of the two sides. Apart from all of these, the technologies used in smart vehicles are quite advanced and face recognition, voice recognition and both of them are equipped with systems that can make adjustments according to these recognitions. For

*Corresponding Author: sahini@ogu.edu.tr
Received: 05.10.2020 Published: 30.09.2020

example; BOSCH has developed a system that provides face recognition [3] in their smart car and that can automatically play favorite tracks according to the recognized face.

Moreover, there has been numerous studies to investigate the performance of machine learning based approach for pedestrian and vehicle recognition. Mainly, the deep learning algorithms produced valuable results in recent years. In a study [4], YOLO-v2 combined with SqueezeNET model, was applied to localize pedestrians by using a GPU graphic card. After conducting experiments with Tensorflow and OpenCV, the detection performance for Advanced Driver Assistance Systems (ADAS) was reported as 75.8% mAP. The potential capability of Mask-R-CNN algorithm [5] was inspected to localize cars, buses and pedestrians on the road to improve the accuracy of a anti-collision warning system. The Long Short-Term Memory (LSTM) neural network model [6] was utilized to produce predictions about pedestrians' red-light violation behaviors. After modelling the LSTM architecture with real traffic data, which includes labelled pedestrians' unexpected crossing intentions, the accuracy of system was established with 91.6% rate in case of testing at one signalized crosswalk. With a different approach [7], the Faster R-CNN was utilized to recognize pedestrian in urban scenarios. In referred system, the 600 samples reserved for training, while 100 samples for testing and 100 samples for validation purposes. Totally, 16000 samples labeled as pedestrian candidates and the result was accounted to 87.75% mAP score when testing on 100 samples.

The negative radiation effects of the Wi-Fi used in roads will be influenced people life much worse than before due to increased number of vehicles on the road. Instead of this, the object detection technology is a better choice that is configured to eliminate the possible collisions. In this regard, we have studied on a deep learning based alert system, which detects pedestrians in front of the car. The proposed study consists of three stages. First, pre-trained FCN-VGG16 [8] convolutional neural network used for road detection. Second, pre-trained SSD (Single Shot MultiBox Detector) Mobilenet v1 [9, 10] convolutional neural network used for human detection. In the final stage, we localize intersection of pedestrian and road by analyzing the segmented regions. If the pedestrian is crossing on the way, the system will produce an alarm.

The rest of the paper is organized as follows. Section II presents the technical information about developed system. Section III shows the findings through experimental study. Finally, in the last section, a conclusion is given about proposed methods.

2. METHODS

2.1. Part1: Human Detection Framework

In this study, the Object Detection API provided by Tensorflow [11] was used for object detection. This API is a framework that utilizes pre-trained models prepared by Google as open source to users. SSD Mobilenet v1 model [9, 10] was used for human recognition process and model was trained on the COCO (Common Objects in Context) dataset. Further, in COCO datasets, there are commonly used daily objects. Depending on the gpu and cpu performance of the system, more powerful and faster models can be used instead of the SSD Mobilenet v1 (refers to Figure 1) model that best suited for our proposed system.

There are two types of neural networks as base network and detection network. The SSD architecture is built on the VGG-16 [12] base network architecture. VGG-16 was chosen as the base network due to its strong performance in high quality image classifications and problem solving skill in learning phase. Inputs are gradually reduced by passing through these layers of VGG-16 like CNN structure.

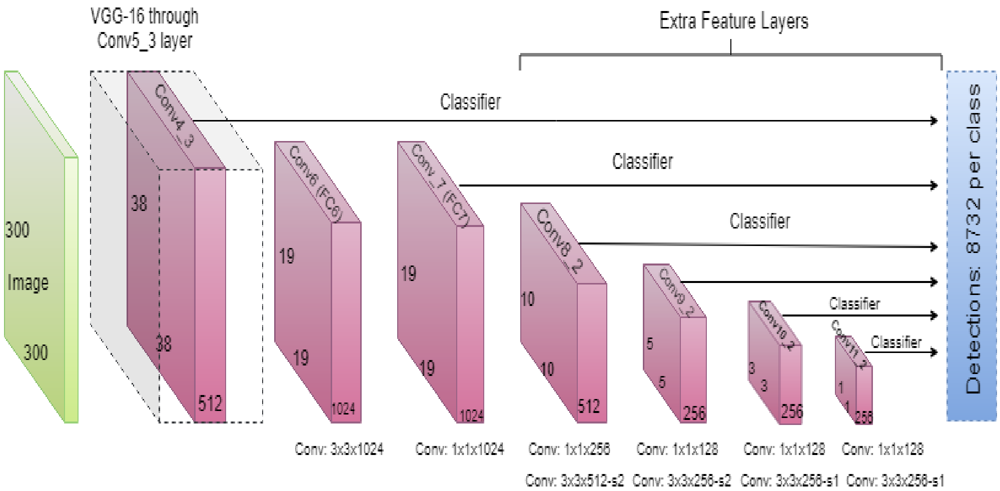

Figure 1. SSD Mobile-Net Structure.

SSD, which stands for Single Shot Detector is using to realize the multi-box detectors. SSD takes only one single shot to detect multiple objects within the image. Regional Proposal Network (RPN) based series take two shots. In general, SSD is much faster than other models. SSD series achieve average 75% Mean Average Precision (mAP) and frequency per second could be from 22(fps) to 59(fps). Accuracy can be increased according to the used feature extractor such as MobileNet [9], ResNet [13], VGG [12] etc. The more details about which detector and what configurations give us the best balance between speed and accuracy is available at the link of [14].

Figure 2. Pedestrian detection with SSD Mobile Net v1 model.

The Figure 2 presents illustrations about the pedestrian detection by using the SSD Mobile Net v1 model.

2.2. Part2: Road Detection Framework

FCN-VGG16 convolutional neural network, which was trained with KITTI dataset, was employed for the segmentation of the road. We used a pre-trained model in our study. (https://junshengfu.github.io/semantic_segmentation/).

VGG16 is a 16-layer convolutional neural network architecture that includes 3x3 filters and performs these filters in convolutional layers to extract low and high features in the image. 3x3 filters obtain more accurate results and feature maps. 1x1 filters can be used in addition to 3x3 convolutional filters.

As well, max pooling reduces the computational cost by reducing the number of parameters and dimension of the data. Discriminative learning and multi-class classification are performed in Fully Connected Layer and Softmax Layer, respectively. As shown in Figure 3, the FCN-VGG16 model were applied to reveal the road region.

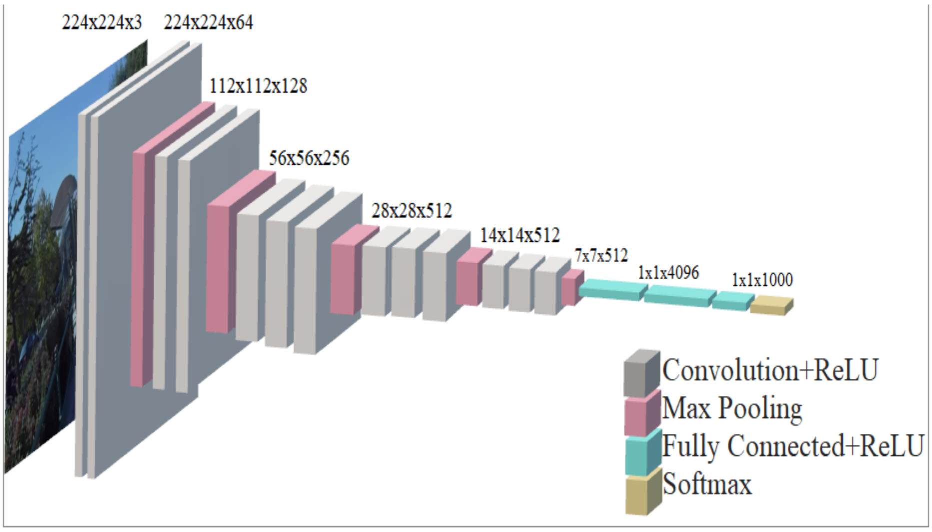

Figure 3. Road detection with FCN-VGG16 Net.

2.3. Proposed Pedestrian Alert System

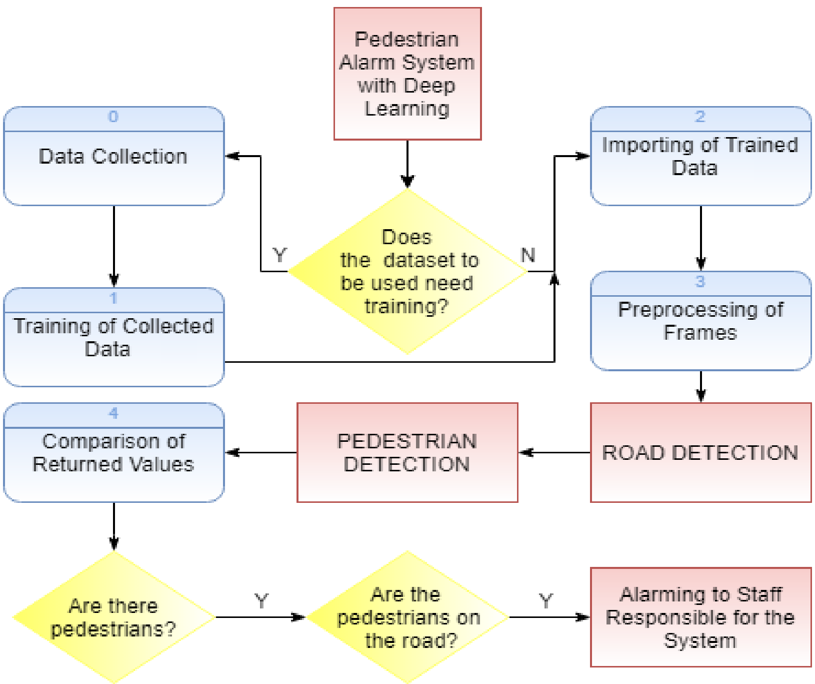

Figure 4. The overall flowchart of proposed system.

Figure 4 shows the he flowchart of the proposed pedestrian alarm system for vehicles. Firstly, the road regions were segmented with respect to the utilized pre-trained FCN-VGG16 model. Then, the pedestrians were detected by using SSD MobileNet v1. Regarding to examples shown in Figure 5, one can emphasize that the deep learning techniques can detect easily and accurately whether the person is on the road or not, when making experiments with not complex traffic scenarios.

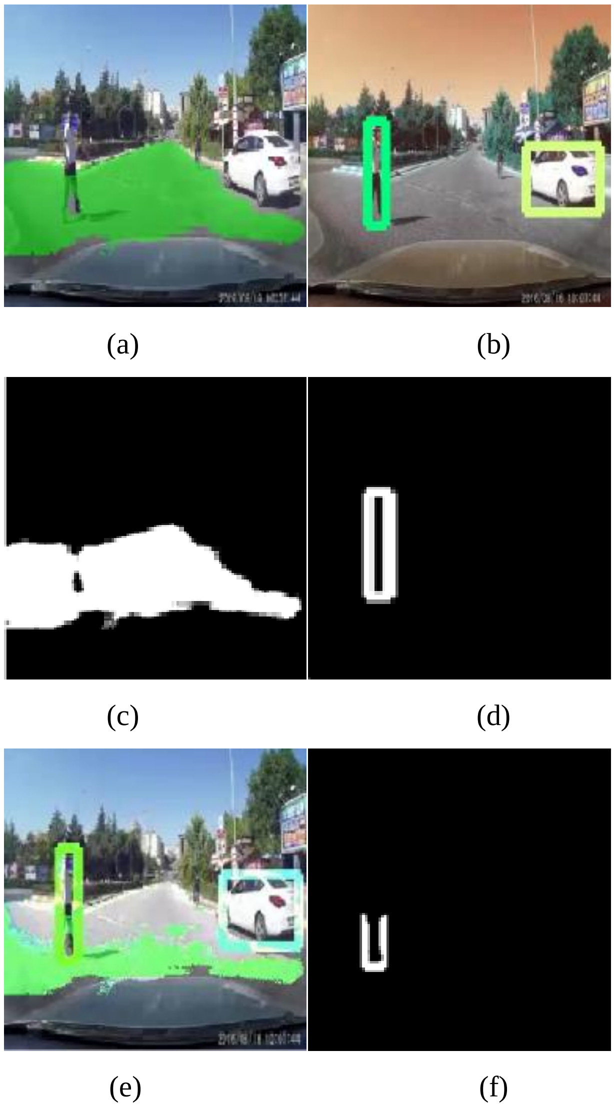

Figure 5. Working stages of the proposed system**.**

a. Generated frame from road segmentation
b. The result frame of the logical operation (OR) between the human detection system and the road segmentation system.
c. The road segmentation mask
d. Human detection system mask
e. Generated frame from human detection system.
f. Intersection of road segmentation mask and human detection mask

3. RESULTS

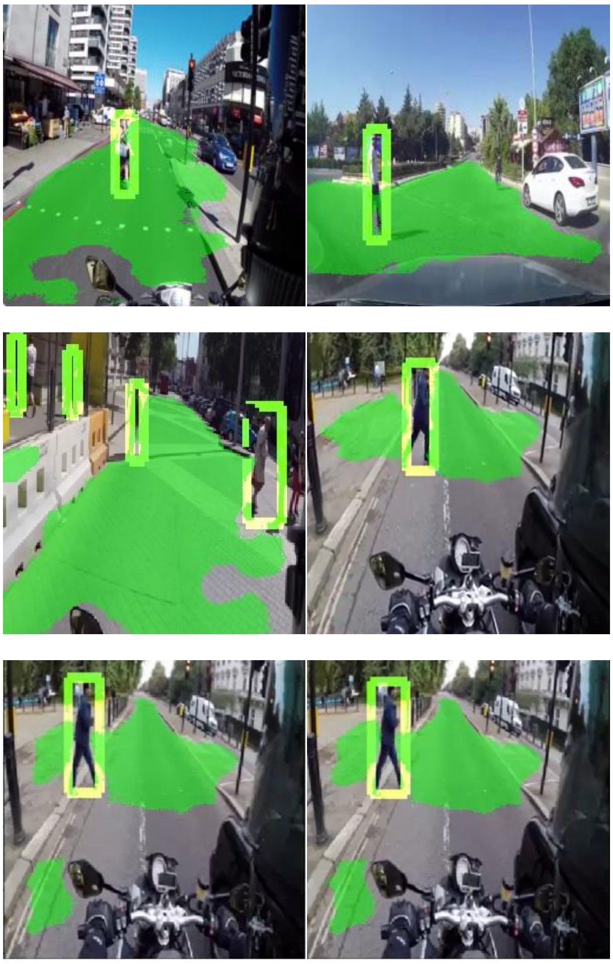

Figure 6. Some visual results.

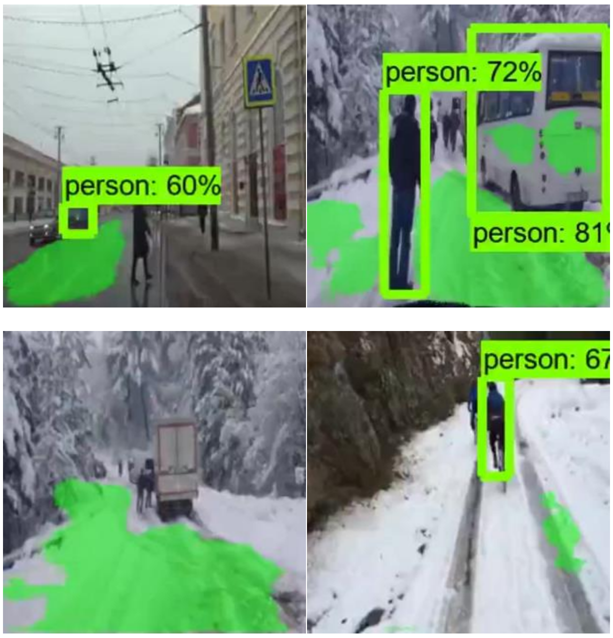

Figure 7. Some visual results in rainy and winter days.

The frame-rate of test videos is 30 and size of images are 450x720. Also, the samples of test video are converted to 256x256 size when estimating predictions in code behind. After empirical evaluations, we have heuristically determined a threshold value in order to produce an alert. By counting the number of pixels in segmented region related to intersection of pedestrian and road, the threshold value is determined as 200. In all experiments, it was observed that this heuristic threshold conformed to different intersections regions between pedestrian and road.

Some visual results related to the system can be analyzed in Figure 6. One can observe that the system detects pedestrians with good accuracy score. Also, the Figure 7 shows the performance of proposed system under rainy and winter days. In case of winter days, the proposed system may collapse due to hidden information about road. However, for rainy days, the system can catch the pedestrian in the road. The drawback of proposed system is that the detection of pedestrians may fail when the vehicle is too fast which may results in blurred images. This problem can be minimized by increasing the number of frames per second or generating predictions with higher resolution of images. One way to overcome such situations is re-training the system under complex conditions such as the distance between pedestrian and vehicle is high. For this purpose, a more sophisticated camera system is required to reduce the risk of collision. Meanwhile, a robust Central Processing Unit (CPU) or Graphics Processing Unit (GPU) is required to detect pedestrian in real time.

The performance of proposed system is determined with accuracy metric. In order to address the robustness of proposed system, we have made experiments with test 30 samples. Samples are a related to one second of a real video. The proposed system can accurately detect 25/30 samples when pedestrian walks across the road. Therefore, the performance is reported to 83.33% accuracy rate. Although the detection accuracy of [6], is higher than the proposed method, but one can say that each method is useful for different tasks.

The test simulations of this study were carried out in the same hardware (Intel core i5-4210U with 2.4 GHz CPU and 4 GB memory) with a software implemented in Tensorflow Python object detection library. To conclude, a more systematic and theoretical analysis is required for correct localization of pedestrians on the road.

4. CONCLUSIONS

In this study, we have developed a pre-collision assist system to generate alert when pedestrian crossing on the road. It can be said that the performance of the proposed system has reached an accuracy rate of 83.33% and the main disadvantage of the proposed system is that it is very difficult to divide the road zones in winter days. As a future study, this project can be improved for detection of animals in the road by re-training. Again, the other functionality of this project can be expressed that the persons entered into restricted areas can be detected for factory site or prohibited military zone. Moreover, the face of detected persons entered in prohibited area can be detected and recorded to track their behaviors and activities. The system can be expanded for different purposes and ideas after re-training the models.

REFERENCES

[1] 'Ford, Https://Www.Ford.Com/Technology/Driver-Assist-Technology/Pre-Collision-Assist/, Accessed 09.27.2019'.

[2] 'Qualcom, Https://Www.Qualcomm.Com/Videos/Honda-V2p-Overview, Accessed 27.09.2019'.

[3] 'Bosch, Https://St-Tpp.Resource.Bosch.Com/Media/Technology_Partner_Programm/10_Public/ Application_Notes/Iss_Ipp_Application_Note_210x280_Final_Lowres_Hyperlinks.Pdf, Accessed 09.27.2019'.

[4] Said YF. and Barr M. Pedestrian Detection for Advanced Driver Assistance Systems Using Deep Learning Algorithms, International Journal of Computer Science and Network Security, 2019; 19, (9), p. 10.

[5] Xu C, Wang G, Yan S, Yu J, Zhang B, Dai S, Li Y and Xu L. Fast Vehicle and Pedestrian Detection Using Improved Mask R-Cnn. Mathematical Problems in Engineering, 2020.

[6] Zhang S, Abdel-Aty M, Yuan J and Li P. Prediction of Pedestrian Crossing Intentions at Intersections Based on Long Short-Term Memory Recurrent Neural Network. Transportation Research Record, 2020; p. 0361198120912422.

[7] Fuentes A, Jun I, Yoon S and Dong S. Pedestrian Detection for Driving Assistance Systems Based on Faster-Rcnn, in, International Symposium on Information Technology Convergence Isitc, 2016.

[8] Long J, Shelhamer E and Darrell T. Fully Convolutional Networks for Semantic Segmentation', in, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015.

[9] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, and Adam H.J.a.p.a. Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications, 2017.

[10] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C-Y and Berg AC. Ssd: Single Shot Multibox Detector, in, European Conference on Computer Vision, Springer, 2016.

[11] Tensorflow, Https://Www.Tensorflow.Org/, Accessed 09.27.2019.

[12] Simonyan K and Zisserman A.J.a.p.a., Very Deep Convolutional Networks for Large-Scale Image Recognition, 2014.

[13] He K, Zhang X, Ren S and Sun J. Deep Residual Learning for Image Recognition, in, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[14] 'Medium, Https://Medium.Com/@Jonathan_Hui/Object-Detection-Speed-and-Accuracy-Comparison-Faster-R-Cnn-R-Fcn-Ssd-and-Yolo-5425656ae359, Accessed 10.03.2019'.