International Journal of Science and Engineering Applications
Volume 11-Issue 12, 250 – 253, 2022, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1112.1001

A Method of Car Driver's Phone Call Recognition Based on Human Joint Points

Bowen Shi
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Wenlong Wang
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Xiaoyuan Wang
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Abstract: The distracted driving state of the car driver has caused serious harm to traffic safety and the personal safety of the driver. The most frequent and common distraction is the distracted driving behavior of the driver on the phone. In order to accurately identify the car driver's phone call is distracting driving behavior. A driver's phone call behavior recognition method based on the joint point information of the driver's body was proposed in this manuscript. After using the OpenPose network structure to extract the joint points of the human body, the driver's angle of the driver's upper and lower arms and the normalized distance between the joint points are calculated and compared to relatively accurate identification of car drivers' phone calls and distracted driving behaviors The experimental results show that the method proposed in this paper has an accuracy rate of 97.23% for the distracted driving state of the car driver's phone call.

Keywords: computer vision; deep learning; OpenPose; phone call recognition; biomechanical distraction

1.Introduction

With the rapid development of the social economy and the improvement of people's living standards, the number of motor vehicles in our country has increased year by year, and the traffic demand has increased rapidly. At the same time, traffic has brought traffic accidents and traffic casualties to people. According to the "China Statistical Yearbook", in 2020, a total of 244,674 traffic accidents occurred nationwide, resulting in 61,703 deaths, 250,723 injuries, and direct property losses of 1,313.61 million yuan.

Behind such a large number of traffic accidents, the most common reason is the driver's distracted driving. A driver's use of a mobile phone while driving a car is 4 times more likely to occur in a car accident. According to statistics, in 2019, there were a total of 33,244 traffic accidents in the United States, of which 2,895 were caused by distracted driving, and 424,000 people were injured. Therefore, how effectively and quickly identify the car driver using a mobile phone to make a phone call is of great significance for effectively and quickly identifying distracted driving of the car driver.

At present, the methods for the detection of car drivers' calling behavior are mainly divided into two categories: one is the detection method based on the mobile phone signal, and the other is the detection method based on machine vision. Ascariz[1] proposed a method to detect driver use of mobile phones through two antennae located inside the vehicle. And Jie Yang[2] also through the signal to do that. Yusuf[3] develops an algorithm to detect driver use of mobile phones. Vamsi[4] detected driver use of mobile phones through distinguishing the position of the camera of driver's phone. Keshav[5] presented a vision based method to detect driver use of mobile phones. XuBeilei[6] proposed a machine-learning-based method for detecting driver cell phone usage. Rafael[7] proposed a vision system to recognize the use of phone. Hoang Ngan[8] presented an advanced deep learning based approach to automatically determine whether a driver is using a cell-phone while ChaoYan[9] presented a novel system which learn and predict pre-defined drivers postures. Shantanu V[10] detected driver use of mobile phones through ECG signal processing aspect. Bensu[11] proposed a novel automated technique towards driver's phone usage violation detection using deep learning algorithms. Youssra[12] proposed an integrated system to automatically detect, track, and report distracted drivers. Xiong[13] proposed a driver's cell phone usage detection algorithm based on deep learning. Kubilay[14] detected mobile phone usage based on YOLOv5 network. Liu[15] proposed a detection algorithm based on extreme gradient boosting (XGBoost) to recognize the use of phone. Benjamin[16] proposes a method to extend such systems by driver posture classification to detect driver cell phone usage.

The detection method based on the mobile phone signal is to determine whether the driver is on the phone by detecting the mobile phone signal. This method has a high false positive rate. The use of mobile phones by people other than the driver in the driving vehicle cannot be distinguished from the use of mobile phones by the driver, and the mobile phone of the car driver is being used but does not affect the normal driving of the car driver. Circumstances have had too much influence on accurately identifying the use of cell phones by car drivers. The second type of method, the detection method based on machine vision, is to monitor the behavior of answering the phone in real-time through visual images. The method is based on the algorithm of a convolutional neural network to detect and recognize the behavior of car drivers on the phone in the face candidate area. This method also has the disadvantage of low robustness.

Based on the shortcomings of the above two methods, this paper proposes a method of calling behavior detection based on the joint points of the car driver's body. In this paper, the body joint points of the car driver are extracted first, and based on the joint points, the car driver's phone call behavior is detected and recognized by the unique behavioral characteristics of the car driver's phone call behavior.

2.Method

2.1Network Architecture

OpenPose[17] is based on a convolutional neural network, improves it, and then realizes the network model of real-time multi-person human joint point detection in a supervised learning environment. OpenPose is widely used. The main architecture of its original network is shown in Figure 1. The input of the network model is an image, and the feature map is obtained through the feature extraction of the OpenPose backbone network VGG19[18]. As the input of the next stage, the network model can be divided into two branches. The upper branch S of the network is used to predict the driver of the car. Part Confidence Map (PCM) of the joint points of the human body, each Confidence Map is a grayscale image, and the position coordinate of the maximum value is the highest probability corresponding to a joint point of the human body; the lower branch L of the network Part Affinity Field (PAF) is used to predict the joint points of the car driver. This branch is a 2D vector field for predicting Part Affinity, which represents the degree of association between two joint points, that is, the car driver. Position information and orientation information between two joint points.

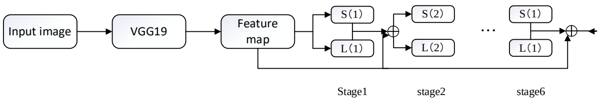

Figure 1. Main Network structure of OpenPose

The upper and lower branch structure of the OpenPose network structure is shown in Figure 2. The network uses a multi-stage network to extract the key information of the driver's human body joints. As shown in the figure, the first stage uses a 3×3 convolution kernel, but after the second stage, a 7×7 convolution kernel that can obtain a large receptive field is adopted, which reduces the computation while retaining the receptive field. The amount of operation calculation has been reduced from 97 to 51. And at the end of each stage, the predicted values from the sub-networks of the upper and lower branches and the initial feature map are connected as the input of the next stage, which can better integrate the deep feature information.

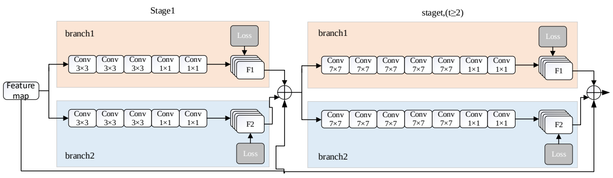

Figure 2. Internal structure of the prediction network

The human body joint point acquisition part of this paper directly uses the model trained by OpenPose. The joint point data model adopts the COCO data set format. As shown in Figure 3, the human body joint points are: 0 nose, 1 neck, 2 left-shoulder, 3 Left-elbow, 4 Left-wrist, 5 Right-shoulder, 6 Right-elbow, 7 Right-wrist, 8 Left-hip, 9 Left-knee, 10 Left-ankle, 11 Right-hip, 12 Right-knee, 13 Right-ankle, 14 Left-eye, 15 Right-eye, 16 left-ear, 17 right-ear.

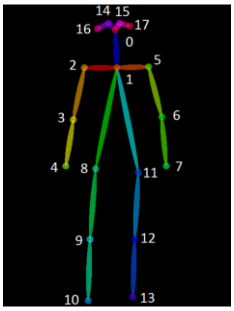

Figure 3. Human joint points in the COCO dataset

2.2 Proposed Model

In the previous section, the position information of the car driver's body nodes is extracted and saved through the OpenPose network. A method for identifying the car driver's phone calls based on body joints was proposed. Since most people are right-handed, this article only considers right-handed car drivers. A general conclusion can be drawn from observation: when a car driver is holding a mobile phone to make a call, he needs to hold the mobile phone with his right hand close to the right ear, so the angle between the right arm will become smaller, and the distance between the joint point of the right wrist and the joint point of the right shoulder will be reduced. will become smaller, and this article uses these two data to determine whether the driver is using his right hand to hold a mobile phone to make a call. The position coordinates of each body node of the driver are shown in formula (1):

$$c_i = [c_x^i, c_y^i], (i \in (0,17)) \qquad (1)$$

The formula for calculating the relative distance between each two body nodes is shown in formula (3):

$$d_{ij} = \frac{\sqrt{\left(c_x^j - c_x^i\right)^2 + \left(c_y^j - c_y^i\right)^2}}{h} \qquad (2)$$

Among them, d_{ij} represents the normalized relative distance between the joint point i and the joint point j, and h represents the relative distance between the left ear and the left eye.

The formula for calculating the included angle α_3 of the right elbow is shown in formula (3):

$$\alpha_3 = \arccos\frac{(c_x^2 - c_x^3)*(c_x^4 - c_x^3) + (c_y^2 - c_y^3)*(c_y^4 - c_y^3)}{\sqrt{(c_x^2 - c_x^3)^2 + (c_y^2 - c_y^3)^2} + \sqrt{(c_x^4 - c_x^3)^2 + (c_y^4 - c_y^3)^2}} \qquad (3)$$

3.Experiment

To verify the effectiveness of the method in this paper, experiments were carried out on a computer with Intel(R) Core(TM) i7-10870H CPU, 16GiB RAM, and CUDA11.1. The "distracted-driver-detection" dataset of the KAGGLE State Farm Distracted Driving Challenge is part of the right-handed mobile phone calling dataset, which uses part of the right-handed mobile phone calling dataset in the KAGGLE State Farm Distracted Driving Challenge dataset to train the model. Parts are shown in Figure 4. The test set consists of 500 images randomly selected from the videos captured by the driver in the real vehicle experiments.

Figure 4. Part of the driver's body nodes in the training dataset

Model detection performance is evaluated based on Precision (P), Recall (R), mean Average Precision (mAP), and FPS. FPS is used to measure the detection efficiency, which represents the number of images that the model can process per second. The mAP value is defined as the mean value of the average precision (Average Precision, AP) of each class, and the AP value corresponds to the area under a certain type of P-R curve. The calculation formulas are shown as follows.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

$$AP = \int_0^1 PdR \tag{6}$$

Among them, TP is the true class, indicating that the correct prediction is correct, FP is the false positive class, indicating that the wrong prediction is correct, and FN is the false negative class, indicating that the correct prediction is wrong.

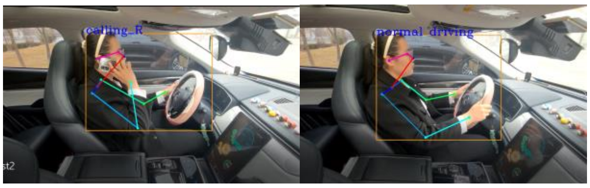

Figure 5. Part of the results

In the table below, the sample numbers and respective detection accuracies of normal driving state and phone call behavior are listed.

Table 1. Model recognition accuracy table

Sample type	Sample number	AP/%
Calling	9800	97.23
Normal driving	6900	97.89

4.Conclusion

In this paper, a body node-based identification method for a driver's phone-calling driving behavior is proposed. At this stage, most of the recognition methods for the driver's calling and driving behavior use the image recognition method, and the accuracy gap between different data sets is too large. In this paper, the OpenPose network is used to extract the body joint points of the car driver. Based on the body joint points, the angle between the driver's forearm and the forearm and the relative distance between the joint points are calculated, and the distracted driving behavior of the driver's phone call is finally determined. After verification, the method proposed in this paper has an accuracy rate of 97.23% for the car driver's answering the phone.

5.Acknowledgments

This study was funded by the Qingdao Top Talent Program of Entrepreneurship and Innovation (Grant No.19-3-2-11-zhc), the Natural Science Foundation of Shandong Province (Grant No. ZR2020MF082), the Foundation of Shandong Intelligent Green Manufacturing Technology and Equipment Collaborative Innovation Center (Grant No. IGSD-2020-012), and the National Key Research and Development Project (Grant No.2018YFB1601500).

6.References

[1] RODRIGUEZ-ASCARIZ J M, BOQUETE L, CANTOS J, et al. Automatic system for detecting driver use of mobile phones [J]. Transportation Research Part C-Emerging Technologies, 2011, 19(4): 673-81.

[2] YANG J, SIDHOM S, CHANDRASEKARAN G, et al. Detecting driver phone use leveraging car speakers; proceedings of the 17th Annual International Conference on Mobile Computing and Networking, MobiCom'11 and Co-Located Workshops, September 19, 2011 - September 23, 2011, Las Vegas, NV, United states, F, 2011 [C]. Association for Computing Machinery.

[3] ARTAN Y, BULAN O, LOCE R P, et al. Driver cell phone usage detection from HOV/HOT NIR images; proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPRW 2014, June 23, 2014 - June 28, 2014, Columbus, OH, United states, F, 2014 [C]. IEEE Computer Society.

[4] PARUCHURI V, KUMAR A. Detecting driver distraction using smartphones; proceedings of the 29th IEEE International Conference on Advanced Information Networking and Applications, AINA 2015, March 25, 2015 - March 27, 2015, Gwangju, Korea, Republic of, F, 2015 [C]. Institute of Electrical and Electronics Engineers Inc.

[5] SESHADRI K, JUEFEI-XU F, PAL D K, et al. Driver cell phone usage detection on Strategic Highway Research Program (SHRP2) face view videos; proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPRW 2015, June 7, 2015 - June 12, 2015, Boston, MA, United states, F, 2015 [C]. IEEE Computer Society.

[6] XU B, LOCE R P. A machine learning approach for detecting driver cell phone usage; proceedings of the Video Surveillance and Transportation Imaging Applications 2015, February 10, 2015 - February 12, 2015, San Francisco, CA, United states, F, 2015 [C]. SPIE.

[7] BERRI R, OSORIO F, PARPINELLI R, et al. A hybrid vision system for detecting use of mobile phones while driving; proceedings of the 2016 International Joint Conference on Neural Networks, IJCNN 2016, July 24, 2016 - July 29, 2016, Vancouver, BC, Canada, F, 2016 [C]. Institute of Electrical and Electronics Engineers Inc.

[8] LE T H N, ZHENG Y, ZHU C, et al. Multiple Scale Faster-RCNN Approach to Driver's Cell-Phone Usage and Hands on Steering Wheel Detection; proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPRW 2016, June 26, 2016 - July 1, 2016, Las Vegas, NV, United states, F, 2016 [C]. IEEE Computer Society.

[9] YAN C, COENEN F, ZHANG B. Driving posture recognition by convolutional neural networks [J]. IET Computer Vision, 2016, 10(2): 103-14.

[10] DESHMUKH S V, DEHZANGI O. ECG-Based Driver Distraction Identification Using Wavelet Packet Transform and Discriminative Kernel-Based Features; proceedings of the 2017 IEEE International Conference on Smart Computing, SMARTCOMP 2017, May 29, 2017 - May 31, 2017, Hong Kong, China, F, 2017 [C]. Institute of Electrical and Electronics Engineers Inc.

[11] ALKAN B, BALCI B, ELIHOS A, et al. Driver cell phone usage violation detection using license plate recognition camera images; proceedings of the 5th International Conference on Vehicle Technology and Intelligent Transport Systems, VEHITS 2019, May 3, 2019 - May 5, 2019, Heraklion, Crete, Greece, F, 2019 [C]. SciTePress.

[12] ELQATTAN Y, MOUSTAFA M N, EL-SHAFEY M H. System for detecting and reporting cell phone distracted drivers; proceedings of the 11th International Symposium on Image and Signal Processing and Analysis, ISPA 2019, September 23, 2019 - September 25, 2019, Dubrovnik, Croatia, F, 2019 [C]. IEEE Computer Society.

[13] XIONG Q, LIN J, YUE W, et al. A deep learning approach to driver distraction detection of using mobile phone; proceedings of the 2019 IEEE Vehicle Power and Propulsion Conference, VPPC 2019, October 14, 2019 - October 17, 2019, Hanoi, Viet nam, F, 2019 [C]. Institute of Electrical and Electronics Engineers Inc.

[14] ATAS K, VURAL R A. Detection of Driver Distraction using YOLOv5 Network; proceedings of the 2nd Global Conference for Advancement in Technology, GCAT 2021, October 1, 2021 - October 3, 2021, Bangalore, India, F, 2021 [C]. Institute of Electrical and Electronics Engineers Inc.

[15] LIU T, ZHOU Z, CHAI C, et al. Non-instinct detection of cellphone usage from lane-keeping performance based on eXtreme gradient boosting and optimal sliding windows [J]. 2022.

[16] WAGNER B, TAFFNER F, KARACA S, et al. Vision Based Detection of Driver Cell Phone Usage and Food Consumption [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(5): 4257-66.

[17] CAO Z, HIDALGO G, SIMON T, et al. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields [J]. Ieee Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(1): 172-86.

[18] Simonyan K ， Zisserman A . Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. arXiv e-prints, 2014.

[19] Kaggle. State farm distracted driver detection[EB/OL]. （2020-11-02） [2021-12-17]. https： //www.kaggle.Com/rightway11/state-farm-distracted-driver-detection.