International Journal of Internet, Broadcasting and Communication Vol.15 No.4 142-148 (2023)
http://dx.doi.org/10.7236/IJIBC.2023.15.4.142

IJIBC 23-4-16

Learning Model for Avoiding Drowsy Driving with MoveNet and Dense Neural Network

Jinmo Yang[1], Janghwan Kim[2], R. Young Chul Kim[3], Kidu Kim[4*]

[1]*Dept. of Physics, Korea University, Republic of Korea*
[2,3]*Dept. of Software and Communication Engineering, Hongik University, Republic of Korea*
[4*]*Telecommunications Technology Association, Republic of Korea*
[1]*yjmd2222@gmail.com,* [2]*janghwan.kim@g.hongik.ac.kr,* [3]*bob@hongik.ac.kr,* [4*]*kdkim@tta.or.kr*

Abstract

In Modern days, Self-driving for modern people is an absolute necessity for transportation and many other reasons. Additionally, after the outbreak of COVID-19, driving by oneself is preferred over other means of transportation for the prevention of infection. However, due to the constant exposure to stressful situations and chronic fatigue one experiences from the work or the traffic to and from it, modern drivers often drive under drowsiness which can lead to serious accidents and fatality. To address this problem, we propose a drowsy driving prevention learning model which detects a driver's state of drowsiness. Furthermore, a method to sound a warning message after drowsiness detection is also presented. This is to use MoveNet to quickly and accurately extract the keypoints of the body of the driver and Dense Neural Network(DNN) to train on real-time driving behaviors, which then immediately warns if an abnormal drowsy posture is detected. With this method, we expect reduction in traffic accident and enhancement in overall traffic safety.

Keywords: *artificial intelligence, drowsiness, posture, single-image detection*

1. INTRODUCTION

Sleep deprivation in modern people leads to drowsy driving and a decrease in body Movement and cognitive abilities [1]. A quick response to traffic emergency is crucial in driving. However, drowsiness hinders the driver's responsiveness to traffic emergencies. As the driver's response time decreases, the risk of accidents increases [2]. From 2019 to 2021, between March and May, there was the number of traffic accidents caused by drowsy driving reached a total of 1,833, and the fatality rate was 2.6%, which is greater than the overall traffic accident fatality rate of 1.4% [3]. To solve this issue, this study proposes a drowsy driving prevention learning model using artificial intelligence techniques to classify drowsy driving and sound a warning message similar to a horn through a sounding device(speaker) in the drowsiness warning application. This is to use MoveNet and Dense Neural Network classification model to develop a safe-driving application for detecting the driver's state and, if the driver is predicted to be in a drowsy state, generate the warning message to wake

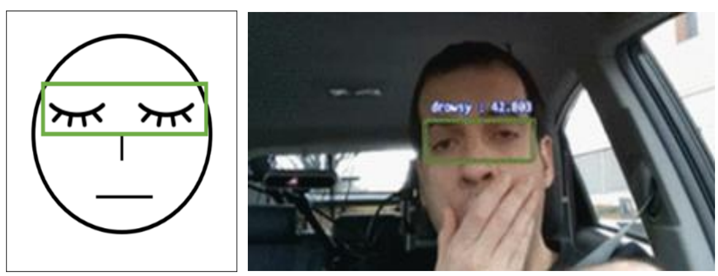
Manuscript Received: September. 14, 2023 / Revised: September. 23, 2023 / Accepted: September. 26, 2023
Corresponding Author: kdkim@tta.or.kr
Tel:*** - **** - ****   Fax: +82-31-5182-9055
Team Manager, AI Digital Infrastructure Team at Telecommunications Technology Association, Korea

the driver. Section 2 discusses relevant research on existing methods for detecting drowsy driving and studies related to preventing drowsy driving. Section 3 presents the posture classification mechanism by use of artificial intelligence and its implementation in the drowsy driving warning application. Section 4 concludes the study and mentions future works.

2. RELATED STUDIES

2.1 Drowsy driving detection through single-image eyes detection

Drowsy driving detection through the single-image eye detection method uses an object detection model, YOLOv3, to train from and detect drowsy eyes to predict drowsy driving and sound an alarm [4]. Of many studies that detect the eyes to predict drowsy driving, this model uses a pre-trained model. A pre-trained computer-vision model is a model that has been trained in advance with large image datasets and recognizes patterns such as the eyes [5]. LabelImg is used to generate the label datasets from the image datasets. Then these preprocessed datasets are used to train the YOLOv3 model [6].



(a) Bounding box (b) Detection of drowsy eyes

Figure 1. Drowsy eyes annotation and model detection

Figure 1 shows the schematic drawing of LabelImg Tool annotation and the detection from a real image. **(a)** The detection region on the schematic is shown. **(b)** The detected object's bounding box and label(identification) are drawn on the real image. There may be an issue where the model incorrectly detects the closed eyes of an awake driver as drowsy driving and sounds the alarm in the awakened state.

2.2 Drowsy driving detection through consecutive-image head posture and eye detection

Drowsy driving detection through consecutive-image head posture and eyes detection method uses facial features to determine head posture and eye blinking to detect drowsy driving [7]. This method uses frames extracted from a video recording of a driver and detects the head and the eyes with the Viola-Jones face detection algorithm [8]. The two steps that follow can be divided into classifying the driving state through the eyes and classifying through the head posture. For eye classification, the Wavelet classification model is trained to classify the eyes as closed or open [9]. If the model detects closed eyes for a certain period, the driver is considered to be in a drowsy state by the eyes.

Now, the eye and the head detection can be combined, and the head posture can be determined to boost the prediction in addition to the drowsiness detection from the eye classification failure. This study uses the fact that head posture changes from the normal position if the driver is in a drowsy state to improve the conventional method of drowsiness detection system that only tracks the eyes. However, the study does not consider the chance that the first frame of head tracking may not be in the normal position. If the head postures of all frames are not in the normal position, left inclination for example, the positions in the first and the succeeding frames are the same, and the state of the driver will not be predicted as the drowsy state.

3. DRIVING POSTURE DETECTION MECHANISM

For safe driving, the drowsy driving prevention learning model analyzes the driver's state and classifies the posture into normal driving posture and drowsy driving posture. Regardless of the closure of the eyes, the difference in the angle between the upper body parts and the lower body parts is a major factor in classifying the posture. The act of closing the eyes applies to both normal driving and drowsy driving. However, there is a significant difference between keeping the head up in a normal driving posture and bowing the head in a drowsy driving posture. Therefore, the accuracy of detecting the posture from a single image will be higher than that of detecting the eye closure from a single image. This performance may be expected to be on par with tracking the eyes for a certain period, which must frequently predict all frames in a video clip.

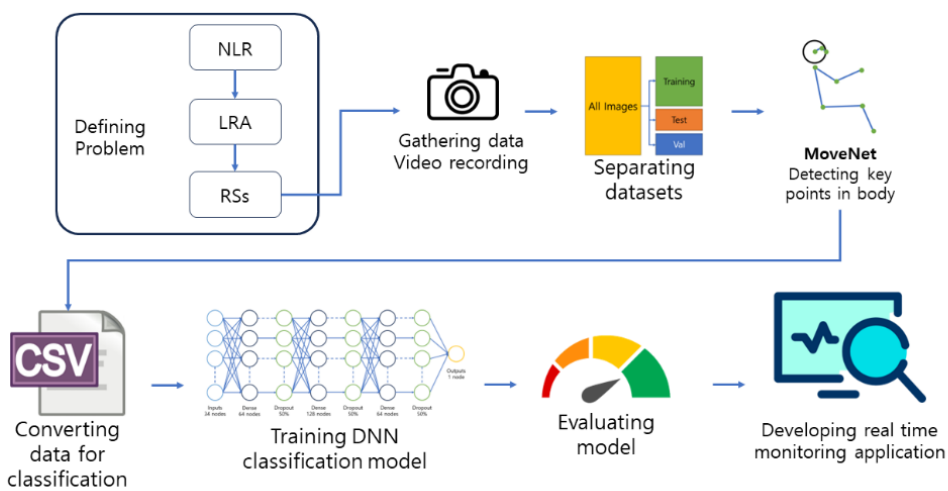


Figure 2. The process of developing a drowsiness warning application

Figure 2 shows the process of developing a drowsiness warning application. Firstly, from natural language requirement (NLR) to linguistic requirement analysis (LRA), the requirements to develop the application are extracted. NLRs present distinctive requirements, but requirement specifications (RSs) contribute to reducing the development time of reaching the application's purpose. By following these RSs, normal driving data and drowsy driving data are collected. The method of extracting body information from the collected data with MoveNet and using these features in the DNN classification model to classify the driving posture is constructed. For this method, video clips of a driver's normal driving and drowsy driving are recorded. Image frames are extracted from the video clips and split into train, validation, and test datasets. With MoveNet, 17 key points in the body are extracted from these datasets and analyzed. To use the extracted key points in the DNN classification model, the key points are converted into respective train, validation, and test CSV files. The data in the train CSV file is used to train the DNN classification model. The data in validation and test CSV files is used to evaluate the performance of this model. This method of classifying driving posture is applied to the drowsiness warning application and the performance is tested.

3.1 Gathering and Preprocessing the Data

The frontal camera of the iPhone 6 is used to gather data. For a duration of 7 minutes, normal driving and drowsy driving posture are captured in a video clip in the size of 720 by 1280. Here, both normal driving and drowsy driving postures are simulated, including handling the steering wheel, operating the turn signal, and head movement. Frames are extracted from the recorded video clip every 100 milliseconds, and labels corresponding to the postures are assigned.

The extracted frame counts are 2,964 for normal driving posture and 1,501 for drowsy driving posture, at the imbalanced ratio of approximately 2:1. Class imbalance hurts training, the most notable tendency being learning in a class of small size [10]. In this study, where a drowsy driving posture must be classified, such likelihood of not being able to classify this posture must be minimized. For creating a balanced dataset, the larger, normal driving data of 2,964 frames are down-sampled to match the smaller, drowsy driving data of 1,501 frames. Also, for training and evaluating performance, the data is divided into train, validation, and test datasets.

3-2. Extracting Body Key Point Location Data using MoveNet

The Drowsiness warning application uses MoveNet and Dense Neural Network classification artificial intelligence model. MoveNet is a model developed by the members of Google Research to analyze images. From a given image, 17 key points of the human body in the image are recognized, and from the horizontal and vertical parts of the points, a total of 34 numbers is obtained [11]. This model has two variations: the faster Lightning model and the more accurate Thunder model, which is still considered fast. Since this study aims to predict drowsiness at a longer interval, the more accurate Thunder model is chosen.

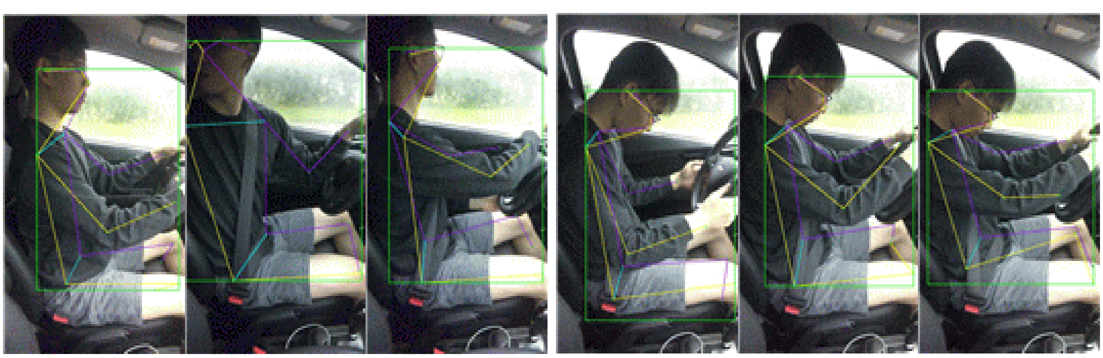


Figure 3. Analyzing of the driving posture using MoveNet

Figure 3 shows the drawing of the extracted body key points from the original frames using MoveNet. 17 key points, such as the eyes, nose, and ears, are plotted and connected to form the skeletons. The train, validation, and test datasets are processed with MoveNet, and the resulting key points are constructed into the corresponding train, validation, and test CSV to be used in the DNN classification model presented in the next subsection. To analyze the angle of the neck in normal driving posture and drowsy driving posture, the angle from the nose, shoulders, and hips, which will be referred to as the NSH angle, is calculated.

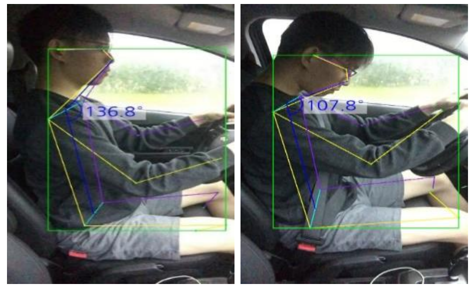


Figure 4. The NSH angle

Figure 4 shows the plotting of the NSH angle along with the body key points on the frames. Then the statistics of the NSH angles calculated from the train data are shown in **Table 1**.

Table 1. The statistics of the NSH angles from the training dataset

Posture	Avg.	Std.	Min.	25%	50%	75%	Max.
Normal driving	142.5	9.64	120.6	137.8	140.5	144.2	195.3
Drowsy driving	87.8	10.2	66.6	81.3	86.7	92.7	119.3

Table 1 shows the statistics of the NSH angles from the normal and drowsy driving posture in the training dataset. The angle of normal driving ranges from a minimum of 120.6° to a maximum of 195.3°, with an average of 142.5°, and the angle of normal driving ranges from 66.6° to 119.3°, with an average of 87.8°. It can be seen that the postures are divided into normal and drowsy driving postures from about 120°.

3-3. Training and evaluating the DNN Classification Model

The DNN Classification model is a deep learning model composed of dense layers. From each dense layer, all neurons are connected to the next layer, receiving a 1-dimensional vector as the input and passing it through the weight to the output [12]. The 34 horizontal and vertical numbers of body key points from the generated CSV file are flattened into a 1-dimensional row and entered at the input of the DNN classification model. The structure of this model is as follows: three dense layers with the activation function of ReLU [13], each layer followed by the dropout layer at a rate of 0.5 to avoid overfitting [14]. The Adam optimizer was used [15].

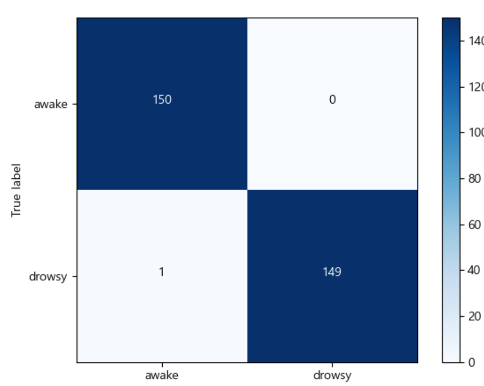


Figure 5. Confusion matrix from the test dataset

Figure 5 shows the plot of the test data confusion matrix from the best weights of the 48th epoch. For the total of 300 data points, the accuracy and the precision of the model were 0.997 and 1.0, exceeding 0.89

accuracy of [4].

3.4 Drowsiness Warning Application

The operation of the drowsy driving warning application, developed by integrating the frontal camera of the iPhone 6 with MoveNet and a DNN classification model, functions as follows. At every 5-second interval, MoveNet extracts essential body key point data from images captured by the iPhone camera. Subsequently, a DNN classification model identifies the posture based on this data. If the identified posture corresponds to a drowsy driving position, an audible warning message is sounded to wake the driver.

To facilitate real-time classification, scenarios depicting both normal and drowsy driving were simulated. The model achieved an average accuracy of 0.963.

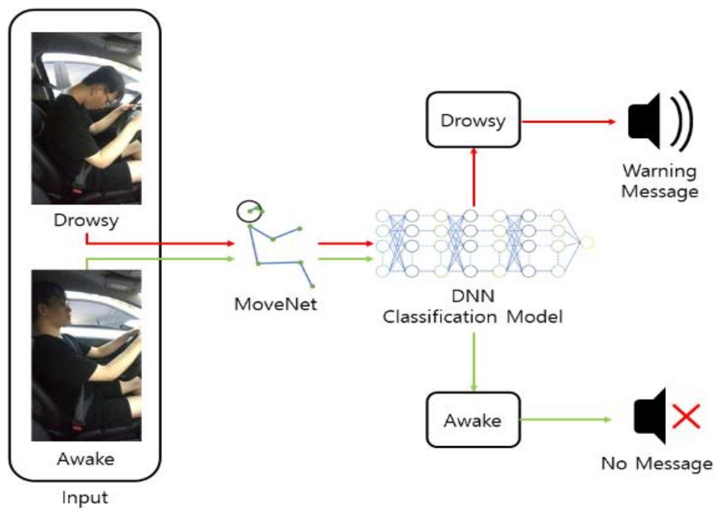


Figure 6. The process of the drowsiness warning application

Figure 6 illustrates the process where the drowsiness warning application classifies normal driving and drowsy driving postures and outputs a warning message for drowsy driving postures. Images containing the driver's posture are entered through the MoveNet and the DNN classification model, resulting in classification as either a drowsy driving or a normal driving posture. In the case of drowsy driving, a warning message is sounded through the output device, while in the case of normal driving, no message is sounded.

4. CONCLUSION

This study uses artificial intelligence techniques to detect the driver's driving state as normal driving or drowsy driving. MoveNet and DNN classification models are used to identify the driver and classify his or her posture, resulting in an accuracy of 0.997 and a precision of 1.0. This exceeds the performance of detecting drowsy driving from eye closure, both from single-image detection and consecutive-image detection. The application developed from this classification process sounds like a warning message to wake the driver if the driver's posture is classified as drowsy driving posture, thereby preventing traffic accidents. However, the main body point information data used in training is gathered from a single individual. The dropout method was used to prevent the model from overfitting to this specific person; nevertheless, the performance may still be limited if other drivers use this application. Furthermore, the training data is extracted from the video clips that are recorded during the day which may affect the application's performance at night times. For future work, we look forward to collecting data from many drivers under different illumination and training the model to improve its generalization performance.

References

[1] Kwak, Seung-Hyun and Min, Byung-Chan, “Effect of sleepiness due to complete sleep deprivation on driving performance,” *Asia-pacific Journal of Multimedia Services Convergent with Art*, Vol. 9, No. 1, pp. 615-623, 2019. DOI: 10.35873/ajmahs.2019.9.1.062

[2] Philip, P. and Akerstedt, T., “Transport and industrial safety, how are they affected by sleepiness and sleep restriction?,” *Sleep medicine reviews*, Vol. 10, No. 5, pp. 347-356, 2006. DOI: 10.1016/j.smrv.2006.04.002

[3] Korea ROAD Traffic Authority, Constantly Drooping Eyelids Due to Allergies... Be Cautious of Drowsy Driving in the Spring Season. https://www.koroad.or.kr/synap/skin/doc.html?fn=BD_202303270808376881.pdf&rs=/upfilePath/koroadNew/synap/result/.

[4] Macalisang, J. R., Alon, A. S., Jardiniano, M. F., Evangelista, D. C. P., Castro, J. C., and Tria, M. L., “Drive-Awake: A YOLOv3 Machine Vision Inference Approach of Eyes Closure for Drowsy Driving Detection,” *2021 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*, pp. 1-5, 2021. DOI: 10.1109/IICAIET51634.2021.9573811

[5] Ikechukwu, V. A., Murali, S., Deepu, R., and Shivamurthy, R. C., “ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images,” *Global Transitions Proceedings*, Vol. 2 No. 2, pp. 375-381, 2021. DOI: 10.1016/j.gltp.2021.08.027

[6] Tzutalin, LabelImg, https://github.com/HumanSignal/labelImg.

[7] Teyeb, I., Jemai, O., Zaied, M., and Ben Amar, C., “A novel approach for drowsy driver detection using head posture estimation and eyes recognition system based on wavelet network,” *The 5th International Conference on Information, Intelligence, Systems and Applications*, pp. 379-384, 2014. DOI: 10.1109/IISA.2014.6878809.

[8] Viola, P., and Jones, M. J., “Robust real-time face detection”, *International Journal of Computer Vision*, Vol. 57, pp. 137–154, 2004. DOI: 10.1023/B:VISI.0000013087.49260.fb

[9] Amar, C. B. and Jemai, O., “Wavelet networks approach for image compression,” *International Journal on Graphics, Vision and Image Processing*, pp. 37-45, 2007.

[10] Kumar, P., Bhatnagar, R., Gaur, K., and Bhatnagar, A., “Classification of imbalanced data: review of methods and applications,” *IOP conference series, materials science and engineering*, Vol. 1099, No. 1, 2021. DOI 10.1088/1757-899X/1099/1/012077

[11] TensorFlow, Next-Generation Pose Detection with Movenet and Tensorflow.Js, https://blog.tensorflow.org/2021/05/next-generation-pose-detection-with-movenet-and-tensorflowjs.html.

[12] Wang, C., Yao, C., Zhao, S., Zhao, S., and Li, Y., “A Comparative Study of a Fully-Connected Artificial Neural Network and a Convolutional Neural Network in Predicting Bridge Maintenance Costs,” *Appl. Sci*, Vol. 12, 2022.

[13] LeCun, Y., Bengio, Y., and Hinton, G., “Deep Learning,” *Nature*, Vol. 521, pp. 436-444, 2015. DOI: 10.1038/nature14539.

[14] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R., “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *Journal of Machine Learning Research*, Vol. 15, No. 56, pp. 1929–1958, 2014. DOI: 10.5555/2627435.2670313

[15] Kingma, D. P., and Ba, J., “Adam: A Method for Stochastic Optimization,” *Conference Paper at the 3rd International Conference for Learning Representations*, 2015. DOI: 10.48550/arXiv.1412.6980