Deep Learning Based Driver Smoking Behavior Detection for Driving Safety

Tzu-Chih Chien, Chieh-Chuan Lin, and Chih-Peng Fan
Department of Electrical Engineering, National Chung Hsing University, Taiwan
Email: {e29013281, tall20you}@yahoo.com.tw, cpfan@dragon.nchu.edu.tw

***Abstract*—According to the previous researches by experts, the smoking behavior in driving will cause three hazards: reducing vision, distracting, and irritating. Thus the risk is extremely high. By detecting whether there is any object of a cigarette, the smoking behavior in driving is recognized, and it will greatly help for driving safety. In this paper, the YOLOv2 deep-learning image based methodology is applied for driver's cigarette object detection. The driver's images are captured by a dual-mode visible light and near-infrared camera, and the developed system judges whether or not there is driver smoking behavior in the day and night conditions. By the YOLOv2 deep learning network, the pre-prepared images of driver smoking behavior are marked, and the cigarette detector is trained to detect the cigarette object when the driver is smoking. In experimental results, the applied deep learning based design performs that the precision is up to 97% and the recall is 98%. Besides, by the proposed design, the average accuracy of cigarette detection is up to 96% during the day condition, and that of cigarette detection is up to 85% during the night condition.**

***Index Terms*—smoking behavior detection, deep learning, YOLOv2, driving safety**

I. INTRODUCTION

Researches and analyses from the US bureau of safety experts reveals that the probability of car accidents for smokers is 1.5 times larger than that for non-smokers. The United Kingdom and Germany experts believe that the 5% of car accidents are related to smoking while driving. According to the US experts, the smoking behavior leads to three driving hazards: reduced vision, distracting, and odors stimulations. People need a high degree of attention and a clear vision while driving, and the smoking behavior may irritate eyes and respiratory tract, and then the vision will be blurred. The acrid smoke also causes coughing and distraction. Specially, the influence of smoke on the vision is very terrible. Based on statistics, when it is at dusk, smoking four cigarettes during a short time will reduce the vision by 20-30%. When a driver is in such circumstances, the risk is very high. Therefore, the development of cigarette detection technology is a very important issue when the driver smoking behavior happens.

In previous related works [1]-[10], the image-based and non image-based technologies were used to detect the smoking behaviors. In [1, 2, 3, 8, 9], the color information of cigarette object or smokes was analyzed, and the smoking condition or smoking behavior were detected effectively. Besides, in [4, 5, 6, 7, 10], the sensor-based (i.e. the non image-based) designs were developed to recognize the smoking behavior. To develop the image-based cigarette detector, in general, the well-known color based designs will be influenced by light disturbance and insufficient light conditions. In addition, the color based designs may not work well at the night condition. Fig. 1 depicts the image-based driver smoking behavior detection system.

In this paper, to overcome the light disturbance and insufficient light conditions in the car by image-based recognition methodologies, the YOLO-based deep learning design is developed to detect the cigarette object when there is the driver smoking condition. The rest of the paper is described as follows. In Section II, the previous color-based detection design is briefly reviewed. In Section III, the design methodology of the proposed YOLO-based cigarette object detector is discussed. The experimental results and comparison are shown in Section IV. Finally, a conclusion is given.

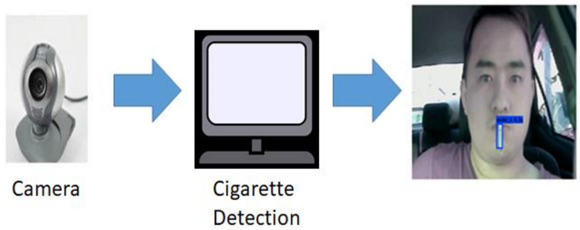

Figure 1. The image-based driver smoking behavior detection

II. PREVIOUS WORK

In [9], the developed color-based method used four major skills, which included the face detection, face boundary detection, mouth positioning, and cigarette detection. Driver facial videos are captured through the camera, and the design applied color information and separated the statistical threshold condition at different daytime. Firstly, the facial region detection was active based on facial features, and the possible facial region was found. Next, based on the symmetry and concentration properties of facial features, the facial boundaries were detected, and the face size was estimated.

Manuscript received December 3, 2019; revised March 6, 2020.

doi: 10.18178/joig.8.1.15-20

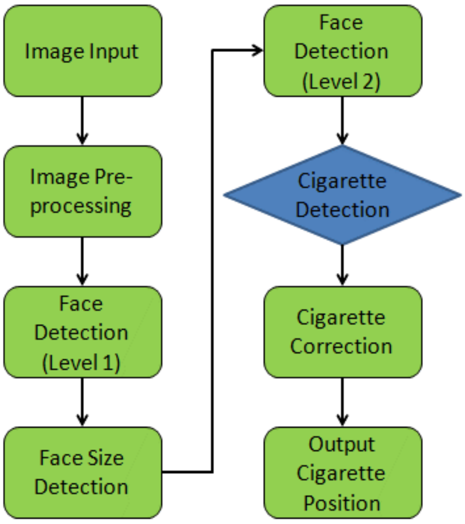

Figure 2. The previous cigarette detection flow in [9]

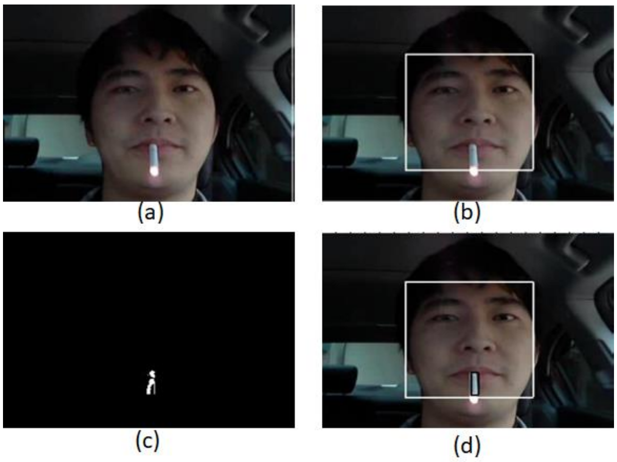

Figure 3. The processing and detection results in [9]

Finally, the system applied the luminance threshold condition to detect the cigarette near the mouth region for driver smoking behavior detection. In [9], the developed algorithm utilized the brightness information for the cigarette detection. The test video sequences were analyzed by statistics, and the threshold value was set to filter out a cigarette, and then the projection method was used to detect the location of the cigarette. However, the developed method in [9] cannot perform well at the night condition or the insufficient light condition in the car. Fig. 2 illustrates the previous color-based cigarette detection flow in [9], and Fig. 3 demonstrates the processing and detection results by the color-based method in [9], where Fig. 3(a) depicts the input image, Fig. 3(b) shows the face location result, Fig. 3(c) reveals the detection of cigarette object, and Fig. 3(d) shows the final result.

To overcome the night time, light disturbance, and insufficient light conditions in the car, the YOLO-based deep learning methodology is developed to detect the cigarette object when there is the driver smoking condition. Fig. 4 depicts the applied cigarette detection technology by the YOLO-based deep learning method.

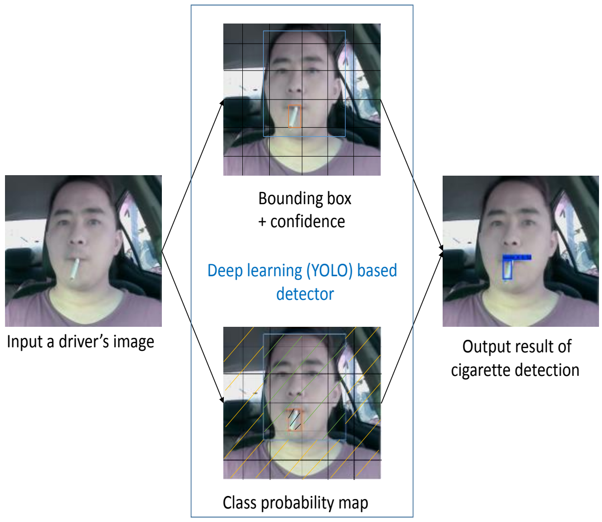

Figure 4. The cigarette detection by YOLO-based deep learning method

III. Proposed Deep-Learning Based Smoking Behavior Detection

A. *Preparing Image Data for Training and Testing Cigarette Object Detector*

To collect the useful image databases for training and testing the deep learning based inference model, for smoking drivers in car, 28 video sequences, which include a total of 5500 images, are captured for various driver smoking behaviors at the various day and night situations. By integrating the image databases as shown in Fig. 5 and Fig. 6, all of the proper images are selected to be labeled. The unsuitable images are removed, and they are not selected to be used for training the proposed deep learning based cigarette detector.

B. *YOLOv2 Based Cigarette Object Detection*

In [11], the deep-learning based YOLO algorithm is very effective for the application of object detections. The traditional object detection methodologies have three processing parts, which include localization of object, extraction of features, and classification of image object. Compared with previous object detection designs, the YOLO based algorithm can directly inference the bounding boxes and class probabilities of image objects by effectively evaluating a deep-learning neural network, and then the YOLO network greatly reduces the processing and computing time.

In several deep-learning based algorithms, the R-CNN (Region Convolutional Neural Networks) method is one of the well-known technologies, and the R-CNN network firstly uses several region proposals, which contain the possible objects, and each region through convolutional neural networks (CNN) will be classified by the network. Finally, by the regression process, the CNN network rectifies the positions of bounding boxes for possible objects. The most significant characteristic by the YOLO based interference is to achieve the object detection directly through end-to-end models, and the YOLO based network utilizes the whole image as input to the neural network for predicting the positions of bounding boxes directly. The bounding box involves the confidence score of image object and the category to which the image object belongs. The computational complexity of the YOLO based network is low, and it is suitable for the real-time applications.

Figure 5. The first set of eight video sequences for deep learning

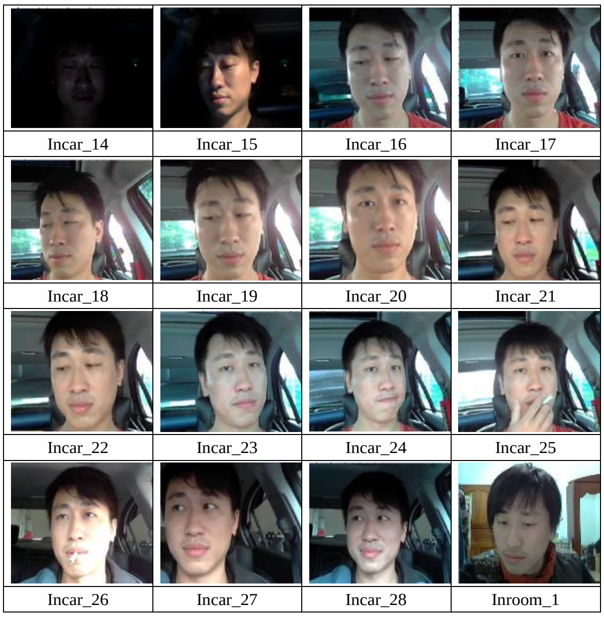

Figure 6. The second set of sixteen video sequences for deep learning

Compared with the YOLOv1 model, the YOLOv2 network in [11] is improved by: (1) batch normalization, (2) high resolution classifier, (3) convolution with anchor boxes, (4) dimension clusters, (5) direct location prediction, and (6) fine-grained features. In our design, the YOLOv2 network is used as a framework to design the proposed cigarette detector. The design process is divided into the training and testing phases. The training phase is decomposed into the following processes: (1) Labeling the cigarette object in images, (2) Architecture configuration of the YOLOv2 network, and (3) Training an inference model. In the testing phase, the trained YOLOv2 inference model is tested by testing images for functional evaluations. The applied process is described as follows:

At Stage 1 - Preparing the labeled cigarette images: Fig. 7 shows that the cigarette object in an image is labeled, and the corresponding bounding box is built by the semi-automatic tool, i.e. LabelImg [12].

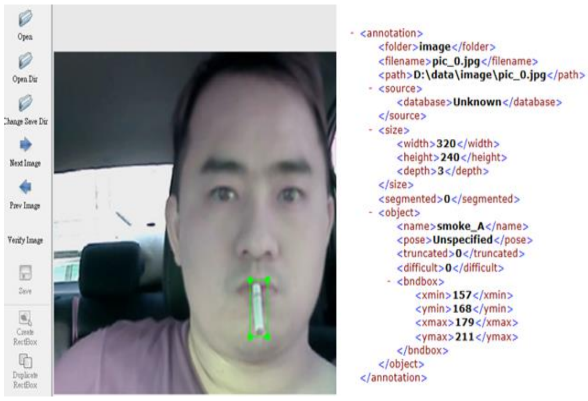

Figure 7. Labeling process for cigarette object

At Stage 2 - Training and testing an inference model by the YOLOv2 network: Fig. 8 demonstrates the training/testing flows to obtain the YOLOv2 inference model. The operational procedure consists of three parts, which include the data pre-processing, training for the inference model, and testing for the inference model. After the required convolution layer and normalizations are set, the weights of YOLOv2 network, which are pre-trained by the Darknet method, are loaded when the pre-processing process is active. Because the last layer of the convolution layer determines the final output, the convolution layer must be re-adjusted and trained to achieve a fine-tuned training model. After the inference model is trained, the testing procedure is followed. In the testing process, the highest final confidence score is computed, and then the detection accuracy and recall rate will be calculated and obtained.

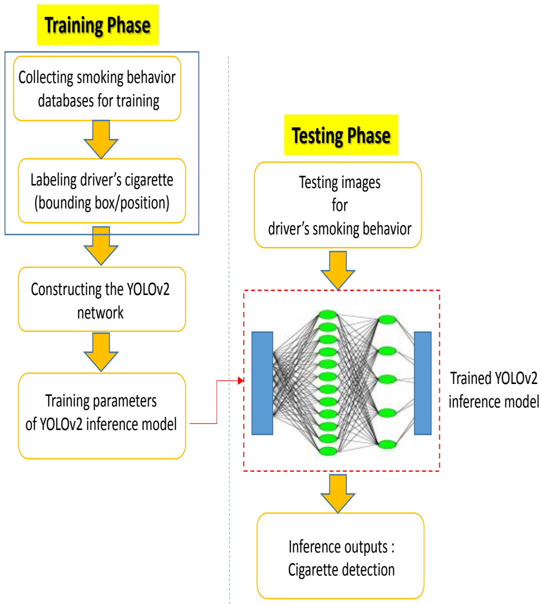

Figure 8. The training/testing flow to obtain the YOLOv2 inference model for cigarette object detection

IV. Experimental Results and Comparisons

In our experiments, a computing sever with two NVIDIA GPU acceleration cards is applied for training the inference model of cigarette detector, and a personal computer, which involves an Intel I7-5700 CPU at 2.7GHz operational frequency, is used for testing the inference model. To prepare the suitable databases, 28 video sequences of smoking drivers, which have a total of 5500 images (i.e. frames), are used to label cigarette objects for different driver smoking behaviors. In the 28 labeled video sequences, 4 video sequences, which includes 2 day-time videos and 2 night-time videos, are selected for testing. Fig. 9 illustrates the selected four driver smoking video sequences to test the inference model. By the experimental results, the proposed deep learning based method performs that the precision is up to 97% and the recall achieves 98%. Fig. 10(a) and 10(b) show the testing results at the day time and the night time, respectively.

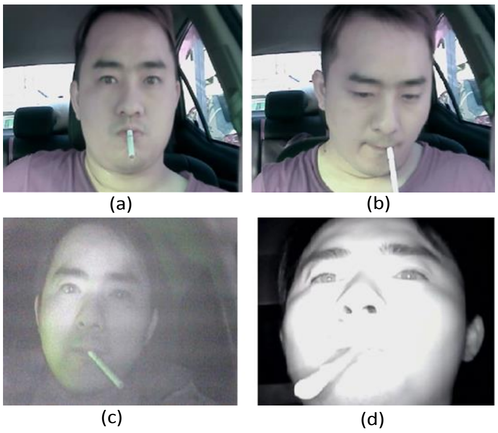

Figure 9. The four driver smoking video sequences for testing the inference model

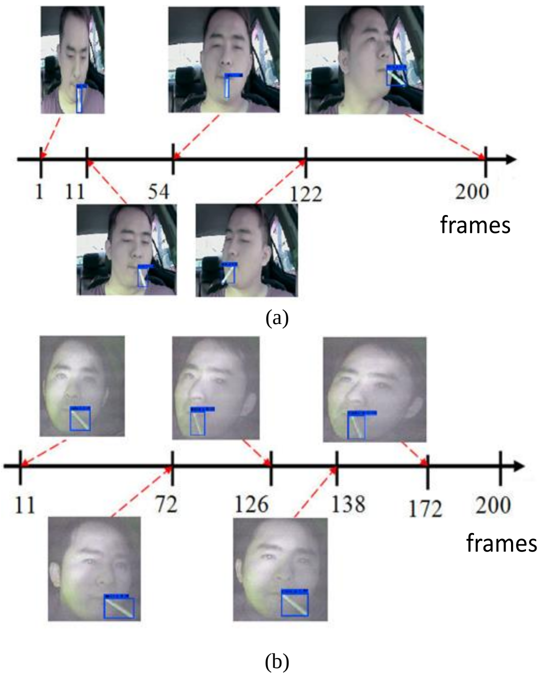

Figure 10. Testing results (a) at the day time (b) at the night time

Besides, when the individual condition is selected for comparison, Fig. 11 and Fig. 12 show the detection results by the proposed method in "Incar_4" and "Incar_5", respectively. Table I and II list that the proposed YOLO-based design has better accuracy rate and less false rate than the color-based detection in [9] at the day time. Similarly, Fig. 13 and Fig. 14 demonstrate the detection results by the proposed method in "Incar_14" and "Incar_15", respectively. Table III reveals that the proposed YOLO-based design also has better accuracy rate and less false rate than the color-based detection in [9] at the reflection-light condition. Specially, in Table IV, the proposed YOLO-based design is effective, and it can detect the cigarette object of driver well at the night condition. However, when the color-based design in [9] is applied for cigarette detection, the color-based method will be fail to detect any cigarette at the night condition. In general, the applied deep learning based method has the average accuracy of cigarette detection is up to 96% during the day condition, and that of cigarette detection is up to 85% during the night condition.

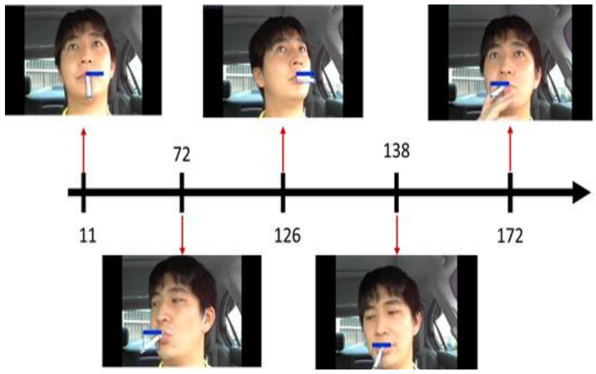

Figure 11. Results by the YOLO-based method with "Incar_4"

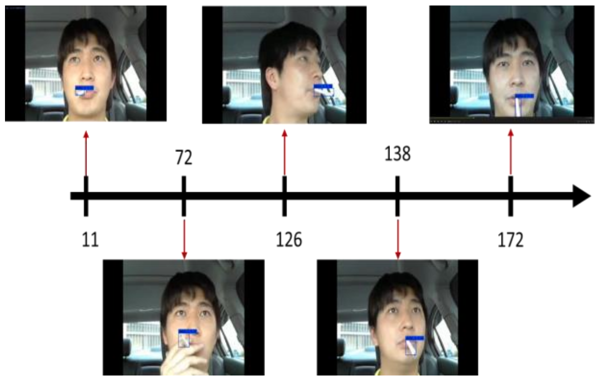

Figure 12. Results by the YOLO-based method with "Incar_5"

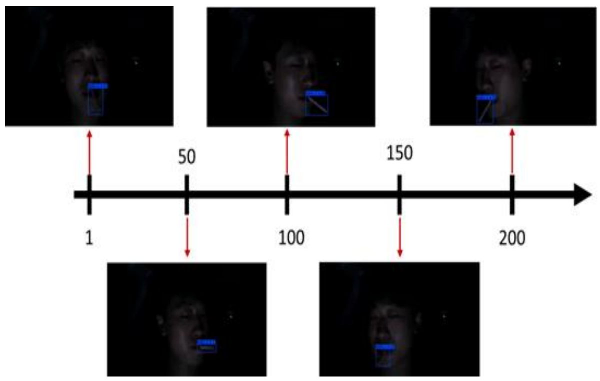

Figure 13. Results by the YOLO-based method with " Incar_14 "

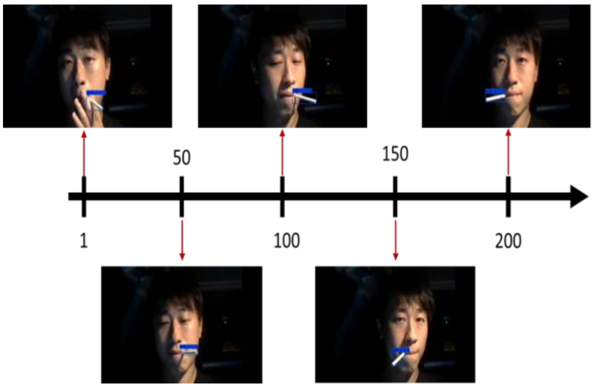

Figure 14. Results by the YOLO-based method with “Incar_15”

TABLE I. PERFORMANCE COMPARISON BETWEEN TWO DETECTION METHODS WITH “INCAR_4” VIDEO SEQUENCE

Number of images	Methods	Number of correct detection	Number of false detection	Accuracy rate	False rate
200	Lin [9] Color-based	147	53	73.5%	26.5%
	Proposed YOLO-based	195	5	97.5%	2.5%

TABLE II. PERFORMANCE COMPARISON BETWEEN TWO DETECTION METHODS WITH “INCAR_5” VIDEO SEQUENCE

Number of images	Methods	Number of correct detection	Number of false detection	Accuracy rate	False rate
200	Lin [9] Color-based	112	88	56.0%	44.0%
	Proposed YOLO-based	190	10	95.0%	5.0%

TABLE III. PERFORMANCE COMPARISON BETWEEN TWO DETECTION METHODS WITH “INCAR_15” VIDEO SEQUENCE

Number of images	Methods	Number of correct detection	Number of false detection	Accuracy rate	False rate
200	Lin [9] Color-based	165	35	82.5%	17.5%
	Proposed YOLO-based	190	10	95.0%	5.0%

TABLE IV. PERFORMANCE COMPARISON BETWEEN TWO DETECTION METHODS WITH “INCAR_14” VIDEO SEQUENCE

Number of images	Methods	Number of correct detection	Number of false detection	Accuracy rate	False rate
200	Lin [9] Color-based	X	X	X	X
	Proposed YOLO-based	170	30	85.0%	15.0%

V. CONCLUSION

In this paper, the YOLOv2 deep learning based detector is developed for the cigarette object detection when there is driver smoking behavior. The developed method recognizes the driver smoking behavior in the day and night conditions. By the YOLOv2 based deep learning network, the cigarette detector is trained to detect the cigarette object for driver’s smoking actions. In experiments, the applied deep learning based method performs that the precision is up to 97% and the recall is 98%. Besides, by the proposed method, the average accuracy of cigarette detection is up to 96% during the day condition, and that of cigarette detection is up to 85% during the night condition.

ACKNOWLEDGMENT

This work was supported by Ministry of Science and Technology, Taiwan, under Grant MOST 107-2218-E-005-019.

CONFLICT OF INTEREST

The authors declare no conflict of interest for this work.

AUTHOR CONTRIBUTIONS

Chieh-Chuan Lin started the research and collected the image dataset; Tzu-Chih Chien finished the deep-learning based simulation and analyzed the data; Chih-Peng Fan wrote the paper; all authors had approved the final version.

REFERENCES

[1] J. P. Stitt and L. T. Kozlowski, “A system for automatic quantification of cigarette smoking,” in *Proc. International Conference of the IEEE Engineering in Medicine and Biology Society*, 2006.

[2] P. Wu, J. W. Hsieh, J. C. Cheng, S. C. Cheng, and S. Y. Tseng, “Human smoking event detection using visual interaction clues,” in *Proc. 20th International Conference on Pattern Recognition*, 2010.

[3] W. C. Wu and C. Y. Chen, “Detection system of smoking behavior based on face analysis,” in *Proc. Fifth International Conference on Genetic and Evolutionary Computing*, 2011.

[4] Y. Patil, P. Lopez-Meyer, S. Tiffany, and E. Sazonov, “Detection of cigarette smoke inhalations from respiratory signals using reduced feature set,” in *Proc. 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2013.

[5] Y. Patil, S. Tiffany, and E. Sazonov, “Understanding smoking behavior using wearable sensors - Relative importance of various sensor modalities,” in *Proc. 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014.

[6] I. T. U. Echebarr á, S. A. Imtiaz, M. Peng, and E. R. Villegas, “Monitoring smoking behaviour using a wearable acoustic sensor,” in *Proc. 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2017.

[7] B. Bhandari, J. C. Lu, X. Zheng, S. Rajasegarar, and C. Karmakar, “Non-invasive sensor based automated smoking activity detection,” in *Proc. 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2017.

[8] A. Filonenko, D. C. Hern ández, and K. H. Jo, “Smoke detection for static cameras,” in *Proc. 21st Korea-Japan Joint Workshop on Frontiers of Computer Vision*, 2015.

[9] J. Z. Lin, “Driver smoking behavior detection with facial video processing,” master thesis, National Chung Hsing University, 2015.

[10] C. A. Cole, D. Anshari, V. Lambert, J. F. Thrasher, and H. Valafar, “Detecting smoking events using accelerometer data collected via smartwatch technology: Validation study,” *JMIR Mhealth Uhealth.*, vol. 5, no. 12, p. e189, Dec. 13, 2017.

[11] YOLO: Real-Time object detection. [Online]. Available: https://pjreddie.com/darknet/yolov2/

[12] LabelImg. [Online]. Available: https://github.com/tzutalin/labelImg

Tzu-Chih Chien was born in Pingtung, Taiwan, in 1982. He received the B.S. degree in Electronic Engineering form National Chin-Yi University of Technology, Taichung, Taiwan, in 2006, and the M.S. degree in Electrical Engineering from National Chung Hsing University, Taichung, Taiwan, in 2019. His research interests include digital video/image processing and VLSI signal processing design.

Chieh-Chuan Lin was born in Taichung, Taiwan, 1985. He received a bachelor's degree in electrical engineering from the Chien Kuo Technology University, Taiwan, in 2007, and the M.S. degree in Electrical Engineering from National Chung Hsing University, Taichung, Taiwan, in 2015. His research interests include digital video/image processing and VLSI signal processing design.

Chih-Peng Fan received the B.S., M.S., and, and Ph.D. degrees, all in electrical engineering, from the National Cheng Kung University, Taiwan, in 1991, 1993 and 1998, respectively. During October 1998 to January 2003, he was a design engineer with N100, Computer and Communications Research Laboratories (CCL), Industrial Technology Research Institute (ITRI), Hsinchu, Taiwan. In 2003, he joined the faculty of the Department of Electrical Engineering, National Chung Hsing University, Taichung city, Taiwan, where he is currently a full Professor. He has more than 100 publications, which include technical journals, book chapters, conference papers, and technical reports. His teaching and research interests include digital image processing and pattern recognition, digital video processing, baseband transceiver design, and VLSI design and FPGA prototype of DSP systems.