

## Retraction Notice

Title of retracted article: Convolutional Neural Network and Bayesian Gaussian Process in Driving Anger Recognition

Author(s): Bowen Cai, Wufei Ma

\* Corresponding author' Email: caibowen@tongji.edu.cn

Journal: Engineering (ENG)  
 Year: 2020  
 Volume: 12  
 Number: 7  
 Pages (from - to): 534-548  
 DOI (to PDF): <https://doi.org/10.4236/eng.2020.127037>  
 Paper ID at SCIRP: 101789  
 Article page: <https://www.scirp.org/journal/paperinformation.aspx?paperid=101789>  
 Retraction date: 2021-05-08

### Retraction initiative (multiple responses allowed; mark with X):

- All authors  
 Some of the authors:  
 Editor with hints from  Journal owner (publisher)  
 Institution:  
 Reader:  
 Other:

Date initiative is launched: 2021-05-08

### Retraction type (multiple responses allowed):

- Unreliable findings  
 Lab error  Inconsistent data  Analytical error  Biased interpretation  
 Other:  
 Irreproducible results  
 Failure to disclose a major competing interest likely to influence interpretations or recommendations  
 Unethical research  
 Fraud  
 Data fabrication  Fake publication  Other:  
 Plagiarism  Self plagiarism  Overlap  Redundant publication \*  
 Copyright infringement  Other legal concern:  
 Editorial reasons  
 Handling error  Unreliable review(s)  Decision error  Other:

Other: ENG does not meet author's publication requirements.

### Results of publication (only one response allowed):

- are still valid.  
 were found to be overall invalid.

### Author's conduct (only one response allowed):

- honest error  
 academic misconduct  
 none (not applicable in this case – e.g. in case of editorial reasons)

\* Also called duplicate or repetitive publication. Definition: "Publishing or attempting to publish substantially the same work more than once."

**History**

Expression of Concern:

yes, date: yyyy-mm-dd

no

Correction:

yes, date: yyyy-mm-dd

no

**Comment:**

ENG does not meet author's publication requirements.

This article has been retracted to straighten the academic record. In making this decision the Editorial Board follows [COPE's Retraction Guidelines](#). Aim is to promote the circulation of scientific research by offering an ideal research publication platform with due consideration of internationally accepted standards on publication ethics. The Editorial Board would like to extend its sincere apologies for any inconvenience this retraction may have caused.

# Convolutional Neural Network and Bayesian Gaussian Process in Driving Anger Recognition

Bowen Cai<sup>1,2\*</sup>, Wufei Ma<sup>3</sup>

<sup>1</sup>The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Shanghai, China

<sup>2</sup>School of Transportation Engineering, Tongji University, Shanghai, China

<sup>3</sup>Department of Computer Science, Rensselaer Polytechnic Institute, Troy, NY, USA

Email: \*caibowen@tongji.edu.cn

**How to cite this paper:** Cai, B.W. and Ma, W.F. (2020) Convolutional Neural Network and Bayesian Gaussian Process in Driving Anger Recognition. *Engineering*, 12, 534-548. <https://doi.org/10.4236/eng.2020.127037>

**Received:** July 2, 2020

**Accepted:** July 25, 2020

**Published:** July 28, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

With the development of motorization, road traffic crashes have become the leading cause of death in many countries. Among roadway traffic crashes, almost 90% of accidents are related to driver behaviors, wherein driving anger is one of the most leading causes to vehicle crash-related conditions. To some extent, angry driving is considered more dangerous than typical driving distraction due to emotion agitation. Aggressive driving behaviors create many kinds of roadway traffic safety hazards. Mitigating potential risk caused by road rage is essential to increase the overall level of traffic safety. This paper puts forward an integrated computer vision model composed of convolutional neural network in feature extraction and Bayesian Gaussian process in classification to recognize driver anger and distinguish angry driving from natural driving status. Histogram of gradients (HOG) was applied to extract facial features. Convolutional neural network extracted features on eye, eyebrow, and mouth, which are considered most related to anger emotion. Extracted features with its probability were sent to Bayesian Gaussian process classifier as input. Integral analysis on three extracted features was conducted by Gaussian process classifier and output returned the likelihood of being anger from the overall study of all extracted features. An overall accuracy rate of 86.2% was achieved in this study. Tongji University 8-Degree-of-Freedom driving simulator was used to collect data from 30 recruited drivers and build test scenario.

## Keywords

Deep Learning, Road Rage, Computer Vision, Pattern Recognition, Dlib, Convolutional Neural Network, Anger Detection, Multidimensional Analysis

## 1. Introduction

Driving anger, also known as driving rage, characterized by feelings of annoyance, fury, or rage, is now becoming a serious traffic psychology issue. Driving anger is a significant contributor to risky driving and motor vehicle crashes, which are the leading causes of roadway morbidity and mortality [1]. Driving anger is initially defined as a specific situation consisting of emotional structures of feelings and thoughts associated with anger produced during driving [2]. Research shows that driving anger can lead to strong acceleration, higher speed and more yellow traffic light crossing [3] [4]. Drivers in the anger state make more errors on the lane keeping and on the traffic rules [5].

Driving anger is the common experienced emotional state on driving, leading to aggressive driving and risky operation. Deffenbacher *et al.* [2] introduced driving anger and proposed 14-item Driving Anger Scale (DAS). Qu *et al.* [6] pointed that approximately 94.4% of all traffic deaths in China are accounted for risky and aggressive driving behaviors. Reason *et al.* [7] initially put forwarded the Driver Behavior Questionnaire (DBQ), and classified into three subscales, violations, errors, and lapses to capture different aspects of driving behaviors. Shi *et al.* [8] revised DBQ contents to fit the actual needs of different studies and additional subscales.

Research on angry driving using questionnaires or driving-simulator experiments to explore the effect of angry driving on driving behavior has been studied for years, but the identification of angry driving has received less attention in previous studies. Some studies investigating the methodologies of driving anger detection are found as follows. Wan *et al.* [9] used drivers' physiological features such as heart rate, skin conductance, respiration rate, and electroencephalogram (EEG) to identify drivers' anger state in the driving process. The receiver operating characteristic (ROC) curve showed that the recognition accuracy of the model is 85.84% and demonstrates that this method can effectively identify a driver's anger state. Wang *et al.* [10] used a factorization model to recognize various driving emotions by extracting skin conductance, blood volume pulse, respiration rate, etc. Katsis *et al.* [11] applied decision trees and Naive Bayesian methods to identify car-racing drivers' emotions, such as stress level, dysphoria, and euphoria with features extracted from electromyography (EMG), electrocardiogram (ECG), electrodermal activity (EDA) and respiration. Fan *et al.* [12] utilized Bayesian network to classify driver emotion using EEG features of power spectrums and bring driver personality as well as traffic situation into analysis.

The above researches used physiological measurement criteria to identify the difference between natural status and driving anger. However, physiological measurement facilities and psychological related methods are not working in real time driving anger detection taking consideration of both inconvenience of equipping on driver themselves during driving state and costly expense of doing so. In addition, any non-real-time method of driver anger recognition is not ideally applicable to future commercial applications. Very a few papers proposed

video analysis on driver anger recognition. Azman *et al.* [13] applied Haar cascade classifier to locate face and recognize driver anger through support vector machine (SVM) in real time from a live video. Although 97% in-door accuracy was reported on 213 trained images with 5 fold cross-validation by Azman *et al.* [13], overfit was the problem it may be faced with for such limited training images. Gao *et al.* [14] developed a real-time non-intrusive monitoring system using linear SVMs to detect anger and disgust of drivers, and achieved 85.5% accuracy for in-car scenario. This paper initiates a real-time non-intrusive method based on HOG classifier, CNN, and Gaussian process to identify driver anger during driving process.

Study combined machine learning and deep learning algorithms to differentiate driving anger from natural driving status using only a camera to capture facial expressions. Histogram of gradients method was used to segregate facial features. Movements of facial features were tracked and analyzed with labeling as natural status or driving anger. VGG is one type of convolutional neural networks, which is commonly used for large scale image processing [15]. A VGG-like convolutional neural network (CNN) was applied to extract facial features from images. Extracted features from CNN were then sent to Bayesian Gaussian process as input to classify anger state from baseline natural status. Outputs of Gaussian process give the judgments of driver status.

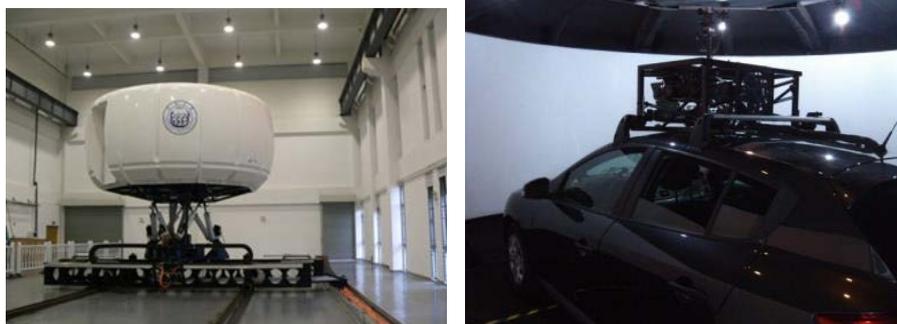
Three innovative points of this study are 1) using CNN and Gaussian process to identify drivers' anger during driving, 2) needing only a camera in the environmental setting with no other accessories as necessary, 3) unlike other studies used whole face to process, this study used HOG to extract facial organs as inputs of CNN, which reduces possible influences of different personal appearance for anger classification. The integrated methods of HOG, CNN & Gaussian process initiate an innovative way to recognize driving anger from natural driving status with high generality from person to person.

## 2. Data Preparation

### 2.1. Apparatus

Driving simulators can create a repeatable and safe environment in which to study driving behavior. This study used a high-fidelity simulator at Tongji University, currently the most advanced in China as shown in **Figure 1**. The simulator consists of a fully instrumented Renault Megane III vehicle, a dome with five projectors that provide a front image view of  $250^{\circ} \times 40^{\circ}$  at  $1000 \times 1050$  resolution and 60 Hz, an 8 degree-of-freedom motion system with an X-Y range of  $20 \times 5$  m, and SCANeR™ studio software to create driving scenarios and to generate displays.

Eye movement data were collected through the Eyelid Video Monitoring System (developed by Tongji University). The camera is placed in front of the vehicle to capture the driver's facial features and track head motion and determine the direction of the driver's sight through the detection and positioning of the



**Figure 1.** Tongji University 8-DoF driving simulator.

driver's pupils at a sampling rate of 60 Hz.

## 2.2. Participants

In this experiment, a total of 30 drivers (20 men and 10 women) aged 21 to 48 years (mean 25.97, SD 6.31) were recruited. The average driving experience of participants was 3.53 years (SD 2.77). All participants were required to have a valid Chinese driver's license, good health, no history of medicine use within the month prior to the experiment, no alcohol consumption within 24 hours, and no beverages with stimulants within 12 hours before the start of the experiment. Before the experiment, participants were required to sign the "Experimental Informed Consent", which described the experiment's requirements and the participants' rights. A cash reimbursement of 100 CNY (approximately 15 USD) was offered to each participant.

## 2.3. Experimental Scenario

The driving course used in the experimental scenario is a two-way four-lane mountain freeway with a total length of 20 km, as shown in **Figure 2(a)**. This mountain freeway, rather than an urban road, was chosen because the mountain freeway is a more monotonous visual scene. The complexity and variety of an urban road scene would subject the driver to distractions that could affect driving behavior and eye movement and thus lead to invalid results when studying driving distraction or driving anger. In order to increase the sense of environmental reality, green grass and trees were built on both sides of the road. There were no vehicles other than the subject vehicle on the road during baseline driving and the driving distraction task.

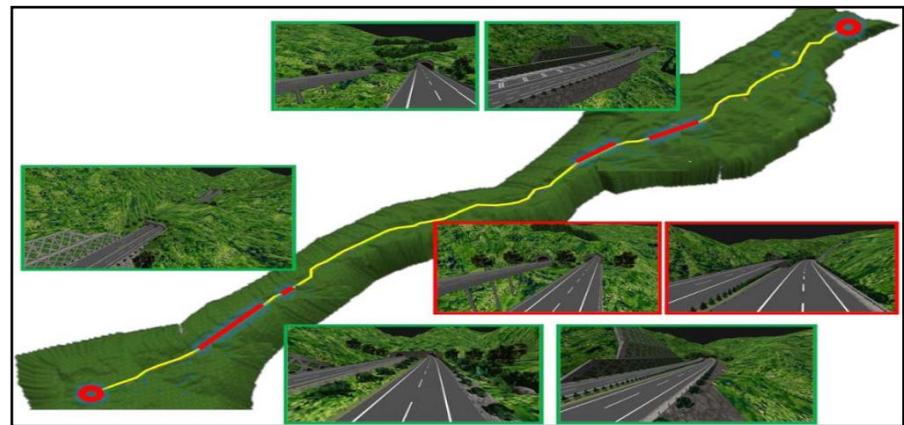
## 2.4. Procedure

The study consists of two parts: baseline driving and driving with anger stimuli.

### 1) Baseline driving

During baseline driving, drivers were asked to drive a designated route, and no anger stimuli were presented. The total driving time in this part of the experiment was about 15 min.

### 2) Driving with anger stimuli



(a)



(b)

**Figure 2.** Experiment scenario. (a) Mountain freeway with 4 tunnels marked in red lines; (b) Driver view of the experiment scenario.

This stage consisted of two parts. First, before driving, participants were asked to read a paragraph intended to induce anger. After the start of driving, some of the background vehicles were set to appear driving slowly or engaging in sudden decelerations or lane changes.

**Inducing Anger.** The most frequently used methods for inducing anger in the lab include film, stressful interviews, punishment, and harassment. Anger-induction methods that include personal contact, such as harassment and interviews, may produce more physiological reactivity [16]. Event recall and imagination tasks (“imagine the event as vividly as possible”) are two more commonly used anger-induction methods in the field of anger research [3] [17].

From the literature on existing anger-induction methods, some hypothesis scenarios were chosen to encourage participants to recall previous experiences of driving anger. These scenarios were selected from the Driving Anger Scale [2]. Participants were asked to select two commonly occurring situations in their daily lives and describe the scenes in detail. After finishing their descriptions, participants were asked to rate their current anger score from 1—*not angry at all* to 5—*extremely angry*. The following situations were used:

- You encountered road construction.
- You were stuck in a traffic jam due to an accident.

- You were driving in heavy rain.
- Someone cut you off and turned the opposite direction.
- Someone stopped suddenly in front of you.
- You were driving behind a large truck that made frequent stops.
- You were driving behind a large truck and you could not see around it.
- A pedestrian walked slowly in a crosswalk even when you had the green light.
- You missed the green light because a car backed out of a driveway into the intersection.
- Someone did not use indicators correctly.
- You were driving in heavy, slow traffic.

**Normal Driving:** Drivers were asked to imagine being in a hurry and needing to drive as rapidly as possible to reach a destination. The speed limit on the road is 100 km/h. Slow driving, sudden deceleration and sudden lane changes may occur in some background vehicles. Drivers were asked to drive along a designated route (with anger stimuli) and report their anger score (from 1—*not angry at all* to 5—*extremely angry*) after each stimulus event. The normal drive lasted about 10 minutes.

### 3. Methodology

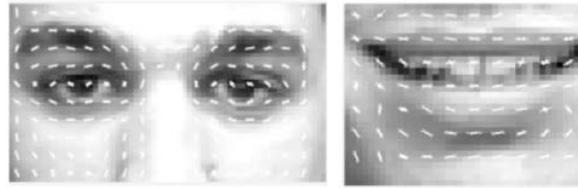
#### 3.1. Histogram of Oriented Gradient

Histogram of oriented gradient (HOG) is a feature descriptor for object detection, which is widely used in computer vision by counting the occurrences of gradient orientation in localized portions of an image. This technique is the gradient-based method that uses overlapping local contrast normalization. Use of HOG in our study is to locate specific facial features in the complicated environment, which can make the followed convolutional neural network classification more robust to noises. Compared to intensity or texture-wise methods, HOG contains more information in facial expression feature extraction [18]. Being as the illumination resistant and gradient-based feature descriptor, it divides the image into cells and calculates the magnitude and angular orientations through gradient filters. HOG calculation determines facial features by separating image into evenly sized and spaced grids. The orientation of the gradient for each pixel at  $(x, y)$  is calculated as Equation (1).

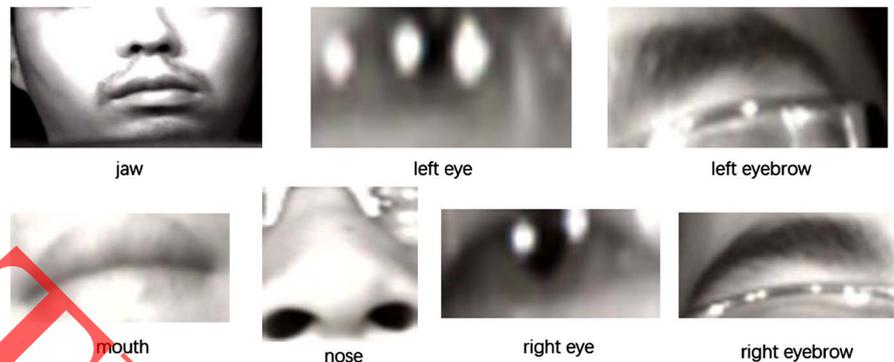
$$\theta_{x,y} = \tan^{-1} \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (1)$$

where  $L$  is the intensity function of the image. These orientations of gradients are then binned into a histogram for each evenly sized and spaced grid, and every grid within the image is concatenated resulting in a HOG description vector. **Figure 3** is the visualization of HOG features on eye and month regions extraction.

With application of HOG, seven facial features: jaw, mouth, nose, left eye, right eye, left eyebrow, right eyebrow were extracted in our study, as shown in



**Figure 3.** Visualization of HOG features on eye and month regions.



**Figure 4.** Extracted seven facial expressions.

**Figure 4.** Multidimensional analysis was applied to cropped features to differentiate anger from baseline.

### 3.2. Convolutional Neural Network

Convolutional neural networks (CNNs) are used mainly in image processing and computer vision. Layers within the CNN are composed of neurons organized into three dimensions: the spatial dimensionality of the input (height and width) and the depth. Neurons within any given layer connect only to a small region of the layer preceding them, which is convoluted to each other. Three types of layers (convolutional layers, pooling layers, and fully connected layers) make up CNNs.

The convolutional layer will determine the output of the neurons that are connected to local regions of the input through the calculation of the scalar product between their weights and the region connected to the input volume. The pooling layer will perform down-sampling along the spatial dimensionality of the given input, further reducing the number of parameters within that activation. The fully connected layers will produce class scores from the activations for classification.

During training, the input to the convolutional neural network is a  $250 \times 250$  RGB image. In this study, the image was preprocessed by converting JPEG content to RGB grids of pixels. Then the RGB grids of pixels were converted into floating-point tensors and the mean RGB value was subtracted from each pixel. The pixel values were then rescaled from the original 0 to 255 to the final [0, 1] interval. Moreover, in order to avoid the overfitting problem, data augmentation methods were applied to generate more training data from existing samples to ensure that in the training process the model would never see the same picture

twice. The rotation range of the pictures was set at 40, meaning that the pictures would be rotated randomly in value from 0 to 40 degrees. Shear and shift range were set to be 0.2, meaning that the pictures would be randomly translated and sheared vertically and horizontally as a fraction of 0.2 of the general size. Feature-wise standardized normalization was also applied to divide inputs by standard deviation of the data set.

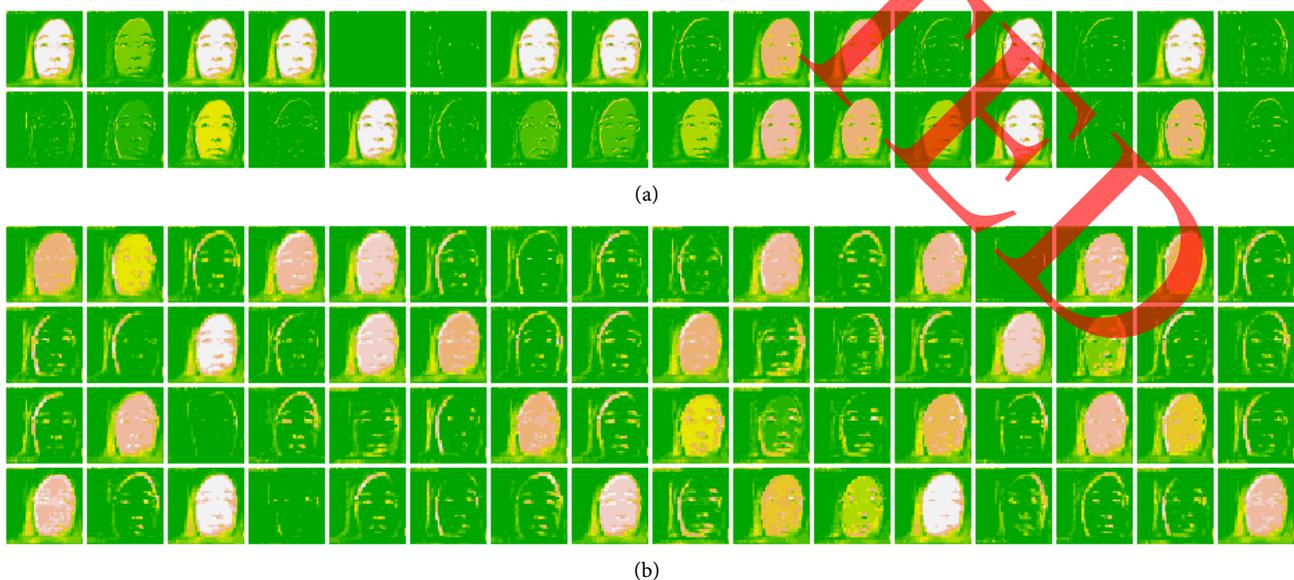
The image was passed through a stack of convolutional layers with filters set to be a small receptive field  $3 \times 3$  to capture the notion of left/right, up/down, center, in the smallest size. The convolution stride was set to be 1 pixel, and the padding was set to be 1 pixel for  $3 \times 3$  convolutional layers. Max-pooling was performed over a  $2 \times 2$  pixel window with stride 2. The activation function for all hidden layers in the convolutional layers was the rectification non-linearity (ReLU) function.

CNN transforms the original input layer by layer using convolutional and down-sampling techniques to produce class scores for classification and regression purposes. **Figure 5** shows the visualization of activations taken from the randomly selected convolutional layers (the first layer and the fourth layer) of the deep learning neural network built in this study.

It is easy to see that the convolutional layers have successfully picked characteristics unique to specific facial features. Different convolutional layers scan different facial expression features and create feature maps through learned filters to summarize the presence of features. The fully connected layer combines all the convolutional layers to produce the final classification score for a certain participant, e.g., whether the driver appears to be experiencing road anger in this case.

### 3.3. Bayesian Gaussian Process

Gaussian process model can be used for binary classification. Let  $y \in \{-1, 1\}$



**Figure 5.** Visualization of convolutional neural network. (a) First layer; (b) Fourth layer.

denote the class label of an input  $x$ . Gaussian process classification models  $p(y|x)$  for given  $x$  by a Bernoulli distribution. The probability of success  $p(y=1|x)$  is related to an unconstrained latent function  $f(x)$  which is mapped to the unit interval by a sigmoid transformation. Probit model  $p(y=1|x)=\Phi(f(x))$  is exclusively used for convenience, where  $\Phi$  denotes the cumulative density function of the standard Normal distribution [19].

For binary classification, the basic idea of Gaussian Process prediction is to place a Gaussian process prior over the latent function  $f(x)$ , which is then squashed through the logistic function to obtain a prior on  $\pi(x) \triangleq p(y=+1|x) = \sigma(f(x))$ . The latent function  $f(x)$  is also known as the nuisance function, which allows a convenient formulation of the model [19].

Inference is divided into two steps. In the first step, we compute the distribution of the latent variable using

$$p(f_*|X, y, x_*) = \int p(f_*|X, x_*, f) p(f|X, y) df$$

and in the second step, a probabilistic prediction is obtained using this distribution over the latent  $f_*$

$$\bar{\pi}_* \triangleq p(y_* = +1|X, y, x_*) = \int \sigma(f_*) p(f_*|X, y, x_*) df_*.$$

## 4. Modeling Procedure and Data Analysis

### 4.1. Modeling Procedure

Deep learning convolutional neural networks are integrated with Gaussian process to distinguish driving anger from natural expression. This study used the combined models because conventional computer vision techniques like HOG can easily capture facial features from complex environments and CNN can classify the status of each extracted facial feature of being natural or anger. The obvious drawback of CNN classification on each feature is that it may misclassify

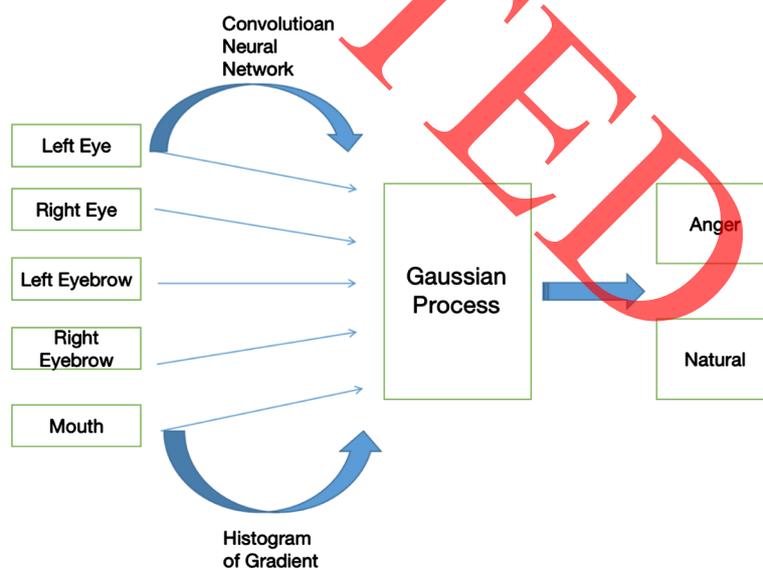


Figure 6. Flow chart of algorithm processing procedure.

expressions because it uses only certain features, instead of the whole face, to make comparisons. Therefore, Gaussian process was used to combine results returned from CNN and classify the overall facial expression with taking consideration into every feature expression. The modeling procedure is summarized in **Figure 6**.

#### 4.2. Convolutional Neural Network

The whole data pool was formed by 3000 facial images extracted from 30 tested drivers labeled either natural or anger. Images of 20 drivers were put into a training set, 5 other drivers' images were put into a validation set, and the remaining 5 drivers were put into a test set to judge the model's accuracy. Eye, eyebrow, and mouth are considered as facial expression features to indicate anger emotion [13]. CNN was built to process eye, eyebrow, and mouth images extracted by HOG from the whole face. Input size was set at  $70 \times 140$  with 3 channels for left and right eyes,  $128 \times 256$  for mouth,  $100 \times 240$  for left and right eye-brows. A rectified linear unit function was set as the activation function. Max-pooling was used to reduce the dimensionality of the representation. Same convolutional neural network structure was used for processing eye, eyebrow and mouth. Four convolutional layers were formed with stacked layers, and two fully connected layers were formed with 512 input filters. In total, 1,765,473 parameters were tuned through a back-propagation process. The CNN architecture is shown in **Figure 7**.

A Graphic Processing Unit Quadro P-4000 was used to train the model in TensorFlow and provide the predicted image labels. The training process was carried out by minimizing the loss function of the binary cross-entropy using mini-batch gradient descent with momentum. Weight decay of 0.0005 and dropout regularization for the first fully connected layers with dropout ratio of 0.5

```
Model: "sequential"
```

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 98, 238, 16)	448
max_pooling2d (MaxPooling2D)	(None, 49, 119, 16)	0
conv2d_1 (Conv2D)	(None, 47, 117, 32)	4640
max_pooling2d_1 (MaxPooling2D)	(None, 23, 58, 32)	0
conv2d_2 (Conv2D)	(None, 21, 56, 64)	18496
max_pooling2d_2 (MaxPooling2D)	(None, 10, 28, 64)	0
conv2d_3 (Conv2D)	(None, 8, 26, 64)	36928
max_pooling2d_3 (MaxPooling2D)	(None, 4, 13, 64)	0
flatten (Flatten)	(None, 3328)	0
dense (Dense)	(None, 512)	1704448
dense_1 (Dense)	(None, 1)	513

```
Total params: 1,765,473
Trainable params: 1,765,473
Non-trainable params: 0
```

**Figure 7.** Architecture of the trained CNN.

were applied to training process. Learning rate was set to be 0.00001 to avoid the missing local minima. During the training process, weights were updated for each batch.

### 4.3. Gaussian Process

Many studies investigated anger expression on whole face. However, faces vary from person to person. Judgment may get influenced by personal looking. This paper used HOG to extract key facial features, such as eye, eye-brow, and mouth as they are considered most related to anger emotion [18]. Convolutional neural network was applied to process extracted features to get probability of each facial feature being anger expression. Output probability for extracted facial features returned from CNN was sent to Bayesian Gaussian process classifier as input. Gaussian process was succeeded to bring all the extracted facial features into analysis to form comprehensive study on the expression of each facial feature to indicate road anger. Gaussian process classifier returned the likelihood of being anger of the whole face by processing probability scores of CNN outputs on each facial feature. The prediction from Gaussian process is probabilistic and threshold of 0.8 was specified to refit the prediction for reducing false positive rate of classifying natural status into anger which may make driver feel annoyed.

### 4.4. Data Analysis

The overall accuracy of facial expression recognition for the integrated model of pattern recognition and convolutional neural network was 86.2%. False positive calls are compressed in purpose by setting higher threshold of 0.8 to classify driving status into road anger. Recall rate measures the fraction of the total amount of relevant instances that were actually retrieved [20]. The cross-accuracy table for pattern recognition procedure is shown in Table 1.

Compared with pattern recognition, CNN has much higher model transferability. Traditional pattern recognition cannot correctly differentiate anger from baseline because pattern recognition uses fixed dimensions as criteria to classify mixed classes. If more dimensions were used, then the problem of overfitting would arise because of curse of dimensionality. Deep learning can ameliorate the problem of dimensionality by modeling the functionality of an artificial neural network with more hidden layers added.

After passing through the CNN and Gaussian process, true positive and true negative rates are 81.2% and 91.3% respectively, where 2734 out of 3000 true-baseline samples were classified correctly as natural status, for the recall rate of

**Table 1.** Cross accuracy for pattern recognition.

	Classified to Anger	Classified to Baseline	Recall Rate
True Anger	2436	564	81.2%
True Baseline	266	2734	91.3%
Accuracy Rate		86.2%	

91.3%. Within true anger, 2436 out of 3000 were correctly classified as real anger with the recall rate of 81.2%, as shown. Using CNN and Gaussian process together, a total of 86.2% accuracy can be achieved for driving anger detection by measuring overall accuracy as the sum of diagonal divided by the total number of pictures in the cross table.

Thousands of statisticians have spent years working to improve the accuracy of methods to detect human emotion. Although deep learning neural networks have achieved relatively high accuracy in classifying anger from baseline, some images are still hard to differentiate, see **Figure 8**. Looking at the misclassified images, even for human eyes, it is often difficult to identify them correctly.

## 5. Conclusions

Aggressive driving behaviors originate from road rage. Road anger detection using only a camera is difficult because the appearance of anger varies from person to person. Additionally, driver road anger is much less observable than typical anger. Typical anger status is often more exaggerated than road rage during driving process. In some cases, anger can be detected by measuring heartbeat and using voice analysis as accessory tools in addition to vision. Nevertheless, if additional devices are required, anger detection becomes impractical in real situations. This paper puts forward an efficient way to identify driver anger with moderate accuracy using a computer vision deep learning neural network and Gaussian process. Even though its overall 86.2% accuracy rate is laboratory-based and more data and field tests are needed to improve the algorithm, the integrated model can automatically recognize human faces and capture facial features in complex and extreme environments, such as noisy backgrounds, strong illumination, backlighting, and others. Accuracy of the CNN integrated model achieves a relatively high accuracy in road rage detection and can win over most of the anger emotion detection models in current use.

A camera with an embedded chip and a local computer with no additional



**Figure 8.** Examples of misclassified images. (a) Anger but classified into Baseline; (b) Baseline but classified into Anger.

processing units is sufficient to run this model. No Internet connection is required to transfer judgment results to the vehicle central control system, which can play a warning message or take over the car if it is equipped with an advanced autonomous driving assistance system.

The advantages of this method are as follows:

First, this study proposes the method to detect road rage by use of camera only, making it more pragmatic in broader use.

Second, specific facial features (e.g. eye, eyebrow, and mouth) are extracted from the whole face. Being as the processed unit of CNN, it excludes the influence of face dissimilarity to large extent since some faces are more likely to be recognized as anger than the others.

Third, integration of HOG, CNN, and Gaussian process helps make the methodology to fit into various complex and extreme environments with HOG to locate face, CNN to extract features, and Gaussian process to proceed facial features and get the classification.

## 6. Discussion

Road anger expression has a different appearance from common anger expression.

Compared to road anger, typical anger is more obvious and easier to detect. Facial expressions for road anger are less exaggerated. Therefore, more effort should be made to distinguish road anger from natural driving status through in-depth minor facial expression analysis. However, identification of road anger from camera recording only is hard. It should be a comprehensive analysis with other parameters included. In future study, vehicle operational data such as speed and acceleration will also be involved because operation conditions are good reflections of rage driving. Furthermore, band watch may be used to collect heart beating rate, blood pressure, and oxygen saturation from drivers. Valuable information collected from vehicle operation, band watch driver status monitoring, along with camera recordings will form the comprehensive research on road rage analysis.

In this study, although data is being collected, classification of different serious levels of driving anger is not covered by this paper, which will be the next research topic of traffic safety analysis in road rage detection with combination of the analysis of vehicle operation parameters and driver status monitoring.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Zinzow, H. and Jeffers, S. (2017) Driving Aggression and Anxiety: Intersections, Assessment, and Interventions. *Journal of Clinical Psychology*, **74**, 43-82.

- <https://doi.org/10.1002/jclp.22494>
- [2] Deffenbacher, J.L., Oetting, E.R. and Lynch, R.S. (1994) Development of a Driving Anger Scale. *Psychological Reports*, **74**, 83-91. <https://doi.org/10.2466/pr0.1994.74.1.83>
- [3] Abdu, R., Shinar, D. and Meiran, N. (2012) Situational (State) Anger and Driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, **15**, 575-580. <https://doi.org/10.1016/j.trf.2012.05.007>
- [4] Roidl, E., Frehse, B. and Höger, R. (2014) Emotional States of Drivers and the Impact on Speed, Acceleration and Traffic Violations—A Simulator Study. *Accident Analysis & Prevention*, **70**, 282-292. <https://doi.org/10.1016/j.aap.2014.04.010>
- [5] Jeon, M., Walker, B.N. and Yim, J.-B. (2014) Effects of Specific Emotions on Subjective Judgment, Driving Performance, and Perceived Workload. *Transportation Research Part F: Traffic Psychology and Behaviour*, **24**, 197-209. <https://doi.org/10.1016/j.trf.2014.04.003>
- [6] Qu, W., Ge, Y., Jiang, C., Du, F. and Zhang, K. (2014) The Dula Dangerous Driving Index in China: An investigation of reliability and validity. *Accident Analysis & Prevention*, **64**, 62-68. <https://doi.org/10.1016/j.aap.2013.11.004>
- [7] Reason, J., Manstead, A., Stradling, S., Baxter, J. and Campbell, K. (2011) Errors and Violations on the Roads: A Real Distinction? *Ergonomics*, **33**, 1315-1332. <https://doi.org/10.1080/00140139008925335>
- [8] Shi, J., Bai, Y., Ying, X. and Atchley, P. (2010) Aberrant Driving Behaviors: A Study of Drivers in Beijing. *Accident Analysis & Prevention*, **42**, 1031-1040. <https://doi.org/10.1016/j.aap.2009.12.010>
- [9] Wan, P., Wu, C., Lin, Y. and Ma, X. (2017) On-Road Experimental Study on Driving Anger Identification Model Based on Physiological Features by ROC Curve Analysis. *IET Intelligent Transport Systems*, **11**, 290-298. <https://digital-library.theiet.org/content/journals/10.1049/iet-its.2016.0127>  
<https://doi.org/10.1049/iet-its.2016.0127>
- [10] Wang, J. and Gong, Y. (2009) Normalizing Multi-Subject Variation for Drivers' Emotion Recognition. 2009 *IEEE International Conference on Multimedia and Expo*, Hilton Cancun, 28 June-3 July 2009, 354-357. <https://doi.org/10.1109/ICME.2009.5202507>
- [11] Katsis, C.D., Goletsis, Y., Rigas, G. and Fotiadis, D.I. (2011) A Wearable System for the Affective Monitoring of Car Racing Drivers during Simulated Conditions. *Transportation Research Part C: Emerging Technologies*, **19**, 541-551. <https://doi.org/10.1016/j.trc.2010.09.004>
- [12] Fan, X., Bi, L. and Chen, Z. (2010) Using EEG to Detect Drivers' Emotion with Bayesian Networks. 2010 *International Conference on Machine Learning and Cybernetics*, Vol. 3, 1177-1181. <https://doi.org/10.1109/ICMLC.2010.5580919>
- [13] Azman, A., et al. (2019) Real Time Driver Anger Detection. In: *Information Science and Applications 2018*, Springer, Berlin, 157-167. [https://doi.org/10.1007/978-981-13-1056-0\\_17](https://doi.org/10.1007/978-981-13-1056-0_17)
- [14] Gao, H., Yüce, A. and Thiran, J. (2014) Detecting Emotional Stress from Facial Expressions for Driving Safety. 2014 *IEEE International Conference on Image Processing*, Paris, 27-30 October 2014, 5961-5965. <https://doi.org/10.1109/ICIP.2014.7026203>
- [15] Cireşan, D., Meier, U., Masci, J., Gambardella, L.M. and Schmidhuber, J. (2011) Flexible, High Performance Convolutional Neural Networks for Image Classification. *Proceedings of the 22nd International Joint Conference on Artificial Intelli-*

*gence*, Barcelona, 16-22 July 2011, 1237-1242.

<https://doi.org/10.1109/IJCNN.2011.6033458>

- [16] Lobbestael, J., Arntz, A. and Wiers, R. (2008) How to Push Someone's Buttons: A Comparison of Four Anger-Induction Methods. *Cognition and Emotion*, **22**, 353-373. <https://doi.org/10.1080/02699930701438285>
- [17] Rusting, C.L. and Nolen-Hoeksema, S. (1998) Regulating Responses to Anger: Effects of Rumination and Distraction on Angry Mood. *Journal of Personality and Social Psychology*, **74**, 790-803. <https://doi.org/10.1037/0022-3514.74.3.790>
- [18] Kumari, J., Rajesh, R. and Kumar, A. (2016) Fusion of Features for the Effective Facial Expression Recognition. 2016 *International Conference on Communication and Signal Processing*, Madras, 6-8 April 2016, 457-461. <https://doi.org/10.1109/ICCSP.2016.7754178>
- [19] Malte, K. and Rasmussen, C.E. (2005) Assessing Approximations for Gaussian Process Classification.
- [20] Hjørland, B. (2010) The Foundation of the Concept of Relevance. *Journal of the American Society for Information Science and Technology*, **61**, 217-237.

RETRACTED