DEEP LEARNING TECHNOLOGY FOR FACE RECOGNITION

by

Bin YUAN *

Key Laboratory of Analytical Mathematics and Intelligent Computing
for Yunnan Provincial Department of Education and School of Mathematics and Economics,
Qujing Normal University, Qujing, China

Original scientific paper
https://doi.org/10.2298/TSCI2503007Y

In China, the rapid development of public transportation network construction has been accompanied by a high incidence of traffic accidents caused by sleep-deprived driving. The monitoring of drivers' sleep-deprived driving and the sending out of early warnings has been identified as a field of research with both important theoretical and practical value. This article proposes a fatigue detection algorithm based on facial recognition information fusion. The algorithm extracts facial feature information and head features from the driver's face and fuses them into facial recognition information parameter feature vectors. A set of feature nodes is thus constructed, and an information fusion feature map is constructed according to the interaction information between nodes. The fatigue status is detected using the label propagation strategy, which is more effective than other benchmark detection algorithms. The fatigue detection algorithm has practical value.

Key words: *fatigue test, face recognition, information fusion, deep learning*

Introduction

The National Bureau of Statistics, as evidenced by data from the China Statistical Yearbook, has released information indicating that over the course of nearly five years, the average number of traffic accidents in China has been approximately 231900 per year. Moreover, the number of individuals who perish in road traffic incidents averages 63243 annually. In general, the number of accidents and the number of fatalities resulting from them have been on the rise year by year.

A review of traffic accident data from the past five years indicates that approximately 20% of all road accidents are attributable to sleep-deprived driving. Furthermore, the proportion of road accidents attributed to sleep-deprived driving may be as high as 50% on some roads. A survey of freight vehicle drivers conducted by relevant Chinese departments revealed that 84% of respondents had an average daily driving time of over 8 hours, with 40% exceeding 12 hours. Furthermore, 64% of freight vehicles are only equipped with one driver [1]. Based on the aforementioned data analysis, It is of great significance to detect the prevalence of sleep-deprived driving among motor vehicle drivers. Safety warnings can be issued in advance when the driver enters a fatigue state to ensure driving safety and reduce the probability

* Author’s e-mail: 501190349@qq.com

of traffic accidents. In order to gradually reduce the personal safety threat and economic losses caused by sleep-deprived driving [2], it is necessary to implement effective measures.

Currently, there is a great deal of interest among scholars at home and abroad in the field of fatigue detection. A significant number of researchers have employed deep learning and computer vision technologies in the field of fatigue detection. Furthermore, the field of fatigue detection has emerged as a significant research frontier [3].

Fatigue driving detection algorithm based on facial feature information fusion

Design of fatigue detection algorithm based on face recognition information fusion

This article proposes a fatigue detection algorithm, face recognition information fusion based fatigue detection algorithm (FRFFDA), which fuses face recognition feature information. This information is extracted and fused, and a feature graph is constructed. Furthermore, label propagation is employed [4]. First, the facial features of the eye and mouth, as well as those of the head, are extracted and then fused into a single feature node, designated x_i. Secondly, a multi-source data information feature graph is constructed. Third, the fatigue level of each node is identified through the label propagation strategy.

In order to address the issue of data analysis related to fatigue detection. This article employs a fatigue detection methodology based on the integration of facial recognition feature information [5]. The algorithm comprises three algorithms, which are in a progressive relationship with one another.

– Active learning algorithm based on information entropy (ALAIE). The set of nodes X is divided into two subsets, X_l and X_u, according to the principle of information entropy. The X_l is pre-labeled, while X_u is unlabeled.
– Feature graph construction algorithm based on KNN (FGCAK). The feature nodes are mapped into a graph in order to construct a fatigue detection feature graph.
– Fatigue detection algorithm based on label propagation strategy (FDALPS). The labels are propagated on the face feature graph for the purpose of fatigue state analysis and recognition.

Active learning algorithm based on information entropy

Active learning is a process whereby a small number of nodes from a given dataset are selected for labeling, employing a specific selection strategy. The ALAIE algorithm selects a small number of nodes and labels them to form a labeled subset X_l, while the remaining nodes become an unlabeled subset X_u.

Fatigue is an overall feeling of sleepiness or lack of energy, and its definition is vague. In academia, there is no precise physical quantity to determine the degree of human fatigue, which makes fatigue detection difficult. Researchers have taken different ways to measure the fatigue state [6]. In this article, we use information entropy to describe this uncertainty of fatigue, adopting the criterion of maximum information entropy to select l nodes that need to be labeled. And then constructing the pre-labeled subset X_l. The total number of extracted features is n^m in a time window of size n^w, and the use of υ_{ij} denotes the value when the i^th feature is in the j^th frame, and $\mathbf{V} = (\upsilon_{ij})n^m \times n^w$ can be used to denote the feature matrix within n^w, which can be expanded to obtain eq. (1):

$$\mathbf{V} = (\upsilon_{ij})n^m \times n^f = \begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1n_w} \\ V_{21} & V_{22} & \cdots & V_{2n_w} \\ \vdots & \vdots & \vdots & \vdots \\ V_n m_1 & V_n m_2 & & V_n m_{n_w} \end{bmatrix} \tag{1}$$

In the specified temporal interval, n^w, the specific calculation of feature nodes x is derived from eq. (2):

$$x = \frac{1}{n^m}(V_1, V_2, ..., V_n m) \tag{2}$$

Next, the feature matrix $\mathbf{V}$ will be normalized, the normalization matrix $V^S = (v_{ij}^S)_n m_{*n} f$, where eq. (3) is used to calculate the standard term v_{ij}^S.

$$v_{ij}^S = \frac{v_{ij} - \min_j \{v_{ij}\}}{\max_j \{v_{ij}\} - \min_j \{v_{ij}\}} \tag{3}$$

This gives the information entropy $h(x)$ of the node x in the time window n^w as shown in:

$$h(x) = \sum_{j=1}^{n^w} \sum_{i=1}^{n^m} p(v_{ij}^S) \left[\ln \frac{1}{p(v_{ij}^S)} \right] \tag{4}$$

$$p(v_{ij}^S) = v_{ij}^S \left(\sum_{i=1}^{n^m} v_{ij}^S \right)^{-1} \tag{5}$$

where $p(v_{ij}^S)$ in eq. (5) describes the distribution of feature v_{ij}, and $h(x) = 0$ when $p(v_{ij}^S) = 0$. After this step, the information entropy of all nodes is obtained. Based on the principle of maximum information entropy, l nodes with the largest entropy among them are selected, and assigned initial labels, these nodes form the pre-labeled subset X_l, and the remaining unlabeled nodes become the subset X_u.

Feature graph construction algorithm based on K-nearest neighbor

This article presents a novel feature graph construction algorithm (FGCAK) based on the K-nearest neighbor (KNN) approach. The algorithm is designed to identify node-to-node relationships. Furthermore, the construction of feature graphs for subsequent use in community discovery is employed to achieve fatigue state identification [7].

Firstly, the feature map is constructed from the feature nodes, and its definition is presented:

$$G = (V, E) \tag{6}$$

where V is the set of feature points and E – the set of edges. In a community, the internal nodes are more densely connected, exhibiting greater mutual information (MI) between them. In contrast, nodes between different communities are more sparsely connected. Consequently, the mutual information between two vertices is of paramount importance for the construction of edges. It is therefore necessary to connect nodes with strong mutual information to one another [8].

In this article, we use affinity distance based on Gaussian function to measure the relationship between nodes, which is defined as shown:

$$s_{ij} = \mathrm{e}^{-\frac{\|x_i - x_j\|^2}{2\sigma^2}} \tag{7}$$

where s_{ij} is the affinity distance between node x_i and node x_j, and the parameter σ is obtained by learning the distance between a node and its neighbor nodes. Inspired by the EFANNA algorithm, for node x_i. It is necessary to find out the k neighbors with the closest affinity distances to it. Then connect x_i to the k nodes, respectively, to form k edges. The process is repeated for all feature nodes, and finally the constructed feature graph is obtained [9].

Fatigue detection algorithm based on label propagation strategy

Once the construction of the KNN facial feature maps has been completed. This section presents a fatigue detection algorithm based on the label propagation strategy, which is employed to identify fatigue states on the feature maps. Similarly to the pre-annotation process. This article employs information entropy to assess the mutual influence between two nodes [10]. Nodes with greater entropy often exert a greater influence on their neighbors than nodes with smaller entropy. With regard to distance, the closer two nodes are, the more likely they are to have similar labels. As a result, the node x_i may contain multiple labels from its neighbors. Consequently, it is essential to calculate the membership of x_i for each label in each propagation process. The label is then assigned to the label set of x_i, and finally, the label with the highest membership is selected as the label for this node [11].

In the multi-round label propagation process, the propagation ability of the neighboring nodes is crucial to the propagation process. In this paper, we use eq. (8) to express the propagation information, which represents the probability that node x_i propagates its label to node x_j.

$$p_{ij} = p_{ij} \log_2 \frac{p_{ij}}{p_i \times p_j} \tag{8}$$

where p_{ij} is the intersection probability of node x_i and node x_j and p_i – the occurrence rate of node x_i, which are computed respectively using eqs. (9) and (10):

$$p_{ij} = \frac{n_{ij}}{N} \tag{9}$$

$$p_{ij} = \frac{N_i}{N} \tag{10}$$

where n_{ij} is the number of identical neighbours of node x_i and node x_j, N – the total number of nodes in the feature graph, and N_i – the number of communities in the graph that contain node x_i and its neighbor nodes [12].

In the propagation process t, each neighbour node of x_i has the ability to propagate to it, and each node possesses a different label. This requires different weights to correspond to these actively propagating nodes and uses membership to describe the close relationship between labels and nodes. In general, the higher the weight of a node, the higher the priority of this node to send tags [13]. During the t^{th} propagation process, eq. (11) is used to define the weight of node x_i for the labels.

In the propagation process, each neighbor node of x_i has the capacity to propagate to it, and each node is assigned a distinct label. This necessitates the assignment of distinct weights to correspond with the actively propagating nodes, and the utilization of membership to delineate the intimate relationship between labels and nodes. In general, the greater the weight of a node, the greater the priority of this node to transmit tags [13]. In the t^{th} propagation process, eq. (11) is employed to determine the weight of node x_i for the labels.

$$\omega(c, x_i) = \sum\nolimits_{j \in N(x)} b_{t-1}(c, x_j) \times \theta(c, x_j) \times I_{ij} \tag{11}$$

where $N(x)$ is the set of neighbouring nodes x_i of node x_j, b – the membership of node x_j to the label, and θ – the discriminant function for determining whether the label c is in the set of labels L_j of node x_j, which is computed by eq. (12):

$$\theta(c, x_j) = \begin{cases} 1, & c \in L_j \\ 0, & \text{otherwis} \end{cases} \tag{12}$$

In this article, we use the membership $b(c, x_i)$ to represent the ability of the label c to influence the node x_i. The higher the degree of membership, the higher probability that the node has the label [14]. It is the ratio of node x_i weight $\omega(c, x_i)$ to the total weight $\omega(x_i)$ for label c, and is derived by eqs. (13) and (14):

$$\omega(x_i) = \sum\nolimits_{c \in L_i} \sum\nolimits_{j \in N(x)} b_{t-1}(c, x_j) \times I_{ij} \tag{13}$$

$$b_t(c, x_i) = \frac{\omega(c, x_i)}{\omega(x_i)} \tag{14}$$

Specifically, $b_0(c, x_i)$ comes from the percentage of labels c in its neighbouring nodes.

Experimental results and analysis of secondary classification

The objective of this article was to design five binary classification experiments with the aim of evaluating the FRFFDA algorithm. Due to the four distinct fatigue states present in the original dataset, In order to perform secondary classification experiments, it was necessary to first perform a preliminary classification [15]. It is necessary to reorganize the dataset in order to facilitate the secondary classification experiments. The new fatigue levels are designated as *normal driving* and *fatigue driving*. This article extracts data with only labels of 0000 and 1000. Furthermore, a continuous time window is employed to evaluate the driver's fatigue state [16].

The efficacy of this experiment is gauged by four metrics: precision, P, recall, R, accuracy, Acc, and $F1$-*score*. These are calculated according to the formulas presented in eqs. (15)-(18).

$$p = \frac{TP}{TP + FP} \times 100\% \tag{15}$$

$$R = \frac{TP}{TP + FP} \times 100\% \tag{16}$$

$$Acc = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \tag{17}$$

$$F1\text{-}score = \frac{2 \times P \times R}{P + R} \times 100\% \tag{18}$$

wherein, for each sample, there are four possible scenarios for their labels.

The terms true positives (TP) and true negatives (TN) are used to describe the number of samples where the values of both the actual and predicted labels are true or false, respectively. False positives (FP) denotes the number of samples where the actual labels are false, but the predicted labels are true, and false negatives (FN) denotes the number of samples where the actual labels are true, but the predicted labels are false [17].

The fatigue community is contingent upon the affinity distance threshold. This experiment also assesses the impact of varying parameters on the final community structure across different iterations. The ε is set to 0.1, 0.2, and 0.3 to ascertain the optimal classification accuracy that can be attained by FRFFDA. It can be demonstrated that the FRFFDA algorithm attains the highest accuracy when ε = 0.3. When the number of iterations is less than 2000, ε = 0.1 and ε = 0.2 yield superior outcomes, as a smaller threshold facilitates the aggregation of nodes in a more expeditious manner [18]. When ε is relatively large, a significant number of nodes remain unclassified until the subsequent stage. In subsequent experiments, the value of ε was set to 0.3.

This experiment employs four additional classical benchmark classification algorithms, including naive bayes (NB), random forest (RF), decision tree (DT), and support vector machine (SVM). Furthermore, the outcomes were contrasted with those of the FRFFDA algorithm proposed in this chapter. The experimental results are presented in tab. 1.

Table 1. Comparison of experimental results between FRFFDA and benchmark algorithm

Methods \ Metrics	P [%]	R [%]	Acc [%]	$F1$-*score* [%]
NB	83.49	77.61	81.96	80.44
RF	86.92	90.92	89.57	88.88
DT	89.00	88.40	88.73	88.69
SVM	88.23	91.42	89.37	89.80
FRFFDA	92.23	94.35	93.31	93.28

As illustrated in tab. 1, the results produced by different algorithms vary considerably. Among the aforementioned algorithms, NB performs less effectively than the other methods, due to the underlying assumption that each feature is independent of the others. In real-world scenarios, this assumption is often invalid [19]. This is particularly evident when the number of features is large or the correlation between features is high. The efficacy of this approach is not optimal. The results indicate that RF and SVM exhibit superior accuracy compared to the other two methods. The $F1$-*score* for NB, RF, DT, and SVM were 80.44%, 88.88%, 88.69%, and 89.80%, respectively. The calculations were performed using the data presented in tab. 1. It can be observed that the accuracy of FRFFDA is 11.35% higher than that of NB and 3.74% higher than that of RF, which represents the most effective benchmark method [20]. One advantage of FRFFDA as a semi-supervised learning method is that it requires less labeled data. Furthermore, the computational complexity of the label propagation process is approximately linear, which effectively reduces the computational burden of the algorithm. In summary, the FRFFDA algorithm demonstrated the most favorable outcomes among the aforementioned methods [21]. The algorithm achieved the highest F1-score of 93.28% and the best accuracy of 93.31%.

Conclusion

To date, there is no unified standard for the criteria for fatigue detection. This is particularly pertinent in the context of fatigue occurring during driving. The occurrence of fatigue is influenced by a multitude of factors, including the mental state of the driver, the driving environment, and other variables. In the current body of research, these two types of fatigue states are the most prevalent. However, in reality, fatigue changes dynamically over time, exhibiting characteristics such as ambiguity and uncertainty [22, 23]. In future work, the challenge will be to collect accurate data from various complex driving environments. This data will then need to be used to model and integrate features of different information sources, types, and data for comprehensive analysis. Furthermore, the efficacy of algorithms must be evaluated in actual driving environments. In order to approach the human body state in real driving environments, it is necessary to. This is a direction that will require further in-depth research in the future.

Acknowledgment

This work was supported by Key Laboratory of Analytical Mathematics and Intelligent Computing for Yunnan Provincial Department of Education. This work was supported in part by the Education Department of Yunnan Province. And in part by the Yunnan Provincial Department of Education of China under Grant 2023J1030, Grant 2023J1028.

References

[1] Shahrndin, N. S. N., Sidek, K. A., A Review of ECG Data Acquisition for Driver Drowsiness Detection, *Compusoft, 9* (2020), 7, pp. 3749-3754
[2] Girvan, M., Newman, M. E. J., Community Structure in Social and Biological Networks, *Proceedings of the National Academy of Sciences, 99* (2002), 12, pp. 7821-7826
[3] Mott, G. E., *et al.*, Efficient Driver Drowsiness Detection at Moderate Levels of Drowsiness, *Accident Analysis and Prevention, 50* (2013), 1, pp. 341-350
[4] Li, Z. J., *et al.*, Online Detection of Driver Fatigue Using Steering Wheel Angles for Real Driving Conditions, *Sensors, 17* (2017), 3, pp. 495-507
[5] Wang, W., *et al.*, A Survey of Zero-Shot Learning: Settings, Methods, and Applications, *ACM Transactions on Intelligent Systems and Technology, 10* (2019), 2, pp. 1-37
[6] Furugori, S., *et al.*, Estimation of Driver Fatigue by Pressure Distribution on Seat in Long Term Driving, *Review of Automotive Engineering, 26* (2005), 1, pp. 53-58
[7] Wei, Y., *et al.*, Multi-Vehicle Detection Algorithm through Combining Harr and Hog Features, *Mathematics and Computers in Simulation, 79* (2019), 8, pp. 130-145
[8] Bila, C., *et al.*, Vehicles of the Future: a Survey of Research on Safety Issues, *IEEE Transactions on Intelligent Transportation Systems, 18* (2017), 5, pp. 1046-1065
[9] Mandal, B., *et al.*, Towards Detection of Bus Driver Fatigue Based on Robust Visual Analysis of Eye State, *IEEE Transactions on Intelligent Transportation Systems, 18* (2017), 3, pp. 545-557

[10] Nie, B. S., *et al.*, Experimental Study on Visual Detection for Fatigue of Fixed-Position Staff, *Applied Ergonomics*, *65* (2017), Nov., pp. 1-11
[11] Yao, S., *et al.*, Eye State Detection Method Based on LBP, *Application Research of Computers*, *32* (2015), 6, pp. 1897-1900
[12] Li, M., *et al.*, Structure-Revealing Low-Light Image Enhancement via Robust Retinex Model, *IEEE Transactions on Image Processing*, *27* (2018), 6, pp. 2828-2841
[13] LeCun Y, *et al.*, Deep Learning, *Nature*, *521* (2015), 7553, pp. 436-444
[14] Fu, X., *et al.*, A Probabilistic Method for Image Enhancement with Simultaneous Illumination and Reflectance Estimation, *IEEE Transactions on Image Processing*, *24* (2015), 12, pp. 4965-4977
[15] Di, W., *et al.*, Driver Eye Feature Extraction Based on Infrared illuminator, *Journal of Transport Information & Safety*, *27* (2009), 1, pp. 79-82
[16] Chen, X., *et al.*, Vehicle Detection in Satellite Images by Hybrid Deep Convolutional Neural Networks, *IEEE Geoscience and Remote Sensing Letters*, *11* (2014), 10, pp. 1797-1801
[17] Gu, J., *et al.*, Recent Advances in Convolutional Neural Networks, *Pattern Recognition*, *77* (2018), 12, pp. 1620-1634
[18] Zhang, K., *et al.*, Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks, *IEEE Signal Processing Letters*, *23* (2016), 10, pp. 1499-1503
[19] Song, F., *et al.*, Eyes Closeness Detection from Still Images with Multi-Scale Histograms of Principal Oriented Gradients, *Pattern Recognition*, *47* (2014), 9, pp. 2825-2838
[20] Zhu, X., Goldberg. A. B., Introduction to Semi-Supervised Learning, *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *3* (2009), 1, pp. 1-130
[21] Graves, A., *et al.*, A Novel Connectionist System for Unconstrained Handwriting Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31* (2008), 5, pp. 855-868
[22] Liu, W., *et al.*, Convolutional Two-Stream Network Using Multi-Facial Feature Fusion for Driver Fatigue Detection, *Future Internet*, *11* (2019), 5, 115
[23] Kuang, M. H., *et al.*, A Hybrid Deep Learning Approach for Sentiment Analysis in Product Review, *Facta Univ.-Ser. Mech.*, *21* (2023), 3, pp. 479-500

Paper submitted: November 23, 2023
Paper revised: May 11, 2024
Paper accepted: May 12, 2024