Driver Emotion Recognition and Regulation Based on Artificial Intelligence

Jiasheng Zhang

{ zhangjiasheng1990810660@outlook.com}

Department of Mathematics, Jinzhong College, Jinzhong, Shanxi, China

Abstract. Emotion recognition is one of the key research topics in contemporary studies. Emotion recognition is one of the important research directions in the current fields of human-computer interaction and emotional computing. The accurate perception and modeling of emotional states have been extensively studied at the algorithm and multimodal integration levels. However, compared to the emotion recognition method itself, its systematic application in driving scenarios is still insufficient. Especially in converting the recognition results into effective emotional regulation intervention strategies to further ensure driving safety, there is still a lack of research. This paper first reviews the methods based on single-modal and multi-modal emotion recognition, then summarizes the influence of different modalities on emotional regulation, identifies the application gap in the intelligent cabin field, and finally designs a scheme to regulate the negative emotions of drivers.

Keywords: Multimodal emotion recognition; Negative emotions; Emotion regulation; Intelligent cockpit.

1 Introduction

With the rapid development of intelligent cabin technology, driver state monitoring is becoming a key path to enhance road traffic safety. The "China Statistical Yearbook 2024" shows that there were approximately 200,000 road traffic accidents across the country in 2023. Studies have shown that the cognitive state and emotional fluctuations of drivers are deep and crucial factors affecting driving behavior [1, 2]. Therefore, through real-time identification and implementation of corresponding dynamic adjustments of the vehicle environment, and regulating the negative emotions of drivers, has become an important research direction for intelligent cabins and human-computer interaction. This is not only a targeted remedy for existing safety shortcomings, but also a paradigm shift from "passive response" to "active care", providing technical possibilities for fundamentally reducing the incidence of traffic accidents.

The current research paths for emotion recognition mainly include those based on multi-layer perceptron classifiers for recognizing voice emotions [3] and driver abnormal facial emotion recognition based on improved residual networks [4] (single-modal emotion recognition) and

ICIAAI 2026, May 22-24, Singapore, Singapore

DOI 10.4108/eai.22-5-2026.2365092

emotion recognition based on facial expressions and electroencephalogram signals (multi-modal emotion recognition) [5]. Although multi-modal methods demonstrate certain advantages in terms of recognition accuracy and stability, most studies still remain at the "emotion recognition" level, with relatively insufficient discussion on how the recognition results can serve subsequent emotion regulation decisions.

Emotion recognition is the prerequisite for emotion intervention and regulation. The ultimate goal of emotion recognition is to achieve timely and accurate control of emotions through them, to avoid adverse consequences caused by the inability to timely regulate emotions. Emotion recognition is the perception layer, while emotion regulation is the decision and action layer. Together, they form a closed-loop system oriented towards driving safety. Therefore, this paper systematically reviews the application status of single-modal and multi-modal emotion recognition methods in intelligent cabins, and on this basis, analyzes the potential impacts of different modalities on emotion regulation, summarizes the key problems and application gaps in current research, and accordingly proposes a design idea for negative emotion regulation of drivers, providing a reference for subsequent emotion intervention research.

2 Driver emotion recognition and detection based on single-modal information

Monomodal refers to a system or method that, during information processing, relies solely on a single type of data (such as facial expressions, speech, and electroencephalogram signals, etc.) for analysis and decision-making.

Li et al. first induced specific emotions in drivers through visual and auditory stimuli in a driving simulator, collected their natural driving state facial video data, and used MTCNN for face detection and alignment in complex driving environments, and extracted effective emotional expression segments to construct a dedicated facial expression dataset for driving scenarios (DEFE), which adopts a two-dimensional emotion annotation system. Subsequently, based on this dataset, the study established baseline classification models for dimensional emotions and discrete emotions, and through action unit analysis, it was found that the driving task significantly inhibited facial emotional expression [6].

Pan Dongjie et al. proposed a model based on the improved RepVGG network for the task of face emotion recognition. By introducing the ECA attention mechanism to enhance feature extraction capabilities and combining Softmax loss and center loss to optimize classification performance. During training, a joint optimizer and dynamic learning rate scheduling strategy were used, and the LAHE technology was employed to preprocess the CK+, FER2013, and RAF-DB datasets to enhance local contrast. Finally, through comprehensive evaluation, it was verified that this method is effective in improving the accuracy, robustness, and applicability of emotion recognition [4].

In practical applications, monomodal data may be interfered with by environmental noise. Emotion, as a complex physiological phenomenon, the recognition results obtained based on monomodal data are often unreliable and the monomodal emotion recognition model has poor generalization ability across different individuals. To improve the accuracy, robustness, and

applicability of emotion recognition, the use of multimodality for emotion recognition has gradually become widely adopted.

3 Driver emotion recognition and detection based on multimodal information

Multimodal emotion recognition refers to a method that integrates information from multiple modalities (such as speech, electroencephalogram signals, facial expressions, etc.) for fusion, thereby comprehensively determining the emotional state of a person. Compared with single-modal emotion recognition, multimodal emotion recognition can capture emotional information more comprehensively, thereby improving the accuracy and robustness of emotion recognition.

Jin Zeyu et al. designed an audio-visual emotion recognition system for driving scenarios. By extracting audio MFCC features and facial visual features, and introducing multi-head attention mechanism for cross-modal fusion to capture local and global correlation information, a convolutional model combined with a two-layer LSTM structure model was adopted to enhance the temporal modeling ability. Finally, through the fully connected layer, emotion classification was achieved, with a focus on the identification of negative emotions during driving. This research was verified on the RAVDESS and eNTERFACE'05 datasets, and through ablation and comparative experiments, the effectiveness and robustness of the fusion method were proved [5].

Zhou et al. proposed a driver emotion recognition method based on multimodal fusion. By integrating three modal data types: speech, facial images, and video sequences, features were extracted using mixed MFCC, HOG, and circular LBP algorithms respectively, and were fused at the feature level. A deep learning model combining CNN, bidirectional LSTM, and a hybrid attention module was constructed to adaptively focus on key features and capture spatio-temporal correlations. The model was verified on a self-built multimodal dataset combining RAVDESS, CK+, and NTHU datasets, achieving high-precision recognition of six emotions with an accuracy rate of 85.52%, and its effectiveness and robustness were proved through comparative experiments and generalization tests with the FER2013 dataset [7].

4 Discussion and Solutions

This study previously discussed and analyzed research based on single-modal emotion recognition and identified its shortcomings, subsequently further explored multi-modal emotion recognition technology. With the development of emotion recognition technology, driver state monitoring systems have been able to effectively identify emotions through multiple modalities, but they have not been well applied in intelligent cabins. After identifying a specific emotion, how should the system effectively regulate the driver's negative emotions? This key aspect still lacks systematic scheme design. Therefore, converting the cutting-edge emotion recognition achievements into a reliable, safe, and user-friendly takeover strategy can further promote the development of intelligent cabins, which is also the core key to improving driving safety and human-machine interaction experience. Based on this, this study designed a

set of solutions to achieve systematic functional integration from emotion perception to intervention response, enhancing driving safety and human-machine collaboration experience.

Li Zhengsheng et al. proposed that the color of ambient lights has a soothing effect on driving emotions [8], and Wang et al. suggested that breathing lights can help reduce the physiological arousal level after emotional stimulation [9], ultimately alleviating the problem of driving safety to a certain extent. If an angry emotion is identified, the main color of the ambient lights should be switched to blue or cold white, and the lights should adopt a slow, regular breathing gradient mode; if a sad or depressed emotion is identified, the main color of the ambient lights should be switched to warm orange or daylight yellow, and the lights should adopt a soft, stable constant-on mode; if an anxious and uneasy emotion is identified, the main color of the ambient lights should be switched to soft green or light gold. The lights should be set to a completely static constant-on state.

Aromatherapy has been widely applied in multiple fields for emotion regulation. Therefore, introducing an aroma system to regulate emotions in emotion regulation is considered by Tang et al. after careful consideration of the popularity, comprehensiveness, and reliability of the adjustment effect [10]. After analysis and comparison, the optimal solution for regulating abnormal emotions such as depression in drivers is the essential oil of sandalwood. Therefore, when an anxious or angry emotion is identified, the aroma system will release the aroma of sandalwood; Dmitrenko et al. in exploring the role of olfactory stimulation in the driving environment found that pleasant smells (such as rose and mint) may shift the driver's mood to a positive one, and when a low or negative emotion is identified, the aroma system will release the aroma of mint [11].

Yu et al. studied that when drivers have negative emotions, empathetic voice interaction can provide emotional support and positive feedback, through emotional resonance and emotional transmission, assisting drivers in achieving cognitive reevaluation, thereby significantly alleviating their negative emotions during driving [12]. Wu et al. concluded that personalized voice assistants can better enhance the trust of drivers, such as using the voice of close people, an obedient voice, and humorous voice, which can provide more emotional support to drivers and help them better regulate their emotions [13].

Therefore, based on the physical adjustment of ambient lights and aroma systems, introducing empathetic voice for assistance, combined with NOV assisted driving, jointly regulates the negative emotions of drivers. This design successfully constructed an "emotion-adaptive" intelligent cabin system (Figure 1) based on multi-modal emotion recognition, using multi-sensory collaborative regulation and intelligent vehicle takeover as means, with the core goal of ensuring driving safety and improving the experience. The ambient lights, aroma system, empathetic voice, and NOV assistance are organically coordinated. The three based on the intervention targets of different emotions and human perception characteristics form a dynamic coordination with a primary and secondary relationship and a time sequence.

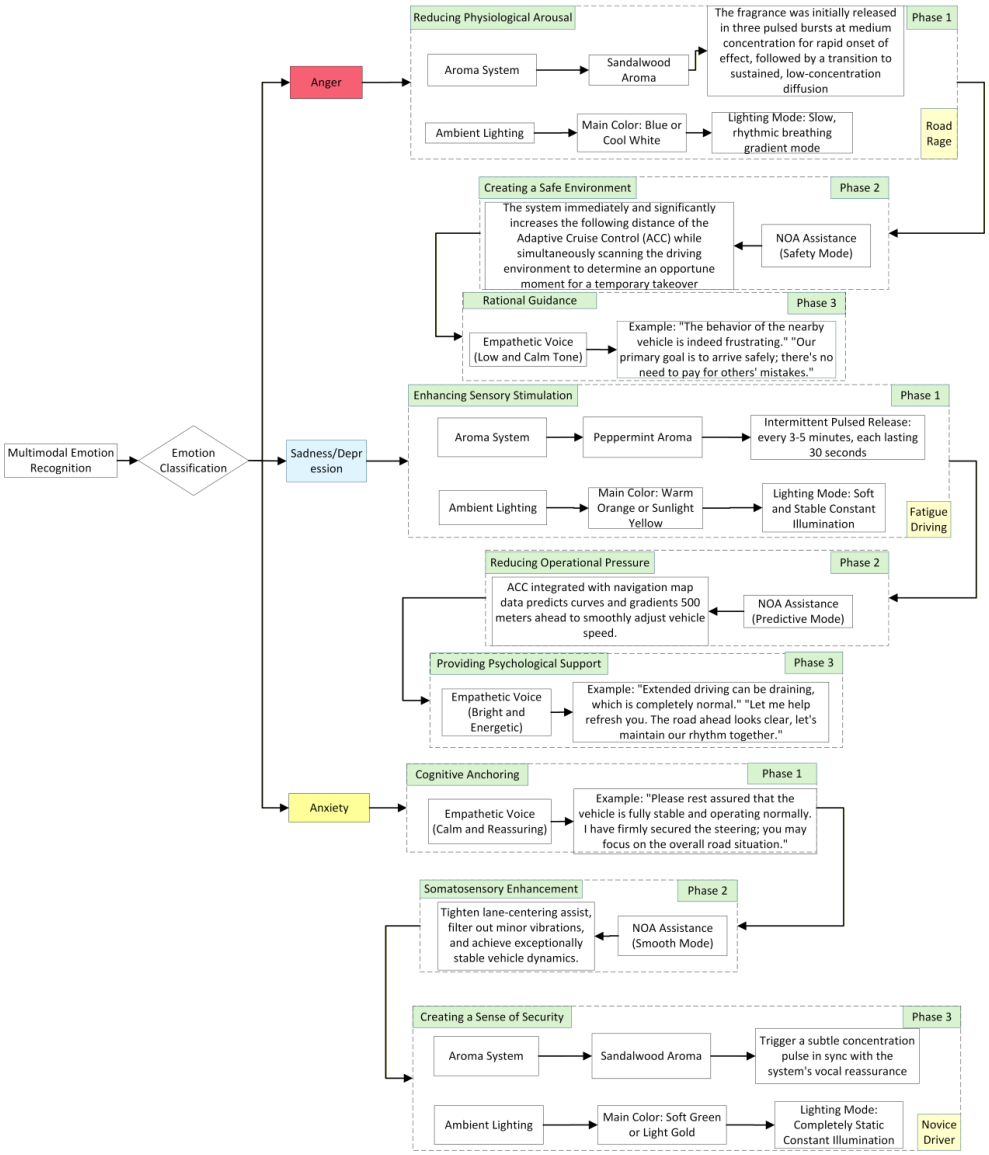

Fig. 1. Flowchart of the negative emotion regulation plan for drivers

5 Conclusion

This article reviews the current research status of single-modal and multi-modal emotion recognition. It summarizes the characteristics of different methods and their recognition effects. Since single modality always lacks diversity, it is impossible to fully understand complex human emotions through a single modality signal. In contrast, the application of multi-modal has more advantages. Then, from the aspects of vision, olfaction, and empathetic

speech, it reviews the influence of this modality on emotion regulation. Finally, based on the above review, a systematic plan is designed to intervene and regulate the negative emotions of drivers. The intelligent cabin field is developing rapidly nowadays. Applying multi-modal emotion recognition to this field can not only further promote the development of the intelligent cabin but also provide a possibility for reducing the incidence of traffic accidents.

References

[1] Dukes, R. L., Clayton, L. T., Miller, T. L., Rodgers, S. E.: Effects of aggressive driving and driver characteristics on road rage. The Social Science Journal. Vol. 38, No. 2, pp. 323–331 (2001)

[2] Vaa, T.: Some comments on the definition of aggression and aggressive driving behaviour. In: Proceedings of the Road Safety on Three Continents Conference, Pretoria, South Africa, pp. 181–195 (2000)

[3] Alnuaim, A. A., Zakariah, M., Shukla, P. K., Alhadlaq, A., Hatamleh, W. A., Tarazi, H., Sureshbabu, R., Ratna, R.: Human–computer interaction for recognizing speech emotions using multilayer perceptron classifier. Journal of Healthcare Engineering. Article 6005446 (2022)

[4] Pan, D., Qin, X., Cui, Y., Wang, X.: Driver abnormal emotion recognition based on an improved residual network. Computing Technology and Automation. Vol. 44, No. 2, pp. 121–127 (2025)

[5] Jin, Z.: Research on an audiovisual fusion-based driver emotion recognition system and its applications. Master's thesis, Sichuan Normal University (2025)

[6] Li, W., Cui, Y., Ma, Y., et al.: A spontaneous driver emotion facial expression (DEFE) dataset for intelligent vehicles. arXiv Preprint. arXiv:2005.08626 (2020)

[7] Zhou, D., Cheng, Y., Wen, L., Luo, H., Liu, Y.: Drivers' comprehensive emotion recognition based on HAM. Sensors. Vol. 23, No. 19, Article 8293 (2023)

[8] Li, Z., Xing, W., He, C., et al.: Design study on the soothing effects of ambient light color on driving emotions. Chinese Journal of Ergonomics. Vol. 27, No. 2, pp. 14–20 (2021)

[9] Wang, J., Lu, J., Xu, Z., Wang, X.: When lights can breathe: Investigating the influences of breathing lights on users' emotion. International Journal of Environmental Research and Public Health. Vol. 19, No. 20, Article 13205 (2022)

[10] Tang, B., Zhu, M., Deng, L., Hu, Z., Chen, B.: Positive regulation of drivers' negative emotions based on olfactory stimulation. Journal of Automotive Engineering. Vol. 13, No. 6, pp. 832–843 (2023)

[11] Dmitrenko, D., Maggioni, E., Brianza, G., et al.: Caroma therapy: Pleasant scents promote safer driving, better mood, and improved well-being in angry drivers. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. ACM, pp. 1–13 (2020)

[12] Zhao, Y., Jiang, K., Yu, Z., Huang, Z.: Emotion-matched in-vehicle assistants: An exploration in regulating drivers' incidental emotions and enhancing takeover performance and situational awareness. Cognition, Technology & Work. Vol. 26, No. 2, pp. 225–245 (2024)

[13] Wu, L., Xue, J., Li, W., Wang, K., Zhang, X., Guo, G.: Toward decreasing the driving risk: Speech-based driver's anger regulation in smart cockpit. IEEE Journal of Radio Frequency Identification. Vol. 6, pp. 764–768 (2022)