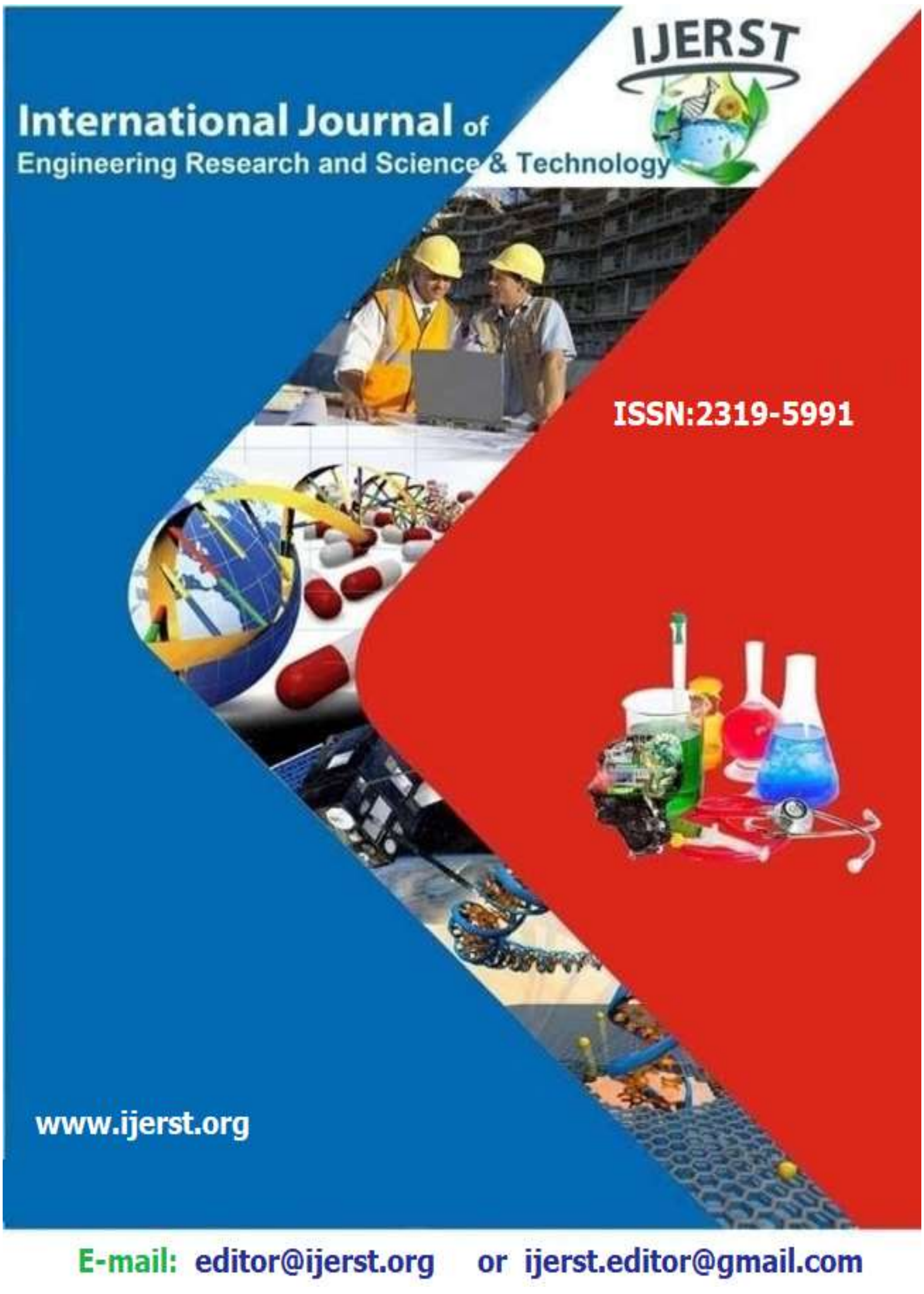


E-mail: editor@ijerst.org or ijerst.editor@gmail.com

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

A MULTIMODAL DRIVER ANGER RECOGNITION METHOD BASED ON CONTEXT AWARENESS

[1]MD.OBAIDULLAH KHAN, [2]MD.ADNAN, [3]G.MAHESH, [4]SYED.FAIZAN, [5]P. VINAY REDDY
[1, 2, 3, 4, 5]Department of Computer Science and Engineering, [5]Assistant Professor
[1, 2, 3, 4, 5]Vijay Rural Engineering College, Manik Bhandar, Nizamabad-503003

Abstract: In modern traffic situations, driver aggression presents a significant risk to road safety, requiring specific real-time emotion recognition technologies. A unique Context-aware Multi-modal driver Emotion identification (CA-MDER) method is supplied to efficaciously address great issues in multi-modal emotion identity, such as individual variability and contextual sensitivity in emotional expression underneath various driving eventualities. This approach combines facial expressions, verbal indicators, and using status signals to facilitate thorough rage identification. It utilizes an attention Mechanism-enhanced Depthwise Separable Convolutional Neural network (AM-DSCNN) for effective feature extraction, together with improved support Vector Machines (SVM) and Random forest (RF) classifiers for reliable classification. A Context-aware Reinforcement learning (CA-RL) paradigm is utilized for adaptive weight allocation, facilitating dynamic decision-level integration across modalities according to contextual importance. Furthermore, CNN-based transfer learning methodologies are hired to enhance model generalization and classification precision. The system demonstrates incredible performance, with an overall accuracy of 91.68% and an F1 score of 90.37%, indicating its proficiency in accurately identifying and categorizing angry emotions across various data modalities. The implementation has full-size promise for improving using safety by facilitating accurate emotion identification via sophisticated deep learning and adaptive fusion techniques.

***"Keywords** - Context-awareness, driving state emotion recognition, emotional expression heterogeneity, multimodal emotion recognition, machine learning".*

1. INTRODUCTION

The fast advancement of human-computer interaction and smart transportation systems has rendered the incorporation of biometric technology into driver emotion detection an important detail for enhancing using safety and experience. Emotion detection algorithms, utilizing multimodal data along with facial expressions, vocal tone, physiological indications (e.g., heart rate, skin conductance), and using behavior styles, are progressively being used to assess drivers' psychological states in real time [1]. Among many emotional states, rage is notably significant due to its profound effect on driving performance and decision-making. Studies indicates that emotional disturbances inclusive of rage may preclude attention, perception, and cognitive functioning, thereby raising the likelihood of aggressive or risky driving behavior [2].

Angry driving, additionally referred to as road rage, has emerged as an everyday and regarding trouble in numerous countries. Studies in China exhibits that over 68% of drivers have faced road rage, positioning fury as one of the most popular emotions while driving. Parkinson [4] asserts that individuals have a markedly higher propensity for rage while using as compared to different daily occasions. Environmental stresses, consisting of traffic congestion, time constraints, and interpersonal disputes while driving, serve as sizable catalysts for driver rage [5]. Furthermore, anger impairs the driver's cognitive processing, modifies their threat belief, prolongs reaction time, and encourages impulsive actions such as tailgating, speeding, and beside the point lane adjustments [6].

Figuring out and addressing driver rage in its first phases is crucial for alleviating its poor consequences. Prompt emotion popularity allows the implementation of treatments together with auditory notifications, adaptive cruise control, or automatic vehicle behavior modifications to mitigate emotional intensity [7]. Unimodal techniques, which depend completely on a single data source, have shown moderate effectiveness but frequently lack resilience because of noise or ambiguity inherent in a single channel. Consequently, new studies is focusing on multimodal emotion identity, which amalgamates

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

many input streams to enhance accuracy and contextual focus [8].

This studies provides a context-aware modeling technique for multimodal driver rage identification in reaction to recent improvements. Its objective is to enhance the reliability of emotion popularity and to advance the protection and adaptableness of smart transportation systems.

2. LITERATURE REVIEW

Comprehending the emotional situation of drivers, in particular rage, is critical for promoting road safety and growing superior in-vehicle technology. numerous research have advanced this field with the aid of examining the influence of driving force rage on conduct and creating emotion detection models making use of several modalities, which include facial expressions, vocal alerts, physiological reactions, and contextual using facts.

Techer et al. [9] examined the effect of rage on driving performance utilizing neurophysiological data. They hired electroencephalographic (EEG) facts to assess the modulation of attention while riding underneath the effect of rage. Their findings emphasised that emotional states, specifically rage, drastically affect attention tiers and driving capacity. The addition of EEG statistics was crucial in presenting goal insights into the cognitive and emotional alterations experienced by furious drivers, therefore facilitating the incorporation of physiological signals into multimodal emotion identification systems.

Precht, Keinath, and Krems [10] applied naturalistic driving data to look at the real impacts of rage on driving force behavior. Their studies proven that anger precipitates competitive driving behaviors, along with increased velocity, much less headway, and frequent lane alterations. This good sized behavioral investigation verified the concrete risks related to using rage, highlighting the necessity for in-vehicle monitoring systems that can discover and cope with feelings in real-time.

Zhong, Hong, and Yuan [11] completed an experimental research to assess the impact of anger on driving behavior. Their research assessed measures which includes response time, acceleration, and steering stability, demonstrating that anger diminishes cognitive control and effects in extra hazardous judgments while driving. This observe substantiates the significance of emotional control in autonomous using systems and affirms the need of including behavioral measures in rage detection frameworks.

Similarly, Lei [12] examined the traits of aggressive using and its effects for traffic safety. The have a look at supplied an extensive exam of the behavioral expressions of rage, together with horn blowing, tailgating, and overtaking, emphasizing their importance in traffic accidents. This research highlights the want of making context-aware algorithms that not only identify emotions but additionally recognise their behavioral implications.

Hasan and Islam [13] proposed a way for emotion detection from Bengali speech through recurrent neural network (RNN) modulation-based classification within the field of voice emotion recognition. Their methodology confirmed the efficacy of RNNs in processing consecutive audio input and categorizing emotional states with commendable precision. This examine is essential for multilingual driver emotion identity, since vocal cues are crucial for identifying non-visual emotional expressions.

Kwon [14] introduced the CLSTM model, a hierarchical convolutional long short-term memory (ConvLSTM) network, for deep feature-based speech emotion identity. This design efficaciously captured spatial and temporal records in audio alerts, rendering it extremely appropriate for actual-time packages which include in-vehicle emotion recognition. The hierarchical structure enhanced the model's capacity to detect nuanced fluctuations in voice tone and pitch, which are essential symptoms of rage and other emotions.

Sun, Fu, and Wang [15] supplied a hybrid model combining decision tree and SVM, which integrates Fisher feature selection for the motive of speech emotion identification. Their method emphasized the selection of the maximum pertinent elements to enhance classification accuracy and limit computing complexity. This approach is well-perfect to the desires of embedded systems in motors, where real-time processing and performance are paramount.

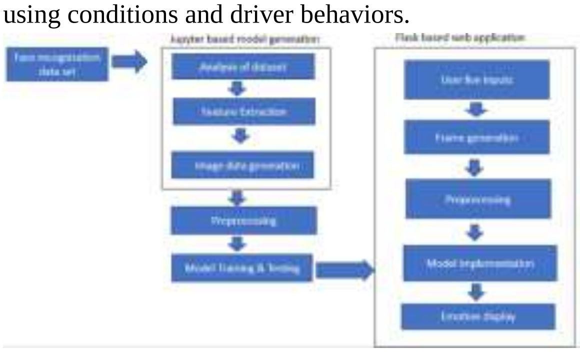
Mao et al. [16] encouraged the usage of convolutional neural networks (CNNs) to extract

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

giant characteristics for voice emotion identity. Their research revealed that CNNs may proficiently discern discriminative characteristics from unprocessed audio inputs without vast pre-processing. The utilization of CNNs in this area has created new possibilities for incorporating deep learning into multimodal emotion identity systems, especially for managing the noisy environment characteristic of using settings.

3. MATERIALS AND METHODS

The proposed Context-aware Multi-modal driver Anger Emotion Recognition (CA-MDER) system amalgamates face traits, speech inputs, and using condition data to provide thorough emotion identification. It employs “Attention Mechanism-Depthwise Separable Convolutional Neural Networks (AM-DSCNN)” for efficient feature extraction and integrates enhanced “Support Vector Machines (SVM) and Random forest (RF) for precise emotion categorization”. “Context-aware Reinforcement learning (CA-RL)” is applied for adaptive weight allocation in decision-level fusion, facilitating dynamic change in keeping with real-time driving instances. The system utilizes CNN-primarily based transfer learning models, consisting of “XceptionNet, ResNet, and DenseNet”, to improve face emotion identification. Data augmentation approaches are employed to enhance the training set's diversity, facilitating improved generalization across numerous using conditions and driver behaviors.



“Fig.1 Proposed Architecture”

The system architecture (fig.1) consists of two primary phases: model introduction with Jupyter and actual-time emotion identity through a Flask-based web application. To start with, a face recognition dataset is subjected to analysis, feature extraction, image facts production, and preprocessing prior to model training and testing. The trained model is subsequently included into a Flask application, where it handles actual-time consumer inputs via creating frames, doing preprocessing, and utilizing the discovered model to identify emotions. The ultimate emotional situation is exhibited, facilitating real-time face emotion identification.

A) Dataset Collection:

The dataset for the emotion detection system is derived from a heterogeneous Face detection Dataset including face images taken underneath various driving situations. This dataset is critical for training the model to identify rage through facial expressions, encouraged by diverse emotional states while driving. The collection guarantees variation and encompasses a wide spectrum of using instances, making sure the version generalizes effectively across diverse drivers and surroundings.

B) Preprocessing Steps:

Normalization: Normalization normalizes input data by normalizing pixel values to the same scale [0, 1] for ensuring that all images have the same image scale. This phase reduces the effect of variation of light or other factors thus, making the model more robust and keeping what features have on an even scale for easy learning.

Augmentation: Data augmentation artificially augments the size and diversity of the data set by such transformations as rotations, flips and crops. Such alterations create additional pattern range, which improves the models generalization capability and mitigates overfitting. Augmentation mimics actual world alterations to enable the model to support many facial orientations and emotions.

Feature Extraction: Feature extraction is extraction of the most salient facial features and includes the eyes, lips, and other regions of importance to identification of anger. This simply reduces the complication of the data, accentuating the maximum sizeable visual indications. Extracted functions increase model overall performance by allowing it to focus on the data that is most important to emotional state categorization.

C) Training and Testing:

The following technique for model training calls for use of appropriate processed and more favorable data to train deep learning algorithms including Convolutional Neural Networks (CNNs). The model has been trained repeatedly with monitoring of its

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

performance on a separate dataset of test set. Evaluation criteria such as accuracy, precision, and keep in mind, a set of guidelines are used to determine can the model generalize to new data. The model is further tested rigorously post training to ensure that it is robust before being integrated.

D) Algorithm:

Attention Mechanism-Depthwise Separable Convolutional Neural Networks (AM-DSCNN): AM-DSCNN is a sub-article. The enhanced emotion identification of the AM-DSCNN is informed by optimal face extraction through depthwise separable convolutions while being computationally efficient. The attention mechanism allows the model to focus on important face areas and, therefore, recognize angry emotions in various surroundings better. This model architecture is improved in terms of feature acquisition at reduced resources consumption.

Support Vector Machines (SVM): Once features are extracted, SVM are used for classification with excellent results for binary classification problems. It defines the best hyperplane that is able to differentiate between furious and not angry emotion, providing impressive precision. This approach combines multimodal data so as to improve rage detection through identifying subtleties of speech and behavior.

Random Forest (RF): RF is used for emotion categorization as it can cope with noisy input and generate exact predictions. It works by generating multiple decision trees and outputs the dominant class, through this making it resistant to overfitting. This classifier is particularly helpful in rage identification because of the range of emotional characteristics which include facial expressions, vocalisations and driving situations.

Context-Aware Reinforcement Learning (CA-RL): CA-RL increases weight distribution in decision-level fusion by considering the contextual antecedents of emotion expression. This approach permits dynamic modification of the effect of each of facial, verbal, or driving states according to the real-time driving environment and is capable of adaptive, context- sensitive rage detection enhancing model efficacy in many situations.

CNN-based Transfer Learning (XceptionNet, ResNet, DenseNet): "The transfer learning with CNN architectures XceptionNet, ResNet, and DenseNet" increases recognition accuracy by using pre-trained models. These models which have been optimized to identify emotions will easily pull precise features from one's face photo, thus speeding up the training process and increasing the model's ability to generalize new, unseen data.

4. EXPERIMENTAL RESULTS

"Fig.2 Home Page and Upload Input Audio"

An ablation experiment was conducted to assess the efficacy of the context-aware module and multimodal fusion in the proposed model, which involved the exclusion of context awareness and the integration of individual modality data. M denotes the multimodal task, F signifies the facial single modality, V indicates the voice single modality, S refers to the driving state single modality, and Awareness pertains to the context-awareness module. The findings of the ablation experiment for each indication are presented in Fig. 3.

M	F	V	S	Awareness	ACC(%)	F1(%)
√	×	√	√	√	72.53	72.12
√	√	×	√	√	88.43	86.75
√	√	√	×	√	87.57	87.31
√	√	√	√	×	86.72	84.91
√	√	√	√	√	91.68	90.37

"Fig.3 Results of Modality Ablation Experiment: Multimodal Dataset"

This work introduced a context-aware multimodal method for detecting driver rage emotions (CA-MDER) which combines facial expressions, vocal cues, as well as vehicle driving status information to

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

increase accuracy. At first, the detection of anger as an emotion was done on a per-individual modality basis. First or we introduced a face emotion identification approach, which used Attention Mechanism Deep Separable Convolutional Neural Network (AM-DSCNN), to boost the accuracy of face emotion identification. This approach highlights critical face features, which indicate emotion with an attention module, then enhances the computing efficiency incorporating deep separable convolution modules.

5. CONCLUSION

The CA-MDER context-aware multi-modal driver anger emotion recognition approach offers a resilient and intelligent alternative to detecting driver's anger based on an amalgamation of facial, vocal, and driving state data. Utilizing the leveraging "Attention Mechanism-Depthwise Separable Convolutional Neural Networks (AM-DSCNN)" for the extraction of facial emotion characteristic, the proposed system ensures high efficiency and low computational overhead. The integration of the better "Support Vector Machines (SVM) and Random Forest (RF)" improves the category accuracy in exceptional emotional phases. Although "Adaptive vector weighted sum (AVWS)" weights provide the wanted adaptive weighting scheme, CA-RL permits the adaptive weight distribution from the choice-level fusion, so that the version dynamically focuses on modalities dependent on real-time driving conditions. Moreover, integrated "CNN based transfer learning models (XceptionNet, ResNet, and DenseNet)", enhanced facial emotion recognition considerably. Data augmentation strategies enhance the dataset size thus enabling the version to study off numerous set of using and emotional situation. The system carried out a notable "accuracy of 91.68% and an F1 score of 90.37%," demonstrating its effectiveness in handling individual variability and contextual challenges in emotion recognition. Universal, CA-MDER proves to be an effective, adaptive, and context-aware solution, able to enhancing road safety through unique and reliable multi-modal emotion detection.

The future scope of the CA-MDER device consists of integrating extra modalities along with physiological indicators (e.g., heart rate, EEG) to similarly enhance emotion recognition accuracy. Enhancing the system with advanced deep learning models like vision Transformers and Graph Neural Networks can improve feature representation. Real-time deployment in diverse driving environments, including excessive weather or excessive-traffic conditions, will increase robustness. Additionally, incorporating continual learning strategies will allow the system to evolve to evolving driver behaviors and emotional patterns over time.

REFERENCES

[1] J. Tang, Z. Ma, K. Gan, J. Zhang, and Z. Yin, ''Hierarchical multimodal fusion of physiological signals for emotion recognition with scenario adaptation and contrastive alignment,'' *Inf. Fusion*, vol. 103, Mar. 2024, Art. no. 102129, doi: 10.1016/j.inffus.2023.102129.

[2] D. Zhou, Y. Cheng, L. Wen, H. Luo, and Y. Liu, ''Drivers' comprehensive emotion recognition based on HAM,'' *Sensors*, vol. 23, no. 19, p. 8293, Oct. 2023, doi: 10.3390/s23198293.

[3] J. Ni, W. Xie, Y. Liu, J. Zhang, Y. Wan, and H. Ge, ''Driver emotion recognition involving multimodal signals: Electrophysiological response, nasal-tip temperature, and vehicle behavior,'' *J. Transp. Eng., A, Syst.*, vol. 150, no. 1, Jan. 2024, Art. no. 04023125, doi: 10.1061/jtepbs.teeng7802.

[4] G. Du, Z. Wang, B. Gao, S. Mumtaz, K. M. Abualnaja, and C. Du, ''A convolution bidirectional long short-term memory neural network for driver emotion recognition,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 7, pp. 4570–4578, Jul. 2021, doi: 10.1109/TITS.2020.3007357.

[5] F. Wang, ''Research on driver anger emotion recognition method based on multimodal fusion,'' M.S. thesis, Jilin Univ., 2023.

[6] X. Yu, ''Driver's anger recognition method based on human-vehicle environment information fusion,'' M.S. thesis, Shandong Univ. Technol., 2021.

[7] P. Wang, ''Study on multimodal recognition method of Driver's anger emotion and mechanism of driving risk under anger emotion,'' M.S. thesis, Chongqing Univ., 2020.

[8] S. Shafaei, T. Hacizade, and A. Knoll, ''Integration of driver behavior into emotion recognition systems: A preliminary study on steering wheel and vehicle acceleration,'' in *Proc. Asian Conf. Comput. Vis.*, G. Carneiro and S. You, Eds.,

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

Cham, Switzerland: Springer, 2019, pp. 386–401, doi: 10.1007/978-3-030-21074-8_32.

[9] F. Techer, C. Jallais, Y. Corson, F. Moreau, D. Ndiaye, B. Piechnick, and A. Fort, ‘‘Attention and driving performance modulations due to anger state: Contribution of electroencephalographic data,’’ *Neurosci. Lett.*, vol. 636, pp. 134–139, Jan. 2017.

[10] L. Precht, A. Keinath, and J. F. Krems, ‘‘Effects of driving anger on driver behavior—Results from naturalistic driving data,’’ *Transp. Res. F, Traffic Psychol. Behav.*, vol. 45, pp. 75–92, Feb. 2017.

[11] M. Zhong, H. Hong, and Z. Yuan, ‘‘Experiment research on influence of angry emotion for driving behaviors,’’ *J. Chongqing Univ. Technol. Natural Sci.*, vol. 25, no. 10, pp. 6–11, 2011.

[12] H. Lei, ‘‘The characteristics of angry driving behaviors and its effects on traffic safety,’’ M.S. thesis, Wuhan Univ. Technol., 2011.

[13] H. Hasan and Md. A. Islam, ‘‘Emotion recognition from Bengali speech using RNN modulation-based categorization,’’ in *Proc. 3rd Int. Conf. Smart Syst. Inventive Technol. (ICSSIT)*, Aug. 2020, pp. 1131–1136. Accessed: Dec. 19, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9214196/

[14] S. Kwon, ‘‘CLSTM: Deep feature-based speech emotion recognition using the hierarchical ConvLSTM network,’’ *Mathematics*, vol. 8, no. 12, p. 2133, Nov. 2020.

[15] L. Sun, S. Fu, and F. Wang, ‘‘Decision tree SVM model with Fisher feature selection for speech emotion recognition,’’ *EURASIP J. Audio, Speech, Music Process.*, vol. 2019, no. 1, p. 2, Dec. 2019, doi: 10.1186/s13636-018-0145-5.

[16] Q. Mao, M. Dong, Z. Huang, and Y. Zhan, ‘‘Learning salient features for speech emotion recognition using convolutional neural networks,’’ *IEEE Trans. Multimedia*, vol. 16, no. 8, pp. 2203–2213, Dec. 2014.

[17] E. M. Albornoz, D. H. Milone, and H. L. Rufiner, ‘‘Spoken emotion recognition using hierarchical classifiers,’’ *Comput. Speech Lang.*, vol. 25, no. 3, pp. 556–570, Jul. 2011.

[18] T. M. Rajisha, A. P. Sunija, and K. S. Riyas, ‘‘Performance analysis of Malayalam language speech emotion recognition system using ANN/SVM,’’ *Proc. Technol.*, vol. 24, pp. 1097–1104, Jan. 2016.

[19] H. Aouani and Y. B. Ayed, ‘‘Speech emotion recognition with deep learning,’’ *Proc. Comput. Sci.*, vol. 176, pp. 251–260, Jan. 2020, doi: 10.1016/j.procs.2020.08.027.

[20] Y. Zhou, J. Li, Y. Sun, J. Zhang, Y. Yan, and M. Akagi, ‘‘A hybrid speech emotion recognition system based on spectral and prosodic features,’’ *IEICE Trans. Inf. Syst.*, vol. E93-D, no. 10, pp. 2813–2821, 2010.