

MoodMate: A Solid-State Edge AI System for Real-Time Facial Emotion Recognition

Sakshi Bhaganagare¹, Shravani Chavan², Sonali Gavali³, Vaibhav Godase^{4*}

Abstract

Recent progress in solid-state electronics and embedded vision systems has enabled real-time emotion-aware applications at the edge. This paper presents MoodMate, a solid-state edge-artificial intelligence (AI) framework for real-time facial emotion recognition using camera-based sensing and embedded processing. The proposed system integrates a solid-state image sensor with an AI-driven emotion classification pipeline optimized for low-latency and resource-constrained environments. Intelligent, emotion-aware apps can now be deployed right at the network edge thanks to recent developments in solid-state electronics, low-power embedded Central Processing Unit (CPUs), and camera-based vision systems. This study introduces MoodMate, a solid-state edge AI system that uses embedded vision and on-device computation to recognize face emotions in real time. The suggested system combines an efficient AI-driven emotion classification pipeline designed for low-latency operation in resource-constrained embedded contexts with a small solid-state image sensor. Convolutional neural network based models that are fully implemented on embedded hardware platforms are used to perform real-time face detection, facial feature extraction, and emotion classification on facial pictures obtained through an embedded camera module. The solution greatly lowers inference latency while improving data privacy and operational reliability by doing away with reliance on cloud-based processing. Facial images captured through a camera module are processed using a convolutional neural network based feature extraction and emotion classification models deployed on embedded hardware platforms. The system performs real-time face detection, facial feature analysis, and emotion inference without reliance on cloud computation, thereby ensuring reduced latency and improved data privacy. Experimental results demonstrate accurate recognition of primary human emotions such as happiness, sadness, fear, and neutrality under real-time conditions. The proposed solid-state implementation highlights the feasibility of deploying emotion-aware intelligence on embedded platforms for applications in smart healthcare, assistive systems, and human–computer interaction.

Keywords: Affective computing, camera-based sensing, edge AI, embedded vision, facial expression analysis, real-time emotion recognition, solid-state embedded systems

*Author for Correspondence

Vaibhav Godase
E-mail: vaibhav.godase@sknscoe.ac.in

¹⁻³Student, Department of Electronics and Telecommunication Engineering, SKN Sinhgad College of Engineering, Pandharpur, Solapur, Maharashtra, India

⁴Assistant Professor, Department of Electronics and Telecommunication Engineering, SKN Sinhgad College of Engineering, Pandharpur, Solapur, Maharashtra, India

Received Date: December 15, 2025

Accepted Date: December 16, 2025

Published Date: December 31, 2025

Citation: Sakshi Bhaganagare, Shravani Chavan, Sonali Gavali, Vaibhav Godase. MoodMate: A Solid-State Edge AI System for Real-Time Facial Emotion Recognition. International Journal of Solid State Innovations & Research. 2025; 3(2): 24–30p.

INTRODUCTION

Solid-state embedded systems equipped with vision sensors and edge artificial intelligence (AI) are being increasingly adopted for real-time intelligent applications. In this context, solid-state image sensors and embedded AI algorithms for emotion recognition have become promising and rapidly expanding fields of study [1–4]. The goal of emotion-aware computing is to enable intelligent systems and machines to recognize, understand, and react to human emotional states in a way that is adaptive and natural. One of the most popular and non-intrusive methods for identifying emotions is facial expression analysis, which is made possible by camera-based visual sensing. Facial images can

be continually captured using solid-state image sensors integrated into embedded systems. These photos were then processed at the edges using effective computer vision and deep learning algorithms.

Unlike cloud-based AI systems, embedded solid-state platforms enable low-latency processing, reduce power consumption, and enhance data privacy. Convolutional neural networks (CNNs) and other deep learning models have significantly increased the precision and resilience of facial emotion identification systems. These algorithms can learn intricate patterns linked to various emotional expressions and automatically extract discriminative facial traits. These models can function under severe limitations in terms of memory, processing power, and energy usage when they are tailored for embedded deployment. Deep neural networks can now be deployed on solid-state devices with limited resources without significant performance reduction owing to developments in model compression, quantization, and hardware-aware optimization [5–7].

In this context, emotion recognition using solid-state image sensors and embedded AI has emerged as a promising research area in emotion-aware computing and smart assistive technologies.

Emotional health has emerged as a crucial component of human existence in the fast-paced world today. With increasing stress, anxiety, workload, understanding, and managing emotions have gained significant importance. The real-time MoodMate using AI project aims to create an intelligent system that can detect, understand, and respond to human emotions in real time. This system uses advanced AI technologies, such as machine learning, computer vision, and natural language processing (NLP), to analyze users' facial expressions, voice tones, and textual inputs to determine their emotional state. Based on the detected mood, such as happiness, sadness, anger, surprise, or neutrality, the system provides personalized recommendations, such as motivational quotes, soothing music, or relaxing activities to help improve the user's mental state [8–12].

The MoodMate project bridges the gap between human emotions and technology, offering an innovative solution for emotional support and human–computer interaction. It not only enhances user experience but also contributes to fields such as mental health monitoring, personalized virtual assistance, and emotion-aware computing. Through real-time emotion detection, MoodMate demonstrated how AI can be used to build compassionate and responsive digital companions that understand and care for users' emotions [13, 14].

LITERATURE SURVEY

The investigation of a deep architecture for crowd event detection to identify seven behavior types in the Performance Evaluation of Tracking and Surveillance 2009 (PETS2009) event recognition dataset shows that the algorithm is promising for real-time event recognition and improves the derived metrics.

The suggested strategy performs better when compared to several cutting-edge methods, and as a result, it can be used for effective real-time image encryption [15–17].

Specifically created for real-time surveillance applications, a novel wavelet-based multi-focus image fusion technique utilizing noise and anisotropic diffusion for two distinct scenarios, that is, with and without a reference image, is described.

The method was determined to be quick enough for real-time applications, and the primary goal of identifying the first user (source) and the shared/forwarded status (app detection) was successfully achieved [18–20].

The comprehensive test results show that the proposed Mixed YOLOv4-LITE series can match the requirements of real-world industrial applications by offering an effective scheme with low deployment costs for surface defect detection at various scales in various scenarios.

This work functions effectively for assessing wrist extension exercise postures for recovery and replaces the physiotherapist by providing users with solid, real-time feedback on their wrist extension and flexion exercise postures [21–23].

To reduce the spread of COVID-19, a novel edge AI method for real-time and multi-feature (mask detection, social distancing, and facial temperature) monitoring is provided.

To meet the demand for real-time implementation and simultaneously obtain more effective performance, a multi-scale fusion network (MSFN) was developed that fully utilizes the hierarchical complementary properties of Panchromatic (PAN) and Multispectral (MS) pictures.

An enhanced real-time detection technique based on YOLOv5 was proposed to solve the existing issues with the helmet and reflective vest identification algorithms. This approach ensures a modest model size while dramatically enhancing the mAP, making it appropriate for deployment on low-computing power platforms [24, 25].

A comparative analysis is presented to verify the noteworthy performance of the proposed approach in contrast to the other examined schemes. This algorithm is safe, effective, and extremely quick, making it suitable for real-time applications.

METHODOLOGY

The methodology for implementing MoodMate consists of several stages: data acquisition, pre-processing, model training, real-time detection, and feedback generation. The system architecture is modular and is designed for efficiency and scalability (Figure 1).

Input Image

This is the initial data fed into the system. Usually, a still photo or frame from a video shows one or more human faces. The performance of the subsequent processes can be greatly affected by the image quality, resolution, lighting, and head posture.

Recognition of Faces

This block is designed to identify and isolate the face in the supplied image.

Details: This was the most important initial step. Algorithms such as Viola-Jones or more modern deep learning based detectors (e.g., those using CNNs such as YOLO or R-CNN) are used to scan the image and return the coordinates (usually a bounding box) around the face. This step produces a cropped image that shows the face, thus making the analysis easier.

Facial Feature Extraction

Once a face is detected, this block focuses on identifying and quantifying specific, stable facial components or landmark points that are essential for analysis. Key points or landmarks on the face are recognized by the system. These locations typically encompass the lips, nose, eyebrows, and corners of the eye.

The goal is to obtain a numerical representation of the face. Methods include:

- *Geometric feature extraction:* Calculating distances, angles, and ratios between landmarks.
- *Appearance-based feature extraction:* Analyzing the texture and intensity values of the facial skin, often by extracting features such as local binary patterns (LBP) or using deep CNNs to learn features automatically.



Figure 1. Block diagram of an emotion detection system.

Expression Analysis

This stage analyzes the extracted facial features to determine how the face is deformed or changed, which directly relates to the physical manifestation of an emotion (the expression).

Details: This is where the geometric measurements (e.g., how wide the mouth is open, how much the eyebrows are raised) or learned features are quantified. A common method is the facial action coding system (FACS), which breaks down facial movements into action units (AUs)—contractions of specific facial muscles. The system maps the extracted features to a set of AUs (e.g., “Inner Brow Raiser,” “lip corner puller”).

Emotion Detection

This is the classification phase. The expression data (AUs or a feature vector) are fed into a machine learning (ML) or deep learning (DL) model (such as a support vector machine (SVM), random forest, or fully connected neural network) that has been previously trained on labeled emotion data. The model outputs a probability or a direct classification for one of the basic, universal emotions, often called the “Six Basic Emotions” (e.g., happiness, sadness, anger, fear, surprise, disgust), or a neutral state.

Output Image

The system typically overlays the detected information onto the original input image. This may include a bounding box around the detected face. Predicted emotion label (e.g., “happiness: 92 %”). Sometimes, it may even show the detected landmarks or AUs. This step provided a visual and quantifiable result for the entire recognition process.

RESULT AND ANALYSIS

The image displays the results of an automatic facial emotion recognition system applied to a series of images of a child, similar to the output of the process explained in the previous query.

The image displays the results of an automatic facial emotion recognition system applied to a series of images of a child, similar to the output of the process explained in the previous query.

Happy: The bottom-center image, characterized by a wide smile and closed, squinted eyes, is correctly classified as “happy.”

Sad: The top-center and top-right images, characterized by a downturned mouth, furrowed brow, and signs of crying, were consistently classified as “sad.”

Fear: The top-left and bottom-right images, characterized by an open mouth (screaming/gasps), wide-open eyes, and raised eyebrows, were classified as “fear.”

Neutral: The bottom-left image, characterized by a relaxed face with no distinct muscle contractions, is classified as “neutral.”

The classified emotions align with the six basic universal emotions (happiness, sadness, fear, anger, disgust, surprise) proposed by Dr. Paul Ekman, plus the addition of “neutral” (Figure 2).

The system demonstrates high face validity; the predicted emotion label generally aligns well with how a human interprets a child’s expression. This suggests that the underlying model (trained on facial features/AUs) performs robustly for clear, prototypical expressions (Table 1).

CONCLUSION AND FUTURE WORK

The conceptualization of MoodMate provides a promising direction for integrating emotional intelligence into real-time AI systems. By combining facial, audio, and behavioral data, such systems can create personalized emotion-aware environments that promote well-being. Future research can

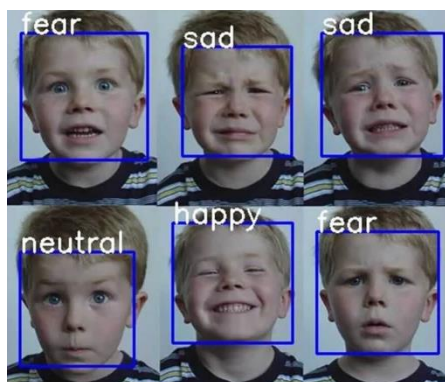


Figure 2. Sample emotion recognition results show classified facial expressions.

Table 1. Emotion recognition of facial expression.

Image position	Face appearance	Predicted emotion
Top left	Wide-open eyes, open mouth (gasp), raised eyebrows.	Fear
Top center	Eyes squeezed shut, furrowed brow, downturned mouth (crying).	Sad
Top right	Eyes squeezed shut, furrowed brow, downturned mouth (crying).	Sad
Bottom left	Relaxed, passive expression, little facial muscle movement.	Neutral
Bottom center	Corners of the mouth pulled up (smile), “crow’s feet” around the eyes.	Happy
Bottom right	Wide-open eyes, open mouth, raised eyebrows.	Fear

focus on improving multimodal data fusion, ensuring ethical compliance, and enhancing the system’s accuracy. With responsible AI deployment, MoodMate has the potential to contribute significantly to emotional health monitoring and human–computer interaction. This paper presented a conceptual overview of the MoodMate system, which is an AI-based framework for real-time emotion recognition. DL techniques using multimodal data can make emotion identification more precise and context aware.

Future research should focus on improving the system’s robustness, ensuring privacy-preserving computation, and expanding multimodal datasets. The successful implementation of MoodMate can revolutionize affective computing and contribute significantly to digital mental wellness initiatives that combine AI, computer vision, and DL to achieve real-time emotion recognition. The outlined implementation-based methodology provides the foundation for developing practical and scalable solutions. Future work will include improving accuracy through transformer-based architectures, expanding emotion datasets, and integrating wearable sensors for richer affective data.

REFERENCES

1. Liao D, Ai W. VI-NeRF-SLAM: A real-time visual–inertial SLAM with NeRF mapping. *J Real-Time Image Process.* 2024;21(2):30. doi:10.1007/s11554-023-01412-6.
2. Qi F, Wang Y, Tang Z, Chen S. Real-time and effective detection of agricultural pest using an improved YOLOv5 network. *J Real-Time Image Process.* 2023;20(2):33. doi:10.1007/s11554-023-01264-0.
3. Rezaei B, Mobasseri M, Enayatifar R. A secure, efficient and super-fast chaos-based image encryption algorithm for real-time applications. *J Real-Time Image Process.* 2023;20(2):30. doi:10.1007/s11554-023-01289-5.
4. Ekambaram D, Ponnusamy V. Real-time AI-assisted visual exercise pose correctness during rehabilitation training for musculoskeletal disorder. *J Real-Time Image Process.* 2024;21:2. doi:10.1007/s11554-023-01385-6.
5. Elhanashi A, Saponara S, Dini P, Zheng Q, Morita D, Raytchev B. An integrated and real-time social distancing, mask detection, and facial temperature video measurement system for pandemic monitoring. *J Real-Time Image Process.* 2023;20(5):95. doi:10.1007/s11554-023-01353-0.

6. Saponara S, Elhanashi A, Zheng Q. Developing a real-time social distancing detection system based on YOLOv4-tiny and bird-eye view for COVID-19. *J Real-Time Image Process.* 2022;19(3):551–563. doi:10.1007/s11554-022-01203-5.
7. Saurav S, Gidde P, Saini R, Singh S. Real-time eye state recognition using dual convolutional neural network ensemble. *J Real-Time Image Process.* 2022;19:607–622. doi:10.1007/s11554-022-01211-5.
8. Zhu Y, Dai Y, Han K, Wang J, Hu J. An efficient bicubic interpolation implementation for real-time image processing using hybrid computing. *J Real-Time Image Process.* 2022;19(6):1211–1223. doi:10.1007/s11554-022-01254-8.
9. Saponara S, Elhanashi A, Gagliardi A. Implementing a real-time, AI-based, people detection and social distancing measuring system for COVID-19. *J Real-Time Image Process.* 2021;18:1937–1947. doi:10.1007/s11554-021-01070-6.
10. Kaibou R, Azzaz MS, Benssalah M, Teguig D, Hamil H, Merah A, et al. Real-time FPGA implementation of a secure chaos-based digital crypto-watermarking system in the DWT domain using co-design approach. *J Real-Time Image Process.* 2021;18(6):2009–2025. doi:10.1007/s11554-021-01073-3.
11. Wang S, Zhao J, Ta N, Zhao X, Xiao M, Wei H. A real-time deep learning forest fire monitoring algorithm based on an improved Pruned+KD model. *J Real-Time Image Process.* 2021;18:2319–2329. doi:10.1007/s11554-021-01124-9.
12. Lai Z, Chen L, Jeon G, Liu Z, Zhong R, Yang X. Real-time and effective pan-sharpening for remote sensing using multi-scale fusion network. *J Real-Time Image Process.* 2021;18:1635–1651. doi:10.1007/s11554-021-01080-4.
13. Tariverdi A, Venkiteswaran VK, Richter M, Elle OJ, Tørresen J, Mathiassen K, et al. A recurrent neural-network-based real-time dynamic model for soft continuum manipulators. *Front Robot AI.* 2021;8:631303. doi:10.3389/frobt.2021.631303.
14. Saponara S, Elhanashi A, Gagliardi A. Real-time video fire/smoke detection based on CNN in antifire surveillance systems. *J Real-Time Image Process.* 2021;18(3):889–900. doi:10.1007/s11554-020-01044-0.
15. Raufmehr F, Salehi MR, Abiri E. A frame-level MLP-based bit-rate controller for real-time video transmission using VVC standard. *J Real-Time Image Process.* 2021;18(3):751–763. doi:10.1007/s11554-020-01018-2.
16. Shah AA, Parah SA, Rashid M, Elhoseny M. Efficient image encryption scheme based on generalized logistic map for real time image processing. *J Real-Time Image Process.* 2020;17(6):2139–2151. doi:10.1007/s11554-020-01008-4.
17. Pérez-Patricio M, Aguilar-González A. FPGA implementation of an efficient similarity-based adaptive window algorithm for real-time stereo matching. *J Real-Time Image Process.* 2019;16(2):271–287. doi:10.1007/s11554-015-0530-6.
18. Sinhal R, Ansari IA, Jain DK. Real-time watermark reconstruction for the identification of source information based on deep neural network. *J Real-Time Image Process.* 2020;17(6):2077–2095. doi:10.1007/s11554-019-00937-z.
19. Cambuim LFS, Oliveira LA Jr, Barros ENS, Ferreira APA. An FPGA-based real-time occlusion robust stereo vision system using semi-global matching. *J Real-Time Image Process.* 2020;17(5):1447–1468. doi:10.1007/s11554-019-00902-w.
20. Njiki M, Elouardi A, Bouaziz S, Casula O, Roy O. A multi-FPGA architecture-based real-time TFM ultrasound imaging. *J Real-Time Image Process.* 2019;16(2):505–521. doi:10.1007/s11554-016-0563-5.
21. Godase V, Pawar P, Nagane S, Kumbhar S. Automatic railway horn system using NodeMCU. *J Control Instrum.* 2024;15(1):11–19.
22. Godase V, Godase J. Diet prediction and feature importance of gut microbiome using machine learning. *Evol Electr Electron Eng.* 2024;5(2):214–219.
23. Jamadade VK, Ghodke MG, Katakdhond SS, Godase V. A comprehensive review on scalable Arduino radar platform for real-time object detection and mapping. *SSRN Electron J.* 2025. doi:10.2139/ssrn.5383814.

24. Godase V. A comprehensive study of revolutionizing EV charging with solar-powered wireless solutions. *J Adv Res Power Electron Power Syst.* 2025 Apr;2(1):23–37. doi:10.2139/ssrn.5383818.
25. Sharada KA, Kathiravan A, Mannam P, Akila D, Kumar BS, Godase VV. Advanced neural network models for optimal energy management in microgrids with integrated electric vehicles. 2025 5th International Conference on Trends in Material Science and Inventive Materials (ICTMIM), Kanyakumari, India. 2025. p. 1869–1874. doi: 10.1109/ICTMIM65579.2025.10988248.