

RESEARCH ARTICLE



Edge AI for Real-Time Facial Emotion Detection: A Raspberry Pi-Based Approach with CNN

Muneera Altayeb^{1,*} , Khaled Abu Al Rub¹ , Ahmad Nazzal¹ and Areen Arabiat¹

¹Department of Communications and Computer Engineering, Al-Ahliyya Amman University, Jordan

Abstract: Forecasts and analysis of human behavior and emotions have received a lot of development in the area of artificial intelligence. There are some potentials that edge computing in the field of human behavior and emotion forecasting and analysis. In this research, an autonomous edge computing system utilizing Raspberry Pi for real-time emotion detection, which can detect seven basic emotions, namely, happiness, anger, sadness, surprise, neutral, fear, and disgust, has been developed. The system extracts several feature sets such as edges, shapes, and textures, as well as landmarks of the face and upper body. A convolutional neural network classifier is applied to these feature sets to classify emotions in real time. A classification model has been trained using a manually developed dataset through the Data Collector class. Real-time testing of the system and evaluation of the system based on several criteria such as F-measure, accuracy, precision, and sensitivity have been performed. It has been found that the proposed system has up to 98% accuracy in real-time emotion detection.

Keywords: convolutional neural network, emotion detection, MediaPipe, Raspberry Pi

1. Introduction

Recent developments in affective computing have led to a rising need for real-time emotion recognition systems for applications in intelligent education, healthcare monitoring, human-computer interaction, and smart environments [1]. Traditional emotion recognition system frameworks based on the cloud usually face challenges related to communication delays, dependency on network infrastructure, and privacy issues due to constant sending of facial data to distant servers. However, edge artificial intelligence (AI) has proven the capability of deep learning models to be applied directly on limited-resource devices while maintaining good performance [2]. Integration of technology into security systems has become vital in the era of digitalization, as machine learning (ML) is a significant development in improving surveillance capabilities. Video cameras have revolutionized research by enabling the acquisition of real-time human faces, which have become an important part of human communication. Emotive facial recognition has drawn interest in the fields of computer science, computer animation, and perceptual sciences. Research focuses on creating robots that can recognize facial expressions and mood states and their application in interviews, surveillance systems, and aggression detection. Methods and algorithms for Facial Expression Recognition (FER) based on deep learning techniques are developed within the fields of computer vision and AI [3]. However, despite significant advances

in edge intelligence and facial analytics, several challenges still remain unresolved. Previous research has proved the possibility of applying deep learning models to resource-limited devices such as Raspberry Pi and embedded Internet of Things (IoT) devices for facial analysis and real-time intelligent monitoring [4].

2. Literature Review

A popular technology for face detection and identification systems, the Raspberry Pi is a low-cost, low-power single-board computer. With cloud storage providing the best performance features, research attempts to enhance these capabilities in security systems, picture capture, and storage. Real-time tracking and monitoring in a variety of industries are made possible by face recognition technologies [5].

Within the sphere of facial emotion recognition, Romero and Armenta [6] have designed an algorithm for real-time facial emotion recognition, which was implemented on Raspberry Pi 3B+. It consists of the use of a convolutional neural network (CNN) for feature extraction and emotion recognition training based on the FER-2013 dataset. The process includes three main stages: face detection, facial feature extraction, and emotional facial expression classification, providing 95% identification accuracy.

Another study carried out by Al-Dabbas et al. [7] deals with the investigation of face recognition using ML and DL algorithms. The researchers apply the Viola-Jones technique, Linear Discriminant Analysis (LDA), Mutual Information (MI), and Analysis of Variance (ANOVA) and achieve a maximum accuracy of classification equal to 96.01%. Venkatesan et al. [8] proposed a

*Corresponding author: Muneera Altayeb, Department of Communications and Computer Engineering, Al-Ahliyya Amman University, Jordan. Email: m.altayeb@ammanu.edu.jo

system that combines DeepFace and AI techniques for real-time emotion recognition. The proposed system relies on face capture, feature extraction, gesture analysis, and emotion detection. The system achieved an accuracy of 92.02% on test data for real-time face recognition and emotion classification.

Rathour et al. [9] provide a fast and reliable method of facial landmark detection and feature extraction for real-time capture of facial images, which is one of the essential prerequisites for emotion recognition systems. The proposed methodology includes three steps: face detection, landmark recognition, and feature extraction using linear Support Vector Machine (SVM), HOG mapper, and OpenCV. By using a Raspberry Pi 3, the proposed system has achieved real-time face detection. Similarly, Jaiswal et al. [10] have proposed a DL model based on CNN for emotion recognition based on images for the purpose of evaluating the model's efficiency. For the evaluation, two benchmark datasets—JAFPE (Japanese Female Facial Expression) and FER-2013—were used. The model was assessed using several performance metrics, including validation accuracy, detection rate, learning rate, validation loss, computational complexity, and processing time. Experimental results showed accuracies of 70.14% and 98.65% on the FER-2013 and JAFPE datasets, respectively.

In addition to emotion recognition, some systems have been designed to enhance monitoring and security in restricted environments by detecting unauthorized activities and generating real-time alerts. This system enables monitoring for restricted and classified zones, identifies disruptions, and alerts the respective team through Python Tkinter. An inventive system for the detection of facial expressions was proposed by Aly [11], where he utilized ResNet50, Temporal Convolutional Network (TCN), and Convolutional Block Attention Module (CBAM). Specifically, special attention is paid to the accuracy of the ResNet50 system. These accuracies are 91.86% for RAF-DB, 91.71% for FER2013, 95.85% for CK+, and 97.08% for Karolinska Directed Emotional Faces (KDEF) databases. In addition, Rathnala et al. have designed real-time face recognition with the help of Python and Arduino.

A real-time face recognition using Python and Arduino was implemented by Rathnala et al. [12] based on three phases. Detecting faces and collecting images was performed in the first phase. In the second phase, the recognizer was trained, and interesting elements were separated, and in the third phase, they were grouped and stored in XML records. Aswin et al. [13] proposed a system for real-time face and emotion recognition that extracts the Mel-Frequency Cepstral Coefficient (MFCC) features after converting the captured images into signals. The SVM then classifies these features for directed emotion classification. The model achieved an overall accuracy of 85% and successfully identified four main emotional states.

Subramanian et al. [14] developed a real-time face and emotion recognition system based on ML and AI. The introduced model consists of three stages: data acquisition, feature detection, and emotion classification. The model was trained using gradient boosting (GB) and achieved an average accuracy of 99%. On the other hand, Gaddam et al. [15] developed a system using OpenCV for face detection and recognition to improve universities' traditional attendance system. The system manages attendance notes, automatically estimates present and absentees, and produces an automated document or spreadsheet. The system uses the OpenCV library, Haar-Cascade for face detection, and Local Binary Pattern Histogram (LBPH) for recognition. Individual student training is conducted, and the system achieves a 97% accuracy rate.

3. Novelty and Contribution

The novelty of the proposed system lies not in the introduction of a new deep learning architecture, but in the development of a practical and efficient edge-based emotion recognition framework that integrates facial and upper-body behavioral cues for real-time analysis on resource-constrained hardware. Unlike conventional approaches that rely primarily on facial expressions extracted from publicly available datasets, the proposed system utilizes a custom dataset of manually captured images representing real operating conditions and combines facial landmarks with upper-body pose information to capture complementary emotional indicators. This design addresses a key limitation of existing systems, where subtle facial expressions, partial occlusions, or pose variations may reduce recognition accuracy.

Beyond emotion classification, this framework has the potential to support a wide range of computer vision tasks. It also provides additional behavioral indicators, such as attention level and emotional variability, which can support a broader evaluation of user engagement. The implementation on a Raspberry Pi platform shows that reliable emotion recognition can be achieved using low-cost edge hardware. These findings demonstrate that real-time emotion recognition can be effectively implemented on low-cost edge devices without relying on computationally demanding models, making the system suitable for practical human-computer interaction scenarios.

4. Research Methodology

The proposed framework focuses on detecting facial expressions and body posture patterns in public environments, enabling improved situational awareness and behavioral analysis. The Python code relies on a systematic approach to collecting labeled emotion data. These data are preprocessed to improve the training process. It is based on a dataset of images representing seven target emotions: anger, happiness, sadness, neutral, surprise, fear, and disgust. The system uses OpenCV to extract 68 facial landmark points and 14 upper-body landmark points as features, followed by a CNN for emotion classification. The overall system workflow is described in Figure 1. Advanced layers, such as dropout techniques, batch normalization, and data augmentation, ensure the model generalizes well to new data. Jupyter, an open-source web application, is used to create and share files containing live code, conditions, visualizations, and narrative texts across different domains.

Figure 2 illustrates the complete architecture of the proposed real-time emotion detection system with body language analysis, integrating both hardware and software components around a Raspberry Pi 4 [16], which serves as the central processing unit. The hardware setup includes a camera module or USB webcam connected via USB/CSI to capture live video streams of facial expressions and body posture, a microSD card for storing the operating system and application data, and a 5V/3A USB-C power supply to ensure stable operation [17]. User interaction is facilitated through a keyboard and mouse connected via USB, while a monitor connected through HDMI displays real-time video, detected emotions, and analysis results. An optional alert system, consisting of an LED or buzzer, is interfaced through GPIO pins [18] to provide notifications when specific conditions are met.

The software workflow begins with live video capture, followed by preprocessing using OpenCV for frame resizing and noise reduction. MediaPipe is then used for detecting facial landmarks and body keypoints, which are subsequently processed

Figure 1
The system's workflow

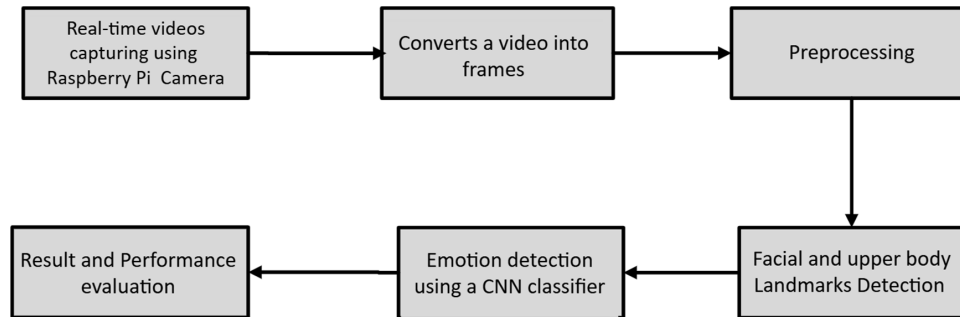
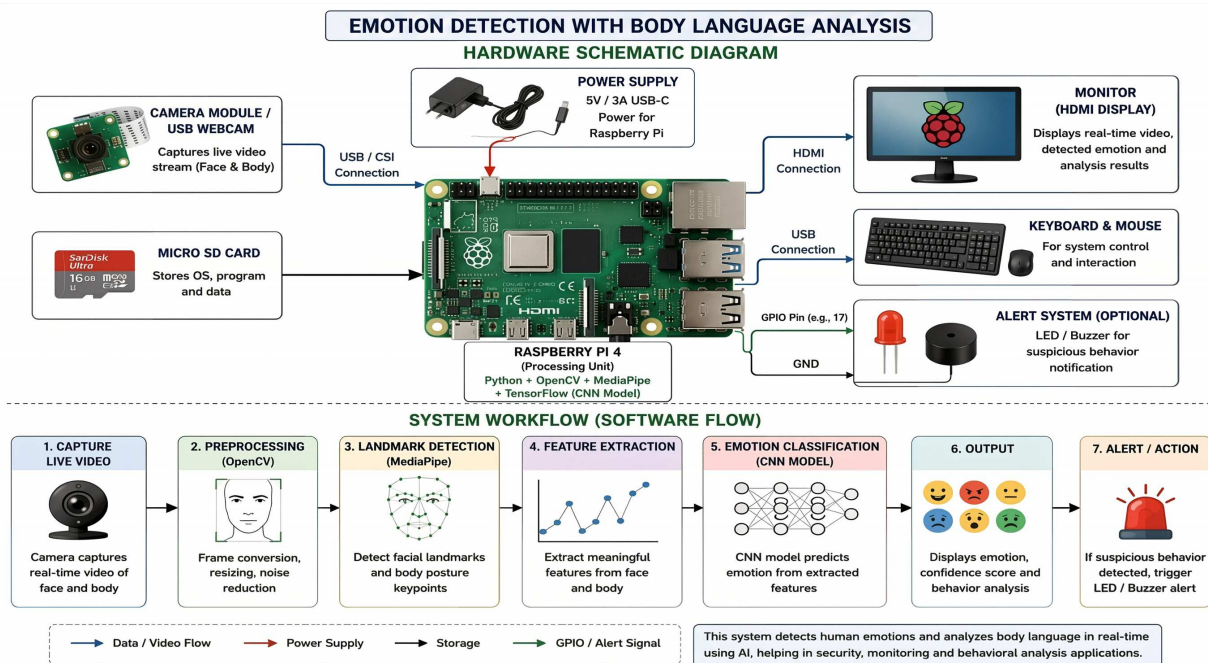


Figure 2
System schematic diagram



for feature extraction. These features are input into a trained CNN model for emotion classification. Finally, the system outputs the detected emotions and confidence scores on the display and triggers an alert if abnormal or predefined behaviors are identified. Overall, the system demonstrates an efficient integration of sensing, processing, analysis, and feedback modules to enable accurate and real-time emotion recognition.

4.1. Data collection and preparation

A structured method for collecting labeled emotional data was developed using the Data Collector class, an object-oriented programming pattern designed to centralize data collection, extraction, and validation. This approach facilitated organized and consistent data acquisition for training the emotion classification model. The collected data were preprocessed and augmented to improve the robustness of the training process. To train the enhanced model for emotion detection, the dataset was created through manual image collection. Pictures were captured using a webcam consisting of 350 RGB images (320×240 pixels), with 50 images for each of the seven emotion categories: angry, disgusted, happy, sad, fearful, neutral, and surprised, as shown in

Figure 3. The Data Collector class in the enhanced code facilitated this process, saving images into labeled directories for each emotion.

4.2. Facial and upper-body landmark detection

The goal of the proposed approach is to automatically detect 68 facial and 14 upper-body landmark points by identifying the precise (x, y) coordinates corresponding to key facial and upper-body structures using MediaPipe. Figure 4 (a) depicts the 68 landmark locations for the face, Figure 4 (b) depicts the indices of the facial landmark points [19], Figure 4 (c) shows the upper-body landmark points, and Figure 4 (d) depicts the indices of the upper-body landmark points [20]. The facial landmark points are organized into two fundamental categories. The mouth, eyes, nose, and eyebrows are among the first landmarks that are directly noticed. These are crucial for determining facial identities based on low-level picture characteristics. The nose, non-extremity points, chin, cheek, left-right eyebrow, and lips are the secondary landmark points; these are crucial for tracking faces and recognizing facial expressions, while the upper-body landmarks are critical for recognizing body language.

Figure 3

A sample of the dataset for the seven emotions: (a) angry, (b) disgusted, (c) happy, (d) sad, (e) fearful, (f) neutral, and (g) surprised

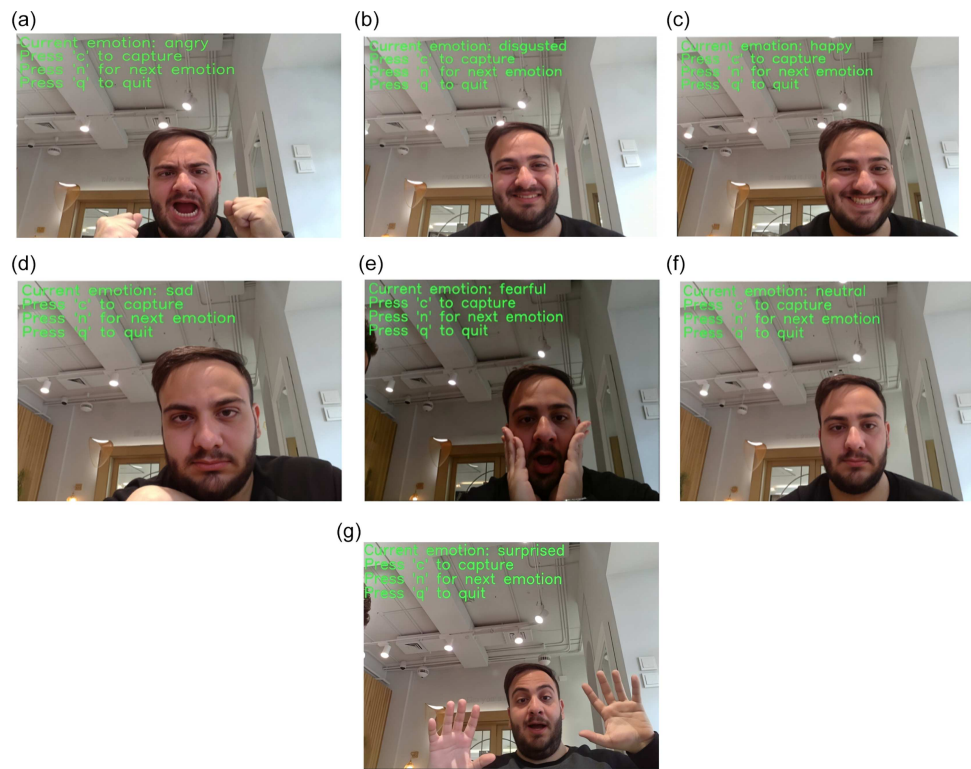
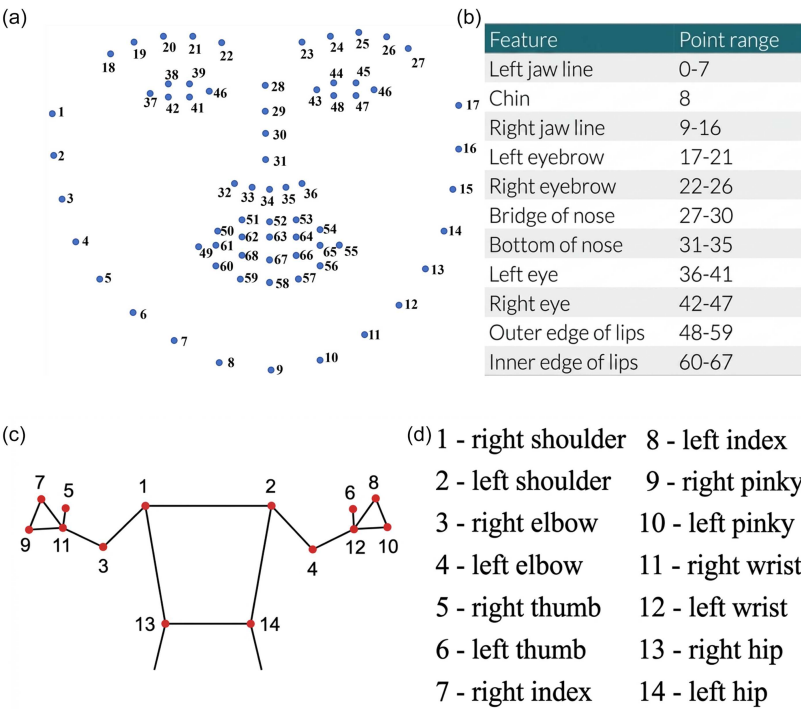


Figure 4

Landmark detection. (a) The 68 facial landmark points, (b) the indices of the facial landmark points, (c) the 14 upper-body landmark points, and (d) the indices of the upper-body landmark points



4.3. Emotion detection

The model interacts with the live video feed, detecting faces and classifying them into angry, happy, sad, neutral, surprised, fearful, or disgusted. However, Figure 5 depicts a script for the emotion detection code.

The Python code script describes the emotion detection. Also, this module evaluates body language by analyzing pose landmarks such as shoulders and hips to determine postural indicators like

openness or defensiveness. Figure 6 illustrates how body posture can be analyzed to detect any suspicious body posture.

4.4. Model training and classification method

4.4.1. Model training

The model is trained using a cyclic approach over 50 epochs. As shown in Figure 7, the loss curves demonstrate the model's learning progress, while the accuracy curves compare performance

Figure 5
Emotion detection code

```
class EmotionBodyDetector:
    def detect_emotion(self, face_landmarks):
        """Analyze facial expression for emotion."""
        if not face_landmarks:
            return "Unknown"

        # Get mouth points for smile detection
        mouth_left = face_landmarks[61].y
        mouth_right = face_landmarks[291].y
        mouth_center = face_landmarks[0].y
        :
        :
        :
        # Calculate mouth curve
        mouth_curve = mouth_left + mouth_right - 2 * mouth_center

        # Determine emotion based on mouth curve
        if mouth_curve < -0.02:
            return "Happy"
        elif mouth_curve > 0.02:
            return "Angry"
        :
        :
        .
```

Figure 6
Python script describes the steps of upper-body landmark detection

```
def detect_body_posture(self, pose_landmarks):
    """Analyze body posture and potential emotional indicators."""
    if not pose_landmarks:
        return "Unknown"

    # Shoulder and hip alignment
    left_shoulder = pose_landmarks[mp_holistic.PoseLandmark.LEFT_SHOULDER]
    right_shoulder = pose_landmarks[mp_holistic.PoseLandmark.RIGHT_SHOULDER]
    left_hip = pose_landmarks[mp_holistic.PoseLandmark.LEFT_HIP]
    right_hip = pose_landmarks[mp_holistic.PoseLandmark.RIGHT_HIP]

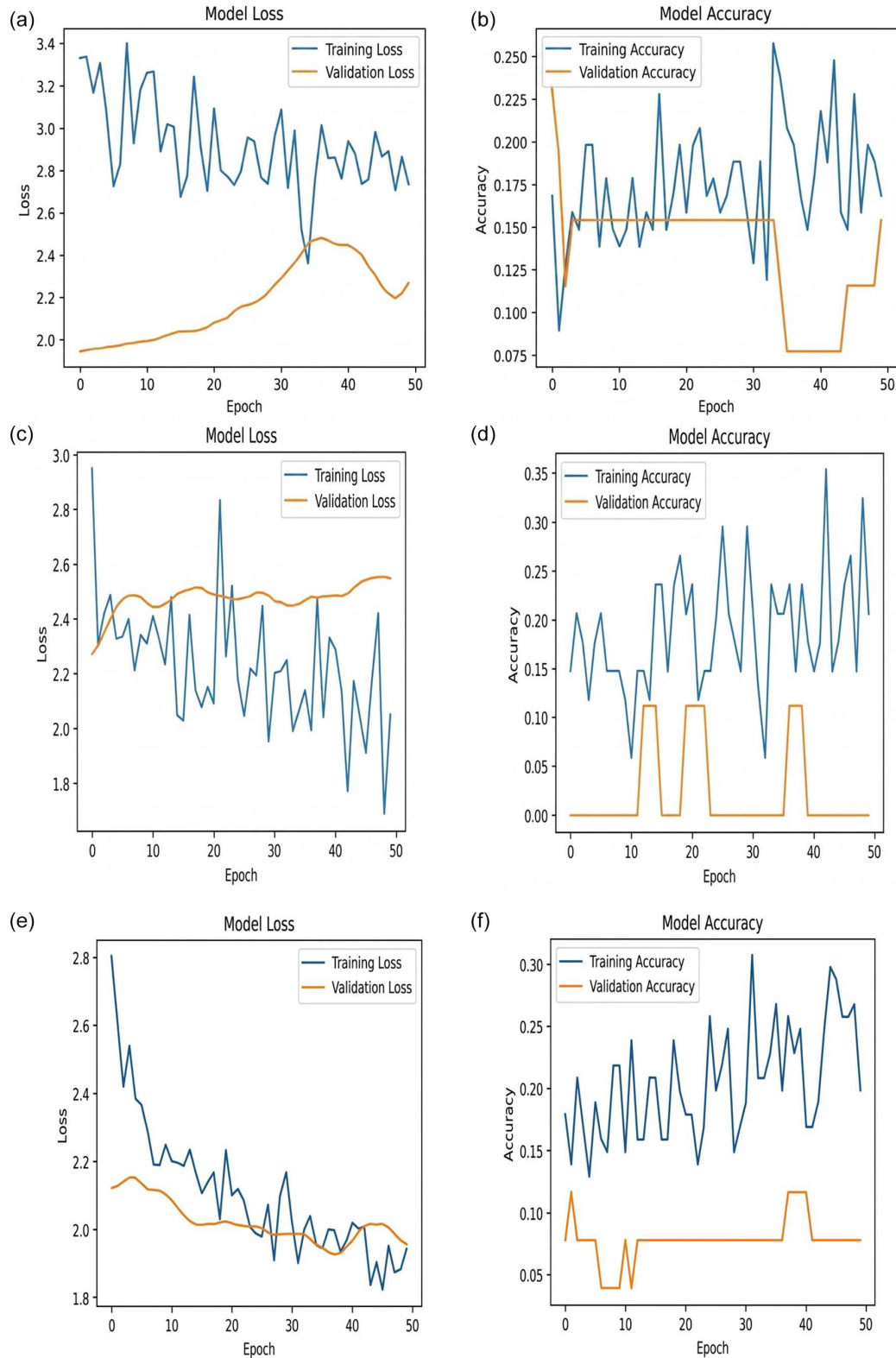
    # Check shoulder and hip alignment
    shoulder_width = abs(left_shoulder.x - right_shoulder.x)
    hip_width = abs(left_hip.x - right_hip.x)

    # Check head tilt
    left_ear = pose_landmarks[mp_holistic.PoseLandmark.LEFT_EAR]
    right_ear = pose_landmarks[mp_holistic.PoseLandmark.RIGHT_EAR]
    head_tilt = abs(left_ear.y - right_ear.y)

    # Assess body Language
    if head_tilt > 0.05:
        return "Uncertain/Stressed"
    elif shoulder_width < 0.15 and hip_width < 0.15:
        return "Closed/Defensive"
    elif shoulder_width > 0.2:
        return "Open/Confident"
    else:
```

Figure 7

Learning progress of the proposed model. (a) Training and validation loss for the first training configuration, (b) training and validation accuracy for the first training configuration, (c) training and validation loss for the second training configuration, (d) training and validation accuracy for the second training configuration, (e) training and validation loss for the third training configuration, and (f) training and validation accuracy for the third training configuration



on the training and validation sets, highlighting its predictive capability [21]. During each epoch, model parameters are updated by aggregating loss, accuracy, and sample counts across all batches. The model assessment criterion for the validation set is validation accuracy, and loss and accuracy are computed without gradients. To improve convergence, the learning rate scheduler is modified, and the optimal model is stored.

4.4.2. Classification method

Classification and pattern recognition are essential activities in data mining and ML. This paper used a CNN, a supervised learning method. In a CNN, there is an arbitrary number of bias nodes, always one in each layer. However, it consists of three layers: the input layer, the hidden layer, and the output layer. In this work, the CNN consists of three convolutional blocks, each comprising a convolution layer, batch normalization, ReLU activation function, max-pooling operation, and dropout regularization. Feature maps extracted by the convolutional layers are flattened and passed to a fully connected layer before the final SoftMax classifier. The network is trained using the Adam optimizer with a learning rate of 0.001 and categorical cross-entropy loss. Table 1 presents the detailed training hyperparameters used during the experiments. Figure 8 describes the real-time detection of happiness and also shows the identification of landmark points for the face and upper body.

5. Results and Performance Evaluation

The proposed system achieved smooth and responsive instant performance in detecting emotions and body language in real time, identifying basic emotions and evaluating body posture according to seven categories: surprised, angry, sad, neutral, fearful, disgusted, and happy. The accuracy for detecting emotions varied between 85% and 98%, as illustrated in Table 1.

Table 1
Training hyperparameters

Value	Parameter
Adam	Optimizer
0.001	Learning rate
50	Number of epochs
ReLU	Activation function
32	Batch size
Categorical cross-entropy	Loss function

5.1. Performance metrics

To evaluate the predictive performance of the model, testing was done after the verification of basic assumptions concerning the model. For this purpose, some assessment metrics were utilized to measure the appropriateness of the proposed models. The confusion matrix is a useful tool that can be used to assess the efficiency of classification and prediction algorithms [22] by determining proportions of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) for calculating accuracy [23]. The efficiency of the proposed approach was evaluated, as shown in Equations (1)–(4) and Table 2, using F-measure, accuracy, precision, and sensitivity [24]:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{F-measure} = \frac{2 \text{ Precision Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (4)$$

To assess the suitability of the proposed CNN-based facial emotion recognition system for edge AI deployment, a set of runtime and resource utilization metrics was collected on the Raspberry Pi 4 platform. As shown in Table 3, the system achieved an average inference time of 78.4 ms per frame, with an average processing rate of 12.7 FPS, demonstrating the capability to perform real-time emotion recognition on resource-constrained hardware. The average face detection and emotion classification times were 23.6 ms and 54.8 ms, respectively, resulting in an overall end-to-end response time of 95.3 ms. These results indicate that the proposed architecture can provide timely emotion predictions while maintaining acceptable latency for real-time applications.

Table 4 presents the resource utilization characteristics of the proposed system during operation. The average CPU utilization was 53.8%, while memory utilization remained at 27.6%, corresponding to approximately 1124 MB of RAM usage. Furthermore, the application storage footprint was only 28.4 MB, confirming the lightweight nature of the developed model and its suitability for deployment on low-cost edge devices. Overall, the obtained results demonstrate that the proposed solution achieves a balance between computational efficiency and recognition performance, making it appropriate for standalone real-time emotion recognition applications on Raspberry Pi.

6. Discussion

However, the experimental findings show that the suggested facial expression recognition technique performs efficiently, yet certain discrepancies occur when considering different types of emotions. For instance, the accuracy for recognizing the sad emotion amounts to 85%. It is noticeably lower than the accuracy of detecting the disgusted emotion, which equals 98%. A possible explanation of this phenomenon could be the fact that faces displaying sad emotions contain less distinctive features. Small frowning, eye narrowing, and lip corner lowering may be confused with neutral or tired faces, thus leading to misclassification. In turn, disgust emotions require more pronounced facial actions, like wrinkling the nose, raising the upper lip, etc., that provide better features for classification.

In order to evaluate the efficiency of the suggested technique, it is possible to perform its comparison with previous research in this field and present it in the form of Table 5. The suggested technique demonstrates an accuracy of 98%, which is higher than the majority of the other methods used. Thus, comparing the proposed model with the SVM algorithm using the MFCC feature set (accuracy – 85%), LDA/ANOVA/MI-based approach (accuracy – 96.01%), LBPH with OpenCV (accuracy – 97%), and TCN-CBAM-ResNet50 (accuracy – 95.85%), it could be concluded that the suggested model shows better results in terms of classification accuracy.

Furthermore, it slightly surpasses several deep learning-based methods, including TCN-CBAM-ResNet50 architectures (95.85%). The comparison confirms that the proposed method

Figure 8
Emotion detection and identification of facial and upper-body landmarks

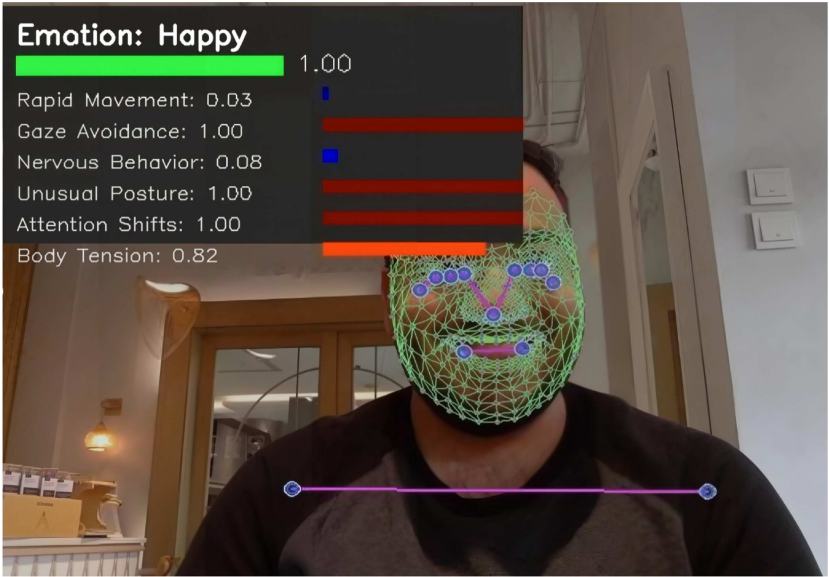


Table 2
Performance matrices of the suggested model

Class	Accuracy	Precision	Sensitivity	F1-score
Happy	0.90	0.88	0.90	0.89
Sad	0.85	0.85	0.85	0.85
Neutral	0.92	0.89	0.92	0.90
Surprised	0.96	0.94	0.96	0.95
Angry	0.94	0.92	0.94	0.93
Fearful	0.94	0.93	0.94	0.93
Disgusted	0.98	0.98	0.98	0.98
Macro (Avg)	0.93	0.91	0.93	0.92
Weighted (Avg)	0.92	0.92	0.92	0.92

Table 3
CNN inference performance on Raspberry Pi 4

Interpretation	Max	Min	Avg	Metric
Time required to classify one frame	102.0	61.0	78.4	Inference time (ms)
Real-time processing capability	16.4	9.8	12.7	FPS
Face localization overhead	35.8	18.4	23.6	Face detection time (ms)
CNN prediction time	66.2	42.6	54.8	Emotion classification time (ms)
End-to-end latency	128.4	74.8	95.3	System response time (ms)

Table 4
Raspberry Pi 4 resource utilization

Interpretation	Max	Min	Avg	Metric
Processor load during inference	87.4	18.2	53.8	CPU utilization (%)
Percentage of memory used	35.3	21.5	27.6	RAM utilization (%)
Runtime memory consumption	1432	978	1124	RAM used (MB)
Application footprint	28.4	28.4	28.4	Storage utilization (MB)

Table 5
Comparison of the proposed model with previous studies

Study/year	Methods/model	Performance metric	Percentage (max)
[7]/2024	LDA, ANOVA, MI	Accuracy	96.01%
[10]/2020	CNN	Accuracy	98.65%,
[11]/2025	TCNs, CBAM, ResNet50	Accuracy	95.85%
[13]/2016	SVM, MFCCs	Accuracy	85%
[14]/2022	GB	Accuracy	99%
[15]/2022	LBPH, OpenCV	Accuracy	97%
Proposed	CNN	Accuracy	98%

is competitive with, and in many cases superior to, existing state-of-the-art facial expression recognition techniques. Overall, the results confirm the effectiveness of the proposed approach while highlighting areas where further development is needed to enhance robustness, generalizability, and practical applicability.

7. Conclusion

This work presents the design process of a real-time emotion detection and recognition system based on Raspberry Pi and CNN. The novelty of this work lies in the idea of building and operating a unique and distinct system for face recognition, detecting suspicious people, and tracking them in real time. The system relies on digital image processing and extracting basic representative features of the 68 facial and 14 upper-body landmark points to detect one of seven basic emotions: surprised, angry, sad, neutral, fearful, disgusted, and happy. The system is fully functional and built from readily available components. In addition to several security and military applications, it is noteworthy that this system can be utilized in a variety of settings, including surveillance systems at retail malls and airports. The system was implemented and tested to detect seven basic emotions, achieving the highest detection accuracy of 98% for disgust and the lowest accuracy of 85% for sadness.

Despite the promising results achieved by the proposed system, several limitations should be acknowledged. The current study was evaluated under controlled experimental conditions and did not comprehensively address several real-world challenges, including illumination variations, facial occlusions, extreme pose changes, and crowded multi-face environments. Therefore, further investigation is required to assess the generalizability and robustness of the proposed approach in unconstrained deployment scenarios. Future research will focus on expanding the evaluation using diverse benchmark datasets and real-world environments while optimizing the system for improved performance and reliability.

8. Future Work

The results show that it is possible to perform real-time emotion recognition of facial expression using an efficient and lightweight approach based on CNNs on a Raspberry Pi. Although the current approach produces promising results, there are some ways to develop this study further. First, one can test different lightweight Vision Transformers and hybrid CNN-Transformer models to see whether they provide any improvements in classification accuracy while preserving efficiency.

One can also consider adding audio and physiological data along with facial expressions in order to have a more

complete emotional state description and make the system more effective in hard conditions. It is also possible to test adaptive methods taking into account the differences between users and environmental changes. Such improvements would expand the range of the applicability of the proposed approach, making it useful in healthcare, assistive technology, intelligent learning environment, human-computer interaction, and social robotics spheres.

Acknowledgement

The authors are grateful to the Department of Communications and Computer Engineering at Al-Ahliyya Amman University for their support.

Ethical Statement

This study was reviewed by the Graduation Projects Committee of the Department of Communications and Computer Engineering at Al-Ahliyya Amman University, and no formal ethical approval was required under institutional regulations. All authors adhere to strict data privacy protocols and the IEEE Code of Ethics.

Conflicts of Interest

The authors declare that they have no conflicts of interest related to this work.

Data Availability Statement

Data are available from the corresponding author upon reasonable request.

Author Contribution Statement

Muneera Altayeb: Conceptualization, Methodology, Software, Resources, Writing – original draft, Supervision, Project administration. **Khaled Abu Al Rub:** Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft. **Ahmad Nazzal:** Validation, Formal analysis, Investigation, Data curation, Visualization. **Areen Arabiat:** Writing – review & editing, Visualization.

References

- [1] Rathour, N., Khanam, Z., Gehlot, A., Singh, R., Rashid, M., AlGhamdi, A. S., & Alshamrani, S. S. (2021). Real-time facial emotion recognition framework for employees of

- organizations using Raspberry-Pi. *Applied Sciences*, 11(22), 10540. <https://doi.org/10.3390/app112210540>
- [2] Aboluhom, A. A. A., & Kandilli, I. (2025). Real-time facial recognition via multitask learning on Raspberry Pi. *Scientific Reports*, 15(1), 28467. <https://doi.org/10.1038/s41598-025-97490-6>
 - [3] Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. *Sensors*, 18(2), 401. <https://doi.org/10.3390/s18020401>
 - [4] Nagaveni, C. R., Monika, K., Rakshitha, Tejaswini., B. G., & Vismaya, G. R. (2026). Automated response and classification of infant's cry and emotions based on deep learning technique. In *2026 International Conference on Emerging Research in Smart Electronics and Machine Informatics*, 1–7. <https://doi.org/10.1109/ECMI68341.2026.11602875>
 - [5] Alhajim, D., Salman, FAHAMA., H., & T.A., M. (2024). Enhancing and securing a real-time embedded face recognition system using Raspberry Pi. *Journal of Al-Qadisiyah for Computer Science and Mathematics*, 16(1), 92–104. <https://doi.org/10.29304/jqcs.2024.16.11438>
 - [6] Romero, A., & Armenta, Á. (2024). Real-time emotion recognition using convolutional neural network: A Raspberry Pi architecture approach. In *Advances in Computational Intelligence. MICAI 2023 International Workshops* (pp. 191–200). https://doi.org/10.1007/978-3-031-51940-6_15
 - [7] Al-Dabbas, H. M., Azeez, R. A., & Ali, A. E. (2024). Two proposed models for face recognition: Achieving high accuracy and speed with artificial intelligence. *Engineering, Technology & Applied Science Research*, 14(2), 13706–13713. <https://doi.org/10.48084/etasr.7002>
 - [8] Venkatesan, R., Shirly, S., Selvarathi, M., & Jebaseeli, T. J. (2023). Human emotion detection using DeepFace and artificial intelligence. *Engineering Proceedings*, 59(1), 37. <https://doi.org/10.3390/engproc2023059037>
 - [9] Rathour, N., Gehlot, A., & Singh, R. (2019). A standalone vision device to recognize facial landmarks and smile in real time using Raspberry Pi and sensor. *International Journal of Engineering and Advanced Technology*, 8(6), 4383–4388. <https://doi.org/10.35940/ijeat.F8957.088619>
 - [10] Jaiswal, A., Raju, A. K., & Deb, S. (2020). Facial emotion detection using deep learning. In *2020 international conference for emerging technology*, 1–5. <https://doi.org/10.1109/INCET49848.2020.9154121>
 - [11] Aly, M. (2025). Revolutionizing online education: Advanced facial expression recognition for real-time student progress tracking via deep learning model. *Multimedia Tools and Applications*, 84(13), 12575–12614. <https://doi.org/10.1007/s11042-024-19392-5>
 - [12] Rathnala, P., Patnaik, M. S. P. K., Sura, S. R., Prasad, B., Rao, N. S. M., & Rayan, D. N. (2022). Design of an efficient face recognition system using deep learning technique. *International Journal of Electrical and Electronics Research*, 10(3), 689–693. <https://doi.org/10.37391/IJEER.100345>
 - [13] Aswin, K. M., Vasudev, K., Shanty, K., & Sreekutty, I. K. (2016). HERS: Human emotion recognition system. In *2016 International Conference on Information Science*, 176–179. <https://doi.org/10.1109/INFOSCI.2016.7845322>
 - [14] Subramanian, B., Kim, J., Maray, M., & Paul, A. (2022). Digital twin model: A real-time emotion recognition system for personalized healthcare. *IEEE Access*, 10, 81155–81165. <https://doi.org/10.1109/ACCESS.2022.3193941>
 - [15] Gaddam, D. K. R., Ansari, M. D., Vuppala, S., Gunjan, V. K., & Sati, M. M. (2022). Human facial emotion detection using deep learning. In *ICDSMLA 2020: Proceedings of the 2nd International Conference on Data Science, Machine Learning and Applications*, 1417–1427. https://doi.org/10.1007/978-981-16-3690-5_136
 - [16] Raspberry Pi Foundation. (n.d). *Raspberry Pi 4 Model B specifications*. Retrieved from: <https://www.raspberrypi.com/products/raspberry-pi-4-model-b/specifications/>
 - [17] IndiaMART. (n.d). *Raspberry Pi 4 Armor case with fan, Raspberry Pi 4 passive aluminum alloy*. Retrieved from: <https://www.indiamart.com/proddetail/raspberry-Pi-4-armor-case-with-fan-raspberry-Pi-4-passive-aluminum-alloy-case-2849826826162.html>
 - [18] Altayeb, M., Al-mandeel, L., & Arabiat, A. (2026). Smart assistive shoes for blind people based on IoT. *Smart Wearable Technology*, 2, A5. <https://doi.org/10.47852/bonviewSWT62028203>
 - [19] Jindal, A., & Priya, R. (2019). Landmark points detection in case of human facial tracking and detection. *International Journal of Engineering and Advanced Technology*, 9(2), 3769–3776. <https://doi.org/10.35940/ijeat.B3367.129219>
 - [20] Edge, Google AI. (2026). *Pose landmark detection guide*. Retrieved from: https://developers.google.com/edge/mediapipe/solutions/vision/pose_landmarker
 - [21] Qetheth, M. K., Al-Zuky, A. A. D., & Hamza, B. H. (2025). Exploring the impact of epoch value variation on human detection system training across diverse distances using the Faster R-CNN technique. In *AIP Conference Proceedings*, 3211(1), 030012. <https://doi.org/10.1063/5.0257225>
 - [22] Owida, H. A., Al-Nabulsi, J., Al Hawamdeh, N., Abuowaida, S., Salah, Z., & Elsoud, E. A. (2024). Deep learning-based kidney disease classification using ResNet50: Enhancing diagnostic accuracy. In *2024 Second Jordanian International Biomedical Engineering Conference*, 126–129. <https://doi.org/10.1109/JIBEC63210.2024.10931915>
 - [23] Altayeb, M., & Arabiat, A. (2025). A sustainable system for predicting appliance energy consumption based on machine learning. *Journal of Environmental Management*, 382, 125434. <https://doi.org/10.1016/j.jenvman.2025.125434>
 - [24] Bilal, H., Tian, Y., Ali, A., Muhammad, Y., Yahya, A., Izneid, B. A., & Ullah, I. (2024). An intelligent approach for early and accurate prediction of cardiac disease using hybrid artificial intelligence techniques. *Bioengineering*, 11(12), 1290. <https://doi.org/10.3390/bioengineering11121290>

How to Cite: Altayeb, M., Abu Al Rub, K., Nazzal, A., & Arabiat, A. (2026). Edge AI for Real-Time Facial Emotion Detection: A Raspberry Pi-Based Approach with CNN. *Smart Wearable Technology*. <https://doi.org/10.47852/bonviewSWT620210477>