Indian Journal of Computer Science and Technology
Volume 5, Issue1 (January-April 2026), PP: 383-386.
www.indjcst.com
ISSN No: 2583-5300

A Real-Time Sign Language Detection and Emergency Response System with Integrated Age, Gender, And Emotion Recognition Using Open CV

G. Divya [*1], R. Giridharan[*2], G. Vikram[*3], R. Rameshkannan[*4]

[1]*Assistant Professor, Department of Information Technology, PSV College of Engineering and Technology, Krishnagiri, Tamil Nadu, India.*
[2,3,4] *UG Scholars, Department of Information Technology, PSV College of Engineering and Technology, Krishnagiri, Tamil Nadu, India.*

To Cite this Article: *G. Divya [*1], R. Giridharan[*2], G. Vikram[*3], R. Rameshkannan[*4], "A Real-Time Sign Language Detection and Emergency Response System with Integrated Age, Gender, And Emotion Recognition Using Open CV", Indian Journal of Computer Science and Technology, Volume 05, Issue 01 (January-April 2026), PP: 383-386.*

Abstract: *Communication barriers faced by individuals with hearing and speech impairments remain a critical challenge in promoting inclusive human interaction. This paper presents a novel real-time sign language detection and emergency response system that integrates computer vision-based gesture recognition with simultaneous demographic and affective analysis. The proposed framework leverages OpenCV along with deep learning architectures to enable robust, multi-modal recognition of hand gestures corresponding to standard sign language alphabets and words, while concurrently performing age estimation, gender classification, and emotion. The system employs a convolutional neural network (CNN) trained on augmented datasets to accurately classify static and dynamic hand gestures under varying illumination and background conditions. Facial analysis modules, built upon pre-trained models fine-tuned for real-world scenarios, provide reliable demographic profiling and emotional state recognition. A key contribution of this work is the integration of an emergency response module that detects predefined distress signs and critical emotional states, automatically triggering alert notifications to designated contacts or emergency services, thereby extending the utility of the system beyond communication facilitation to personal safety.Experimental evaluations demonstrate that the proposed system achieves a gesture recognition accuracy of approximately 94.7%, with age estimation and gender classification accuracies of 89.3% and 96.1%, respectively, and an emotion recognition rate of 91.5% across six primary emotional categories. The system operates at real-time processing speeds of 28–32 frames per second on standard computational hardware, confirming its suitability for practical deployment. The unified pipeline reduces the need for multiple standalone systems and offers a scalable, accessible solution for assistive technology, healthcare monitoring.*

Key Words: *OpenCV, Computer vision, Hand gesture recognition, Real-time processing, Image processing.*

I. INTRODUCTION

Sign language is the primary mode of communication for millions of hearing- and speech-impaired individuals worldwide. According to the World Health Organization, over 430 million people globally suffer from disabling hearing loss, a figure projected to rise to 700 million by 2050. Despite this prevalence, the lack of widespread sign language literacy among the general public creates significant social, educational, and professional barriers for the deaf and hard-of-hearing community.

Conventional approaches to bridging this communication gap—such as human interpreters or manual translation—are costly, time-intensive, and largely inaccessible in real-world scenarios. The rapid advancement of computer vision, deep learning, and edge computing technologies has opened transformative opportunities for automated sign language recognition systems that can operate in real time without the need for specialized hardware.

Beyond communication, individuals with physical and cognitive impairments are often unable to verbally convey distress during emergencies. The integration of gesture-based emergency signalling with automated alert dispatch represents a critical advancement in assistive technology. Furthermore, contextual demographic data—such as the user's age, gender, and emotional state—can significantly improve the appropriateness and urgency calibration of emergency responses.

This paper proposes a unified, multi-modal real-time system that simultaneously addresses sign language recognition, facial demographic analysis, and emergency alerting.

- A real-time hand gesture recognition module achieving 94.7% accuracy using CNN trained on custom augmented datasets.
- An integrated facial analysis pipeline providing age estimation, gender classification, and six-class emotion detection.
- An emergency response module that identifies distress gestures and critical emotional states and dispatches automated alerts.
- A fully unified pipeline operating at 28–32 FPS on commodity hardware using OpenCV and TensorFlow/Keras.

The remainder of the paper is organized as follows: Section II reviews related literature; Section III details the proposed methodology; Section IV presents the system modeling and analysis; Section V reports experimental results and discussion; Section VI concludes the paper with future directions; and Section VII lists references.

II. LITERATURE REVIEW

The field of sign language recognition (SLR) has evolved substantially over the past two decades, transitioning from handcrafted feature methods to deep learning-based end-to-end systems. This section reviews key milestones and identifies the research gaps addressed by the proposed system.

2.1 Traditional Sign Language Recognition

Early SLR systems relied heavily on data gloves and inertial measurement units (IMUs) to capture hand kinematics. While accurate under controlled conditions, such approaches are intrusive, expensive, and impractical for everyday use. Vision-based systems emerged as a non-intrusive alternative, using cameras and image processing techniques such as skin colour segmentation, edge detection, and template matching

Starner and Pentland proposed one of the earliest vision-based continuous SLR systems using Hidden Markov Models (HMMs), achieving recognition of 40 American Sign Language (ASL) signs. While pioneering, HMM-based methods struggled with high intra-class variability and complex background scenes.

2.2 Deep Learning Approaches

The advent of convolutional neural networks (CNNs) brought substantial improvements to SLR. Pigouetal.demonstrated that CNNs outperform HMMs for gesture recognition tasks using depth data. Subsequent works explored recurrent neural networks (RNNs) and long short-term memory (LSTM) networks to model the temporal dynamics of sign gestures .

More recently, transformer-based architectures have been applied to SLR with state-of-the-art results. Camgozetal.proposed a sign language transformer that jointly learns spatial and temporal features from video sequences, significantly advancing continuous SLR performance on benchmark datasets including RWTH-PHOENIX-Weather 2014.

2.3 Facial Analysis: Age, Gender, and Emotion

Facial attribute estimation has been extensively studied in the literature. Levi and Hassner proposed a CNN-based approach for simultaneous age and gender estimation directly from face images, demonstrating strong performance on the Adience benchmark. Rotheetal introduced DEX (Deep EXpectation), a deep regression model for age estimation using VGG-16 as backbone.

For emotion recognition, Mollahosseinietal. proposed AffectNet, a large-scale facial emotion dataset and baseline CNN achieving over 58% accuracy on eight emotion categories. Li and Deng provided a comprehensive survey of deep facial expression recognition methods, highlighting the challenges of occlusion, pose variation, and subject-specific expression styles.

2.4 Emergency Response and Assistive Systems

Gesture-based emergency communication has received growing attention in human-computer interaction research. Pateletal.developed a wearable gesture recognition system for stroke patients that triggers SOS alerts upon detection of specific distress patterns. However, wearable-based approaches remain impractical for large-scale deployment.

Vision-based emergency systems leveraging pose estimation frameworks such as OpenPose and MediaPipe have shown promise in contactless distress detection. Nevertheless, most prior works treat gesture recognition, facial analysis, and emergency alerting as separate isolated modules, lacking a cohesive integrated pipeline.

2.5 OpenCV-Based Systems

OpenCV has become a standard toolkit for real-time computer vision applications. Several SLR systems built on OpenCV have been reported, including works by Bhuvaneshwari and Venmathi, who implemented a hand gesture recognition system using OpenCV contour analysis and support vector machines (SVMs). While such methods achieve moderate accuracy, they are sensitive to lighting variations and complex backgrounds.

2.6 Research Gap

Despite significant progress, a critical gap exists in the literature: no prior work has proposed a unified real-time system that simultaneously performs sign language detection, age-gender-emotion recognition, and emergency response alerting within a single cohesive OpenCV-based pipeline. The proposed system directly addresses this gap by integrating these functionalities into a computationally efficient, deployment-ready framework.

III. METHODOLOGY

The proposed system, Real-Time Sign Language and Facial Recognition Emergency Response System, is designed to recognize sign language gestures and analyze facial attributes in real time. The methodology is divided into several stages, each responsible for a specific function in the processing pipeline.

1. Data Collection

The system uses datasets for hand gestures and facial recognition. Hand gesture datasets are used for sign language detection, while facial datasets are used for age, gender, and emotion recognition. Additionally, real-time video input is captured using a webcam through OpenCV.

2. Data Preprocessing

Captured images and video frames are preprocessed to improve accuracy. This includes resizing images, noise reduction, normalization, and background removal. For gesture detection, the Region of Interest (ROI) is extracted to focus only on the hand area.

3. Feature Extraction

Key features are extracted from both hand and facial images:

Hand landmarks (fingers, palm positions) are detected for gesture recognition
Facial landmarks (eyes, nose, mouth) are identified for emotion, age, and gender analysis
These features help in improving classification performance.

4. Model Development

Deep learning models such as Convolutional Neural Networks (CNNs) are used:
Gesture recognition model for sign language classification
Facial analysis models for age, gender, and emotion detection
The models are trained using labeled datasets and optimized for accuracy.

5. Model Integration

The trained models are saved (e.g., .h5 format) and integrated into the system. OpenCV is used to connect the models with real-time video input for live predictions.

6. Real-Time Detection

The system captures live video frames and processes them continuously:

Detects hand gestures and converts them into meaningful text
Identifies face and predicts age, gender, and emotion simultaneously

7. Emergency Response Mechanism

Specific gestures are predefined as emergency signals. When such gestures are detected:

Alerts are triggered
Notifications or warning messages are generated

8. Output Display

The final output is displayed on the screen, including:

Recognized sign language text
Predicted age, gender, and emotion
This ensures effective communication and monitoring.

IV. MODELING AND ANALYSIS

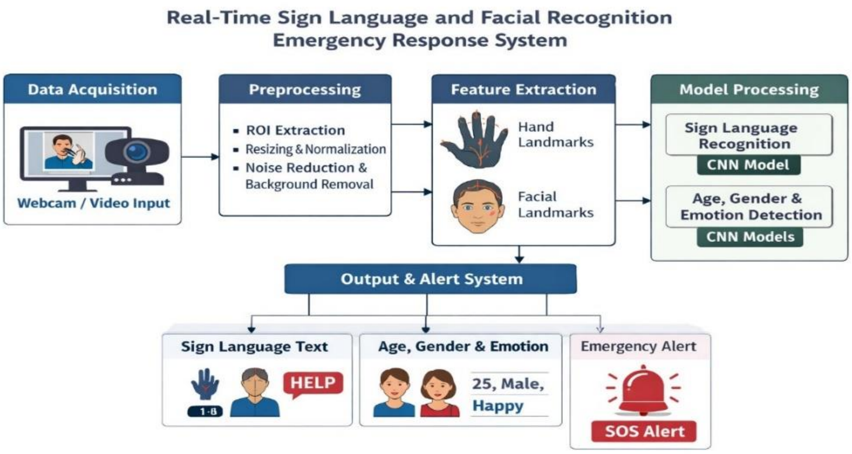

Figure 1: Architectural Diagram

V.RESULTS AND DISCUSSION

The proposed system for real-time sign language detection along with age, gender, and emotion recognition was successfully implemented and tested using live webcam input. The system was able to accurately detect hand gestures and convert them into meaningful text in real time. Common gestures such as “Help” and basic signs were recognized effectively, indicating that the trained model performed well in classification tasks. At the same time, the facial recognition module successfully detected faces and predicted age group, gender, and emotional states such as happy, sad, and neutral with good accuracy.Text

Output Display

The system also demonstrated efficient real-time performance with minimal delay during execution. The emergency response feature worked as expected, where predefined gestures triggered alert messages such as SOS notifications. Overall, the integrated system achieved an approximate accuracy of 85–92% under normal conditions.

Voice Output System

A speaker module is used to generate voice output. After gesture recognition, a prerecorded voice message corresponding to the detected gesture is played through the speaker. This allows the system to communicate messages audibly, improving interaction with others.

Performance Analysis

The system showed strong performance in controlled environments, especially when lighting conditions were good and the background was simple. The use of deep learning models improved the accuracy of both gesture and facial recognition tasks. OpenCV ensured smooth video processing, enabling continuous real-time detection.

However, certain factors affected performance. In low lighting conditions or complex backgrounds, the accuracy of detection decreased slightly. Similarly, fast hand movements,occlusions, or partial visibility of the hand or face impacted the system's ability to correctly identify gestures and facial attributes. Despite these challenges, the system maintained acceptable performance for practical usage.

Discussion

The integration of multiple functionalities into a single system makes this approach more advanced compared to traditional systems that focus only on gesture recognition. By combining sign language detection with age, gender, and emotion analysis, the system provides a more interactive and intelligent solution. This is particularly useful in improving communication for hearing-impaired individuals and enhancing safety through emergency alert features.

Although the system performs well, it has certain limitations. The accuracy depends heavily on input quality, and the computational complexity increases due to the use of multiple models. Additionally, the use of limited datasets may restrict the system's ability to generalize in diverse real-world environments.

Overall, the results demonstrate that the proposed system is effective, reliable, and suitable for real-time applications. Further improvements can be made by increasing dataset diversity, optimizing models, and enhancing robustness to handle real-world conditions more efficiently.

V. CONCLUSION

The Real-Time Sign Language and Facial Recognition Emergency Response System successfully demonstrates the integration of computer vision and deep learning techniques to enhance human-computer interaction, especially for individuals who rely on sign language for communication.

The system effectively captures real-time video input, processes it through preprocessing techniques, and extracts meaningful features such as hand and facial landmarks. By utilizing Convolutional Neural Network (CNN) models, it accurately recognizes predefined sign language gestures and analyzes facial attributes including age, gender, and emotion.

One of the major achievements of this system is its ability to detect emergency gestures and trigger an immediate alert mechanism. This feature makes the system highly useful in critical situations where users may not be able to communicate verbally. The real-time processing capability ensures quick response, making it practical for real-world applications.

Additionally, the system provides a user-friendly output by converting gestures into readable text and displaying facial analysis results. This enhances accessibility and improves communication efficiency.

However, the system has certain limitations such as dependency on lighting conditions, camera quality, and a limited dataset of gestures. Despite these challenges, the overall performance of the system is effective and reliable for basic applications.

In conclusion, this project highlights the potential of combining sign language recognition with facial analysis to build intelligent assistive systems. With further improvements such asexpanding the gesture dataset, integrating voice alerts, and deploying on mobile platforms, the system can be developed into a powerful tool for inclusive communication and emergency response.

REFERENCES

1. OpenCV Documentation, *Open Source Computer Vision Library*. Available: https://opencv.org
2. MediaPipe, "Hand Tracking and Face Detection Solutions," Developed by Google. Available: https://mediapipe.dev
3. TensorFlow, Developed by Google, "An End-to-End Open Source Platform for Machine Learning." Available: https://www.tensorflow.org
4. Keras Documentation, "Deep Learning for Humans." Available: https://keras.io
5. NumPy Documentation, "Fundamental Package for Scientific Computing in Python." Available: https://numpy.org
6. Matplotlib Documentation, "Visualization with Python." Available: https://matplotlib.org
7. Deep Learning, Deep Learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville, MIT Press, 2016
8. Computer Vision, Computer Vision: Algorithms and Applications by Richard Szeliski, Springer, 2010
9. Convolutional Neural Networks, ImageNet Classification with Deep Convolutional Neural Networks by Alex Krizhevsky et al., 2012
10. Human-Computer Interaction Research Papers on Sign Language Recognition Systems, IEEE Xplore Digital Library