

Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients

Saviour Inyang ^{1,✉}, Iniabasi Udo ², Imeh Umoren ³, Augustine Inyang ⁴, Felix Njoku ⁵

1. Cybersecurity, Federal University of Technology, Ikot Abasi, NGA

2. Computing, Topfaith University, Mkpatak, NGA

3. Computer Science, Federal University of Technology, Ikot Abasi, NGA

4. Business Administration, Topfaith University, Mkpatak, NGA

5. Software Engineering, Federal University of Technology, Ikot Abasi, NGA

Received: June 09, 2026 | Review began: July 19, 2026 | Review ended: August 01, 2026 | Published: August 19, 2026

© Copyright 2026

This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

This study presents the design and implementation of an Image-Based Sign Language Recognition System aimed at facilitating emergency communication with hearing-impaired patients. The system integrates deep learning and computer vision techniques to recognize real-time emergency-specific hand gestures such as Accident, Pain, Cold, Fever, and Help. A custom dataset comprising 250 manually captured images was developed to train and evaluate the model, with 50 images per class to ensure balanced representation. The hybrid architecture employs YOLOv8 for gesture detection and a Convolutional Neural Network (CNN) for classification. YOLOv8 efficiently localizes hand regions within video frames, while the CNN interprets the extracted gestures into corresponding emergency labels. The system was implemented using Python, OpenCV, and TensorFlow, and tested under various lighting conditions, hand orientations, and background environments to simulate real-world scenarios. Experimental results demonstrate that the integrated YOLOv8-CNN framework achieved a training accuracy of 99% and a testing accuracy of 93.5%, with high recall (0.927), F1-score (0.933), and precision (0.933) across all gesture classes. Performance analysis shows that YOLOv8 attained a mean average precision (mAP@50 = 0.984, mAP@50-95 = 0.915), indicating excellent detection capability, while the CNN maintained consistent classification accuracy across categories. The results confirm the system's robustness, reliability, and real-time responsiveness, demonstrating its potential to enhance emergency communication efficiency for hearing-impaired individuals in healthcare settings.

Categories: Bioinformatics, Machine Learning (ML), Medical Expert systems

Keywords: emergency communication, deep learning, assistive technology, gesture recognition, object detection

Introduction

Sign language involves the application of hand gestures and facial expression components to communicate thoughts and emotions. While sign language is an important means of communication for individuals with speech and hearing impairments, people who are able to speak and hear typically have little need to use it, others who are able to speak and hear rarely have to use sign language [1]. Although the process of learning sign language can be a difficult task, which limits the number of people learning to use sign language for communication, different technologies based on artificial intelligence have been suggested for solving the problems of sign language learning. Hearing-impaired individuals find it hard to communicate without the use of sign language [2]. Notwithstanding, the Americans with Disabilities Act legally mandates effective communication for individuals who are deaf or hard of hearing, but many healthcare systems

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

continue to fall short in fully meeting these standards [3]. Miscommunication frequently occurs when emergency medical staff struggle to grasp and respond appropriately to the needs of hearing-impaired patients. This issue may be largely driven by limited knowledge of sign language and insufficient availability of trained interpreters, leading to misunderstandings that may endanger patient safety [4]. Moreover, many emergency response guidelines overlook the inclusion of visual communication tools tailored to those with hearing impairments; by doing so, healthcare practitioners often depend on family members or friends to bridge communication gaps, a practice that risks inaccuracies in information transfer and leaves essential health concerns inadequately discussed [5]. Again, technological limitations significantly impact communication as well; although certain tools exist to assist interactions, they sometimes fail to meet the demands of real-time usage [6] during emergencies in the healthcare system for hearing-impaired patients [7]. Such current systems struggle to accurately interpret and convert sign language gestures into text or spoken words with the speed required in urgent scenarios. Additionally, gaps in training among emergency personnel exacerbate the issue, as they often receive little to no guidance on effectively interacting with deaf or hearing-impaired individuals [8]. Several studies have explored the areas of sign language recognition and communication technologies customized for individuals with impaired hearing, with a keen focus on model development, overall system assessments, and increased accessibility. Some studies [9-12], for instance, carried out meticulous studies into machine learning and deep learning approaches related to sign language interpretation, outlining contemporary approaches and applications while significantly excluding the deployment of real-time recognition systems. In contrast, Kothadiya et al. [13], Abid et al. [14], and Orovwode et al. [15] managed to develop pragmatic systems with impressive accuracies between 90% and 98.68% in the areas of sign language recognition or learning systems. However, despite these new breakthroughs, no single research considered in any of these works was related to image-based sign language recognition systems specifically customized for emergency communication applications among hearing-impaired patients. This presents a substantial research gap and the context in which this current research attempts to intervene by designing an intelligent-based recognition model built to establish rapid and accurate communication in emergency healthcare applications among hearing-impaired individuals.

Background

Emergency Communication Systems

Emergency Communication Systems constitute critical communication infrastructures designed to ensure the accurate and reliable exchange of information during emergency situations [16]. These systems are the backbone of disaster management, allowing coordination among first responders, governments, and impacted populations to reduce risks, save lives, and restore normalcy [17]. Artificial intelligence methods can enhance the proper flow of emergency communication when dealing with hearing-impaired patients through the conversion of spoken language into accessible formats such as text or sign language [18], while also identifying environmental emergency sounds that they may not perceive. In numerous emergency situations, especially within hospitals, delays in comprehension or communication can result in misdiagnosis, extended wait times, and preventable safety hazards for patients who are deaf or hard of hearing.

Assistive Technologies in Emergencies Communication

Assistive technologies are very useful for hearing-impaired individuals. They help individuals get information and stay safe during emergencies [19]. Tools like video relay services help with communication between responders, real-time captioning applications also help them understand what is happening, and vibrating alert systems give them emergency alerts [20]. These tools allow hearing-impaired individuals to communicate and get the information they need during emergencies. Speech-to-text software, augmented reality navigation aids, and adaptive mobility devices empower people with communication or physical disabilities to act swiftly [21].

Deep Learning for Recognition Systems

Deep learning is a subset of artificial intelligence that has been utilized over the years in recognition systems, especially when utilizing images. These deep learning models enable computers to interpret and understand visual data from images and videos [22]. Deep learning models can be applied to building systems that can be employed in object

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

detection, image classification, facial recognition, and scene understanding by extracting features from images, which include edges, lines, and shapes that can be applied for the classification of images into predefined categories. Convolutional neural networks (CNNs) are a type of deep learning model that has been highly effective when employing them for image processing and recognition systems. CNNs use convolutional and pooling layers to extract relevant features from images, which are then used in a wide range of image-based recognition tasks, including object detection, image classification, and facial recognition. CNNs are particularly effective at learning hierarchical representations of images and can be trained on large datasets to achieve high levels of accuracy [23]. YOLO (You Only Look Once) is another type of deep learning model that is employed for real-time object detection. It detects signs in an image by predicting bounding boxes and class probabilities directly from full images [24]. YOLO divides the input image into a grid and has each grid cell predict a fixed number of bounding boxes, along with objectness scores and class probabilities, all from a shared convolutional network. This unified, end-to-end architecture makes inference very fast (often 45+ frames per second) and, therefore, suitable for real-time applications such as video surveillance, autonomous driving, and robotics. Nevertheless, this study will utilize CNN and YOLO for sign language recognition in order to enhance emergency communication response among hearing-impaired patients.

Materials And Methods

The proposed methodology of this study is presented in Figure 1: Image-Based Sign Language Recognition System for Emergency Communication, which is aimed at addressing critical communication barriers between hearing-impaired patients and medical personnel by leveraging CNN and YOLO for real-time gesture translation. The system captures emergency-specific sign gestures (e.g., "Pain," "Fever," "Accident," "Help," and "Cold") with a camera, where YOLO rapidly detects and localizes hand regions in the input frames, minimizing computational overhead and enabling real-time processing [25]. The gesture is then classified by a CNN trained on a diverse dataset of annotated emergency signs, ensuring robustness to variations in lighting, angles, and hand sizes. Key challenges include dataset diversity, requiring augmentation for skin tones, backgrounds, and gesture dynamics; model optimization to balance speed and accuracy; and mitigating obstruction or motion blur in chaotic emergency environments [26-28]. Ethical considerations have been considered at various stages of the process of system development to facilitate the responsible usage of sign language data [29]. To protect the privacy of the users, all sign language images used in this project have been anonymized by removing any identifiable information from them and using them only for research purposes. The problem of bias mitigation has been addressed by designing the training dataset with signers having different hand shapes, skin tones, lighting conditions, and background variations, which will help in generalizing the model for different users. Also, the dataset is balanced among different gesture classes to avoid the problem of class imbalance and prediction bias. By converting sign language gestures into text in real time, the suggested system makes the communication process more efficient between healthcare professionals and patients. At the same time, there is a risk of miscommunication because of poor camera quality and lighting conditions.

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

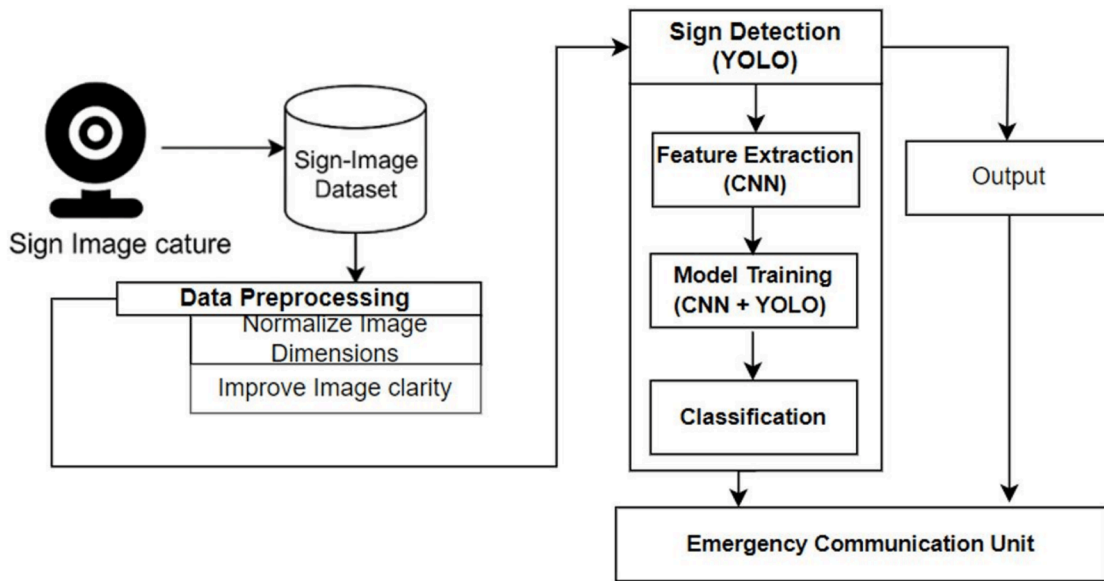


FIGURE 1: Proposed Framework for Sign Language Recognition System

CNN, Convolutional Neural Network; YOLO, You Only Look Once

Key components of the proposed framework

The following constitutes the description of the proposed framework:

1. Sign Image Capture and Dataset: This component captures sign language images through a camera. These images are stored in a folder, which is used as a dataset for training and detection processes.
2. Data Preprocessing: Before feeding the images into the detection model, preprocessing steps are applied, which include:
 - i. Normalize Image Dimensions: Rescale images to a fixed size.
 - ii. Improve Image Clarity: Apply filters or contrast enhancement.

Nevertheless, normalization can be represented in Equation 1:

$$X' = \frac{x - \mu}{\sigma} \quad (1)$$

where x is the original pixel value, μ is the mean, and σ is the standard deviation of the pixel values.

3. Sign Detection (using YOLO): YOLO is a real-time object detection algorithm. It detects signs in an image by predicting bounding boxes and class probabilities directly from full images. YOLO divides the image into an $S \times S$ grid, and for each grid cell, it predicts the bounding boxes and confidence scores.
4. Feature Extraction (using CNN): CNNs are used to extract spatial features from images through layers such as convolution, pooling, and activation functions using the following convolution operation in Equation 2.

$$(I * K)(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) K(x - m, y - n) \quad (2)$$

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

where I represents the input image, K represents the convolution kernel (filter), (x, y) represents the spatial coordinates, and (m, n) represents the indices used in the convolution operation.

The convolution operation between an input image I and kernel K is vital in image processing and CNNs for tasks such as feature extraction, blurring, or sharpening. This defines how each pixel at spatial coordinates (x, y) is calculated by summing the element-wise product of the image and the flipped kernel [30].

5. Model Training (CNN + YOLO): The extracted features are fed into the combined CNN and YOLO model to train the system for accurate detection and classification [31].

6. Classification: The output of the trained model is used to classify the detected signs into predefined emergency categories using a SoftMax layer.

7. Emergency Communication Unit: The final output from the classification is sent to the emergency communication unit for immediate action, enabling a timely response to critical situations.

System design

The design of the proposed deep learning-based system for sign language detection presents the structural and behavioral model that defines how various components interact to collect, process, and analyze sign language through the use of flowcharts, input/output design, and algorithms for the model training process. The design shows a comprehensive blueprint that guides the implementation of the system, ensuring accuracy, modularity, and maintenance.

Flowchart design

The flowchart in Figure 2 shows the process through which the sign language gestures will be captured, processed, and translated.

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

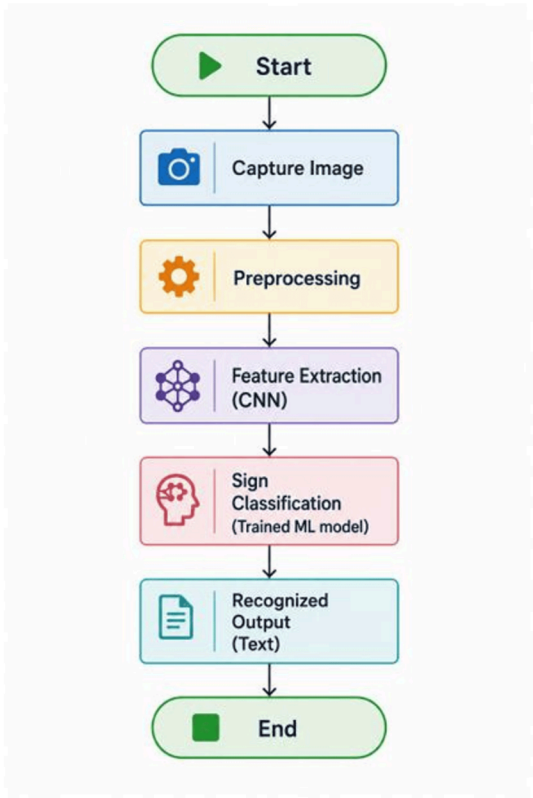


FIGURE 2: Flowchart of the Proposed System
CNN, Convolutional Neural Network

Input/output design

The input design for the proposed system in Figure 3 begins with a webcam, which serves as the primary input device to be used for capturing hand gestures in real time. The captured hand image is then passed into a trained deep learning model, which analyzes the visual features of the hand. Based on this analysis, the model then classifies the hand gesture as sign language. This setup ensures a simple and effective process for identifying sign language.

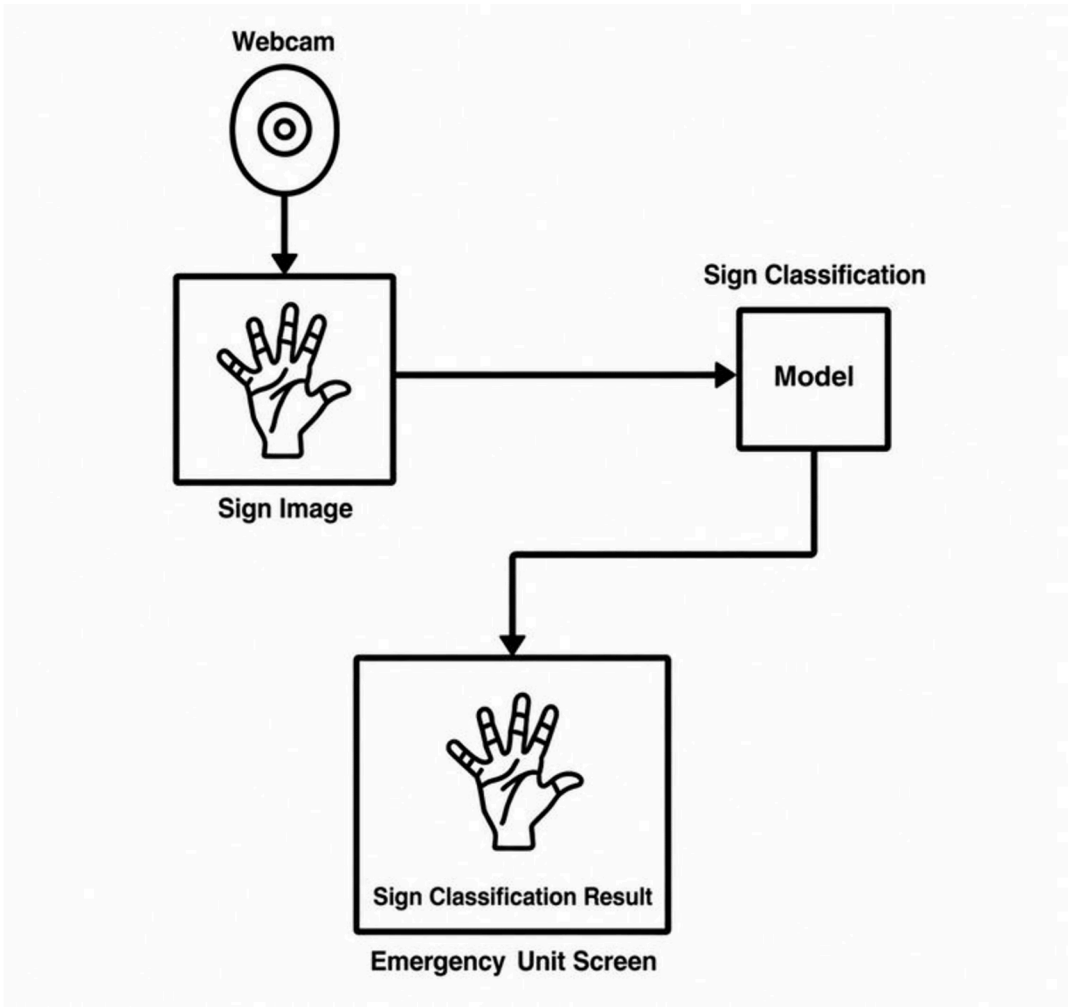


FIGURE 3: Input/Output Design

Algorithm/pseudocode

Tables 1-3 are used to present the algorithms that were utilized in the development and implementation of the model in this study.

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

CNN Algorithm for Feature Extraction
Step 1: Start
Step 2: Input: Preprocessed and cropped hand image (from YOLO)
Step 3: Resize the input image to 224 × 224 pixels
Step 4: Normalize pixel values to the range 0–1
Step 5: Pass the image into the CNN model: --- Apply Convolution Layer 1 --- Apply ReLU Activation --- Apply Max Pooling --- Repeat for multiple convolution-pooling layers
Step 6: Flatten the final pooled feature maps
Step 7: Pass through Fully Connected (Dense) Layers
Step 8: Apply Dropout to prevent overfitting
Step 9: Use Softmax Activation in the final layer to classify into gesture labels: Accident, Pain, Fever, Cold, and Help

TABLE 1: CNN Algorithm for Feature Extraction

CNN, Convolutional Neural Network; YOLO, You Only Look Once

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

YOLO-Based Hand Detection Algorithm
Step 1: Start
Step 2: Input: Raw webcam frame
Step 3: Resize the frame to the standard input size expected by YOLOv8 (e.g., 640 × 640)
Step 4: Divide the image into an S × S grid
Step 5: For each grid cell: Predict B bounding boxes, confidence scores, and class probabilities
Step 6: Apply Non-Maximum Suppression to remove overlapping/duplicate detections
Step 7: Select the bounding box with the highest confidence score
Step 8: Crop the region of interest (ROI) around the detected hand gesture
Step 9: Output: Cropped image passed to the CNN
Step 10: End

TABLE 2: YOLO-Based Hand Detection Algorithm

CNN, Convolutional Neural Network; YOLO, You Only Look Once

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

Integrated YOLOv8 + CNN Gesture Recognition Algorithm
Step 1: Start
Step 2: Capture a real-time frame from the webcam
Step 3: Apply YOLOv8 Detection: --- Detect hand gestures --- Crop the bounding box around the detected gesture
Step 4: Preprocess the cropped image
--- Resize to 224 × 224 --- Normalize pixel values
Step 5: Feed the image to the CNN model
Step 6: CNN extracts spatial features
Step 7: CNN classifies the gesture into one of five emergency categories
Step 8: Display/Output: --- Show the gesture label on the screen --- Send the label to the emergency communication interface
Step 9: Repeat Step 2 for next frame
Step 10: End

TABLE 3: Integrated YOLOv8 + CNN Gesture Recognition Algorithm

CNN, Convolutional Neural Network; YOLO, You Only Look Once

Results And Discussion

Experimental setup

The implementation of the Image-Based Sign Language Recognition System was undertaken with the aim of developing an efficient and effective means of communication for hearing-impaired patients during emergency conditions. This study was carried out through the incorporation of cutting-edge technology in computing and deep learning to achieve the automatic recognition and translation of sign language in real time. The system is implemented through a multi-component architecture that consists of a deep learning model used for gesture recognition, real-time image capturing and processing using the Python programming language, TensorFlow as the machine learning tool, OpenCV for image capture and processing, and the use of YOLO for object detection and gesture localization. The implementation phase involved developing a working prototype that can detect, classify, and recognize emergency-related signs from gestures captured through the camera interface. In addition, the implementation of this system was done such that image preprocessing, feature extraction, gesture detection, and classification could all be done within one workflow. The developed prototype was then tested under controlled settings and used to measure the performance of the system and determine how efficient it was in aiding communication between hearing-impaired patients and other healthcare professionals.

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

Input dataset

The dataset used in this study comprises a total of 250 manually captured images representing five distinct emergency sign language gestures, which include “Accident,” “Pain,” “Cold,” “Fever,” and “Help,” which are presented in Figure 4 and Figure 5, where the dataset is in [32]. Each class contains 50 images, ensuring balanced representation across categories. The dataset was created through controlled webcam recordings involving different users, hand orientations, and background environments to simulate real-world emergency scenarios. The rationale behind this dataset’s size and distribution is to provide sufficient variability and data volume for training a deep learning model [33] that can accurately distinguish between similar hand gestures under diverse conditions. This approach aligns with best practices in sign language dataset design, where visual consistency and class balance are crucial for robust classification performance [34].

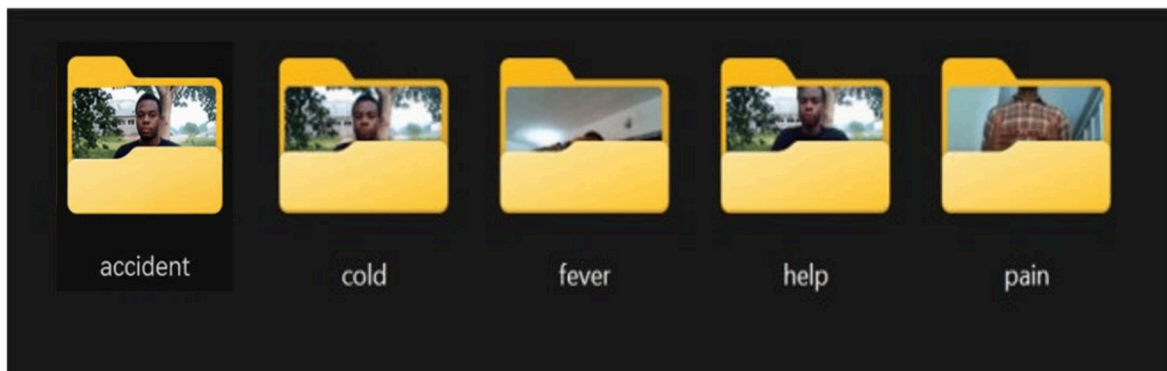


FIGURE 4: Snapshot of Dataset Groups

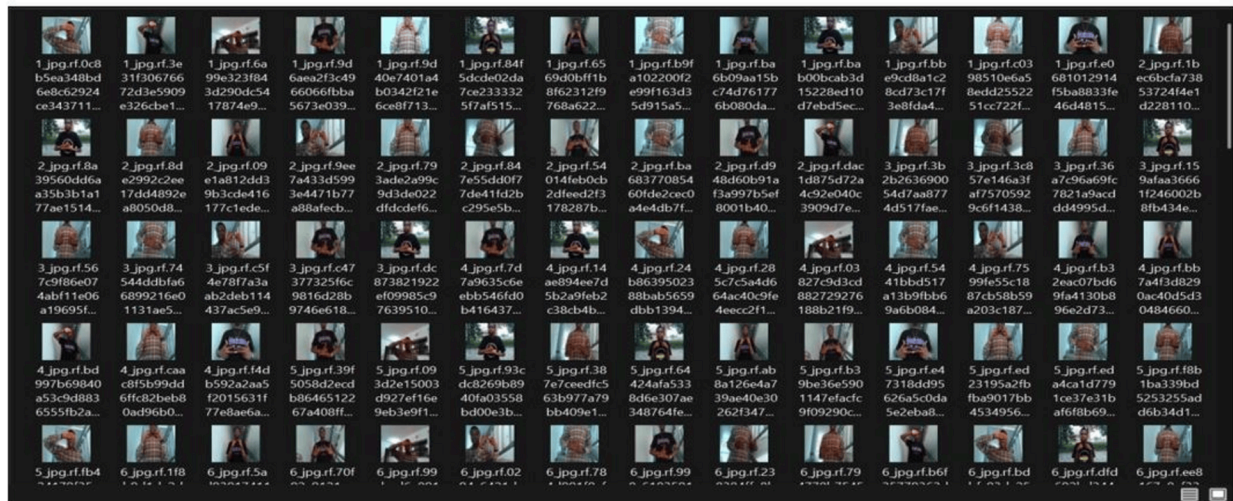


FIGURE 5: Snapshot of Images Classes: Accident, Cold, Fever, Help, and Pain

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

System testing

The procedures used to evaluate the functionality and performance of the hybrid detection-classification model are described in this section. Each component of the system was individually validated and tested for integration in real-time emergency sign language recognition.

Testing objectives

The primary objective of the testing phase was to ensure that the system could reliably identify emergency sign language gestures under varying lighting conditions, hand orientations, and user backgrounds. Additionally, the testing assessed the system's robustness in detecting signs in real time, its latency between detection and classification, and its ability to map gestures to corresponding emergency labels. Testing also confirmed system integrity, responsiveness, and usability from both technical and user-centered perspectives.

Sign language detection model

YOLOv8 was utilized for gesture detection. The model was fine-tuned on a custom dataset of emergency signs to detect hand regions within an image frame. YOLO processed input frames from a webcam and generated bounding boxes around regions of interest where a hand gesture was detected. These regions were then cropped, resized, and passed to the CNN model. The detection performance was optimized through adjustments in anchor boxes, confidence thresholds, and non-max suppression to minimize false positives and improve localization accuracy.

Deep learning model

The classification model employed was a custom CNN constructed with multiple convolutional and pooling layers, followed by fully connected layers. Dropout layers were used to prevent overfitting, and the Adam optimizer with a learning rate of 0.001 was applied for gradient-based optimization. The CNN was trained using a labeled dataset of cropped gesture images corresponding to predefined emergency classes such as "Help," "Pain," "Fever," and "Accident." Training was conducted over 20 epochs with a batch size of 8, and categorical cross-entropy served as the loss function.

How images were fed into the model

The real-time workflow starts with webcam image acquisition. Each frame is passed through YOLOv8 for detection. When a gesture is detected, YOLO extracts the bounding box, crops the image, and resizes it to a standard input shape of 224 × 224 pixels. The preprocessed image is normalized and fed into the CNN model, which outputs a predicted label. This prediction is then formatted and displayed as text, an alert, or both. The entire inference cycle operates with minimal latency, enabling near-instant feedback.

Model accuracy

Model performance was assessed using standard classification metrics over the test set (30% split from the entire dataset). The trained CNN was evaluated for its ability to classify cropped hand gestures into predefined emergency labels. The dataset included variations in background, hand size, and gesture orientation, ensuring the model's generalizability in real-world environments. The following metrics, which include accuracy, precision, and F1-score, were computed from the confusion matrix in order to evaluate model performance. Figure 6 and Figure 7 represent the CNN and YOLO performance during training based on the individual signs that were considered in this study, respectively.

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

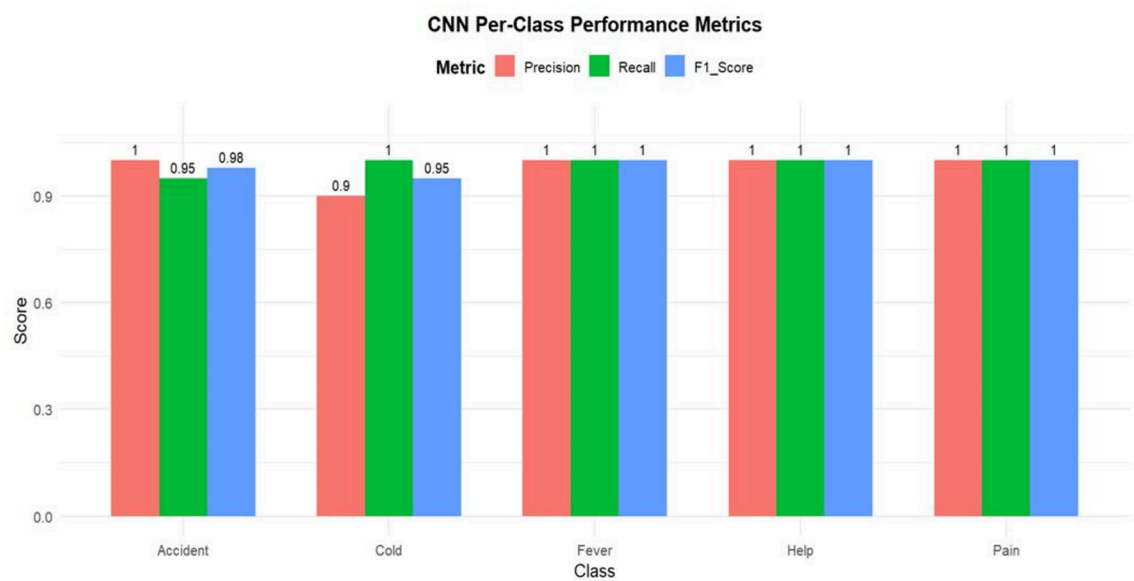


FIGURE 6: CNN Performance
CNN, Convolutional Neural Network

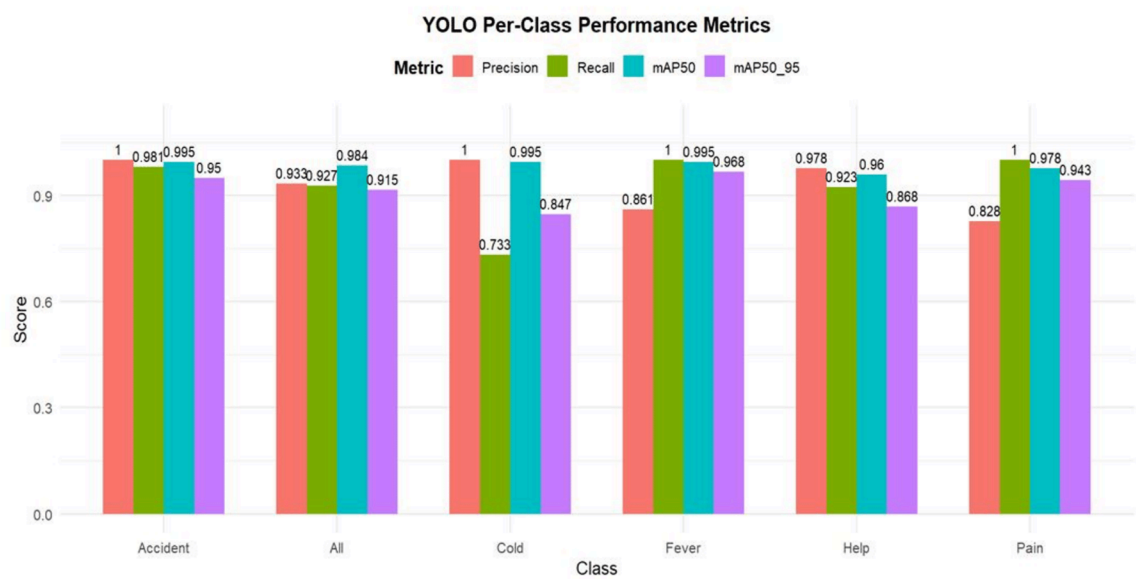


FIGURE 7: YOLO Performance
YOLO, You Only Look Once

Figure 6 and Figure 7 illustrate the comparative performance of the CNN and YOLO algorithms employed for the image processing system of the sign language recognition technique for emergency communication with hearing-impaired patients. Figure 6 reveals that the precision, recall, and F1 scores of the CNN approach for all classes of gestures (Accident, Cold, Fever, Help, and Pain) remain high, with a range of 0.90 to 1.00. This analysis reveals that the CNN method performs

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

almost flawlessly, with a high level of accuracy for gesture recognition after the detection of the hand sign. On the other hand, Figure 7 demonstrates the performance of YOLO on the detection task of the study. Not only does YOLO involve spatial detection, but it also shows precision, recall, mean average precision (mAP)@50, and mAP@50-95 metrics. Its performance in detection accuracy appears quite strong since all the values are greater than 0.90 for precision. A slightly lower value of 'Cold' recall at 0.733 does indicate some inconsistencies in detection accuracy for the various types of gestures. Yet, the YOLO model works quite well, with a high value of mAP of 0.984 for mAP@50 and 0.915 for mAP@50-95. Additionally, the overall summary performance of the developed sign language recognition system, which was evaluated during the training and testing phases, is presented in Table 4 using a held-out portion of the dataset (30% of the total 250 images). The trained CNN classifier, integrated with YOLOv8 for gesture detection, achieved a testing accuracy of 93.5%, which is consistent with its performance during the training and validation phases. No confidence interval was calculated. The model demonstrated robustness in accurately identifying emergency gestures like "Help," "Pain," "Cold," "Fever," and "Accident" across varying lighting conditions and hand orientations.

Phase	Accuracy	Recall	F1-Score	Precision
Training	0.99	0.99	0.99	0.99
Testing	0.935	0.927	0.933	0.933

TABLE 4: Summary for the Overall Training and Testing

Furthermore, Figure 8 depicts different real-time images that show the live testing accuracy using the webcam on the laptop to detect the effectiveness of the sign language recognition based on the trained model.

How to cite this article:



FIGURE 8: Real-Time Sign Detection

Conclusions

The image-based sign language recognition system built for this study is efficient for the recognition of hand emergency-related signs in a real-time environment by using YOLOv8 for image recognition and CNN for classification purposes. This cutting-edge system minimizes indirect interpretations by humans, provides a quick means of translating gestures into words for better understanding of life-critical emergencies, and facilitates easy accessibility for hearing-impaired patients for emergency communications. This system's ability to perform activities in a real-time environment signifies its use for life-critical healthcare applications that require swift and accurate performance. This study's impact on healthcare-related communication aids, supportive healthcare assistance, and AI-assisted gesture recognition technology systems signifies a supportive boundary for AI-assisted intelligent healthcare communications. Certain limitations found in this study include limitations in gesture recognition sets, dataset diversity, and environmental dependency that require further improvements. Further research should primarily focus on increasing datasets for gesture recognition, using diverse user datasets for improvements, designing user-centered interface designs for better accessibility and efficiency, and further improving and streamlining the system for ideation with portable device restrictions. Additional improvements for continuous gesture recognition techniques, 3D motion capture techniques, and further innovations for sign-to-speech translation may further enhance the system's recognition abilities and impact for swift communication for hearing-impaired individuals in life-critical emergencies within a short span of time.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

Concept and design: Saviour Inyang, Iniabasi Udo, Imeh Umoren, Felix Njoku

Acquisition, analysis, or interpretation of data: Saviour Inyang, Imeh Umoren, Augustine Inyang, Felix Njoku

Drafting of the manuscript: Saviour Inyang, Imeh Umoren, Augustine Inyang, Felix Njoku

Critical review of the manuscript for important intellectual content: Saviour Inyang, Iniabasi Udo

Supervision: Saviour Inyang

Disclosures

Human subjects: Consent was obtained or waived by all participants in this study. Professor Imeh Jonah Umoren issued approval (approval number not applicable). The study received ethical clearance prior to commencement. All participants were informed about the objectives and procedures of the research and provided informed consent before participating. Participants were advised of their right to withdraw from the study at any stage without penalty. All images collected for the Image-Based Sign Language Recognition System were anonymized to protect participants' privacy and confidentiality, and no images were shared with third parties. Data collection was conducted in a safe environment with no foreseeable risk of physical or psychological harm. The study adhered to the ethical principles of respect for persons, beneficence, and justice and was conducted solely to support healthcare communication between hearing-impaired patients and healthcare providers. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The datasets (and/or code) supporting this study are available from the corresponding author upon reasonable request.

References

1. Mannava V, Gnanasagar V: [Portable sign language translator for emergency response teams](#). International Journal of Scientific Research & Engineering Trends. 2020, 6:1203-1207.
2. Dere MD, Dere RO, Adesina A, Yauri AR: [SmartCall: A real-time, sign language medical emergency communicator](#). 2022 5th Information Technology for Education and Development (ITED), Abuja, Nigeria. 2022, 1-6. [10.1109/ITED56637.2022.10051420](#)
3. Engelman AA: [Addressing disparities in emergency communication with the deaf and hard-of-hearing: Cultural competence and preparedness for first responders \[dissertation\]](#). Berkeley: University of California, Berkeley. (2012). Accessed: October 17, 2025: <https://escholarship.org/uc/item/4j843811>.
4. Tannenbaum-Baruchi C, Feder-Bubis P, Aharonson-Daniel L: [Communication barriers to optimal access to emergency rooms according to deaf and hard-of-hearing patients and health care workers: A mixed-methods study](#). Academic Emergency Medicine. 2025, 32:246-259. [10.1111/acem.15037](#)
5. Akeely YY, Alenezi AQ, Albishr NN, Almutairi BA, Alotaibi NF, Almansour RA, Sabi MA: [Communication challenges while dealing with a deaf patient in the emergency department and suggested solutions](#). Cureus. 2022, 14:e31091. [10.7759/cureus.31091](#)
6. Inyang S, Umoren I: [From text to insights: NLP-driven classification of infectious diseases based on ecological risk factors](#). Journal of Innovation Information Technology and Application (JINITA). 2023, 5:154-165. [10.35970/jinita.v5i2.2084](#)
7. Lyons G, Normandin PA: [Strategies to improve emergency department care of the deaf and hard of hearing patient](#). Journal of Emergency Nursing. 2023, 49:489-494. [10.1016/j.jen.2023.02.007](#)
8. Gugenheimer J, Plaumann K, Schaub F, Di Campli San Vito P, Duck S, Rabus M, Rukzio E: [The impact of assistive technology on communication quality between deaf and hearing individuals](#). CSCW '17: Proceedings of the 2017 ACM Conference on Computer

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

- Supported Cooperative Work and Social Computing. 2017, 669-682. [10.1145/2998181.2998203](https://doi.org/10.1145/2998181.2998203)
9. Al-Qurishi M, Khalid T, Souissi R: [Deep learning for sign language recognition: Current techniques, benchmarks, and open issues](#). IEEE Access. 2021, 9:126917-126951. [10.1109/ACCESS.2021.3110912](https://doi.org/10.1109/ACCESS.2021.3110912)
10. Papastratis I, Chatzikonstantinou C, Konstantinidis D, Dimitropoulos K, Daras P: [Artificial intelligence technologies for sign language](#). Sensors. 2021, 21:5843. [10.3390/s21175843](https://doi.org/10.3390/s21175843)
11. Yenisari E, Yavuz S: [Deep learning-based sign language recognition using efficient multi-feature attention mechanism](#). IEEE Access. 2025, 13:126684-126699. [10.1109/access.2025.3586096](https://doi.org/10.1109/access.2025.3586096)
12. ZainEldin H, Gamel SA, Talaat FM, et al.: [Silent no more: a comprehensive review of artificial intelligence, deep learning, and machine learning in facilitating deaf and mute communication](#). Artificial Intelligence Review. 2024, 57:188. [10.1007/s10462-024-10816-0](https://doi.org/10.1007/s10462-024-10816-0)
13. Kothadiya D, Bhatt C, Sapariya K, Patel K, Gil-González A-B, Corchado JM: [Deepsign: sign language detection and recognition using deep learning](#). Electronics. 2022, 11:1780. [10.3390/electronics11111780](https://doi.org/10.3390/electronics11111780)
14. Abid MA, Pillai N, Ahammed N: [Greek sign language recognition using machine learning](#). International Journal of Advanced Computer Research. 2024, 14:1-17. [10.19101/ijacr.2023.1362022](https://doi.org/10.19101/ijacr.2023.1362022)
15. Orovwode H, Ibukun O, Abubakar JA: [A machine learning-driven web application for sign language learning](#). Frontiers in Artificial Intelligence. 2024, 7:1297347. [10.3389/frai.2024.1297347](https://doi.org/10.3389/frai.2024.1297347)
16. Carreras-Coch A, Navarro J, Sans C, Zaballos A: [Communication technologies in emergency situations](#). Electronics. 2022, 11:1155. [10.3390/electronics11071155](https://doi.org/10.3390/electronics11071155)
17. Deepak GC, Ladas A, Sambo YA, Pervaiz H, Politis C, Imran MA: [An overview of post-disaster emergency communication systems in the future networks](#). IEEE Wireless Communications. 2019, 26:132-139. [10.1109/mwc.2019.1800467](https://doi.org/10.1109/mwc.2019.1800467)
18. Marcolino MS, Oliveira LFRD, Valle LR, et al.: [Innovations in deaf health care communication: systematic review of sign language recognition systems](#). Journal of Medical Internet Research. 2026, 28:e70417. [10.2196/70417](https://doi.org/10.2196/70417)
19. Știrbu OI, Adochiei IR, Paraschiv RV, Nicolescu ȘT, Adochiei FC, Serîțan GC: [Assisting deaf and hard-of-hearing people in critical situations: alleviating stress and enhancing safety](#). 6th International Conference on Nanotechnologies and Biomedical Engineering. Sontea V, Tiginyanu I, Railean S (ed): Springer, Cham; 2024. 92:50-58. [10.1007/978-3-031-42782-4_6](https://doi.org/10.1007/978-3-031-42782-4_6)
20. Zhang H, Zhang R, Sun J: [Developing real-time IoT-based public safety alert and emergency response systems](#). Scientific Reports. 2025, 15:29056. [10.1038/s41598-025-13465-7](https://doi.org/10.1038/s41598-025-13465-7)
21. Mahmoudi-Dehaki M, Nasr-Esfahani N, Vasan S: [The transformative role of assistive technology in enhancing quality of life for individuals with disabilities](#). Assistive Technology Solutions for Aging Adults and Individuals With Disabilities. Pandey R, Maurya P, Prajapati J, Halder R, Tyagi K (ed): IGI Global Scientific Publishing, Hershey, PA; 2025. 45-72. [10.4018/979-8-3693-6308-9.ch003](https://doi.org/10.4018/979-8-3693-6308-9.ch003)
22. Li C, Li X, Chen M, Sun X: [Deep learning and image recognition](#). 2023 IEEE 6th International Conference on Electronic Information and Communication Technology (ICEICT), Qingdao, China. 2023, 557-562. [10.1109/ICEICT57916.2023.10245041](https://doi.org/10.1109/ICEICT57916.2023.10245041)
23. Iqbal S, Qureshi AN, Li J, Mahmood T: [On the analyses of medical images using traditional machine learning techniques and convolutional neural networks](#). Archives of Computational Methods in Engineering. 2023, 30:3173-3233. [10.1007/s11831-023-09899-9](https://doi.org/10.1007/s11831-023-09899-9)
24. Inthiyaz S, Ahamad SKH, Krishna AS, Bhargavi V, Govardhan D, Rajesh V: [YOLO \(you only look once\) making object detection work in medical imaging on convolution detection system](#). International Journal of Pharmaceutical Research. 2020, 12:312. [10.31838/ijpr/2020.12.02.0003](https://doi.org/10.31838/ijpr/2020.12.02.0003)
25. Etuk U, Umoren I, Umoren O, Inyang S: [Realtime-based system for facemask detection using PCA, with CNN and COCO model](#). Journal of Information Systems and Informatics. 2024, 6:1114-1131. [10.51519/journalisi.v6i2.759](https://doi.org/10.51519/journalisi.v6i2.759)
26. Guesmi A, Hanif MA, Ouni B, Shafique M: [Physical adversarial attacks for camera-based smart systems: Current trends, categorization, applications, research challenges, and future outlook](#). IEEE Access. 2023, 11:109617-109668. [10.1109/access.2023.3321118](https://doi.org/10.1109/access.2023.3321118)
27. Doan LMT, Shahhosseini K, Verma S, et al.: [Bridging modalities with AI: a review of AI advances in multimodal biomedical imaging](#). Communications Engineering. 2026, 5:30. [10.1038/s44172-026-00602-x](https://doi.org/10.1038/s44172-026-00602-x)
28. Chen J, Seng KP, Smith J, Ang LM: [Situation awareness in AI-based technologies and multimodal systems: Architectures, challenges and applications](#). IEEE Access. 2024, 12:88779-88818. [10.1109/access.2024.3416370](https://doi.org/10.1109/access.2024.3416370)

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>

29. Miah MA, Faruque M, Akter S, Jahan I: [Ethical considerations in AI and information technology privacy and bias](#). International Law Policy Review Organizational Management. 2024, 1:33-37. [10.13140/RG.2.2.18218.53440](#)
30. Wiatowski T, Bölcskei H: [A mathematical theory of deep convolutional neural networks for feature extraction](#). IEEE Transactions on Information Theory. 2018, 64:1845-1866. [10.1109/tit.2017.2776228](#)
31. Khan K, Aleem A: [A hybrid approach combining YOLO and deep neural networks for efficient lung disease detection](#). 2025 3rd International Conference on Communication, Security, and Artificial Intelligence (ICCSAI), Greater Noida, India. 2025, 1204-1211. [10.1109/ICCSAI64074.2025.11064147](#)
32. [Kaggle: Emergency Sign Language](#). (2026). Accessed: July 16, 2026: <https://doi.org/10.34740/KAGGLE/DSV/18091941>.
33. Inyang S, Essien D, Ezichi N: [Comparative analysis of machine learning models for irrigation technique classification in precision agriculture](#). International Journal of Applied Information Systems. 2025, 12:37-46. [10.5120/ijais2025452020](#)
34. Antonowicz P, Kasperek D, Podpora M: [Sign language recognition—dataset cleaning for robust word classification in a landmark-based approach](#). IEEE Access. 2025, 13:81877-81888. [10.1109/access.2025.3566338](#)

How to cite this article:

Inyang S, Udo I, Umoren I, et al. (August 19, 2026) Real-Time Sign Language Recognition System for Enhancing Emergency Response Among Hearing-Impaired Patients . Cureus J Comput Sci 3 : es44389-026-00229-1. DOI <https://doi.org/10.7759/s44389-026-00229-1>