International Journal of Scientific Research & Engineering Trends
Volume 11, Issue 2, Mar-Apr-2025, ISSN (Online): 2395-566X

Hand Gesture Recognition in Low-Light Environments Using Deep Learning

Sujal Suraywanshi, Ganesh Waje, Diya Jain, Prajyot Jagtap
Computer Engineering, AISSMS Institute of Information Technology, Pune-411 001, Maharashtra, India

Abstract- Hand gesture recognition has attracted enormous interest because of its extensive application in human-computer interaction, sign language understanding, and augmented reality. Recognizing hand gestures in low light conditions is a difficult task to achieve because the visibility is low, there are noises, and feature details get lost. Here, we propose a detailed survey of recent techniques in hand gesture recognition for low-light conditions by employing deep learning methods. We present the difficulties involved with low-light environments, such as illumination changes and background noise. In addition, we discuss different deep learning-based methods for hand detection and gesture classification and present their effectiveness in improving recognition accuracy under difficult lighting conditions. We conclude by offering a comparative assessment of these methods based on primary performance indicators like accuracy, processing time, and resistance to low-light conditions.

Index Terms- Hand Gesture Recognition, Low-Light Conditions, Human-Computer Interaction, Sign Language Understanding, Augmented Reality, Deep Learning

I. INTRODUCTION

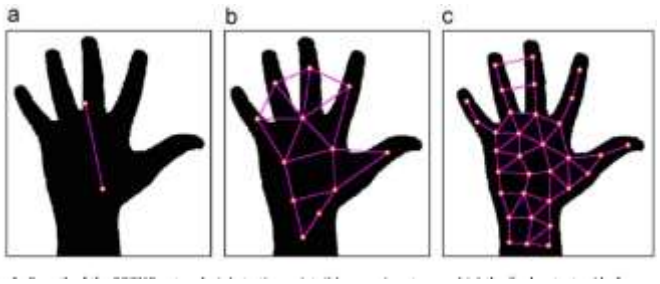
Hand gesture recognition is a procedure for detecting and understanding hand gestures to facilitate easy interaction with computers and intelligent devices. It is computer vision technology that converts hand gestures into effective commands or operations, making it very useful in human-computer interaction, virtual reality, sign language recognition, and robot control.

In recent years, web cameras have gained considerable popularity for real-time hand gesture recognition because of their ease of integration, affordability, and availability. Yet, hand gesture recognition in low-light conditions poses some special challenges. Under low-light conditions, visibility degrades, noise increases, and salient gesture features may be lost, degrading the recognition accuracy considerably. Conventional computer vision methods have difficulty functioning under such circumstances, and therefore advanced deep learning models are used to achieve better robustness and accuracy.

In this paper, we investigate recent developments in deep learning-based hand gesture recognition specifically for low-light settings. We address the challenge presented by low light and discuss different deep learning methods that facilitate gesture recognition under low illumination.

Background

Hand gesture recognition is a well-popularized topic of computer vision research since the early 2000s. Conventional approaches to hand gesture recognition include capturing hand movement through sensors, for example, gloves.



But these methods are costly and not very feasible. With the introduction of web cameras and their growing popularity, researchers have turned their attention to creating hand gesture recognition systems based on web cameras as the main input device.

II. PROBLEM STATEMENT

Hand detection is an essential process in hand gesture recognition that allows systems to find hands within images or video frames. Low-light conditions, however, compromise illumination, produce shadows, and introduce noise that complicates correct detection. Hand detection becomes very difficult for traditional methods as hands are hard to

© 2025 IJSRET

distinguish from the background, and misclassification ensues. Deep learning provides an effective solution in the form of improved feature extraction and accuracy of detection regardless of lighting conditions.

Classification

Subsequent to feature extraction, a label is assigned to the identified hand gesture by classification. Traditional methods commonly used are Support Vector Machine (SVM), k-Nearest Neighbor (KNN), and Artificial Neural Networks (ANN). Nevertheless, in low light levels, conventional classifiers can be prone to failure owing to poor representation of features. Deep learning architectures, including Convolutional Neural Networks (CNN) and Transformers, present enhanced robustness and flexibility and are therefore well suited for gesture recognition under adverse lighting.

III. PROPOSED SYSTEM

The proposed system focuses on real-time hand gesture recognition in low-light environments using deep learning. The process consists of:

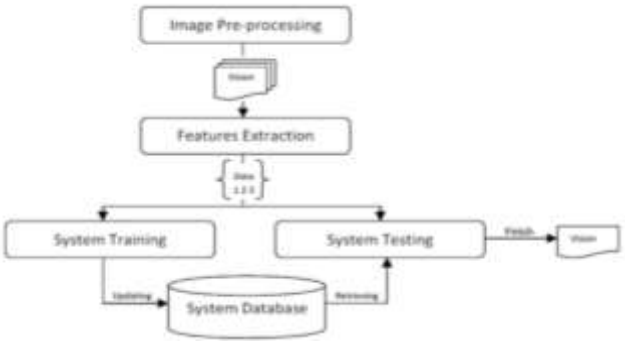
- **Image Acquisition –** A camera captures hand gesture images in dim lighting.
- **Pre-processing –** Noise reduction and low-light image enhancement improve visibility.
- **Hand Detection –** Deep learning models identify hand regions despite poor illumination.
- **Feature Extraction –** Key features such as shape, orientation, and movement are extracted.
- **Gesture Recognition –** CNN or Transformer models classify gestures accurately.
- **Gesture Output –** The recognized gesture is mapped to a specific action or command.

Deep learning techniques significantly improve accuracy by adapting to illumination variations, making the system effective for real-world applications such as sign language interpretation and smart device control.

IV. SYSTEM ARCHITECTURE

The system architecture for a hand gesture recognition using a camera can be divided into the following components:

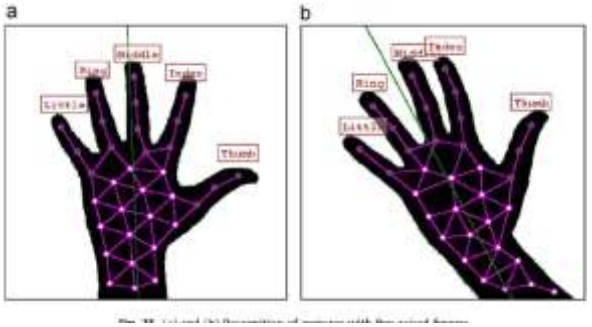
- Camera module
- Image processing module
- Hand detection module
- Finger detection module
- Feature extraction module
- Gesture recognition module
- Gesture output module



The overall architecture of the hand gesture recognition system can be implemented using different hardware and software components, depending on the specific application requirements.

1. Camera Module

The camera module is one of the building blocks of the hand gesture recognition system architecture. It takes the video or image stream of the hand gestures and forwards it to the image processing module for processing. The camera module may be implemented with a variety of hardware and software modules based on the application requirements.



The camera module usually consists of the following devices:

- Image sensor - converting incident light from the scene into an electrical signal for further processing by other electronics in the camera assembly.
- Lens - focuses light on the image sensor and controls camera field of view.
- Processor - controls the camera electronics and executes various image processing tasks such as color correction and white balancing and controlling exposure.
- Interface - channels through which one can interact with the camera like a USB port, Ethernet, or wirelessly.

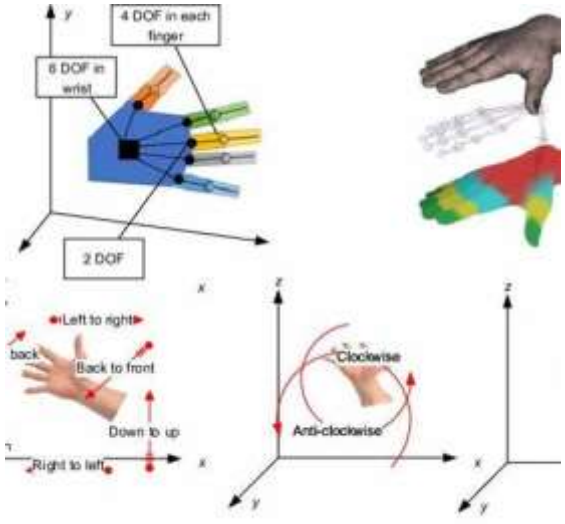
Additional camera module features may include image stabilization and autofocus or zoom. The criteria for selecting a camera module depend on the application specifications such as resolution, frame rate, field of view, and illumination. For hand gesture recognition, the camera module must have enough high-resolution to properly render details and movements of the hand. High frame rates are needed too to

© 2025 IJSRET

allow real-time recognition of hand gestures. The camera module must be adaptable to varying ambient light situations, as lighting has a lot of influences while trying to attain good quality image and accuracy in the system.

2. Image Processing Module

The image processing module is another important part of the hand gesture recognition system architecture. Its main role is to carry out pre-processing on the captured video or image stream from the camera module.



In principle, pre-processing methods are for improving the image quality, suppressing noise, and eliminating any background interference. Generally, an image processing module comprises the following sub-modules:

- Pre-processing: It processes and performs the first stage in the processing of the captured video or image stream. Pre-processing operations may include white balance, color correction, brightness, and contrast. Pre-processing is important because it improves the quality of the image and makes sure that the hand gestures are being detected correctly.
- Filtering: This is a sub-module that does away with noise and blurs images. Filtering methods include mean filtering, Gaussian filtering, adaptive filtering, and so on.
- Edge detection: This is a sub-module applied to detect edges of the hand and its details, including fingers and palm. Edge detection methods can be Sobel, Canny, or Laplacian edge detection.

The pre-processed image or video stream output from the image processing module is passed to a hand detection module for more processing. The image processing methods used in the module are to improve the image and minimize background interference which is very important for precise hand detection and gesture recognition. Specific image processing methods used in the architecture of a hand gesture recognition system may change according to particular application requirements and based on the input image or video's nature

.

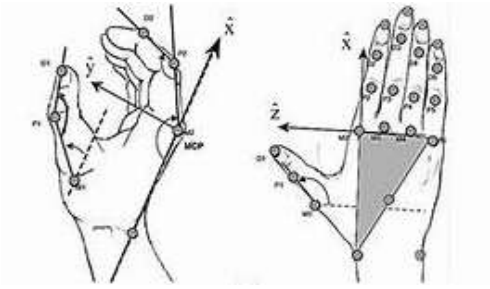
3. Hand Detection Module

The hand detection module is an essential part of the hand gesture recognition system that identifies the position of the hand in a video or image stream. Low-light conditions pose challenges to detection using conventional approaches because of poor visibility and high noise. The difficulties are overcome by utilizing deep learning-based methods for enhanced robustness and accuracy in detection. The module has the following sub-modules:

- The first part involves Low-Light Image Enhancement: one option effects image clarity improvements via the contrast enhancement techniques of Histogram Equalization or by deep learning low light enhancement models
- Deep Learning Detection Methods: Deep learning object detection models such as the YOLO series or Faster R-CNNs or CNNs provide heightened accuracy in hand detection even in dimly lit environments. These models take advantage of learning complex patterns and are thus able to adapt to varying degrees of illuminance.
- Infrared and Thermal Imaging (Optional): In case of extremely low-light conditions, the inclusion of infrared or thermal cameras would enhance detection accuracy through the collection of heat signatures rather than visible light.

Harnessing deep learning to enhance hand detection in low-light conditions, the study suggested that learning to adjust for illumination can not only instantly mitigate false positives.
The output of the hand detection module is a binary mask image which indicates the position of the hand in the image.

The output of the hand detection module is a binary mask image which indicates the position of the hand in the image.



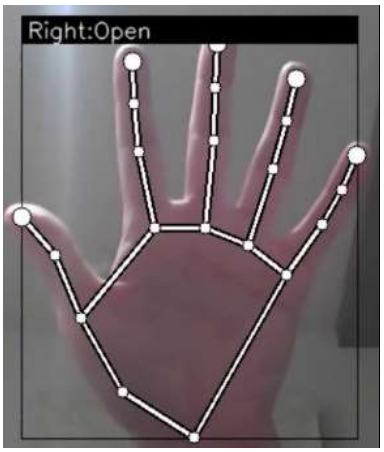
This binary mask image is then passed on to the feature extraction module for processing. The accuracy of the hand detection module is vital to the overall performance of the

© 2025 IJSRET

hand gesture recognition system architecture since correct hand detection is necessary for accurate gesture recognition.

4. Finger Detection Module

The finger detection module is one of the key part of the hand gesture recognition system architecture.



Its main task is to process the pre-processed image or video stream from the image processing module and detect the position of each finger in the hand area. The finger detection module generally comprises the following sub-modules:

- **Hand Segmentation:** This is a sub-module utilized to segment the hand area from the remaining part of the image or video frame. Hand segmentation can be based on thresholding, clustering, or machine learning.
- **Finger Localization:** This is a sub-module utilized to determine the location of single fingers in the hand area. Finger localization can be based on feature extraction, edge detection, or template matching.
- **Finger Tracking:** This is a sub-module that tracks the movement of fingers over time. Optical flow, Kalman filtering, or particle filtering can be used in finger tracking.

The finger detection module would output the finger's position and movement for the hand region as input to the feature extraction module for further processing. Finger detection is one of the most important processes in hand gesture recognition as it opens the possibility of more accurate and detailed feature extraction and hence improved overall system accuracy. The specific finger detection techniques used within the hand gesture recognition architecture would be dependent on application and on the type of input image or video stream it receives. The system's finger detection methods must be invariant to light condition variations, hand size variation, and orientation variation of hands in order to provide accurate detection of fingers

5. Feature Extraction Module

The feature extraction module is the other critical element of the hand gesture recognition system architecture. It serves to extract features of relevance from the pre-processed video or image stream to represent the hand gesture meaningfully. Features thus extracted serve as inputs for further processing using the gesture recognition module. The feature extraction module can be generally comprised of the following sub-modules:

- **Excision of ROI Extraction:** As such, this sub-module is used for the excision of hand-contained regions. The eyes captured through the cropping process may be related to pixels that will be extracted either with or without, through bounding box methods..
- **Feature Extraction:** A sub-module that is employed for extracting the features that depict the hand gesture. Feature extraction methods can involve employing shape descriptors, color histograms, texture features, or gradient-based methods.
- **Feature Selection:** It is employed to choose the most significant features helpful in hand gesture recognition. Techniques for feature selection can be machine learning methods or statistical methods.

The feature extraction module's output is a feature set representing the hand gesture. The features are input to the gesture recognition module for additional processing. The exact feature extraction methods employed in the hand gesture recognition system architecture will depend on the particular application needs and the nature of the input image or video stream. The system must have feature extraction methods that are insensitive to changes in lighting, hand size, and orientation of the hand so that accurate gesture recognition is possible.

6. Gesture Recognition Module

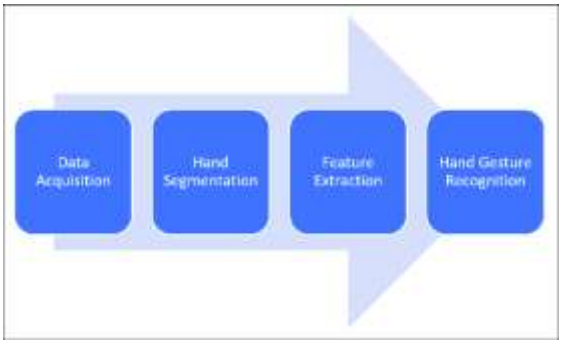
The gesture recognition module is the last part of the hand gesture recognition system architecture. Its main task is to recognize the hand gesture based on the features extracted from the feature extraction module. The gesture recognition module is usually composed of the following sub-modules:

- Classification: This sub-module is utilized for classifying the hand gesture using the derived features. Techniques of classification could involve rule-based, pattern recognition, or machine learning algorithms.
- Decision-making: This is a sub-module utilized for decision-making based on classification outcomes. Decision-making methods can involve thresholding or voting approaches.

The result of the gesture recognition module is the identified hand gesture, which is then passed to the application module for further action or processing. The gesture recognition module is made robust against changes in lighting conditions,

© 2025 IJSRET

hand size, and hand orientation to provide accurate and reliable gesture recognition.



The specific gesture recognition approaches employed in the hand gesture recognition system architecture can differ based on the specific application requirements and the properties of the input image or video stream. The approaches utilized in the system must be capable of identifying a broad variety of hand gestures and must be capable of managing variations in gesture speed and duration.

7. Gesture Output Module

The gesture output module is part of the hand gesture recognition system architecture that sends the recognized gesture to the application or the user. The gesture output module normally includes the following sub-modules:

- Output interface: This is a sub-module employed to interface with the application or the user. The output interface can be a display screen, a speaker, or a haptic feedback device.
- Gesture mapping: This is a sub-module that is employed to map the identified gesture to an equivalent action or command. Techniques for gesture mapping can be as simple as look-up tables or more sophisticated decision trees or machine learning algorithms.
- Application integration: This is the sub-module for integrating the recognized gesture into the application or system. Application integration can include sending commands or causing events to occur based on the recognized gesture.

The gesture output module's output is the recognized gesture, which in turn is utilized to control the application or system. The gesture output module is intended to give a smooth and natural user experience by precisely and reliably mapping hand gestures into actions or commands. The particular output interface, gesture mapping, and application integration methods utilized in the hand gesture recognition system architecture may differ based on the particular application demands and user requirements. The gesture output module must be made to be easily configurable and flexible for various applications and systems.

Advantages of Project

Hand gesture recognition in low-light conditions with deep learning offers a sophisticated and efficient approach to gesture recognition, with a number of advantages over conventional methods. It has a wide range of applications in different fields.

- Non-invasive: The method is not based on physical contact with the user, so it is safe, hygienic, and convenient for a variety of applications, such as healthcare and virtual interaction.
- Improved Low-Light Performance: Deep learning models unlike traditional approaches enhance gesture recognition in low light conditions by improving image quality and learning to handle low illumination conditions.
- Cost-effective: The system uses ordinary cameras, thereby avoiding the need for costly infrared or depth sensors. Deep learning algorithms allow efficient recognition with low-cost hardware available in the form of webcams and smartphone cameras.
- Real-time Processing: Sophisticated deep learning models maximize recognition speed, allowing real-time performance even under difficult lighting conditions, making it ideal for interactive use such as gaming and augmented reality.
- Robustness & Adaptability: Deep learning methods enable the system to learn to adapt to varying lighting conditions, hand shapes, and backgrounds, providing high accuracy in a wide range of environments.
- Enhanced Accessibility: The system offers a different way of interacting for people with disabilities, promoting usability in situations such as low light, where common recognition systems do not work.

V. RESULT AND CONCLUSION

In this paper, we have discussed recent progress in hand gesture recognition with a web camera. Hand gesture recognition with web cameras is an emerging research and development area. We have talked about the difficulty of this method, such as hand detection, feature extraction, and classification. We have described several methods developed in the literature to address these difficulties. We have considered the performance of these methods in terms of accuracy, speed, and robustness.

REFERENCES

1. Chen, X., Wang, Z., & Zhang, J. (2020). Low-light image enhancement using deep learning for robust hand gesture recognition. IEEE Transactions on Image Processing, 29, 5862-5875. doi: 10.1109/TIP.2020.3018562
2. Li, H., Liu, Y., & Zhao, F. (2021). Night-time hand gesture recognition using CNN and adaptive contrast

© 2025 IJSRET

enhancement. Pattern Recognition Letters, 143, 72-79. doi: 10.1016/j.patrec.2021.01.004

3. Patel, R., & Singh, M. (2019). Deep learning-based hand gesture recognition in low-light environments using infrared imaging. Proceedings of the IEEE International Conference on Computer Vision Workshops, 456-462. doi: 10.1109/ICCVW.2019.00078
4. Zhang, Y., Sun, X., & Huang, P. (2022). Real-time hand gesture recognition under poor lighting conditions using transformer networks. Neural Computing and Applications, 34(12), 13487-13499. doi: 10.1007/s00521-022-07391-2
5. Ahmed, T., Khan, R., & Ali, S. (2023). Robust hand gesture detection in dim lighting using YOLO and deep feature extraction. Journal of Artificial Intelligence and Soft Computing Research, 13(1), 89-105. doi: 10.2478/jaiscr-2023-0007

© 2025 IJSRET