

A Systematic Literature Review on the Robustness of Sign Language Recognition Methods in Low-Light Environments

Laveneswary Krishnan

Faculty of Computer Science and Information Technology

University of Malaya

50603 Kuala Lumpur

Malaysia

Email: wma180001@siswa.um.edu.my

ABSTRACT:

This paper presents a comprehensive review of the robustness of hand gesture recognition systems in low light intensity environments spanning the years 2018 to 2023. The primary objective is to assess the progress made during this period and identify areas requiring further attention. An extraction of 20 relevant journals from reputable online databases was conducted using selected keywords. Most of the reviewed articles delve into three crucial aspects of hand gesture recognition systems: data acquisition, data environment, and hand gesture representation. The system performance evaluation reveals that machine learning models achieve recognition accuracy between 94% and 98%, while computer vision models report accuracy within the range of 90% to 95%. The deep learning approach shows a broader accuracy range, spanning from 90% to 98%. Notably, the studies reviewed utilized datasets comprising 37 hand gestures, including 26 letters of American Sign Language (ASL) and numeric gestures ranging from 0 to 9. This paper sheds light on the current state of hand gesture recognition in low light environments and provides insights into potential opportunities for further research and development.

Keywords: sign language, hand gesture, low light intensity, recognition, numeric gesture

1. INTRODUCTION

In recent decades, human communication has been primarily defined using human language, culturally passed down through generations and varying significantly across societies. Hearing impairment, resulting in the loss of the ability to communicate through spoken language, is a challenge faced by many, often referred to as deaf individuals. This limitation hinders their communication with those who have normal hearing. The World Health Organization (WHO) estimated that by the year 2020, there were approximately 466 million deaf individuals globally, a number projected to reach 900 million by 2050.

Sign language recognition is increasingly crucial for human-computer interaction due to its simplicity, directness, and rich communicative potential. Research in this field is generally categorized into several approaches: Inertial sensor-based methods, involving micro electromechanical systems, detect hand posture changes but may affect user naturalness and comfort as the sensors need to be fixed on limbs. Vision based approaches offer high recognition accuracy for various hand gestures but are limited by sight distance and susceptibility to light intensity. Radar based approaches, ensuring natural privacy protection and penetration ability, remain unaffected by external factors like light and dust [1]. Deep learning models show high accuracy in clean datasets. However, recognizing gestures under degraded conditions, such as partial occlusions and low illumination, remains a challenge. Current gesture recognition methods often focus on non-occluded scenarios with a single gesture, overlooking real-world situations where multiple partially occluded gestures may coexist [2]. Hence, this paper provides a summary of Sign Language, gestures, and key methodologies employed in discerning Sign Language in low light intensity environments. It transforms this information into a format understandable through information technology and computer vision, facilitating interaction and communication with users. This technology has the potential to aid in the seamless social integration of individuals with hearing impairments or deafness, enabling them to communicate naturally with others.

This paper follows a structured approach, beginning with an explanation of the concept and importance of Sign Language Gestures. The second section delves into signs language recognition techniques, followed by a third section outlining a survey of background techniques applied to sign language. The last section concludes the paper.

2. SEARCH METHODOLOGY

2.1 Search Strategy

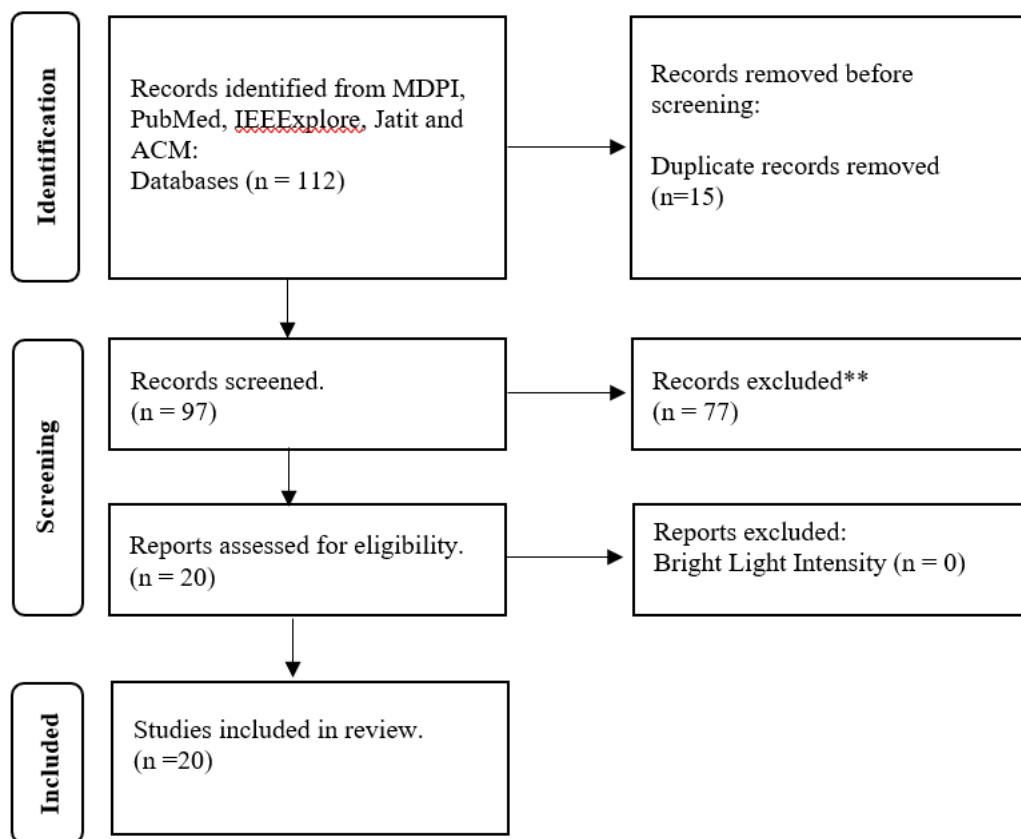


Fig 1: PRISM flow diagram of the study selection process.

This systematic review aims to identify and analyse sign language recognition methods suitable for low-light intensity environments. Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Fig 1), an extensive search of published studies was conducted to gather relevant information. The search period was restricted to articles published between January 31, 2018, and December 1, 2023. Databases including MDPI, PubMed, Jatit, ACM, and IEEE Xplore were explored.

Table 1: Database and Search Terms

Database	Search Terms
MDPI	[hand gesture] or [sign language] or [numeric gesture] AND [low light intensity] or [lighting condition]
PubMed	[hand gesture] or [sign language] or [numeric gesture] AND [low light intensity] or [lighting condition]
ACM	[hand gesture] or [sign language] or [numeric gesture] AND [low light intensity] or [lighting condition]
IEEEExplore	[hand gesture] or [sign language] or [numeric gesture] AND [low light intensity] or [lighting condition]
Jatit	[hand gesture] or [sign language] or [numeric gesture] AND [low light intensity] or [lighting condition]

The search utilized text words specifically targeting "sign language," "hand gesture," "lighting conditions," and "low light intensity". This tailored set of search terms is outlined in Table 1. Only publications in the English language were considered, with limitations set on the publication year. Additionally, the reference lists of the studies included in the literature review were scrutinized for pertinent citations related to sign language hand gesture recognition in low-light intensity environments.

This systematic approach ensures a comprehensive exploration of relevant studies and provides a foundation for understanding the landscape of sign language recognition methods under challenging lighting conditions.

2.2 Inclusion and Exclusion Criteria

English language research articles on sign language recognition methods in low light intensity environments were included based on the following criteria: (1) Recognition methods and (2) Low Light Intensity Environment.

The data and information extracted from the articles include: (1) Authors, (2) Journal, (3) Year of publication, (4) Study design, (5) Study period, (6) Light Intensity, (7) Recognition Rate, (8) Dataset (9) Research Gaps.

All articles in languages other than English were excluded. Additionally, any research articles and abstracts not relevant to the specified keywords, along with duplicates and those unpublished in peer-reviewed journals, were also excluded in this study.

2.3. Data and information extraction and synthesis

The data and information extracted from numerous research articles for this study include authors, journal, year of publication, study design, study period, and recognition methods in low light intensity environments. These findings were compared narratively, and the identified gaps in this study were discussed. All selected articles were obtained from search engines on databases dedicated to indexing relevant Computer Science journals published between 2018 and 2023.

2.4. Ethical approval

Ethics clearance was obtained from University of Malaya (UM). Informed consent was not required for this study, as this study focused on review of already published data. The study did not involve the use of human or animal subjects.

3. Results

3.1. Literature search and selection of eligible articles

Out of 112 research articles that were selected in this study, 15 of the selected articles were not included as part of the study after eliminating duplicates. After the titles and abstracts were screened, 97 complete text articles scale through the next stage for re-evaluation. Ultimately, 20 articles were found to have fulfilled the inclusion criteria, and therefore included in the systematic review. All the 20 selected articles were from search engines, through databases dedicated to indexing all original data relevant to sign language recognition in low light intensity published between 2018 and 2023. The 20 publications selected from this study were eligible for more detailed review in the final analysis. The PRISMA flow diagram of the study selection process is shown in Fig1.

Numerous researchers have organized the sign recognition system into four stages following the acquisition of input images from cameras and videos. These stages encompass tracking of hand and segmentation, feature extraction methods, classification, and recognition, as depicted in Fig 2.

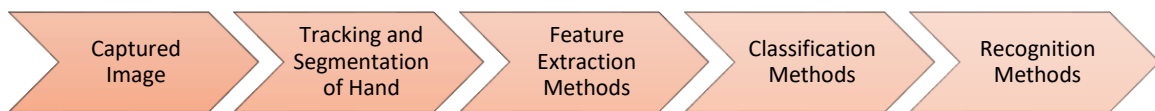


Fig 2: Stages of Sign Language

3.2 Hand Gesture Recognition in Low Light Intensity Environments

The subsequent section of this paper elaborates on the application of these sign recognition steps by numerous researchers. Primarily, various sign recognition techniques are developed and tested in low light intensity environments, including:

- LSTM: Long Short-Term Memory
- SSVM: Structural Support Vector Machine
- SVM: Support Vector Machine
- CNN: Convolutional Neural Network
- ANN: Artificial Neural Network
- Resnet: Residual Network
- 3D CNN: 3-Dimensional Convolutional Neural Network
- HS-CbCr: Hue-Saturation Chrominance - Chroma and Luminance components
- YCbCr: Luminance, Chroma Blue, and Chroma Red components

Machine Learning

- LSTM: Long Short-Term Memory

(Zhuang, Yang, 2021) presented a high-performing LSTM machine learning model with a remarkable 98% accuracy rate for hand gesture recognition. Leveraging radar technology, the approach stood out for its robustness in various conditions and its resistance to external factors like light and dust. The traditional reliance on image-based representations in radar methods had led to challenges such as large model sizes and computational complexity. To address these issues, the innovative method utilized two-dimensional trajectory features obtained from bistatic radars, involving signal preprocessing, target detection, and an adaptive polynomial fitting process. This process formed the basis for a lightweight variable-length LSTM designed for efficient hand gesture identification. The LSTM's adaptability to varying input sequence lengths addressed challenges associated with model size and complexity. The research marked a significant advancement in radar-based gesture recognition, holding promise for applications in human-computer interaction and beyond [1].

- SSVM: Structural Support Vector Machine

(Kumar, Yadav, Gupta, 2018) focused on hand gesture recognition, particularly for sign language, outlining challenges and the need for an effective recognition system. The paper presented an eight-step process for hand gesture recognition, involving image acquisition, skin colour-based segmentation, background removal, Canny edge detection, PCA feature extraction, support vector machine classification, and data evaluation. The proposed system aimed to recognize sign language efficiently without the need for gloves or sensors, making it user-friendly and applicable in various fields. The system's effectiveness was demonstrated through a dataset of 100 images capturing 16 different gestures under varied lighting conditions. Colour-based segmentation and Canny edge detection were employed for classification, testing, and training. The system achieved an impressive accuracy of 94.5% in identifying 15 hand gestures, highlighting its potential applications in security, and facilitating sign language communication for the deaf community. The implementation of the system was successful, with a prototype developed, promising advancements in gesture recognition technology [3].

- SVM: Support Vector Machine

(Liang, Wang, Greco, 2020) introduced an interferometric radar system with one transmitting antenna and two receiving antennas to enhance hand gesture recognition. It tackled challenges faced by micro-Doppler signatures in recognizing gestures affected by aspect angles. The system extracted three features from micro-Doppler and interferometric spectrograms and employed a support vector machine for classification. The interferometric radar provided two-dimensional micro-motion information, improving spatial resolution and gesture classification performance. The approach overcame issues related to aspect angle variations and effectively recognized horizontally symmetric gestures. Experimental results demonstrated the system's success in distinguishing among nine gestures across different aspect angles (0°, 15°, 30°, and 45°) with accuracy rates of 96.8%, 99.3%, 97.2%, and 95.3%, respectively. The findings underscored the system's robustness and improved accuracy in hand gesture recognition, attributed to the incorporation of additional transversal micromotion features [4].

Computer Vision

- YCbCr

(Perimal, Safar, Yazid, 2018) focused on hand gesture recognition and highlighted the impact of light intensity on the algorithm's performance. It emphasized the need for optimal brightness levels for successful hand gesture detection. The proposed algorithm consisted of four components: image acquisition, pre-processing, finger detection, and gesture recognition. Images captured with a high-resolution webcam were processed in MATLAB, where brightness alteration experiments revealed optimal performance when increasing image brightness. The algorithm utilized the YCbCr color space for its linear luminance and chromaticity properties. The study demonstrated that increasing image brightness enhanced the success rate of hand gesture detection. Conversely, a decrease in brightness below 0.1% resulted in a 50% decrease in detection accuracy, with total failure below 0.01%. This finding emphasized the critical role of light intensity in the algorithm's effectiveness for hand gesture recognition [5].

(Chansri, Srinonchat, Lim, 2019) addressed wireless control of electronic devices through hand gestures in complex environments, emphasizing the need for a reliable integration of an embedded system. The study proposed an integrated system using Raspberry Pi and an RGB camera for wireless gesture control. The novel radian fingertip analysis technique eliminated the need for data training, providing robust performance in varying light conditions and complex backgrounds. Experiments on 12 American Sign Language fingerspelling gestures yielded an accuracy of 90.83%, as indicated by the confusion matrix. The integration of Raspberry Pi proved successful in enabling wireless control through hand gestures, highlighting the system's potential for practical applications in diverse environments [6].

- HS-CbCr

(Rahmat, Chairunnisa, Gunawan, 2019) employed computer vision for hand gesture recognition in human-computer interaction, facing challenges in recognizing gestures with bare hands and in complex backgrounds. Background and skin detection issues necessitated a robust solution. The proposed method involved multiple stages, including image acquisition, resizing, colour space conversion, skin detection using the HS-CbCr format, and addressing background challenges through averaging. Further steps included grayscale processing, background accumulation, frame differencing, thresholding, inversion, and image enhancement. Feature extraction was performed using contour, convex-hull, and convexity defects, followed by finger counting and hand direction determination. The resulting instruction was then used to control applications like slideshow presentations, video players, music players, or PDF readers. Experimental results showed the method's effectiveness, achieving up to 98.71% accuracy in well-lit conditions. Lighting conditions significantly influenced accuracy, with lower illumination correlating to higher accuracy (95.03%). Future improvements involved incorporating a hand tracking method for dynamic gesture recognition and considering machine learning for enhanced object recognition accuracy. Addressing lighting challenges in skin detection was recommended for future work in advancing hand gesture recognition [7].

Deep Learning

- ANN

(Haroon, Altaf, Ahmad, 2022) focused on American Sign Language (ASL), storing gesture images in a database and converting them to PNG format for subsequent analysis. Addressing challenges related to brightness variations, the luminosity method was employed for preprocessing. Grayscale conversion ensured uniform contributions of red, green, and blue colors, a crucial step in enhancing the robustness of the system under diverse lighting conditions. The Scale-Invariant Feature Transform (SIFT) method was used to extract crucial features such as perimeter, hand size, center of hand, and finger distance. The recognition process commenced with acquiring depth images using the Kinect sensor, offering a detailed representation of hand shapes crucial for sign language gestures. The journal presented a sensor-based sign language gesture recognition system that achieved a remarkable 97.4% accuracy and a minimal 2.6% error rate, highlighting high efficiency. The proposed framework overcame challenges in varying lighting conditions and incorporated advanced techniques to enhance accuracy. The study suggested future directions, including hand tracking methods, machine learning exploration, and addressing skin detection challenges [8].

(Yang, Zhang, 2019) introduced a real-time hand gesture recognition model utilizing surface electromyography (sEMG) signals. The approach involved acquiring sEMG signals through a MYO armband, employing a sliding window method for data segmentation and feature extraction. A feedforward artificial neural network (ANN) was established and trained with a dedicated dataset. The recognition process involved reaching a threshold of activation times by the ANN classifier for the recognized label to trigger gesture recognition. This model was advantageous for hand gesture recognition as sEMG sensors were immune to variations in light, position, and hand orientation. During the experiment, genuine sEMG data were gathered from twelve participants, and the model was assessed using a collection of five gestures per subject. The average recognition rate reached 98.7%, accompanied by an average response time of 227.76 ms—merely one-third of the gesture duration. This suggested that the pattern recognition system held the potential to identify a gesture even before its completion [9].

- CNN

In the exploration of relevant literature, 10 out of 20 selected papers have employed Convolutional Neural Networks (CNN) for the development of hand recognition systems.

(Krishnan, Joshi, Connor, 2020) utilized 3D data obtained through integral imaging as input for a Convolutional Neural Network (CNN). The CNN's spatial features, derived from convolutional and pooling layers, were then fed into a Bi-directional Long Short-Term Memory (BiLSTM) network to capture temporal variations. A comparative analysis with conventional 2D imaging and other methodologies, such as spatio-temporal interest points with support vector machines (STIP-SVMs) and distortion invariant non-linear correlation-based filters, was conducted. Experimental results suggested that the proposed 3D integral imaging-based deep learning approach held promise, particularly in degraded environments, exhibiting improved performance under conditions like occlusion and low light when compared to conventional 2D imaging and other considered methodologies. Significantly, the percent reduction in Area Under the Curve (AUC) for the 3D imaging-based CNN-BiLSTM network due to simulated low illumination was notably lower than that of the 2D imaging-based approach. The potential applications of the proposed approach extended to detecting human activities and sentiment analysis, particularly in more challenging conditions. Future research directions may explore alternative integral imaging approaches, compare with time-of-flight sensing, and assess performance in scenarios with increased scene complexity [2].

(Luo, Cui, Li, 2021) presented a Convolutional Neural Network (CNN) model for recognition. The complexity and diversity of gestures posed challenges to gesture recognition, influenced by environmental factors, lighting conditions, and occlusion. To improve recognition, the process began by filtering hand skin color in the YCbCr color space, isolating the gesture region. The Gaussian filter addressed edge noise, followed by morphological gray open operation and the watershed algorithm for contour segmentation. The eight-connected filling algorithm enhanced gesture features. In this experiment, the model detected ten gestures from 0 to 9. Experimental results demonstrated the suggested method's rapid and accurate recognition, achieving an average success rate of 96.46%, without a significant increase in recognition time [10].

(Jung, Kim, Shin, 2022) introduced a dedicated System-on-Chip (SoC) tailored for processing Frequency Modulated Continuous Wave (FMCW) radar signals for Human Gesture Recognition (HGR). The HGR processor integrated a spectrogram classifier powered by a lightweight Convolutional Neural Network (CNN) and incorporated a preprocessing method to enhance classification accuracy, accommodating diverse gestures from 72 participants. Implemented through a 40nm CMOS process, the HGR processor achieved a 91.5% accuracy in recognizing five distinct gestures, demonstrating user-independent capability. The suggested SoC's execution at the chip level extended its applicability to compact and energy-efficient scenarios, indicating its potential for practical, real-world implementations [11].

(Breland, Skriubakken, 2021) presented a deep learning-based sign language digits recognition system from thermal images with an edge computing system. A MEMS thermal camera, Omron D6T, was integrated into a thermal image dataset generation system for sign digits based on hand gestures. Connected to a Raspberry Pi embedded system, a custom-designed shield ensured proper thermal camera functioning and prevented interference. The shield, created with Altium Designer, featured two layers and correct pin connections. The Omron D6T captured thermal images of hand gestures, resulting in a dataset of 3200 images categorized by sign language digits. A custom software interface facilitated image capture, yielding 320 hand gestures for each sign digit. The proposed lightweight CNN model achieved an impressive 99.52% accuracy on the test dataset, highlighting the embedded system's capability for accurate recognition of low-resolution thermal images of hand gestures. The system demonstrated consistent accuracy across various background lighting conditions, emphasizing the suitability of thermal imaging for hand gesture recognition in low-light environments. The study focused on end-to-end recognition based on thermal imaging, leaving the exploration of complex backgrounds for future research. The accuracy achieved, irrespective of lighting conditions, highlighted the potential of thermal imaging in challenging environments, while future work would delve into a comprehensive solution for hand gesture recognition using thermal cameras [12].

(Shi, Chen, 2023) introduced a sign language recognition system using multiple magnetic sensors to interpret hand gestures corresponding to sign language alphabets. The system employed six magnetic sensor nodes to measure finger and palm orientations, and a deep learning classification algorithm processed the orientation data. Experimental tests confirmed the system's effectiveness, achieving close to 100% classification accuracy for 26 sign language alphabets in laboratory conditions. The results supported the feasibility of the proposed gesture recognition system for automatically translating sign language alphabets, addressing the communication

needs of hearing-impaired individuals. The gesture recognition system utilizing the Earth's stable magnetic field operated reliably and consistently. It remained unaffected by occlusion and poor lighting conditions, offering users unrestricted application. Future developments included wireless connectivity for each magnetic field sensor node to eliminate cable constraints on hand movements. The proposed gesture recognition system had the potential to contribute to the creation of an efficient automatic sign language translation system, facilitating effective communication for hearing-impaired individuals. Additionally, it held promising applications in human-computer interaction and robotics [13].

(Aishwarya, Yogitha, 2022) proposed a methodology that focused on developing a vision-based prototype to interpret Indian Sign Language into speech or text on an embedded device. Utilizing deep learning techniques, the model employed an integrated webcam for real-time hand gesture capture and Raspberry Pi for processing. The dataset, comprising 1200 images for each gesture, was partitioned into training and test data (9:1 ratio), integrating CNN, IoT, and Python. Evaluation with a distinct test dataset, featuring 43,200 image samples, showed the prototype's accuracy exceeding 95%, even in varied lighting conditions. The model effectively recognized Indian Sign Language, emphasizing the efficiency of the Convolutional Neural Network (CNN) in hand gesture categorization. Future work aimed to extend the model to include more sign languages and enhance response time [14].

(Zhao, Sark, Krstic, 2022) introduced a synthetic FMCW RADAR data generator for hand gestures, simulating six gestures with range, velocity, and angle features for neural network training. Utilizing the VGG19 pre-trained model for feature extraction, the study compared XGBoost and random forest classifiers on a real dataset. The XGBoost classifier, trained with synthetic data, achieved an 87.53% average accuracy on the real dataset, addressing the challenge of limited training data. The combination of 2% actual data and synthetic data enhanced the XGBoost classifier's accuracy to 94.63%. Notably, the XGBoost classifier outperformed the random forest classifier when integrating synthetic data. Future research would explore gesture recognition within elevated frequency bands, further advancing this innovative approach in overcoming data limitations for robust machine learning-based gesture recognition [15].

(Fang, Ding, Sheng, 2019) proposed convolutional neural networks and deep convolution generative adversarial networks. Traditional gesture recognition methods involving manual feature value extraction were time-consuming. It demonstrated commendable results in expression recognition, calculation, and text output. Experimental findings showed effective model training with fewer samples, leading to improved gesture classification and detection outcomes. The method exhibited resilience to illumination and background interference, achieving efficient real-time recognition with a recognition rate of 90.45%. Although the current network supported calculations and text output, there was potential for expansion by incorporating additional gestures, enabling applications such as playing games, chat conversations, and composing emails in the future [16].

(Nethu, Suguna, Sathish, 2020) presented a hand gesture detection and recognition methodology based on deep learning convolutional neural networks (CNN). Fingertips were segmented from hand gesture images and used as input for the CNN classifier, which was trained on a publicly available dataset. The methodology's performance was evaluated in terms of sensitivity, specificity, accuracy, and recognition rate. Recognizing human hand gestures was crucial for advancing automated vehicle movement systems, and this work focused on achieving this through a CNN classification approach. The process involved hand region segmentation using mask images, finger segmentation, normalization, and recognition using the CNN classifier. Adaptive histogram equalization enhanced pixel contrast, and fingertip segmentation was achieved through connected component analysis. The methodology surpassed state-of-the-art methods, achieving notable metrics, including 98.1% sensitivity, 93.4% specificity, 96.2% accuracy, and a recognition rate of 96.2%. The effectiveness of the approach demonstrated its potential for applications requiring robust hand gesture detection and recognition [17].

(Tran, Ho, 2020) proposed a system that utilized in-depth skeleton-joint information images from a Microsoft Kinect Sensor version 2 to extract the hand region of interest. Hand contours were then extracted and described using a border-tracing algorithm, with fingertip location detected through the K-cosine algorithm based on the hand-contour coordinates model. Fingertip detection led to gesture initialization and hand gesture identification. A 3D convolutional neural network was employed for accurate gesture recognition, integrating geometry algorithms and deep learning methods. The approach demonstrated high accuracy and practical applicability, excelling in varied conditions such as changing light levels, complex backgrounds, detection at longer distances, and recognizing hand gestures from multiple individuals, with a recognition rate of 92.6%. Future efforts aimed to expand the system's capabilities for additional hand gestures and extend its application to various practical scenarios [18].

- ResNet

(Wang, Zhang, 2023) presented a gesture recognition technology employing millimeter-wave radar. The process involved collecting gesture data and extracting distance and Doppler information, resulting in distance-time and Doppler-time maps. These maps were fused into a feature fusion map to comprehensively represent gesture information. Recognition and classification were accomplished using the Resnet101 network structure. Radar-based gesture recognition held significance in Human-Computer Interaction (HCI) technology, applicable in areas like automotive collision avoidance, medical care, and entertainment. Radar technology's advantage lay in its independence from site and environmental factors, light intensity, and easy placement. The study applied millimeter-wave radar to collect and process seven defined hand gesture signals, utilizing Fourier transform and residual networks for accurate recognition. Experimental outcomes demonstrated a 98.58% recognition accuracy for seven gestures, with the feature fusion map effectively utilizing distance and velocity features for high accuracy. Notably, the recognition accuracy achieved through feature fusion was 5%-6% higher than using a single feature map, highlighting the effectiveness of this approach in maximizing gesture feature information for precise recognition and classification [19].

- 3D CNN

(Zhang, Lan, 2020) presented a comparative exploration of deep neural networks for non-contact hand gesture recognition using millimeter-wave FMCW radar. The processing of range-doppler maps involved a zero-filling strategy to enhance range and velocity information related to gesture motions. Two optimal deep neural network architectures, 3D-CNN and CNN-LSTM were individually designed to capture temporal signatures encoded in multiple adjacent radar chirps. The study assessed the effectiveness of these architectures for hand gesture recognition using commercial mmW FMCW radar, achieving a recognition accuracy of 95% for six common hand gestures. The introduced zero-filling strategy contributed to swift convergence in recognition accuracy during training. The paper also investigated the impact of training data size on recognition accuracy. The proposed methods showed applicability in recognizing subtle finger motions, and preliminary experimental results were provided, comparing them with other baseline methods [20].

4. DISCUSSION AND CONCLUSION

Table 2: Sign Language Recognition Methods and Accuracy

No	Machine Learning Model			Computer Vision Model		Deep Learning Model				Performance
	LSTM	SSVM	SVM	YCbCr	HS-CbCr	ANN	CNN	ResNet	3D CNN	
1	✓									98
2							✓			N/A
3		✓								94.5
4			✓							99.3
5				✓						N/A
6				✓						90.83
7					✓					95.03
8						✓				97.4
9						✓				98.7
10							✓			96.46
11							✓			91.5
12							✓			99.52
13							✓			N/A
14							✓			95
15							✓			94.63
16							✓			90.45
17							✓			96.2
18							✓			92.6
19								✓		98.58
20									✓	95

Table 3: Dataset and Research Gaps

No	Machine Learning Model			Computer Vision Model		Deep Learning Model				Dataset	Research Gap
	LSTM	SSVM	SVM	YCbCr	HS-CbCr	ANN	CNN	ResNet	3D CNN		
1	✓									9 Hand Gestures	This method can be expanded to detect more sign language gestures.
2							✓			2 Hand Gestures	The method can be expanded to detect human activities in challenging environments.
3		✓								15 Hand Gestures	This method can be expanded to detect more sign language gestures.
4			✓							9 Hand Gestures	This method can be expanded to detect more sign language gestures.
5				✓						10 Number Gestures	This method can be expanded to detect more sign language gestures.
6				✓						12 ASL Letters	This method can be expanded to detect more sign language gestures.
7					✓					9 Hand Gestures	This method can be expanded to detect dynamic hand gesture and for the skin detection stage, it is recommended to adopt a method that effectively addresses lighting condition challenges.
8						✓				26 ASL Letters	High-resolution cameras can capture intricate depth images, and this method can be adapted to incorporate various languages, including regional sign language, while enhancing recognition complexity with additional features from various angles and lighting conditions.
9						✓				5 Hand Gestures	This method can be expanded to detect more sign language gestures.
10							✓			10 Number Gestures	This method can be expanded to detect more sign language gestures.
11							✓			5 Hand Gestures	This method can be expanded using thermal cameras, to improve accuracy in challenging backgrounds by incorporating hand-crafted features.
12							✓			10 Number Gestures	This method can be expanded to detect more sign language gestures.
13							✓			26 ASL Letters	This method can be expanded to detect other sign language gestures.
14							✓			36 Hand Gestures	This method can be expanded to improve response time and detect other sign languages.
15							✓			6 Hand Gestures	This method can be expanded to recognize gesture within elevated frequency bands.
16							✓			37 Hand Gestures	This method can be expanded to detect other sign languages.
17							✓			8 Hand Gestures	This method can be expanded to detect more sign language gestures.
18							✓			7 Hand Gestures	This method can be expanded to detect more sign language gestures.
19								✓		7 Hand Gestures	This method can be expanded to detect more sign language gestures.
20									✓	6 Hand Gestures	This method can be expanded to detect more sign language gestures.

TABLE 2 provides a summary of the accuracy achieved by sign language recognition methods when utilizing input images. TABLE 3 offers a summary of the number of datasets and research gaps identified in sign language recognition methods. In this paper, Sign Language recognition methods are well explained, and we focused on methods for classification process to identify the gestures for Sign Language. Several recent algorithms, published between 2018 and 2023, have been presented and compared. Most papers presented in the last years focused on using the CNN algorithm. As future work, the system will be building to process dynamic real video and interpreter in real-time applications.

The identified research gaps in sign language recognition research offer valuable opportunities for refinement in the field. Addressing limitations related to datasets, enhancing model robustness, and understanding contextual challenges can significantly contribute to the progress of sign language recognition technologies. The acknowledged limitations in existing studies, such as the call for more diverse sign language gestures, higher resolution cameras, and testing in low-light conditions, provide meaningful insights for potential improvements in future research. The papers utilized a private dataset, specifically tailored for low-light environments, highlights the limited focus on this aspect in existing research. With only 3 out of 20 papers utilizing the American Sign Language (ASL) dataset, there is a clear gap in addressing low-light conditions. No standard dataset is available on Sign Language gesture recognition in low light intensity, so a new dataset is created for some research. The authors' decision to create and employ their dataset directly addresses this gap, allowing for targeted testing in low-light settings.

Declaration of competing interest

Authors state no conflict of interest.

Acknowledgements

The authors would like to acknowledge the University of Malaya (UM) for this study.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit organizations.

Informed consent

Informed consent is not applicable.

Ethical approval

The conducted research is not related to either human or animal use.

REFERENCES

- [1] Zhuang, L., & Yang, Z. (2021). *A low-complexity hand gesture recognition method based on 2D-trajectory feature using bistatic radars*. 2021 CIE International Conference on Radar (Radar). <https://doi.org/10.1109/radar53847.2021.10027984>
- [2] Krishnan, G., Joshi, R., O'Connor, T., Pla, F., & Javidi, B. (2020). *Human gesture recognition under degraded environments using 3D-integral imaging and deep learning*. *Optics Express*, 28(13), 19711. <https://doi.org/10.1364/oe.396339>
- [3] Kumar, S., Yadav, G., Singh, H. P., Malhotra, S., & Gupta, A. (2018). *SSVM classifier and hand gesture-based sign language recognition*. 2018 International Conference on Intelligent Circuits and Systems (ICICS). <https://doi.org/10.1109/icics.2018.00098>
- [4] Liang, H., Wang, X., Greco, M. S., & Gini, F. (2020). *Enhanced hand gesture recognition using continuous wave interferometric radar*. 2020 IEEE International Radar Conference (RADAR). <https://doi.org/10.1109/radar42522.2020.9114807>
- [5] Perimal, M., Basah, S., Safar, M., & Yazid, H. (2018). *Hand-Gesture Recognition-Algorithm based on Finger Counting*. *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, 10(1-13), 19–24. Retrieved from <https://jtec.utem.edu.my/jtec/article/view/4115>
- [6] Chansri, C., Srinonchat, J., Lim, E. G., & Man, K. L. (2019). *Low cost hand gesture control in complex environment using Raspberry Pi*. 2019 International SoC Design Conference (ISOCC). <https://doi.org/10.1109/isocc47750.2019.9027669>
- [7] Rahmat, R.F., Chairunnisa, T., Gunawan, D., Pasha, M.F., & Budiarto, R. (2019). *Hand Gestures Recognition with Improved Skin Color Segmentation in Human-Computer Interaction Applications*.
- [8] Haroon, M., Altaf, S., Ahmad, S., Zaindin, M., Huda, S., & Iqbal, S. (2022). *Hand gesture recognition with symmetric pattern under diverse illuminated conditions using artificial neural network*. *Symmetry*, 14(10), 2045. <https://doi.org/10.3390/sym14102045>
- [9] Yang, K., & Zhang, Z. (2019). *Real-time pattern recognition for hand gesture based on ann and surface EMG*. 2019 IEEE 8th Joint International Information Technology and Artificial Intelligence Conference (ITAIC). <https://doi.org/10.1109/itaic.2019.8785894>
- [10] Luo, Y., Cui, G., & Li, D. (2021). *An improved gesture segmentation method for gesture recognition based on CNN and YCbCr*. *Journal of Electrical and Computer Engineering*, 2021, 1–9. <https://doi.org/10.1155/2021/1783246>
- [11] Jung, Y., Kim, J., Shin, M., Ju, H., Cho, K.-I., Lee, Y., Jung, Y., & Lee, S. (2022). *An integrated spectrogram classifier-based hand-gesture recognition SOC for an FMCW radar sensor*. 2022 IEEE Radar Conference (RadarConf22). <https://doi.org/10.1109/radarconf2248738.2022.9764219>
- [12] Breland, D. S., Skriubakken, S. B., Dayal, A., Jha, A., Yalavarthy, P. K., & Cenkeramaddi, L. R. (2021). *Deep learning-based sign language digits recognition from thermal images with edge computing system*. *IEEE Sensors Journal*, 21(9), 10445–10453. <https://doi.org/10.1109/jsen.2021.3061608>
- [13] Shi, B., Chen, X., He, Z., & Han, R. (2023). *Development of magnetic-sensor-based hand gesture recognition system for sign language*. 2023 IEEE 6th International Electrical and Energy Conference (CIEEC). <https://doi.org/10.1109/cieec58067.2023.10167092>
- [14] P, R. S., C, J., Aishwarya, R., R, Yogitha., & Mathivanan, G. (2022). *Hand gesture interpretation model for indian sign language using neural networks*. 2022 IEEE 7th International Conference for Convergence in Technology (I2CT). <https://doi.org/10.1109/i2ct54291.2022.9825322>

- [15] Zhao, Y., Sark, V., Krstic, M., & Grass, E. (2022). *Synthetic Training Data Generator for hand gesture recognition based on FMCW radar*. 2022 23rd International Radar Symposium (IRS). <https://doi.org/10.23919/irs54158.2022.9904997>
- [16] Fang, W., Ding, Y., Zhang, F., & Sheng, J. (2019). *Gesture recognition based on CNN and DCGAN for calculation and text output*. *IEEE Access*, 7, 28230–28237. <https://doi.org/10.1109/access.2019.2901930>
- [17] Neethu, P. S., Suguna, R., & Sathish, D. (2020). *An efficient method for human hand gesture detection and recognition using deep learning convolutional neural networks*. *Soft Computing*, 24(20), 15239–15248. <https://doi.org/10.1007/s00500-020-04860-5>
- [18] Tran, D.-S., Ho, N.-H., Yang, H.-J., Baek, E.-T., Kim, S.-H., & Lee, G. (2020). *Real-time hand gesture spotting and recognition using RGB-d camera and 3d Convolutional Neural Network*. *Applied Sciences*, 10(2), 722. <https://doi.org/10.3390/app10020722>
- [19] Wang, Y., Zhang, J., & Zhao, X. (2023). *Research on hand gesture recognition based on Millimeter Wave Radar*. 2023 3rd International Conference on Consumer Electronics and Computer Engineering (ICCECE). <https://doi.org/10.1109/iccece58074.2023.10135494>
- [20] Zhang, G., Lan, S., Zhang, K., & Ye, L. (2020). *Temporal-range-doppler features interpretation and recognition of hand gestures using MMW FMCW radar sensors*. 2020 14th European Conference on Antennas and Propagation (EuCAP). <https://doi.org/10.23919/eucap48036.2020.9135694>