



A Hybrid Approach to Arabic Sign Language Recognition by HSV Space and Deep Learning Classification

Noor Fadel Hussain^{1*}, Hiba Al-Khafaji²

¹Information Security, College of Information Technology, University of Babylon, noor.fadel@uobabylon.edu.iq, Babil, Iraq.

²Software, College of Information Technology, University of Babylon, hibamj.alkhafaji@uobabylon.edu.iq, Babil, Iraq.

*Corresponding author email: noor.fadel@uobabylon.edu.iq ; mobile: 07501921586

نهج هجين للتعرف على لغة الإشارة العربية باستخدام مساحة الالوان وتصنيف التعلم العميق (HSV)

Accepted: 27/10/2025

Published: 31/12/2025

ABSTRACT

Understanding and automatically recognizing sign language is crucial for equal opportunities for deaf and hard-of-hearing individuals in society. In the world of Arabic sign language communication, people share their ideas, thoughts, and feelings and engage in meaningful daily interactions. While artificial intelligence and computer vision technologies have improved, studies focused on Arabic sign language processing have come first and continue to be a research frontier compared to other world languages. This research proposed a methodology to leverage the HSV color space model for initial segmentation to eliminate background noise to refine the input data. This process segments the hand or semantic sign area. Then, the data is passed onto a Convolutional Neural Network (CNN) for precise sign classification. This method's novelty is the coordinated interplay of differential data simplification through color space manipulation and the profound representational power of the CNN for improved overall system performance for hood sign recognition.

The research also seeks to fill an existing gap in the field of Arabic signal processing by presenting an integrated model, opening up broad horizons for applications in the fields of education, healthcare, and smart government services. The proposed method showed promising results, as the results of the used criteria ranged between Accuracy 95.3%, Precision 93.5%, Recall 92.8%, and F1-score 93.1%.

Background:

Recognizing sign languages has always been a vital topic in societies, especially with the recent massive expansion of the deaf and mute community. It serves as the main form of communication for over 70 million deaf and hard-of-hearing individuals globally. Thus, the development of simple, accurate, and effective methods that utilize readily available resources deserves attention. Among the various forms of sign language, sign language and Arabic are especially significant, as Arabic is predominantly spoken in the Middle East and North Africa. Yet, the challenges that arise from the communication gap between this population and the hearing members of society continue to affect various areas of the deaf and hard-of-hearing individuals educational, social, and professional for educational, social, and, professional integration. Hence, creating automated systems that can comprehend Arabic Sign Language is an essential development in helping socially integrate and bridge communication gaps [1].

Materials and Methods:

The proposed hand gesture recognition system relies on a multi-stage methodology aimed at high-resolution image processing and efficient feature extraction. The methodology begins with the data collection phase,



where a set of images and videos representing letters and words in Arabic Sign Language is collected on the Kaggle Dataset.

Results:

The proposed Arabic Sign Language recognition system was developed through two principal stages: (1) segmentation using the HSV color space with experimentally defined thresholds, and (2) classification via a Convolutional Neural Network (CNN). The segmentation stage, relying on the HSV model with the defined lower range ($H=0, S=30, V=60$) and upper range ($H=20, S=150, V=255$), demonstrated effectiveness in isolating the hand region from complex backgrounds, contributed significantly to noise reduction and the removal of irrelevant small objects, and enhanced the clarity of the extracted hand region prior to classification.

Conclusion:

Recognizing sign languages has always been a vital topic in societies, especially with the recent massive expansion of the deaf and mute community. Therefore, it is important to focus on easy and highly accurate techniques that can be relied upon with relatively available resources. The proposed methodology is highly flexible and scalable, as the HSV color space boundaries can be modified or the CNN architecture optimized to suit research requirements and the characteristics of different datasets.

Key words:

Sign language, Arabic Sign language, HSV, convolutional neural network.

INTRODUCTION

Recognizing sign languages has always been a vital topic in societies, especially with the recent massive expansion of the deaf and mute community. It serves as the main form of communication for over 70 million deaf and hard-of-hearing individuals globally. Thus, the development of simple, accurate, and effective methods that utilize readily available resources deserves attention. Among the various forms of sign language, sign language and Arabic are especially significant, as Arabic is predominantly spoken in the Middle East and North Africa. Yet, the challenges that arise from the communication gap between this population and the hearing members of society continue to affect various areas of the deaf and hard-of-hearing individuals educational, social, and professional for educational, social, and, professional integration. Hence, creating automated systems that can comprehend Arabic Sign Language is an essential development in helping socially integrate and bridge communication gaps [1].

Over the past few decades, the advancements in AI and computer vision have significantly advanced the creation of systems that can recognize and interpret signs and gestures. However, the predominant focus is still on globally spoken languages, like English and American Standard Sign Language, while Arabic is still in the nascent stages of development. This illustrates the need for innovative models that can address the nuances of the Arabic language and its diverse dialects [2].

Difficulties in recognizing Arabic Sign Language. Language recognition systems encounter numerous obstacles, as follows:

- Underdeveloped benchmark databases for Arabic signs focused recognition. The vast differences in signed languages across the Arab world.
- Problematics of isolating the arm or hand from the scene, more so in uncontrolled scenarios.



RELATED WORK

[12] This study focused on building a hybrid system based on two sensors: a Kinect V2 to capture body and arm motion, and a Leap Motion to capture fine finger movements with high accuracy. The research involved classifying signals using classical machine learning methods, specifically Support Vector Machines (SVMs). Accurately distinguishing between signals that differ only minimally during finger movements was facilitated by sensor feature fusion. While this exemplifies the value of using multiple modalities for sign language recognition, it introduces costs and complexities, as evidenced by having to two separate devices.

[13] For this research, the Google MediaPipe framework was used for hand feature detection and a combination of a CNN for spatial feature extraction and an LSTM for temporal feature extraction. No specific accuracy was stated. However, the research is problematic because it has not only a complex data processing requirement, but also, likely, sensitivity to signal variation.

[14] This research developed a real-time system for recognizing Arabic alphabetic characters. It only used pre-trained CNNs (e.g. VGGNet) and transfer learning for sign recognition, from still images captured via a webcam. Its implementation for isolated letter recognition is simple and performs quite well, but the system is limited in that it does not process words or sentences.

[15] This study addressed the toughest issue of all: understanding complete sentences in sign language instead of just individual signs. They utilized a Kinect sensor, which provided 3D skeletal tracking data of the joints in the body. Then, they implemented recurrent neural networks, specifically long short-term memory networks, which work with time series data and are able to learn and retain information over long periods, to translate gestures into sign language sentences in real time.

[16] This study conducted research in the area of computer vision and used an ordinary RGB camera to capture videos. They manually extracted features from the frames first and then used convolutional neural networks to learn features directly from images. The authors used a fuzzy logic-based decision-making system to increase system accuracy in ambiguous conditions and to increase system robustness in the presence of inter-user variability.

**Table 1: Comparison of related work**

Study	Techniques used	Accuracy	Limitations	Main Benefits and Contributions
[12]	Kinect V2, Leap Motion, SVM	95% - 98%	Expensive system (two sensors), non-portable, operates in a controlled environment.	Combining two sensors to compensate for each other's weaknesses and achieve very high accuracy for isolated signals.
[13]	Google Media Pipe with CNN and LSTM	--	requires complex data processing and is sensitive to signal variation	Combining CNN and LSTM to improve the accuracy of Arabic gesture recognition
[14]	Webcam, Transfer Learning (VGGNet), CNN	96% - 98%	For static images (single frames) only, does not handle motion or sentences.	A simple, low-cost, and effective system for real-time isolated alphabetic character recognition.
[15]	Kinect (3D Skeletal Data), RNN-LSTM	85% - 90%	Lower accuracy than isolated signal systems. Complexity in training sentences, depends on the accuracy of Kinect joint tracking.	Addressing the problem of recognizing continuous sign language (sentences) using advanced LSTM models, which is a greater challenge than isolated signals.
[16]	Computer Vision (RGB Camera), CNN, Fuzzy Logic	92% - 95%	Performance depends on good lighting and background, difficulty recognizing continuous signals.	Integrating artificial intelligence techniques (deep learning and fuzzy logic) to improve robustness and generalization.
Our	HSV With CNN	95.3	Performance depends on good lighting	Integrating artificial intelligence techniques (deep learning and HSV) to improve robustness detect hand and generalization classification with clutter backgrounds.

