AUT Journal of Electrical Engineering
10.22060/EEJ.2026.24533.5725

MPSL: A Balanced, Large-Scale Image Database for Persian Static Hand Gesture Recognition

Mahin Moghbeli Damaneh[1], Farahnaz Mohanna[2], Pouria Jafari[3*]
[1] Phd Student, Communication Engineering Department, Electrical and Computer Engineering Faculty, University of Sistan and Baluchestan, Zahedan Iran
[2] Associate Professor, Communications Engineering Department, Electrical and Computer Engineering Faculty, University of Sistan and Baluchestan, Zahedan Iran
[3] Assistant Professor, Electrical and Electronic Engineering Department, Electrical and Computer Engineering Faculty, University of Sistan and Baluchestan, Zahedan, Iran

Abstract: Hand gesture recognition is a pivotal area in artificial intelligence, enabling seamless human–machine interaction across diverse applications such as home appliances, electronic devices, and public transportation systems. This paper introduces the Moghbeli Persian Sign Language (MPSL) database, a comprehensive image-based resource designed to advance Persian sign language recognition research. The database comprises 107,630 images organized into 47 classes, including 37 Persian alphabet characters and 10 numerical digits (0–9). The images were extracted from 1–2-minute videos captured using a mobile phone camera and involve 18 participants (seven women and eleven men) fluent in Persian sign language. Data acquisition was performed under varying lighting conditions and imaging environments to enhance dataset diversity and robustness. Each class contains 2,290 images, resulting in a balanced large-scale database suitable for training and evaluation purposes. To validate the proposed database, four deep learning and hybrid recognition models were evaluated. The experimental results demonstrate that all evaluated methods achieve high recognition performance on the Moghbeli Persian Sign Language database. Among the evaluated methods, the hybrid model combining Oriented FAST and Rotated BRIEF, Gabor filters, and a Convolutional Neural Network achieved the highest average accuracy of 99.94% while requiring only 241,343 trainable parameters. The obtained results confirm that the proposed database provides sufficiently discriminative visual information for reliable

* pjafari@ece.usb.ac.ir

recognition of Persian sign language characters and numerical digits. This database therefore represents a valuable benchmark resource for future research in Persian sign language recognition.

Keywords: Hand Gesture Recognition, Moghbeli Persian Sign Language Database, Deep learning, Convolution Neural Network.

1. Introduction

Hand gestures constitute a prominent visual medium for transferring information between humans and machines [1]. The recognition of hand gestures has diverse applications, including in home appliances, electronic devices, automobiles, and public transportation systems [2]. Substantial research has been done in the domain of automatic hand gesture recognition, employing both non-intelligent and intelligent methodologies to extract features from hand gesture images. Non-intelligent methods exhibit certain limitations, such as a mismatch between the extracted features and the original image and the time inefficiency in identifying the optimal descriptive features [3].

Z`Conversely, intelligent methods, particularly those leveraging deep learning architectures, demonstrate a robust capability to extract features that effectively describe the shape and characteristics of the region of interest within an image, following their processing through various components of the intelligent algorithm [4]. Recently, the intelligent methods based on deep learning architectures have been widely used to recognize and classify hand gestures in sign language [5]. Among the most important methods based on the deep learning, we can mention auto-encoder networks, the Long Short-Term Memory (LSTM) neural networks [6] and convolutional neural networks (CNNs) [7-9]. The field of hand gesture recognition in sign language benefits from several image datasets. For instance, the French Belgian Sign Language (FBSL) database comprises videos of the Belgian sign language conversation [5]. The Massey database contains 2,524 hand gesture images for the English sign language, which are classified into 36 classes with alphabets and numerical digits [10]. The American Sign Language database, called ASL, has 87,000 hand gesture images for 29 American English alphabets [11]. The ASL Alphabet database has 60,000 hand gesture images showing 24 American English alphabets, a space character, a delete character, and a special default image for no symbols [12].

Each country has developed its own sign language database, tailored to its unique linguistic requirements. Examples include the British Sign Language (BSL) in the United Kingdom [13], Indian Sign Language (Indian SL) in India [14, 15], Arabic Alphabet Sign Language (ArASL) in Arab countries [16], Mexican Sign Language (Mexican SL) in Mexico [17] and Bangla Sign Language (Bangla SL) in Bangladesh [18].

A review of the available literature and public repositories indicates that existing Persian sign language resources are either limited in scale, focused on a subset of characters or words, or designed for hardware-based recognition systems. Therefore, a large-scale and balanced image-based database covering Persian alphabets and digits is still required for the development and evaluation of Persian sign language recognition systems.

The main contributions of this study include the development of the Moghbeli Persian Sign Language (MPSL) database, which contains 107,630 images of Persian sign language gestures collected under various acquisition conditions, along with the construction of a balanced database covering all 37 Persian alphabet gestures and 10 numerical gestures. Additionally, a statistical analysis of the database was conducted using Principal Component Analysis (PCA) to investigate the distribution characteristics of the collected samples. The proposed database was also compared with existing Persian sign language resources and publicly available repositories. Finally, a benchmark evaluation of several recognition methods was performed on the database to demonstrate its suitability for Persian sign language recognition research.

The structure of this paper is as follows: Chapter 2 discusses related work in the field. The proposed Moghbeli Persian Sign Language database for the hand gesture recognition is presented in Chapter 3. Sign language recognition methods are reviewed in Chapter 4 and the simulation results and their evaluation are given in Chapter 5. The paper is concluded in Chapter 6.

2. Previous Work

This chapter presents recent research related to hand gesture recognition for alphanumeric classes in sign language is presented. The methods of recognition, feature extraction, and hand gesture classification in sign language are divided into three categories: traditional, deep learning-based, and hybrid. A combination of convolutional neural networks (CNNs) with traditional method of feature extraction was employed for real-time hand gesture recognition. The highest accuracy was at 99.59% on the LaRED database, comprising 81,000 images of 27 hand gestures [6]. The deep-learning-based BSL recognition system was introduced using pre-trained CNNs such as GoogleNet, SqueezeNet, and VGGNet for classification. The highest accuracy was 90.07% using the VGGNet on 1,950 images of 15 hand gestures from the BSL database [13]. The Speeded Up Robust Features (SURF) descriptor and the directional histogram were used to extract features from the segmented hand gestures. Subsequently, the features were classified by using the CNN and Support Vector Machine (SVM) to recognize hand gestures in the Indian Sign Language (ISL) database, comprising 36,000 images of 36 hand gestures. The system achieved the highest accuracy of 99.17% for the ISL database [14].

A method for the recognition of Arabic sign language was presented, comprising pre-processing, feature extraction, and classification. Following the segmentation of the hand from the background of the hand gesture image, features of the segmented hand were extracted. Thereafter, the supervised machine learning techniques were employed to classify 28 Arabic alphabets in 2800 hand gesture images from the Arabic Alphabet Sign Language Recognition System (ArASLRS) database. The classification accuracy of the method was 99.02% with the KNN classifier [16]. Furthermore, the method achieved a classification accuracy of 99.50% for the Quranic Sign Language (QSL) database, which comprises 840 images of 14 hand gestures. The hand gesture images were prepared using the Microsoft Kinect sensor. Following pre-processing, the input image was entered into Hidden Markov Models (HMMs) for feature extraction and classification. The recognition accuracy was obtained at 99% for the 82 hand gestures in the Mexican SL database. The number of used images was not mentioned [17].

10.22060/EEJ.2026.24533.5725

In a method, the input images were initially converted into the HSV color spaces and binary masks were applied to identify the area of hand skin. The skin area was enhanced through the application of morphological operations, after which it was separated from the image background using thresholding. Subsequently, the boundaries of the hand gestures were identified with the watershed algorithm and extracted by applying a bitwise mask. Subsequently, the training data was augmented using seven methods. Finally, these images were entered into the CNN-based network, called BenSignNet, for feature extraction and classification. The performance accuracy of the method was reported as 94.00%, 99.60%, and 99.60% respectively for the BSL, Khulna University Bengali Sign Language (KU-Bd SL) and Ishara-Lipi databases [18].

A network employed consisted of six Restricted Boltzmann Machines (RBMs) with 500 hidden neurons each. Initially, six input images were fed into the six RBMs, and the outputs from these RBMs were subsequently integrated into another RBM for the final hand gesture classification. Various hand gesture datasets were utilized in the experiments, and the highest recognition accuracy achieved was 99.31% for the Massey database [19].

A method was employed to recognize sign language using depth images, with the hand area separated by thresholding in the depth images. The hand features were then extracted using PCANet, with these extracted features entered into SVM for classification. The average recognition accuracy of the method was 88.70% for 60,000 images of 24 hand gestures from the ASL Alphabet database [20]. Another method was presented based on a CNN for the recognition of hand gestures in the ASL database. The recognition accuracy of the method was reported as 99.99% [21]. Various machine learning algorithms were presented for recognizing hand gestures on the MNIST sign language database. These algorithms include the Naive Bayes, KNN, random forest, XGBoost, support vector classifier, logistic regression, stochastic gradient descent classifier, and CNN. The CNN achieved the highest accuracy of 91.41% on the MINST database. The MINST database comprises 24,000 images of 20 hand gestures. Both KNN and random forest demonstrated satisfactory performance, with accuracies of 80.46% and 84.43%, respectively. Conversely, the Naive Bayes exhibited the lowest accuracy, at 46.85% [22].

10.22060/EEJ.2026.24533.5725

A real-time sign recognition network was designed by applying a lip movement controller which used the LSTM recurrent neural network in conjunction with KNN to classify signs. The experimental results demonstrated that the network achieved an average accuracy of 99.44% for 26 alphabets of the ASL-26 database, with 100 hand gesture images for each alphabet [23].

A CNN was introduced with fewer parameters and higher recognition accuracy compared to other CNNs, such as VGG-11 and VGG-16. The recognition accuracy of this CNN was 96.99% on 2,150 hand gesture images from the Indian SL-43 database [24]. An Arabic sign language recognition system was presented using hand and body key points extracted from video frames using OpenPose. A 3D CNN was employed to classify these key points with an accuracy of 88.89% for 1,120 images of 28 hand gestures from the Arabic SL database [25]. A hand gesture recognition system was introduced based on high-resolution thermal imaging, which is independent of light conditions [26]. A CNN was used to classify hand gestures and achieved a recognition accuracy of 98.81% for 14,400 images of 10 hand gestures from the ASL-A database. A method was presented that included five steps: pre-processing, hand segmentation, extraction of hand key points, and classification to recognize hand gestures. The method performance evaluation on the MNIST, and ASL databases showed the accuracy of 93.20%, and 91.70% respectively [27]. A real time system was implemented for the Bangla SL-A recognition using deep learning models such as Detectron2, EfficentDet, and YOLOv7, and a Jetson Nano edge device. The deep learning models were trained using the Okkhornama database. The mean average accuracy of the recognition system on 3920 images of 49 hand gestures from the Bangla SL database was reported at 94.91% by using the YOLOv7 [28]. A three-stage neural network was developed based on the spatial attention method. This involved utilizing the self-attention from the database image and then multiplying the feature vector with the attention map in order to highlight the effective features of the database image for the hand gestures recognition in the sign language. The spatial attention method assisted the neural network in recognizing the effective parts of the hand gesture images and using them for more accurate recognition of the hand gestures. The accuracy of the network for the Senz3D, Kinematic, and NTU databases was reported as 99.67%, 99.75%, and 99.46%,

10.22060/EEJ.2026.24533.5725

respectively [29]. The YoloV4-Tiny network was trained on the Bangla SL database, which included 49 classes of the standard Bengali sign language symbols. This resulted in an average accuracy of 99.70% in the hand gesture recognition task [30]. Three enhanced deep learning models of the network were presented by combining the YOLOv5x architecture with the attention algorithm for the sign language recognition task. The models included the Squeeze-and-Excitation (SE) module, the Convolutional Block Attention Module (CBAM), and their combination. The SE module assisted in the selection of important features, while the CBAM module facilitated the analysis of these features in two dimensions, thereby enhancing the accuracy of the deep learning architectures. Two sign language databases, namely Massey and Bangla SL-92, were employed to assess the efficacy of these models. The highest accuracy of 98.90% was observed on the Massey database with the SE-YOLOv5x model [31].

A recent study in [32] introduced a hybrid three-stream recognition framework combining ORB descriptors, Gabor filters, and a parallel CNN architecture for static sign language recognition. Evaluated on benchmark datasets including Massey, ASL, and ASL Alphabet, the method achieved over 99% accuracy by integrating complementary geometric, texture, and deep visual features. Due to its proven robustness and generalizability, this framework is adopted in the present study as a baseline recognition model for evaluating the proposed MPSL database. An overview of the recent research pertaining to hand gesture recognition in sign language is presented in Table 1.

Table 1 An overview of the recent research related to the hand gesture recognition in the sign language.

Number	Method	Database	Images No.	Number of classes	Accuracy %
1.	FFCNN (2018) [6]	Massey	2524	36	98.05
		LaRED	81000	27	99.59
		OUHANDS	3000	10	98.75
2.	VGGNet (2022) [13]	British SL	1950	15	90.07
3.	SVM (2022) [14]	Indian SL	36000	36	99.17
4.	Supervised ML & KNN (2021) [16]	QSL	840	14	99.50
		ArASL	1400	28	99.02
5.	Hidden Markov (2022) [16]	Mexican SL	-	82	99.00
6.	BenSignNet (2022) [18]	BSL	12160	38	94.00
		KU-Bd SL	1500	30	99.60
		Ishara-Lipi	1800	36	99.60

10.22060/EEJ.2026.24533.5725

7.	RBM (2018) [19]	Massey	2524	36	99.31
		ASL Fingerspelling A	131000	24	98.13
		NYU	81000	36	90.01
		ASL Fingerspelling	130000	24	97.56
8.	PCANet, & SVM (2019) [20]	ASL	60000	24	88.70
9.	CNN (2020) [21]	ASL Alphabet	60000	24	99.90
		Massey	2515	36	99.92
		Sign Language Digits	2180	10	99.90
		ASL	87000	29	99.99
10.	CNN (2022) [22]	MNIST	24000	24	91.44
11.	LSTM, and KNN (2022) [23]	ASL-26	2600	26	99.44
12.	CNN (2021) [24]	Indian SL-43	2150	43	99.66
		ASL	30000	10	100
13.	3D CNN (2021) [25]	Arabic SL	1520	28	98.39
14.	CNN (2021) [26]	ASL-A	14400	10	98.81
15.	landmark detection, & CNN (2022) [27]	ASL	87000	29	91.70
		MNIST	24000	24	93.20
16.	YOLOv7 Tiny (2023) [28]	Bangla SL-A	3920	49	94.91 (mAP)
17.	Self-attention CNN (2023) [29]	senz3D	3200	11	99.67
		Kinetic	1400	10	99.75
		NTU	1000	10	99.64
18.	YOLOv4 (2023) [30]	Bangla SL	14745	49	99.70
19.	YOLOv5x, SE, &CBAM attention modules (2023) [31]	ASL	9072	36	98.90
		Bangla SL-46	9821	46	97.60
20.	Hybrid three-stream framework (ORB + Gabor + parallel CNN) (2023) [32]	Massey	25200	36	99.92
		ASL	87000	26	99.81
		ASL Alphabet	23400	26	99.80

3. Moghbeli Persian Sign Language (MPSL) Database

To date, there has not been a database available containing Persian sign language letters that can be utilized for the classification and recognition of Persian sign language. In this context, we introduce a valid, accurate, standard, and comprehensive database showcasing static hand movements for 37 letters of the Persian alphabet and 10 numbers from 0 to 9.

Therefore, a series of 1-2 minute videos for each of the 47 hand movements, recorded against a simple background, were captured from seven women and eleven men fluent in Persian sign language. These videos were made using a mobile phone camera equipped with a 108-megapixel main lens, an 8-megapixel ultra-wide lens, a 2-megapixel macro lens, and a 2-megapixel depth sensor. To enhance diversity in the dataset, the images were taken under varying lighting conditions and at a distance of 1.5

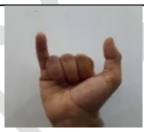
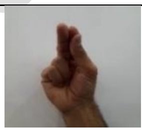
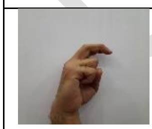
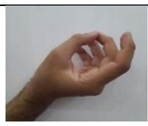
10.22060/EEJ.2026.24533.5725

to 2 meters from the camera, with participants not wearing any coverings such as gloves. From these videos, 2,290 images were extracted for each character, resulting in a comprehensive database of Persian sign language, named Moghbeli Persian Sign Language (MPSL). This database comprises a total of 107,630 images representing 47 Persian alphabet characters and numerical digits. Table 2 shows a selection of hand gesture images from MPSL database. In this table, the meaning of each hand movement in Persian and its equivalent in English is provided in parentheses.

It is important to note that the simplicity of the background in the MPSL database does not detract from its comprehensiveness. This is because the hand regions can be segmented offline from other videos of sign language before being inputted into the proposed recognition network. Additionally, the hand gestures used to represent numbers from 0 to 9 in Persian sign language are identical to those used in English and Arabic sign languages.

A comparison of the MPSL database with 30 well-known sign language databases is presented in Table 3. Among these databases, MPSL, ASL, ASL Fingerspelling A, and ASL Fingerspelling databases have the highest numbers of hand gesture images. Notably, MPSL surpasses all others in terms of the total number of hand gesture images. It is important to highlight that the ASL-10 database comprises only 10 characters with 30,000 images sourced from the ASL database. Similarly, the ASL Alphabet database includes 60,000 images derived from the ASL Fingerspelling A database.

Table 2: Selected images of 47 Persian sign language characters in the MPSL database along with their equivalent phonetics in English (in parentheses).

 ب (b)	 ای (i)	 او (u)	 اُ (o)	 اِ (e)	 اَ (a)	 آ (ā)
 خ (kh)	 ح (h)	 چ (ch)	 ج (j)	 ت (t)	 ث (s or th)	 پ (p)
 ش (sh)	 س (s)	 ژ (zh)	 ز (z)	 ر (r)	 ذ (z or dh)	 د (d)

AUT Journal of Electrical Engineering
10.22060/EEJ.2026.24533.5725

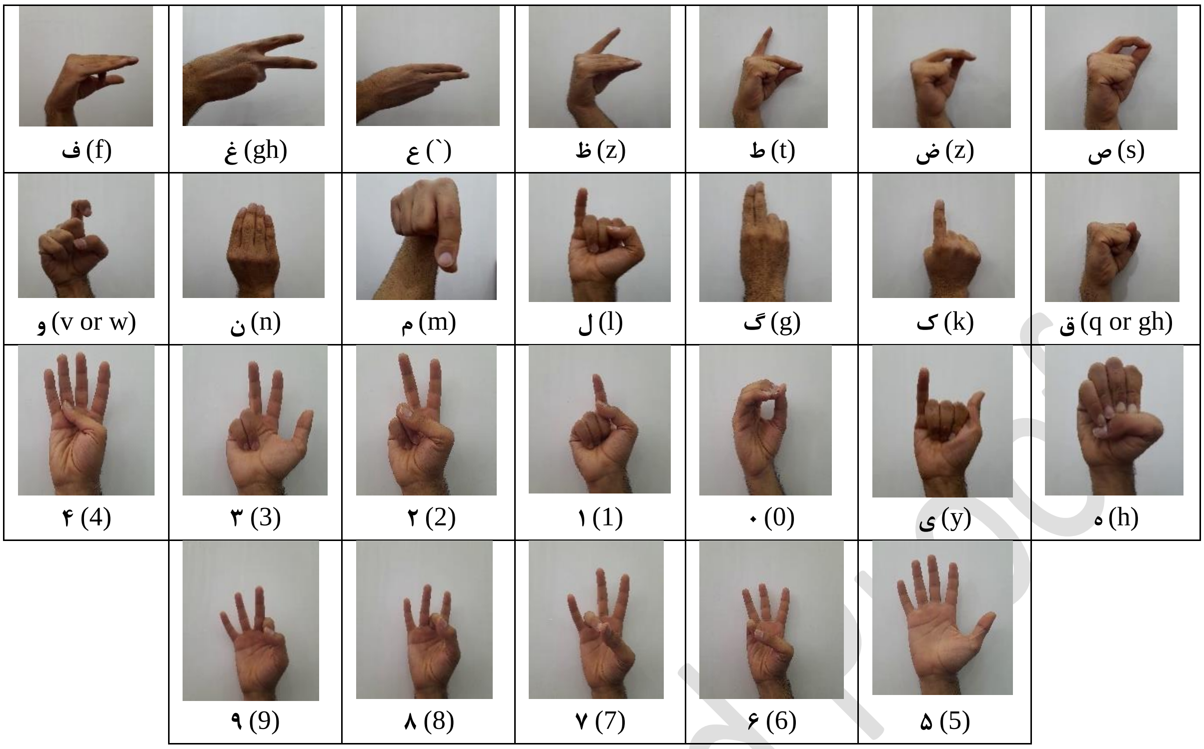
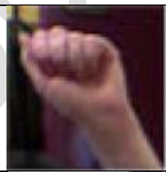
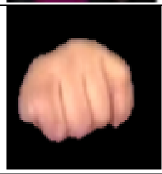


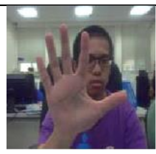
Table 3: The MPSL comparison with the well-known sign language databases

Number	Database	Images Color Space	Images No.	Class No.	Description	Sample Image
1	MPSL	RGB	107630	47	-18 different subjects. -Different lightening, rotation, and scaling. -Simple background.	

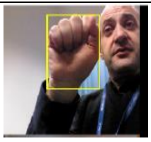
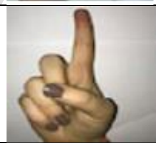
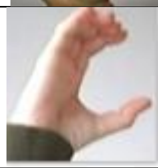
AUT Journal of Electrical Engineering
10.22060/EEJ.2026.24533.5725

2	Massey	RGB, Gray, and Binary	2524	36	-5 different subjects. -Different illumination, rotation, and scaling. -Simple background.	
3	ASL	RGB	87000	29	-5 different subjects. -Different illumination, rotation, and scaling. -Simple background.	
4	ASL Alphabet	RGB, and Depth	60000	24	-5 different subjects. -Slight rotation. -Non-simple background.	
5	ASL Fingerspelling	RGB, and Depth	130000	24	-Non-simple background.	
6	ASL Fingerspelling A	RGB, and Depth	131000	24	-Non-simple background.	
7	ASL-26	RGB	2600	26	-100 different subjects. -Simple background.	
8	ASL-10	RGB	30000	10	-Simple background.	
9	ASL-A	RGB, and Gray	14400	10	-30 different subjects. -Simple background.	 (b)
10	Mexican SL	RGB, and Binary	-	82	-20 different subjects. Non-simple background.	

10.22060/EEJ.2026.24533.5725

11	British SL	Gray	1950	15	-4 different subjects. -Simple background.	
12	Indian SL	RGB, Gray, and Binary	36000	36	-3 different subjects. -Simple background.	
13	Indian SL-43	RGB	2150	43	-50 different subjects. -Different illumination, and scaling. -Non-simple background.	
14	ArASL	RGB	2800	28	-Different illumination. -Simple background.	
15	Arabic ASL	RGB, and Depth	1120	28	-40 different subjects. -Different illumination. -Simple background.	
16	QSL	RGB	840	14	-Different illumination, and rotation. -Simple background.	
17	BSL	RGB	12160	38	-320 different subjects. -Non-simple background.	
18	Bangla SL	RGB	14745	49	-Non-simple background.	
19	Bangla SL-46	RGB	9821	46	-Non-simple background.	
20	Bangla SL-A	RGB	3920	49	-11 different subjects. -Different illumination. -Non-simple background.	
21	LaRED	RGB, Depth, and Binary	81000	27	-10 different subjects. -Different rotation. -Non-simple background.	

10.22060/EEJ.2026.24533.5725

22	OUHANDS	RGB, Depth, Gray, and Binary	3150	10	-20 different subjects. -Non-simple background.	
23	KU-Bd SL	RGB	1500	30	-33 different subjects. -Non-simple background.	
24	NYU	RGB	81000	36	-50 different subjects. -Simple background.	
25	Ishara-Lipi	RGB	1800	36		
26	Sign Language Digits	RGB	2180	10	-218 different subjects. -Simple background.	
27	MNIST	RGB	24000	20	-Different rotation. -Simple background.	
28	Senz3D	RGB, and Depth	3200	11	-4 different subjects. -Different rotation. -Non-simple background.	
29	Kinetic	RGB, and Depth	1400	10	-14 different subjects. -Non-simple background.	
30	NTU	RGB, and Depth	1000	10	-10 different subjects. -Non-simple background.	

3.1. Critical Gaps in Existing Persian Sign Language (PSL) Resources

Although several studies and open-source projects related to Persian Sign Language (PSL) have recently been reported, most of them are designed for specific applications and do not provide a comprehensive benchmark database for character-level recognition. Some available repositories focus on a limited number of Persian words and contain only a small number of images. For example, the

Persian-Sign-Language-Detection repository [33] is designed to recognize a few isolated Persian words and does not include the complete Persian alphabet or digits. Similarly, the Iranian-Sign-Language-ISL repository [34] contains images of Persian alphabet gestures but does not include numerical gestures and covers fewer classes.

Several other studies, such as PenSLR [35] and Sign-Language-Glove [36], rely on wearable sensors and electronic gloves for sign language recognition. Although such systems can provide useful motion information, they require additional hardware and are fundamentally different from conventional vision-based recognition systems.

Furthermore, video-based corpora such as the ZEI Corpus [37, 38] and ISLR101 [39, 40] are mainly intended for continuous sign language analysis and isolated word recognition. These databases are not specifically designed as balanced image datasets for recognizing Persian alphabets and digits. To provide a clearer comparison, Table 4 summarizes the characteristics of the currently available Persian sign language resources and compares them with the proposed MPSL database.

As shown in Table 4, the MPSL database contains 107,630 images representing all 37 Persian alphabet characters and numerical digits, and 10 digits. Each class contains 2,290 images collected from multiple participants under different illumination conditions. To the best of our knowledge, MPSL is currently the largest publicly documented image-based Persian sign language database that simultaneously provides complete coverage of Persian alphabets and digits in a balanced structure suitable for deep learning applications. Representative samples of the database are publicly available through the following repository: (https://github.com/nnislab-cmd/MPSL-Dataset/). The repository provides sample images, class labels, and the dataset structure to facilitate future research on Persian sign language recognition.

10.22060/EEJ.2026.24533.5725

Table 4: Comparative analysis of the proposed MPSL database against existing Persian Sign Language resources.

Number	Dataset Name	Total Images / Videos	No. of Classes	Full Alphabet Coverage?	Core Modality
1.	MPSL (Proposed)	107,630 Images	47 (37 Letters and symbols + 10 Digits)	Yes	Vision-Based (Static Gestures)
2.	Persian-Sign-Language-Detection [33]	~Dozens of Images	4 Words	No	Vision-Based (Real-time Demo)
3.	PenSLR [35]	0 (Sensor Signals Only)	<15 Isolated Words	No	Hardware-Based (Digital Gloves)
4.	ZEI Corpus [37, 38]	25.7 Hours of Video	Continuous	No	Vision-Based (Continuous Video)
5.	ISLR101 [39, 40]	4,614 Videos	101 Signs	No	Vision-Based (Isolated Words)
6.	Iranian-Sign-Language-ISL [34]	~1,600 Images	32 (Letters Only)	No (Missing Digits)	Vision-Based (Static Gestures)
7.	Sign-Language-Glove [36]	Time-Series Data	Unspecified	No	Hardware-Based (Glove Sensors)

The reviewed literature demonstrates that considerable progress has been achieved in sign language recognition using both conventional machine learning and deep learning techniques. However, existing Persian Sign Language resources remain limited in terms of dataset scale, class coverage, modality, and public availability. Furthermore, most available studies focus either on a restricted vocabulary of isolated words or on sensor-based acquisition systems. Consequently, a large-scale, balanced, image-based benchmark covering both Persian alphabet gestures and numerical gestures is still lacking. The proposed MPSL database has been developed to address this gap and provide a comprehensive benchmark for future Persian Sign Language recognition research.

3.2. Statistical Analysis of the MPSL Database

To provide a better understanding of the structure and complexity of the proposed database, Principal Component Analysis (PCA) was applied to a randomly selected subset containing 15% of the MPSL image samples. Since the primary objective of this analysis was to visualize the overall data distribution and variability, a representative random subset was considered sufficient for generating the two-

10.22060/EEJ.2026.24533.5725

dimensional projection. PCA projects the high-dimensional image space into a lower-dimensional representation while preserving the maximum possible variance of the original data.

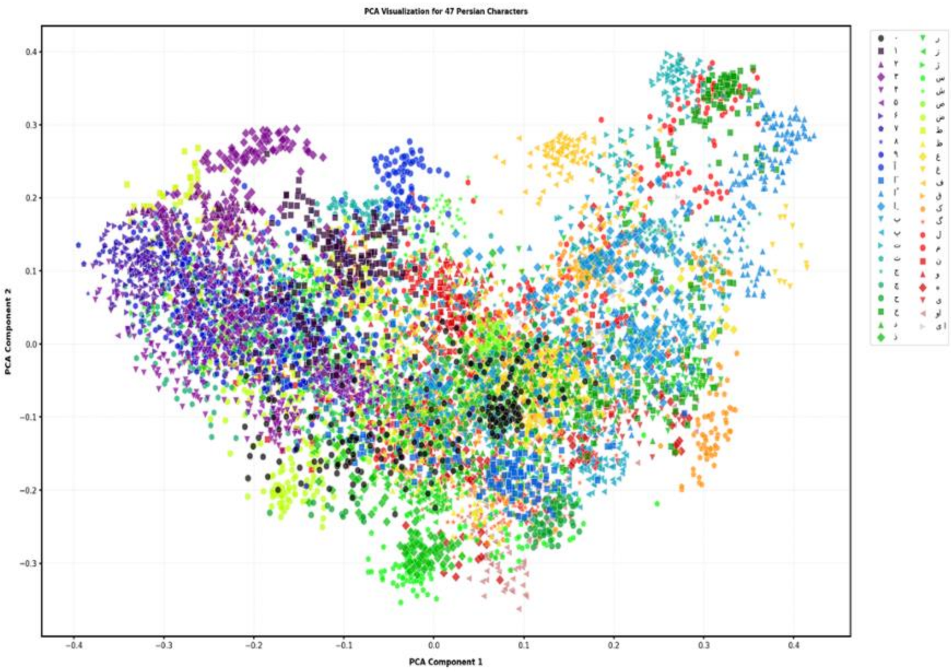


Figure 1. Two-dimensional visualization of the MPSL database using Principal Component Analysis (PCA). Each point represents an image sample projected onto the first two principal components.

Fig. 1 illustrates the distribution of the database samples in the two-dimensional space formed by the first two principal components (PC1 and PC2). It can be observed that the projected samples exhibit a considerable degree of overlap among different classes. This indicates that the Persian sign language characters are not linearly separable when represented only by the first two principal components.

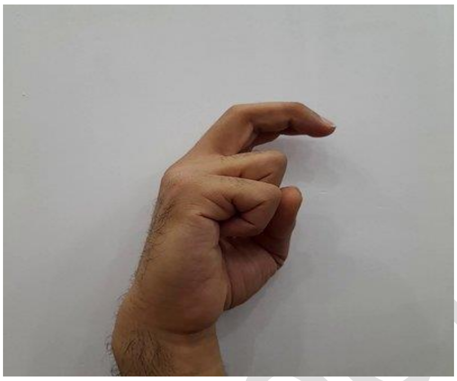
This behavior is expected because many Persian sign language gestures differ only in subtle variations of hand shape, finger position, or gesture configuration. Consequently, a substantial amount of discriminative information remains embedded in the original high-dimensional image space and cannot be fully represented through a two-dimensional linear projection.

The PCA visualization also reveals that several visually similar sign characters tend to occupy neighboring regions in the projected space. Fig. 2 presents representative examples of two Persian sign language letters with similar visual structures. Such observations demonstrate the inherent complexity of the dataset and the challenges associated with accurate Persian sign language recognition. Overall,

10.22060/EEJ.2026.24533.5725

the PCA analysis provides additional statistical insight into the distribution and structural characteristics of the MPSL database.

a- The letter ح (h)



b- The letter خ (kh)

Figure 2. Example of two Persian sign language characters with similar visual structures that contribute to overlapping regions in the PCA projection.

4. Sign language recognition methods

The recognition framework employed in this study is based on our previously published architecture presented in [32]. Since the primary objective of the present work is the introduction and validation of the MPSL database rather than the development of a new classification architecture, the baseline framework of [32] was retrained from scratch using the MPSL dataset and utilized as a baseline recognition model for benchmarking purposes. The training process was performed using the Adam optimizer with a learning rate of 0.001, a batch size of 64, and 50 training epochs. Interested readers are referred to [32] for a detailed description of the network architecture, feature extraction process, and design rationale.

In recent years, numerous methods for detecting hand movements have been introduced, leveraging artificial intelligence and deep neural networks. In this article, we evaluate the presented database using four methods renowned for their optimal performance and high accuracy. Each of these methods employs convolutional neural networks (CNNs) [21, 22, 24, 32].

Among these, the hybrid ORB–Gabor–CNN framework [32] is employed in this study as a baseline model for evaluating the proposed MPSL database. This model demonstrates the highest accuracy and

10.22060/EEJ.2026.24533.5725

efficiency on the proposed database. The general structure of this method is depicted in Fig. 3. This hybrid CNN method comprises three independent feature extraction streams, each dedicated to processing a hand gesture image.

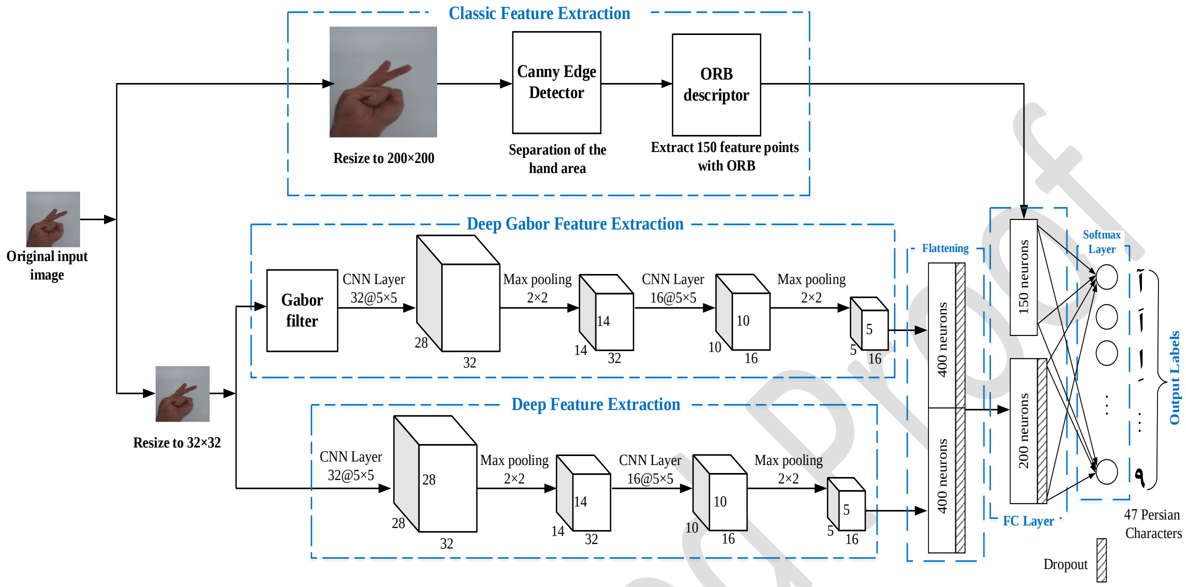


Figure 3: The hybrid CNN method to recognize the hand gesture in the Persian sign language [32].

This method employs a three-stream network for feature extraction from hand gesture images. The first stream uses classic feature extraction, involving resizing the image, applying Canny edge detection, and extracting key points with the ORB descriptor, followed by K-means clustering to produce a feature vector. The second stream uses Gabor filter for texture feature extraction, which is then processed by a CNN. The third stream applies a similar CNN directly to the resized image. The outputs from the CNNs are flattened, subjected to dropout techniques to prevent overfitting, and combined with ORB descriptor features. These combined features are fed into a softmax output layer for hand gesture classification. Table 5 presents the configuration of this hybrid ORB–Gabor–CNN framework with dropout rate denoted by P.

Table 5: Configuration of the hybrid ORB–Gabor–CNN baseline model used in this study.

CNN Model Contents	Details
First convolution layer	32@5×5×3,Relu function, input size 32×32
First max pooling layer	32@2×2

10.22060/EEJ.2026.24533.5725

Second convolution layer	16@5×5×32, Relu function
Second max pooling	Kernel with size of 2×2
Dropout after flattening layer	P=63%
Fully connected layer + Dropout	200 neurons, Relu function, P=23%
Output layer	The neurons number depends on the classes number in the database, Softmax

5. Simulation Results

To demonstrate the performance of the desired classification methods in the proposed Persian sign language recognition, four methods were applied. Table 6 demonstrates that the proposed MPSL database can be effectively utilized by a wide range of recognition architectures with substantially different model capacities. Although all evaluated networks achieve high recognition performance, noticeable differences can be observed in their computational complexity. In particular, the hybrid ORB–Gabor–CNN framework [32] achieves the highest average accuracy while requiring significantly fewer trainable parameters than the high-capacity deep model. These results indicate that the proposed database provides sufficient discriminative information for both conventional and deep learning-based recognition approaches.

Table 6: Comparison of Recognition Results for Different Methods Applied to the MPSL Database

Network	Average accuracy%	Average sensitivity%	Average specificity%	Average F-score%	Parameters No.
Medium-Scale Standard Architecture [21]	99.82±0.13	99.78±0.10	99.72±0.19	99.70±0.10	3,128,815
VGG-Based [22]	98.90±0.10	98.40±0.40	98.90±0.10	98.20±0.35	2,182,385
High-Capacity Deep Model [24]	99.85±0.17	99.79±0.25	99.75±0.13	99.70±0.22	67,250,910
hybrid CNN method [32]	99.94±0.06	99.88±0.12	99.94±0.06	99.88±0.12	241,343

The results are shown based on the accuracy, sensitivity, specificity, and F_Score criteria, which are respectively defined by Eqs. (1) to (4).

$$\text{Accuracy} = 100 \times (TN + TP)/(TP + TN + FP + FN) \quad (1)$$

$$\text{Sensitivity} = 100 \times TP/(TP + FN) \quad (2)$$

10.22060/EEJ.2026.24533.5725

$$\text{Specificity} = 100 \times TN / (TN + FP) \quad (3)$$

$$\text{F\_score} = 100 \times 2TP / (2TP + FP + FN) \quad (4)$$

The accuracy of the classification structure was evaluated using the numbers of true positives (*TP*), false positives (*FP*), true negatives (*TN*), and false negatives (*FN*) in the hand gesture recognition for the sign language database images. The evaluation metrics were selected to provide a comprehensive assessment of classification performance. Accuracy reflects the overall recognition capability of the model, while sensitivity and specificity quantify the ability to correctly identify positive and negative samples, respectively. In addition, the F-score combines precision and recall into a single indicator, providing a balanced measure of classification quality. The simultaneous use of these metrics reduces the risk of drawing conclusions based solely on overall accuracy and enables a more reliable comparison between different recognition models.

Furthermore, 60% of each sign language database image was randomly selected for training, 30% for testing, and 10% for validation. The experiments were conducted on a system with an Intel(R) Core(TM) i7-10870H CPU @ 2.20 GHz, 500 GB storage space, 16 GB installed RAM, and a GeForce RTX 3070 GPU. The software utilized includes the Windows 10 Pro 2021 operating system, Python version 3.10, TensorFlow Library version 2.0.1, Keras Library version 2.0.3, and OpenCV library version 3.2. This setup ensured a robust environment for executing and validating the proposed method, confirming its efficacy in recognizing static hand gestures across different sign language databases.

Due to the high accuracy and fewer parameters of the hybrid CNN method [32], this structure is used to compare the proposed database with other widely used databases such as ASL, ASL Alphabet, and Massey. The characteristics of these databases are presented in Table 7. The test results of the hybrid CNN method on the selected databases were evaluated over 10 executions in Table 8.

Table 7: Characteristics of the sign language databases.

Image Database	Image Size	Images No.	Classes No.
MPSL	200×200	107630	47
Massey	400×400	2524	36
ASL	50×50	23400	26
ASL Alphabet	200×200	87000	29

10.22060/EEJ.2026.24533.5725

Table 8: Sign language recognition rate on the MPSL, Massey, ASL, and ASL Alphabet databases by the hybrid CNN method.

Evaluation Criterion \ Image Dataset	MPSL	Massey	ASL	ASL Alphabet
Average of Accuracy	99.94%±0.06	99.90%±0.1	99.85%±0.15	99.91%±0.09
Average of Sensitivity	99.88%±0.12	99.85%±0.15	99.76%±0.24	99.82%±0.18
Average of Specificity	99.94%±0.06	99.92%±0.08	99.81%±0.19	99.80%±0.2
Average of F-Score	99.88%±0.12	99.79%±0.21	99.78%±0.22	99.88%±0.12
Test time per test image (sec.)	1.85×10^{-3}	2.1×10^{-3}	2.5×10^{-3}	1.9×10^{-3}

The sign language recognition accuracy of the proposed network on the MPSL database is also reported in Table 9 according to the various number of the ORB key points, which shows as the number of the ORB key points goes higher, the sign language recognition accuracy also goes higher.

Table 9: Sign language recognition accuracy by the hybrid CNN method for the MPSL database with various number of the key points.

Best accuracy	Average accuracy	Standard Deviation	ORB key points No.
99.70%	99.55%	0.45	50
99.80%	99.66%	0.30	100
100%	99.94%	0.06	200
100%	99.96%	0.04	300
100%	99.98%	0.02	400

The ORB key point’s matrix, irrespective of its size, undergoes K-means clustering with 150 clusters. Consequently, the K-means algorithm consistently produces a 1×150 feature vector from the optimal ORB key points selected from each hand gesture image.

6. Conclusion

Every automatic sign language recognition system typically involves three key stages: hand gesture identification, hand gesture feature extraction, and classification. In this study, we introduced a comprehensive and systematically collected database containing static hand gestures corresponding to 37 Persian alphabet characters and 10 numerical digits. Some series of 1-2 minute videos captured using a mobile phone camera were recorded for each of the 47 hand movements, using a simple background

10.22060/EEJ.2026.24533.5725

and involving seven women and eleven men fluent in Persian sign language. To ensure diversity in the dataset, the images were taken under various lighting conditions and at a distance of 1.5 to 2 meters from the camera, with participants not wearing any coverings such as gloves. From these videos, 2,290 images were extracted for each character, culminating in the creation of the Moghbeli Persian Sign Language (MPSL) database. This database encompasses a total of 107,630 images representing 47 alphanumeric classes, including 37 Persian alphabet characters and 10 numerical digits.

To validate the proposed MPSL database, four different deep learning architectures were evaluated. The experimental results demonstrate that the proposed database can effectively support Persian sign language character and digit recognition across different recognition models while maintaining consistently high performance.

The obtained results further indicate that the high recognition rates are not a consequence of trivial class separability. The PCA analysis performed on a randomly selected subset of the MPSL database revealed considerable overlap among classes in a reduced two-dimensional representation, suggesting that the underlying classification problem is inherently nonlinear. Nevertheless, the consistently strong performance achieved by multiple deep learning architectures demonstrates that the proposed database contains sufficiently discriminative visual information for reliable Persian sign language recognition. Moreover, the comparable results obtained across different network architectures confirm the quality, balance, and suitability of the MPSL database as a benchmark resource for future research in Persian sign language recognition.

Disclosures

The authors have no conflicts of interest.

Data Availability Statement

A representative subset is provided for demonstration and verification purposes. The complete dataset will be publicly released following the completion of the ongoing research studies associated with the database. Representative samples of the MPSL database, including class labels and dataset structure, are publicly available at: https://github.com/nnislab-cmd/MPSL-Dataset/

References

[1] G. Benitez-Garcia, J. Olivares-Mercado, G. Sanchez-Perez, K. Yanai, Ipn hand: A video dataset and benchmark for real-time continuous hand gesture recognition, in: 2020 25th international conference on pattern recognition (ICPR), IEEE, 2021, pp. 4340–4347.
[2] L.P. Berg, J.M. Vance, Industry use of virtual reality in product design and manufacturing: a survey, Virtual reality, 21 (2017) 1–17.
[3] S.S. Rautaray, A. Agrawal, Vision based hand gesture recognition for human computer interaction: a survey, Artificial intelligence review, 43 (2015) 1–54.
[4] L. Nanni, S. Ghidoni, S. Brahnam, Handcrafted vs. non-handcrafted features for computer vision classification, Pattern recognition, 71 (2017) 158–172.
[5] J. Fink, B. Frénay, L. Meurant, A. Cleve, Lsfb-cont and lsfb-isol: Two new datasets for vision-based sign language recognition, in: 2021 International Joint Conference on Neural Networks (IJCNN), IEEE, 2021, pp. 1–8.
[6] S.F. Chevtchenko, R.F. Vale, V. Macario, F.R. Cordeiro, A convolutional neural network with feature fusion for real-time hand posture recognition, Applied Soft Computing, 73 (2018) 748–766.
[7] L. Zhang, Q. Tian, Q. Ruan, Z. Shi, A simple and effective static gesture recognition method based on attention mechanism, Journal of Visual Communication and Image Representation, 92 (2023) 103783.
[8] H. Noori, Persian Car License Plate Recognition Using Deep Convolutional Neural Networks, AUT Journal of Electrical Engineering, 57(2 (Special Issue)) (2025) 295–316.
[9] M. Jadidi, A. Mohammadi, Deep Learning for Recognition of Digital Modulations: A Detailed Study, AUT Journal of Electrical Engineering, 54(1) (2022) 67–78.
[10] A.L. Barczak, N.H. Reyes, M. Abastillas, A. Piccio, T. Susnjak, A new 2D static hand gesture colour image dataset for ASL gestures, (2011).
[11] M.A. Rahim, J. Shin, K.S. Yun, Hand gesture-based sign alphabet recognition and sentence interpretation using a convolutional neural network, Annals of Emerging Technologies in Computing (AETiC), Print ISSN, (2020) 2516–0281.
[12] A. Sharma, A. Mittal, S. Singh, V. Awatramani, Hand gesture recognition using image processing and feature extraction techniques, Procedia Computer Science, 173 (2020) 181–190.
[13] H. Hameed, M. Usman, A. Tahir, K. Ahmad, A. Hussain, M.A. Imran, Q.H. Abbasi, Recognizing british sign language using deep learning: A contactless and privacy-preserving approach, IEEE Transactions on Computational Social Systems, 10(4) (2022) 2090–2098.
[14] S. Katoch, V. Singh, U.S. Tiwary, Indian Sign Language recognition system using SURF with SVM and CNN, Array, 14 (2022) 100141.
[15] G. Joshi, S. Singh, R. Vig, Taguchi-TOPSIS based HOG parameter selection for complex background sign language recognition, Journal of Visual Communication and Image Representation, 71 (2020) 102834.
[16] G. Tharwat, A.M. Ahmed, B. Bouallegue, Arabic sign language recognition system for alphabets using machine learning techniques, Journal of Electrical and Computer Engineering, 2021(1) (2021) 2995851.
[17] C.O. Sosa-Jiménez, H.V. Ríos-Figueroa, A.L. Solís-González-Cosío, A prototype for Mexican sign language recognition and synthesis in support of a primary care physician, IEEE Access, 10 (2022) 127620–127635.

10.22060/EEJ.2026.24533.5725

[18] A.S.M. Miah, J. Shin, M.A.M. Hasan, M.A. Rahim, Bensignnet: Bengali sign language alphabet recognition using concatenated segmentation and convolutional neural network, Applied Sciences, 12(8) (2022) 3933.
[19] R. Rastgoo, K. Kiani, S. Escalera, Multi-modal deep hand sign language recognition in still images using restricted Boltzmann machine, Entropy, 20(11) (2018) 809.
[20] W. Aly, S. Aly, S. Almotairi, User-independent American sign language alphabet recognition based on depth image and PCANet features, IEEE Access, 7 (2019) 123138–123150.
[21] M.M. Rahman, M.S. Islam, M.H. Rahman, R. Sassi, M.W. Rivolta, M. Aktaruzzaman, A new benchmark on american sign language recognition using convolutional neural network, in: 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), IEEE, 2019, pp. 1–6.
[22] S. Bhushan, M. Alshehri, I. Keshta, A.K. Chakraverti, J. Rajpurohit, A. Abugabah, An experimental analysis of various machine learning algorithms for hand gesture recognition, Electronics, 11(6) (2022) 968.
[23] C.K. Lee, K.K. Ng, C.-H. Chen, H.C. Lau, S.Y. Chung, T. Tsoi, American sign language recognition and training method with recurrent neural network, Expert Systems with Applications, 167 (2021) 114403.
[24] S. Sharma, S. Singh, Vision-based hand gesture recognition using deep learning for the interpretation of sign language, Expert Systems with Applications, 182 (2021) 115657.
[25] M.A. Bencherif, M. Algabri, M.A. Mekhtiche, M. Faisal, M. Alsulaiman, H. Mathkour, M. Al-Hammadi, H. Ghaleb, Arabic sign language recognition system using 2D hands and body skeleton data, IEEE Access, 9 (2021) 59612–59627.
[26] D.S. Breland, A. Dayal, A. Jha, P.K. Yalavarthy, O.J. Pandey, L.R. Cenkeramaddi, Robust hand gestures recognition using a deep CNN and thermal images, IEEE Sensors Journal, 21(23) (2021) 26602–26614.
[27] H. Ansar, N. Al Mudawi, S.S. Alotaibi, A. Alazeb, B.I. Alabduallah, M. Alonazi, J. Park, Hand gesture recognition for characters understanding using convex Hull landmarks and geometric features, IEEE Access, (2023).
[28] S. Siddique, S. Islam, E.E. Neon, T. Sabbir, I.T. Naheen, R. Khan, Deep learning-based bangla sign language detection with an edge device, Intelligent Systems with Applications, 18 (2023) 200224.
[29] A.S.M. Miah, M.A.M. Hasan, J. Shin, Y. Okuyama, Y. Tomioka, Multistage spatial attention-based neural network for hand gesture recognition, Computers, 12(1) (2023) 13.
[30] N. Begum, R. Rahman, N. Jahan, S.S. Khan, T. Helaly, A. Haque, N. Khatun, Borno-net: a real-time Bengali sign-character detection and sentence generation system using quantized Yolov4-Tiny and LSTMs, Applied Sciences, 13(9) (2023) 5219.
[31] N.F. Attia, M.T.F.S. Ahmed, M.A. Alshewimy, Efficient deep learning models based on tension techniques for sign language recognition, Intelligent systems with applications, 20 (2023) 200284.
[32] M.M. Damaneh, F. Mohanna, P. Jafari, Static hand gesture recognition in sign language based on convolutional neural network with feature extraction method using ORB descriptor and Gabor filter, Expert Systems with Applications, 211 (2023) 118559.
[33] Shayansyn, Persian-Sign-Language-Detection: Real-time Persian Sign Language Recognition Repository, in, GitHub Repository, 2021.
[34] Salar96, Iranian-Sign-Language-ISL-Dataset Repository, in, GitHub Repository, 2022.
[35] Persian Sign Language Team, PenSLR: Persian Sign Language Recognition Project, in, GitHub Repository, 2021.
[36] Amirparsa-Sal, Sign-Language-Glove-Time-Series-Clustering Repository, in, GitHub Repository, 2023.
[37] Y. Choubsaz, O. Crasborn, S. Siyavoshi, F. Soleimanbeigi, Documenting the Use of Iranian Sign Language (ZEI) in Kermanshah, in: Proceedings of the LREC2022 10th workshop on the representation and processing of sign languages: Multilingual sign language resources, 2022, pp. 37–41.

[38] Institute for German Sign Language and Communication of the Deaf, ZEI Corpus: Hamburg Corpus of Persian Sign Language, in, University of Hamburg, 2021.
[39] H. Ranjbar, A. Taheri, ISLR101: an Iranian Word-Level Sign Language Recognition Dataset, IEEE Access, (2025).
[40] H. Ranjbar, ISLR101: Isolated Persian Sign Language Recognition Codebase, in, GitHub Repository, 2025.