

Arabic and American Sign Languages Alphabet Recognition by Convolutional Neural Network

Shroog Alshomrani¹, Lina Aljoudi¹, Muhammad Arif^{1*}

¹ Department of Computer Science, College of Computer and Information Systems, Umm Alqura University, Kingdom of Saudi Arabia

* Corresponding author's e-mail: syedmarif2003@yahoo.com

ABSTRACT

Hearing loss is a common disability that occurs in many people worldwide. Hearing loss can be mild to complete deafness. Sign language is used to communicate with the deaf community. Sign language comprises hand gestures and facial expressions. However, people find it challenging to communicate in sign language as not all know sign language. Every country has developed its sign language like spoken languages, and there is no standard syntax and grammatical structure. The main objective of this research is to facilitate the communication between deaf people and the community around them. Since sign language contains gestures for words, sentences, and letters, this research implemented a system to automatically recognize the gestures and signs using imaging devices like cameras. Two types of sign languages are considered, namely, American sign language and Arabic sign language. We have used the convolutional neural network (CNN) to classify the images into signs. Different settings of CNN are tried for Arabic and American sign datasets. CNN-2 consisting of two hidden layers produced the best results (accuracy of 96.4%) for the Arabic sign language dataset. CNN-3, composed of three hidden layers, achieved an accuracy of 99.6% for the American sign dataset.

Keywords: convolutional neural network, deep learning, American sign language, Arabic sign language.

INTRODUCTION

Communication in a broader sense is the perception of a person about an expression from another person. It is essential to understand the physical and psychological needs of the people. Communicating by speaking a language is the fastest way for a human, but signs are also used in expressing some words such as yes, no, etc., either with the head movement or the hand gesture. Non-linguistic and Sign-based communication are developed for silent or faraway communication methods. In 1620, Juan Pablo Bonet set a dictionary of sign language and letters to communicate silently [1]. The first development of sign language to help deaf and people with hearing impairment was by Martha's Vineyard, who created their sign language because they had recessive genes that made them suffer from deaf people in this town [2]. Although there were many attempts to define

sign language protocols, the most crucial work was by Charles Michel De L'Epee, a French Priest who established the first school for the deaf and was free of charge in Paris [3]. He made great effort and progress in transforming the French language into a sign language dictionary with basic ideas and concepts instead of just letters. Then, in 1800, Thomas Hopkins Gallaudet developed American Sign Language (ASL) that was inspired by French Sign language [4]. Other languages have also developed their sign language, and the research in sign languages has expanded. Each country defined its sign language [5]. Arabic Sign Language was developed based on American and British Sign Language and the shapes of nature. The Arabic language contains many dialects, including Saudi, Egyptian, Libyan, Lebanese, Iraqi, Moroccan, etc. These countries developed sign languages according to their country dialects and customs [6]. Sign language is the best way of communication for

deaf and hard-of-hearing people. Still, few people know it, especially ordinary people who believe that they do not need to learn it unless a relative whose hearing is impaired. Sign language varies according to the original language, and there is no international sign language like English which makes it even more difficult to communicate between different languages. According to the World Health Organization, the number of deaf people is on the rise, reaching 430 million people, and it is estimated that this number will rise to 700 million people by 2050 [7]. This number is enormous, and those people need to communicate with others. Sign language is the language expressed using different hand gestures, body movements, and facial expressions. These gestures refer to a sentence, word, letter, number, idea, or concept. As a result, it is difficult to expand the sign language to include all languages because there are similarities between gestures that explain various expressions, making communication between languages hard. Many techniques have appeared to communicate with deaf people, using sensors and gloves

or through images of gestures. Machine learning and deep learning have been effective in extracting features from images, videos, and sounds. These methods can overcome the communication barriers between ordinary and deaf people. The American sign language alphabet gestures are illustrated in Figure 1, and Arabic sign language alphabet gestures are shown in Figure 2.

This paper has focused on identifying sign language gestures that correspond to letters in American and Arabic languages using the Convolutional Neural Network CNN algorithm, a deep learning algorithm.

RELATED WORK

The Block diagram in Figure 3 shows the flow of a typical sign language identification system. Multiple techniques are used to capture the data for sign language recognition, such as cameras, data gloves [8, 9], and sensors like motion sensors, EMG sensors, or EEG sensors [10, 11,



Fig. 1. American sign alphabets

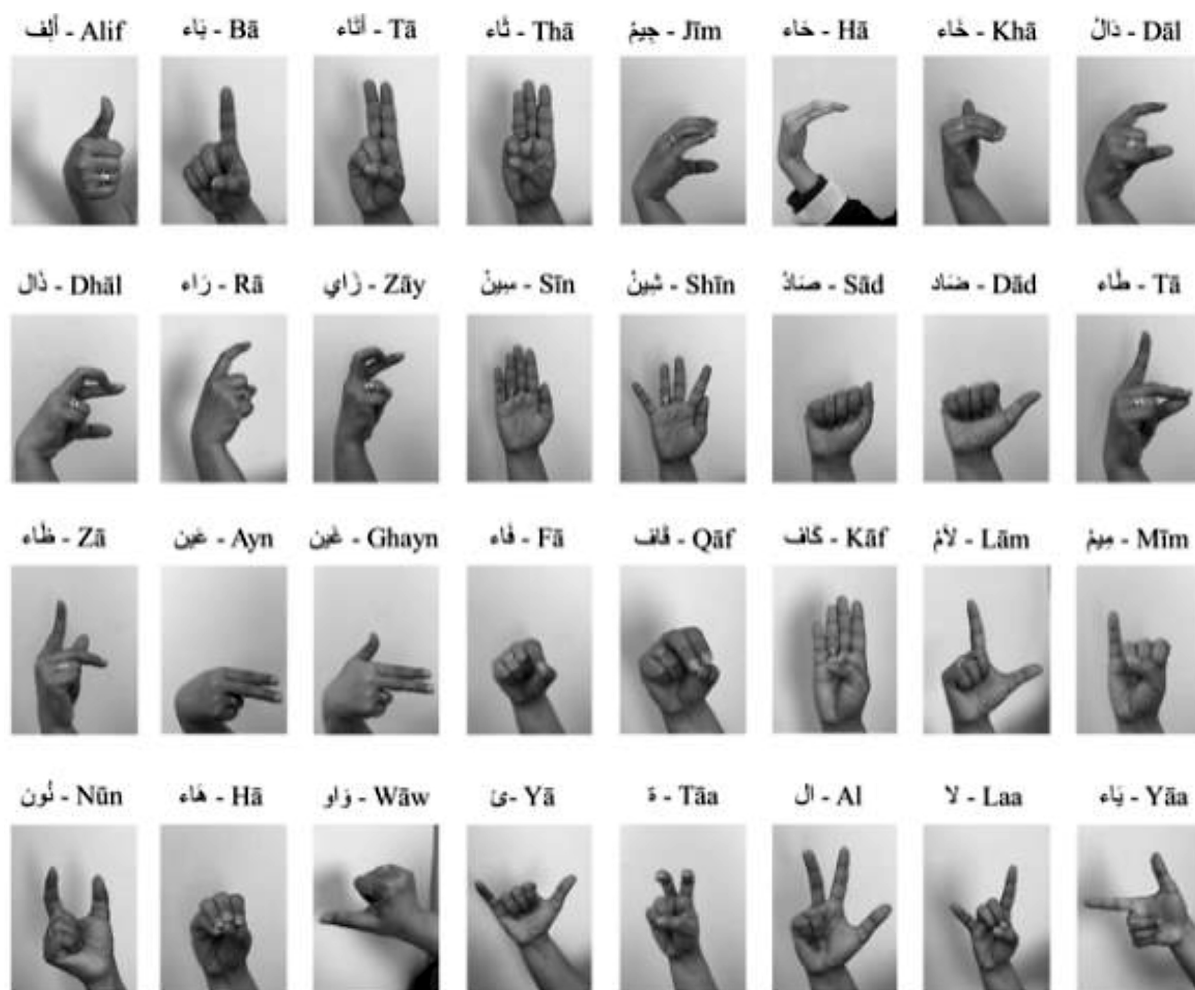


Fig. 2. Arabic sign alphabets

12]. The data may include an image, a signal, or a video stream. Preprocessing step is required to remove any unwanted noise or signal by appropriate filters, cropping or resizing images, etc. In the case of images, sometimes segmentation is also applied to remove unwanted background. In the next step, relevant features are extracted from one-dimensional signals (a time series from a sensor) or two-dimensional images. Fast Fourier transform [10, 11, 13], statistical features [8, 14], or wavelet transform [15, 16] are common methods to extract the useful features from the sensor signals or images. In case of images, Scale-invariant feature transform (SIFT) [17], Histograms of Oriented Gradients (HOG) features [18, 19], or speeded up robust features (SURF) [20, 21] are common methods to extract features. Features reduction techniques like principal component analysis (PCA) [22, 23], linear discriminant analysis (LDA) [12, 24], or independent component analysis (ICA) citealqattan2017towards, tharwat2020independent can be applied to improve the computation cost and to

remove the irrelevant features. Various classifiers are used in the literature to classify the sign alphabets in different languages. These classifiers include Artificial neural network (ANN) [25, 26, 8], Support vector machines (SVM) [26, 27, 28], hidden Markov model (HMM) [29, 30], tree-based classifiers [31, 32], K-nearest neighbor (KNN) [33] etc.

Chuan et al. [34] used a palm-sized leap motion sensor to capture the sign gestures. They applied the Support vector machine (SVM) and k-nearest neighbor (KNN) algorithms to classify 26 alphabets in American sign language. Two faculty members including a deaf person, collected the data using the leap motion sensor. Classification accuracy was 61.95% by KNN and 83.39% by the kernel SVM classifier. Abbas et al. [35] collected hand gesture images by the smartphone camera. Classification accuracy for all the alphabets was 92.5% using the SVM classifier.

Deep learning architectures [36, 37] are not new in machine learning research and are getting popular in various real-life applications. For

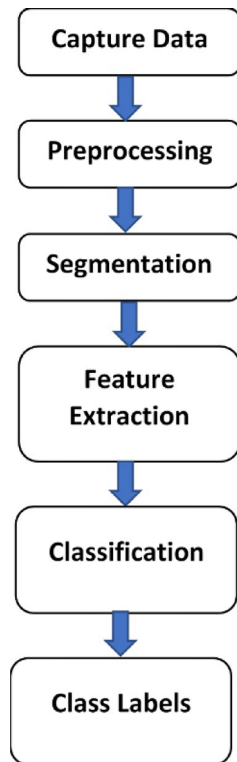


Fig. 3. Flow Diagram of sign language identification

example, in supervised learning, convolutional neural networks [38] and recurrent neural networks [39] are widely used.

Flow diagram of automatic sign language classification using deep learning methods is shown in Figure 7. Feature extraction and classification are combined in deep learning architectures. For example, Convolutional layers in CNN extract the essential features from the images. Noor Tubaiz et al. [40] proposed an Arabic sign language detection system of 40 sentences using data collected from two DG5-VHand data sensor gloves, in addition to a camera used to record the 40 sentences. For classification, a modified K-Nearest Neighbor (MKNN) is used and achieved an accuracy of 98.9%. Depth sensors were used to capture the upper part of the body to recognize hand gestures showing Arabic sign language [41]. The dataset consists of five words with 143 signs examined by ten people. They used the support vector machine (SVM) classifier with two kernels (linear and radial) applied on four SVM models with different parameter settings. The result showed that the SVM with linear kernel had the highest accuracy of 97.059%. Reema et al. [42] applied a Support vector machine (SVM) on a dataset collected from 30 persons and for each person, 30 gestures of Arabic sign language alphabets (ArSL). They noted that each letter had an accuracy

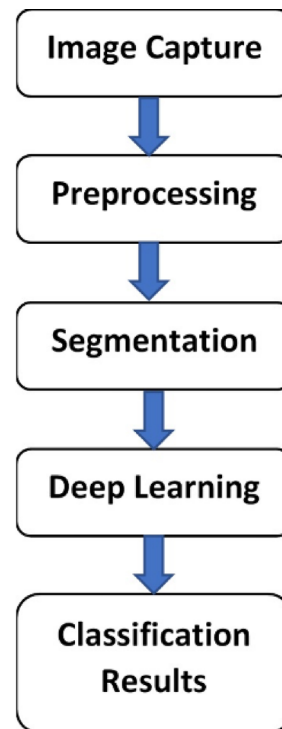


Fig. 4. Flow Diagram for sign language identification using deep learning architectures

that varies according to the hand gesture type. The accuracy of SVM for all alphabets was 63.5%.

Deep learning with a convolutional neural network is widely used for sign language recognition. Bheda et al. [43] presented the CNN classifier to detect the American sign language (ASL) alphabet and digits from 0 to 9. They used a self-generated dataset by taking pictures from five people with different skin colors and different lights. Their CNN architecture consists of three groups of convolutional layers, max-pool layer, and dropout layer. Afterward, two other groups of fully connected and dropout layers are used. The classification accuracy was 82.5%.

Qing Geo et al. [44] proposed a two-stream CNN (2S-CNN) or Inception-ResNetv2 model that combines features of Inception and ResNet. Inception-ResNetv2 can extract features better and avoid overfitting. They trained the model on the ImageNet dataset and applied it to the American sign language ASL dataset. The 2S-CNN classifier gave the best classification accuracy of 92%.

Md Asif et al. [45] proposed a posture learning framework consisted of a convolutional layer and the concept of pooling with capsules network routing for sign language recognition. They have used the Alexnet pre-trained CNN model on the Kaggle American sign language dataset. The proposed framework gave a high classification accuracy of

Table 1. Description of CNN-2 used for ASL dataset

Layer	No. of filters	Padding	Activation shape	Activation size
Input image	-	-	(28,28,1)	784
Conv2d 1(s=1)	128	same	(28, 28, 128)	100,352
MaxPooing2d 1(s=2)	-	same	(14, 14, 128)	25,088
Conv2d 2(s=1)	64	same	(14, 14, 64)	12,544
MaxPooing2d 2(s=2)	-	same	(7, 7, 64)	3,136
Flatten 1	-	-	(3136,1)	3136
Dense 1	-	-	(512,1)	512
Drouput 1	-	-	(512,1)	512
Dense 2 (softmax)	-	-	(24,1)	24

Table 2. Description of CNN-3 used for ASL dataset

Layer	No. of filters	Padding	Activation shape	Activation size
Input image	-	-	(28,28,1)	784
Conv2d 1(s=1)	128	same	(28, 28, 128)	100,352
MaxPooing2d 1(s=2)	-	same	(14, 14, 128)	25,088
Conv2d 2(s=1)	64	same	(14, 14, 64)	12,544
MaxPooing2d 2(s=2)	-	same	(7, 7, 64)	3,136
Conv2d 3(s=1)	32	same	(7, 7, 32)	1,568
MaxPooing2d 3(s=2)	-	same	(4, 4, 64)	1,024
Flatten 1	-	-	(3136,1)	3136
Dense 1	-	-	(512,1)	512
Drouput 1	-	-	(512,1)	512
Dense 2 (softmax)	-	-	(24,1)	24

Table 3. Description of CNN-2 used for ArSL dataset

Layer	No. of filters	Padding	Activation shape	Activation size
Input image	-	-	(32,32,1)	1,024
Conv2d 1(s=1)	128	same	(32, 32, 128)	131,072
MaxPooing2d 1(s=2)	-	same	(16, 16, 128)	32,768
Conv2d 2(s=1)	64	same	(16, 16, 64)	16,384
MaxPooing2d 2(s=2)	-	same	(8, 8, 64)	4,096
Flatten 1	-	-	(4096,1)	4096
Dense 1	-	-	(512,1)	512
Drouput 1	-	-	(512,1)	512
Dense 2 (softmax)	-	-	(32,1)	32

Table 4. Description of CNN-3 used for ArSL dataset

Layer	No. of filters	Padding	Activation shape	Activation size
Input image	-	-	(32,32,1)	1,024
Conv2d 1(s=1)	128	same	(32, 32, 128)	131,072
MaxPooing2d 1(s=2)	-	same	(16, 16, 128)	32,768
Conv2d 2(s=1)	64	same	(16, 16, 64)	16,384
MaxPooing2d 2(s=2)	-	same	(8, 8, 64)	4,096
Conv2d 3(s=1)	32	same	(8, 8, 32)	2,048
MaxPooing2d 3(s=2)	-	same	(4, 4, 32)	512
Flatten 1	-	-	(4096,1)	4096
Dense 1	-	-	(512,1)	512
Drouput 1	-	-	(512,1)	512
Dense 2 (softmax)	-	-	(32,1)	32

99% in sign language recognition. Kshitij Bantupalli and Ying Xie [46] used a customized sign language dataset for the video streams. The recorded videos were divided into images. CNN detects spatial features, whereas temporal features are extracted by Recurrent Neural Network (RNN). The classification accuracy obtained was between 90% to 93%. Rangel et al. [47] presented a CNN architecture inspired by Densely Connected Convolutional Neural Networks (DenseNet) to classify the RGB images of signs. For training the CNN classifier, they used more than 50,000 images of American sign language from the previous works and photos from the web camera. The training dataset is further increased by the data augmentation technique to enhance the diversity. Their model achieved a classification accuracy of 90.3%. Ghazanfar et al. [48] developed a convolutional neural network (CNN) to predict Arabic sign language. The dataset contains 54,000 images of ARSL letters [41]. The classifier gave high accuracy of 97.6%. Elson et al. [49] used a pulse-coupled neural network to deal with pose variations in 3D object recognition. Capturing image features from two different viewing angles are combined to generate optimized 3D features. They achieved a classification accuracy of 96% on a small dataset of 250 dynamic gestures in Arabic sign language.

MATERIALS AND METHODS

We have used two datasets, American sign language (ASL) [50] and Arabic sign language (ArSL) [51]. Examples from these datasets are shown in Figure 1 and Figure 2. The following sections describe both datasets.

American sign language (ASL) dataset

This dataset comprises 34727 images having a 28x28 dimension. These images are in grayscale between 0–255. The dataset contains gestures of the English letters A-Z without the letters J and Z. Because the gestures for these two letters, J and Z, are animated and not static. The distribution of classes is shown in Figure 5. The distribution of images is not equal for all the alphabets. The data is divided randomly into 60% training set (20283 images), 20% validation set (7172 images), and 20% testing set (7172 images).

Arabic sign language (ArSL) dataset

For Arabic sign language, a dataset named ArSL2018 [52] is used. This dataset consists of fifty-four thousand forty-nine gesture images for the 32 Arabic alphabets. The distribution of classes is shown in Figure 6. The dataset is collected from a group of participants of different ages. The images are resized to 32x32 pixels because they were not of the same size. Images are in grayscale between 0-255. After removing the noisy images, the remaining 41280 images are used. The distribution of classes is shown in Figure 6. The distribution is equal for all the alphabets. The dataset is divided randomly into 70% training set (28896 images), 30% testing set (12480 images). The training set is further split into training (20227) and validation (8669 images) sets.

Preprocessing of the Datasets

The data augmentation technique is used to increase the size of the training set. It generates

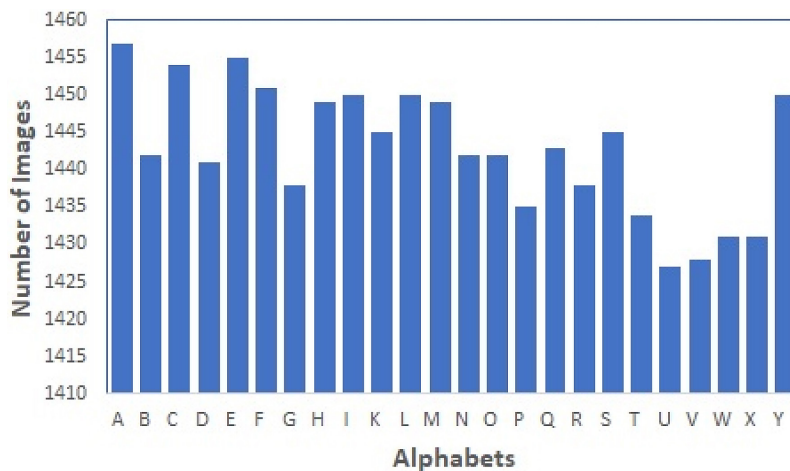


Fig. 5. Distribution of classes of ASL

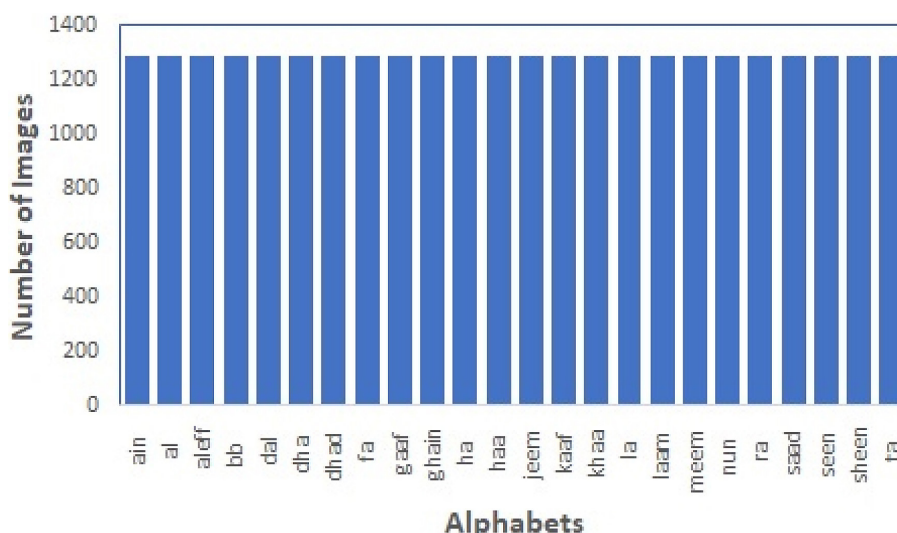


Fig. 6. Distribution of classes of ArSL

variability to improve the generalization power of the model and minimizes the overfitting. Data augmentation is done by shifting, flipping, and zooming the images. Since the images are captured from one direction Horizontal Flip that reverses rows of the pixels of an image helps the model become insensitive to the direction of image capture [50]. After data augmentation, images are normalized by dividing them by 255.

Similarity of sign gestures in Arabic sign language and American sign language

There are some similarities between the gestures of the alphabets in English and Arabic sign languages. In American sign language, alphabets A, M, and S have similar signs. Similarly, (C and

O) and (N and E) also have similar signs (figure 1). In the Arabic sign language, "Dhal" and "Zay" signs are identical, "Fa" and "Qaf" signs are also similar (Figure 2 and Figure 7). The alphabets "Dal" and "Dhal" have a similar pointing to the right. Some alphabets in both sign languages are identical in sound and signs, such as the alphabet "Lam" and "L" in Figure 7a, "Sad" and "S" Figure 7b "Ya" and "Y" Figure 7c.

Convolutional Neural Network

A convolutional neural network (CNN) is used to detect objects, shapes and edges by a sequence of filters (kernel) consisting of trainable parameters which convolve on the input images to extract the features in it. American sign

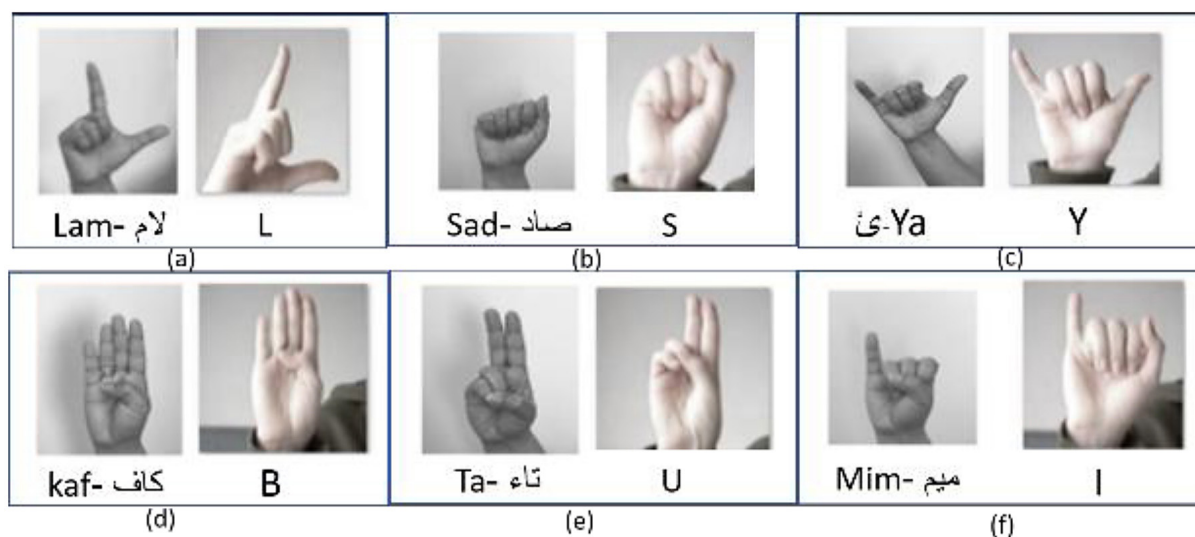


Fig. 7. Similarity of gestures in ASL and ArSL

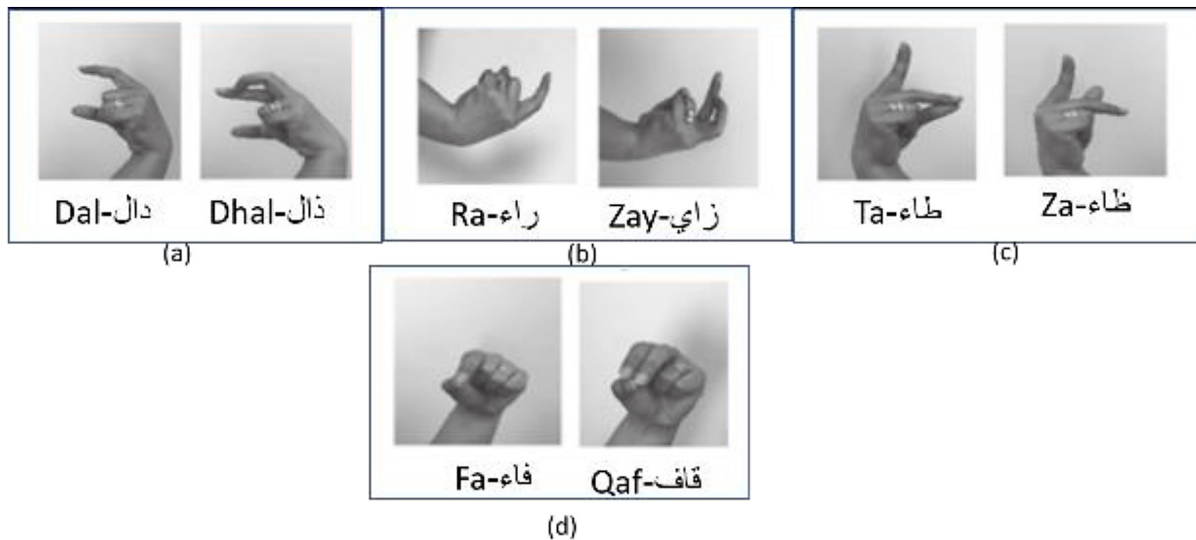


Fig. 8. Similarity of gestures in ArSL [42]

language (ASL) images are 28x28 pixels and convert them to grayscale with 28x28x1. Two CNN architectures are used, CNN-3 and CNN-2 models. Settings of CNN-2 are described in Table 3. It has two Convolutional layers with 128 and 64 neurons and two max-pooling layers. The pool size is 3x3 for the first convolutional layer and 2x2 for the second convolutional layer. Settings of CNN-3 are described in Table 2. It has 3 Convolutional layers with 128, 64, 32 neurons and three max-pooling layers. The pool size of the first Convolutional layer is 3x3, 2x2 for the other two Convolutional layers. Arabic sign language (ArSL) images size are of different sizes. All the images are resized to 32x32 pixels. Images are converted to gray color. For this dataset, two CNN architectures, CNN-2 and CNN-3,

same as the ASL dataset. Settings of both CNN-2 and CNN-3 are explained in Table 1 and Table 4.

RESULTS AND DISCUSSIONS

Two CNN architectures are applied to the two datasets. Different hyperparameter settings such as learning rate, epochs number, batch size, and optimizers (Adam and SGD) are tried to find the best settings. Figures 9 and 10 show the classification accuracy changes to different learning rates for the ASL and ArSL datasets, respectively. For ASL dataset, classification accuracy varies slightly. It then decreases for a higher learning rate (Figure 9) whereas, for ArSL dataset, the change in the classification accuracy to various learning rates is not significant (Figure 10).

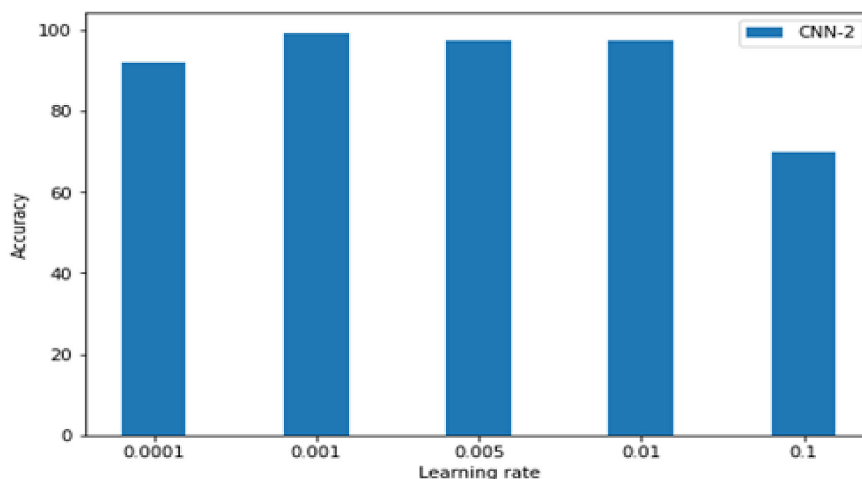


Fig. 9. Relationship between learning rate and accuracy using CNN-2 for ASL

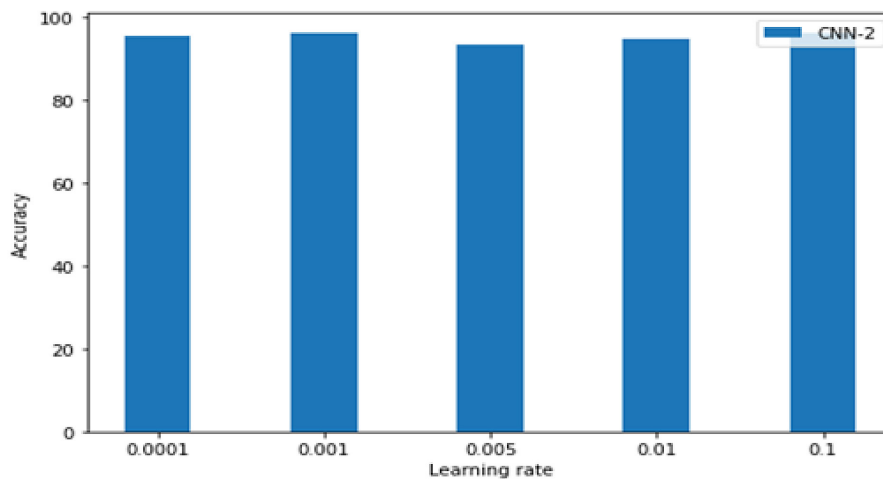


Fig. 10. Relationship between learning rate and accuracy using CNN-2 for ArSL

Table 5 shows results on the ASL dataset for CNN-3 and CNN-2 architectures with different settings. The CNN-2 architecture gave the best results with a learning rate of 0.001, batch size of 200, and Adam optimizer. Classification accuracy is 99.4%, with a recall of 99.37% and precision of 99.47%. The CNN-3 architecture performed slightly better than CNN-2 with a learning rate of 0.001, batch size of 200, and Adam optimizer. Classification accuracy is 99.6%, with a recall of 99.6% and precision of 99.61%.

This table concludes that a learning rate of 0.001 and batch size of 200 is optimal for both classifiers. The confusion matrix for CNN-3 architecture is shown in Figure 11. It can be seen from the figure that few classes are confused with

some other classes like Class 2 (Alphabet C) is misclassified as class 13 (Alphabet O) in 21 instances. Signs of alphabet C and O are somewhat similar, with only fingers touching with thumb or not. Similarly, alphabet M is confused with Alphabet A, Alphabet E, and Alphabet S. These shapes have many similarities, as shown in Figure 1. Performance analysis of both CNN-2 and CNN-3 architectures for Arabic sign language dataset for different settings are described in Table 6. For CNN-2, the learning rate of 0.001, batch size of 200, and Adam optimizer produced the best results. Classification accuracy for this setting is 96.4%, with a precision of 96.3% and recall of 96.7%. Classification accuracy for CNN-3 architecture with similar settings also produced the

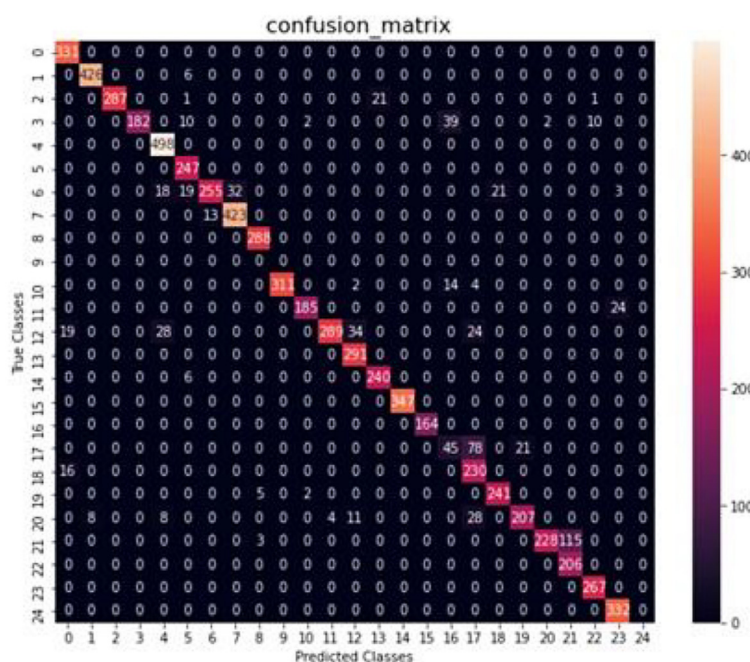


Figure 11. Confusion matrix for ASL dataset

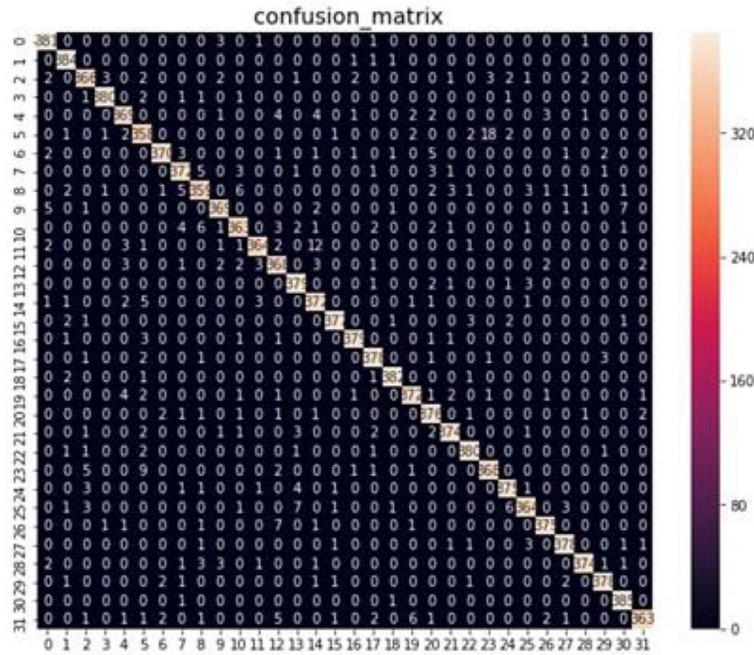


Figure 12. Confusion matrix for ArSL dataset

Table 5. Experimental results for American Sign Alphabet Dataset (ASL)

Experiment English sign	Architecture	Number of layers	Learning rate	Batch size	Optimizer	Epoch	Accuracy	F1 score	Recall	Precision
1	CNN-2	2	0.001	200	Adam	50	99.4%	99.42%	99.37%	99.47%
2	CNN-2	2	0.005	200	Adam	50	97.52%	97.6%	97.43%	97.69%
3	CNN-2	2	0.1	200	SGD	50	69.92%	63.90%	53.10%	81.01%
4	CNN-2	2	0.01	210	Adam	50	97.5%	97.5%	97.39%	97.66%
5	CNN-2	2	0.0001	200	Adam	60	92.12%	90.44%	85.6%	96.12%
6	CNN-3	3	0.001	200	Adam	50	99.6%	99.6%	99.6%	99.61%
7	CNN-3	3	0.1	250	SGD	50	97.3%	97.3%	96.9%	97.88%
8	CNN-3	3	0.01	210	Adam	50	93.92%	93.88%	93.32%	94.47%
9	CNN-3	3	0.005	200	Adam	50	96.07%	96.11%	95.85%	96.39%
10	CNN-3	3	0.0001	200	Adam	50	78.64%	76.42%	67.99%	87.68%

Table 6. Experimental results for Arabic Sign Alphabet Dataset (ArSL)

Experiment Arabic sign	Architecture	Number of layers	Learning rate	Batch size	Optimizer	Epoch	Accuracy	F1 score	Recall	Precision
1	CNN-2	2	0.001	200	Adam	50	96.4%	96.6%	96.7%	96.3%
2	CNN-2	2	0.01	210	Adam	50	94.8%	94.97%	94.7%	95.2%
3	CNN-2	2	0.1	200	SGD	50	96.17%	96.49%	95.93%	97.07%
4	CNN-2	2	0.005	200	Adam	50	93.38%	93.71%	92.95%	94.51%
5	CNN-2	2	0.0001	200	Adam	50	95.52%	95.52%	94.71%	97.23%
6	CNN-3	3	0.001	200	Adam	50	96.4%	96.4%	96.2%	96.7%
7	CNN-3	3	0.01	400	Adam	50	95.4%	95.3%	95.4%	95.4%
8	CNN-3	3	0.0001	200	Adam	50	94.04%	94.43%	92.45%	96.60%

same performance (96.4%). The confusion matrix for CNN-2 architecture with Adam optimizer is shown in Figure 12. All of the alphabets are identified correctly, with few miss-classification for

some alphabets. Alphabet Jeem is confused with alphabet Laa (14 instances), although the shape of signs is different. Similarly, the alphabet Ra is also confused on few instances with the alphabet

Table 7. Performance comparison of our Arabic/American sign language recognition models to others

Author	Sign Language	Accuracy
Reema Alzohairi et al [42]	Arabic	63.5%
Abbas Muhammad Zakariya and Rajni Jindal [35]	Arabic	92.5%
Ours (CNN-2 model)	Arabic	96.4%
Chinf-Hua Chuan et al [34]	American	83.39%
Vivek Bheda and N.Dianna Radpour [43]	American	82.5%
Ours (CNN-3 model)	American	99.6%

Saad. A performance comparison between the classification accuracy of our CNN architectures and published results is given in Table 7. The table shows that CNN-2 and CNN-3 architectures with the settings mentioned above are better than the published results on these datasets.

CONCLUSIONS

Classification models based on CNN architectures are proposed in this paper for Sign language recognition to make communication between ordinary people and deaf people easier. These models are applied to two datasets: the American sign language and the Arabic sign language. After trying different settings, CNN architectures with a learning rate of 0.001, batch size of 200, and Adam optimizer produced the best results. As a result, the classification accuracy of 96.6% is achieved in the Arabic sign language dataset and 99.6% in the American sign language dataset. In the future, we will try to test these architectures on larger datasets of people. Moreover, our future considerations are time and space complexity optimization such that these architectures can be used in mobile phones.

REFERENCES

- Juan Pablo Bonet J.B. Reduccion de las letras y arte para enseñar a hablar a los mudos. Madrid: Francisco abarca de angulo. Recuperado de biblioteca digital hispanica. Biblioteca nacional de espana; 1620.
- Groce N.E. Everyone here spoke sign language: Hereditary deafness on Martha’s Vineyard. Harvard University Press; 1985.
- Luzerne Ray L. The abbe’de l’eepe. American Annals of the Deaf and Dumb. 1848;1(2):69–76.

- Shaw E., Delaporte Y. A historical and etymological dictionary of American Sign Language. Gallaudet University Press; 2014.
- Brentari D. Sign languages. Cambridge University Press; 2010.
- Abdel-Fattah M.A. Arabic sign language: a perspective. Journal of deaf studies and deaf education. 2005;10(2):212–221.
- World Health Organization (WHO). 2021. Deafness and hearing loss.
- Kumar P., Gauba H., Roy P.P., Dogra D.P. A multimodal framework for sensor based sign language recognition. Neurocomputing. 2017;259:21–38.
- Ahmed M.A., Zaidan B.B., Zaidan A.A., Salih M.M., Bin Lakulu M.M. A review on systems-based sensory gloves for sign language recognition state of the art between 2007 and 2017. Sensors. 2018;18(7):2208.
- Savur C., Sahin F. American sign language recognition system by using surface emg signal. In 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE 2016, 002872–002877.
- Luqman H., Mahmoud S.A., et al. Transform-based arabic sign language recognition. Procedia Computer Science. 2017;117:2–9.
- AlQattan D., Sepulveda F. Towards sign language recognition using eeg-based motor imagery brain computer interface. In 2017 5th International Winter Conference on Brain-Computer Interface (BCI), IEEE 2017, 5–8.
- Hou J., Li X.Y., Zhu P., Wang Z., Wang Y., Qian J., Yang P.L. Signspeaker: A real-time, high-precision smartwatch-based sign language translator. In The 25th Annual International Conference on Mobile Computing and Networking 2019, 1–15.
- Kolivand H., Joudaki S., Sunar M.S., Tully D. An implementation of sign language alphabet hand posture recognition using geometrical features through artificial neural network (part 2). Neural Computing and Applications. 2021;1–23.
- Ravi S., Suman M., Kishore P.V.V., Eepuri K.K. Sign language recognition with multi feature fusion and ann classifier. Turkish Journal of Electrical Engineering and Computer Science. 2018;26(6):2871–2885.
- Oyeniran O.A., Oyeniyi J.O., Sotonwa K.A., Ojo A.O. Review of the application of artificial intelligence in sign language recognition system. International Journal of Engineering and Artificial Intelligence. 2020;1(4):20–25.
- Awwad S., Idwan S, Gharaibeh H. Real-time sign languages character recognition. International Journal of Computer Applications in Technology. 2021;65(1):36–44.
- Joshi G., Singh S., Vig R. Taguchi-topsis based hog

- parameter selection for complex background sign language recognition. *Journal of Visual Communication and Image Representation*. 2020;71:102834.
19. Tyagi A., Bansal S. Feature extraction technique for vision-based indian sign language recognition system: A review. *Computational Methods and Data Engineering*. 2021;39–53.
 20. Zhu Z., Jiang X., Zhang J. Sign language video classification based on image recognition of specified key frames. In *International Conference on Multimedia Technology and Enhanced Learning*. 2020;371–381.
 21. Baumgartner L., Jauss S, Maucher J., Gottfried Zimmermann G. Automated sign language translation: The role of artificial intelligence now and in the future. *CHIRA*. 2020;170–177.
 22. Sahoo J.P., Ari S., Patra S.K. Hand gesture recognition using pca based deep cnn reduced features and svm classifier. In *2019 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS)*, IEEE 2019, 221–224.
 23. Diana Alejandra Contreras Alejo and Francisco Javier Gallegos Funes. Recognition of a single dynamic gesture with the segmentation technique hs-ab and principle components analysis (pca). *Entropy*. 2019;21(11):1114.
 24. Mohamed Deriche, Salihu O Aliyu, Mohamed Mohandes. An intelligent arabic sign language recognition system using a pair of lms with gmm based classification. *IEEE Sensors Journal*. 2019;19(18):8067–8078.
 25. Rabeet Fatmi, Sherif Rashad, Ryan Integlia. Comparing ann, svm, and hmm based machine learning methods for american sign language recognition using wearable motion sensors. In *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, 2019;0290–0297. IEEE,
 26. Nigus Kefyalew Tamiru, Menore Tekeba, and Ayodeji Olalekan Salau. Recognition of amharic sign language with amharic alphabet signs using ann and svm. *The Visual Computer*, pages 1–16, 2021.
 27. Xianwei Jiang, Suresh Chandra Satapathy, Longxiang Yang, Shui-Hua Wang, and Yu-Dong Zhang. A survey on artificial intelligence in chinese sign language recognition. *Arabian Journal for Science and Engineering*. 2020;1–36.
 28. Quinn M., Olszewska J.I. British sign language recognition in the wild based on multi-class svm. In *2019 Federated Conference on Computer Science and Information Systems (FedCSIS)*, IEEE 2019, 81–86.
 29. Guo D., Zhou W., Li H., Wang M. Online early-late fusion based on adaptive hmm for sign language recognition. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*. 2017;14(1):1–18.
 30. Roy P.P., Kumar P., Kim B.G. An efficient sign language recognition (slr) system using camshift tracker and hidden markov model (hmm). *SN Computer Science*. 2021;2(2):1–15.
 31. Kumar P., Saini R., Roy P.P., Dogra D.P. A position and rotation invariant framework for sign language recognition (slr) using kinect. *Multimedia Tools and Applications*. 2018;77(7):8823–8846.
 32. Dong C., Leu M.C., Yin Z. American sign language alphabet recognition using microsoft kinect. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 2015;44–52.
 33. Malik M.S.A., Kousar N., Abdullah T., Ahmed M., Rasheed F., Awais M. Pakistan sign language detection using pca and knn. *International Journal of Advanced Computer Science and Applications*. 2018;9(54):78–81.
 34. Ching-Hua Chuan, Eric Regina, and Caroline Guardino. American sign language recognition using leap motion sensor. In *2014 13th International Conference on Machine Learning and Applications*, IEEE 2014, 541–544.
 35. Zakariya A.M., Jindal R. Arabic sign language recognition system on smartphone. In *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, IEEE 2019, 1–5.
 36. Hosseini M.P., Lu S., Kamaraj K., Slowikowski A., Venkatesh H.C. Deep learning architectures. In *Deep learning: concepts and architectures*. Springer; 2020;1–24.
 37. Calin O. *Deep Learning Architectures*. Springer; 2020.
 38. Gu J., Wang Z., Kuen J., Ma L., Shahroudy A., Shuai B., Liu T., Wang X., Wang G., Cai J., et al. Recent advances in convolutional neural networks. *Pattern Recognition*. 2018;77:354–377.
 39. Alam M., Samad M.D., Vidyaratne L., Glandon A., Iftekharuddin K.M. Survey on deep neural networks in speech and vision systems. *Neurocomputing*. 2020;417:302–321.
 40. Tubaiz N., Shanableh T., Assaleh K. Glove-based continuous arabic sign language recognition in user-dependent mode. *IEEE Transactions on Human-Machine Systems*. 2015;45(4):526–533.
 41. Almasre M.A., Al-Nuaim H. Comparison of four svm classifiers used with depth sensors to recognize arabic sign language words. *Computers*. 2017;6(2):20.
 42. Alzohairi R., Alghonaim R., Alshehri W., Aloqeely S., Alzaidan M., Behir O. Image based arabic sign language recognition system. In *International Journal of Advanced Computer Science and Applications (IJACSA)*. 2018;9(3).

43. Bheda V., Radpour D. Using deep convolutional networks for gesture recognition in american sign language. arXiv preprint arXiv:1710.06836; 2017.
44. Gao Q., Ogenyi U.E., Liu J., Ju Z., Liu H. A two-stream cnn framework for american sign language recognition based on multimodal data fusion. In UK Workshop on Computational Intelligence. Springer; 2019;107–118.
45. Jalal M.A., Chen R., Moore R.K., Mihaylova L. American sign language posture understanding with deep neural networks. In 2018 21st International Conference on Information Fusion (FUSION), IEEE 2018, 573–579.
46. Bantupalli K., Ying Xie. American sign language recognition using deep learning and computer vision. In 2018 IEEE International Conference on Big Data (Big Data), IEEE 2018, 4896–4899.
47. Daroya R., Peralta D., Naval P. Alphabet sign language image classification using deep learning. In TENCON 2018-2018 IEEE Region 10 Conference, IEEE 2018, 646–650.
48. Latif G., Mohammad N., Roaa A., Rawan A., Alghazo J., Khan M. An automatic arabic sign language recognition system based on deep cnn: An assistive system for the deaf and hard of hearing. International Journal of Computing and Digital Systems. 2020;9(4):715–724.
49. Elons A.S., Abull-Ela M., Tolba M.F. A proposed pcnn features quality optimization technique for pose-invariant 3d arabic sign language recognition. Applied Soft Computing. 2013;13(4):1646–1660.
50. Ghazanfar L., Jaafar A., Nazeeruddin M., Roaa A., Rawan A. Arabic alphabets sign language dataset (arasl). Mendeley Data, 2018;1.
51. Tecperson. Sign language mnist, 2017.
52. Latif G., Mohammad N., Alghazo J., Roaa A., Rawan A. Arasl: Arabic alphabets sign language dataset. Data in brief. 2019;23:103777.