Al-Salam Journal for Engineering and Technology
Journal Homepage: http://journal.alsalam.edu.iq/index.php/ajest
e-ISSN: 2790-4822p-ISSN: 2958-0862

Iraqi Sign Language Translator system using Deep Learning

Raja'a M. Mohammed[1,2*], Suhad M. Kadhem[1]

[1] Computer Science Department, Collage of Sciences for Women, University of Baghdad, Iraq.
[2] Computer Science Department, University of Technology, Baghdad, Iraq.

*Corresponding Author: Raja'a M. Mohammed

DOI: https://doi.org/10.55145/ajest.2023.01.01.0013
Received October 2022; Accepted December 2022; Available online January 2023

ABSTRACT: The deaf and mute use sign language by moving their hands, faces, and bodies to talk to each other or normal people. Sign language is non-verbal communication, which is the process of communication by sending and receiving messages without words between people. The number of deaf people is increasing in the world and Iraq in particular, in addition to the problems in communicating with the world and the difficulty of learning sign language by deaf and hard-of-hearing families, there must be other ways to help the deaf communicate efficiently with ordinary people and learn sign language easily. One such method is to use artificial intelligence to create translation software and recognize hand gestures. This paper presents a computer program that can translate Iraqi sign language into Arabic (text). First, the translation starts with capturing videos to make up the dataset, the proposed system uses a convolutional neural network (CNN) to classify sign language based on its features to impute the meaning of the sign. The accuracy of the part of the proposed system that translates sign language into Arabic text is 99% for words sign.

Keywords: Sign language, Iraqi sign language, Video processing, Deep learning, CNN

2. RELATED WORK

Many studies on sign language translation and redefinition have been applied recently; these studies use various input devices, specialized features, and machine learning algorithms. The following are the most recent works on sign language translation: In [11], Hidden Markov Models (HMM) are used to recognize the dataset containing 20 isolated words from the Standard Arabic sign language. In [12], Leap Motion Controller (LMC) is used as the input device to make the dataset. the 12 features are selected from twenty-three features produced from the device, then the classification is done using Naive Bayes (NB) and Multilayer Perceptron (MLP). In [13], the Microsoft Kinect Sensor Depth Camera was used to extract the 3D information of the features. Using the Hidden Markov Model classifier, the 40 signs from the standard ArSL were used for training the system. In [14], Arabic Sign Language system translation is provided. The proposed system begins with extracting the morphological features from the input images for the 3 Arabic alphabets. Finally, the Artificial Neural Network (ANN) is used for classification. In [15], the Microsoft Kinect sensor is used for automatic Arabic sign language recognition in Real-Time. This system used the Dynamic Time Warping algorithm to compare signs to classify the 30 isolated sign words from standard ArSL. In [16], The Arabic sign language recognition system is proposed for alphabet classification. The preprocessing begins with segmenting the skin color, the Hull Convex color can be applied to the skin color segments, and the convexity defect draws the three points for each sign. Finally, the distance between these three points is calculated to obtain the vector of the features; different classification methods use this vector. In [17, 18, 19], Convolutional Neural Networks (CNN) are based on deep learning architecture and standard Arabic sign alphabetic letters used in the proposed gesture recognition system. In [20] designed, a recognition system for American sign language using CNN with 125 words sign.In [21], develop a recognition system for American sign language using CNN with standard American sign alphabetic letters. In [22], the application of Iraqi sign language was presented for smartphone devices to translate Iraqi sign language into what it means in classical Arabic and vice versa based on the Standard Qatari Sign Language dictionary.
This paper will introduce the Application of the Iraqi sign language translation system; vision-based is the type of this system. The major contribution of the proposed method is building the first fully-labeled dataset for the Iraqi sign language, which includes common words of the Iraqi sign language.

*Corresponding author: cs.19.93@grad.uotechnology.edu.iq

3. THE PROPOSED SYSTEM

This section presents the design of an automatic translation system from Iraqi sign language to Arabic text. The system consists of two primary stages: data pre-processing and classification (translation to test). The Proposed Iraqi Sign Language Translation System (IRSLT) is clarified in fig.1.

3.1 Capturing and Data pre-processing

Data preprocessing is a crucial stage in Machine Learning since it helps enhance the data quality and thus facilitates extracting useful features from it. To construct a translation system, a dataset is required, and as we have seen in earlier studies, the dataset was either public or private. The public dataset is available in Arabic sign language and contains only Arabi sign letters. Consequently, the Iraqi Sign Language (IrSL) dataset was created from scratch using the Iraqi Sign Language Dictionary [23]. The words sign used for a dataset in the proposed system are shown in table 1. The initial step of this system is videos captured through the camera. Then apply the following steps:

1. Selecting frames: The proposed system uses visual information with the modification of the length of the video and selecting specific numbers of the video frames, as shown in figure 2.
2. Converting to Grayscale
3. Noise Removal with Gaussian Filter as shown in fig 3
4. Brightness and Contrast Adjustment using gamma correction
5. Aggregation frames in one Image by collecting all of the video frames into a single image, as shown in fig 4

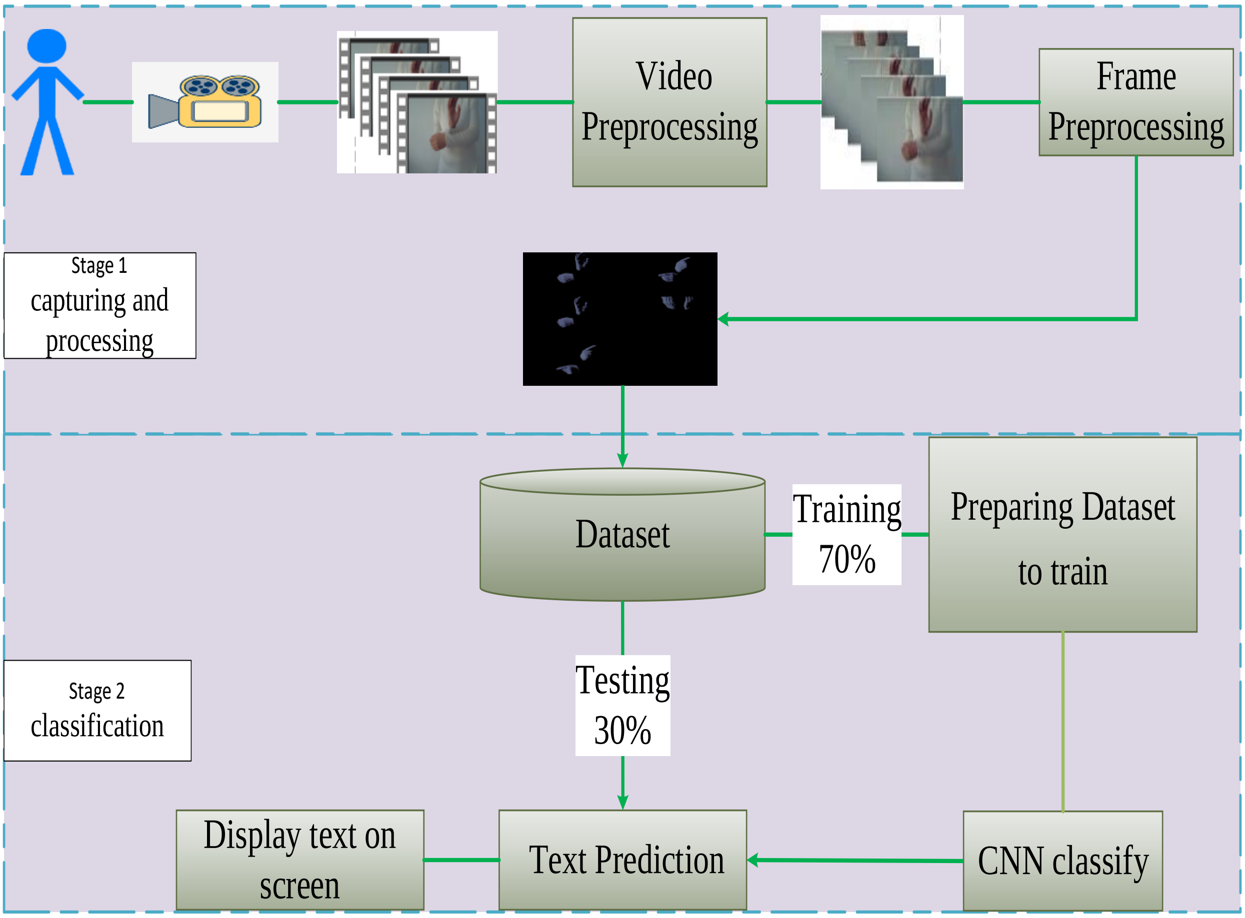

Figure 1. The Main Framework for the Proposed Model

Table1. Words in Irsl used in the proposed dataset

Words in English	Words	Words in English	Words	Words in English	Words
number one	رقم واحد	gloves	قفازات	Shower	الاستحمام
mortar	هاون	noon	الظهر	a cup of coffee	كوب من القهوة
number eleven	رقم احد عشر	work hours	دوام	falafel	فلافل
Mail	بريد	dolma eater	اكلة الدولمة	necklace	قلادة
envelope or message	ظرف او رسالة	a knife	سكين	Diamonds	ماس
marriage	زواج	Spread clothes on the rope	نشر الملابس على الحبل	Air Freshener	معطر جو
hour	ساعة	laundry	غسل ملابس	dish washing	غسل الصحون
Multiplication sign	علامة الضرب	take, buy, or own	يأخذ او يشتري او يمتلك	number four	رقم اربعة
Friend	صديق	glass	قدح	brother and sister	اخ او اخت
night	ليل	number three	رقم ثلاثة	number six	رقم ستة
fork	شوكة	needle	ابرة	Rectangle	مستطيل
number two	رقم اثنان	scissors	مقص	Travel abroad	سفر خارج البلاد
notebook	دفتر	loofah	ليفة	stuffed food	محشي

FIGURE 2. video preprocessing (selected frames)

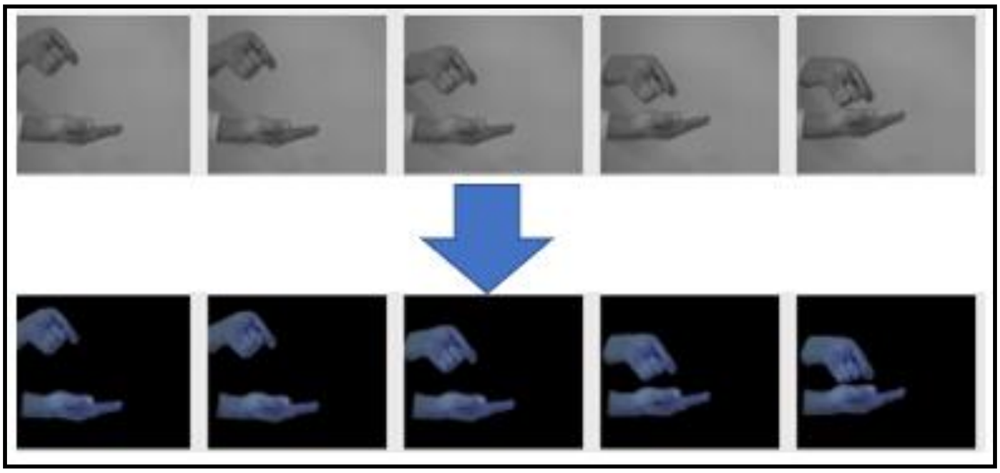

FIGURE 3. Apply the Gaussian smoothing process to a grayscale image

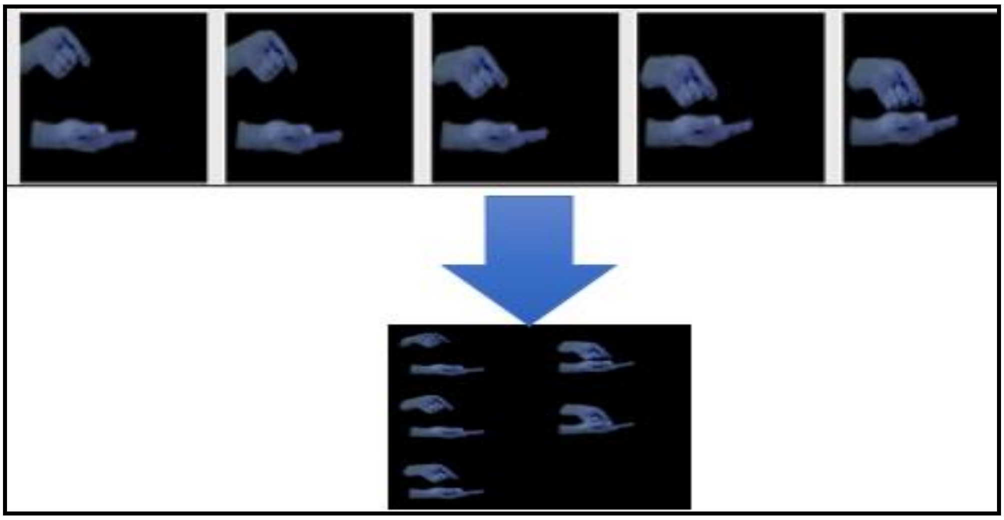

FIGURE 4. collects frames of video into a single image

3.2 Classification (translation to test)

the classification stage consists of describing the process of preparing the dataset and developing classification models:

1. Preparing dataset: It is necessary to improve the hand gesture images to use them in the classification phase. This step consists of the following sub-steps:
   a) The dataset is split into the training set and testing set. Seventy percent of the dataset is included in the training set, while the remaining thirty percent is reserved for the testing phase. Seventy percent of the dataset is divided into validation sets comprising twenty percent and training sets containing eighty percent. During the training phase, the validation procedure evaluates the model.
   b) Data Normalization: Normalization is a commonly used procedure that seeks to rescale the input data to balance the data within the same range of values in pre-processing.
   c) Data Augmentation: Data augmentation improves the performance and output of machine learning models by generating new and distinct training dataset samples. The proposed system used the Rotating, shifting sigmoid, and rescale.

2. Classification model: The most crucial aspect of supervised machine learning is classifier selection. Recently, deep learning techniques have shown better performance than machine learning approaches.

Their structures are adaptable to the difficulty of the problems. At this step, CNN's layers are put together. The CNN model is divided into two phases, the first stage is implemented for feature extraction purposes, and the second stage is for classification purposes, as shown in Fig. 5

Here are the layers of the feature extraction phase:

a) A convolution layer with 32 and 64parameters, a kernel size of (3*3), and a ReLU function.
b) A pooling layer that uses Max Pooling and a pooling size of (2*2)

The second phase of the CNN model is for classifying and has several layers; the following layers are used:

a) Dropout (0.25)
b) Flatten layer.
c) Fully connected layer with 512 nodes.
d) Dropout layer (0.5).
e) A fully connected layer with 39 output nodes (the number of classes) and a SoftMax activation function.

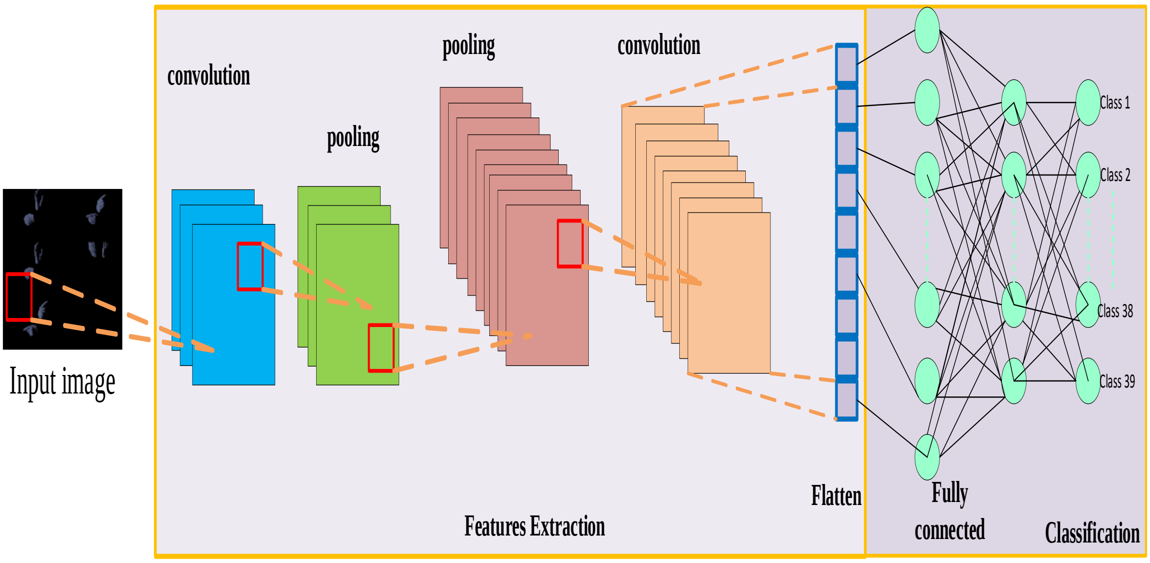

FIGURE 5. CNN model

4. EXPERIMENTAL RESULTS

The Iraqi Sign Language Recognition system was implemented with Python 3.7, an Intel(R) Core (TM) i7-8565U CPU @ 1.80GHz x 1.99GHz, Windows 10 Home Basic (64-bit), 8GB RAM, and an iPhone Camera 12.2 MP. the dataset contains 390 IrSL sign word videos. Each video is 3 seconds long. During preprocessing, the 34,38,46, 54, and 56 frames are chosen to form the final CNN-trained frame. Following the augmentation phase, the dataset contains 3000 images for training and 117 images for testing. For 50 epochs, the system was trained using the Adam optimizer and the loss function. Table 2 shows the values of Training and validation accuracy values, Training and validation loss of the dataset; as illustrated, the values are closed, indicating that your model can generalize.

Table 2. Training and validation Accuracy, Training, and validation loss of words dataset

Training Accuracy	Training Loss	validation accuracy	validation Loss
0.9964	0.0068	0.9979	0.0037

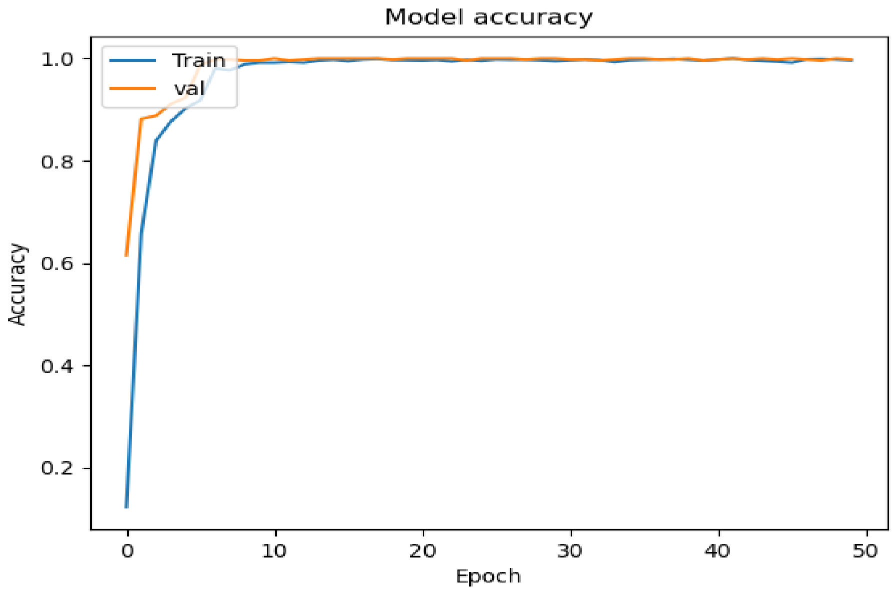

FIGURE 6. Accuracy of training

Fig. 6. shows that the training accuracy set started at 15%, and the validation accuracy set started at 65%. The number of samples causes this significant difference. The training accuracy starts to go up slowly, and by epoch 5, it is close to the verification accuracy. Then, it gets stable from epoch ten until the end of the training phase.

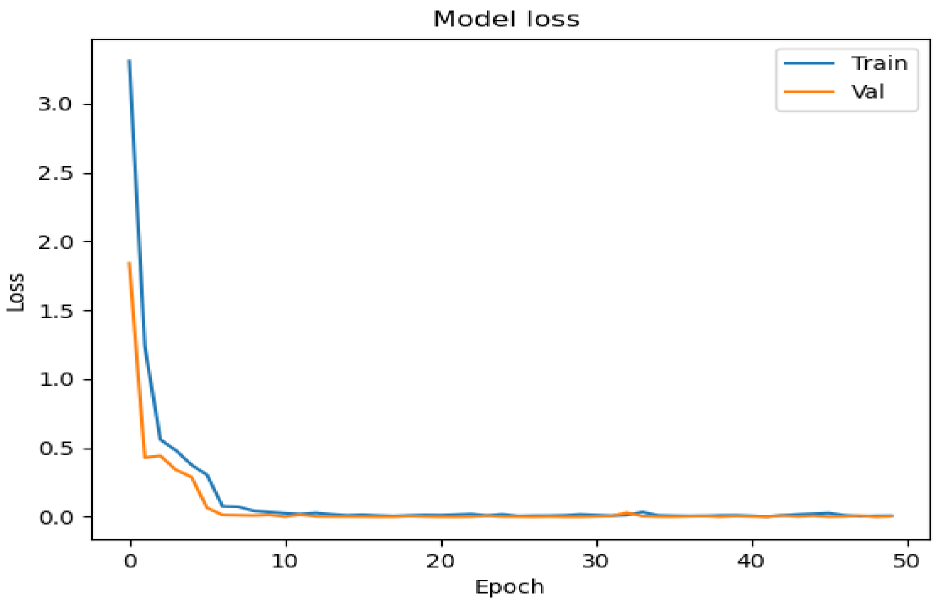

FIGURE 7. Loss of training words dataset

Fig.7 shows that the training set began with a high loss, started at 4, and the validation loss set started at 1; the training loss is reduced to less than 1.0 between epochs 4 and 6 and then became nearly steady between epochs 10 and 50.

Table 3. recall, precision, and F-score measurements

	Precision	recall	f1-score
Macro average	100	100	100

Weighted average	99	100	100

Table 3 shows the recall, precision, and F-score measurements and the average rate for each metric.

Table 4. Comparison between the proposed system and other work

Paper no.	Input device	dataset	Classify method	Accuracy
[11]	camera	20 words of ArSL	HMM	82%
[12]	LMC	Alphabet of ArSL	(MLP) (NP)	99%
[13]	Microsoft Kinect	40 signs from standard Arabic sign language	HMM	90%
[14]	camera	3 from the Alphabet of Arabic sign language	ANN	73%
[15]	Microsoft Kinect	30 words from regular ArSL	Dynamic time warping matching	97.58%
[17.18.19]	camera	Alphabet of ArSL	CNN	97.6% 97.2% 90%

As mentioned in this paper, each country has its unique sign language because of its culture and traditions. So, Iraq's sign language differs from other Arab countries' sign languages to evaluate the proposed system against previous studies. The comparisons focused on the dataset and the classification model. Table 4, The studies that used LMC and Microsoft Kinect as the input devices get good accuracy even with traditional classification methods. The studies that used Deep learning also got good accuracy; all the previous studies used the Arabic sign language, but there is no study of Iraqi sign language, and there is no dataset for Iraqi sign language, so this paper produced an Iraqi sign language translation system.

5. CONCLUSIONS

This paper provides an automatic Iraqi sign language translation system for the deaf/mute community; the system was designed to facilitate communication between deaf/mute and hearing individuals. People can use the proposed method to learn sign language and help deaf or mute people when needed. Thirty-nine signs in Iraq were changed to use the system. Using a webcam, the capturing process focused on the hand area to reduce the time required for processing. Because all sign words are continuous, the process step saves all frame sign videos in the videos. With the CNN model, the features are extracted and put into groups. The CNN model was made to help people who are deaf or mute talk to each other. The system was able to translate Iraqi sign language into Arabic text with a 99% accuracy rate. The recommendations for future work This proposed system has significantly improved predicting Iraqi sign language. The proposed method can be expanded by adding specific enhancements to the database due to the following limitations: The dataset was captured at a single location and with a single background by a single participant. Insufficient modifications in lighting and noise levels were implemented, and the system can be improved by mixing the different types of signs, such as combining hand and lip movements.

6. References

[1] S. M, "Real-time sign language gesture (word) recognition from video sequences using CNN and RNN”, *Intelligent Engineering Informatics. Springer, Singapore, pp. 623-632, 2018.*

[2] R. M. M. ., S. M. K., "A Review on Arabic Sign Language Translator Systems*", Journal of Physics: Conference Series, vol. 1818, no. 1, 2021*.

[3] S. a. S. V. A., "A convolutional neural network to classify American Sign Language fingerspelling from depth and color image", *Expert Systems, vol. 34, no. 3, 2017.*

[4] T. Bezdan and N. B. Džakula, "Convolutional Neural Network Layers and Architectures", *International Scientific Conference On Information Technology And Data Related Research, Pp. 445-451, 2019*.

[5] A. Khan, "survey of the recent architectures of deep convolutional neural networks.," *Artificial Intelligence Review, vol. 53, no. 8, pp. 5455-5516. 2020.*

[6] Gholamalinezhad, Hossein and H. Khosravi, "Pooling Methods in Deep Neural Networks, a Review", *arXiv preprint arXiv, 2020.*

[7] R. Yamashita, M. N. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into Imaging, 2018*.

[8] Albelwi, N. R and Y. M. Alginahi, "Real-time Arabic sign language (arsl) recognition." *International Conference on Communications and Information Technology*, 2012.

[9] ZakiAbdo and Mahmoud, "Arabic sign language recognition", *International Journal of Computer Applications, vol. 89, no. 20, pp. 19-26, 2014.*

[10] Amin and Omar, "HMM-based automatic Arabic sign language translator using Kinect", *Tenth International Conference on Computer Engineering & Systems IEEE, 2015.*

[11] Eqab, Abdulla and T. Shanableh, "Android mobile app for real-time bilateral Arabic sign language translation using leap motion controller", *International Conference on Electrical and Computing Technologies and Applications IEEE, 2017.*

[12] Luqman, Hamzah and S. A. Mahmoud, "Transform-based Arabic sign language recognition", Procedia Computer Science, pp. 2-9, 2017.

[13] A. M. Ahmed, "Arabic sign language translator," Journal of Computational Science, pp. 1522-1537, 2019.

[14] L. Elsiddig and M. Mohamed.," Translating Arabic Sign Language ARSL To Text Using Artificial Neural Networks," pp. 43-47, 2017.

[15] A.-G. A.-R. Abdel-Samie, F. A. Elmisery, A. M.Brisha, and A. H. Khalil, “Arabic Sign Language Recognition Using Kinect Sensor”, *Res. J. Appl. Sci. Eng. Technol., vol. 15, no. 2, pp. 57–67, 2018*.

[16] M. M. Kamruzzaman, "Arabic sign language recognition and generating Arabic speech using convolutional neural network", *Wireless Communications and Mobile Computing, 2020.*

[17] R. A. Kadhim and M. Khamees, "A Real-Time American Sign Language Recognition System using Convolutional Neural Network for Real Datasets", *TEM Journal, vol. 9, no. 3, pp. 937-943, 2020.*

[18] R. A. Alawwad, O. Bchir, and M. M. B. Ismail, "*Arabic Sign Language Recognition using Faster R-CNN*", International Journal of Advanced Computer Science and Applications, vol. 12, no. 3, 2021.

[19] R.EL.Rwelli; O.R. Shahin; A.I.Taloba, “Gesture-based Arabic Sign Language Recognition for Impaired People base on convolution neural network.” IJACSA, 2021.

[20] A Sunitha Nandhini, Shiva Roopan D, Shiyaam S, Yogesh S “Sign Language Recognition Using Convolutional Neural Network", *Journal of Physics: Conference Series” 2021, doi: 10.1088/1742-6596/1916/1/012091*.

[21] *R.S. Sabeenian, S. Sai Bharathwaj, M. Mohamed Aadhil,* “Sign Language Recognition Using Deep Learning and Computer Vision, Jour of Adv Research in Dynamical & Control Systems, Vol. 12, 05-Special Issue, 2020

[22] M. A. Alobaidy and S. K. Ebraheem, “Application for Iraqi sign language translation on Android system,” no. October 2020, pp. 5227–5234, 2021, DOI: 10.11591/ijece.v10i5.pp5227-5234.