DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

Automatic Translation From Iraqi Sign Language to Arabic Text or Speech Usign CNN

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

[1]*Computer Science Department, Collage of Sciences for Women, University of Baghdad, Baghdad, Iraq*

[2]*Computer Science Department, University of Technology, Baghdad, Iraq*

[1]*cs.19.93@grad.uotechnology.edu.iq*, [2]*110102@uotechnology.edu.iq*

Abstract— *Sign language (SL) is Non-verbal communication and a way for the deaf and mute to communicate without words. A deaf and mute person's hands, face, and body shows what they want to say. Since the number of deaf and dumb people is increasing, there must be other ways to learn sign language or communicate with deaf and dumb people. One of these ways is using advanced technology to produce systems that help the deaf/dumb, such as creating recognition and sign language translators. This paper presents an application that works on the computer for machine translation of Iraqi sign language in two directions from sign language to Arabic language (text/speech) and from Arabic language(text) to Iraqi sign language. The proposed system uses a Convolution Neural Network (CNN) to classify sign language based on its features to predicate the sign meaning. The sign language to Arabic language(text/speech) part of the proposed system has an accuracy of 99.3% for letters.*

Index Terms—*Sign Language, Iraqi Sign Language, Deep Learning, CNN,* gtts.

I. INTRODUCTION

Sign language (SL) is a type of communication used by the deaf and mute community in which an idea is represented via the use of the hand, face, and body. Sign language is not a collection of random signs; it is the deaf community's natural language. It is the deaf society's sole mode of communication. Sign language recognition can be broken down into two distinct categories: the manual sign and the non-manual sign. Non-manual signs include a person's facial expressions, lip patterns, body posture, and head position, whereas manual signs are associated with the hands. Hand gestures, often known as hand signs, are crucial to any gesture-based language. At the same time, non-manual signs comprise several other forms, including facial expressions and physical motion [1][2]. The World Federation of the Deaf estimates over 72 million deaf persons worldwide. Over 80% of them reside in poor countries. They collectively employ around 300 distinct sign languages [3]. As in the spoken language, SL varies from one country to another. SL is not international, but there are several sign languages, for example, American, British, Arabic, and others[4].

Sign language has different types: Isolated sign, which is one sign gesture; a continuous sign which offers a complete clause sign [5]. All sign language recognition systems contain four primary stages: dataset collecting, preprocessing, features extracting, and classifying signs images or videos. The dataset is gathered using various capturing devices, such as a webcam, a phone camera, or colorful gloves[6]. Microsoft Kinect sensor was developed in 2010 as a set of sensors to be used as an input device for the XBOX gaming console. It has three sensors: RGB, sound, and depth [7]. Electronic Glove: a glove with built-in sensors that recognize hand gestures. People who are deaf or hard of hearing must wear a glove connected to sensors that collect data [8]. Leap Motion Controller is a small, simple device for tracking hands and fingers. It is equipped with three Infrared lamps that emit light with a

Received 30/October/2022; Accepted 11/February/2023

 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

wavelength of 850 nm and two cameras that capture the reflected light in this spectrum[9]. Sign language recognition systems can be classified as Sensor-based and Vision-based. Sensor-based systems usually use specialized equipment such as the Microsoft Kinect sensor, Electronic Glove, and Leap motion controller. In contrast, Vision-based systems rely only on standard cameras and image-processing techniques to interpret gestures [10].

Learning sign language can be challenging for hearing individuals, just as learning oral languages can be challenging for deaf/mute people. It is conceivable to build translation systems and recognition of sign language to enhance communication between various groups as a solution to this issue. Additionally, the existence of these systems makes it easier to learn sign language, which may be done without having to read dictionaries or attend pricey sign language schools.

Many studies in this field have been presented over the years.Artificial intelligence methods ranging from machine learning to deep learning were used.All studies focused on the unified Arabic sign language or the languages of Arab countries such as Saudi Arabia, Qatar, and Jordan, and there are no studies or research on Iraqi sign language. This paper will introduce the Application of the Iraqi sign language translation system; vision-based is the type of this system. The system works on a dataset that includes the Iraqi alphabetic sign language, which is the base of the language with which can be formed different words sign. The following are the main contributions of the current work:

1. Building the first dataset for the Iraqi sign language.
2. Develop, train, and evaluate a deep learning model which translates the Iraqi sign language into Arabic (text/ speech) language.

II. RELATED WORK

Many types of research have been done on sign language translation; most of these works are focused on utilizing the input device, selected features, and machine learning algorithms. The following are the most recent works:

In[11], Youssif et al. introduced an automatic Arabic Sign Language (ArSL) recognition system based on the dataset containing 20 isolated words from the Standard Arabic sign language. The hand was represented by a model consisting of the palm, the five fingers as ridges at finer scales and the fingertips as even finer scale blobs. In the classification, Hidden Markov Models (HMM) are used and achieved an accuracy of 82.22%.

In [12], Leap Motion Controller (LMC) is used as the data acquisition stage returns twenty-three features for each frame of data. However, in this system, select 12 features that Naive Bayes (NB) and Multilayer Perceptron (MLP) are used for classification, The accuracy of the system was 99% and 98%.

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 system has been trained to recognize 40 signs from the standard Arabic sign language, The accuracy of the system was 90%. In [14], Translating the Arabic Sign Language system is presented by using an Artificial Neural Network(ANN). This system uses the morphological features extracted from the captured sign of 3 alphabets Arabic.

In[15], the Real-Time system is presented for an automatic Arabic sign language recognition system based on the Kinect sensor using the Dynamic Time Warping algorithm to compare signs to recognize 30 isolated words from standard Arabic sign language signs. For signer-dependent online case, the system achieves recognition rate of 97.58%. On the other hand, for signer-independent online case, the system achieves a recognition rate of 95.25%.

In[16], various classification methods are used, such as Stochastic Gradient Descent, Random Forest, Logistic Regression, K-NN, Decision Tree, SVC, and Linear SVC. for the recognition of the Arabic sign language alphabet. Skin color segmentation is used as preprocessing; based on the skin

Received 30/October/2022; Accepted 11/February/2023

ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

segmentation results, the Hull Convex color can be applied. The Hull convex step is followed by drawing the convexity defect three points for each sign, then calculating the distance between these three points to produce the features; the vector includes the number of error points, the spaces, and the different locations of the end of the defect.

In [17], introduced a system to identify dynamic, isolated Arabic gestures by using one or both hands with facial expressions. Arabic video signals were used as input. Each sign was treated as a secret gesture. The features are extracted using the intensity histogram and integrated with the Gray Level Co-occurrence Matrix (GLCM) features, and Euclidean distance is used to classify them. The experimental results show that the proposed system recognizes signs with a accuracy of 95.8%.

In [18], introduced the system for Arabic sign language gesture recognition. Experiments were performed using the ArSL2018 dataset. The CNN was used for classification, and SMOTE oversampling method on the ArSL2018 dataset used for improved the results. The proposed system has a recognition rate of 97%.

In [19], developed a system for recognition Arabic sign language ,the hand movements in the dataset were captured using DG5-V hand glove with wearable sensors, for classification used the CNN model is trained an tested Arabic sign language alphabetic. The proposed system has a recognition rate of 90%.

In [20], design a recognition system for American sign language using CNN with 125 words sign. In [21], design a recognition system for American sign language using CNN with standard American sign alphabetic letters.

In [22], 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.

III. PROPOSED SYSTEM

This section includes the IrSL architecture proposed for classifying gestures in Iraqi Sign Language. In addition, this section explains the IrSL dataset and the preprocessing techniques that were used to it. The Proposed system consists of two phases machine translation from Iraqi sign language to Arabic text and translation from Arabic text to sign language: first phase consists of capturing and processing, classification (translation from sign language to Arabic text), and converting text into a sound. The Proposed Iraqi Sign Language Translation IRSLT system is clarified in the *Fig. 1*.

A. Machine Translation from Sign Language to Arabic Text

This section presents how the proposed system work for Machine Translation from Iraqi sign language to Arabic Text or speech. Consists of the following subsection:

a) Capturing Images

The dataset for Iraqi Sign Language (IrSL) was created from scratch using the Iraqi Sign Language Dictionary [23]. The initial step of this system is image capture through the camera, withe the decimation of the image is 1820x720 ,in IrSL.in *Fig. 2* shown samples of IrSL letters in the generated dataset.

Received 30/October/2022; Accepted 11/February/2023

ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

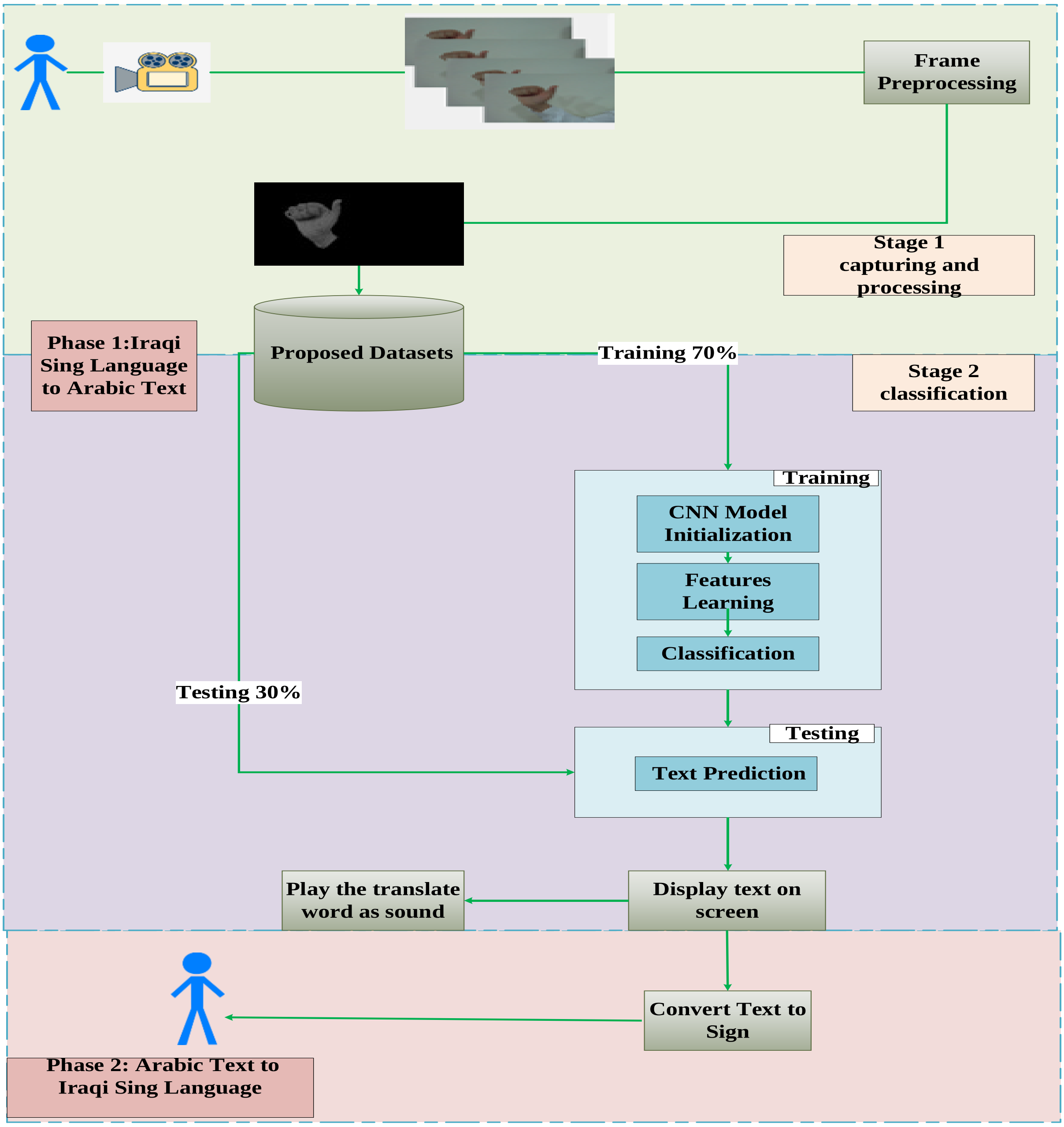

FIG. 1. THE MAIN FRAMEWORK FOR THE PROPOSED MODEL.

FIG. 2. SAMPLES OF IRSL LETTERS IN THE GENERATED DATASET.

Received 30/October/2022; Accepted 11/February/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

b) Frame Preprocessing

Frame preprocessing includes hand skin mask, converting to grayscale, noise removal with a gaussian filter, gamma correction to make the images more contrasty. The frame preprpcessing is clarified in the *Fig. 3.*

→ Hand skin mask: In this step, the hand skin region is detected. First, convert the RGB image to a BGR image to easily covert to HSV, then determine the skin region and apply the threshold to separate the hand region from the background.

→ Gaussian filter: it has a different kernel size; by experiment, the size of 5x5 will be chosen to blur the edge of the hands, the Gaussian filter has been used to reduce noise in the all images.

→ Brightness and Contrast Adjustment: all the images in the datasets employed in the proposed system was created from scratch, owing to the fact that many cameras and video recorders do not accurately capture brightness, because they must be corrected, the gamma correction function is employed to adjust the image's luminance.

→ Converting to Grayscale: Generally, grey-scale pixels are stored in an 8-bit integer array, resulting in 256 distinct shades of grey ranging from black to white. Converting a color image to a greyscale one minimizes the following levels' computing load.

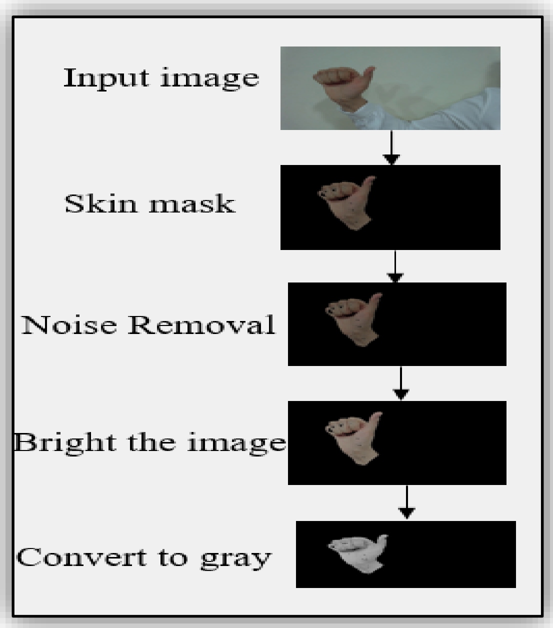

FIG. 3. THE PREPRESSING STEP OF PROPOSED SYSTEM.

c) Classification (translation from Iraqi sign language to Arabic text)

This section will describe the process of preparing the dataset and developing classification models:

A- Preparing dataset

It is necessary to improve the hand gesture images to use them in the classification phase. This stage consists of the following steps:

- Data Labeling: The initial step of this phase is to label the dataset images. This procedure involves assigning a specific class label to each image based on the letter sign. All images of a particular class are kept in a separate folder with the class label's name. This technique is

Received 30/October/2022; Accepted 11/February/2023
 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

necessary for training since the filename is regarded as the class label when the dataset is trained for the CNN model.

- Split Dataset: the dataset is split into the training set and testing set. Seventy per cent of the dataset is included in the training set, while the remaining thirty per cent is reserved for the testing phase. Seventy per cent of the dataset is divided into validation sets comprising twenty per cent and training sets containing eighty percent. During the training phase, the validation procedure evaluates the model.
- 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 preprocessing.to apply data normalization used the following equation:
- normalized data= (input data / 255.0) [24]
- Data Augmentation: Data augmentation improves the performance and output of machine learning models by generating new and distinct training dataset samples. To apply data augmentation, use the following methods:

→ Rotating the images within 5 degrees on the left.
→ Rotating the images within 3 degrees on the right.
→ add random noise to the images.
→ Apply random shifts to the images.
→ Apply sigmoid to the images.
→ Apply log to the images.
→ change the brightness of the image with (100/255).
→ change the contrast of the images to 1.5.
→ rescale the images with (0.2,99.8).
→ Apply the gamma to the images.

B- 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 stage, CNN's layers are put together. The CNN model is divided into two stages, the first stage is implemented for feature extraction purposes and the second stage is for classification purposes as show in *Fig. 4*.

This proposed CNN consists of three convolution layers with three max pooling to learn input images, explain in *Fig. 2 and Fig. 3*. The batch size is adjusted as 64 (by experiment), and the ADAM optimization method is used.

The 32 kernels of size 3x3 are convolved in the first convolutional layer. Following the max-pooling layer with filters of size 2x2 A max-pooling layer with a size 2x2 filter is utilized to reduce the size of feature maps. The pooling layer explicitly computes the maximum value in each neighborhood at various points.

As a result of the pooling layer, the network's parameters and computation would considerably decrease, the second convolution layer applies convolutions with 64 kernels of size 3x3, followed by a max-pooling layer with filters of size 2x2. the third convolution layer applies convolutions with 128 kernels of size 3x3, followed by a max-pooling layer with filters of size 2x2.

Two Dropout layers with (0.25 and 0.5), dropout layers are important in training CNNs because they prevent overfitting on the training data.

Two fully connected (FC) layers make up the categorization module. The last FC layer's output has different numbers of neurons depend on the types of datasets are used with SoftMax activation function. The suggested CNN is trained using the back-propagation approach. The error

Received 30/October/2022; Accepted 11/February/2023

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

propagation approach adaptively updates the weights and bias in the convolutional and FC layers. As a result, the classification result is fed back to us.

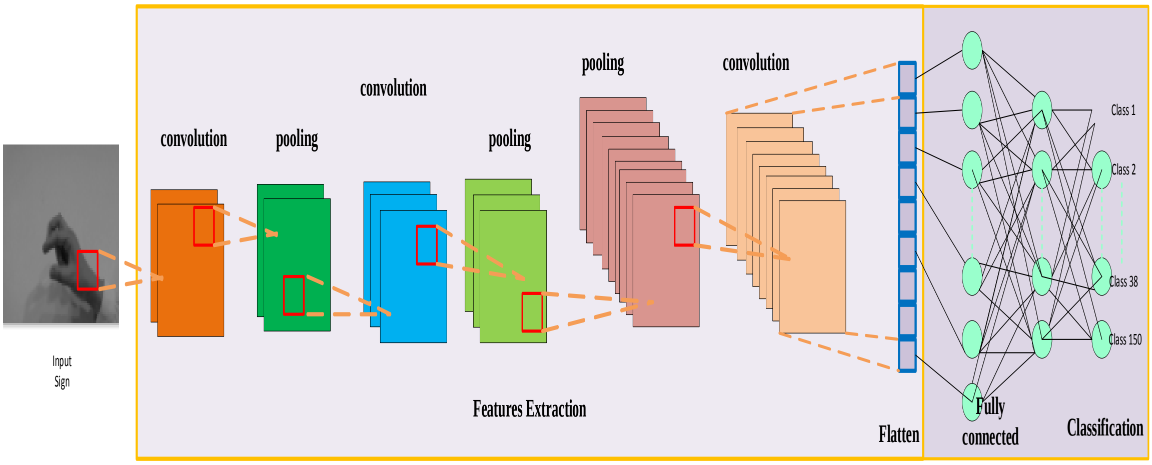

FIG. 4. THE ARCHITECTURE OF THE PROPOSED CNN MODEL.

d) converting the predicted text into speech.

In addition to showing as text on a screen, translating hand gestures can be carried out by pronouncing the motions into speech. The process of converting from text to speech is accomplished through the use of the gtts(Google Text-to-Speech) library[25], which requires the presence of the Internet because the library operates in online mode.

B. Translation from Arabic text to sign language

The input in this option is an Arabic text, and its output is an equivalent sign language translation as a video. The videos and images from an Iraqi sign language dataset are saved in a folder containing files with Iraqi sign language names. The system searches within this folder for the names of photos or movies that match the text entered. As a result, the image or video on the screen is presented.

IV. PERFORMANCE MEASURES

An essential part of any project is to test your machine learning algorithm. Your model may give you good results when using a metric like an accuracy score. The Classification accuracy use to measure how well our model does most of the time. However, this is not enough to truly judge our model. Several metrics depend on the standard measure used in image classification.

The Confusion Matrix visualizes the Actual and Predicted values. It is a table-like structure as shown in *Fig. 5* that indicates the performance of our Machine Learning classification mode [26].

		Actual values	
		1	0
Predicted values	1	TP	FP
	0	FN	TN

FIG. 5. CONFUSION MATRIX OF A BINARY CLASSIFICATION PROBLEM.

Received 30/October/2022; Accepted 11/February/2023

Iraqi Journal of Computers, Communications, Control & Systems Engineering (IJCCCE), Vol. 23, No. 2, June 2023

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

The components of the Confusion Matrix it shows the different ways Actual vs Predicted values can be combined. Let's go over them one by one[26].

a. TP: True Positive (TP): The values that were positive and were predicted to be positive were called True Positive.

b. FP: False Positive (FP): Values that were actually negative but were wrongly thought to be positive by someone else.

c. False Negative (FN): The values were positive but were thought to be negative.

d. True Negative (TN): The negative values that were also predicted to be negative are called True Negative.

Many metrics are based on the Confusion Matrix as following [27]:

1. Accuracy

Accuracy is often the first choice when evaluating an algorithm's performance in classification issues. It is the percentage of correctly classified data to the total number of observations. The accuracy is calculated as follows:

$$\text{Accuracy} = \frac{TP+TN}{(Tp+Tn+FP+FN)} \quad \text{……….(1)}$$

2. Precision

It merely indicates “the amount of relevant data items.” How many of the positive observations predicted by an algorithm are positive. The precision is calculated as follows:

$$\text{Precision} = \frac{TP}{(Tp+FP)} \text{............ (2)}$$

3. The Recall

The fraction of True Positive items is divided by the total positively classified units (row sum of the actual positives). False Negative elements are those that the model labels as negative but are positive. The Recall is calculated as follows:

$$\text{Recall} = \frac{TP}{(Tp+Fn)} \text{..............(3)}$$

4. F1-score

F1-score is the harmonic mean of precision and recall. It combines precision and recall into a single number using the following formula:

5. $$F1 - score = 2x \frac{\text{Precision x Recall}}{\text{Precision+ Recall}} \text{.........(4)}$$

V. RESULTS AND DISCUSSION

The experiment was conducted using Keras libraries, TensorFlow, and anaconda with the Python programming language. a CNN was built and trained to distinguish new samples it had never seen before, which is the goal of this stage.

A. Training Model

To train the proposed system, we used the iraqi sign language dataset , the dataset consist of 100 images for each letter in dataset (28 letters) , after applied the image augmentation the dataset consist of 280000 image. The proposed IrSLT system uses several layers of convolution layers, pooling layers, activation layers, flatten layers, and fully connected layers to complete the CNN model. Finally, in the following parameters of CCN are initialized:

→ set the first dropout ratio to 0.25.

→ set the number of units in the first fully connected layer to 512 units.

Received 30/October/2022; Accepted 11/February/2023
 ISSN (Print) 1811-9212 *ISSN (Online) 2617-3352*

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

→ set the second dropout ratio to 0.5.
→ set the number of units in the second fully connected layer to 28. (the number of alphabetic in IrSL.
→ Set the validation set to 0.2.
→ Set the batch size to 64.
→ Set the epochs= to 50.
→ Set the optimizer with Adam optimizer.

It can be observed from Table I and *Fig. 6* that the training accuracy is less than the validation accuracy. This is due to the number of samples in both sets; the number of samples in the training set is four times more than in the validation set, which will decrease accuracy.

TABLE I. TRAINING AND VALIDATION ACCURACY, TRAINING AND VALIDATION LOSS OF LETTERS DATASET

Training Accuracy	Training Loss	Validation Accuracy	Validation Loss	Testing Accuracy
98.13%	0.0540	99.11%	0.0269	99.3

In *Fig. 6,* the training accuracy begins with a lower value than the validation accuracy. When it reaches an approximation of epoch 10, the training accuracy begins to go the accuracy of the validation and continues to close until it reaches a point of stability near the end of the epochs.

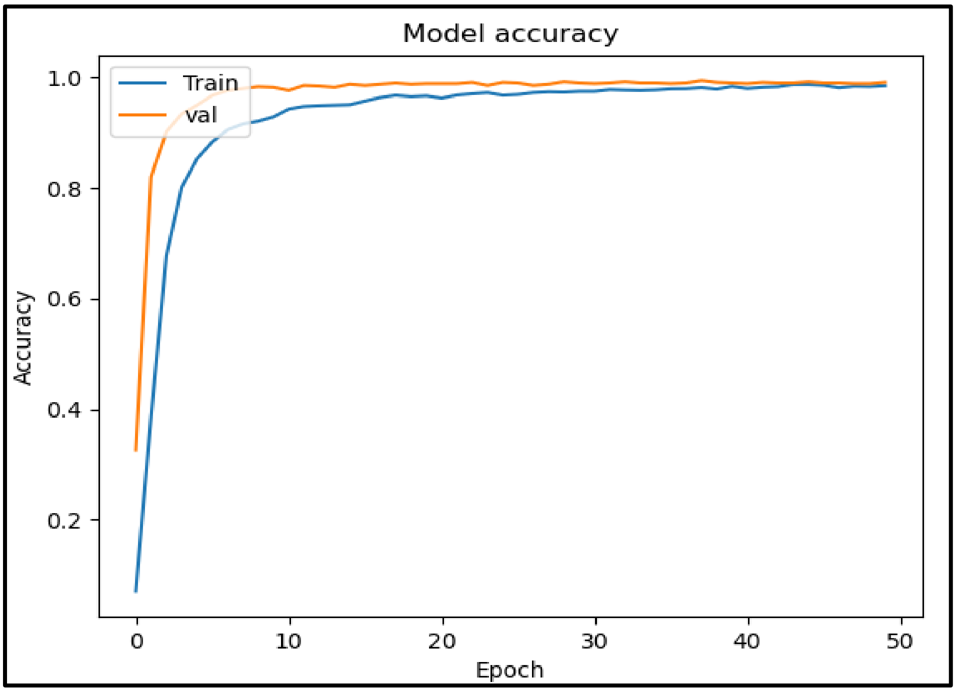

FIG. 6. ACCURACY OF TRAINING LETTERS DATASET.

Received 30/October/2022; Accepted 11/February/2023

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

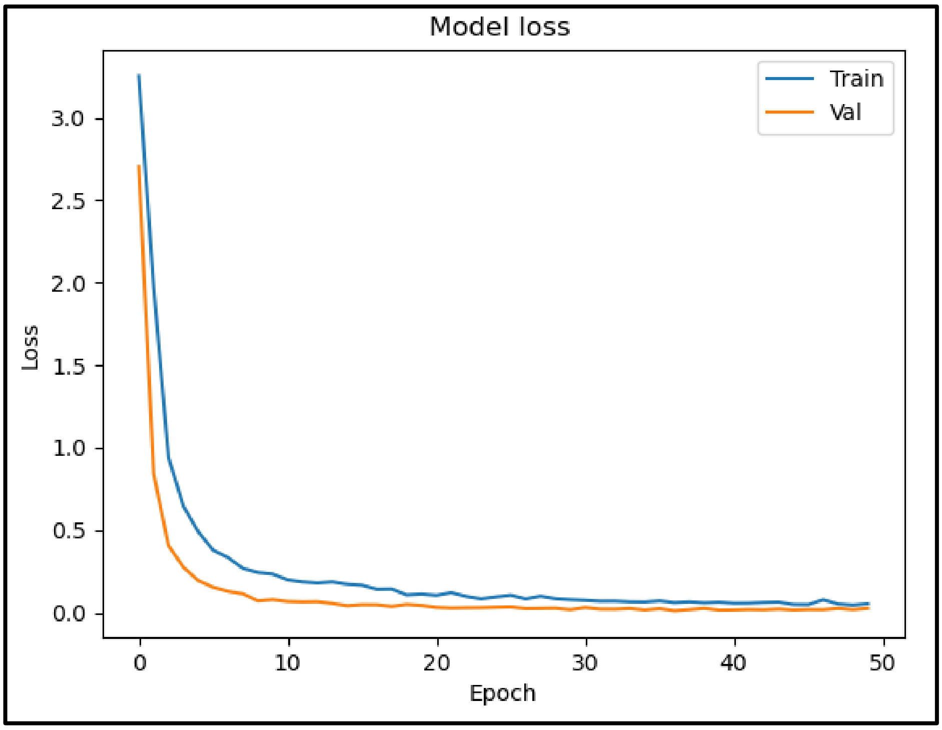

FIG. 7. LOSS OF TRAINING LETTERS DATASET.

The *Fig. 7* shows that the training set began with a high loss, reduced to less than 1.0 between epochs 4 and 6, and then became nearly steady between epochs 20 and 50.

Fig. 8 illustrates the confusion matrix for this dataset's testing phase. Table II displays the accuracy for each letter as well as the overall accuracy. Whereas Table III shows the recall, precision, and F-score measurements, as well as the average rate for each metric.

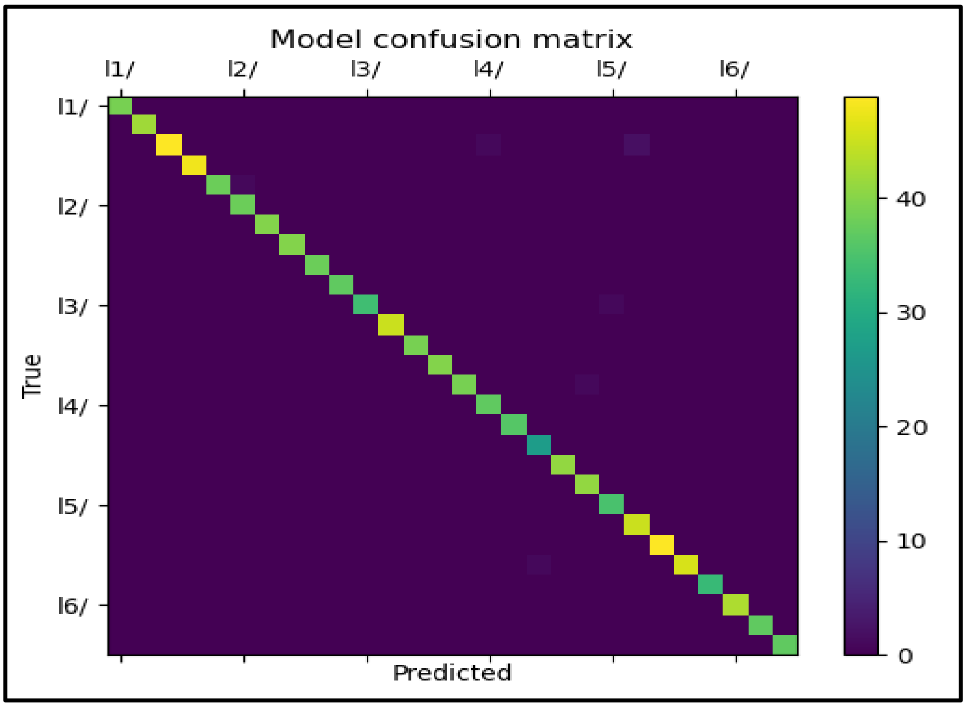

FIG. 8. CONFUSION MATRIX OF TRAINING LETTERS DATASET.

Received 30/October/2022; Accepted 11/February/2023

 ISSN (Print) 1811-9212 ISSN (Online) 2617-3352

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

TABLE II. TESTING ACCURACY OF EACH LETTER IN DATASET

Letters	Accuracy	Letters	Accuracy	Letters	Accuracy	Letters	Accuracy
الاف	99.20	دال	100	ضاد	99.32	كاف	99.20
باء	99.38	ذال	99.50	طاء	98,3	لام	99.54
تاء	99.55	راء	98.49	ظاء	99.10	ميم	98.71
ثاء	99.54	زاي	98.53	عين	99.37	نون	99.30
جيم	98.21	سين	99.29	غين	100	هاء	99.31
حاء	98.42	شين	99.97	فاء	98.43	واو	98.10
خاء	99.11	صاد	100	قاف	99.87	ياء	100

TABLE III. PRECISION, RECALL AND F1-SCORE MEASUREMENTS RESULTS OF LETTERS DATASET

Letters	Precision	recall	f1-score
الاف، باء،ثاء،خاء، دال،ذال،راء،سين،شين ،صاد،ظاء،غين،لام، نون،هاء،واو،ياء	1.00	1.00	1.00
تاء	1.00	0.94	0.97
جيم،ضاد،زاي	1.00	0.97	0.99
حاء،طاء،قاف	0.97	1.00	0.99
عين،كاف	0.96	1.00	0.98
فاء	0.98	1.00	0.99
ميم	1.00	0.98	0.99
Average	0.99	0.99	0.99

B. Comparison with alternative Methods

As mentioned in this paper, each country has its unique sign language because each has its own culture and traditions. So, Iraq's sign language is different from other Arab countries' sign languages to evaluate the proposed system against previous studies. The comparisons focused on the alphabet and the CNN model.

TABLE IV. COMPARISON OF THE RESULTS OF THE PROPOSED SYSTEM AND OTHER PREVIOUS STUDIES

Paper number	Input device	Methods used	Dataset used	Recognition rate (%)
[11]	camera	Correlation coefficient	20 signs	85.67
[12]	LMC	NB MLP	Arabic Letters	99 98
[13]	Microsoft Kinect	HMM	40 words	90
[14]	camera	ANN	3 Arabic letters	73
[15]	Microsoft Kinect	—	30 words	97.85
[17]	Camera	CNN	ArSL2018	97.6
[18]	Camera	CNN	ArSL2018	97.2
[19]	Camera	CNN	ArSL2018	90
Theproposedsystem	Camera	CNN	ArSL2018	99.9
The proposed system	camera	CNN	The proposed dataset	99.3

Received 30/October/2022; Accepted 11/February/2023

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

In Table IV, observes and compares related works that used the camera as an input device [11,14] with different machine learining techniques, the proposed system; it should be noted that the accuracy of the suggested approach is significantly greater. This is due to the employment of appropriate data processing techniques in the proposed system.

This is also true for systems [12,13,15], those systems used sensors devices as input devices, which get high accuracy because it produces different features such as numbers like angle and distance, but still, the proposed method is also superior since deep learning is used in the proposed system for feature extraction and classification.

CNN served as the classification model with the Arabic sign Language stander dataset (ArSL2018) in the systems [18,19], those systems have been approximated to the accuracy of the proposed system, but the proposed system accuracy is higher because better processing techniques are used,

The proposed system is implemented using the Arabic sign language standard dataset (ArSL2018). Also, the proposed system's accuracy was higher than all the previous studies that work the same dataset and used deep learning.

VI. CONCLUSIONS

In this paper, we present an application for machine translation of Iraqi Sign Language to facilitate communication between deaf /mute people the ordinary people. The system works on translating the Iraqi sign language into Arabic text and translating from Arabic text into Iraqi sign language This method eliminates the need for the signer to wear any equipment while executing the signs. A database was built from scratch based on the Iraqi sign language dictionary. The images were captured using a webcam. The database contains 280000 images in gray color representing 28 class, and the image dimensions are 1820 * 720. 2.The process of capturing focused on the hand area to reduce the preprocessing time and get good results in classification.Through the presentation of the previous works, it was found that the use of deep learning is better than other methods of machine learning, not only through high accuracy but also speed, as when using deep learning, it eliminates the stage of feature extraction.The proposed system uses CNN to train and classify images. The proposed system achieved an accuracy rate of 99.3 using the proposed Iraqi sign language database, 99.9 using the standard Arabic sign language database. 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.

REFERENCES

[1] I Papastratis, C. Chatzikonstantinou, D. Konstantinidis, K. Dimitropoulos, and P. Daras," Artificial Intelligence Technologies for Sign Language". Sensors (Basel). 2021 Aug 30;21(17):5843. doi: 10.3390/s21175843. PMID: 34502733; PMCID: PMC8434597.

[2] M. Mukushev, A. Sabyrov, A. Imashev, K. Koishybay, V. Kimmelman, and A. Sandygulova, "Evaluation of manual and non-manual components for sign language recognition," *Lr. 2020 - 12th Int. Conf. Lang. Resour. Eval. Conf. Proc.*, no. May, pp. 6073–6078, 2020.

[3] U. Nations, "International Day of Sign Languages | United Nations," United Nations, Sep. 2021. Accessed: Aug. 31, 2022. [Online]. Available: https://www.un.org/en/observances/sign-languages-day.

[4] R. Thakare, P. K. Pathan, M. Lokhande, and N. Waje, "Dynamic and Static Gesture Recognition System Using Moments," *Int. J. Adv. Eng. Res. Sci.*, vol. 4, no. 2, pp. 69–72, 2017, doi: 10.22161/ijaers.4.2.14.

[5] S. Fakhfakh and Y. Ben Jemaa, "Gesture recognition system for isolated word sign language based on key-point trajectory matrix," *Comput. y Sist.*, vol. 22, no. 4, pp. 1415–1430, 2018, doi: 10.13053/CyS-22-

Received 30/October/2022; Accepted 11/February/2023

DOI: https://doi.org/10.33103/uot.ijccce.23.2.9

4-3046.

[6] M. Maraqa, F. Al-zboun, M. Dhyabat, and R. A. Zitar, "Recognition of Arabic Sign Language (ArSL) Using Recurrent Neural Networks," vol. 2012, no. February, pp. 41–52, 2012.

[7] D. Mart, "MSc Thesis VIBOT Sign Language Translator using Microsoft Kinect XBOX 360TM", Department of Electrical Engineering and Computer Science Computer Vision Lab University of Tennessee (Knoxville- USA), 2012.

[8] A. Z. Shukor, M. F. Miskon, M. H. Jamaluddin, F. Bin Ali Ibrahim, M. F. Asyraf, and M. B. Bin Bahar, "A New Data Glove Approach for Malaysian Sign Language Detection," *Procedia Comput. Sci.*, vol. 76, no. Iris, pp. 60–67, 2015, doi: 10.1016/j.procs.2015.12.276.

[9] L. E. Potter, J. Araullo, and L. Carter, "The leap motion controller: A view on sign language," *Proc. 25th Aust. Comput. Interact. Conf. Augment. Appl. Innov. Collab. OzCHI 2013*, pp. 175–178, 2013, doi: 10.1145/2541016.2541072.

[10] R. M. Mohammed and S. M. Kadhem, "A Review on Arabic Sign Language Translator Systems," *J. Phys. Conf. Ser.*, vol. 1818, no. 1, 2021, doi: 10.1088/1742-6596/1818/1/012033.

[11] A. A.A, A. Elsayed, and H. Hamdy, "Arabic Sign Language (ArSL) Recognition System Using HMM," *Int. J. Adv. Comput. Sci. Appl.*, vol. 2, no. 11, 2011, doi: 10.14569/ijacsa.2011.021108.

[12] M. Mohandes, S. Aliyu, and M. Deriche, "Arabic sign language recognition using the leap motion controller," *IEEE Int. Symp. Ind. Electron.*, no. November 2015, pp. 960–965, 2014, doi: 10.1109/ISIE.2014.6864742.

[13] O. Amin, H. Said, A. Samy, and H. K. Mohammed, "HMM based automatic Arabic sign language translator using Kinect," *Proc. - 2015 10th Int. Conf. Comput. Eng. Syst. ICCES 2015*, no. December, pp. 389–392, 2016, doi: 10.1109/ICCES.2015.7393081.

[14] L. ELSIDDIG and M. MOHAMED, "Translating Arabic Sign Language ARSL To Text Using Artificial Neural Networks," no. September, pp. 43–47, 2017, doi: 10.15224/978-1-63248-131-3-17.

[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, doi: 10.19026/rjaset.15.5292.

[16] M. Y. Ftichi, A. Benabbou, and K. Abbad, "Arabic Sign Language Alphabet Recognition Methods Comparison, Combination and Implementation," *Adv. Intell. Syst. Comput.*, vol. 756, no. 4, pp. 518–524, 2019, doi: 10.1007/978-3-319-91337-7_46.

[17] G. Latif, N. Mohammad, R. AlKhalaf, R. AlKhalaf, J. Alghazo, and M. A. Khan, "An automatic arabic sign language recognition system based on deep CNN: An assistive system for the deaf and hard of hearing," *Int. J. Comput. Digit. Syst.*, vol. 90, no. 4, pp. 715–724, 2020, doi: 10.12785/ijcds/090418.

[18] A. A. Alani and G. Cosma, "ArSL-CNN: A convolutional neural network for arabic sign language gesture recognition," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 22, no. 2, pp. 488–499, 2020, doi: 10.11591/ijeecs.v22.i2.pp488-499.

[19] R.EL.Rwelli;O.R. Shahin;A.I.Taloba, "Gesture basedArabic Sign Language Recognition for Impaired People bace on convolution neural network." IJACSA, 2021.

[20] C. Series, "Sign Language Recognition Using Convolutional Neural Network," 2021, doi: 10.1088/1742-6596/1916/1/012091.

[21] S. E. E. Profile, "Sign Language Recognition Using Deep Learning and Computer Vision," no. June, 2020, doi: 10.5373/JARDCS/V12SP5/20201842.

[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.

[23] T. M. of L. and S. Affairs, *Iraqi Sign Language Dictionary*. The Ministry of Labour and Social Affairs and United Nations Children's, 2000.

[24] P. Jamal, M. Ali, R. H. Faraj, P. J. M. Ali, and R. H. Faraj, "1-6 Data Normalization and Standardization: A Technical Report," *Mach. Learn. Tech. Reports*, vol. 1, no. 1, pp. 1–6, 2014, [Online].Available: https://docs.google.com/document/d/1x0A1nUz1WWtMCZb5oVzF0SVMY7a_58KQulqQVT8LaVA/edit#.

[25] P. N. D., " gTTS documentation" 2018,. https://gtts.readthedocs.io/en/v2.0.0/.

[26] K. M. Ting, "Confusion Matrix," Encycl. Mach. Learn. Data Min., No. October, Pp. 260–260, 2017, Doi: 10.1007/978-1-4899-7687-1_50.

[27] M. Grandini, E. Bagli, And G. Visani, "Metrics For Multi-Class Classification: An Overview," Pp. 1–17, 2020, [Online]. Available: Http://Arxiv.Org/Abs/2008.05756.

Received 30/October/2022; Accepted 11/February/2023
