

Sign Language Semantic Translation System using Ontology and Deep Learning

Eman K. Elsayed¹, Doaa R. Fathy²
Mathematical and computer science
Al-Azhar University
Cairo, Egypt

Abstract—Translation and understanding sign language may be difficult for some. Therefore, this paper proposes a solution to this problem by providing an Arabic sign language translation system using ontology and deep learning techniques. That is to interpret user's signs to different meanings. This paper implemented ontology on the sign language domain to solve some sign language challenges. In this first version, simple static signs composed of Arabic alphabets and some Arabic words started to translate. Deep Convolution Neural Network (CNN) architecture was trained and tested on a pre-made Arabic sign language dataset and on a dataset collected in this paper to obtain better accuracy in recognition. Experimental results show that according to the pre-made Arabic sign language dataset the classification accuracy of the training set (80% of the dataset) was 98.06% and recognition accuracy of the testing set (20% of the dataset) was 88.87%. According to the collected dataset, the classification accuracy of the training set was 98.6% and Semantic recognition accuracy of the testing set was 94.31%.

Keywords—Deep Learning (DL); ontology; sign language translation

I. INTRODUCTION

Sign language considered the only way for communication between deaf, hearing-impaired and normal people. According to the World Health Organization (WHO), there are about 466 million people around the world have disabling hearing loss, and more than 28 million of these are Americans, 13 million people within Egypt across all age groups [1]. People use sign language gestures as a means of non-verbal speech to express their thoughts and emotions.

Sign language has two types of gestures: static gestures and dynamic gestures [2]. Due to Arabic language variations (delicates), Arabic Sign Language has multiple country variants and dialects. It differs from one country to another and sometimes differs within the same country. Deaf and hearing-impaired people have all the right to speak to each other even for different sign languages. Sign languages need an intelligent system to translate sign language to another based on natural languages. It is hard for most people who are not interested in sign language to communicate without an interpreter. Automatic sign language translation systems are important for solving these problems.

Sign Language Recognition Systems can be developed by using either sensor-based technique or image-based technique [3]. Many types of research focused on sign language translation to text and spoken language and vice versa. But

this is not easy to be done by machines since it depends on the natural language processing and image recognition. On the other hand, the traditional way of translation needs a translator that specializes in sign language may not be present in all situations, especially when the deaf continues with the person hearer. Translation of Arabic sign language (ArSL) is facing many challenges; the lack of linguistic studies on ArSL, sign language is assumed to be a universal language.

Deep Learning [4] algorithms have been successfully applied to image recognition problems. Deep learning involves neural networks with more than one hidden layer, has been used successfully in face recognition, speech recognition and natural language processing problems [5, 6]. Deep learning successfully implemented for human gesture recognition in recent years.

Ontology is a formal representation of knowledge; it handles a set of concepts in a specific domain, their relationships, and properties [7]. WordNet is a linguistic resource containing not only words of the targeted language but also synsets and semantic conceptual relations between them. These relations provide semantic information about concepts and their original words. These concepts and their relations are exploited to improve Information Retrieval, Text Classification, and Text Summarization Technically. Arabic WordNet lexicon provides a good semantic structure for computing the semantic similarity between words.

The contribution of this paper is: To enhance sign language translation using the power of semantic web technologies (i.e. ontologies) and deep learning. Ontology used in building multi-sign language ontology (MSLO) and deep CNNs used in Arabic sign language feature extraction and recognition processes. As a case study, the proposed method applied in Arabic sign language semantic translation system.

The rest of this paper is organized as follows; the next section presents some related works. Then the proposed method will present in Section III. Subsection A in Section III presents the proposed Multi Sign Language Ontology (MSLO) and subsection B presents the details of the proposed "Semantic Deep Learning". Section IV presents the application of the proposed method on a case study in Arabic Sign Language, while Section V presents the analysis of results. Finally, conclusion and future works.

II. RELATED WORKS

In recent years, several research projects in Arabic static sign language translation systems have been developed. Author in [8] presented an automatic visual technique that translates individual Arabic sign to text word. Geometric features of the hands are employed to formulate the feature vector. Euclidean distance classifier is applied for the classification stage. A dataset of 30 isolated words of the hearing-impaired children was developed. The system has a recognition rate of 97%.

The SVM method and HOG descriptors were used to recognize Arabic Sign Language alphabets [9]. This system extracts the HOG descriptor and transfer features to One Versus All soft-margin (SVM). The proposed system has reached an accuracy of 63.5% for Arab Alphabet signs

Hidden Markov Model (HMM) technique successfully recognizes the alphabet and numbers signs of Arabic sign language to text or speech [10]. The proposed algorithm uses hand geometry and the different shape of a hand in each sign for classifying letters shape by using HMM.

Also, there exist many reported research systems related to Arabic sign language recognition based on deep learning. Authors in [11] proposed a new system based on the CNNs, fed with a real dataset, this system recognized automatically numbers and letters of Arabic sign language. To validate the system, they done a comparative study that shows the effectiveness and robustness of their proposed method compared to traditional approaches based on k-nearest neighbors (KNN) and support vector machines (SVM).

ElBadawy et al. [12] developed the system for Arabic sign language recognition. In this system, CNN was used to recognize 25 gestures from the Arabic sign language dictionary. The system achieved 98% accuracy for observed data and 85% average accuracy for new data.

Although a variety of methods have been proposed in recent years to recognize static hand gestures, most of them worked only on translating text to sign or sign to text. They did not take care of the semantics of the translated text. In another side, some researches work on sign language translation using Ontology. Author in [13] presents a proposed system for semantically translating Arabic text to Arabic Sign-Writing in the jurisprudence of prayer domain using ontology as a semantic web technology. The system designed to translate Arabic text by applying Arabic Sign Language (ArSL) grammatical rules as well as semantically looking up the words in Ontology. Author in [14] focuses on the development of information technology for Ukrainian Sign Language translation based on grammatically augmented ontologies.

III. THE PROPOSED ONTOLOGY AND DEEP LEARNING METHOD

The proposed method is divided into two main parts: building ontology-based and combining ontology with deep learning.

A. Multi Sign Language Ontology (MSLO)

The proposed Multi Sign language Ontology (MSLO) is linguistic Ontology takes the advantages of WordNet and multilingual WordNet on sign language domain. But analysis of the nature of sign language leads us to create some different relations to solve some of sign language challenges.

MSLO extends the WordNet relations to deal with sign language. The sign language synsets are created in correspondence with the WordNet synsets, and semantic relations are imported from the corresponding synsets. It assumed that if there are two synsets and relations are holding between them, the same relations hold between the corresponding synsets in sign language. Generally, sign express word meaning (concept), it could be said that it is important to sign a concept rather than a word. MSLO solves some sign language challenges. The following three examples of these challenges:

1) Each sign means a bag of words in each language (synonyms). For example, Fig. 1 presents the sign of "car" in English language. The same sign could represent the words: Vehicle, automobile, transportation, and auto. The same sign also used to represent the words "سيارة", "مركبة", "حافلة", "عربية" in the Arabic language.

2) The same meaning has different signs and vice versa. In MSLO, the two words related semantically to each other may have different signs. For example, the word "bad" in English has the sign (a) in Fig 2. The word "سيئ" in Arabic has the sign (b) in Fig. 2.

Also, one sign may use to represent different words such as; sign in Fig. 3, represents the word "لا", the Character "ب" and number "1" in the Arabic language. This challenge solved in MSLO ontology, where sign gesture interprets based on its class and domain.



Fig. 1. The Word "CAR" in Sign Language.



Fig. 2. (a) "سيئ" Gesture Sign (b) Bad Gesture Sign.



Fig. 3. Gesture Sign of Word "لا", Character "ب" and Number "1".

3) Each person has his-self way to sign any word in any language. MSLO takes into consideration personality signs.

Relations were created in MSLO to solve these sign language challenges. Properties used to describe a class/concept or individuals are:

a) Object properties used for the relationship between words:

- Arabic_meaning
- English_meaning
- French_meaning
- Has_the_same_sign; to relate the concepts which have the same gestures signs.

b) Data property that used to describe words are:

- DL-Label; to assign the word to its label in the deep learning model.
- Arabic_sign
- English_sign
- French_sign

The URL address of each sign gesture image is saved as a data property value.

c) Annotation property (label) also used to add all possible meanings of the words.

Deaf, hearing impaired and normal users can interact with the ontology via interaction system to develop and update his personal information (My-Dictionary class) and its corresponding signs.

Fig. 4 (screenshot from protégé) shows Part of MSLO Ontology. The main concepts are: English_WORDNET,

French_WORDNET, and Arabic_WORDNET. Each language consists of sub-classes. Each subclass is a bag of words in one domain. There is a class for personal information in sign language to represent personal data; like name, address and birth date. Also, Fig. 4 shows an example of two English words (Good and Bad) with their meanings in Arabic and French languages are related to each other in antonym relation. Their sign language gestures' are also related in antonym relation. The English word Good with its meanings in Arabic and French languages are related to each other in synonym relation. Their sign language gestures' are also related in synonym relations.

B. Combining Ontology with Deep Learning

The link between ontology and deep learning is being explored. Author in [15] addresses the extraction of OWL information from raw text with deep learning. Author in [16] applies deep learning to Ontology extraction, obtaining encouraging results.

Deep learning output needed to be enriched by semantic, so the proposed method considered merging ontology with deep learning. Using ontology as a semantic layer in deep learning produces Semantic deep learning (SDL).

Convolutional Neural Network (CNN) [17] is the most popular algorithm used for implementing the deep learning technique. A CNN is composed of several layers: Convolutional Layer, Pooling Layer, and Fully-Connected Layer. In this method, Ontology layer is inserted to add semantic information to deep CNN classification output.

Fig. 5 shows a diagram of CNN with 4 hidden layers. The input layer has three neurons and each hidden layer has four neurons. Hidden layers followed by the output layer, the semantic layer, and the semantic output layer.

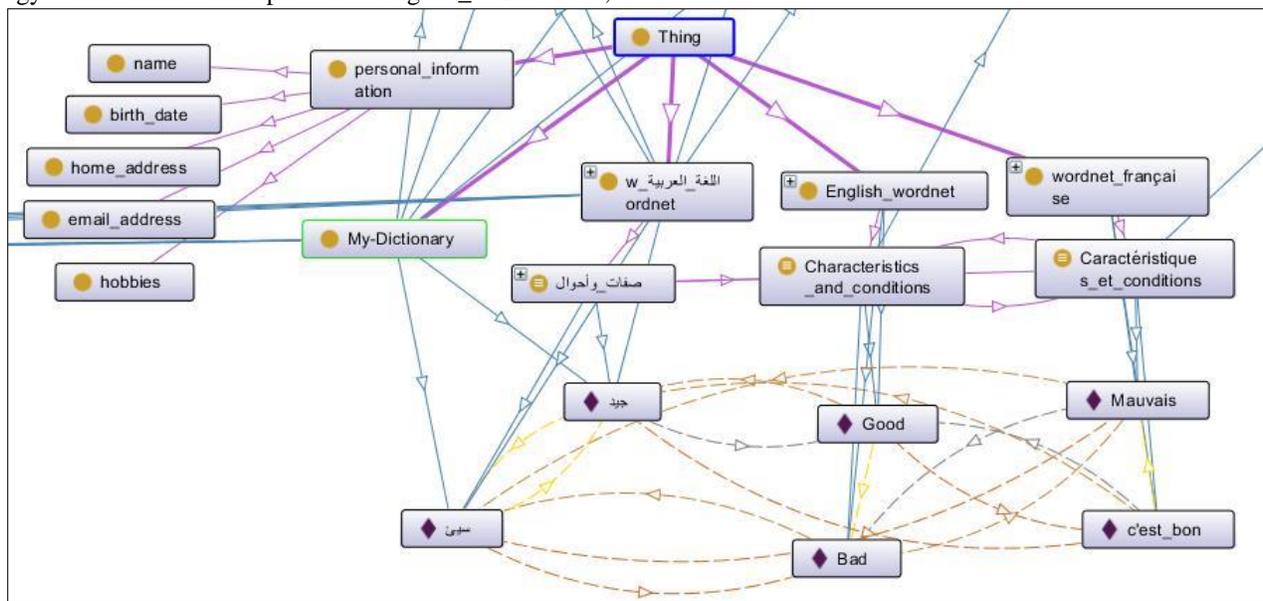


Fig. 4. Part of Proposed (MSLO).

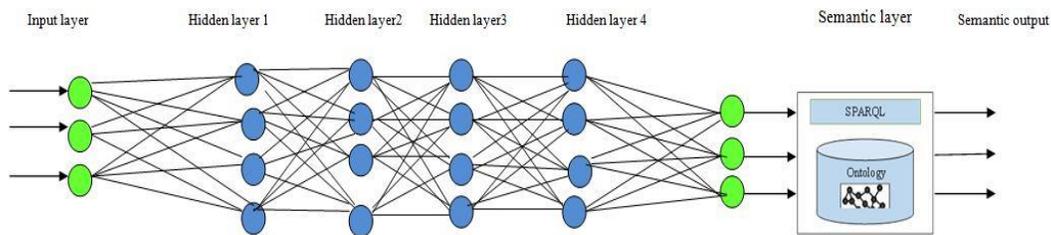


Fig. 5. Semantic Deep Learning Layers (SDL).

Semantic Deep Learning (SDL) layers are as the following:

- Convolutional layer composed of groups of neurons act as kernels. These kernels are associated with a small region of the image known as a receptive field, then convolving them (applying Convolution operation) with a specific set of weights. Where multiply each element of the filter (weights) with corresponding receptive field elements. Different sets of features extracted by sliding convolutional kernel on the image with the same set of weights. Convolution operation may categorize into different types based on the type and size of filters.
- Activation layer also called a non-linear layer; the convolution layer output is generally followed by a non-linear function called activation function to introduce non-linearity into the model. Thus, features generated by the convolution layer are transformed into another space by the activation function, and the data can be better classified. In this paper, the activation functions used are (Rectified Linear Units) ReLU.
- Batch normalization (BN) layer, is the most widely used to enhance generalization [18] and avoid over. Also, is added after a convolutional layer and its ReLU activations. Batch Normalization achieves the same accuracy with fewer training steps thus accelerating the training process.
- Dropout layer also used to avoid overfitting. The key idea of Dropout is to randomly drop units (along with their connections) during the training phase. The reduction in the number of parameters in each step of training has an effect on regularization.
- Pooling layer is used on one hand to reduce the spatial dimensions of the representation and to reduce the amount of computation done in the network. The most used pooling layer has filters of size 2 x 2 with a stride 2. This reduces the thoughts to a quarter of its original size. In this CNN max pooling is applied, which retains the maximum value within the local neighborhood of the sliding window.
- Fully connected layer, interpret these feature representations and perform high-level reasoning. Each neuron from a fully connected layer is linked to each output of the previous layer. The operations behind a convolutional layer are the same as in a fully connected layer.

- Loss layer is used to penalize the network for deviating from the expected output. This is normally the last layer of the network. Various loss function exists: softmax is used for predicting a class from several different classes, where sigmoid cross-entropy is used to predict multiple independent probabilities (from the [0, 1] interval).
- Semantic layer maps the class label to a formal representation of its meaning. When a set of words share the same fundamental meaning, they will have one label. In this layer the result of the previous layer "DL class label" used in searching Ontology to get all semantic meanings. The output of deep learning (class label) named "DL_label" represents as data property value in MSLO ontology. So each class label has its corresponding data property value to retrieve all semantic data of this class.

In this research, it was started by an image augmentation technique to adapt the training set with such images. Data augmentation has a large impact on improving the DL accuracy by reducing overfitting and increasing generalization capacity of the network [19]. Image data-augmentation increases the size of the training dataset by applying small geometric and photometric transformations.

As shown in Fig. 6, the CNN model of this paper consists of 5 convolutional layers each one followed by a max-pool layer with a filter of shape 2 x 2 with stride 2. Batch normalization and dropout layers are employed in all convolutional layers. Fully-connected (FC) layer follows the convolutional layers. It has 1024 channels where is a soft-max layer to perform the classification over the number of classes. Dropout regularization and batch normalization are used also in the fully-connected layer. The dropout rates of 0.25 and 0.5 were set for the convolutional-pooling layers and the fully-connected layer, respectively.

C. Implementation

Proposed SDL method applied by Python using Keras. Tensorflow was chosen as a backend because it is mainly used in classification. Tensorflow is an open-source software library for deep learning created and maintained by Google. Anaconda used because of its flexibility to work with any operating system. Also, it has Spyder editor, a powerful python editor that has all the functionalities needed to write and run the codes. It also has the option to create an environment to install the packages. Also, Python OpenCV library used for image processing, and view the results in the 'IPYTHON' console. The MSLO implemented using Protégé as Ontology editor.

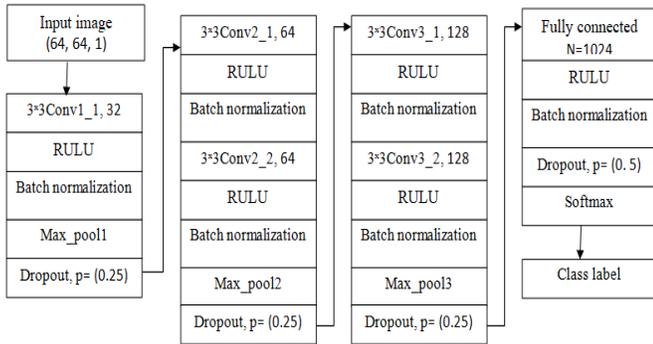


Fig. 6. Architecture of Deep Learning layers in SDL.

IV. THE CASE STUDY ON ARABIC SIGN LANGUAGE AND ITS SEMANTIC TRANSLATION SYSTEM

A. Arabic Signs Dataset

In this paper, samples of 10 Arabic words in Arabic sign language, their Arabic concepts, their corresponding's English and French signs collected. Images are real and were picked using a mobile camera. Dataset images were not captured in a controlled setting; it was especially prone to differences in light, skin color, and other differences in the environment. A snapshot of a few images from these samples is given in Fig. 7.



Fig. 7. Snapshot of the Dataset.

The proposed system applied also to the pre-made ArSL2018 dataset in [20]. The Arabic Sign Language dataset (ArSL2018) is compiled of 54,049 images for the 32 Arabic sign language sign and alphabets collected from 40 participants in different age groups. The dataset contains grayscale 64 * 64 dimensions.

About 80% of the dataset used in training and 20% used in testing the system.

B. Pre-Process Phase

The pre-processing phase consists of many main operations, which applied to the Arabic singe language images after it is taken from the camera. That is to enhance the image. These operations applied to the training set, testing set and on any new images to be recognized. The pre-process operations are as the following:

- Cropping: cropped out the hand object and discarded all the other unnecessary elements in the image.
- Resizing: The images were in different dimensions; therefore, each image resized to a common resolution of 64 × 64 pixels.

- Converting: All the images were converted into grayscale before feeding into the model.
- Also, the collected dataset images smoothed by convolving with a Gaussian filter and median blur to reduce noise in the images.

C. CNN Training

The input consists of standard RGB images of similar size 64 x 64 pixels. Then the dataset was augmented where the images are randomly rotated 0 to 30 degrees. Images were randomly sheared in a range of 0.2-degrees, width and height shift in a range of 0.1 degrees, zoomed in a range of 0.2 degrees and images were horizontally flipped. 200 sign images of Arabic gestures were used in the training phase. The classification accuracy of the collected training set was 98.6 %.

Also, to evaluate CNN model performances, (42960 images) about 80% of (ArSL2018) sign images of Arabic gestures were used to measure the training rate of the CNN architecture. The classification accuracy of this training set was 98.06%.

D. Testing Phase

The input of the system is the image of the sign language static gesture that needs to be recognized. The saved network parameter is loaded to recognize this sign and produced its corresponding label. Then the ontology layer used to produce all available meanings of this input sign. The method used for all test cases and tabulated the obtained accuracies in Section V.

Also, the input could be a text (Arabic, English or French word) which users type to select all respective signs using ontology-based.

E. Connection Between DL and MSLO Ontology for Semantic Translation

After loading CNN and recognize image gesture the output used as input of the semantic layer. SPARQL is used to predict semantically the corresponding meaning and the corresponding sign gesture in another language. SPARQL is used to access the RDFs inside the ontology by using a SELECT statement as a query. SPARQL query has a standard syntax and depends on using variables that contain the predicate, subject, and objects for RDF.

Also, the input of the system could be a text (Arabic, English or French word) which users type to select their respective signs. The process of retrieving and displaying relevant images is based on the user's queries. SPARQL query used in searching ontology to get meaning that represents this word and gets the URL sign image of this meaning.

Some of SPARQL queries used in searching and retrieving data from MSLO:

- SPARQL query to translate Arabic meaning to its corresponding Arabic sign language.

```
SELECT ?arb ?o
WHERE { ?arb rdfs:label ?x.
?arb sign:arabic_sign ?o
```

```
FILTER (regex(str(?x), """"+arabictext+""""))
bind( strafter(str(?arb),str(sign:)) as ?arb )
```

- SPARQL query to translate Arabic sign to English sign.

```
SELECT ?o ?s ?z
WHERE { ?o sign:label ?x.
?o sign:english_meaning ?s.
?s sign:english_sign ?z
FILTER (regex(str(?o), """"+label+""""))
bind( strafter(str(?s),str(sign:)) as ?s ) }
```

V. ANALYSIS OF RESULTS

The proposed DL model was tested on images that had not seen before. Data (11089 images) about 20% of the pre-made dataset used in testing. The recognition rate is defined by the ratio of the numbers of correct sign recognition to the total number of testing signs. Recognition accuracy was 88.87%. Table I shows the performance of trained network on testing data.

Also, Table I shows Arabic letter name, letter name in English script, the number of trained and tested samples, correctly recognized, and misrecognized samples for each Arabic letter.

- The total number of training samples is 42960.
- The total number of testing samples is 11089.
- The total number of correct recognized signs is 9855 about 88.87% of testing samples.
- The total number of incorrectly recognized signs is 1234 about 11.13% of testing samples.

Also, system performance is evaluated by the rate of successful semantic sign language recognition of new signs such that all possible meanings are listed.

The semantic translation rate in percentages in the case of collected 10 Arabic signs is showed in Table II. Without using Ontology, it might be difficult to find meaningful connections among synonyms signs, words in one language and words in different languages.

Table II shows the number of training and testing samples for each word, incorrectly recognized samples of each sign gestures and results with traditional deep learning model and when using SDL.

Also, Table II shows the number of correctly recognized and misrecognized samples for each Arabic word and the corresponding accuracy. The results also show that while most of the words signs are recognized with 100% accuracy, it can be seen that the system recognition accuracy becomes less when gives the input signs for the words "شكرا" and "سيء".

Overall, out of the 88 Arabic sign images used for testing 83 images were recognized to correct meanings and the remaining 5 images were misrecognized resulting. The recognition accuracy was checked as per the correctness of all testing gestures made was 94.31 %.

The results obtained when recognize sign gesture using deep learning model were only the label that represents the word. But when recognize using deep learning and Ontology, the results were all available meanings of this sign gesture.

TABLE. I. TRAINING AND TESTING PREMADE DATASET SAMPLES

Letter Name in English Script	Letter Name in Arabic Script	Training Samples	Testing Samples	Correct Recognized	Incorrect Recognized	Letter Name in English Script	Letter Name in Arabic Script	Training Samples	Testing Samples	Correct Recognized	Incorrect Recognized
Alif	أ (الف)	1035	637	631	6	Za	ظ (طاء)	1378	345	345	0
B a	ب (باء)	1432	359	259	100	Ain	ع (عين)	1691	423	357	66
Ta	ت (تاء)	1471	367	332	35	Ghin	غ (غين)	1582	395	389	6
Thaa	ث (ثاء)	1413	353	294	59	Fa	ف (فاء)	1564	391	261	130
Jeem	ج (جيم)	1242	310	295	15	Qaf	ق (قاف)	1364	341	223	118
Haa	ح (حاء)	1211	315	311	4	Kaaf	ك (كاف)	1419	355	232	123
Khaa	خ (خاء)	1268	339	290	49	Laam	ل (لام)	1464	368	340	28
Dal	د (دال)	1308	326	319	7	Meem	م (ميم)	1412	353	325	28
Zal	ذ (ذال)	1265	317	313	4	Nun	ن (نون)	1506	313	213	100
Ra	ر (راء)	1328	331	317	14	Ha	ه (هاء)	1274	318	244	74
Zay	ز (زاي)	1100	274	269	5	Waw	و (الواو)	1096	275	275	0
Seen	س (سين)	1311	327	325	2	Ya	ي (ياء)	1377	345	344	1
Sheen	ش (شين)	1205	302	300	2	Taa	ة (تاء)	1433	358	303	55
Sad	ص (صاد)	1516	379	324	55	Al	ال (ال)	1075	268	268	0
Dhad	ض (ضاد)	1336	334	236	98	Laa	لا (لا)	1397	349	299	50
T a	ط (طاء)	1452	364	364	0	Yaa	يا (ياء)	1035	258	258	0
		20893	5634	5179	455			22067	5455	4676	779

TABLE. II. NUMBER OF SAMPLES, RECOGNITION RATE (%) AND RESULTS PER WORD

Arabic Words	Label	Training Samples	Testing Samples	Incorrect Recognized	Result with Traditional DL	Results of Proposed Method	
						Synonyms	Words in Same sign
سئ	Bad	16	5	1	Bad	الرديء, السوء, السيء, الطالح	
سيارة	Car	20	8	0	Car	سيارة, سيارة, عربة, حافلة, مركبة	
جيد	Good	30	22	0	Good	مُتَقَنٌ, جيد, حسن	حرف الألف
مرحبا	Hello	19	7	0	Hello	مَرْحَبًا, مرحبا	الرقم خمسة
أحبك	I_love_you	17	9	0	I_love_you	أحبك, احبك	
الحب	Love	22	7	0	Love	حُبٌ, وَجْدٌ, قبول	الاعتذار
لا	No	20	9	0	No	لا, نفي, اعتراض	حرف الباء, الرقم واحد
صورة	Photo	29	6	0	Photo	تَمَثَّل , خَيَال , شَبَهٌ شَبِيه , صُورَةٌ , نَمُوج , مَثِيل , مَجَاز , مِثَال , نَظِير , صُورَةٌ فُوتُوغَرافِيَّة , تَصْنُويِرَةٌ	
شكرا	Thanks	11	8	4	Thanks	اِمْتِنَانٌ , عِزْفَانٌ بِالْإِحْسَانِ , التَّنَاءُ الجَمِيلِ , شكرا	
نعم	Yes	16	7	0	Yes	نعم , موافقة , قَبُول , أجل , بلى	
		200	88	5			

VI. CONCLUSION AND FUTURE WORKS

In this paper, the proposed Ontology with Deep Learning (SDL) method bridges the gap between Ontology and deep learning thus taking advantages of both. The proposed Multi-Sign Language Ontology (MSLO) first version to solve some sign language challenges. It extends the WordNet properties to deal with sign language. Some relations were added in MSLO to be compatible with multiple sign languages.

A Pre-made Arabic Sign Language dataset is used to evaluate the built CNN architecture. The training classification accuracy was 98.06 and the recognition accuracy was 88.87%.

To check the feasibility of the proposed method, it was applied to a semantic translation system for translating Arabic static signs to its meanings as Arabic text and other sign languages. Where CNN trained model used in the recognition process with adding the semantic layer. Collected signs of 10 Arabic gestures and their meanings in English and French sign languages used in training and testing the system. Best results were obtained; semantic translation accuracy was 94.31%. Also, the results compared when using deep learning only and when using ontology with deep learning. The SDL output presents all available meanings of the input sign gesture.

As future work, the SDL will enhance to cover dynamic real video in real-time applications. The system will convert to an Android to produce a free long-life mobile application for deaf and speaking impaired people.

REFERENCES

- [1] M.M. Reda, N.G. Mohammed ,and R.A.A.A. Seoud. SVBiComm: Sign-Voice Bidirectional Communication System for Normal, "Deaf/Dumb" and Blind People based on Machine Learning. 1st International Conference on Computer Applications & Information Security (ICCAIS) ,IEEE, pages 1-8, 2018.
- [2] S. Saqib and S. A. R. Kazmi. Repository of static and dynamic signs. International Journal of Advanced Computer Science and Applications(IJACSA), 8(11),2017.
- [3] N. B. Ibrahim, M. M. Selim and H. H. Zayed. An automatic arabic sign language recognition system (arslrs). Journal of King Saud University-Computer and Information Sciences, 30(4): 470-477,2018.
- [4] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. Nature 521(7553): 436-444, May 2015.
- [5] W. Wang, J. Yang, J. Li. S. Xiao, and D. Zhou. Face recognition based on deep learning. In International Conference on Human Centered Computing, Springer, Cham, pages 812-820, 2014.
- [6] K. Noda, Y. Yamaguchi, K. Nakadai, H. G. Okuno, and T. Ogata. Audio-visual speech recognition using deep learning. Applied Intelligence, 42(4): 722-737,2015.
- [7] A. A. Abu-Shareha and M. Rajeswari. A review on ontology-based label extraction from image data. Journal of Theoretical & Applied Information Technology, 71(2): 268-280,2015.
- [8] N.B. Ibrahim, B. Nada, Mazen ,M.M Selim ,and H. H. Zayed. An automatic arabic sign language recognition system (ArSLRS). Journal of King Saud University-Computer and Information Sciences, 30(4) : 470-477,2018.
- [9] R. Alzohairi, R.Alghonaim, W.Alshehri, S.Aloqeely, M.Alzaidan and O.Bchir. Image based arabic sign language recognition system. International Journal of Advanced Computer Science and Applications (IJACSA), 9(3), 2018.
- [10] M. Abdo, A. Hamdy, S. Salem, and E. M. Saad. Arabic alphabet and numbers sign language recognition. International Journal of Advanced Computer Science and Applications (IJACSA), 6(11):209-214,2015.
- [11] S. Hayani, M. Benaddy, O. El Meslouhi, and M. Kardouchi. Arab Sign language Recognition with Convolutional Neural Networks. In 2019 International Conference of Computer Science and Renewable Energies (ICCSRE), IEEE, pages 1-4,2019.
- [12] M. ElBadawy, A. S. Elons, H. A.Shedeed, and M. F Tolba. Arabic sign language recognition with 3D convolutional neural networks. In 2017 Eighth International Conference on Intelligent Computing and Information Systems (ICICIS) ,IEEE, pages 66-71, 2017.
- [13] A. M. Almasoud, and H. S. Al-Khalifa. SemSignWriting: A Proposed Semantic System for Arabic Text-to-SignWriting Translation. Journal of Software Engineering and Applications, 5(08): 604-612, 2012.
- [14] O. Lozynska, and M. Davydov. Information technology for Ukrainian Sign Language translation based on ontologies. ECONTechMOD: an international quarterly journal on economics of technology and modeling processes , 4(2): 13-18, 2015.
- [15] G. Petrucci, C. Ghidini , and M. Rospocher. Ontology learning in the deep. In European Knowledge Acquisition Workshop, Springer, Cham, pages 480-495, 2016.
- [16] P. Hohenecker and T. Lukasiewicz. Deep learning for ontology reasoning. arXiv preprint arXiv:1705.10342 ,2017.
- [17] J. Zhang and C. Zong. Deep Neural Networks in Machine Translation: An Overview. IEEE Intelligent Systems, 5(30):16-25,2015.
- [18] S.Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 ,2015.
- [19] L. Perez and J. Wang. The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:1712.04621, 2017.
- [20] G. Latif, N. Mohammad, J. Alghaz, R. AlKhalaf and R. AlKhalaf. ArASL: Arabic Alphabets Sign Language Dataset. Data in Brief, Volume 23, Article 103756, 2019.