English Section

Using Deep Learning Technology for Real-Time Sign Language Detection and Recognition at Public Libraries in Egypt : An Experimental Study

Dr. Mo’men S. Elnasharty
Library and Information Science Dept
Data Science & AI Dept. Faculty of Graduate Studies of Statistical Research, MDS Candidate. Cairo University
Navigator001@cu.edu.*eg*

Abstract : *Sign Language is the most expressive way for communication between deaf and hearing-impaired people, where information is majorly conveyed through the hand gestures. Despite this, one of the most issues that facing people and public domains that are outside the deaf community is understanding sign language, they are need an interpreter to understand this language. Technologies including deep learning and computer vision aid this issue by providing several technical apps and other platforms. In this paper, The study proposes a real-time sign language detection model using TensorFlow and OpenCV. The main aim of this study is developing model that recognize sign language, and that can be used offline. Through this model, the study can detect a common words and sentences that are expressed by hand gestures, like among American and Arabic sign languages like good (thumbs up), bad (thumbs down), thanks, and live long. This study is belonging to the experimental and applicable research. The results of this study were that the model succeed to detect the hand gestures and recognize it with accuracy 100 %.*

Keywords: *Deep Learning - AI - Public Libraries – Image Recognition – Sign Language.*

تاريخ الاستلام: 21 يوليو 2023 تاريخ القبول: 5 أغسطس 2023

1. Introduction

The necessity of providing access to information is at the heart of the mission of libraries and library associations around the world. many library associations over world like IFLA, ALA, BL emphasize this principle, to support development, learning, creativity, and innovation by provided by high quality libraries services should serve all Kids, young, adult, and special needs Users.

IFLA indicated there are a many of reasons most libraries have not considered focusing particular attention on the provision of services to persons who are deaf. people who are deaf from birth or from an early age often have difficulty reading and tend to not use libraries. Consequently, libraries and deaf people have mostly been unaware of each other [Day, 2000].

However, in recent years, People with special needs become one of the most important groups that have receivedattention as library users. In 2017, IFLA conducted a surveyof all types of libraries, to obtain a snapshot ofthe formal policies, and practical assistance in support of information access – notably through technology – for people with special needs. they covered 470 Libraries from 92 countries, the number of public libraries has been coveredin this report is 167, with 30% of the total libraries, about 30% from these public Libraries have formal policies for accessibility to people with special needs. about 40% public libraries (82 public libraries) have services were adapted to help the deaf people [Bolt & Wyber, 2017].

The primary issue involved with the provision of ser- vices to deaf people is that communication often requires additional effort, knowledge, patience, and (where available)technological aids, additionally individual differences [Day, 2004].

According to Walaa [Mostafa, 2017], many Libraries in Egypthave been lacked to methods, technology, and professional staff to communicate with deaf people.

Deaf People (hard hearing) use Sign Language as their primary means to express their ideas and thoughts with theirown community and with other people with hand and body gestures.

Sign Language is the most expressive way for communication between deaf and hearing-impaired people, where information is majorly conveyed through the hand gestures (Sense, 2023). Despite this, one of the most issues that facing people and public domains that are outside the deaf community is understanding sign language, they are need an interpreter to understand this language.

This type of gesture-based language allows people to convey ideas and thoughts easily overcoming the barriers caused by difficulties from hearing issues. this Language has its own vocabulary, meaning, and syntax which is different from thespoken language or written language [Rastgoo. 2021].

The differences between spoken language and sign languages are:

- Spoken language is a language produced by articulate sounds mapped against specific words and grammatical combinations to convey meaningful messages.
- Sign language uses visual hand and body gesturesto convey meaningful messages [McInnes, 1993].

Hand gestures for sign language can be classified as two classes: static and dynamic. The static gesture is used for alphabet and number representation, and consists of hand gestures, whereas the latter includes motion of hands, head, or both. Whereas the dynamic gesture is used for specific concepts. it includes words, sentences, etc. Natural gestures and formal cues are the two types of sign language:

- The natural cue: is a manual (hand-handed) expression agreed upon by the user (conventionally), recognized to be limited in a particular group (esoteric), and a substitute for words used by a deaf person (as opposed to body language).
- A formal gesture is a cue that is established deliberately and has the same language structure as the community's spoken language [Suthar, 2021].

However, static hand gesture recognition is simpler than dynamic hand gesture recognition, but both recognition is important to the human community [Kodandaram, 2021].

According to statistics from the World Federation of the Deaf. there are 72 million deaf persons worldwide, 80 of these people live in developing countries. There are somewhere between 138 and 300 different types of Sign Language us around globally today [Rastgoo. 2021].

Arabic sign languages (ARSLs) are still in their developmental stages. Only in recent years has there been an awareness of the existence of communities consisting of individuals with disabilities; the Deaf are not an exception. Arab Deaf communities are almost closed ones. Interaction between a Deaf community and a hearing one is minimal and is basically concentrated around families with deaf members, relatives of the deaf, and sometimes play friends and professionals. According to The Deaf Unit Cairo, there are approximately 1.2 million deaf and hard of hearing individuals in Egypt aged five and older [Abdel-Fattah, 2005].

Egypt has its own sign language "Egyptian Sign Language" which used by members of the deaf community in Egypt, although that, there are no official statistics on the number of deaf people or the number of people who use Egyptian Sign Language as their primary language. A study by the World Health Organization In 2007 places the prevalence of hearing loss in Egypt at 16.02 across all age groups. Adding to that, Egyptian Sign Language is not formally recognized by the government [LibGuides, 2023].

2. Problem Formulation

The study problem's is a realistic issue facing community organizations like libraries, Services offices, and other institutions that deal with the public when are visited by a special need user, while these organizations don't have the qualifier staffs to communicate with these users, and don't have a method or means to know their needs or recognize their questions. people who can listen and speak find it challenging to understand the gestures of the sign language.

In the Arab society the view held disability is still one of accommodation rather than assimilation [Abdel-Fattah, 2005].

Another problem standing behind this study is the sign languages are not- universal, every country has a special sign language, In Arabic, hearing learners of sign language vernaculars have considerable difficulty in grasping the idea of not signing for every word in an utterance as one would say it in the spoken variety [Abdel-Fattah, 2005].

This would be a problem to teach sign language to the deaf as there is a limited number of sign language interpreters exit today.

There are 113 schools for the deaf in Egypt, but the standard of education is low. Some use English Sign Language ESL and others are oral. There is little internet information about the schools; the Deaf Unit in Cairo reports it uses ESL and Arabic [LibGuides, 2023].

At the same time, there is no survey about certified sign language interpreters for the Egyptian deaf population.
This problem can be solved by sign language recognition, that is aims to convert sign language into text or speech making, communication easy for both worlds.
This conversion can be done with the help of deep learning techniques and machine learning algorithms. Today Computer Vision and Deep Learning have gained a lot of popularity and much research have builds models based it. In Deep Learning Convolution Neural Networks (CNN) is the most popular neural network algorithm which is a widely used algorithm for Image/Video tasks.

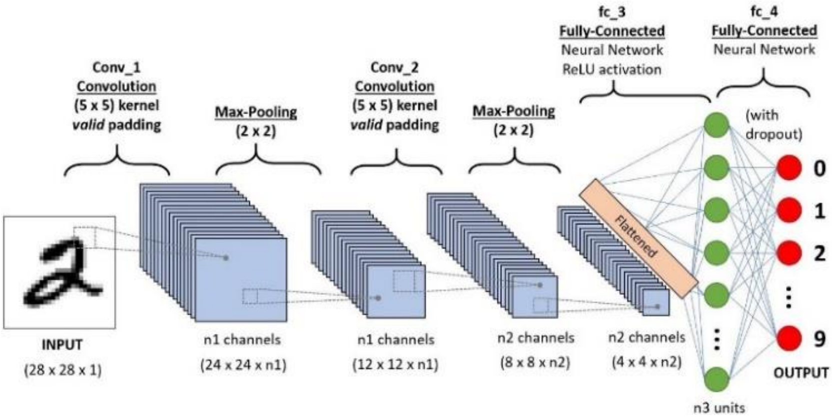

Figure 1. Deep Learning CNN. [Varshini, 2020].

The study uses Deep Learning Computer Vision to recognize the hand gestures by building Deep Neural Network architectures (Convolution Neural Network Architectures) where the model will learn to recognize the hand gestures images over at real time (Figure1).

The design of a sign language translator is quite challenging despite many research efforts during the last few decades for many reasons, most important is signs have significantly different appearances for different signers and different viewpoints.

3. Objectives

The main aim of this study is to contribute to the field of real time sign language recognition and translation to text by developing a system that recognize sign language, and that can be used offline. A vision-based approach was developed to obtain the data from the signer. One of the main features of this study is the ability of the system to identify and recognize the words included in local dataset (researcher customized dataset).
This paper focus on dynamic sign language hand gestures by using Deep Neural Networks (DNN) for recognizing the hand gestures which includes: The four common statements between Arabic and English Sign Language: (Thumbs up – Thank you – Live long – Thumbs down).
By created a convolution neural networks classifier that can classify the 4 statements, this neural network trained by TensorFlow.
The various advantages of building a real time sign language recognition system by Tenser-Flow is translate hand gestures to text, and this will be helpful for using in specific public domains such as public libraries, airports, post offices, or hospitals, to enables inter-communication between normal people and special needs people. in the future this model can be used with embedded devices/standalone applications/web applications having fewer resources.

4. Literature Review

Sign language involves the usage of different parts of the body, such as fingers, hand, arm, head, body, and facial expression [Cheok, 2019]. There are five main parameters in sign language, which are hand-shape, palm orientation, movement, location, and expression/non-manual signals.

One of the main challenges is developing a sign language recognition system to translate the sing words or sentences into the text or voice to facilitate the communication between deaf people and the hearing majority in a real-world situation. The current works in sign language recognition use datasets including videos of only one sign or images of only one character. This situation needs to develop systems that can be applied in a real chat or conversation between a deaf person and a hearing one. To do this, the proposed systems need to be efficient and intelligent enough to split the input videos, including some characters, words, or sentences, into separate characters, words, or sentences. Another main challenge in this area is the high difference between sign languages of different countries [Rastgoo. 2021].

Many applications benefit from sign language recognition advantages such as translation systems, interpreting services, video remote human interpreting, human–computer interaction [Tang, 2015] [Yang, 2015] [Wang, 2016] [Deng, 2017] [Zheng, 2017] [Supancˇicˇ, 2018].

With the advent of deep learning in recent years, many research efforts have been conducted to sign language recognition [Asadi-Aghbolaghi, 2017] [Guo, 2017] [Ferreira, 2019] [Lim, 2019].

One of the early efforts on hand and gesture recognition is dated back to 1987, where a hand gesture interface is proposed by [Zimmerman, et al 1987]. to estimate the hand position and orientation using the magnetic flux sensors of a glove However, there are different challenges to address in visual sign language recognition (e.g., inter and intra subject variability, illumination conditions, partial occlusions, different points.

Over recent years, deep learning methods outperformed previous state-of-the-art machine learning techniques in different areas, especially in Computer Vision and Natural Language Processing [Voulodimos, et al, 2018].

The most significant deep learning models used in computer vision problems are Convolutional Neural Network (CNN) [Li, 2012].

One of the main goals of deep models is to avoid the need to build/extract features. Deep learning allows computational models of multiple processing layers to learn and represent data with multiple levels of abstraction to imitate the human brain mechanism and implicitly capture the complex structures of large-scale data [Rastgoo. 2021].

Deep learning contains a wealthy group of methods, including neural networks, hierarchical probabilistic models, and a variety of unsupervised and supervised feature learning algorithms, one of the important factors that contributed to the huge boost of deep learning is the advent of large-scale, high- quality, and publicly available labeled datasets along with the capability of parallel GPU computing [Rastgoo. 2021].

Convolutional Neural Networks (CNN) is a special architecture of Artificial Neural Networks (ANN), proposed by [LeCun, 1988].

CNN uses some features of the visual cortex, one of the most popular uses of this architecture is image classification, for example, Facebook uses CNN for automatic tagging algorithms, Amazon — for generating product recommendations and Google — for search through among users' photos, instead of the image, the computer sees an array of pixels, for example, if image size is 300 x300, in this case, the size of the array will be 300x300x3. Where 300 is width, next 300 is height and 3 is RGB channel values, the computer is assigned a value from 0 to 255 to each of these numbers. his value describes the intensity of the pixel at each point [Wu, 2017].

One of the strongest techniques was be developed at sign language recognition by deep learning model was TensorFlow frameworks [Rastgoo. 2021] [TensorFlow, 2023] TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library and is also used for machine learning applications such as neural networks. It is used for both research and production at Google. TensorFlow can be written in C++, Python, and CUDA [TensorFlow, 2023].

Another most tech used for these tasks is OpenCV (Open-Source Computer Vision Library) is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel, it was later supported by Willow Garage. The library is cross-platform and free for use under the open-source license [OpenCV, 2023].

Keras, is another tech using for creating real time sign languages recognition systems, Keras is an open-source neural-network library written in Python. It can run on top of TensorFlow, Microsoft Cognitive Toolkit, R, Theano, or PlaidML. Designed to enable fast experimentation with deep neural networks, it focuses on being user- friendly, modular, and extensible [Keras, 2023].

Generally, vision-based, and glove-based approaches are the two main categories considered for sign language recognition models. While the vision-based models use the captured video data of the signers for different signs, the glove- based models employ some mechanical or optical sensors attached to a glove to use the electrical signals for hand pose detection. Vision-based models provide more natural and real systems based on the information that human can sense from the surroundings. The Focusing will be on vision-based models [Rastgoo. 2021] (Figure 2).

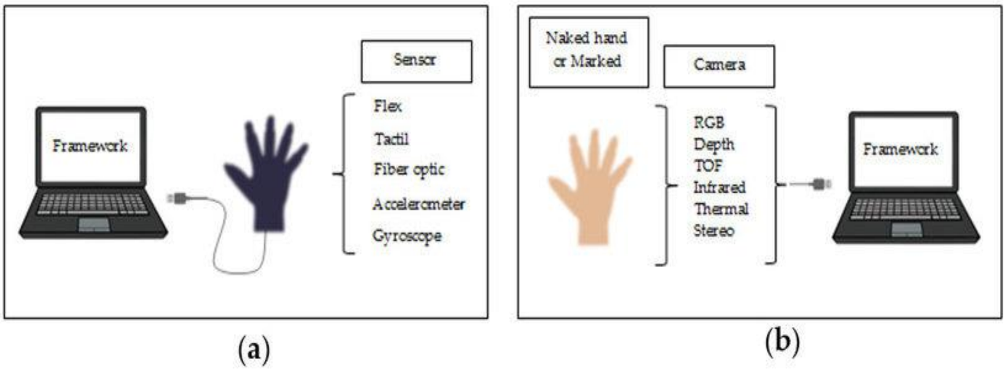

Figure 2: vision-based, and glove-based approaches. [Oudah, 2020]

There are some visual input data modalities in sign language recognition area in recent years. RGB and depth input data are two common types of input data used in sign language recognition models. While RGB images or videos include the high-resolution contents, depth inputs have the accurate information related to the distance between the im- age plane and the corresponding object in the image. Some of the models use the advantages of two input modalities, simultaneously.

There are different devices to record the input data modalities and camera is the most common one of them. Different cameras support different formats and qualities of input data.

Two forms of input data, static or dynamic, are used in sign language recognition models to extract the necessary features. Many deep based models have been proposed to use the still or sequence inputs in recent years. While dynamic inputs include the sequential information that could be useful to improve the sign language recognition accuracy.

The most famous datasets, including videos, for sign language recognition was be mentioned in [Rastgoo. 2021].

Deep sign language recognition models categorized based on the datasets used for evaluation has been descried by [Rastgoo. 2021].

[Nandy et al, 2010] proposed an approach for detecting and recognizing Indian Sign Language gestures from gray-scaled images. In their approach, a video source that contains signing gestures is converted into gray-scaled frames, from which features are extracted by applying a directional histogram. Lastly, clustering is used to classify the signs, ac- cording to their features, into one of the pre-defined classes. The authors achieved a 100 %. sign recognition rate in their study and concluded that the 36-bin histogram method was more accurate than the 18-bin histogram method.

[Mekala et al, 2011] proposed neural network architecture for recognizing and tracking sign language and generating real- time text from the video stream. The system architecture consists of different phases, such as framing, pre-processing images, extracting the feature based on hand movement and hand position, etc. These hand attributes are represented using a point of interest (POI) of the hand. By using this method 55 different features were extracted, which were used as an input for the neural network architecture proposed by the authors, consisting of CNN layers that predicted the signs. I their study, they trained and tested the model on the English alphabet, from A to Z, and claimed to achieve a 100 % recognition rate and 48 % of noise immunity.

[Shivashankara & Srinath, 2018] developed a system based on American Sign Language (ASL) for recognizing signs/gestures. The authors proposed a model in which they used YCbCr, which optimized the performance of skin color clustering. This model was used for image pre- processing. To recognize the gesture, the centroid of the hand was identified in the pre-processed image; then, the peak offset was found to identify the gesture. The overall accuracy of this model was 93.05 %.

5. Methodology

This study is belonging to the experimental and applicable research, is an implementation project for designing a model to recognition the common sentences that using of sign language at Arabic and American sign language in Egyptian deaf community. to create this model, it is necessary to go through the following phases:

1) Collecting Data.
2) pre-processing Data
3) model construction.
4) model training.
5) model testing.
6) model evaluation.

This model is based on Artificial Neural network by using CNN, TensorFlow, and OpenCV libraries. starting by which web camera is used for capturing images of hand gestures. then the study labels images, to divide it into two parts, one for model training, and other for testing.

6. Data

Many large training datasets for Sign Language are available on [Kaggle, 2023]. The one used in general is called "Sign Language MNIST" [MNIST, 2023] and is a public-domain free-to-use dataset with information for around 1,000 images of each of 24 American Sign Language Letters. But this datasets for static sign languages which depended on using gesture for alphabet and number representation, not for Dynamic Sign language which representing a common words or sentences; so, this study is directing to collecting these signs by recording hand gesture by camera. this is meaning, this study will process the sign language captured in real-time to recognize the sign language words.

The Researcher created dataset for training and testing. The dataset was collected from researcher himself. The dataset was collected using OpenCV and webcam, to five topic (hand gestures) that express to 5 words these are (good) (thumbs up) (Figure 6), bad (thumbs down) (Figure 4), thanks (Figure 5), and live long (Figure 3).

Figure 3. Live Long

20 images have been collected for every topic, and all images have a resolution of 1920 X1080, with an average length of 2 s. The images are recorded using an OpenCV webcam on a laptop or desktop with moderate configurations. This dataset was created under natural conditions—without any extra brightness, orientation, Kinect sensor, background adjustments, gloves, and take shoot from different angles, with fast and slow momentum of the hands of the person.

Figure 5. Thanks sign

Figure 6. Thumbs up Sign

Figure 4. Thumbs down Sign

Furthermore, the study did not use any external help from sensor or smart gloves to detect hand movements, as shown in Figure 3. The distance between the user and the laptop, correctly positioned, is about 30 cm. This distance can be exceeded by 32 to 35 cm. about technology utilized is:

1) Interface: Jupyter notebook for inserting python libraries in a notebook format, it is typically a python code where study can easily estimate our data sets model in one single notebook.
2) Operating System Environment: Windows 10.
3) Software: Python (3.7.4), Anaconda (2019-0.7), IDE (Jupyter), NumPy (version 1.16.5), cv2 (openCV) (version 3.4.2), TensorFlow (version 2.11.0), Github, Virtual Studio (2022), CUDA (10.1 and CuDNN (7.6) (For NIVIA GPU for faster training model).
4) Hardware Environment: RAM- 16GB, GRAPHIC CARD – 6GB, ROM-1060TB.
5) TensorFlow: It is an open-source artificial intelligence package that builds models using data flow graphs. It enables developers to build large-scale neural networks with several layers. TensorFlow is mostly used for classification, perception, comprehension, discovery, prediction, and creation.
6) Object Detection API: It is an open-source Tensor- Flow API to locate objects in an image and identifyit.
7) OpenCV: OpenCV is an open-source, highly optimized Python library targeted at tackling computer vision issues. It is primarily focused on real-time applications that provide computational efficiency for managing massive volumes of data. It processes photos and movies to recognize items, people, and even human handwriting.
8) LabelImg: LabelImg is a graphical image annotation tool that labels the bounding boxes of objectsin pictures.

7. Pre-processing Data

Main aim of pre-processing is an improvement of the image data that reduce unwanted deviation or enhances image features for further processing.

Pre-processing is also referred as an attempt to capture the important pattern which express the uniqueness in data without noise or unwanted data which includes cropping, resizing and gray scaling.

7.1. Cropping

Cropping refers to the removal of the unwanted parts of an image to improve framing, accentuate subject matter or change aspect ratio.

7.2. Labelling Images

study used LabelImg [Boesch, 2023] to annotate images. study used LabelImg to label every image, study is drawing a rectangle across the sign language gesture. This will allow us to separate each gesture into different categories. Once the annotations for images are done, study saves the images in our training and test folders, (Figure 7).

The XML (eXtensible Markup Language) files will be created for each data image. XML files consist of the coordinates of gesture bounding boxes, category, filename, etc. for each sign language gesture inside the image.

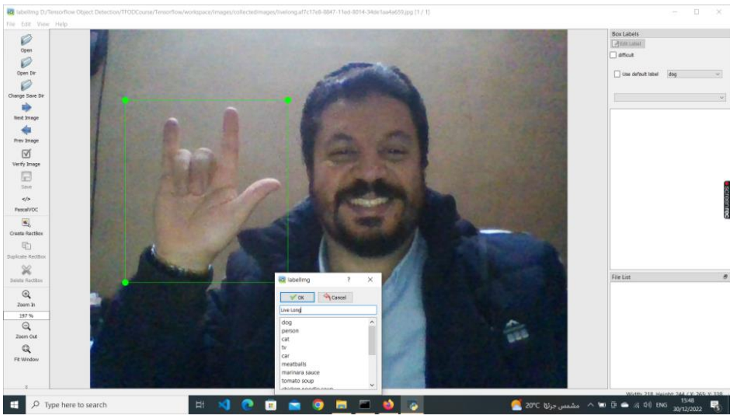

Figure 7. Detection to Sign and Labelling by LabelImg.

8. Model

There are two types of SLR systems – isolated SLR and continuous SLR. In isolated SLR, the system is trained to recognize a single gesture. Each image is labelled to rep- resent an alphabet, a digit, or some special gesture. Continuous SLR is different from isolated gesture classification. In continuous SLR, the system can recognize and translate whole sentences instead of a single gesture.

In this proposed model TensorFlow object detection API is used to solve a real-time problem of sign language detection.

This model uses pipeline that takes input through a web camera from a user who is signing a gesture and then by extracting different frames of video, it generates sign language possibility for each gesture.

The Main Steps to Creating Study Model are:

1) Collect images for deep learning using your web-cam and OpenCV.
2) Label images for sign language detection usingLabelImg.
3) Setup TensorFlow Object Detection pipeline configuration.
4) Use transfer learning to train a deep learning model.
5) Detect sign language in real time using OpenCV

First stage was dividing dataset into 2 files, first one for training model with 70 % images, second for testing model with 30 % images.

Next stage is TF (TensorFlow) Record generation, where TensorFlow API accepts datasets in TF-Record file format, which is a sequence of binary format strings. To record helps in better performance of import pipeline and in the training time of the model (Figure 8).

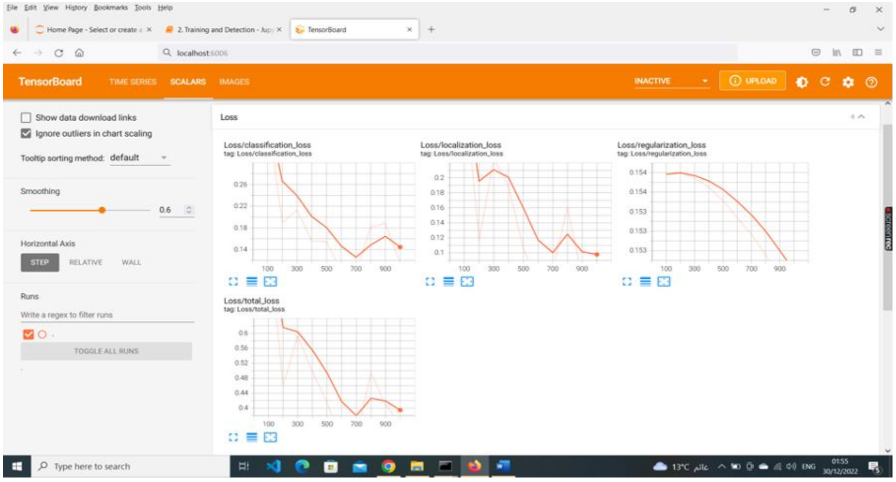

Figure 8. TensorFlow object detection workflow

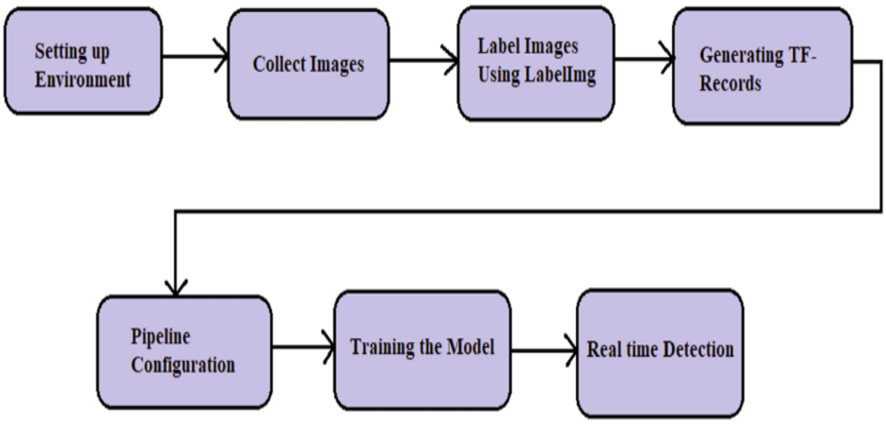

Figure 9. Training Model

After creating the TFR dataset, study used the SSD- mobilenet-v2-pnlite-320x320-coco17 model which was pre- trained on the coco dataset. Configuration of a model is necessary to modify model-related features i.e., hyperparameters, loss function, etc. so that study can train the model to detect the objects that study is interested in.

Now lastly, study did the training of the model using trans- fer learning in a terminal window. The Study trained our model for 2000 training steps for better results. The model was saved at several checkpoints. In our case, the total checkpoints created are 3. After training for a few hours, the total loss gets down to 0.242. Now study load our trained model from checkpoints and load the object detection model (Fig. 9).

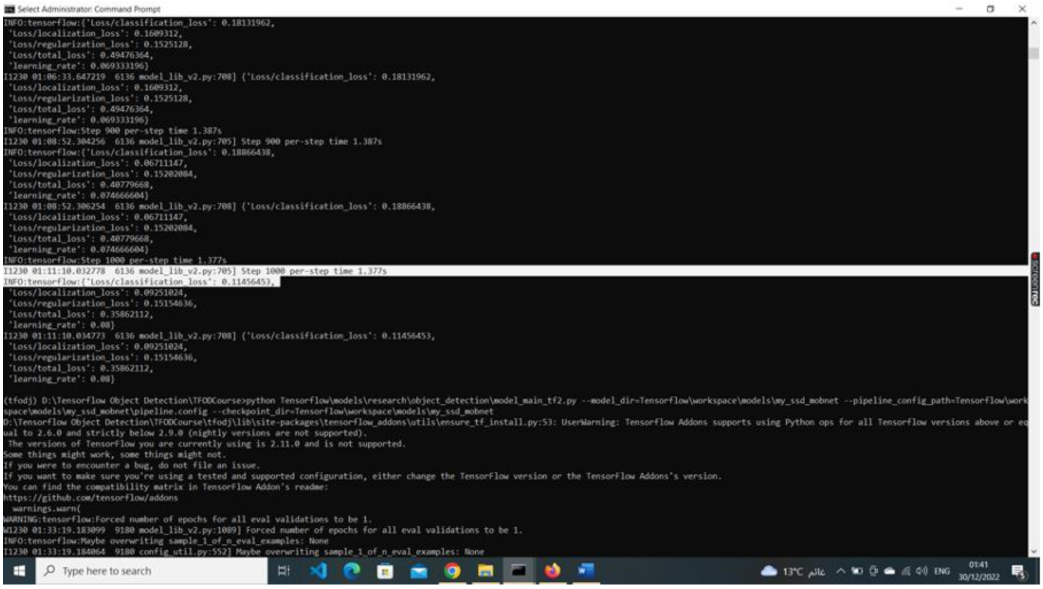

Figure 10. Evaluating The Model inside of Tensor board

After completing the training process for the proposed model, study evaluated this model by verifying the recall and precision rate of the model’s output, using the confusion matrix. The model’s validation rate in recall function was about 50 percent, and the model precision rate was approximately 50 percent. Tensor board is used to create the Confusion Matrix and determine the perception ratio (Figure 10).

9. Results

The model has been proposed in this paper works to detect real-time sign and hand gestures for common sentence among American and Arabic sign language. Study trained our model for 4 gestures used in sign language. study has toked 20 images for each sign, then has divided them into train and test files. The study has toked images from several angles to get a better result in the training process for high accuracy. The experimental results show the capability of model can be detect and recognize the sign in image with 98 % to 100 % accuracy rate. The TensorFlow object detection API provides better accuracy and performance for real time sign language detection and recognizing, the result of the study, and accuracy of model for detecting the sign and recognizing it, is shown in (figure 11,12).

Figure 12. The model is detecting & recognising (Live Long) Sign, with accuracy rate 100 percent.

Figure 11. The model is detecting & recognizing (Thank you) Sign, with accuracy rate 100 percent.

10.Conclusion

The main objective of this paper is to build a real- time sign language detection model using TensorFlow and OpenCV. This model is to decrease the conversation gap between deaf people and the acute majority. Through this model, study detected 4 common gestures between Arabic and American sign languages that's "Thumbs Up", "Thumbs Down" "Thanks", "Live long". Accuracy rate of the pro- posed Model to detect and recognize for each sign language gestures has been between 98 % to 100 %. This accuracy rate can be increased using a large dataset with each gesture captured with different angles.

11.References

Abdel-Fattah. Mahmoud A. (2005). Arabic sign language: a perspective. Journal of deaf studies and deaf education, 10(2):212–221. Available a: DOI:10.1093/deafed/eni007.

Asadi-Aghbolaghi, Maryam, Albert Clape´s, Marco Bellantonio, Hugo Jair Escalante, V´ıctor Ponce-Lo´pez, Xavier Baro´, Isabelle Guyon, Shohreh Kasaei, and Sergio Escalera. (2017). Deep learning for action and gesture recognition in image sequences: A survey. In Gesture Recognition, pages 539–578. Springer, 2017.

Boesch, Gaudenz. (2023). "LabelImg for Image Annotation". Available at: https://viso.ai/computer-vision/labelimg-for-image-annotation/

Bolt, Nancy & Wyber, Stephen. (2017) "Library service to people with disabilities". IFLA. Available at:https://www.ifla.org/wp-content/uploads/2019/05/assets/hq/topics/libraries-development/documents/report_to_medd_on_library_service_to_people_with_disabilities_final_2_14_18.pdf

Cheok, Ming. Jin, Zaid Omar, and Jaward Mohamed Hisham. (2019). A review of hand gesture and sign language recognition techniques. International Journal of Machine Learning and Cybernetics, 10(1):131–153.

Day, John Michael. (2004) "Guidelines for library services to deaf people". IFLA. Available at: https://www.ifla.org/wp-content/uploads/2019/05/assets/hq/publications/professional-report/62.pdf

Deng, Xiaoming, Shuo Yang, Yinda Zhang, Ping Tan, Liang Chang, and Hongan Wang. (2017). Hand3d: Hand pose estimation using 3d neural network. arXiv preprint arXiv:1704.02224.

Ferreira, Pedro M, Jaime S Cardoso, and Ana Rebelo. (2019). On the role of multimodal learning in the recognition of sign language. Multimedia Tools and Applications, 78(8):10035–10056.

Guo, Hengkai, Guijin Wang, Xinghao Chen, and Cairong Zhang. (2017). Towards good practices for deep 3d hand pose estimation. arXiv preprint arXiv:1707.07248.

Kaggle. (2023). https://www.kaggle.com/

Keras. (2023). https://keras.io/

Kodandaram. Satwik Ram, N. Kumar, and Sunil Gl. (2021). "Sign language recognition". Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12:994–1009, 08.

L,i Zewen, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. IEEE transactions on neural networks and learning systems.

LeCun, Y., Touresky, D., Hinton, G., & Sejnowski, T. (1988, June). A theoretical framework for back-propagation. In Proceedings of the 1988 connectionist models summer school (Vol. 1, pp. 21-28).

Libguides: Deaf statistics: Africa amp; the Americas. Africa amp; the Americas - Deaf Statistics – Lib Guides at Gallaudet University Library.

LibGuides. (2023). Deaf statistics Africa the Americas. https://libguides.gallaudet.edu/c.php?g=773916p=5553200.

Lim, Kian Ming, Alan Wee Chiat Tan, Chin Poo Lee, and Shing Chi- ang Tan. (2019). Isolated sign language recognition using convolutional neural network hand modelling and hand energy image. Multimedia Tools and Applications, 78(14):19917–19944.

McInnes. John M and Treffry, Jacquelyn A. (1993). "Deaf-blind infants and children: A developmental guide". Available at: https://www.jstor.org/stable/10.3138/j.ctt13x1pr2.

Mekala. Priyanka, Ying Gao, Jeffrey Fan, and Asad Davari. (2011). Real-time Sign language recognition based on neural network architecture. In 2011 IEEE 43rd Southeastern symposium on system theory, pages 195–199. IEEE, 2011.

MNIST. (2023). https://www.kaggle.com/datasets/hojjatk/mnist-dataset

Mostafa, Walaa M. (2017). "Information services in deaf libraries at Egypt". Cybrarians Journals. Available at:http://journal.cybrarians.info/index.php?option=com_content&view=article&id=777%3Awmegahed&catid=303%3Areviews&Itemid=80

Nandy. Anup, Jay Shankar Prasad, Soumik Mondal, Pavan Chakraborty, and Gora Chand Nandi. (2010). Recognition of isolated Indian sign language gesture in real time. In International conference on business administration and information processing, pages 102–107. Springer.

OpenCV. (2023) https://opencv.org/

Oudah, Munir & Al-Naji, Ali Abdulelah & Chahl, Javaan. (2020). Hand Gesture Recognition Based on Computer Vision: A Review of Techniques. Journal of Imaging. 6. 73. 10.3390/jimaging6080073.

Rastgoo, Razieh, Kiani, Kourosh, and Escalera, Sergio. (2021). "Sign language recognition: A deep survey". Expert Systems with Applications, 164:113794, 2021. Available at: https://doi.org/10.1016/j.eswa.2020.113794

Sense (2023). "Sign language". Available at: https://www.sense.org.uk/information-and-advice/communication/sign-language/

Shivashankar, a Ss and S Srinath. (2018). American sign language recognition system: an optimal approach. International Journal of Image, Graphics and Signal Processing, 11(8):18.

Supancˇicˇ, James Steven, Gregory Rogez, Yi Yang, Jamie Shotton, and Deva Ramanan. (2018). Depth-based hand pose estimation: methods, data, and challenges. International Journal of Computer Vision, 126(11):1180–1198.

Suthar, Jay. Parikh. Devansh, Sharma. Tanya, Patel. Avi (2021). "Sign Language Recognition for Static and Dynamic Gestures". Global Journal of Computer Science and Technology Volume XXI, Issue II, Version 1, Year 202. Available at: https://globaljournals.org/GJCST_Volume21/1-Sign-Language-Recognition-for-Static.pdf

Tang, Ao, Ke Lu, Yufei Wang, Jie Huang, and Houqiang Li. (2015). A real- time hand posture recognition system using deep neural networks. ACM Transactions on Intelligent Systems and Technology (TIST), 6(2):1–23.

TensorFlow. (2023). https://www.tensorflow.org/learn

Varshini, Chava & Hruday, Gurram & Chandu, Guguloth & Sharif, Shaik. (2020). Sign Language Recognition. International Journal of Engineering Research and. V9. 10.17577/IJERTV9IS050781.

Voulodimos, Athanasios, Nikolaos Doulamis, Anastasios Doulamis, and Eftychios Protopapadakis. (2018). Deep learning for computer vision: A brief review. Computational intelligence and neuroscience, 2018.

Wang, Xingyou, Weijie Jiang, and Zhiyong Luo. (2016). Combination of convolutional and recurrent neural network for sentiment analysis of short texts. In Proceedings of COLING 2016, the 26th international conference on computational linguistics: Technical papers, pages 2428–2437.

Wu, Jianxin. (2017). to convolutional neural networks. National Key Lab for Novel Software Technology. Nanjing University. China, 5(23):495.

Yang, Yezhou, Yi Li, Cornelia Fermuller, and Yiannis Aloimonos. (2015). Robot learning manipulation action plans by"

watching" unconstrained videos from the world wide web. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 29.

Zheng, Lihong, Bin Liang, and Ailian Jiang. (2017). Recent advances of deep learning for sign language recognition. In 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), pages 1–7. IEEE.

Zimmerman, Thomas G, Jaron Lanier, Chuck Blanchard, Steve Bryson, and Young Harvill. (1986). A hand gesture interface device. ACM Sigchi Bulletin, 18(4):189–192

تطبيق تقنية التعلم العميق للتعرف على لغة الإشارة في الوقت الفعلي داخل المكتبات العامة المصرية : دراسة تجريبية

د. مؤمن سيد النشرتي
مدرس علوم البيانات والمعلومات
قسم المكتبات والوثائق والمعلومات
كلية الآداب – جامعة القاهرة

تاريخ الاستلام: 21 يوليو 2023 **تاريخ القبول: 5 أغسطس 2023**

المستخلص:

تعد لغة الإشارة من أكثر الطرق تعبيراً للتواصل بين الصم وضعاف السمع، حيث يتم نقل المعلومات بشكل رئيسي من خلال إيماءات اليد. وعلى الرغم من ذلك، فإن فهم لغة الإشارة تعد من أحد أكثر المشكلات التي تواجه الأشخاص والمجالات العامة خارج مجتمع الصم، في ظل الحاجة إلى مترجم لفهم هذه اللغة. وعلي هذا تأتي التقنيات الحديثة كتقنيات التعلم العميق ورؤية الكمبيوتر كأحد وسائل التي تساعد على حل هذه المشكلة وذلك من خلال توفير العديد من التطبيقات التقنية والأنظمة الأساسية الأخرى في هذا البحث، تقترح الدراسة نموذجًا للكشف عن لغة الإشارة في الوقت الفعلي باستخدام نماذج للتعلم العميق وهي TensorFlow وOpenCV

ويعد الهدف الرئيسي من هذه الدراسة هو تطوير نموذج يتعرف على لغة الإشارة، من خلال هذا النموذج، يمكن الكشف عن الكلمات والجمل الشائعة التي يتم التعبير عنها بإيماءات اليد، كما هو الحال بين لغات الإشارة الأمريكية والعربية مثل إشارة جيد، وإشارة سيء، وإشارة الشكر، وإشارة الدعاء بطول العمر تنتمي هذه الدراسة إلى المنهج التجريبي والقابل للتطبيق وكانت نتائج هذه الدراسة أن النموذج نجح في الكشف عن إيماءات اليد والتعرف عليها بدقة 100٪.

الكلمات المفتاحية: الذكاء الاصطناعي، التعلم العميق، لغة الإشارة، المكتبات العامة، التعرف على الصور.