IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14714

Real Time Sign Language Recognition using Machine Learning Techniques

Seif ELduola F.

El Haj Technology University, Computer Science and Information Technology

Abstract: Sign language is a good visual communication aid for those with auditory disabilities. This language is also prevalent for those with speech impairment. However, the general populous have little knowledge on sign language, and often find difficulty communicating with someone who is primarily versed in sign language. Our goal is to build a system that can provide robust hand sign-language gesture recognition for SL. This can be of extensive help in public places, especially sign language that isn't often universally understood by the majority of people. This chose four signs used worldwide and prepared a data set for these signs and performed the necessary processing for them, then we chose the SVM algorithm to classify these data, and the algorithm showed a high classification efficiency that reached 99%.

I. INTRODUCTION

Many people are suffering from hearing loss, speaking impairment or both. A partial or complete inability to hear in one or both ears is known as hearing loss. On the other hand, mute is a disability that impairs speaking and makes the affected people unable to speak. If deaf-mute happens during childhood, their language learning ability can be hindered and results in language impairment, also known as hearing mutism. These ailments are part of the most common disabilities worldwide [1]. Statistical report of physically challenged children during the past decade reveals an increase in the number of neonates born with a defect of hearing impairment and creates a communication barrier between them and the rest of the world [2]. Sign language recognition aims to recognize meaningful movements of hand gestures and is a significant solution in intelligent communication between the deaf community and hearing societies. In last decade lot of efforts had been made by research community to create sign language recognition system which provide a medium of communication for differently-abled people and their machine translations help others having trouble in understanding such sign languages. Computer vision and machine learning can be collectively applied to create such systems.

1.1 Problem Statement

- Dumb people use hand signs to communicate, hence normal people face problem in recognizing their language by signs made.
- Hence there is no system in Sudan which recognizes the different signs and conveys the information to the normal people.

1.2 Objective of the Study

This study formulates the following specific objectives:

- To build a machine learning model that will be able to classify the various hand gestures using SVM machine learning algorithm.
- To use model to detect real time sign language and convert real time sign language from video into text.

Scope and limitation

Machine learning techniques will be used to recognize 4 signs for deaf and dump in real time.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14714

II. LITERATURE REVIEW

Machine Learning (ML)

Is a field of inquiry devoted to understanding and building methods that 'learn', that is, methods that leverage data to improve performance on some set of tasks. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so. [5] Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, speech recognition, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks. [6]

Supervised Machine Learning

Supervised learning is the types of machine learning in which machines are trained using well "labelled" training data, and on basis of that data, machines predict the output. The labelled data means some input data is already tagged with the correct output.

In supervised learning, the training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly. It applies the same concept as a student learns in the supervision of the teacher.

Supervised learning is a process of providing input data as well as correct output data to the machine learning model. The aim of a supervised learning algorithm is to find a mapping function to map the input variable(x) with the output variable(y).

In the real-world, supervised learning can be used for Risk Assessment, Image classification, Fraud Detection, spam filtering, etc.

In supervised learning, models are trained using labelled dataset, where the model learns about each type of data. Once the training process is completed, the model is tested on the basis of test data (a subset of the training set), and then it predicts the output.

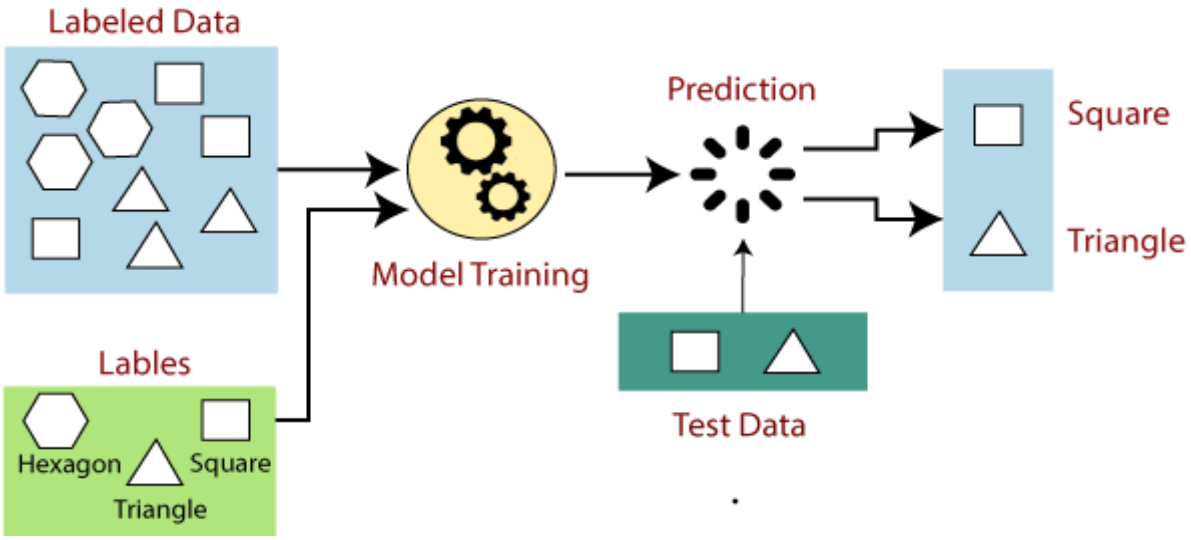

Fig 2: The working of supervised learning

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14714

Support vector machine algorithm(SVM)

A Support Vector Machine (SVM) is a supervised machine learning algorithm that can be employed for both classification and regression purposes. SVMs are more commonly used in classification problems and as such, this is what we will focus on in this post.

SVMs are based on the idea of finding a hyperplane that best divides a dataset into two classes, as shown in the image below.

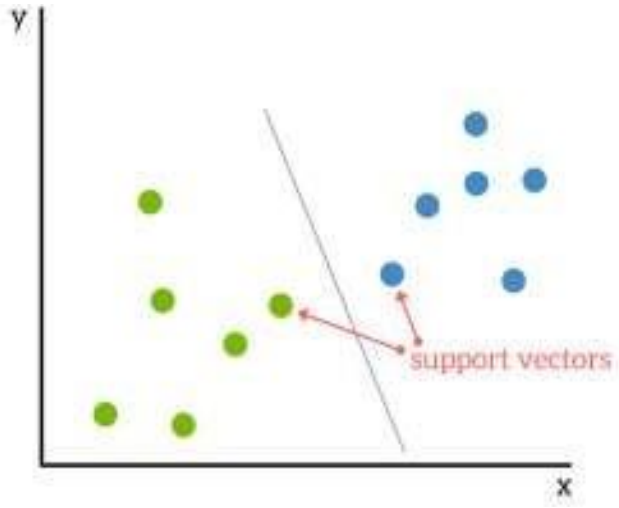

Fig 4:Support Vectors

Support vectors are the data points nearest to the hyperplane, the points of a data set that, if removed, would alter the position of the dividing hyperplane. Because of this, they can be considered the critical elements of a data set.

2.1 Hyperplane

Hyperplane as a line that linearly separates and classifies a set of data.

Intuitively, the further from the hyperplane our data points lie, the more confident we are that they have been correctly classified. We therefore want our data points to be as far away from the hyperplane as possible, while still being on the correct side of it.

So when new testing data is added, whatever side of the hyperplane it lands will decide the class that we assign to it.
The distance between the hyperplane and the nearest data point from either set is known as the margin. The goal is to choose a hyperplane with the greatest possible margin between the hyperplane and any point within the training set, giving a greater chance of new data being classified correctly.

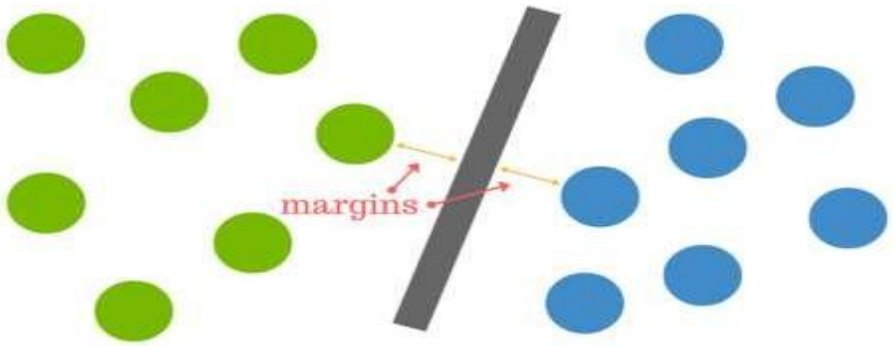

Fig 5: Hyperplane

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025
DOI: 10.17148/IJARCCE.2025.14714

2.2 SVM Uses

SVM is used for text classification tasks such as category assignment, detecting spam and sentiment analysis. It is also commonly used for image recognition challenges, performing particularly well in aspect-based recognition and color-based classification. SVM also plays a vital role in many areas of handwritten digit recognition, such as postal automation services.

2.3 Sign language recognition (SLR)

Is a multidisciplinary research area that involves natural language processing, linguistics, pattern matching, computer vision and machine learning [7]. The final goal of sign language recognition is to develop methods and algorithms in order to build a sign language recognition system (SLRS) capable of identifying already produced signs, decoding their meaning and producing some sort of textual, or visual output in another language that the intended receiver can understand.

2.4 Related Work

Research in sign language recognition has gained a notable interest since many years ago. Nowadays, modern technologies in handheld and smart devices facilitates a lot many processes in computer vision tasks. Also, various programming languages become rich of off-the-shelf packages and source codes, in particular those of developing mobile apps. Most researchers have been following one of three approaches: sensor-based gloves, 3-D skeletons, or computer vision. The first two approaches neglect facial expressions which play a huge rule in sign language recognition. On the other hand, computer-vision systems are capable of capturing the whole gesture, not to mention their mobility that differentiates them from glovebased systems. Reference [14] proposed a computer vision approach for continuous American sign language recognition (ASL). He used a single camera to extract two- dimensional features as input of the Hidden Markov Model (HMM) on a dataset of 40 words collected in a lab. He followed two approaches for the camera position: desk- mounted cam with a word recognition accuracy of 92% and wearable capmounted cam with an accuracy of 98%. Another computer vision system developed by Dreuw et al. was able to recognize sentences of continuous sign language independent of the speaker, described in Reference [15]. He employed pronunciation and language models in sign language with a recognition algorithm based on the Bayes‟ decision rule. The system was tested on a publicly available benchmark database consisting of 201 sentences and 3 signers, and they achieved a 17% word error rate. Reference [16] also proposed a computer vision system that uses hand and face detection for classifying ASL alphabets into four groups depending on the hand‟s position. The system used the inner circle method and achieved an accuracy of 81.3%. Reference [17] followed a different approach for classifying ASL alphabets. They used the Leap Motion controller to compare the performance of the Naive Bayes Classifier (NBC) with a Multilayer Perceptron (MLP) trained by the backpropagation algorithm. An accuracy of about 98.3% was achieved using NBC, while MLP gave an accuracy of about 99.1% In the past 20 years, deep learning have been used in sign language recognition by researchers from all around the world. Convolutional Neural Networks (CNNs) have been used for video recognition and achieved high accuracies last years. B. Garcia and S. Viesca at Stanford University proposed a realtime ASL recognition with CNNs for classifying ASL letters, described in Reference [18] . After collecting the data with a native camera on a laptop and pre-training the model on GoogLeNet architecture, they attained a validation accuracy of nearly 98% with five letters and 74% with ten. CNNs have also been used with Microsoft Kinect to capture depth features. Reference [19] proposed a predictive model for recognizing 20 Italian gestures. After recording a diverse video dataset, performed by 27 users with variations in surroundings, clothing, lighting and gesture movement, the model was able to generalize on different environments not occurring during training with a cross-validation accuracy of 91.7%. The Kinect allows capture of depth features, which aids significantly in classifying ASL signs. Reference [20] also used a CNN model with kinect for recognizing a set of 50 different signs in the Flemish Sign Language with an error of 2.5%. however, this work considered only a single person in a fixed environment. There‟s another version of CNNs called 3D CNNs, which was used by Reference [20] to recognize 25 gestures from Arabic sign language dictionary. The recognition system was fed with data from depth maps. The system achieved 98% accuracy for observed data and 85% average accuracy for new data. Computer vision systems face two major challenges: environmental concerns (e.g. lighting sensitivity, background) and camera‟s position. Most previous systems lack the diversity of data and capturing the whole gesture.

III. METHODOLOGY

There are several ways for recognizing gestures, which includes sensor-based and vision-based systems. Sensor-equipped devices capture numerous parameters such as the trajectory, location, and velocity of the hand in the sensor-based approach. On the other hand, vision-based approaches are those in which images of video footages of the hand gestures are used.[8] The steps followed for achieving the sign language recognition are:

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14714

Here are the steps regularly found in machine learning project:

- Collect dataset
- Split the data into labels
- Pre-process the data
- Divide the data into training and testing sets
- Train the SVM algorithm
- Make some predictions
- Evaluate the results of the algorithm

3.1 Dataset

A camera has been used to record hand gestures and the obtained videos have been converted into image frames. The frames have been passed through a pre-processing phase using open computer vision (openCV) library which reads the video frame by frame. These frames have been sent to media-pipe which locates landmarks in each frame. The feature extraction module has used these landmarks to extract features. Then the extracted feature set has been passed to train classifiers.

3.1.1 OpenCV

A huge open-source library for computer vision, machine learning, and image processing, OpenCV plays a major role in real-time operations today. the main aim of using OpenCV was real-time applications for computational efficiency. Upon integration with other libraries, such as NumPy, Python can process the OpenCV array structure for analysis. Identifying image patterns and its several features needs use of vector space and carrying out mathematical operations on these features.

3.2 MediaPipe

MediaPipe is a framework that enables developers for building multi-modal(video, audio, any times series data) cross-platform applied ML pipelines. MediaPipe has a large collection of human body detection and tracking models which are trained on a massive and most diverse dataset of Google. As the skeleton of nodes and edges or landmarks, they track key points on different parts of the body.

ML pipeline at its backend consisting of two models working dependently with each other: a) Palm Detection Model b) Land Landmark Model.

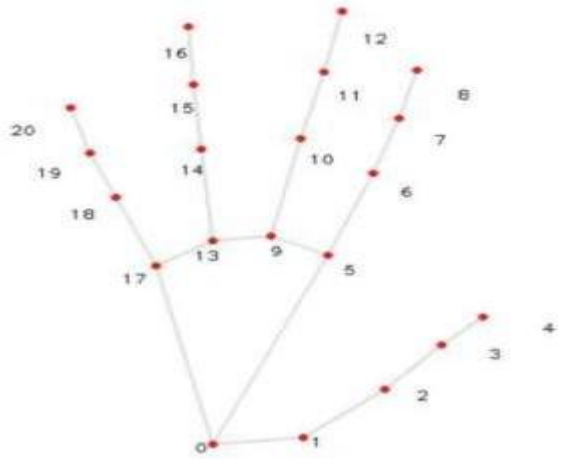

Figure 5: Landmark generation on hand and finger from 0 to 20

3.3 Classification

We proceed to the classification stage once the feature detection and extraction process is finished. It involves classification using a Support Vector Machine (SVM).

3.4.1. Support Vector Machine

The Support Vector Machine (SVM) is a supervised model that can solve both linear and non-linear problems for

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14714

classification and regression problems. It operates on the idea of decision planes that specify boundaries for decisions.

For this classification, we have used SVM. We have passed cvm file to the SVM as feature vectors for the classification and recognition of SL signs. The training is done using a total of 1000 images. After the training is completed, the performance of the classifier is checked on the testing set which has a total of 933 images, and its performance is evaluated on various parameters like accuracy.

3.4 Prediction using Machine Learning Algorithm

Predictive analysis of different sign languages are performed using Support Vector Machine (SVM). The details of the analysis are discussed in the result section. SVM is effective in high dimensional spaces. In the case where the number of samples are greater than the number of dimensions, SVM performs effectively. SVM is a cluster of supervised learning methods capable of classification, regression and outliers detection.

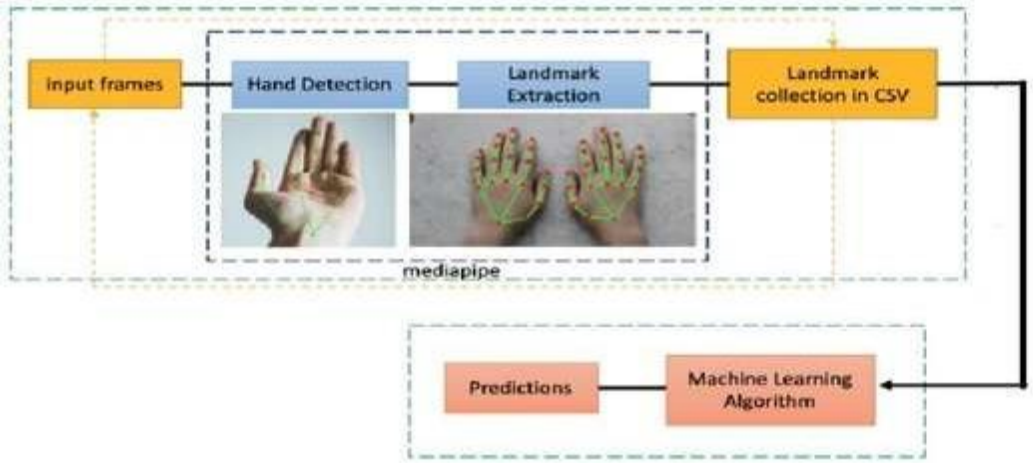

Figure 6: Proposed architecture to detect hand gestures and predict sign language finger-spellings.

4.1 Proposed System

Our proposed system is sign language recognition system using SVM which recognizes various hand gestures by

1. Capturing video and converting it into frames.
2. Then the hand pixels are segmented and the image it obtained and sent for comparison to the trained model.
3. Thus our system is more robust in getting exact text labels of letters.

Results

With an average accuracy of 99% in most of the sign language dataset using Media Pipe‟s technology and machine learning, our proposed methodology show that MediaPipe can be efficiently used as a tool to detect complex hand gesture precisely. Although, sign language modeling using image processing techniques has evolved over the past few years but methods are complex with a requirement of high computational power. Time consumption to train a model is also high. From that perspective, this work will provides us new insights into this problem. Less computing power and the adaptability to smart devices makes the model robust and cost-effective. Training and testing with various sign language datasets show this framework can be adapted effectively for any regional sign language dataset and maximum accuracy can be obtained. Faster real-time detection demonstrates the model‟s efficiency better than the present state-of-arts. In the future, the work can be extended by introducing word detection of sign language from videos using Media pipe and different algorithms.

IV. CONCLUSION AND RECOMMENDATIONS

We have developed and studied a desktop application that translates a real time video based sign language into text, we recommend that:

1- The system can be developed as android application in the future.

2- To develop the system in the future to translate texts into sign language using different algorithms and techniques.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 7, July 2025

DOI: 10.17148/IJARCCE.2025.14714

REFERENCES

[1]. Hasan, M.M., Srizon, A.Y., Sayeed, A. and Hasan, M.A.M., 2020, November. Classification of sign language characters by applying a deep convolutional neural network. In 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT) (pp. 434-438). IEEE.

[2]. Adeyanju, I.A., Bello, O.O. and Adegboye, M.A., 2021. Machine learning methods for sign language recognition: A critical review and analysis. Intelligent Systems with Applications, 12, p.200056.

[3]. Luger, G.F., 2005. Artificial intelligence: structures and strategies for complex problem solving. Pearson education.

[4]. Hi'ovská, K. and Koncz, P., 2012. Application of Artificial Intelligence and Data Mining Techniques to Financial Markets. Economic Studies & Analyses/Acta VSFS, 6(1).

[5]. Alpaydin, E., 2020. Introduction to machine learning. MIT press.

[6]. Abu-Mostafa, Y.S., Magdon-Ismail, M. and Lin, H.T., 2012. Learning from data vol. 4: AMLBook New York. NY, USA.

[7]. . Cooper, H.; Holt, B.; Bowden, R. Sign Language Recognition. In Visual Analysis of Humans; Moeslund, T., Hilton, A., Krüger, V., Sigal, L., Eds.; Springer: London, UK, 2011. https://www.ibm.com/design/ai/basics/ml/ Pedregosa F, Varoquaux G, Gramfort A,

[8]. https://link.springer.com/article/10.1007/s42979-021-00815-1

[9]. https://www.sciencedirect.com/topics/neuroscience/neural-networks

[10]. https://towardsdatascience.com/supervised-unsupervised-and-deep-learning- aa61a0e5471c.

[11]. https://www.ibm.com/cloud/learn/recurrent-neural-networks.

[12]. Siami-Namini S, Tavakoli N, Namin AS. The performance of lstm and bilstm in forecasting time series. In: 2019 IEEE International Conference on Big Data (Big Data), 2019; p. 3285–292. IEEE.

[13]. Dreuw, P., Rybach, D., Deselaers, T., Zahedi, M., and Ney, H. Speech recognition techniques for a sign language recognition system. Hand, 2007.

[14]. Arabic Sign Language Recognition, International Journal of Computer Applications (0975 – 8887) Volume 89 – No 20, 2014.

[15]. Arabic Sign Language Recognition using the Leap Motion Controller M. Mohandes, S. Aliyu and M. Deric.

[16]. B.Garcia, S. Viesca, “Real-time American Sign Language Recognition with Convolutional Neural Networks” 2016.

[17]. L. Pigou, S. Dieleman, P. Kindermans, B. Schrauwen. “Sign Language Recognition using Convolutional Neural Networks”.

[18]. Verschaeren, R.: Automatische herkenning van gebaren met de microsoft kinect (2012).

[19]. A. S. Elons ; Howida A. Shedeed ; M. F. Tolba ”Arabic sign language recognition with 3D convolutional neural networks” 2017.

