

FINAL YEAR PROJECT | DEPARTMENT OF INFORMATION TECHNOLOGY

**PROJECT SYNOPSIS
ON
INTEGRATION OF ASL DETECTION SYSTEM IN A DRIVE THROUGH**

Submitted in Partial Fulfillment of Requirements for the Award of Degree of Bachelor
of Technology in Information Technology

Submitted To



**KIET Group of Institutions,
Ghaziabad (U.P)**

Submitted By

Abhineet Sharma (1802913007)

Anupam Chopra (1802913038)

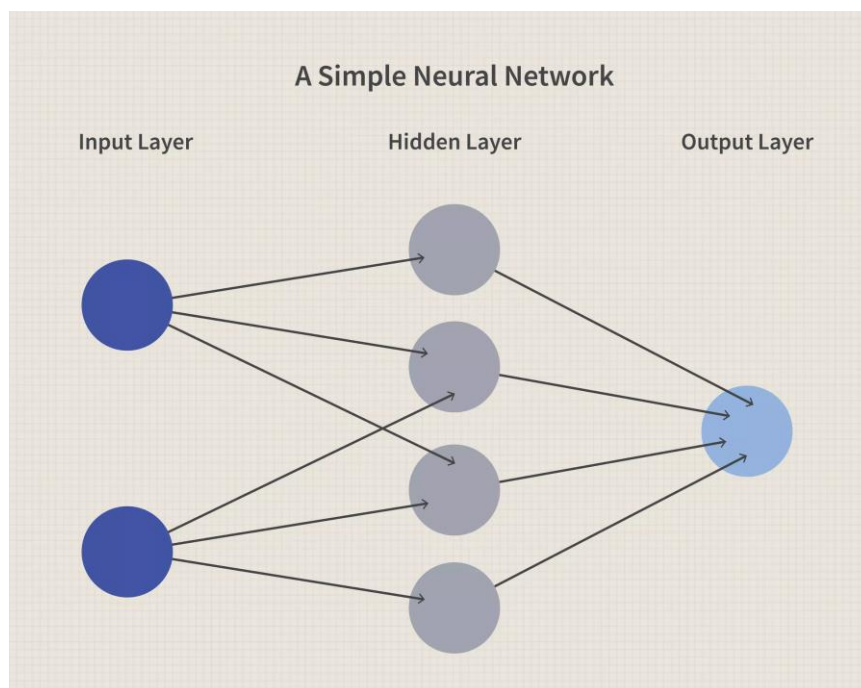
Under the Guidance of Professor Shivangi Tyagi

**DEPARTMENT OF INFORMATION TECHNOLOGY,
KIET GROUP OF INSTITUTIONS, GHAZIABAD, UTTAR PRADESH**

**(AFFILIATED TO DR. A.P.J ABDUL KALAM TECHNICAL UNIVERSITY,
LUCKNOW, UTTAR PRADESH, INDIA)**

Abstract

A neural network is a series of algorithms that endeavours to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates. In this sense, neural networks refer to systems of neurons, either organic or artificial in nature. Neural networks can adapt to changing input; so, the network generates the best possible result without needing to redesign the output criteria. The concept of neural networks, which has its roots in artificial intelligence, is swiftly gaining popularity in the development of trading systems.[1]



Source: Investopedia

The paraphernalia involves the integration of a Digital Signage with the Machine Learning Algorithm based on Neural Networks and Computer Vision.

Digital signage is a sub-segment of electronic signage. Digital displays use technologies such as LCD, LED, projection and e-paper to display digital images, video, web pages, weather data, restaurant menus, or text.

Introduction

American Sign Language (ASL) is a complete, natural language that has the same linguistic properties as spoken languages, with grammar that differs from English. ASL is expressed by movements of the hands and face. It is the primary language of many North Americans who are deaf and hard of hearing, and is used by many hearing people as well.



ASL originated in the early 19th century in the [American School for the Deaf \(ASD\)](#) in [West Hartford, Connecticut](#), from a situation of [language contact](#). Since then, ASL use has propagated widely by schools for the deaf and Deaf community organizations. Despite its wide use, no accurate count of ASL users has been taken. Reliable estimates for American ASL users range from 250,000 to 500,000 persons, including a number of [children of deaf adults](#). ASL users face stigma due to beliefs in the superiority of [oral language](#) to sign language.

ASL emerged as a language in the [American School for the Deaf \(ASD\)](#), founded by Thomas Gallaudet in 1817,^{[6]:7} which brought together [Old French Sign Language](#), various [village sign languages](#), and [home sign](#) systems; ASL was created in that situation by [language contact](#). ASL signs have a number of [phonemic](#) components, such as movement of the face, the torso, and the hands. ASL is not a form of [pantomime](#) although [iconicity](#) plays a larger role in ASL than in spoken languages. English [loan words](#) are often borrowed through [fingerspelling](#), although ASL grammar is unrelated to that of English

ASL has verbal [agreement](#) and [aspectual](#) marking and has a productive system of forming agglutinative [classifiers](#). Many linguists believe ASL to be a [subject–verb–object \(SVO\)](#) language. However, there are several alternative proposals to account for ASL word order.^[2]

Feasibility Study

Drive thru sales represent 70% of fast-food sales which generates billions of dollars for the industry each month. (The New York Times)

According to the NPD Group, a leading market research company, 57% of hamburger fast food customers use the drive-thru lane, 40% with Mexican QSRs, and 38% of chicken fast food customers went straight to drive through lanes.[\[3\]](#)

Average Drive – Through Salary Structure

Crew Member: \$13.39 per hour (40 Salaries)

Crew Leader: \$14.32 per hour (6 Salaries)

Assistant Manager: \$15.49 per hour (5 Salaries)

Maximum Shifts Allowed: 3

Minimum Work shift Hours: 20 hours per Week

Total Amount Spent per Week: \$13,979.4

Total Amount Spent per Week (PPP): \$4193.82

Amount in Indian Rupees (PPP): ₹ 306,830.36 per Week

Cost of the Proposed Drive – Through

Webcam: ₹ 12,115.00

Basic Software Development: \$30 per hour (\$4,800 Total)

Digital Signage: ₹ 4,700 per Board

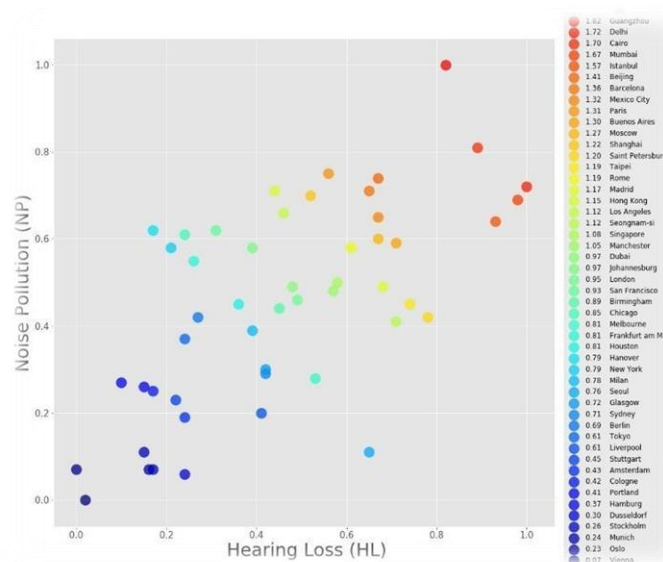
Total Amount: ₹ 368,035 (One Time Investment) [\[4\]](#)

Objective

The basic ideology behind this project was to make any customer-based utility service more friendly for the people who are disabled (deaf and mute specifically)

Since they are taught sign – language as an alternative for the standard way of interacting with others to overcome the barriers of communication, a lot of them face problems in day – to – day life like: ordering food at a “Drive – thru” or having a conversation with an attendant who is oblivious to their predicament and whose incompetence might be a bit irritating for the person who is disabled.

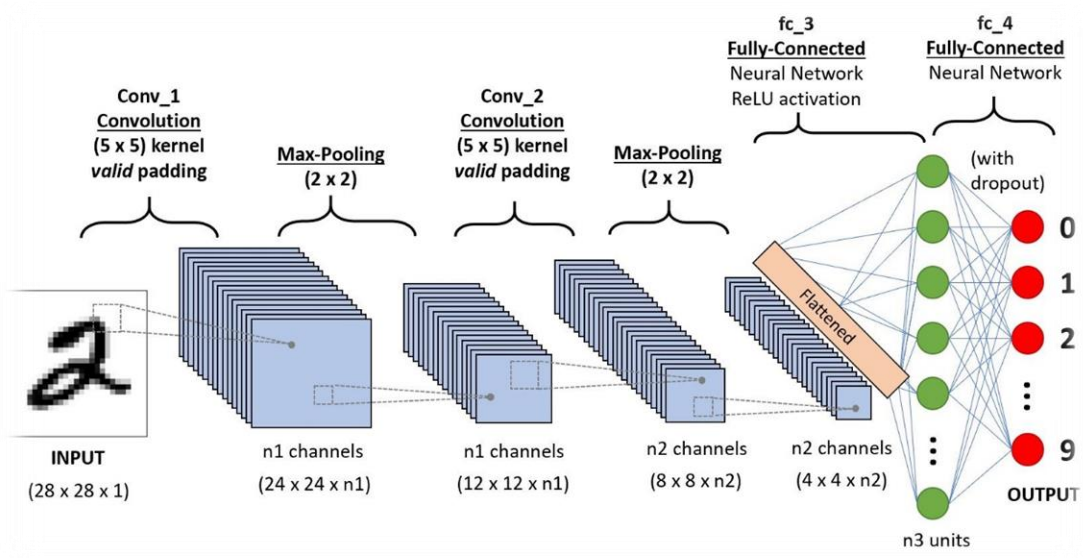
So, the ASL Detection is based on Image Processing and Object Detection integrated with a Computer webcam or any other video detecting accoutrements in order to help both; the customer and the retailer to have a better engagement whilst having a conversation.



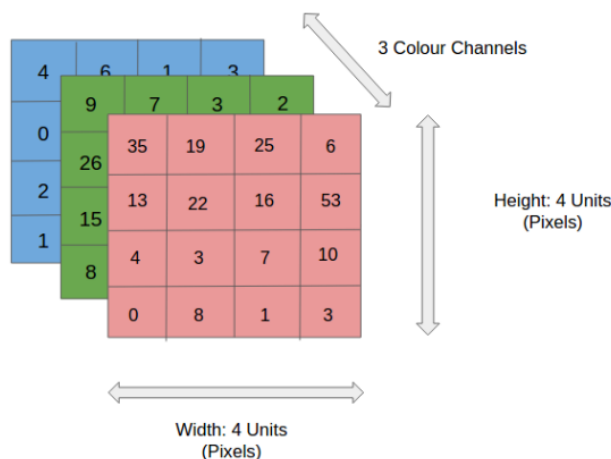
1. Around 466 million people worldwide have disabling hearing loss, and 34 million of these are children.
2. It is estimated that by 2050 over 900 million people will have disabling hearing loss.
3. 60% of childhood hearing loss is due to preventable causes.
4. 1.1 billion young people (aged between 12–35 years) are at risk of hearing loss due to exposure to noise in recreational settings.
5. Unaddressed hearing loss poses an annual global cost of US\$ 750 billion. Interventions to prevent, identify and address hearing loss are cost-effective and can bring great benefit to individuals.[5]

Technology

A **Convolutional Neural Network (ConvNet/CNN)** is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.[6]



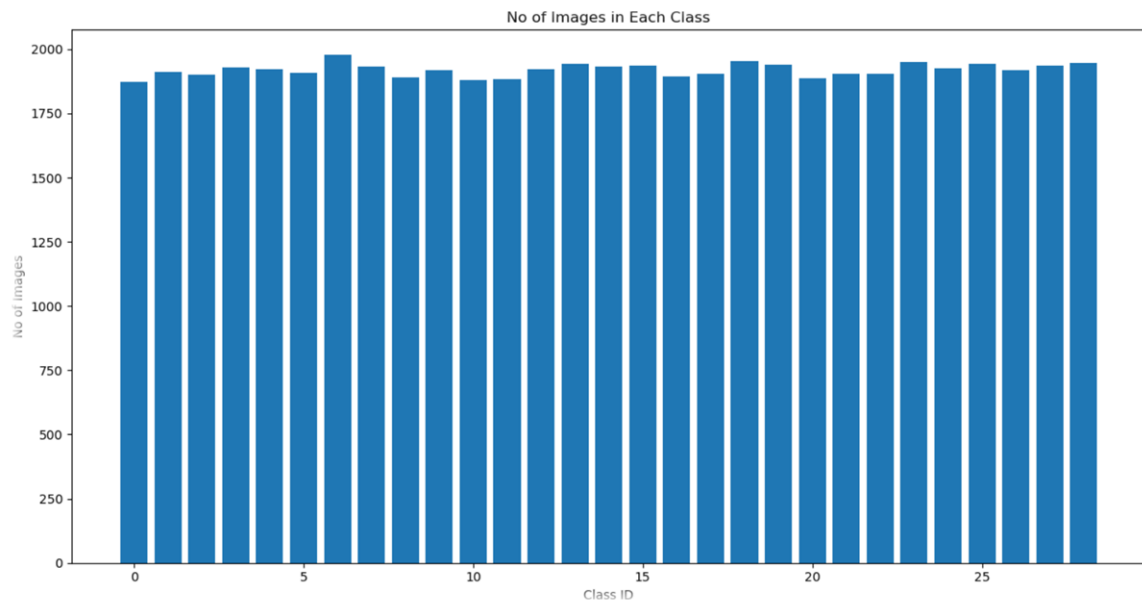
A ConvNet is able to **successfully capture the Spatial and Temporal dependencies** in an image through the application of relevant filters. The architecture performs a better fitting to the image dataset due to the reduction in the number of parameters involved and reusability of weights. In other words, the network can be trained to understand the sophistication of the image better.



You can imagine how computationally intensive things would get once the images reach dimensions, say 8K (7680×4320). The role of the ConvNet is to reduce the images into a form which is easier to process, without losing features which are critical for getting a good prediction. This is important when we are to design an architecture which is not only good at learning features but also is scalable to massive datasets

Planning of Work

The data set used in the project was downloaded from Kaggle which contains about 87,000 images and approximately 1600 images per Alphabet.



The images used in the project are split in a ratio of 80:20 for both Training and Validation Set

The Training set is one – hot encoded to convert categorical values to numeric in order to make the dataset more nuanced and generic.[\[7\]](#)

Label Encoding

Food Name	Categorical #	Calories
Apple	1	95
Chicken	2	231
Broccoli	3	50



One Hot Encoding

Apple	Chicken	Broccoli	Calories
1	0	0	95
0	1	0	231
0	0	1	50

The Model follows a sequential architecture i.e., it is built layer – by - layer using the Sequential API.

First, the model is fed a Gray – Scale Image which is pre – processed by using the cvtColor () function provided by OpenCV.

During the pre – processing stage, the image is zero –centered.[\[8\]](#)

References

1. Investopedia.com | Neural Network | By: James Chen | Reviewed By: Michael J. Boyle
2. Wikipedia | American Sign Language | Introduction
3. The New York Times | QSR
4. Quora.com | By: Tyler Hook
5. Who.int | Hearing Loss | Facts
6. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way by: Sumit Shah.
7. Text Detection using Neural Networks | OPENCV Python by Murtaza Hassan.
8. Stack Overflow | Zero – Centering an Image | By: Balraj Ashwath