International Journal For Technological Research In Engineering
Volume 7, Issue 9, May-2020 *ISSN (Online): 2347 - 4718*

SURVEY ON SIGN POSTURES RECOGNITION SYSTEM AND THE WORKING OF TENSORFLOW AND OBJECT DETECTION

Maitri Radhesh[1], Thejaswini S[2], Kanchana R K[3], Poorvika K Gowda[4]
[1,2,3,4]Students, Dept. of Computer Science and Engineering, Vidya Vikas Institute of Engineering and Technology, Mysore, Karnataka, India

Abstract: In our daily life, we come across many people who are differently abled such as deaf, dumb, blind, etc… There is a lot of difficulty that these people face tin order communicate to the other people, not of their type. For this purpose, there exists a special language called sign language, which helps them to communicate with others. But, on the contrary, normal people find it difficult to understand this language and also the hand gestures made by these people. The idea behind the project is to use the concepts of Tensor Flow, Android Studio and Artificial Neural Networks (ANN) to train the system. The Normal Camera which is present on our phones is used to take a video or images of the gestures and that will be used as input to the TensorFlow. It detects the hand gesture made by the user and displays the corresponding words on the screen, or also the output is produced as a voice. The unique methodology used in the proposed system replaces the web camera in the image processing technique by android smart phones which provide more accurate results than the previously proposed systems.
Keywords: Tensor Flow, Artificial Intelligence (AI), Flex Sensors, Machine Learning (ML), Artificial Neural Networks (ANN), Android Technology, Digital Processing, Object Detection, Java, Automatic Memory Management.

I. INTRODUCTION

Sign languages are natural languages that use different means of expression for communication in everyday life. More particularly, it is the only means of communication for the hearing impaired. Thus, it provides replacement for speech among deaf and mute people. Several research works are going on sign language in order to make the communication between a deaf person and a normal person easy. The sign language is quite a complicated language as its composed of a system of conventional gesture, mimic, hand sign and figure spellings. Hence, the main purpose was to introduce an embedded system which helps them to communicate with normal people and also to access the internet. The system converts sign language into text and speech. The main motive of the project is to bridge the gap of communication between deaf and normal people. For this purpose, a lot of research and understanding was put together to analyse what are all the technologies and techniques that are required in order the successfully complete the project.

This paper is completely based on the complete research on all the possible technologies and techniques that we can use in our project to achieve success.

II. LITERATURE SURVEY

This paper [1] highlights a virtual talk given to the specially aided people. The speaker has given a clear picture about the various hand gestures that the deaf and dumb people use and what exactly do they mean. It was an open conference which was held for all the people which could help the normal people to understand their language is a better way.

This paper [4] gives information about the sign language used by the Americans in their nation. It threw light on the small difference in gestures that the Americans used. This entire system was based on Hough Transform and Neural Networks Expert Systems along with its various Applications. We focused our attention towards the data related to NNS, its introduction, methodology, modules, etc...

In this paper [5], the focus is on one of the currently existing systems which is the use of Flex Sensors. These are the special gloves used by the specially aided people in order to communicate. As shown in Fig-1, the system takes the input figure as a value, and the controller starts to react with speech. Each flex sensor holds a unique voice which is an advantage.

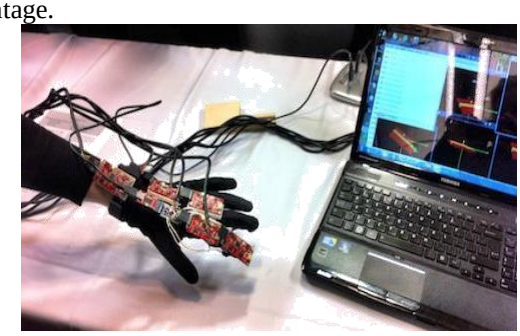


Fig- 1: Usage of Flex Sensor Gloves

The paper [6], is a set of two papers. One paper gives a gist on what kind of methodology can be expected for the gesture recognition. The other paper highlights the various literature reviews done on the various papers and what was their learning from those papers. The study of the methodology gave a huge in-site into the working of the idea.

This paper [7], gives information on the other existing system. The system takes the input figure as a value, and the controller starts to react with speech. In this system, the work is done only for some alphabets and not for the words or sentences, and the accuracy obtained is very low.

In this paper [8], an image processing algorithm is presented which is an interpretation of the American Sign Language (ASL). ASL is one of the sign languages which is used by the majority of the deaf community. The process involves the detection of hand motion, tracking the hand location based on the motion and classification of signs using adaptive clustering of stop positions, simple shape of the trajectory, and matching of the hand shape at the stop position.

This paper [9], highlights the thought that computer recognition of sign language is an important research problem for enabling communication with hearing impaired people. This paper introduces an efficient and fast algorithm for the identification of the number of fingers opened in a gesture, representing an alphabet of the American Sign Language. The concept of Finger Detection is accomplished based on the concept of Boundary Tracing and Finger Tip Detection. This system does not require the hand to be perfectly aligned to the camera or use any of the special markers or input gloves on the hand.

This paper [10], contains a survey of the various image texture analysis techniques. Three broad classes of methods are discussed namely, pixel-based, local-feature based and region-based. The pixel based models include grey level co-occurrence matrices, difference histograms and energy-measures. The local feature-based models mostly rely on edges as local features and include Marr's primal sketch model and a generalization of co-occurrence matrices. Region-based models include a region-growing model and a topographic model which treats the texture image as a digital terrain model.

This website [13] and [14], give an insight to Artificial Intelligence. Because of the research on this, we got a few insights into Machine Learning.

Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving.

Its various problems are also highlighted well.

- Reasoning, Problem Solving
- Knowledge Representation
- Planning
- Learning
- Natural Language Processing
- Perception
- Motivation and Manipulation
- Social Intelligence
- General Intelligence.

It has various applications in various fields:

- Healthcare
- Automotive
- Finance and Economics
- Cyber-Security
- Government
- Law related Professions
- Video Games
- Military
- Audit
- Advertising
- Art

In this website [15], a detailed study on Machine Learning was done.

Introduction: Machine learning (ML) is the study of computer algorithms that improve automatically through experience. It is seen as a subset of artificial intelligence. Machine learning algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as email filtering and computer vision, where it is difficult or infeasible to develop conventional algorithms to perform the needed tasks.

Machine learning is closely related to computational statistics, which focuses on making predictions using computers. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning. In its application across business problems, machine learning is also referred to as predictive analytics.

Also, the complete history and the relationships with various fields was well described.

- Relation to Artificial Intelligence: This was when they found 'neural networks'. This was built to bridge the gap between AI and ML.
- Relation to Data Mining: this keeps its focus on 'prediction properties'.
- Relation to Optimization: They focus on how to minimize the 'loss function'.
- Relation to Statistics: They were used to draw
- 'inferences' from the samples.

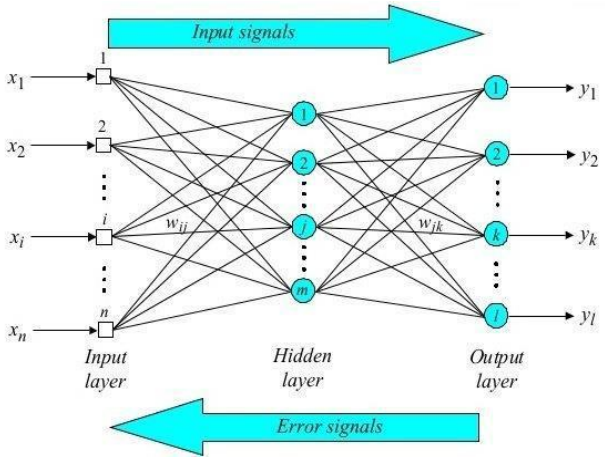
There are various types of approaches to ML, such as;

- Supervised Learning: They are built on mathematical models and their types include active learning, classification and regression.
- Unsupervised Learning: They are used for 'density estimation'.
- Semi-Supervised Learning
- Reinforcement Learning: It deals with ML applications in the real environment.
- Self-Learning
- Feature Learning: Learning happens during training.
- Sparse Dictionary Learning: It's an included feature to use sparse matrix.
- Anomaly Detection: Identification of rare items.

- Robot Learning
- Association Rule Learning: Used for understanding relationships.

The various models of ML are:

- Artificial Neural Networks: Inspired by Biological Neural Networks.



- Decision Trees: It's a predictive model
- Super Vector Machines: Related to supervised learning
- Regression Analysis: Large variety of statistical methods to estimate the relationship between input variables and their associated features. There are 3 types namely, linear, ridge and non-linear.
- Bayesian Networks: directed acyclic graphical model
- Genetic Algorithms: a search algorithm and heuristic technique that mimics the process of natural selection.

The various applications of ML are further highlighted.
This gave us an idea that there exists a technology called TensorFlow which is based on ML.
This paper [16], gives a complete insight into the complete structure of TensorFlow. It begins with for what is TensorFlow used which is followed by its history.
Background & Motivation: to deal with this, the previously existing system and their drawbacks should be well known.

The previously existing system is called 'DistBelief'. DistBelief uses the parameter server architecture. DistBelief uses the DAG structure and knowledge of the layers' semantics to compute gradients for each of the model parameters, via backpropagation. This system found its own flexible requirements such as:

- Defining new layers: For efficiency, we implemented DistBelief layers as C++ classes.
- Refining the training algorithms: Many neural networks are trained using stochastic gradient descent (SGD), which iteratively refines the parameters of the network by moving them in the direction that maximally decreases the value of the loss function.
- Defining new training algorithms: DistBelief workers follow a fixed execution pattern: read a batch of input data and the current parameter values, compute the loss function (a forward pass through the network), compute gradients for each of the parameter (a backward pass), and write the gradients back to the parameter server. This pattern works for training simple feed-forward neural networks, but fails for more advanced models, such as recurrent neural networks, which contain loops; adversarial networks, in which two related networks are trained alternately; and reinforcement learning models, where the loss function is computed by some agent in a separate system, such as a video game emulator.

TensorFlow was designed in-order to overcome all these limitations.

Design Principles: We designed TensorFlow to be much more flexible than DistBelief, while retaining its ability to satisfy the demands of Google's production machine learning workloads. TensorFlow provides a simple dataflow-based programming abstraction that allows users to deploy applications on distributed clusters, local workstations, mobile devices, and custom-designed accelerators.

The core designing principles of TensorFlow are:

- Dataflow graphs of primitive operators: Both TensorFlow and DistBelief use a dataflow representation for their models, but the most striking difference is that a DistBelief model comprises relatively few complex "layers", whereas the corresponding TensorFlow model represents individual mathematical operators (such as matrix multiplication, convolution, etc.) as nodes in the dataflow graph.
- Deferred execution: A typical TensorFlow application has two distinct phases: the first phase defines the program as a symbolic dataflow graph with placeholders for the input data and variables that represent the state; and the second phase executes an optimized version of the program on the set of available devices. By deferring the execution until the entire program is available, TensorFlow can optimize the execution phase by using global information about the computation.
- Common abstraction for heterogeneous accelerators: In addition to general-purpose devices such as multicore CPUs and GPUs, special-purpose accelerators for deep learning can achieve significant performance improvements and power savings.

Also, the related works such as single machine frameworks, batch dataflow systems and parameter servers are highlighted.

TensorFlow Execution Model: TensorFlow uses a single dataflow graph to represent all computation and state in a

machine learning algorithm, including the individual mathematical operations, the parameters and their update rules, and the input pre-processing. The entire execution is given in Fig-3.

The various dataflow elements listed are well explained.

- Tensors
- Operations
- Stateful Operations: Variables
- Stateful Operations: Queues

Also, the different methods of execution are explained in detail.

- Partial and concurrent execution
- Distributed execution
- Dynamic control flow

All these executions can be seen clearly in the Fig-2.

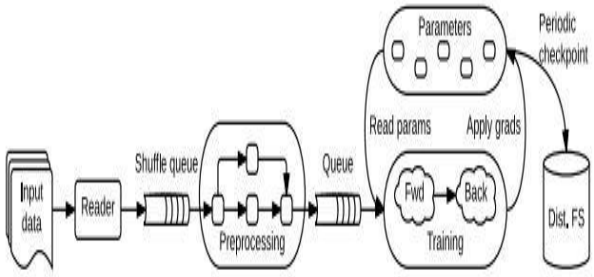


Fig- 3 A schematic TensorFlow dataflow graph for a training pipeline, containing subgraphs for reading input data, reprocessing, training, and check-pointing state

At the last, a list of case studies are briefly mentioned.

From this, we got an idea as-to what is TensorFlow exactly.

TensorFlow can help you build neural network models to automatically recognize images. These are typically Convolutional Neural Networks (CNN). There are two approaches to TensorFlow Image (Object) Recognition:

- Classification—train the CNN to recognize categories like cats, dogs, cars, or anything else. The system classifies the image as a whole, based on these categories. See our in-depth guide on TensorFlow Image Classification.
- Object Detection—more powerful than classification, it can detect multiple objects in the same image. It also tags the objects and shows their location within the image. In this article, we focus on the object detection approach in TensorFlow [20].

Also, the studies on other sites gave us an understanding on how the objects are recognized in TensorFlow [21] [22].
This website [18] gave a detailed information about what exactly Android is.

Android is a comprehensive open source platform designed for mobile devices. Android is revolutionizing the mobile space. It is a truly open Platform that separates the hardware from the software that runs on it. This allows for a much larger number of devices to run the same applications and creates a much richer ecosystem for developers and consumers. Android is a comprehensive platform, which means it is a complete software stack for a mobile device. For developers, Android provides all the tools and frameworks for developing mobile applications quickly and easily. The Android SDK is all you need to start developing for Android.
After the study on the properties of Android, we decided to use it as our platform for the project because of its flexibility and compatibility. Also for a fact that, it will be easy also at the user end to understand the application.
This paper [2], gave us a view on Digital Image Processing.
We could analyse the back-end of the project easily.
A Digital image is a representation of a two dimensional image as a finite set of digital values, called picture elements or pixels. [2]
The processing is divided into two stages called the Video Processing and the Image Processing.

Video Processing:
Video Processing is the particular case of signal processing in particular image processing which often employs video filters and where the input and output signals are video files or video streams.

Image Processing:
Image refers to a two-dimensional function f(x,y) where x and y denotes spatial coordinates and the value of f at the point (x,y) is proportional to the brightness of the image at that point. [2]
Image Processing is sub-divided into 4 stages namely: Image Capture, Pre-Processing Image, Image Segmentation and Image Classification.

- Image Capture: Capturing the image using the Android Camera.
- Pre-Processing Image: Conversion of colour captured in the image or the system point of view from grey to binary for further processing.
- Image Segmentation: Segmentation of image to such a form that it becomes convenient of the Tensor Flow to identify the image similarities.
- Image Classification: The ability to finally classify which images indicate what symbol or gesture.

This website [19], gives us information about java which we chose our coding language. Here, we got our idea on the advantages of programming in java because of its flexibility and compatibility with Android. The various in-built functions helped us carry out our work better. Its various implementations, performances, Automatic Memory Management features blew our minds off.

III. CONCLUSION

This was a complete research done in-order to understand what are the various concepts and technologies that we can use to create our project. We aimed at making the project user-friendly because of which we chose Android as our working-platform. Also, as Java was the most compatible

programming language with Android and also that it was easy to understand, we decided to use Java. With all of these, the dataset was of maximum importance and also its classification and analysis for which we chose TensorFlow. Also, for the fact that TensorFlow works on the concepts of Machine Learning and Artificial Intelligence, a detailed study on these concepts was a necessity.

REFERENCES

[1] Vikram Sharma M, Virtual Talk for deaf, mute, blind and normal humans, Texas instruments India Educator's conference, 2013.
[2] Digital Image Processing, Rafael Gonzalez & Richard E Woods, Addison-Wesley, 2002.
[3] Indian Sign Language Research and Training Centre, a boon for the deaf and dumb in India
[4] Q. Munib, et al, American Sign Language (ASL) recognition based on Hough Transform and Neural Networks Expert Systems with Applications, 2007.
[5] P. S Rajan and Balakrishnan G, Real time Indian Sign Language recognition system to aid deaf and dumb people, IEEE 13th International Conference on Communication Technologies, 2011, pp. 737-742.
[6] Students of VVIET, Mysuru, Karnataka, India, Sign Language and Gesture Recognition for Deaf and Dumb People, IEEE IRJET volume 06, issue March 2019, pp. 3399-3402, 6951-6955.
[7] J. Kim et al, Bi0-channel sensor fusion for an automatic sign language recognition in the 8th IEEE International Conference-2008.
[8] Indian Sign Language Recognition based on Histograms of Oriented Gradient, Neha V. Tavari, Prof. A. V. Deorankar
[9] Image processing system for interpreting motion in Sign Language, by C. Charayaphan A.E. Marble
[10] Finger Detection for Sign Language Recognition by, Ravikiran J, Kavi Mahesh, Suhas Mahishi, Dheeraj R, Sudheender S, Nitin V Pujari.
[11] Image Texture Analysis Techniques, by Larry S. Davis
[12] https://en.wikipedia.org/wiki/Sign_language
[13] https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp
[14] https://en.wikipedia.org/wiki/Artificial_intelligence
[15] https://en.wikipedia.org/wiki/Machine_learning#Sparse_dictionary_lea rning
[16] TensorFlow: A System for Large-Scale Machine Learning, November 2–4, 2016 • Savannah, GA, USA ISBN 978-1-931971-33-1
[17] https://www.tensorflow.org/about
[18] https://en.wikipedia.org/wiki/Android_(operating_system)
[19] https://en.wikipedia.org/wiki/Java_(programming_language)
[20] https://missinglink.ai/guides/tensorflow/tensorflow-image-recognition-object-detection-api-two-quick-tutorials/
[21] https://www.edureka.co/blog/tensorflow-object-detection-tutorial/
[22] https://www.tensorflow.org/lite/models/object_detection/overview

