ISSN: 2037-4445

Hand Gesture Recognition

[1]**Ms. Kavitha. C**
Assistant Professor
Information technology
Manakula Vinayagar Institute of Technology
Puducherry,India

[2]**Dhanuja Sri. P**
UG Scholar
Information Technology
Manakula Vinayagar Institute of Technology
Puducherry,India

[3]**Suvedha. M**
UG Scholar
Information Technology
Manakula Vinayagar Institute of Technology
Puducherry,India

[4]**Suriya Priya. K**
UG Scholar
Information Technology
Manakula Vinayagar Instiute of Technology
Puducherry,India

Corresponding author mail id: dhanujasriofficial@gmail.com

Abstract—

Over nine billion people worldwide are blind and deaf, and 2.78 percent of Indians are physically handicapped (deaf and dumb). Since that communication is a vital part of human life, this causes them several issues. The deaf-dumb have their own visual language known as sign language. Between normal and deaf or dumb individuals, sign language is a natural means of communication. Normal individuals occasionally struggle to read the signs correctly and comprehend what they are trying to express. In this paper we propose a software system using CNN to capture transform hand gestures into audio and text so that both regular people and handicapped people may understand the recognized gesture. The proposed software system takes pictures of hand gestures with a camera and pre-processes them to extract characteristics. The CNN model is then used to classify the hand movements using these characteristics. The audio output is obtained that can be heard through a speaker. This makes the project a complete solution for deaf and dumb people because commands can be wirelessly heard and displayed.

Keywords —Hand gesture, Machine Learning, CNN, Translation

I. INTRODUCTION

Hand gesture Recognition system can translate sign language into written or spoken language in real-time.

The hand gesture recognition of sign language can help individuals who are deaf or hard of hearing communicate with people who do not know sign language. It can also be useful for people who are learning sign language, as it provides real-time feedback on their signing.

Some Hand Gesture Recognition system use machine learning algorithms to improve the accuracy of the translation over time.

Overall, a Hand gesture Recognition system for sign language has the potential to significantly improve communication and accessibility for people who use sign language as their primary means of communication.

Every year, Millions of people suffer the problem of traumatic brain injuries. This project can be useful for speechless patients whose half body is paralyzed and who are not able to speak but are able to move their fingers. This system is done using a camera and CNN model in Python which is a type of computer vision application that involves recognizing and classifying hand gestures in real-time using a camera feed. This technology has a wide range of potential applications, from human-computer interaction to sign language interpretation to industrial automation.

The first step is to collect a dataset of images of hand gestures, along with corresponding labels indicating the gesture being made. The collected data is then preprocessed to prepare it for use in training the CNN model. The next step is to train a CNN model on the preprocessed dataset to minimize the training loss. By using the CNN algorithm, features are mapped and summarized by the internal three layers of the algorithm, at last classes are formed based on their features and image is predicted this is the internal function of the algorithm which gives the output as audio.

Finally, the trained model is used to perform real-time hand gesture recognition on a camera feed. This involves capturing a continuous stream of video frames from the camera, processing each frame to detect and classify any hand gestures, and outputting the results to the user.

II. RELATED WORK

Software created by Anshal Joshi, Heidy Sierra, and Emmanuel Arzuaga called "American Sign Language Translation Using Edge Detection and Cross Correlation" employs a library of visual gestures as training data for the gesture identification stage. There are two translation paradigms offered: complete words or sentences and the English alphabetic letters. Edge detection and image segmentation are used to analyse the image, and a cross-correlation coefficient-based approach is used to identify gestures. A serious flaw in the hardware is the consequence, which demonstrated susceptibility to the environment. Since a backdrop with many items and glare could add unnecessary structures during the edge detection, resulting in a misclassification of the sign, the environment must be constant [1].

The "American Sign Language Recognition Using Deep Learning and Computer Vision" by Kshitij Bantupalli and Ying Xie's. The goal of their project was to create a vision-based application that translates sign language into text to help signers and non-signers communicate. Video sequences were used in their suggested model to extract temporal and spatial characteristics. The drawback was while testing with various faces, motions, and skin tones led to a decline in accuracy. Videos have to be edited to just show neck-level motions [3].

A Virtual Sign Language Translator on Smartphones by Yun-Jung Ku, Min-Jen Chen, and Chung-Ta King, records signers and turns their hand gestures into text. It identifies the positions of the hands and the bones of the fingers by combining a classification model with two previously trained models. The categorization model is used to extract the skeleton's features, which are then linked to the specific words that are used. A significant limitation is the requirement to employ its capabilities to handle more general actions. Application of improvements is required for categorization in various datasets. The framework uses a smartphone's regular camera, which might cause problems with gesture detection [5].

By Tianming Zhao, Jian Liu, Yan Wang, Hongbo Liu, and Yingying Chen, they developed “Towards Low-Cost Sign Language Gesture Recognition Using Wearables” which used the PPG and motion sensors present in typical wearables are used in this research to differentiate between fine-grained finger motions, making it the first low-cost sign language gesture recognition system. When there are significant body movements that the traditional motion sensor-based systems cannot manage, our technology successfully recognises sign language motions. This resulted in Process Delay, Energy Consumption, Skin Tone Impact, and Impact of Intense Body Movements.[7]

A “Low-Cost Data Glove, Dynamic Gestures Recognition” by Francesco Pezzouli, Dario Corona, and Maria Letizia Corradini proposed a

in-depth examination of the Talking Hands prototype's ability to recognise dynamic motions is presented in this publication. It employs the polynomial fitting, cubic spline fitting, and fast Fourier transform (FFT) algorithms to extract the characteristics of the signs. The gestures are recognised by a variety of classifiers, some of which achieve 100% accuracy on a dataset of ten dynamic gestures. For an effective implementation in many application sectors, the gesture dataset was too tiny. This project ignored the information regarding the fingers' bending and solely used data from the IMUs [8].

Rivista Italiana di Filosofia Analitica Junior Vol 14, No. 2 (2023)
ISSN: 2037-4445

III. EXISTING SYSTEM

In Existing system, the researcher presents a paper of a smart multi-modal glove system that features on-board classification and is based on a commercially avail-able glove. It combines 16 resistive sensors spanning the hand's degrees of freedom with an accelerometer on the back of the hand. To demonstrate the capabilities of the glove, it performs rolling time-series classification of 24 letters and words from American Sign Language (ASL) [8]. The vocabulary also includes both static poses and dynamic motions. The current work builds on past research and develops an embedded learning system that leverages both strain and acceleration sensing to perform real-time pose and gesture classification.

Key contributions include sensorizing a commercially available strain-sensitive glove for ease of fabrication and adding an accelerometer, developing a neural network pipeline to detect time-series events, and conducting preliminary experiments using a vocabulary of 24 ASL words and letters. Classification performance results evaluated on unseen sessions of wearing the glove, using offline segmented examples or online rolling predictions.

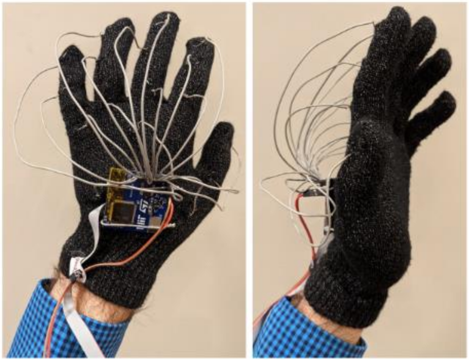

Fig4.1 Existing system

Fig 4.2 Pipeline Process to create a neural network that can classify segmented examples or streaming real-time data

In this paper the researcher increased the capabilities by adding on-board signal processing and machine learning, enhances the ability to capture dynamic motions, simplifies fabrication by using fewer sensors, and explores a new application domain.

The presented device can be a stand-alone wearable system rather than one that streams sensor data to a laptop. It also omits adding pressure-sensing infrastructure in favour of a more streamlined glove based on the commercial knit. It adds an accelerometer to capture orientations and motions relative to the world, rather than only sensing in a hand-centric frame. Finally, a new learning pipeline and experimental paradigm demonstrate applicability to real-time gesture and pose detection.

This system had several drawbacks, such as:

Limited Range of Motion: Smart gloves that use wires may have a limited range of motion due to the constraints of the wires. This can be particularly problematic for activities that require a lot of movement, such as sports or dancing.

Inconvenience: Wires can be inconvenient and cumbersome to wear, especially if they get tangled or twisted. They can also get caught on objects, which can be dangerous.

Durability Issues: The wires used in smart gloves can be prone to breaking or fraying over time, which can cause the gloves to stop functioning properly.

Complexity: Smart gloves that use wires can be complex and difficult to set up and use. They may require specialized software or hardware to work properly, which can be intimidating for some users.

Limited Flexibility: Wires may restrict the flexibility and dexterity of the wearer's hands, which can be a significant drawback for tasks that require fine motor skills.

Cost: Smart gloves that use wires can be expensive, particularly if they are designed for specialized applications such as medical or industrial use. This can make them inaccessible for many users.

IV. PROPOSED SYSTEM

In the proposed machine learning method, we create a Gesture recognition system to convert the hand action to text or audio for our research, we used the hand action Dataset, which includes both live and existing hand action images.

The proposed system of hand gesture recognition using a camera and Python in CNN is an image classification system that uses a convolutional neural network (CNN) to recognize hand gestures captured by a camera. The system involves several stages, including data collection, data pre-processing, model building, training, validation, testing, and deployment.

A. Collecting hand gestures datasets lively using camera

The first stage involves collecting a dataset of hand gesture images. This dataset can either be created from scratch or obtained from an existing database. The dataset should contain a variety of hand gestures captured from different angles and lighting conditions to ensure that the model is robust to variations in the environment.

B. Pre-processing the collected dataset

The second stage involves pre-processing the dataset. Techniques such as image resizing, normalization, and contrast stretching can be used to enhance the quality of the images.

C. Splitting the dataset to recognize hand gestures in real time

The third stage involves splitting the dataset into training, validation, and testing sets. The training set is used to train the CNN model, while the validation set is used to tune the hyperparameters of the model. The testing set is used to evaluate the performance of the model.

D. Building CNN model for the datasets

The fourth stage involves building the CNN model. A CNN model is built in Python using Jupyter Notebook. The architecture of the model can vary depending on the complexity of the hand gestures being recognized.

E. Training and validating the dataset CNN model

The fifth stage involves training the CNN model. Techniques such as data augmentation and dropout can be used to improve the performance of the model. It also involves validating the CNN model. The hyperparameters of the model, such as the learning rate, batch size, and optimizer, can be adjusted to improve the accuracy of the model.

F. Testing the developed model

Test the CNN model: The seventh stage involves testing the CNN model using the testing set. The performance of the model is evaluated using metrics such as accuracy, precision, recall, and F1 score.

G. Final system

Deploy the CNN model: The final stage involves deploying the CNN model to recognize hand gestures in real-time using a camera. This stage involves integrating the CNN model with the camera, capturing images of hand gestures, and classifying them using the trained model.

Jupyter Notebook is used throughout the process to facilitate data analysis, model building, and performance evaluation. It provides an interactive and flexible environment for data exploration and experimentation, making it an ideal tool for hand gesture recognition using a camera and Python in CNN and output is given as audio through speaker or text.

Overall, the proposed system of hand gesture recognition using a camera and Python in CNN is a complex system that involves several stages, but it has the potential to recognize hand gestures accurately and robustly in a variety of environments in real time.

V. TRANSLATING HAND GESTURE INTO AUDIO

As mentioned in the section above, our proposed system of hand gesture recognition using a camera and Python in CNN involves collecting and pre-processing a dataset of hand gesture images through camera, building, and training a CNN model, and deploying the model to recognize hand gestures in real-time and provide output through audio or text.

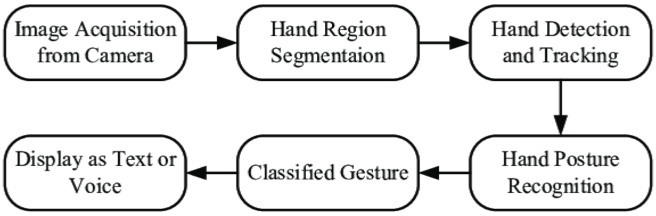

Fig. 1. Action to audio or text

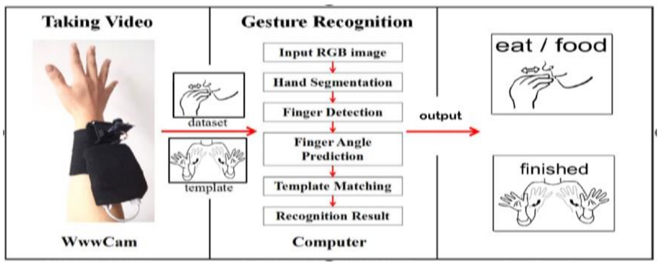

Fig. 2. Architecture Diagram of Proposed System

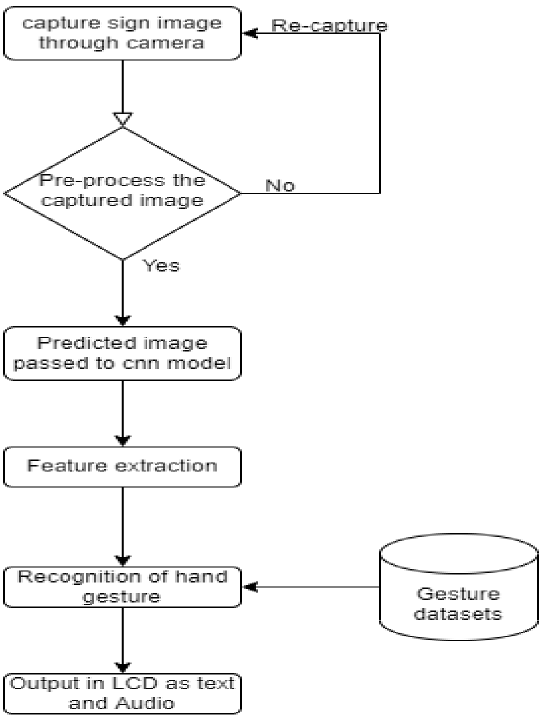

Fig. 2. Flow Diagram of Proposed System

VI. RESULTS AND DISCUSSION

In hand gesture recognition using a camera and CNN model in Python, a dataset is used to train and validate the CNN model. The dataset is a collection of images of hand gestures, where each image is labelled with the corresponding hand gesture class.

The dataset is typically divided into three subsets: a training set, a validation set, and a testing set. The training set is used to train the CNN model, while the validation set is used to tune the hyperparameters of the model, such as the learning rate, batch size, and optimizer. The testing set is used to evaluate the performance of the trained CNN model.

During training, the CNN model learns to recognize the features that are relevant for each hand gesture class. It does this by minimizing the loss function, which measures the difference between the predicted class and the true class of each image in the training set. The weights of the model are adjusted during training to minimize the loss function.

After training, the performance of the model is evaluated on the testing set. Metrics such as accuracy, precision, recall, and F1 score are used to evaluate the performance of the model. If the performance is not satisfactory, the hyperparameters of the model can be tuned using the validation set, and the training process can be repeated until the desired performance is achieved.

The dataset plays a crucial role in the performance of hand gesture recognition using a camera and CNN model in Python. A high-quality dataset with enough images for each hand gesture class is essential for achieving high accuracy in hand gesture recognition.

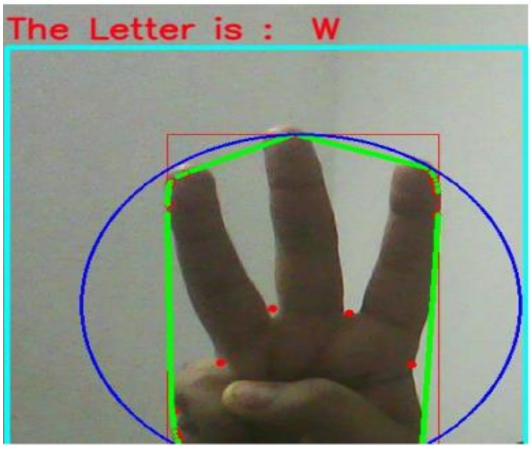

Fig. 3. Screen shot of Hand Gesture Input

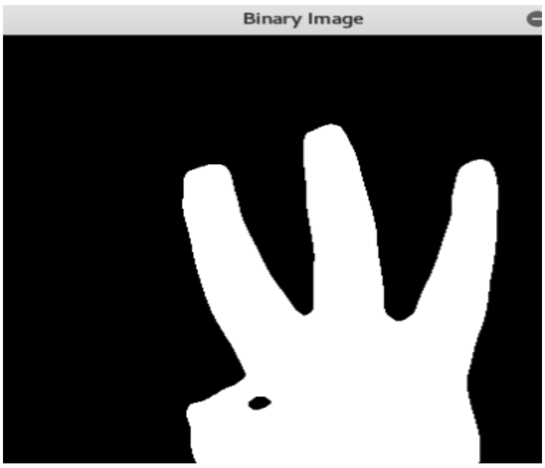

Fig 4. Binary image of Letter “W”

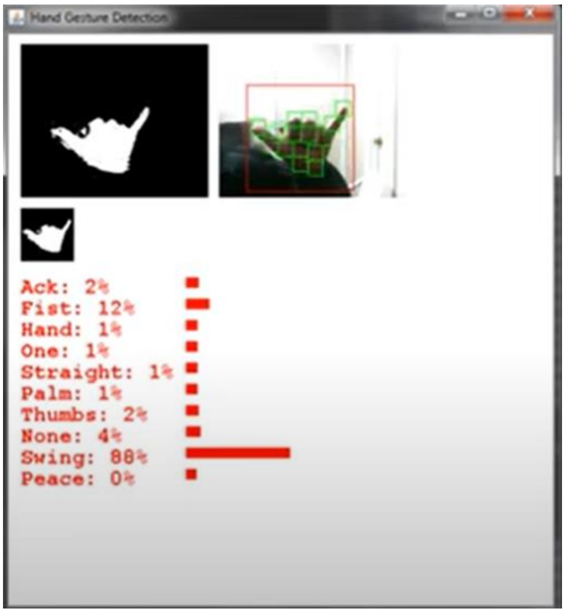

Fig. 5. UI screen shot with Hand Gesture

It is now feasible to properly recognise and categorise hand gestures in real-time by utilising computer vision and machine learning techniques. This opens new possibilities for communication, human-computer interaction, and virtual reality. Although there are still issues to be resolved, such as precise hand identification and real-time picture processing, technological and algorithmic advancements are making hand gesture recognition using a camera more practical and efficient than ever.

VII. CONCLUSIONS

This paper presents a CNN model and camera system written in Python are presented. This system is a strong and adaptable tool with a wide variety of uses. Hand gesture recognition using a camera and CNN model in Python is a reliable and flexible tool for a wide range of applications. With the use of computer vision and machine learning algorithms, hand gestures can now be accurately identified and categorised in real-time. New avenues for interaction between people and computers, virtual reality, and communication are therefore made possible. The use of a camera for hand gesture detection is now more feasible and effective than ever, despite the fact that there are still certain problems to be overcome, such as exact hand identification and real-time photo processing.

Overall, hand gesture detection utilising a camera and CNN model in Python has the potential to completely change how we engage with technology and one another.

VIII. FUTURE WORK

The proposed work of Hand Gesture Recognition can be enhanced by increasing the diversity and size of the training dataset by applying various data augmentation techniques. This can include adding noise, rotation, scaling, or changing lighting conditions to the existing dataset. Augmentation helps the model generalize better to different hand poses, lighting conditions, and backgrounds.in the system and by also providing extra details about the position and motion of the hand and enhance the precision and accuracy of gesture recognition by utilising pre-trained CNN models (like VGG net, ResNet, etc). Incorporate user interaction and feedback mechanisms into the system to refine the recognition results by allowing users to correct misclassified gestures or provide additional context to improve the system's understanding and thereby improving communication between deaf and dumb people and non-sign language speakers. Recognition of hand Gesture can also be used to operate other devices, such as mobiles, PCs, and home appliances.

REFERENCES

[1] Joshi A, Sierra H, Arzuaga E. American sign language translation using edge detection and cross correlation. In2017 IEEE Colombian Conference on Communications and Computing (COLCOM) 2017 Aug 16 (pp. 1-6). IEEE.

[2] Mahesh, M., Jayaprakash, A. and Geetha, M., 2017, September. Sign language translator for mobile platforms. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1176-1181). IEEE. -3

[3] Bantupalli, K. and Xie, Y., 2018, December. American sign language recognition using deep learning and computer vision. In *2018 IEEE International Conference on Big Data (Big Data)* (pp. 4896-4899). IEEE.

[4] Shaikh, F., Darunde, S., Wahie, N. and Mali, S., 2019, July. Sign language translation system for railway station announcements. In *2019 IEEE Bombay section signature conference (IBSSC)* (pp. 1-6). IEEE.

[5] Ku, Y.J., Chen, M.J. and King, C.T., 2019, November. A virtual sign language translator on smartphones. In *2019 Seventh International Symposium on Computing and Networking Workshops (CANDARW)* (pp. 445-449). IEEE.

[6] Ku, Y.J., Chen, M.J. and King, C.T., 2019, November. A virtual sign language translator on smartphones. In *2019 Seventh International Symposium on Computing and Networking Workshops (CANDARW)* (pp. 445-449). IEEE.

[7] Zhao, T., Liu, J., Wang, Y., Liu, H. and Chen, Y., 2019. Towards low-cost sign language gesture recognition leveraging wearables. *IEEE Transactions on Mobile Computing*, *20*(4), pp.1685-1701.

[8] Pezzuoli, F., Corona, D. and Corradini, M.L., 2020, September. Dynamic gestures recognition through a low-cost data glove. In *2020 IEEE International Conference on Human-Machine Systems (ICHMS)* (pp. 1-3). IEEE.

[9] Jeny, J.R.V., Anjana, A., Monica, K., Sumanth, T. and Mamatha, A., 2021, June. Hand Gesture Recognition for Sign Language Using Convolutional Neural Network. In *2021 5th International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 1713-1721). IEEE.

[10] Sonawane, P., Shah, K., Patel, P., Shah, S. and Shah, J., 2021, February. Speech to Indian sign language (ISL) translation system. In *2021 international conference on computing, communication, and intelligent systems (ICCCIS)* (pp. 92-96). IEEE.

[11] Tripathi, K. and Nandi, N.B.G., 2015. Continuous Indian sign language gesture recognition and sentence formation. *Procedia Computer Science*, *54*, pp.523-531.

[12] San Segundo, R., Gallo, B., Lucas, J.M., Barra-Chicote, R., D'Haro, L.F. and Fernandez, F., 2009. Speech into sign language statistical translation system for deaf people. *IEEE Latin America Transactions*, *7*(3), pp.400-404.

[13] Kim, T. and Kim, S., 2016, August. Sign language translation system using latent feature values of sign language images. In *2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)* (pp. 228-233). IEEE.

[14] Yash Jhunjhunwala, Pooja Shah, Pradnya Patil and Jyoti Waykule," Sign language to speech conversion using Arduino" , International Journal of Recent Innovation in Engineering and Research, pages 37-43, January 2017

[15] Hina Shaheen, Tariq Mehmood, " Talking Gloves: Low-Cost Gesture Recognition System for Sign Language Translation", IEEE Region 10 Symposium, 2018

[16] G.A. Rao, K. Syamala, P.V.V. Kishore, A.S.C.S. Sastry, Deep convolutional neural networks for sign language recognition, in: 2018 Conference on Signal Processing and Communication Engineering Systems (SPACES), Vijayawada, 2018,

[17] O. Koller, S. Zargaran, H. Ney, et al., Deep sign: enabling Robust statistical continuous sign language recognition via hybrid CNN-HMMs, Int. J. Comput. Vis. 126 (2018)

[18] S. Subburaj , S. Murugavalli , Survey on sign language recognition in context of vision-based and deep learning, ScienceDirect Article Journal, 2022

[19] Laura Dipietro, Angelo M. Sabatini and Paolo Dario ―A Survey of Glove-Based Systems and Their Applications‖, IEEE Transactions on Systems, Man and Cybernetics—Part C: Applications and Review, Vol. 38, No. 4, pp. 461-482, July 2008.

[20] Shoaib Ahmed.V Department of Electronics and Communication C. Abdul Hakeem College of Engineering and Technology Melvisharam, Vellore, Tamil Nadu 632 509, India submitted the thesis report of ―Hand Gesture Recognition and Voice Conversion System for Differentially Able Dumb People‖, 2012.

[21] Venkatraman.S and T.V. Padmavathi, Speech For The Disabled, Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I, IMECS 2009, March 18 - 20, 2009, Hong Kong.

[22] K. Rajeswari, E.Jeevitha, M.Meenakshi, "Voice for the voiceless", IEEE, August 2011.