Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6- Issue 2, August 2023
Paper : 39

MENTAL HEALTH TRACKER FOR DUMB USING SIGN LANGUAGE

Dr.Soma prathibha
Department of Information Technology Engineering
Sri Sairam Engineering College
Chennai, India
prathibha.it@sairam.edu.in

Jothika S
Department of Information Technology Engineering
Sri Sairam Engineering College
Chennai, India
sec20it047@sairamtap.edu.in

Dharshini S
Department of Information Technology Engineering
Sri Sairam Engineering College
Chennai, India
sec20it060@sairamtap.edu.in

Abstract **- In accordance with Statista, on average, 14% of Indians suffer from mental illness in the country. In the view of the WHO, there is no single formal definition of mental health. Mental health refers to the mental processes, sentiments, and social well-being of an individual; it influences how he or she feels and thinks, and behaves. The state of cognitive and behavioral wellness is designated by the term mental health. People without any deficiency itself not feel comfortable and not considering anyone as their comfortable place to disclose their thoughts and mental suffering. In this situation, we need to give so much care to people with deficiencies. The primary objective of our project is to monitor the mental health of people who do not talk. In this project, We designed a system that would be helpful for the caretaker of the dumb to under their sign language and convert them into text format, and with that text output, we can track the mental health of the dumb. Our proposed system focuses on converting sign language into a human-readable format using a network of convolutional neurons. The text output we got from the previous process will be analyzed to maintain the mental health of the user using NDLK classifiers. The additional thing is the text output can be converted into other languages using google trans and speech using the gTTS python library. With our system, we are able to make the unspeakable speak.**

Keywords - Sign Recognition, Mental Tracker, Alarm, Hand Gestures, Convolutional Network of Neurons.

I. INTRODUCTION

In this busy and stressful world, Everyone is suffering from mental health issues. It can be because of their surroundings, people around them, or because of their own thinking. Most people are not discussing their mental issues with their close ones though. Normal people themselves are not disclosing their mental state, it's much more difficult in the case of people with disabilities such as dumb. Our proposed system will definitely be helpful in analyzing the mental health of the dumb. Many suicides can be prevented if we are conscious of the mental health of the people around us. Sentimental analysis which is also called Opinion exploration is an approach to handling natural language that recognizes the emotional tone behind the body of the text. This sentimental analysis is useful in many fields such as social media monitoring, Customer support ticket analysis, Brand monitoring, and reputation management where it tells the organization how people feel about their products. In the same way, this sentimental analysis will be useful in detecting the mental health condition of dumb. Nowadays, Mental health issue is increasing among people. Disclosing our thoughts with our close ones is very much important to overcome this issue. In doing so one can easily recover from this problem. But the problem is that people are not mostly disclosing their problems. Our project will mainly be helpful in knowing the mental state of the people. With that knowledge, we can able to help the affected to overcome that.

II. LITERATURE SURVEY

In this literature review, we have categorized the existing systems that are developed for the wellness of dumb people.

Convolutional Neural Network

In [1], the author developed a model with a Convolutional Neural Network algorithm for predicting hand gestures of the dumb and deaf. In this model, the hand gestures of the dumb and deaf people will be taken as input, and then based on the pre-trained data set accuracy will come for the given input.

In [2], the author developed a CNN-based model for converting sign language into text format. In addition, this system also identifies emotion which is helpful in identifying

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6- Issue 2, August 2023
Paper : 39

the third component of sign language. This paper is very much helpful in constructing our model.

[1]	2022	They used a Convolutional Neural Network to recognize various types of human gestures from deaf and dumb people.	IEEE Xplore
[2]	2022	This method of converting sign language to text is based on CNN. This system focuses primarily on emotion regulation features to support the conversion of the third component of sign language recognized as a stable gesture, which is a multi-class classifier algorithm that predicts the stability of sign language gesture	IEEE Xplore

Table 2.1 Categorization of Convolutional Neural Network based Research Works

Machine Learning

In [3], the author constructed a system that is helpful for recognizing static sign languages. Here, the author compares the static sign language input with the collected dataset and finds out the matches based on the already existing data.

[3]	2020	They improved vision-based sign language recognition using Otsu's segmentation technique.	IEEE Xplore

Table 2.1 Categorization of Machine Learning based Research Works

Open CV

In [4], the author developed a system for recognizing Indian sign language gestures. The target regions are identified using OpenCV's skin segmentation feature, and the landmarks for the hand and their key points are then stored in a NumPy array. Keras, TensorFlow, and LSTM are used to teach the system. Finally, a webcam is used to evaluate the model in real time.The target areas are identified using OpenCV's skin segmentation feature, and the landmarks of hand and important aspects of landmarks are saved in an array in NumPy. Keras, TensorFlow, and LSTM are used to train the system. Finally, the model is tested in real-time using a webcam.

[4]	2022	Through deep learning models, they improved real-time sign language detection and recognition	IEEE Xplore

Table 2.2 Categorization of OpenCV based Research Works

3D-CNN

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6- Issue 2, August 2023
Paper : 39

In [5], the author worked on a real-time Chinese sign language recognition system. For the collection of the dataset, the author used the RGB camera and made a dataset consisting of 500000 video samples. The three-dimensional convolutional neural network (3D-CNN) along with optical flow processing is based on total variation regularization and L1-norm robust (TV-L1) used for improving the accuracy rate. The comparison of this system is made with the hidden Markov model (HMM) and recurrent neural network (RNN) in the work.

[5]	2021	They created a Sign Language Recognition System, which includes an interface for human interaction, a module for motion detection, detection of hand and head module, and a mechanism for video gathering..	IEEE Xplore

Table 2.3 Categorization of 3D-CNN based Research Works

Image Processing

In [6], the paper attempts to recognize 24 English alphabets. The author used image processing to extract different hand gestures from the background. Finally, the convolutional neural network model is tested on the custom data set created and real-time hand gestures performed by people of different skin tones.

[6]	2021	They used a Convolution Neural Network to test both custom and real-time hand gesture data from people of various skin tones.	IEEE Xplore

Table 2.4 Categorization of Image Processing based Research Works

Recurrent Neural Network

In [7], the author used RNN and CNN-based models for automated sign language recognition. Here, they used a combination of these two deep learning algorithms to improve the accuracy as they are using a video-based dataset. The system was developed in the manner that it first recognizes the sign language, converts them into text, and then converts them into speech format using Text-To-Speech API with python.

[7]	2021	They translated the recognized real-time video sign language data into the relevant text, then into speech for subsequent communication	IEEE Xplore

Table 2.5 Categorization of Recurrent Neural Network based Research Works

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6- Issue 2, August 2023
Paper : 39

Study Paper

In [8], the author made research work on how different technologies are helpful in sign language recognition. Based on their research work they concluded that sign-based communication is better than motion-based communication for the dumb and deaf.

The proposed system of [9] discussed different ways that can be used for sign language recognition for instance censored-based recognition and image-based recognition. They mainly talked about various image or vision-based sign language recognition systems comprising feature extraction and classification.

[8]	2020	Various SLR systems including the features of extraction and classification were described. This may be a comprehensive introduction to Automated Gesture Recognition and Sign Language Interpretation (ILS).	IEEE Xplore
[9]	2022	They provided a concise overview of the various examination work carried out to date in the area of Sign Language Recognition.	Journal of Trends in Computer Science and Smart Technology

Table 2.6 Categorization of Review Papers

Sensor-Based System

In [10], the author worked on the wearable sign language interpreter. Six imu items (IMUs) were positioned on the tips of each finger and the back of the hand to record the movements of human hands and digits as well as their orientation. They built their system using the LSTM method, which is an artificial neural network (ANN).Six IMU were placed on the back of the hand palm and each fingertip to record the movements and orientation of human hands and fingers. They built their system using an LSTM algorithm, which is an artificial neural network.

[10]	2020	To get over the communication barrier for the deaf and dumb, they discovered a portable ASL interpreter.	IEEE Xplore

Table 2.7 Categorization of Sensor-Based Research Works

III. METHODOLOGY

In our proposed solution, we have collected video data collection of 20,000 video frames. This dataset includes all their hand gestures, alphabets, and numbers. In Pre-processing phase, we split the video into frames, removed all noise, and then convert it to gaussian blur format. After training in Convolutional Neural Networks, it will display the sequence of prediction models for each frame extracted from the video sequence of gestures. Here in RNN, its extension LSTM will be trained using the sequence of models generated by CNN for feature representation. LSTM will be used to train these features. Long-Short Term Memory generates a Model file that will be used to predict temporal images. Using GoogleTrans and gTTS, the output text can be in multiple languages as well as in speech format. We can track their mental health using that output text, such as anger, sadness, and happiness.

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6- Issue 2, August 2023
Paper : 39

If the person is not feeling well, the alarm notification is displayed to notify the caretaker of their mental health. To keep the database of their daily dialogue in text format up to date. This will be useful for tracking their mental health on a regular basis. We make it an Android application, integrating the deep learning module, and going to launch it on the Google Play store.

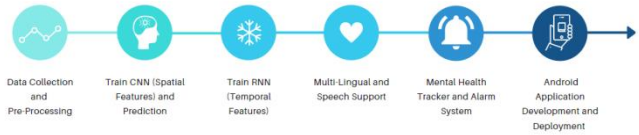


Fig 3.1 Workflow chart

Application working process

The user can enable camera access and point their camera towards the dumb people. It recognizes their hand gesture and then shows the result to them as text and speech. They can also track their mental health progression using this application.

IV. RESULT AND DISCUSSION

To enhance the image quality, we use media pipe and OpenCV to identify the motion of the hand in the video dataset. By using OpenCV, we apply various filters and make some comparisons on them. Among various filters, Gaussian Blur is finalized and that can be applied to all frames for training. Because it highlights the low color intensity into high and blurs the background. The filter comparison results are shown below

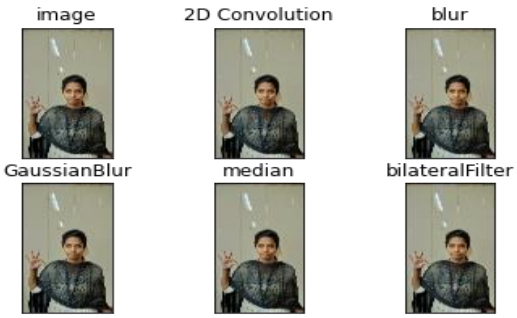


Fig 4.1 Filters Differentiation

As of now, we completed the training for the CNN-LSTM module. We got a training accuracy of 99.98%. The layers of the CNN are shown below

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6- Issue 2, August 2023
Paper : 39

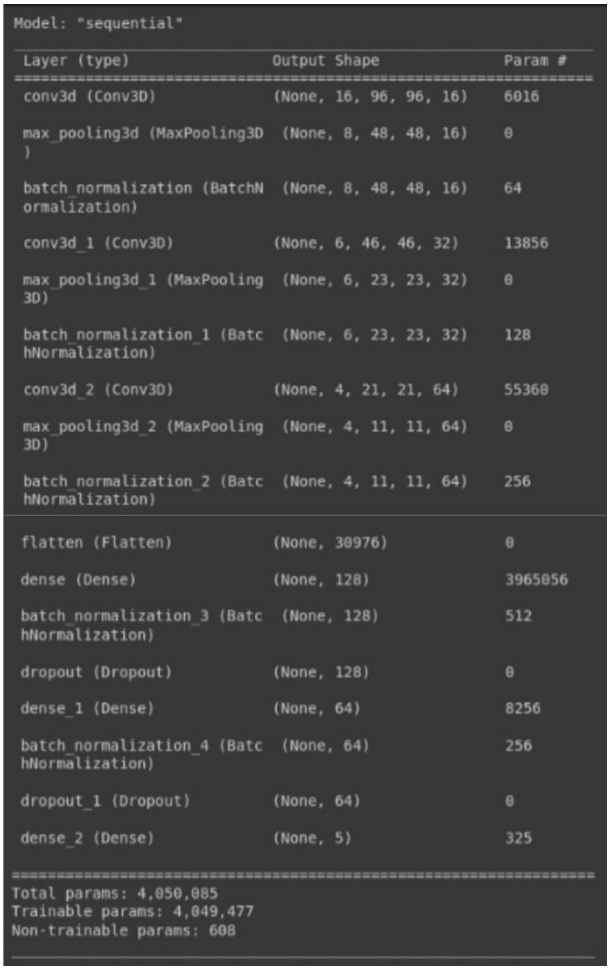
```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 148, 148, 100)     2800

max_pooling2d (MaxPooling2D  (None, 74, 74, 100)       0
)

conv2d_1 (Conv2D)            (None, 72, 72, 100)       90100

max_pooling2d_1 (MaxPooling  (None, 36, 36, 100)       0
2D)

flatten (Flatten)            (None, 129600)            0

dropout (Dropout)            (None, 129600)            0

dense (Dense)                (None, 50)                6480050

dense_1 (Dense)              (None, 2)                 102

=================================================================
Total params: 6,573,052
Trainable params: 6,573,052
Non-trainable params: 0
_________________________________________________________________
None
```

Fig 4.2 CNN Module Diagram

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv3d (Conv3D)              (None, 16, 96, 96, 16)    6016

max_pooling3d (MaxPooling3D  (None, 8, 48, 48, 16)     0
)

batch_normalization (BatchN  (None, 8, 48, 48, 16)     64
ormalization)

conv3d_1 (Conv3D)            (None, 6, 46, 46, 32)     13856

max_pooling3d_1 (MaxPooling  (None, 6, 23, 23, 32)     0
3D)

batch_normalization_1 (Batc  (None, 6, 23, 23, 32)     128
hNormalization)

conv3d_2 (Conv3D)            (None, 4, 21, 21, 64)     55360

max_pooling3d_2 (MaxPooling  (None, 4, 11, 11, 64)     0
3D)

batch_normalization_2 (Batc  (None, 4, 11, 11, 64)     256
hNormalization)

flatten (Flatten)            (None, 30976)             0

dense (Dense)                (None, 128)               3965056

batch_normalization_3 (Batc  (None, 128)               512
hNormalization)

dropout (Dropout)            (None, 128)               0

dense_1 (Dense)              (None, 64)                8256

batch_normalization_4 (Batc  (None, 64)                256
hNormalization)

dropout_1 (Dropout)          (None, 64)                0

dense_2 (Dense)              (None, 5)                 325

=================================================================
Total params: 4,050,085
Trainable params: 4,049,477
Non-trainable params: 608
_________________________________________________________________
```

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6- Issue 2, August 2023
Paper : 39



```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv3d (Conv3D)             (None, 16, 96, 96, 16)    6016

 max_pooling3d (MaxPooling3D  (None, 8, 48, 48, 16)    0
 )

 batch_normalization (BatchN  (None, 8, 48, 48, 16)    64
 ormalization)

 conv3d_1 (Conv3D)           (None, 6, 46, 46, 32)     13856

 max_pooling3d_1 (MaxPooling  (None, 6, 23, 23, 32)    0
 3D)

 batch_normalization_1 (Batc  (None, 6, 23, 23, 32)    128
 hNormalization)

 conv3d_2 (Conv3D)           (None, 4, 21, 21, 64)     55360

 max_pooling3d_2 (MaxPooling  (None, 4, 11, 11, 64)    0
 3D)

 batch_normalization_2 (Batc  (None, 4, 11, 11, 64)    256
 hNormalization)

 flatten (Flatten)           (None, 30976)             0

 dense (Dense)               (None, 128)               3965056

 batch_normalization_3 (Batc  (None, 128)              512
 hNormalization)

 dropout (Dropout)           (None, 128)               0

 dense_1 (Dense)             (None, 64)                8256

 batch_normalization_4 (Batc  (None, 64)               256
 hNormalization)

 dropout_1 (Dropout)         (None, 64)                0

 dense_2 (Dense)             (None, 5)                 325

=================================================================
Total params: 4,050,085
Trainable params: 4,049,477
Non-trainable params: 608
_________________________________________________________________
```

Fig 4.3 CNN-LSTM Module Diagram

V. CONCLUSION

In this busy world, everyone is very stressed and has mental pressure. Our proposed system is used to track the mental health of dumb. This will be upgraded with some features in the future as available for physically-challenged people. We hope this will be the better solution for tracking the mental health of the dumb. Sign language is converted into text and it can be available in many languages and also in speech. So it can be helpful for the person to understand what he/she speaks. Even the illiterate can also use this application very easily because of the speech availability.

VI. REFERENCES

[1] Thylashri, S., et al. "Human Sign Language Recognition Using Deep Learning." 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES). IEEE, 2022.

[2] Gupta, Vibhu, Mansi Jain, and Garima Aggarwal. "Sign Language to Text for Deaf and Dumb." 2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2022.

[3] Datta, Pushan Kumar, et al. "Creation of Image Segmentation Classifiers for Sign Language Processing for Deaf and Dumb." 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO). IEEE, 2020.

[4] Deep, Aakash, et al. "Realtime Sign Language Detection and Recognition." 2022 2nd Asian Conference on Innovation in Technology (ASIANCON). IEEE, 2022.

[5] Zhao, Kai, et al. "Real-time sign language recognition based on a video stream." International Journal of Systems, Control, and Communications 12.2 (2021): 158-174.

[6] Raval, Jinalee Jayeshkumar, and Ruchi Gajjar. "Real-time sign language recognition using computer vision." 2021 3rd International Conference on Signal Processing and Communication (ICPSC). IEEE, 2021.

[7] Sonare, Babita, et al. "Video-based sign language translation system using machine learning." 2021 2nd International Conference for Emerging Technology (INCET). IEEE, 2021.

[8] Vaidhya, G. K., and C. A. S. Preetha. "A Comprehensive Study on Sign Language Recognition for Deaf and Dumb people." Journal of Trends in Computer Science and Smart Technology 4.3 (2022): 163-174.

[9] Nimisha, K. P., and Agnes Jacob. "A brief review of the recent trends in sign language recognition." 2020 International

Journal of Current Research in Engineering and Science
Bi-Annual Online Journal (ISSN : 2581 - 611X)

Volume 6- Issue 2, August 2023
Paper : 39

Conference on Communication and Signal Processing (ICCSP). IEEE, 2020.

[10] Chong, Teak-Wei, and Beom-Joon Kim. "American sign language recognition system using wearable sensors with the deep learning approach." The Journal of the Korea institute of electronic communication sciences 15.2 (2020): 291-298.