International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0217
e ISSN: 2584-2854
Volume: 03
Issue: 04 April 2025
Page No: 1335 - 1341

Two Way Sign Language Detection and Communication System

Nandhika B[1], Preethika L[2], Sanjeetha P M[3], Sarumathi S[4], Sahana S[5], Dr. D Nithya[6]
[1,2,3,4,5]Student, Computer Science and Engineering, School of Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India.
[6]Associate Professor, Computer Science and Engineering, School of Engineering, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore, Tamil Nadu, India.
Email ID: *21106ueo48@avinuty.ac.in*

Abstract
The project focuses on developing an innovative travel website designed to enhance user experience through personalized travel plans and comprehensive budget management. The platform allows users to input their desired destination, trip duration, and estimated budget, generating a detailed day-by-day itinerary that covers activities, accommodations, dining, and travel suggestions between cities. Incorporating advanced algorithms like Dijkstra's Algorithm for finding the shortest path between cities, Kalman Filters for improving GPS data accuracy and providing smoother route transitions, and the Knapsack Problem Algorithm for selecting the best combination of travel activities within budget constraints, the platform ensures efficiency and value. Users can choose from pre designed packages or customize every aspect of their travel plans. Real-time expense tracking enables accurate budget management, and the platform provides a clear summary of spending at the end of the trip. This project aims to offer a seamless, tailored travel planning experience, helping users create memorable trips while maintaining control over their budget.
Keywords: *Budget Management; Customized Travel Plan; Enhanced User Experience.*

1. Introduction
Fifteen percent of people worldwide have some kind of disability. 466 million individuals, or more than 5% of the population, are deaf. The World Health Organization (WHO) estimates that this might increase to 500 million people by 2050, which would be around 2.7 times the population in 2000. Speech and hearing impairments affect at least 70 million people [1]. It can have a significant impact on a person's communication skills. In order to help impaired people communicate with the outside world and improve their independence and quality of life, we can employ the Sing Language Recognition (SLR) technology. Although the public may not understand sign language, disabled persons utilize it to communicate their thoughts and feelings to regular people [2]. Industry 5.0, the latest evolution of industrial revolutions, represents a paradigm shift in manufacturing and production, intertwining cutting-edge technologies with human ingenuity [3]. While Industry 4.0 focused on automation and digitization, Industry 5.0 takes a step further by emphasizing the collaboration between humans and advanced technologies. In this context, our research in real-time SLR holds significant relevance. By developing a Convolutional Neural Network (CNN)-based model capable of detecting and interpreting sign Real-time linguistic gestures have brought us into line with the fundamental ideas of Industry 5.0, which holds that technology not only simplifies procedures but also fosters inclusivity and interpersonal communication. In line with Industry 5.0's goal of utilizing technology to improve human skills and promote a more inclusive, effective, and cooperative industrial landscape, our research helps individuals with speech and hearing impairments communicate more effectively [4]. Our application is a concrete illustration of this collaborative and human-centered approach, improving communication accessibility and fostering a more inventive and socially conscious industrial landscape as Industry 5.0 envisions a future in which technology acts as a facilitator rather than a substitute. A smooth cooperation between humans and machines is encouraged by Industry 5.0. Within the framework of

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0217
e ISSN: 2584-2854
Volume: 03
Issue: 04 April 2025
Page No: 1335 - 1341

our suggested system, it emphasizes. The contribution of this research lies in its ability to enhance the standard of living for those with disabilities by giving them access to a larger audience for communication, there by promoting inclusion and equality. Additionally, the research paves the way for the development of assistive technologies that make everyday activities such as education, employment and social interactions more accessible and equitable for people With disabilities by using the sign language as the primary form of communication. The suggested technique trains a CNN using a sizable dataset of SLR (which we gathered from Kaggle [5] and our own surveys). One kind of a deep neural network for visual data interpretation is a CNN [6]. CNN uses body language and sign language to extract features. For those with distinct sign patterns, this increases recognition accuracy. Additionally, the system has a monitoring module that use computer vision techniques to identify environmental changes and react appropriately [7]. Accuracy, reaction speed, and user happiness are just a few of the indicators we used to assess the suggested system's effectiveness. We have shown our system's efficacy in enabling impaired people by contrasting it with other methods. To communicate with other people and environment. In this paper, we have also completed the sign-to-text conversation with the accuracy of over 90% [8].

2. Literature Review

Researchers have used a variety of methods in SLR systems. To address the issue of sign language gesture recognition, the creators of [2] employed a number of strategies. In order to understand sign language gestures, transfer learning approaches modify pre-trained models from other fields, such as object identification, face recognition, natural language processing, etc. This method aids in lowering the quantity of labeled data needed for SLR model training. Large datasets are necessary for hyper-tuned CNNs to function at their best. However, the size and diversity of SLR datasets are frequently constrained, which can result in poor generalization and overfitting of the model. Researchers can expand their dataset's size and diversity by employing data augmentation techniques like image rotation and mirroring. A camera could be used by the system to record Three models were utilized by the authors in [6] for classification: 3D CNN, CNN and LSTM combined, and object detection based on the YOLO v5 algorithm for hand gesture detection. These models demonstrate the outstanding recognition of dynamic gestures by the object detection model. However, The CNN model's accuracy is 82%, and they utilized less Dataset to finish their work. Furthermore, the models were unable to assist the blind. It will be able to assist blind individuals if it is transformed to speech recognition. The Bag of Visual Words (BOVW) concept was used by the authors of [8] to demonstrate how to recognize Indian sign language letters and numbers. In addition to histogram-based sign mapping, they employed segmentation based on skin color and background subtraction. Lastly, the categorization was done using CNN and Support Vector Machine (SVM). For convenience, they also created their Graphical User Interface (GUI). This study employed a proprietary dataset to recognize Indian sign language from over 36,000 images. Binary edges on the dataset mask and smart edges were created, and Speeded Up Robust Features (SURF) was used to extract features. They obtained accuracy rates of 99.17% and 99.64%, respectively, using SVM and CNN [9]. The CNN approach was utilized in [10] to identify sign language in order to communicate with others. The writers would probably explore the wide range of assistive technology and how they affect the accessibility of communication. They can go over how SLR systems have changed throughout time, examining the advantages and disadvantages of different strategies used by scholars in this area. They could get important knowledge about the effectiveness of various methods by looking at earlier research projects; they might even be able to highlight developments like Deep Learning and Neural Networks (NN), which have completely changed the accuracy of SLR systems. 8,958 photos were used in the testing. Their research has an accuracy rate of 89.1%. The SVM model was employed by the authors of [11]. The literature review might also explore the topic of datasets utilized in studies on gesture recognition, emphasizing the necessity of extensive and varied

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0217
e ISSN: 2584-2854
Volume: 03
Issue: 04 April 2025
Page No: 1335 - 1341

datasets designed for particular scenarios, such as emergencies. By examining how these datasets have been used in research on gesture recognition, communication support, and emergency response, they may make comparisons with datasets from other sign languages. The authors might highlight how their video dataset addresses a significant need in the field and clarify the originality of their contribution by reviewing previous research. For testing, they used 824 pictures. Their model's accuracy is 90%.

The SLR was presented by the authors in [12] in an effort to close the gap in communication between non-hearing and hearing-impaired individuals. The input was extracted from the library of Indian sign language gestures or ROBITA and preprocessing was done to get rid of unnecessary artifacts. The literature study would most likely make links between earlier studies that used CNNs for the SLR. This could entail going over various model setups, datasets, and notable findings from these investigations. 356 photos were used in their study. They used an ML-CNN with encoder and achieved an accuracy of 87.50%. The writers of [13] talked about how sign language facilitates communication between the general public and those who are hard of hearing. For continuous gesture sequences, or continuous SLR, they offered a modified LSTM model. They were able to recognize a number of related gestures during their investigation. They used 3,150 photographs in total, and their accuracy rates were 89.5% and 72.3%, correspondingly.

3. Research Methodology

The work's main goal is to recognize different sign actions. Our goal is to create a CNN model that can accurately identify sign behaviors while maintaining minimal parameter values and high accuracy. Ultimately, the technology is able to identify and convert hand gestures into sign language. The ensuing subsections contain information about data collecting, preprocessing, training, etc.

3.1 Data Collection

Data from nine distinct classes has been gathered by us. We first gathered hand gesture data for training, after which we scaled every picture. Prior to using them for training, we resized the original photos from big pixels to 190×190 pixels. [6] and [9] are the sources from which the testing data is chosen. Nine distinct classes—Washroom, Call, Food, Happiness, Helping, Medicine, Pain, Sadness, and Thief—were represented by those images. A few sample photos from various classes in the dataset is displayed in Figure 1. Different lighting conditions, weather conditions and times of day were used to get the photos.

3.2 Data Description

First, we had to learn the types of sign languages that individuals with disabilities used to communicate. Once we understood, we gathered and saved the pictures. We have gathered the pictures from many sources. As shown in Table I, our Figure 1: Various Classes of Images from the Dataset

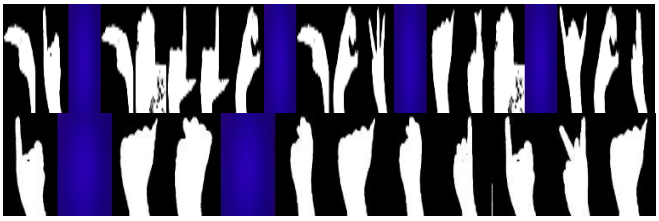

Figure 1 Various Classes of Images from The Dataset

Dataset has a total size of 6,182. We resized the photos after gathering them. To assist the deaf and disabled, we have employed sign languages to help the disabled people who are deaf.

Table 1 Training and Testing Images Count for Each Class

Class Name	Total Images	For Training	For Validation
Call	704	560	112
Food	945	655	159
Happiness	602	389	153
Helping	679	514	105
Washroom	584	395	129
Medicine	807	608	157
Pain	624	460	165
Sadness	510	387	113
Thief	727	639	133
Total	6,182	4,607	1,226

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0217
e ISSN: 2584-2854
Volume: 03
Issue: 04 April 2025
Page No: 1335 - 1341

3.3 Models

We have mostly used supervised Machine Learning (ML) techniques for the inquiry. In our research, we have employed the CNN model.

3.3.1 Data Preprocessing

TensorFlow and Keras were utilized for the back-end development in this paper. Our RGB input photographs have been transformed to grayscale. We have Data collection for CNN model training was the first step in the inquiry. After that, hashing was used to find and eliminate duplicate photos from the dataset as part of the data cleansing process, creating a clean dataset. After that, this dataset was divided into training and test sets. Techniques for preparing data are used to improve their quality and training appropriateness. To avoid biases brought on by differences in the data distribution, data normalization has been done. The CNN model was created with The evaluation criteria include the Adam optimizer, accuracy, and the categorical cross entropy loss function. The model was trained using a batch size of 64 and 40 epochs, with 20% of the data being used for validation. The model's performance has been assessed following training on a distinct test dataset to evaluate its capacity for prediction.

3.3.2 Data Normalization

ase the consistency of the learning algorithm's performance, data normalization in machine learning refers to the act scaling and transforming the input data to a standard range or distribution. Normalization of the training data set should be followed by the same change for the validation and test data sets. This ensures that the learning process is not distorted by the distribution of the data and that the normalization is the same for all data subsets. Equation 1 contains the formula that we used for normalization [13].

3.3.3 Data Training and Validation

We have sent the preprocessed input photos to our suggested model for training and testing after completing all of the previously mentioned processes. The probability that the image will fit into one of the categories is determined by the prediction layer. Consequently, the output is normalized, and the sum of all the values in every class is equal to 1. We used the Sigmoid and Relu functions to achieve this.

3.4 Proposed Methodology

In order to train the CNN models, the study started with data collection. After that, a clean dataset was produced by employing hashing to find and eliminate duplicate photos from the dataset. After that, this dataset was divided into training and test sets. To improve the quality and usefulness of the data for training, preprocessing techniques are used. To avoid biases brought on by differences in the data distribution, data normalization has been done.

The CNN model was created with accuracy, categorical cross-entropy loss function, and the Adam optimizer as evaluation criteria. A batch size of 64 and 40 epochs were used to train the model, and 20% of the data was used for validation. Following training, the model's predictive power was assessed by evaluating its performance on a different test dataset.

3.5 Model Building

We have built a CNN model that is sequential. In our model, every layer has a single input and output. The whole network is created by stacking each of these layers on top of the others. The system is founded on a vision. Since naked hands are used to depict each sign, it removes the issue with interacting with any artificial technologies. The suggested CNN model is displayed in Figure 2. The suggested strategy uses a deep CNN model, which is a type of feed forward NN, to change the network's parameters and lower the value of the cost function [14]. There are 199,799 parameters in the constructed model. The suggested SLR architecture is displayed in Figure 3.

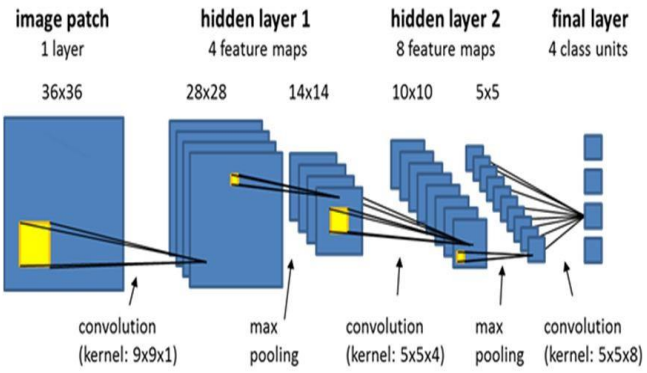

Figure 2 Proposed CNN Model

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0217
e ISSN: 2584-2854
Volume: 03
Issue: 04 April 2025
Page No: 1335 - 1341

This study used the convolution layer, which has a kernel size of three and a variety of characteristics. Rely is the activation function that has been chosen. It also makes advantage of the Soft Max function. In the Two dropout layers were employed in the suggested model. It is used to filter data as it moves forward over the network. We used the Max Pooling technique to compress the feature map for optimization. It determines the maximum or greatest value present in every feature map patch. Here, a flattening layer is used to reduce multidimensional input to one dimension. The thick layer is then used to generate the completely connected layers. The majority of the training choices were kept at their default defaults, except for batch size (64), drop rate (0.2, 0.5), epochs (40), validation percentage (20%), and testing percentage (80%). Because there weren't many training photos, we increased the training batch size from the standard 132 to 32. How many iterations were made depends on how many photographs were used in a batch? We have included the complete validation dataset in the validation batch size because additional training rounds will ultimately produce more dependable results.

3.6 Model Summary

The suggested model's overall overview is provided in Figure 4. Details of three Convolution layers, three Max-pooling layers, three Dropout layers, one Flatten layer, two Dense layers, shape and parameters are presented in this figure. We have compared to other well-known models, this one uses fewer layers 11 in all. It took roughly four minutes per epoch in the Jupyter Notebook to train our suggested CNN model using the images.

4. Experimental Results and Analysis

The results of the suggested approach are explained in this part along with a comparison to previous models. Using the chosen features from the feature selection procedure, we have assessed the suggested models since they performance that was somewhat improved.

4.1 Evaluation

After training the model, we assessed its implementation. As shown in Figure 5, the validation accuracy is 0.9167 and the training accuracy is 0.9917. Figure 6 illustrates the training and validation loss graphs, indicating a decline in both training and validation loss.

4.2 Comparative Results

The effectiveness of the classification model has been evaluated using the CNN technique. A comparison between the suggested model and the current models is given in Table II. We have compared the detection, dataset size, accuracy, etc. Table II demonstrates how the suggested model performs better than the current ones. Particularly when considering systems designed to help those with disabilities, the suggested system's accuracy of 91.67% is quite astounding. Technology can help many disabled persons reclaim their independence. A system with high accuracy can comprehend and react to their demands or directions efficiently, enhancing their quality of life. This can involve activities like communicating, using the Internet to get information, or even controlling smart home appliances. In these circumstances, precision is essential to guarantee a prompt and trustworthy reaction when required. For those with disabilities, high-precision assistive technologies can enhance social interactions and encourage inclusion. Communication with friends and family, access to job and educational possibilities, and involvement in all facets of society can all be facilitated by this.

4.3 Real-Time Sign Language Detection

By implementing the proposed method, the real-time detection has been able to detect the sign actions correctly. Figure 7 shows some demos from which we can conclude that we are

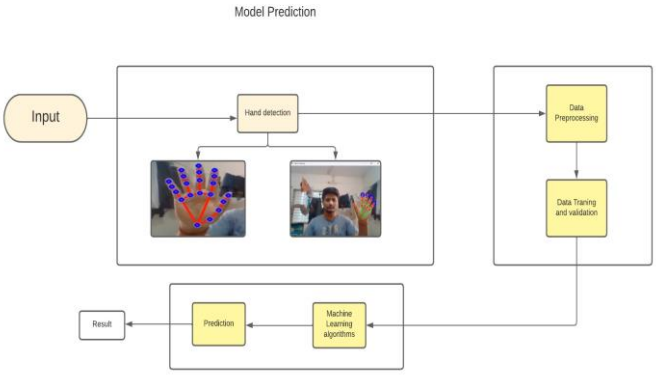

Figure 3 Proposed Architecture for the Sign Language Recognition

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0217
e ISSN: 2584-2854
Volume: 03
Issue: 04 April 2025
Page No: 1335 - 1341

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 95, 95, 90)        2520

 max_pooling2d (MaxPooling2D  (None, 47, 47, 90)       0
 )

 conv2d_1 (Conv2D)           (None, 24, 24, 80)        64880

 max_pooling2d_1 (MaxPooling  (None, 12, 12, 80)       0
 2D)

 conv2d_2 (Conv2D)           (None, 6, 6, 70)          50470

 max_pooling2d_2 (MaxPooling  (None, 3, 3, 70)         0
 2D)

 dropout (Dropout)           (None, 3, 3, 70)          0

 flatten (Flatten)           (None, 630)               0

 dense (Dense)               (None, 128)               80768

 dropout_1 (Dropout)         (None, 128)               0

 dense_1 (Dense)             (None, 9)                 1161

=================================================================
Total params: 199,799
Trainable params: 199,799
Non-trainable params: 0
```

Figure 4 Convolution Layers, Three Max-Pooling Layers, Three Dropout Layers

Table 2 Comparison with the Existing Models

Existing Works	Model Name	Accuracy	Image	Detection
[6]	CNN	82.00%	2,060	Yes
[9]	CNN	89.1%	8958	No
[10]	SVM	90.00%	824	Yes
[11]	CNN	87.50%	356	No
[12]	LSTM	89.5%	3150	No
Proposed Method	CNN	91.67%	6,182	Yes

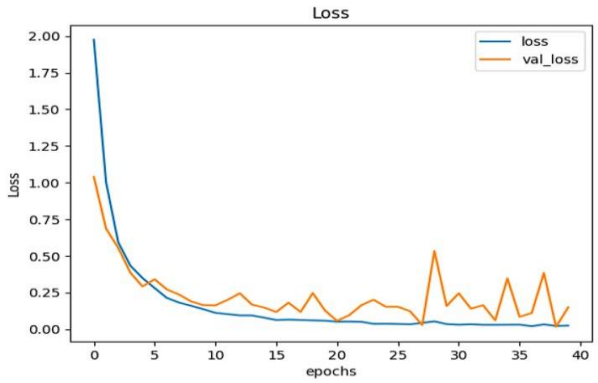

Figure 5 Evaluation of the Training and Validation Loss

5. Summary of the Proposed Model

Able to clearly detect the signs. In the context of sign language recognition, it may mean that the model overfits the specific data in the training set and does not generalize very well for the new signs.

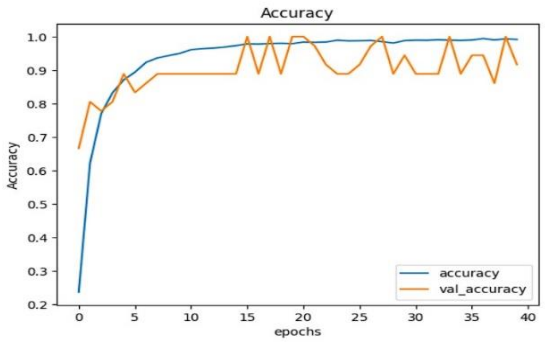

Figure 6 Evaluation of the Training and Validation Accuracy

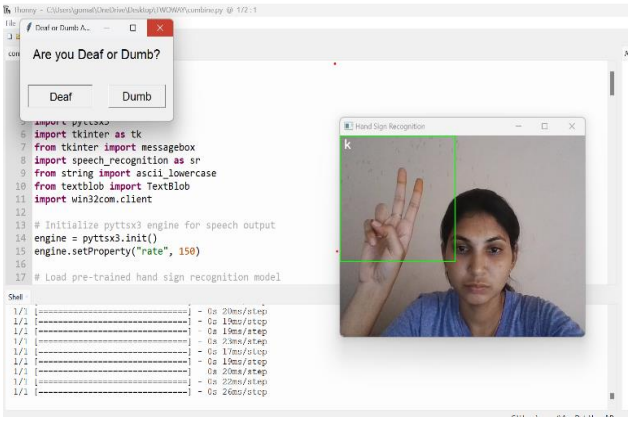

Figure 6 Real-Time Sign Language Detection

Conclusion

In order to assist the disabled, we have concentrated on sign language in this article. By using the CNN model, we were able to recognize sign language with greater accuracy. Our precision remains using real-time sign language detection at 91.67%. The suggested method has helped to improve sign detection accuracy. However, there are still many issues that need more research. In the future, we'll attempt to use a range of statistical moments to build utterance level representations. Furthermore, as technology develops further, voice-to-text conversion becomes our next crucial goal. By

International Research Journal on Advanced Engineering and Management
https://goldncloudpublications.com
https://doi.org/10.47392/IRJAEM.2025.0217
e ISSN: 2584-2854
Volume: 03
Issue: 04 April 2025
Page No: 1335 - 1341

expanding into voice recognition, sign language users and those who rely on spoken language will be able to converse seamlessly, closing another communication gap. We are committed to meeting the needs of people with disabilities, and we are well-positioned to keep developing cutting-edge solutions that promote empowerment, understanding, and inclusion. Accurately translating the speech of people with various accents, communication styles, or linguistic impairments may be challenging for our suggested method. Speech recognition systems may encounter difficulties due to accents, dialects, and differences in speech patterns. Concerns about security and privacy may arise from the technology, particularly if it continuously records or listens to human interactions.

References

[1]. K. K. Podder, M. E. H. Chowdhury, A. M. Tahir, Z. B. Mahbub, A. Khandakar, M. S. Hossain and M. A. Kadir, “Bangla Sign Language (BdSL) Alphabets and Numerals Classification Using a Deep Learning Model," Sensors, vol. 22, no. 2, pp. 574(1-18), 2022.

[2]. A. Mannan, A. Abbasi, A. R. Javed, A. Ahsan, T. R. Gadekallu and Q. Xin, “Hypertuned Deep Convolutional Neural Network for Sign Language Recognition," Computational Intelligence and Neuroscience, vol. 2022, Article ID 1450822, pp. 1-10, 2022.

[3]. F. Tanrisever and K. A. W. Voorbraak, “Crowdfunding for Financing Wearable Technologies," 2016 49th Hawaii Int. Conf. System Sciences (HICSS), pp. 1800-1807, 2016.

[4]. S. C. Das, M. B. Alam, M. S. A. Moon and M. S. Mia, “An Application Programming Interface to Recognize Emotion using Speech Features," 2022 4th International Conference on Sustainable Technologies for Industry 4.0 (STI), pp. 1-6, 2022.

[5]. A. Mavi, “27 Class Sign Language Dataset,â Kaggle, [Online]. Available: https://www.kaggle.com/datasets/ardamavi/27-class-sign-languagedataset. [Accessed: 26-Oct-2023].

[6]. Q. M. Areeb, Maryam, M. Nadeem, R. Alroobaea and F. Anwer, “Helping Hearing-Impaired in Emergency Situations: A Deep LearningBased Approach," IEEE Access, vol. 10, pp. 8502-8517, 2022.

[7]. A. Halder and A. Tayade, “Real-time Vernacular Sign Language Recognition using MediaPipe and Machine Learning," Int. J. Research Publication and Reviews, vol. 2, no. 5, pp. 9-17, 2021.

[8]. N. Mohamed, M. B. Mustafa and N. Jomhari, “A Review of the Hand Gesture Recognition System: Current Progress and Future Directions," IEEE Access, vol. 9, pp. 157422-157436, 2021.

[9]. Y. Obi, K. S. Claudio, V. M. Budiman, S. Achmad and A. Kurniawan, “Sign language recognition system for communicating to people with disabilities," Procedia Computer Science, vol. 216, pp. 13-20, 2023.

[10]. V. Adithya and R. Rajesh, “Hand gestures for emergency situations: A video dataset based on words from Indian sign language," Data in Brief, vol. 31, pp. 1-7, 2020.

[11]. G. A. Prasath and K. Annapurani, “Prediction of sign language recognition based on multi layered CNN," Multimedia Tools and Applications, vol. 82, pp. 29649-29669, 2023.

[12]. A. Mittal, P. Kumar, P. P. Roy, R. Balasubramanian and B. B. Chaudhuri, “A Modified LSTM Model for Continuous Sign Language Recognition Using Leap Motion," IEEE Sensors Journal, vol. 19, no. 16, pp. 70567063, 2019.

[13]. Z. H. Nayem, I. Jahan, A. A. Rakib and M. S. Mia, “Detection and Identification of Rice Pests Using Memory Efficient Convolutional Neural Network," 2023 International Conference on Computer, Electrical & Communication Engineering (ICCECE), pp. 1-6, 2023. +