IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

"A Hand gesture enabled service for the disabled using Deep Learning"

Atharva Bansode, Faraj Mulani, Fazila Sayed, Iramsabah Syed,

Prof-Ishwari Raskar

Trinity College of Engineering and Research,Pune.

Abstract: We often encounter people facing ailments such as hearing loss and blindness. You have trouble interacting with others. All the technologies developed so far are sensor-based and do not offer generalized solutions. This article describes a new technique for sensorless virtual communication. An image processing technique called histogram of gradients (HOG) and an artificial neural network (ANN) were used to train the system. A webcam is used to capture images of various gestures that are used as inputs to Mat Lab. The software recognizes the images and identifies core pending narrations to be played using the narration kit. This document describes two-way communication between deaf, blind, and sighted people. This means that the proposed system can convert sign language into text and speech.

Keywords: HOG (Histogram of Gradient), ANN (Artificial Neural Network), Mat lab, Voice replay kit.

I.INTRODUCTION

Machine learning has changed the dimension of how society thinks and perceives the world. The range of applications supported by machine learning is multidimensionally expanding. All three of his forms of machine learning, including supervised learning, unsupervised learning, and reinforcement learning, are intriguing newer ways, of impacting human life and reshaping the way modern businesses operate, increase. Among the various applications of machine learning, computer vision is a field that is attracting a research community to harness the potential of technological developments to advance society. Computer vision is an area of active research and development with advances in the field of computer vision technology. Researchers use computer vision to classify still images, and many algorithms have been proposed for this. Similarly, computer vision plays an important role in medical sensing and diagnostics. There is an essential need for such a system in society so that deaf people and people who use sign language can communicate in their daily lives. Computer vision allows machines to recognize hand gestures. Anyone can interpret their needs through these gestures if they know what they want to express with their gestures. In this article, we have developed a robust system that can predict words using hand signals.

This allows them to use this system more accurately after determining which sign recommends the best possible words, you can choose to make the system hassle-free and user-friendly. Especially tailored exactly to your needs. A common approach uses a modeled convolutional neural network (CNN) to recognize characters.

II. PROBLEM STATEMENT

Creation of a character recognition system that supports people with disabilities and enables appropriate communication using deep learning technology and algorithms.

III.LITERATURE SURVEY

Machine learning is not a new technology that has just emerged. Since the 1970s, attempts have beenmade to make the world a better place with the help of artificial intelligence and machine learning. kings showed curve-invariant poses using boundary histograms. A camera was used to protect the information image. The skin shade detection channel is used by the clustering procedure to determine the boundaries of each collection in the clustered image using the computed common shape. The image is split into multiple networks and the boundaries are normalized. Boundaries are called harmonic magnitude chains, which divide the image into multiple regions N according to explicit edges and are used in the form of histograms. Neural networks, MLP and dynamic programming, and DP adjustment were used for the classification process. Despite using characteristic harmonic magnitude histograms and harmonic magnitude FFTs, many analyzes were performed at different highlight positions. Convolutional neural networks have an established role in image recognition, as demonstrated by several researchers in the past. The concrete contribution of CNNs to medical disease diagnosis by relating scanned images to the presence or absence of disease is an exceptional application that has proven efficient and reliable. Rectified Linear Unit (ReLU) is one of the most robust activation functions used in image processing. ReLU is one of the most popular nonlinear activation functions trusted by researchers in deep learning projects. The experiment used 26 static positions of American Sign Language

1. **Paper Name: Real-Time Recognition of Indian Sign Language**

   Author: Muthu Mariapppan, Dr.Gomathi V

Abstract: - The real-time sign language recognition system is developed for recognizing the gestures of Indian Sign Language (ISL). Generally, sign languages consist of hand gestures and facial expressions. For recognizing the signs, the Regions of Interest (ROI) are identified and tracked using the skin segmentation feature of Open CV. The training and prediction of hand gestures are performed by applying a fuzzy c-means clustering machine learning algorithm. Gesture recognition has many applications such as gesture-controlled robots and automated homes, game control, Human-Computer Interaction (HCI), and sign language interpretation. The proposed system is used to recognize real-time signs. Hence it is very muchuseful for hearing and speech-impaired people to communicate with normal people.

2. **Paper Name:- Hand sign recognition-based communication system for speech-disabledpeople**

Author: Mr. Vishal P. Bhujbal, Dr.K.K.Warhade

Abstract: - According to the Census of India 2011. In India, 70 million people have a disability, among of that 18per cent of people are speaking and hearing impaired. That means India is a country that has many people having this kind of disability. These people experience the problem to participate in society and the enjoyment of equal rights and opportunities. Because they don't have the power to express feelings in the form of words and sentences. So,people try to deal with this problem using different techniques. In this paper, a smart hand sign interpretation system using a smart glove is proposed to reduce the communication gap between speech-impaired people and normal people. This wearable system utilizes five flex-sensors, a 3-axis accelerometer, one Bluetooth module, and a 16*2 LCD. In this system, the processor collects data from the 5 flex sensors and accelerometer. Further processor matches the data which is received from the sensors and the previously saved data. If the data matches with the saved data, then the assigned meaning for that data will be displayed on the LCD screen and also send to the Android mobile through Bluetooth. Android mobile app can convert this into voice. So, this system can convert sign language into a voice in a very simple way.

3. Paper Name: A Video-Based Indian Sign Language Recognition System (INSLR)Using Wavelet Transform and Fuzzy Logic

Author name: P. V. V. Kishore and P. Rajesh Kumar

Abstract: This paper proposes a complete skeleton of an isolated Video Based Indian Sign Language Recognition System (INSLR) that integrates various image processing techniques and computational intelligence techniques to deal with sentence recognition. The system is developed to improve communication between hearing-impaired people and normal people promising them better social prospects. A wavelet-based video segmentation technique is proposed which detects shapes of various hand signs and head movements in a video-based setup. Shape features of hand gestures are extracted using elliptical Fourier descriptions which to the highest degree reduces the feature vectors for an image. Principle component analysis (PCA) still minimizes the feature vector for a particular gesture video and the features are not affected by scaling or rotation of gestures within a video which makes the system more flexible. Features generated using these techniques make the feature vector unique for a particular gesture. Recognition of gestures from the extracted features is done using a Sugeno-type fuzzy inference system which uses linear output membership functions.

Finally, the INSLR system employs an audio system to play the recognized gestures along with text output. The system is tested using a data set of 80 words and sentences by 10 different signers. The experimental results show thatour system has a recognition rate of 96per cent.

4. Paper Name: Smart Tutoring System for Arabic Sign Language Using Leap Motion Controller

Author::- Heba Fasihuddin, Shatha Alsolami,Seham Alzahrani

Abstract: This paper presents a smart tutoring system for Arabic Sign Language (ArSL). Sign language is one of the main approaches to communication for people with hearing impairment. Many people are willing to learn sign language and support this segment of society; however, learning this language requires some effort and assistance. Tools that are used to support sign language learners and specifically ArSL are limited and insufficient. Hence, the development of a tool that is capable of training and assessing ArSL learners becomes a necessity. We proposed smart tutoring for ArSL based on using the leap motion's hand-tracking technology. This system aims to assist non-disabled learners who want to learn sign languages, such as undergraduates specializing in hearing disabilities, parents of kids with hearing impairment, or any interested subject. The system allows learners to practice ArSL at different levels and self-assess themselves. As it utilizes the re- cent technology of the leap motion controller, it can detect and track hand and finger movements and consequently assess the position and movement accuracy. Machine learning techniques, specifically the K- Nearest Neighbor algorithm was applied for classification and sign recognition. The preliminary prototype was developed and testedin terms of users' acceptance. The outcomes show satisfactory and promising results.It is expected that the proposed system will contribute to enriching the learning process of ArSL and consequently support an important segment of our community.

5. Paper Name: Deep Convolutional Neural Networks for Sign Language Recognition

Author: G.Anantha Rao, K.Syamala, P.V.V.Kishore, A.S.C.S.Sastry.

Abstract:- Extraction of complex head and hand movements along with their constantly changing shapes for recognition of sign language is considered a difficult problem in computer vision. This paper proposes the recognition of Indian sign language gestures using a powerful artificial intelligence tool, convolutional neural networks (CNN). Selfie mode continuous sign language video is the capture method used in this work, where a hearing-impaired person can operate the SLR mobile application independently.

Due to the non-availability of datasets on mobile selfie sign language, we initiated the creation of the dataset with five different subjects performing 200 signs in 5 different viewing angles under various background environments. Each sign occupied 60 frames or images in a video. CNN training is performed with 3 different sample sizes, each consisting of multiple sets of subjects and viewing angles. The remaining 2 samples are used for testing the trained CNN. DifferentCNN architectures were designed and tested with our selfie sign language data to obtain better accuracy in recognition. We achieved a 92.88per cent recognition rate compared to other classifier models reported on the same dataset

IV.METHODOLOGY

The current work has been graphically summarized using Figure 1. The implementation includes the classification of classes that people using this application would like to predict. (A task is to recognize a hand gesture and classify it as a specific task, which is a predictor that needs to be classified. Figure 1 mainly illustrates the pipeline concept used to train a deep neural network.(The initial convolutional neural network weights are trained by changing specific feature values, i.e.,specific feature values extracted from the images in the dataset will be the model that has been trained. After all the images have been passed to the model, we update their weights and other parameter values that the model depends on. A test data set was used to build confidence and reliability in the developed application. After several tests and voting initiatives, the best and most accurate modelsare considered. (The process of optimization is iterative. Models are trained and tested iteratively to find the best model for inference Dataset. To train the proposed model, we first prepared a custom dataset related to American Sign Language consisting of combinations of various charactersand their corresponding labels. Therefore, the models were trained on the same dataset. We maintain MNIST images from the Kaggle open repository.CSV file to perform analysis. (Using a snapshot of the training CSV file I created a script to transform an image consisting of various hand symbols and given pixel values using np. Create an array function and generate the corresponding image from it. Similarly, for the test CSV, an image was created using the pixel values from the given set of CSV files, and the corresponding image was generated and saved in the test folder Training photos. To perform good and best predictions, we need to build a reliable model, so we created a dataset that fits the process. (Including specific steps such as binary image creation , background removal, and edge detection to create a specific dataset. This wasdone to ensure that the responses were more appropriate, but this can be generalized to similar gestures for similar types of responses from the system. The dataset used for modeling consists of 23,826 images totaling 28*28 pixels, with different captions for each set of images from A to Z.

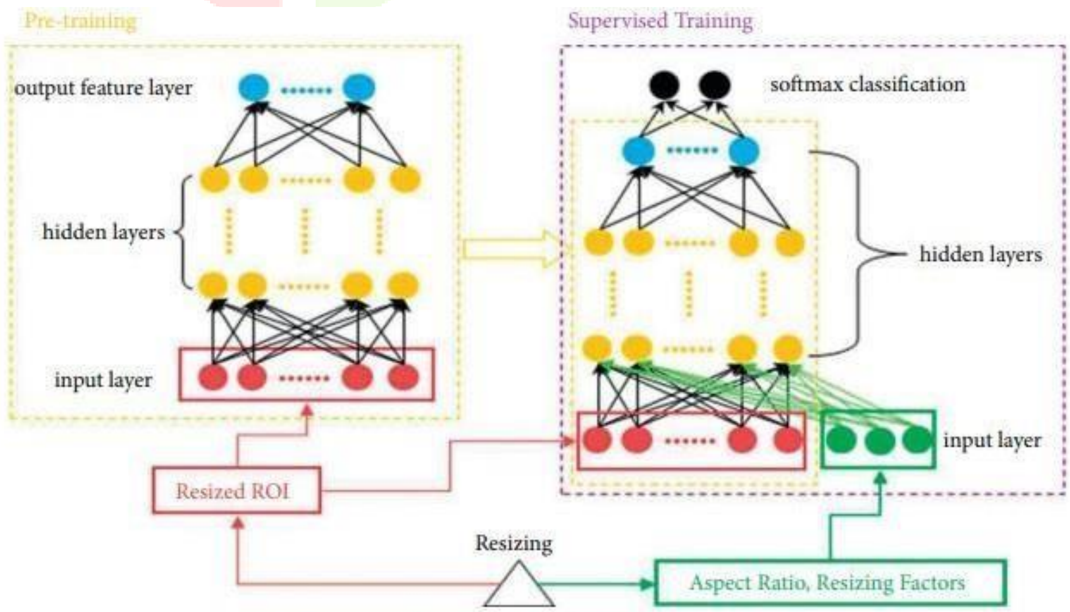
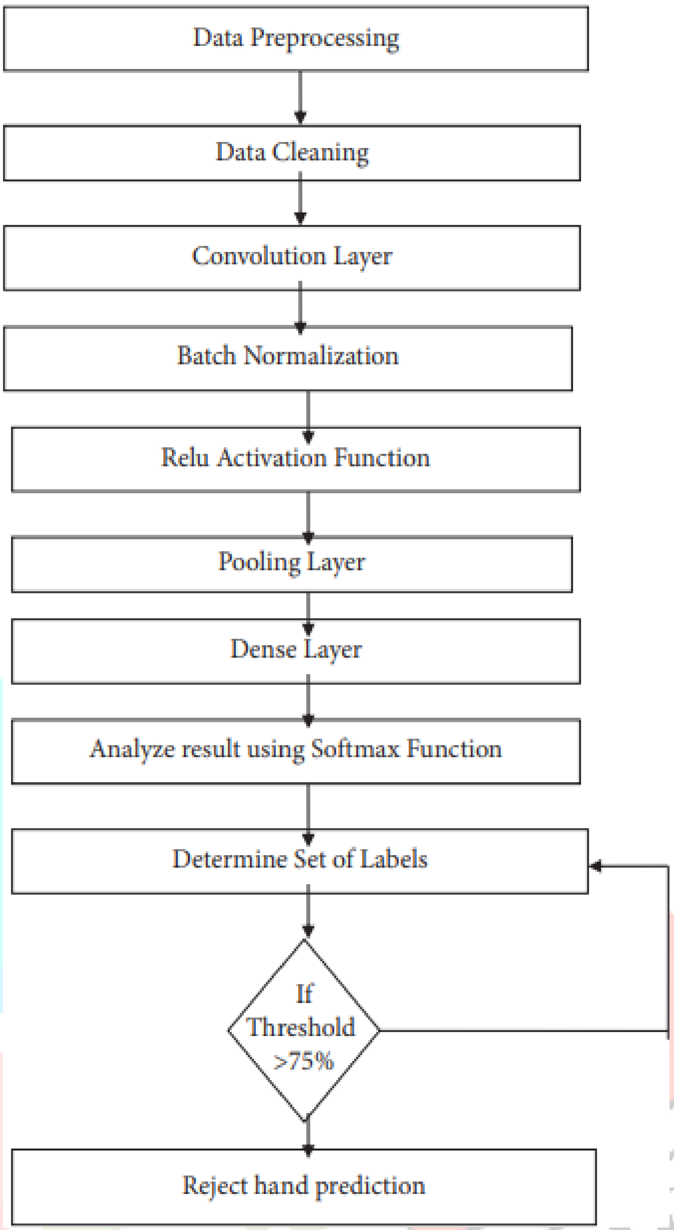


FIGURE 1: Flow chart for the proposed system.

V.IMPLEMENTATION

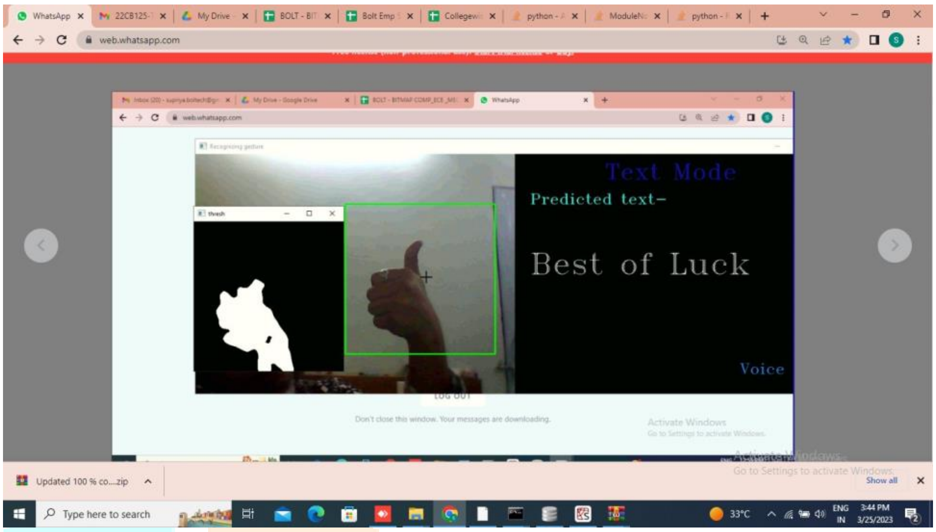
This project work was carried out in a multi-layered approach. To conduct this research work, we created five layers, including fully connected layers to perform an analysis of the data. The block diagram in Figure shows the experimental process flow. The first step, the data preprocessing step, cleanses the data to find more accurate data to analyze the information. Data cleansing is necessary to ensure accuracy and reliability in estimating the correct answer. Various data cleansing methods were used for this purpose. The data is then passed to the next shift for training. (The training layer involves the extraction of features that are more important for decision-making. Training is performed based on these selected features. In the current work, 32, 64, 128, and Batch normalizationis performed given the 256 sizes of the layer and the convolution image. Batch normalization provides fast and accurate processing. It is then passed to the ReLU activation function to classify and train the neurons. This is then passed to a max pooling layer to find the largest possible pixels that can contribute to image training. It is then transferred to the fully connected dense layer of the network and the SoftMax function analyzes the results. After the training phase, the model is trained on different image sets to recognize the characters required to recognize human gestures in live feeds. So, you must pass the image to the model according to the frame. The model then recognizes each frame passed to it and within it, the regions allocated or passed. This is the region of interest (ROI). In our case, we take the ROI regions and convert these multiple image sets into an array of pixel values. An array of pixel values is converted to float values mapped to it. After obtaining a series of image float values transferred to it based on intensity values, the model is trained to identify these specific regions with a certain probability. That is, it recognizes all letters of the alphabet with a certain probability value. (The proposed model was trained for 100 epochs and kept a batch size of 32 to make better and more accurate predictions on the data. Similarly, to test the results, images weresent to the neural network Predict a specific outcome based on the model passed in and trained. The results are produced as data labels or words as output representing the person holding hands. Generalization also involves building models to better predict outcomes. To achieve this goal, a trained model is generated on the desired set of 28*28-pixel images using the hand signature MNIST data set. Our work allowed us to create images and create a framework for predicting the hand gestures experienced by the system. To unbiased the results, we used a CNN that included image extraction with edge detection , denoising, and corner detection, and used these layers to map the results to show sample images of regions of interest from the datasets used.



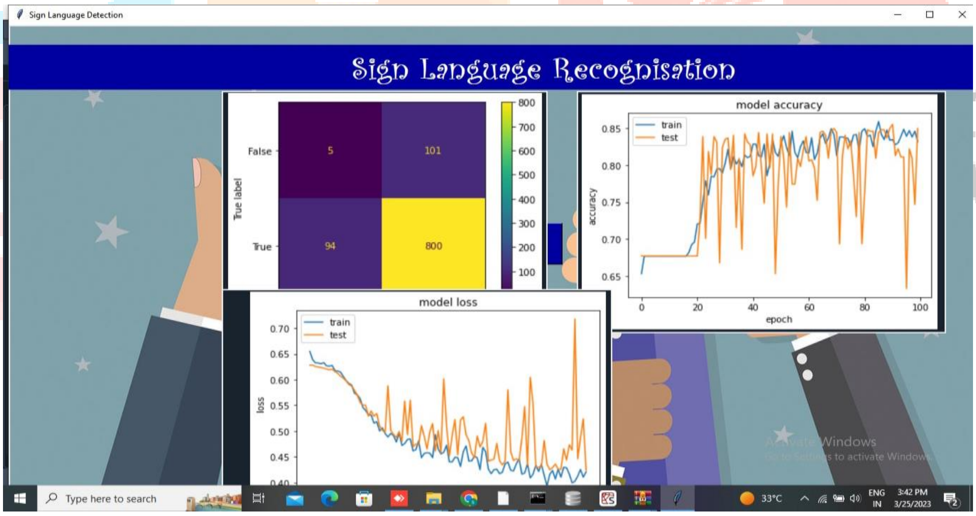
Block Diagram.

VI.EXPERIMENTAL RESULTS

In the proposed work, using the OpenCV [56] library cv2.putText, we proposed a model to predict the most probable character using minimum threshold selection. We chose this threshold as 95% in the current case to represent these values that determine the model's outcome to be delivered as a predicted gesture. Overall, we captured the most likely characters predicted by the model. Figures 10and 11 show typical responses of the system to hand gestures entered as inputs to the system. The proposed system works with 91.07% accuracy in predicting gestures. To achieve this accuracy, we have created 2 separate folders, based on training and test set size which is obtained by creating the custom dataset having 23,826 training and 2,668 test set images of different labeled folders. Furthermore, we specified those characters where we have used CNN models to train on those sets of images. (It was trained with 100 epochs with a batch size of 32 to obtain that accuracy for those images which have been augmented with hand images by using certain convolution and other linearoperations to increase the set of images. As those images are 28 ∗ 28 pixels, we had 05 layers followed by a fully connected neural network layer in which all five layers consist of one convolution layer which extracts convolved features from the images; then, those features are passed to the batch normalization layer to smoothen the images. Later, they are passed to the max pooled layer which will grab those higher-valued features that mostly affect our images, which are finally passed to the activation function. (In our case, this ReLU is used to obtain a specific set of weights for training a model to obtain specific weights). Similarly, the model represents 100 times the weights Trained for 100 epochs, model.



Output 1



Output 2

VII. TEST CASES

Test Case ID	Test Case	Test Case I/P	Actual Result	Expected Result	Test case criteria(P/F)
001	Enter The Wrong username or password click on submit button	Username or password	Error comes	Error Should come	P
002	Enter the correct username and password click on submit button	Username and password	Accept	Accept	P

Testcase 1

Test Case ID	Test Case	Test Case I/P	Actual Result	Expected Result	Test case criteria(P/F)
001	Enter the number in username, middle name, last name field	Number	Error Comes	Error Should Comes	P
001	Enter the character in username, middle name, last name field	Character	Accept	Accept	p
002	Enter the invalid email id format in email id field	Kkgmail,com	Error comes	Error Should Comes	P
002	Enter the valid email id format in email id field	kk@gmail.com	Accept	Accept	P
003	Enter the invalid digit no in phone no field	99999	Error comes	Error Should Comes	P
003	Enter the 10 digit no in phone no field	9999999999	Accept	Accept	P

Testcase 2

Test Case ID	Test Case	Test Case I/P	Actual Result	Expected Result	Test case criteria(P/F)
001	Store csv File	csv file	csv file store	Error Should come	P
002	Parse the csv file for conversion	parsing	File get parse	Accept	P
003	Attribute identification	Check individual Attribute	Identify Attributes	Accepted	P
004	Weight Analysis	Check Weight	Analyze Weight of individual Attribute	Accepted	P
005	Tree formation	Form them-Tree	Formation	Accepted	P
006	Cluster Evaluation	Check Evaluation	Should check Cluster	Accepted	P
007	Algorithm Performance	Check Evaluation	Should work Algorithm Properly	Accepted	P
008	Query Formation	Check Query Correction	Should check Query	Accepted	P

Testcase 3

VIII.CONCLUSION

In this paper, a neural network-based method for the automatic recognition of fingerprints in Indian Sign Language is presented. Identify characters by extracting features from hand shapes. We usedskin color-based segmentation to extract the hand region from the image. In this work, a new shape feature based on image distance transform is proposed. Features extracted from character imagesare used to train a feedforward neural network to recognize characters. The method is implemented entirely using digital image processing technology, so users do not need to carry special hardware devices to obtain hand shape features. The proposed method has a smaller amount of calculationthan the existing methods, and the accuracy is very high.

IX.REFERENCES

[1] Z. Gong, P. Zhong, and W. Hu, “Diversity in machine learning,” IEEE Access, vol. 7,pp.64323–64350, 2019.

[2] R. Sharma and K. Davuluri, “Design patterns for machine learning applications,” in Proceedings of the 2019 3rd Inter-national Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019.

[3] V. Wiley and T. Lucas, “Computer vision and image processing: a paper review,” International Journal of Artificial Intelligence Research, vol. 2, no. 1, pp. 28–36, 2018.

[4] M. I. Lakhal, H. Çevikalp, S. Escalera, and F. Ofili, “Recurrent neural networks for remote-sensing image classification,” IET Computer Vision, vol. 12, no. 7, pp. 1040–1045, 2018.

[5] J. Thevenot, M. B. Lopez, and A. Hadid, “A survey on computer vision for assistive medical diagnosis from faces,” IEEE Journal of Biomedical and Health Informatics, vol. 22,no. 5, pp. 1497–1511, 2018.

[6] A. Juneja, S. Juneja, A. Soneja, and S. Jain, “Real-time object detection using CNN- based single shot detector model,” Journal of Information Technology Management, vol. 13,no. 1, pp. 62–80, 2021.

[7] V. S. Kulkarni and S. D. Lokhande, “Appearance-based recognition of American sign language using gesture segmentation,” International Journal on Computer Science andEngineering IJCSE, vol. 2, no. 3, pp. 560–565, 2010.

[8] Z. Wang, S. Wu, C. Liu, and S. Wu, “The regression of MNIST dataset based on convolutional neural network,” in Proceedings of the 2020 International Conference on Advanced Machine Learning Technologies and Applications, Jaipur, India, 2020.

[9] A. Thakur and D. S. Saini, “Correlation processor-based sidelobe suppression for polyphase codes in radar systems,” Wireless Personal Communications, vol. 115, 2020.

[10] S. Zhao, W. Tan, S. Wen, and Y. Liu, “An Improved algorithm of hand gesture recognition under the intricate pond,” International Conference on Intelligent Robotics and Applications, pp. 786-787, 2008.9

[11] J. Sancho Rieger, S. Waldert, T. Pistohl, et al., “Vision-based gesture recognition usingneural networks approaches a review,” International Journal of Human-Computer Interaction, vol. 32, no. 1, pp. 480–494, 2008.

[12] M. Hagen, A. E. Ayadi, J. Wang, N. Vasiloglou, and E. Afshar, “Optimizing training data for image classifiers,” in Proceedings of the KDD 2019 Workshop on Data Collection, Curation, and Labeling for Mining and Learning, 2019.

[13] B. Shubankar, M. Chowdhary, and M. Priyaadharshini, “IoT device for disabled people IoT device for disabled people,” Procedia Computer Science, vol. 165, pp. 189–195, 2020.

[14] J. Gu, Z. Wang, J. Kuen, et al., “Recent advances in convolutional neural networks,” Pattern Recognition, vol. 77,2006.

[15] N. Sharma, V. Jain, and A. Mishra, “An analysis of convolutional neural networks for image classification,” ProcediaComputer Science, vol. 132, pp. 377–384, 2018.

[16] A. F. Agarap, “Deep learning using rectified linear units (ReLU),” 2018, https://arxiv.org/abs/1803.08375.

[17] M. Yasen and S. Jusoh, “A systematic review on hand gesture recognition techniques,challenges, and applications,” PeerJ Computer Science, vol. 20, 2019.

[18] J. Liu, K. Furusawa, T. Tateyama, Y. Iwamoto, and Y. Chen, An Improved Hand GestureRecognition with Two-Stage Convolution Neural Networks Using a Hand Color Image and its Pseudo-Depth Image Graduate School of Information Science and Engineering, Ritsumeikan University, Kyoto, Japan,2019.

[19] W. Aly, S. Aly, and S. Almotairi, “User-independent American sign language alphabet recognition based on depth image and PCANet features,” IEEE Access, vol. 7, pp. 123138–123150,2019.