Static Gesture Recognizer Using Hybrid Neural Network

R.Vasumathi[1], J.Rizwana[2]

PG Student [EST], Dept. of EEE, Rajalakshmi Engineering College, Thandalam, Tamilnadu, India[1]

Associate Professor, Dept. of EEE, Rajalakshmi Engineering College, Thandalam, Tamilnadu, India[2]

ABSTRACT: The main objective of this paper proposes an Embedded System employing Hybrid Neural Network (NN) for an Efficient Static Recognizer. The Hybrid Neural Network consists of Active Contour Model (ACM) and Convolutional Neural Network (CNN) in which the input data of hand sign is pre-processed and segmented using ACM and that image is feed forward to CNN classifier to classify the image for free-air hand gesture recognition without the use of hand gloves or sensors to offer a more efficient way of gesture segmentation as gesture boundaries. The system is developed to detect the hand signs in American Sign Language (ASL), and its convenience will be verified through simulations and the input is converted from sign to text and vice versa. The image is described in four languages such as English, Tamil, Hindi and Telugu and there is a need in designing an efficient human-computer interface. The possibility of the system in which the hand shapes or poses of static gestures is being captured by a camera, to improve the way of conveying information's between the deaf and the non-deaf.

KEYWORDS:Active Contour Model (ACM); Hybrid model; Convolutional Neural Network (CNN)

I.INTRODUCTION

Gesture recognition is the pattern of understanding and interpreting relevant movements of the arms, face, hands or sometimes head. There is of great need in designing an efficient human-computer interface. The potential for an application is the natural way of human machine interaction has been in research for engineers and scientists. Today the industry is working on different application to make interactions more natural, easy and convenient without wearing any sensors, hand gloves, etc.,. It is to recognize specific human gestures and process them to reduce the communication gap. Sign Language are used to convey some meaningful information. The different types of ideas and actions are used to differentiate the gestures that includes shapes, orientation and figure pattern. But the hand gesture recognition is a challenging task, the major task in this paper is a recognition and classification of gestures.

The image capture circuit is fixed on FPGA which consists of an on-board camera, so that the posture images which are mapped to the lower dimensional map of neurons in the (Self Organizing Map) SOM and Hebbian is single feed forward neural network trained with Hebbian algorithm [6]. The neural network is well suited for real-time systems because of their fast response and computational times which are because of their parallel processing architecture. The human–computer interaction for smart environment applications using fuzzy hand posture and gesture models [1]. The system transforms preprocessed data detected hand into a fuzzy hand posture feature model by using fuzzy neural network and based on this model determines actual hand posture applying fuzzy inference. Real-Time Hand Gesture Detection and Recognition Using Bag-of-Features and Support Vector Machine Techniques [7].

The hand gestures is fed through the multiclass Support Vector Machines (SVM). This helps in building a grammar which has detected hand gesture and its key points. A real-time posture recognition using haar-like topological feature describes initially detecting the region of hand by a statistical method based on haar like feature and color segmentation technique in real time [3]. Rapid hand posture recognition using adaptive histogram template of skin and hand edge contour in three steps namely hand segmentation, feature extraction and posture recognition [5]. Vision based hand gesture recognition in which the hands are used directly as an input device for providing natural HCI [9]. It has been evolved from text based interfaces through 2D graphical based interfaces, to fully fledged multi participant virtual environment (VE) system rely on suitable controlled lab setting but not generalized to arbitrary settings. The static hand

sign recognition using linear projection methods explains shape matching algorithm method such as linear projection methods for static hand sign recognition in Malaysian sign language [8].

A real time hand gesture recognition based on haar wavelet representation which consists of a code word scheme based on features of hand gestures and the use of new measurement metric with a penalty score which enhances recognition accuracy [11]. Artificial neural networks for real-time optical hand posture recognition using a color-coded glove describes optical capture of the image of hand in the frame of video camera which is being processed inside an immersive virtual reality workspace by use of specially coded color glove [4]. The hand sign uses a feed forward ANN trained using supervised hebbian algorithm for input data preprocessing and data classification in competitive Artificial Neural Network (ANN) [10]. The pose of hand model is estimated with an Unscented Kalman Filter (UKF) in which it permits higher frame rates providing higher accuracy to minimize the geometric error [2].

II.FLOW DIAGRAM

The flow diagram representation of the Hybrid NeuralNetwork (ACM_ CNN) architecture is shown in Fig 1.

1. InputImage: The hand gesture images are captured and collected in the data folder, so that the image of hand sign is taken as input during simulation. The image is preprocessed, segmented and classified using hybrid neural network, the sign to text and vice versa is performed. The conversion of languages such as English, Tamil, Hindi and Telugu are being used in this system.

2. Pre-processing and Segmentation: The input image of hand sign is cropped in certain size and the threshold value is set according to the image of hand. The image is preprocessed and segmented by converting Red Green Blue (RGB) to gray and the pixels are assigned.

3. CNN Classifier: The hand sign is recognized and it is classified using CNN for saving memory as well as time. CNN has N number of layers and in which each layer has many neurons that respond to different combinations of inputs from the previous layers.

4. Hand Gesture Output: Finally, the hand gesture is recognized, which is the output of the proposed hybrid neural network model.
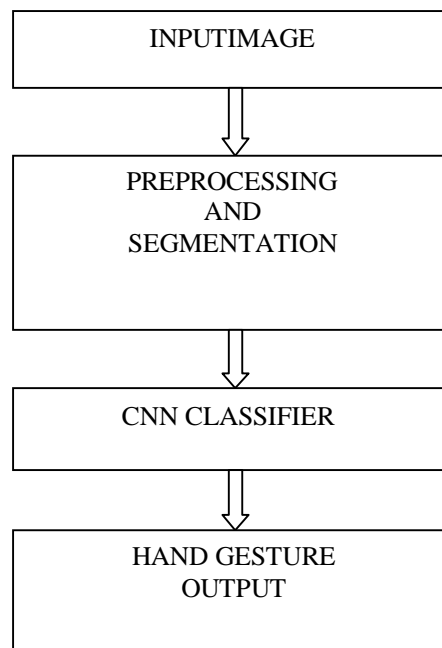

Fig. 1. Flow diagram of Hybrid Architecture

III. ARCHITECTURE OF HYBRID NEURAL NETWORK

The block diagram representation of the Sub model of Hybrid Neural Network (ACM_ CNN) is shown Fig 2. The proposed hybrid neural network model is a composition of Active Contour Model (ACM) and Convolutional Neural Network (CNN). This is to realize for much better prediction system. ACM can be used for preprocessing and segmentation of data. Convolutional neural network is used as a classifier. The objective of hybrid model is to reduce error of each model and to obtain globally optimal efficient performance. The hybrid neural network model has better properties than single model, it is able to interpolate and extrapolate much more accurately is easier to analyze and interpret the data. This model reject input noise and provides better results for the system.

1. Active Contour Model (ACM): Active contour algorithm was first introduced in 1987 by Kass et al to gain popularity and to represents an object boundary as a parametric curve. Active contour is the model widely used in image segmentation which is a fundamental task in image analysis responsible for partitioning an image into multiple sub-regions based on a desired feature.
Active contour detects the static hand poses from the camera is composed into image preprocessing and segmentation. The aim of preprocessing is to revamp the image data that extracts unwanted distortions and enhances some image features important for further processing. The image of hand sign is cropped and the Red Green Blue (RGB) is converted to gray by means of preprocessing. The image segmentation is the process of splitting a digital image into multiple segments (sets of pixels, also known as super-pixels). The goal of segmentation is to simplify and/or change of an image into edges that is easy to analyze. Image segmentation is typically used to locate boundaries and objects such as lines, curves, etc., in images. The image segmentation is the process of assigning a label to every pixel in an image of the hand such that pixels with the same label share certain characteristics.

2.Convolutional Neural Network (CNN): A Convolutional Neural Network (CNN, or ConvNet) is a feed-forward artificial neural network in which the connectivity pattern of the hand gesture, whose individual neurons are arranged in a way that they respond to overlapping regions covering the visual field. Convolutional neural networks are often used in image recognition systems to reduce the error rate and sharing weights. The hand sign is recognized and its area is calculated using CNN. The connections of CNN have numeric weights that are tuned during the training process, so that a trained network will correctly respond when presented with a hand sign image to recognize. The network consists of multiple layers of cluster-detecting neurons. Each layer has many neurons that matches to different combinations of inputs from the previous layers. CNNs are used in variety of areas, including pattern and image recognition, natural language processing, and video analysis. In CNNs, the weights of the convolutional layer is used for classification of image are determined during the training process of gesture recognition. The network structures of CNNs lead to savings in memory requirements, computation complexity and, at the same time, give better performance for applications where the input has local correlation such as image and speech. The advantages of using CNN in this project for image recognition are ruggedness to shifts and deformation in the image, easier and better training.
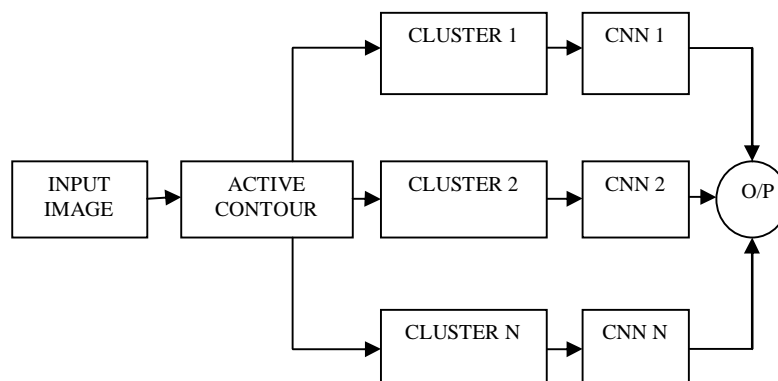

Fig.2. Sub Model of Hybrid NN (ACM_CNN) Architecture

IV.METHODOLOGY

1. Image Acquisition: The image is captured in real time by using web camera the gesture is captured and recognized.

2. ImagePreprocessing: The image is preprocessed to get suitable input image to recognize the gesture some preprocessing techniques are image conversion and image size conversion. The image conversion is performed at color level such as RGB to gray level. The image can be converted based on the requirement.

3. Image Segmentation: The segmentation process is used to detect or segment the origin of interest in an image. Segmentation operation can be done by threshold method. Threshold can be fixed based on the required object by segmentation. In proposed methodology, segmentation is used to detect and segment level of hand. The skin color or hand color is given as the input for thresholding. The segmented hand is converted to binary image.

Fig.3. Original hand image for Have a nice day

4. Feature Extraction: The extracted features are described as follows:

- Extrema Points - The feature is extracted based on the extrema points in edge of the object. For every binary object will have an extrema points.

- Area - Area can be calculated for the white pixels in the binary image. The white pixels represents the hand image.

- Centroid - Centroid point has been calculated from an object center points.

5. Testing Phase: The test image of the hand gesture is proposed by processing and segmentation. The extracted features of the test image is given to the CNN phase and the input features are compared with stored features in a databases. It classifies the gesture based on features.

6. Training Phase: The set of samples are taken for each and every gestures and their features are stored in the neural network system. Likewise all the gestures are trained and their features are stored in neural network.

V. RESULT AND DISCUSSION

The experiments in hybrid NN were carried out to test performance in comparison with that of convolutional neural network. ACM divides the input data into clusters, in order to improve the accuracy. Each cluster has CNN stages. The product of each CNN with ACM output is found to produce the recognition output. All the simulation results were generated using the MATLAB (version 2013a). The hybrid model is an effective and accurate. The figures shows the image of pre-processing and segmentation in which it is described within 256x256 pixel values as programmed. The segmentation of hand and pre-processing result is shown below:
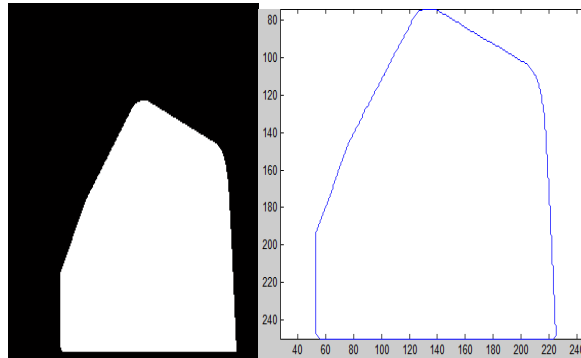
Fig. 4. Pre-processed image of Have a nice day

The Fig. 3 shows the original image of the hand gesture for the sign 'Have a nice day'. Fig. 4 shows the pre-processed image and the x axis and y axis represents the pixels of range within 250*250. The actual colour image Red Green Blue (RGB) is converted into grey. The x axis and y axis shows the values of pixels according to the area which is calculated for the image one is 13342 using the threshold value of 0.58 as shown in Fig. 5.
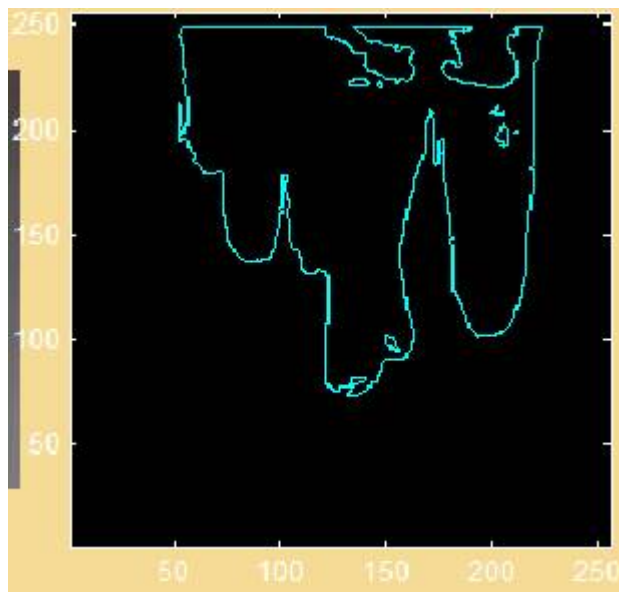

Fig.5. Segmented image of Have a nice day

Edges of the image is processed by means of active contour and classified using convolutional neural network. From Fig 6, it is easy to convey the information about the gesture sign to the deaf people that the sign is 'God Bless You' in 4 languages such as English, Tamil, Hindi and Telugu.

Conversion of Languages

Have a nice day!

English

ஒரு நல்ல நாள்
இருக்கிறது

Tamil

आपका दिन शुभ हो

Hindi

ఒక మంచి రోజు

Telugu

Fig.6. Conversion of Languages

VI.CONCLUSION

The suggested hybrid model of neural network can be used in hand gesture recognition systems. Each network is trained and tested using image data. An experimental result shows that hybrid model output is performed with accurate sign detection and displayed in 4 languages such as English, Tamil, Hindu and Telugu for communication between deaf and non-deaf.

REFERENCES

[1]   A. R. Varkonyi-Koczy and B. Tusor, "Human–Computer Interaction For Smart Environment Applications Using Fuzzy Hand Posture And Gesture Models," May2011.
[2]   B. Stenger, P. R. S. Mendonca, and R. Cipolla, "Model-Based 3D Tracking Of An Articulated Hand," in Proc. Brit. Mach. Vis. Conf. (BMVC), vol. 1. Sep. 2001, pp. 63–72.
[3]   C. Chuqing and L. Ruifeng, "Real-Time Hand Posture Recognition Usingharr-Like and Topological Feature," in Proc. Int. Conf. Mach. Vis.HumanMach. Inter., 2010.
[4]   F. Malric, A. E. Saddik, and N. D. Georganas, "Artificial Neuralnetworks For Real-Time Optical Hand Posture Recognition Using A Color-Codedglove," in Proc. IEEE Int. Conf. Jul. 2008, pp. 105–110.
[5]   G. Tofighi, S. A. Monadjemi, and N. Ghasem-Aghaee, "Rapid Handposture Recognition Using Adaptive Histogram Template Of Skin And Hand Edge Contour," in Proc. 6th. (MVIP), Oct. 2010, pp. 1–5.
[6]   Hiroomi Hikawa, Member, IEEE, and Keishi Kaida, "Novel FPGA Implementation of Hand Sign Recognition System with SOM–Hebb Classifier" Ieee Trans. Circuits And Syst. For Video Technology, Vol. 25, No. 1, January 2015.
[7]   N. H. Dardas and N. D. Georganas, "Real-time hand gesture detectionand recognition using bag-of-features and support vector machine techniques," IEEE Trans. Instrum. Meas., vol. 60, no. 11, pp. 3592–3607, Nov. 2011.
[8]   N. Gamage, K. Y. Chow, and R. Akmeliawati, "Static hand sign recognition using linear projection method," in Proc. 4th Int. Conf. Auto. Robots Agents (ICARA), Feb. 2009, pp. 403–407.
[9]   P. Garg, N. Aggarwal, and S. Sofat, "Vision based hand gesture recognition," World Acad. Sci., Eng. Technol., vol. 49, no. 1, pp. 972–977,2009.
[10]   S. Oniga, A. Tisan, D. Mic, A. Buchman, and A. Vida-Ratiu, "Hand postures recognition system using artificial neural networks implemented inFPGA," in Proc. 30th Int. Spring Seminar Electron. Technol., 2007, pp. 507–512.
[11]   W. K. Chung, X. Wu, and Y. Xu, "A realtime hand gesture recognition based on haar wavelet representation," in Proc. IEEE Int. Conf. Robot. Bio mimetics, Feb. 2009, pp. 336–341.