Journal of Theoretical and Applied Information Technology
15th March 2020. Vol.98. No 05

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ENHANCEMENT OF SINGLE-HANDED BENGALI SIGN LANGUAGE RECOGNITION BASED ON HOG FEATURES

[1]TASNIM TABASSUM, [2]IQBAL MAHMUD, [3]MD. PALASH UDDIN, [4]EMRAN ALI, [5]MASUD IBN AFJAL, [6]ADIBA MAHJABIN NITU

[1] B. Sc. Student (Session 2014), Hajee Mohammad Danesh Science and Technology University, Dinajpur, Bangladesh

[2] Software Engineer, Samsung R&D Institute, Bangladesh

[3] PhD Candidate, Deakin University, Geelong, VIC 3220, Australia and Assistant Professor, Hajee Mohammad Danesh Science and Technology University, Department of Computer Science and Engineering, Dinajpur, Bangladesh

[4] Assistant Professor, Hajee Mohammad Danesh Science and Technology University, Department of Computer Science and Engineering, Dinajpur, Bangladesh

[5] Faculty Member, Hajee Mohammad Danesh Science and Technology University, Department of Computer Science and Engineering, Dinajpur, Bangladesh

[6] Associate Professor, Hajee Mohammad Danesh Science and Technology University, Department of Computer Science and Engineering, Dinajpur, Bangladesh

E-mail: [1]towmony@gmail.com, [2]i.mahmud494@gmail.com, [3]palash_cse@hstu.ac.bd, [4]emran.cse@hstu.ac.bd, [5]masud@hstu.ac.bd, [6]nitu.hstu@gmail.com

ABSTRACT

Deaf and dumb people usually use sign language as a means of communication. This language is made up of manual and non-manual physical expressions that help the people to communicate within themselves and with the normal people. Sign language recognition deals with recognizing these numerous expressions. In this paper, a model has been proposed that recognizes different characters of Bengali sign language. Since the dataset for this work is not readily available, we have taken the initiative to make the dataset for this purpose. In the dataset, some pre-processing techniques such as Histogram Equalization, Lightness Smoothing etc. have been performed to enhance the signs' image. Then, the skin portion from the image is segmented using YCbCr color space from which the desired hand portion is cut out. After that, converting the image into grayscale the proposed model computes the Histogram of Oriented Gradients (HOG) features for different signs. The extracted features of the signs' are used to train the K-Nearest Neighbors (KNN) classifier model which is used to classify various signs. The experimental result shows that the proposed model produces 91.1% accuracy, which is quite satisfactory for real-life setup, in comparison to other investigated approaches.

Keywords: *Deaf and Dumb, Bengali Sign Language Recognition, Skin Segmentation, HOG Features, Bengali Sign Language Dataset*

1. INTRODUCTION

The incredible invention of language is what draws a distinction between humans and all other organisms on earth. Humans are solely the one species that has developed an advanced method of communication between individuals [1]. Nevertheless, apart from the normal individuals, the individuals that are deprived of the ability to hear and speak the only possible alternative form of communication is the use of signs and consequently sign language was constructed. Sign language is an intrinsic language that is systemic and rule-based, accompanied by discrete vocabularies of unrestrained standard symbols and grammatical formations [2]-[5]. For the hearing and speech impaired individuals, it is nearly impossible to effortlessly communicate with people who are not familiar with sign language. For this reason a system can be set, that will convert sign language into conventional form of communication either in text or in audio [2]. Sign language recognition is compulsory for apprehending a human oriented interactive system that can implement an interaction

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

identical to regular communication. A myriad of works on sign language recognition system have been proposed beforehand, which generally make use of probabilistic models, for instance Hidden Markov Models (HMM) and Artificial Neural Network (ANN) [6]. This paper proposes a model that works with HOG features and KNN classifier and then it compares the performance of different Bengali sign language recognition models. The objective of the comparison is to obtain the most prudent sign language recognition method for constructing the Bengali Sign Language (BdSL). According to WHO there are 360 million persons in the world with disabling hearing loss and 1 million are dumb [7]. They lack the ability to fluently communicate with normal people. Sign language aids them to communicate with others and lead their lives with less difficulty. Sign Language Recognition is needed to construct a system that can assist people understand sign language and can be incorporated into daily lives using digital devices.

The arena of sign language research can be classified into two types; first one is a vision-based computer that implements computer vision and the second one is built on the sensor data. For this reason, it is important to know how to make use of the services of the system effectively [8]. As sign language is not a universal language, it is as complex and diverse as any spoken language. There is a variety of sign language that is used in different regions of the world. The American Sign Language (ASL) is regarded as the International Sign Language. For various sign languages a number of researchers worked on systems for the recognition of sign languages. Sign language recognition and sign-to-text transformation are less advanced field comparing to other arenas of computer science application. Moreover, recognizing Bengali signs is lesser improved region. However, recent breakthroughs have been made in this arena and research is emerging only. In 1977, a mechanical hand was built up that was fit for spelling words utilizing letters in order [9]. However, the initial model could not form letters requiring wrist movement. A robotic hand was designed by 1992 that could smoothly produce letters received from text phones. A computer-controlled electro-mechanical finger hand, called Ralph, was finally developed in 1994 [10]. Input from different sources could be accepted by the robotic hand. Takahashi and Kishino's early approach to sign language recognition in 1991 relied on users wearing wired gloves, which extracted only hand shapes and was therefore limited to finger spelling or static gestures [11]. In 2002, Ryan Patterson developed a simplistic hand glove that senses the hand movements of signed alphabets and then transfers the data wirelessly to a portable device which shows the text on the display [9]. Starner and Pentaland's approach in 1995 featured real-time video recognition of American Sign Language using Hidden Markov Models [12]. The signers had to wear specially colored gloves for hand tracking in this approach. Later in 2007, Dreuw et. al. proposed an independent speech recognition system based on a single webcam without any use of a single glove [9] [8]. In 2009 Kelly et. al.'s more recent approach incorporates non-manual features, i.e., head movement as opposed to its previous approach, which focused primarily on manual features [13]. A single webcam was the system and the user wore colorful guides to constantly recognize sign language. ASL was converted into written and spoken English through the Sign2conversion system. CopyCat, a game intended for deaf children that uses 2D camera and cabled gloves to develop their working memory and language skills while playing [14][9]. Currently, Dreuw has launched and implemented a larger project called Sign-Speak [15]. In addition to all of the above researches, many others are initiated by students and scientific experts all over the world to build the bridge between the handicapped and the others.

Nearly 3 million deaf people live in Bangladesh, according to WHO. Bengali sign language is the first language these people. The user community of Bengali sign language is the largest community among Bangladesh's language-based minority communities. They are not able to communicate easily with ordinary people. A service that will interpret sign language into our standard form of communication can be provided to enable communication between deaf and silent people using sign language and those without understanding of sign language. In recent years, a number of researches has been conducted to develop recognition of Bengali sign language. The video clips of different gestures of sign language as input and sound are analyzed in a system that has been developed by Pavel et. al. [16]. By analyzing captured images from input video frames, angles of various parts of the hand with body were calculated manually and stored in a database with corresponding audio meanings. Najeefa Nikhat Choudhury and Golam Kayas have proposed a system that recognizes Bengali sign language using an approach based on a computer vision [17]. Neural networks were also used to train individual signs in that research [17]. Their initial objective is

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

to recognize isolated signs that are conveyed by hand movement [17]. Kinect was used to track the user's hands movement. M. V. D. Prasad, P. V. V. Kishore, E. Kiran Kumar, D. Anil Kumar developed a fusion based edge operator for Indian sign language recognition which works with the canny segmentation, elliptical fourier descriptors and morphology Segmentation [18]. Ashok Kumar Sahoo, Kiran Kumar Ravulakollu proposed a model based on KNN classifier and artificial neural network for Indian sign language which works with 95.30% accuracy for single handed signs [19]. Angur M. Jarman et. al. developed a Bengali sign language fingertip algorithm [20] to identify 46 hand gestures, including 9 for 11 vowels, 28 for 39 consonants and 9 for 9 numerals, based on the pronunciation similarity [20]. The captured image was initially re-sized and transformed into a binary format, using only top-most, left-most, and right-most white pixels to crop the region of concern. Using a fingertip finder algorithm, the positions of the fingertips were found. They demanded 88.69 percent accuracy in the recognition of Bengali sign language. Kaushik Datta et. al. proposed an interface between Bengali text and the Bengali sign language. The input text is scanned and rearranged in their system (if necessary) in accordance with the Bengali sign language grammar rule [21]. The method used grammar ids to process the input text and then concatenate and demonstrate the corresponding signs in a custom media player [21]. The system has been designed to convert Bengali sign language used in India's West Bengal [21]. Biswajit Sarkar et. al. instituted to convert Bengali Text to Sign, a multimedia-based translator [22]. Most recent work is carried out by Md Sanzidul Islam et. al. where they built a model for the recognition of Bengali Sign Language (BdSL) digits in deep learning approach [23]. Fahmid Nasif Arko has proposed a methodology that utilizes a model that uses YCbCr algorithm to detect skin color of all kinds and uses the Bag of features for feature extraction and Support Vector Machine (SVM) for training and evaluation [24]. The model used both male and female hand gestures to use their own dataset of Bengali sign language. The average precision was 86 percent. This paper proposes a model for Bengali sign language recognition that works with Histogram of Oriented Gradient (HOG) features and K-Nearest Neighbors classifier on a self-generated dataset. The reasons of using ANN, KNN and SVM in the comparative study are due to their promising adaptations in various application domains such as speech recognition [25] [26], face recognition [27], HCR [28], object recognition [29], diseases detection [30]-[32], remote sensing [33]-[36] and so on.

The rest of this paper is prepared into the following sections. Section 2 discusses the proposed Bengali sign language recognition model with the detail explanation of the constituent steps. In Section 3, we focus on the experimental setup and result analysis of the proposed model whereas Section 4 summarizes the explanations and accomplishes the paper.

Figure 1: Dataset prepared for recognizing single handed Bengali signs

2. PROPOSED METHEDOLOGY

Sign language recognition is a pattern recognition problem in machine learning and image processing that deals with various categories of hand postures for different signs. In this manner, it is more likely a category classification problem rather than a straight forward pattern recognition problem. In Bengali vocabulary, there are 50 alphabets including all the vowels and consonants. All these alphabets can be expressed using 35 different Bengali signs. These signs can be expressed using a single hand or both hands. This work deals with single handed mechanism for recognizing Bengali alphabets' signs. The key problem when developing a Bengali sign language recognition model is that there is no public dataset available. For this reason, we have prepared a dataset in this work through going to some local schools of deaf and dumb. The prepared dataset contains 1400 images for 35 Bengali alphabets' signs, 40 instances for each sign. We had managed to get 8 deaf/dumb people to take pictures from and taken 5 pictures of different instances for each sign. For convenience, we have taken background dissimilar to skin color as this work deals with skin color segmentation. The background is usually set as white or light sky blue. This is constant throughout capturing all the images. An overview of the dataset prepared is given in Fig. 1:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

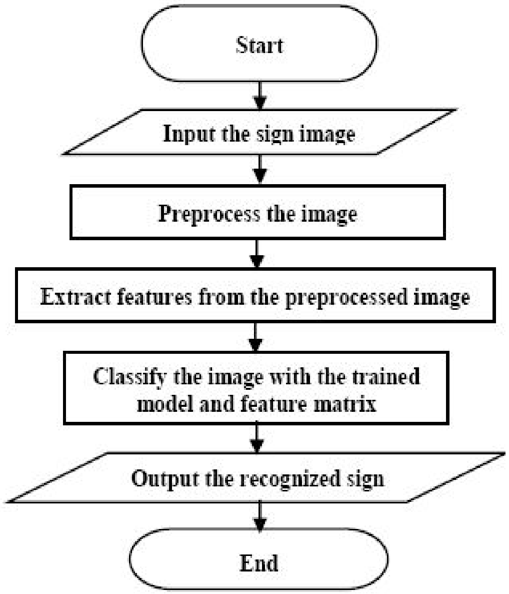


Figure 2: Flowchart of the proposed model

However, the proposed model is composed of two parts which are (a) making suitable feature matrix from the images for training and testing the classifier and (b) training a classifier with features for testing new instances. Now, to make suitable feature matrix from the image-set this work proposes some well-studied preprocessing techniques for image processing on the images. As this work discusses about different methodologies and their accuracies too several different preprocessing techniques have been used throughout this work. The flow chart of the proposed model is given in the Fig. 2.

2.1 Pseudo-code for the Proposed Model

This work proposes a model that works in three phases: preprocessing, training and testing. The procedures are discussed below:

2.1.1 Preprocess (*image_set, image_label*)

i. Resize each image to 255×255×3 for making all images of the same resolution
ii. Process the each image using histogram equalization and lightness smoothing
iii. Perform skin segmentation on the images using YCbCr color space and biggest blob analysis for finding hand portion
iv. Crop out the desired hand portion
v. Normalize and resize each image again to 255×255×3 color image
vi. Convert the image into gray-scale image
vii. Split the image-set randomly into 55:45 ratio for training and testing purpose

2.1.2 Training (*training_set, training_label*)

i. Extract the HOG features from all of the images of training set to create the feature matrix
ii. Generate the feature vector using the feature matrix and training label
iii. Train the KNN classifier with the feature vector

2.1.3 Testing (*test_set, test_label*)

i. For each image in the test set:
   a. Extract the HOG features from the image
   b. Classify the image using the trained model
   c. Output the recognized sign
   d. Compare with the real sign of the image
ii. Find out the accuracy of the model from the recognized sign and real sign of the image

The total architecture of the proposed model can be described by the following block diagram of Fig. 3.

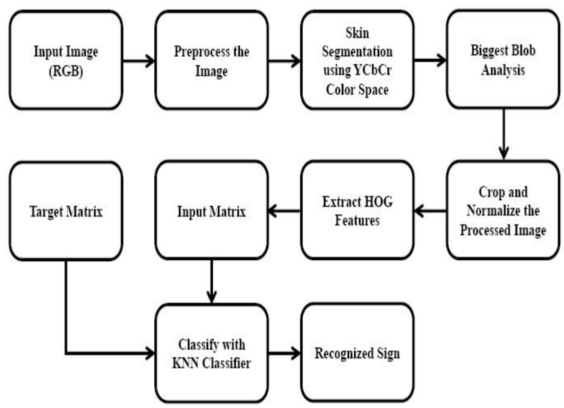


Figure 3: Architecture of the Proposed Model

2.2 Detailed Discussion of the Proposed Model

The operational steps are described in this section. The steps are as follows:

2.2.1 Preprocessing

The accuracy of any image processing model prominently depends on the pre-processes done on the image. This work uses some straightforward preprocessing techniques in image processing. The images in the dataset are in RGB color space. As the images in the dataset are taken through different cameras and different setups of light exposure, appropriate preprocessing is necessary in this dataset. At first, the images are resized into 255×255×3 for making all the images of the same resolution. Different people have different skin tone, so the skin color may vary in different images as well as the light intensity at a particular image

Journal of Theoretical and Applied Information Technology
15th March 2020. Vol.98. No 05

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

gives different skin color. As a consequence, lightness smoothing and histogram equalization are performed on the images for normalizing the skin color.

2.2.2 Skin Segmentation and Biggest Blob Analysis

The next part of the proposed model is skin segmentation for finding out the hand segment from the images. This paper uses YCbCr color space for segmenting the hand portion from the background. As all the images in the dataset are in RGB, so at first the images are converted into YCbCr color space, where Y is the luminance information presented in the image, Cb defines the distance of any color component from blue value and Cr defines the distance of any color component from red value [37]. The equation for converting RGB to YCbCr color space is given below:

$$Y' = 16 + (65.481.R' + 128.553.G' + 24.966.B')$$
$$Cb = 128 + (-37.797.R' - 74.203.G' + 112.0B')$$
$$Cr = 128 + (112.0.R' - 93.786.G' - 18.214.B')$$

YCbCr color space is independent of any luminance information so that it gives better performance in skin segmentation. The corresponding skin color clusters are:

$$Y > 80$$
$$77 < Cb < 127$$
$$133 < Cr < 173$$
$$\text{Where, } Y, Cb, Cr = [0,255]$$

Then, the model identifies the biggest blob in the image by biggest blob analysis which gives the hand segment from the image removing the background. This work sets the background to black and the hand section retains the same color as the preprocessed image which this work denotes as binary coded color image.

2.2.3 Cropping and Normalizing

After the image segmentation, the image is cropped for desired portion of the hand. This work uses center cropping as the basis. Keeping the wrist at middle position of the image the hand segment is cropped out from wrist to the finger tips and 50 pixels towards wrist to elbow section of the hand to find out the shape of the hand. Then the cropped image is normalized. As different persons' hands have different thickness, it is necessary to normalize the cropped image. Then the image is again resized to 255×255×3 pixels.

2.2.4 Feature Extraction

Before extracting features, the images are converted into gray-scale images. This is denoted as binary coded gray image (Gray (BW)). From that gray image HOG features are extracted in this work. Gradient magnitude and gradient direction of an input image is calculated by HOG feature which is basically is a descriptor block [38]. Depending on the direction of the edges the HOG feature characterizes shape of an object. This work divides the window into 8×8 size blocks where every block is composed of 2×2 cells. Histogram of oriented gradient for each cell is then calculated. The combined histogram entrances form the descriptor blocks is referred to as Histogram of Oriented Gradient (HOG) descriptors. The steps in HOG feature extraction are:

a. Gradient Computation

One dimensional directive for horizontal and vertical masks gives the gradient of an image **I**,

$$\mathbf{D}_X = [-1\ 0\ 1] \ and\ \mathbf{D}_Y = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$$

Where, $\mathbf{D}_X$ and $\mathbf{D}_Y$ are horizontal and vertical masks respectively and get the X and Y derivatives using following process,

$$\mathbf{I}_X = \mathbf{I} * \mathbf{D}_X \ and\ \mathbf{I}_Y = \mathbf{I} * \mathbf{D}_Y$$

The degree of the gradient is,

$$|G| = \sqrt{\mathbf{I}_X{}^2 + \mathbf{I}_Y{}^2}$$

The orientation of the gradient is given by,

$$\theta = arctan \frac{\mathbf{I}_X}{\mathbf{I}_Y}$$

b. Orientation Binning

The second step consists of creating the cell histograms for the images. Based on the values found in the gradient computation, each pixel calculates a weighted vote for an orientation based histogram channel. The cells themselves are rectangular. The histogram channels are evenly spread over 0° to 180° or 0° to 360°, depending on whether the gradient is unsigned or signed.

c. Descriptor Blocks

In order to change the lightness, the gradient strength should be regionally normalized, which needs alignment of the cells together into larger spatially connected blocks. The HOG descriptor is then the vector of the elements of the normalized cell histograms from all of the block regions. These blocks generally overlap which means that each cell contributes more than once to the ultimate descriptor. For the entire block a normalization factor is then computed and all histograms within this block are normalized according to this factor. Once the normalization step is done, all the histograms will be concatenated in a single feature

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

vector. There are several approaches for block normalization. The normalization factor used in this work can be described as, let **v** be the non-normalized vector containing all histograms in a given block, $[\![V_K]\!]$ be its *k*-norm for *k*=1, 2 and *e* be some small constant. Then, the normalization factor *f* is:

$$L1\text{-}norm\text{: } f = \frac{\mathbf{v}}{[\mathbf{v}]+e}$$

$$L2\text{-}norm\text{: } f = \frac{\mathbf{v}}{\sqrt{[\mathbf{v}]^2+e^2}}$$

2.2.5 Classifying with the KNN classifier

K nearest neighbors (KNN) is a straightforward machine learning approach that stores all the available cases and classifies new instances based on the similarity measures that is the distance function [39]. In the training phase, an instance is classified by a majority vote of its neighbors, with the instance being assigned to the class most common amongst its *K* nearest neighbors measured by a distance function. This work uses Euclidian distance as similarity measure which is denoted by,

$$d(\mathbf{X}, \mathbf{Y}) = \sqrt{\sum_{i=1}^{k} (x_i - y_i)^2}$$

It runs through the dataset computing the distance between **X** and each of the training observations. The *K* points in the training data closest to **X** are called set A. *K* is usually odd to prevent tie situations. It then measures the conditional probability for each class.

$$\boldsymbol{P}(y = j|\mathbf{X} = x) = \frac{1}{k}\sum_{i \in A} I(y^{(i)} = j)$$

Finally, **X** is assigned to the class with the largest probabilistic value. KNN searches the learned training observations for the *K* instances that most closely resemble the new instance and assigns to the most common class. However, electing the ideal value for *K* is best done by first examining the dataset. A large *K* value is more exact as it reduces the overall noise but there is no guarantee. Therefore, 5-fold cross-validation is done on the training set to determine a good *K* value. The optimal *K* value is found at 7. This produces the best result in this dataset.

3. EXPERIMENTAL RESULT ANALYSIS

3.1 Dataset Description and Experimental Setup

The images of different signs were taken in the prepared dataset. From the dataset this work computed the feature vector. The training and testing ratio in the dataset was set as 55:45 ratio. The detail of the dataset is given in Table 1:

Table 1: Dataset description

Dataset	Signs	Number of Images	Pixel value	Image Type
Training Set	অ-ঔ-ক-হ	770	Unit8	JPG
Testing Set	অ.ঔ.ক.হ	630	Unit8	JPG

All the images in the dataset were in RGB color space. The images were processed with the proposed preprocessed approaches. For the analysis of other models, further processing was done on the dataset as needed for the respective models. Fig. 4 shows the step by step transformation of the dataset images using the proposed model. The normalized image was then resized again to 255×255×3. Then, the image was converted to grayscale. From the gray image HOG features were extracted. These HOG features were used to train the classification model that is the KNN classifier in this case. The extracted HOG features are shown in Fig. 5.

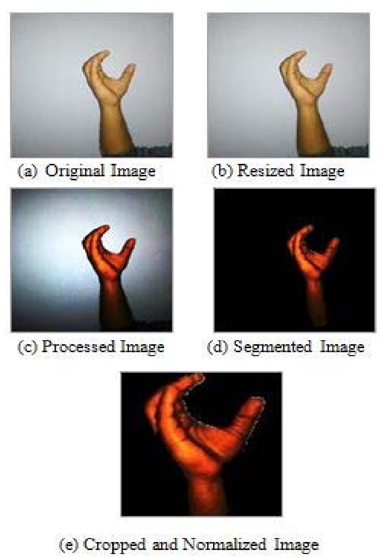


Figure 4: Pre-processing of the dataset images

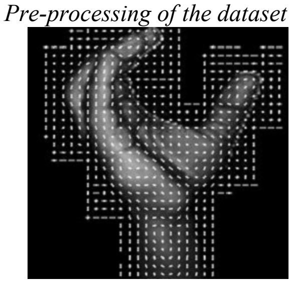


Figure 5: Extracted HOG features

Journal of Theoretical and Applied Information Technology
15th March 2020. Vol.98. No 05

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

To compare the accuracy of this model with other models some further processing was done as needed for corresponding models. Some of these processing includes converting the image to binary image and then extracting HOG and SURF features from the image. Detecting Canny edges from the image and then extracting HOG, SURF and line features from the image. The extracted features from the processed images are shown in the Fig. 6.

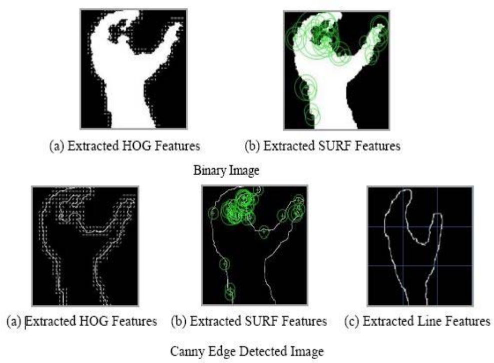


Figure 6: Different features on different image types

3.2 Classification Performance and Evaluation

The proposed work uses SVM (Support Vector Machine), ANN (Artificial Neural Network), KNN (K-Nearest Neighbors) and Category Classifier algorithms on the processed dataset to find out the most dependable classifier and processed data for Bengali sign language recognition. A classical multiclass classification method called one-versus-one (oVo) or pairwise decomposition is used for the SVMs [40]. It evaluates all possible pairwise classifiers and thus induces a $k*(k-1)/2$ individual binary classifiers. Applying each classifier to a test example would give one vote to the winning class. A test example is labeled to the class with the most votes. For the kernel, this work uses only the linear kernel. The artificial neural network used in this work is a two layered model. First is the input layer which takes the input from the database and second is the output layer which forms the outputs. There is three hidden layer of size 100×100×100. In this paper, there are 35 categories of outputs for each sign. The neural network consists of N numbers of input perceptron and there are 35 output perceptron. The category classifier uses Bag of Features where the Bag is generated from the extracted SURF features and K-means clustering of the features [20]. The accuracy of these models on the binary image-set is given in Table 2 while the accuracy of these models for canny edge detected images for various features is given in Table 3. However, the accuracy of Category Classifier with SURF features for binary coded gray, binary and canny edge detected image is given in Table 4 while the accuracy of the proposed model and proposed feature extraction method with other classification models are given in Table 5.

Table 2: Test results for binary images

Name of the Signs	No. of Test Inputs	Models Used					
		SVM		KNN		ANN	
		Features Used					
		HOG	Line	HOG	Line	HOG	Line
অ	18	12	N/A	16	N/A	0	N/A
আ	18	12	N/A	16	N/A	14	N/A
ই	18	14	N/A	18	N/A	18	N/A
ঈ	18	10	NA	9	NA	0	NA
ঋ	18	8	N/A	18	N/A	10	N/A
এ	18	6	N/A	10	N/A	10	N/A
ঐ	18	12	N/A	14	N/A	6	N/A
ও	18	18	N/A	18	N/A	12	N/A
ঔ	18	18	N/A	18	N/A	18	N/A
ক	18	14	N/A	14	N/A	10	N/A
খ	18	14	N/A	14	N/A	14	N/A

গ	18	8	N/A	15	N/A	14	N/A
ঘ	18	12	N/A	12	N/A	10	N/A
ঙ	18	6	N/A	12	N/A	12	N/A
চ	18	12	N/A	14	N/A	0	N/A
ছ	18	14	N/A	18	N/A	6	N/A
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
হ	18	18	N/A	18	N/A	15	N/A
Accuracy		**82.7%**	**N/A**	**86.5%**	**N/A**	**69.8%**	**N/A**

Table 3: Test results for canny edge detected images

Name of the Signs	No. of Test Inputs	Models Used					
		SVM		KNN		ANN	
		Features Used					
		HOG	Line	HOG	Line	HOG	Line
অ	18	10	8	12	8	0	12
আ	18	12	10	10	10	8	18
ই	18	18	10	4	2	16	17
ঈ	18	18	18	16	18	14	18
ঋ	18	4	9	14	8	6	6
এ	18	2	1	4	4	8	8
ঐ	18	6	5	10	8	0	14
ও	18	14	16	8	12	0	18
ঔ	18	12	12	18	14	18	14
ক	18	8	6	10	10	10	18
খ	18	18	16	18	18	6	18
গ	18	10	6	12	12	8	18
ঘ	18	16	14	16	12	6	14
ঙ	18	2	4	12	12	6	15
চ	18	10	8	14	8	8	18
ছ	18	15	6	18	8	18	6
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
হ	18	11	14	12	18	12	2
Accuracy		**73.5%**	**72.1%**	**79.6%**	**76.3%**	**65.5%**	**68.7%**

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 4: Test results for category classifier

Name of the Signs	**No. of Test Inputs**	**Category Classifier**		
		Bag of Features		
		Binary	**Edge**	**Gray (BW)**
অ	18	4	4	6
আ	18	18	18	18
ই	18	8	6	10
ঈ	18	6	6	7
ঋ	18	16	14	17
এ	18	8	6	10
ঐ	18	10	12	12
ও	18	12	10	10
ঔ	18	13	8	13
ক	18	18	18	18
.	.	.	.	.
.	.	.	.	.
Accuracy		**75.6%**	**74.5%**	**79.2%**

Table 5: Test results binary coded gray image

Name of the Signs	**No. of Test Inputs**	**Models Used**		
		SVM	**KNN**	**ANN**
		Features Used		
		HOG	**HOG**	**HOG**
অ	18	15	**15**	2
আ	18	18	**18**	16
ই	18	14	**14**	18
ঈ	18	14	**18**	0
ঋ	18	10	**10**	12
এ	18	14	**14**	16
ঐ	18	12	**12**	10
ও	18	18	**18**	13
ঔ	18	14	**12**	18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ক	18	14	**14**	10
খ	18	14	**18**	14
গ	18	8	**8**	14
ঘ	18	18	**18**	10
ঙ	18	16	**16**	12
চ	18	18	**18**	0
ছ	18	18	**18**	4
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
হ	18	18	**18**	15
Accuracy		**89.8%**	**91.1%**	**73.5%**

From the results shown in the aforementioned tables, this is evident that the most accuracy is gained from KNN classifier with extracted HOG features from binary coded gray images. It produces 91.1% accuracy over the test set. Multiclass SVM (OvO) also gives accuracy quiet close to the KNN classifier and the required time for SVM is faster to process the test set. So, SVM model can easily replace the KNN model with some further kernel setup. But for dependency KNN model is most appreciated. The accuracy of these models is shown in Fig. 7. Based on the experimental results it can be said that the proposed model gives much better result than any of the other discussed models. The reason that gray images gives better result in every case is that the binary and edge images become ambiguous for different signs. But the gray image contains the crucial information to detect different signs with more accuracy. The accuracy of the proposed model can be shown in the confusion matrix of Fig. 8. Although the proposed model gives better accuracy, it is also prone to error because of the hand shapes for many signs' are quite similar. And to completely remove the ambiguity is very hard. With proper training set and combining the features discussed in this work, we can arrive at a better approximation of the result. Some of the results found by the proposed model are given in Table 6:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

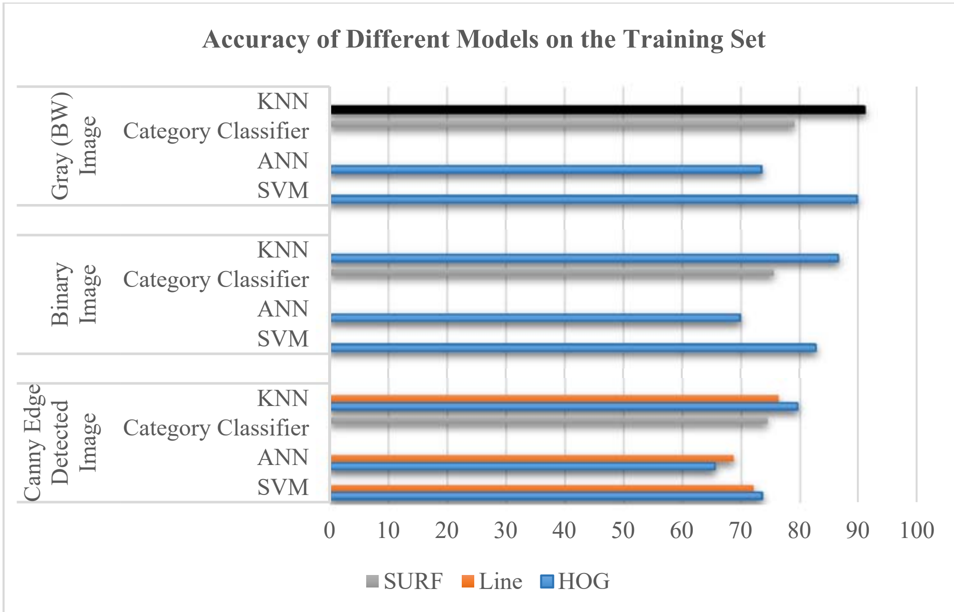


Figure 7: Comparative Results of Different Models

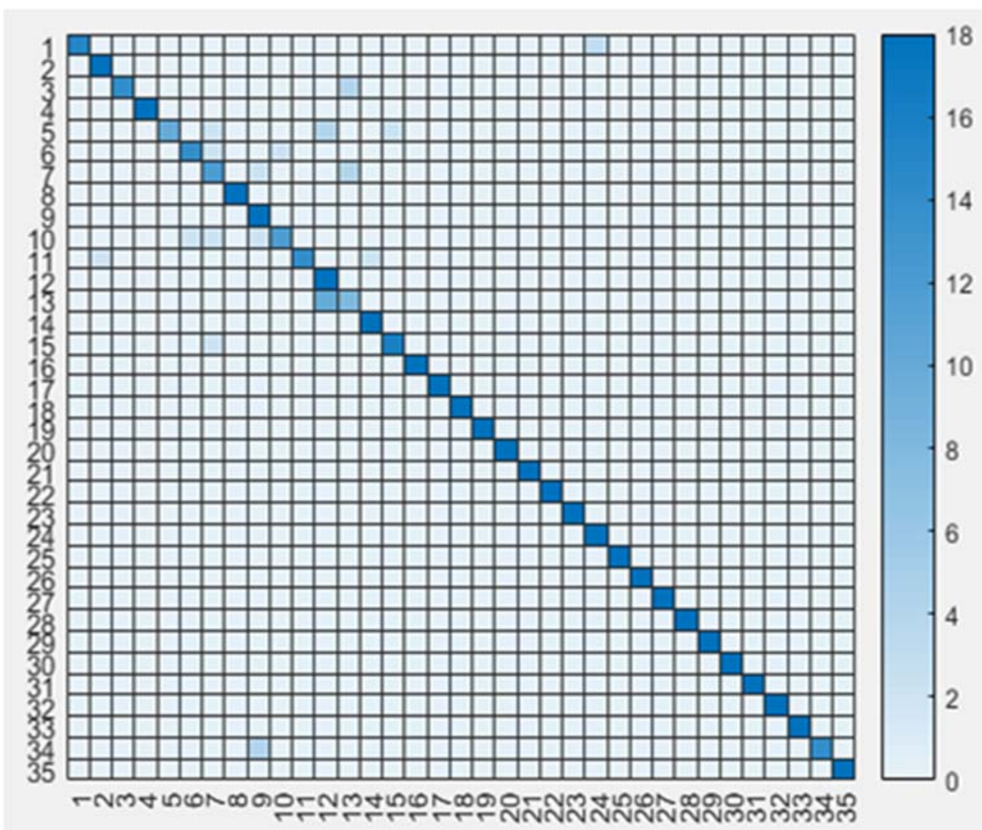
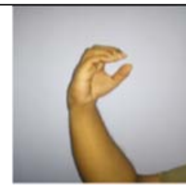
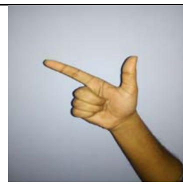


Figure 8: Confusion Matrix of the Proposed Model

Journal of Theoretical and Applied Information Technology
15th March 2020. Vol.98. No 05

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 6: Recognition results

Image	Recognized Sign	Image	Recognized Sign
	অ		ক
	থ		জ

4. CONCLUSION

Sign language recognition is a very under-developed area in computer science and technology. Deaf and dumb people are under-appreciated in every area because of the communication gap between them and normal people. This work is for those people to contribute in the national economy of the nation. This work involves around a method that works with specially processed gray scale images to extract HOG features and then training it with KNN classifier to classify any unknown instances. The proposed model works with 91.1% accuracy. In real time, the proposed model also works with a great accuracy. In future, the proposed model can be implemented to recognize signs except alphabets. Real time motion signs can also be detected using this model with concatenating the feature vector.

REFRENCES:

[1] J. M. Jose, "The Difference between Animal and Human Communication", [Online]. Available: https://owlcation.com/stem/The-difference-between-animal-and-human-communication.

[2] Sutarman, M. A. Majid and J. M. Zain, "A Review on the Development of Indonesian Sign Language Recognition System", *Journal of Computer Science*, vol. 9, pp. 1496-1505, 2013.

[3] O. Mauricio, Z. Israel and P. Jesús, "Language Interpreter for the Deaf Core Using a Virtual Reality Data-Glove", *Journal of Theoretical and Applied Information Technology*, vol. 33, no. 2, pp. 135-141, 2011.

[4] Sutarman, M. A. Majid and E. I. Sela, "Performance Evaluation of Combined Consistency-Based Subset Evaluation and Artificial Neural Network for Recognition of Dynamic Malaysian Sign Language", *Journal of Theoretical and Applied Information Technology*, vol. 95, no. 11, pp. 2489-2496, 2017.

[5] I. Mahmud, T. Tabassum, M. P. Uddin, E. Ali, A. M. Nitu and M. I. Afjal, "Efficient Noise Reduction and HOG Feature Extraction for Sign Language Recognition", *IEEE International Conference on Advancement in Electrical and Electronic Engineering - (ICAEEE 2018)*, 22-24 November 2018, Bangladesh.

[6] M. J. Cheok, Z. Omar and M. H. Jaward, "A review of hand gesture and sign language recognition techniques"m *International Journal of Machine Learning and Cybernetics*, vol. 10, 2017.

[7] K. Yousaf, Z. Mehmood, T. Saba, A. Rehman, M. Rashid, M. Altaf and Z. Shuguang, "A Novel Technique for Speech Recognition and Visualization Based Mobile Application to Support Two-Way Communication between Deaf-Mute and Normal Peoples", *Wireless Communications and Mobile Computing*, vol. 2018, 2018.

[8] Lang and Simon, "Sign Language Recognition with Kinect", *Institute for Informatik Freie Universitat Berlin, Berlin: Freie Universitat Berlin,* 2011.

[9] Parton and B. Sue, "Sign Language Recognition and Translation: A Multi-disciplined Approach from the Field of Artificial Intelligence." *Journal of Deaf Studies and Deaf Education,* vol. 11, 2005.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[10] D. L. Jaffe, "Ralph: A fingerspelling hand" [Online]. Available: https://web.stanford.edu/group/rrd/TTran/ttralph.html

[11] T. Takahashi and F. Kishino, "Hand Gesture Coding based on Experiments using a Hand Gesture Interface Device", *ACM SIGCHI Bulletin*, vol. 23, issue. 2, pp. 67-74, 1991.

[12] T. E. Starner, "Visual Recognition of American Sign Language Using Hidden Markov Models", *Massachusetts Institute of Technology, Cambridge, Dept. of Brain and Cognitive Sciences,* 1995.

[13] D. Kelly, J. R. Delannoy, J. M. Donald and C. Markham, "Automatic recognition of head movement gestures in sign language sentences." *China-Ireland Information and Communications Technologies Conference.* 2009, Ireland.

[14] K. Weaver, H. Hamilton, Z. Zafrulla, H. Brashear, T. Starner, P. Presti and A. Bruckman, "Improving the language ability of deaf signing children through an interactive American Sign Language-based video game", *International Conference of the Learning Sciences*, 2010, Illinois.

[15] J. Forster, C. Schmidt and H. Ney, "Scientific Understanding and Vision Based Technological Development for Continuous Sign Language Recognition and Translation." *SignSpeak Project, Annual Public Report*, *Human Language Technology and Pattern Recognition Group RWTH Aachen University, Aachen, Germany*, [Online]. Available: http://www.signspeak.eu/

[16] D. S. Hossain Pavel, T. Mustafiz, A. I. Sarkar and M. Rokonuzzaman, "Geometrical Model Based Hand Gesture Recognition for Interpreting Bengali Sign Language Using Computer Vision", *International Conference on Computer and Information Technology (ICCIT)*, 2003.

[17] N. N. Choudhury and G. Kayas, "Automatic Recognition of Bangla Sign Language", *Ph.D. Dissertation*, *BRAC University*, Bangladesh, [Online]. Available: http://dspace.bracu.ac.bd/xmlui/handle/10361/2387, 2012.

[18] A. K. Sahoo and K. K. Ravulakollu, "Vision based Indian Sign Language Character Recognition." *Journal of Theoretical and Applied Information Technology,* vol. 67, no. 3, pp. 770-780, 2014.

[19] M. V. D. Prasad and P. V. V. Kishore, E. Kiran Kumar, D. Anil Kumar, "Indian Sign Language Recognition System using New Fusion based Edge Operator." *Journal of Theoretical and Applied Information Technology,* vol. 88, no. 3, pp. 574-584, 2016.

[20] A. M. Jarman, S. Arshad, N. Alam and M. J. Isalm, "An Automated Bengali Sign Language Recognition System based on Fingertip Finder Algorithm", *International Journal of Electronics & Informatics*, vol. 4, no. 1, pp. 1-10, 2015.

[21] K. Datta, B. Sarkar, C. D. Datta, D. Sarkar, S. J. Dutta, I. D. Roy, A. Paul, J. U. Molla, and A. Paul, "Interface Between Bangla Text and Sign Language", 2009.

[22] B. Sarkar, K. Datta, C. D. Datta, D. Sarkar, S. J. Dutta, I. D. Roy, A. Paul, J. U. Molla, and A. Paul, "A translator for Bengali text to sign language." *2009 Annual IEEE India Conference*, 2009, India.

[23] S. Islam, S. S. S. Mousumi, A. S. A. Rabby, S. A. Hossain, and S. Abujara, "A Potent Model to Recognize Bengali Sign Language Digits Using Convolutional Neural Network", *Procedia Computer Science,* vol. 143, pp. 611-618, 2018.

[24] F. N. Arko, N. Tabassum, T. R. Trisha, F. Ahmed, D. J. Uddin, "Bangla Sign Language Interpretation Using Image Processing", *B .Sc. Thesis, BRAC University*, *Bangladesh*, 2017.

[25] A. K. Roy, A. K. Ghosh, A. S. Asa, M. P. Uddin, M. R. Islam and M. I.Afjal, "Bengali consonants-voice to text conversion using machine learning tool", *International Journal of Research in Computer Engineering and Electronics*, vol. 6 issue. 2, 2017.

[26] S. M. Syfullah, Z. B. Zakaria, M. P. Uddin, M. F. Rabbi, M. I. Afjal and A. M. Nitu, "Efficient Vector Code-book Generation using K-means and Linde-Buzo-Gray (LBG) algorithm for Bengali Voice Recognition", *IEEE International Conference on Advancement in Electrical and Electronic Engineering-(ICAEEE 2018)*, Bangladesh, 2018.

[27] M. A. A. Mamun, M. A. Islam, M. I. Afjal, M. A. A. Mamun, U. A. M. E. Ali and M. P. Uddin, "Feature Extraction Using Uniform Pattern based Standard Deviation for Face Recognition", *HBRP Advancement in Image Processing and Pattern Recognition*, vol. 1 issue. 2, pp. 32-38, 2018.

[28] M. Rahman, I. Mahmud, M. P. Uddin, M. I. Afjal, M. A. Habib and F. Kabir "Handwritten Bengali Character Recognition through Geometry Based Feature Extraction", *Journal of Theoretical and Applied Information*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Technology, (IF=0.67), vol. 97, no. 23, pp. 3570-3582, 2019.

[29] A. C. Roy, K. Hossin, M. P. Uddin, M. A. A. Mamun, M. I. Afjal and A. M. Nitu, "Detection and Classification of Geometric Shape Objects for Industrial Applications", *HBRP Advancement in Image Processing and Pattern Recognition*, vol. 1 issue. 2, pp. 11-19, 2018

[30] M. H. O. Rashid, M. A. Mamun, M. A. Hossain and M. P. Uddin, "Brain Tumor Detection Using Anisotropic Filtering, SVM Classifier and Morphological Operation from MR Images", *IEEE International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, Bangladesh, 2018.

[31] S. M. M. Hasan, M. A. Mamun, M. P. Uddin and M. A. Hossain, "Comparative Analysis of Classification Approaches for Heart Disease Prediction", *IEEE International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, Bangladesh, 2018.

[32] M. F. Rabbi, M. P. Uddin, M. A. Ali, M. F. Kibria, M. I. Afjal, M. S. Islam and A. M. Nitu, "Performance Evaluation of Data Mining Classification Techniques for Heart Disease Prediction", *American Journal of Engineering Research (AJER)*, vol. 7, issue. 2, pp-278-283, 2018.

[33] M. P. Uddin, M. A. Mamun and M. A. Hossain, "Feature Extraction for Hyperspectral Image Classification", *IEEE 5th Region 10 Humanitarian Technology Conference (R10HTC)*, pp. 379-382, Bangladesh, 2017.

[34] M. P. Uddin, M. A. Mamun and M. A. Hossain, "Improved Feature Extraction Using Segmented FPCA for Hyperspectral Image Classification", *IEEE 2nd International Conference on Electrical & Electronic Engineering (ICEEE)*, Bangladesh, 2017.

[35] M. P. Uddin, M. A. Mamun and M. A. Hossain, "Segmented FPCA for Hyperspectral Image Classification", *IEEE 3rd International Conference on Electrical Information and Communication Technology (EICT)*, Bangladesh, 2017.

[36] M. P. Uddin, M. A. Mamun and M. A. Hossain, "Effective feature extraction through segmentation-based folded-PCA for hyperspectral image classification", *International Journal of Remote Sensing (IJRS)*, vol. 40, 2019- Issue. 18, pp. 7190-7220, 2019.

[37] T. M. Mahmoud, "A new fast skin color detection technique", *International Journal of Computer and Information Engineering*, vol. 2, no. 7, pp. 2254-2358, 2008.

[38] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005, USA.

[39] S. Sayad, "K Nearest Neighbors - Classification", [Online]. Available: https://www.saedsayad.com/k_nearest_neighbors.htm

[40] S. Knerr, L. Personnaz, G. Dreyfus, "Single-layer learning revisited: a stepwise procedure for building and training neural network", *Neurocomputing Algorithms, Architectures and Applications*, vol. 68, 1990.