Visual Recognition of Bengali Sign Language using Artificial Neural Network

Md. Abdur Rahim
Lecturer, Department of CSE
Pabna University of Science and Technology, Pabna, Bangladesh

Tanzillah Wahid
Senior Lecturer,
Department of CSE, Uttara University, Dhaka, Bangladesh

Md. Khaled Ben Islam
Lecturer, Department of CSE Pabna University of Science and Technology, Pabna, Bangladesh

ABSTRACT

This paper presents an overview of visual recognition of Bengali Sign Language. In this paper we learn and detect a sequence of sign words and recognize the sign language that are understandable to the deaf and hearing impaired people to help normal people understand the meaning of these words. The research discusses the characteristics of the human sign languages, the requirements and difficulties behind visual sign recognition, how to deal with others persons and the different techniques used in the sign language recognition. The project consists of two major parts, namely the learning part and the detection part. The system takes the sign images as its input. First sign images are learnt by the proposed system. When a sign image is given for recognition, the detection part identifies the image with the help of previously learned images. For learning and detection we have used back propagation algorithm of Artificial Neural Network. We believe that this research will be of much help to express their thoughts and feelings between the deaf people and the normal people.

Keywords

Sign Language, ANN, Back Propagation Algorithm.

1. INTRODUCTION

Many people are hearing impaired in our country. Sign languages are used by the hearing impaired people to mainly communicate with other peoples of their kind, and also to some extent with their friends, families as well. Many of these people are intelligent enough to share their opinions and work with rest of the population to move ahead our country. They participate in different development organizations in different sectors. But there is a communication gap between the hearings impaired and/or deaf people and the rest of the people in our country. They can communicate with their family members as they also learn sign language to some extent in the process of observing and interacting with the disabled person in the family day after day. They can communicate with other people who know sign language. But, it is difficult for normal people to communicate with them when hearing impaired and/or deaf people work with normal people in different organization and/or development sectors. The communication between the hearing impaired people and the normal people are not well defined. It is important to communicate with hearing impaired and deaf people in an understandable way to share their opinions, to solve their problem and help them to be a part our normal day to day life. In this project we work on the Bengali Sign Language that helps both the hearing impaired & deaf people and the normal people to express their thoughts and feelings and to give them a better medium for communication and exchanges of ideas

[2]. This project follows the Bengali Sign Language dictionary to create the understandable sentences using the easy words. First, we will learn the images of these words sign language and store it. Then we can take as input a sign language image and compare it with our stored sign language words to recognize what the input image represents in sign language. This project has been done to make the people to recognize the sign language with an easy way and also to communicate with others.

2. SIGN LANGUAGE

A sign language (also signed language) is a language which, instead of acoustically conveyed sound patterns, uses visually transmitted sign patterns (manual communication, body language and lip patterns) to convey meaning— simultaneously combining hand shapes, orientation and movement of the hands, arms or body, and facial expressions to fluidly express a speaker's thoughts. Sign languages commonly develop in deaf communities, which can include interpreters, friends and families of deaf people as well as people who are deaf or hard of hearing themselves [4].In Bangladesh the method of teaching the hearing impaired and/or deaf children started more than three decades ago. Lip reading, use of sign language and speech training has been followed in teaching them. Though the basis of sign language is the same, yet every country has developed its own techniques according to its own language. Teachers who are working in this particular fields are trained in different countries where the follow their own methods and techniques, specially the sign language. So it is not expected that all of them will follow the signs. More over the hearing impaired children of the different parts of the country use their own sign language according to their needs and local conditions, but no step was undertaken for the hearing impaired in Bangladesh [2]. Considering all these facts and prevailing conditions of this particular field, steps have been taken by the Department of Social Services to develop a uniform sign language dictionary keeping in mind the language, social and cultural values and way of living in this country.

In this recognition system, sentences of the form "personal pronoun, verb, adjective, personal Pronoun" are to be recognized. This sentences such as I eat rice (Avwg fvZ LvB) structure emphasis the need for a distinct grammar for Bengali Sign Language recognition and allows a large variety of meaningful sentences to be randomly generated using words from each class. Table 1 shows the words chosen for each class [3]. At first a naive eye was used to avoid ambiguities in the selected signs, but this was shortly subsumed by the Table1.

TABLE 1. BSL VOCABULARY USED

পদ প্রকরণ (Parts of Speech)	শব্দ তালিকা (Words List)
সর্বনাম (Pronoun)	আমি/আমার (I/My), তুমি(You), আমরা/আমাদের(Our), তারা/তাদের (They) etc.
ক্রিয়া (Verb)	খাওয়া (Eat), খেলা করা(Play), ঘুমানো(Sleep), যাওয়া(Go), পড়া(Read), জানা (Know) etc.
বিশেষ্য (Noun)	বই(Book), টেবিল (Table), স্কুল (School), ভাত (Rice), পছন্দ করা (Like) etc.
বিশেষণ (Adjectives)	লাল (Red), নীল (Blue), সাদা (White), হলুদ (Yellow) etc.

2.1 Sign Languages' Relationships with Oral Languages

A common misconception is that sign languages are somehow dependent on oral languages, that is, that they are oral language spelled out in gesture, or that they were invented by hearing people. Hearing teachers in deaf schools, such as Thomas Hopkins Gallaudet, are often incorrectly referred to as "inventors" of sign language. Manual alphabets (finger spelling) are used in sign languages, mostly for proper names and technical or specialized vocabulary borrowed from spoken languages. The use of finger spelling was once taken as evidence that sign languages were simplified versions of oral languages, but in fact it is merely one tool among many. Finger spelling can sometimes be a source of new signs, which are called lexicalized signs. On the whole, deaf sign languages are independent of oral languages and follow their own paths of development. For example, British Sign Language and American Sign Language are quite different and mutually unintelligible, even though the hearing people of Britain and America share the same oral language. Similarly, countries which use a single oral language throughout may have two or more sign languages; whereas an area that contains more than one oral language might use only one sign language. South Africa, which has 11 official oral languages and a similar number of other widely used oral languages, is a good example of this. It has only one sign language with two variants due to its history of having two major educational institutions for the deaf which have served different geographic areas of the country [5].

3. BASIC CONCEPT OF NEURAL NETWORK

In general neural network are simply mathematical techniques designed to accomplish a variety of task. Neural networks can be configured in various arrangements to perform a range of tasks including pattern recognition, data mining, classification and process modeling. The later is the primary interest in this context and although the types and topologies of neural networks vary greatly in the field, by far the most commonly used type of neural network, particularly in process control, is feed-forward, back-propagation neural network [6]. Our basic computational element model neuron is often called a node or unit. It receives input from some other units, or perhaps from an external source. Each input has an associated weight w, which can be modified so to model synaptic learning. The unit computes some function f of weighted sum of its inputs:

$$y_i = f(\sum_j w_{ij} y_j) \quad \dots\dots\dots\dots\dots\dots\dots\dots (1)$$

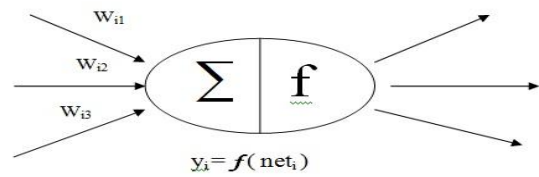Its output, in turn, can serve as input to other units.

Fig. 1. A Simple Artificial Neuron

- The weighted sum $\sum_j w_{ij} y_j$ is called the net input to unit i, often written net_i.
- Note that w_{ij} refers to the weight from unit j to unit i (not the other way around).
- The function f is the unit's activation function. In the simplest case, f is the identity function, and the unit's output is just its net input. This is called a linear unit.
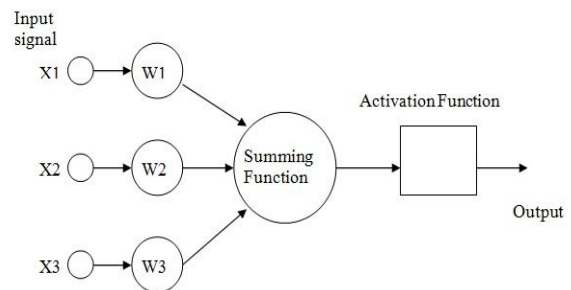

Fig. 2. Model of Computing Neuron

The architecture of the 3-layer neural network besides processing an input and an output layer also has an intermediary layers called hidden layer. The computational units of the hidden layer aids in performing useful intermediary computation before directing to the input to the output layer. The input layer neurons are linked to the hidden layer neutron and the weight on these links is referred to as input hidden layer weight. The hidden layer neurons are linked to the output layer neurons and weights. On these links are referred to as hidden-output layer weights. A model of multilayer network also called three-layer is shown below:
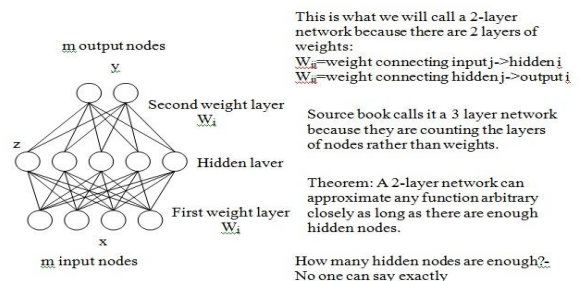

Fig. 3. Model of Multilayer Network

4. SIGN RECOGNITION USING BACK PROPAGATION ALGORITHM

The benefit of the middle-hidden layer allows ANN to develop its own internal representation of this mapping. Such a rich and complex internal representation capability allows the hierarchical network to learn any mapping and not just linearly separable ones. Les us consider the three-layer network with input layer having 'l' nodes, hidden layer having 'm' nodes and an output layer with 'n' nodes. We consider sigmoidal functions for activation functions for the hidden and output layers and linear activation function for input layer [7]. The number of neurons in the hidden layer may choose to lie between 1 and 21. The basic algorithm loop structure is given as

> Initialize the weights
>
> > Repeat
> >
> > For each training pattern
> >
> > Train on that pattern
> >
> > End
>
> Until the error is acceptably low

Algorithm 1 (Sign Language Detection Algorithm)

Step 1:

Normalize the inputs and outputs with respect to their maximum values. It is proves that the neural networks work better if input and outputs lie between 0-1. for each training pair, assume there are 'l' inputs given by $\{I\}_r \atop l \times 1$ and 'n' outputs $\{O\}_o \atop n \times 1$ in a normalized form.

Step 2:

Assume the number of neurons in the hidden layer to lie between $l < m < 2l$

Step 3:

[V] Represents the weights of synapses connection input neurons and hidden neurons and [W] represent weight of synapses connecting hidden neurons and output neurons. Initialize the weights to small random values usually from -1 to 1. For general problems, λ can be assumed as 1 and the threshold values can be taken as zero.

> [V]=[random weights]
>
> [W]=[random weights]

$$[\Delta V]^o = [\Delta W]^o = [O] \ldots\ldots\ldots (2)$$

Step 4:

For the training data, presents one set of inpus and outputs. Present the pattern to the input layer $\{I\}_r$ as inputs to the input layer. By using linear activation function, the output of the input layer may be evaluated as

$$\{O\}_I \atop l \times 1 = \{I\}_I \atop l \times 1 \quad \ldots\ldots(3)$$

Step 5:

Compute the inputs to the hidden layer by multiplying corresponding weights of the synapses as $$\{I\}_H \atop m \times l = [V]^T \atop m \times l \{O\}_I \atop l \times 1 \ldots\ldots\ldots(4)$$

Step 6:

Let the hidden layer units evaluate the output using the sigmoidal function as

$$\{O\}_H = \left\{ \frac{1}{(1 + e^{-I_{ni}})} \right\} \ldots\ldots\ldots(5)$$

Step 7:

Compute the inputs to the output layer by multiplying corresponding weight of synapses as

$$\{I\}_O \atop n \times 1 = [W]^T \atop n \times m \{O\}_H \atop m \times 1 \ldots\ldots(6)$$

Step 8:

Let the output units evaluate the output using sigmoidal function as

$$\{O\}_O = \left\{ \frac{1}{(1 + e^{-I_{oi}})} \right\} \ldots\ldots\ldots(7)$$

> The above is the network output.

Step 9:

Calculate the error and the difference between the network output and the desired output as for the ith training set as

$$E^P = \frac{\sqrt{\sum (T_j - O_{Oj})^2}}{n} \ldots\ldots\ldots(8)$$

Step 10:

> Find {d} as

$$\{d\} = \{(T_k - O_{ok})O_{ok}(1 - O_{ok}) \atop n \times 1 \ldots\ldots(9)$$

Step 11:

Find [Y] matrix as

$$[Y]_{mxn} = \{O\}_{mx1\ H}\{d\}_{1xn} \ldots (10)$$

Step 12:

Find

$$[\Delta W]_{mxn}^{t+1} = \alpha[\Delta W]_{mxn}^{t} + \eta[Y]_{mxn} \ldots (11)$$

Step 13:

Find

$$\{e\}_{mx1} = [W]_{mxn}\{d\}_{nx1} \ldots (12)$$

$$\{d^*\} = \{e_i(O_{Hi})(1-O_{Hi}) \ldots (13)$$

Step 14:

Find

$$[\Delta V]_{1xm}^{t+1} = \alpha[\Delta V]_{1xm}^{t} + \eta[X]_{1xm} \ldots (14)$$

Step 15:

$$[V]^{t+1} = [V]^{t} + [\Delta V]^{t+1} \ldots (15)$$

$$[W]^{t+1} = [W]^{t} + [\Delta W]^{t+1} \ldots (16)$$

Step 16:

Find error rate as

$$\text{Error rate} = \frac{\sum E_p}{nset} \ldots (17)$$

Step 17:

Repeat steps 4-16 until the convergence in the error rate is less than the tolerance value.

5. RESULTS AND DISCUSSION

Bengali Sign Language recognition is not a simple problem since a new image of a sign seen in the recognition phase is usually different from the image seen by the learning phase. Although a sign image has unique features for sign language recognition there are several sign uses by the human. The sign image depends on viewing conditions and environment. In addition the sign image changes according to the expressions. Bengali Sign Language recognition project, which is flexible and efficient, should be able to solve the problems. This section reveals results obtained by the image detector that has been developed. In this paper the detector has been tested on images that contain sign image and every image contains only one sign. The standard size of the image is 128 x 128 or 256 x 256. A simple approach to extracting the information contains

in an image of sign is somehow to capture the variation in collection of sign image. To capture sign image is to simplify scan in physical image samples and make computerized copy of that image. The sign images are collected from different persons. From each person some samples are captured and these are scanned through a scanner. Figure shows some samples of the sign images [1].

Fig. 4. Sample Sign Imges

In order to train the difficult stages of the research, the Back propagation algorithm requires to be fed with positives examples- that are images of sign. Our developed matlab-script is capable of conditioning the sign images into positives examples were constructed. The script does the following:

- Opens a sign images (images with in gray scale)
- Display the image and places a bounding box around the sign and learn the image

- Open the tested sign image which we do not understand
- Detect the image if the image is the same with the sign image then display match it with its grammar else no match found.
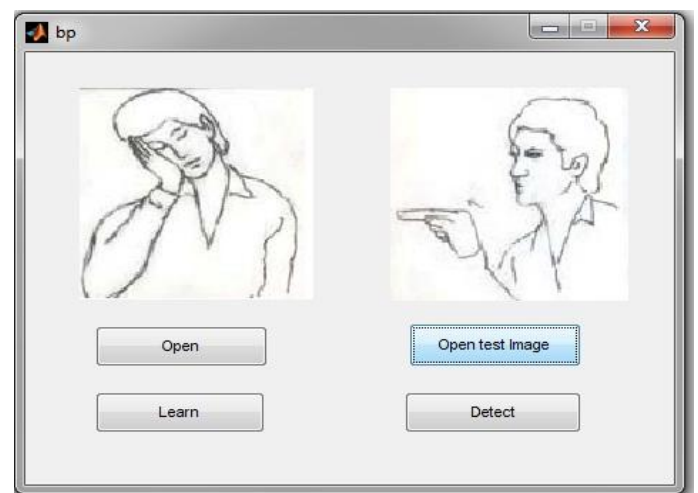

Fig. 5. Learn and Detect the Sign Image

The following table shows overall sign words learning and detection rate:

TABLE 2. DETECTION RATE OF THE PROJECT

Number of Sign word images learned	No of tested images	Detected Images	Undetected images	Detection Rate
2000	2000	1980	20	99%

It is expected from sign language recognition system to acquire high accurate recognition rate while detection rate should be too low. In the table-II, the detection rate is 99%. We detect the sign word one after another and recognize the sign language. From the experimental result, it is seen that the project satisfies all requirements to recognize the sign language.

6. CONCLUSION AND FUTURE IMPROVEMENT

6.1 Conclusion

In this study, we presented an off-line Bengali Sign Language recognition project which is based on artificial neural network using Back propagation algorithm. The purpose of this research is to implement Bengali Sign Language recognition. Through use of Back propagation algorithm low error rates were achieved on both learning an image and testing image. The accuracy of the system is 99% by using Back propagation detection algorithm.

6.2 Future Improvement

It is obvious that the result of this Bengali Sign recognition system but due to some limitation we were not able to implement all portion of the project yet. The main improvements will pursue the performances, recognize of multiple images. What is more the following are few points that we are going to address the near future [8].

- Multiple Gestures: we limited our gesture vocabulary to a simple set of some gesture, but there are many interesting application that involve more gestures.

- The learning process could be improved by introducing the advantage of decision trees into the learning via Backpropagation.
- Another interesting work would be to try other kind of features. For this, many solutions can be tried.
- Let Backpropagation algorithm tells how many and which examples are classified correctly after each new feature is added.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] Mohammad Osiur Rahman and Hassan Basri, "Real Time Road Sign Recognition System Using Artificial Neural Networks for Bengali Textual Information Box", European Journal of Scientific Research, 2009.

[2] Bangladesh National Federation Of the Deaf, "Bengali Sign Language Dictionary" Published 1994, Re-print 1997

[3] Thad Starner and Alex Pentland, Visual Recognition of American Sign Language Using Hidden Markov Model, NJ: IEEE Press

[4] Sign Language, History of Sign Language," Search in Wikipedia, the free encyclopedia"

[5] Different between Sign Language and Oral Language, " Search in Wikipedia, the free encyclopedia".

[6] R Beale and T Jackson," Neural Computing: An Introduction"

[7] S.RAJASEKARAN, G.A. IJAYALAKSHMI, "Neural Networks, Fuzzy logic and Genetic Algorithm Synthesis and Application.

[8] Jeffrey C. Liter and Heinrich H. Bulthoff, " An Introduction to Object Recognition", Max-Planck-Institute fur biologsche Kybernetik, Germany November 1996.