Recognizing Bengali Sign Language Gestures for Digits in Real Time using Convolutional Neural Network

Khalil Ahammad[*1], Jubayer Ahmed Bhuiyan Shawon[#2], Partha Chakraborty[*3], Md. Jahidul Islam[*4], Saiful Islam[#5]

Department of Computer Science and Engineering
[*]*Comilla University, Cumilla-3506, Bangladesh*
[#]*Bangladesh Army International University of Science and Technology, Cumilla, Bangladesh*
[1]khalil@cou.ac.bd
[2]shawon01821@gmail.com
[3]partha.chak@cou.ac.bd
[4]jahidul479@gmail.com
[5]defiant.saif@gmail.com

***Abstract*—Recognizing sign language gestures for different languages has been found as a promising field of research that explores the possibility of communication by interpreting various signs and translating them into text or speech. Establishing a better communication way between deaf-mute people and ordinary people is the prime objective of this research arena. There are many existing Sign Language Recognition (SLR) systems throughout the world and these SLR systems are implemented using various methods, tools and techniques with a view to achieving better recognition accuracy. This research work aims at applying the concept of Convolutional Neural Network (CNN) for recognizing Bengali Sign Language gesture images for digits only in real time. Bengali sign language images for digits are collected from different individuals and the CNN model is trained with these images after performing several pre-processing tasks i.e. resizing to a specific dimension, converting these RGB images to the gray scale images, finding the equivalent binary images and rotating the images into different degrees both in left and right direction. The experiment is conducted using two major techniques. Firstly, the model has been trained with the dataset containing the equivalent binary images of the row images collected directly from different individuals. Secondly, the dataset is enriched by rotating all the images into 3°, 6°, 9°, 12°and 15°in both left and right directions. After applying the rotation technique, the recognition accuracy is found to be increased significantly. The maximum recognition accuracy of the proposed CNN model with the dataset without image rotation technique is 94.17% whereas the recognition accuracy is 99.75% while including the rotated images in the dataset.**

***Index Terms*—Bengali Sign Language Recognition, Convolutional Neural Network, Image Recognition, Real Time Recognition**

I. Introduction

WHILE there are many different types of gestures, the most structured sets belong to the sign languages. In sign language, each gesture already has an assigned meaning, and strong rules of context and grammar may be applied to make recognition tractable [1]. It is one of the most natural means of exchanging information for the hearing impaired. It is a kind of visual language via hand and arm movements accompanying facial expressions and lip motions [2]. Very few people understand sign language. Moreover, contrary to popular belief, it is not an international language. Obviously, this further complicates communication between the deaf community and the hearing majority. The alternative of written communication is cumbersome because the deaf community is generally less skilled in writing a spoken language [10]. Sign language recognition aims to provide an efficient and accurate mechanism to translate sign language into text or speech. The formation of Bengali Sign Language is structurally dissimilar from other nations' sign languages. Generally, both hands are used to accomplish the Bengali Sign Language [3]. Deaf and Mute(DM) people suffer from hearing and speech impairment and use sign language to express their feelings. In social activities, the communication between the DM and the general people is hard because usually, the sign language is not understandable by the general people. Only a few general people who have learned the sign language can understand and translate it for the general ones. The DM people also cannot understand what the general people say as well as the lip reading too [3]. Considering the necessity of Bangladeshi sign language recognition (BdSL), it becomes one of the challenging topics in the field of computer vision and machine learning. Previously, few approaches tried to resolve this issue and received a respectable acceptance. But most of the research works have constraints and the least accuracy in the performance which arises the necessity to introduce a new method that simultaneously received continuous data from deaf people and produce a successful result with significant performance. In previous works, researchers captured data from static images which is a major constraint in the recog-

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

nition system. So, considering all the constraints from the previous works, Convolution Neural Network (CNN) has been applied with multiple hidden layers, and neurons are built in a network. It will execute the training and testing data set over this network and achieve the least error rate which enhances the recognition of Bengali Sign Language. This method will improve the performance and accuracy rate in the field of Bengali Sign Language Recognition [4]. The purpose of this work is to contribute to the field of automatic sign language recognition. This model focuses on the recognition of the signs or gestures in real time also [11].

II. Related Work

The recognition of sign language has inaugurated with the help of electronics devices at the end of 1990. During last two decades, researchers introduced new algorithms to enhance the method of recognition and developed robust approaches. But still, now the researchers are working on optimizing the recognition rate and reducing the error factor. The main approach is to consider a static image from the continuous video processing or the sample still images. In future, hand movements, body movements and facial expression will be introduced [4]. Some previous works on sign language recognition focuses on finger-spelling recognition and isolated sign recognition. This work uses a layered neural network, recurrent neural network, instance-based learning (IBL), dynamic programming matching (DPM), or rule-based matching (RBM). Previous works also focused on continuous sign recognition. Continuous dynamic programming matching (CDPM), rule-based matching (RBM), or the hidden Markov model (HMM) were used to recognize signed words from the inputted gestures of signed sentences. The main approach is to decide whether the gesture is represented by one hand or two hand [5]. Following a similar path to early speech recognition, many previous attempts at machine sign language recognition concentrate on isolated signs or finger-spelling. Space does not permit a thorough review but in general, most attempts either relied on instrumented gloves or a desktop-based camera system and used a form of template matching or neural nets for recognition [1]. From previous work, Haar-like feature-based cascaded classifiers have been used to detect the probable hand area from the captured image frames. Skin color-based segmentation method is used to extract hand signs and then extracted hand images are converted to binary images that are used as either training or testing images. Finally, the K Nearest Neighbor classifier is used to recognize both. This system cannot properly segment the hand area if some objects rather than hands have skin like colors and cannot properly distinguish some signs [6]. To implement BdSLR, they have exploited feature extraction along with the NCL algorithm for training that is capable enough to perform good recognition. BdSLR has been implemented involving different numbers of individual NNs in NCL. The limitation is, it is not real-time recognition and don't have enough data set [7]. A framework was showed for the Bengali hand sign recognition method using a support vector machine. Here, the first images are converted to HSV color space. Subsequently, features are extracted from the segmented image by Gabor filter and then KPCA is applied to reduce the dimensionality. Finally, SVM is used to recognize the hand sign. The limitation includes the ability not to work with a larger data set and only work for a single image, can't work for video clips [3]. In the Bengali sign language, very few research works have been established in recent years. To enhance the technique of Bengali sign language recognition, it is recommended to research other sign languages which have advanced techniques on hand gesture detection and modified classification [4].

III. Proposed Model

As Bengali Sign language recognition so, the main task is to propose a system that recognize images more accurately. Image recognition technology has a great potential of wide adoption in various industries. In fact, its not a technology of the future, but its already present. Such corporations and startups like Tesla, Google, Uber, Adobe Systems etc. heavily use image recognition.

There are different types of methods used for image classification. The proposed system used Convolution Neural Network for this research. There are some specific reasons to choose this model in this research. Neural Network and Deep learning are required to understand CNN properly.

The proposed system used a feed forward artificial neural network. Feed forward ANN means the output from one neuron always goes to another neuron as input. The flow of data is never back to the previous neuron or layer. The system was built by using a CNN with different layers and that can classify Bengali sign digits accurately.

A. Proposed Methodology

The proposed system for recognizing Bengali sign language (Digit) mainly has three parts:

1) Data collection/Acquisition
2) Data preprocessing and
3) Training and testing by CNN

The block diagram of proposed system is given in fig 1:

1) Data Collection: Data acquisitionăis the process of sampling signals that measure real world physical conditions and converting the resulting samples into digital numeric values that can be manipulated by a computer. In that state the images of BdSL (Bangladeshi sign language) digits were captured manually from 180 persons and 160 images were used in this proposed system. The proposed system used 9 signs of Bengali sign language (Digit 0 to Digit 9 except for Digit 5) as input to train the proposed CNN model where each of the sign has 160 samples that makes total collection of samples. The proposed model can recognize all these 9 signs in real time also.

In order to test the model for different Bengali digit signs, the signs have to be shown on the specific frame using web cam. The images were captured at a specified distance (typically 1.5 -2 ft) [12] between camera and signer. The signs used to train the CNN model is shown in fig 2.

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

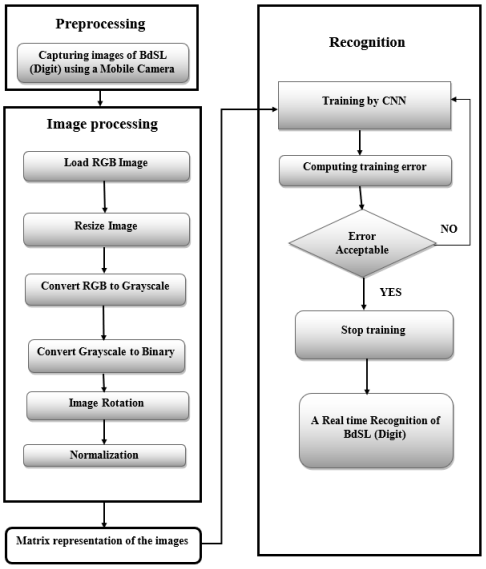

Fig. 1. System architecture

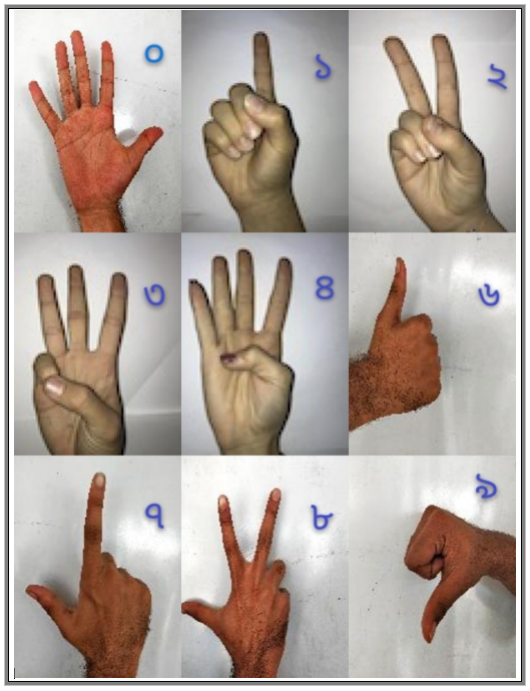

Fig. 2. Captured Bengali Digit Sign

2) Pre-processing: Data pre-processing is a data mining technique and it is the first stage for the proposed system. In case of the proposed system data pre-processing converts the raw images to an understandable format for training the proposed CNN model because raw images can have noise, inconsistent data etc. Pre-processing include Image conversion, cleaning, rotating, normalization etc. Pre-processing cleans the arbitrary images into common shape or form that makes appropriate to feed into classifiers [8].

Different step of data pre-processing is discussed below:

Step 1: At first, all the images were loaded to python code defining a specific path. For example:

path1 =r"D:\NoRo"

path2 =r"D:\NoRoOut"

Step 2: Then all the images (RGB) were resized into a specific shape 64*64 = length*width. This is shown in the following figure:

Fig. 3. RGB image (size=64*64)

Step 3: Converting the RGB images into gray scale image and the resultant image is shown in the following figure:

Fig. 4. Gray Scale image (size=64*64)

Step 4: Converting Gray scale images into binary image and the binary image is shown in the following figure:

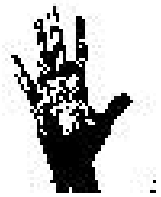

Fig. 5. Binary image (size=64*64)

Step 5: Rotating all images into different angles for increasing the training and validation accuracy. The following figure reflects the action:

Fig. 6. Binary image, Left Rotation=12(size=64*64)

Step 6: Then all the 64*64 images are converted to NumPy array of flatten image or all images converted into matrix form.

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

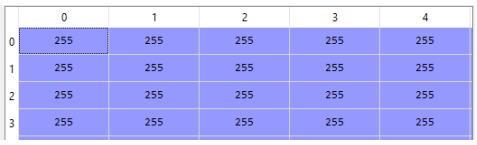

Fig. 7. Matrix form of image

Suppose 4 images = 4 rows in a matrix and (64*64) column size.

Step 7: As the model work to predict 9 different labels or classes then all NumPy images are labeled or classified to a specific class based on their similarity such as all images of digit 0 is located into the same label using python list. Suppose 3 images are representing 3 different classes.

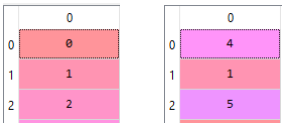

Fig. 8. Representing images into label and Shuffling of 10 images

Step 8: Then the images are shuffled for making the model more accurate, the network will learn in a better way and it wont memorize the data. For example, 10 images are shuffled in a form and a partial part is shown.

Step 9: After shuffling, the next part is to splitting the training and testing data and label. Here splitting of 10 images is given below. Split size=0.25

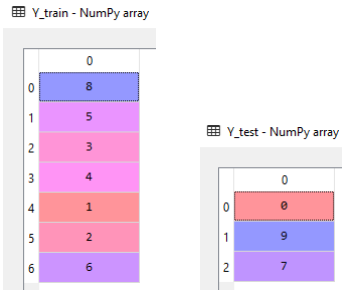

Fig. 9. Train and Test Label

Step 10: The final task is to feed the data to the proposed CNN model

Thats how preprocessing is done for all the images those were collected manually from different persons.

3) Proposed CNN model: After the preprocessing, all the images are fed to the model for training and testing purposes. The CNN model uses different layers such as Convolutional layer, Maxpooling layer, Dense layer etc. for recognizing different labels of images. CNN model learns knowledge from calculating weight to each layer and finally the maximum probability holder class is the resultant output. Thats how the model recognizes different Bengali sign language (Digit).

B. Main Processing

The main processing step works as the core decision making module.

1) Keras was used as a front-end and TensorFlow as a back-end in an anaconda (spyder) environment.
2) Python 3.7 installed with most recent anaconda.
3) Different Keras libraries were used for implementing CNN model and OpenCV, pillow for image preprocessing.
4) OpenCV was also used for accessing webcam at Real time processing.
5) For generating curve, matplotlib was used.
6) NumPy was used for generating NumPy array of images.
7) OS was used to give access to the system.
8) Scikit (scikit-learn) was used for shuffling data and label and also used for splitting data into training and testing set.

CNN is a Feed forward neural network and it uses different layer for identifying different classes accurately. It is good to use because it gives more accuracy using less connected layer and it is also efficient for training different images.

When the training started it gave more errors or loss at beginning, as the number of epochs increased, the loss of training data decreased gradually and the system can learn in a better way compared to before. The accuracy level became stable after performing a number of epochs and then the model gave an approximate accuracy to identify different signs.

1) The Proposed CNN model: Different layers were used for this proposed model. CNN uses convolution layer, polling layer and fully connected layer to identify different images. Keras sequential model was used for performing different operations.

Convolution Layer: The first layer of the proposed CNN model is a convolution layer. Convolution 2D was used for working with images. A convolution layer uses different filters or masks for convolving an image.

Activation Relu: An Activation function is a nonlinear transformation and its positioned between layers and determine the neuron would fire or not. Relu activation function was used to the proposed CNN model. Relu means Rectified Linear Unit(Relu).

Max-Pooling Layer: Then Max-pooling layer was used. There are different pooling layers but max-pooling was used because it reduces the image size by taking maximum pixel values. It uses different size of window to perform this task but mostly 2*2 is used for pooling.

Flatten Layer: The next step is to flattening the output of pooling layer that means the 2D images now converts into 1D images of pixel for training purposes. It is a linear operation. Flatten layer is a must because it is an input layer for CNN.

Dense Layer or Fully connected layer: Dense layer is a fully connected layer that means the output of flatten layer is now fully connected with some nodes to determine or produce an output class that consist of the maximum weight. This proposed model used 2 dense layers, one is the hidden

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

layer and others produces the output using Softmax activation function. Softmax is used for identifying classes of different level (number of class = 1 to infinity).

Dropout Layer: A drop out layer is used to drop some connections between nodes for avoiding over-fitting. A dropout layer reduces redundant or unnecessary connection at the time of training and that does not affect the accuracy but reduce the computation time.

2) Calculating Parameter: Convolution 1: The 1st layer in CNN is a convolution layer. We used kernel size = 9, stride = 1 and 64 filters in 1st layer. The input filter = 1. So, the number of parameters in this layer

$$(3x3)x1+1x64=640$$

In the 1st layer it has 64 filters as output goes to the next Convolution layer. One thing must be kept in mind that Relu layer cant change the parameter number because it only changes pixel values.

Convolution 2: In this layer the input filter = 64 So, number of parameters in this layer

$$(3x3)x64+1x64=36928$$

Convolution 3: In this layer the input filter is also 64 and kernel size, number of filters are same so the number of parameters will be same as Convolution 2. Number of parameters 36928 Dense 1: In Dense layer 1, got an input of 12544 nodes and output is 128 nodes So, number of parameters

$$(12544+1)128=1605760$$

Dense 2: In this final layer, got 128 nodes as input and 9 nodes as output. So, number of parameters

$$(128+1)9=1161$$

So, The total number of trainable parameters

$$640+36928+36928+1605760+1161=1681417$$

This is all about the summary of this model. CNN works actually how neurons works. It is step by step procedure to identify different labels. An overview of this full CNN model is illustrated in Fig 10.

Testing with existing images in the system the model works very well and gives an accuracy that is shown in Fig 11

From fig 11 it can be seen that all the prediction is accurate for this five images and test loss is very less. Test accuracy is outstanding and it is 99.74%.

Testing with new images that is not loaded at the time of training the model can predict accurately and it is shown in fig 12 and 13.

When The model was checked for digit 0 it can accurately identify the image represented as class 0 and in class 0 the images were checked for digit 0 as well as it was checked in the same way for digit 1 to 9 except 5.

Thats how a new image can actually be predicted through this CNN model. The proposed CNN model can identify new images properly and the error probability is very less as well as it gives different result for some scenario.

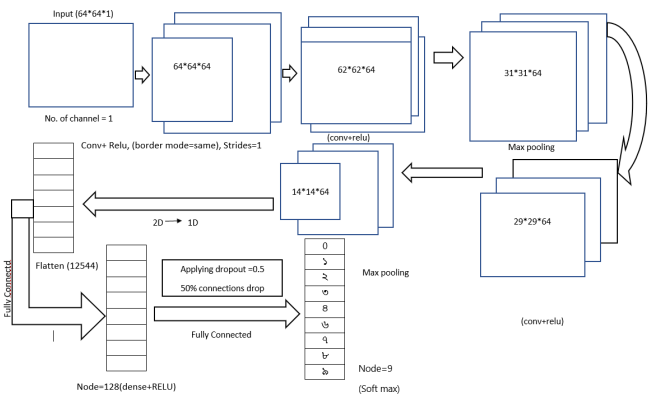

Fig. 10. An overview of this proposed CNN model for recognizing Bengali sign for digits

```
Test Loss: 0.014039364638987917
Test accuracy: 0.9974747474747475
(5, 64, 64, 1)
[[4.9642267e-21 0.0000000e+00 1.0566927e-25 5.3891174e-26 1.6102198e-32
  4.2197250e-22 1.0000000e+00 1.7275867e-09 1.1088874e-19]
 [2.7654533e-37 0.0000000e+00 0.0000000e+00 7.7035334e-23 1.0000000e+00
  0.0000000e+00 0.0000000e+00 0.0000000e+00 0.0000000e+00]
 [1.9100133e-37 2.5374187e-25 0.0000000e+00 0.0000000e+00 3.2585176e-34
  4.6179666e-12 1.0000000e+00 4.5581586e-29 3.6885034e-24]
 [0.0000000e+00 0.0000000e+00 0.0000000e+00 0.0000000e+00 0.0000000e+00
  1.0000000e+00 5.3696186e-35 0.0000000e+00 7.1684114e-31]
 [3.3558132e-21 0.0000000e+00 3.4851134e-38 2.3683863e-31 2.9283331e-28
  8.0487749e-26 5.8136102e-20 1.0000000e+00 2.1451180e-31]]
[6 4 6 5 7]
```

Fig. 11. Testing with images that already loaded in the system

3) Comparing different scenarios: The proposed system produced some output at the implementation time. Some of these are shown below:

Using no rotation: When The system used only the binary images to train the proposed CNN model, the model was found to perform less accurately which can predict well but not with all the cases. This approach is given in the following figure: 14

From fig 14 it can be seen that some pixels value doesnt match when checking how the model works. The model found some error. For minimizing the loss error, the proposed CNN model used dropout layer after max-pooling layer. This procedure is shown in the following figure: 15

In this case error was reduced a little bit but still needs to improve the accuracy level so, the proposed system used rotated image to train the model and got an outstanding result.

Using rotation: This time the CNN model used rotation of the images at the preprocessing time and the loss error was

```
(64, 64)
(1, 64, 64, 1)
[[9.9996901e-01 3.8864724e-14 1.2961416e-14 5.5885720e-11 2.4422188e-05
  7.7225844e-11 8.2322773e-14 2.4945468e-09 6.5502086e-06]]
[0]
```

Fig. 12. Testing with New Bengali digit sign = '0'

```
(64, 64)
(1, 64, 64, 1)
[[4.5075438e-14 9.9999166e-01 1.3104261e-08 5.4665373e-16 1.4435141e-14
  1.0020306e-06 7.3387314e-06 1.7401149e-10 4.6043974e-10]]
[1]
```

Fig. 13. Testing with New Bengali digit sign = '1'

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

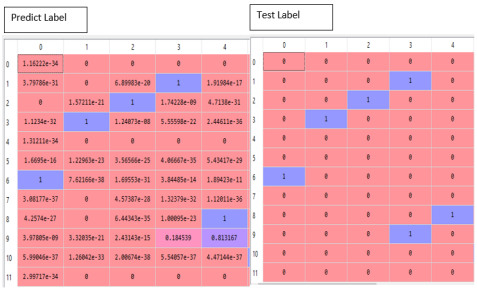

Fig. 14. Checking Model (using no rotation)

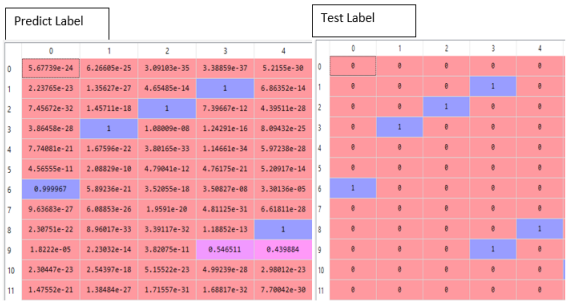

Fig. 15. Checking Model (using dropout after max-pooling)

reduced at a high level as well as the accuracy was better than previous cases. This procedure is shown in Fig 16.

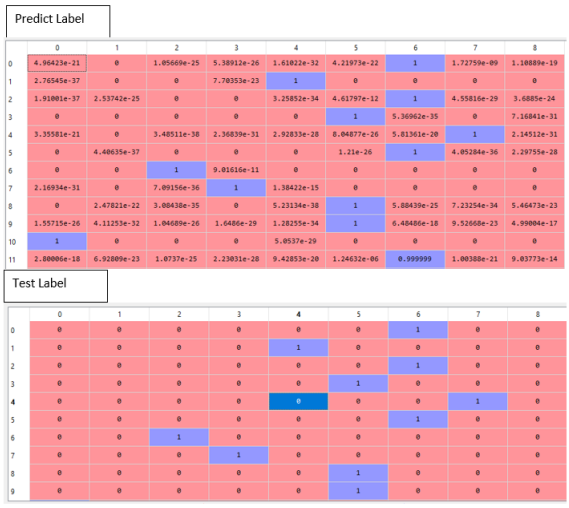

Fig. 16. Checking Model (using rotation on images)

IV. Experimental Result Discussion

Sign language can be recognized by different machine learning algorithms. Comparing to other sign languages, Bengali sign language recognition system is rare because a smaller number of experts are in practice in that field. So, collecting the data set was much more challenging. The data was captured manually from different individuals and all the images were pre-processed before training and testing by the proposed CNN model and finally the result was outstanding. The model was trained for different cases and different accuracy was found. The accuracy of the model is given bellow:

A. Training the model using images with no rotation

Bengali sign language images were used for training CNN model. At the time of pre-processing the model was trained with binary images not using rotation on them. 160 images were used for representing each class. So, total images used for training the model was 160*9 = 1440. A validation accuracy of 93.33% was found.

Finally a training accuracy of 99.81% and validation accuracy of 93.33% was found. A validation accuracy Vs training accuracy curve is illustrated in Fig 17.

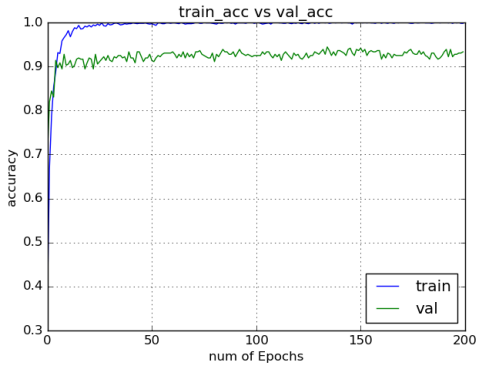

Fig. 17. Training accuracy Vs Validation accuracy curve (without rotation)

From fig 17. It is clear that validation accuracy is less compared to training accuracy. If less differences are found between training and validation accuracy, performance of system will increase more.

A validation loss Vs training loss curve is illustrated on Fig 18

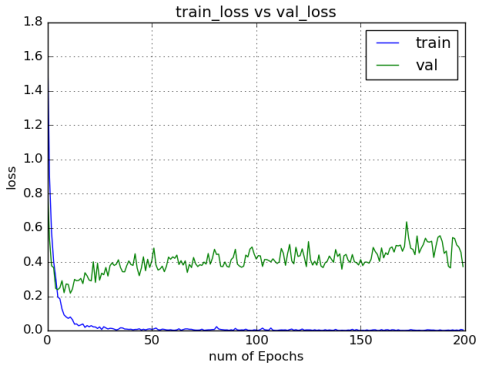

Fig. 18. Training loss Vs Validation loss curve (without rotation)

From fig 18, it is observed that the validation loss is high and it needs to be optimized. So, a solution was found to

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

increases the accuracy of the model. The new approach is that rotation on the input images can be used to increase the accuracy of CNN model.

When dropout layer was used after max-pooling layer, accuracy was slightly better than the previous accuracy seen in Fig 19. The accuracy curve and loss curve is also given in Fig 20 and Fig 21

```
Epoch 198/200
 - 50s - loss: 0.0185 - acc: 0.9926 - val_loss: 0.2748 - val_acc: 0.9472
Epoch 199/200
 - 51s - loss: 0.0111 - acc: 0.9954 - val_loss: 0.2731 - val_acc: 0.9528
Epoch 200/200
 - 51s - loss: 0.0081 - acc: 0.9963 - val_loss: 0.2661 - val_acc: 0.9417
```

Fig. 19. Accuracy gained, using dropout after max-pooling layer

Accuracy found in that case is 94.17% that is slightly better than the previous accuracy.

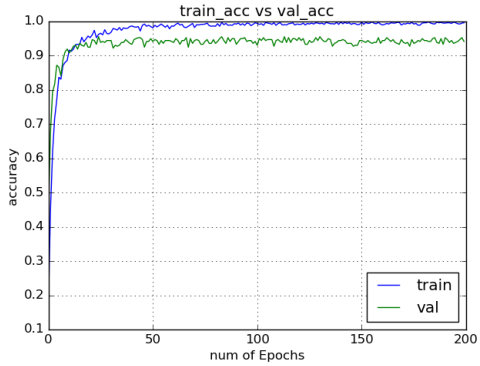

Fig. 20. Training accuracy Vs Validation accuracy curve (using dropout after max-pooling layer)

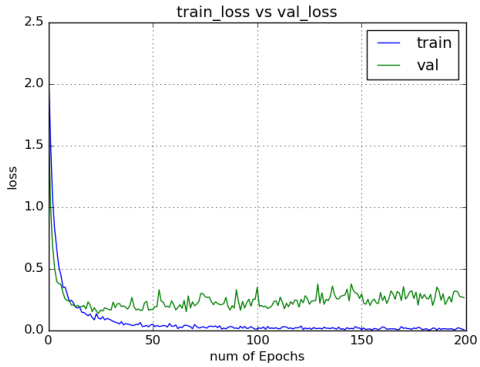

Fig. 21. Training loss Vs Validation loss curve (using dropout after max-pooling layer)

The accuracy is not the desired one so, a new feature was used to the data set at the time of preprocessing and it gives the best result compared to other analysis.

B. Training the model using images with rotation

Rotation is defined as a circular movement around a center or point. Rotation can be applied in different axes. The proposed system used 2D rotated images for training the model. During pre-processing time, 2D images were rotated in different directions like 0ř, 3ř, 6ř, 9ř, 12ř, and 15ř in both left and right directions. Python Image Library (PIL) was used for rotating images. It was decided to train the model with some images that contains rotation on them and an accuracy was found that is maximum comparing with all other cases. This gives the validation accuracy of 99.75% for 200 epoch

Difference between these two accuracies is very nominal. So, this time the CNN model can recognize images more accurately. The accuracy and loss curves are shown in fig 22 and 23 for this model.

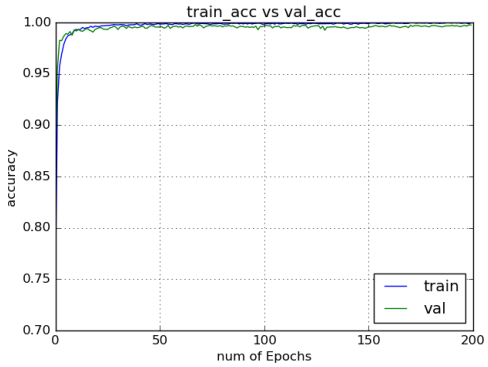

Fig. 22. Training accuracy Vs Validation accuracy curve (with rotation on images)

From fig 22 it is clear that difference between training and validation accuracy is very nominal. From fig : 23, it can be seen that training loss and validation loss is very small. So, the CNN model is learning very well and when test images are compared to training images the recognition level is very high.

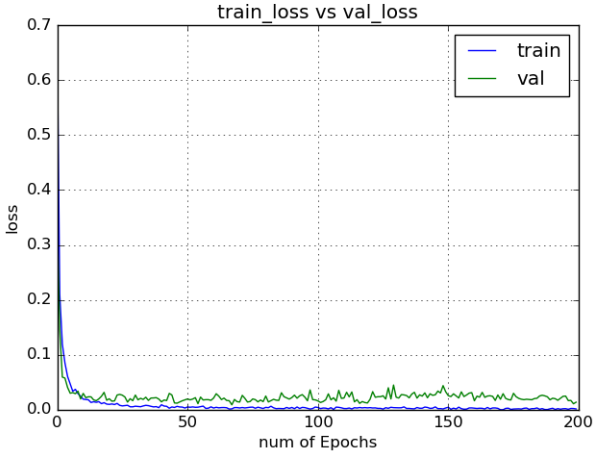

Fig. 23. Training loss Vs Validation loss curve (with rotation on images)

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

C. Real Time Recognition

When recognizing a sign language that has been presented to the system, the image frames are captured that construct the sign. The hand postures are segmented from these initial image frames. These frames containing the hand signs are preprocessed and then converts the hand sign to binary images [9].

Extracted binary hand signs are trained or recognized by comparing with pre-trained binary images of hand sign using the proposed CNN model. When extracted binary hand sign is matched against the pre-stored training hand signs then the system recognizes the specific sign using the proposed CNN classifier. This proposed model was checked for real time recognition of Bengali sign language (Digit) and it performs quite better. A small amount of data set is used for training the proposed model thats why some of the signs are not stable in real time processing. This proposed system can recognize all signs that was trained using CNN. Here, some examples are shown in Fig 24.

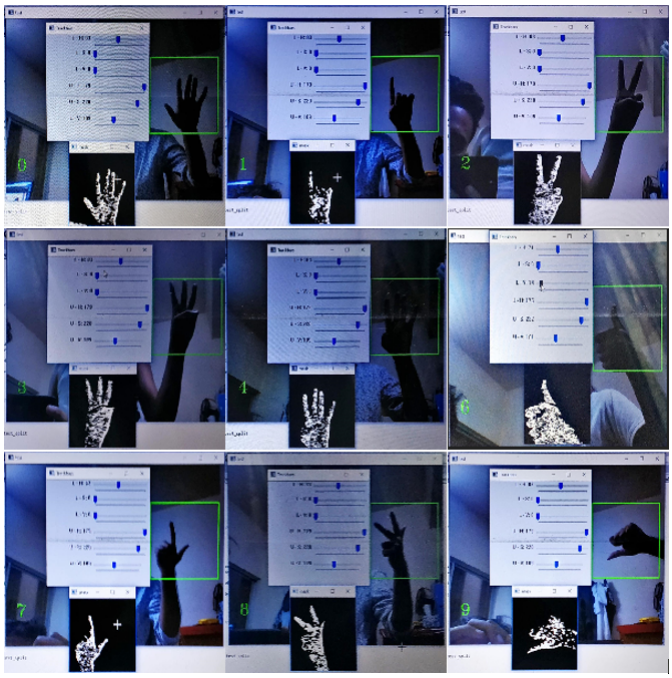

Fig. 24. Real time recognition of Bengali Sign Language (Digit) using CNN

V. Discussion

Bangladesh is an underdeveloped country beset with many major problems and it is not possible to focus on all of them due to a lack of resources and keen motivation. Speech and hearing-impaired people are now facing negligence and problems due to their inability to communicate with normal people. But in recent times many researchers have come forward to help them with their research to find a way to communicate with them. The proposed system has a similar goal to help hearing-impaired people. Many other approaches were taken for sign language recognition where many other models were different than CNN based model. But the proposed CNN model has worked comparatively effectively than those models for recognizing sign language gestures.

A. Conclusion

Sign language is an important way for deaf and dumb people to communicate daily. But it is not possible for people who dont know sign language. This research work has successfully collected Bengali sign language (Digit) data set in captured images from different individuals. After the data was collected it went through pre-processing steps to transform the images in such a format that the computer or machine understands them. Then the processed data is being fed to the proposed CNN model for training purposes. After training is done and got three different accuracy's but the best accuracy was found for the rotated image dataset. The model also works well with real-time recognition for Bengali digits from 0 to 9 except for 5 but some are not stable due to the smaller data set.

B. Evaluation of Research Questions

At the beginning of the research, there were three questions that needed to be answered and the best effort was given to do so. The research work was done in such a way that the output of the experiment would contain some key information regarding the questions.

The first question was about CNN being better at generating output for Bengali Sign Language (Digits) for given data sets better than any other models. It was keenly observed that there were difficulties and challenges while implementing the CNN model. But all of those challenges were met and the model successfully generated output for the given data sets.

The second question was about the model being able to recognize Bengali Sign Language (Digits). From the research that was conducted it was observed and proved that the model was able to recognize all of the Bengali Signs (Digits) after processing all the data in the model that was collected. Although some of them were slow but still they worked.

The third question focused on the effectiveness of the proposed model over the society. As the proposed model is finally implemented, all the data sets were processed and the model was able to recognize the signs so it is safe to say that this research work would help the impaired people and in living in a difficult society where the means of expressing themselves are limited.

C. Future Work

Bengali Sign Language has embarked on the journey of many dedicated researchers to follow the path in creating a new system of communication and introducing it to society. Finding a new way of communication for deaf and dumb people, extending the research to introduce facial expression to the system, building mobile applications, etc. are now intriguing topics many researchers involved in Sign Language. To enhance Bengali Sign Language even more with the proposed system the following work scopes should be attempted:

https://sites.google.com/site/ijcsis/
ISSN 1947-5500

1) **Including all signs as data to the proposed system:** Right now, the proposed system can only recognize digits as they are only given as inputs.

But training the model with all the letters and symbols of the Bengali Language assigns data inputs would broaden the possibility of this model being even more successful and goal achieving.

1) **Building Mobile Applications:** It is a well-known fact that mobile phones have become an integral part of human life. So, if it is possible to build a mobile platform-based application that will be easily accessible to everyone then communicating in a normal way will be much simpler for speech and hearing-impaired people.
2) **Introducing Facial Expression to the system:** If facial expression and body language from different people can be taken as data inputs to train the model to recognize them and what they mean then the system would be much more effective and efficient in the benefit of sign language recognition.

REFERENCES

[1] Starner, Thad, Joshua Weaver, and Alex Pentland. "Real-time american sign language recognition using desk and wearable computer based video." IEEE Transactions on pattern analysis and machine intelligence 20.12 (1998): 1371-1375.
[2] Wang, Chunli, Wen Gao, and Shiguang Shan. "An approach based on phonemes to large vocabulary Chinese sign language recognition." Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on. IEEE, 2002.
[3] Uddin, Md Azher, and Shayhan Ameen Chowdhury. "Hand sign language recognition for Bangla alphabet using Support Vector Machine." Innovations in Science, Engineering and Technology (ICISET), International Conference on. IEEE, 2016.
[4] Yasir, Farhad, et al. "Sift based approach on bangla sign language recognition." Computational Intelligence and Applications (IWCIA), 2015 IEEE 8th International Workshop on. IEEE, 2015.
[5] Sagawa, Hirohiko, and Masaru Takeuchi. "A method for recognizing a sequence of sign language words represented in a japanese sign language sentence." Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on. IEEE, 2000
[6] Jasim, Mahmood, Tao Zhang, and Md Hasanuzzaman. "A real-time computer vision-based static and dynamic hand gesture recognition system."International Journal of Image and Graphics 14.01n02 (2014): 1450006.
[7] Karmokar, Bikash Chandra, Kazi Md Rokibul Alam, and Md Kibria Siddiquee. "Bangladeshi sign language recognition employing neural network ensemble." International journal of computer applications 58.16 (2012).
[8] Yasir, Farhad, et al. "Bangla Sign Language recognition using convolutional neural network." Intelligent Computing, Instrumentation and Control Technologies (ICICICT), 2017 International Conference on. IEEE, 2017.
[9] Rahaman, Muhammad Aminur, et al. "Real-time computer vision-based Bengali Sign Language recognition." Computer and Information Technology (ICCIT), 2014 17th International Conference on. IEEE, 2014.
[10] Van Herreweghe, M.: Prelinguaal dove jongeren en nederlands: een syntactisch onderzoek. Universiteit Gent, Faculteit Letteren en Wijsbegeerte (1996)
[11] Poppe, R.: A survey on vision-based human action recognition. Image and Vision Computing 28(6), 976990 (2010)
[12] Rajam, P. Subha, and G. Balakrishnan. "Real time Indian sign language recognition system to aid deaf-dumb people."ă2011 IEEE 13th International Conference on Communication Technology. IEEE, 2011.

https://sites.google.com/site/ijcsis/
ISSN 1947-5500