TELKOMNIKA Telecommunication, Computing, Electronics and Control
Vol. 18, No. 3, June 2020, pp. 1292~1301
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v18i3.15311

Automated Bangla sign language translation system for alphabets by means of MobileNet

Tazkia Mim Angona, A. S. M. Siamuzzaman Shaon, Kazi Tahmid Rashad Niloy, Tajbia Karim, Zarin Tasnim, S. M. Salim Reza, Tasmima Noushiba Mahbub
Department of Information and Communication Technology, Bangladesh University of Professionals, Bangladesh

Article Info

Article history:

Received Jan 16, 2020
Revised Feb 7, 2020
Accepted Feb 28, 2020

Keywords:

Accuracy
Bangla sign language (BSL)
CNN
Convolution
MobileNet

ABSTRACT

Individuals with hearing and speaking impairment communicate using sign language. The movement of hand, body and expressions of face are the means by which the people, who are unable to hear and speak, can communicate. Bangla sign alphabets are formed with one or two hand movements. There are some features which differentiates the signs. To detect and recognize the signs, analyzing its shape and comparing its features is necessary. This paper aims to propose a model and build a computer systemthat can recognize Bangla Sign Lanugage alphabets and translate them to corresponding Bangla letters by means of deep convolutional neural network (CNN). CNN has been introduced in this model in form of a pre-trained model called "MobileNet" which produced an average accuracy of 95.71% in recognizing 36 Bangla Sign Language alphabets.

Corresponding Author:

S. M. Salim Reza,
Department of Information and Communication Technology (ICT),
Bangladesh University of Professionals (BUP),
Mirpur Cantonment, Dhaka-1216, Bangladesh.
Email: salim4419@gmail.com

1. INTRODUCTION

Communication is one of human society's cornerstones. Human communicates by the means of spoken language primarily. The persons who are unable to speak and hear, however, cannot communicate this way. Therefore, many cultures have developed their own language for communicating with the speaking and hearing impairedpeople. All the letters and words are represented visually to the impaired person. This language is called sign language. Sign language mostly consists of specific gestures made with one or two hands where the position of the palm and placement of fingers combined make up letters of the language. Each alphabet has a predetermined pattern of hand and finger position. The variant of sing language for Bangla language is referred to as the Bangla Sign Language (BSL). Figure 1 shows the hand sign images for BSL vowels and Figure 2 shows Bangla numerical signs.

Learning the sign language can be quite a daunting task to an able person because speaking as a means of communication becomes their nature. As a result, most people of this country do not know sign language of Bangla language. As such with hearing and speech impaired people hardly becomes a full functioning member of the society. Therefore, an automated process that can translate the sign language into natural spoken language is necessary. As every language consists of alphabets, recognition and translation of alphabets is the first step towards a complete real-time translation system.

Creating a computer system that can recognize BSL alphabet sign from an image and translate it to Bangla letter for Bangla speaking people based on CNN [1], is the objective of this paper. We present a simple and compact scheme for recognition of BSL in this paper. MobileNet V1 has been used based on transfer learning. MobileNet is a class of lightweight deep convolutional neural network which are vastly smaller in size and faster in performance. It is fine-tuned to adapt to our problem. Ishara-Lipi dataset [2] was used to train and evaluate our model. This approach also requires less storage than other existing solutions. We have built a model for desktop computer, but due to low memory usage, it can also be extended to mobile devices as well.

Contribution of this work can be written as follows:
- BSL alphabets recognition model has been proposed and implemented.
- A reliable desktop application has been built which can translate 36 BSL alphabets.

Figure 1. BSL vowels hand sign [3]

Figure 2. BSL numericals [4]

2. RELATED WORKS

Various machine learning techniques with image processing have already been used by various researchers for BSL recognition. Some of these methods are discussed briefly here. In paper [5], the authors proposed a method using ANN for translating the BSL alphabet to corresponding Bangla letter. For image processing operations, Matlab tools were applied. A total of 828 images were collected for training and testing. The system can recognize 36 BSL alphabets. The accuracy of the system is 80.902% [5].

In paper [6], the authors developed a model for BSL recognition by image processing. Initially skin color is detected using YCbCr skin detection algorithm. Extracted features from images were fed into SVM classifier for training and testing. They collected images of 15 Bangla consonant hand signs. Each class of signs has 52 or 56 images of 29 different bare hands. The 95% average accuracy has been received for training set and 86% has been received for evaluation set. For the letters 'ঙ' and 'ট' it got very low accuracy [6].

CNN approach was taken by Rony et al. for single-hand BSL character Recognition. This system of BSL interpreter has been designed with the purpose of providing real time communication. This system is trained by using Tensorflow Inception-v3 Model and tested by applying CNN. It uses a camera to detect sign in real time. They achieved an average recognition accuracy of 92.85% [7].

Another method has been proposed by Shanta et al. to detect BSL using SIFT and CNN. This method extracts feature descriptors of ROI (Region of Interest) using SIFT, then clusters the features using K-means clustering to get clustered descriptors, uses Bag of interest (BOI) to represent clustered features as histograms and then input the histograms in CNN. This method was done both with and without SIFT where using SIFT yields higher accuracy on every sign. Using SIFT this method can detect 38 hand sign for 51 letters with 88.52% accuracy on average [8].

BSL Recognition Using Deep CNN has been introduced by Hossen et al. This system is made to detect static hand sign from images taken under various lighting conditions, different background and positions of the hand. This system can detect 37 BSL alphabets with validation accuracy of 84.68% [9]. Islam et al. proposed a model to identify BSL digits using CNN. This model uses multi-layer CNN with 2 sub layers and ADAM optimizer. A flatten and dense layer was used after max-pooling layer and on the final output layer softmax activation has been used. Ishara-Lipi dataset [8] was used and the model has shown 94.88% accuracy for digit 0 to 9 [10].

Recognition of BSL alphabets by CNN was proposed by Islam et al. They applied their model on Ishara-Lipi dataset [2] which contains 36 BSL characters. The model had multi layered CNN and used Adam optimizer with 0.001 learning late. On this dataset, the model achieved 92% accuracy [11]. A BSL recognition system offered by Hasan et al. which applied ANN for training and classification and Freeman

Chain Code was used for feature extraction. YCbCr was used as skin detection algorithm. The system could recognize 20 alphabets and 96.5% average accuracy rate was achieved. But this system can detect only one handed signs [12].

Another system developed by Yasir et al. to recognise two hand sign. ANN was used to train the system and YCbCr algorithm for skin detection. Two different techniques namely PCA and LDA were used and compared as feature extractor where PCA stands for principal component analysis and LDA stands for Linear Discriminant Analysis. Images of 15 Bangla alphabets were collected from 22 people. ANN classifier with LDA produced an excellent result. The result is much better than ANN with PCA or raw image [13].

A new approach was proposed by Hasan et al. Hybrid color model (RGB-YCbCr-HSV) was used for skin detection. A combination of PCA and FLDA (fisher's linear discriminant analysis) was introduced to extract features. They adopted a new classifier LSVM. The system can identify 10 Bangla numbers [14]. Yasir et al. introduced a SIFT (System Invariant Feature Transform) based approach for two handed sign alphabets recognition. After normalization of input images geometrical based SIFT method is used for extracting features. SVM was applied as classifier. Comparing between classifiers K-NN (experimental results of Rahaman et al. [5]) and SVM recognition rates, k-NN classifies better as the dataset was small [15].

Karmokar et al. introduced a BSL recognition system which recognizes 47 alphabets of Bangla Sign Language by NNE (neural network ensemble). A method combining feature extraction and NNE was proposed. A total 235 samples (5 for each alphabet) were used for training of the system. NCL algorithm was employed for faster training. The experimental result shows NCL with feature extraction gives an approximate accuracy of 93%. Varying lighting, presence of skin colored objects, quality of the image causes it to face problem in recognizing the alphabets [16].

Aziz et al. approached a different method for recognition of BSL. The process of recognition is based on dynamic calibration of skin as well as geometric hashing. Input image is converted into binary image. This process of conversion is done by dynamic color-based skin threshold and mean shift segmentation algorithm. Hash table was trained with 1147 samples of 37 alphabets. The hand sign is identified by hashing [17]. For working with images of different backgrounds, a model was proposed by Haque et al. which identifies two handed signs using PCA algorithm and ANN. The mechanism was to pre-process images and use PCA for feature extraction. K-Nearest Neighbor was chosen as classifier. It can recognize 26 Bangla alphabets with an average accuracy rate of 77.846%. The mechanism can give satisfactory results for input images with different backgrounds [18].

In paper [19], the authors developed a BSL recognition system.The system is basially a real time computer vision based system where hand area is detected from the input image using the classifier namely Haar-like feature-based cascaded classifier [20].Using K-Nearest Neighbor classifier, hand sign is extracted by the proposed system from Hue and Saturation values of HSV color model where HSV stands for Hue, Saturation, and Value. HSV color model is frequently used instead of the RGB color model in application programs such as graphics and paint programs. The system can recognize 30 consonants with an accuracy of 94.75% and 6 vowels with an accuracy of 98.17%. The system can give output in real time. The system has some problems recognizing the similar hand signs. "র" with 'ল' and 'ফ' with 'ঝ' can be the examples of similar hand signs [19]. Hoque et al proposed a detection system namely Real Time Bangla Sign Language Detection. Faster R-CNN has been used in this proposed system for detecting hand sign and different background of ROI where R-CNN stands for Region-Convolutional Neural Network. This model is designed to detect BSL from images with different background and lighting in real time. It can detect 10 letters [21].

3. BACKROUND

An overview of fine-tunning and MobileNet V1 is given in this section.

3.1. Fine-tunning

Fine-tunning is linked with transfer learning. When we use knowledge that was gained from one task and apply it to a new task, that's transfer learning. Fine-tunnig is a process where it takes an already trained model and then tunes or tweaks that model to perform a second task. It allows to take the advantage of the gained knowledge without having to develop it from scratch as building a model from scratch is quite difficult. By replacing the top layer with a classifer, a pre-trained model can be easily fine-tuned. The pre-trained weights can be used to speed up convergence rather than using random weights [1].

3.2. MobileNet V1 architecture

MobileNets are a class oflow latency, low power model that can be used in classification, detection and other tasks CNN are good for. As they are small in size, these are considered to be great deep learning models to be used in computer and mobile devices. It is built on filters that can be separated in depth. This gives the models an opportunity to decrease the number of parameters needed for convolutional operations, thus reducing the model's size. They are pre-trained on ImageNet [22] dataset [23].

The architecture followed a pattern which was easy to replicate.

- The filter size of each convolutional layer that isn't Pointwise convolution-3x3
- Stride 2 is used for downsampling in intermediate convolution.
- The output is flattened by a Global Pooling layer [23].

A substitute of the standard convolutional filters in MobileNetis depth-wise separable filter which consists of two layers. Those are depth-wise and point-wise convolution. Figure 3 illustrates the comparison between standard and depth-wise (dw) separable filter. MobileNet has 13 convolution blocks. Figure 4 describes the MobileNet architecture. A batch normalization and Relu layer is after each convolution layer. Lastly there is a global average pooling layer and a fully connected layer which feeds into Softmax function for classification [23].

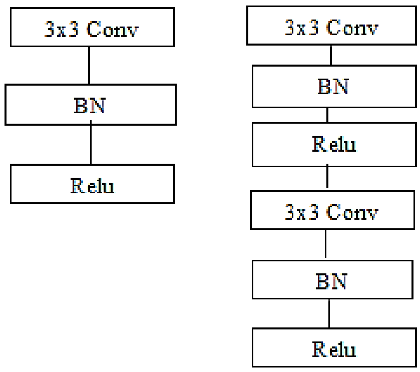

Figure 3. Standard and Depth-wise (dw) separable convolution

Type / Stride	Filter Shape	Input Size
Conv / s2	3 × 3 × 3 × 32	224 × 224 × 3
Conv dw / s1	3 × 3 × 32 dw	112 × 112 × 32
Conv / s1	1 × 1 × 32 × 64	112 × 112 × 32
Conv dw / s2	3 × 3 × 64 dw	112 × 112 × 64
Conv / s1	1 × 1 × 64 × 128	56 × 56 × 64
Conv dw / s1	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 128	56 × 56 × 128
Conv dw / s2	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 256	28 × 28 × 128
Conv dw / s1	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 256	28 × 28 × 256
Conv dw / s2	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 512	14 × 14 × 256
5× Conv dw / s1	3 × 3 × 512 dw	14 × 14 × 512
Conv / s1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw / s2	3 × 3 × 512 dw	14 × 14 × 512
Conv / s1	1 × 1 × 512 × 1024	7 × 7 × 512
Conv dw / s2	3 × 3 × 1024 dw	7 × 7 × 1024
Conv / s1	1 × 1 × 1024 × 1024	7 × 7 × 1024
Avg Pool / s1	Pool 7 × 7	7 × 7 × 1024
FC / s1	1024 × 1000	1 × 1 × 1024
Softmax / s1	Classifier	1 × 1 × 1000

Figure 4. MobileNet architecture [23]

4. PROPOSED SYSTEM DESCRIPTION

The research work is described in the following sub-sections.

4.1. Dataset preparation

We used BSL's first complete isolated charactersdataset, Ishara-Lipi dataset [2]. The dataset contains 30 consonants and 6 vowels of BSL characters. Table 1 shows the properties of the dataset. The dataset holds 36x50=1800 images in total. All images are cropped and resized by 128x128 pixels. The images were taken with same background and in same lighting conditions. The images are in .JPEG format and they are converted into binary image before feeding it into the model. Figure 5 shows some samples of the dataset. The 36 BSL alphabets (অ, আ, ই, উ, এ, ও, ক, খ, গ, ঘ, চ, ছ, জ, ঝ, ট, ঠ, ড, ঢ, ত, থ, দ, ধ, ণ, প, ফ, ব, ভ, ম, য়, র, ল, স, হ, ড়,ং, ঃ) are assigned as the class labels of (0,1,2……, 36) in the dataset. 10% data were separated from the dataset for final testing of our system.

Table 1. Dataset properties

Data	Number of class	Sets per class	Total number of data
Vowels	6	50	300
Consonants	30	50	1500

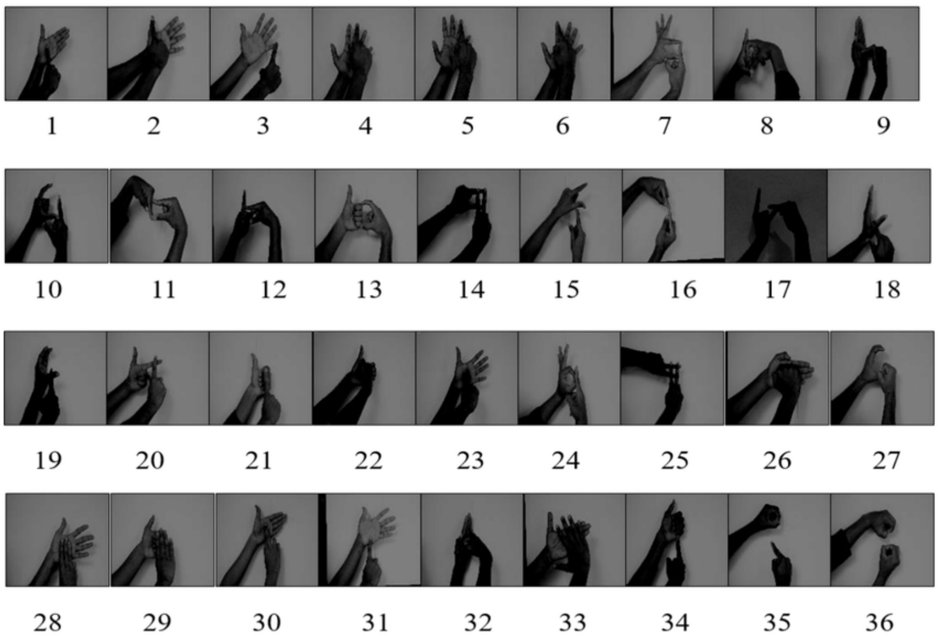

Figure 5. BSL Characters Dataet Samples [10]

Using stratified K fold cross validation, the entire dataset was divided into train and test set. For the purpose of training, the value of K was set to 5. The data was split into 5 subsets and the model was trained on 5-1=4 subsets. So the last subset was hold for testing and it was done for each of the subsets. In this case, all the samples had a chance to be in the test set in each iteration, thus building a more reliable system.

4.2. Methodology

The proposed method uses Google's pretrained model MobileNet version 1 which has a good accuracy on ImageNet dataset [23]. The transfer learning [24] approach was taken to fine tune the model to our specific classification problem. We trained MobileNet for 50 epochs with prior weights of ImageNet [22] dataset on our training set. All pre-trained weights can be used to initialize entire network weights in order to reduce training time [25]. The images were preprocessed for the model. The last 6 layers were not trained from the model which had the fully connected or dense layer. We introduced a dense layer with 36 class labels which is the number of our classes and 'softmax' as activation function instead of these 6 layers. The proposed model used ADAM optimizer where learning rate was 0.001 and cost function was categorical cross entropy.

The convolutional layers extract the features of the images. From the last convolution layer “conv_pw_13_relu” 50176 features were extracted for each image. During the training process, the extracted features from trainset are used to train the dense layer. Dense layer acts as a classifier which classifies the images into the corresponding class labels. After training, the model was evaluated by test set. The model predicted the class labels of the images belonging to test set. Figure 6 illustrates the block diagram of our proposed approach.

The proposed model is trained using stratified k fold cross validation scheme. The dataset of 1800 images is shuffled and split uniformly into 5 folds. For testing, one-fold is set aside and the other four folders are used for training. This is repeated five times in order to use each fold as test data. The measures for evaluation were Precision, Recall, F1-score and support. Precision is the ratio of positive observations that were predicted accurately to total predicted observations that were positive. F1-score is the average of precision and recall. Recall is the ratio of positive events that were accurately predicted to all the observations in the class. Support shows the number of samples in the classes.

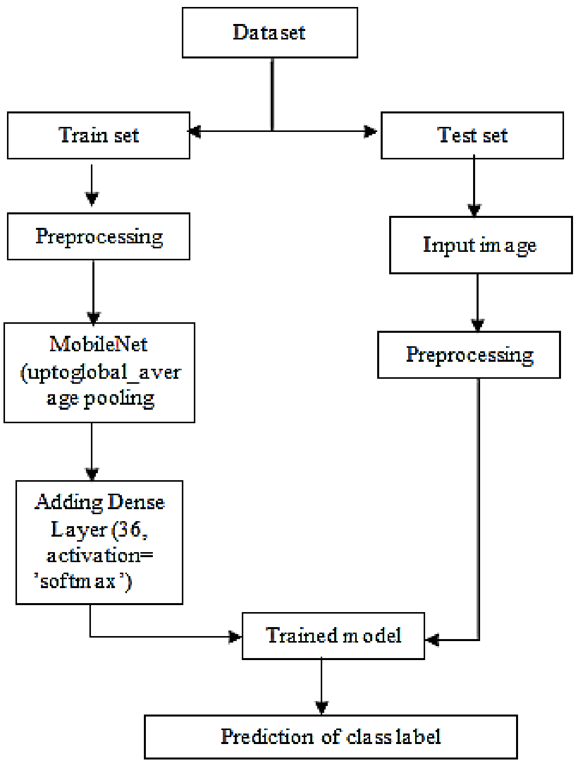

Figure 6. Our proposed model

5. RESULT AND DISCUSSION

The testing set was 20% of the whole dataset. The model has been trained for 50 epochs. The average accuracy for 36 alphabets of the proposed model is 95.71%. The evaluation of the model for each iteration is given in Table 2. A more detailed report is shown in Table 3. In each iteration, the report shows the following measures: precision, recall and F1-score. Average precision and recall the model returned was 96%. The classes with highest precision were ‘ক’ and ‘দ’. The classes with the lowest precision were ‘স’ and ‘ড’. As some images were blur and complex in shape the model couldn’t recognize them correctly. The model got confused between some hand signs such as ‘স’ and ‘গ’ as they are similar in shape, so these classes had a low accurate prediction rate.

Table 4 presents a comparison among existing models and our proposed model. Our proposed architecture gave better performance than most of these existing models. Using conventional CNN, Islam et al. used Ishara-Lipi dataset [1] which is the same dataset that we used. They built their own CNN model which gave an average accuracy of 92.00%. Our proposed model MobileNet gave a better accuracy rate than that.

Table 2. Model evaluation result

Iterations (For 5-fold)	Accuracy (%)
1^{st}	95.71
2^{nd}	96.39
3^{rd}	92.00
4^{th}	96.94
5th	97.54
Average	95.71

Table 3. Average result for each iteration

Iterations (For 5 fold)	Precision	Recall	F1-score	Support
1^{st}	0.96	0.96	0.96	840
2^{nd}	0.97	0.96	0.96	831
3^{rd}	0.93	0.92	0.92	825
4^{th}	0.97	0.97	0.97	817
5^{th}	0.98	0.98	0.98	812
Average	0.96	0.96	0.96	4125

Table 4. Comparison among existing and our proposed model

Models	Methodology	Dataset	Class	Average accuracy
Rahman et al. [5]	Muli-Layer Neural Network with image Processing	23x36=828 samples	36 alphabets	80.90%
Uddin et al. [6]	YCbCr algorithm and SVM Classifier	800 samples	15 consonants	86%
Rony et al [7]	CNN	1900 samples	38 alphabets, 3 special character	92.85%
Islam et al. [11]	CNN	1800 samples (Ishara-Lipi dataset [1])	36 alphabets	92%
Proposed model	MobileNet	1800 samples (Ishara-Lipi dataset [1])	36 alphabets	95.71%

6. DEVLOPING THE DESKTOP APPLICATION

The proposed architecture MobileNet was integrated in a computer vision based application. A user interface was created with the help of Python Tkinter. The homescreen and transition from homescreen of the application is shown in Figure 7. The application takes a hand sign image as input and the image is fed into the trained model. It classifies the image and returns the corresponding Bangla letter in real time. The Bangla letter is displayed in the screen in text format. It also shows the top 5 probabilsitc class of the input image. Here, the class having the highest probability is selected as the class of that particular hand sign.

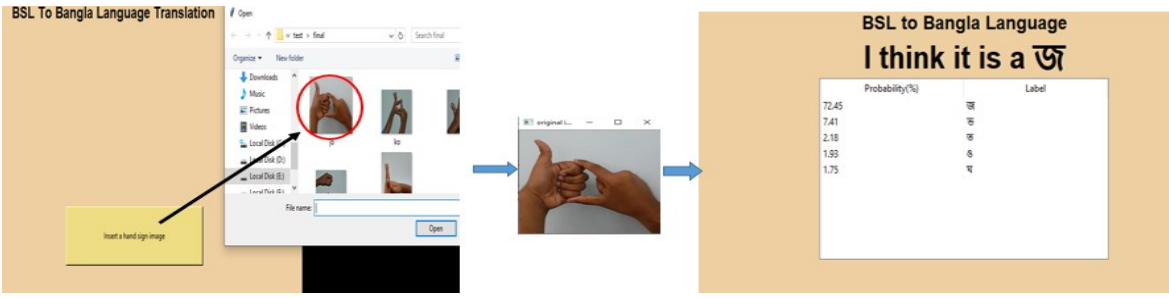

Figure 7. User interface of BSL desktop translation system for alphabets

7. CONCLUSION AND FUTURE WORK

Without the creation of sign language hearing and speech impaired people would forever be unable to communicate with other people and even each other. Whilst sign language can enable them to express themselves, lack of knowledge of sign language among the general people creates a gap of communication. The paper presents a quick and compact mehod for translation of BSL alphabets into Bangla Language.

The model could recognize 36 alphabets with an average accuracy of 95.71%. It also takes less computational time, resulting in less power consumption which is necessary nowadays. It was used to build an automated translation system for computer. This research is the first step towards making a system that can be used in day-to-day situations to communication as a system needs to recognize the alphabets first before it can be trained to start recognizing words.

By using images of higher resolution, the model could be enhanced. This model cannot detect BSL characters in different backgrounds and lighting conditions. This approach can be improved using a more versatile dataset. In future, it can be used to recognize all the BSL characters (51 alphabets). This approach is the basic step in building a more accurate BSL recognition system in real time. We have built a desktop translation system and due to low storage it can be extended to mobile devices.

REFERENCES

[1] I. Goodfellow, Y. Bengio and A. Courville, "Convolutional Networks," *Deep Learning, MIT Press*, 2016.

[2] M. Sanzidul Islam, S. Sultana SharminMousumi, N. A. Jessan, A. Shahariar Azad Rabby, and S. Akhter Hossain, "Ishara-Lipi: The First Complete MultipurposeOpen Access Dataset of Isolated Characters for Bangla Sign Language," *International Conference on Bangla Speech and Language Processing (ICBSLP)*, pp. 1-4, 2018.

[3] T. Hoque and F. Kabir, "Automated Bangla Sign Language Translation System : Prospects, Limitations and Applications," Conference: *International Conference on Informatics, Electronics and Vision (ICIEV)*, pp. 856-862, 2016.

[4] A. M. Jarman, S. Arshad, N. Alam, and M. J. Islam, "An Automated Bengali Sign Language Recognition System Based on Fingertip Finder Algorithm ISSN : 2186-0114," *Int. J. Electron. Informatics*, vol. 4, no. 1, 2015.

[5] M. A. Rahman, "Recognition of Static Hand Gestures of Alphabet in Bangla Sign Language," *IOSR J. Comput. Eng*., vol. 8, no. 1, pp. 07-13, 2012.

[6] J. Uddin, F. N. Arko, N. Tabassum, T. R. Trisha, and F. Ahmed, "Bangla sign language interpretation using bag of features and Support Vector Machine," *3rd Int. Conf. Electr. Inf. Commun. Technol. EICT,* pp. 1-4, 2017.

[7] A. J. Rony, K. H. Saikat, M. Tanzeem, and F. M. Rahat Hasan Robi, "An effective approach to communicate with the deaf and mute people by recognizing characters of one-hand bangla sign language using convolutional neural-network," *4th Int. Conf. Electr. Eng. Inf. Commun. Technol. iCEEiCT*, pp. 74-79, 2018.

[8] Sultana Shanta, Shirin &Taifur Anwar, Saif&Rayhanul Kabir, Md. (2018), "Bangla Sign Language Detection Using SIFT and CNN," *Conference: International Conference on Computing, Communication and Networking Technologies (ICCCNT),* pp. 1-6, 2018.

[9] M. A. Hossen, A. Govindaiah, S. Sultana, and A. Bhuiyan, "Bengali sign language recognition using deep convolutional neural network*," Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR),* pp. 369-373, 2018.

[10] M. S. Islam, S. S. S. Mousumi, A. K. M. S. Azad Rabby, S. A. Hossain, and S. Abujar, "A potent model to recognize bangla sign language digits using convolutional neural network," *Procedia Comput. Sci*., vol. 143, pp. 611-618, 2018.

[11] M. S. Islam, S. S. S. Mousumi, A. S. A. Rabby, and S. A. Hossain, "A Simple and Mighty Arrowhead Detection Technique of Bangla Sign Language Characters with CNN," *International Conference on Recent Trends in Image Processing and Pattern Recognition* vol. 1035, pp. 429-437, 2019.

[12] M. M. Hasan, M. Khaliluzzaman, S. A. Himel, and R. T. Chowdhury, "Hand sign language recognition for Bangla alphabet based on Freeman Chain Code and ANN," *4th International Conference on Advances in Electrical Engineering (ICAEE)*, pp. 749-753, 2017.

[13] R. Yasir and R. A. Khan, "Two-handed hand gesture recognition for Bangla sign language using LDA and ANN," *The 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, pp. 7-11, 2014.

[14] S. M. K. Hasan and M. Ahmad, "A new approach of sign language recognition system for bilingual users," *International Conference on Electrical & Electronic Engineering (ICEEE*), pp. 33-36, 2015.

[15] F. Yasir, P. W. C. Prasad, A. Alsadoon, and A. Elchouemi, "SIFT based approach on Bangla sign language recognition IEEE 8th International Workshop on Computational Intelligence and Applications (IWCIA), pp. 35-39, 2016.

[16] B. ChandraKarmokar, K. Md. RokibulAlam, and M. Kibria Siddiquee, "Bangladeshi Sign Language Recognition Employing Neural Network Ensemble," *Int. J. Comput. Appl.,* vol. 58, no. 16, pp. 43-46, 2012.

[17] K. E. Aziz, A. Wadud, S. Sultana, M. A. Hussain, and A. Bhuiyan, "Bengali Sign Language Recognition using dynamic skin calibration and geometric hashing," *6th Int. Conf. Informatics, Electron*. Vis. *7th Int. Symp. Comput. Med. Heal. Technol. ICIEV-ISCMHT*, pp. 1-5, 2017.

[18] P. Haque, B. Das, and N. N. Kaspy, "Two-Handed Bangla Sign Language Recognition Using Principal Component Analysis (PCA) and KNN Algorithm," *2nd Int. Conf. Electr. Comput. Commun. Eng. ECCE*, pp. 1-4, 2019.

[19] M. A. Rahaman, M. Jasim, M. H. Ali, and M. Hasanuzzaman, "Real-time computer vision-based Bengali sign language recognition," *17th Int. Conf. Comput. Inf. Technol. ICCIT*, pp. 192-197, 2003.

[20] Jasim, Mahmood & Zhang, Tao &Hasanuzzaman, Md., "A Real-Time Computer Vision-Based Static and Dynamic Hand Gesture Recognition System," *International Journal of Image and Graphics*, vol. 4, 2014.

[21] O. B. Hoque, M. I. Jubair, M. S. Islam, A. F. Akash, and A. S. Paulson, "Real Time Bangladeshi Sign Language Detection using Faster R-CNN," *Int. Conf. Innov. Eng. Technol. ICIET*, 2018.

[22] L. Fei-Fei, J. Deng, and K. Li, "ImageNet: Constructing a large-scale image database," *J. Vis*., vol. 9, no. 8, pp. 1037-1037, 2010.
[23] A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," 2017, [Online], Available: https://arxiv.org/abs/1704.04861
[24] D. Rathi, "Optimization of Transfer Learning for Sign Language Recognition Targeting Mobile Platform," *International Journal on Recent and Innovation Trends in Computing and Communication*, pp. 198-203, 2018.
[25] I. Kemendikbudet al., "WEMOTE –Word Embedding based Minority Oversampling Technique for Imbalanced Emotion and Sentiment Classification*," PloS One*, vol. 8, no. 3, pp. 1-16, 2018.

GRAPHIES OF AUTHORS

Tazkia Mim Angona received the B.Sc in Information and Communication Technology from Bangladesh University of Professionals. She is working on several research projects currently. Her research interests include Computer Vision, Pattern Recognition, Artificial Intelligence, Machine Learning and data mining.

A. S. M. Siamuzzaman Shaon received the B.Scand in Information and Communication Technology from Bangladesh University of Professionals. His research interests include Computer Vision, Pattern Recognition and Artificial Intelligence.

Kazi Tahmid Rashad Niloy received the B.Sc in Information and Communication Technology from Bangladesh University of Professionals. His research interests include Computer Vision and Machine Learning.

Tajbia Karim is serving as a lecturer in Information and Communication Technology (ICT) department at Bangladesh University of Professionals (BUP). Previously she worked in Electrical and Electronic Engineering department of American International University-Bangladesh. Besides being an academician, she is voluntarily working as a student advisor of the ICT department students enrolled in 2019 at BUP. Till now she has published two journal and three conference papers. She has a diversified research interest including Electronics, Signal processing, Data science etc.

Zarin Tasnim is working as a Lecturer in the Department of Information and Communication Technology (ICT) under the Faculty of Science and Technology (FST), Bangladesh University of Professionals (BUP). She achieved B.Sc. (Engineering) degree in Electrical, Electronic and Communication (EECE) Engineering from Military Institute of Science and Technology (MIST), Dhaka, Bangladesh in 2014 and M.Sc. (Engineering) degree in Telecommunications and Wireless Systems from the University of Liverpool, UK in 2019. She has one national journal and one international conference publications.

S. M. Salim Reza is an Assistant Professor of the Department of ICT, Bangladesh University of Professionals (BUP), Dhaka, Bangladesh. He is a Professional Member of IEEE and many other professional organizations. He is a keen planner & implementer with proficiency in managing academic and administrative activities entailing operational policies/norms, faculty appraisal/computer science training. He has more than 50 research papers in ISI/Scopus Indexed journal and many reputed International conferences.

Tasmima Noushiba Mahbub, currently studying in BSc in Information and Communication Technology at Bangladesh University of Professionals, is working on several research projects. Her research interests include IoT, Cyber Security, Signal Processing and Machine Learning.