Developing a Bangla Currency Recognizer for Visually Impaired People

Hasan Murad
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
1305040.hm@ugrad.cse.buet.ac.bd

Nafis Irtiza Tripto
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
nafisirtiza@cse.buet.ac.bd

Mohammad Eunus Ali
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh
eunus@cse.buet.ac.bd

ABSTRACT

Deep learning based assistive technologies for the visually impaired and blind people have gained increasing attention from various research communities in recent years. In this paper, we have developed a camera-based automatic currency recognizer for Bangladeshi bank notes that assists visually impaired people in Bangladesh. We have exploited the deep learning architecture MobileNet for classification of bank notes. We have evaluated the performance of our model using a novel dataset consisting of nearly 8000 images of Bangladeshi bank notes. To verify the effectiveness and efficacy of the proposed solution, we have developed a mobile Android application, and evaluated and validated the application with the users from a blind community. The validation shows that our proposed system is robust and highly effective with heterogeneous environment.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing design and evaluation methods**;

KEYWORDS

AI for good, Bangla currency recognizer, deep learning, blind people.

ACM Reference Format:
Hasan Murad, Nafis Irtiza Tripto, and Mohammad Eunus Ali. 2019. Developing a Bangla Currency Recognizer for Visually Impaired People. In *Tenth International Conference on Information and Communication Technologies and Development (ICTD '19), January 4–7, 2019, Ahmedabad, India.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3287098.3287152

ICTD '19, January 4–7, 2019, Ahmedabad, India

ACM ISBN 978-1-4503-6122-4/19/01…$15.00
https://doi.org/10.1145/3287098.3287152

1 INTRODUCTION

Advances in Artificial Intelligence (AI) are beginning to make the lives of visually impaired people much richer by allowing them to access to the physical world through computer vision technologies. These technologies that facilitate accessibility, safety, and improved quality of life for visually impaired people have profound social impact [6]. In this study, we aim to solve currency recognition problem as it is a major difficulty faced by visually impaired people in Bangladesh. A recent study by the World Health Organization[1] reports that 253 million people are visually impaired globally, in which 36 million are blind and 217 million have moderate to severe vision impairment. In Bangladesh, 0.8 million people are blind[2]. While interacting with the environment, visually impaired people face a number of challenges in their daily life. Currency recognition is one of the most ubiquitous problem faced by visually impaired people due to increased amount of financial transactions occurring face to face.

According to the American Foundation for the Blind, one way that a blind person can identify a paper currency is to fold each denomination in different ways. Although the idea of folding the currency seems promising, it needs others' help to organize bills. People with 100% visual impairment require notes with different sizes and tactile marks that they can be felt easily by touching a bank note. Individuals with low vision need contrast colours and large fonts to identify currency. However, these techniques have been proven ineffective as they fail to distinguish among different bank notes.

Although there are a number of studies on camera-based currency recognition that have been published in literature [1–3, 5], most of these studies are restricted to specific environment or standard background. A common restriction

[1] http://www.who.int/blindness/en/
[2] https://www.theguardian.com/world/2010/sep/28/bangladesh-volunteers-childhood-blindness-treatment

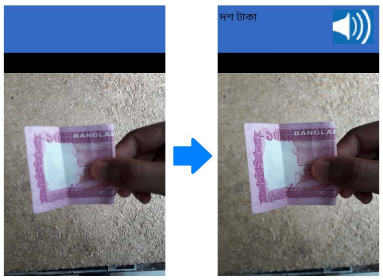

Figure 1: A sample execution of our application. In the left figure, the user opens the app and holds the currency in front of camera. In the right figure, the value of the currency is shown with a voice output.

is that the full note must be visible without wrinkles, occlusion, etc. In context of Bangla currency recognition, Rahman et al. [7] had 89.4% accuracy in homogeneous background using ORB, and with the restriction of the full scan of a note, Jahangir et al. [5] achieved 98.57% accuracy. However, imposing these restrictions in camera based currency recognition for visually impaired people is not feasible.

Currently, there are several smartphone applications available for recognizing currencies of different countries; example as LookTel Money Reader[3], IDEAL Currency Identifier[4], etc. However, they do not have the interface for Bangla currency recognition. Moreover, these apps are not totally satisfactory if we consider the background effect and low end phone configuration.

Therefore, the main goal of this research is to develop a Bangla currency recognizer using a deep learning based approach that can recognize different Bangla bank notes with high accuracy and high reliability in a unconstrained environment. To accomplish this goal, we have first created a novel Bangladeshi currency dataset consisting of more than 8000 images. Then, we have built a model based on the MobileNet deep learning architecture. We have also developed an Android application to aid visually impaired people in practical scenario. To test the viability of our system, we have conducted a feasibility analysis in a blind society of Bangladesh and evaluated the performance of our application in real environment. Experimental results involving blind people in real scenarios validate the high efficiency and efficacy our solution.

2 PROPOSED SOLUTION

In this Section, we provide a basic overview of the proposed solution of our problem. Since there is currently no publicly

[3]http://www.looktel.com/

[4]https://play.google.com/store/apps/details?id=org.ideal.currencyid&hl=en

Figure 2: Bangladeshi currency dataset of different background, illumination and condition

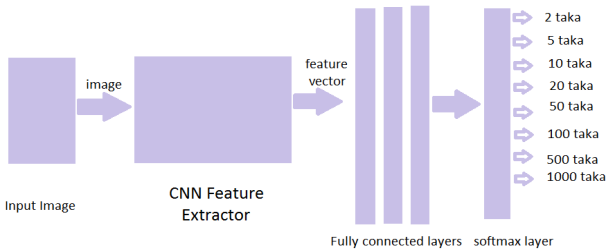

Figure 3: Architecture of our model

available dataset of Bangladeshi bank note images, we have created a Bangladeshi currency dataset that contains more than 8000 images. There are currently eight different categories of Bangladeshi paper currency (2, 5, 10, 20, 50, 100, 500 & 1000 Taka). Our dataset has equal number of images for each category. Figure 2 shows some sample images from the dataset.

There are two distinct parts in our proposed architecture. The first part is the automatic feature extraction process and the later one is the classification task. Figure 3 represents a high level overview of our proposed architecture. For sufficient, and distinctive feature extraction, we use MobileNet [4] which is smaller and more efficient model than VGG, Inception, ResNet. The classification part consists of fully connected dense layers and one softmax layer, which generates a probability between 0 to 1 for each category of bank note.

The MobileNet architecture only accepts input image of size 224×224 pixels. Therefore, we resize input image to 224×224 pixels and normalize each pixel value between 0 to 1. To ensure high accuracy and to avoid overfitting,

we perform different data augmentation. We choose online data augmentation because of its effectiveness to reduce overfitting. We also apply different augmentation techniques like rotation, zooming, shifting and shearing. In our training process, we utilize the fine-tune method since it is the most appropriate training approach for small scale dataset. In this approach, we initialize our model with pre-trained weights and fine tune the entire model.

3 APPLICATION DEVELOPMENT

In this Section, we discuss the requirement analysis part of our Android application. We also present the application design and implementation process.

Requirement Analysis

In order to conduct feasibility study and requirement analysis of our desired application, we visit Bangladesh Visually Impaired People's Society (BVIPS). It is an autonomous organization dedicated for the interest of visual impaired community. We have conducted rigorous face to face interview with three male and two female members from visually impaired community in BVIPS. During the interview, we focus on the technology engagement of the visually impaired people and identify unique requirements for our application from visually impaired people.

Since capturing image is a difficult task for blind people, they prefer a video mode based application where the value of the currency can be identified from the video streams of the camera. Blind people usually navigate any application by using TalkBack application. Therefore, we need to label all the setting options so that TalkBack or built-in accessibility features can operate with our application properly. Nested setting options are difficult to navigate for blind people. The number of clicks to access any feature should be at minimal level. Therefore, we need to implement all setting options in single depth. The application should also notify the recognition result with a clear and non-robotic sound for visually impaired people.

Since we need to develop a faster currency recognizing application within the standard application size, we utilize MobileNet which is a light weighted model and suitable for mobile device. We also need to ensure that no internet connection is necessary for currency recognition.

Application Design and Implementation

To ensure comprehensibility of our application, we develop a real time currency recognizer. The users only need to open the application and then place it in front of a note. As blind people can comfortably use Android phone with the help of assisting technologies, we develop our application in android platform. Our application works in offline mode, and recognizes currency in video mode. The application collects image frame from the video and recognizes the corresponding currency note. After identifying the note, we notify the user with a clear sound in both Bangla and English language.

4 EXPERIMENTS AND RESULTS

We analyze the performance of our proposed architecture with Bangladeshi currency dataset. We divide the dataset into three subset: training(80%), validation(10%), and testing(10%). In each subset, there is an equal number of images for each class.

We use Keras framework with Tensorflow as a background to implement all methods for training, tuning, and testing. For optimization, we use Adam optimizer as it is one of the most effective optimization techniques. To evaluate the performance of our model after each iteration, categorical accuracy is used as desired performance metric and categorical cross entropy as loss function. We attach dropout layer after each dense layer to ensure the regularization. The dropout rate is set to 0.3. We choose a batch size of 32.

We train our model by using MobileNet as feature extractor.As we train the model for 100 epochs, the size of augmented training dataset is 100×6400 or 640000. For the mentioned parameter setting, we get convergence of our model after 30 epoch on average. We achieve maximum accuracy of 99.81% in validation and 99.80% in testing.

5 EVALUATION

We evaluate our application by using two different approaches. First approach is to evaluate with individual tester to find any limitation of our application. In the next step, we distribute our application among the blind users and get their feedback.

Evaluation with Testers

We perform an alpha testing on our application to find out any defect or false recognition before releasing the application for blind people. The performance is measured in terms of true recognition of each note. We randomly select a specific group consisting of ten people. We provide them an early version of Bangla currency recognizer. As proper lighting plays an important role while capturing an image, they test the application in different lighting conditions. They also use different phone cameras to conduct the test. We provide a summary of the result in Table 1.

The testers investigate our application in practical scenarios and get high accuracy in most of the cases. The only fault with our application that has been found so far is recognizing background as a valid note as our dataset does not cover all the variations of background. Therefore, some testers report these errors as false classifications. If we extend the size of our dataset, we can minimize these errors. We show some background images classified as notes in Figure 4.

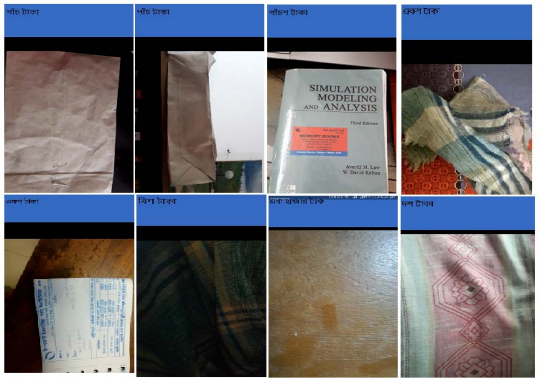

Figure 4: Background images classified as bank notes.

Table 1: Testing summary for individual user

Tester ID	Number of Test	Right	Wrong
1	50	49	1
2	50	47	3
3	50	48	2
4	50	49	1
5	50	49	1
6	50	48	2
7	50	47	3
8	50	50	0
9	50	48	2
10	50	50	0

Evaluation with Visually Impaired People

We have provided the first version of Bangla currency recognizing application to Bangladesh Visually Impaired People's Society (BVIPS). To introduce our application, we deliver them a general demonstration and training. In that training program, they learn how to use our application to recognize currency. They cannot focus the bank note properly by using their phone camera at the beginning. Most of the time, the paper note remains out of the camera range. After several trials, they manage to focus appropriately.

BVIPS takes the responsibility to distribute Bangla currency recognizer to the blind people. They were excited to share the application to the visually impaired community. In Figure 5, we see that they have shared in social media to contact with them to collect the application[5]. They have also agreed to provide regular feedback to us about the performance of the application.

We have also conducted an interview with visually impaired people in BIVPS to investigate the usefulness, simplicity and accuracy of the application. They have provided a grade between 0 to 5 for each of the criterion. We provide a summary in Table 2. The application helps them to instantly recognize paper notes during financial transaction. They appreciate the simplicity of our application design and are satisfied with high recognition accuracy.

[5]https://goo.gl/pukZoT

Table 2: Feedback summary from participants

Real User ID	Usefulness	Simplicity	Accuracy
1	5	4.5	5
2	5	5	4.5
3	5	4.5	5
4	5	5	5
5	5	5	4.5

Figure 5: A screenshot taken from a Facebook post by BVIPS.

6 CONCLUSION

In this work, we have developed a mobile based Bangla currency recognizer application for visually impaired people. For this purpose, we have created a novel Bangladeshi currency dataset. We have utilized a recent deep learning based model called MobileNet and achieved 99.8% accuracy in testing. We have also evaluated our mobile application with visually impaired people and the application has received much appreciation from the community. We have also established an effective communication channel with BVIPS for regular evaluation, feedback, and improvement of the application.

REFERENCES

[1] Seung-Hoon Chae, Jong Kwang Kim, and Sung Bum Pan. 2009. A study on the Korean banknote recognition using RGB and UV information. In *Communication and Networking*. Springer, 477–484.

[2] P.M. Farahabadi H. Hassanpour. 1992. Using Hidden Markov Models for paper currency recognition. In *Expert Systems with Applications*, Vol. vol. 5. pp. 10105–10111.

[3] Faiz M Hasanuzzaman, Xiaodong Yang, and YingLi Tian. 2011. Robust and effective component-based banknote recognition by SURF features. In *Wireless and Optical Communications Conference (WOCC), 2011 20th Annual*. IEEE, 1–6.

[4] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. 2017. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *CoRR* abs/1704.04861 (2017). arXiv:1704.04861 http://arxiv.org/abs/1704.04861

[5] N. Jahangir and A. R. Chowdhury. 2007. Bangladeshi banknote recognition by neural network with axis symmetrical masks. In *2007 10th international conference on computer and information technology*. 1–5. https://doi.org/10.1109/ICCITECHN.2007.4579423

[6] Roberto Manduchi and James Coughlan. 2012. (Computer) vision without sight. *Commun. ACM* 55, 1 (2012), 96–104.

[7] Mohammad M. Rahman, Bruce Poon, M. Ashraful Amin, and Hong Yan. 2014. Recognizing Bangladeshi Currency for Visually Impaired. In *Machine Learning and Cybernetics*, Xizhao Wang, Witold Pedrycz, Patrick Chan, and Qiang He (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 129–135.