Miami University

Department of Electrical and Computer Engineering

Convolutional Neural Networks based bangla handwritten character recognition from CMATERdb

Authors
Md Rakibul Ahasan

Advisor
Dr. Chi-Hao Cheng

Contents

1 Bangla Character Recognition

1.1 Project Objective

In Bangla language there are 50 characters where 11 are vowel and 39 is consonant. Proposal is to recognize the characters from the list of hand written images by convolutional neural network (CNN). The convolutional neural network has to build from scratch. High level details are as below.

1. I need to use convolutional neural network however the programming language can be any. I use python for my project implementation.
2. I can use any of the online bangla language dataset to train my CNN model. I used a reduced version of CMATERdb dataset. This dataset contains 240 images for each of the character for training and 60 images per character for testing.
3. I need to evaluate my model performance by checking it through a handwritten bangla character image. I have taken those images from one of my fellow colleague and evaluated accordingly.

1.2 Data Set Details

CMATERdb is a very popular dataset for bangla handwritten character recognition. It's very huge dataset though I will not use the full set of CMATERdb due to my computational resource constraints. Instead I will use a reduced version of CMATERdb. Details of this dataset is in below table:

Table 1.1: Dataset details used in this paper

Type	DataSetDirectory	Folder Count	ImagePerFolder	TotalSample
Train	/Dataset/Train	50	240	12000
Test	/Dataset/Test	50	60	3000

All the image size is $32 * 32 * 3$. A sample of those images is given in Figure 1.1

Figure 1.1: Bangla character visualization from CMRETdb

1.3 CNN architecture and Parameters Used In This Project

A CNN is widely used in object recognition and I will use CNN to train the data set and the classify the handwritten image. The structure of my CNN neural network is in Figure 1.2. Here each of the convolution layer is constructed with Conv2D of Keras api. Keras Conv2D is a 2D Convolution Layer where I use the parameter 'filters' to defines the number of filter the convolution layer will learn form, the 'kernel_size' defines the window size which extract features from the image. Here 'activation' refers to the activation function we want for CNN after performing the convolution. I use 'relu' as my activation function. The fist convolution layer also define the input_shape of the CNN network.

```
classifier=Sequential()

classifier.add(Conv2D(filters=128, kernel_size=(3,3), activation='relu', input_shape=(40,40,3)))
classifier.add(MaxPooling2D(pool_size=(2,2)))
classifier.add(Dropout(.2))

classifier.add(Conv2D(filters=64, kernel_size=(3,3), activation='relu'))
classifier.add(MaxPooling2D(pool_size=(2,2)))
classifier.add(Dropout(.2))

classifier.add(Conv2D(filters=32, kernel_size=(3,3), activation='relu'))
classifier.add(MaxPooling2D(pool_size=(2,2)))
classifier.add(Dropout(.2))

classifier.add(Flatten())

#classifier.add(Dense(units=64, activation='relu'))
#classifier.add(Dropout(.2))

classifier.add(Dense(units=128, activation='relu'))
classifier.add(Dropout(.2))

classifier.add(Dense(units=50, activation='softmax'))

classifier.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])
```

Figure 1.2: CNN Architecture used in this project

In some of layer I use max pooling to reduce the spatial dimensions of the output volume instead of average pooling. The main reason I used max pooling because it significantly improve the computational performance. I have also use another important layers which is called 'dropout'. This layer preventing the over-fitting case. During learning of data there are lots of noise which will deviate the outcome. The dropout layer will reduce that effect by removing noise data during the training process.There are some other important training parameter of CNN. I use the 'adam' optimizer, model evaluation metrics is 'accuracy', the number of epoc is 150 and steps per epoc is 375 image. Moreover there are another important layer those are 'Flatten' and 'Dense'. In a layer flatten function pulled out the feature map into a single column, later that passed to the fully connected layer. The dense layer is classify the image based on the output from convolutional layer. Some important details of CNN network which is used is in 1.2.

Table 1.2: Data Augmentation

ImageDataGenerator	Training Approach 1	Training Approach 2
rescale	1./255	1./255
shear_range	0.2	0.2
rotation_range	25	25
width_shift_range	2	NotUsed
height_shift_range	2	NotUsed
zoom_range	[0.7,1.0]	NotUsed
Training_accuracy	80.46	90.19
Validation_accuracy	90.6	93.83

To enhance the diversity of training data set I perform popular data augmentation library from python. This library is called 'ImageDataGenerator'. There are different parameter within the

ImageDataGenerator has an impact on the CNN output. My parameter selection is determined by which combination will give the highest validation accuracy. Details are in Table 1.2 which shows in training approach 2 with 'rescale', 'shear_range', and 'rotation_range' parameter the validation accuracy achieved is 93.83.

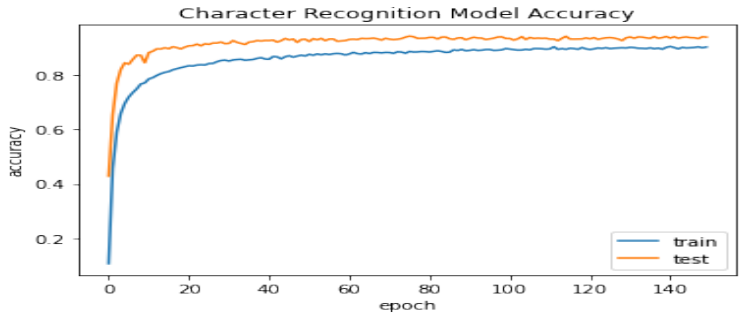

Figure 1.3: Training accuracy vs Validation accuracy

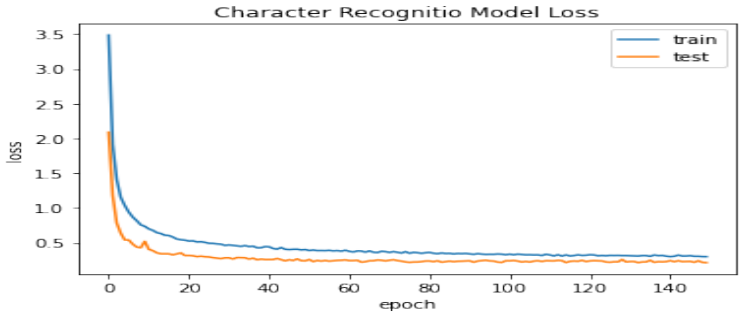

Figure 1.4: Training loss vs Validation loss

After training the CNN model achieved 90.19 accuracy in the training dataset and 93.83 during validation. Details are in Figure 1.3. Figure 1.4 specifically detailing the loss incurred during the training and validation.

I am satisfied with the performance and I saved the model as "*bangla_character_recognition.json*" and it weights under "*bangla_character_recognition.h5*" file. Those file is present in the project directory */ECE529_Project_Bangla_Handwritten_Character_Recognition* and this model python file name is *529_Model_Train.ipnyb*.

1.4 Model Performance Evaluation

In this section I will be describing how to evaluate the performance of our trained model in hand written bangla language. To do that I took help from one of my fellow colleague who written two bangla character both in paper and using computer visual. Hence my total sample of bangla character is four.

Below are the steps needs to follow to evaluate model performance:

1. Need to initiate *529_Bangla_Character_Recognition* python program and make sure the json file "*bangla_character_recognition.json*" of model and "*bangla_character_recognition.h5*" weight file is loaded.

2. *determine_character*(*res*), *single_prediction*, and *guess_the_character* fuctions need to declare.

3. lastly *guess_the_character* funtion need to run and it will take one single image at a time and show the bangla language character prediction result.
4. Note: Need to change the image file name in the *guess_the_character* funtion one by one. The syntax is something like that */SinglePrediction/*1.*jpg*.

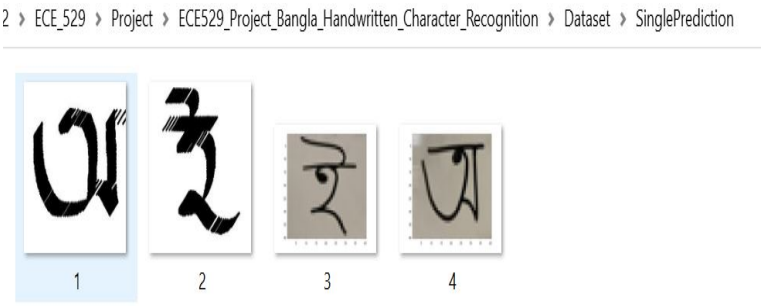

Figure 1.5: Sample images in SinglePrediction folder

Figure 1.5 is giving the view of four handwritten sample image in this project.Figure 1.6 is showing the final result of my project. Out of 4 hand written image one image is rightly predicted, two image has close prediction and remaining image is wrongly predicted.

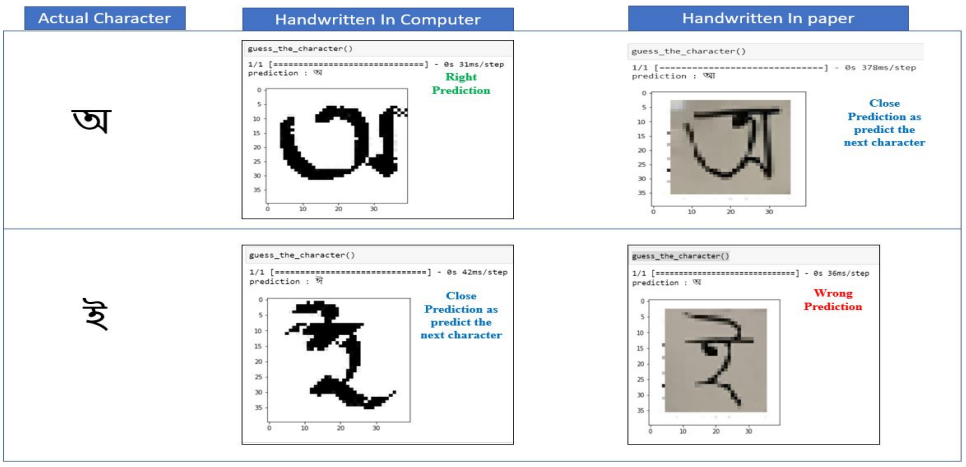

Figure 1.6: Model performance over handwritten data

2 Conclusion

I am very much delighted to work on project which is based on my native language and I siccerely thanks Dr. Chi-Hao Cheng for his direction. I have learned many things from this project below has the details of it:

1. I understand the details of convolutional neural network much better now and what are parameters are important designing a convolutional neural network.
2. I have now hands on experience of building a CNN model from scratch.
3. I have achieved sufficient accuracy in training and validation 90.19 and 93.83 respectively.
4. I know now how to train a CNN model using a dataset and used it later for any application building.
5. I learn how to build function in python and understand how it can improve the efficiency.
6. At last I have validate my model with a practical scenario of handwritten bangla language and understand how to analyze the outcome.