International Journal of Modern Electronics and Communication Engineering (IJMECE) ISSN: 2321-2152
Volume No. - 8, Issue No. - 5, September, 2020

Enhance Sensitivity Of Bangla Handwritten Digit Recognition Using Ten Layered D-CNN Model

R.Rani1
Associate Professor
Department of ECE,
Vemu Institute of Technology,
Chittoor , India.
E-mail:ranikps30@gmail.com

Abstract: This algorithm is hardly ever used in Knowledge of handwritten digits like bangle. This mission proposes a deep convolution neural network (D-CNN) primarily based Bangla hand written digits recognition. This D-CCN has seven layers. Mainly three convolution layers, three pooling layers and thoroughly related layer. Deep convolution neural network has these days received recognition due to the fact of its improved Performance over the typical computer learning algorithms. However, it has been very not often used on cognizance of bangle handwritten digit. The proposed method can decorate layers to ten or twelve layers via using deep CNN architecture for recognizing bangle handwritten digits with excessive sensitivity/specificity.

I. INTRODUCTION

The principle challenge in manually written character characterization is to manage the large assortment of handwriting styles by various essayists in various languages. A portion of the intricate penmanship contents contain various styles for composing words. In some different cases, they are cursive and some of the time the characters are associated with one another (e.g., English, Bangladeshi and Arabic). These difficulties are as of now perceived by numerous specialists in the field of Natural Language Processing (NLP).Handwritten character recognition is more difficult comparing to printed forms of characters. This is because characters written by different people are not identical and varies in different aspects such as size and shape. The similarities in different character shapes, the overlaps, and the interconnections of the **NEIGHBORING** characters further complicate the character recognition problem. Therefore, here in our project we are able to recognize hand written CHARACTERS OF DIFFERENT STYLES BY USING D-CNN METHOD. WE USED BANGLA LANGUAGE CHARACTERS AS EXAMPLE TO WORK WITH OUR METHODS.

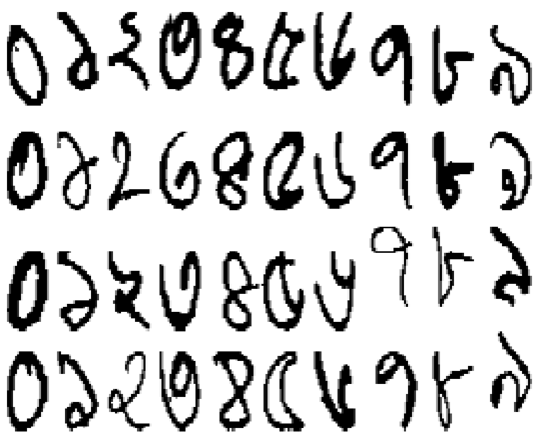

FIG1 : BANGLA LANGUAGE DIGITS USED AS REFERENCES

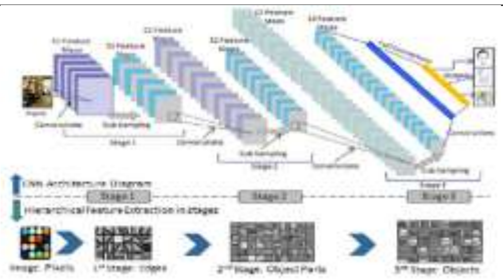

FIG:2 DEEP CNN ARCHITECTURE

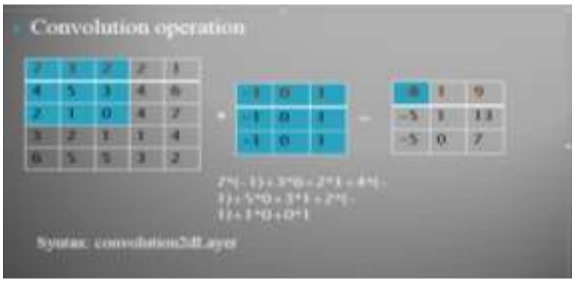

FIG 1.3 CONVOLUTION OPERATION

.

Fig 1.4 Average polling operation

II. LITERATURE REVIEW

C. Saha, R. H. Faisal, and M. M. Rahman: Optical Character Recognition (OCR) especially for handwritten characters is an important task for its numerous applications in daily life including data digitizing, robotics vision, helping visually disabled people and many more. However, Bangla Handwritten Character Recognition (HCR) is rarely explored despite Bangla being one of the mostly spoken languages over the world. For classifying Bangla basic characters, compound characters and digits various feature descriptors and classification algorithms can be used. This project provides a comparative study of different Local Binary Pattern (LBP) based feature descriptors on Bangla basic characters, compound characters and digits. Bangla compound characters and Bangla digits respectively have showed reasonable accuracies of different LBP based feature descriptors.

N. Das, S. Basu, R. Sarkar, M. Kundu, and M. Nasipuri: Appropriate feature set for representation of pattern classes is one of the most important aspects of handwritten character recognition. The effectiveness of features depends on the discriminating power of the features chosen to represent patterns of different classes. However, discriminatory features are not easily measurable. Investigative experimentation is necessary for identifying discriminatory features. In the present work we have identified a new variation of feature set which significantly outperforms on handwritten Bangla alphabet from the previously used feature set. 132 numbers of features in all viz. modified shadow features, octant and centroid features, distance based features, quad tree based longest run features are used here.

III. EXISTING METHOD:

Bangla handwritten digit recognition using an improved deep convolutional neural network architecture: This technique uses 7 layered D-CNN model.

- It consists of 3 Convolution Layers, 3 Pooling Layers and 1 fully connected Layer.
- It gives an accuracy of 80% - 97%
- It provides sensitivity of 0.01 to 0.5 based on test data.

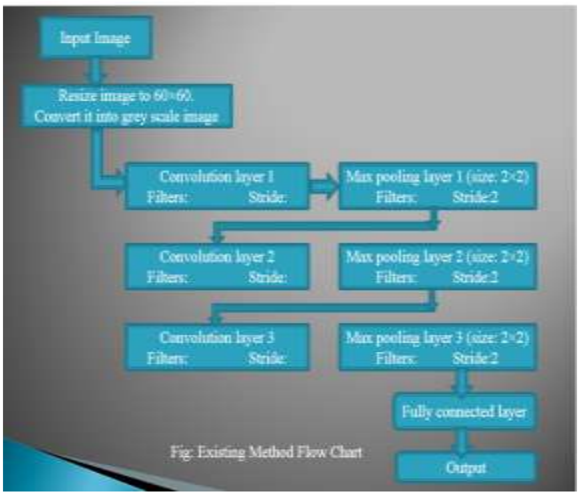

Fig 3.1: Existing Method Flow Chart

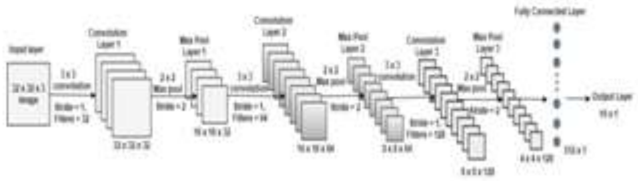

Fig 3.2 : Block Diagram of Existing Method

3.1 Convolutional Neural Network

A convolutional neural network (CNN or ConvNet) is one of the most popular algorithms for deep learning, a type of machine learning in which a model learns to perform classification tasks directly from images, video, text, or sound. CNNs are particularly useful for finding patterns in images to recognize objects, faces, and scenes. They learn directly from image data, using patterns to classify images and eliminating the need for manual feature extraction.

Applications that call for object recognition and computer vision such as self-driving vehicles and face-recognition applications rely heavily on CNNs. Depending on your application; you can build a CNN from scratch, or use a pertained model with your dataset.

CNNsUses

Using CNNs for deep learning has become increasingly popular due to three important factors:

- CNNs eliminate the need for manual feature extraction the features are learned directly by the CNN.
- CNNs produce state-of-the-art recognition results.
- CNNs can be retrained for new recognition tasks, enabling you to build on pre-existing networks.

Working Principle of CNN

A convolutional neural network can have tens or hundreds of layers that each learn to detect different features of an image. Filters are applied to each training image at different resolutions, and the output of each convolved image is used as the input to the next layer. The filters can start as very simple features, such as brightness and edges, and increase in complexity to features that uniquely define the object. CNNs perform feature identification and classification of images, text, sound, and video.

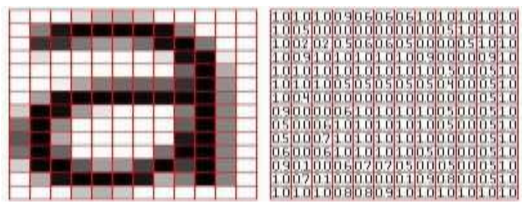

Fig 3.3: Representation of image as a grid of pixels

The human brain processes a huge amount of information the second we see an image. Each neuron works in its own receptive field and is connected to other neurons in a way that they cover the entire visual field. Just as each neuron responds to stimuli only in the restricted region of the visual field called the receptive field in the biological vision system, each neuron in a CNN processes data only in its receptive field as well. The layers are arranged in such a way so that they detect simpler patterns first (lines, curves, etc.) and more complex patterns (faces, objects, etc.) further along. By using a CNN, one can enable sight to computers.

3.2 Advantages:

1. The main advantage of CNN compared to its predecessors is that it automatically detects the important features without any human supervision.

2. CNN is also computationally efficient.

3.3 Disadvantages:

1. Sensitivity is low.
2. Performance is low because ReLu layer which we use

The multidimensional space is mapped into a space of fewer dimensions by transforming the original space using a linear transformation via a principal component analysis.

IV. PROPOSED METHOD

Enhance Sensitivity/Specificity of Handwritten Bangla Digit Recognition Using Ten layered D-CNN Model:

It is the extension of existing method.

1. Here D-CNN Model consists of 12 layers in total including input and output layers.
2. They are 3 convolution layers, 3 batch normalization layers, 3 average polling layers, 1 fully connected layer.
3. Also we have ReLu activation function and softmax layer to speed up the process.
4. It produces an accuracy of 97% to 99%.
5. It produces sensitivity of 0.5 to 0.8 and in some cases we also acquire sensitivity to 1 based on test data.

Fig 4.1: Block diagram for proposed method

The functions which are used in this discussed below:

Convolution Layer

The convolution layer is the core building block of the CNN. It carries the main portion of the network's computational load. This layer performs a dot product between two matrices, where one matrix is the set of learnable parameters otherwise known as a kernel, and the other matrix is the restricted portion of the receptive field. The kernel is spatially smaller than an image, but is more in-depth. This means that, if the image is composed of three (RGB) channels, the kernel height and width will be spatially small, but the depth extends up to all three channels.

Batch Normalization Layer

A batch normalization layer normalizes each input channel across a mini-batch. To speed up the training of convolution neural networks and reduce the sensitivity to network initialization, it is used between convolution layer and nonlinearities, such as ReLU layers, The layer first normalizes the activations of each channel by subtracting the mini-batch mean and dividing by mini-batch standard deviation. Then the layer shifts the input by a learnable offset β and scales it by a learnable scale factor. Syntax: batch normalization Layer ('Name', Value)

4.1 Average pooling layer

An average pooling layer performs down-sampling by dividing the input into rectangular pooling regions and computing the average values to each region.

On two-dimensional feature maps, pooling is typically applied in 2×2 patches of the feature map with a stride of (2, 2). Average pooling involves calculating the average for each patch of the feature map. This means that each 2×2 square of the feature map is down sampled to the average value in the square. Because the down sampling operation halves each dimension, we will expect the output of pooling applied to the 6×6 feature map to be a new 3×3 feature map. Syntax: averagePooling2dLayer (pool Size, Name, Value)

4.2 Fully ConnectedLayer

A fully connected layer multiplies the input by a weight matrix and then adds a bias vector. Fully connected layers connect every neuron in one layer to every neuron in another layer. It is in principle the same as the traditional multi-layer perceptron neural network (MLP). The flattened matrix goes through a fully connected layer to classify the images. Syntax: fully Connected Layer.

4.3 SoftmaxLayer

A softmax layer applies a softmax function to the input. For classification problem, a softmax layer and then a classification layer must follow the final fully connected layer. The softmax function is the output unit activation function after the last fully connected layer for multi-class classification problems. The softmax function is also known as the normalized exponential and can be considered the multi-class generalization of the logistic sigmoid function. Syntax: softmax Layer (‘Name’, Name)

4.4 ClassificationLayer

A classification layer computes the cross entropy loss for multi-class classification problems with mutually exclusive classes. The layer infers the number of classes from the output size of the previous layer. For example, to specify the number of classes *K* of the network, include a fully connected layer with output size *K* and a softmax layer before the classification layer. Syntax: classification Layer (Name, Value)

4.5 Non-LinearityLayers

Since convolution is a linear operation and images are far from linear, non-linearity layers are often placed directly after the convolutional layer to introduce non-linearity to the activation map. There are several types of non-linear operations, the popular ones being:

4.5.1 Sigmoid:

The sigmoid non-linearity has the mathematical form $\sigma(\kappa) = 1/(1+e^{-\kappa})$. It takes a real- valued number and “squashes” it into a range between 0 and 1.

However, a very undesirable property of sigmoid is that when the activation is at either tail, the gradient becomes almost zero. If the local gradient becomes very small, then in back propagation it will effectively “kill” the gradient. Also, if the data coming into the neuron is always positive, then the output of sigmoid will be either all positives or all negatives, resulting in a zigzag dynamic of gradient updates for weight. A sigmoid function is a bounded, differentiable, real function that is defined for all real input values and has a non-negative derivative at each point. A sigmoid "function" and a sigmoid "curve" refer to the same object. In general, a sigmoid function is monotonic, and has a first derivative which is bell shaped. A sigmoid function is constrained by a pair of horizontal asymptotes as x\rightarrow \pm\infty.

The sigmoid function is convex for values less than 0, and it is concave for values more than 0. Because of this, the sigmoid function and its affine compositions can possess multiple optimal. It is easy to understand and apply but it has major reasons which have made it fall out of popularity.

- Vanishing gradient problem
- Secondly, its output isn’t zero centered. It makes the gradient updates go too far in different directions. 0 < output < 1, and it makes optimization harder.
- Sigmoid saturate and kill gradients.
- Sigmoid have slow convergence.

4.5.2 Tanh:

Tanh squashes a real-valued number to the range [-1, 1]. Like sigmoid, the activation saturates, but unlike the sigmoid neurons its output is zero centered. Hence optimization is easier in this method hence in practice it is always preferred over Sigmoid function. But still it suffers from Vanishing gradient problem.

4.5.3 ReLU- Rectified Linear units:

It has become very popular in the past couple of years. It was recently proved that it had 6 times improvement in convergence from Tanh function. It’s just R(x) = max (0, x) i.e. if x < 0, R(x) = 0 and if x >= 0, R(x) = x. Hence as seeing the mathematical form of this function we can see that it is very simple and efficient. A lot of times in Machine learning and computer science we notice that most simple and consistent techniques and methods are only preferred and are best. Hence it avoids and rectifies vanishing gradient problem. Almost all deep learning Models use ReLU nowadays. Most Deep Learning applications right now make use of ReLU instead of Logistic Activation functions for Computer Vision, Speech Recognition and Deep Neural Networks etc. Some of the ReLU variants include: Soft plus (Smooth RELU), Noisy ReLU, Leaky ReLU, Parametric ReLU and Exponential ReLU(ELU).

V. EXPERIMENTAL RESULTS:

5.1.1: Existing Method Results:

Fig 5.1.1: Input Image

Fig 5.1.2: Accuracy and loss for seven layered D-CNN

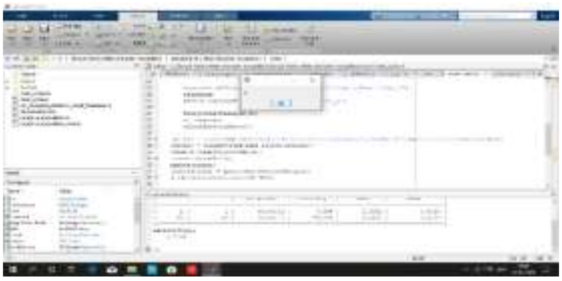

Fig 5.1.3: Sensitivity of seven layered D-CNN

Fig 5.1.4: Output displayed in message box

5.2 Proposed Method Results

Fig 5.2.1: Input image for ten layered architecture

. **Fig 5.2.2:** Accuracy, loss and output displayed in message box for ten layered architecture

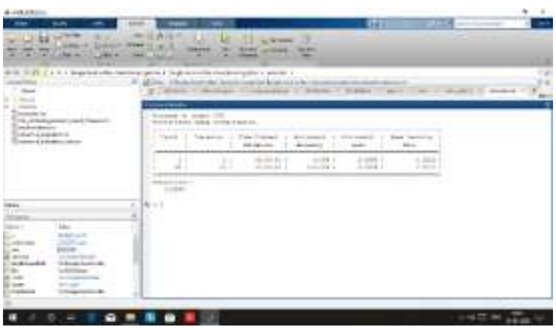

Fig 5.2.3: Sensitivity for ten layered architecture

VI. CONCLUSION AND FUTURE SCOPE:

6.1. CONCLUSION:

A ten layered D-CNN model is proposed in this project for Bangla handwritten isolated digits, which provides up to 99% accuracy. However, using data augmentation to increase the amount of data and a deeper network may increase the performance significantly which will be addressed in future works.

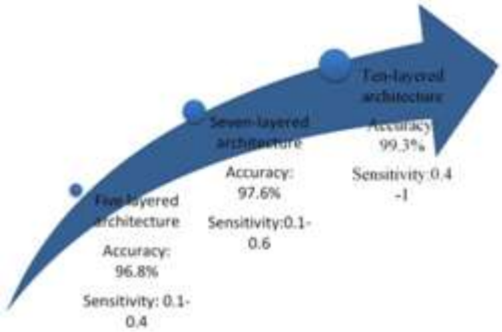

Fig: Accuracy and Sensitivity Results for different D-CNN architectures

Hence this project can be used in Number plate recognition, data digitizing, and automatic key information extraction, smart education, helping visually disabled people, computer vision and many other sectors.

6.2. FUTURE SCOPE:

- In future this technique can be applied to any language digit recognition which can be handwritten form or printable form.
- It can also be used to help differently abled people by converting the recognized digits into audios and videos forms further.
- Braille (tactical writing system) can also be made in any language with the help of this technique to help visually impaired peoples.
- As unknown languages can be converted to our own languages using this technique peoples need not spend much time to learn other language.
- With the help of D-CNN technique we can get maximum accuracy and sensitivity as layers can be extended to any number in future

REFERENCES

[1] ChandrikaSaha, Rahat Hossain Faisal and Md. Mostafijur Rahman: "Bangla handwritten digit recognition using an improver deep convolution neural network architecture" 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February, 2019.

[2] S. Basu, N. Das, R. Sarkar, M. Kundu, M. Nasipuri, and D. K. Basu, "Handwritten bangla alphabet recognition using an mlp based classifier," arXiv preprint arXiv: 1203.0882,2012

[3] N. Das, B. Das, R. Sarkar, S. Basu, M. Kundu, and M. Nasipuri, "Handwritten bangla basic and compound character recognition using mlp and svm classifier," arXiv preprint arXiv: 1002.4040, 2010.

[4] C. Saha, R. H. Faisal, and M. M. Rahman, "Bangla handwritten character recognition using local binary pattern and its variants." In press,2018.

[5] U. Pal, B. Chaudhuri, and A. Bela¨ıd, "A complete system for bangla handwritten numeral recognition," IETE journal of research, vol. 52, no. 1, pp. 27–34,2006.

[6] G. F. Simons, Ethnologue: Languages of the world. Sil International, 2017.

[7] R.Rani,"A Novel Approach on Improve Image De-blurring on RGB Images"in IJMECE,VOL:7,ISSN:2321 2152,Issue no:3,May 2019.

AUTHOR'S BIOGRAPHIES

R.Rani, Associate professor, in Department of Electronics & Communication Engineering at Vemu Institute of Technology,P.Kothakota..Her areas of interest are Embedded Systems, Image Processing. Her Publications includes in 10 International Journals / conferences.