Deep Convolutional Neural Networks for Handwritten Kannada Numerals Recognition

Ramesh G*, Srihari W†, Srinidhi G‡ and Champa H.N§

Department of CSE
UVCE
Bangalore, India
Email: *rameshmg6308@gmail.com, †srihariw1999@gmail.com, ‡srinidhig73@gmail.com, §champahn@yahoo.co.in

***Abstract*—Kannada is an official language of India mostly spoken by people of Karnataka in Southwestern India. This paper describes the recognition/classification of Handwritten Kannada numerals using Deep Convolutional neural network(Deep-CNN). The occurrence of handwritten text is abundant. Convolutional neural networks (CNN) have been known for it's computationally efficient way of extracting features. For the extraction of features a series of convolution and pooling operations is performed. We have used the new Kannada MNIST dataset consisting of 60,000 samples of isolated handwritten numerals and a Handwritten dataset created by non-native users of the language called Dig-MNIST. The proposed system achieves an accuracy of 98.24% on the Kannada MNIST dataset and 86.85% on the Dig-MNIST dataset.**

***Index Terms*—Computer Vision, Convolutional Neural Network, Deep Learning, Handwritten Numerals recognition, Kannada Numerals.**

I. Introduction

Deep learning has been widely used for the recognition of handwritten characters as well as numerals. For humans, the recognition of Handwritten digits is very easy but it is difficult for machines to recognize handwritten digits. Using Deep Learning and Machine Learning methods, this task can be achieved with the models having very high accuracy. Unlike traditional methods which involve different preprocessing steps, deep learning automatically identifies the features. Therefore, deep learning mostly depends on the data and hence can be applied to solve different kinds of problems, including handwritten characters and numerals recognition. There has been a huge progress in the field of Computer Vision with the application of deep learning in it. A Convolutional Neural Network (ConvNet/CNN) is one of the types of deep neural networks which is mainly used to analyze images. It takes images as input and extracts different features from the image. It contains one input layer, one output layer and one or many hidden layers. CNNs are very useful and feasible in dealing with patterns involving spatial arrangements, thus useful in recognition of handwritten characters and digits. One of the main advantages of a Convolutional Neural Network (CNN) when compared to the traditional machine learning techniques is that it consequently distinguishes the significant features in the image with no human oversight. The MNIST dataset is a standard dataset used worldwide consisting of 60,000 samples for training and 10,000 for testing. Kannada language is the official language of the state of Karnataka,India, spoken by over 50 million people all over the world. Similar to the MNIST dataset, the Kannada-MNIST dataset was introduced which contains the Kannada-MNIST dataset as well as Dig-MNIST dataset. The Dig-MNIST is a very challenging dataset when compared to the Kannada-MNIST as the dig-MNIST was created with the help of volunteers that were non-native users of the language. The aim of this paper is to provide a classifier that achieves a very high accuracy on the Kannada-MNIST as well as Dig-MNIST datasets without further pre-processing.

II. Related Work

The recognition of Offline Kannada Handwritten Numerals was implemented by Sharma *et al.,* [1]. In this paper, the author has used a quadratic classifier based scheme for the recognition of offline handwritten Kannada numerals and achieved an accuracy of 90.34% with 2300 data samples. Here the author has given the idea of dimensional feature extraction using block segmentation.

Further down, the recognition of Handwritten Devanagiri Numerals using the SVM classifier was implemented by Jangid *et al.,* [2]. In this paper, the author has used density and background directional distribution features for the zones, in which the numerals were already divided and have used SVM classifier with RBF kernel for classification.

Data Clustering is considered as an intuitive methodology for discovering similarities with regards to data and placing comparative information into gatherings. Here the author Mamatha *et al.,* [3] has used a clustering algorithm for classification called the K-Means algorithm. The proposed method is experimented on 1,000 examples of Handwritten Kannada numerals and have achieved 96.00% accuracy.

A multi-layered neural network architecture was chosen for better accuracy by the authors Majumdar *et al.,* [4]. For the feature selection, the authors use pixel and shape based features such as hole position, curve position, and block-wise corner position. The author has accomplished an accuracy of 97.20% on the dataset consisting of 10,500 images of Handwritten numerals.

Here the authors Gati *et al.,* use the Skip CNN model for the classification of Kannada numerals [5]. The author without any additional preprocessing has accomplished an accuracy of 85.02% on the dig-MNIST and 97.53% on the Kannada-MNIST.

III. Problem Statement

This exertion is coordinated to settle the "Classification/Recognition of Handwritten Kannada numerals ", by utilizing Convolutional neural networks, a dynamic procedure to yield dependable outcomes. The exertion is centred around the identification of these numerals from different wellsprings of writing. The work is likewise pointed towards breaking down the inadequacies in different conventional AI and OCR procedures that are right now existing as the answer for the difficult proclamation.

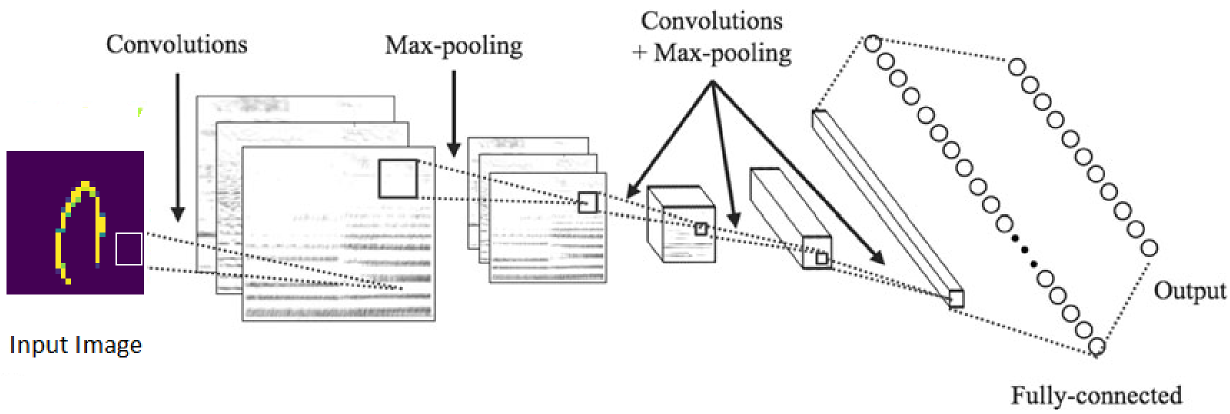

Fig 1 Architecture diagram for the proposed model

IV. Network Architecture

The Convolutional neural network is used both as a feature extractor as well as a classifier. The ConvNet architecture used in this paper is a custom built network inspired by the AlexNet and VGGNet which are very deep network architectures and use several convolutional layers followed by one max pooling layer [6]. The size of the filter used in this architecture is 3x3 , having a stride size of 1,and a max pooling layer of size 3x3. In Figure I, the architecture of the model is illustrated in detail. Zero Padding of size 1x1 is applied to the input of each convolution layer and after every two convolution layers the total number of filters for the convolution layers is increased by a factor of 2 going from 32 to 256. Dropout is applied after every two convolution layers to avoid over-fitting. It has a Fully Connected (FC) layer with 512 neurons and the activation function for every layer is the ReLU activation function except for the final output layer which has the softmax as its activation function.

V. Experimental Results and Discussion

A. Data Set

Our emphasis is precisely on the recognition of digits present in the Kannada-MNIST dataset as well as on the more challenging Dig-MNIST dataset. Both the datasets used in this paper are compiled by [7].

The main Kannada-MNIST dataset has a training set which comprises of 60,000 images converted to grayscale, each of the size 28 x 28 and a testing set which comprises of 10,000 images converted to grayscale with the same size and are divided equally among 10 classes.

The Dig-MNIST dataset contains 10,240 grayscale images of size 28 x 28 organized to give a more challenging dataset. This dataset was prepared with the help of people who were non-native users of the language and had trouble in writing these digits. Fig. 2 represents the digits from 0 – 9 in Kannada.

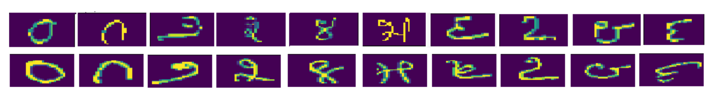

Fig. 2. Sample Dataset: Kannada MNIST and Dig MNIST

B. Results

The Kannada MNIST dataset contains 60,000 images belonging to 10 classes for training and 10,000 images for testing. The Dig-MNIST dataset comprises of 10,240 images belonging to 10 classes over which the trained model is evaluated. The grayscale images of size 28 X 28 was the input to our model as displayed in Figure 1 and trained using the Adams optimizer having a learning rate(alpha) equal to 0.0001. The batch size of 64 was used for training. A dropout of 0.25 was added after every two convolutional layers and a dropout of 0.50 was added after the Fully Connected(FC) layer to avoid overfitting during training. The training is stopped after 50 epochs. In figure 5, the Accuracy Vs Epochs graph of the model during training is shown and in figure 6, Loss Vs Epochs graph of the model during training is shown.

We accomplished an accuracy of 98.24% on the Kannada-MNIST dataset and an accuracy of 86.85% on the Dig-MNIST showing an improvement to [7] and [5]. The classification report is used to measure the quality of predictions from a classification algorithm. It shows the main classification metrics such as F1-Score, Precision and Recall. The classification report for the Kannada-MNIST is shown in Table 1 and the classification report for the Dig-MNIST is shown in Table 2.

As we can see in the classification report(Dig-MNIST), the precision for class-5 (87.75%), class-6 (78.93%) and class-9 (83.47%) were higher when compared to [7] and [5]. The Normalized confusion matrix for the Dig-MNIST is shown in Figure 3 and for the Kannada-MNIST is shown in Figure 4. Table 3 shows the comparison of accuracy between the baseline accuracy, [5] and the proposed model when trained on the Kannada-MNIST dataset and evaluated on the Kannada-MNIST and table 4 shows the comparison of accuracy between the baseline accuracy, [5] and the proposed model when evaluated on the Dig-MNIST and we have achieved better results in both the cases.

TABLE I
CLASSIFICATION REPORT OF KANNADA-MNIST DATASET

Class	PRECISION	RECALL	F1-SCORE
0	0.984663	0.963	0.973711
1	0.967742	0.99	0.978744
2	0.99501	0.997	0.996004
3	0.985015	0.986	0.985507
4	0.968841	0.995	0.981746
5	0.997912	0.956	0.976507
6	0.982915	0.978	0.980451
7	0.986815	0.973	0.979859
8	0.975586	0.999	0.987154
9	0.981113	0.987	0.984048
Accuracy	0.9824	0.9824	0.9824
Macro Avg	0.982561	0.9824	0.982373
Weighted Avg	0.982561	0.9824	0.982373

TABLE II
CLASSIFICATION REPORT OF DIG-MNIST DATASET

CLASS	PRECISION	PRECISION	F1-SCORE
0	0.81758	0.8447266	0.8309318
1	0.968254	0.7744141	0.8605534
2	0.879201	0.9453125	0.9110588
3	0.965921	0.8857422	0.9240958
4	0.98188	0.8466797	0.9092816
5	0.877551	0.9658203	0.9195723
6	0.789366	0.7539063	0.7712288
7	0.802139	0.8789063	0.8387698
8	0.82247	0.9365234	0.8757991
9	0.834766	0.8535156	0.8440367
Accuracy	0.868555	0.8685547	0.8685547
Macro Avg	0.873913	0.8685547	0.8685328
Weighted Avg	0.873913	0.8685547	0.8685328

	0	1	2	3	4	5	6	7	8	9
0	0.0963	0.0032	0	0.0002	0	0	0	0	0.0003	0
1	0.0007	0.099	0	0.0002	0	0	0	0	0	0.0001
2	0.0003	0	0.0997	0	0	0	0	0	0	0
3	0.0003	0	0	0.0986	0	0.0001	0.0001	0.0009	0	0
4	0	0	0	0.0001	0.0995	0	0.0002	0.0001	0.0001	0
5	0	0	0.0002	0.0006	0.003	0.0956	0	0	0.0006	0
6	0	0	0	0	0	0	0.0978	0.0003	0.0001	0.0018
7	0.0002	0.0001	0.0003	0.0004	0.0001	0	0.0012	0.0973	0.0004	0
8	0	0	0	0	0	0.0001	0	0	0.0999	0
9	0	0	0	0	0.0001	0	0.0002	0	0.001	0.0987

Fig. 3. Normalized confusion-matrix : Kannada-MNIST

	0	1	2	3	4	5	6	7	8	9
0	0.084473	0.002051	0.002344	0	9.77E-05	0.000977	9.77E-05	0.001074	0.00146	0.00742
1	0.015625	0.077441	0.001758	0.00088	0	0.000488	0	0.002637	9.77E-05	0.00107
2	0.00166	0	0.094531	0.00068	9.77E-05	0.001074	0	0.001172	0.00049	0.00029
3	0.000293	0	0.001563	0.08857	0.000293	0.002734	0.000293	0.006152	9.77E-05	0
4	9.77E-05	0	0.002734	9.77E-05	0.084668	0.003711	0.001074	0.000488	0.00566	0.00146
5	0	0	9.77E-05	0.00029	0.000488	0.096582	0	0	0.00254	0
6	0.000391	0	0.001953	0.00088	0.000586	0.001074	0.075391	0.009766	0.00557	0.00439
7	0.000488	0.000195	0.001465	0.00029	0	0.000879	0.006543	0.087891	0.00186	0.00039
8	0.000293	9.77E-05	0.000586	0	0	0.002344	0.000879	0.000293	0.09365	0.00186
9	0	0.000195	0.000488	0	0	0.000195	0.01123	9.77E-05	0.00244	0.08535

Fig. 4. Normalized confusion-matrix : Dig-MNIST

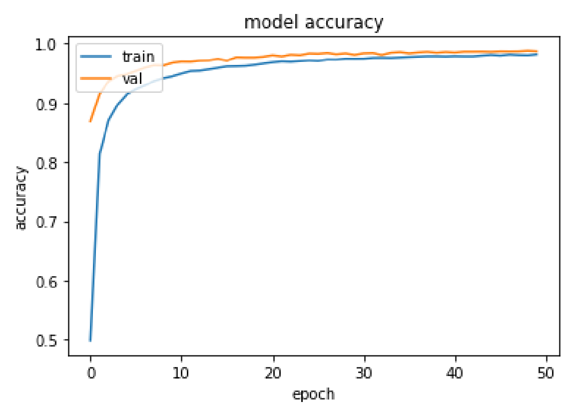

Fig. 5. Accuracy Vs Epochs

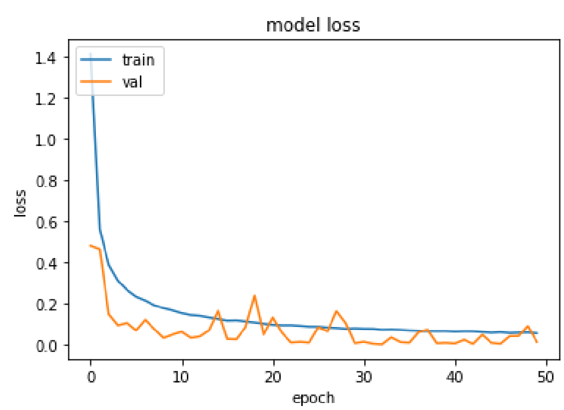

Fig. 6. Loss VS Epochs

TABLE III
COMPARISON TABLE :- KANNADA-MNIST

Author	Training Set	Testing Set	Accuracy
Prabhu *et al.,* [7] 2019	60,000	10,000	96.85%
Gati *et al.,* [5] 2019	60,000	10,000	97.53%
Proposed Method	60,000	10,000	98.24%

TABLE IV
COMPARISON TABLE :- DIG-MNIST

Author	Testing Set	Accuracy
Prabhu *et al.,* [7] 2019	10,240	76.17%
Gati *et al.,* [5] 2019	10,240	85.02%
Proposed Method	10,240	86.85%

VI. CONCLUSION

In this paper, Deep CNN was used to achieve a very high accuracy on the Kannada-MNIST dataset as well as the more difficult and challenging Dig-MNIST dataset. 98.24% on the Kannada MNIST dataset and 86.85% on the Dig-MNIST dataset is achieved without any additional preprocessing. This approach can be further used to achieve similar results on other languages which share the same characteristics of Kannada.

REFERENCES

[1] N. Sharma, U. Pal, and F. Kimura, "Recognition of handwritten kannada numerals," in *9th International Conference on Information Technology (ICIT'06)*. IEEE, 2006, pp. 133–136.

[2] M. Jangid, R. Dhir, R. Rani, and K. Singh, "Svm classifier for recognition of handwritten devanagari numeral," in *2011 International Conference on Image Information Processing*. IEEE, 2011, pp. 1–5.

[3] H. Mamatha, K. S. Murthy, A. Veeksha, P. S. Vokuda, and M. Lakshmi, "Recognition of handwritten kannada numerals using directional features and k-means," in *2011 International Conference on Computational Intelligence and Communication Networks*. IEEE, 2011, pp. 644–647.

[4] A. Majumdar and B. Chaudhuri, "A mlp classifier for both printed and handwritten bangla numeral recognition," in *Computer Vision, Graphics and Image Processing*. Springer, 2006, pp. 796–804.

[5] E. S. GATI, B. D. NIMO, and E. K. ASIAMAH, "Kannada-mnist classification using skip cnn," pp. 245–248, 2019.

[6] Y. Han, J. Kim, and K. Lee, "Deep convolutional neural networks for predominant instrument recognition in polyphonic music," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 25, no. 1, pp. 208–221, 2016.

[7] V. U. Prabhu, "Kannada-mnist: A new handwritten digits dataset for the kannada language," *arXiv preprint arXiv:1908.01242*, 2019.