Word Transcription of MODI Script to Devanagari Using Deep Neural Network

Vimal Patel[1], Brijesh Chauhan[2], Atharv Joshi[3], Ronak Kulhari[4], Madhura Shirodkar[5]

[1,2,3,4]*Students, Department of Electronics and Telecommunications*
[5]*Professor, Department of Electronics and Telecommunications*
Xavier Institute of Engineering

---***---

Abstract -Many historical documents and letters are written in MODI script. Study of "Shivakalin" and "Peshvekalin" era documents is almost impossible without the knowledge of MODI script. This work aims to bridge the gap between Devanagari and MODI Script by developing a system to map the recognized MODI characters to its Devanagari equivalent. Our dataset would comprise 57 different classes of MODI Script characters. We has solved the problem of loss of information in CNN as the model goes deeper using Dense Net-121. The various approaches for feature extraction usually used include moment invariant, affine moment invariant, chain code histogram, intersection junction and for character classification. Deep Neural Networks on the other hand do not require any feature to be explicitly defined, instead they work on the raw pixel data to generate the best features and use them to classify the inputs into different classes. Hence, we propose a deep learning architecture for character recognition. Dense Net-121 uses little pre-processing compared to other image classification algorithms. This means that the network learns the filters which in traditional algorithms were hand-engineered. The system aims to provide a good recognition rate by implementing DCNN.

***Key Words*:**Siamese Neural Network, CNN, Devanagari, HOCR

*1.*INTRODUCTION *(Size 11, Times New roman)*

Modi is an ancient script. Crores of Modi documents. Origin: 12th century and used uptil the 20th century. Modi lipi converter is a low resource solution Many historical documents and letters are written in MODI script. Study of "Shivakalin" and "Peshvekalin" era documents is almost impossible without the knowledge of MODI script. This work aims to bridge the gap between Devanagari and MODI Script by developing a system to map the recognized MODI characters to its Devanagari equivalent, provides authors with most of the formatting specifications needed for preparing electronic versions of their papers. All standard paper components have been specified for three reasons: (1) ease of use when formatting individual papers, (2) automatic compliance to electronic requirements that facilitate the concurrent or later production of electronic products, and (3) conformity of style throughout a conference proceedings. Margins, column widths, line spacing, and type styles are built-in; examples of the type styles are provided throughout this document and are identified in italic type, within parentheses, following the example. Some components, suchas multi-leveled equations, graphics, and tables are notprescribed, although the various table text styles are provided. The formatter will need to create these components, incorporating the applicable criteria that follow.

2. LITRATURE SURVEY

A brief review of the work done on character recognition of MODI script is presented as below.

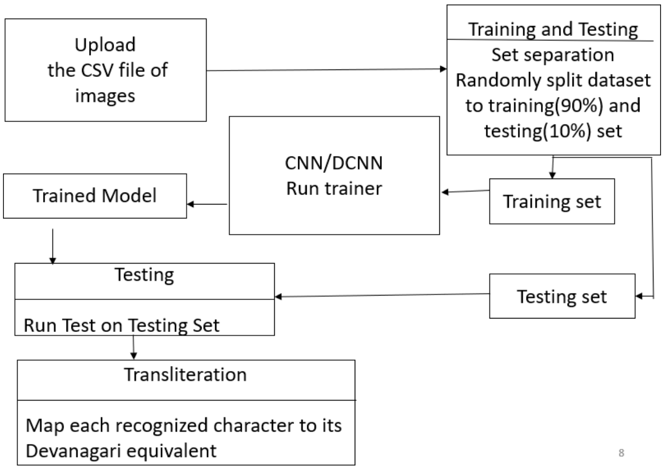
Kulkarni Sadanand et al. [1] analyzed different methodologies adopted for optical character recognition of MODI Characters. The paper mostly focuses on methods for segmentation and feature extraction along with their accuracies. It concludes that structural similarity approach was lacking for reliable and correct implementation of HOCR in MODI Script compared to the other methods used.

Savitri Chandure et al. [2] used the concept called chain code histogram and Intersection junction feature extraction methods. The neural and non-neural approaches; BPN (Back Propagation Neural Network), KNN and SVM were used to train, test and classify Devanagari and MODI vowels distinctly. They noted that the recognition rate or accuracy rate for MODI vowels is relatively less and this is due to the misclassification of similar vowels. The paper concludes that the system performance can be increased by taking larger training data.

Sanjay Gharde et al. [3] proposed a hybrid approach for extraction of feature by uniting Moment Invariant and Affine Moment Invariant for extraction of feature. The preprocessing performed on the data samples collected from various writers included linearization, thinning, noise removal, normalization and resizing. Overall near about 18 features corresponding to each numeral were used for classification. The system provided a good recognition rate of 89.72% by applying SVM as a classification method. The method has been applied on the dataset of handwritten. samples collected by ANESP means Automatic Numeral Extraction and Segmentation Program.

3. PROPOSED METHODOLOGY

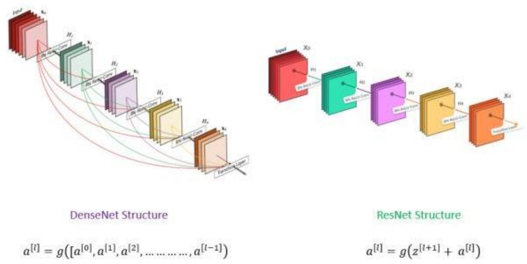
International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 04 | April-2022Impact Factor: 7.185 ISSN: 2582-3930



We purpose DenseNet-121 as our final model .
DenseNet was developed specifically to improve the declined accuracy caused by the vanishing gradient in high-level neural networks. In simpler terms, due to the longer path between the input layer and the output layer, the information vanishes Before reaching its destination.

Dense Net Architecture:

Dense Net falls in the category of classic networks.
This image shows a 5-layer dense block with a *growth rate* of k = 4 and the standard ResNet structure.



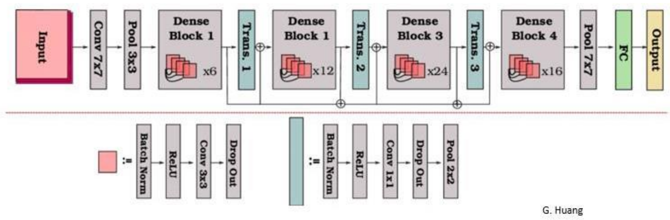
Sources: DenseNet Structure - G. Huang, Z. Liu and L. van der Maaten, "Densely Connected Convolutional Networks," 2018; Resnet Structure - [*Missinglink.ai*](Missinglink.ai)

An output of the previous layer acts as an input of the second layer by using *composite function operation*. This composite operation consists of the convolution layer, pooling layer, batch normalization, and non-linear activation layer.
These connections mean that the network has L(L+1)/2 direct connections. L is the number of layers in the architecture.
The DenseNet has different versions, like DenseNet-121, DenseNet-160, DenseNet-201, etc. The numbers denote the number of layers in the neural network. The number 121 is computed as follows:

DenseNet-121:5+(6+12+24+16)*2 = 121
5- Convolution and Pooling Layer
3- Transition layers (6,12,24)
1- Classification Layer (16)
2- Dense Block (1x1 and 3x3 conv)

Dense Blocks and Layers:

Be it adding or concatenating, the grouping of layers by the above equation is only possible if feature map dimensions are the same. What if dimensions are different? The DenseNet is divided into Dense Blocks where a number of filters are different, but dimensions within the block are the same. *Transition Layer* applies batch normalization using downsampling; it's an essential step in CNN.
Let's see what's inside the DenseBlock and transition layer.:



Source: G. Huang, Z. Liu and L. van der Maaten, "Densely Connected Convolutional Networks," 2018.

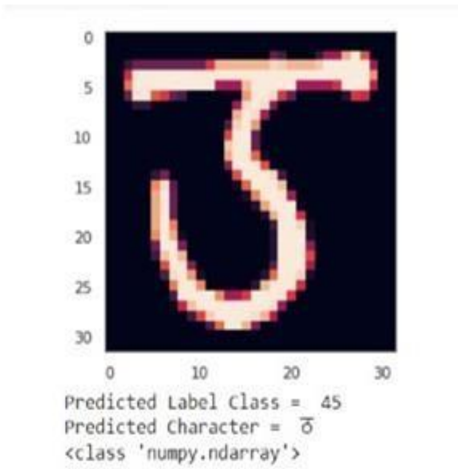
Data Augmentation

We have used Data Augmentation to make our data realistic by changing the Zoom range, Height, Width , Rotation.

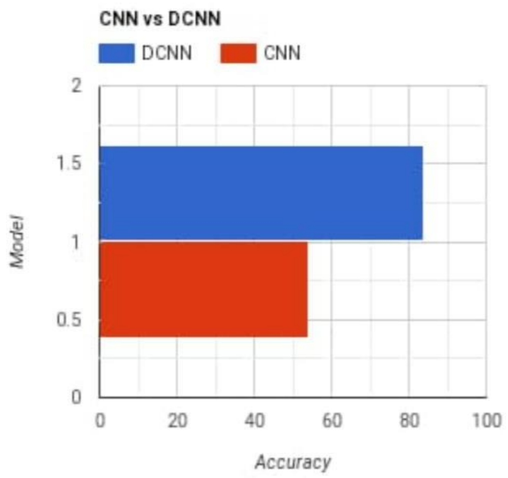
Data augmentation: a technique to increase the diversity of your training set by applying random (but realistic) transformations, such as image rotation etc,.

4. Results

This presents the results on number of handwritten words and characters to show the validity of the proposed algorithm. First, segmented results of words are presented and analyzed. Second, character recognition results are discussed and finally, word recognition results are validated and discussed.



At first we have used the CNN and we got an accuracy of 54% and with DenseNet121 we got an accuracy of 80%



5. CONCLUSIONS

In this system, character recognition of the MODI script was investigated. 'MODI' lipi is a cursive form of ‘Marathi’ which is the main language spoken in the state of Maharashtra in western India.
Thus we could transliterate Modi characters into its Devanagari characters but we could not transliterate Modi words to Devanagari words due to the unavailability of the script dataset.
At first we have used the CNN and we got an accuracy of 54% and with DenseNet121 we got an accuracy of 88% .

ACKNOWLEDGEMENT

This research was supported/partially supported by Xavier Institute of Engineering. We thank our friends and Professers from Xavier Institute of Engineering who provided insight and expertise that greatly assisted the research.
We thank Prof. Madhura Shirodker for assistance with the whole project.

REFERENCES

[1]. A. K. Sadanand, L. B. Prashant, R. M. Ramesh and L. Y. Pravin, ”Impact of zoning on Zernike moments for handwritten MODI character recognition,” International Conference on Computer, Communication and Control (IC4), Indore.
[2]. S. L. Chandure and V. Inamdar, ”Performance analysis of handwritten Devanagari and MODI Character Recognition system,” International Conference on Computing, Analytics and Security Trends (CAST).
[3]. S. S. Gharde and R. J. Ramteke, ”Recognition of characters in Indian MODI script,
” International Conference on Global Trends in Signal Processing”, Information Computing and Communication (ICGTSPICC), Jalgaon.
.[4]. Pankaj A. Patil, ”Character Recognition System for Modi Script”, International Journal of Computational Engineering Research (IJCER), Jalgaon.
[5]. Solley Joseph, Jossy George “Handwritten Character Recognition of MODI Script using Convolutional Neural Network Based Feature Extraction Method and Support Vector Ma-chine Classifier” 2020 IEEE 5th International Conference on Signal and Image Processing.
[6]. Asish Chakrapani, Sukalpa Chanda, Umapada Pal, and David Doermann “One-Shot Learning- Based Handwritten Word Recognition” (2020)
[7] https://www.pluralsight.com/guides/introduction-to-densenet-with-tensorflow