Design of Machine Learning Model for Character Recognition

Rituraj Jain[1], Swapnil Parikh[2], Yohannes Bekuma[3], Jaishri Gothania[4]

[1] *Department of Electrical and Computer Engineering, Wollega University, Nekemte, Ethiopia.*
[2] *Department of Computer Science & Engineering, Parul University, Gujarat, India.*
[3] *Department of Electrical and Computer Engineering, Wollega University, Nekemte, Ethiopia.*
[4] *Department of Computer Science & Engineering, Lingaya's Vidyapeeth, Faridabad, Haryana, India.*

Abstract: Hand written character recognition or text recognition is the ability of the system to identify character or text automatically. Documents can be in any form such as documents, photographs, touch-screens or other devices. Each character will have its own feature sets. The classification or identification of characters are done based on proper selection of feature sets. Feature selection is the major step for any classification process. The machine learning model is created based on feature sets. Logistic regression model is used in this model to identify different characters. 97.33% accuracy is achieved using logistic regression model compared to existing work.

Keywords: Accuracy, Character Recognition, Machine Learning, Logistic Regression

1. INTRODUCTION

Character recognition is very much important for different applications. It has numerous amounts of applications, such as read forms, bank cheque amounts, reading postal address and many more. Moreover, Optical character Recognition plays an important role for digital libraries, entering image textual information into computers by digitization, image restoration and recognition methods. Text detection and recognition can help visually impaired people for reading and can make their life easier. It can automatically read cheque signature. Text recognition can also be used for automatic document scanning process. Text detection and recognition can also be used to read video captions from web pages. Still, lots of works are going on to solve critical problems like segmentation and extraction of text from natural scenes.

The paper is organized as follows. Section II presents literature review for text detection and recognition. Section III presents methodologies adopted in this work. Section IV describes Results and Discussions. Section V deals with Conclusions. Section VII highlights about future scopes.

II LITERATURE REVIEW

Text detection and recognition have vital applications. Text Detection deals with finding exact text area from given image, where, recognition deals with converting obtained text into characters and words [1].
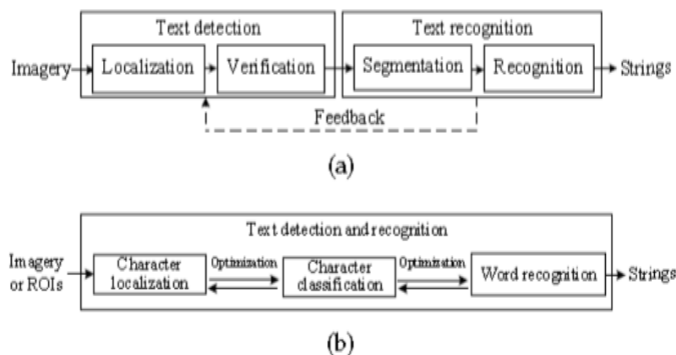

Fig. 1. (a) Stepwise Method (b) Integrated Method

Convolution Neural Network (CNN) based classifier is used for text recognition by Elagouni et. al. [2]. The stepwise methods to detect texts of arbitrary orientation from natural images are used in literature [3]. Wang and Belongie used integrated methods where segmentation stage can be avoided and replaced by word recognition and matching stage [4]. Histogram of Oriented Gradients (HOG) features with nearest neighbour classifier is used for character recognition. It can obtain nearest word from set of characters [5].
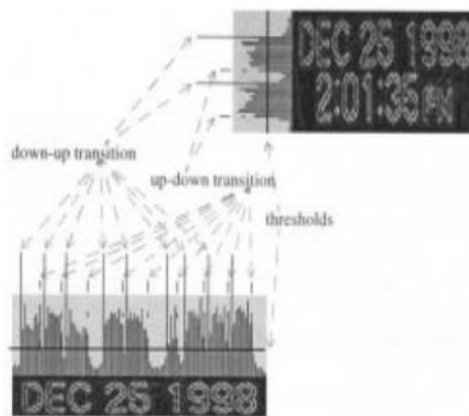

Fig. 2. Projection Profile

Connected component analysis method forms graph of connected points based on colour or edge features from binary images [6]. [7] extracts text in the form of connected components. [8] used efficient connected component extraction methods instead of using filtering approach. Morphological feature extraction and region of interest is

376

identified in reference [9]. Unsupervised Learning approach with 8x8 Gray scale patches of image for feature learning is used in reference [10]. Supervised Learning approach is used in references [11] and [12] respectively.

Fig. 3 Binarization of Text

Constraints on edge area as well as area, height and width are detected in [13] for classification. Wavelet transform is used in reference [14]. Different features like variance of stroke of width, difference between contrast of text and background as well as aspect ratio are classified used k-means based classifier [15].

Fig. 4 Character Segmentation of Degraded Character

Kim et. al. used adapting thresholding for binarization [16]. Mathematical Morphology operator based adaptive thresholding is used in reference [17]. In [18] Optical Character Recognition Module is used. Support Vector Machine (SVM) is used for multiclass classification of character [19]. Graphical approach of character recognition is used by Wachenfeld et. al. [20].

Table 1 Comparison of methods used for Text Detection and Recognition

Method	Features of Method
Stepwise Methods	Separate Detection and Recognition Module Less Computation cost Apt for large number of word detection Divided into localization, classification, segmentation and recognition stages
Integrated Methods	Detection and recognition modules are not separated. Can avoid segmentation Small lexicon Increase in lexicon size makes recognition difficult

Table 2 Comparison of approaching for processing stages of Text Detection and Recognition

Processing Stage	Approach for Pre-processing	Features
Text Detection and Localization	Connected Component Analysis	Graph based Methods High Speed Not efficient for noisy images
	Region Based Methods	Window based approach Less Speed
Classification	Supervised Approach	It has training phase Features are colours, size, projection profile
	Unsupervised Approach	It does not have training phase Wavelet, stroke width, contrasts are some features
Segmentation	Binarization	Simple algorithm Converts Gray scale image to binary image

	Character Segmentation	Converts text into multiple sets of single characters Complex algorithm
Text Recognition	Character Recognition	Independent of lexicon Divide text into cut-outs of single characters
	Word Recognition	Identifies word from text image Recognized small no. of text form lexicons

In the above tables different methods and processing stages are explained.

III METHODOLOGIES

In the above literature review different methods of classification, segmentation and text recognition methods are discussed. In existing model stepwise methods, integrated methods are used for text detection and recognition. Processing stages are classified as text detection and localization, classification, segmentation and text recognition. Both supervised and unsupervised models are used for classification. Segmentation methods used binarization and character segmentation. Text recognition specifically dealt with character recognition and word recognition. All these methods are proven less efficient compared to proposed model. Python libraries are used to implement the linear regression model.

In this proposed work logistic regression model is used for character recognition, which is not adopted in the above literature discussion. MNIST dataset is used for this work. Since the dataset is pre-processed, each character in the dataset has consistent height and width. Each digit is reshaped uniquely before using inside the machine learning model. Training data set for both X axis and testing dataset for both X domain are divided into 75 to 25 ratios. Logistic Regression model is imported using Sklearn library.

In this model Scaling intercept is 1. Maximum iterations used is 10000. Multiclass classification is set as auto. Penalty value is set as 12. LBFGS solver is used for this model. Limited Memory BFGS algorithm is for optimization. It is very popular parameter optimization technique for machine learning.

This model with little modifications is applicable for other standard or customised database. Logistic regression is a statistical model which uses logistic function to model a binary dependent variable.
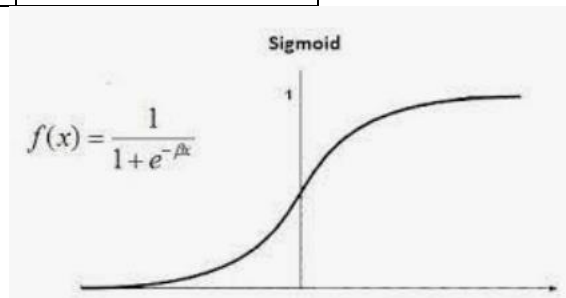

$$f(x) = \frac{1}{1 + e^{-\beta x}}$$

Fig. 5. Logistic Regression

Logistic regression model is very easy to implement, interpret and efficient to create model.

This model cannot be best fitted if no. of observations is less than number features. It will create overfitting problem.

IV RESULTS AND DISCUSSIONS

Since MNIST data set is used for this work, the general shape of data was (1797, 64). Individual data reshaped with (8,8) dimension. In this work logistic regression-based model is used to recognize character using MNIST dataset. 75% data is used for training and 25% data is used for testing in this dataset. Sample of Training dataset is shown in the figure below.
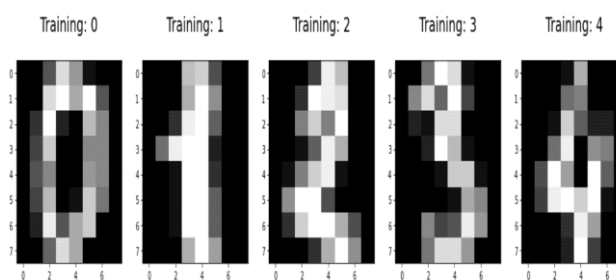

Fig. 6. Training Data Set for MNIST
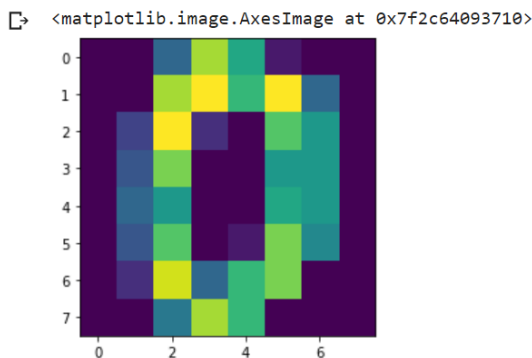
RGB value of the digit is shown in figure 7.

Fig. 7. RGB Dataset

The shape of the x_train data is (1347,64). The shape of y_train data is (1347,). The shape of x_test data is (450,64). The shape of y_test data is (450,).

In fig. 8. Logistic Regression model parameters are mentioned. 1000 iterations have been done for best results. In this model Scaling intercept is 1. Maximum iterations used is 10000. Multiclass classification is set as auto. Penalty value is set as 12. LBFGS solver is used for this model. Limited Memory BFGS algorithm is for optimization. It is very popular parameter optimization technique for machine learning.

```
LogisticRegression(C=1.0, class_weight=None, dual=False, fit_intercept=True,
                   intercept_scaling=1, l1_ratio=None, max_iter=10000,
                   multi_class='auto', n_jobs=None, penalty='l2',
                   random_state=None, solver='lbfgs', tol=0.0001, verbose=0,
                   warm_start=False)
```

Fig. 8. Logistic Regression Model Parameters

The model identified the character with 97.33% accuracy. It is much better than the models discussed in literature.
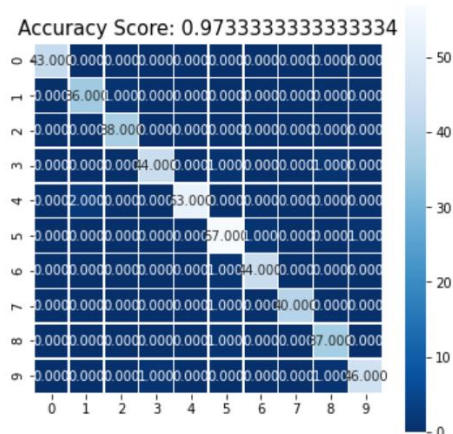

Fig. 9. Accuracy Score through Confusion Matrix

This is a better model compared to another existing model.

V CONCLUSIONS

In this work Logistic Regression based Machine Learning Model is used to get more better result for character recognition. This Machine learning feature extraction and model produce 97.33% accuracy in the result, which is clearly visible using confusion matrix.

VI FUTURE SCOPES

This work can be extended with advanced machine learning algorithms with ensemble learning method. The logistic regression model is tested with existing pre-processed dataset. The same model can be used for customised dataset with minute modifications.

REFERENCES

[1] Qixiang Ye, David Doermann, "Text Detection and Recognition in Imagery: A Survey" IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014

[2] K. Elagouni, C. Garcia and P. Sbillot," A Comprehensive Neural-Based Approach for Text Recognition in Videos using Natural Language Processing," in Proc. ACM Conf. Multimedia Retrieval, 2011.

[3] C. Yao, X. Zhang, X. Bai, W. Liu, Y. Ma, and Z. Tu," Detecting Texts of Arbitrary Orientations in Natural Images," in Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition, pp.1083-1090, 2012.

[4] K. Wang, S. Belongie," Word Spotting in the Wild", in Proc. European Conference on Computer Vision, pp. 591- 604, 2010.

[5] L. Neumann, J. Matas," On Combining Multiple Segmentations in Scene Text Recognition," in Proc. IEEE Int'l Conf. Document Analysis and Recognition, pp. 523- 527, 2013.

[6] A.K. Jain and B. Yu," Automatic Text Location in Images and Video Frames," Pattern Recognition, vol. 31, no. 12, pp. 2055-2076, 1998.

[7] S.M. Hanif, L. Prevost, P.A. Negri," A Cascade Detector for Text Detection in Natural Scene Images," in Proc. IEEE Int'l Conf. Pattern Recognition, pp. 1-4, 2008.

[8] H. Koo, D.H. Kim," Scene Text Detection via Connected Component Clustering and Non-text Filtering," IEEE Trans. Image Processing, vol. 22, no. 6, pp. 2296-2305, 2013.

[9] J. Lee, P. Lee, S. Lee, A. Yuille and C. Koch," AdaBoost for Text Detection in Natural Scene," in Proc. IEEE Int'l Conf. Document Analysis and Recognition, pp. 429-434, 2011.

[10] A. Coates, B. Carpenter, C. Case, S. Satheesh, B. Suresh, T. Wang, D. J. Wu, Andrew Y. Ng," Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning," in Proc. IEEE Int'l Conf. Document Analysis and Recognition, pp. 440- 445, 2011.

[11] R. Lienhart and A. Wernicke," Localizing and Segmenting Text in Images and Videos," IEEE Trans. Circuits System on Video Technology, vol. 12, no. 4, pp. 256-268, 2002.

[12] M. Li, C. Wang," An Adaptive Text Detection Approach in Images and Video Frames," in Proc. Int'l Joint Conf.

Neural Network, pp. 72-77, 2008.

[13] H. Li, D. Doermann, and O. Kia," Automatic Text Detection and Tracking in Digital Video," IEEE Trans. Image Processing, vol. 9, pp. 147-156, 2000.

[14] A. Mosleh, N. Bouguila, A. Ben Hamza," Image Text Detection Using a Bandlet-Based Edge Detector and Stroke Width Transform," in Proc. British Machine Vision Conference, pp. 1-2, 2012.

[15] W. Kim and C. Kim," A New Approach for Overlay Text Detection and Extraction from Complex Video Scene," IEEE Trans. Image Processing, vol. 18, no. 2, pp. 401-411, 2009.

[16] J.J. Weinman, Z. Butler, D. Knoll, J. Feild," Toward Integrated Scene Text Reading," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 3, no. 2, pp. 375- 387, 2014.

[17] T. Phan, P. Shivakumara, B. Su and C.L. Tan," A Gradient Vector Flow-Based Method for Video Character Segmentation," in Proc. IEEE Int'l Conf. Document Analysis and Recognition, pp. 1024-1028, 2011.

[18] S. Nomura, K. Yamanak, O. Katai, H. Kawakami, T. Shiose," A Novel Adaptive Morphological Approach for Degraded Character Image Segmentation," Pattern Recognition, vol. 38, no. 11, pp. 1961-1975, 2005.

[19] D. Chen, J.M. Odobez, H. Bourlard,"Text Detection and Recognition in Images and Video Frames," Pattern Recognition, vol. 37, no. 3, pp. 596-608, 2004.

[20] K. Sheshadri, S.K. Divvala," Exemplar Driven Character Recognition in the Wild," in Proc. British Machine Vision Conference, pp. 1-10, 2012.