Dizitalization of Handwritten Using Signature Features And KNN Algorithm

Ms. S. Priya[1] , Mrs. N. Surya[2]
[1]Dept of Computer Science and Engineering
[2]Assistant Professor, Dept of Computer Science and Engineering
[1, 2] Panimalar Institute of Technology, Tamil Nadu, India.

Abstract- *There is a growing need for recognition of digits manuscripts for use in various situations. As we move towards digital life, it is very much needed to automatically detect handwritten alphabets and digits. This will be very much helpful in areas like converting manual scripts to digital. There were various algorithms like SVM, Optimum Path Forest applied in OCR technology. Where as here we apply enhanced Knn algorithm to improve the efficiency of Recognition of Handwritten Digits. And we develop android application to automatically convert handwritten characters to digital.*

Keywords- Machine learning techniques, Pattern Recognition, Digital Image Processing, Computer Vision, OCR Applications, Optimum-Path Forest, OPF.

I. INTRODUCTION

Recognition of Handwritten Digits converting manual scripts to digital. It automatically detect handwritten alphabets and digits. This will be very much helpful in areas like converting manual scripts to digital. Many tasks with which computers fail humans considerably they do better. Many of these tasks, in which computers are related to nature interpretive and multiprocessing of the brain. A simple way to characterize well the difference between the computer and man would compare the computer, which is a serial machine, with our brain. The main characteristic is the learn things. Optical Character Recognition – OCR allows a machine to automatically recognize a character through a optical mechanism.

The OPF (Optimum- Path Forest) standard classifier was created with the aim of allying efficiency in the training, effectively in the data classification stage, is a Framework and also considered a classifier of simple mathematical approach based on concepts of Graph Theory has been widely used in pattern recognition applications Neural classifiers tend to perform significantly better than other types of classifiers on this basis. Convolutional algorithms lead the records of better grading rates. The most successful job uses an association of convolutional networks together with the increase in the training base using distortions elastic, having obtained an error rate of 0.27%. This paper proposes a new approach to extraction and classification of manuscripts based on the extraction and classification of manuscripts based on the extraction of characteristics using the signature of the characters and proposes classification of handwritten digits using the Optimum- Path Forest(OPF) learning error rate equivalent to or less than current classifiers.

II. SYSTEM ANALYSIS

EXISTING SYSTEM

Existing OCR technology use SVM and Optimum Path Forest algorithms to convert handwritten digits to Digital automatically.

Disadvantages

- High Computational cost
- Less Accuracy

PROPOSED SYSTEM

We propose enhanced Knn algorithm to convert handwritten digits / alphabets to digital.

Advantages

- Less Computational cost
- High Accuracy

III. FEASIBILITY STUDY

During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. There are some understanding of the major requirements for the system is essential. Two key considerations involved in the feasibility analysis are,

- ECONOMICAL FEASIBILITY

- TECHNICAL FEASIBILITY

Economic feasibility:

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available.

Technical feasibility:

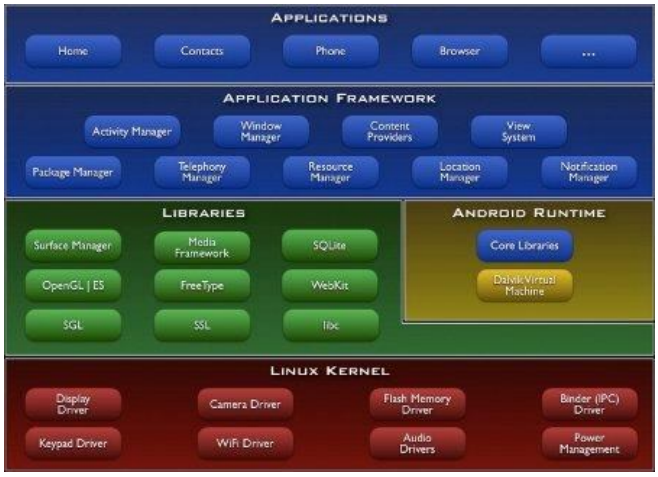
This study is carried out to check the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a model requirement, as only minimal or null changes are required for implementing the system.

IV. LANGUAGE SPECIFICATION

Android:

Android is a Linux-based operating system and is developed by the Open Handset Alliance led by Google. It has a large community of developers that extend the functionality of the devices. Google's Andy Rubin, says that as of December 2011 there are over 700,000 Android devices activated every day. Primarily it is written in a customized version of Java by developers

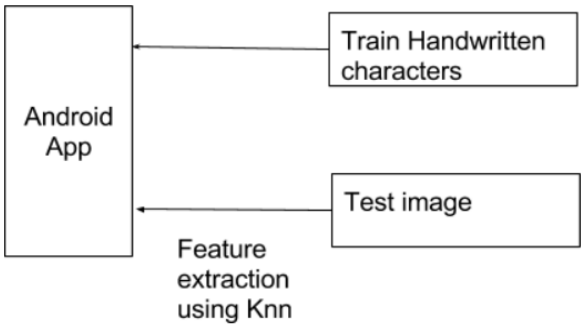
Design



Features

The Android Emulator default home screen
Current features and specifications.

V. SYSTEM DESIGN



VI. MODULES

1. Training dataset
2. Capture image
3. Apply Knn algorithm
4. OCR for text and numbers

1.1 Training dataset:

We train the datasets to convert handwritten to digital.

2.1 Capture image:

Capture the handwritten alphabets and convert digital.

3.1 Apply Knn algorithm

Apply enhanced Knn algorithm to improve the efficiency of Recognition of Handwritten Digits.

4.1 OCR for text and numbers

Optical Character Recognition is the recognition of printed or written text characters by a computer.

VII. RESULT

This experiment is conducted by taking various samples of training data for each feature. The following image shows the recognition of handwritten digits.

VIII. CONCLUSION

Therefore after implementation of KNN using various features, it is found that template matching performs the best with an accuracy of 97% The combination of more than one feature is beyond the scope of the experiment. It can be considered as an extension of this work.

REFERENCES

[1] D. Keysers, J. Dahmen, T. Theiner, and H. Ney. Experiments with an Extended Tangent Distance. In Proc. 15th Int. Conf. on Pattern Recognition, volume 2, pages 38–42, Barcelona, Spain, September 2000.

[2] L. Bottou, C. Cortes, J. S. Denker, H. Drucker, I. Guyon, L. Jackel, Y. Le Cun, U. Mu¨ller, E. Sa¨ckinger, P. Simard, and V. N. Vapnik. Comparison of Classifier Methods: A Case Study in Handwritten Digit Recognition. In Proc. of the Int. Conf. on Pattern Recognition, pages 77–82, Jerusalem, Israel, October 1994.

[3] G. Mayraz and G. Hinton. Recognizing Handwritten Digits Using Hierarchical Products of Experts. IEEE Trans. Pattern Analysis and Machine Intelligence, 24(2):189–197, February 2002.

[4] B.Schölkopf.Support Vector Learning.Oldenbourg Verlag, Munich, 1997.

[5] L.-N. Teow and K.-F. Loe. Handwritten Digit Recognition with a Novel Vision Model that Extracts Linearly Separable Features. In Proc. CVPR 2000, Conf. On Computer Vision and Pattern Recognition, volume 2, pages 76–81, Hilton Head, SC, June 2000.

[6] L.-N. Teow and K.-F. Loe. *Robust Vision-Based features and Classification Schemes for Off-Line Handwritten Digit Recognition.* Pattern Recognition, 35(11):2355–2364, November 2002.

[7] P. Simard, Y. LeCun, J. Denker, and B. Victorri, "An Efficient Algorithm for Learning Invariances in Adaptive Classifiers," Proc. Int'l Conf. Pattern Recognition (IAPR '92), 1992.

[8] D. Keysers, C. Gollan, and H. Ney. *Local Context in Non-linear Deformation Models for Handwritten Character Recognition*. In ICPR 2004, 17th Int. Conf. on Pattern Recognition, volume IV, pages 511– 514, Cambridge, UK, August 2004.

[9] C.-L. Liu, K. Nakashima, H. Sako, and H. Fujisawa. *Handwritten Digit Recognition: Benchmarking of State-of-the-Art Techniques*. Pattern Recognition, 36(10):2271–2285, October 2003.

[10] P. Simard. *Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis.* In 7th Int. Conf. Document Analysis and Recognition, pages 958–962, Edinburgh, Scotland, August 2003.