A Survey on Optical Handwriting Recognition System using Machine Learning Algorithms

Omkar Kumbhar
BE (Electronics and Telecommunication)
Modern Education Society's
College of Engineering
Pune-411001, India

Ajinkya Kunjir
BE (Computer Science)
Modern Education Society's
College of Engineering
Pune-411001, India

ABSTRACT
The recent advancement and development in technical fields and logistics has tackled the challenges of science and have delivered a good impact ratio by avoiding issues faced by fully or partially challenged and impaired people. The purpose of this paper is to provide easy facilities and improved services to a wide range of customers from school going children to teachers, professors, home tutors, visually impaired, and literary scholars. The title says it all, the concepts and theoretical background in the paper explains how the handwriting scripted in the image can be converted into textual information using several machine learning algorithms which follow the principles of supervised and unsupervised learning respectively. With the help of microcontrollers, feature extraction material, trainers and image processing concepts the proposed idea can be implemented feasibly. Machine learning algorithms such as Support Vector Machine, Random forest and neural nets algorithms have been compared and evaluated on basis of precision, recall, accuracy and other performance parameters. On comparing the mentioning algorithms it has been observed that 'SVM' algorithm outclasses the comparison test and gives out the best result in terms of accuracy, latency and robustness. On observing the overall effects of algorithms and conversions between handwritten text into textual information (maybe pdf's). It has been concluded that the work can be further implemented on platforms such as e-learning, in which the application can be deployed on internet for android and IoS users with the help of mobile computing principles. Students in the classroom can just click a picture of what's written on the board by the professor and the application will convert the text written on the image into a PDF and save it in the internal memory of your gadget.

General Terms
Image Processing and Machine Learning Algorithms

Keywords
Image Processing, Image Recognition, Machine Learning, Support Vector Machine, Supervised, Unsupervised, Clustering.

1. INTRODUCTION
People nowadays quite frequently write small notes of a piece of paper or stick the 'sticky notes' on their walls to remind them to do some task on a specific day. These notes range from a small meeting schedule, meet a friend, get milk from the shop and other daily routine movements. Some people in this digital age might wonder why is paper been used more than digital documents. The answer is simple: an effective replacement for paper has not been invented yet. Paper is still the winner when it comes to affordability, re-usability and reliability. Just like paper, even digital documents have some advantages such as fast searching (also called as Indexing), less storage in any smart gadget and many more other functionalities. Being able to search a huge database of written notes saves time and money, two important things in the business world. To achieve these advantages people need an intermediate to convert the handwritten text into digital text. Hence, we propose an efficient optical handwriting recognition system which uses machine learning algorithms to scan the handwriting and convert the text into some document and store it in the internal database. The use of machine learning algorithms is well suited for this task because they allow for variation and impreciseness in the data. The same character may be written in various different ways, and there is no well-defined mapping between a bitmap image of a letter and the character it represents. The ability of supervised machine learning algorithms is explained in detail in this paper. Karan Magiya et.al in their paper "Multipurpose real time handwriting recognition" stated the use and ability of neural networks back propagation algorithm used for conversion if handwritten text into digital text using feature extraction technique [1].

1.1 Existing Technology
Nowadays many smart phones following the trend of 3G and 4G are equipped with ability to recognize the letters and alphabets written on the screen. The technique used for racing the movements of the stylus on the screen of mobile phones is called as 'Graph transformation network'. These continuous movements are then marked and graphed to make a symbol. But the major drawback of this technique is that there is no way to recognize the text written or printed on the digital images. The secondary drawback are poor accuracy of recognition and less reliability of the feature extraction. A much more efficient system which avoids all these drawbacks and performs with higher accuracy is stated in this paper.

2. THEORETICAL BACKROUND
2.1 Related Work in Image Processing and Handwriting Recognition
K. Magiya et.al proposed a project which used image processing tools for getting data to be sorted for a neural network. They used kirsch detectors for feature extraction which made sure neural networks could work faster on less data. The neural network used was focused at reducing MSE(mean squared error) by using the Back propagation algorithm. They used a high definition camera to take pictures which was fed to the image processor after some initial filters. An FPGA was used for feature extraction in because of its simplicity and real time application. This data was then given to the neural network trainer which implemented BPN and decision to classify them took place here. Error correction was considered for grammatical errors and spelling mistakes. The data generated from the system was stored in ASCII format depending upon the three proposed applications – "Voice Stick", extracting details of check and digital classroom. D.

Ciresan et.al in their article proposed a deep neural network (DNN) instead of using computer vision and traditional machine learning algorithms to get close to human recognition tasks of traffic sign recognition and handwritten digits [2]. Only winner neurons are trained in this with the help of a GPU (graphics processing unit) which makes the processing faster. They used this DNN on MNIST data base to get near human accuracy on digits database whereas on Traffic signs they did better than humans in accuracy. Taking inspiration from deep column neural networks in the cerebral cortex, they implemented a multi column DNN or MCDNN which got them lower error rate than any other algorithm. They also implemented this DNN on various other data bases and outperformed most of them. M. Elleuch et.al in their paper "A New Design Based-SVM of the CNN Classifier Architecture and Dropout of Offline Arabic Handwriting" gave out the result when the mixed two techniques of classification for recognizing offline Arabic Handwriting [3]. A CNN(convolutional Neural Network) was used for extracting features and used an SVM classifier for recognition. Their model worked better than conventional standard CNN classification model on HACDB and IFN/ENIT databases. By mixing capabilities of CNN and SVM they proposed an effective system for recognition of Arabic handwriting. The images are passed to a convolutional neural network which extracts features till the fifth layer and gives it to SVM consisting of RBF neurons for classification. A technique of randomly removing a unit from FCL layer to prevent over fitting was applied, which was known as dropout. This hybrid of CNN and SVM succeeded in reducing error classification rate on HACDB data base consisting of 66 classes. A Hybrid of SVM classifier with a convolutional neural Network was implemented and tested on MNIST digits database. This theory was first published in 2011 in the paper "A novel hybrid CNN–SVM classifier for recognizing handwritten digits", by the authors Xiao-Xiao Niu and Ching Y. Suen [4].The focus was not only on achieving high accuracy but also to find high reliability. CNN was used for automatic optimization of features and in accordance with SVM lead to high reliability. Achievement of High reliability was by using a rejection rule. M. Zimmermann et.al in their IEEE transaction paper on pattern analysis and machine intelligence described the working and implementation of an offline system using grammar based recognition of handwritten sentences [5]. R. Plamondon et.al in their IEEE article of pattern analysis and machine intelligence contributed to M. Zimmermann's offline grammar based handwriting recognition system and wrote a comprehensive survey on online and offline handwriting recognition systems which also used optical concepts and variations of supervised learning algorithms [6].

3. PROPOSED SYSTEM

We propose a system which uses a machine learning classification algorithm for Handwriting Recognition used for different applications in various sectors. The proposed theory aims at image acquisition via a high resolution camera and applying image processing techniques for feature extractions using either of the options – FPGA based image processors or OpenCV or Matlab on single board computers. The Machine learning algorithms selected for comparison and classification are SVM, Back propagation and Random Forrest. Best out of the three classification techniques selected will be used for implementing the Handwriting recognition and storing them in text format.

4. METHODOLOGY

The methodology of the proposed system illustrates the scanning of a handwritten text by using image acquisition techniques. The handwritten text can consist of alphabets, numbers, special characters and different symbols as well. The image acquisition can be done using high definition cameras with superior resolution quality. After the image acquisition, the next step is image processing which is can be done by many different approaches by using OpenCV or by using MATLAB over SBCs(single board computers). After this phase of processing the processed image is given to the trainer. The machine learning algorithms are called as 'Image Trainer's', which also receives the processed image as input and then the image is toned and groomed using error detection and correction techniques using microcontrollers. The derived text is then converted into ASCII format which can further be converted into other formats as and when needed. The figure given below depicts all the mentioned steps in details.

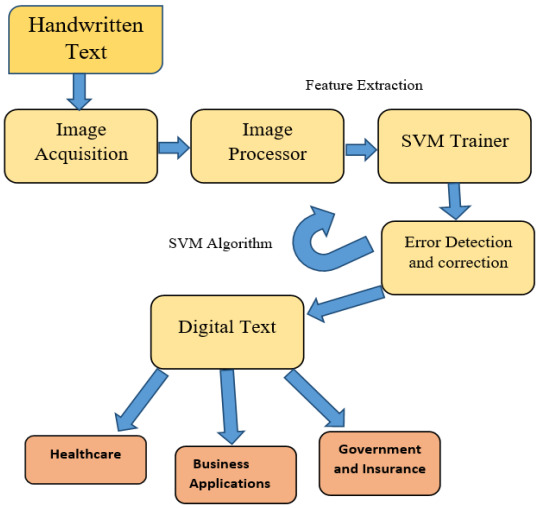

Fig 1: Working of the proposed system

4.1 Applications of Digital Text

- **Healthcare and Hospitals**: Hospitals especially in India are flooded with patients with different ailments, doctors running around, and utter chaos 24/7. There is lots of paper work involved from patient history, prescriptions, legal documents to name a few. A lot of time goes in documenting them. But what this system of health care lacks, especially the poor ones is a lack of centralized data base which can be used by other hospitals in future. So, if a patient has to give his medical history, it can be documented in one hospital, scanned and uploaded on a centralized data base and used by other hospitals for reference if the need arises. This saves time and effort involved in retaking medical history if the patient can't give specific details, or if the patient misplaces the medical file. The clerks trusted of managing the documentation can save a lot of type "scanning" the documents instead of "Typing" them.
- **Small Scale Business**: In developing countries, most of the record keeping of goods and commodities of shopkeepers, traders and business men happen in record books. A better way of storing this documentation is by using the proposed system along with modern tools like excel sheets.

This helps in saving valuable time spent and a safer documentation which can be stored in a few megabytes instead of fat record books. This will help in contributing to a paperless business practice.

- **Government, Education, Banking and Insurance**: All the mentioned sectors are riddled with paperwork which can make or break the smooth functioning of these sectors. Cheque details extraction is one of the applications of OCR. Taking relevant information from various forms related to insurance would help save valuable time, also the added advantage of information retrieval from stored documents becomes much easier than searching from tons of paper work. Education sector uses a lot of paper work related to admissions data, transcripts and forms, therefore an OCR system will make many of these processes easier and cost effective to these sectors.

5. COMPARISON OF MACHINE LEARNING ALGORITHMS

There are several algorithms or networks proposed by various researchers for speech and handwriting recognition. The selection of these algorithms depends on various factors such as accuracy, redundancy, Training time, rejection ratio and error identification/rectification. The algorithms can be supervised, unsupervised and reinforced. The algorithms selected for comparison are SVM, Neural Networks and Random Forest algorithm.

5.1 SVM (Support Vector Machine)

Support vector machine, also called as SVM is a supervised learning algorithm which is used for classification and regression purposes in fields of data science and machine learning. SVM uses a method of data scattering. A SVM model is a representation of data examples in a hyperplane. Here the data examples are represented as points and hyperplane is a two dimensional plane as shown in the figure given below. SVM performs extremely well on small and limited datasets as compared to huge data sets. The method used for classifying the unknown data examples is called as the 'Gap Trick', in which the points are classified by referring the distance between the classes of the data points.

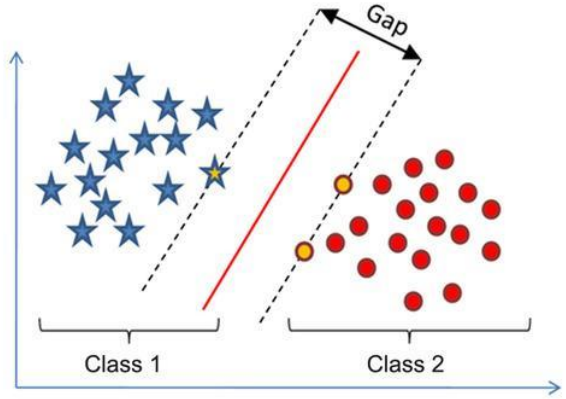

Fig 2: Classification by SVM

5.2 Random Forest

One of the best ensemble learning method is 'Random forest' or 'Random trees'. Random forest algorithm is a variation of decision tree which is used for classification, regression and other mining techniques. The working of random forest algorithm with steps for classification and prediction are described in the flowchart given below. At first, for each tree the training data is selected from the dataset chosen for implementation.

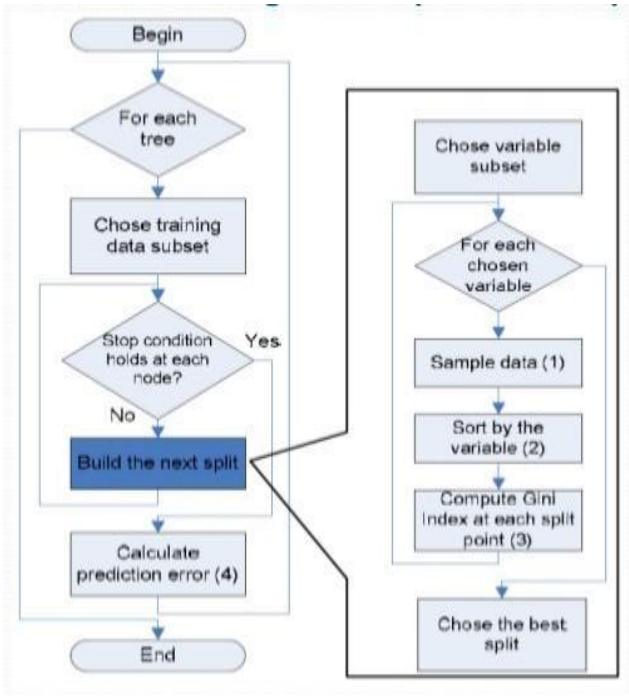

Fig 3: Steps of Random Forest Algorithm

Then next, the condition is checked for the split by choosing variables and sorting them further by computing the Gini indexes of each split point. The best split is selected and the prediction error is computed then after.

5.3 Back Propagation Method

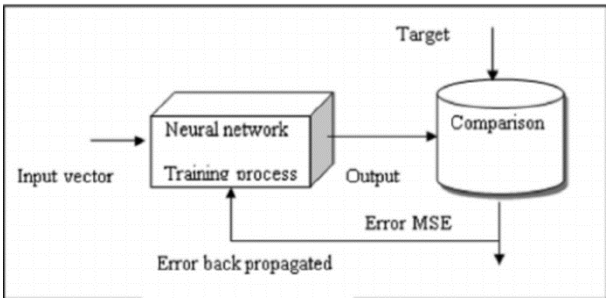

Fig 4: Block Process of Back Propagation method

Rich Caruana et.al in their comprehensive survey derived the results of comparison of supervised learning algorithms [7]. The algorithms selected for comparison and performance evaluation were SVM, Logistic regression, Naïve Bayes, random forest, decision trees, bagged trees and boosted trees. The algorithms were evaluated on factors such as accuracy, efficiency, precision, F-Score, Recall, ROC curves, Kernel tests, Lift, average precision, break even pint(BEP), squared error and cross entropy. The algorithms were evaluated on a dataset with 11 attributes from UCI repository. The results obtained and the conclusions derived from the comparison and performance evaluation proved that algorithms such as neural nets (back p trees, boosted trees, propagation) performed well on datasets without calibration. As on the

other side, algorithms such as SVM, Bagged trees, Logistic regression performed excellent with calibration. Even the best models fail at some time while the algorithms which are not so efficient work more accurately when calibrated using several methods. SVM was selected as the best algorithm to work with calibration and performs accurately with efficient results when used with image recognition and processing techniques.

6. WHEN ML MEETS IMAGE PROCESSING

The three dimensional and four dimensional images, also called as 3D and 4D images play an important role in fields of computer technology such as medical diagnosis, agriculture, military, cryptography, segmentation and other encoding-decoding applications. The recent research related to image processing an restoration led to the evolution of 'Computational Photography'. Computational photography (CP) aims at enhancing current imaging techniques and also provides new techniques to capture and store our surroundings. CP has provided a huge scope of improvisation in fields of scientific imaging, biomedical imaging, astrological imaging and have proven to be impactful in other areas of image processing. Machine learning algorithms, which are categorized by supervised, unsupervised and reinforcement learning algorithms help advancing the concepts of image processing. Daniel Rueckert et.al in their paper "Machine Learning meets medical imaging", mentioned the use of machine learning algorithms in imaging technique such as segmentation which was used for quantification of temporal changes such as atropy in Alzheimer's disease [8]. 'STATLOG', a concept which was studied and engineered in early 90's derived a theory of supervised learning algorithms which were compared on basis of efficiency and other algorithmic parameters [7]. Supervised learning algorithms such as SVM, bagged trees, boosted trees, Naïve Bayes algorithms were compared and evaluated for implementing in image recognition and processing systems. Machine learning algorithms prove to be a valuable asset and a catalyst in isolating the letters or text hidden in the images and helps in conversion of image media into textual formats such as documents, text files, and PDF's.

7. CONCLUSION

The paper describes the proposed system which is used for optical handwriting recognition in various application in different sectors of image processing and computational photography. The paper also describes about the image processors such as MATLAB, OpenCV and FPGA's used for feature extraction after image acquisition. The machine learning algorithms selected for comparison are SVM, Back propagation and Random Forest. These algorithms are compared and evaluated on parameters such as precision, recall, efficiency, accuracy and other measurable factors. The SVM algorithm outclassed the test and was proceeded for further implementation of handwriting recognition system. The proposed work can be further enhanced by converting the system into a mobile application which can be used in schools, colleges and other organization to scan and then convert the image displayed on board into textual document which can be stored in internal or external storage of the smart device.

8. REFERENCES

[1] K. Magiya, A. Joshi, R. Nagpal, S. Yadav, V. Rao, "MULTIPURPOSE REAL TIME HANDWRITING RECOGNITION - An Introduction to Optical Character recognition", International Journal of Medical and Computer Science, 2014.

[2] Dan Cireşan, Ueli Meier, Juergen Schmidhuber, "Multi-column Deep Neural Networks for Image Classification", CVPR 2012, p. 3642-3649.

[3] M. Elleuch, R. Maalej, M. Kherallah, "A New Design Based-SVM of the CNN Classifier Architecture with Dropout for Offline Arabic Handwritten Recognition", Procedia Computer Science, Volume 80 Issue C, June 2016, p. 1712-1723.

[4] Xiao-Xiao Niu, Ching Y. Suen, "A novel hybrid CNN-SVM classifier for recognizing handwritten digits", Pattern Recognition, Volume 45 Issue 4, April, 2012, p. 1318-1325.

[5] M. Zimmermann, J.C. Chappelier, H. Bunke, "Offline grammar based recognition of handwritten sentences," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 5, pp. 818–821, 2006.

[6] R. Plamondon, S.N. Srihari," On-line and off-line handwriting recognition: A comprehensive survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63–84, 2000.

[7] Rich Caruana, Alexandru Niculescu-Mizil, "An Empirical Comparison of Supervised Learning Algorithms", 23rd international conference on Machine learning, Pittsburgh, 2006.

[8] Daniel Rueckerts, Robin Wolz, Paul Aljabar, "Machine learning meets medical imaging : Learning and discovery of clinically useful information from images", London 2014.