International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:3, No:10, 2009

A Cognitive Model of Character Recognition Using Support Vector Machines

K. Freedman

***Abstract*—**In the present study, a support vector machine (SVM) learning approach to character recognition is proposed. Simple feature detectors, similar to those found in the human visual system, were used in the SVM classifier. Alphabetic characters were rotated to 8 different angles and using the proposed cognitive model, all characters were recognized with 100% accuracy and specificity. These same results were found in psychiatric studies of human character recognition.

***Keywords*—**Character recognition, cognitive model, support vector machine learning.

I. Introduction

OPTICAL character recognition (OCR) may seem simplistic due to the fact that humans make, recognize, and read hundreds of characters every day. However, this ability that humans have is not as easily modeled or replicated as once thought. Just thinking about the variability in handwriting between different people and the fact that humans can still read what is written is astonishing. Much research has been put into developing character recognition computer algorithms however none have achieved the level of performance that humans process.

OCR machines have been available since the 1950's and due to the continued research in this field and the growing complexity of digital processing techniques, the performance of the character recognizers gradually have increased. The significance and practicality of these inventions range from helping people with disabilities such as the blind (reading aid) to helping workers and companies deal with the enormous amount of paper/forms generated on a daily basis [1]. Psychologists are also interested in OCR for the purpose of trying to model the cognitive processes that go on in the brain when humans recognize characters. Studies conducted by psychologists have provided a wealth of data about how humans perform at various recognition tasks. If models are then generated that fit that data, theories about how the human mind works can then be proposed.

There are four major subdivisions of OCR which specifies the type of input: fixed font character recognition, on-line character recognition, handwritten character recognition, and script recognition. The model being proposed here is a form of fixed font character recognition. The approaches to making a character recognition system can be classified as either statistic pattern recognition or syntactic/linguistic. Both types have good and bad qualities about them and some researchers have proposed using a hybrid technique. However it is our goal to model a human cognitive phenomenon and so the complexities involved with making a robust character recognition system are therefore alleviated. The theory, trends, and meaning behind the data are most scientifically significant and desirable.

The classification scheme used here relies on certain features being extracted from the test data. Feature detection OCR's can be described as either template matching and correlation techniques or as feature analysis and matching techniques. The template matching and correlation technique essentially takes the input character and compares it to a database of standard prototypes. It has been noted however that this form of character recognition does not explain how humans recognize characters. The feature analysis and matching techniques extract features from the input character and compare them to the feature descriptions of a set of ideal characters. This technique is more common and is used in the present study. It is also noted that human reasoning is better represented using this method compared to template matching [1].

Psychologists and cognitive scientists who want to model human cognition in a parallel distributed processing (PDP) scheme generally use a neural network methodology. Neural networks generally work by using a training dataset to determine the weights/gains of the each neuron. Once the weights are optimized using an algorithmic approach, a test dataset is used to determine the classifying ability of the structure [2]. However neural networks are not the only type of classifier available. Support vector machines are another type of classification algorithm that is also used in the field of Artificial Intelligence. The mechanics of how neural networks and SVM's classify a given dataset differ. Neural networks work by minimizing the misclassification of data in the training set. SVM's on the other hand maximize the distance between the decision boundary and the most similar samples from each class allowing for better performance on unseen data [3].

II. Psychological Phenomena

Character recognition studies on humans have found a number of interesting cognitive phenomena. One of the

K. Freedman is with Drexel University, Philadelphia, PA 19104 USA (corresponding author phone: 215-396-675; e-mail: kjf32@ Drexel.edu).

International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:3, No:10, 2009

widely held and proven findings in such studies is that humans can accurately recognize specific letters irrespective of the angle of orientation [4]. The basic methodologies of such tests include volunteers sitting with their heads on a head/chin rest to keep their vision straight. A screen is kept at a constant distance from the volunteer and displays the test characters. The character could be oriented on the screen at different rotational angles unknown to the volunteer. The volunteer would then have to respond indicating whether they recognized the character in the allotted time [5]. It is the goal of this study to develop a cognitive model using support vector machines which is able to replicate the observed psychological data.

III. The Cognitive Model

Similar to the way neural networks are implemented, support vector machines (SVM) require a training dataset and a testing dataset. The positive class in the training dataset was a rotated letter and the negative class was randomly generated non-letters. Since the output of a SVM is binary, an individual SVM structure was made for each letter in the alphabet.

SVM's classify data based on features that were extracted from a given image. For N-features, an N-dimensional feature space is generated. During training, a decision boundary is created that separates the two clusters corresponding to each class of data. The decision boundary was found using the radial basis function offered by MATLAB™. Figure 1 shows a 2-dimensional projection of the feature space and how the decision boundary separates the 2 clusters.

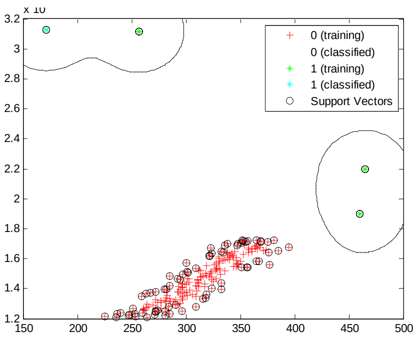

Fig. 1 A 2-Dimensional projection of the feature space. The four enclosed points are various rotations of the letter 'A'. Red points: randomly generated data.

A. *Feature Extractions*

A total of ten features were extracted from 1400 x 1400 pixel images. The image was loaded into MATLAB™ in order to detect various characteristics of the image. For each image, 10 features are extracted and make up a single point in the multi-dimensional space.

The features were calculated for a given binary image as described below:

Feature 1: diagonal summation of the image.

Feature 2: diagonal summation of the transposed image.

Feature 3: The kurtosis of the column-wise summation of the image.

Feature 4: The skewness of the column-wise summation of the image.

Feature 5: through Feature 10: The image is segmented into three horizontal and three vertical strips. The number of binary positive elements within a given strip was summed.

IV. Results and Discussion

In the final SVM, features 3 and 4 were removed for two reasons. First, the likelihood of humans using such complex feature extraction techniques is highly unlikely, and secondly, classification was successful without them.

Input data consisted of each letter of the alphabet rotated in at eight different angles (0, 45, 90, 135, 180, 225, 270, and 315). Therefore the testing data consisted of 208 (26*8) input characters that needed to be correctly classified. The results were successful in both accuracy and specificity. Every letter was correctly recognized without regard to the angle of orientation and no letters were falsely recognized.

In order to further validate the SVM model, the effects of noise was determined (Figure 2). Since this model is not optimized for robust OCR and is meant to model cognition, human data is needed for validation. It was found in a human study that error rate does indeed have the same sigmoidal curve seen in Figure 2 indicating the presence of a critical noise threshold [6]. If robust OCR is the goal, it is suggested that new features be implemented as well as optimization of the training dataset.

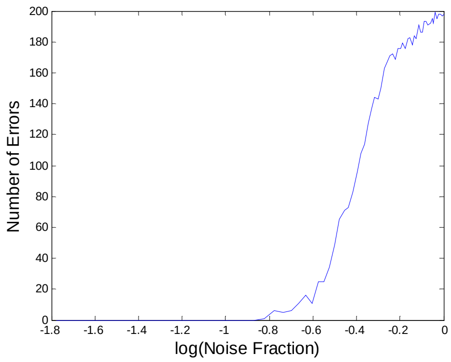

Fig. 2 Effects of noise on the SVM model. Number of errors is plotted against log (noise factor).

The theory behind SVM's may not be so different to how the human brain works. It is well known that the human visual system has groups of cells that act as simple feature detectors such as straight lines. These cells then connect to higher levels of visual processing [7]. The same process is very similar to the way SVM's work. Simple features are detected and converge into a space where a decision is made about what the image represents. It can also be argued that the model is distributed due to the need for multiple classifiers (one for each letter).

REFERENCES

[1] V. K. Govindan, "Character recognition- a review," *Pattern Recognition*, vol. 23, no. 7, pp. 671–683, 1990.
[2] J. Cao, M. Ahmadi, and M. Shridhar, "Recognition of handwritten numerals with multiple feature and multistage classifier," *Pattern Recognition*, vol. 28, no. 2, pp. 153–160, 1995.
[3] A. Shoeb, H. Edwards, J. Connelly, B. Bourgeois, S. Treves, and J. Guttag, "Patient-specific seizure onset detection," *Epilepsy & Behavior*, vol. 5, no. 4, pp. 483–498, 2004.
[4] M. Corballis, J. Zbrodoff, L. Shetzer, and P. Butler, "Decisions about identity and orientation of rotated letters and digits," *Memory and Cognition*, vol. 6, no. 2, pp. 98–107, 1978.
[5] A. Koriat, J. Norman, and R. Kimchi, "Recognition of rotated letters: extracting invariance across successive and simultaneous stimuli," *Journal of Experimental Psychology*, vol. 17, no. 2, pp. 444–458, 1991.
[6] D. Parish and G. Sperling, "Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination," *Vision Research*, vol. 31, no. 7, pp. 1399–1415, 1991.
[7] P. Series, J. Lorenceau, and Y. Fregnac, "The "silent" surround of V1 receptive fields: theory and experiments," *Journal of Physiology*, vol. 97, pp. 453–474, 2003.