

A Simple Systematic Approach to Mood Invariant Handwriting Analysis Using SVM Classifier

Syeda Asra

Visveswaraya Technological University, Belgaum:
Department of Computer science & Engineering
Appa Institute of Engineering & Technology
Gulbarga, India
asrascholar@gmail.com

Dr.Shubhangi D.C

Visvevaraya Technological University, Belgaum:
Department of PG Regional Centre, Gulbarga, India
shubhangidc@vtu.ac.in

Abstract—Handwriting is actually brain writing. The subconscious mind comes into play, which reveals the traits .A systematic approach is proposed towards understanding the Human behaviors through handwriting. The feature vector is built by considering the distance between words, orientation and alignment of the words right or left. The data set was developed using 200 samples from people belonging to different works of life. This system developed is mood invariant since the samples were collected by different people at different point of time and unconstrained hand writing analysis. The handwriting of 100 adults was submitted for graphological analysis. The graphologist's answers to questions on the person's personality, her description of his character and her assessment of his inclination towards past ,present and current were checked by the person's own answers to the questionnaire, by the personality descriptions in the case-sheets, and by the results of the Progressive Matrices Test . Further SVM classifier was used. The results were checked and compared with graphologist. The result as high as 95%were obtained.

Keywords—*cursive handwriting ;human behaviour; graphology ;personality;*

I. INTRODUCTION

To understand this concept, one needs to understand the difference between a trait and a state. A trait is a relatively permanent individual characteristic. For example, most of known people who are outgoing, friendly, confident, or shy when we describe these people, we use these traits[1] to better understand their personality; to better understand who they are. A state, on the other hand, is a temporary change in one's personality. Examples of states might be angry, depressed, fearful, or anxious. We typically use states to describe a person's reaction to something. To describe person's behavior traits are to be considered.

In order to quantify these traits, handwriting samples are considered. In the proposed work cursive handwritten samples were analyzed. Handwriting is brain writing. Cursive

handwriting necessitates physical sequential strokes. To form just one letter massive regions in the brain are activated, including areas of thinking, language, and temporary information storage and management [2]. Today, keyboarding is in, the common core standards and no longer requires students to learn cursive, and some schools are dropping the teaching of cursive, dismissing it as an ancient skill [3]. Scientists are discovering that learning cursive is an important tool for cognitive development, particularly in training the brain to learn functional specialization, [4] that is capacity for optimal efficiency.

Handwriting reveals hundreds of elements of the person's "personality and character" which includes glimpses into the subconscious mind, intellect, energy, fears, motivations, imagination, integrity, aptitudes, etc. There are over 100 individual traits revealed and an unlimited number of combinations. But, to accommodate all these traits in one paper is not plausible. Hence, in this work proposed, I dots with eighteen classes are considered for analysis.

II. RELATED WORK

Margins [5] in handwriting reflect an individual's personality [6]. Margins could either be narrow or wide with each showing certain personality traits about the writer. Narrow left margin indicate writer is attached to his past than to his future. Wide left margin indicates writer leaves behind ones past and continue moving. A wide right margin indicates writer may be afraid to take a future step. A narrow right margin indicates writer may be willing to take a forward step and that he may not be experiencing uncertainties at that period of his life. Balanced sizes indicate the writer balanced person when it comes to risk taking. In addition to margin the other work carried out was on placing f I dots [7]. Since the placing f I dots reveal different traits. A dot placed exactly on the stem reveals the person is very keen in his work; he pays attention to minute details as well. The dots on the left and right indicate he lives in the past and future respectively. The handwriting of 100 adults was submitted for graphological analysis. The graphologist's answers to questions on the patient's personality,

her description of his character and her assessment of his inclination towards past, present and current were checked by the person's own answers to the questionnaire, by the personality descriptions in the case-sheets, and by the results of the Progressive Matrices Test. In this work carried, the performance was as high as 95%. The feature extraction was done using Zernike moments [8].

III. METHODOLOGY

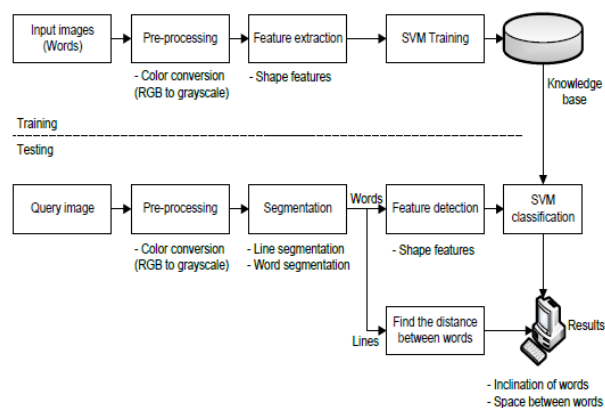


Fig.1. Architecture design of the proposed system

- Initially the work begins with the pre-processing of the given input image and their by performing the color conversion as RGB to gray scale. Following the pre-processing feature extraction and SVM training are performed in the training phase. Having the trained features in knowledge base the testing works with segmentation and then performs the feature detection and SVM classification for the given query image and finally shows the results.
- Images are provided the network. Words can be of any length, complete or incomplete, left or right inclined. Color conversion of all the images from RGB color to gray scale is done in preprocessing step. Following this the shape features of all these images are extracted and trained using SVM algorithm. And these feature vectors are stored in database known as Knowledge base.
- Testing query image is carried out for the unknown data set. The distance between words is calculated finally results are classified as "Left inclined, complete words", "Left inclined, incomplete words", "Right inclined, complete words".

A. Image Acquisition and Database Creation

Data samples of 500 in number were collected from people belonging to different works of life both equally from males & females. In order to make the results time and mood invariant

the samples were collected in different days at different points of time. A4 paper with black ball point pen with a hard surface was used for writing. Three paragraphs were given to write. The images were scanned using laser jet scanner with a resolution of 2528x3507 pixels and 300dpi.

B. Noise Removal & Image Handwriting Pre-Processing

Area based noise removal algorithm is used. OTSUREC [9] returns a set of thresholds for the input image using the multi-level otsu algorithm. OTSUREC computes a set T of thresholds from the input image I employing the multi-level Otsu algorithm. The multi-level Otsu algorithm consists in finding the threshold that minimizes the input image intra-class variance. Then, recursively, the Otsu algorithm is applied to each image region until total thresholds are found.

C. Shape features

- Area – Actual number of valid (pixel value=1) pixels in the region.
- Perimeter –Perimeter is the distance around the boundary for region. Function region props will find the perimeter through calculating distance among each adjoining pair of pixels around border of the region. When the image has any discontinuous regions, region props will give unexpected results.
- Form factor – is computed by utilizing Area and Perimeter. Form factor = $4 \times \pi \times \text{Area} / \text{Perimeter}^2$
- Major axis length – This will give the length for major axis of the ellipse which has same normalized second central moments as the region. The 2-D input label matrices will only support this property.
- Minor axis length – This will give the length of the minor axis of the ellipse which have the same normalized second central moments as the region. The 2-D input label matrices will only support this property.
- Roundness – is computed using Area and Major axis length.
- Roundness = $4 \times \text{Area} / (\pi \times \text{Major axis length}^2)$
- Compactness – is computed as $\text{Compactness} = 4 \times \text{Area} \pi / \text{Major axis length}$
- Density – is formulated by utilizing Bounding Box. A smallest rectangle which contains the region can be defined as Bounding Box, a 1-by-Q *2 vector, where Q is the number of image dimensions.
- Black pixels each line – Considering each row and each column of the binary image the number of white pixels can be computed. And inverting this value with the width of the image gives us the number of black pixels present in each line. By efficient utilization of content-based image features such as

shape features in our proposed system, the semantic gap can be minimized.

D. SVM Classifier

We have used SVM [10] model to perform the classification. All the 10 extracted features such as texture, shape orientation of the I dots from the image are the input pattern for SVM network. Once trained for each image we want to analyze, the output of the net is a value between 1to 16, that indicates the input pattern belongs to the desired personality profile.

IV. EXPERIMENTAL RESULTS

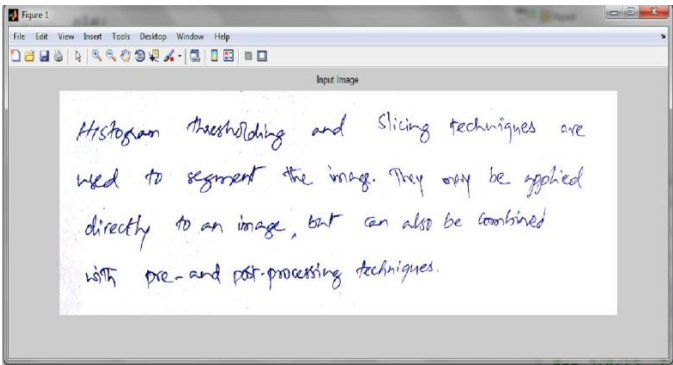


Fig.1. Input image

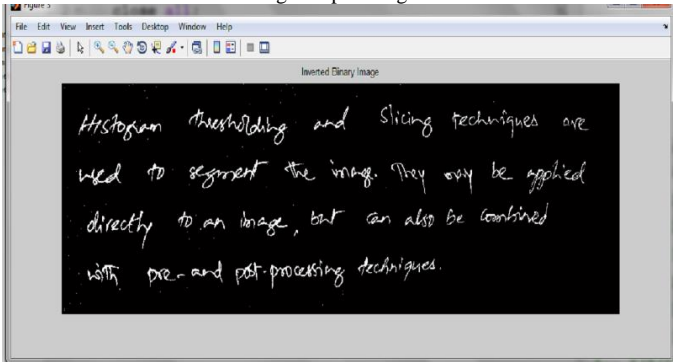


Fig.2. Binary image

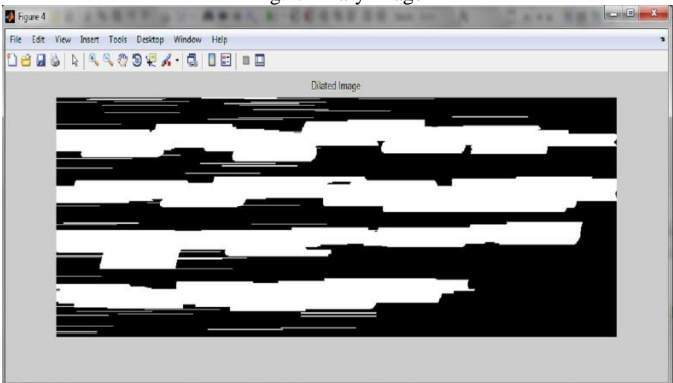


Fig.3. Dilated image

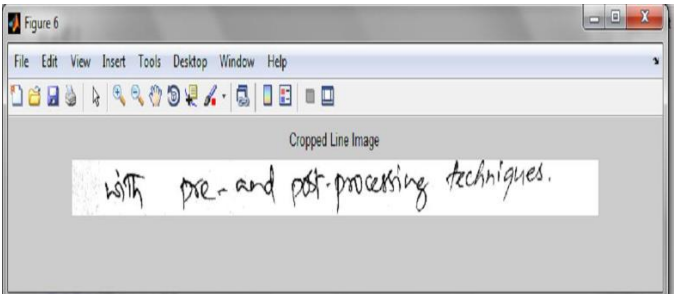


Fig.4. Cropped line image

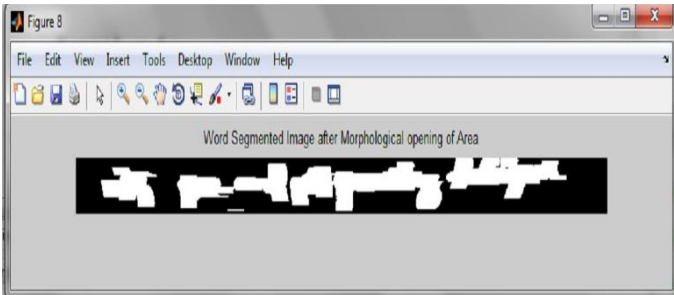


Fig.5. Word segmentation

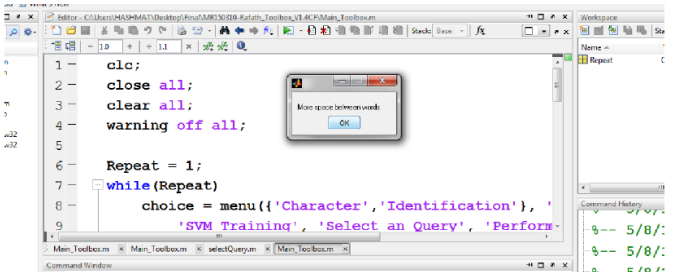


Fig.6. More space between words



Fig.7. Right inclined and complete words

Table I
Inclination and Spacing between words and its Psychological Analysis

S.No	Personality Trait Identification	
1.	Right inclined	writer reacts strongly to emotional situations
2.	Right inclined complete words	writer has a balanced emotional state

S.No	Personality Trait Identification	
3.	Close space between the words	indicates a symbol of cooperation
4.	Large gap through words	shows an individual being restful alone
5.	Incomplete words	Loses energy at the end
6.	Complete words	Sustain energy till the end of the task
7.	i dot is an arc that opens at the left side	it is a sign of the neurotic, untrustworthy personality. He shies away from the system.

V. RESULTS AND DISCUSSION

Table I shows different styles of handwriting with respect to alignment spacing between words and their specific trait associated with it. The work is ongoing for different styles in which character O can be written and it reveals many interesting traits. The system developed can be used in recruitments, so that right person gets the right job.

ACKNOWLEDGMENT

I would like to thank my parents my husband my kids and my guide for giving me encouragement in all kind of endeavors directly or indirectly. I further extend my thanks to the famous psychologist Gordon Allport and above all the Almighty.

REFERENCES

- [1] Syeda Asra, 2 Dr.Shubhangi D.C,” Personality Trait Identification - A Survey”, IJCSN International Journal of Computer Science and Network, Volume 3, Issue 2, April 2014.
- [2] Slape, L. Cursive Giving Way to Other Pursuits as Educators Debate Its Value. The Daily News, Feb. 4, 2012.
- [3] Lisa garber, “How Cursive Hand writing Uniquely helps Brain Development” A News Letter ,January 2013.
- [4] James, Karin H. an Atwood, Thea P. (2009).The role of sensorimotor learning in the perception of letter-like forms: Tracking the causes of neural specialization for letters. Cognitive Neuropsychology.26 (1), 91-100.
- [5] Syeda Asra, and Dr. Shubhangi DC,”Specific Trait Identification in Margins Using Hand Written Cursive Text”. Int. Journal of Engineering Research and Applications , ISSN: 2248-9622, Vol. 6, Issue 4, (Part - 4) April 2016, pp.89-94.
- [6] Syeda Asra and Dr.Shubhangi DC,” Personality Trait Identification –A Survey” IJCSN International Journal of Computer Science and Network, Volume 3, Issue 2, April 2014 ISSN (Online) : 2277-5420
- [7] Syeda Asra and Dr.Shubhangi DC,” Personality Trait Identification Using I Dots” International Referred Journal of Scientific Research in Engineering”(accepted and awaiting for publishing)
- [8] Evangelos Sariyanidi ; Department of Control Engineering, Istanbul Technical University, Turkey ; Volkan Dağlı ; Salih Cihan Tek ; Birkan Tunç more authors” Local Zernike Moments: A new representation for

face recognition” Page(s):585 – 588 Published in: 19th IEEE International Conference on Image Processing (2012).

- [9] Jun Zhang : Image Segmentation Based On 2D Otsu Method With Histogram Analysis ,IEEE Intl Conf. On Computer Science And Software Engineering 2008.
- [10] Guangjun Shi ; Sch. of Autom., Beijing Inst. of Technol.,Beijing, China; Xiangyang Xu ; Yaping Dai “SIFT Feature Point Matching Based on Improved RANSAC Algorithm”, Page(s):474 – 477 Aug. (2013).
- [11] Syeda Asra, Dr.Shubhangi D.C,” Personality Trait Identification Using Unconstrained Cursive and Mood Invariant Handwritten Text”, IJ. Education and Management Engineering, 2015, 5, 20-31 Published Online October 2015 in MECS.