International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-12 Issue-10, September 2023

Text Independent Writer Verification using Spatial Domain Technique under Varying Writing Conditions

Sharada L. Kore, Shaila D. Apte

Abstract: This work presents the results of text-independent writer verification under various environmental conditions, including writing instrument, ink width, paper quality, image size, image type, scanning device, and resolution. It is observed that the style of a writer and the kind of writing instrument significantly used affect the handwriting. Additionally, the quality of handwritten samples has a significant impact on the performance of a writer verification system. The objective of our work is to verify a writer under changing conditions. To improve the verification rate, directional information is captured from the image contours, and the histogram of chain code is utilised as a feature. The system is tested on our dataset of 261 writers, each with two samples, using different writing instruments. A total of 50% of the samples are used to train the system, and the system is then tested on the remaining 50% of the samples. The system is tested using chain code. The total error rate (TER) is 0.76 %. The method works even under poor resolution.

Keywords: Chain Code, Handwriting Identification, Pattern Recognition, Person Authentication, Writer Verification

I. INTRODUCTION (HEAD 1)

Writer identification and verification have great importance in the forensic area for determining identity in conjunction with the intentional aspects of a crime. The other applications of writer identification are biometric security systems, historic document analysis, postal address interpretation, and bank cheque recognition. Most handwriting identification experts rely almost entirely on manually intensive techniques today. Although some literature is available on prototype toolsets for document examination, there does not exist any tool that has completely automated the handwriting identification process, which gives 100% accuracy under varying writing conditions [12].

Writer verification involves a one-to-one comparison to determine whether the same person wrote the two samples. The main challenges are high within writer variations and low between writer variations.

Manuscript received on 08 July 2023 | Revised Manuscript received on 05 August 2023 | Manuscript Accepted on 15 September 2023 | Manuscript published on 30 September 2023.
*Correspondence Author(s)
Sharada L. Kore*, Ph.D. Research Scholar, Bharati Vidyapeeth Deemed University College of Engineering, Pune (M.H.), India, Email ID: sharadakore@gmail.com
Shaila D. Apte, Anubhuti Solutions, Pune (M.H.), India, Email ID: sdapte@rediffmail.com

Handwriting is influenced by physiology, training and other behavioural factors and affected by writing conditions, writing instrument, writing position, writing surface, adequacy of standards, alternative styles, physical conditions and mental state, etc [13]. The type of writing instrument used decides an ink width variation in handwritten documents [4] [10]. The time required and accuracy obtained in a writer identification and verification system depend on the feature extraction method. The quality of a handwritten image, which depends on the image acquisition device, scanning resolution, image type, paper quality, etc., plays a vital role in deciding the performance of a writer verification system [14]. Therefore, it is necessary to design a generalised model for a writer verification system that achieves high accuracy and speed under varying writing conditions.

Our work employs various methods to verify a person's handwriting under different writing conditions and resolutions. The Freeman chain code represents the directional information of each pixel on the boundary of a contour [8]. The technique is based on the hypothesis that the contour of a handwritten sample encapsulates the writing style of its writer [5].

The survey of existing methods is presented in Section II. The objective of our work is to verify a writer under different writing conditions and different ink widths. It could not be evaluated on existing data sets [7]; hence we have created new dataset and discussed in section III. The proposed feature extraction methods are explained in Section IV. Results are presented in section V. Section VI, deals with the conclusion.

II. LITERATURE SURVEY

The survey of existing methods is presented in [1] [6] [9] [11]. Said et al. employed the multichannel Gabor filtering technique for writer identification and verification [15]. Zhang et al. measured the individuality of handwritten characters through identification and verification models [16]. Schlapbach and Bunke showed that writer-specialised handwriting recognizers can be used for writer identification and verification [17]. He et al. treat Chinese writer identification using wavelet-based Generalized Gaussian Density (GGD) methods [18]. Bensefia et al. and Pervouchine and Leedham used local features for writer identification and writer verification [19], [20]. Bulacu and Schomaker employed the probability density functions of bitmaps and/or connected-component contours, analyzing both the texture and allograph level approaches [2], [3].

*Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)*

III. DATASET

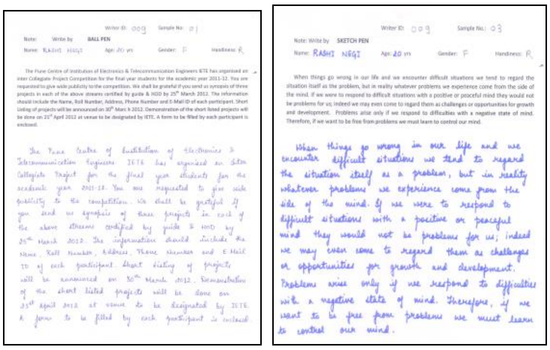
We have created three different datasets comprising a total of 261 writers from various age groups and genders. The dataset includes four samples, each from 261 writers, including writing with a Ball pen and a sketch pen. Samples were scanned using a professional HP scanner at three different resolutions: 96, 200, and 300 dpi, with an 8-colour level.

The subjects are asked to write a given text of 6 lines using a ballpoint pen and a sketch pen on A4-size plain paper. The given text differs for different sets of writers. The dataset comprises a total of 1236 samples, collected under varying writing conditions. Out of 1236 samples, 124 samples are scanned at a resolution of 96 dpi and stored in a computer as PNG images. The remaining 1112 samples are scanned at a resolution of 300 dpi and saved as JPEG type. Out of 1112 samples, 192 samples are scanned at 200 dpi and saved as JPG type. The details of the dataset are given in table1.

Table-I: Details of Our Newly Created Dataset at the Page Level

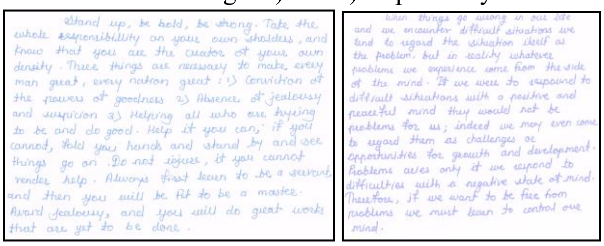
Dataset	Image type	Scanning resolution	No. of writers	Size of dataset
Set 1	PNG	96 dpi	31	124
Set 2	JPG	200 dpi	48	192
Set 3 (includes writers of dataset 2)	JPG	300 dpi	230	920

The scanned handwritten sample pages of the same writer, written with different writing instruments, are shown in Figure 1.

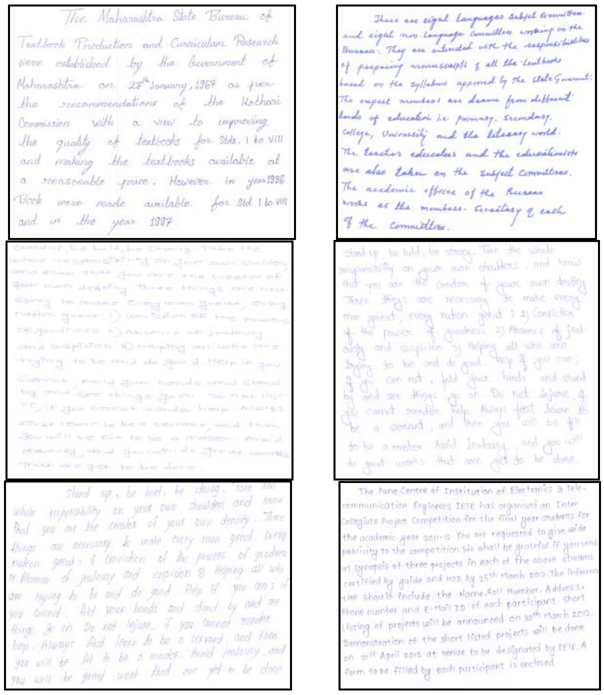


[Fig.1 a) Scanned Handwritten Samples of Same Writer Using Ball Pen and Sketch Pen]

Separation of printed text and handwritten text is done manually. The samples of the same writer and different writers are shown in Fig. 2a) and b) respectively.



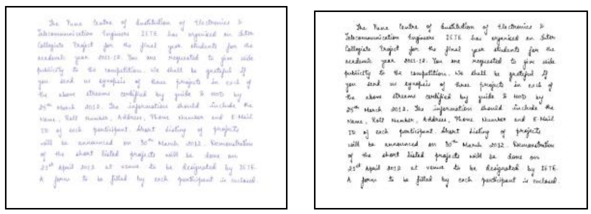
[Fig.2 a) Same Writer Samples]



[Fig.2 b) Different Writer Samples]

IV. FEATURE EXTRACTION

The present work uses the chain code feature. The histogram of the chain code is calculated to capture the directional information of the writer. First the color image is converted to a gray scale and then it is binarized using Otsu's thresholding using MATLAB as shown in figure 3 a) and b) respectively.



(a) Handwritten image (b)Binary Image of fig 3a)

[Fig.3: (a) Handwritten Image and (b) Binary Image of (a)]

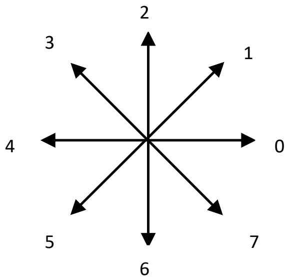
The boundary detection algorithm is used to detect the contours, as shown in fig. 4.

[Fig.4: Boundary Detection in Binary Image. Red trace represents the Boundaries, and Green Star Represent the Starting of the Boundary Pixel]

The contour is represented using the Freeman chain code. The histogram of this chain code is used as a

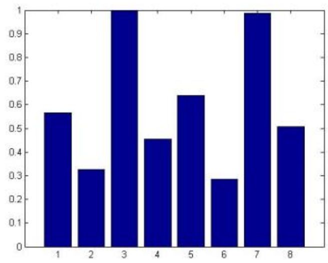
Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

feature. The chain code represents the directional component for the boundary. The direction is represented for eight connected components. The boundary pixels in the original binary image are then labelled by their respective codes as shown in fig.5.



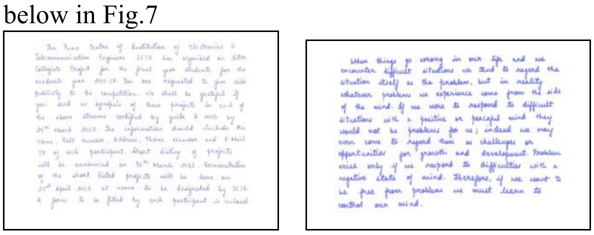
[Fig.5. Directions of Boundary Pixels for 8 Connected Components]

We then proceed to extract features from the newly formed image. To make the chain code independent of the starting point, the normalized histogram of chain code is used as feature. The normalized histogram of handwritten sample image shown in fig3 a) is shown below in Fig. 6.

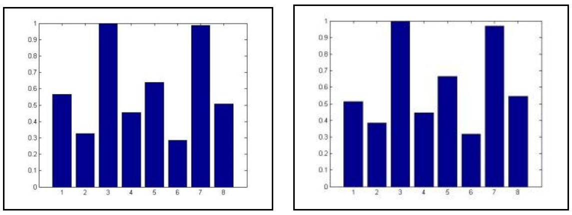


[Fig.6: Normalized Histogram of Image Shown in Fig.3 a)]

The histogram has 8 bins; hence, the size of the feature vector is eight elements. The method is applied to all images in the dataset. The dataset includes two samples with a ball pen and two samples with a sketch pen. The two histograms of two handwritten photos of the same writer are shown below in Fig.7



(a) writing with a ball pen (b)writing with a sketch pen

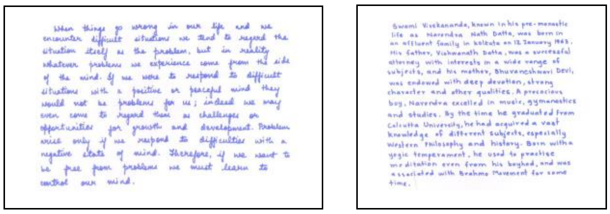


(c) histogram of image in (a) (d) histogram of image in (b)

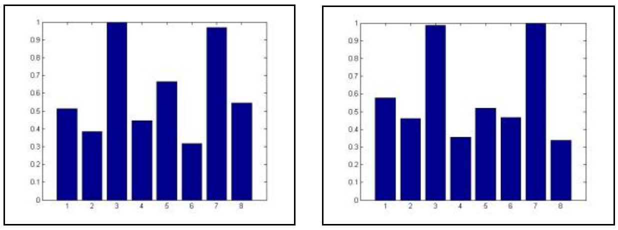
Figure 7 a) handwritten image of writer 009 showing writing with a ball pen, (b) handwritten image of writer 011 showing writing with a sketch pen, (c) histogram of the image in (a), and (d) histogram of the image in (b).

Histograms of two handwritten images from different writers are shown below in Fig. 8.

Referring to Figs. 7 and 8 show that the two histograms of handwritten images from the same writer, using a ballpoint pen and a sketch pen, are identical. In contrast, the two histograms of handwritten images from different writers are distinct. The chain code can be used to discriminate between the handwritten document images.



(a)image by writer 009 (b)Image by writer 011



(c) Histogram of image in (a) (d) Histogram of image in (b)

[Fig.8 a) Handwritten Image of Writer 009 (b) Handwritten image of Writer 011 (c) Histogram of Writer 009 with ball Pen and (d) Histogram of Writer 011 with Sketch Pen]

V. EXPERIMENTAL DETAILS AND RESULTS

The system operates in two modes: training mode and testing mode. The training mode evaluates a writer-specific threshold value, which is based on the false acceptance rate (FAR) and false rejection rate (FRR). The initial threshold is calculated as follows.

The maximum distance within-writer class and the minimum distance between-writer class were calculated for each writer in the dataset. The average of the maximum and minimum distances is considered the initial threshold.

The results of the above four experiments are presented in this section. The initial threshold value is updated in this mode to minimise the error between the false acceptance rate (FAR) and the false rejection rate (FRR). Based on the trial-and-error method, the system is trained with 1000 iterations using a threshold step size of 0.001. The system was trained with 50% samples of the total dataset.

Table-II: Effect of Scanning Resolution on Initial and Final Threshold Values of Different Writers

Writer ID	Resolution			
	Initial Threshold value		Final Threshold value	
	200 dpi	300 dpi	200 dpi	300 dpi
0 1 8	0.1752	0.1704	0.1412	0.1399
0 1 9	0.1502	0.1659	0.0912	0.0899
0 2 0	0.0812	0.0873	0.0707	0.0653
0 2 1	0.0911	0.0889	0.0816	0.0874
0 2 7	0.0908	0.0889	0.0908	0.0889

The results of testing for four different experiments are presented in Table 1. Fifty per cent of the total dataset is used for testing the feature performance. The two

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

samples were given to the system, and the decision on whether the same person wrote both samples was made based on the saved threshold value for that particular writer. If the distance between the two test samples is less than the final threshold, then the decision is that the same person wrote both samples; otherwise, they were written by two different persons. The system is tested on 261 writers, with two samples from each writer.

Table-III: Testing Results of Chain Code Method on Different Datasets

Dataset	Image Type	Resolution (dpi)	No. of writers	% Verification Rate
Set 1	PNG	96	31	97.7
Set 2	JPG	200	48	96.33
Set 3	JPG	300	230	99.24

Table-IV: Testing Results of Chain Code Method on Dataset of Same Writers for No. of Writers=48 Writers, Image Type: PNG

Resolution	% Verification Rate
200	96.33
300	95.96

VI. CONCLUSION

Referring to Tables 3 and 4, we can conclude that the normalised histogram of the chain code is a feature that discriminates the writers. The total error rate (TER) is 0.76% using the chain code. The method is independent of text and achieves a good accuracy of 99.24% under varying writing conditions, ink width variations, and scanning resolutions. The process is tested using low-resolution images with a resolution of 96 dpi. Referring to Table 4, it is clear that the method yields good results for low-resolution images. The technique used also works on cursive handwriting. We are conducting experiments with 1000 writers and a varying number of discriminating characters.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This independence ensures that the research is conducted with objectivity and without any external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Author's Contributions:** The authorship of this article is contributed equally to all participating individuals.

REFERENCES

1. R. Plamondon and S. Srihari, "On-line and off-line handwriting recognition: a comprehensive survey," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63–84, 2000. http://www.cedar.buffalo.edu/papers/articles/Online_Offline_2000.pdf
2. M. Bulacu and L. Schomaker, "Text-independent writer identification and verification using textural and allographic features," IEEE Trans. Pattern Analysis and Machine Intelligence, Special Issue - Biometrics: Progress and Directions, IEEE Computer Society, vol. 29, no. 4, pp. 701-717, April 2007. https://doi.org/10.1109/TPAMI.2007.1009
3. L. Schomaker and M. Bulacu, "Automatic writer identification using connected-component contours and edge-based features of upper-case western script," IEEE Trans. Pattern Analysis and Machine Intelligence, IEEE Computer Society, vol. 26, no. 6, pp. 787-798, June 2004. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=c656946b26f7501b74eac6fd4cfac751e3ebb734
4. A. Brink, J. Smit, M. Bulacu, and L. Schomaker, "Writer identification using direct ink-trace width measurements," Pattern Recognition(45), 2012, pp. 162–171, doi: http://doi.org/10.1016/j.patcog.2011.07.005.
5. I. Siddiqi and N. Vincent, "Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features," Pattern Recognition (43), 2010, pp. 3853-3865, doi: http://doi.org/10.1016/j.patcog.2010.05.019
6. M. Sreeraj and S. M. Idicula, "A survey on writer identification schemes," International Journal of Computer Applications (0975 – 8887), vol. 26, no.2, pp. 23-33, July 2011.
7. U. V. Marti and H. Bunke, "The IAM-Database: An English Sentence Database for Offline Handwriting Recognition," Int. J. Document Analysis and Recognition, 2002, vol. 5, no. 1, pp. 39-46.
8. I. Siddiqi and N. Vincent, "A set of chain code-based features for writer recognition," 10th International Conference on Document Analysis and Recognition (ICDAR 09), pp. 981-985, 2009, doi http://doi.org/10.1109/ICDAR.2009.136 .
9. S. Srihari and G.Leedham, "A survey of computer methods in forensic document examination," Proc. of 11th Conference of the International Graphonomics Society(IGS 03), Nov. 2003, pp. 278 - 281.
10. I. Siddiqi and N. Vincent, "Stroke width independent feature for writer identification and handwriting classification," 2008.
11. S.L.Kore and S.D.Apte, "The current state of art-the handwriting a behavioural biometric for writer identification and verification", International Conference on Advanced Computing, Communications and Informatics (ICACCI), Aug 3-5, Chennai, Tamil Nadu, ACM Digital Library, pp.,2012. 978-1-4503-1196-0/12/08
12. S. Gupta, "Automatic person identification and verification," International Institute of Information Technology, Hyderabad, India, Ph.D. Thesis, 2007, March 2008.
13. M. L. Bulacu, "Statistical pattern recognition for automatic writer identification and verification," Ph.D. Thesis, 2007.
14. S. N. Srihari, "Individuality of persons," Phd Thesis, 2001.
15. H. Said, T. Tan, and K.Baker, "Personal identification based on handwriting, " Pattern Recognition, vol. 33, no.1, pp.149-160, Jan.2000. http://dx.doi.org/10.1049/ic:19980678
16. B. Zhang, S.N. Srihari, and S. Lee, "Individuality of handwritten characters," Proc. Seventh Int'l Conf. Document Analysis and Recognition, Aug. 2003. https://doi.org/10.1109/TPAMI.2007.1009
17. A. Schlapbach and H. Bunke, "A Writer Identification and Verification System Using HMM Based Recognizers," Pattern Analysis and Applications, vol. 10, no. 1, Feb. 2007. https://dl.acm.org/doi/abs/10.5555/2736757.2736810
18. Z. He, B. Fang, J. Du, Y. Yan Tang, and X. You, "A Novel Method for Off-Line Handwriting-Based Writer Identification," Proc. Eighth Int'l Conf. Document Analysis and Recognition, 2005. http://dx.doi.org/10.1109/ICDAR.2005.27
19. A. Bensefia, T. Paquet, and L. Heutte, "A Writer Identification and Verification System," Pattern Recognition Letters, vol. 26, pp. 2080-2092, 2005. https://doi.org/10.1016/j.patrec.2005.03.024
20. V. Pervouchine and G. Leedham, "Extraction and Analysis of Forensic Document Examiner Features Used for Writer Identification," Pattern Recognition,

*Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)*

vol. 40, 2007. https://doi.org/10.1016/j.patcog.2006.08.008

AUTHOR'S PROFILE

Sharada L. Kore received a Master of Engineering degree in Electronics from Shivaji University, Kolhapur, Pune, India, in 2000. She registered for a Ph.D. at Bharati Vidyapeeth Deemed University, Pune, India, in 2009. She has 13 years of experience in teaching. Currently, she is an associate professor in the E&TC department. at Bharati Vidyapeeth's College of Engineering, Pune, India. This space is for writing your short resume.

Shaila D. Apte is currently working as a Professor at Rajarshi Shahu College of Engineering, Pune. She has formerly been an Assistant Professor at Walchand College of Engineering, Sangli, for 27 years; a member of the Board of Studies at Shivaji University; and a Principal Investigator for a research project sponsored by the Armament Research and Development Establishment (ARDE), New Delhi. She has 28 years of extensive teaching experience in Electronics Engineering.