

REFERENCES

- [1] Hand DJ (2007), "*Principles of data mining*", *Drug safety*, Vol.30 (7), pp.621-622
- [2] Tan, PN, (2007), "*Introduction to data mining*", Pearson Education India, pp.1-769.
- [3] Witten IH, Frank E, Hall MA and Pal CJ (2016), "*Data Mining: Practical machine learning tools and techniques*", pp.1-654.
- [4] Chen MS, Han J and Yu PS, (1996), "*Data mining: an overview from a database perspective. IEEE Transactions on Knowledge and data Engineering*", Vol.8 (6), pp.866-883.
- [5] Berry MJ and Linoff GS (2009), "*Data mining techniques*", John Wiley & Sons.
- [6] Raval KM (2012), "*Data Mining Techniques*", *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol.2(10), pp.439-442.
- [7] Ali SM and Tuteja MR (2014), "*Data Mining Techniques*", *International Journal of Computer Science and Mobile Computing*, Vol.3 (4), pp. 879-883.
- [8] Hegland M (2001), "*Data mining techniques*", *Acta numerica*, Vol.10, pp.313-355.
- [9] Shekhar S, Zhang P, Huang Y and Vatsavai RR (2003), "*Spatial data mining*", pp.1-23.
- [10] Li D, Wang S et al (2015), "*Spatial data mining*", Springer Berlin Heidelberg.
- [11] Shekhar S, Kang J and Gandhi V (2009), "*Spatial data mining. In Encyclopedia of Database Systems*", pp. 2695-2698.
- [12] Giannotti F, Nanni M, Pinelli F and Pedreschi D (2007), "*Trajectory pattern mining*", In *Proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 330-339.
- [13] Al-Ma'adeed S, Mohammed E, AlKassis D and Al-Muslih F (2008), "*Writer identification using edge-based directional probability distribution features for Arabic words*, *IEEE / ACS International Conference on Computer Systems and Applications*, pp.582-590.
- [14] Bensefia A, Paquet T and Heutte L (2005), "*A writer identification and verification system*", *Pattern Recognition Letters*, Vol.26 (13), pp.2080-2092.
- [15] Maarse F and Thomassen A (1983), "*Produced and perceived writing slant: differences between up and down strokes*", *Acta Psychologica*, Vol.54 (1-3), pp.131-147.

- [16] Crettez JP, (1995), “*A set of handwriting families: style recognition*”, Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, Canada, pp. 489–494.
- [17] Namboodiri AM and Jain AK (2004), “*Online handwritten script recognition*”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.26(1), pp.124-130.
- [18] Awaida SM and Mahmoud SA (2012), “*State of the art in off-line writer identification of handwritten text and survey of writer identification of Arabic text*”. Educational Research and Reviews, Vol.7 (20), pp.445-463.
- [19] Al-Dmour A and Abu Zitar R (2007), “*Arabic writer identification based on hybrid spectral-statistical measures*”, Journal of Experimental and Theoretical Artificial Intelligence, Vol. 19(4), pp. 307–332.
- [20] Bulacu M and Schomaker L (2007), “*Text-independent writer identification and verification using textural and allographic features*”, IEEE transactions on pattern analysis and machine intelligence, Vol.29 (4), pp.701-717.
- [21] Bulacu M, Schomaker L and Brink A (2007), “*Text-independent writer identification and verification on offline arabic handwriting*”, Ninth International Conference on Document Analysis and Recognition, Vol. 2, pp. 769-773.
- [22] Kannan RJ, Prabhakar R and Suresh RM (2008), “*Off-line cursive handwritten Tamil character recognition*”, International Conference on Security Technology (SECTECH'08), pp. 159-164.
- [23] Al-Maadeed S (2012), “*Text-dependent writer identification for arabic handwriting*”, Journal of Electrical and Computer Engineering, pp.1-8.
- [24] Jayanthi SK and Rajalakshmi D (2011), “*Writer identification for offline Tamil handwriting based on gray-level co-occurrence matrices*”, Third International Conference on Advanced Computing (ICoAC), pp. 187-192.
- [25] Saranya K and Vijaya MS (2013), “*Text dependent writer identification using support vector machine*”, International Journal of Computer Applications, Vol.65 (2), pp.6-11.
- [26] Karunakara K and Mallikarjunaswamy BP (2011), “*Writer identification based on offline handwritten document images in Kannada language using Empirical Mode Decomposition method*”, Vol.30(6), pp.31-36.

- [27] Patil V and Shimpi S (2011), "*Handwritten English character recognition using neural network*", Elixir Computer Science Engineering, Vol.41, pp.5587-5591.
- [28] Feng J and Zhu Y (2006), "*Text independent writer identification based on Gabor filter and SVM classifier*", In Sixth International Symposium on Instrumentation and Control Technology: Signal Analysis, Measurement Theory, Photo-Electronic technology and Artificial Intelligence, pp. 63570Z-63570Z.
- [29] Said HES, Peake GS, Tan TN and Baker KD (1998), "*Writer Identification from Non-uniformly Skewed Handwriting Images*", In BMVC, pp. 1-10.
- [30] He Z, You X and Tang YY (2008), "*Writer identification using global wavelet-based features. Neurocomputing*", Vol.71 (10), pp.1832-1841.
- [31] Schlapbach A, Kilchherr V and Bunke H (2005), "*Improving writer identification by means of feature selection and extraction*", Eighth International Conference on Document Analysis and Recognition, pp. 131-135.
- [32] Dhandra BV, Mukarambi G and Hangarge M (2012), "*Writer identification by texture analysis based on Kannada handwriting*", Int. J. Comm. Netw. Secur, Vol.1 (4), pp.80-85.
- [33] Schapire R (1998), "*Theoretical Machine Learning Technical Report*", Princeton University, www.cs.princeton.edu/~schapire.
- [34] Ikonomakis M, Kotsiantis S and Tampakas V (2005), "*Text classification using machine learning techniques*", WSEAS transactions on computers, Vol.4 (8), pp.966-974.
- [35] Mitchell TM (1997), "*Machine Learning*", McCraw-Hill, Boston.
- [36] Nasrabadi NM (2007), "*Pattern recognition and machine learning*", Journal of electronic imaging, Vol.16 (4), p.049901.
- [37] AI Hasan M, Chaoji V, Salem S and Zaki M (2006), "*Link prediction using supervised learning. In SDM06: workshop on link analysis, counter-terrorism and security*", pp.1-10.
- [38] Hastie T, Tibshirani R and Friedman J (2009), "*Unsupervised learning. In The elements of statistical learning*", pp. 485-585.
- [39] Le QV (2013), "*Building high-level features using large scale unsupervised learning*", IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8595-8598.
- [40] Zhu X and Goldberg AB (2009), "*Introduction to semi-supervised learning. Synthesis lectures on artificial intelligence and machine learning*", Vol.3 (1), pp.1-130.

- [41] Kodratoff Y (2014), "*Introduction to machine learning*", Elsevier, pp.1-298.
- [42] Ratsch G (2004), "*A brief introduction into machine learning*", Friedrich Miescher Laboratory of the Max Planck Society, pp.1-6.
- [43] Furey TS, Cristianini N, Duffy N, Bednarski DW, Schummer M and Haussler D (2000), "*Support vector machine classification and validation of cancer tissue samples using microarray expression data*", *Bioinformatics*, Vol.16(10), pp.906-914.
- [44] Soman KP, Diwakar S and Ajay V (2006), "*Data mining: theory and practice*", PHI Learning Pvt. Ltd, pp.1- 486.
- [45] Burges CJ (1998), "*A tutorial on support vector machines for pattern recognition. Data mining and knowledge discovery*", Vol.2 (2), pp.121-167.
- [46] Vapnik VN (1999), "*An overview of statistical learning theory*", *IEEE transactions on neural networks*, Vol.10 (5), pp.988-999.
- [47] Soman KP, Loganathan R and Ajay V (2009), "*Machine Learning with SVM and other Kernel Methods*", PHI, India, pp.1- 486.
- [48] Er MJ, Wu S, Lu J and Toh HL (2002), "*Face recognition with Radial Basis Function neural networks*", *IEEE transactions on neural networks*, Vol.13 (3), pp.697-710.
- [49] Ripka S (1999), "*An algorithm for selecting a good value for the parameter c in Radial Basis Function interpolation*", *Advances in Computational Mathematics*, Vol.11(2-3), pp.193-210.
- [50] Cristianini N and Shawe-Taylor J (2000), "*An introduction to support vector machines and other kernel-based learning methods*", Cambridge university press, pp.1-189.
- [51] Wang L. et.al (2005), "*Support vector machines: theory and applications, Springer Science & Business Media*", pp.1-431.
- [52] Cristianini N and Scholkopf B (2002), "*Support vector machines and kernel methods: the new generation of learning machines*", *Ai Magazine*, Vol.23 (3), pp.31-41.
- [53] Platt J.C, Cristianini N, Shawe-Taylor J (2007), "*Large margin DAG's for Multiclass classification*", *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA, USA, Page 547-553.
- [54] Weston J and Watkins C (1998), "*Multi-class support vector machines*", Technical Report CSD-TR-98-04, Department of Computer Science, Royal Holloway, University of London, Egham, UK.

- [55] Crammer K and Singer Y (2001), "*On the Algorithmic Implementation of Multi-class Kernel-based Vector Machines*", Journal of Machine Learning Research, MIT Press, Cambridge, MA, USA. Vol.2, pp. 265-292.
- [56] Schmidhuber J (2015), "*Deep learning in neural networks: An overview*", Neural networks, vol.61, pp.85-117.
- [57] Yegnanarayana B (2009), "*Artificial neural networks*", PHI Learning Pvt. Ltd, pp.1-476.
- [58] He X and Xu S (2010), "*Artificial neural networks*", *Process Neural Networks: Theory and Applications*, pp.20-42.
- [59] Kişi O (2004), "*River flow modeling using artificial neural networks*", Journal of Hydrologic Engineering, Vol.9 (1), pp.60-63.
- [60] Olden JD, Joy MK and Death RG (2004), "*An accurate comparison of methods for quantifying variable importance in Artificial Neural Networks using simulated data*", Ecological Modelling, Vol.178 (3), pp.389-397.
- [61] Abraham A (2005), "*Artificial neural networks*", Handbook of measuring system design.
- [62] Lawrence S, Giles CL and Tsoi AC(1997), "*Lessons in Neural Network Training: Overfitting May be Harder than Expected*", Proceedings of the Fourteenth National Conference on Artificial Intelligence, AAAI-97, AAAI Press, Menlo Park, California, pp. 540-545.
- [63] LeCun Y, Bengio Y and Hinton G (2015), "*Deep learning*", Nature, Vol.521 (7553), pp.436-444.
- [64] Collobert R and Weston J (2008), "*A unified architecture for natural language processing: Deep neural networks with multitask learning*", In Proceedings of the 25th international conference on Machine learning, pp. 160-167.
- [65] LeCun Y (2015), LeNet-5, "*Convolutional neural networks*", <http://yann.lecun.com/exdb/lenet>.
- [66] Zhang W, Itoh K, Tanida J and Ichioka Y (1990), "*Parallel distributed processing model with local space-invariant interconnections and its optical architecture*", Applied Optics, Vol.29(32), pp. 4790-4797.
- [67] Matsugu M, Mori K, Mitari Y and Kaneda Y (2003), "*Subject independent facial expression recognition with robust face detection using a convolutional neural network*", Neural Networks, Vol.16 (5-6), pp.555-559.

- [68] Van den Oord, A Dieleman S and Schrauwen B (2013), “*Deep content-based music recommendation*”, In advances in neural information processing systems, pp. 2643-2651.
- [69] Aghdam HH and Heravi EJ (2017), “*Guide to convolutional neural networks: a practical application to traffic-sign detection and classification*”, Springer, pp.1- 282.
- [70] Ciresan DC, Meier U, Masci J, Maria Gambardella L and Schmidhuber J (2011), “*Flexible, high performance convolutional neural networks for image classification*”, In IJCAI Proceedings-International Joint Conference on Artificial Intelligence, Vol. 22(1), pp.1237–1242.
- [71] Krizhevsky A, Sutskever I and Hinton GE (2012), “*Imagenet classification with deep convolutional neural networks*”, In advances in neural information processing systems, pp. 1097-1105.
- [72] Kannan RJ and Prabhakar R (2009), “*A Comparative study of Optical Character Recognition for Tamil Script*”, European Journal of Scientific Research, Vol.35 (4), pp.570-582.
- [73] Leng WY and Shamsuddin SM (2010), “*Writer identification for Chinese handwriting*”, International Journal of Advanced Soft Computing Application, Vol.2 (2), pp.142-173.
- [74] Varga T (2006), “*Off-line cursive handwriting recognition using synthetic training data*”, pp.1-124.
- [75] Shahabi F and Rahmati M (2006), “*Comparison of Gabor-based features for writer identification of Farsi/Arabic handwriting*”, In Tenth International Workshop on Frontiers in Handwriting Recognition, Suvisoft, pp.1-7.
- [76] Blankers V, Niels R and Vuurpijl L (2007), “*Writer identification by means of explainable features: shapes of loop and lead-in strokes*”, In Proc. of BNAIC, pp. 17-24.
- [77] Revathi SV and Vijaya MS (2014), “*Predicting the Identity of a Person using Aggregated Features of Handwriting*”, Advances in Image and Video Processing, Vol. 2(6), pp. 23-33.
- [78] Dunn DF and Higgins WE (1993), “*Optimal Gabor-filter design for texture segmentation*”, IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 5, pp. 37-40.
- [79] He Z, Tang YY and You X (2005), “*A contourlet-based method for writer identification*”, IEEE International Conference on Systems, Man and Cybernetics, Vol. 1, pp. 364-368.
- [80] Bulacu M, Schomaker L and Vuurpijl L (2003), “*Writer identification using edge-based directional features*”, 1, pp.1-6.

- [81] Khan S and Lewbel A (2007), “*Weighted and two-stage least squares estimation of semiparametric truncated regression models*”, *Econometric Theory*, Vol.23 (2), pp.309-347.
- [82] Sulaimon Mutiu O (2015), “*Application of Weighted Least Squares Regression in Forecasting*”, *International Journal of Recent Research in Interdisciplinary Sciences (IJRRIS)*, Vol. 2(3), pp.45-54.
- [83] Maura Stokes, Fang Chen, and Funda Gunes (2014), “*An Introduction to Bayesian Analysis with SAS/STAT® Software*”, SAS400-2014.
- [84] Schölkopf B, Smola A and Muller KR (1997), “*Kernel principal component analysis*”, In *International Conference on Artificial Neural Networks*, pp.583-588.
- [85] Selvakumar Raja S and John M (2010), “*A Novel Method of Tamil Character Recognition Using Slope Method*”, *IUP Journal of Science & Technology*, Vol.6 (2), pp. 7-14.
- [86] Bharath A and Madhvanath S (2012), “*HMM-based lexicon-driven and lexicon-free word recognition for online handwritten Indic scripts*”, *IEEE transactions on pattern analysis and machine intelligence*, Vol.34 (4), pp.670-682.
- [87] LeCun Y, Kavukvuoglu K and Farabet C (2010), “*Convolutional networks and applications in vision*”, *International Symposium on Circuits and Systems*, pp. 253–256
- [88] Long J, Shelhamer E and Darrell T (2015), “*Fully convolutional networks for semantic segmentation*”, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431-3440.
- [89] Zeiler MD, Krishnan D, Taylor GW and Fergus R (2010), “*Deconvolutional networks*”, In *Computer Vision and Pattern Recognition*, pp. 2528–2535
- [90] Ms. Sonali, B. Maind, Ms. Priyanka Wankar, “*Research Paper on Basic of Artificial Neural Network*”, *International Journal on Recent and Innovation Trends in Computing and Communication*, ISSN: 2321-8169, Vol.2(1), pp.96-100

LIST OF PUBLICATIONS

Papers published in International Conference Proceedings

1. Supervised Learning Approach for Tamil Writer Identity Prediction Using Global and Local Features, *International Conference on IT and Intelligent Systems (ICITIS'2013)*, *International Conference Proceedings of PSRC, Penang (Malaysia)*, August 28-29, 2013
2. Analysis of Tamil Character Writings and Identification of Writer Using Support Vector Machine, *IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)*, 2014

Papers published/accepted in International Journals

1. Discovering Tamil Writer Identity Using Global and Local Features of Offline Handwritten Text, *International Review on Computers and Software (IRECOS)*, Italy, Vol. 8 No 9, Page 2080 – 2087, 2013
2. Prediction of Writer Using Tamil Handwritten Document Image Based on Pooled Features, *World Academy of Science, Engineering and Technology, International science index*, Vol. 9, No. 6, Page 1481 – 1487, 2015
3. Detection of a Person Using Descriptive Features of Tamil Handwriting and Pattern Learning, *International Journal of Applied Engineering Research*, ISSN 0973-4562, Volume 10, Number 21 (2015) pp 41902-41909, **(Scopus Indexed)**
4. Bayesian Linear Regression Co-efficients for SVM linear kernel to identify writers, *Journal of Advanced Research in Dynamical and Control Systems (JARDCS)*, Vol. 10, 01-Special Issue, 2018, **(Scopus Indexed)**
5. Hybrid Linear Kernel With PCA In SVM Prediction Model of Tamil Writing Pattern, *International Journal of Simulation Systems, Science & Technology*, **(Accepted - Scopus Indexed)**

Papers in review - International Journals

1. Linear Kernel with Weighted Least Square Regression Co-efficient for SVM Based Tamil Writer Identification, *Malaysian Journal of Computer Science (MJCS)*
2. An Efficient Convolutional Neural Network Based Classifier to Predict Tamil Writer, *International Journal of Technology Intelligence and planning*

Appendix - A
Sample Tamil Handwritings (Paragraphs)

நீங்கள் என்ன நினைக்கிறீர்களோ அது நடக்கிறது
என்பது நினும். திடீரென நிகழ்கின்றதென்றே அனைத்து
பற்றிய உங்கள் எண்ண அலைகள் மூலியது போகாமல்
அந்நேரம் கொண்டு இந்நேரம், உங்கள் எண்ணம்
நினைவேனும்.

நினைவாகவே நினைவிழ் பணிகளாகிணை எப்படி
உருவாக்கிவந்து என்மதநீ நன் கொண்டு அனுபவங்கள்
நிலம் கண்டு கொண்டுள்ள பானம். இவா இந்த அளப்பாட
அந்நேரம் கண்டுநீ கண்டு நிறுபுதற்கு கண்டுபிடிக்கவேண்டி
பழையபாட இந்நேரம் இரகசியம் அருகிடு தெரிய
உந்தது சிற உருபுகளாகிடு முன் நமநீ சம்பவம் அது
பணம் வேண்டும் என்கு அனை பற்றிய அந்நேரம்
தினமும் கிண்கு தடவை சந்தம் போடும் பழக்க வேண்டும்
ஏற்கெனவே பணம் கைகிடு உந்தாயிற்று என்மதாகம்
பாவிதீது உணர் வேண்டும்.
ஏதாவது பொருணை அகைய உருமீயினால்
நமது செயல். நினைவு ஆகிய நினைவிகள் அனைந்தும்
அனைப்பற்றியதாகவே இருக்க வேண்டும்.

Appendix – B

Sample Datasets

Normalized dataset for character level writer identification

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	1	01:02.5	02:01.1	03:01.2	04:01.3	5:01	06:01.3	07:01.1	8:01	9:01	10:01	11:01	12:01	13:01	14:04.2	15:01.4	16:01.0	17:01.0	18:01.0	19:01.0	20:01.0
2	1	01:01.4	02:01.1	03:01.1	04:01.6	5:01	06:01.5	07:01.6	8:01	9:01	10:01	11:01	12:01	13:01	14:03.8	15:01.6	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
3	1	01:02.5	02:01.7	03:01.6	04:02.0	5:01	06:02.1	07:02.4	8:01	9:01	10:01	11:01	12:01	13:01	14:04.3	15:01.4	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
4	1	01:01.4	02:01.5	03:01.2	04:02.0	5:01	06:02.0	07:03.4	8:01	9:01	10:01	11:01	12:01	13:01	14:04.6	15:01.8	16:01.2	17:01.0	18:01.0	19:01.1	20:01.0
5	1	01:01.7	02:01.3	03:01.2	04:01.5	5:01	06:01.5	07:01.3	8:01	9:01	10:01	11:01	12:01	13:01	14:04.5	15:01.5	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
6	1	01:02.1	02:02.7	03:02.3	04:01.6	5:01	06:01.5	07:02.8	8:01	9:01	10:01	11:01	12:01	13:01	14:03.5	15:01.6	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
7	1	01:01.4	02:01.1	03:01.1	04:01.9	5:01	06:01.8	07:01.4	8:01	9:01	10:01	11:01	12:01	13:01	14:04.1	15:01.6	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
8	1	01:02.1	02:01.4	03:01.3	04:01.8	05:01.7	06:01.8	07:01.3	8:01	9:01	10:01	11:01	12:01	13:01	14:04.5	15:01.4	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
9	1	01:01.7	02:01.1	03:01.1	04:01.5	5:01	06:01.5	07:02.0	8:01	9:01	10:01	11:01	12:01	13:01	14:03.4	15:02.0	16:01.2	17:01.0	18:01.0	19:01.1	20:01.0
10	1	01:02.8	02:01.4	03:01.5	04:01.0	5:01	6:01	07:01.7	8:01	9:01	10:01	11:01	12:01	13:01	14:03.8	15:01.7	16:01.1	17:01.0	18:01.0	19:01.1	20:01.0
11	1	01:02.8	02:01.1	03:01.2	04:02.3	05:01.7	06:02.4	07:02.7	8:01	9:01	10:01	11:01	12:01	13:01	14:04.8	15:01.6	16:01.1	17:01.0	18:01.0	19:01.1	20:01.0
12	1	01:03.9	02:01.4	03:01.7	04:01.9	5:01	06:01.7	07:01.5	8:01	9:01	10:01	11:01	12:01	13:01	14:01.9	15:03.3	16:01.8	17:01.3	18:01.1	19:01.1	20:01.0
13	1	01:01.7	02:01.3	03:01.2	04:01.7	05:02.3	06:01.6	07:01.5	8:01	9:01	10:01	11:01	12:01	13:01	14:04.4	15:01.5	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
14	1	01:02.5	02:01.1	03:01.2	04:01.8	05:01.7	06:01.8	07:02.4	8:01	9:01	10:01	11:01	12:01	13:01	14:04.3	15:01.7	16:01.1	17:01.0	18:01.0	19:01.1	20:01.0
15	1	01:02.5	02:01.4	03:01.4	04:02.4	5:01	06:02.3	07:01.6	8:01	9:01	10:01	11:01	12:01	13:01	14:04.1	15:01.6	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
16	1	01:02.5	02:01.1	03:01.2	04:01.8	5:01	06:01.9	07:02.1	8:01	9:01	10:01	11:01	12:01	13:01	14:04.3	15:01.8	16:01.1	17:01.0	18:01.0	19:01.1	20:01.0
17	1	01:01.7	2:01	03:01.0	04:01.4	5:01	06:01.4	07:02.1	8:01	9:01	10:01	11:01	12:01	13:01	14:04.6	15:01.6	16:01.1	17:01.0	18:01.0	19:01.1	20:01.0
18	1	01:02.5	02:01.3	03:01.3	04:01.3	5:01	06:01.2	07:01.6	8:01	9:01	10:01	11:01	12:01	13:01	14:02.1	15:02.1	16:01.2	17:01.0	18:01.0	19:01.1	20:01.0
19	1	01:02.5	02:01.3	03:01.3	04:01.5	05:01.7	06:01.5	07:01.3	8:01	9:01	10:01	11:01	12:01	13:01	14:04.1	15:01.3	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
20	1	01:03.2	02:01.3	03:01.4	04:02.2	5:01	06:02.1	07:01.7	8:01	9:01	10:01	11:01	12:01	13:01	14:04.2	15:01.6	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
21	1	01:02.8	02:01.3	03:01.4	04:02.6	05:02.3	06:02.5	07:02.0	8:01	9:01	10:01	11:01	12:01	13:01	14:04.8	15:01.5	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
22	1	01:02.5	2:01	03:01.1	04:02.6	05:02.3	06:02.6	07:03.9	8:01	9:01	10:01	11:01	12:01	13:01	14:04.7	15:01.8	16:01.2	17:01.0	18:01.0	19:01.1	20:01.0
23	1	01:03.2	02:01.4	03:01.6	04:02.0	5:01	06:01.9	07:02.0	8:01	9:01	10:01	11:01	12:01	13:01	14:03.5	15:01.7	16:01.1	17:01.0	18:01.0	19:01.1	20:01.0
24	1	01:03.2	02:01.1	03:01.3	04:01.8	5:01	06:01.7	07:02.7	8:01	9:01	10:01	11:01	12:01	13:01	14:02.3	15:03.8	16:03.0	17:01.3	18:01.5	19:01.1	20:01.0
25	1	01:02.5	02:01.3	03:01.3	04:01.3	5:01	06:01.3	07:01.7	8:01	9:01	10:01	11:01	12:01	13:01	14:04.2	15:01.4	16:01.0	17:01.0	18:01.0	19:01.1	20:01.0
26	1	01:02.8	02:01.3	03:01.4	04:01.8	5:01	06:01.6	07:01.4	8:01	9:01	10:01	11:01	12:01	13:01	14:02.5	15:02.0	16:01.1	17:01.0	18:01.0	19:01.1	20:01.0

Normalized dataset for word level writer identification

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	1	01:01.3	02:01.5	03:01.0	04:01.4	05:02.6	06:01.5	07:03.1	08:02.0	09:01.1	10:01.0	11:02.0	12:02.1	13:03.1	14:03.2	15:02.9	16:02.2	17:01.3	18:01.3	19:01.0	20:01.0
2	1	01:01.8	2:01	03:01.0	04:01.4	05:02.2	06:01.4	07:02.2	08:02.0	09:01.1	10:01.0	11:01.8	12:02.0	13:03.0	14:03.1	15:02.4	16:01.6	17:01.3	18:01.2	19:01.0	20:01.0
3	1	01:01.2	2:01	03:01.0	04:01.3	05:02.4	06:01.3	07:02.2	08:01.9	09:01.1	10:01.0	11:01.8	12:02.0	13:03.0	14:03.3	15:02.2	16:01.5	17:01.0	18:01.0	19:01.0	20:01.0
4	1	01:02.0	02:01.4	03:01.1	04:01.3	05:02.2	06:01.3	07:01.6	08:02.0	09:01.2	10:01.0	11:01.9	12:01.8	13:03.2	14:02.7	15:01.6	16:01.2	17:01.0	18:01.0	19:01.0	20:01.0
5	1	01:01.5	02:01.3	03:01.0	04:01.1	05:01.2	06:01.2	07:01.3	08:01.5	09:01.9	10:01.0	11:01.6	12:01.1	13:03.3	14:02.5	15:01.6	16:01.1	17:01.2	18:01.0	19:01.0	20:01.0
6	1	01:01.5	02:01.5	03:01.1	04:01.3	05:01.6	06:01.3	07:01.7	08:02.1	09:01.1	10:01.0	11:02.1	12:02.6	13:03.0	14:02.4	15:02.6	16:01.9	17:01.0	18:01.0	19:01.0	20:01.0
7	1	01:01.1	02:01.4	03:01.0	04:01.2	05:01.6	06:01.2	07:02.1	08:01.7	09:01.2	10:01.0	11:01.7	12:01.9	13:03.2	14:02.8	15:02.0	16:01.3	17:01.0	18:01.0	19:01.0	20:01.0
8	1	01:01.1	02:01.2	03:01.0	04:01.2	05:01.7	06:01.3	07:02.5	08:01.8	09:01.1	10:01.0	11:01.8	12:02.3	13:03.7	14:02.9	15:02.5	16:01.7	17:01.1	18:01.0	19:01.0	20:01.0
9	1	01:01.2	02:01.6	03:01.0	04:01.2	05:01.6	06:01.3	07:01.6	08:01.9	09:01.2	10:01.0	11:01.9	12:01.8	13:03.3	14:02.4	15:01.8	16:01.2	17:01.0	18:01.0	19:01.0	20:01.0
10	1	01:01.2	02:02.0	03:01.1	04:01.3	05:02.1	06:01.3	07:02.3	08:01.6	09:01.2	10:01.0	11:01.6	12:01.1	13:03.2	14:02.9	15:02.2	16:01.5	17:01.0	18:01.0	19:01.0	20:01.0
11	1	01:01.5	02:01.2	03:01.0	04:01.2	05:02.0	06:01.3	07:02.0	08:01.8	09:01.1	10:01.0	11:01.8	12:02.0	13:03.0	14:03.1	15:02.3	16:01.5	17:01.0	18:01.0	19:01.0	20:01.0
12	1	01:02.0	02:01.6	03:01.2	04:01.3	05:01.8	06:01.3	07:02.1	08:01.8	09:01.1	10:01.0	11:01.7	12:01.6	13:03.1	14:02.4	15:02.2	16:01.5	17:01.1	18:01.1	19:01.0	20:01.0
13	1	01:01.5	02:01.8	03:01.1	04:01.2	05:01.6	06:01.2	07:01.8	08:01.8	09:01.1	10:01.0	11:01.8	12:01.9	13:03.0	14:02.4	15:01.8	16:01.2	17:01.2	18:01.1	19:01.0	20:01.0
14	1	01:01.6	02:01.8	03:01.2	04:01.2	05:01.3	06:01.2	07:01.6	08:02.0	09:01.1	10:01.0	11:02.0	12:03.8	13:02.5	14:01.9	15:01.9	16:01.3	17:01.1	18:01.0	19:01.0	20:01.0
15	1	01:01.7	02:01.4	03:01.1	04:01.4	05:01.7	06:01.4	07:01.8	08:02.0	09:01.2	10:01.0	11:02.0	12:01.7	13:03.2	14:02.2	15:02.4	16:01.6	17:01.6	18:01.5	19:01.1	20:01.0
16	1	01:01.7	02:01.6	03:01.1	04:01.2	05:01.6	06:01.2	07:01.4	08:01.8	09:01.2	10:01.0	11:01.8	12:02.0	13:02.9	14:02.4	15:01.6	16:01.1	17:01.0	18:01.0	19:01.0	20:01.0
17	1	01:02.1	02:01.5	03:01.2	04:01.2	05:01.7	06:01.2	07:01.4	08:01.8	09:01.2	10:01.0	11:01.8	12:01.4	13:03.3	14:02.7	15:01.4	16:01.1	17:01.0	18:01.0	19:01.0	20:01.0
18	1	01:01.5	02:01.4	03:01.1	04:01.2	05:01.6	06:01.2	07:01.4	08:01.8	09:01.2	10:01.0	11:01.8	12:02.2	13:03.0	14:02.9	15:01.4	16:01.1	17:01.0	18:01.0	19:01.0	20:01.0
19	1	01:01.7	02:01.5	03:01.1	04:01.2	05:01.5	06:01.3	07:02.0	08:01.7	09:01.1	10:01.0	11:01.6	12:01.4	13:03.4	14:02.3	15:02.6	16:01.8	17:01.0	18:01.0	19:01.0	20:01.0
20	1	01:01.9	02:01.9	03:01.3	04:01.2	05:01.2	06:01.2	07:01.5	08:01.6	09:01.2	10:01.0	11:01.6	12:01.5	13:03.4	14:02.2	15:01.4	16:01.1	17:01.0	18:01.0	19:01.0	20:01.0
21	1	01:01.5	02:01.4	03:01.1	04:01.3	05:01.7	06:01.3	07:01.9	08:02.0	09:01.2	10:01.0	11:02.0	12:02.3	13:03.1	14:02.4	15:02.4	16:01.6	17:01.0	18:01.0	19:01.0	20:01.0
22	1	1:01	02:01.5	03:01.0	04:01.5	05:02.2	06:01.5	07:01.7	0												

Normalized dataset for paragraph level writer identification

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.2	09:04.9	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8
2	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.2	09:04.8	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8
3	1	01:01.3	02:01.3	03:01.4	04:01.0	05:01.4	06:01.0	07:01.4	08:01.3	09:04.6	10:04.7	11:01.3	12:01.3	13:04.7	14:04.7	15:04.7	16:04.7	17:04.7	18:04.7	19:04.7	20:04.7
4	1	01:01.5	02:01.5	03:01.0	04:01.0	05:01.0	06:01.0	07:01.4	08:01.5	09:04.6	10:04.5	11:01.5	12:01.5	13:04.5	14:04.5	15:04.5	16:04.5	17:04.5	18:04.5	19:04.5	20:04.5
5	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.2	09:04.8	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8
6	1	01:01.1	02:01.1	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.1	09:04.9	10:04.9	11:01.1	12:01.1	13:04.9	14:04.9	15:04.9	16:04.9	17:04.9	18:04.9	19:04.9	20:04.9
7	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.2	09:04.8	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8
8	1	01:01.3	02:01.3	03:01.0	04:01.0	05:01.0	06:01.0	07:01.3	08:01.3	09:04.7	10:04.7	11:01.3	12:01.3	13:04.7	14:04.7	15:04.7	16:04.7	17:04.7	18:04.7	19:04.7	20:04.7
9	1	01:01.3	02:01.3	03:01.0	04:01.0	05:01.0	06:01.0	07:01.3	08:01.3	09:04.7	10:04.7	11:01.3	12:01.3	13:04.7	14:04.7	15:04.7	16:04.7	17:04.7	18:04.7	19:04.7	20:04.7
10	1	01:01.5	02:01.6	03:01.4	04:01.0	05:01.4	06:01.0	07:01.8	08:01.6	09:04.3	10:04.4	11:01.5	12:01.6	13:04.4	14:04.4	15:04.5	16:04.4	17:04.5	18:04.4	19:04.4	20:04.4
11	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.2	08:01.2	09:04.8	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8
12	1	01:01.4	02:01.4	03:01.0	04:01.0	05:01.0	06:01.0	07:01.3	08:01.4	09:04.7	10:04.6	11:01.4	12:01.4	13:04.6	14:04.6	15:04.6	16:04.6	17:04.6	18:04.6	19:04.6	20:04.6
13	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.2	09:04.9	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8
14	1	01:01.3	02:01.3	03:01.0	04:01.0	05:01.0	06:01.0	07:01.2	08:01.3	09:04.8	10:04.7	11:01.3	12:01.3	13:04.7	14:04.7	15:04.7	16:04.7	17:04.7	18:04.7	19:04.7	20:04.7
15	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.2	09:04.9	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8
16	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.2	09:04.9	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8
17	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.2	09:04.8	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8
18	1	01:01.3	02:01.3	03:01.4	04:01.0	05:01.4	06:01.0	07:01.4	08:01.3	09:04.6	10:04.7	11:01.3	12:01.3	13:04.7	14:04.7	15:04.7	16:04.7	17:04.7	18:04.7	19:04.7	20:04.7
19	1	01:01.5	02:01.5	03:01.0	04:01.0	05:01.0	06:01.0	07:01.4	08:01.5	09:04.6	10:04.5	11:01.5	12:01.5	13:04.5	14:04.5	15:04.5	16:04.5	17:04.5	18:04.5	19:04.5	20:04.5
20	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.2	09:04.8	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8
21	1	01:01.1	02:01.1	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.1	09:04.9	10:04.9	11:01.1	12:01.1	13:04.9	14:04.9	15:04.9	16:04.9	17:04.9	18:04.9	19:04.9	20:04.9
22	1	01:01.2	02:01.2	03:01.0	04:01.0	05:01.0	06:01.0	07:01.1	08:01.2	09:04.8	10:04.8	11:01.2	12:01.2	13:04.8	14:04.8	15:04.8	16:04.8	17:04.8	18:04.8	19:04.8	20:04.8

Normalized dataset for Paragraph level writer identification – 422 features

	QQ	OR	OS	OT	OU	OV	OW	OX	OY	PA	PB	PC	PD	PE	PF	PG
1	00406:2.5035	00407:3.9774	00408:3.209	00409:1.9912	00410:1.5333	00411:2.9528	00412:4.4063	00413:3.1351	00414:3	00416:2.7231	00417:2.9209	00418:2.4933	00419:2.6512	00420:2.0488	00421:3.2822	00422:3.5
2	00406:2.6454	00407:3.5564	00408:1.806	00409:3.3009	00410:2.28	00411:3.2362	00412:5	00413:3.1892	00414:1	00416:3.4	00417:2.6271	00418:1.8533	00419:3.0465	00420:3.0488	00421:2.9387	00422:3.0
3	00406:2.4468	00407:2.7143	00408:2.7015	00409:2.6283	00410:3.16	00411:2.4803	00412:1.8438	00413:2.973	00414:2.8734	00416:3.1538	00417:2.1977	00418:2.7333	00419:3.2093	00420:3.0244	00421:1	00422:2.7
4	00406:2.5603	00407:3.3158	00408:2.1343	00409:4.8938	00410:3.1333	00411:2.0394	00412:1.1563	00413:3.4865	00414:3.8608	00416:2.9077	00417:1.2938	00418:2.8133	00419:2.6512	00420:2.8537	00421:3.5031	00422:3.4
5	00406:2.6738	00407:4.5789	00408:2.8209	00409:2.0265	00410:1	00411:4.1181	00412:3.75	00413:2.3514	00414:2.4177	00416:4.4462	00417:1.7684	00418:2.1733	00419:3.1395	00420:3.5366	00421:3.8712	00422:2.6
6	00406:2.5887	00407:2.1128	00408:2.1045	00409:3.1239	00410:2.5733	00411:2.2283	00412:3.0313	00413:3.6486	00414:2.8987	00416:1.8923	00417:2.1977	00418:3.0533	00419:2.3953	00420:2.6585	00421:3.1595	00422:4.0
7	00406:2.305	00407:3.6466	00408:1.9851	00409:2.3451	00410:1.9067	00411:3.5197	00412:3.9688	00413:2.2432	00414:2.6203	00416:3.1538	00417:1.8136	00418:1.48	00419:3.5581	00420:3.0244	00421:3.7485	00422:2.8
8	00406:2.9291	00407:2.5038	00408:1.597	00409:2.6637	00410:3.88	00411:1.9764	00412:2.4688	00413:3	00414:2.0633	00416:2.4769	00417:3.6441	00418:3.24	00419:2.3488	00420:1.8049	00421:3.184	00422:2.7
9	00406:2.5887	00407:3.8872	00408:3.2985	00409:2.8053	00410:3.0267	00411:1.4409	00412:1.8438	00413:2.3514	00414:2.152	00416:3.6462	00417:3.1695	00418:1.9867	00419:1.9302	00420:3.5366	00421:3.0123	00422:2.2
10	00406:2.5035	00407:2.8346	00408:3.6567	00409:3.3009	00410:3.8267	00411:1.7559	00412:1.3438	00413:3.4595	00414:2.0127	00416:3.5231	00417:2.2881	00418:3.8267	00419:2.4884	00420:2.0976	00421:4.1411	00422:3.8
11	00406:2.4468	00407:1.4812	00408:2.3134	00409:2.7345	00410:2.6267	00411:4.4331	00412:3.8438	00413:2.0541	00414:3.1772	00416:1.6462	00417:2.8531	00418:4.12	00419:1.4419	00420:2.2927	00421:3.3067	00422:3.4
12	00406:3.6099	00407:3.8872	00408:1.7761	00409:1.8142	00410:2.8933	00411:2.7638	00412:2.4063	00413:4.2162	00414:3.2532	00416:2.7538	00417:2.2429	00418:2.92	00419:3.3721	00420:1.9512	00421:3.4785	00422:3.1
13	00406:2.9291	00407:2.6541	00408:1.209	00409:3.2301	00410:3.7467	00411:2.1339	00412:2.5938	00413:2.9189	00414:3.2025	00416:2.9077	00417:3.1469	00418:3.0533	00419:1.7442	00420:1.8293	00421:4.0675	00422:2.8
14	00406:2.5035	00407:3.6466	00408:3.0299	00409:3.6903	00410:1.9333	00411:1.063	00412:3.2813	00413:3.5946	00414:1.6329	00416:3.8308	00417:1.9944	00418:2.36	00419:2.6977	00420:3.7805	00421:3.0123	00422:2.5
15	00406:2.6738	00407:3.0752	00408:1.4179	00409:2.7699	00410:5	00411:3.0157	00412:1.75	00413:1.7027	00414:3.5063	00416:1.8615	00417:1.9492	00418:3.8	00419:3.8837	00420:2.439	00421:3.135	00422:2.6
16	00406:2.5035	00407:3.9774	00408:3.209	00409:1.9912	00410:1.5333	00411:2.9528	00412:4.4063	00413:3.1351	00414:3	00416:2.7231	00417:2.9209	00418:2.4933	00419:2.6512	00420:2.0488	00421:3.2822	00422:3.5
17	00406:2.6454	00407:3.5564	00408:1.806	00409:3.3009	00410:2.28	00411:3.2362	00412:5	00413:3.1892	00414:1	00416:3.4	00417:2.6271	00418:1.8533	00419:3.0465	00420:3.0488	00421:2.9387	00422:3.0
18	00406:2.4468	00407:2.7143	00408:2.7015	00409:2.6283	00410:3.16	00411:2.4803	00412:1.8438	00413:2.973	00414:2.8734	00416:3.1538	00417:2.1977	00418:2.7333	00419:3.2093	00420:3.0244	00421:1	00422:2.7
19	00406:2.5603	00407:3.3158	00408:2.1343	00409:4.8938	00410:3.1333	00411:2.0394	00412:1.1563	00413:3.4865	00414:3.8608	00416:2.9077	00417:1.2938	00418:2.8133	00419:2.6512	00420:2.8537	00421:3.5031	00422:3.4
20	00406:2.6738	00407:4.5789	00408:2.8209	00409:2.0265	00410:1	00411:4.1181	00412:3.75	00413:2.3514	00414:2.4177	00416:4.4462	00417:1.7684	00418:2.1733	00419:3.1395	00420:3.5366	00421:3.8712	00422:2.6
21	00406:2.5887	00407:2.1128	00408:2.1045	00409:3.1239	00410:2.5733	00411:2.2283	00412:3.0313	00413:3.6486	00414:2.8987	00416:1.8923	00417:2.1977	00418:3.0533	00419:2.3953	00420:2.6585	00421:3.1595	00422:4.0
22	00406:2.305	00407:3.6466	00408:1.9851	00409:2.3451	00410:1.9067	00411:3.5197	00412:3.9688	00413:2.2432	00414:2.6203	00416:						

Appendix - C

Sample Code

Pre-processing – Character and Word Text Images

```
%reading the image
i=imread('D:\PhD\matlab file\DATASET\K001.jpg');
%converting rgb image into gray scale
c=rgb2gray(i);
%imshow(c);title('GrayScale image');
%Global thresholding the gray level image
level=graythresh(c);
%convert gray scale image to binary
bw=im2bw(c,level);
%thining
%p1=imcomplement(bw);
%new=bwmorph(p1,'thin',1);
%g=imcomplement(new);
%Removal of noise
n=medfilt2(c);
%edge detection using sobel filters
so=edge(c,'sobel');
%dilation of the image
se = strel('line',1,1);
di = imdilate(bw,se);
%calling function "No of black pixel"
[m]= myennoblkSize(so);
%n=sprintf('No of black pixel: %d',m);
%disp(n);
disp(m);
%calling Function "length"
[len]=myenlength(so);
%disp(len)
%calling fuction "edge_x"
[vertical_edges] = edge_x(so);
%disp(so);
%calling function "edge hinge"
[eh] = edgehinge(so, 4);
disp(eh);
%label conn comp
l = bwlabel(so,8);
disp(l);
%number of rows & col
[r c]=size(so);
disp(r); disp(c);
```

```

%skew angle, slope angle etc...
word = preProcessing(i);
disp(word);
subplot(2,3,1);
imshow(i); title('Original image');
subplot(2,3,2);
imshow(bw);title('Binary image');
subplot(2,3,3);
imshow(n);title('Noiseless image');
subplot(2,3,4);
imshow(so);title('Edge Detection');
subplot(2,3,5);
imshow(di);title('Dilation');

```

Pre-processing – Paragraph Text Images

```

clc;
clear all;
close all;
warning off all;
[file,path]=uigetfile('*.jpg','Select an image');
i=imread(strcat(path,file));
figure('name','Original image','numbertitle','off');
imshow(i);
imm=rgb2gray(i);
i1 = rgb2gray(i);
i2 = edge(i1,'canny',0.1);
figure('name','Closed image ','numbertitle','off');
%figure('name','Edge image ','numbertitle','off');
imshow(i2);
% trthh
se = strel('square',2);
i3 = imdilate(i2,se);
figure('name','Filtered image','numbertitle','off');
%figure('name','Dilate image','numbertitle','off');
imshow(i3);
i4 = imfill(i3,'holes');
%imshow(i3);
s1 = regionprops(i4, 'BoundingBox');
i5=i4;
figure;imshow(i4);
for i=1:size(i4,1)
    for j=1:size(i4,2)
        if i4(i,j)==1
            for i1=1:5
                i5(i,j+i1)=1;
            end
        end
    end
end

```

```

        end
    end
end
figure;imshow(i4);title('Box Bounding Image');
hold on;
s = regionprops(i5, 'BoundingBox');
for i=1:size(s,1)
    t1=s(i).BoundingBox;
    rectangle('Position',round( s(i).BoundingBox),'LineWidth',2,'Edgecolor','b');
    hold on;
    pause(0.1)
end
subplot(2,3,1);
imshow(i),title('Original image');
subplot(2,3,2);
imshow(i3); title('Closed image');
subplot(2,3,3);
imshow(i2);title('Filtered image');
subplot(2,3,4);
imshow(i4);title('Box Bounding Image');
subplot(2,3,5);
imshow(i5);title('Final normalised Text lines image');

```

Feature Extraction – Character and Word Text Images

```

myFolder = 'D:\PhD\matlab file\DATASET\';
jpegFiles = dir('D:\PhD\matlab file\DATASET\*.jpg');
for k = 1:length(jpegFiles)
    baseFileName = jpegFiles(k).name;
    fullFileName = fullfile(myFolder, baseFileName);
    i= imread(fullFileName);
    %pause(0.5);
    %imshow(i); % Display image.
    %drawnow; % Force display to update immediately.
    %converting rgb image into gray scale
    c=rgb2gray(i);
    %imshow(c);title('GrayScale image');
    %Removal of noise
    n=medfilt2(c);
    %Global thresholding the gray level image
    level=graythresh(n);
    %convert gray scale image to binary
    bw=im2bw(n,level);
    %thining
    p1=imcomplement(bw);
    new=bwmorph(p1,'thin',1);
    thi=imcomplement(new);

```

```

%edge detection using sobel filters
so=edge(c,'sobel');
%dilation of the image
se = strel('line',1,1);
di = imdilate(bw,se);
%Length Of an image
[r c]= size(so);
len=0;
l1=c;
l2=0;
for j=1:c
    for k=1:r
        if so(k,j)==1
            if(k<l1)
                l1=k;
            end
        end
        %if b(j,k)==1 && b(j,k+1)==0
        % l2=k;
    %end

    end
end
for j=c:1
    for k=r:1
        if so(k,j)==1
            if(k>l2)
                l2=k;
            end
        end
        %if b(j,k)==1 && b(j,k+1)==0
        % l2=k;
    %end

    end
end
len=l1-l2;
%disp(len);
%Height of an image
[r c]= size(n);
hei=0;
l1=c;
l2=0;
for j=1:r
    for k=1:c
        if so(j,k)==1
            if(k<l1)

```

```

        l1=k;
        end
    end
    %if b(j,k)==1 && b(j,k+1)==0
    % l2=k;
    %end
end
end
for j=r:1
    for k=c:1
        if so(j,k)==1
            if(k>l2)
                l2=k;
            end
        end
        %if b(j,k)==1 && b(j,k+1)==0
        % l2=k;
        %end
    end
end
    hei=l1-l2;
    %disp(hei);
%area
a=len*hei;
%disp(a);
%number of black pixel
[m]= myennoblkSize(so);
%str=sprintf('\nNumber Of black Pixel: %d',m);
%disp(m);
%entropy of gray values
ent= entropy(so);
%str=sprintf('\tEntropy of gray value: %f ',ent);
%disp(ent);
%moment invaraint
%[M]=feature_vec(so);
% mome=sprintf('\n\tMoment invariant: %f',M);
%disp(mome);
[M1 M2 M3 M4 M5 M6 M7]=feature_vec(so);
%disp(M1);disp(M2); disp(M3);disp(M4);disp(M5);disp(M6);disp(M7);
%loops
[loops, pixels] = getLoops(so);
%disp(loops);
%junction
[junctions, pixels] = getJunctions(thi);
%disp(junctions);
%aspect ratio

```



```

ap=c/r;
%disp(ap);
lhbemv=sprintf('%d %d %d %d %d %8.4f %8.4f %8.4f %8.4f %8.4f %8.4f %8.4f %8.4f
%8.4f',len,hei,a,loops,junctions,ap,ent,M1,M2,M3,M4,M5,M6,M7);
%disp(lhbemv);
%slant image,slope etc..
%[word] = preProcessing(i);
%disp(word);
%edge hinge
eh = edgehinge(so, 4);
%disp(eh);
%to write into notepad
%sm=0;cl=4;
%fileid=fopen('Z:\mphil\matlab file\Datasets\trail.dat','a');
%fprintf(fileid,'%d %d:%d %d:%d %d:%d %d:%d %d:%d %d:%0.4f %d:%0.4f %d:%0.4f
%d:%0.4f %d:%0.4f %d:%0.4f %d:%0.4f %d:%0.4f\r\n',cl,sm+1,len, sm+2,hei,
sm+3,a, sm+4,loops, sm+5, junctions, sm+6, ap, sm+7,ent, sm+8,M1, sm+9,M2, sm+10,M3,
sm+11,M4, sm+12,M5, sm+13,M6, sm+14,M7);
%fclose(fileid);
axa=[num2str(len) ',' num2str(hei) ',' num2str(a) ',' num2str(m) ',' num2str(loops) ','
num2str(junctions) ',' num2str(ap) ',' num2str(word.lowerBaseline) ','
num2str(word.upperBaseline) ',' num2str(word.ascenderBaseline) ','
num2str(word.descenderBaseline) ',' num2str(word.angleSlope) ',' num2str(word. angleSlant) ','
num2str(ent) ',' num2str(M1) ',' num2str(M2) ',' num2str(M3) ',' num2str(M4) ',' num2str(M5) ','
num2str(M6) ',' num2str(M7)];
ReportGeneratorAdd(axa,'D:\PhD\matlab file\DATASET\bn.csv');
%axa=[horzcat(eh)];
%ReportGeneratorAdd(axa,'Z:\mphil\matlab file\Datasets\hori.arff');
end

```

Feature Extraction – Paragraph Text Images

```

clc;
clear all;
close all;
warning off all;
[file,path]=uigetfile('10 writer 15 each\*.jpg','Select an image');
i=imread(strcat(path,file));
figure('name','Original image','numbertitle','off')
%imshow(i);
imm=rgb2gray(i);
i1 = rgb2gray(i);
imor=imresize(i1,[256,256]);
i2 = edge(i1,'canny',0.1);
figure('name','Edge image image','numbertitle','off')
imshow(i2)

```

```

% trthh
se = strel('square',2);
i3 = imdilate(i2,se);
figure('name','Dilate image','numbertitle','off')
imshow(i3)
i4 = imfill(i3,'holes');
imshow(i3);
s1 = regionprops(i4, 'BoundingBox');
i5=i4;
figure;imshow(i4);
for i=1:size(i4,1)
    for j=1:size(i4,2)
        if i4(i,j)==1
            for i1=1:30
                i5(i,j+i1)=1;
            end
        end
    end
end
figure;imshow(i5);
figure;imshow(i4);
hold on;
s = regionprops(i5, 'BoundingBox');
t2=1;
for i=1:size(s,1)
    t1=s(i).BoundingBox;
    rectangle('Position',round( s(i).BoundingBox),'LineWidth',2,'Edgecolor','b');
    hold on;
    pause(0.1)
    y(i)=round(t1(2));
end
y1=sort(y);
for i=1:size(y1,2)
    y2=find(y==y1(i));
    for j=1:size(y2,2)
        y3(i+(j-1),:)=round( s(y2(j)).BoundingBox);
        i=i+(j-1);
    end
end
tmp=1;
for i=1:size(s,1)
    t1=y3(i,:);
    if t1(4)>30
        tm2(tmp,:)=t1;
        tmp=tmp+1;
    end
end

```

```

end
figure;imshow(i4);
hold on;
s = regionprops(i5, 'BoundingBox');
for i=1:size(tm2,1)
%   t1=s(i).BoundingBox;
    rectangle('Position',tm2(i,:), 'LineWidth',2,'Edgecolor','b');
    hold on;
    pause(0.1)
%   y(i)=round(t1(2));
end
k=imor;
k1=double(k);
img=imHistogram(k1);
GLCM2 = graycomatrix(k1,'Offset',[2 0;0 2]);
out = GLCM_Features1(GLCM2,0);
f1=(out.contr);
f2=(out.corr);
f3=(out.corrp);
f4=(out.cprom);
f5=(out.cshad);
f6=(out.dissi);
f7=(out.energ);
f8=(out.entro);
f9=(out.homom);
f10=(out.homop);
f11=(out.maxpr);
f12=(out.sosvh);
f13=(out.savgh);
f14=(out.svarh);
f15=(out.senth);
f16=(out.dvarh);
f17=(out.denth);
f18=(out.inf1h);
f19=(out.inf2h);
f20=(out.homom);
f21=(out.indnc);
f22=(out.idmnc);
f23=(out.autoc);
X=[f1,f2,f3,f4,f5,f6,f7,f9,f10,f11,f12,f13,f14,f15,f16,f17,f18,f19,f20,f21,f22,f23];
img_out_disp=mean(gabor(k));
C = mean(ContCode(k,2));% directional features
d=GGD16(k);
pfilt = '9-7';
dfilt = 'pkva';
nlevs = [0, 0, 4, 4, 5];

```

```

kr=imresize(i1,[512,512]);
y = pdfbdec(double(kr), pfilt, dfilt, nlevs);
c1=GGD16(y{1});
test=[X,img_out_disp,C];
on=1;
if on==1
for it=1:150
    img =imread(strcat('10 writer 15 each\',int2str(it),'.jpg'));
    i=imresize(img,[256 256]);
    imm=rgb2gray(i);
i1 = rgb2gray(i);
imor=i1;
i2 = edge(i1,'canny',0.1);
se = strel('square',2);
i3 = imdilate(i2,se);
i4 = imfill(i3,'holes');
s1 = regionprops(i4, 'BoundingBox');
i5=i4;
% figure;imshow(i4);
%figure;imshow(i4);
for i=1:size(i4,1)
    for j=1:size(i4,2)
        if i4(i,j)==1
            for i1=1:30
                i5(i,j+i1)=1;
            end
        end
    end
end
end
%figure;imshow(i5);
%figure;imshow(i4);
k=imor;
k1=double(k);
img=imHistogram(k1);
GLCM2 = graycomatrix(k1,'Offset',[2 0;0 2]);
out = GLCM_Features1(GLCM2,0);
f1=(out.contr);
f2=(out.corm);
f3=(out.corrp);
f4=(out.cprom);
f5=(out.cshad);
f6=(out.dissi);
f7=(out.energ);
f8=(out.entro);
f9=(out.homom);
f10=(out.homop);

```

```

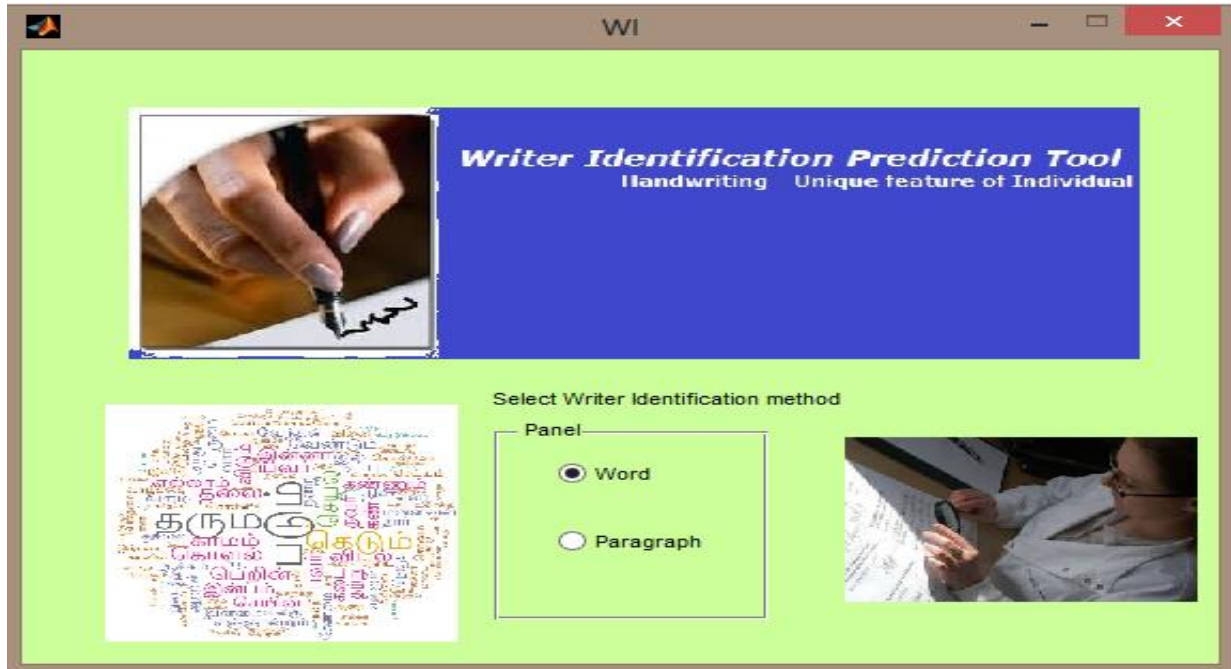
f11=(out.maxpr);
f12=(out.sosvh);
f13=(out.savgh);
f14=(out.svarh);
f15=(out.senth);
f16=(out.dvarh);
f17=(out.denth);
f18=(out.inf1h);
f19=(out.inf2h);
f20=(out.homom);
f21=(out.indnc);
f22=(out.idmnc);
f23=(out.autoc);
X=[f1,f2,f3,f4,f5,f6,f7,f9,f10,f11,f12,f13,f14,f15,f16,f17,f18,f19,f20,f21,f22,f23];
img_out_disp=mean(gabor(k));
% directional features
C = mean(ContCode(k,2));
%GGD FEATURE
d=GGD16(k);
pfilt = '9-7';
dfilt = 'pkva';
nlevs = [0, 0, 4, 4, 5];
kr=imresize(i1,[512,512]);
y = pdfbdec(double(kr), pfilt, dfilt, nlevs);
%CONTOURLET GGD FEATURE
%c1=GGD16(y{1});
train(it,:)=[X,img_out_disp,C];
it
clear y
end
%save trs1.xls train;
end
%load trs;
for i=1:size(train,1)
    for j=1:size(train,2)
        if isnan(train(i,j))==1
            train(i,j)=1;
        end
    end
end
end
[eivec, prodata, eival] = princomp(train);
[fo, fdx] =sort(eival,'descend');
sc1=train(:, fdx(1:222));
sc2=test(:, fdx(1:222));
train1=xlswrite('D:\PhD\matlab file\Datasets\paragraph\dataset\datasetII.xlsx',train);
sc11=xlswrite('D:\PhD\matlab file\Datasets\paragraph\dataset\datasetIII.xlsx',sc1);

```

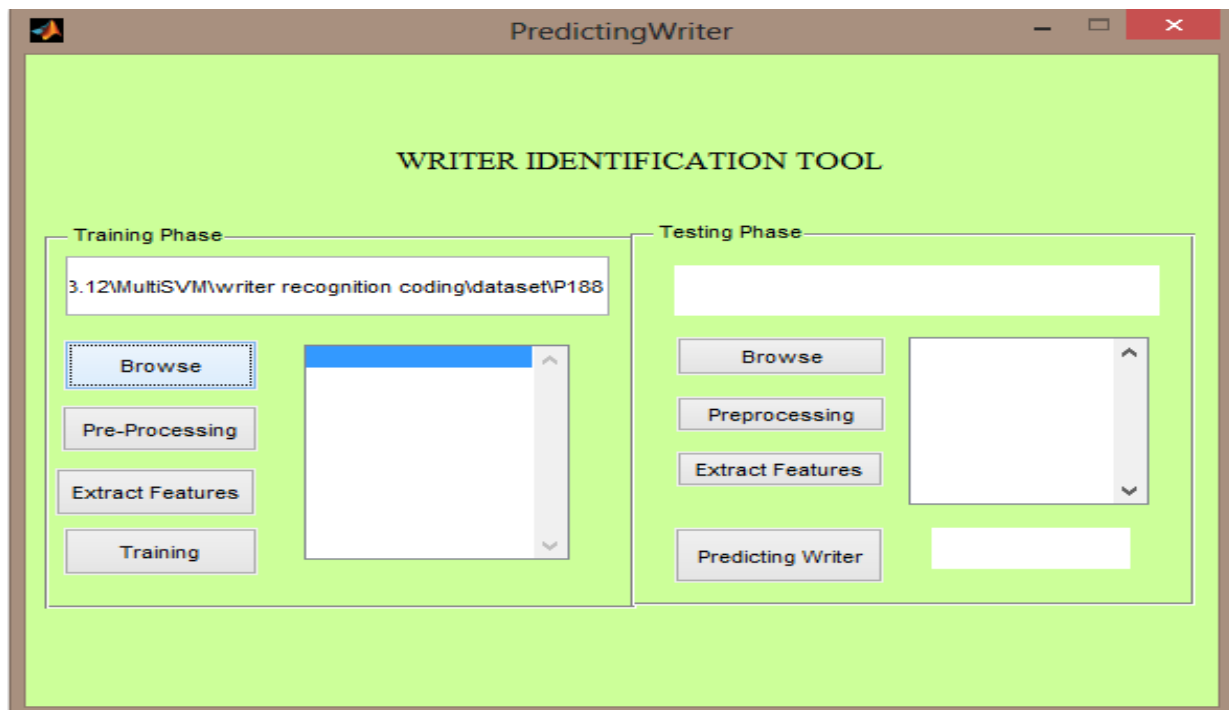
Appendix - D

Sample Screen Shots - Writer Identification Tool

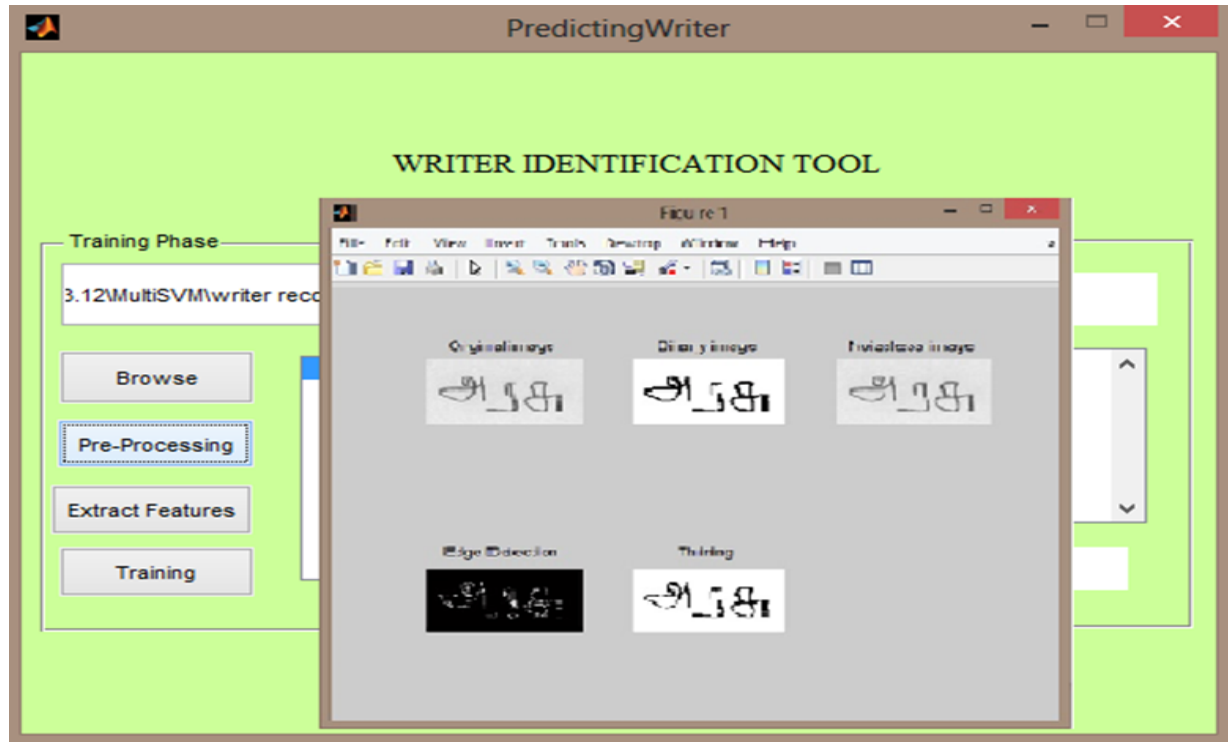
Writer identification tool design – word text image



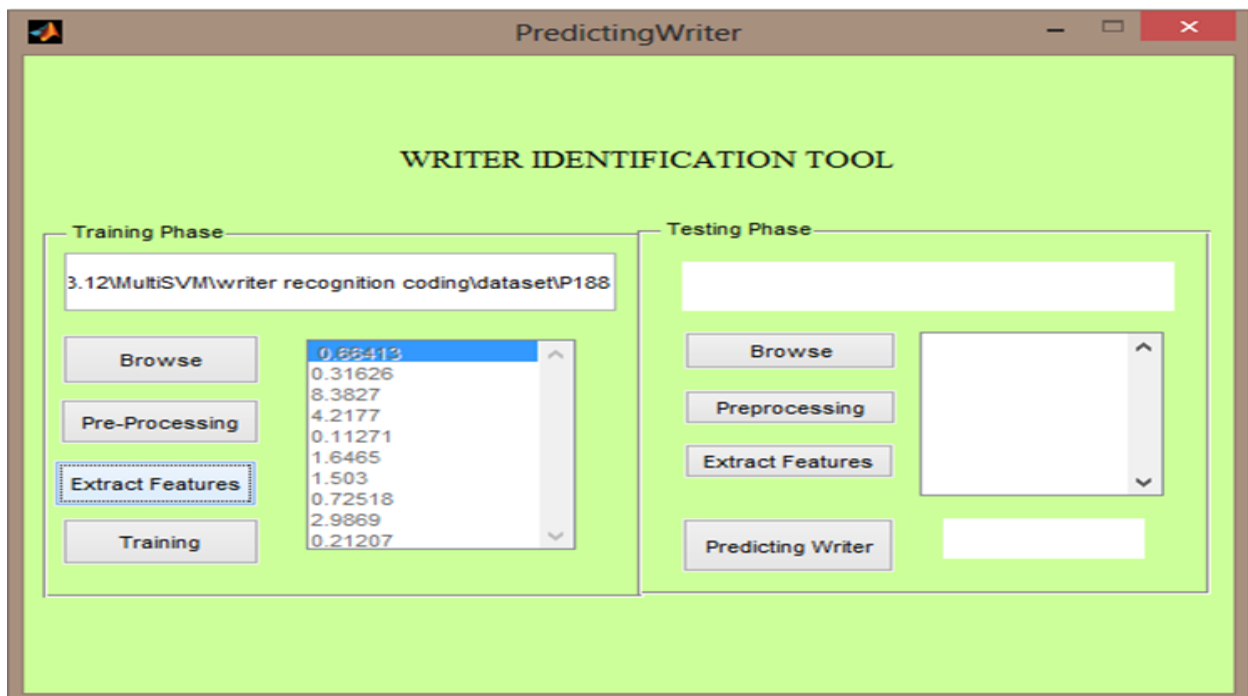
Loading dataset – Training Phase



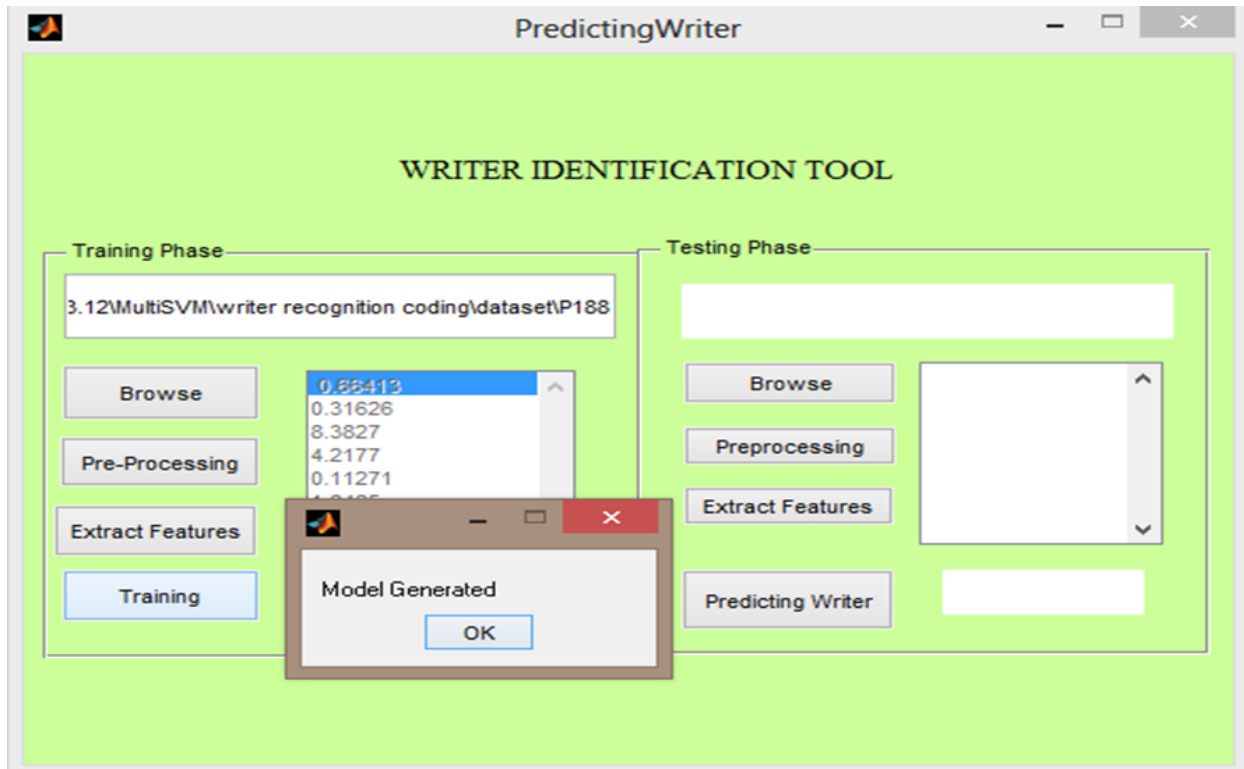
Preprocessing - Training Phase



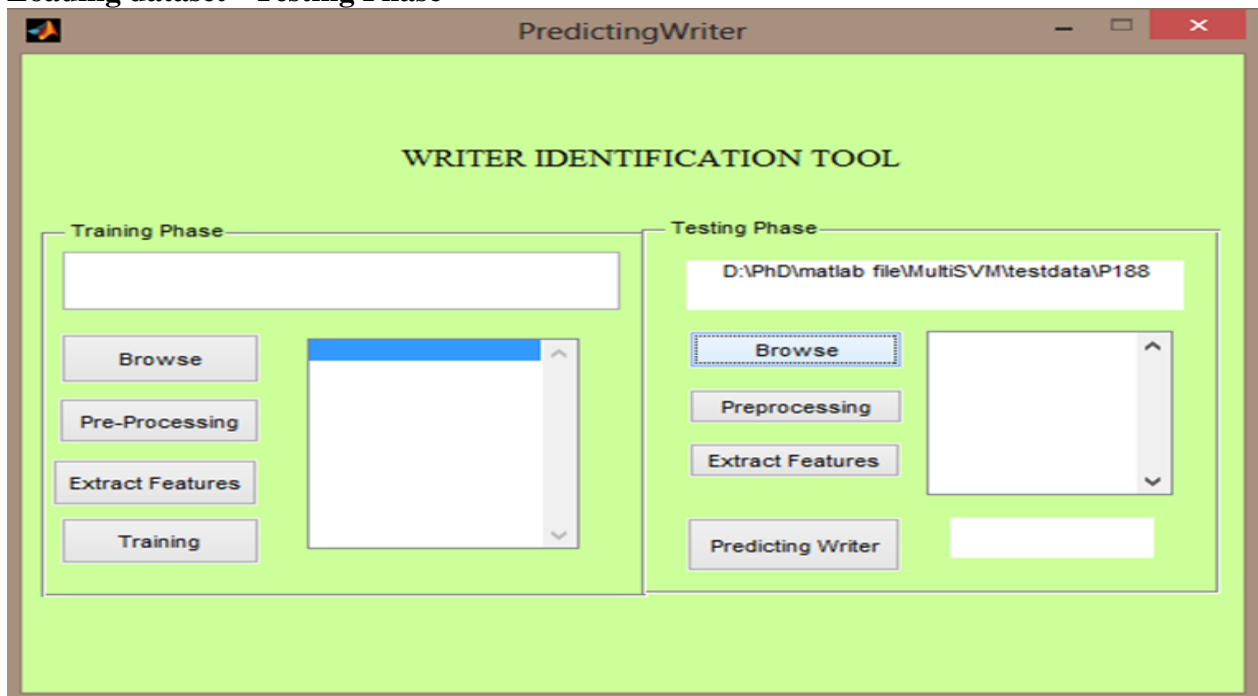
Feature extraction - Training Phase



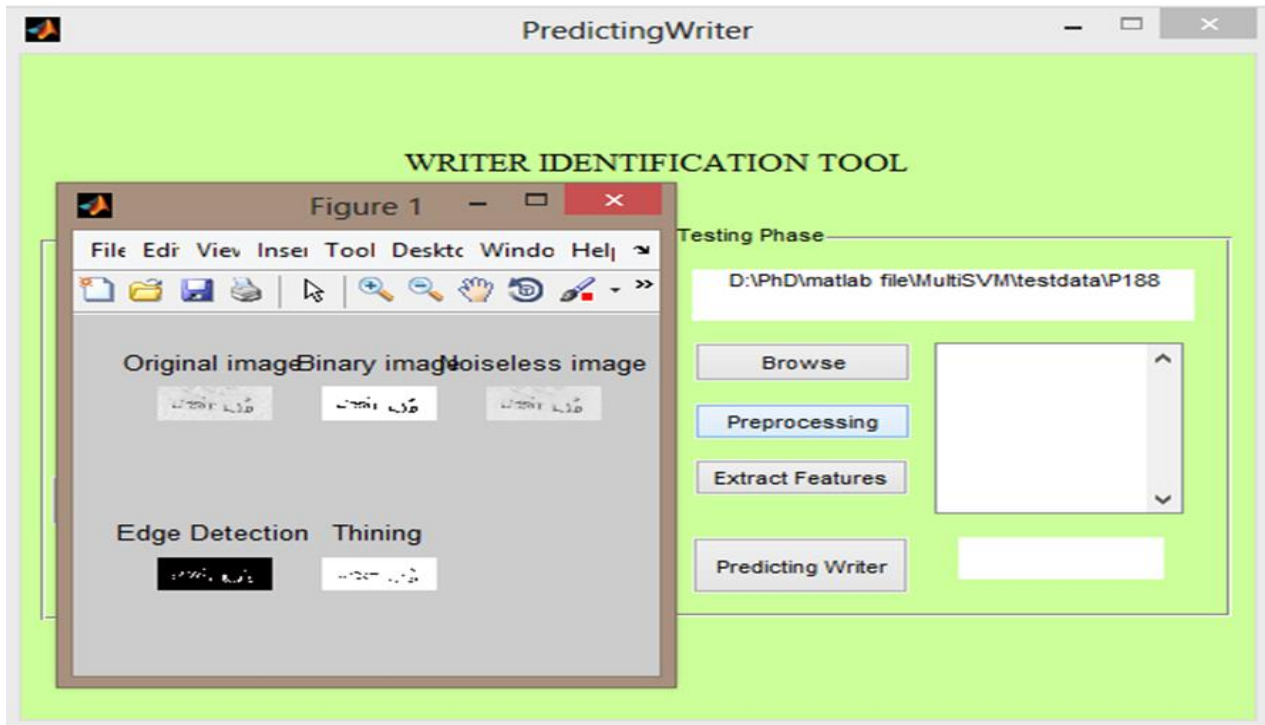
Model generated - Training Phase



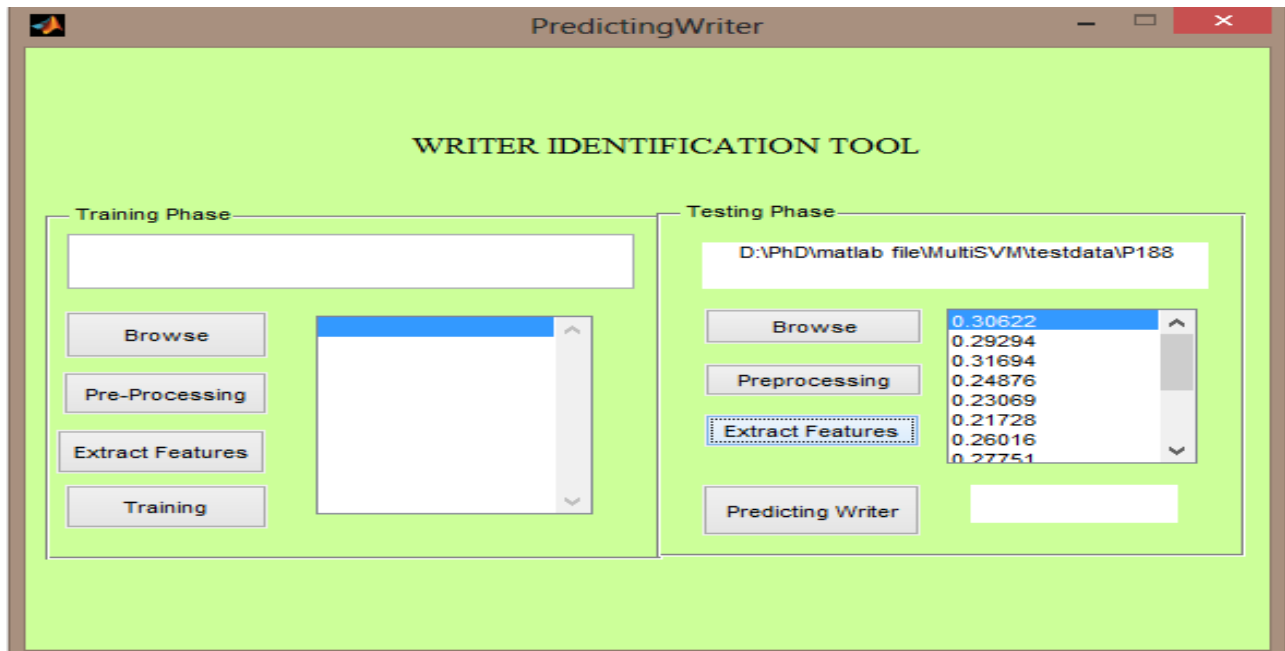
Loading dataset - Testing Phase



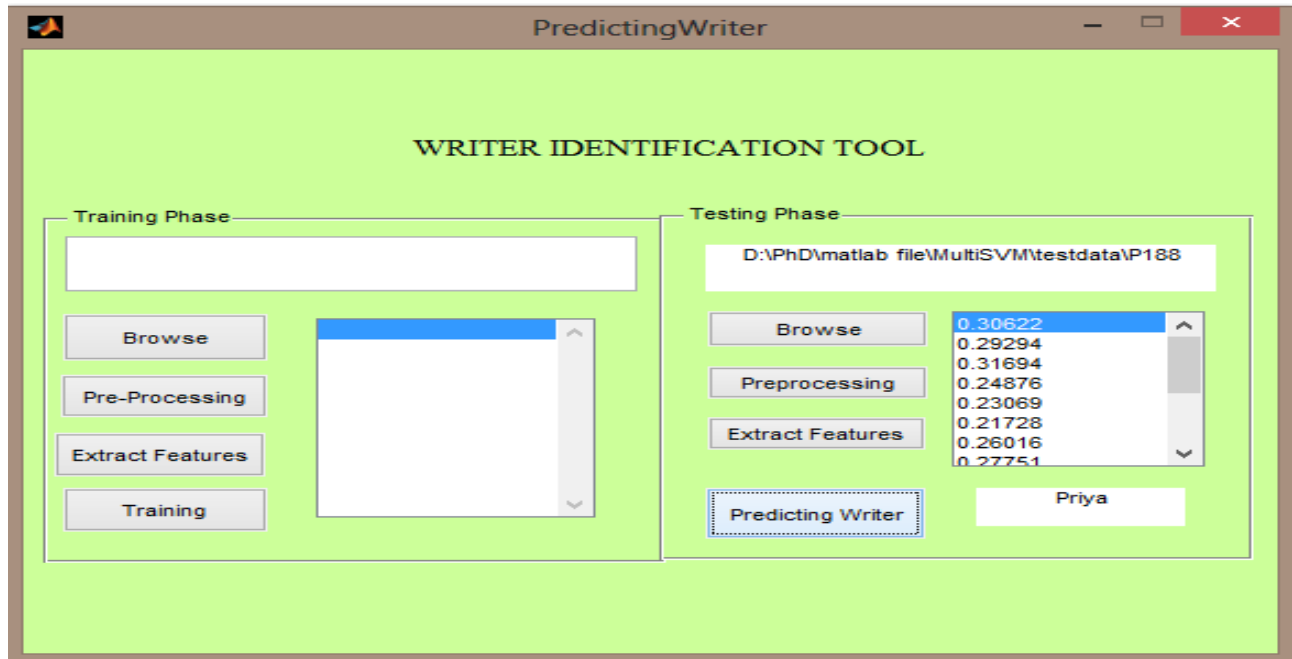
Preprocessing – Testing Phase



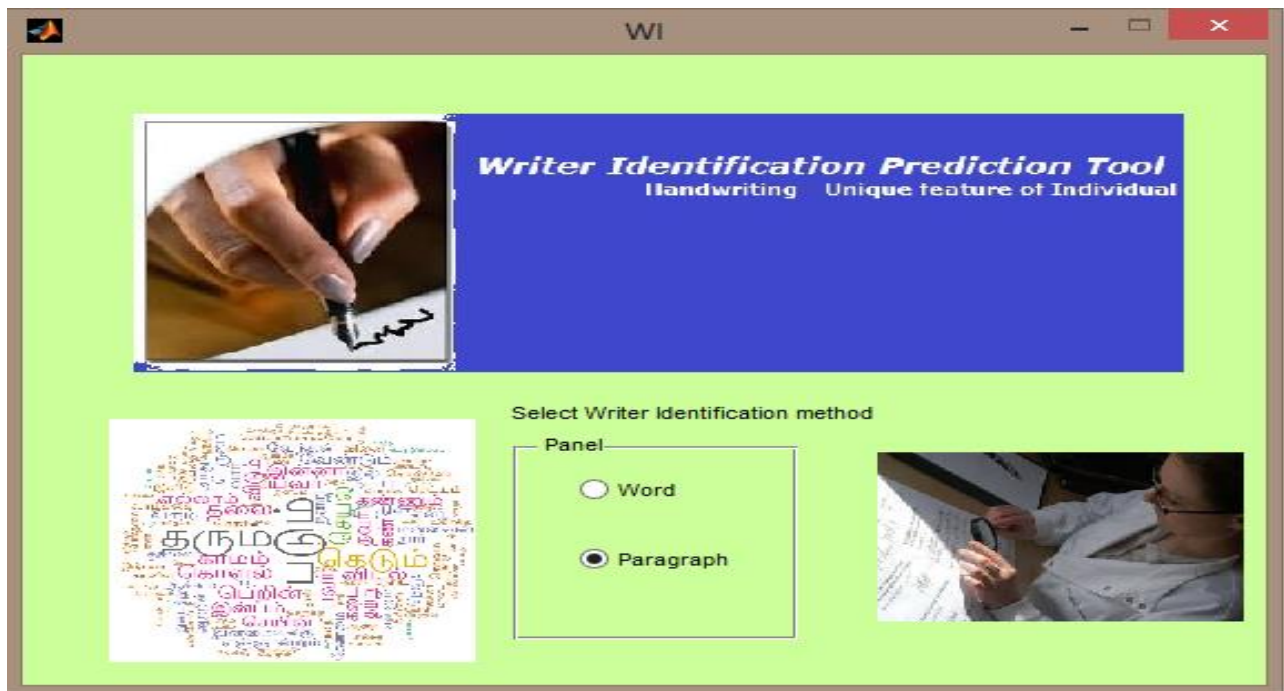
Feature Extraction - Testing Phase



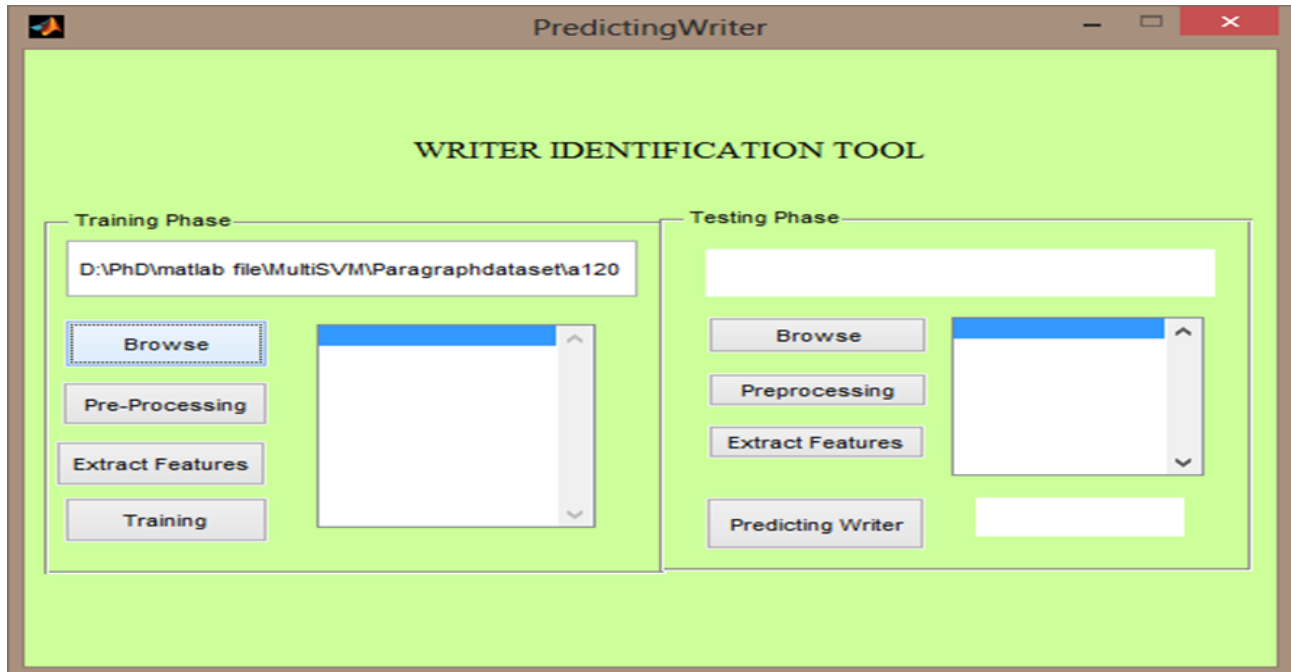
Predicting Writer using word text image



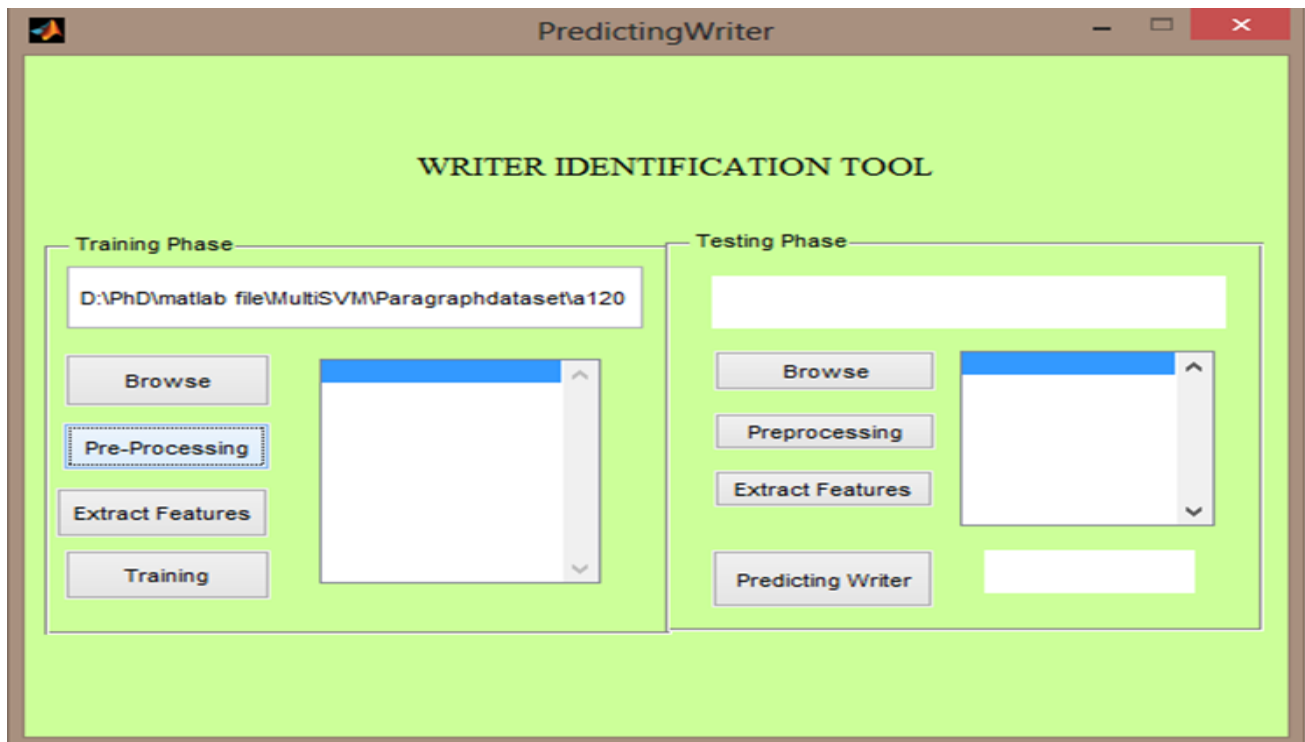
Writer identification tool - paragraph text image



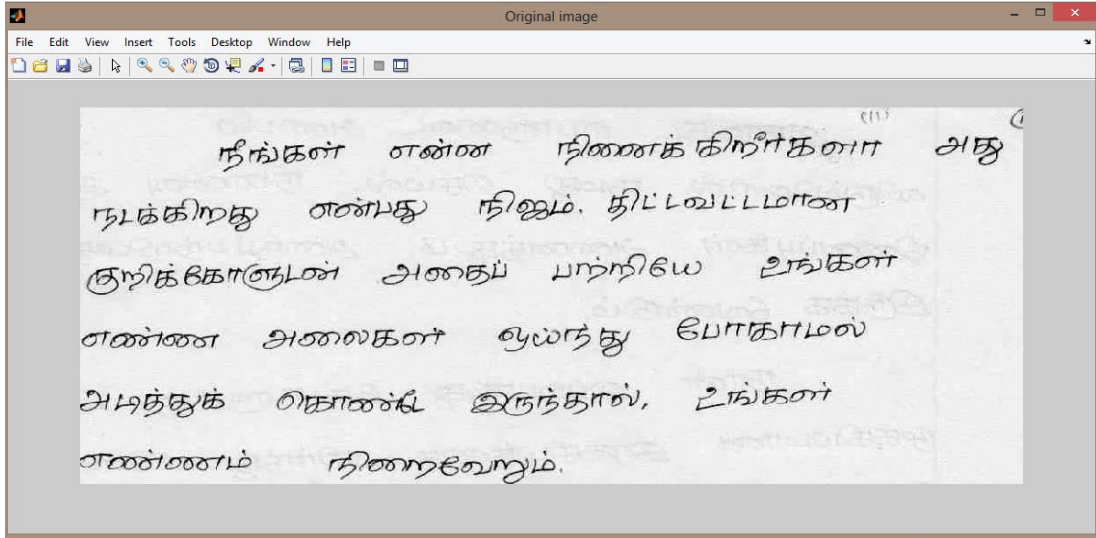
Loading dataset - Training Phase



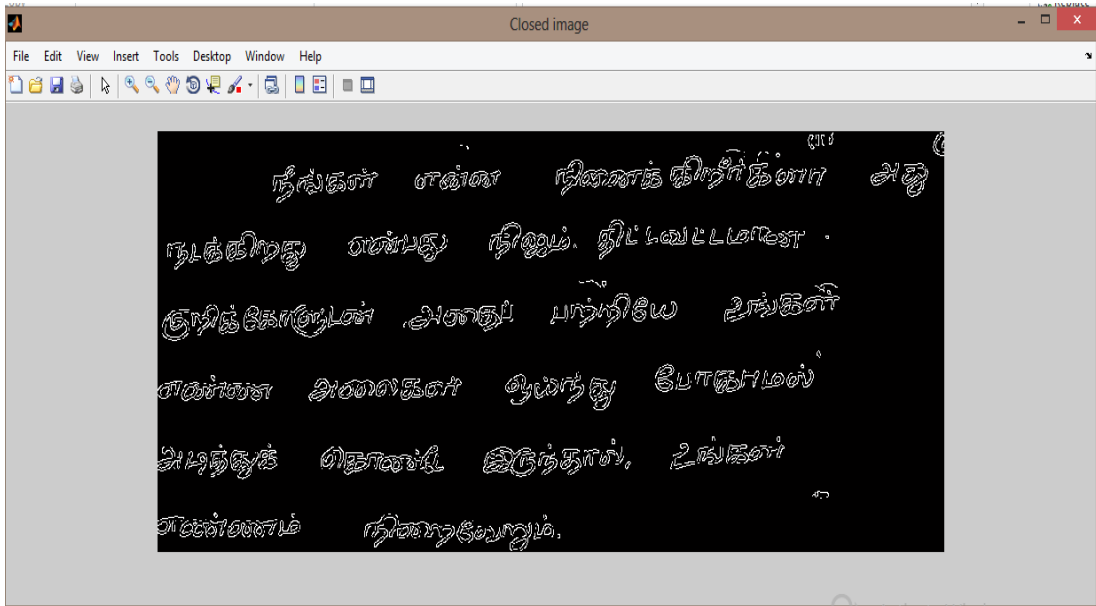
Preprocessing – Training Phase



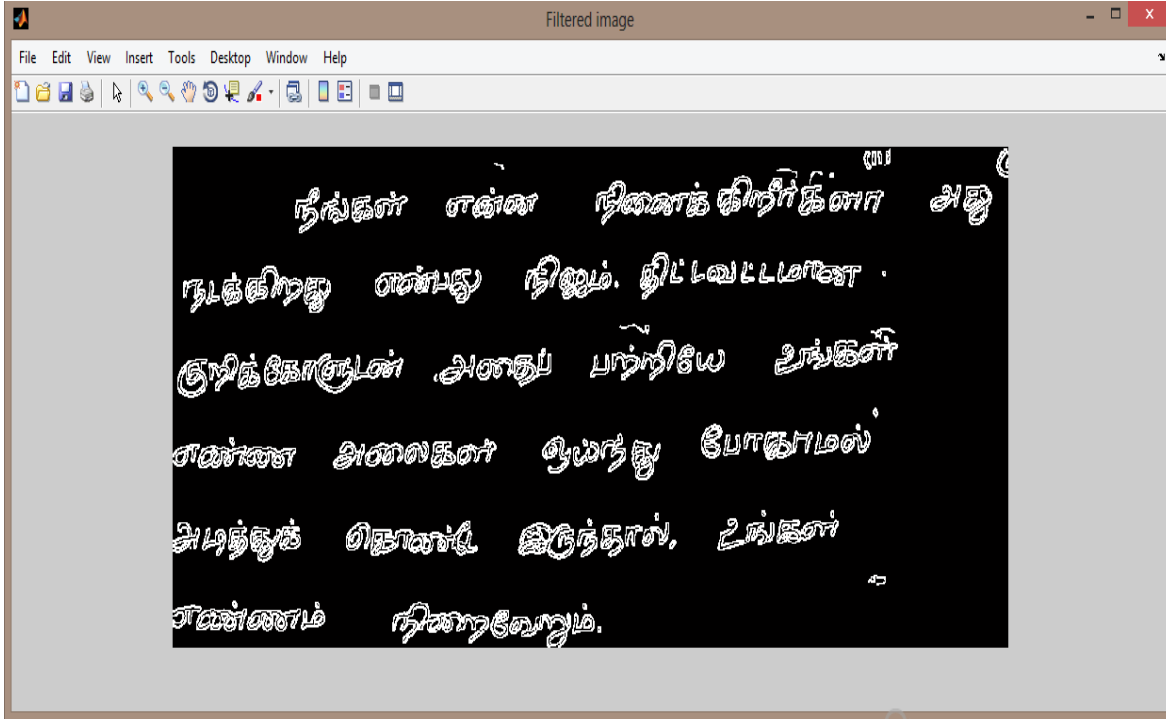
Original image



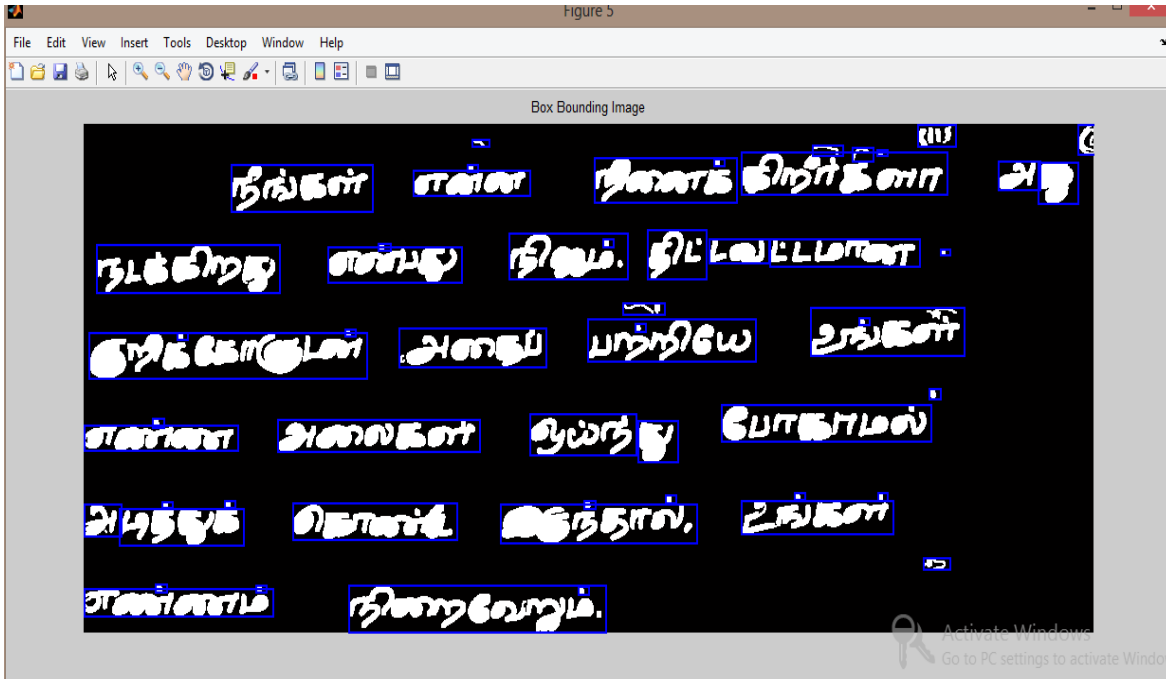
Dilation



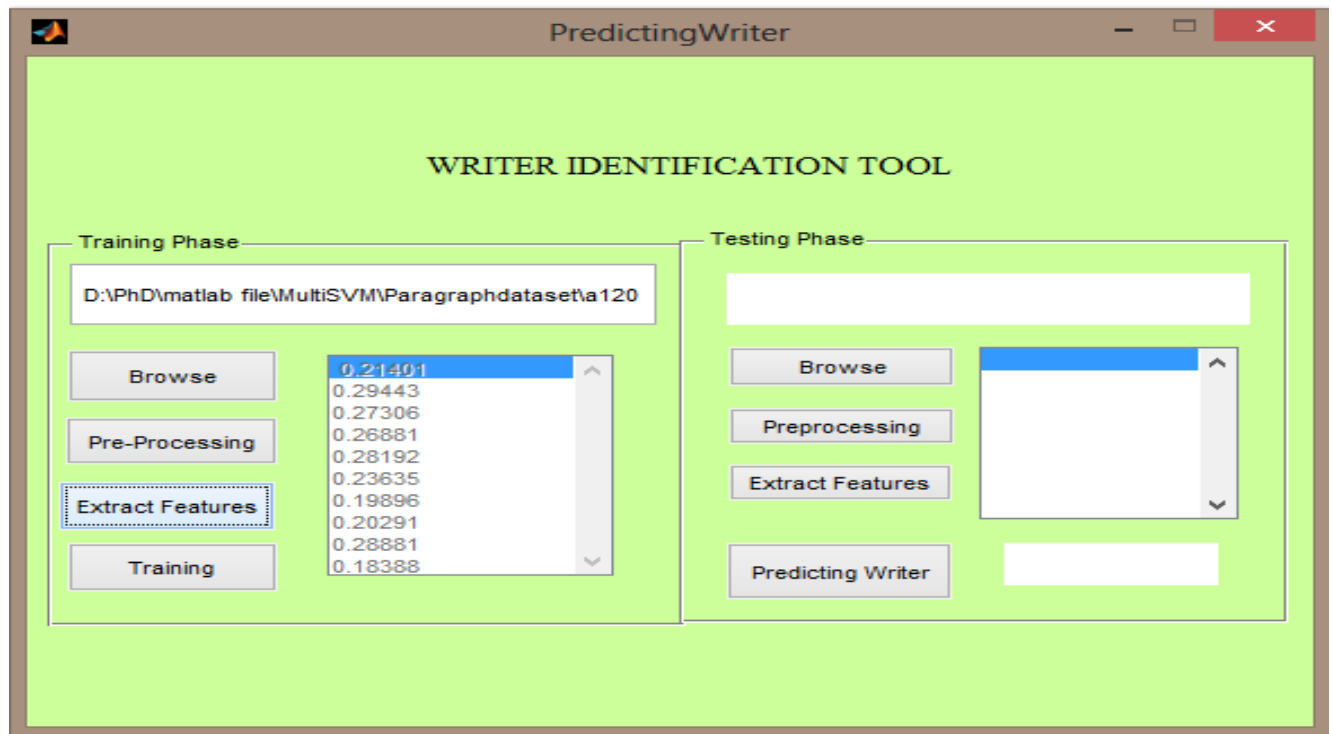
Edge detection



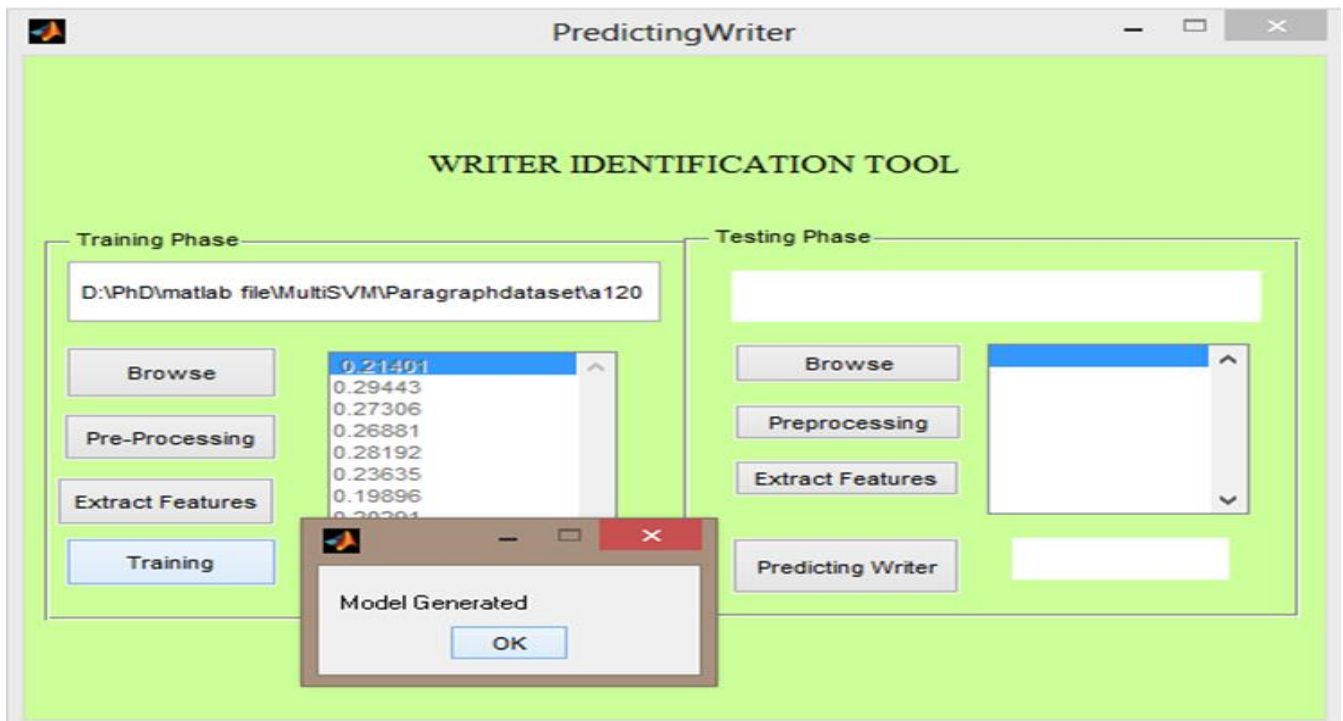
Box bounding



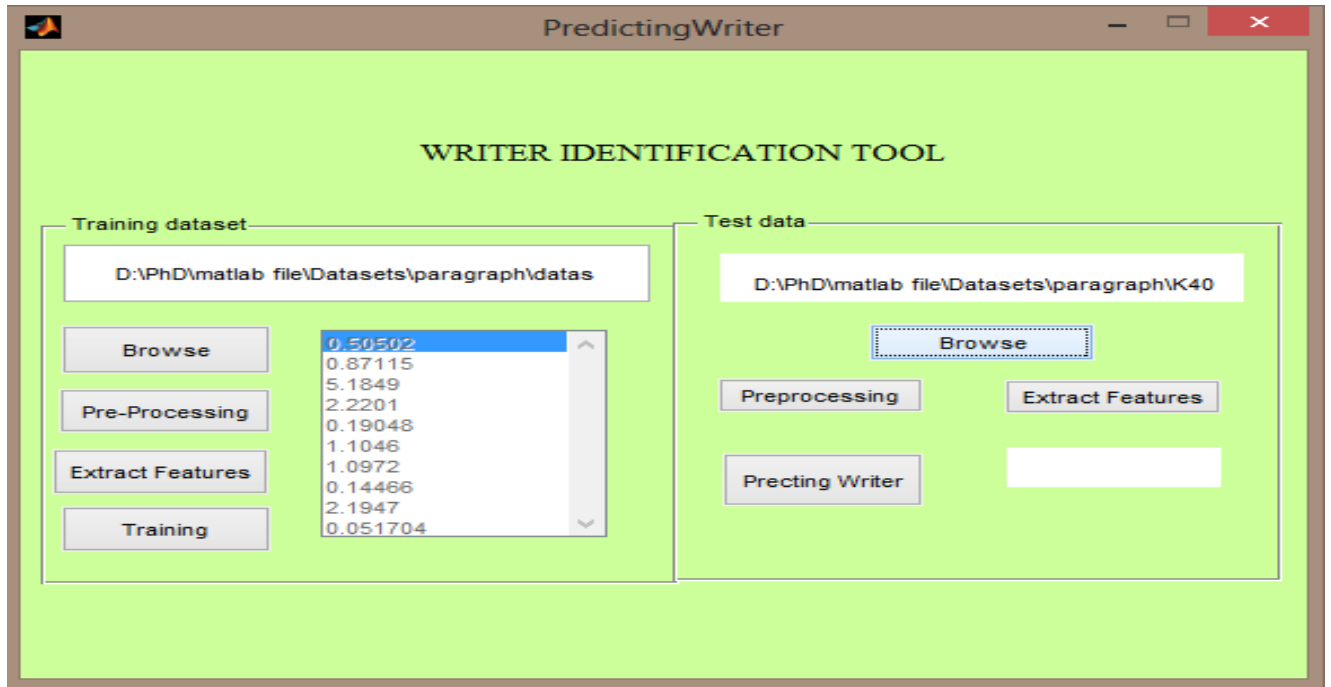
Feature extraction - Training Phase



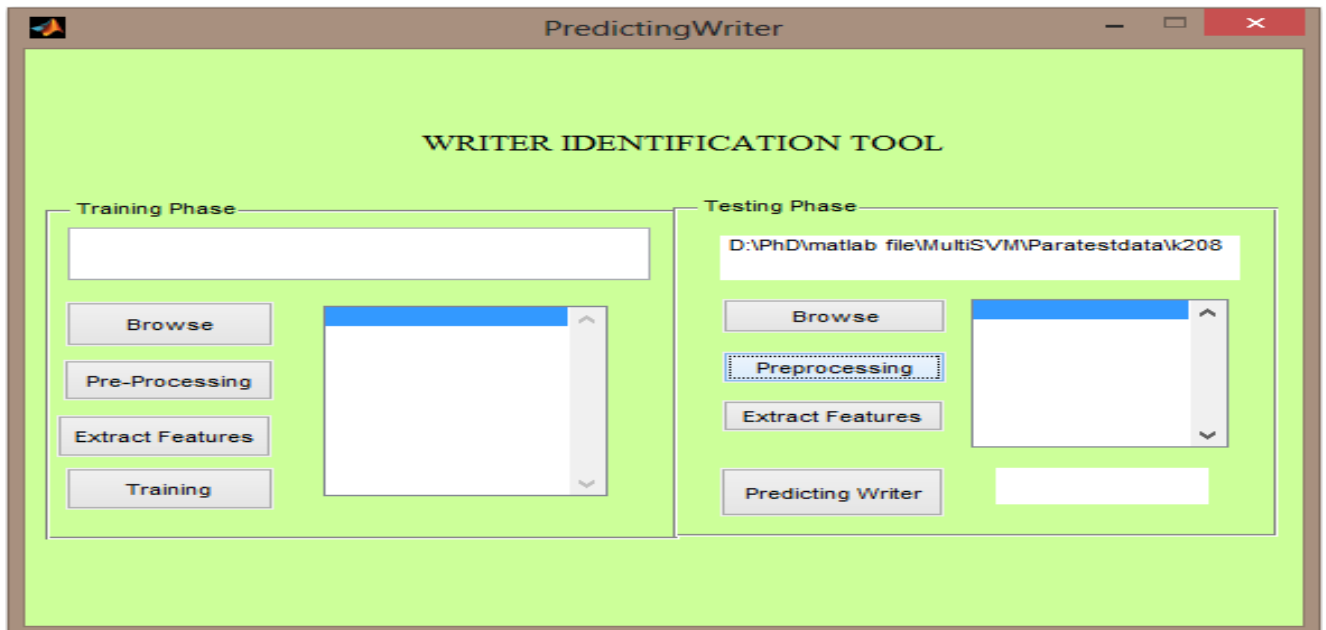
Model generated - Training Phase



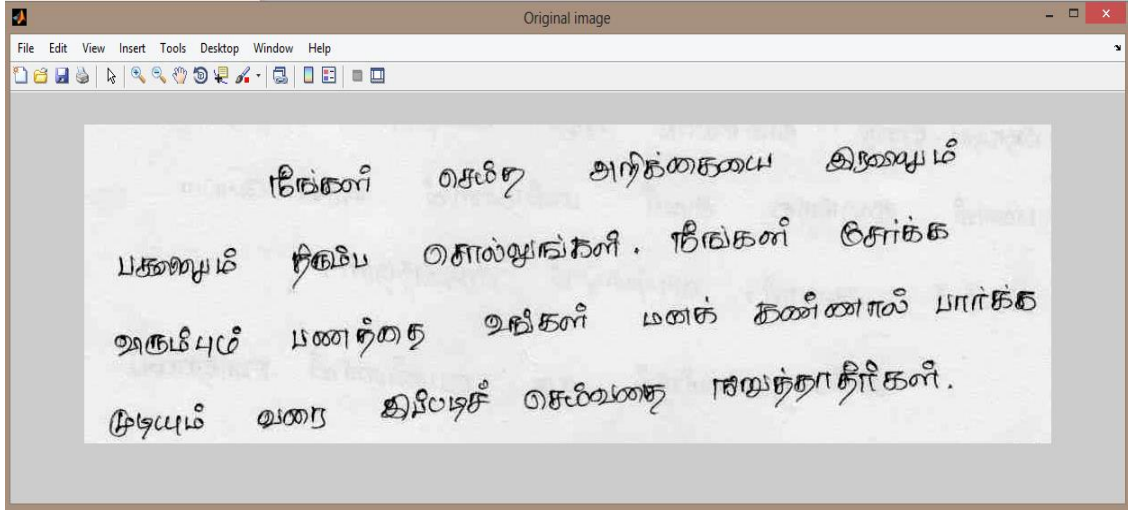
Loading dataset - Testing Phase



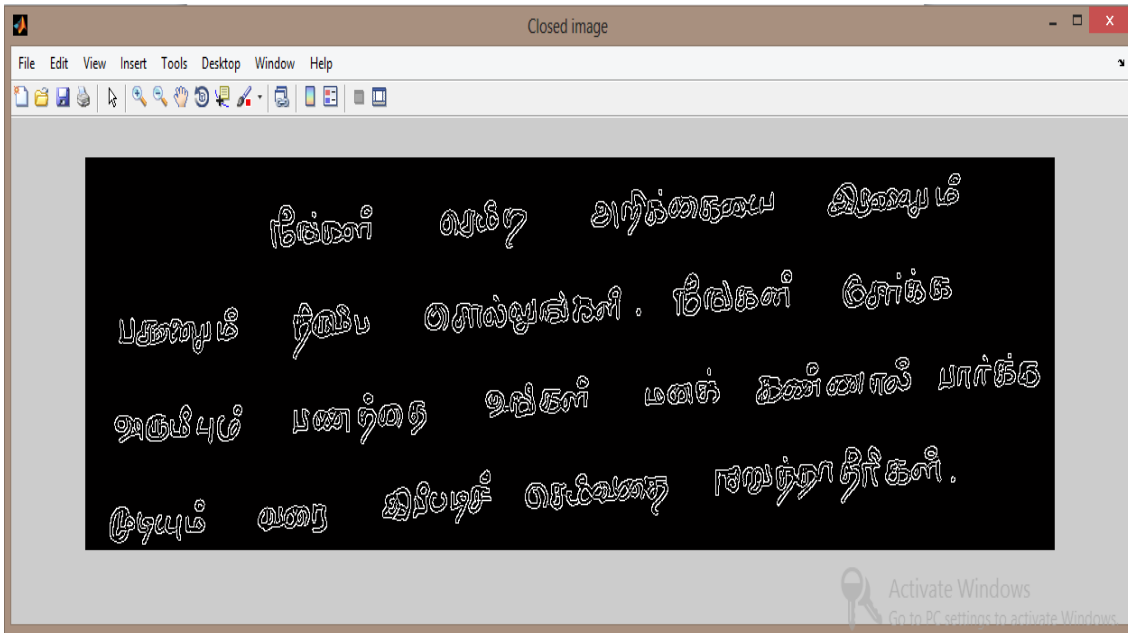
Preprocessing - Testing Phase



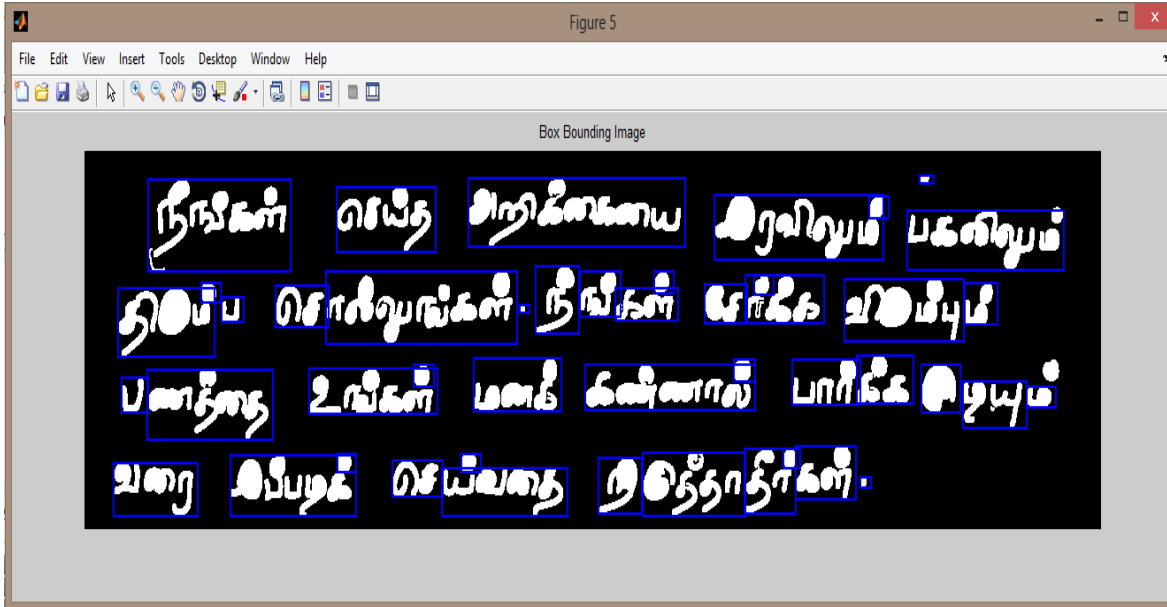
Original Image



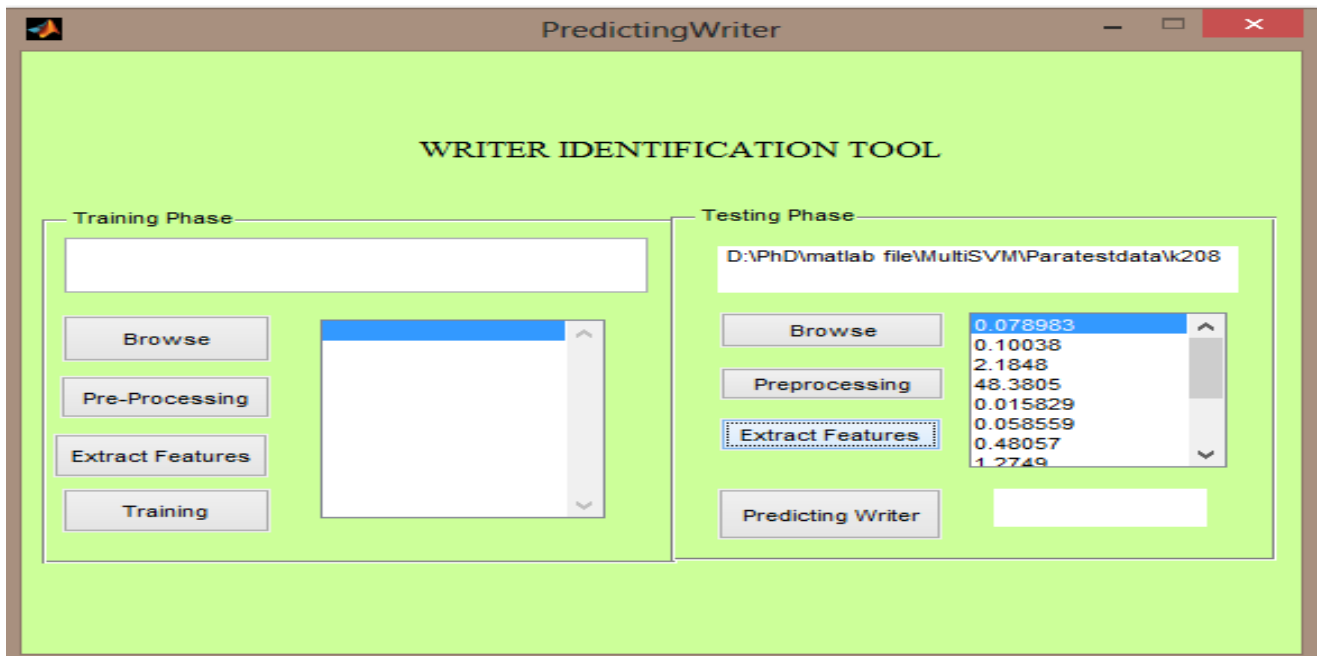
Dilation



Boxbounding



Feature Extraction – Testing Phase



Predicting Writer using paragraph text image

