
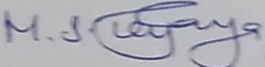


## CERTIFICATE

This is to certify that the thesis, entitled "PERFORMANCE ENHANCEMENT OF CLASSIFIERS FOR TAMIL WRITER IDENTIFICATION THROUGH MODIFIED SVM LINEAR KERNEL WITH PARAMETER ESTIMATION AND DEEP LEARNING" submitted to Bharathiar University, in partial fulfillment of the requirements for the award of the Degree of **Doctor of Philosophy in Computer Science** is a record of original research work done by **Mrs. T. Thendral** during the period of October 2011 to September 2018 of her research in the Department of Computer Science at PSGR Krishnammal College for Women, Coimbatore under my supervision and guidance and the thesis has not formed the basis for the award of any Degree/Diploma/Associateship/Fellowship or other similar title to any candidate of any University.

 05/09/2018

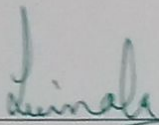
Countersigned

 05/09/2018

**Signature of the Guide**  
**Dr. M. S. VIJAYA**, M.Sc., M.Phil., Ph.D.  
Associate Professor and Head  
Department of Computer Science (PG)  
P.S.G.R. Krishnammal College for Women  
COIMBATORE - 641 004.

**Head of the Department**

**Dr. M. S. VIJAYA**, M.Sc., M.Phil., Ph.D.  
Associate Professor and Head  
Department of Computer Science (PG)  
P.S.G.R. Krishnammal College for Women  
COIMBATORE - 641 004.



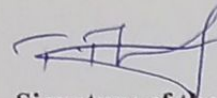
**Principal**

**PRINCIPAL**  
PSGR KRISHNAMMAL COLLEGE FOR WOMEN  
COIMBATORE - 641 004.

DECLARATION

## DECLARATION

I, **T. Thendral** hereby declare that the thesis, entitled "**PERFORMANCE ENHANCEMENT OF CLASSIFIERS FOR TAMIL WRITER IDENTIFICATION THROUGH MODIFIED SVM LINEAR KERNEL WITH PARAMETER ESTIMATION AND DEEP LEARNING**", submitted to Bharathiar University, in partial fulfillment of the requirements for the award of the **Degree of Doctor of Philosophy in Computer Science**, is a record of original and independent research done by me during the period of October 2011 to September 2018 under the Supervision and Guidance of **Dr. (Mrs.) M. S. Vijaya M.Sc., M.Phil., Ph.D**, Associate Professor and Head, Department of Computer Science (PG) at PSGR Krishnammal College for Women, Coimbatore and it has not been formed the basis for the award of any Degree / Diploma / Associateship / Fellowship or other similar title to any candidate of any University.



**Signature of the Candidate**

## GENINUNENESS OF PUBLICATION

This is to certify that the PhD candidate **Mrs. T. Thendral** working under my supervision has published research articles in the following refereed journals.

1. "Detection of a Person Using Descriptive Features of Tamil Handwriting and Pattern Learning", International Journal of Applied Engineering Research, (**Scopus Indexed**), ISSN 0973-4562, Vol 10(21), 2015, pp 41902-41909.
2. "Bayesian Linear Regression Co-efficients for SVM linear kernel to identify writers", Journal of Advanced Research in Dynamical and Control Systems (JARDCS), (**Scopus Indexed**), Vol. 10, 01-Special Issue, 2018

The contents of the publication incorporate part of the results in his/her thesis.

*M. S. Vijaya* 05/09/2018

Countersigned

*M. S. Vijaya* 05/09/2018

Head of the Department

**Dr. M. S. VIJAYA**, M.Sc., M.Phil., Ph.D.  
Associate Professor and Head  
Department of Computer Science (PG)  
P.S.G.R. Krishnammal College for Women  
COIMBATORE - 641 004.

*Jurnal*

Principal

**PRINCIPAL**  
PSGR KRISHNAMMAL COLLEGE FOR WOMEN  
COIMBATORE - 641 004.

Research Supervisor

**Dr. M. S. VIJAYA**, M.Sc., M.Phil., Ph.D.  
Associate Professor and Head  
Department of Computer Science (PG)  
P.S.G.R. Krishnammal College for Women  
COIMBATORE - 641 004.



பாரதியார் பல்கலைக்கழகம்  
**BHARATHIAR UNIVERSITY**  
COIMBATORE - 641 048, TAMILNADU, INDIA.

State University | Re-accredited with "A" Grade by NAAC | Ranked 14<sup>th</sup> among Indian Universities by MHRD-NIRF

**CERTIFICATE OF PLAGIARISM CHECK**

1	Name of the Research Scholar	T. THEENDRAL
2	Course of study	M.Phil., / Ph.D.,
3	Title of the Thesis / Dissertation	PERFORMANCE ENHANCEMENT OF CLASSIFIERS FOR TAMIL WRITER IDENTIFICATION THROUGH MODIFIED SVM LINEAR KERNEL WITH PARAMETER ESTIMATION AND DEEP LEARNING
4	Name of the Supervisor	DR. MS VIJAYA
5	Department / Institution/ Research Centre	DEPARTMENT OF COMPUTER SCIENCE, PSGR KRISHNAMMAL COLLEGE FOR WOMEN
6	Acceptable Maximum Limit	30 %
7	% of Similarity of content Identified	08%
8	Software Used	URKUND
9	Date of verification	29/08/2018

Report on plagiarism check, items with % of similarity is attached

*M. S. Vijaya*  
29/8/18

Signature of the Supervisor

(Seal)

Dr. M. S. VIJAYA, M.Sc., M.Phil., Ph.D.  
Associate Professor and Head  
Department of Computer Science (PG)  
P.S.G.R. Krishnammal College for Women  
COIMBATORE - 641 004.

*T. Theendral*

Signature of the Scholar

*M. S. Vijaya*  
29/8/18

R&D Director (BU) / Head of the Department

Dr. M. S. (Seal) VIJAYA, M.Sc., M.Phil., Ph.D.  
Associate Professor and Head  
Department of Computer Science (PG)  
P.S.G.R. Krishnammal College for Women  
COIMBATORE - 641 004.

*V. Rajendran*  
20/8/18

University Librarian (BU)

Arignar Anna Central Library  
Bharathiar University  
Coimbatore - 641 049

Research Coordinator (BU)

## ACKNOWLEDGEMENT

The work presented in this thesis would not have been possible without the contribution of many people. I take this opportunity to extend my sincere gratitude and appreciation to all those who made this Ph.D thesis possible.

I am grateful to **Smt. R. Nandini**, Chairperson, PSGR Krishnammal College for Women for providing the opportunity to pursue the research.

I have been fortunate to be inspired by **Dr. (Mrs.) N. Yesodha Devi, M.Com., Ph.D**, Secretary, PSGR Krishnammal College for Women for her warm support and all the amenities provided for the conduct of study.

I am deeply indebted to **Dr. (Mrs). S. Nirmala, MBA., M.Phil., Ph.D**, Principal, PSGR Krishnammal College for Women for her encouragement and motivation. The institution's impressive accomplishment in advancing research has motivated me in taking up this research work.

I place my sincere gratitude to my guide, **Dr. (Mrs.) M.S. Vijaya M.Sc., M.Phil., Ph.D**, Associate Professor and Head, Department of Computer Science (PG), PSGR Krishnammal College for Women. She has been judicious and provided strong support, during the journey of this research. I am indebted to her for the constant encouragement, valuable suggestions and resourceful ideas. Words are inadequate to quantify her immense knowledge and understanding of the subject which helped me to have better grasp of the theoretical aspects. Her insights into several topics have shaped the progress of this research. I thankfully acknowledge and value the amount of time she has spent editing my papers, discussing my research ideas, listening to my problems and taking away all my doubts and worries. Without her perseverance, fore thought and intelligence, I would not have been able to complete this research.

I am greatly indebted to the all the faculty members, lab assistants, librarians and supporting staff of the Department of Computer Science, PSGR Krishnammal College for Women for their inspiring technical help and valuable suggestions rendered at various stages of this research.

I feel immense pleasure in expressing my profound sense of gratitude to my husband, **Mr.T.Suresh**, my sons **Master S.Ashwin** and **Master S.Kaniyan**. My research and this thesis would have never been complete without their constant support, direction and patient collaboration. I thank them for the trust they have on me and their confidence about the usefulness of my work.

Last but not the least, I would like to thank my family, my parents for supporting me morally throughout my life. They have offered me encouragement, drive, passion and ideas, as well as critiques in the process towards the completion of this research work. Without them, this research would not have been complete to its perfection.

**T.THENDRAL**

## LIST OF FIGURES

<b>FIGURE NO.</b>	<b>TITLE</b>	<b>PAGE NO.</b>
1.1	Pattern Matching Approach	14
1.2	Machine Learning Approach	15
2.1	Maximum Margin Classifier	30
2.2	Linear Classifier with Training Error	31
2.3	Nonlinear Classifier	32
2.4	Input Space and Feature Space	33
2.5	Data which Require Linear SVM Formulation with Soft Margin	39
2.6	Data which Require Non Linear Classifier	47
2.7	Non-Linear Mapping into Feature Space	48
2.8	Prediction using One Against All SVM	52
2.19	Prediction using One Against One SVM	53
2.10	Prediction using DAGSVM	54
2.11	Artificial Neuron Model	58
2.12	Multi-Layer Perceptron Neural Network	60
2.13	Example of DNNs	64
2.14	Typical CNN Architecture	66
3.1	Proposed Writer Identification Model	71
3.2	Multifaceted Calligraphies in Tamil	72
3.3	Text Dependent Character Images – Sample Data	73
3.4	Text Dependent Word Images – Sample Data	73
3.5	Text Dependent Paragraph Images – Sample Data	74
3.6	a) Before Noise Removal	76
	b) After Noise Removal	76

<b>FIGURE NO.</b>	<b>TITLE</b>	<b>PAGE NO.</b>
3.7	a) Non- Binarized Image	76
	b) Binarized Image	76
3.8	a) Original Image	77
	b) Edge Detected Image	77
3.9	a) Original Image	78
	b) Thinned Image	78
3.10	Edge Detection	79
3.11	Dilated Image	79
3.12	Box Bounded Image	80
3.13	Length of the Word	82
3.14	Height of the Word	83
3.15	Length from Baseline to Upper Edge	83
3.16	Length from Baseline to Lower Edge	84
3.17	Ascender and Descender Baseline	84
3.18	Junction of the Character	84
3.19	Loops of the Character	85
3.20	Calculation of GLCM	92
4.1	Learning SVM with Linear Kernel	111
4.2	Classification using Linear Kernel SVM Model	111
4.3	Comparative Results of Accuracy (Character)	114
4.4	Comparative Results of Accuracy (Word)	117
4.5	Comparative Results of Accuracy (Paragraph)	120
5.1	Performance Comparison of Linear SVM and WLK - SVM Based Models	128
5.2	ROC for TWINC Dataset	128
5.3	ROC for TWINW Dataset	128



<b>FIGURE NO.</b>	<b>TITLE</b>	<b>PAGE NO.</b>
5.4	ROC for TWINP Dataset	128
5.5	Performance Comparison of Linear SVM, WLK – SVM and BLK - SVM Based Models	133
5.6	ROC for TWINC Dataset	133
5.7	ROC for TWINW Dataset	133
5.8	ROC for TWINP Dataset	133
5.9	Performance Comparison of Linear SVM, WLK – SVM, BLK – SVM and PCK – SVM Based Models	140
5.10	ROC for TWINC Dataset	140
5.11	ROC for TWINW Dataset	140
5.12	ROC for TWINP Dataset	140
6.1	Process Flow of CNN Based Writer Identification Model	144
6.2	Recognition Rate of Character Text Image Identification	149
6.3	Error rate of Character Text Image Identification	149
6.4	Recognition Rate of Word Text Image Identification	150
6.5	Error Rate of Word Text Image Identification	150
6.6	Recognition Rate of Paragraph Text Image Identification	151
6.7	Error Rate of Paragraph Text Image Identification	151
6.8	Comparative Analysis of ANN and CNN	152
7.1	Opening Screen of the Tool	155
7.2	Components of the Tool	156

## LIST OF TABLES

<b>TABLE NO.</b>	<b>TITLE</b>	<b>PAGE NO.</b>
1.1	Summary of Literature Survey	21
3.1	Feature Vector of a Single Word Image	89
3.2	Count of Features Before and After PSO Feature Selection	103
4.1	Profile of the Datasets	110
4.2	Results of SVM with Linear Kernel (Character)	112
4.3	Results of SVM with RBF Kernel (Character)	113
4.4	Results of SVM with Polynomial Kernel (Character)	113
4.5	Consolidated Results of all Three Models (Character)	114
4.6	Results of SVM with Linear Kernel (Word)	115
4.7	Results of SVM with RBF Kernel (Word)	116
4.8	Results of SVM with Polynomial Kernel (Word)	116
4.9	Consolidated Results of all Three Models (Word)	117
4.10	Results of SVM with Linear Kernel (Paragraph)	118
4.11	Results of SVM with RBF Kernel (Paragraph)	119
4.12	Results of SVM with Polynomial Kernel (Paragraph)	119
4.13	Consolidated Results of all three Models (Paragraph)	120
5.1	Results of SVM with WLK Kernel by Tuning C-Regularization Parameter	127
5.2	Performance Comparison of Linear SVM and WLK - SVM Based Models	127
5.3	Results of SVM with BLK Kernel by Tuning C-Regularization Parameter	131
5.4	Performance Comparison of Linear SVM, WLK – SVM and BLK - SVM Based Models	132
5.5	Results of SVM with PCK Kernel by Tuning C-Regularization Parameter	138
5.6	Performance Comparison of Linear SVM, WLK – SVM, BLK – SVM and PCK – SVM Based Models	139

6.1	Results of CNN Based Models	146
6.2	Results of ANN Based Models	147
6.3	Performance Comparison of ANN and CNN	148

## LIST OF ABBREVIATIONS

ANNs	Artificial Neural Networks
BLK	Bayesian Linear Kernel
BLR	Bayesian Linear Regression
BP	Back Propagation
BPNN	Back Propagation Neural Network
CNN	Convolutional Neural Network
CNNC	Convolutional Neural Network Characters
CNNP	Convolutional Neural Network Paragraphs
CNNW	Convolutional Neural Network Words
CRM	Customer Relationship Management
DCT	Discrete Cosine Transform
DNN	Deep Neural Network
EMD	Empirical Mode Decomposition
FFNN	Feed Forward Neural Network
FN	False Negatives
FP	False Positives
FS	Feature Selection
FSS	Feature Set Search
GAs	Genetic Algorithms
GGD	Generalized Gaussian Density
GIS	Geographic Information Systems
GLCM	Gray Level Co-occurrence Matrix
GSCM	Grey Scale Co-Occurrence Matrix
HMM	Hidden Markov Model
HPP	Horizontal Projection Profile
KDD	Knowledge Discovery in Databases
K-NN	K-Nearest Neighbor
LDC	Linear Discriminant Classifier
MAX	Maximum Value

MDA	Multiple Discriminant Analysis
MLP	Multilayer Perceptron
NNs	Neural Networks
PCA	Principal Component Analysis
PCK	Principal Component Kernel
PCR	Principal Component Regression
PDFB	Pyramidal Directional Filter Bank
PDFs	Probability Distribution Functions
PNN	Probabilistic Neural Network
RBF	Radial Basis Function
ROC	Receiver Operating Characteristic
SIANN	Space Invariant Artificial Neural Networks
SSMs	Spectral–Statistical Measures
SVM	Support Vector Machine
THWI	Tamil Handwriting Writer Identification
TN	True Negatives
TP	True Positives
WED	Weighted Euclidean Distance
WLK	Weighted Least Kernel
WLS	Weighted Least Square