CERTIFICATE

This is to certify that the thesis entitled "**GRAPH BASED SEGMENTATION AND DEEP LEARNING FOR PHONEME PATTERN CLASSIFICATION IN TAMIL CONTINUOUS SPEECH**" 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 **B.R. Laxmi Sree** during the period January 2012 to September 2019 of her study 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 of any candidate of any university.

M. S. Vijaya

Countersigned

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.

M. S. Vijaya

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

I, **B.R. Laxmi Sree** hereby declare that the thesis, entitled **"GRAPH BASED SEGMENTATION AND DEEP LEARNING FOR PHONEME PATTERN CLASSIFICATION IN TAMIL CONTINUOUS SPEECH"** 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 work done by me during January 2012 to September 2019 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 at PSGR Krishnammal College for Women, Coimbatore and it has not formed the basis for the award of any Degree / Diploma / Associateship / Fellowship or other similar title to any candidate in any university.

Signature of the Candidate

CERTIFICATE OF GENUINESS OF THE PUBLICATION

This is to certify that the Ph.D. candidate Mr./Ms. **B.R. Laxmi Sree** working under my supervision has published a research article in the refereed journal named **International Journal of Speech Technology** with Volume No. **22** Page Numbers **143-156** and year of publication **2019** published by **Springer.** The contents of the publication incorporate part of the results presented in his/her thesis.

Countersigned

Principal/Head of the Dept./Director
(College/University/Institute)
(with seal)

PRINCIPAL
PSGR KRISHNAMMAL COLLEGE FOR WOMEN
COIMBATORE - 641 004.

Research Supervisor
(with 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.

பாரதியார் பல்கலைக்கழகம்

BHARATHIAR UNIVERSITY

COIMBATORE - 641 046, TAMILNADU, INDIA.

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

CERTIFICATE OF PLAGIARISM CHECK

1	Name of the Research Scholar	B.R. LAXMI SREE
2	Course of study	~~M.Phil.,~~ / Ph.D.,
3	Title of the Thesis / Dissertation	GRAPH BASED SEGMENTATION AND DEEP LEARNING FOR PHONEME PATTERN CLASSIFICATION IN TAMIL CONTINUOUS SPEECH
4	Name of the Supervisor	Dr. (Mrs.) M.S. VIJAYA
5	Department / Institution/ Research Centre	PSGR KRISHNAMMAL COLLEGE FOR WOMEN, COIMBATORE DEPT. OF COMPUTER SCIENCE
6	Acceptable Maximum Limit	30 %
7	% of Similarity of content Identified	0 %
8	Software Used	URKUND
9	Date of verification	03/09/2019

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

M.S. Vijaya

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.

Signature of the Scholar

M.S. Vijaya

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

(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.

V. Rajendran 3/9/19

University Librarian (BU)

University Librarian
Arignar Anna Central Library
Bharathiar University
Coimbatore - 641 046.

Research Coordinator (BU)

ACKNOWLEDGEMENT

First and foremost I thank the God Almighty for giving me the strength, knowledge, ability, opportunity to undertake this research work and to persevere and complete it successfully.

I extend my thanks to **Dr. (Mrs.) R. Nandini, Chairperson,** PSGR Krishnammal College for Women, Coimbatore for having given me the opportunity to undertake this research work in this esteemed institution..

I express my whole hearted thanks to **Dr. (Mrs.) N. Yesodha Devi, M.Com, M.Phil, Ph.D., Secretary,** PSGR Krishnammal College for Women, Coimbatore for her continuous motivation and encouragement.

My heartfelt thanks to **Dr. (Mrs.) S. Nirmala, MBA, M.Phil., Ph.D., Principal,** PSGR Krishnammal College for Women, Coimbatore for her kind support and for all the resources provided.

I place on record my deep sense of gratitude to my research supervisor **Dr. (Mrs.) M.S. Vijaya, M.Sc., M.Phil., Ph.D.,** Associate Professor and Head, Department of Computer Science, PSGR Krishnammal College for Women, Coimbatore for her sustained interest and advice that have contributed to a great extent for the completion of this work. I am thankful for her appropriate guidance, insightful suggestions and support in the completion of this research work. Her patience and immense knowledge has helped me during the tenure of research and writing of thesis.

I wish to thank my friend **Dr. (Mrs.) M. Suguna** for her timely help, words of encouragement and suggestion during the course of study. I also extend my thanks to fellow colleagues at Department of Computer Science, Dr. G.R. Damodaran College of Science, Coimbatore for their continuous support and encouragement. My sincere thanks to my research colleagues for the motivation and guidance rendered. I would also thank the faculty members of Department of Computer Science (PG) for their co-operation and greeting.

Finally, I place on record my deep sense of gratitude to my husband **Mr. D. Raguram**, beloved parents, father-in-law, son, daughter and my sisters for supporting me in successfully completing this work. This would not have been possible without their unwavering and unselfish love and support given to me at all times.

B.R. Laxmi Sree

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

2D	-	Two Dimensional
AI	-	Artificial Intelligence
ANFIS	-	Adaptive Neuro-Fuzzy Inference System
ANN	-	Artificial Neural Network
ATIS	-	Airline Travel Information System
AUC	-	Area Under Curve
BBDBN	-	Bernoulli - Bernoulli DBN
BSNLF	-	Blind Segmentation using Non-linear filters
CART	-	Classification and Regression Trees
CD	-	Context Dependent
CD-DBN	-	Deep Belief Networks pretrained with Contrastive Divergence
C-MMSE	-	Cepstral-Minimum Mean Squared Error
CMU	-	Carnegie Mellon University
CSR	-	Continuous Speech Recognition
CT	-	Computed Tomography
CWT	-	Continuous Wavelet Transform
DBN	-	Deep Belief Network
DNN-HMM	-	Deep Neural Network – Hidden Markov Model
DP	-	Deleted Phonemes
DTW	-	Dynamic Time Warping
DWT	-	Discrete Wavelet Transform
DWTFS	-	Discrete Wavelet Transform Feature Set
EIS	-	Evolving Intelligent System
EP	-	Expected number of Phonemes
ESPNet	-	Self Evolving and Parameter adaptation Network
FFT	-	Fast Fourier Transform
FIR	-	Finitive Impulse Response

FN	-	False Negatives
FNN	-	Fuzzy Neural Network
FOS	-	Factor of Safety
FP	-	False Positives
GBDBN	-	Gaussian - Bernoulli DBN
GD	-	Gradient Descent
GMM	-	Gaussian Mixture Model
GPU	-	Graphical Processing Unit
HFCC-E	-	Human Factor Cepstral Co-efficients – Equivalent Rectangular Bandwidth
HMM	-	Hidden Markov Model
ID3	-	Iterative Dichotomiser 3
IIR	-	Infinite Impulse Response
IL	-	Indian Lanugages
IP	-	Inserted Phonemes
K-nn	-	K-Nearest Neighbour
LDA	-	Linear Discriminant Analysis
LDC	-	Linguistic Data Consortium
LFB	-	Low Frequency Band
LFCC	-	Linear Frequency Cepstral Co-efficients
LPC	-	Linear Predictive Coding
LSTM	-	Long Short Term Memory
LVCSR	-	Large Vocabulary Continuous Speech Recognition
LVSR	-	Large Vocabulary Speech Recognition
MC	-	Mis-Classified Phonemes
McSLM	-	Metacognitive Scaffolding Learning machine
MDR	-	Multimedia Document Recognition
MFCC	-	Mel- Frequency Cepstral Co-efficients
MLP	-	Multi Layer Perceptron

MMF	-	Maximum Mutual Information
MMI	-	Maximum Mutual Information
MRI	-	Magnetic Resonance Imaging
MSE	-	Mean Square Error
MWP-ACE	-	Mixed Wavelet Packet Advanced Combinational Encoder
NLP	-	Natural Language Processing
NMPSO	-	New Method Particle Swarm Optimization
NN	-	Neural Network
NUSDWT	-	Non-uniform segmentation using DWT
PCA	-	Principle Component Analysis
pClass+	-	Parsimonious Classifier+
PER	-	Phoneme Error Rate
PLP	-	Perceptual Linear Prediction
PSO	-	Particle Swarm Optimization
RAST	-	Rapid Annotation using Subsystem Technology
RBM	-	Restricted Boltzmann Machine
ReLU	-	Rectified Linear Unit
RIVMcSLM	-	Recurrent Interval-Valued McSLM
RM	-	Resource Management
RMSE	-	Root Mean Square Error
RNN	-	Recurrent Neural Network
ROC	-	Receiver-Operating Characteristic Curve
SBC	-	Subband based Cepstral Parameter
SCARF	-	Segmental Conditional Random Fields
SGPSO	-	Second Generation Particle Swarm Optimization
SMOTE	-	Synthetic minority over-sampling technique
ST2Class	-	Scaffolding Type-2 classifier
STFT	-	Short Term Fourier Transform

SVM	-	Support Vector Machine
TI	-	Texas Instruments
TN	-	True Negatives
TP	-	True Positives
TPSO	-	Temperature controlled Particle Swarm Optimization
VAD	-	Voice activity detection
VE	-	Voting Expert
VOP	-	Vowel Onset Point
WER	-	Word Error Rate
WMSE	-	Weighted Mean Square Error

LIST OF SYMBOLS

θ	-	Phase spectrum
ω	-	Phase shift
τ	-	Group delay function
S	-	Speech signal
s(n)	-	da
ϕ	-	Wavelet function
m	-	Resolution level while applying wavelet transform
$\mathbb{G}$	-	Graph (Multigraph)
$\mathbb{V}$	-	Set of vertices of graph G
$\mathbb{E}$	-	Set of edges of graph G
cut()	-	Degree of dissimilarity between two graphs
w(u,v)	-	Similarity between two nodes *u* and *v* in a Graph/weight of edge *(u,v)*
Ncut()	-	Normalized cut – measure of disassociation between two subgraphs
assoc()	-	Measure of association between two subgraphs
Nassoc()	-	Measure of normalized association between two subgraphs
F	-	Set of feature vectors of speech S
ζ	-	Distance factor/node distance
W	-	Weight matrix of graph
D	-	Diagonal matrix where each d_i represents the total weight i^{th} node
ε_i	-	i^{th} value in the sorted list of eigenvalues
E_i	-	i^{th}eigen vector

Y	-	Filtered speech
b_0, b_1, b_2	-	Co-efficients of second order filter
db2	-	Daubachies wavelet
θ	-	Model parameters
V	-	Number of visible nodes in RBM
H	-	Number of hidden nodes in RBM
E()	-	Energy function of neurons
w_{ij}	-	Connection weight between i^{th} node in visible layer to j^{th} node in hidden layer of an RBM
b_i	-	Bias of i^{th} visible neuron
a_j	-	Bias of j^{th} hidden neuron
$\mathcal{N}$	-	Gaussian function
$p()$	-	Conditional probability distribution
Δw_{ij}	-	Change in weight parameter for connection between i^{th} visible node and j^{th} hidden node
$\langle v_i h_j \rangle_{cd}$	-	Measured frequency of visible units for given training data
$\langle v_i h_j \rangle_{rd}$	-	Measured frequency of visible units with reconstructed data
Σ	-	Sigmoid function
x_i	-	Position of i^{th} particle in PSO
M	-	Population size in PSO
T	-	Time/iteration in PSO
$v_i(t)$	-	Velocity of i^{th} particle at time t in PSO
p_g	-	Global position in PSO
r_1, r_2, r_3	-	Uniformly distributed random variables in PSO

c_1	-	Local acceleration co-efficient
c_2	-	Global acceleration co-efficient
ω	-	Inertia weight
v_{min}	-	Lower limit for velocity of particles in PSO
v_{max}	-	Upper limit for velocity of particles in PSO
$\bar{P}$	-	Geometric centre of the particle swarm
T	-	Time frequency
c_3	-	Geometric centre acceleration co-efficient
ω_1	-	Initial intertia
ω_2	-	Final intertia
MAXITER	-	Maximum number of iterations in PSO
Iter	-	Current iteration in PSO
h()	-	Temperature function
I_{min}	-	Lower bound vector of decision variables in PSO
I_{max}	-	Upper bound vector of decision variables in PSO
P	-	Population in PSO
L	-	Depth of DBN/Number of layers in DBN
L-1	-	Number of RBMs forming DBN
W_i	-	Connection weight of i^{th} RBM
B_i	-	Biases of hidden layer of i^{th} RBM
N_i	-	Number of neurons in i^{th} layer of DBN
$\aleph_i$	-	DBN built using i^{th} particle in the population
R	-	RBM
q	-	Number of decision parameters in PSO

n_i	-	Number of biases in i^{th} layer
C	-	Number of output classes
$o_i^{\hat{}}$	-	Desired output
$y_i^{\hat{}}$	-	Observed output
m	-	Number of training samples
δ	-	Error
Γ	-	Learning rate
β_n	-	Influence term of n^{th} class
$\mathbb{C}_n$	-	Number of instances of class n in training dataset
I	-	Total number of instances in training dataset
O	-	Sequence of observations
w_i	-	i^{th} word
$P(w_i)$	-	Prior probabilities
$P(O\|w_i)$	-	Likelihood
B_i	-	Co-efficients of hyperplane
X_i	-	Variables or datapoints of hyperplane
$K(x, X^i)$	-	Kernel function on input vector x and support vector X^i
gamma	-	Influence of individual sample
P_i, p(I)	-	Probability of i^{th} class