

# **Sparse Spectrum Enabled Swarm Decomposition for Brain Signal Processing**

**Ph.D. Thesis**

by

**Shailesh Vitthalrao Bhalerao**



**Department of Biosciences and Biomedical Engineering  
INDIAN INSTITUTE OF TECHNOLOGY INDORE**

**MARCH 2025**





# **Sparse Spectrum Enabled Swarm Decomposition for Brain Signal Processing**

**A Thesis**

*Submitted in partial fulfillment of the  
requirements for the award of the degrees  
of*  
**Doctor of Philosophy**

*by*

**Shailesh Vitthalrao Bhalerao**



**Department of Biosciences and Biomedical Engineering  
INDIAN INSTITUTE OF TECHNOLOGY INDORE**

**OCTOBER 2024**

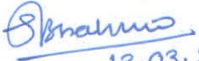




## INDIAN INSTITUTE OF TECHNOLOGY INDORE

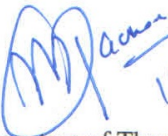
I hereby certify that the work which is being presented in the thesis entitled **Sparse Spectrum Enabled Swarm Decomposition for Brain Signal Processing** in the partial fulfillment of the requirements for the award of the degree of **Doctor of Philosophy** and submitted in the **Department of Biosciences and Biomedical Engineering, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from August 2020 to September 2024 under the supervision of Dr. Ram Bilas Pachori, Professor, Indian Institute of Technology Indore, Indore, India.

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

  
13.03.2025


Signature of the student with date  
(Shailesh Vitthalrao Bhalerao)

-----  
This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

  
13.03.2025

Signature of Thesis Supervisor with date  
(Prof. Ram Bilas Pachori)

-----  
Shailesh Vitthalrao Bhalerao has successfully given his Ph.D. Oral Examination held on  
11.03.2025.

  
13.03.2025

Signature of Thesis Supervisor with date  
(Prof. Ram Bilas Pachori)



## ACKNOWLEDGEMENTS

I would like to take this opportunity to express my heartfelt gratitude to a number of persons who in one or the other way contributed by making this time as learnable, enjoyable, and bearable. At first, I would like to thank my supervisor **Prof. Ram Bilas Pachori**, who was a constant source of inspiration during my work. Without his constant guidance and research directions, this research work could not be completed. His continuous support and encouragement has motivated me to remain streamlined in my research work.

I am thankful to **Prof. Neminath Hubballi** and **Prof. Kiran Bala**, my research progress committee members for taking out some valuable time to evaluate my progress all these years. Their valuable comments and suggestions helped me to improve my work at various stages. I am also grateful to Head of Department of Biosciences and Biomedical Engineering for his help and support.

My sincere acknowledgement to **Prof. Suhas Joshi**, Director, Indian Institute of Technology Indore for providing me the opportunity to explore my research capabilities at Indian Institute of Technology Indore.

I would like to appreciate the fine company of my dearest colleagues and friends especially, Kritiprasanna Das, Nabasmita Phukan, Ashok Mahato, Vivek Kumar Singh, Aditya Nalwaya, Achinto Mondal, Krishkant Sharma, Shiva Reddy, and Amishi Vijay. I am also grateful to the Institute Staffs for their unfailing support and assistance, and the All India Council for Technical Education for funding the PhD research.

I remember my late father, Vitthalrao Ukanda Bhalerao, whose memory continues to inspire me. I would like to express my heartfelt respect and gratitude to my mother for her unwavering love, care, and support throughout my life. I am also deeply grateful to my wife, Aarti, as this thesis would not have been possible without her encouragement and assistance.

Last but not the least, I thank the God Almighty for giving me the opportunity, strength and ability to perform this research work and complete it satisfactorily.

*Shailesh Vitthalrao Bhalerao*



*Dedicated*  
*to*  
*My Parents and Teachers*





## ABSTRACT

Brain signal analysis has a crucial role in the investigation of neuronal activity for diagnosis of neurological diseases and brain-computer interfaces (BCI) applications. The electroencephalogram (EEG) and magnetoencephalogram (MEG) are the most efficient brain signals that assist in the diagnosis of neurological diseases and also play an essential role in all neurosurgery related to the brain. Furthermore, it is a valuable tool for clinicians and researchers and has the potential to develop advanced BCIs and help in partial paralysis rehabilitation. The clinicians visually study these brain signal recordings to determine the manifestation of abnormalities in the brain. Hence adequate and accurate brain signal analysis of these signals is always a major challenge for researchers. The complex and non-stationary nature of EEG and MEG signals demands advanced signal analysis techniques for proper interpretation and extraction of brain information. In this perspective, adaptive nonstationary signal-processing algorithms like swarm decomposition (SWD) and iterative filtering provide useful signal representation for EEG and MEG, enabling easy interpretation and feature extraction. Consequently, the primary focus of this thesis is to propose a novel extension of the univariate SWD-based nonstationary signal-processing technique to decompose multichannel signals for brain signal analysis, and to further apply these techniques in the development of neurological disease detection and BCI frameworks.

In this thesis, we have proposed various signal processing techniques, including swarm sparse decomposition method (SSDM), multivariate SSDM (MSSDM), enhanced SSDM (ESSDM), and clustering SSDM (CSSDM) for robust nonstationary brain signal analysis. These techniques adopt newly designed optimized sparse spectrum representation and filter banks for efficient extraction of components from multi-channel nonstationary signals and time-frequency resolution improvement. Additionally, we have also developed a sleep apnea disease detection framework and BCI frameworks based on these techniques, such as cognitive visual object recognition using EEG and MEG signals, EEG-based motor imagery recognition, EEG-based imagined speech task detection, and EEG-based upper limb movements detection. We have also designed and studied different feature techniques like Riemann's correlation-assisted fusion feature, low-dimensional joint time-frequency deep

feature, time-frequency graph spectral, fused time-frequency graph features along with the decomposition method. The classification performance of proposed frameworks have shown better results than the state-of-art methods.

**Keywords:** Swarm sparse decomposition method, Sparse spectrum, Electroencephalogram, Magnetoencephalogram, Brain disease diagnosis, Brain-computer interface, and Machine learning.

# Contents

<b>List of Figures</b>	<b>v</b>
<b>List of Tables</b>	<b>xii</b>
<b>List of Abbreviations</b>	<b>xiii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Brain signal processing . . . . .	1
1.2 Applicability of recorded EEG and MEG signals . . . . .	3
1.2.1 Clinical application . . . . .	3
1.2.1.1 Sleep analysis in sleep apnea disorder . . . . .	3
1.2.2 Brain–computer interface and rehabilitation applications . . . . .	5
1.2.2.1 Cognitive visual objects analysis . . . . .	5
1.2.2.2 Imagined speech decoding . . . . .	6
1.2.2.3 Motor imagery BCI . . . . .	7
1.3 Nonstationary signal analysis techniques . . . . .	8
1.4 Motivation . . . . .	12
1.5 Objectives . . . . .	14
1.6 Thesis contributions . . . . .	15
1.7 Organization of the thesis . . . . .	18
<b>2 SSDM Method for Robust Nonstationary Signal Analysis with EEG-based Sleep Apnea Detection</b>	<b>21</b>
2.1 Introduction . . . . .	22
2.2 Proposed methodology . . . . .	25

2.2.1	SFT spectrum coefficients generation and computation . . . . .	25
2.2.2	Spectrum smoothing . . . . .	29
2.2.3	SWD-based SFT spectrum boundary estimation and band selection .	30
2.2.3.1	Signal description . . . . .	30
2.2.3.2	Initialization of swarm model parameters . . . . .	31
2.2.3.3	SWD-based spectrum sensing and estimation . . . . .	31
2.2.4	Performance measure . . . . .	34
2.3	Results and Discussions . . . . .	35
2.3.1	Example 1: AM nonstationary signal . . . . .	35
2.3.2	Example 2: FM nonstationary signal . . . . .	39
2.3.3	Study on real EEG signals . . . . .	41
2.4	Summary . . . . .	52

### **3 Multivariate SSDM for Multivariate Signal Analysis with Cognitive Visual Object Detection from Multichannel EEG-MEG Signals** **55**

3.1	Introduction . . . . .	56
3.2	Multichannel dataset . . . . .	59
3.3	Proposed methodology . . . . .	59
3.3.1	Preprocessing . . . . .	60
3.3.2	Multivariate extension of swarm-sparse decomposition . . . . .	61
3.3.2.1	Multivariate extension of SSDM . . . . .	62
3.3.2.2	Ablation study of the SSDM algorithm . . . . .	64
3.3.3	Feature extraction and selection . . . . .	70
3.3.3.1	Common spatial filter on Teager energy . . . . .	70
3.3.3.2	Correntropy spectral density . . . . .	71
3.3.3.3	Rényi's entropy . . . . .	71
3.3.3.4	Sparse entropy . . . . .	71
3.4	Results and discussion . . . . .	73
3.5	Summary . . . . .	80

<b>4</b>	<b>Multivariate SSDM-based Joint Time-Frequency Analysis with Improved Imagined Speech Decoding for Intuitive BCI</b>	<b>83</b>
4.1	Introduction . . . . .	84
4.2	Dataset and experimental paradigm used . . . . .	87
4.3	Proposed methodology . . . . .	88
4.3.1	MSSDM-based multivariate time-frequency representation . . . . .	88
4.3.2	Feature formulation and classification . . . . .	93
4.3.2.1	Canonical correlation analysis . . . . .	94
4.3.2.2	Hellinger distance-based correlation . . . . .	95
4.4	Results . . . . .	95
4.5	Discussions . . . . .	104
4.6	Summary . . . . .	107
<b>5</b>	<b>Enhanced SSDM with Multi-Class Motor Imagery-based EEG-BCI System</b>	<b>109</b>
5.1	Introduction . . . . .	109
5.2	Dataset and experimental paradigm . . . . .	113
5.3	Proposed methodology . . . . .	113
5.3.1	Enhanced sparse swarm decomposition method . . . . .	115
5.3.2	Adoption of SHO for optimal threshold parameter tuning . . . . .	117
5.3.3	Comparison with existing other methods . . . . .	118
5.3.4	Feature formulation and classification . . . . .	120
5.3.5	Evaluating classifier metrics and validation scheme . . . . .	122
5.4	Results and discussions . . . . .	124
5.4.1	Performance comparison with specific-subject benchmarks . . . . .	126
5.4.2	Performance comparison with cross-subject benchmarks . . . . .	129
5.4.3	Interpretability of the learned relevant fused features representations versus classification performance . . . . .	131
5.5	Summary . . . . .	134

<b>6</b>	<b>Clustering SSDM with MI-EEG based Upper Limb Movement Recognition for BCI Application</b>	<b>137</b>
6.1	Introduction . . . . .	138
6.1.1	Existing baseline approaches . . . . .	138
6.2	Proposed methodology . . . . .	139
6.2.1	Dataset used . . . . .	140
6.2.2	Clustering sparse swarm decomposition method . . . . .	140
6.2.2.1	DBSC algorithm with CMI measure . . . . .	141
6.2.3	Feature formulation and classification . . . . .	142
6.3	Results and Discussions . . . . .	143
6.3.1	Effect of optimal clustering channel-based features . . . . .	143
6.3.2	Effect of rhythm-based features . . . . .	146
6.4	Summary . . . . .	147
<b>7</b>	<b>Conclusions and future scope</b>	<b>149</b>
7.1	Conclusions . . . . .	150
7.2	Future directions . . . . .	151
	<b>List of Publications</b>	<b>173</b>

# List of Figures

1.1	The structure of the thesis. . . . .	20
2.1	The systematic block diagram of the proposed SSDM method to obtain ESR. . . . .	26
2.2	Plots of (a) coefficients magnitude ( $ c_m $ ) versus frequency (Hz) with sparse spectrum estimation and (b) spectral analysis of FFT versus SFT coefficients after SGF (averaging filter) in case multicomponent AM signal analysis. . . . .	28
2.3	Visual description of (a) detected boundaries in the SFT spectrum and (b) filter bank (FB1-FB5) generated for $s_I(n)$ multicomponent signal. . . . .	36
2.4	Plots of $s_I(n)$ : (a) TFR of input AM signal, (b) HHT-based TFR, (c) SWD-based TFR, and (d) SSDM-based TFR. . . . .	36
2.5	Visual description of (a) detected boundaries in the SFT spectrum and (b) filter bank (FB1-FB5) generated for $s_{II}(n)$ multicomponent signal. . . . .	38
2.6	Plots of $s_{II}(n)$ : (a) TFR of input AM signal with AWGN (-10dB), (b) HHT-based TFR, (c) SWD-based TFR, and (d) SSDM based TFR. . . . .	39
2.7	Visual description of (a) detected boundaries in the SFT spectrum and (b) filter bank (FB1-FB3) generated for $s_{III}(n)$ multicomponent signal. . . . .	40
2.8	Plots of $s_{III}(n)$ : (a) TFR of input FM signal, (b) HHT-based TFR, (c) SWD-based TFR, and (d) SSDM-based TFR. . . . .	41
2.9	Visual description of (a) detected boundaries in the SFT spectrum and (b) filter bank (FB1-FB3) generated for $s_{IV}(n)$ multicomponent signal. . . . .	41
2.10	Plots of $s_{IV}(n)$ : (a) TFR of input FM signal with AWGN (-10dB), (b) HHT-based TFR, (c) SWD-based TFR, and (d) SSDM based TFR. . . . .	42
2.11	Plots of (a) epoch corresponding to apnea patient and (b) epoch corresponding to non-apnea subject. . . . .	44

2.12	Schematic representation of (a) the proposed sleep apnea recognition model based on the SSDM method, (b) CNN unit block, and (c) BiLSTM unit block.	47
3.1	The timing scheme of the experimental paradigm for the visual object database. . . . .	59
3.2	Block diagram of the proposed approach for visual object recognition using the MSSDM method. . . . .	60
3.3	Channel-specific multivariate analysis from raw EEG channels [EEG005 (FC2), EEG042 (T8), EEG003 (FC1), EEG041 (T7), and EEG036 (F4) for cognitive visual objects. The highlighted yellow boxes in the plots demonstrate mutual patterns of the imagined stimulus responses. . . . .	62
3.4	Filter bank structure for MSSDM-based swarm filter bank of 4-channel EEG and MEG signals with selected OCs. . . . .	64
3.5	MSSDM decomposition of 4-channel MEG signal and selected OCs (OC1, OC2, OC3, OC4, OC5, and OC6). . . . .	65
3.6	Decomposition of multivariate of 3-channel synthetic signal with similar frequency modes using MSSDM method in (a) FFT without SGF, (b) FFT with SGF, (c) SFT without SGF, and (d) SFT with SGF approaches. . . . .	67
3.7	Plots of (a) nonstationary multicomponent AM signal ( $s_{c1}(n)$ ) and corresponding spectral analysis using (b) FFT without SGF, (c) FFT with SGF, (d) SFT without SGF, and (e) SFT with SGF approaches. . . . .	68
3.8	Plots of (i) raw multivariate MEG signals and (ii) their multivariate TFRs of the $\alpha$ - $\beta$ rhythms from three selected channels (FC3, T8, and F4 of subject 1) using proposed MSSDM method for visual object classes (a) HB, (b) AB, (c) HF, (d) AF, and (e) IO, respectively. . . . .	72
3.9	Plots of averaged classification accuracies of 5-class visual imagery for the 16 subjects from the ( $\alpha$ - $\beta$ ) rhythm-based FF features using (a) MSSDM, (b) USSDM, and (c) BPF approaches. . . . .	76
3.10	Overall performance using different feature schemes. . . . .	78
4.1	Experimental paradigm for IMS recording. . . . .	88



4.2	Block diagram of the proposed approach for IMS recognition using MSSDM method. . . . .	89
4.3	Plots of (a)-(f) raw EEG signals and their spectral analysis using FFT with SFT coefficients, (g)-(h) MSSDM-based decomposed CAOCs and its combined visual interpretation, and (i) MSSDM-based swarm filter bank structure of 3-channel (F3, F4, and P3) IMG-EEG signals for the “Right” class (subject 2) in Coretto dataset. . . . .	92
4.4	Proposed architecture for deep features extraction. . . . .	94
4.5	The JTF representation of the IMS word 5-class data using proposed MSSDM for 15 subjects (BCI Competition dataset). The highlighted yellow boxes in the plots demonstrate similar imagined stimulus responses with extracted $\theta$ and $\alpha$ bands. . . . .	97
4.6	Plots of: (i)-(iii) decomposed CAOCs from 3-channel EEG signals and (iv) their JTF analysis for IMS classes (a) Hello, (b) Help me, (c) Stop, (d) Thank you, and (e) Yes, respectively. . . . .	98
4.7	Subject-wise evaluation of 10-fold classification accuracy performed by two different feature correlation schemes, (a)-(c) HDC and (b)-(d) CCA based on Coretto dataset and BCI Competition dataset, respectively. The error bars denote the standard deviation of values $< 1\%$ . . . . .	99
4.8	Averaged confusion matrix of the classification across all subjects in (a) the Coretto dataset and (b) the BCI Competition dataset. . . . .	99
4.9	Cross-validation of average accuracy plotted against each CF value for four selected deep features in case of (a) CCA with Coretto dataset, (b) HDC with Coretto dataset, (c) CCA with BCI Competition dataset, and (d) HDC with BCI Competition dataset. . . . .	101
4.10	k-fold cross validation across all subjects in CCA approach on (a) the Coretto dataset and (b) the BCI Competition dataset. . . . .	105
4.11	Comparison of improvement of averaged decoding accuracies (in %) during imagination across all subjects. . . . .	106

5.1	Block diagram of the proposed framework for MI-EEG recognition using ESSDM method. . . . .	114
5.2	Experimental paradigm for MI-EEG recording. . . . .	114
5.3	Plots of (a) SWD-based decomposed OCs with different threshold parameter selection and (b) obtained TFRs using HSA. . . . .	116
5.4	Plots of (a)-(b) convergence curves, and (c) fitness values for each iteration showing the best selection of threshold parameters ( $P_{th} = 0.18$ and $StdTh = 0.15$ ) for optimal mode extraction using ESSDM. . . . .	119
5.5	Plots of (a) nonstationary AM input signal and decomposition results using (b) EMD, (c) Q-VMD, (d) SWD, and (e) proposed ESSDM method. . . . .	121
5.6	The extracted FTFG features based adjacency matrix from TF image for four MI-EEG classes. . . . .	122
5.7	Averaged confusion matrix across all subjects using (a) ESSDM-FTFG-ELM, (b) ESSDM-FTFG-CNet, and (c) ESSDM-FTFG-CapsNet on the BCI IV-2a dataset. . . . .	125
5.8	The classification performance of cross-subject benchmarks ( $ACC_{avg}$ , $\kappa$ ) on the BCI IV-2a and III-4a datasets. . . . .	126
5.9	Average accuracy obtained in specific-subject (a) bi-class and (b) multi-class on IV-2a, and cross-subject (c) and (d) for IV-2a and III-4a, respectively. . . . .	126
5.10	The effect of a sub-clustered factor in (a) bi-class specific-subject and (b) cross-subject performance on the IV-2a dataset. . . . .	127
5.11	Performance comparison with the proposed method with the benchmark baseline models. . . . .	132
6.1	Block diagram of the proposed automated CSSDM-based upper-limb movement recognition from MI-EEG sensors data. OCk denotes the oscillatory component. . . . .	139
6.2	The extracted TFGS features from TF images for (a) EE, (b) EF, (c) HC, (d) HO, (e) FS, (f) FP, and (g) rest MI-EEG classes. . . . .	142

6.3	Performance comparison of proposed CSSDM-based models for optimal channel selection in (a) all rhythms and (b) rhythm-specific. . . . .	145
6.4	The obtained confusion matrix using (a) CSSDM-TFGS-SAViT and (b) CSSDM-TFGS-ConvNet models across all subjects. . . . .	146

# List of Tables

2.1	Summary of related work on sleep apnea detection techniques. . . . .	24
2.2	The mathematical expression of synthesized test signals. . . . .	35
2.3	Comparison of the ESR values computed for reconstructed monocomponent signals of $s_I(n)$ using HHT, SWD, and SSDM methods. . . . .	37
2.4	Comparison of the ESR values computed for reconstructed monocomponent signals of $s_{II}(n)$ using HHT, SWD, and SSDM methods. . . . .	38
2.5	Comparison of the ESR values computed for reconstructed monocomponent signals of $s_{III}(n)$ using HHT, SWD, and SSDM methods. . . . .	40
2.6	Comparison of the ESR values computed for reconstructed monocomponent signals of $s_{IV}(n)$ using HHT, SWD, and SSDM methods. . . . .	42
2.7	Extracted the number of apnea and non-apnea epochs from EEG signals corresponding to every subject. . . . .	44
2.8	The statistical analysis of most significant features chosen using student t-test for SSDM, SWD, and HHT methods. . . . .	45
2.9	Architectures of the CNN (Alexnet) and BiLSTM. . . . .	46
2.10	Feature specific cross-validation performance obtained using the SSDM, SWD, and HHT methods. . . . .	49
2.11	Performance comparison with the methods available in the literature. . . . .	50
2.12	The most significant five feature fusion chosen using student t-test for SSDM, SWD, and HHT methods. . . . .	50

3.1	Comparison of the obtained ESR values of $x(n)$ for AM reconstructed monocomponent signals $s_1(n)$ , $s_2(n)$ , $s_3(n)$ , $s_4(n)$ , and $s_5(n)$ using FT (without SGF), FT (with SGF), SFT (without SGF), and SFT (with SGF) methods. Note: '-' indicates the blank ESR value due to a non-used signal in the respective channel. . . . .	69
3.2	Classification accuracy (in %) of combined $(\alpha\text{-}\beta)$ rhythms-based fusion features computed from extracted OCs using MSSDM, USSDM, and BPF with Riemann's correlation analysis. $Q$ denotes the correlation coefficient value. The best accuracy is marked in boldface. . . . .	74
3.3	Feature-specific average classification performance obtained from subject A01-16. The best accuracy is marked in boldface. . . . .	78
3.4	Performance comparison against state-of-the-art methods available in the literature. . . . .	79
4.1	Hyper-parameter configuration of four different models. . . . .	94
4.2	Architecture of used pre-trained neural networks with obtained deep features. . . . .	96
4.3	Average accuracy score (in %) obtained from reduced features (in %) for four different pre-trained neural networks using CCA scheme and SVM classifier on the Correto dataset. . . . .	100
4.4	Average accuracy score (in %) obtained from reduced features (in %) for four different pre-trained neural networks using HDC scheme and SVM classifier on the Correto dataset. . . . .	102
4.5	Average accuracy score (in %) obtained from reduced features (in %) for four different pre-trained neural networks using CCA scheme and SVM classifier on the BCI Competition dataset. . . . .	103
4.6	Average accuracy score (in %) obtained from reduced features (in %) for four different pre-trained neural networks using HDC scheme and SVM classifier on the BCI Competition dataset. . . . .	104
4.7	Performance comparison with the other state-of-the-art methods. . . . .	106
5.1	Parameters initialization SHO optimizer for threshold parameter tuning . . .	117

5.2	Assessment of ESR and MSE metrics for AM reconstructed monocomponent signals 1 <sup>st</sup> , 2 <sup>nd</sup> , 3 <sup>rd</sup> , 4 <sup>th</sup> , and 5 <sup>th</sup> using EMD, SWD, Q-VMD, and proposed ESSDM method. (Note: '-' represents missing values for non-constructed components). . . . .	120
5.3	Average accuracy score (in %) and kappa ( $k$ ) in both MI-EEG labeled scenarios on the BCI IV-2a dataset. The best individual values are marked in bold. Notation NR denotes non-reported values. . . . .	130
5.4	Cross-subject classification accuracy in cross-subject scenario on the BCI IV-2a and III-4a datasets. . . . .	131
6.1	The obtained classification performance (in %) for different optimal channel selection of cross-subject on the BNCI Horizon 7-class database. . . . .	143
6.2	The cross-subject performance (in %) of obtained features from $\theta$ , $\alpha$ , and $\beta$ rhythms with distinct cortical regions on on the BNCI Horizon 7-class database. . . . .	145
6.3	Performance comparison with the existing baseline models. . . . .	147

## List of Abbreviations and Acronyms

**ACC:** Accuracy

**ACCavg:** Average accuracy

**AHI:** Apnea-hypopnea index

**ALS:** Amyotrophic lateral sclerosis

**ANN:** Artificial neural network

**AM:** Amplitude modulated

**AWGN:** Additive white Gaussian noise

**BPF:** Band-pass filter

**BoF:** Bag of features

**BCI:** Brain-computer interface

**BNL:** Batch normalization layer

**CAOC:** Channel-aligned oscillatory component

**CCA:** Canonical correlation analysis

**CCM:** Covariance-based connectivity measure

**CSD:** Correntropy spectral density

**CF:** Critical features

**CL:** Convolution layer

**CMI:** Canonical correlation analysis-based mutual information

**ConvNet:** Convolutional neural network

**CNN-LSTM:** Convolutional neural network-long short-term memory

**CNet:** Compact CNN

**CSA:** Central sleep apnea

**CSSBP:** Common spatial-spectral boosting pattern

**CSSSP:** Common sparse spectral-spatial pattern

**CV:** Cross-validation

**CSP:** Common spatial filter

**CSSDM:** Clustering sparse swarm decomposition method

**DBSC:** Density-based spatial clustering

**DTCWT:** Dual-tree complex wavelet transform

**DFBCSP:** Discriminative filter bank common spatial pattern

**DFT:** Discrete Fourier transform

**DOL:** Dropout layer

**DWT:** Discrete wavelet transform

**ECG:** Electrocardiogram

**ECoG:** Electrocorticogram

**EEG:** Electroencephalogram

**ELM:** Extreme learning machine



**EMD:** Empirical mode decomposition

**EEMD:** Ensemble empirical mode decomposition

**ESSDM:** Enhanced SSDM

**EMG:** Electromyogram

**EOG:** Electrooculogram

**EVD:** Eigenvalue decomposition

**EWT:** Empirical wavelet transform

**FBSE:** Fourier-Bessel series expansion

**FAWT:** Flexible analytic wavelet transform

**FM:** Frequency modulated

**fMRI:** Functional magnetic resonance imaging

**Fused-Lasso:** Fused least absolute shrinkage and selection operator

**fNIRS:** Functional near-infrared spectroscopy

**FTFG:** Fused TF graph

**GA:** Genetic algorithm

**GNB:** Gaussian Naïve Bayes

**HDC:** Hellinger distance-based correlation

**HFD:** Higuchi and Katz's fractal dimension

**HJ:** Hjorth

**HPF:** High-pass filter

**HHT:** Hilbert-Huang transform

**IA:** Instantaneous amplitude

**IF:** Instantaneous frequency

**IMF:** Intrinsic mode function

**IMG-EEG:** Imagined speech electroencephalogram

**IMS:** Imagined speech

**JIA:** Joint instantaneous amplitude

**JIF:** Joint instantaneous frequency

**JTFA:** Joint time-frequency analysis

**JTFR:** Joint time-frequency representation

**JTFDF:** Joint time-frequency deep feature

**KNN:** K-nearest neighbors

**LASSO:** Least absolute shrinkage and selection operation

**LDA:** Linear discriminant analysis

**LOOCV:** Leave one out cross validation

**LSTM:** Long-short term memory

**LIS:** Locked-in syndrome

**MDM:** Minimum distance to mean

**ME:** Movement execution

**MEG:** Magnetoencephalogram

**MI:** Motor imagery

**MI-EEG:** Motor imagery electroencephalogram

**MCCM:** Maximum linear cross-correlation-based connectivity measure

**mFBTRCA:** Multi-class filter bank task-related component analysis

**MDM:** Minimum distance to mean

**MS-AMF:** Multi-scale fusion attention mechanism

**MSFT:** Metric-based spatial filtering transformer

**MPL:** Max-pooling layer

**MSSDM:** Multivariate swarm-sparse decomposition method

**NIRS:** Near-infrared spectroscopy

**NDI:** Normalized Dunn index

**OC:** Oscillatory component

**ONC:** Optimal neuronal components

**OS:** Optimum sample points

**OSA:** Obstructive sleep apnea

**OSPs:** Optimum sample points

**PLSP:** Peak latency time points

**PET:** Positron emission tomography

**PLTP:** Phase locking tracking potential

**PSG:** Polysomnography

**PSD:** Power spectral density

**PPG:** Photoplethysmography

**RASM:** Rational asymmetry

**RD:** Riemannian distance

**RCFF:** Riemann's correlation-assisted fusion feature

**RE:** Renyi entropy

**RDM:** Representational dissimilarity matrix

**RLDA:** Regularized linear discriminant analysis

**RASM:** Rational asymmetry

**RLU:** Rectified linear unit

**RSA:** Representational similarity analysis

**SAViT:** Self-attention vision transformer

**SFT:** Sparse Fourier transform

**SPoC:** Source power comodulation

**SVM:** Support vector machine

**SHO:** Selfish herd optimization

**SSA:** Singular spectrum analysis

**SpO2:** Oxygen saturation

**SSDM:** Swarm sparse decomposition method

**SWD:** Swarm decomposition

**TF:** Time-frequency

**TFR:** Time-frequency representation

**TFGS:** Time-frequency graph spectral

**TQWT:** Tunable-Q wavelet transform

**TSLDA:** Tangent space linear discriminant analysis

**USSDM:** Univariate SSDM

**VMD:** Variational mode decomposition

**VSA:** Visual stimulus activation

**WT:** Wavelet transform

**WeiRD:** Weighted robust distance

**WPT:** Wavelet packet transform

**WeiRD:** Weighted robust distance

# Chapter 1

## Introduction

This chapter provides a brief overview of the thesis, which includes objectives and key contributions of the work. First, we discuss the need for brain signal processing, followed by the motivation for using electroencephalogram (EEG) and magnetoencephalogram (MEG) signals in neurological disease diagnosis and brain-computer interface (BCI) applications. Next, we present an overview of nonstationary signal processing techniques and a related key literature survey with its advantages and shortcomings. Additionally, we have discussed the swarm decomposition (SWD) technique with its applications. In the subsequent sections, the objectives and contributions of the thesis are listed in an enumerated manner. Finally, the organization of the thesis is outlined.

### 1.1 Brain signal processing

Brain signals capture the brainwave patterns information from the human brain, which measures the both passive or active mental state [1]. Brain signals capture the information that is processed by millions of neurons present in the brain which resemble neural activities. Analyzing these brain signals helps in understanding and monitoring sensory and motor activities. This brain signal analysis became a significant tool in the investiga-

tion of neuronal activities for diagnosing various neurological diseases and developing BCI applications, ultimately enhancing patient care and treatment outcomes [2]. With the emerging technologies, brain signals are analysed using different signal modalities such as EEG, MEG, functional magnetic resonance imaging (fMRI), and positron emission tomography (PET). Among these signal modalities, the EEG and MEG are the most efficient brain signals to analyse brainwave patterns that assist in the diagnosis of neurological diseases and also play an essential role in all neurosurgery related to the brain. EEG and MEG are the most studied potential brain signals to measure brain activity because of their noninvasive nature and ability to offer enhanced fine-grained analysis by identifying spatial, temporal, and spectral components [3, 4]. EEG records electrical activity from electrodes placed on the scalp by capturing voltage fluctuations from ionic currents, while MEG measures the magnetic fields generated by electrical activity in the brain's neurons. In fMRI and MRI cases, it measures changes in blood flow and oxygen levels in the brain, while PET involves injecting a tracer isotope that emits gamma rays, which are detected by a scanner. However, PET and MRI-based applications are limited compared to EEG and MEG due to their low temporal resolution and show inability to capture various mental activities and brain disorders effectively [5, 6]. Therefore, it has become a valuable tool for clinicians and researchers and has the potential to develop advanced BCIs and help in partial paralysis rehabilitation. The clinicians visually study these brain signal recordings to determine the manifestation of abnormalities in the brain. However, EEG and MEG suffer from artifacts, high complexity, and variations introduced by different brain patterns among individuals, which may lead to erroneous interpretations. This, coupled with numerous challenges such as signal quality, the nonstationary nature of EEG and MEG signals, common artifacts, and feature dimensionality, presents significant challenges for EEG and MEG-based diagnosis of neurological diseases and the development of BCI systems. Therefore, adequate and accurate analysis

of these signals remains a major challenge for researchers and it demands advanced signal analysis techniques for proper interpretation and extraction of useful brain information.

## **1.2 Applicability of recorded EEG and MEG signals**

EEG and MEG signals serve as effective tools for monitoring brain functional activities associated with a wide range of pathological conditions. These non-invasive techniques are highly effective in detecting neural changes linked to various neurological disorders, such as sleep apnea, epilepsy, and other cognitive impairments. They also play a pivotal role in BCI applications by enabling direct communication between the brain and external devices. Extensive research has demonstrated the utility of EEG and MEG in assessing human mental health, monitoring clinical conditions, and capturing complex cognitive activities such as imagination, thought processes, and attention states. These tools provide insights into both healthy and disordered brain functions, offering valuable applications in diagnosis, therapy, and cognitive enhancement. A brief description of specific applications of EEG and MEG signals studied in this thesis is provided below:

### **1.2.1 Clinical application**

#### **1.2.1.1 Sleep analysis in sleep apnea disorder**

Sleep apnea is a severe sleep disorder which is characterized by recurring interruptions in breathing during sleep [7]. It involves progressive disruptions of normal breathing patterns through repeated episodes of apnea (complete pauses) or hypopnea (partial pauses) throughout sleep. Sleep apnea affects 2–7% of the population and primarily causes disrupted sleep, excessive daytime drowsiness, and cognitive difficulties in severe cases, it can lead to serious chronic conditions such as cardiovascular disease, cerebrovascular issues,



diabetes, hypertension, coronary heart disease, and stroke [8, 9, 10]. It is majorly categorized as obstructive sleep apnea (OSA), central sleep apnea (CSA), and mixed sleep apnea (MSA). Among these, OSA is the most common and severe. In most cases, OSA results from physical airway blockages and is more severe than CSA, which is caused by brain signal disruptions, and MSA, a combination of both. Thus, OSA is the primary focus for diagnosis and treatment due to its significant impact on public health. The severity of sleep apnea is assessed using the apnea-hypopnea index (AHI), which measures the combined number of apneas and hypopneas per hour of sleep. An  $AHI < 5$  is indicated as a controlled OSA state. An AHI between 5 and 15 indicates OSA, while an AHI from 15 to 30 represents moderate OSA. An  $AHI > 30$  is indicative of severe OSA. Detecting OSA is challenging due to the need for advanced clinical tools, trained professionals, and overnight monitoring, which complicates early diagnosis and timely treatment. Overnight polysomnography (PSG) is the primary assessment tool for diagnosing OSA which involves the recording of multiple physiological signals such as EEG, electrocardiogram (ECG), electromyogram (EMG), airflow, snoring, and oxygen saturation [11, 12]. However, PSG is expensive, time-consuming, and intrusive, and does not provide suitability to patients due to disrupting natural sleep patterns while recording. This has led to a growing demand for cost-effective and simpler diagnostic methods that can automate sleep apnea detection. Recent research focuses on automated sleep apnea diagnosis using single modalities such as airflow, oxygen saturation ( $SpO_2$ ), and photoplethysmography (PPG). These signals are easy to collect but may lack precision, especially in certain medical scenarios involving patients with irregular breathing patterns [8, 9, 10, 13]. Among these modalities, EEG signals found effective modality in sleep staging and apnea detection which provide accurate classification of sleep states and apnea events. EEG offers distinct advantages in scenarios where other signals may not work, such as during surgeries or in cases of irregular breathing. It captures brain

activity and provides reliable information about sleep stages and related neural patterns, which can help in accurately diagnosing sleep apnea and evaluating sleep quality.

## **1.2.2 Brain–computer interface and rehabilitation applications**

### **1.2.2.1 Cognitive visual objects analysis**

In visual object decoding and reconstruction, neural signals reveal how the human brain processes and interprets natural visual stimuli [3]. These dynamic brain activities cover the human perception of the visual world, which is influenced by the characteristics of external visual stimuli. Among other signal modalities, EEG and MEG are the most widely used tools to study dynamic visual cognitive processes in the human brain [3, 4, 5, 14]. These tools facilitate a broad assessment of brain activity with millisecond-level time resolution and allow researchers to extract complex spatiotemporal dynamics non-invasively in clinical practices. In visual object decoding, MEG and EEG visual stimulus activation (VSA) patterns demonstrate the utmost discriminative cognitive analysis due to their multivariate oscillatory nature [3, 14]. However, high noise in the recorded EEG-MEG signals and subject-specific variability make it extremely difficult to classify the subject’s cognitive responses to different visual stimuli response. Also, in real-world environments, complex visual object recognition faces significant challenges. These include channel selection, smoothing, data reduction, indeterminate scalp regions, and the nonstationary nature of EEG and MEG signals [3, 14]. These issues limit mutually informative features across channels and consequently degrade the classification performance in visual object detection. Therefore, robust multivariate visual analysis is required to decode visual neural patterns from visual object stimuli, and further automated classification methods are needed to accurately classify different visual object classes from the extracted neural patterns of EEG and MEG signals.

### 1.2.2.2 Imagined speech decoding

The imagined speech (IMS) involves imagery of speaking where individuals feel they are generating speech without moving their articulatory movements [15, 16]. It is known to be the most effective and direct communication interface for understanding thoughts. Since speech is the primary means of interaction among humans, IMS becomes a natural and instinctive process of brain communication. In the last few decades, IMS has developed as a common BCI paradigm because it allows for effective communication without external stimuli and requires less training [17, 18, 19]. During the IMS, the brain waves associated with the motor cortex, left superior temporal lobe (Wernicke's area), and left inferior frontal lobe (Broca's area) show increased activation and exhibit more distinctive brain patterns associated with motor imagery (MI) tasks. It is also an endogenous BCI paradigm and established a new BCI communication channel for people who cannot speak due to motor disabilities or neurological issues [19, 20]. Various health issues can lead to a loss of verbal communication which is triggered by injuries or neurodegenerative conditions that affect speech motor production, speech articulation, and language processing [20, 21]. Some neurodegenerative conditions such as stroke trauma and amyotrophic lateral sclerosis (ALS) can lead to locked-in syndrome (LIS) where patients are unable to communicate due to total motor control loss. Therefore, IMS-based BCI serves as an assistive technology that offers a new way to communicate with these patients. Additionally, IMS-based BCI technologies develop a bridge between the brain and the external world enabling a two-way communication interface that reads signals from the brain and transforms them into the desired cognitive tasks. This allows for a thought-to-speech interface to be developed so that individuals unable to speak due to motor disabilities can utilize their brain signals for communication without any physical movement. Several acquisition techniques have been investigated both invasive and noninvasive to acquire brain signals produced during the

speech imagining process. These techniques include MEG, fMRI, functional near-infrared spectroscopy (fNIRS), electrocardiogram (ECOG), and EEG. In certain conditions, invasive methods like ECOG offer higher classification accuracy than non-invasive ones (MEG, fMRI, fNIRS, EEG) for IMS decoding but face limitations in classifying a wider range of IMS classes for practical BCI applications [16, 18, 19, 20, 21, 22]. Among the mentioned techniques for IMS recognition, EEG is the most widely used technique due to its high temporal resolution, affordability, and portability. However, existing EEG-based IMS BCI systems face challenges in real-life scenarios mainly because interpreting EEG signals is difficult due to their low signal-to-noise ratio and nonstationary characteristics. IMS offers a promising paradigm due to its intuitive nature and the potential for multiclass scalability in feasible BCI applications.

### 1.2.2.3 Motor imagery BCI

MI involves the mental simulation of motor activities, which induces changes in activity within the corresponding motor cortex [23, 24]. In MI process, a person imagines a specific movement and the motor cortex exhibits activity as if the movement were actually being performed even though no physical movement takes place. This brain activity can be recorded and used to control biomechatronic devices. For example, imagining a left-arm movement can generate signals that control the movement of the left arm in a full-body exoskeleton. One of the major aspects of MI is that it activates brain areas similar to those involved in the execution of the actual movement, including the sensory and motor cortices [25, 26]. Additionally, the temporal characteristics of these activations are similar to those observed during physical movement [27, 28]. Unlike other BCI paradigms like steady-state visual evoked potentials (SSVEPs) or P300, MI does not require external stimuli or devices, which provides an advantage in terms of user comfort. An MI-based BCI

(MI-BCI) enables the brain to communicate with and control external devices by capturing and interpreting the brain signals generated during the mental imagining of a specific limb movement such as the upper limb and lower. MI-BCI has gained significant importance due to their wide-ranging applications in fields such as neurorehabilitation and neuroprosthetics [29, 30, 31]. In neurorehabilitation, MI-BCI has proven beneficial for neurological patients, especially stroke patients, who frequently experience long-term residual motor impairments in their upper limbs. MI-BCI can help these patients to regain motor function in their upper limbs, which is critical for performing daily tasks. MI-BCI systems have been developed using various brain signal recording techniques, such as EEG, MEG, fMRI, and near-infrared spectroscopy (NIRS) [25, 26, 29, 30, 31, 32, 33, 34]. Among these, EEG is the most commonly used because it is non-invasive and portable. However, since EEG signals are complex and dynamic in nature, the system needs to extract specific MI patterns and interpret these signals for BCI implementation. Another key challenge with MI-BCI is that it requires more user training to generate associated MI brain patterns, and only a limited range of movements can be detected through EEG signals which makes MI-BCI more challenging.

### **1.3 Nonstationary signal analysis techniques**

Nonstationary signal analysis techniques are valuable tools for modeling and interpretation of the time-varying nature of real-life signals recorded in several areas, namely speech synthesis, radar, telecommunication, and biomedicine [35], which are characterised by extreme nonstationary nature. In biomedical applications, the primary motivation for studying multicomponent nonstationary signals is to enhance signal denoising and component extraction from biomedical signals such as ECG [1], EMG [6], EEG [36], MEG [3]. It is observed that the nonstationary biomedical signal that includes spectral content carries inherent time-

varying characteristics of the analyzed signal, which introduces difficulty in analysing real signals. Therefore, an effective decomposition method is required to identify and extract the real-time biological signals that are characterised by extreme nonstationary nature.

In the literature, numerous nonstationary analysis techniques have been proposed to adopt the efficient decomposition of nonstationary multicomponent signals. The majority of such techniques comprise discrete Fourier transform (DFT) based filter bank methods [37, 38] which are not well-suited for analyzing the temporal and spectral characteristics of such signals. The short-time Fourier transform (STFT) [39] employs a moving window approach for time-frequency (TF) representation, but it suffers from a trade-off between time and frequency resolutions due to its fixed window design. To enhance time-frequency representation (TFR) of nonstationary signals, wavelet transform (WT) was proposed which overcomes the drawbacks of STFT. WT extracts signals into multiple oscillatory levels and provides better TF resolution by preserving their transient characteristics [40]. However, WT works on a multi-scale basis, it suffers from issues related to nonstationary signal analysis due to the predetermined filter bank structure and is also inefficient in determining instantaneous amplitude (IA) and instantaneous frequency (IF) from the decomposed sub-band signals. Furthermore, a novel method, empirical mode decomposition (EMD), was proposed by Huang et al. [41] for the analysis of nonlinear and nonstationary signals. The generated intrinsic mode functions (IMFs) are functions of the posterior meaning of physical signal to estimate IA and IF functions. However, the extracted IMFs suffer from a mode mixing problem and cannot effectively reconstruct the actual monocomponent from the signal.

Further, WT-based techniques have been developed, including the flexible analytic wavelet transform (FAWT) [42], tunable-Q wavelet transform (TQWT) [43], and empirical wavelet transform (EWT) [44], wavelet packet transform (WPT) [45], contourlet trans-

form [46], dual-tree complex wavelet transform (DTCWT) [47] along with its hybrid versions [44, 48, 49]. FAWT improves the traditional WT which offers a more flexible framework for analyzing nonstationary signals and provides, better TF localization [42]. TQWT adopts tunable quality factors to control adjustable bandwidths based on the oscillatory behavior of the signal [50]. This provides desirable decomposition performance and makes it suitable for analyzing nonstationary signals with varying frequency content. EWT decomposes signals into narrow sub-bands through empirically designed signal adaptive wavelet filter banks which provide more tailored analysis according to the signal's inherent characteristics [44, 48]. Similarly, the WPT offers improved signal decomposition with accurate extraction of sub-bands by enabling flexible frequency band selection through customizable wavelet filter banks [45]. Synchrosqueezed WT (SWT) [49] provides improved TFR by combining WT with reallocation methods in the TF plane. These methods found better suitability to analyse linear amplitude modulated (AM) and frequency modulated (FM) signals. However, these WT-based techniques have certain limitations such as analyzed signal length, predefined filter banks, optimal window length, predefined set of basis functions, and non-adaptive nature. Additionally, close frequency signal components are difficult to differentiate due to the introduction of cross-terms [51].

Recent advancements in multi-scale adaptive decomposition, such as variational mode decomposition (VMD) [52], Fourier-Bessel series expansion (FBSE) [53] and its variants [36, 54, 55, 56], ensemble EMD [57], compact EMD [58], eigenvalue decomposition (EVD) and its variants [51, 59, 60], adaptive Fourier decomposition [61], and singular spectrum analysis (SSA) [62], have already proven their potential to analyze multi-component nonstationary biomedical signals. These methods provide a useful investigation to analyze the features, particularly the weak spectral characteristics in nonlinear and nonstationary signals. However, the majority of these methods have been demonstrated to be superior in

addressing both mode mixing elimination and adaptive decomposition. Yet, the challenge of tuning multiple preset parameters across a wide range severely restricts their potential uses.

Recently, multivariate signal decomposition techniques have gained attention [63, 64, 65] for multivariate analysis of multichannel data. The developed multivariate analysis methods consist of multivariate VMD [63], multivariate EMD [64], multivariate iterative filtering [66], and multivariate EWT [65], which provide accurate estimations of the mutual components shared among the channels.

Among them, the recently proposed technique, i.e., SWD [67], has received a lot of attention for its efficiency in processing stationary and nonstationary signals. SWD is an iterative and data-driven decomposition method inspired by swarm intelligence. It is specially designed by the extended concept of swarm filtering. Numerous experiments have revealed its supremacy in resolving the problem of mode mixing [68, 69] over other state-of-the-art methods [41, 70]. It has been proven to be the best decomposition method for synthetic and natural signals [2, 67]. However, SWD method in the literature are limited by their performance in terms of the trade-off between optimization of hyper-tuning and decomposition performance [68, 71]. To ensure optimal SWD decomposition, it is imperative to meticulously choose the accurate threshold settings for each signal prior to decomposition. A few significant attempts in selecting SWD thresholds using metaheuristic optimization algorithms have been explored for TF analysis [68, 71]. Additionally, the univariate SWD decomposition method shows the limitation when dealing with the simultaneous processing of multivariate signals from multiple channels [67]. Therefore, there is a need for improvement related to individual threshold optimization in SWD to improve mode aliasing issues and multi-channel processing capability.



## 1.4 Motivation

In recent decades, nonstationary multicomponent signal analysis has played a significant role in brain signal processing, particularly with EEG and MEG signals in neurocognitive science [35, 72]. These signals are characterized by their complex and time-varying nature. It requires advanced signal analysis techniques for effective denoising and component extraction related to meaningful brain information. Although many nonstationary analysis methods have been proposed, they introduced significant limitations such as predefined filter banks, optimal window length, and non-adaptive processing [44, 48, 68, 72]. Among them, the recently introduced technique, SWD found better suitability to analyse nonstationary brain signals, however, it suffers from challenges related to pre-defined threshold parameter selection, making it difficult to achieve satisfactory decomposition results with the recommended settings [67]. Moreover, its limitations in univariate processing can result in mutual information loss in multichannel data [65].

Therefore, there is a need to improve SWD-based nonstationary signal processing techniques to decompose multichannel signals for improved brain signal analysis, and to further apply these techniques in the development of automated neurological disease detection and BCI frameworks. The proposed univariate SSDM (USSDM) and its novel extensions, multivariate SSDM (MSSDM), enhanced SSDM (ESSDM), and clustering SSDM (CSSDM), are effective nonstationary techniques for robust analysis of nonstationary brain signals. It extracts meaningful oscillatory modes from multichannel EEG and MEG signals and computes significant features adaptively for improved classification of brain neurological diseases and BCIs. The main motivations for the development of these signal processing techniques and related automated classification frameworks are as follows:

Sleep apnea is a chronic sleep disorder that can severely impact the physical, mental, and social well-being of patients, and in severe cases, it can lead to serious cardiovascular

and psychological complications [10]. While sleep apnea is not typically curable, timely treatment can help reduce its symptoms [8, 13]. Reliable automatic frameworks can assist in the early diagnosis and treatment of sleep apnea, helping to prevent further complications.

BCI technologies provide an effective solution to patients with neuromuscular diseases by providing an alternative communication path between the brain and external devices. MI movement and IMS detection are widely used in the development of BCI frameworks and demonstrate their effectiveness in various practical BCI applications [15, 17, 18, 19, 20]. MI has developed a new intuitive paradigm in the BCI and has become an active rehabilitation tool for the treatment of paralytic stroke patients [34, 73]. MI is a cognitive process in which the subject imagines the motory action rather than performing the actual action, e.g., right-hand or left-hand. In active rehabilitation training, MI-BCI constructs a direct information transfer pathway, which converts EEG brain activity signals into MI commands to control prosthesis instead of traditional brain and muscle control pathways [25, 26]. Also, MI-based EEG has multi-class scalability, thus showing the possibility of building an extensible BCI system.

Similarly, in the recent decade, IMS has developed advanced cognitive communication tools, serving as an intuitive paradigm within BCI technology [15, 17]. Against imagination of control action like in MI-based BCI, IMS is a new intuitive paradigm that conducts the internal pronunciation of words without any physical movement or audio output. IMS exhibits scalability across multiple classes, indicating the potential for constructing a scalable BCI system. This paradigm is especially well-suited for constructing communication systems because of its intuitive nature, making it particularly beneficial for individuals facing physical challenges like apraxia and dysarthria. In these MI studies [24, 25, 26, 28, 29, 30, 31, 32, 33] and IMS [15, 17, 18, 19, 20], EEG is widely utilized for its ability to be recorded in a non-invasive manner and offer improved fine-grained analysis

by identifying spectral, spatial, and temporal components underlying EEG signal. However, in an MI and IMS-based BCI system, accurately recognizing discriminative neural patterns from EEG signals and classifying them is challenging due to limited spatial resolution, intra-channel interference, and the nonstationary and nonlinear nature of the signals. Thus, it is challenging to develop a reliable EEG-based BCI recognition model.

Visual object decoding-based BCI has emerged as a transformative approach in cognitive neuroscience which enables intuitive interaction with technology and enhances rehabilitation for individuals with visual processing deficits [74]. In this context, neural signals represent how the human brain interprets natural visual stimuli [3]. Among various modalities, EEG and MEG serve as key tools for studying dynamic visual cognition due to their millisecond-level time resolution and non-invasive nature [3, 4, 5, 14]. However, high noise levels in EEG and MEG signal recordings and significant subject-specific variability during the visual cognitive decoding process make it difficult to decode visual activation patterns and consequently make it challenging to classify cognitive responses to different visual stimuli. Moreover, real-world visual recognition is hindered by issues such as channel selection, smoothing, data reduction, ambiguous scalp regions, and the nonstationary characteristics of EEG and MEG signals, which limit the extraction of mutually informative features and ultimately degrade classification performance [3, 14]. To overcome these challenges, robust multivariate visual analysis and advanced automated classification methods are essential for accurately decoding visual neural patterns from EEG and MEG signals.

## **1.5 Objectives**

In this thesis, we aim to achieve the following objectives:

**Objective 1:** To propose the SSDM for robust nonstationary signal analysis and develop a

sleep apnea detection framework from EEG signals

**Objective 2:** To propose MSSDM for multivariate signal analysis and develop a classification framework for cognitive visual object analysis from multichannel EEG-MEG signals

**Objective 3:** To propose an MSSDM-based joint TF analysis method and develop a framework for IMS enhancement for the intuitive BCI application

**Objective 4:** To propose an ESSDM for optimal oscillatory mode extraction and to classify multi-class MI movement for EEG-BCI framework

**Objective 5:** To propose CSSDM for automated recognition of upper limb movements from non-homogeneous EEG signals for BCI application

## 1.6 Thesis contributions

The works present in this thesis focus on developing a novel extension of the univariate SWD-based nonstationary signal-processing technique to decompose multichannel signals for brain signal analysis, and to further apply these techniques in the development of neurological disease detection and BCI frameworks. In this thesis, we will discuss four different SWD-based decomposition methods along with five different frameworks that have been proposed for the diagnosis of sleep apnea disease and BCI frameworks such as cognitive visual object recognition, MI recognition, IMS task detection, and EEG-based upper limb movements detection. The contributions of the thesis are as follows:

**Contribution I:**

A novel USSDM has been proposed based on the improved sparse Fourier transform (SFT) spectrum estimation model and designed new swarm filter banks for efficient mode extraction and TF improvement. The proposed SSDM method delivers optimal estimation of boundary frequencies in the SFT spectrum, resulting in improved filter banks. The performance of the proposed method has also been tested for the nonstationary AM and FM signals with additive white Gaussian noise (AWGN) at different signal-to-noise ratio (SNR) levels and compared with existing SWD and Hilbert-Huang transform (HHT) methods. Further, an automated system of sleep apnea detection has been developed using the proposed SSDM-based formulated new fusion feature and TFR image feature and tested on convolutional neural network (CNN) and bidirectional long short-term memory (BiLSTM) on the publicly available EEG database. The proposed SSDM method delivers substantial improvement in TF analysis. Our developed sleep apnea detection model could be a vital aid in clinical solutions.

**Contribution II:**

A new adaptive multivariate extension of SSDM i.e. MSSDM has been proposed to explore the multivariate analysis of nonstationary EEG-MEG signals and further used to improve the decoding accuracy of VSA pattern classification. This method decomposes multichannel EEG and MEG data into multivariate oscillatory modes. These modes are then grouped based on their mean frequency to obtain rhythms. Then new Riemann's correlation-assisted fusion feature (RCFF) has been computed from the features, namely, Renyi's entropy (RE), sparse entropy (SE), common spatial filter (CSF) on Teager energy (CSPTE), and correntropy spectral density (CSD) from computed rhythms of EEG, MEG, and combined EEG-MEG multichannel data. Five different machine learning are then developed to classify cognitive visual objects using these novel MSSDM-RCFF features

and improve the VSA pattern classification.

**Contribution III:**

A novel adaptive MSSDM has been proposed to reveal the multivariate joint oscillatory structure of multichannel EEG signals and delivered channel-aligned oscillatory components (CAOCs). Joint TF multivariate features are generated based on joint IF (JIF) and joint IA (JIA) functions computed from the extracted CAOCs. The new low-dimensional joint TF deep feature has been computed using four different pre-trained neural networks (ELM, AlexNet, SqueezeNet, and ResNet), and further two feature mapping strategies have been employed to generate the most discriminant features. This work also developed a multichannel IMS task detection framework based on the MSSDM, joint TF deep features, and a support vector machine (SVM) from cross-channel EEG signals for BCI application. The proposed framework has been evaluated using an IMS-EEG dataset.

**Contribution IV:**

This work proposed a novel decomposition technique, namely, the ESSDM based on selfish herd optimization (SHO), modified swarm filter bank, and SFT spectrum to solve the issue of choice of uniform decomposition and threshold parameters in existing SWD. In this approach, new fitness function criteria has been designed based on the Kullback–Leibler divergence (KLD) distance from the spectral kurtosis of all modes to select threshold parameters that optimize decomposition effect, avoid excessive iterations, and provide fast convergence with optimal modes. It also introduced a new subspace clustered-based fused TF graph (FTFG) feature from TF information, which is computed from extracted optimal modes using ESSDM. Furthermore, the performance of ESSDM has been tested by developing a multi-class MI electroencephalogram (MI-EEG) recognition model using obtained

FTFG features and three computationally efficient classifiers and demonstrated an efficient MI-EEG recognition model for feasible BCI application.

**Contribution V:**

In this work, the CSSDM is proposed to extract homogeneous spectral characteristics across nonhomogeneous multichannel EEG data with significant channel selection for efficient decomposition. CSSDM, a novel approach proposed to address the limitation of processing nonhomogeneous signals, such as EEG, extends the capabilities of existing SWD. CSSDM addresses issues in processing nonhomogeneous EEG signals in existing SWD by clustering them into homogeneous sets with density-based spatial clustering (DBSC) and CCA-assisted mutual information (CMI) criteria. This work also developed an automated upper limb movement recognition framework based on the CSSDM method using nonhomogeneous cross-channel EEG signals. Further, the TF graph spectral (TFGS) features are formulated from extracted modes using CSSDM. The developed framework has been validated and outperformed against baseline models. It demonstrated an efficient upper limb BCI tool for patients with neuromuscular diseases.

## 1.7 Organization of the thesis

This thesis is organized into seven chapters. The title and short summary of the chapters are structured as follows:

**Chapter 1 (Introduction)**

This chapter describes brain signal processing, applicability of recorded EEG and MEG signals, and nonstationary signal processing methods. In the later part, the objectives, motivations, contributions, and organization of the thesis are discussed.

**Chapter 2 (SSDM method for robust nonstationary signal analysis with EEG-based**

**sleep apnea detection)**

This chapter has proposed the SSDM for robust nonstationary signal analysis. The SSDM-based framework is developed for automatic sleep apnea detection using EEG signals.

**Chapter 3 (Multivariate SSDM for multivariate signal analysis with cognitive visual object detection from multichannel EEG-MEG signals)**

This chapter has proposed MSSDM for multivariate signal analysis and developed a classification framework for cognitive visual object analysis from multichannel EEG-MEG signals.

**Chapter 4 (Multivariate SSDM-based joint time-frequency analysis with improved imagined speech decoding for intuitive BCI)**

In this chapter, we have presented an MSSDM-based joint TF analysis method. The introduced method is used to construct a framework for IMS enhancement for the intuitive BCI application.

**Chapter 5 (Enhanced SSDM with multi-class motor imagery-based EEG-BCI system)**

This chapter has proposed an ESSDM for optimal oscillatory mode extraction from synthetic and real signals. The ESSDM-based classification framework has been developed to classify multi-class MI movements for the EEG-BCI framework.

**Chapter 6 (Clustering SSDM with MI-EEG based upper limb movement recognition for BCI application)**

In this chapter, the CSSDM has been proposed to extract homogeneous spectral characteristics across nonhomogeneous multichannel EEG data. This work also developed a CSSDM-based automated upper limb movement recognition framework using nonhomogeneous cross-channel EEG signals.

**Chapter 7 (Conclusions and future scope)**



This chapter provides a summary and findings of the thesis. This chapter also listed the information of the future work.

The structure of the thesis is illustrated as follows:

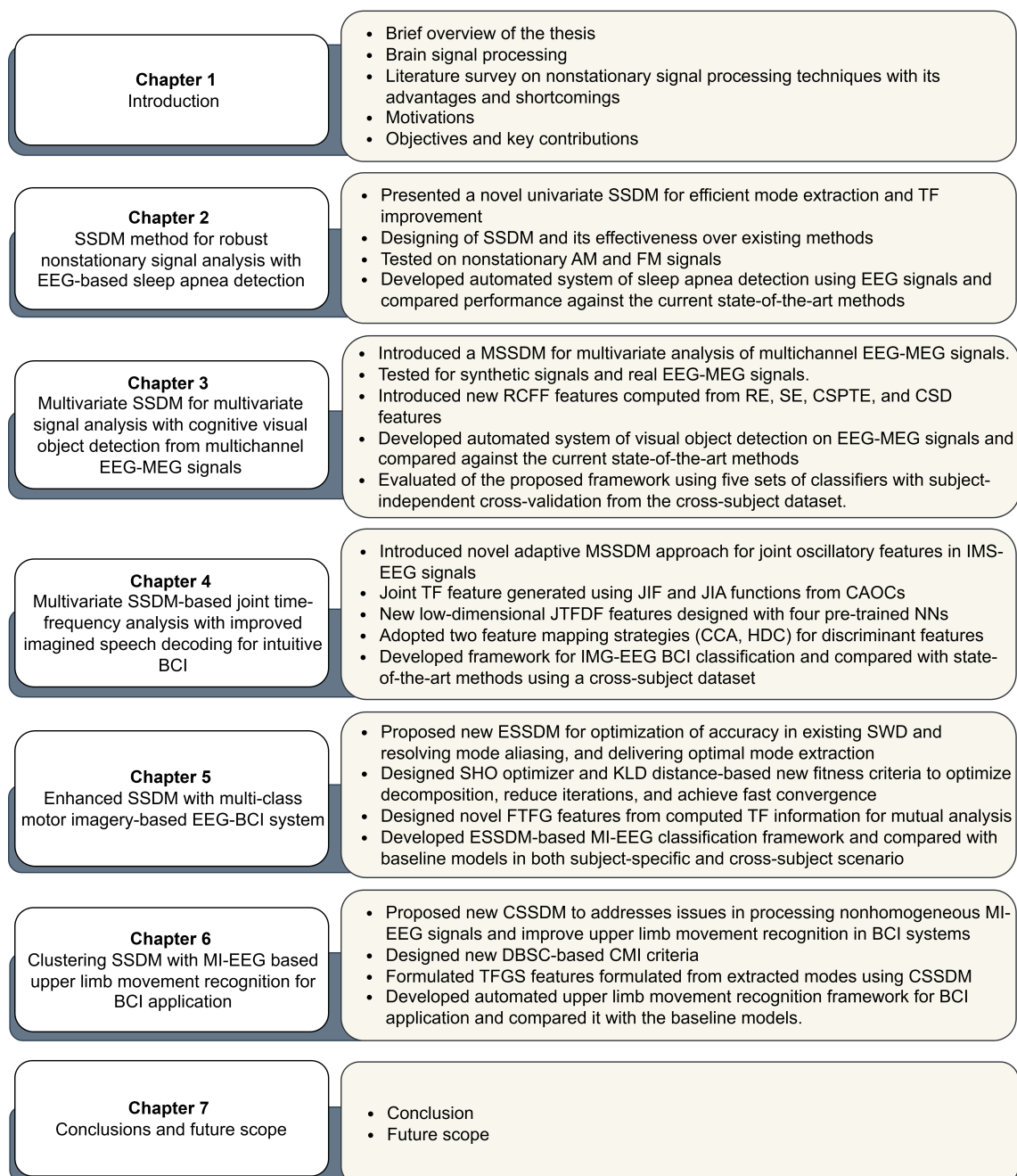


Figure 1.1: The structure of the thesis.

## **Chapter 2**

# **SSDM Method for Robust Nonstationary Signal Analysis with EEG-based Sleep Apnea Detection**

In this chapter, we present a novel USSDM for efficient mode extraction and TF improvement based on an improved SFT spectrum estimation model and specially designed swarm filter banks. The designing of SSDM and its effectiveness over existing methods are discussed in detail in this chapter. Furthermore, we have developed two frameworks to validate the effectiveness of our proposed SSDM approach in detecting sleep apnea disorder. The first framework focuses on enhancing sleep apnea detection using the newly formulated fusion feature which is computed from oscillatory components (OCs) using SSDM and employing a BiLSTM classifier for classification. In the second framework, we explore the efficacy of the TFR image feature which is computed from various rhythms extracted through SSDM-based OCs, and test them on multiple CNN models. The classification performance of the proposed frameworks is illustrated in detail in later sections.

## **2.1 Introduction**

Over the last few decades, nonstationary multicomponent signal analysis in the TF domain has played a significant role in several areas, namely speech synthesis, biomedicine, radar, telecommunication [35, 72]. The primary motivation for studying multicomponent signals is to enhance signal denoising and component extraction from biomedical signals such as ECG [1], EMG [6], EEG [36]. It is observed that the nonstationary signal that includes spectral content carries inherent time-varying characteristics of the analyzed signal, which introduces difficulty to analyse real complex signals. Therefore an effective decomposition method is required to identify and extract the real-time biological signals that are characterised by extreme nonstationary nature. In the literature, numerous nonstationary analysis methods have been proposed to adopt the efficient decomposition of nonstationary multicomponent signals. The majority of such methods comprise of DFT-based filter bank [37, 38], WT [40], FAWT [42], FBSE [53] and its variant [48, 55, 56], TQWT [43], EVD and its variants [51, 59], EWT and its hybrid variants [44, 48]. Amongst these methods, many are based on atomic decomposition and energy distribution on quadratic TFRs. These methods found better suitability to analyse linear AM and FM signals. However, they have certain limitations such as analyzed signal length, predefined filter banks, optimal window length, predefined set of functions, and non-adaptive nature. Additionally, close frequency signal components are difficult to differentiate due to the introduction of cross-terms [51]. Further, EMD [41] was introduced to analyze nonstationary signals. It decomposes a signal into its underlying OCs, known as IMFs. However, mode mixing remains an issue in the extracted IMFs. To overcome this issue, the successive variants were introduced as an ensemble EMD [57], multivariate EMD [64], and compact EMD [58], which have been disclosed by a great deal of literature [57, 58, 64, 70]. Recently, the SWD technique [67] has been introduced for processing stationary and nonstationary signals. It has proven to

be highly effective in overcoming mode mixing issues and for processing real-time signals [2, 68]. However, it suffers from pre-defined threshold selection, and it is difficult to obtain satisfactory decomposition results with the recommended choice.

To resolve the aforementioned limitations, we hereby propose a new technique, SSDM for signal decomposition and TF analysis. Benefiting from the merits of the SFT spectrum and SWD, the proposed SSDM method adopts an effective scale-space-based approach for boundary estimation and segmentation on the SFT spectrum to analyze nonstationary multicomponent signals into the OCs. It is inspired by recent applications in sparsity-based biomedical signal processing problems [75, 76, 77, 78, 79, 80]. To promote the sparsity more strongly, a new penalty function, i.e., fused least absolute shrinkage and selection operation (fused-LASSO) based signal approximator function, was used to design a non-convex SFT spectrum estimation [81, 82]. It leads to the new formulation of optimization solutions to the SFT spectrum estimation and converges the SFT spectrum with compact and real coefficients. The segmented SFT spectrum becomes spectral leakage-free and provides better localization in the frequency domain with more sharp energy peaks.

To evaluate effectiveness, the proposed SSDM method has been tested on the nonstationary AM signals, nonstationary FM signals, and real EEG signals. Further, the proposed method employed to classify sleep apnea disease from EEG signals. In the sleep apnea classification work, we have used a publicly available database [85, 86]. The detection of sleep apnea-hypopnea events is a complicated, time-consuming, and laborious process. In the literature, several approaches have been reported for the diagnosis of sleep apnea events based on different biomedical signals, including ECG, oxygen saturation (SpO<sub>2</sub>), photoplethysmogram (PPG), airflow signal, and nasal pressure signal, and achieved satisfactory results [7, 87, 88]. However, despite advantages, it carries certain predominant limitations such as lagging in multi-sleep states and stages, non-uniform deviation in the presence of

Table 2.1: Summary of related work on sleep apnea detection techniques.

Authors, year, and reference	Methodology	EEG database	Feature	Participants	Validation techniques	Performance (%)
Ahmed et al. (2016) [83]	KNN	MIT-BIH database	Spectral energy in beta band	05	Subject dependent, leave-one out	Acc = 82.28% SEN = 90.58% SPE = 77.72%
Soha et al. (2016) [84]	KNN	St. Vincent's University Hospital	Entropy	05	Subject dependent	Acc = 87.64% SEN = 89.02 % SPE = 86.27%
Shahnaz et al. (2016) [11]	KNN, SVM	MIT-BIH database	Delta band power ratio	14	Subject dependent, 10-folds	Acc = 87.03% SEN = 89.62 % SPE = 84.43%
Bhattacharjee et al. (2018) [12]	VMD-KNN	MIT-BIH database	Rician modeling	15	Subject independent, leave-one out	Acc = 88.49% SEN = 95.43% SPE = 81.56%
Almuhammadi et al. (2015) [9]	SVM, ANN, LDA and Naive Bayes	MIT-BIH database	Energy, variance	16	Subject dependent	Acc = 97.14% SEN = 97.01% SPE = 97.26%
Tanvir et al. (2020) [13]	Full connected CNN	St. Vincent's University Hospital	Temporal variation	25	Subject dependent	Acc = 88.22% SEN = 85.52% SPE = 88.97%
Taran et al. (2017) [8]	PSO-LSSVM	MIT-BIH database	Hermite coefficients	16	Subject dependent	Acc = 98.82% SEN = 98.66% SPE = 99.03%
Mahmud et al. (2021) [10]	VMD, full connected CNN, BiLSTM	MIT-BIH database	Multi-model temporal	16	Subject-independent cross-validation, leave-one-out	Acc = 95.54% SEN = 95.35% SPE = 93.00%

Acc: Accuracy, SEN: Sensitivity, SPE: Specificity, KNN: K-nearest neighbors, SVM: Support vector machine, PSO: Particle swarm optimization, LDA: Linear discriminant analysis, VMD: Variational mode decomposition, LSSVM: Least squares support vector machine, CNN: Convolutional neural network, and BiLSTM: Bidirectional long short-term memory.

other abnormalities conditions like irregular breathing and heart variability [89]. Whereas EEG analysis presented in the studies [8, 10, 11, 12, 83, 89] shows more distinguish sleep apnea-hypopnea index with the severity in a patient coherence over the other signals. A comprehensive review of sleep apnea detection and classification techniques based on EEG signals is presented in Table 2.1. Comparing the other research work, it is notable that the researchers have explored work based on the subject-dependent cross-validation approach and feature complexity for sleep apnea detection. Even though these proposed methods were delivered with improved classification accuracy, but method's performance suffers as the subject-dependent cross-validation approach is accommodated [8, 9]. In these approaches, the performance of methods was tested from the same subject in the training and testing sets. Therefore, the objective of the proposed work is to overcome these issues and improve performance related to an automated sleep apnea detection system on the EEG data. The presented work highlights the main contributions in the following aspects:

1. SSDM: A new decomposition method has been proposed based on the improved SFT

spectrum to modify SWD-related modes.

2. The proposed method has also been tested for the nonstationary AM and nonstationary FM signals with AWGN at different signal-to-noise ratio (SNR) levels.
3. The new feature fusion (time, frequency, and TF domains) and TFR image features have been explored, which give a multi-scale analysis of the EEG signals and improved the accuracy of sleep apnea classification.
4. It compared a developed automated system of sleep apnea detection against the current state-of-the-art methods with subject-independent cross-validation from the cross-subject dataset.

## **2.2 Proposed methodology**

The proposed SSDM method is an adaptive and highly efficient TF analysis method for decomposing nonstationary multicomponent signals into a finite number of dominant OCs. The block diagram of the proposed methodology of SSDM is shown in Fig. 2.1. It comprises all important steps performed to implement the proposed work, and their description is given in the following sub-sections.

### **2.2.1 SFT spectrum coefficients generation and computation**

First, the SFT method is applied to nonstationary multicomponent signals and various spectrum coefficients are generated. The forward SFT and its adjoint operator can be implemented in the frequency domain, and the SFT constraint is performed on the time domain model. Then, the iteratively re-weighted least-square basis estimation is used to estimate the SFT model.

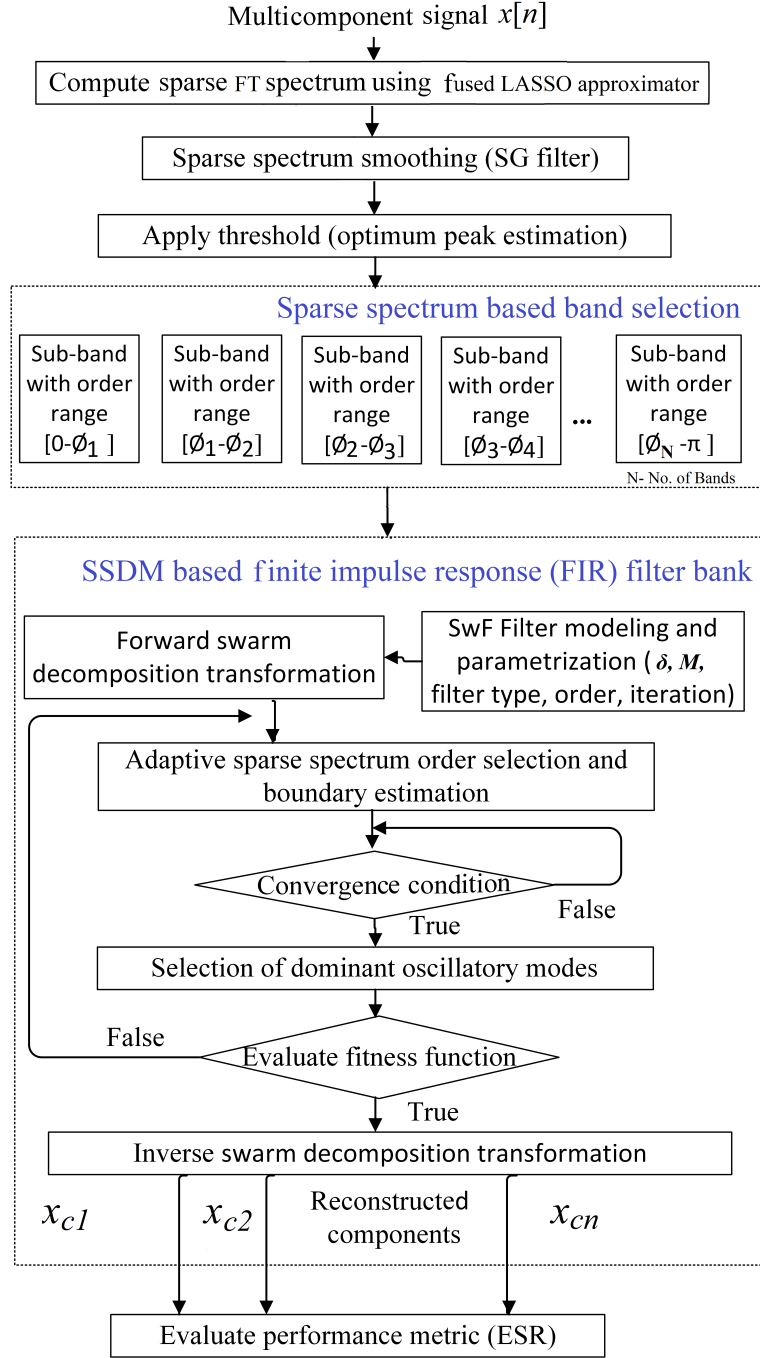


Figure 2.1: The systematic block diagram of the proposed SSDM method to obtain ESR.

Let the forward FT and its adjoint of signal  $x(n)$  can be expressed by,

$$x(n) = \sum_{m=0}^{N-1} c(m) e^{\frac{j2\pi mn}{N}} \quad (2.1a)$$

$$c(m) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-\frac{j2\pi mn}{N}}, \quad (2.1b)$$

where,  $x$  is the time-space domain data,  $c$  denotes the corresponding FT coefficients, and  $\hat{x}$  represents the recovered data using the adjoint FT. In this SFT model, the extracted component is independent at each frequency; thus, it can be estimated in parallel independently. This work mainly concentrates on the reformulation of SFT in the TF domain. The SFT coefficients  $c_m$  of  $m$ -order are computed via the following Eqs. (2.2) and (2.3), respectively. The SFT representation can be rewritten as [82],

$$x = L^H c \quad \Leftrightarrow \quad c = Lx, \quad (2.2)$$

where  $L$  and  $L^H$  are the  $L_1$ -norm of vector for FT and inverse FT (IFT) kernels.

To calculate  $L$  and  $L^H$ , the iterative shrinkage algorithm [80][81] is used in Eq. (2.3). It is used to get an optimized frequency-dependent SFT coefficients.

$$\psi(x) = \|c - F^{-1} [LF[x]]\|_1 + \mu \|x\|_1 \quad (2.3)$$

Where  $F$  and  $F^{-1}$  represent the forward and IFT along the temporal axis. This model uses a matrix inversion to update the model basis weights in each iteration and converge with an optimized SFT spectrum basis. Here, it uses an iterative least square algorithm, and  $\mu$  is the penalty factor, which regulates the SFT model's sparseness and the data misfit.

Even though the SFT can provide a high-resolution FT model, it is sensitive to spiky and high-amplitude noise. In Eq. (2.3), the inappropriate penalty function  $\mu$ ,  $L_1$ -norm of the data misfit indicates the distribution of noise in the TF domain. However, the sparse constraint is performed only along the slope or curvature direction, which cannot guarantee the sparseness along the temporal direction. In addition, the computational burden increases linearly with the increasing basis orders and the number of coefficients. In order to improve the performance of SFT model estimation in Eq. (2.3) and hold its original high efficiency, Tibshirani et al. [90] proposed an iterative fused-LASSO signal approximator based penalty function in Eq. (2.4). Instead of the  $L_1$  term, Eq. (2.3) can be extended by adding an extra



penalty term. This penalty function implemented in the time domain can guarantee the robustness and sparseness of the proposed method. Therefore, the SFT spectrum is spectral leakage-free and more localized in the frequency domain. The SFT coefficients  $\tilde{c}_m$  of time domain data  $x$  are computed as follows:

$$\psi(\tilde{c}_m) = \|c_m - F^{-1}[LF[x]]\|_1 + \lambda_1 \|F[x]\|_1 + \lambda_0 \|LF^{-1}[c_m]\|_1, \quad (2.4)$$

where  $\lambda_0$  and  $\lambda_1$  denote the penalty factors which regulate model's sparsity and it is set as 0.01.

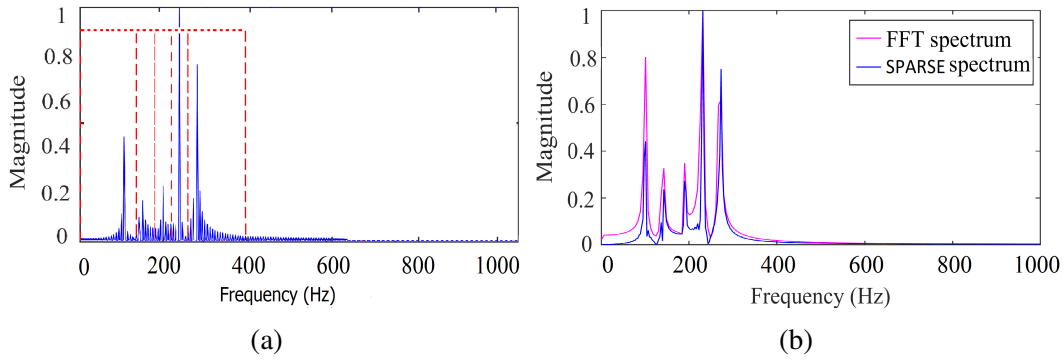


Figure 2.2: Plots of (a) coefficients magnitude ( $|c_m|$ ) versus frequency (Hz) with sparse spectrum estimation and (b) spectral analysis of FFT versus SFT coefficients after SGF (averaging filter) in case multicomponent AM signal analysis.

Fig. 2.2 shows the obtained plot of magnitude of the SFT coefficient ( $|c_m|$ ) versus frequency (Hz) using SFT spectrum technique corresponding to the multicomponent AM signal. The red dotted lines denote boundaries detected (BD) via swarm filter bank (FB1-FB5), and the thick blue line denotes a SFT spectrum of the signal. It is observed that each basis function is represented by a band of frequencies in Fig. 2.2 (a). It is noted that, as compared to the Fourier spectrum, the SFT spectrum provides a unique and compact spectral representation with improved spectral resolution. Due to sparsity, more energy is concentrated at the centre of SFT coefficients. Fig. 2.2 (a) shows the SFT coefficients

against magnitude. Here are five SFT spectrum peaks in Fig. 2.2 (b), representing five different monocomponents of AM nonstationary signals with 1000 samples. A signal can be reconstructed by optimum peak estimation in the SFT spectrum and provide an accurate grouping of the similar spectrum based on the range of SFT coefficients corresponding to the monocomponent signal.

### 2.2.2 Spectrum smoothing

To remove the unwanted peaks due to noise and estimate the highest peak, the energy spectrum must be filtered with a smoothing filter. The Savitzky-Golay filter (SGF) is carefully chosen and acts as a low-pass filter (LPF) to optimize decomposition performance while selecting the height of peaks [91]. In the SGF, the threshold in peak selection ( $P_{th}$ ) parameter controls the SWD performance and decides the number of component extractions. To perform fine decomposition level,  $P_{th}$  is selected around 0.1. The other two parameters, i.e., SGF length ( $SGF_{ln}$ ) and the SGF degree ( $SGF_d$ ) are selected to achieve a specific smoothing level of the SGF. It is experimentally chosen with values 15 and 2, respectively, to improve the SWD efficiency and performance. The criteria for dominant frequency  $\omega_{dom}^q$  estimation in each iteration can be derived from Eqs. (2.5) and (2.6).

$$\omega_{dom}^q = \arg_{\omega} \max \left( S'_{x_k}(\omega) > P_{th} \right) \quad (2.5)$$

$$S'_{x_k}(\omega) = \text{SGF} (S_{x_k}(\omega)) \quad (2.6)$$

Where  $S$  denotes the FT of the discrete data sequence  $x_k(n)$  at  $k$ th iteration, which will use the SGF.  $S'_{x_k}(\omega)$  is the energy spectral density (ESD) of  $x_k(n)$  after SGF.

### 2.2.3 SWD-based SFT spectrum boundary estimation and band selection

In the close-frequency multicomponent signals, spectral estimation is complex due to the overlapped nature of the SFT spectrum. The accurate number of multicomponent signals is not reconstructed correctly. Therefore, the SFT coefficient separability problem is required to be optimized with an appropriate selection of range of SFT coefficients, which can be accurately mapped to the monocomponent signal. However, there is no specific criteria to decide the best boundary condition for appropriate band selection with corresponding SFT coefficients. Therefore, in this step, the SWD method is used to design an approach based on adaptive sliding analysis window-based bandpass filtering for accurate boundary detection while optimizing the swarm model's parameters and responses.

In the swarm filtering process, the SFT spectrum is initially segmented into an approximated  $N$  number of fixed boundary segments. The border of each segment can be calculated by considering the middle of two progressive local maxima, which is expressed by Eq. (2.7).

$$W_i = \frac{Fc_i + Fc_{i+1}}{2}, \quad (2.7)$$

where  $Fc_i$  and  $Fc_{i+1}$  are two frequencies and the  $N$  number of boundaries set is denoted by  $i = 1, 2, \dots, N - 1$ . The algorithmic execution steps are described in detail in the following sub-sections.

#### 2.2.3.1 Signal description

The iterative swarm filter is initially fed by the discrete sequence of the multicomponent input signal,  $x_k(n) = s_{\text{multi}}(n)$  [67]. Where  $s_{\text{multi}}(n)$ , a nonstationary multicomponent signal which is combined with a set of monocomponent  $s_{\text{mono}, l}(n)$ , i.e.,

$$s_{\text{multi}}(n) = \sum_{l=1}^{|\Omega|} s_{\text{mono}, l}(n) = \sum_{l=1}^{|\Omega|} \omega_L(n - (l-1)L) \cos(\omega_l n), \forall l = 1, \dots, |\Omega|, \quad (2.8)$$

where  $\omega_L(n)$  is a periodic Hanning window with length  $L$  and  $|\Omega|$  is the cardinality of  $\Omega$ .  $|\Omega| = \{\omega_l = \frac{2\pi l}{16} \text{radian}; l = 1, 2, \dots, 7\}$  represents set of frequencies.

### 2.2.3.2 Initialization of swarm model parameters

Regarding the initial conditions of the swarm filter, the swarm hunting parameters  $\delta$  (virtual time interval) and  $M$  (swarm members) are preset to control swarm behavior. In each iteration of Swf filtering, the swarm filter parameters  $\delta$  and  $M$  are tuned to get an accurate reconstruction of the OCs. The relationships between these swarm filter parameters and each spectral content are estimated using a genetic algorithm, as given by the following equation.

$$M(\hat{\omega}) = \lfloor 33.46 \hat{\omega}^{-0.735} - 29.1 \rfloor \quad (2.9a)$$

$$\delta(\hat{\omega}) = -1.5\hat{\omega}^2 + 3.454 \hat{\omega} - 0.01, \quad (2.9b)$$

where  $\hat{\omega}$  is the normalized frequency.

### 2.2.3.3 SWD-based spectrum sensing and estimation

In this step, iterative execution of the swarm filter is carried out to find OCs from the residue of a multicomponent input signal. The OCs are extracted by selecting the frequency band which possess the highest amplitude peak through the ESD. With each iteration, the current extracted prominent OCs is subtracted from the rest of the input for further oscillatory modes processing. This process continues until the residual signal does not possess any spectral content of sufficient observable energy. The extraction of dominant OCs is op-

timized with the computation of the fitness function value for each swarm member through the following steps;

I. Initialize each swarm member's velocity vector  $V_i(n)$  and position vector  $P_i(n)$  with integer random values uniformly distributed. According to the swarm-prey hunting mechanism, set the initial value of the prey  $P_i(0)$  with a symmetrical hunting position pattern with a random integer value as follows:

$$p_i(0) = p_{\text{prey}}(0) + d_{\text{cr}} \cdot \left(i - 1 - \frac{M}{2}\right), \forall_i = 1, \dots, M \quad (2.10a)$$

$$v_i(0) = 0, \forall_i = 1, \dots, M, \quad (2.10b)$$

where  $M$  denotes the number of members.

II. Calculate the driving force  $F_{\text{Dr}, i}^n$  and cohesive force  $F_{\text{Coh}, i}^n$  as,

$$F_{\text{Dr}, i}^n = p_{\text{prey}}(n) - p_i(n-1) \quad (2.11)$$

$$F_{\text{Coh}, i}^n = \frac{1}{M-1} \sum_{j=1, j \neq i}^M f(p_i(n-1) - p_j(n-1)) \quad (2.12)$$

Where,  $F_{\text{Dr}, i}^n$  is the attractive force which is induced by prey for members.  $F_{\text{Coh}, i}^n$  is the attractive/repulsive force that is induced among swarm members.  $f(d)$  is cohesion force contribution of the  $j$ th member to the  $i$ th member at a distance  $d$  and it is set for function  $f(d)$  with  $d = (4, 4)$ .  $d_{\text{cr}}$  is the minimal critical distance among members without affecting individual performance by others and set  $d_{\text{cr}} = 0.5, 1$ , and  $2$ .

III. In this step,  $\delta$  and  $M$  parameters are updated as per Eqs. (2.9a) and (2.9b). In every iteration of swarm-prey hunting, each swarm member updates its current state velocity  $V_i(n)$  and position  $P_i(n)$  which are shown in the following equations.

$$v_i(n) = v_i(n-1) + \delta(F_{\text{Dr}, i}^n + F_{\text{Coh}, i}^n) \quad (2.13a)$$

$$p_i(n) = p_i(n-1) + \delta v_i(n), \quad (2.13b)$$

where  $\delta$  controls swarm hunting flexibility in the time interval  $[0, 1]$ .

IV. Calculate the best position of each member in the best local position vector and estimate the fitness function value accordingly with the best global position and iterate for further process.

V. The output  $y(n)$  of SWD is derived by summation of swarm member's positions as,

$$y(n) = \beta \sum_{i=1}^M p_i(n), \quad (2.14)$$

where  $\beta$  is a weighted factor.

Here, the SWD output is measured with swarm member's positions from the trajectory of the swarm; instead, the mean value of the member's positions is discarded [67]. In this adaptive filter, the dominant oscillatory mode locates the spectral information of the analyzed signal. The obtained sub-band signal possess centre frequencies.

VI. For each successive iteration, the following condition is used to select dominant OCs by checking the energy of residual signal, as given below:

$$X_{k+1}(n) = X_k(n) - X'_k(n), \quad \text{if } S'_{x_k}(\omega) < P_{\text{th}} \quad (2.15)$$

Where  $X_{k+1}(n)$  is the obtained residual signal, when the extracted filter signal  $X'_k(n)$  is subtracted from input signal  $X_k(n)$  at  $k$ th iteration. If the difference between consecutive iterations is less than the  $P_{\text{th}}$  value 0.1, the iteration is stopped, else it repeats the process.

VII. Check convergence condition

The iterative algorithm SWD is stopped if it satisfies the convergence criteria for the

residual signal which is indicated in Eq. (2.16).

$$\arg \min_{\delta, M} \sum_k \{|Y_{\delta, M}(k)| - |X(k)|\}^2, \quad (2.16)$$

wherein  $|Y_{\delta, M}(k)|$  and  $|X(k)|$  are the amplitude of FT of the output  $|Y_{\delta, M}(k)|$  and input signal  $x(n)$ .

The above steps are summarised in the pseudo-code of the SSDM method is illustrated in Algorithm 2.1.

---

**Algorithm 2.1** Pseudo code of SSDM method

---

Input: Nonstationary multicomponent signal  $x(n)$

Output: Sub-bands decomposition after SSDM method

Initialization:  $P_{th} = 0.1$ ;  $k = 0$ ;  $M = 20$ ;  $\delta = 0.01$ ;  $c = 0$ ;  $j = 0$

1. Apply SFT method to given input  $x(n)$  to compute coefficients.
  2. To estimate the highest peak in the SFT spectrum, smooth the energy spectrum using the SGF as per Eq. (2.6).
  3. Calculate local maxima ( $L$ ) from the obtained SFT spectrum coefficients ( $|c_m|$ ) using fixed boundary estimation and segmentation of the signal  $x(n)$ .  
 $L = \max(c_m)$ , as  $m=1, 2, 3, \dots, N$
  4. Once maxima are obtained, segmentation of SFT spectrum using adaptive boundaries estimation is performed in Eq. (2.7).
  5. After the segmentation of the SFT spectrum, all bands are passed through an iterative swarm filter bank.  
 $y_i(n) = \text{SwF}(x_i(n), M, \delta)$ ,  $i=1, 2, 3, \dots, N$ ; SwF: Swarm filter
  6. Iterative execution of swarm filter is carried out to find OCs.
  7. Evaluate fitness function using Eq. (2.17).
  8. For the reconstruction of the signal, inverse SFT is applied to each sub-band of the decomposed signal.
  9. Evaluate performance metric, error-to-signal ratio (ESR).
- 

## 2.2.4 Performance measure

To validate the proposed SSDM method, performance metric-ESR is used to measure the reconstruction efficiency of the algorithm, which is calculated in Eq. (2.17). The criteria are to be chosen to select a small value that decides the accurate reconstruction.

$$\text{ESR} = \frac{\sum_n |S_{\text{mono}, l}(n) - \tilde{s}_{\text{mono}, l}(n)|^2}{\sum_n |s_{\text{mono}, l}(n)|^2}, \quad (2.17)$$

where,  $s_{\text{mono}, l}(n)$  and  $\tilde{s}_{\text{mono}, l}(n)$  are the reference and extracted components, respectively.

Table 2.2: The mathematical expression of synthesized test signals.

Test signals	Model parameters
$s_I(n) = \sum_{l=1}^M s_{\text{multi}, l}^{\text{AM}}(n) = \sum_{l=1}^M A_l(1+a_l n) \cos\left(\frac{2\pi f_l}{f_s} n + \phi_l\right) \quad (2.18)$	$M = 5, f_s = 20 \text{ KHz}, A_1 = 2, A_2 = 1, A_3 = 0.9, A_4 = 3, A_5 = 2.5, f_1 = 100 \text{ Hz}, f_2 = 140 \text{ Hz}, f_3 = 190 \text{ Hz}, f_4 = 230 \text{ Hz}, f_5 = 270 \text{ Hz}, \phi_1 = 0, \phi_2 = \pi, \phi_3 = 0, \phi_4 = \pi/2, \phi_5 = \pi/3, a_1 = 0.0042, a_2 = 0.0037, a_3 = 0.0029, a_4 = 0.0037, a_5 = 0.0033, w(n)$ represents AWGN, and $n = 0, 1, 2, \dots, 399$ .
$s_{II}(n) = \sum_{l=1}^M s_{\text{multi}, l}^{\text{AM}}(n) + w(n) = \sum_{l=1}^M A_l(1+a_l n) \cos\left(\frac{2\pi f_l}{f_s} n + \phi_l\right) + w(n) \quad (2.19)$	
$s_{III}(n) = \sum_{l=1}^M s_{\text{multi}, l}^{\text{FM}}(n) + w(n) = \sum_{l=1}^M A_l \cos\left(\left(\frac{2\pi f_l}{f_s} \left(1 + \frac{\beta_l}{f_s} n\right)\right) n + \phi_l\right) \quad (2.20)$	$M = 3; f_s = 20 \text{ KHz}, f_1 = 500 \text{ Hz}, f_2 = 1200 \text{ Hz}, f_3 = 2000 \text{ Hz}, A_1 = 3, A_2 = 2, A_3 = 1, \phi_1 = 0, \phi_2 = \pi, \phi_3 = 0, \beta_1 = 4, \beta_2 = 3, \beta_3 = 3/2, w(n)$ represents AWGN, and $n = 0, 1, 2, \dots, 1999$ .
$s_{IV}(n) = \sum_{l=1}^M s_{\text{multi}, l}^{\text{FM}}(n) + w(n) = \sum_{l=1}^M A_l \cos\left(\left(\frac{2\pi f_l}{f_s} \left(1 + \frac{\beta_l}{f_s} n\right)\right) n + \phi_l\right) + w(n) \quad (2.21)$	

## 2.3 Results and Discussions

To prove the efficacy of the proposed method over the existing HHT and SWD, we have tested the synthetic nonstationary signals (AM and FM signals) and real EEG signals of sleep apnea disorder. The performance of the proposed method has been evaluated by measured values of ESR and TFR analysis. In this work, Hilbert spectral analysis (HSA) has been employed for obtaining TFR [41]. The performance analysis is detailed in the following sub-sections.

### 2.3.1 Example 1: AM nonstationary signal

In case-I, the synthetic AM signal  $s_I(n)$  uses five constant close frequencies with amplitude variation, represented in Eq. (2.18). In this, we have adopted frequency bands



[95-105 Hz], [135-145 Hz], [185-195 Hz], [225-235 Hz], and [265-275 Hz]. The detected boundaries in the SFT spectrum and SSDM-based filter bank for multicomponent AM signals have been shown in Fig. 2.3. The red dotted lines denote boundaries detected in the SFT spectrum of the signal and it shows that the boundaries are optimally detected in the SFT spectrum. Fig. 2.3 show the detected spectrum  $s_{\text{mono}, 1}^{\text{AM}}(n)$ ,  $s_{\text{mono}, 2}^{\text{AM}}(n)$ ,  $s_{\text{mono}, 3}^{\text{AM}}(n)$ ,  $s_{\text{mono}, 4}^{\text{AM}}(n)$ , and  $s_{\text{mono}, 5}^{\text{AM}}(n)$ , which are corresponding to the original AM monocomponents. The obtained TFRs of HHT, SWD and SSDM are shown in Fig. 2.4. In Fig. 2.4 (b), it

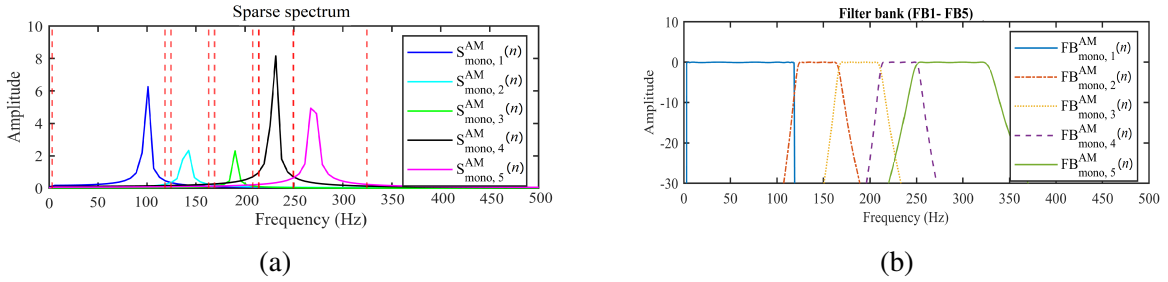


Figure 2.3: Visual description of (a) detected boundaries in the SFT spectrum and (b) filter bank (FB1-FB5) generated for  $s_1(n)$  multicomponent signal.

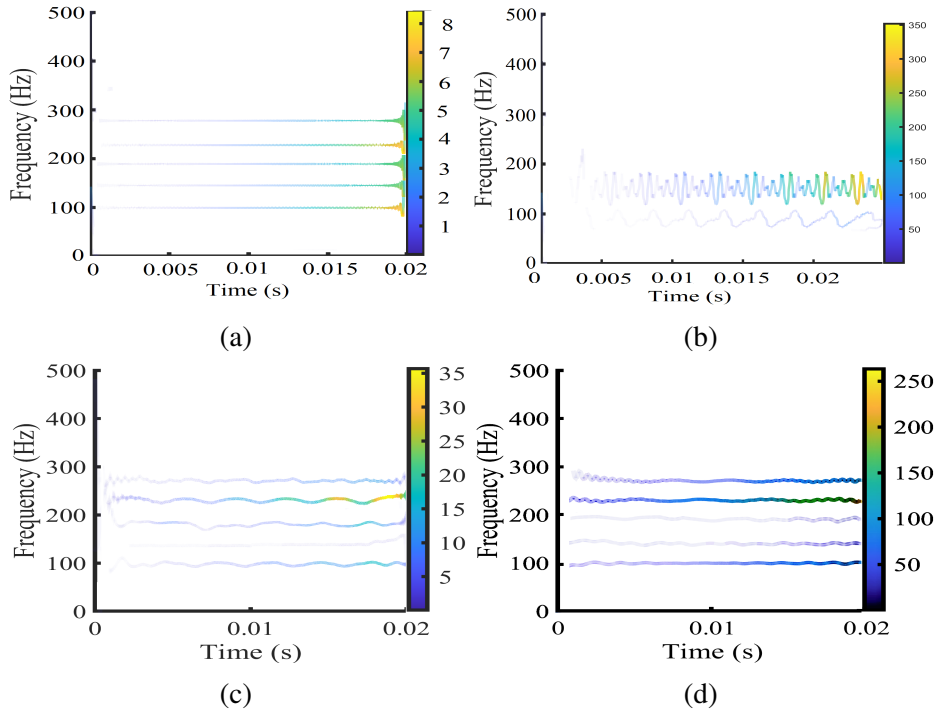


Figure 2.4: Plots of  $s_1(n)$ : (a) TFR of input AM signal, (b) HHT-based TFR, (c) SWD-based TFR, and (d) SSDM-based TFR.

Table 2.3: Comparison of the ESR values computed for reconstructed monocomponent signals of  $s_I(n)$  using HHT, SWD, and SSDM methods.

Methods	SSDM method (Order = 185 and $n = 399$ )	SWD method ( $n = 399$ )	HHT method ( $n = 399$ )
ESR <sub>1</sub> (1st)	0.0947	0.1201	0.6210
ESR <sub>2</sub> (2nd)	0.0830	0.1207	1.5402
ESR <sub>3</sub> (3rd)	0.0791	0.1482	-
ESR <sub>4</sub> (4th)	0.0296	0.0545	-
ESR <sub>5</sub> (5th)	0.0872	0.1273	-

Note: '-' indicates the blank ESR value due to non-constructed component.

shows that the TFR obtained by HHT is overlapped and non-distinguishable. The HHT fails to decompose the signal into the original monocomponents due to the mode mixing issue. We hardly find two monocomponents in the obtained TFR with overlapping nature. In Fig. 2.4 (c), it can be observed that SWD accurately replicates five monocomponents and shows better ability to the mode mixing against the cost of the accurate selection of thresholds  $P_{th} = 0.01$  and  $StD_{th} = 0.05$ . Whereas the proposed method demonstrates all five converged components with five modes. The SSDM method optimally detects boundary frequencies in the SFT domain. This is due to the fact that the SFT spectrum provides better frequency resolution over FT. Here, it achieved the computational efficiency by optimizing threshold value to  $P_{th} = 0.1$  and  $StD_{th} = 0.5$ . Table 2.3 indicates the performance of SSDM, SWD, and HHT decomposition methods. The values of ESR for the SSDM method are the lowest as compared to SWD and HHT methods. The value of ESR of every extracted component for the SWD method is slightly inferior, whereas ESR for the HHT method was the highest. The obtained ESR for the reconstructed monocomponent signals  $x_{c1}$ ,  $x_{c2}$ ,  $x_{c3}$ ,  $x_{c4}$  and  $x_{c5}$  using the SSDM is proved to be better than the SWD method. The values of ESR<sub>3</sub>, ESR<sub>4</sub>, and ESR<sub>5</sub> for HHT are kept blank because reconstructed components are not identifiable.

In case II, the synthetic AM signal  $s_{II}(n)$  is added with uniformly distributed AWGN of SNR (-10 dB to 30 dB), represented in Eq. (2.19). The intention behind this case is to validate the effectiveness of the proposed SSDM method in the presence of noise interference. Fig. 2.6 (b) shows that HHT-based TFR suffers from interference terms. Thus, no actual reconstruction of the components is obtained. Fig. 2.6 (c) shows the extracted components  $x_{c1}$ ,  $x_{c2}$ ,  $x_{c3}$ ,  $x_{c4}$ , and  $x_{c5}$  in case of SWD method. SWD demonstrates effective five modes.

Whereas the proposed SSDM method showed better decomposition with the reconstruction of all five components with better ESR values. Due to optimum boundary estimation in the SFT spectrum, the SSDM method can accurately recover all five sub-band components in the obtained TFR, which is shown in Fig. 2.6 (d). The detected boundaries in the SFT spectrum and filter bank using the proposed method are illustrated in Fig. 2.5. From Table 2.4, it is observed that SSDM performs accurate reconstruction of all components with the obtained  $ESR_1$ - $ESR_5$  against the SWD and HHT. The values of  $ESR_2$ ,  $ESR_3$ ,  $ESR_4$ , and  $ESR_5$  for HHT are kept blank because reconstructed components are not identifiable.

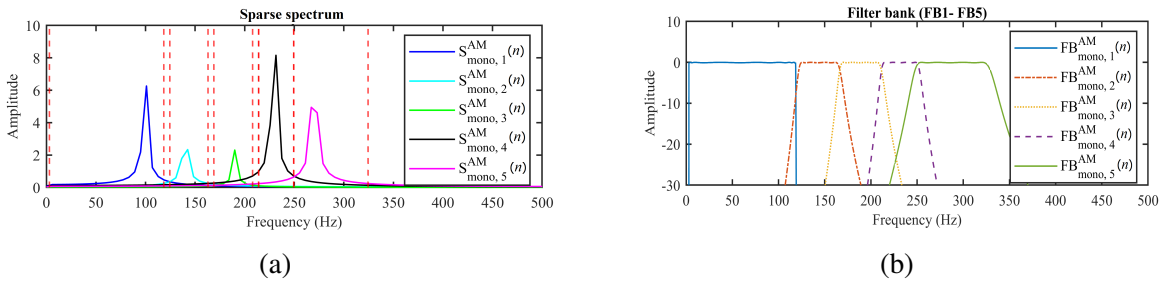


Figure 2.5: Visual description of (a) detected boundaries in the SFT spectrum and (b) filter bank (FB1-FB5) generated for  $s_{II}(n)$  multicomponent signal.

Table 2.4: Comparison of the ESR values computed for reconstructed monocomponent signals of  $s_{II}(n)$  using HHT, SWD, and SSDM methods.

Methods	SSDM method (Order = 185, $n = 339$ )	SWD method ( $n = 399$ )	HHT method ( $n = 399$ )
SNR = 30 dB			
$ESR_1$ (1st)	0.0822	0.2105	1.498
$ESR_2$ (2nd)	0.0837	0.3068	-
$ESR_3$ (3rd)	0.0709	0.2233	-
$ESR_4$ (4th)	0.0450	0.2700	-
$ESR_5$ (5th)	0.0901	0.3598	-
SNR = -10 dB			
$ESR_1$ (1st)	0.0904	0.2075	1.845
$ESR_2$ (2nd)	0.0953	0.1438	-
$ESR_3$ (3rd)	0.0815	0.1513	-
$ESR_4$ (4th)	0.0551	0.0588	-
$ESR_5$ (5th)	0.0962	0.2039	-

Note: '-' indicates the blank ESR value due to non-constructed component.

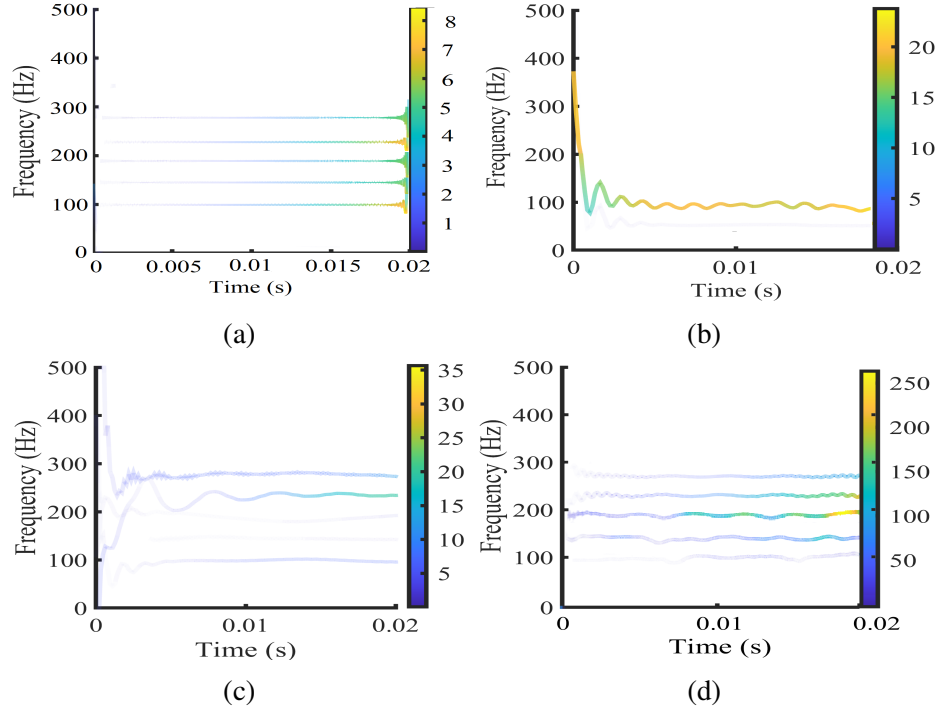


Figure 2.6: Plots of  $s_{II}(n)$ : (a) TFR of input AM signal with AWGN (-10dB), (b) HHT-based TFR, (c) SWD-based TFR, and (d) SSDM based TFR.

### 2.3.2 Example 2: FM nonstationary signal

In case III, the synthetic FM signal  $s_{III}(n)$  consists of three sinusoids with different amplitudes and three linear variable frequencies, represented in Eq. (2.20). In this case, the frequencies of the first, second, and third components of the signal  $s_{III}(n)$  vary from 500 Hz to 900 Hz, 1200 Hz to 1900 Hz, and 2000 Hz to 2600 Hz, respectively. The obtained TFRs of HHT, SWD, and SSDM are shown in Fig. 2.8. It showed that the proposed method reconstructs all sub-band components accurately. The obtained TFR of SSDM was found better than the SWD and HHT methods. In the HHT case, the IF components are clearly visible at low frequencies, but at high-frequency, IF components are hard to identify due to poor visibility. Whereas, from Fig. 2.8 (c), it is observed that the TF resolution of SWD is better than HHT but inferior to the SSDM method. In SWD, the two components  $x_{c1}$  and  $x_{c2}$  are adequately recovered from obtained four OCs, but fail to decompose the third component. Fig. 2.7 shows the construction of filter banks (FB1-FB3) and detected frequency boundaries using the SSDM method. It shows the detected spectrum  $s_{mono, 1}^{FM}(n)$ ,

$s_{\text{mono}, 2}^{\text{FM}}(n)$ , and  $s_{\text{mono}, 3}^{\text{FM}}(n)$ , which are corresponding to the original FM monocomponents. Table 2.5 indicates the performance of SSDM, SWD, and HHT decomposition methods. The values of ESR of the SWD method are slightly inferior in the case of 1st and 2nd components as compared to the SSDM method. Other side, the values of ESR of the HHT method were the highest and could decompose two components with poor ESR values. The first component could not be decomposed in HHT, so it kept a blank value. Whereas the proposed method outperformed the SWD and HHT and delivered the lowest ESR values.

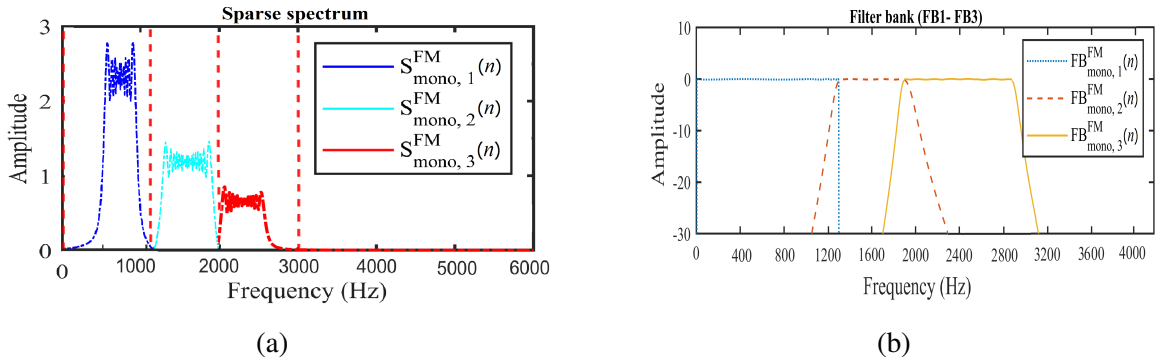


Figure 2.7: Visual description of (a) detected boundaries in the SFT spectrum and (b) filter bank (FB1-FB3) generated for  $s_{\text{III}}(n)$  multicomponent signal.

Table 2.5: Comparison of the ESR values computed for reconstructed monocomponent signals of  $s_{\text{III}}(n)$  using HHT, SWD, and SSDM methods.

Methods	SSDM method (Order = 185, $n = 1999$ )	SWD method ( $n = 1999$ )	HHT method ( $n = 1999$ )
ESR <sub>1</sub> (1st)	0.0932	0.1608	-
ESR <sub>2</sub> (2nd)	0.1148	0.2802	2.8620
ESR <sub>3</sub> (3rd)	0.1820	0.8620	6.3152

Note: '-' indicates the blank ESR value due to non-constructed component.

In case IV, the synthetic FM signal  $s_{\text{IV}}(n)$  is added with uniformly distributed AWGN of SNR (-10 dB to 30 dB), represented in Eq. (2.21). The obtained TFR and Table 2.6 indicate that the HHT could not reconstruct the component clearly in the presence of noise interference. In the SWD case, the obtained OCs modes are five and the reconstructed three components are visible in Fig. 2.10 (c), respectively. However, it is affected by noise interference and produces 2nd and 3rd components distorted, which is shown in Fig. 2.10 (d). In this case, the proposed SSDM method shows supremacy to decompose components accu-

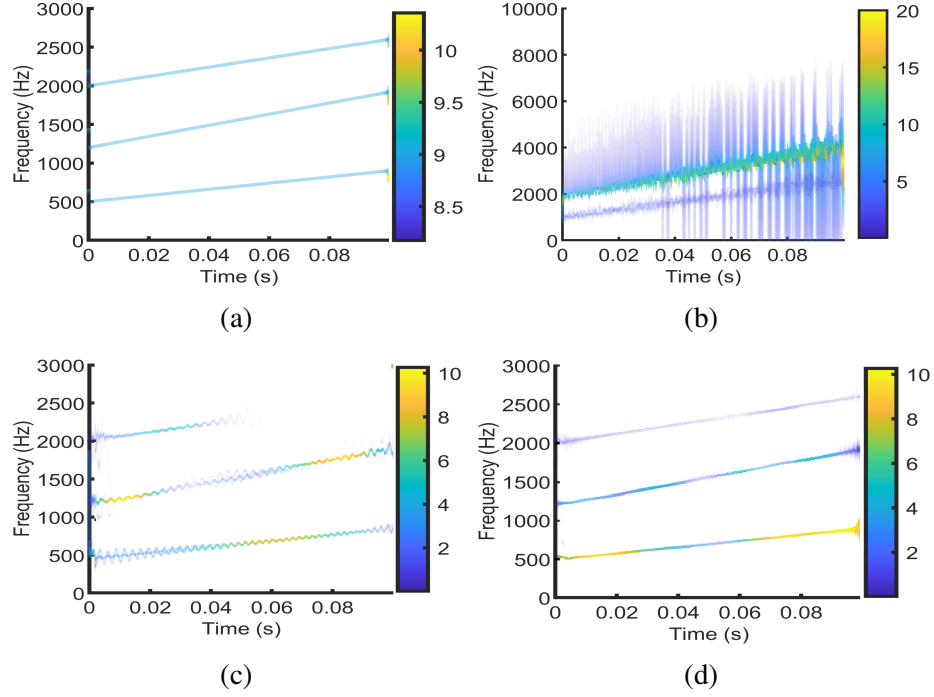


Figure 2.8: Plots of  $s_{III}(n)$ : (a) TFR of input FM signal, (b) HHT-based TFR, (c) SWD-based TFR, and (d) SSDM-based TFR.

rately over HHT and SWD methods. It reconstructed all three components with improved TFR, which is shown in Fig. 2.10 (d).

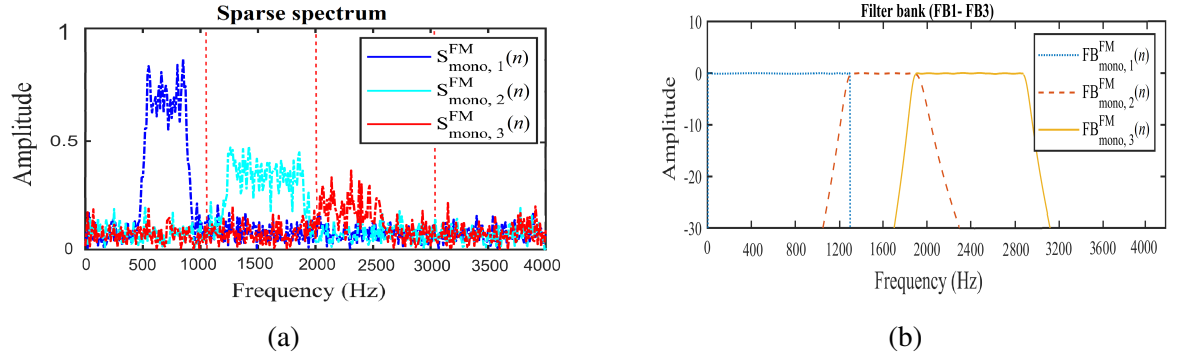


Figure 2.9: Visual description of (a) detected boundaries in the SFT spectrum and (b) filter bank (FB1-FB3) generated for  $s_{IV}(n)$  multicomponent signal.

### 2.3.3 Study on real EEG signals

The effectiveness of the proposed SSDM method has been tested for real EEG datasets corresponding to patients having sleep apnea disorder. Our proposed framework in-

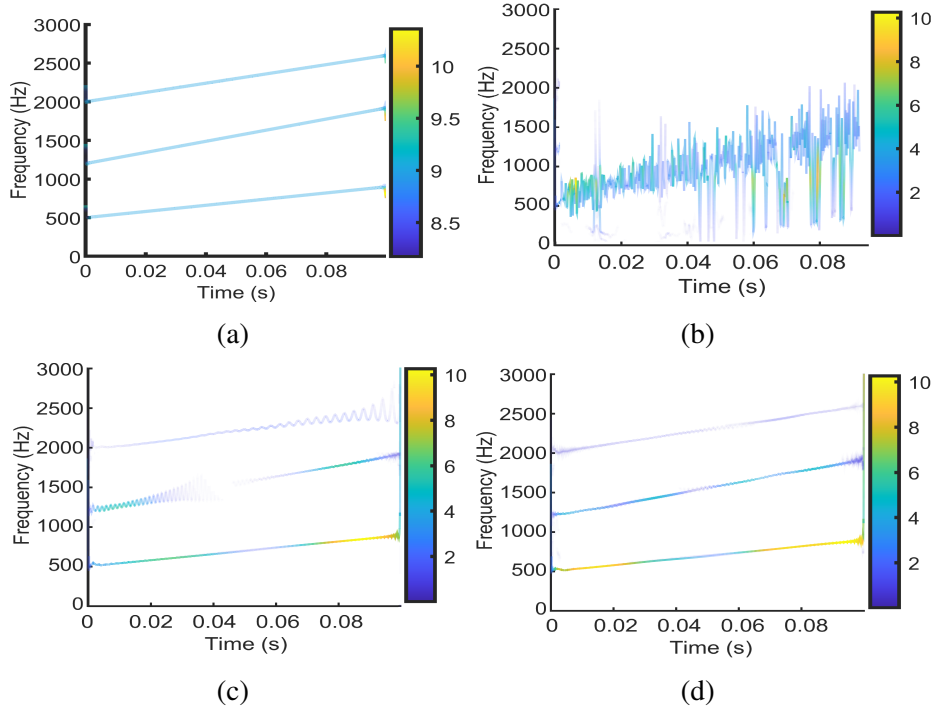


Figure 2.10: Plots of  $s_{IV}(n)$ : (a) TFR of input FM signal with AWGN (-10dB), (b) HHT-based TFR, (c) SWD-based TFR, and (d) SSDM based TFR.

Table 2.6: Comparison of the ESR values computed for reconstructed monocomponent signals of  $s_{IV}(n)$  using HHT, SWD, and SSDM methods.

Methods	SSDM method (Order=185, $n = 1999$ )	SWD method ( $n = 1999$ )	HHT method ( $n = 1999$ )
SNR = 30 dB			
ESR <sub>1</sub> (1st)	0.1105	0.2084	8.6405
ESR <sub>2</sub> (2nd)	0.1308	0.2960	-
ESR <sub>3</sub> (3rd)	0.2130	0.8902	-
SNR = -10 dB			
ESR <sub>1</sub> (1st)	0.1580	0.1506	-
ESR <sub>2</sub> (2nd)	0.2075	0.8345	-
ESR <sub>3</sub> (3rd)	0.2410	0.3108	-

Note: '-' indicates the blank ESR value due to non-constructed component.

cludes a machine learning approach for sleep apnea disorder classification using the MIT-BIH polysomnographic database [85, 86], which is given at <https://archive.physionet.org/physiobank/database/slpdb/>. It contains various physiological signals such as EEG, electrooculogram (EOG), and EMG signals of 16 male subjects (mean age 43) with a different apnea-hypopnea index (AHI). It has a sampling rate of 250

Hz, and 30 seconds frame duration is annotated by an expert. Here, we have used specifically all 14 patient's EEG signals of the 10-second epochs with no overlap for C3/O1, C4/A1, and O2/A1 cross channels. Table 2.7 shows the extracted number of apnea and non-apnea epochs with EEG signals corresponding to every subject. The apnea labels are considered based on all the annotated apnea-hypopnea events and were marked by experienced sleep experts with two broad categories: apnea and non-apnea events (normal breathing events). In our study, we have selected 4020 epochs of obstructive sleep apnea (OSA) events and 3978 epochs of normal breathing events from the overnight sleep data of 14 patients for 80 hours of recording.

EEG signals are always contaminated by various artifacts such as muscle, corneo-retinal-dipole, eyelid-related artifacts, and powerline interference. The frequencies with the largest variations between sleep apnea and non-apnea instances occur below 40 Hz. In the pre-processing, the EEG data is down-sampled to 128 Hz, and then the tenth-order Butterworth band-pass filter (BPF) is used with a pass-band frequency of 2 Hz-45 Hz to remove various artifacts. Also, a 60 Hz notch filter is employed to remove powerline interference. In our study, 10-second segments without artifacts are considered for feature extraction and analysis. The signals are segmented into non-overlapping segments of 10 seconds, and segmentation gave rise to a new sleep apnea database of 12060 apnea sub-frames and 11934 non-apnea sub-frames with a fixed length. Further, different SSDM-based rhythms are extracted from sub-frame EEG data than the whole duration frame. The rhythms are computed based on the mean frequency from the decomposed OCs [66]. Here, EEG epochs are chosen for efficient mode selection based on a high data fidelity constraint with delta ( $\delta$ : 0.1–4 Hz), theta ( $\theta$ : 4–8 Hz), and alpha ( $\alpha$ : 8–13 Hz) rhythms. These bands show a better correlation to sleep apnea discrimination. Figures 2.11 (a) and 2.11 (b) show the visual analysis of a 10-second epoch corresponding to two different classes, i.e., apnea and non-apnea subjects. After pre-processing, these rhythms are processed for feature extraction.

In literature, several features are reported for the classification of apnea and non-apnea states [8, 9, 10, 11, 12, 83, 92]. Here, we have adopted multiple domain feature parameters-Hjorth (HJ: activity, mobility, complexity) [93], Higuchi and Katz's fractal dimension measures (HFD) [94] in time domain, rational asymmetry (RASM) [95] in frequency domain



Table 2.7: Extracted the number of apnea and non-apnea epochs from EEG signals corresponding to every subject.

Subject ID	01	02	03	04	14	16	32	37	48	59	60	61	66	67x	Total epochs
Apnea epochs	168	236	402	224	268	146	124	70	274	256	122	208	166	56	4020
Non-apnea epochs	162	232	398	219	265	138	119	68	271	251	120	208	166	52	3978

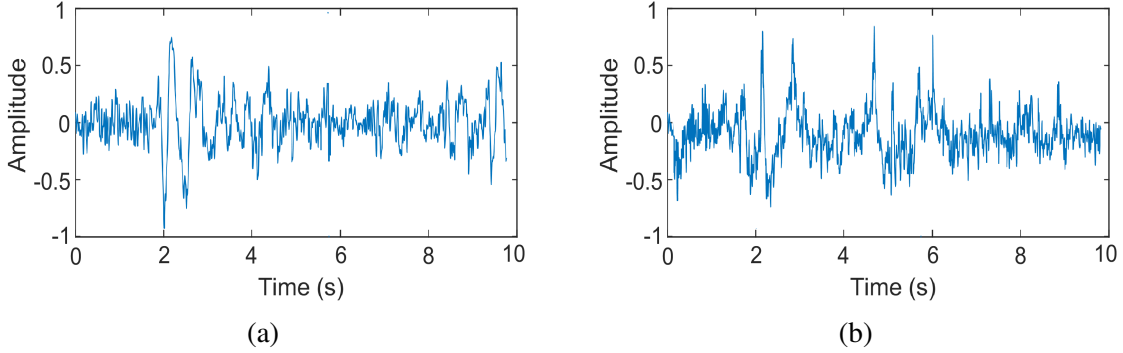


Figure 2.11: Plots of (a) epoch corresponding to apnea patient and (b) epoch corresponding to non-apnea subject.

and RE [96] in the TF domain. The features (HJ and HFD), RASM, and RE signify the variability, amplitude, and complexity of apnea and control EEG epochs. Here, RASM features were calculated from power values of frequency sub-bands. Whereas RE features were calculated from the TFR of SSDM bands, specifically from the delta, theta, and alpha bands. These TFR-based RE features were computed from decomposed SSDM bands incorporating information from both the time and frequency domains. These features are selected based on the better exploration of EEG characteristics such as variability, shape, complexity, and amplitude. It helps to obtain an optimal set of EEG features that accurately discriminate apnea events.

From the SSDM method, 24 EEG features were calculated for four EEG rhythms ( $\delta$ ,  $\theta$ ,  $\alpha$ , and  $\beta$ ) corresponding to three C3/O1, C4/A1, and O2/A1 cross channels, result in the feature vector of dimension 72. Then, the best 26 features were selected among 72 features using feature ranking based on the probabilistic ( $p$ )-values of the student t-test [66]. Similarly, 24 features for SWD and 18 features for HHT were selected by considering criteria ( $p < 0.05$ ). The statistical analysis of the most significant features chosen using the student

t-test for SSDM, SWD, and HHT methods, is presented in Table 2.8. The lowest  $p$ -values indicate that the significant features generated from the SSDM method show better discriminative features as compared to the features generated from SWD and HHT methods. Here, features have been annotated with convention (EEG channel location-feature type-EEG rhythm). In our work, we have presented two approaches for feature extraction. In the first approach, feature level fusion is introduced, which is extracted from multiple domain feature values for classification, inspired by Khateeb et al. work [95]. It is experimented with in different combinations – individually and in a fusion way. The most significant feature combinations are selected iteratively with the better-concatenated result. In the second approach, TFR as an image feature has been extracted from each decomposed EEG band. Here, the decomposed band signal has been transformed into TFR using HSA [41]. These TFR image features contain a total of 23868 samples for C3/O1, C4/A1, and O2/A1 cross channels. Every channel contributes 7956 TFR features. A total of 11934 TFR features belong to each class.

Table 2.8: The statistical analysis of most significant features chosen using student t-test for SSDM, SWD, and HHT methods.

Feature (Location- feature-EEG rhythm)	Proposed method			SWD method			HHT method		
	Apnea class (mean $\pm$ std)	Non- apnea class (mean $\pm$ std)	$p$ -value	Apnea class (mean $\pm$ std)	Non-apnea class (mean $\pm$ std)	$p$ -value	Apnea class (mean $\pm$ std)	Non- apnea class (mean $\pm$ std)	$p$ -value
C3/O1-RE- $\delta$	10.41 $\pm$ 9.41	6.20 $\pm$ 6.09	9.31e-24	8.45 $\pm$ 12.32	4.62 $\pm$ 3.42	4.62e-18	12.36 $\pm$ 12.14	7.72 $\pm$ 8.79	2.42e-12
C4/A1-HJ- $\delta$	8.79 $\pm$ 7.99	6.51 $\pm$ 6.41	9.77e-22	7.12 $\pm$ 8.36	5.24 $\pm$ 6.42	2.82e-16	10.46 $\pm$ 9.30	6.94 $\pm$ 5.90	4.31e-11
C4/A1-RE- $\delta$	14.28 $\pm$ 13.08	9.14 $\pm$ 8.84	3.37e-20	12.68 $\pm$ 10.82	8.20 $\pm$ 0.24	9.74e-16	10.08 $\pm$ 12.61	8.64 $\pm$ 10.37	3.79e-11
C3/O1-HJ- $\delta$	15.12 $\pm$ 16.38	8.56 $\pm$ 8.36	1.28e-18	16.45 $\pm$ 14.77	9.45 $\pm$ 10.38	7.83e-11	15.42 $\pm$ 22.03	6.92 $\pm$ 7.12	5.83e-10
C3/O1-RASM- $\delta$	9.98 $\pm$ 9.20	6.99 $\pm$ 6.89	1.05e-18	13.45 $\pm$ 12.09	7.28 $\pm$ 8.64	4.13e-16	12.53 $\pm$ 10.30	8.09 $\pm$ 6.12	9.73e-10
O2/A1-RE- $\delta$	13.98 $\pm$ 13.18	8.17 $\pm$ 7.87	3.23e-18	11.65 $\pm$ 15.14	7.38 $\pm$ 8.52	3.65e-14	3.72 $\pm$ 4.67	1.86 $\pm$ 2.36	4.36e-12
C4/A1-HFD- $\delta$	12.25 $\pm$ 10.95	7.12 $\pm$ 6.92	4.07e-16	14.31 $\pm$ 16.24	9.17 $\pm$ 8.50	8.45e-13	27.40 $\pm$ 30.86	18.93 $\pm$ 16.17	1.77e-10
C4/A1-RASM- $\delta$	9.24 $\pm$ 8.54	7.07 $\pm$ 6.97	1.06e-17	10.02 $\pm$ 9.37	6.24 $\pm$ 5.07	6.46e-12	6.34 $\pm$ 5.39	3.57 $\pm$ 2.97	3.02e-9
C3/O1-HFD- $\delta$	14.65 $\pm$ 13.22	8.32 $\pm$ 8.02	8.25e-16	22.14 $\pm$ 28.02	12.93 $\pm$ 11.28	3.16e-12	15.06 $\pm$ 16.08	7.82 $\pm$ 10.68	7.86e-8
O2/A1-HJ- $\delta$	12.25 $\pm$ 11.23	7.91 $\pm$ 7.71	2.19e-16	10.48 $\pm$ 12.78	8.20 $\pm$ 9.64	2.76e-11	26.36 $\pm$ 28.78	19.24 $\pm$ 18.08	6.81e-7
C4/A1-RE- $\theta$	8.23 $\pm$ 7.63	5.91 $\pm$ 6.01	8.87e-16	14.75 $\pm$ 16.91	8.25 $\pm$ 8.15	3.19e-10	7.41 $\pm$ 8.09	4.21 $\pm$ 6.46	2.58e-7
C3/O1-RE- $\theta$	4.24 $\pm$ 3.74	3.12 $\pm$ 2.82	4.68e-15	9.41 $\pm$ 11.60	5.65 $\pm$ 4.18	1.54e-10	10.48 $\pm$ 9.39	5.09 $\pm$ 7.54	6.23e-6
C3/O1-RASM- $\theta$	2.86 $\pm$ 2.06	3.43 $\pm$ 3.23	5.15e-15	6.02 $\pm$ 8.66	4.31 $\pm$ 5.82	6.12e-9	36.16 $\pm$ 42.21	17.48 $\pm$ 14.57	5.37e-4
O2/A1-HFD- $\theta$	2.98 $\pm$ 2.18	0.49 $\pm$ 0.39	7.40e-15	13.46 $\pm$ 12.89	7.53 $\pm$ 10.02	4.87e-8	7.43 $\pm$ 8.63	5.69 $\pm$ 3.02	2.38e-3
C4/A1-HJ- $\theta$	3.02 $\pm$ 3.32	0.51 $\pm$ 0.53	7.67e-14	8.92 $\pm$ 11.53	5.42 $\pm$ 4.45	4.06e-6	13.28 $\pm$ 11.46	8.06 $\pm$ 10.19	1.45e-2
O2/A1-RASM- $\delta$	12.23 $\pm$ 10.93	7.22 $\pm$ 7.02	1.08e-13	4.35 $\pm$ 5.32	2.89 $\pm$ 2.06	3.62e-5	2.85 $\pm$ 3.54	1.57 $\pm$ 1.78	1.08e-2
C3/O1-RE- $\alpha$	6.25 $\pm$ 5.65	4.28 $\pm$ 4.18	6.43e-12	24.04 $\pm$ 28.26	16.12 $\pm$ 15.20	3.79e-5	32.57 $\pm$ 20.68	15.82 $\pm$ 12.63	8.99e-2
C4/A1-RE- $\alpha$	8.21 $\pm$ 7.67	5.10 $\pm$ 4.80	7.66e-12	36.14 $\pm$ 50.31	19.34 $\pm$ 15.26	8.24e-4	22.80 $\pm$ 18.72	12.89 $\pm$ 14.60	3.72e-2
C3/O1-HJ- $\theta$	5.53 $\pm$ 4.73	3.76 $\pm$ 3.56	5.65e-11	13.88 $\pm$ 12.72	9.60 $\pm$ 8.52	9.26e-3	2.75 $\pm$ 2.53	1.05 $\pm$ 1.5	1.31e-1
C4/A1-RASM- $\theta$	4.25 $\pm$ 3.35	0.12 $\pm$ 0.02	1.92e-11	8.63 $\pm$ 9.27	5.87 $\pm$ 6.08	8.22e-2	6.21 $\pm$ 7.05	5.41 $\pm$ 4.60	1.87e-1
C3/O1-HFD- $\theta$	3.23 $\pm$ 2.55	1.78 $\pm$ 1.48	2.67e-09	6.28 $\pm$ 8.19	3.28 $\pm$ 5.42	5.03e-2	1.74 $\pm$ 1.03	0.92 $\pm$ 1.56	1.05e-1
C4/A1-HJ- $\theta$	4.58 $\pm$ 3.68	4.29 $\pm$ 4.09	6.44e-10	32.41 $\pm$ 26.92	17.63 $\pm$ 18.03	4.21e-2	24.21 $\pm$ 20.75	15.08 $\pm$ 18.26	9.38e-1
C3/O1-HJ- $\alpha$	18.12 $\pm$ 19.92	10.51 $\pm$ 10.41	1.58e-07	15.75 $\pm$ 13.54	10.93 $\pm$ 13.79	8.75e-1	10.86 $\pm$ 12.36	8.29 $\pm$ 8.72	8.52e-1
C4/A1-HJ- $\alpha$	24.12 $\pm$ 32.35	13.55 $\pm$ 15.25	4.03e-06	28.41 $\pm$ 48.09	20.47 $\pm$ 18.07	5.38e-1	5.79 $\pm$ 4.92	3.21 $\pm$ 4.07	2.86e-1
C4/A1-RASM- $\alpha$	8.25 $\pm$ 6.75	5.67 $\pm$ 5.47	8.51e-05	12.28 $\pm$ 15.83	9.91 $\pm$ 11.62	8.17e-1	3.58 $\pm$ 4.07	2.14 $\pm$ 2.84	6.25e-1
C3/O1-HFD- $\alpha$	3.86 $\pm$ 2.46	0.93 $\pm$ 0.83	1.40e-05	4.63 $\pm$ 4.72	3.29 $\pm$ 5.03	1.43e-1	2.29 $\pm$ 2.17	1.83 $\pm$ 1.56	1.47e-1

std: Standard deviation.

Further, the obtained feature fusion and TFR have been tested with the pre-trained clas-

sifier networks, namely CNN (AlexNet, ResNet50, VGG-16, ShallowNet, and LeNet) [97] and BiLSTM [98], respectively. Figs. 2.12 (b) and 2.12 (c) illustrate the frameworks of implemented CNN and BiLSTM classifiers, respectively. The selection of CNN for the TFR feature is because of its proven potential for image classification, multi-object analysis, and face detection [99]. The CNN is a robust classifier that employs a pooling kernel to optimize the number of parameters and discriminate disturbed input data systematically. It effectively aggregates interchannel and inter frequency-band information without downsampling the feature map. All classifiers configuration is set 70% data for training the network and 30% for testing the network. Similarly, BiLSTM has been employed to test feature fusion. Reason to select BiLSTM to find its suitability to investigate features with appropriate model parameters and counteract the over-fitting problem with fewer hidden layers [100][101]. Table 2.9 presents the architecture details of the CNN (AlexNet) and BiLSTM models.

Table 2.9: Architectures of the CNN (Alexnet) and BiLSTM.

CNN (Alexnet) model				BiLSTM model			
Layer	Number of filters	Size of feature map	Number of strides	Layer	Memory cell	Activation	Parameters
Image input layer	$227 \times 227 \times 3$			Batchnorm-0			4
1st CLs	96	$11 \times 11 \times 96$	[4 4]	RNN-1	128	sigmoid	31824
RLU layer				Batchnorm-1			380
BNL layer				RNN-2	1008	sigmoid	54200
MPL layer	96	$5 \times 5 \times 96$	[2 2]	Batchnorm-2			300
2nd CLs	128	$5 \times 5 \times 128$	[1 1]	RNN-3	64	sigmoid	28420
RLU layer				Batchnorm-3			260
BNL layer				RNN-4	32	sigmoid	16380
MPL layer	128	$3 \times 3 \times 128$	[2 2]				
3rd CLs	384	$3 \times 3 \times 384$	[1 1]	Batchnorm-4			120
RLU layer				Dense layer-1			8460
4th CLs	192	$3 \times 3 \times 192$	[1 1]	RNN-5	16	sigmoid	3820
RLU layer				Batchnorm-5			70
5th CLs	128	$3 \times 3 \times 128$	[1 1]	Dense layer-2			280
RLU layer				Dense layer-3	2	softmax	28
MPL layer	128	$1 \times 1 \times 128$	[2 2]				
1st fully connected layer		$1 \times 4096$					
DOL layer (50%)							
2nd fully connected layer		$1 \times 4096$					
RLU layer		$1 \times 4096$					
DOL layer (50%)		$1 \times 4096$					
3rd fully connected layer		$1 \times 1000$					
Softmax layer		$1 \times 1000$					
Output layer		$1 \times 1000$					

RLU: Rectified linear unit, BNL: Batch normalization layer, MPL: Max-pooling layer, DOL: Dropout layer, CLs: Convolution layers.

Table 2.10 shows the feature-specific classification performance which is obtained us-

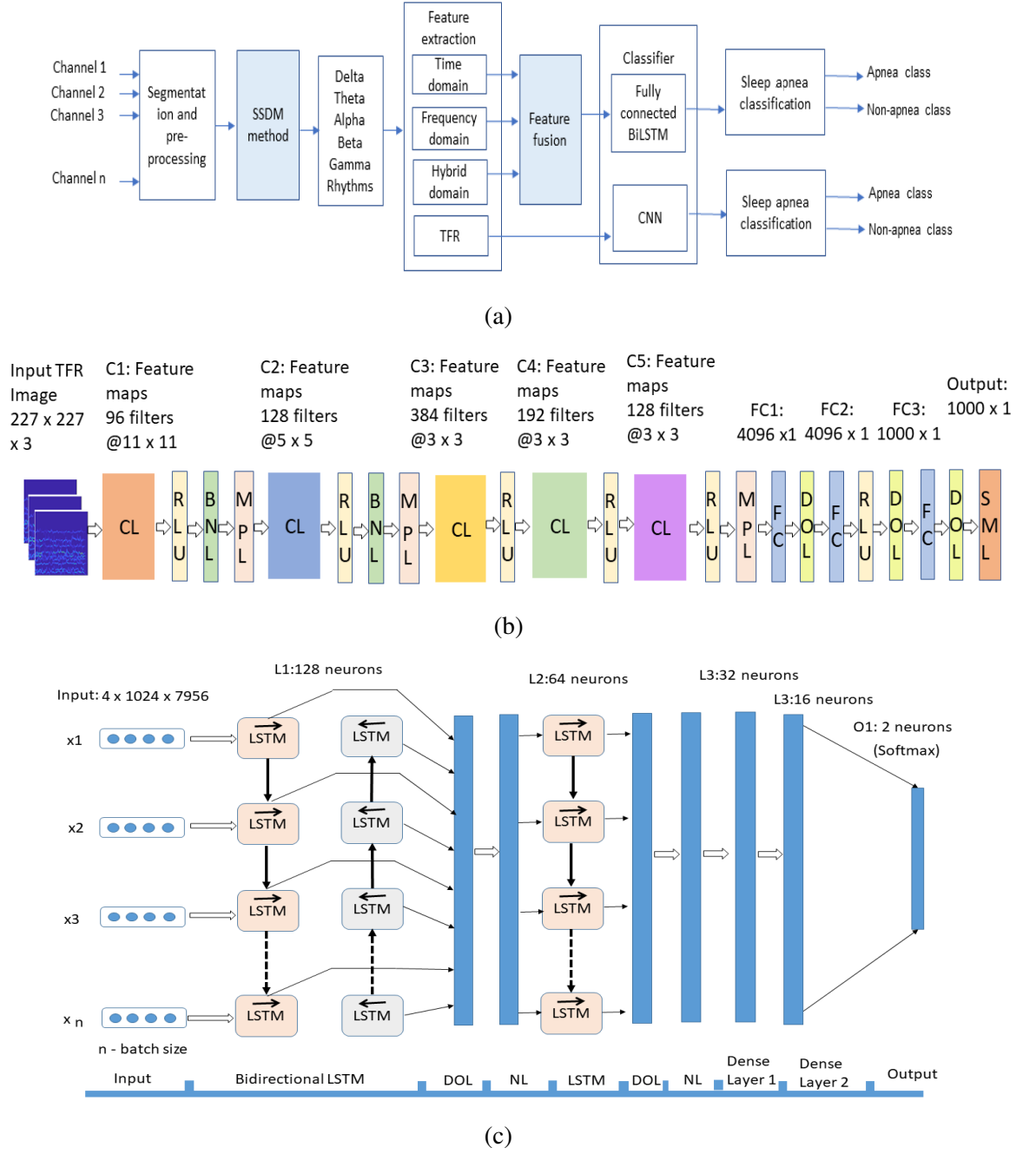


Figure 2.12: Schematic representation of (a) the proposed sleep apnea recognition model based on the SSDM method, (b) CNN unit block, and (c) BiLSTM unit block.

ing the SSDM, SWD, and HHT methods. For cross-validation purposes, we have adopted subject-independent analysis [10]. It employed leave-one-out cross-validation in which training and testing data were considered from the cross-subject dataset. This process was repeated for all subjects, and average results were reported. In order to evaluate the

proposed method, four performance parameters were considered: accuracy (ACC), sensitivity (SEN), specificity (SPE), and Fisher F1-score. Here, we have compared the proposed SSDM-based frameworks (SSDM-FF-BiLSTM and SSDM-TFR-CNN) with SWD-based frameworks (SWD-FF-BiLSTM and SWD-TFR-CNN) and HHT-based frameworks (HHT-FF-BiLSTM and HHT-TFR-CNN ) for respective extracted features in the sleep apnea recognition application. Table 2.10 demonstrates that the proposed SSDM-TFR-CNN (AlexNet) and SSDM-FF-BiLSTM methods outperformed the existing methods, HHT and SWD. The highest accuracy among all frameworks were 96.24% and 95.86% in the case of TFR and feature fusion (HJ and RE) respectively. The SWD-FF-BiLSTM produced a second-best accuracy score where the accuracy score was 94.36% in the case of feature fusion (HFD, HJ, and RE). Whereas HHT-based frameworks, HHT-FF-BiLSTM and HHT-TFR-CNN (AlexNet), reported the performance among the frameworks were 89.90% and 88.29% in the case of feature fusion (HFD, HJ, and RE) and TFR feature. The inferior result reported was 86.40% in the case of HHT-FF-BiLSTM with the feature fusion (HFD and HJ).

In the second approach, TFRs obtained from decomposed EEG bands of SSDM, SWD, and HHT were fed to different existing CNN architecture models, namely CNN (AlexNet, ResNet50, VGG-16, LeNet, and ShallowNet). Here, we have tested CNN architecture with adaptive momentum (Adam) optimizer. Table 2.10 shows that the SSDM-TFR-CNN (AlexNet) found the suitability to sleep apnea classification and deployed overall best classification accuracy of 96.24% comparatively. For SSDM-TFR-CNN (AlexNet), sensitivity, specificity, and  $F1$ -score of 97.60%, 96.72%, and 0.9820 are obtained, respectively. Additionally, SSDM-TFR-CNN (VGG16) and SSDM-TFR-CNN (ResNet) frameworks delivered better classification performance with an accuracy score of 96.08% and 94.26%, but it shows slightly inferior performance as compared to the SSDM-TFR-CNN (AlexNet). The SWD-TFR-CNN (VGG-16) delivered a better accuracy score of 93.89% over the HHT-TFR-CNN (AlexNet), where the accuracy was reported at 88.29%. From this comprehensive analysis, it must be noted that the TFR feature and feature fusion have a significant contribution to the apnea classification problem. The obtained highest classification accuracies using the proposed method for the classification of TFR images with CNN-Alexnet classi-

fier (SSDM-TFR-CNN) and feature fusion (HJ and RE) with BiLSTM classifier (SSDM-FF-BiLSTM) are 96.24% and 95.86%, respectively.

Table 2.10: Feature specific cross-validation performance obtained using the SSDM, SWD, and HHT methods.

Feature Type	Feature used	Classifier	Proposed method				SWD method				HHT method			
			ACC (%)	SEN (%)	SPE (%)	F1-score	ACC (%)	SEN (%)	SPE (%)	F1-score	ACC (%)	SEN (%)	SPE (%)	F1-score
Individual feature	HFD	BiLSTM	92.53	90.54	89.73	0.9425	91.67	89.72	87.51	0.9261	87.41	88.54	88.14	0.8491
	HJ	BiLSTM	93.75	94.20	94.75	0.9613	91.85	92.16	92.76	0.9241	88.96	85.82	84.54	0.8690
	RASM	BiLSTM	92.60	89.45	92.73	0.9350	90.60	86.03	90.14	0.8980	87.20	85.68	83.21	0.8613
	RE	BiLSTM	<b>94.20</b>	95.12	95.50	0.9680	<b>94.03</b>	94.64	94.32	0.9406	<b>89.70</b>	90.38	90.59	0.8990
	(HFD, RE)	BiLSTM	93.42	93.65	92.44	0.9460	91.20	91.54	90.78	0.9420	88.45	89.39	89.78	0.9034
	(HJ, RE)	BiLSTM	<b>95.86</b>	96.50	95.90	0.9742	94.24	94.90	93.14	0.9512	89.83	90.41	90.67	0.9065
Fusion feature	(RASM, RE)	BiLSTM	93.05	94.07	94.14	0.9552	91.11	92.08	92.57	0.9286	89.05	88.02	86.67	0.8629
	(HFD, HJ, RE)	BiLSTM	94.72	95.63	95.40	0.9625	<b>94.36</b>	95.30	94.33	0.9580	<b>89.90</b>	91.67	90.89	0.9085
	(HFD, HJ)	BiLSTM	92.24	90.79	92.17	0.9406	90.27	90.86	89.72	0.9105	88.64	86.50	88.90	0.8752
	(HFD, HJ, RASM)	BiLSTM	92.52	91.12	93.84	0.9358	90.36	90.73	92.81	0.9276	86.40	85.25	86.92	0.8544
	(HFD, HJ, RASM, RE)	BiLSTM	93.40	94.78	93.41	0.9450	92.89	92.24	90.16	0.9315	89.23	90.07	89.55	0.8871
TFR feature	TFR	CNN (ResNet)	94.26	96.18	95.58	0.9674	92.82	92.90	91.86	0.9304	87.14	84.39	86.47	0.8650
	TFR	CNN (AlexNet)	<b>96.24</b>	97.58	96.72	0.9820	93.24	93.85	90.54	0.9286	<b>88.29</b>	85.60	89.17	0.8979
	TFR	CNN (LeNet)	93.50	94.40	93.62	0.9382	91.30	92.32	90.49	0.9050	84.02	83.14	82.45	0.8460
	TFR	CNN (VGG-16)	<b>96.08</b>	95.85	96.76	0.9645	<b>93.89</b>	93.71	91.50	0.9328	86.71	85.92	88.20	0.8701
	TFR	CNN (Shal-lowNet)	93.12	93.18	92.46	0.9350	90.86	90.59	89.75	0.9089	84.08	82.15	83.94	0.8389

Table 2.11 shows performance comparison of the proposed SSDM method with already reported methods in the literature for the automated classification of sleep apnea disorder using EEG signals. In our work, the comparison was made with the same EEG apnea dataset [85]. The comparison is considered based on the subject-dependent cross-validation approach and features complexity. In Table 2.11, the mentioned work in [11, 12, 13, 83, 84] already achieved less classification accuracy as compared to the already reported work [8, 9, 10]. Whereas the work mentioned in [8, 9] achieved the highest accuracy for sleep apnea disorder classification at approximately 98%. But method's performance suffers as the subject-dependent approach is accommodated. On the other side, Mahmud et al. [10] have employed subject-independent testing and validation scheme for extracted features using VMD and BiLSTM and yielded average classification accuracy of 95.54% with 50% overlapping in the training set. In extension to this, our work delivers competent classification

Table 2.11: Performance comparison with the methods available in the literature.

Authors, year, and reference	Methodology	Classification performance		
		SEN (%)	SPE (%)	ACC (%)
Ahmed et al. (2016) [83]	Temporal variation of spectral energy in beta band, KNN	90.58	77.72	82.28
Soha et al. (2016) [84]	Entropy, KNN	89.02	86.27	87.64
Shahnaz et al. (2016) [11]	Delta band power, KNN	89.62	84.43	87.03
Bhattacharjee et al. (2018) [12]	Rician modeling, KNN	95.43	81.56	88.49
Almuhammadi et al. (2015) [9]	IIR Butterworth BPF, (SVM, ANN, LDA and Naive Bayes)	97.01	97.26	97.14
Tanvir et al. (2020) [13]	Temporal variation, FCNN	85.52	88.97	88.22
Taran et al. (2017) [8]	Hermite coefficients, PSO-LSSVM	98.66	99.03	98.82
Mahmud et al. (2021) [10]	VMD, FCNN, BiLSTM	95.35	93.00	95.54
<b>Proposed framework-1</b>	<b>SSDM-TFR-CNN (AlexNet)</b>	<b>97.58</b>	<b>96.72</b>	<b>96.24</b>
<b>Proposed framework-2</b>	<b>SSDM-FF-BiLSTM</b>	<b>96.50</b>	<b>95.90</b>	<b>95.86</b>

IIR: Infinite impulse response, KNN: K-nearest neighbors, BPF: Band-pass filter, ANN: Artificial neural network, VMD: Variational mode decomposition, FCNN: Fully convolutional neural network, PSO-LSSVM: Particle swarm optimization-least squares SVM, and BiLSTM: Bidirectional long short-term memory.

Table 2.12: The most significant five feature fusion chosen using student t-test for SSDM, SWD, and HHT methods.

Feature	Proposed method			SWD method			HHT method		
	Apnea class (mean±std)	Non-apnea class (mean±std)	<i>p</i> -value	Apnea class (mean±std)	Non-apnea class (mean±std)	<i>p</i> -value	Apnea class (mean±std)	Non-apnea class (mean±std)	<i>p</i> -value
(HJ, RE)	4.57±5.02	2.53±1.36	2.39e-21	6.48±6.49	3.82±4.14	3.40e-18	4.59±3.50	2.18±2.45	2.91e-11
(HFD, HJ, RE)	6.78±8.60	4.60±3.41	1.68e-18	9.60±8.52	4.98±4.23	5.21e-16	6.86±4.34	3.36±4.54	3.57e-10
RE	8.29±8.02	5.20±5.84	4.60e-15	7.58±6.09	4.72±3.60	4.20e-12	4.60±3.72	2.28±2.79	1.49e-9
HJ	5.12±5.45	3.68±3.36	1.54e-14	3.72±4.12	2.54±1.83	2.56e-10	4.68±5.82	2.38±1.06	2.65e-8
(HFD, HJ, RASM, RE)	3.52±6.38	1.56±1.49	1.09e-10	4.58±4.03	2.71±1.97	2.10e-8	1.47±1.27	1.46±0.97	1.49e-6

std: Standard deviation.

results with subject-independent cross-validation over these presented works. The advantage of our proposed work over these existing approaches is that we have used our proposed SSDM for extracting TFR and feature fusion. It is tested with subject subject-independent cross-validation scheme for effective apnea events discrimination.

In Table 2.11, the comparison shows that the proposed SSDM method-based classification outperformed others and achieved the higher average classification accuracies on the feature fusion and TFR feature separately with SSDM-FF-LSTM and SSDM-TFR-CNN (Alexnet) classification frameworks are 95.86% and 96.24%, respectively. Here, we ar-

gue that the proposed SSDM method provides more distinctive and significant features that boost sleep apnea classification performance by enhancing interpretability and inter-channel information. It deploys the most discriminate pattern of EEG variations for effective apnea events discrimination. The statistical significance of the most significant five features obtained from proposed SSDM, SWD, and HHT methods are shown in Table 2.12. It showed that feature improvement is proven by performing a statistical t-test where a  $p$ -value of ( $< 0.05$ ) is obtained for apnea classification. Here, we have measured the computational time complexity to extract features (FF and TFR) by the SSDM method and classifier separately. The average computation time taken for feature extraction and classification is 0.84 s and 0.273 s. The experimentation was performed in MATLAB R2021b with Intel(R) Core(TM) i7-2600 CPU at 3.40 GHz. This computational time of feature extraction by the SSDM method is solely considered based on the course iteration parameters of the filter bank ( $P_{th} = 0.1$ ,  $M = 20$ , and  $StD_{th} = 0.5$ ). Despite using subject-independent cross-validation, the proposed SSDM method significantly delivers higher performance that verifies suitability for practical applications, especially in the sleep apnea detection of EEG signals.

To highlight the limitation of the proposed work, we have identified many potential confounding factors that we have not attempted to consider in this study and further presented future remedies. Firstly, even though the proposed framework showed suitability to diagnose sleep apnea over the existing approaches [10, 11, 12, 83, 92], here we have restricted our study to binary classification to suspected OSA patients on a single publicly available EEG database [85]. For clinical significance, detecting a higher number of sleep stages and severity is crucial to understand the complexity associated with OSA problems dealing with machines [7, 102, 103, 104]. Therefore, in future work, we can extend the proposed framework to a sleep-related multi-classification model and accommodate the other public and private sleep apnea-related databases with a subject-independent cross-validation approach. The second potential confounding factor is finding the most discriminative features (handful features) in place of the CNN model, which establishes an exact relation to sleep hypnotic apnea event. This can reduce the CNN classifier training time and enhance practicability of the automated medical system for diagnosis of the sleep-related disorders. From the literature [88, 105, 106], authors have used multi-model signals such as ECG, EMG, PPG, and



EOG for OSA classification, and it achieved satisfactory results. In the present work, EEG signals are solely used to diagnose apnea events which may cause deviations with other neurological diseases and suffer the classification accuracy of sleep apnea. Therefore, we could verify and enhance the performance of the developed sleep apnea method when applying EEG signals together with other cross-signal modalities, which will be investigated in our future work.

## **2.4 Summary**

In this work, our proposed TF analysis method, SSDM, has demonstrated a noteworthy improvement in the TF analysis of nonstationary signals. The contributions of this work are motivated by SWD and fused-LASSO-based SFT spectrum estimation techniques for decomposing nonstationary multicomponent signals into a finite number of OCs. The proposed method has been applied to nonstationary multicomponent AM and FM signals and real EEG signals and achieves improved TFR and signal decomposition over the other two state-of-the-art methods, HHT and SWD. Further to test potential, the proposed method is applied to EEG signals of sleep apnea detection and developed two novel feature schemes: Fusion feature and TFR. Using these obtained features, the proposed SSDM-FF-LSTM and SSDM-TFR-CNN classification frameworks are employed for sleep apnea detection and tested with a subject-independent cross-validation approach.

In the SSDM-TFR-CNN framework, TFR images obtained by SSDM of apnea and non-apnea classes are given to five pre-trained CNN models. The performance of five pre-trained networks AlexNet, ResNet50, LeNet, SwallowNet, and VGG-16, are compared. The results obtained through these networks show that AlexNet delivers the best results with less training and testing time. From comparative results, the proposed framework has reported the best classification performance with an accuracy of 96.24%, sensitivity of 97.60%, specificity of 96.72%, and F1-score of 0.9820.

In the SSDM-FF-LSTM framework, 72 multiple domain features are extracted from decomposed SSDM bands of sleep apnea EEG signals. The most significant features are computed with the student t-test and fed to the BiLSTM classifier. The SSDM-FF-LSTM

framework employed the feature fusion (HJ and RE), which provided better results as compared to all the feature combinations with accuracy, sensitivity, specificity, and F1-score equal to 95.86%, 96.50%, 95.90%, and 0.9742, respectively. The obtained comparison shows that the proposed SSDM method-based classification framework outperformed other SWD-FF-LSTM and HHT-FF-LSTM frameworks. From this comprehensive analysis, it must be noted that the TFR feature and feature fusion have a significant contribution to the apnea classification problem. The obtained highest classification accuracies using proposed methods for the classification of TFR images with CNN-Alexnet classifier (SSDM-TFR-CNN) and feature fusion (HJ and RE) with BiLSTM classifier (SSDM-FF-BiLSTM) are 96.24% and 95.86% against the reported classification approaches for sleep apnea classification. Therefore, the proposed SSDM method can be a powerful tool for sleep apnea investigation of EEG signals. In the future, we will further improve accuracy corresponding to the sleep-related multi-classification model to determine the sleep apnea types and severity based on different biopotential signals with EEG.



## **Chapter 3**

# **Multivariate SSDM for Multivariate Signal Analysis with Cognitive Visual Object Detection from Multichannel EEG-MEG Signals**

In this chapter, a novel adaptive MSSDM is proposed for the multivariate analysis of nonstationary EEG-MEG signals. This method decomposes multichannel EEG and MEG data into multivariate oscillatory modes which determine channel-aligned common modes across all channels. Consequently, we conducted an ablation experiment to examine how various elements of our proposed MSSDM approach impact the multivariate analysis of nonstationary signals (EEG and MEG) and improve reconstruction performance. The designing of MSSDM, along with its effectiveness tested on synthetic and real EEG-MEG signals, is discussed in detail in this chapter. Furthermore, we developed classification frameworks to validate the effectiveness of our proposed MSSDM approach in decoding visual object categories for BCI applications. These frameworks focus on enhancing visual object detection using newly formulated RCFF computed from rhythms derived from multivariate CAOCs of EEG, MEG, and combined EEG-MEG data. This approach not only extracts strongly correlated class-specific multivariate information across multichannel data

but also facilitates effective multivariate analysis of EEG-MEG signals for improved visual object decoding. The classification performance of the proposed frameworks is explored in the subsequent sections.

## **3.1 Introduction**

Over the recent decades, substantial improvement in multivariate pattern analysis of MEG and EEG has placed fundamental importance on recognition of visual objects through neural activation patterns. To identify objects and object categories from brain responses, MEG and EEG signals are recorded in response to visual stimuli i.e. images from different categories are presented to the participants in experimental trials while their brain activities are recorded [3]. Researchers have explored several experimental techniques for recognizing visual objects, including EEG [5], MEG [14], ECOG, and fMRI [3, 14]. Among these studies, MEG and EEG are the widely used techniques for visual recognition systems because they can be non-invasive in nature and ability to offer enhanced fine-grained analysis by identifying spatial, temporal, and spectral components underlying object category discrimination [3, 14].

In order to investigate the effectiveness of using EEG and MEG signals for visual recognition, many research studies have been carried out and promising results have been reported [3, 4, 5, 14]. Cichy et al. have introduced a feature i.e., peak latency time points (PLSP) in [3] that is based on the 306-channel MEG of categorical representation of objects in the human visual response. This study uses PLSP with linear discriminant analysis (LDA) frameworks to identify spatial and temporal MEG components that best discriminate object categories by obtaining a classification accuracy of 68.75% on subsets of brain responses. Further, Cichy et al. [14] have extended the potential of the EEG and MEG multichannel data combinable to multivariate analysis of visual object problems. This study demonstrated that the extracted representational similarity analysis (RSA) features from combined MEG and EEG signals enhance the co-relation between temporal dynamics and category information in recognition of visual objects and deliver the highest classification performance than in MEG and EEG signals alone. In recent works, [5] has introduced a visual ob-

ject classification approach by adopting an informative representational dissimilarity matrix (RDM) as novel self-similarity measures on a 306-channel MEG. It includes multivariate analysis based on the RDM feature, which is computed on the 100 time points of the MEG signal and tested using four different supervised classifiers namely, LDA, support vector machine (SVM) with linear kernel, Gaussian Naïve Bayes (GNB), and weighted robust distance (WeiRD). This study also incorporates multivariate noise normalization in the pre-processing stage for classification performance analysis. The study demonstrates the highest average accuracy across different classifiers: LDA (74.5%), SVM (74%), WeiRD (85%), and GNB (75%). Furthermore, Kong et al. [4] have employed a shallow CNN network to project visual activation response into a 74-channel EEG-based learned RDM feature and achieved a classification accuracy of around 65.6% in a 5-class visual category. However, in the above visual object classification approaches, the computed features from raw EEG or MEG multichannel signals show strong dependency on the formulation of the classification model, thus it may suffer in computing the most significant features with strong relevance to the visual object category [3, 4, 5, 14].

In real-world environments, complex visual object recognition faces significant challenges, such as the selection of channels, smoothing and data reduction, the undeterminable scalp regions, and the nonstationary nature of EEG and MEG signals, which limit mutually informative features across channels and consequently degrade the classification performance of visual object detection. These mutual features show multivariate modulated oscillations which correspond to the matching of the oscillatory component modes with similar multiscale spectral content across multiple channels. Therefore, robust multivariate visual analysis is urgently demanded to detect the most significant features with optimizing channels and computational efficiency. In order to get improved fine-grained characteristics of the temporal dynamics of visual neural activity, the primary motivation for this work to study nonstationary multicomponent EEG-MEG signals is to enhance signal denoising and separation of spectral components to compute multi-spectral features for improving multivariate analysis of visual objects. However, the relation that EEG-MEG in the context of signal decomposition-based rhythm extraction for multivariate analysis has not been investigated.

To resolve the aforementioned limitations, we have presented a novel extension of the SSDM, termed as MSSDM, to classify VSA patterns based on multivariate EEG-MEG multichannel signal data. In the proposed MSSDM, a mode alignment approach is used to determine the orders of effective OC modes in multivariate EEG-MEG signals so that the same frequency components across different channels can be extracted from different OCs and can establish spectral correlation among the different channels. Using multivariate swarm filtering and SFT spectrum, the MSSDM method employs an adaptive scale-space approach to estimate spectral boundaries in the SFT to obtain aligned OCs across different EEG-MEG channels. In addition, the approach utilizes a spectrum optimization technique, fused-Lasso for SFT spectrum estimation [81]. The multi-scale features obtained from these effective OCs extract enhanced spectral dynamics of neural activity related to the common or unique aspects of visual object representation. As per our understanding, MSSDM decomposition-specific oscillatory rhythms-based multi-scale features have not been attempted for the recognition of visual objects. The novelty of the present work is to extract MSSDM-based multi-scale features from EEG-MEG signals and use a suitable classification model for the automated classification of visual objects. The key contributions of the presented work are given as follows:

1. MSSDM: A new adaptive decomposition method has been proposed to explore the multivariate analysis of nonstationary EEG-MEG signals to improve VSA classification performance.
2. The proposed method has also been tested for synthetic signals and real EEG-MEG signals.
3. The new RCFF features have been computed from the features, namely, RE, SE, CSPTE, and CSD, which give a multi-scale analysis of EEG, MEG, and combined EEG-MEG multichannel data and improved VSA classification.
4. It compared a developed automated system of visual object detection against the current state-of-the-art methods with subject-independent cross-validation from the cross-subject dataset.

## 3.2 Multichannel dataset

In this study, we have evaluated our proposed technique on the publicly available human EEG and MEG dataset [14] with responses to 92 images that were used as visual stimuli. The dataset contains 74-electrode EEG and 306-electrode MEG responses from 16 subjects (eight females with a mean age  $23.87 \pm 4.5$  years and eight males with a mean age  $24.37 \pm 4.1$  years) viewing each of 92 stimuli. In our work, we have employed five categories (12 images per category), namely, human body (HB), human face (HF), animal body (AB), animal face (AF), and inanimate objects (IO) from all 380-channel for further multivariate analysis. Data were recorded at a sampling rate of 1000 Hz, and online filtering between 0.03 Hz and 330 Hz was applied. Fig. 3.1 illustrates the timing scheme of the paradigm that was used in the study. Every trial extracts visual stimulus between -100 ms and +900 ms (around 1 s). In total, each participant completed 2820 trials, including 30 trials of 94 visual objects.

## 3.3 Proposed methodology

The proposed MSSDM-based classification framework for visual object recognition using EEG and MEG multichannel signals is presented in Fig. 3.2 and detailed in the following section.

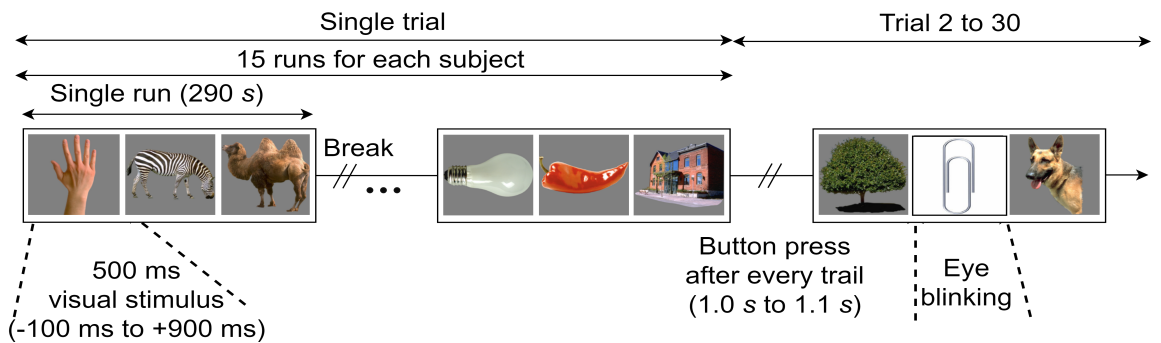


Figure 3.1: The timing scheme of the experimental paradigm for the visual object database.



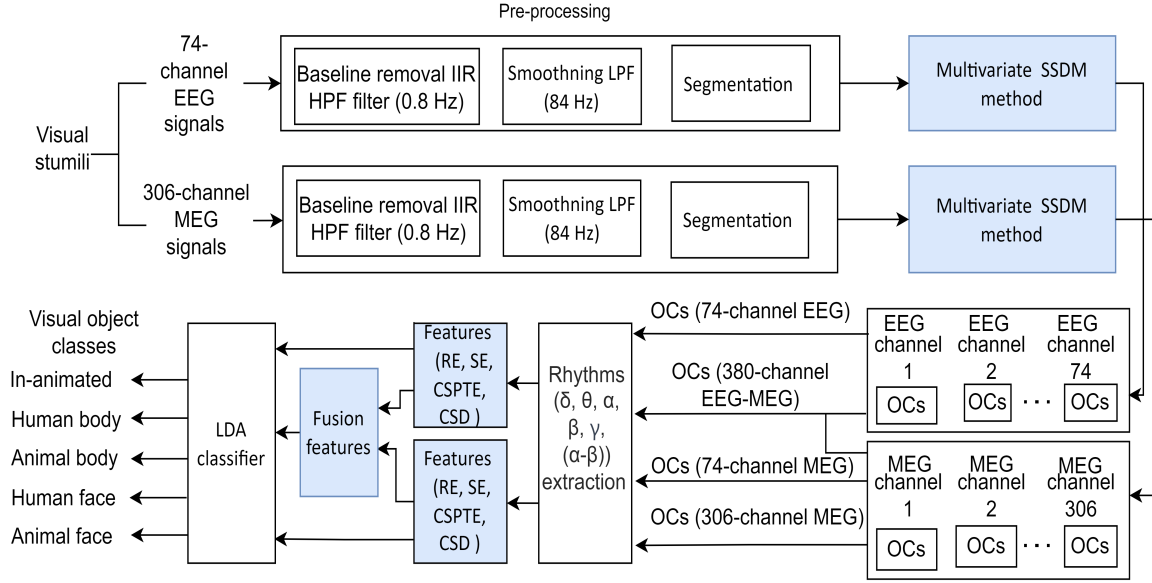


Figure 3.2: Block diagram of the proposed approach for visual object recognition using the MSSDM method.

### 3.3.1 Preprocessing

The preprocessing stage mainly comprises the implementation of filtering and segmentation techniques. In filtering, it comprises two filters (smoothing LPF (30 Hz) and baseline removal infinite impulse response (IIR) high-pass filter (HPF) (0.8 Hz)) which filter out noise and artifacts from EEG and MEG signals. A Matlab MNE toolbox was initially used to extract and clean EEG and MEG data from each participant. In segmentation, multichannel EEG and MEG signal epochs of one-second duration are employed to decompose adaptive OCs. During the segmentation process, the data samples were cut into epoch sizes of 1000 ms windows of time-locked post-stimulus responses. MEG and EEG signals are acquired for 72 minutes using 380-channel at a sampling rate of 1000 Hz for each image. In our work, we have considered the smallest epoch size of 30 among the 16 subjects' epochs. Further, the 30 considered epochs from the 380-channel of EEG and MEG are concatenated to construct the raw pattern vectors separately as well as combined. Each vector contains 2220 ( $30 \times 74$ ) patterns, 9180 ( $30 \times 306$ ) patterns, and 10800 ( $30 \times 380$ ) patterns, which correlate with EEG (74-channel), MEG (306-channel), and combined EEG-MEG (306-channel) for each class.

Figure 3.3 shows the TF analysis which is computed using USSDM and HSA [41] from the visual stimuli responses of the three visual object categories: HB, HF, and AB from the five different EEG channels [EEG005 (FC2), EEG042 (T8), EEG003 (FC1), EEG041 (T7), and EEG036 (F4)] for subject 6. In the TF plots, the highest activation during the visual cognitive tasks was observed in the 8-30 Hz rhythm with imagined stimulus response during the timeframe (-100 to 900 ms) from the onset of imagination. These TF responses exhibit discriminative neural activity patterns due to the cognitive imagination tasks (highlighted in Fig. 3.3). From Fig. 3.3, it is demonstrated that the channel-wise approach is unable to recover multivariate modulated oscillation characteristics or any meaningful mutual information corresponding to the matching OCs with similar multiscale spectral content for visual object classes. On the other side, analyzing joint patterns with an appropriate selection of similar channel characteristics exhibits significance in enhancing multivariate analysis for cognitive visual objects. For instance, selecting channels (FC2, T8, FC1, T7), (FC2, T8, FC1), and (FC2, T8, FC1, T7, F4) show significant to enhance mutual patterns for the visual object classes HB, HF, and AB, respectively. The obtained significant multivariate patterns for 3-class imagined data are demonstrated in Fig. 3.3. Thus, we can infer that, due to the interdependent nature of the channels, raw MEG and EEG patterns may not be able to distinguish class-specific mutual information. In order to strongly correlate mutual multivariate information across channels, the MSSDM method is used to get channel-aligned multi-spectral features for the enhancement of multivariate analysis for visual objects.

### **3.3.2 Multivariate extension of swarm-sparse decomposition**

The proposed MSSDM is a novel extension version of the SSDM method to decompose multichannel nonstationary signals into channel-aligned common OCs across all channels. The proposed MSSDM delivers a set of common OCs that exhibits minimum collective bandwidth while fully reconstructing all input channels. With the exceptional ability of noise suppression and adaptive decomposition, the proposed MSSDM method leads to desirable properties such as prominent superiority in mode mixing, mode alignment, and avoidance of extraneous mode selection [63, 64]. In MSSDM, we extend the optimiza-

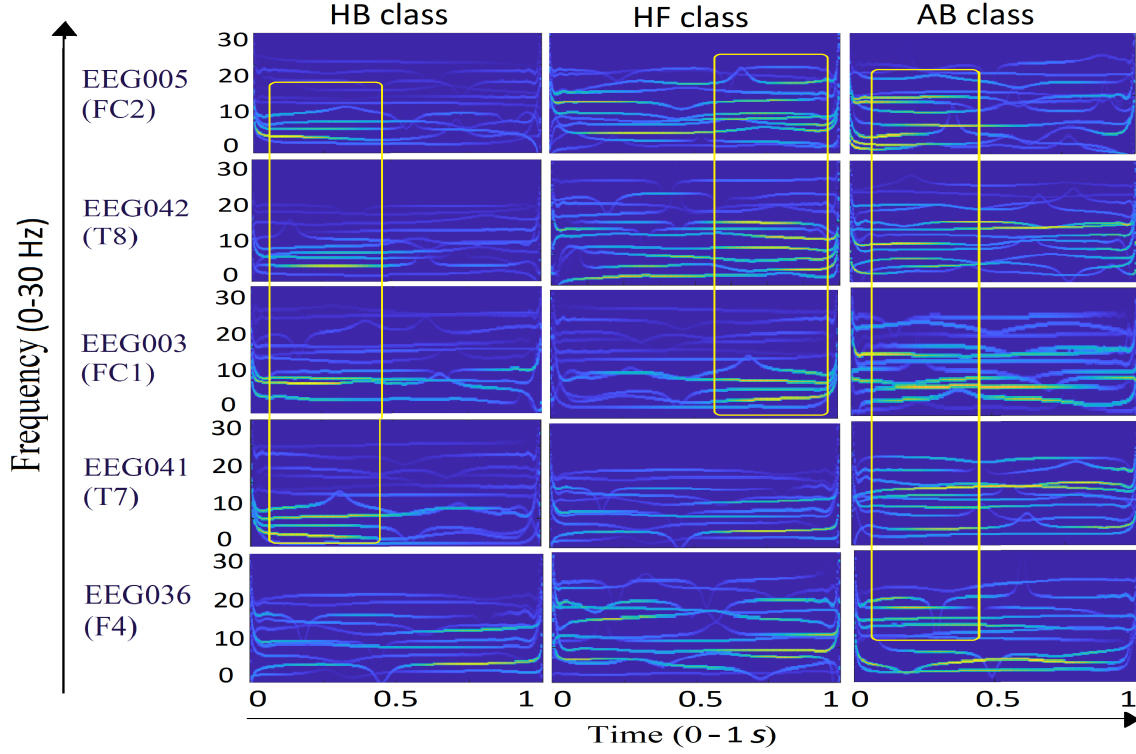


Figure 3.3: Channel-specific multivariate analysis from raw EEG channels [EEG005 (FC2), EEG042 (T8), EEG003 (FC1), EEG041 (T7), and EEG036 (F4)] for cognitive visual objects. The highlighted yellow boxes in the plots demonstrate mutual patterns of the imagined stimulus responses.

tion problem directly in the Lasso-based SFT domain through a multivariate swarm filtering optimization approach and Bhattacharya distance [107] based on convergence criteria. The proposed MSSDM method is efficient because it can generate multivariate modulated oscillations from input data without requiring additional user-defined parameters, such as threshold selection, filter order, or the number of iterations.

The mathematical formulation of the proposed multivariate SSDM model is presented in the following section.

### 3.3.2.1 Multivariate extension of SSDM

In order to apply the concept of SWD to multichannel signals, it is necessary to extract the common OC mode for each channel, which locates the common spectral information of the analyzed signal. In MSSDM, the OCs are extracted by selecting the common frequency band from multichannel signals which possess the highest amplitude peak through the ESD.

Therefore, the prime task of MSSDM is to extract matched OCs from multichannel signals so that the obtained JIF at each oscillatory level possesses common centre frequencies with aligned compact bandwidth. This extension of SSDM delivers the most common multivariate modulated oscillation in the multichannel data due to the following reasons: i) MSSDM extracts common modes from multichannel data to accurately extract the original signal with the least common minimum bandwidth. ii) Contrary to the univariate SWD approach, the MSSDM adopts the mode-alignment property to obtain multivariate modulated oscillations, which correspond to the matching of modes with similar frequency content across multiple channels. Due to the mode alignment problem in the univariate SWD filtering approach, it adopts channel-specific mode reconstruction with univariate modulated oscillation which is invariant to the other channels and introduces difficulties in the multivariate analysis. Therefore to extend the multivariate analysis to find matched oscillatory signals, in MSSDM, we have adopted common boundary estimation in the obtained SFT spectrum by computing the following mean spectrum magnitude  $\hat{\psi}(s)$  of multivariate signals from  $N$  number of channels:

$$\hat{\psi}(s) = \frac{1}{N} \sum_{n=1}^N (\|z_n - F^{-1} [\mathcal{L}F [z_n]]\|_1 + \mu_n \|z_n\|_1) \quad (3.1)$$

Where  $z_n$  represents the FT spectrum of the individual channel.  $F$  and  $F^{-1}$  represent the forward FT and IFT.  $\mathcal{L}$  is the  $L_1$ -norm of vector for FT kernel. Then adaptive SSDM-based filter banks are formed by applying the boundary estimation technique to the obtained mean spectrum. Further, the obtained common filter bank is applied to all channel data to get aligned modes. These OCs are the narrow-band components that deliver common frequency in every oscillatory level. In this work, the convergence criteria (defined in Eq. 2.16) is modified in the SSDM to design an improved filter bank. The modified Bhattacharya distance ( $B$ ) [107] based convergence criteria for the obtained residual signal is given as follows:

$$B = -\ln \sum_{k=1}^d \sqrt{(|y_{\delta,M}[k]| \cdot |z[k]|)^2} < 0.1 \quad (3.2)$$

Wherein  $|y_{\delta,M}[k]|$  and  $|z[k]|$  are the amplitude of FT of the output  $|y_{\delta,M}(k)|$  and input signal  $z[k]$ .  $d$  is the number of extracted OCs.

The MSSDM-based decomposition result of MEG 4-channel (MEG0111-0114) multivariate signals corresponding to the AF visual object category is shown in Fig. 3.5. Through the obtained common decomposed modes OC1-OC5, the proposed MSSDM method has demonstrated its ability to detect common or joint OC modes which having the same spectral content across the multiple channel data. The spectral analysis using Welch power spectral density (PSD), as shown in Fig. 3.4, illustrated the mode alignment of the selected OCs with MSSDM filter banks. A similar spectral response found in the PSD plot clearly demonstrates that OCs have consistently detected oscillations over multiple channels and are properly aligned in cases of MEG signals. In contrast to multivariate decomposition, a univariate decomposition will not align similar OC modes in the same numbers when decomposing the multichannel signal separately [63].

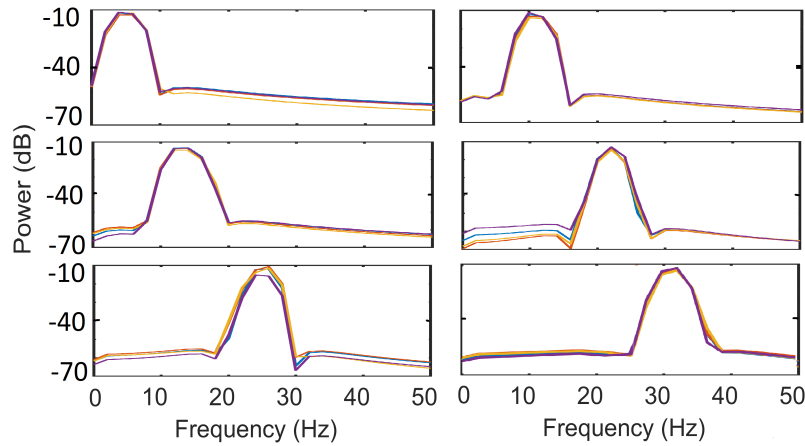


Figure 3.4: Filter bank structure for MSSDM-based swarm filter bank of 4-channel EEG and MEG signals with selected OCs.

The proposed MSSDM is detailed in Algorithm 3.1.

### 3.3.2.2 Ablation study of the SSDM algorithm

To verify mode alignment and denoising ability of MSSDM, we have conducted an ablation experiment to examine how the various elements of our proposed MSSDM approach affect the multivariate analysis of nonstationary signals (EEG and MEG). To demonstrate

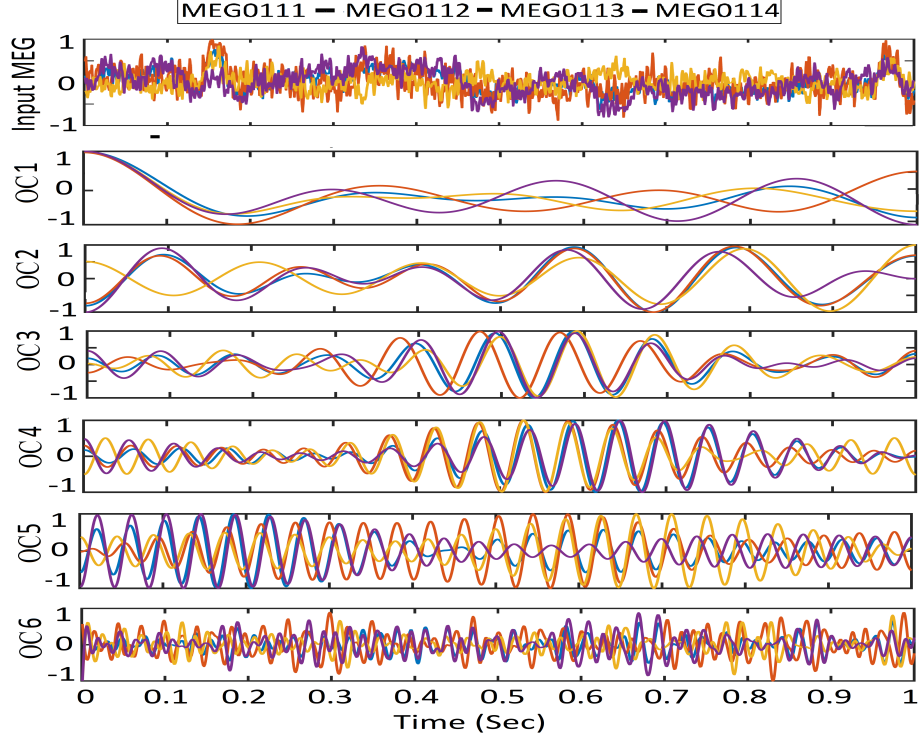


Figure 3.5: MSSDM decomposition of 4-channel MEG signal and selected OCs (OC1, OC2, OC3, OC4, OC5, and OC6).

the significance of the elements of the proposed MSSDM, we have assessed the four approaches: (i) FT spectrum without SGF, (ii) FT spectrum with SGF, (iii) SFT spectrum without SGF, and (iv) SFT spectrum with SGF. In this proposed ablation paradigm, the synthetic signal  $x(n)$  has been designed for 3-channel multivariate signals. Each multivariate channel signal consists of three nonstationary AM monocomponent signals of three sinusoids with different frequencies and three linear variable amplitudes, which is constructed as,  $x(n) = [s_{c1}(n); s_{c2}(n); s_{c3}(n)]$ . Here,  $s_{c1}(n)$ ,  $s_{c2}(n)$ , and  $s_{c3}(n)$  represent the three channels and they are designed as,  $s_{c1}(n) = s_1(n) + s_2(n) + s_3(n) + w(n) = 1.2 \cos(2\pi f_1(1 + \beta n)n + \pi/2) + 2 \cos(2\pi f_2(1 + \beta n)n + \pi/4) + 1.4 \cos(2\pi f_3(1 + \beta n)n) + w(n)$ ,  $s_{c2}(n) = s_2(n) + s_3(n) + s_4(n) + w(n) = 2 \cos(2\pi f_2(1 + \beta n)n + \pi/4) + 1.4 \cos(2\pi f_3(1 + \beta n)n) + 3 \cos(2\pi f_4(1 + \beta n)n + \pi) + w(n)$ , and  $s_{c3}(n) = s_2(n) + s_3(n) + s_5(n) + w(n) = 2 \cos(2\pi f_2(1 + \beta n)n + \pi/4) + 1.4 \cos(2\pi f_3(1 + \beta n)n) + 3 \cos(2\pi f_5(1 + \beta n)n + \pi/2) + 1.8 \cos(\omega(1 + \beta n)n + \pi/2) + w(n)$ . To inherit the complexity of EEG signal characteristics, the test synthetic signal is designed with a selec-

---

**Algorithm 3.1** MSSDM iterative filtering for multichannel data

---

Input: Nonstationary multichannel or multivariate signal  $z(k)$   
Output: OCs  $y_{it}(k)$  after MSSDM decomposition method  
Initialization:  $P_{th} = 0.1$ ;  $it = 0$ ;  $M = 20$ ;  $\delta = 0.01$ ;  $c = 0$ ;  $j = 0$   
**Repeat** Smooth the energy spectrum of  $z_{it}(k)$   
Compute SFT spectrum's extremes for all channels  $N_c$  by Eq. (2.4)  
Find the mean spectrum by Eq. (3.1)  
Estimate common boundary detection in spectrums by Eq. (2.7)  
 $it \leftarrow it + 1$   
**Until**  $N_{it} < N_c$  ;  $N_c$  represent number of channels.  
**Repeat**  
**Repeat**  
Smooth the energy spectrum of  $z_{it}(k)$  using SGF filter by Eq. (2.6)  
Estimate the spectral coefficient with the highest energy peak  
Calculate swarm parameters  $M$  and  $\delta$  by Eqs. (2.9a) and (2.9b), respectively  
**Repeat**  
 $y_j \leftarrow \text{SwF}(z_j(k), M, \delta)$  ; SwF: Swarm filter  
 $j \leftarrow j + 1$   
 $B_{it+1}(k) \leftarrow B_{it}(k)$   
**Until**  $B_{it} < 0.1$   
 $y_{it}(k) \leftarrow y_j(k)$   
 $z_{it+1}(k) \leftarrow z_{it}(k) - y_{it}(k)$   
 $it \leftarrow it + 1$ ,  $j = 0$   
 $y(k) \leftarrow z_{it+1}(k)$   
**Until**  $E_{\psi(z_{it+1})}(\omega) > P_{th}$   
 $it \leftarrow it + 1$   
**Until**  $N_{it} < N_c$   
 $y_{it}(k) \leftarrow y(k)$

---

tion of sinusoids in the frequency range [1-60 Hz]. The parameters are  $f_1 = 10$  Hz,  $f_2 = 20$  Hz,  $f_3 = 30$  Hz,  $f_4 = 40$  Hz,  $f_5 = 50$  Hz, and sampling frequency  $F_s$  is 1 kHz. The signal-to-noise ratio of white Gaussian noise  $\omega(n)$  is -10 dB.  $\omega$  is discrete carrier angular frequency and  $a$  is the modulation index.

In the ablation study, these approaches used to design the proposed MSSDM have been evaluated using spectrum analysis, decomposition results, and reconstruction performance, which are shown in Fig. 3.7, Fig. 3.6, and Table 3.1, respectively. Signal reconstruction efficiency is evaluated using the performance metric, error to signal ratio (ESR). From spectrum analysis mentioned in Fig. 3.7 (e), it shows that, when compared to the other approach versions, the proposed MSSDM, which combines SFT spectrum and SGF filter, performs the improved frequency resolution and optimally detects boundary frequencies in the sparse domain. From Figs. 3.7 (b), 3.7 (c), and 3.7 (d), it shows that the spectrum obtained is overlapped and the obtained spectral boundaries are non-distinguishable, which introduces

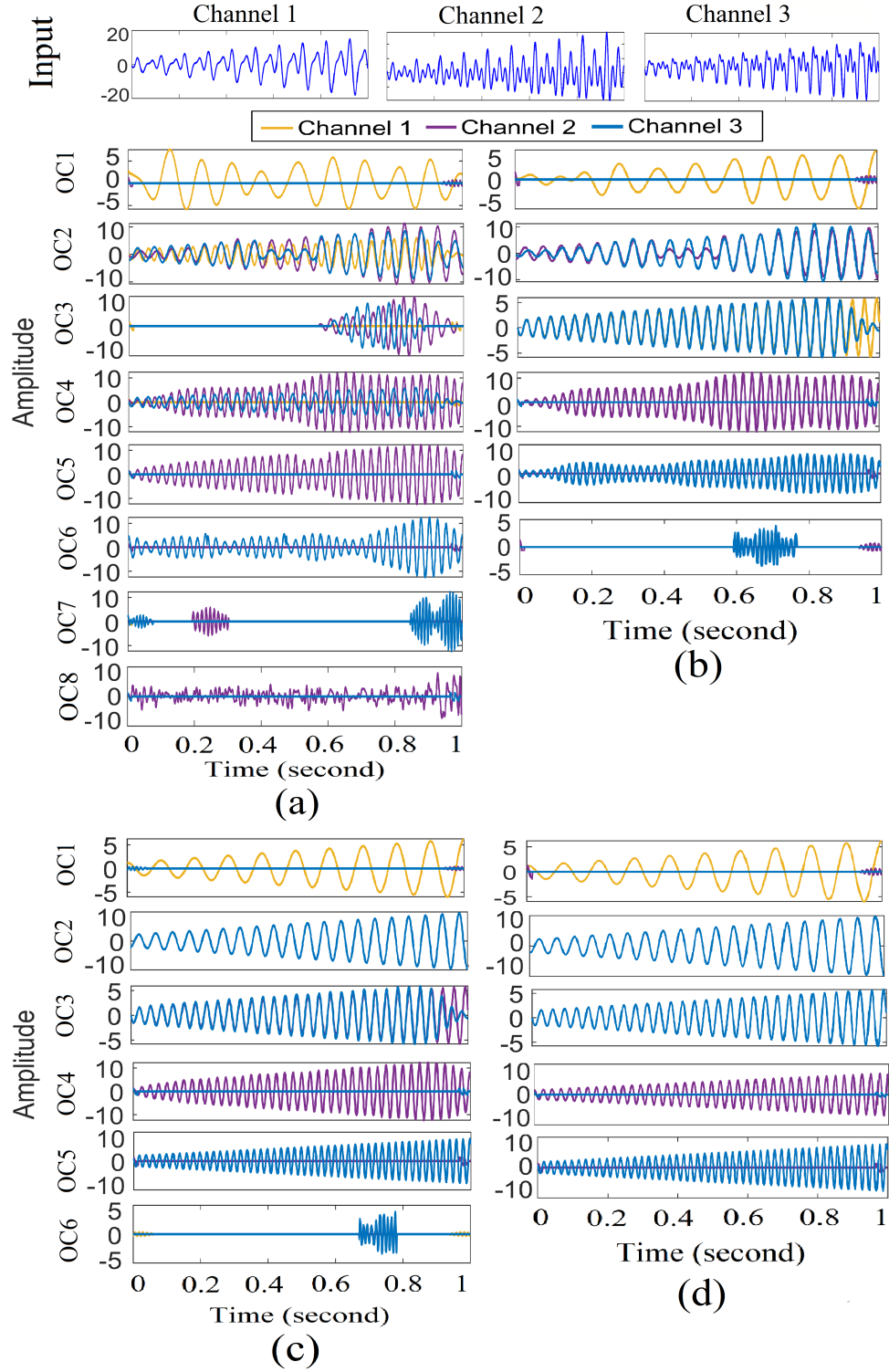


Figure 3.6: Decomposition of multivariate of 3-channel synthetic signal with similar frequency modes using MSSDM method in (a) FFT without SGF, (b) FFT with SGF, (c) SFT without SGF, and (d) SFT with SGF approaches.



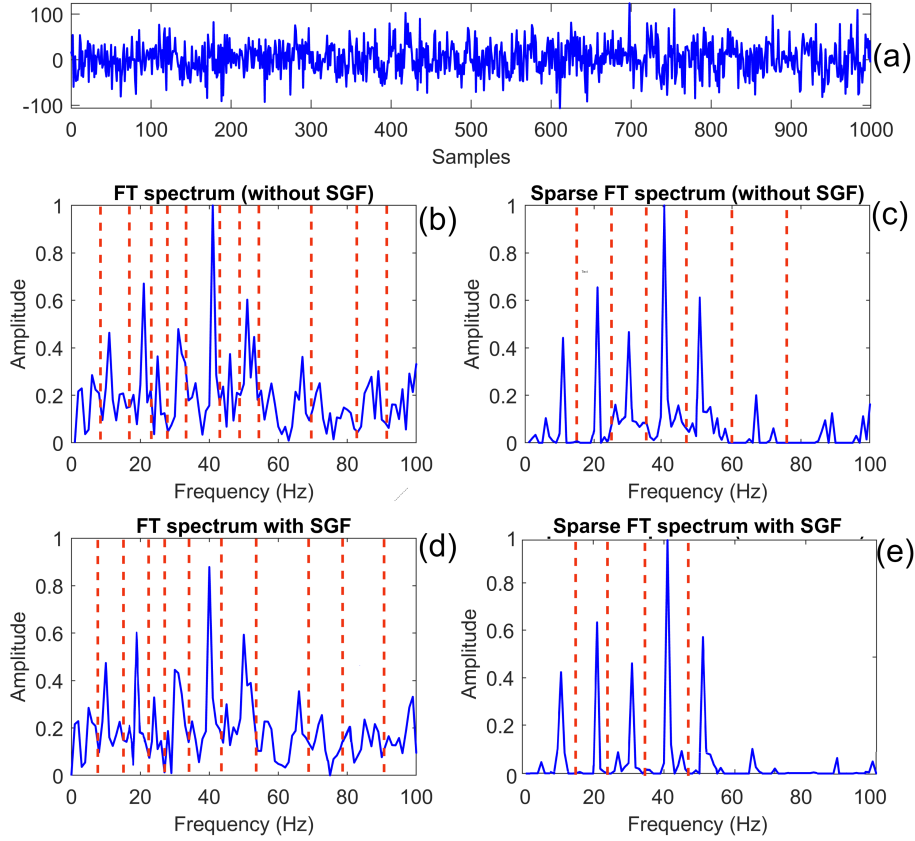


Figure 3.7: Plots of (a) nonstationary multicomponent AM signal ( $s_{c1}(n)$ ) and corresponding spectral analysis using (b) FFT without SGF, (c) FFT with SGF, (d) SFT without SGF, and (e) SFT with SGF approaches.

difficulty to reconstruct channel-aligned components. The performance of our approach suddenly declines when all two components (SFT and SGF) are taken away, demonstrating the significance of these parts, which is shown in Fig. 3.7 (b). Table 3.1 indicates the performance comparison of the obtained ESR values of  $x(n)$  for AM reconstructed monocomponent signals  $s_1(n)$ ,  $s_2(n)$ ,  $s_3(n)$ ,  $s_4(n)$ , and  $s_5(n)$  using FT (without SGF), FT (with SGF), SFT (without SGF), and SFT (with SGF) approaches. The lowest values of ESR of every reconstructed monocomponent signal using the MSSDM method (SFT and SGF approach) prove the efficient reconstruction of components as compared to other approaches. Whereas the reconstruction efficiency is poor in the case of other approaches as the value of ESR was found inferior (highest). It is also demonstrated in Fig. 3.6 indicating the decomposition performance of all approaches. It is quite easy to make an erroneous judgment that the extracted modes from the proposed MSSDM (SFT and SGF) are channel-aligned

in terms of their joint frequency content, which are shown in Fig. 3.6 (d). It is clear that the extracted modes OC2 and OC3 from channels 1, 2, and 3, are aligned with common time-varying frequency components, i.e. 20 Hz and 30 Hz which are locally present with single-mode OC2 and OC3, respectively. Rest extracted components are locally associated with respective modes of channels. The proposed method shows better ability to the mode mixing with the use of SFT and SGF and demonstrates all five converged components with five modes with the lowest ESR (shown in Fig. 3.6 (d)). Whereas other approaches show inferior decomposition performance and extract components with non-alignment mode and overlapping nature (shown in Figs. 3.6 (a), 3.6 (b), and 3.6 (c)). Overall, the above test result demonstrates the effectiveness of MSSDM in separating effective multivariate modulated oscillations from multichannel data while verifying mode alignment. The above descriptions highlight the relative significance of each element and their interdependence in designing the proposed MSSDM method for multivariate analysis to get mutual features across channels and attain the best classification performance for visual cognitive analysis.

Table 3.1: Comparison of the obtained ESR values of  $x(n)$  for AM reconstructed mono-component signals  $s_1(n)$ ,  $s_2(n)$ ,  $s_3(n)$ ,  $s_4(n)$ , and  $s_5(n)$  using FT (without SGF), FT (with SGF), SFT (without SGF), and SFT (with SGF) methods. Note: '-' indicates the blank ESR value due to a non-used signal in the respective channel.

Channel	Signal	FT (with SGF)	FT (without SGF)	SFT (with SGF)	SFT (without SGF)
$s_{c1}$	ESR ( $s_1(n)$ )	0.3491	0.0915	0.0082	0.0078
	ESR ( $s_2(n)$ )	0.2960	0.1653	0.0219	0.0062
	ESR ( $s_3(n)$ )	0.0964	0.0948	0.0052	0.0058
	ESR ( $s_4(n)$ )	—	—	—	—
	ESR ( $s_5(n)$ )	—	—	—	—
$s_{c2}$	ESR ( $s_1(n)$ )	—	—	—	—
	ESR ( $s_2(n)$ )	0.1480	0.0931	0.0728	0.0058
	ESR ( $s_3(n)$ )	0.3150	0.1148	0.0254	0.0097
	ESR ( $s_4(n)$ )	0.1840	0.2910	0.1926	0.0284
	ESR ( $s_5(n)$ )	—	—	—	—
$s_{c3}$	ESR ( $s_1(n)$ )	—	—	—	—
	ESR ( $s_2(n)$ )	0.1548	0.1820	0.0352	0.0045
	ESR ( $s_3(n)$ )	0.0904	0.1285	0.0125	0.0097
	ESR ( $s_4(n)$ )	—	—	—	—
	ESR ( $s_5(n)$ )	0.3071	0.4028	0.0125	0.0096

### 3.3.3 Feature extraction and selection

In literature, several features are reported for the classification of visual objects [3, 14, 108]. Here, we have adopted multiple domain feature parameters: RE [109], CSD [110], SE, and CSPTE from the  $\alpha$ - $\beta$  rhythms of 1 s multichannel EEG-MEG epochs. These selected features proved a strong exploration of spectral variability, amplitude, and complexity of real signals [109, 110, 111]. To enhance feature selection, a novel fusion feature approach is presented based on Riemann's correlation [112] by finding the best-correlated features. Among the employed features, the two newly proposed feature extraction techniques are explained as follows:

#### 3.3.3.1 Common spatial filter on Teager energy

In this feature, we have employed a multi-class CSPTE which is computed from MSSDM modes  $y_c^N$  of total sample  $N$  with  $c^{\text{th}}$  class. It is common to explore EEG signals using CSP features, but CSP with Teager energy (TE) is a new feature approach to solving multi-class visual categorization problems [111]. In our work, EEG-MEG variance in a given beta frequency band can be determined by its TE. Several multi-class extensions of the CSP technique are presented which include the pair-wise schemes (PW), one-versus-rest scheme (OVR), intra-classifier specific CSP, and divide-and-conquer scheme [111]. Here we have adopted the OVR approach using a binary CSP algorithm for multi-class problems. Using the eigenvalue decomposition-based optimization method, the binary CSP algorithm diagonalizes the TE vector values from the two covariance matrices simultaneously. A multi-class CSP algorithm with a one-versus-rest scheme is used to compute the TE that discriminates one class from the rest, then concatenate the features to form a feature vector  $F_i = [f_i^1 f_i^2 f_i^3 \dots f_i^c]^T$ . The extracted CSPTE feature ( $f_i$ ) is defined as,

$$f_i = \log \left( \frac{\text{var}(\omega^T y_c^n)}{\sum_{n=1}^{2m} \text{var}(\omega^T y_c^n)} \right) \quad (3.3)$$

Where  $f_i$  is the extracted binary class CSPTE feature in the selected band after the projection of spatial filters  $\omega$  [113].  $Y_c = \{y_c^1, y_c^2, y_c^3, \dots, y_c^n, \dots, y_c^N\}$  are obtained  $n^{\text{th}}$  rhythm-based OCs of class ( $c = 1, 2, 3, 4, 5$ ) of total sample  $N$ .  $T$  indicates matrix transpose. In our work,

pairing selection parameter  $m = 2$  is used to formulate a significant feature subset of the spatial filter.

### 3.3.3.2 Correntropy spectral density

CSD is a nonlinear correlation function capable of extracting multi-spectral domain characteristics of nonstationary signals, which may be contaminated by noise [110]. The feature  $\tilde{F}(\omega)$  is computed as,

$$\tilde{F}(\omega) = \sum_{q=-(f-1)}^{f-1} P_c(q) e^{-j\omega q} \geq 0 \quad (3.4)$$

where,

$$P_c(q) = E [k(y_i(n), y_{i+1}(n - q))] \quad -(f - 1) \leq q \leq f - 1$$

$P_c$  represents the cross-correntropy function of two computed rhythm-based OCs  $y_i$  of size  $n$  of  $i$ th channel.  $E$  is the expectation,  $m$  is the time-lag between spectral signal,  $f$  is sample length, and  $k(\cdot)$  is Gaussian kernel. The advantage of the CSD function is to improve the spectral resolution of higher-order statistical moments present in multivariate signals with few design parameters, thus potentially separating the highly correlated channel feature data.

### 3.3.3.3 Rényi's entropy

RE explores the mutual spectral variability information across two obtained rhythm-based OCs  $y_i$  of  $i$ th channel. It is given as follows [109]:

$$H_\infty(y_i|y_{i+1}) \stackrel{def}{=} -\log \sum_n \max((y_{i+1}|y_i)p(y_i)) \quad (3.5)$$

### 3.3.3.4 Sparse entropy

The SE feature is introduced to enhance the ability of the entropy feature to extract joint significant features using a sparse filter bank and an adaptive windowing tech-

nique. Essentially, it is a multi-task filtering model that combines SFT spectrum optimization with Lasso penalties ( $\|U\|_1$ ,  $\|U\|_2$ , and  $\|u_t - u_{t+1}\|_1$ ) and entropy feature vectors  $H^{(t)} = -\sum_{i=1}^N P(y_i) \cdot \log_2(P(y_i))$  computed from the obtained rhythm-based OCs  $y_i$  of  $i^{\text{th}}$  channel to enhance the temporal characteristics of the features. The multi-class feature set ( $F_i$ ) is formulated as follows [114]:

$$F_i = \arg \min_F \frac{1}{2} \sum_{t=1}^T \|H^{(y_i)} u_t - v\|_2^2 + \beta_1 \|U\|_1 + \beta_2 \|U\|_2 + \beta_3 \sum_{t=1}^T \|u_t - u_{t+1}\|_1 \quad (3.6)$$

Where penalty parameters ( $\beta_1$ ,  $\beta_2$ , and  $\beta_3$ ) are set at 0.1.  $u_t$  denotes the learned projection vector at the  $T^{\text{th}}$  sliding window.

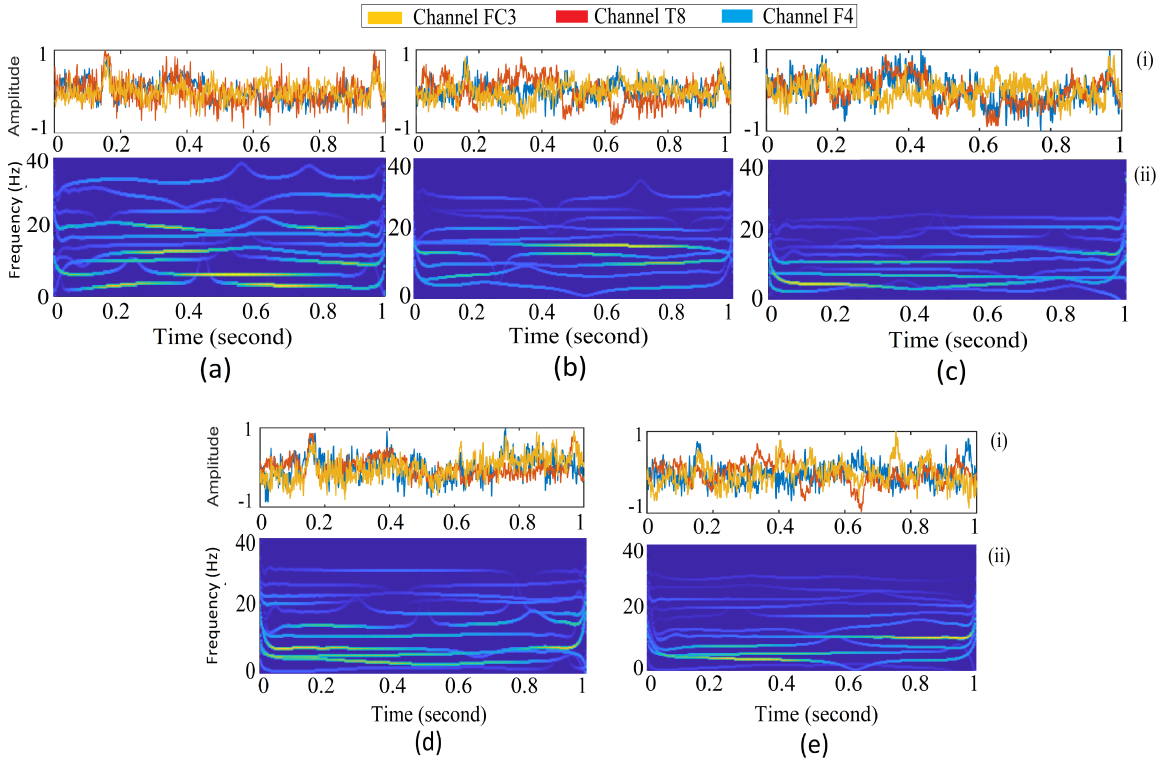


Figure 3.8: Plots of (i) raw multivariate MEG signals and (ii) their multivariate TFRs of the  $\alpha$ - $\beta$  rhythms from three selected channels (FC3, T8, and F4 of subject 1) using proposed MSSDM method for visual object classes (a) HB, (b) AB, (c) HF, (d) AF, and (e) IO, respectively.

### 3.4 Results and discussion

To prove the effectiveness of the proposed MSSDM-based visual object recognition framework, we have tested SSDM-based features from the EEG signals (74-channel), MEG signals (74-channel), MEG signals (308-channel), and combined EEG-MEG signals (380-channel) of the visual stimulus of the MEGEEG92 Objects Dataset [14], which is given at [https://figshare.com/collections/MEGEEG92\\_Objects\\_Dataset/4182587/1](https://figshare.com/collections/MEGEEG92_Objects_Dataset/4182587/1). The five different visual object classes namely, HB, HF, AB, AF, and IO from the EEG-MEG dataset are classified using LDA and validated using performance metrics such as ACC, SEN, SPE, and F1-score. LDA classifier has been chosen based on its proven capability to investigate features with appropriate parameters and counteract over-fitting problems with less computational complexity in VSA decoding applications [5]. To prove its potential against the existing methods [3, 4, 5, 14], the proposed method has been tested for its mode-alignment property by aligning common frequency scales across multiple EEG-MEG data. Also, we have compared the proposed method with USSDM approach (channel-specific analysis), and direct rhythms (delta, theta, alpha, beta, and gamma) analysis which is computed using a BPF [115] from raw EEG signals. In our work, the 1 s epoch of multichannel data is decomposed using MSSDM into different OCs. Further, rhythms: Delta ( $\delta$ : 0.1–4 Hz), theta ( $\theta$ : 4–8 Hz), alpha ( $\alpha$ : 8–13 Hz), beta ( $\beta$ : 13–30 Hz), gamma ( $\gamma$ : 30–80 Hz), and combined alpha and beta ( $\alpha$ - $\beta$ : 8–30 Hz), are computed on mean frequency from the decomposed MSSDM modes. In our work, features are extracted from the combined  $\alpha$ - $\beta$  rhythms.  $\alpha$ - $\beta$  rhythm selection depends exclusively on experimentation and delivers the highest performance. Figure 3.8 shows the multivariate TFRs of the  $\alpha$ - $\beta$  rhythms from three selected channels (FC3, T8, and F4 of subject 1) using the proposed MSSDM method for the 5-class visual object MEG data. It exhibited discriminative temporal and spectral characteristics in the obtained  $\alpha$ - $\beta$  rhythms related to the different visual object classes. In order to extract the most discriminative features from decomposed OCs, a novel feature scheme has been implemented in which fusion features are computed by finding the most correlated features from the normalized features (RE, SE, CSPTE, and CSD) using Riemann’s correlation. The objective is to formulate the most significant VSA features set by separating and eliminating non-discriminative features in subject-specific

# CHAPTER 3. MULTIVARIATE SSDM FOR MULTIVARIATE SIGNAL ANALYSIS WITH COGNITIVE VISUAL OBJECT DETECTION FROM MULTICHANNEL EEG-MEG SIGNALS

without compromising classification performance.

Table 3.2: Classification accuracy (in %) of combined ( $\alpha$ - $\beta$ ) rhythms-based fusion features computed from extracted OCs using MSSDM, USSDM, and BPF with Riemann's correlation analysis.  $Q$  denotes the correlation coefficient value. The best accuracy is marked in boldface.

Subject	MSSDM-FF-LDA (74-channel EEG)						MSSDM-FF-LDA (74-channel MEG)						MSSDM-FF-LDA (306-channel MEG)						MSSDM-FF-LDA (380-channel EEG-MEG)					
Subject	(74-channel EEG)						(74-channel MEG)						(306-channel MEG)						(380-channel EEG-MEG)					
$Q \rightarrow$	> 0.4	> 0.5	> 0.6	> 0.7	> 0.8	> 0.9	> 0.4	> 0.5	> 0.6	> 0.7	> 0.8	> 0.9	> 0.4	> 0.5	> 0.6	> 0.7	> 0.8	> 0.9	> 0.4	> 0.5	> 0.6	> 0.7	> 0.8	> 0.9
Subject 1	61.84	63.35	63.98	<b>64.56</b>	63.84	65.41	67.35	70.00	74.02	<b>75.75</b>	72.77	67.98	71.95	78.73	82.05	<b>88.34</b>	77.73	72.59	76.68	79.69	83.49	<b>90.25</b>	82.85	77.39
Subject 2	69.78	72.55	73.24	<b>74.92</b>	71.78	70.27	72.55	75.41	76.76	<b>84.12</b>	78.17	73.04	77.43	80.50	85.17	<b>89.13</b>	83.48	77.98	82.60	85.86	89.95	<b>92.81</b>	89.00	83.16
Subject 3	68.03	70.69	71.33	<b>72.99</b>	70.03	68.57	70.69	73.45	77.70	<b>79.47</b>	76.28	71.25	77.82	78.44	82.96	81.25	<b>83.65</b>	76.09	81.42	83.63	87.63	86.61	<b>89.20</b>	81.12
Subject 4	63.25	65.70	66.29	<b>72.87</b>	65.25	63.88	65.70	68.26	72.18	<b>73.86</b>	71.06	66.38	72.35	72.90	77.09	<b>84.12</b>	75.88	70.88	75.67	77.72	81.41	<b>84.09</b>	80.90	75.57
Subject 5	66.96	69.60	<b>75.38</b>	74.89	68.96	67.54	69.60	72.33	<b>78.87</b>	78.28	75.13	70.16	76.60	77.23	84.18	89.98	80.20	74.94	80.16	82.35	86.42	<b>89.13</b>	85.53	79.88
Subject 6	62.53	64.95	65.57	<b>72.07</b>	64.53	63.15	64.95	67.52	71.40	<b>80.04</b>	70.25	65.60	71.53	72.07	76.26	<b>82.01</b>	75.04	70.07	74.81	76.88	<b>81.56</b>	80.12	79.98	74.69
Subject 7	69.37	72.07	71.28	<b>72.89</b>	71.37	69.87	72.07	74.88	77.60	<b>82.38</b>	77.72	72.60	79.35	79.97	82.89	<b>90.02</b>	83.00	77.52	83.01	85.25	85.03	<b>90.24</b>	81.23	82.65
Subject 8	62.11	64.56	63.89	<b>67.25</b>	64.11	65.96	64.56	67.12	69.55	<b>76.27</b>	69.81	68.56	71.05	71.64	74.30	<b>78.21</b>	74.55	73.19	74.36	76.41	75.48	<b>83.42</b>	79.48	78.05
Subject 9	64.08	66.56	<b>68.54</b>	66.83	66.08	67.96	66.56	69.18	74.64	<b>76.78</b>	71.93	70.60	73.23	75.24	<b>81.20</b>	79.18	76.84	75.41	75.13	80.24	<b>82.51</b>	80.12	81.90	80.38
Subject 10	64.78	67.32	66.61	<b>76.12</b>	66.78	68.70	63.43	65.93	72.52	<b>80.18</b>	72.70	71.37	69.76	71.71	78.91	<b>85.71</b>	77.66	76.22	71.59	76.48	80.17	<b>83.04</b>	82.78	81.26
Subject 11	70.12	72.83	<b>73.11</b>	70.47	72.12	74.22	68.62	71.28	79.63	<b>83.85</b>	78.55	77.12	75.51	77.57	85.25	<b>88.91</b>	83.87	82.35	77.46	82.68	81.56	87.96	<b>89.43</b>	87.81
Subject 12	62.19	64.60	64.90	67.52	<b>68.19</b>	64.21	60.87	63.28	70.67	82.54	<b>74.30</b>	66.70	66.97	68.82	76.89	<b>89.99</b>	79.31	71.25	68.71	73.39	82.13	<b>87.29</b>	84.59	75.94
Subject 13	64.27	66.78	67.05	<b>76.61</b>	66.27	68.17	62.92	65.38	73.03	<b>76.91</b>	72.14	70.82	69.22	71.13	82.44	<b>89.66</b>	77.07	75.64	71.02	75.83	78.43	<b>89.21</b>	82.14	80.64
Subject 14	64.21	66.75	67.03	<b>68.57</b>	66.21	64.83	62.90	65.35	73.00	<b>74.67</b>	72.11	67.33	69.15	71.10	79.41	<b>81.24</b>	77.00	71.94	70.99	75.80	81.98	80.25	<b>82.11</b>	76.66
Subject 15	60.56	62.90	<b>66.99</b>	65.82	62.56	61.25	59.27	61.57	<b>72.98</b>	72.69	68.13	63.62	65.22	67.00	72.18	<b>84.98</b>	72.76	67.96	66.89	71.41	81.28	<b>87.15</b>	77.57	72.43
Subject 16	62.78	65.20	65.76	64.28	64.78	63.41	61.43	63.80	71.62	<b>73.25</b>	70.54	65.86	66.87	69.45	79.90	<b>86.71</b>	75.34	70.36	68.58	74.00	84.92	<b>90.97</b>	80.31	74.99
Average accuracy	64.99	67.53	68.25	<b>70.54</b>	67.24	66.71	65.84	68.42	74.32	<b>78.25</b>	73.22	69.61	72.13	73.72	79.26	<b>86.05</b>	78.34	74.02	74.94	78.60	81.37	<b>86.42</b>	83.06	78.91
Accuracy (10-fold CV)	64.19	66.98	68.93	<b>68.36</b>	66.41	60.71	66.50	67.05	70.25	<b>77.58</b>	71.54	67.92	71.24	67.08	74.56	<b>84.36</b>	77.37	67.36	75.69	77.03	78.95	<b>84.98</b>	82.54	72.81
Using USSDM approach																								
Subject 1	56.80	58.55	59.07	60.46	58.06	56.86	62.54	65.06	67.41	68.9	<b>69.75</b>	63.16	67.54	70.20	74.24	<b>75.98</b>	72.98	68.17	71.74	74.61	<b>78.13</b>	75.01	71.99	67.28
Subject 2	68.11	67.43	68.77	65.87	<b>68.98</b>	65.31	67.45	70.10	74.10	<b>75.81</b>	72.62	67.89	72.76	75.64	79.99	<b>81.83</b>	78.40	73.26	77.35	80.39	84.19	<b>86.32</b>	77.36	72.31
Subject 3	66.40	65.68	67.00	<b>67.79</b>	65.07	63.71	65.68	68.25	72.20	<b>73.83</b>	70.88	66.19	73.09	73.67	77.92	76.29	<b>78.57</b>	71.46	76.20	78.28	<b>82.02</b>	80.46	77.54	70.51
Subject 4	55.41	61.03	62.23	<b>66.89</b>	60.62	59.35	61.02	63.41	67.05	<b>68.65</b>	66.01	61.68	67.93	68.46	72.39	<b>74.08</b>	71.26	66.57	70.80	72.73	76.18	<b>78.16</b>	70.32	65.69
Subject 5	58.66	64.68	<b>68.01</b>	66.78	64.09	60.71	64.68	67.22	<b>73.32</b>	72.82	69.83	65.17	71.96	72.55	79.10	<b>78.51</b>	75.35	70.37	75.04	77.09	80.92	<b>77.02</b>	74.38	69.42
Subject 6	54.78	60.38	61.57	<b>62.31</b>	59.92	56.76	60.33	62.77	66.33	<b>67.86</b>	65.24	60.92	67.16	67.72	71.61	<b>73.26</b>	70.45	65.79	69.99	71.98	<b>76.32</b>	74.44	69.50	64.90
Subject 7	60.77	66.95	66.90	65.78	<b>68.54</b>	62.81	66.95	69.56	72.07	65.00	<b>72.20</b>	67.45	74.52	75.10	77.82	<b>77.72</b>	<b>78.94</b>	72.81	77.68	79.78	79.56	<b>81.98</b>	70.59	71.84
Subject 8	54.41	60.01	59.97	<b>62.51</b>	59.55	59.32	60.01	62.39	64.60	<b>68.1</b>	64.85	63.72	66.76	67.31	69.75	<b>73.48</b>	70.01	68.76	69.62	71.55	70.61	<b>77.54</b>	69.08	67.86
Subject 9	56.13	61.85	64.35	<b>68.94</b>	61.36	61.08	61.83	64.28	<b>69.34</b>	67.63	66.80	65.57	68.76	70.68	76.26	<b>74.37</b>	72.14	70.80	70.30	75.11	<b>77.21</b>	74.45	71.15	69.84
Subject 10	56.75	62.57	62.53	<b>63.28</b>	62.02	61.75	58.93	61.28	67.36	<b>68.91</b>	67.53	66.29	65.52	67.37	74.09	<b>75.79</b>	72.91	71.57	67.00	71.60	73.01	<b>74.78</b>	71.93	70.62
Subject 11	68.44	67.64	67.93	64.19	<b>69.64</b>	66.73	63.72	66.19	74.00	<b>77.9</b>	72.98	71.66	70.89	72.82	75.38	<b>77.48</b>	<b>78.78</b>	77.35	72.46	77.35	76.34	<b>76.45</b>	77.74	76.33
Subject 12	60.70	60.05	60.29	<b>62.73</b>	63.39	59.63	56.54	58.81	65.66	68.33	<b>69.07</b>	61.95	62.88	64.65	72.21	<b>75.14</b>	74.52	66.89	64.29	68.71	73.11	<b>74.15</b>	73.57	65.98
Subject 13	62.73	62.04	62.30	<b>61.61</b>	61.54	63.32	58.44	60.74	67.85	<b>71.47</b>	67.00	65.79	64.99	66.80	74.61	<b>78.58</b>	72.35	71.03	66.45	70.97	73.40	<b>77.56</b>	71.37	70.08
Subject 14	56.25	62.02	62.27	<b>63.7</b>	61.52	60.20	58.47	60.72	67.82	<b>69.38</b>	67.01	62.52	64.98	66.77	74.58	<b>76.30</b>	72.33	67.53	61.74	65.89	70.12	<b>82.12</b>	71.38	66.60
Subject 15	53.05	58.42	<b>62.26</b>	61.16	58.12	56.88	55.03	57.18	<b>67.83</b>	66.62	63.31	59.08	61.22	62.90	67.81	<b>73.25</b>	68.33	63.80	58.12	62.05	66.72	<b>72.29</b>	67.43	62.93
Subject 16	55.00	60.54	61.10	<b>62.49</b>	60.17	58.89	57.04	59.24	<b>66.56</b>	68.04	65.53	61.16	58.08	60.32	67.73	<b>69.26</b>	65.47	61.13	59.57	64.29	72.00	<b>73.84</b>	69.80	65.15
Average accuracy	59.02	62.49	63.53	<b>64.16</b>	62.66	60.83	61.17	63.58	63.45	<b>69.95</b>	68.16	64.39	67.44	68.94	74.09	<b>75.71</b>	73.24	69.21	69.27	72.65	75.74	<b>77.29</b>	72.19	68.58
Accuracy (10-fold CV)	58.12	57.34	60.36	<b>63.02</b>	57.50	57.79	60.23	58.34	60.28	<b>66.10</b>	62.55	61.17	66.40	63.26	70.39	<b>74.54</b>	67.20	65.74	68.21	66.66	71.95	<b>75.69</b>	66.24	65.16
Using BPF approach																								
Subject 1	51.32	49.50	48.65	<b>50.12</b>	49.12	45.29	50.10	48.03	51.76	<b>59.32</b>	50.82	48.20	57.13	58.10	59.21	<b>62.01</b>	57.13	51.86	52.21	53.43	56.32	<b>60.78</b>	54.76	50.12
Subject 2	57.65	61.80	67.12	64.28	<b>67.32</b>	59.85	61.83	65.65	66.67	65.45	<b>66.78</b>	62.21	66.70	69.32	73.28	<b>74.97</b>	71.82	67.14	70.91	73.68	77.13	<b>79.09</b>	70.86	66.27
Subject 3	56.21	60.17	65.40	<b>66.15</b>	63.51	58.35	60.16	63.89	67.59	<b>69.09</b>	64.93	60.63	66.96	67.49	71.40	<b>69.87</b>	<b>71.99</b>	65.46	69.81	71.72	<b>75.16</b>	73.7	71.05	64.59
Subject 4	46.90	55.90	60.72	<b>61.52</b>	59.15	54.37	55.87	59.35	61.40	<b>64.28</b>	60.47	56.50	62.21	62.71	66.30	<b>67.89</b>	65.28	60.99	64.84	66.63	69.78	<b>71.63</b>	64.41	60.19
Subject 5	49.65	59.26	66.42	<b>65.16</b>	62.58	55.59	52.89	62.94	67.20	<b>68.06</b>	63.99	59.68	65.95	66.48	<b>72.51</b>	67.19	65.06	64.45	68.77	70.65	74.18	<b>75.86</b>	68.16	63.58
Subject 6	46.37	55.34	60.80	<b>62.90</b>	58.46	51.96	55.24	58.79	60.75	<b>63.52</b>	59.74	55.79	61.50	62.07	65.59	<b>67.11</b>	64.52							

based on different Riemann's correlation coefficient factors ( $Q$ ) from computed rhythms using MSSDM, USSDM, and BPF methods. Here FF features are formulated with a selection of different correlation values ( $>0.4$ ,  $>0.5$ ,  $>0.6$ ,  $>0.7$ ,  $>0.8$ , and  $>0.9$ ) and all remaining features are eliminated. From Table 3.2, it is shown that FF features with a correlation value ( $> 0.7$ ) demonstrated the best classification results for five visual object classes corresponding to four channel selection schemes: EEG (74), MEG (74), MEG (306), and EEG-MEG (380). In the study, a subject-specific and cross-subject 10-fold cross-validation scheme was conducted, along with a statistical t-test ( $p < 0.05$ ). It employed leave-one-out cross-validation in feature vectors from the cross-subject dataset considered in a training-testing configuration of 80%-20%. This process was repeated for all subjects and average results were reported. To test the computed FF features, different MSSDM-based classification frameworks have been developed such as the MSSDM method with LDA (MSSDM-FF-LDA) in with the USSDM method with LDA (USSDM-FF-LDA) and the BPF method with LDA (BPF-FF-LDA) in four channel selection schemes. Table 3.3 shows the feature-specific average classification performance, which is obtained from all 16 subjects using the MSSDM, USSDM, and BPF methods. Here classification has been carried out over the 10 times repetition, and average accuracies were reported for four-channel selection schemes of MEG and EEG sensor data.

As shown in Table 3.3, the proposed classification framework MSSDM-FF-LDA has achieved the highest average accuracy of 86.42% among all frameworks in the case of channel scheme of EEG-MEG (380-channel) with fusion feature (RE, CSPTE, SE, CSD). The obtained sensitivity, specificity, and F1-score are 87.08%, 83.11%, and 84.89%, respectively. The performance of MSSDM-SE-LDA slightly deteriorates in the SE feature when using combined EEG-MEG channels. We have achieved approximately the same average accuracy of 84.06%, sensitivity rate of 82.65%, specificity of 88.15%, and F1-score of 80.50%. In contrast, MSSDM-RE-LDA reported relatively low performance with 77.90% for the RE features with combined EEG-MEG channels, respectively. It is noted that the introduction of new features: SE, CSPTE feature, and FF with LDA, have delivered the highest average accuracy for visual object classification in the case of the EEG-MEG (380-channel) scheme. Fig. 3.9 presents an evaluation of classification accuracies by us-



**CHAPTER 3. MULTIVARIATE SSDM FOR MULTIVARIATE SIGNAL ANALYSIS WITH COGNITIVE VISUAL OBJECT DETECTION FROM MULTICHANNEL EEG-MEG SIGNALS**

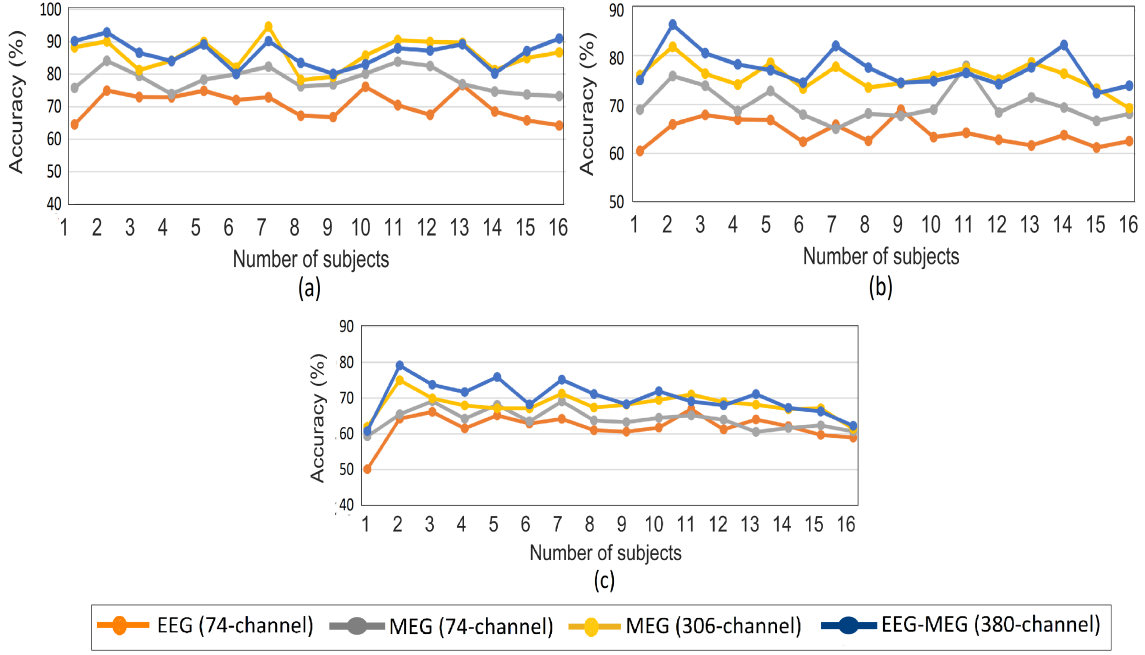


Figure 3.9: Plots of averaged classification accuracies of 5-class visual imagery for the 16 subjects from the  $(\alpha\text{-}\beta)$  rhythm-based FF features using (a) MSSDM, (b) USSDM, and (c) BPF approaches.

ing FF features obtained from the MSSDM, USSDM, and BPF in all 16 subject-specific cases. In our analysis, we have computed the FF features based on Riemann's correlation index ( $Q$ ) with a threshold value ( $Q > 0.7$ ) from features (RE, SE, CSPTE, and CSD). FF feature selection is intended to reduce the complexity of the feature while improving performance. In the subject-specific, the MSSDM-FF-LDA framework demonstrated the highest average accuracy (86.42%) within the EEG-MEG (380-channel) scheme. However, the USSDM-FF-LDA and BPF-FF-LDA classification frameworks achieved comparatively lower average accuracies of 77.29% and 69.96%, respectively. In 10-fold cross-validation, the MSSDM-FF-LDA framework achieved the highest performance with an accuracy of 84.98%.

In the case of MEG or EEG data, the proposed classification frameworks show slightly inferior performance than the combined EEG-MEG channel data. In MEG (306-channel) scheme, performance is slightly improved with the proposed MSSDM-FF-LDA comparatively and delivered accuracy, sensitivity, specificity, and F1-score of values 83.05%, 85.80%, 83.04%, and 83.69%, respectively. Whereas the obtained lowest accuracy rates for

USSDM-FF-LDA and BPF-FF-LDA classification frameworks are 75.71% and 68.34%, respectively. Also, the performance of the LDA classifier with new features (SE and CSPTE) in evaluating the proposed method is noteworthy as the obtained average accuracy rates are 80.98% and 81.97%, respectively. The sensitivity and specificity rates with the CSPTE feature are found higher for subjects 2, 11, 12, and 13 (only an average of 86.05% of VSA classes are truly classified), but it maintains a high average F1-score (82.21%) for all these subjects (shown in Table 3.3). For subjects (2-9,11-13), the LDA classifier with FF achieved very good sensitivity rates (more than 85.80%) with high specificity rates.

In similar cases, MEG (74-channel) and EEG (74-channel) schemes, also provided good sensitivity and specificity rates for most of the subjects. The highest average accuracy, sensitivity, specificity rates, and F1-score for SE feature with LDA (MSSDM-SE-LDA) for MEG (74-channel) scheme are 75.85%, 72.40%, 70.01%, and 69.49%, respectively. Whereas in the case of EEG (74-channel), the performance of features computed from 74-channel EEG signals falls significantly in evaluating the proposed method for all subjects, as the MSSDM-SE-LDA framework achieved the lowest classification performance with average accuracy, sensitivity, specificity rates, and F1-score for RE features are 69.65%, 66.38%, 65.91%, and 66.859%, respectively. However, it is clear from Table 3.3 that the sensitivity of the proposed method for EEG (74-channel) is comparatively less for all subjects. For those subjects, visual object detection is difficult to detect using the obtained MSSDM features because of non-discriminative overlapping multivariate modulated oscillations and highly contaminated artifacts with a short visual stimulation response. Also, overall classification performance using different feature schemes has been presented in Table 3.3 and Fig. 3.10.

In Table 3.4, we compare the proposed MSSDM method with existing methods for classifying visual objects using EEG signals and MEG signals already reported in the literature [3, 4, 5, 14]. This table illustrates the accuracy of classification across subjects and the statistical analysis based on the different experimental conditions of the proposed method. In our work, the comparison was considered on the databases [3, 4, 5, 14] with the same experimental conditions. For classifying visual objects, researchers have explored mostly different machine learning or deep learning methods to compute different discriminative features for raw EEG or MEG signals. The Classification method proposed by Cichy et al. [3]

### CHAPTER 3. MULTIVARIATE SSDM FOR MULTIVARIATE SIGNAL ANALYSIS WITH COGNITIVE VISUAL OBJECT DETECTION FROM MULTICHANNEL EEG-MEG SIGNALS

Table 3.3: Feature-specific average classification performance obtained from subject A01-16. The best accuracy is marked in boldface.

Method	Channel	CSD				CSPTE				RE				SE				FF			
		ACC (%)	SEN (%)	SEP (%)	F1-score	ACC (%)	SEN (%)	SEP (%)	F1-score	ACC (%)	SEN (%)	SEP (%)	F1-score	ACC (%)	SEN (%)	SEP (%)	F1-score	ACC (%)	SEN (%)	SEP (%)	F1-score
MSSDM	EEG (74)	63.24	63.42	58.43	57.28	64.17	62.19	61.13	58.68	67.07	68.19	64.67	66.12	<b>69.65</b>	66.38	65.91	66.85	<b>70.35</b>	74.03	72.29	71.76
	MEG (74)	69.31	77.01	68.85	68.03	<b>74.40</b>	75.42	73.56	71.14	71.31	69.28	65.72	70.04	<b>75.85</b>	72.40	70.01	69.49	<b>78.25</b>	79.06	77.87	80.28
	MEG (306)	72.49	79.10	56.40	76.57	<b>81.97</b>	82.53	83.88	82.21	72.56	79.72	74.12	70.29	<b>80.98</b>	78.16	75.97	78.34	<b>83.05</b>	85.80	83.04	83.69
	MEG-EEG (380)	79.52	91.21	86.81	83.85	<b>83.11</b>	52.34	40.02	41.50	77.90	79.36	74.06	74.59	<b>84.06</b>	82.65	88.15	80.50	<b>86.42</b>	87.08	83.11	84.89
USSDM	EEG (74)	59.34	55.09	56.08	57.28	54.43	55.63	55.87	51.89	54.43	57.15	55.40	58.43	<b>60.97</b>	63.49	60.51	59.92	<b>64.16</b>	65.12	62.46	63.01
	MEG (74)	60.88	58.19	57.46	55.27	64.01	68.62	60.82	61.79	<b>66.10</b>	64.82	62.15	64.87	65.70	69.45	61.71	63.72	<b>69.95</b>	68.95	66.39	68.31
	MEG (306)	64.70	67.43	65.32	64.92	<b>71.23</b>	68.71	69.32	68.30	70.51	75.51	70.06	67.57	<b>73.57</b>	72.51	68.68	70.99	<b>75.71</b>	73.12	72.37	74.21
	MEG-EEG (380)	70.85	93.75	45.41	43.12	<b>71.61</b>	61.16	57.47	49.45	69.24	54.27	42.83	44.41	<b>75.77</b>	70.12	67.47	69.45	<b>77.29</b>	73.68	79.89	76.84
BPF	EEG (74)	50.38	57.30	57.48	55.79	33.01	43.18	34.70	35.44	53.28	50.84	44.31	51.36	<b>52.36</b>	49.30	51.69	50.68	<b>61.92</b>	59.66	55.50	57.58
	MEG (74)	58.99	46.08	47.90	45.66	48.76	47.05	46.64	44.58	<b>64.14</b>	62.09	59.64	60.52	54.28	53.69	50.67	52.39	<b>64.23</b>	61.89	58.15	59.92
	MEG (306)	64.09	58.59	60.48	58.22	<b>65.92</b>	63.73	61.78	61.58	65.51	66.03	62.73	64.06	61.77	60.58	59.71	60.86	<b>68.34</b>	64.58	61.22	64.42
	MEG-EEG (380)	63.52	58.19	57.65	56.10	<b>64.93</b>	60.22	60.80	60.05	64.63	59.15	60.60	59.02	<b>67.63</b>	64.63	61.18	64.63	<b>69.96</b>	66.29	63.89	69.66

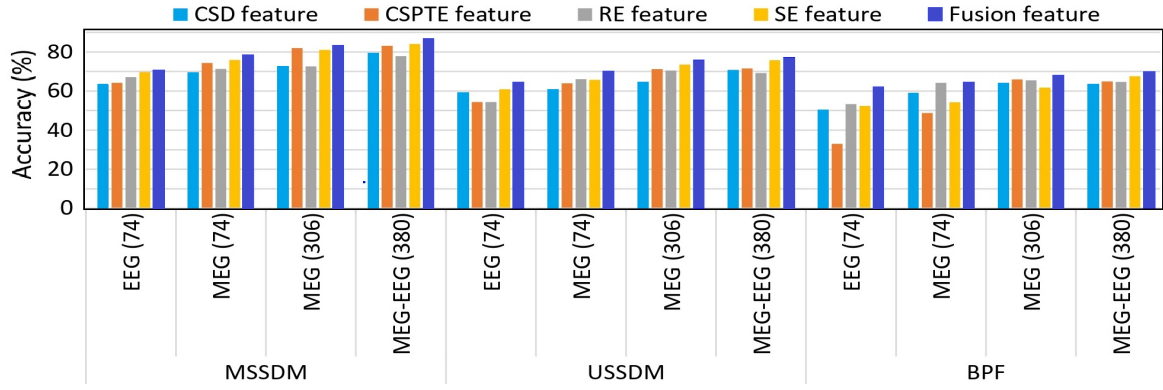


Figure 3.10: Overall performance using different feature schemes.

achieved the lowest classification accuracy of 68.75% for 6-class visual objects using phase locking tracking potential (PLTP) feature with SVM classifier from MEG (306-channel) signals. Similar work mentioned in [5] has delivered the 5-class performance of 74.5%, 74%, 85%, and 75% for LDA, SVM, GNB, and WeiRD-based supervised classification framework with RDM features from 306-MEG channel scheme, respectively. In contrast, our suggested MSSDM-FF-LDA framework using EEG 306-channel yields an enhanced average classification accuracy of 80.71%. Against the above approach, the classification model proposed by Kong et al. [4] has achieved a competent accuracy of 65.60% using a CNN-based machine learning approach. The mentioned studies have delivered poor classification performance even though they have utilized a deep learning-based approach for classification purposes. In comparison, our proposed MSSDM-FF-LDA framework achieves an improved average classification accuracy of 69.36% on EEG 74-channel. However, the above

# CHAPTER 3. MULTIVARIATE SSDM FOR MULTIVARIATE SIGNAL ANALYSIS WITH COGNITIVE VISUAL OBJECT DETECTION FROM MULTICHANNEL EEG-MEG SIGNALS

Table 3.4: Performance comparison against state-of-the-art methods available in the literature.

Authors, year, and reference	Methodology	Participants, stimulus, channels, classes	Cross-validation technique	Classification performance				Significance level
				ACC (%)	SEN (%)	SPE (%)	F1-score	
Cichy et al. (2014) [3]	PLTP, SVM	16, 100 to 1200 ms, MEG 306-channel, 6-class	NR	68.75	NR	NR	NR	$p < 0.05$
Cichy et al. (2017) [14]	Peak latency, RSA, SVM	16, -100 to 900 ms, EEG/MEG 380-channel, 5-class	Subject-independent, 10-fold	$\approx 75$	NR	NR	NR	NR
		16, MEG 306-channel, 5-class	Subject-independent, 10-fold	$\approx 74$	NR	NR	NR	NR
		16, MEG 74-channel, 5-class	Subject-independent, 10-fold	$\approx 72$	NR	NR	NR	NR
		16, EEG 74-channel, 5-class	Subject-independent, 10-fold	$\approx 68$	NR	NR	NR	NR
Guggenmos et al. (2018) [5]	RDM, LDA	16, 50 to 550 ms, MEG 306-channel, 6-class	Pearson distance	74.5	NR	NR	NR	$p < 0.05$
	RDM, SVM	16, 50 to 550 ms, MEG 306-channel, 6-class	Pearson distance	74	NR	NR	NR	$p < 2-16$
	RDM, WeiRD	16, 50 to 550 ms, MEG 306-channel, 6-class	Pearson distance	85	NR	NR	NR	$p < 2-16$
	RDM, GNB	16, 50 to 550 ms, MEG 306-channel, 6-class	Pearson distance	75	NR	NR	NR	$p = 0.012$
Kong et al. (2020) [4]	RDMs, CNN	16, EEG 74-channel, 5-class	Euclidean distance	65.60	NR	NR	NR	
BPF	BPF-FF-LDA (EEG 74-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	61.92	59.66	55.50	57.58	$p < 0.45$
	BPF-FF-LDA (MEG 74-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	64.23	61.89	58.15	59.92	$p < 0.45$
	BPF-FF-LDA (MEG 306-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	68.34	64.58	61.22	64.42	$p < 0.45$
	BPF-FF-LDA (EEG/MEG 380-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	69.96	66.29	63.89	69.66	$p < 0.45$
USSDM	USSDM-FF-LDA (EEG 74-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	64.16	65.12	62.46	63.01	$p < 0.05$
	USSDM-FF-LDA (MEG 74-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	69.95	68.95	66.39	68.31	$p < 0.05$
	USSDM-FF-LDA (MEG 306-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	75.71	73.12	72.37	74.21	$p < 0.05$
	USSDM-FF-LDA (EEG/MEG 380-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	77.29	73.68	79.89	76.84	$p < 0.05$
Proposed framework-1	MSDDM-FF-LDA (EEG 74-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	70.35	74.03	72.29	71.76	$p < 0.01$
Proposed framework-2	MSDDM-FF-LDA (MEG 74-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	78.25	79.06	77.87	80.28	$p < 0.01$
Proposed framework-3	MSDDM-FF-LDA (MEG 306-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	86.05	85.80	83.04	84.79	$p < 0.01$
Proposed framework-4	MSDDM-FF-LDA (EEG/MEG 380-channel)	16, -100 to 900 ms, 5-class	Subject-independent, 10-fold	86.42	87.08	83.11	84.89	$p < 0.01$

Note: NR: Not reported, PLTP: Phase locking tracking potential.

studies are limited to EEG-based visual cognitive analysis. On the other side, Cichy et al. [14] have employed raw extracted features using RSA and SVM for different channel selection schemes i.e. EEG (74-channel), MEG (306-channel) and combined EEG-MEG (380-channel) to improve visual object classification and yielded highest average classification performance. In extension to this, our work delivers competent VSA classification results when EEG and MEG channels are selected separately or combined EEG-MEG channels. Additionally, we have compared the proposed method with the USSDM approach (channel-specific analysis) and direct rhythms analysis which is computed using a BPF. [115]. The comparative performance is demonstrated in the Table 3.4.

Overall, we state that the proposed MSSDM method provides more significant features that enhance interpretability and inter-channel information that boosts the performance of visual object classification. These obtained features deploy the most discriminate multivariate modulation pattern of EEG-MEG data for effective visual object discrimination. Despite using subject-independent cross-validation, the proposed MSSDM method significantly delivers higher performance that verifies suitability for practical applications, especially in the visual object classification of EEG and MEG signals. The proposed MSSDM method effectively captures homogeneous spectral characteristics in multichannel with channel-aligned mode extraction and improves joint multivariate mutual features related to visual cognitive analysis. However, it requires precise parameter tuning in the designed SFT spectrum estimation and multivariate filter bank.

### **3.5 Summary**

This work proposes an integrated approach using a novel multivariate SSDM and RCFF feature for visual imagery multiclass decoding. A novel extension of SSDM is developed for multivariate analysis of nonstationary multichannel EEG-MEG signals as well as performance improvement. We have also used multivariate TFR for analysis of multichannel EEG-MEG signals and enhanced underlying visual stimulus activation patterns for visual object recognition. SFT spectrum-assisted MSSDM delivers an optimized OC mode with a high spectral resolution, and thus could potentially exhibit highly correlated multi-spectral

modulation among all channels. It has been demonstrated that these novel fusion features were very effective in distinguishing 5-class visual objects and improved classification performance in a 1 s short-duration epoch. The performance of the method has been evaluated using different visual object classification frameworks based on five classifiers with a subject-independent 10-fold cross-validation approach.

The analysis results of the MEGEEG92 Objects Dataset demonstrate that the proposed MSSDM-FF-LDA framework has shown distinguished performance and achieved average recognition accuracy of 70.35%, 78.25%, 83.05%, and 86.42% for EEG (74-channel), MEG (74-channel), MEG (306-channel), EEG-MEG (380-channel) respectively. Overall, it is indicated that the proposed MSDDM method with combined EEG-MEG data can be a useful tool for the analysis of visual objects instead of using EEG and MEG separately. As the future unfolds, the prospective work involves the development of a methodology based on a multivariate TF approach to enhance VSA classification performance with a broader range of cognitive visual object classes.



## **Chapter 4**

# **Multivariate SSDM-based Joint Time-Frequency Analysis with Improved Imagined Speech Decoding for Intuitive BCI**

In this chapter, we present a novel MSSDM-based approach to reveal the multivariate joint oscillatory features across multichannel EEG signals. The JTF multivariate features are generated based on the JIF and JIA functions from the extracted CAOCs. The formulation of these JTF multivariate features is detailed in this chapter. Additionally, we also designed an end-to-end framework for imagined speech task detection for BCI applications that employs JTFDF from four pre-trained neural networks and an SVM classifier on cross-channel IMS-EEG signals. This work also introduced a computationally efficient framework for a feasible MI-EEG-based BCI model. The detailed evaluation of the proposed classification framework on computed most discriminant JTFDF features using two feature mapping strategies is given in later sections of the chapter.



## 4.1 Introduction

In the recent decade, IMS has developed advanced cognitive communication tools, serving as an intuitive paradigm within BCI technology [15, 17]. Against imagination of control action like in MI-based BCI, IMS is a new intuitive paradigm that conducts the internal pronunciation of words without any physical movement or audio output. Additionally, IMS exhibits scalability across multiple classes, indicating the potential for constructing a scalable BCI system. This paradigm is especially well-suited for constructing communication systems because of its intuitive nature, making it particularly beneficial for individuals facing physical challenges like apraxia and dysarthria. Researchers have investigated several experimental techniques for IMS that include EEG, MEG, and fMRI [19, 116]. In these studies [15, 17, 18, 19, 20], EEG is widely utilized for its non-invasiveness and affordability, offering improved fine-grained analysis by identifying spectral, spatial, and temporal components underlying speech categories. Over the past few years, several studies have been conducted to assess the potential of EEG to extract informative features for automated IMS recognition systems. Dasalla et al. [15] have used a CSP feature to classify imagined vowels from IMG-EEG signals. The obtained features have been classified using SVM and the proposed system achieved a maximum accuracy of 78%. In the study [17], the use of the discrete wavelet transform (DWT) and a random forest (RF) classifier was explored for the classification of vowels and words during IMS and achieved 22.72% and 19.60% accuracy for 5-vowel and 6-word classification, respectively. García-Salinas et al. [18] have employed a bag of features (BoF) approach for the detection of 5-class imagined words and reported an accuracy of 68.9%. In further work, the two different feature connectivity approaches, i.e., a covariance-based connectivity measure (CCM) and a maximum linear cross-correlation-based connectivity measure (MCCM) have been utilized by Qureshi et al. in [21] to determine the most significant features for the enhancement of IMS classification performance. It explored phase-only data from EEG signals with four different channel selection strategies and achieved a maximum accuracy of 40.30% (CCM) and 87.90% (MCCM) for 5-class using an extreme learning machine (ELM) classifier. In similar work, Saha et al. [16] adopted a hybrid framework that combined LSTM and CNN networks to

formulate concatenated deep features and delivered an accuracy of 83.42% in 2-class IMS classification. Recent work by Cooney et al. in [20] delivered an IMS decoding performance of 32.35% accuracy for 5-class imagined vowel classification using independent component analysis with Hessian approximation. To recognize more IMS classes, Lee et al. [19] have attempted to classify 13-class using CSP features and a shrinkage regularized linear discriminant analysis (RLDA) classifier and achieved an average classification accuracy of 38.67% (in Correto dataset) and 37.33% (in International BCI Competition dataset). Furthermore, Lee et al. [22] attempted a Siamese neural network (SNN) encoder-based end-to-end classification framework using SVM and deep feature and delivered 31.40% accuracy for 6-class EEG imagined word speech.

However, most of the above-demonstrated research studies in IMS recognition have predominantly extracted handcrafted features utilizing deep learning techniques from the multichannel EEG signals. Handcrafted approaches to feature extraction strongly rely on how classification models are designed to extract mutual feature characteristics across multiple EEG channels, thus introducing computational challenges for real-time deployable BCI systems. In addition, the majority of studies have primarily employed channel-specific analysis, i.e., univariate approaches to extract features, and have not explored joint channel analysis for mutual feature extraction across multivariate EEG data [65]. These methods suffer in computing the most distinctive channel-aligned common features across multichannel EEG signals with optimal multi-class scalability in the IMS, thus leading to inadequate classification performance. Therefore, an adaptive multivariate signal analysis technique is needed to extract the joint oscillatory features from cross-channel with optimizing channels and computational efficiency.

Joint TF analysis (JTFA) methods have recently proven to be powerful tools for multichannel multivariate signal analysis. In particular, JTFA methods adopt JIF and bandwidth to model the joint oscillatory structure of multichannel signals and give enhanced joint TF localization. Numerous multivariate approaches have been developed for JTFA across various real-time applications [63, 64, 65, 117], Recently, an adaptive and localized TFR-based method has been implemented with the help of modified SWD [67] and SFT spectrum to enhance TF localization for EEG signals. Although some of these existing JTFR methods

have reported significant results in TF analysis, but they also introduce scope for improvement due to cross-term interference and computational cost for the multi-channel multivariate EEG signals. However, there have been very limited reported studies on IMS, and to the best of our knowledge, JTFR-based techniques for multivariate IMS data have not been studied.

In our proposed work, we have designed an end-to-end framework using the proposed MSSDM (described in Chapter 3) to extract an ensemble of CAOCs from multichannel IMS-EEG signals for IMS enhancement. The proposed MSSDM explores the joint oscillatory structure of multichannel EEG signals and develops the joint TF plane for multivariate IMS-EEG signals by computing JIF and JIA functions from signal adaptive scales. MSSDM incorporates a mode alignment approach to estimate the orders of effective CAOCs modes to establish spectral correlation in multivariate IMS-EEG signals by finding common frequency components across channels. In MSSDM, an adaptive scale-space approach is employed by designing multivariate SwF banks and SFT spectrum optimization to deliver an ensemble of CAOCs modes by estimating channel-aligned spectral boundaries across channels. To enhance spectrum estimation, a new sparse basis optimization model has been designed by employing the fused-Lasso technique. This optimized SFT model offers efficient spectral boundary estimation and ensures mode alignment, thus leading to improved TF localization.

To automate feature extraction, we have introduced a robust JTFDF using different pre-trained neural networks for efficient recognition of IMS. The multi-scale JTFDF features are derived from MSSDM-based JTF images by employing a different deep architecture to extract channel-aligned common JIF and JIA information present in the generated CAOCs. To design an efficient and feasible IMS recognition model, we have investigated feature optimization by considering a few-layer deep neural network feature map architecture to avoid time complexity and most discriminant features have been computed using two well-known feature mapping techniques, CCA and HDC. Finally, the effectiveness of the developed MSSDM-based classification models was assessed on the obtained JTFDF features using SVM classifier [118] in 10-fold cross-validation on cross-subject analysis. The novelty of this study is to present a novel MSSDM-based JTF analysis method and low-computational

JTFDF feature for enhancement of the IMG-EEG recognition model, making it suitable for feasible BCI applications. The key contributions are summarized as follows:

1. A novel adaptive MSSDM has been proposed to reveal the multivariate joint oscillatory structure of multichannel EEG signals and deliver CAOCs from cross-channel IMS-EEG signals for IMS enhancement.
2. Joint TF images are generated based on JIF and JIA functions computed from the extracted CAOCs.
3. The new low-dimensional JTFDF features have been computed using four different pre-trained neural networks, and two feature mapping strategies (CCA and HDC) have been employed to generate the most discriminant features for improved IMG-EEG BCI classification.
4. The performance of the proposed MSSDM-based IMG-EEG classification framework is compared with existing state-of-the-art methods with a cross-subject dataset.

## **4.2 Dataset and experimental paradigm used**

In our work, we have assessed the performance of a proposed method on two publicly available datasets for IMS. The first dataset, the Coretto dataset [17], consists of EEG signals for 15 subjects during IMS using 6-channel (C3, C4, P3, P4, F3, and F4) located in Wernicke's area, which is important for the analysis of imagined cognitive responses. The dataset consists of 11 classes, including 6 Spanish words and 5 Spanish vowels. Here, we have selected six Spanish words used to evaluate the proposed work: "arriba" (up), "abajo" (down), "derecha" (right), "izquierda" (left), "adelante" (forward), and "atrás" (backward) to control external devices intuitively. The second dataset, International BCI Competition dataset [119] (Track 3) at (<https://osf.io/pq7vb/>), consisted of EEG signals for 15 subjects during IMS using 64-channel. The subjects were asked to imagine speaking 5 different commands: "Hello", "Help me", "Stop", "Thank you", and "Yes". These commands are considered useful for patients as they can be used to control various devices and systems.

Figure 4.1 depicts the timing scheme of the recording paradigm utilized in this study. In our experimentation, we have employed k-fold cross-validation with 25 trails and 46 trails according to imagined word 6 and 5 classes from both datasets. During the pre-processing stage, data acquisition was sampled at 1024 Hz and filtered out with a BPF with a pass-band frequency range of 0.01 Hz to 80 Hz. To filter out artifacts, we have employed two filters: an infinite impulse response HPF (0.8 Hz) and a low-pass smoothing filter (30 Hz). Further in segmentation, 4 s windows are applied to generate time-locked post-stimulus epochs for IMS analysis.

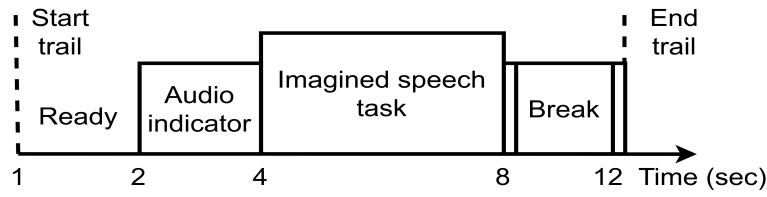


Figure 4.1: Experimental paradigm for IMS recording.

## 4.3 Proposed methodology

The objective of this study is to develop an automated MSSDM-based IMS recognition system that can efficiently extract multivariate features across multichannel EEG signals to classify IMG-EEG classes. Figure 4.2 depicts the proposed methodology framework with detailed descriptions, followed by brief explanations of all involved methods in subsequent subsections.

### 4.3.1 MSSDM-based multivariate time-frequency representation

Multichannel EEG signals are decomposed into CAOC using the MSSDM method (described in Chapter 3). To analyze the CAOCs obtained from the MSSDM method in the TF domain, we determine the multivariate TF coefficients. These coefficients are used to create multivariate TF images by computing the JIA and JIF from the extracted CAOCs. The steps involved in computing multivariate TF coefficients are briefly outlined in this section.

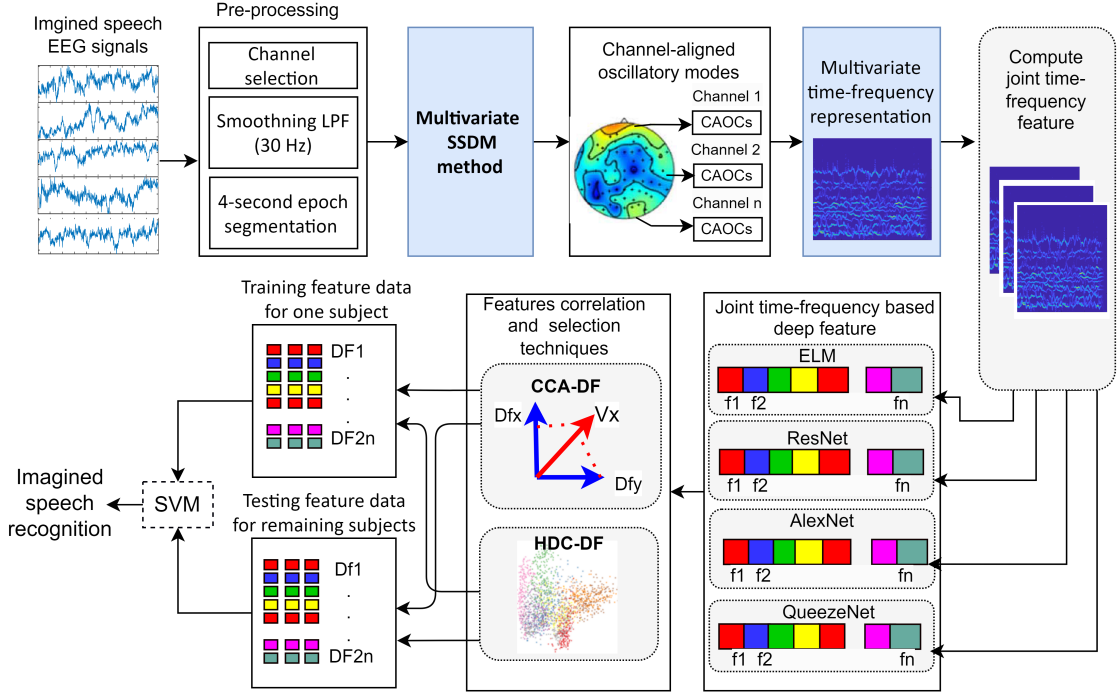


Figure 4.2: Block diagram of the proposed approach for IMS recognition using MSSDM method.

The set of CAOCs obtained after applying MSSDM corresponding to  $N$ -channel from nonstationary signal  $y(t)$  can be expressed as follows:

$$y(t) = \begin{bmatrix} y_{1, \text{CAOC}_1}(t) & y_{1, \text{CAOC}_2}(t) & \cdots & y_{1, \text{CAOC}_Q}(t) \\ y_{2, \text{CAOC}_1}(t) & y_{2, \text{CAOC}_2}(t) & \cdots & y_{2, \text{CAOC}_Q}(t) \\ \vdots & \vdots & \ddots & \vdots \\ y_{N, \text{CAOC}_1}(t) & y_{N, \text{CAOC}_2}(t) & \cdots & y_{N, \text{CAOC}_Q}(t) \end{bmatrix} \quad (4.1)$$

where  $y_{n, \text{CAOC}_q}(t)$  is the  $q^{\text{th}}$  CAOC corresponding to  $n^{\text{th}}$  channel and  $Q$  is the number of CAOCs. In order to obtain the IA and IF functions of a multichannel nonstationary signal  $y(t)$ , the Hilbert transform (HT) [41] has been applied to each of the extracted CAOCs across each channel and its analytic signal representation is given as,

$$y_{+(k, \text{CAOC}_Q)}(t) = y_{k, \text{CAOC}_Q}(t) + iH(y_{k, \text{CAOC}_Q}(t)) \quad (4.2)$$

where,  $H(\cdot)$  is HT operator. The computed IA and IF of each extracted CAOC for  $y(t)$  for

the corresponding individual channels is expressed by Eqs. (4.3) and (4.4), respectively.

$$A_{k, \text{CAOC}_Q}(t) = \sqrt{(y_{k, \text{CAOC}_Q}(t))^2 + (H(y_{k, \text{CAOC}_Q}(t)))^2} \quad (4.3)$$

$$F_{k, \text{CAOC}_Q}(t) = \frac{d}{dt}[\phi_{k, \text{CAOC}_Q}(t)] \quad (4.4)$$

The multivariate IA ( $A(t)$ ) and IF ( $F(t)$ ) for signal  $y(t)$  are given as,

$$A(t) = \begin{bmatrix} A_{1, \text{CAOC}_1}(t) & A_{1, \text{CAOC}_2}(t) & \cdots & A_{1, \text{CAOC}_Q}(t) \\ A_{2, \text{CAOC}_1}(t) & A_{2, \text{CAOC}_2}(t) & \cdots & A_{2, \text{CAOC}_Q}(t) \\ \vdots & \vdots & \ddots & \vdots \\ A_{N, \text{CAOC}_1}(t) & A_{N, \text{CAOC}_2}(t) & \cdots & A_{N, \text{CAOC}_Q}(t) \end{bmatrix} \quad (4.5)$$

$$F(t) = \begin{bmatrix} F_{1, \text{CAOC}_1}(t) & F_{1, \text{CAOC}_2}(t) & \cdots & F_{1, \text{CAOC}_Q}(t) \\ F_{2, \text{CAOC}_1}(t) & F_{2, \text{CAOC}_2}(t) & \cdots & F_{2, \text{CAOC}_Q}(t) \\ \vdots & \vdots & \ddots & \vdots \\ F_{N, \text{CAOC}_1}(t) & F_{N, \text{CAOC}_2}(t) & \cdots & F_{N, \text{CAOC}_Q}(t) \end{bmatrix} \quad (4.6)$$

It has been observed that optimized filter banks based on MSSDM result in CAOCs that exhibit distinguished IFs at each oscillatory level and are well-separable with time instant. To determine the JIA ( $A_N^{\text{multi}}(t)$ ) and JIF  $F_N^{\text{multi}}(t)$  for each oscillatory level, the IA and IF of CAOCs across different  $N$ -channel are combined, which are computed as follows:

$$A_N^{\text{multi}}(t) = \sqrt{\sum_{k=1}^N [(A_{k, \text{CAOC}_N}(t))^2]} \quad (4.7)$$

$$F_N^{\text{multi}}(t) = \frac{\sum_{k=1}^N [A_{k, \text{CAOC}_N}(t)]^2 F_{k, \text{CAOC}_N}(t)}{\sum_{k=1}^N [A_{k, \text{CAOC}_N}(t)]^2} \quad (4.8)$$

The computation of multivariate TF coefficients at each oscillatory level is given as follows:

$$\text{TF}_N^{\text{multi}}(A, F) = A_N^{\text{multi}}(t) \delta[F - F_N^{\text{multi}}(t)] \quad (4.9)$$

Lastly, multivariate TF images are formed by combining all computed joint oscillatory lev-

els across  $N$ -channel, which is expressed as,

$$\text{TF}^{\text{multi}}(A, F) = \text{TF}_N^{\text{multi}}(A, F) \quad (4.10)$$

The overall flow of JTFDF formulation is given in algorithm 4.1.

---

**Algorithm 4.1** Computation of MSSDM-based JTFDF features from multichannel EEG data

---

Input: Nonstationary multivariate or multichannel signal  $z(k)$   
Output: JTFDF features from CAOCs modes using MSSDM  
**// Estimation of common boundaries**  
**Repeat**  
Optimise and generate SFT spectrums by Eq. (2.4)  
Spectrum smoothing and selection using SGF by estimation highest ESD  $s_{it}(k)$   
Obtain extremes of SFT spectrum across cross-channel  $N_c$   
Compute mean spectrum by Eq. (3.1)  
Find spectral boundaries in channel-aligned mean spectrum by Eq. (2.7)  
 $it \leftarrow it + 1$   
**Until**  $N_{it} < N_c$  ;  $N_c$  - number of channels  
**// Apply multivariate swarm filter banks and extract CAOCs**  
**Repeat**  
**Repeat**  
Smooth spectrum by estimation highest ESD  $z_{it}(k)$   
Estimate spectral boundaries in the mean SFT spectrum  
Select swarm filter parameters  $\delta$  and  $M$  using GA by Eqs. (2.9a) and (2.9b), respectively  
**Repeat**  
 $y_{j+1} \leftarrow \text{SwF}(z_j(k), M, \delta)$  ; SwF: Swarm filter  
 $j \leftarrow j + 1$   
 $B_{it+1}(k) \leftarrow B_{it}(k)$   
**Until**  $B_{it} < 0.1$   
 $s'_{it}(k) \leftarrow y_j(k)$   
 $z_{it+1}(k) \leftarrow z_{it}(k) - s'_{it}(k)$   
 $it \leftarrow it + 1, j = 0$   
 $y_0(k) \leftarrow z'_{it}(k)$   
**Until**  $F_{Bz_{it}} < P_{th}$  ;  $P_{th}$  is set to 0.1  
 $C_{it} \leftarrow C_{it+1}$   
**Until**  $N_{it} < N_c$   
 $Y_{it}(k) \leftarrow y_0(k)$   
**// Estimation of JTFDF features**  
Computation of IA and IF using HT by Eqs. (4.3 and 4.4), respectively  
Determine the JIA and JIF across  $Q$ -channel by Eqs. (4.7 and 4.8), respectively  
Find multivariate JTF coefficients by  $N_c$  by Eq. (4.9)  
Apply deep neural network on JTF images to get JTFDF  
 $T_{it} \leftarrow T_{it+1}$   
**Until**  $T_{it} < \text{Max}_{tr}$  ;  $\text{Max}_{tr}$  - maximum number of trails

---

To demonstrate the ability of the MSSDM method to identify and align common mode



oscillations present in the multivariate signals cross-channel, IMS-EEG signals have been tested. Figure 4.3 illustrates MSSDM-based channel-aligned decomposed CAOCs from 3-channel (F3, F4, and P3) IMG-EEG signals for the “Right” class (subject 2) in Correto dataset. It is observed that the extracted oscillatory modes from channels F3, F4, and P3, are strictly aligned with common time-varying frequency components and are locally present with individual single-mode, shown in Fig. 4.3 (h). In addition, we have plotted MSSDM-based swarm filter bank structures that were used to extract common or joint oscillatory modes for different cross-channel signals. The use of Welch power spectral density (PSD) in spectral analysis, as depicted in 4.3 (i), demonstrated the alignment of the selected CAOCs with the MSSDM filter bank. The PSD plots have exhibited a similar spectral response, indicating that the CAOCs detected oscillations across multiple channels consistently and were aligned correctly across cross channels corresponding to the respective CAOCs. Thus, we can state that the proposed MSSDM-JTFDF is an effective approach to extracting multivariate modulated oscillations across multichannel data while simultaneously confirming mode alignment.

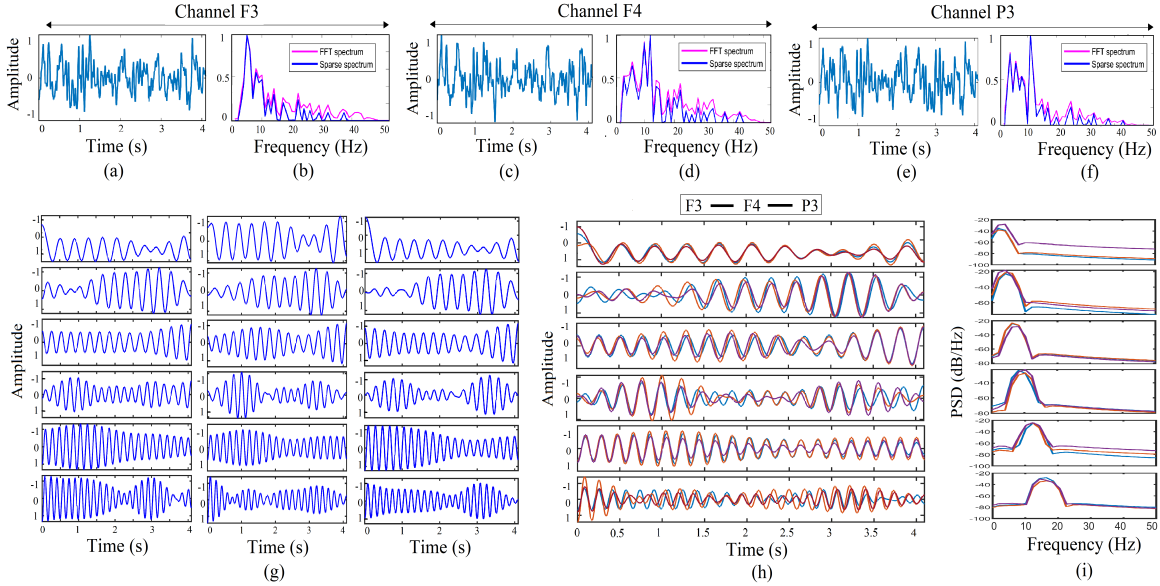


Figure 4.3: Plots of (a)-(f) raw EEG signals and their spectral analysis using FFT with SFT coefficients, (g)-(h) MSSDM-based decomposed CAOCs and its combined visual interpretation, and (i) MSSDM-based swarm filter bank structure of 3-channel (F3, F4, and P3) IMG-EEG signals for the “Right” class (subject 2) in Correto dataset.

### 4.3.2 Feature formulation and classification

In the realm of BCI applications, designing suitable features from multivariate signals is an important stage for constructing an efficient IMG-EEG classification model. In our study, a new feature approach has been introduced that employs CNN deep features and feature mapping techniques to extract significant features from multivariate JTF images for efficient IMS classification. Utilizing deep learning, it is proven higher accuracy and more flexibility by automatically uncovering the underlying hidden patterns within different sets of IMS class images. These JTFDF features are derived from the MSSDM-based JTF images in a categorical way using different pre-trained neural networks (NNs): ELM, ResNet, AlexNet, and SqueezeNet, which enhance channel-aligned common JIA and JIF information present in the obtained CAOCs. To extract deep learning features, CNN-estimated filters were applied to process the JTF images of the size of  $875 \times 756 \times 3$  to learn the feature attributes of the JTFR image and generated discriminative features to improve the performance of multi-class classification. To understand how different pre-trained learning models affect the complexity of computed features, we have selected four learning models that model dense layer formulation in varied complexity [120, 121, 122, 123]. Figure 4.4 illustrates the architectural configuration for the extraction of deep features. The selection of ELM [123], SqueezeNet [122], and AlexNet [121] is due to their shorter training duration and fewer dense layers, whereas ResNet [120] exhibits a comparatively very complex dense layer. Table 4.1 lists the hyper-parameters of the used deep architecture. It is observed that the used network delivers high-dimensional deep learning features that demonstrate incompatibility to a feasible BCI system. Therefore, it is necessary to map high-dimensional deep learning feature space into low-dimensional deep learning feature space by mapping the most correlated features among the class. Thus, the new low-dimensional JTFDF features have been computed by the use of two feature mapping techniques (CCA and HDC) to generate the most discriminant features for improved IMG-EEG BCI classification. Among them, the HDC-based feature correlation method is new and was used first time for a feature reduction approach. A detailed description of these feature-mapping techniques is provided below.

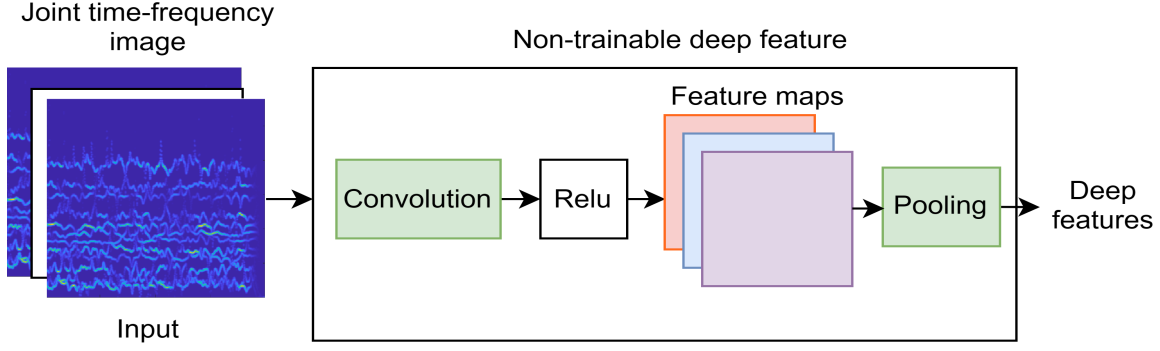


Figure 4.4: Proposed architecture for deep features extraction.

Table 4.1: Hyper-parameter configuration of four different models.

Model	Network configuration
ResNet	2-D COV: 64@1×1×64 (1), 64@3×3×64 (3), 256@1×1×64 (4), 64@1×1×256 (3), 64@1×1×64 (2), 128@1×1×256 (2), 512@1×1×128 (4), 1024@1×1×256 (6), 2048@1×1×512 (3), filters number = 6, BN with channels: 64 (7), 128 (8), 256 (16), 2-D MP: 3×3 (1), ReLU: 124
AlexNet	2-D COV: 96@11×11×3 (1), 128@5×5×48 (1), 384@3×3×256 (1), 192@3×3×192 (1), 128@3×3×192 (1), filters number = 5, BN with channels: 5 (2), 2-D MP: 3×3 (3), ReLU: 5
SqueezeNet	2-D COV: 64@3×3×3 (1), 16@1×1×64 (1), 64@1×1×16 (2), 64@3×3×16 (2), 16@1×1×128 (1), 64@1×1×16 (2), 64@3×3×16 (2), 32@1×1×128 (1), 128@1×1×32 (2), 128@3×3×32 (2), 32@1×1×256 (1), 128@1×1×32 (3), 128@3×3×32 (2), 48@1×1×256, 192@1×1×48 (2), 48@1×1×384 (1), 192@1×1×48 (3), 64@1×1×384 (1), 256@1×1×64 (2), 256@3×3×64 (2), 256@1×1×64 (1), 256@3×3×64 (1), 1000@1×1×512 (1), 2-D MP: 3×3 (3), DC (2 inputs): 8, ReLU: 26
ELM	2-D COV: 384@3×3×256 (1), filters number = 3, BN with channels: 5 (1), 2-D MP: 5×5 (1), ReLU: 1

COV: Convolution, BN: Batch normalization, ReLU: Rectified linear activation function, MP: Max pooling, DC: Depth concatenation.

#### 4.3.2.1 Canonical correlation analysis

The CCA is an effective feature selection method to measure functional connectivity in the deep features by employing maximally auto-correlated and mutually un-correlated features from feature space of different classes [124]. To find the most correlated feature sets  $[U, V]$ , CCA is used to determine the feature weight matrix  $w_n$  for  $n$ -class by maximizing the correlation between two feature sets  $x$  and  $y$  within the class by solving the following optimization problem, which is given as follows:

$$\max_{(W_x, W_y)} \rho(U, V) = \frac{W_x^T C_{xy}^T W_y}{\sqrt{(W_x^T C_{xx} W_x)(W_y^T C_{yy} W_y)}} \quad (4.11)$$

Where  $[U, V]$  represents the canonical pair for  $n$ th class.  $C_{xx}$  and  $C_{yy}$  indicate the autocovariance matrices of feature sets  $x$  and  $y$ .  $C_{xy}^T$  is the cross-covariance matrix computed from

the feature sets  $x$  and  $y$ . Here we have considered canonical auto-correlation coefficient  $\rho^2 = [0 \ 1]$  with respect to eigenvectors  $W_x$  and  $W_y$  for feature sets  $x$  and  $y$ .

#### 4.3.2.2 Hellinger distance-based correlation

Hellinger distance-based similarity measure [125] is a new feature reduction approach to measure the degree of similarity between two probability distributions, which is computed from two different feature sets  $x$  and  $y$  within a given class. The Helinger distance  $H(x, y)$  is given as,

$$H(x, y) = \sqrt{1 - \sum (\sqrt{p(x)} - \sqrt{p(y)})^2} \quad (4.12)$$

Where  $p(x)$  and  $p(y)$  represent the probability distributions of the feature sets  $x$  and  $y$ . In our work, we considered different correlation factors (CF) in the range of 0.3 to 0.8 to compute the feature space. After this step, both the CCA and HDC-based reduced features were fed separately to the SVM classifier [118] to obtain IMS class recognition.

## 4.4 Results

To assess the efficacy of the proposed MSSDM-JTFDF, we have tested channel-aligned multi-scale JTFDF features on a designed end-to-end classification model for the enhancement of IMS based on IMS-EEG signals. The performance of IMS-EEG multi-class classification was investigated using two public datasets: the 6-class Coretto dataset [17] and the 5-class International BCI Competition dataset [119], focusing on word imagination speech. To study the effectiveness of the proposed MSSDM-JTFDF, channel-aligned multi-scale JTFDF features have been tested by designing an end-to-end classification framework for IMS-EEG signals-based IMS enhancement. Under experimentation, IMS-EEG signals were examined using various electrode configurations, including six electrodes (C3, C4, P3, P4, F3, and F4) [17] and 64 electrodes [119], respectively. Each of these EEG signals was analyzed within a 30-second window and further segmented into 4-second epochs with imagined activation stimuli. Further, HT [41] was used to transform multivariate TF im-

ages by computing multivariate TF coefficients from extracted CAOCs that were extracted using MSSDM. To strongly correlate mutual multivariate information, we have explored joint TF plots across every cross-channel and computed JTF images for each class. Figure 4.6 shows the decomposition results and their JTF analysis for 5-class IMS from 3-channel EEG signals in the case of BCI Competition dataset. For each word imagination speech class, the number of multivariate JTFR images generated is given in Table 4.2. Figure 4.5 demonstrates the obtained multivariate JTF images across all subjects based on extracted MSSDM-based CAOCs modes for each word imagination speech class (Coretto dataset). In response to word-imagination speech tasks, these plots exhibit JTF analysis for discriminative brain activity patterns. Notably, in the JTF plots, the most pronounced activation during the word imagination tasks was noted in the theta ( $\theta$ : 4–8 Hz) and alpha ( $\alpha$ : 8–13 Hz) rhythms, corresponding to imagined stimulus response within the time interval (0.5 s - 2.0 s) from the onset of imagination.

Table 4.2: Architecture of used pre-trained neural networks with obtained deep features.

Dataset	Pre-train classification model	Size of input layer	Number of layers	Parameter (in millions)	Trainable parameter	Feature dimension obtained from single JTF image	Number of JTF images per class	Feature space size	Total feature size using CCA with CF = 0.7	Total feature size using HDC with CF = 0.7
Coretto DB	AlexNet	875x756x3	16	0.40	4096	2048	225	460800	276480	165888
	ResNet-18	875x756x3	174	2.08	32768	16384	225	3686400	2211840	1327104
	SqueezeNet	875x756x3	66	0.81	15360	7680	225	1728000	1036800	622080
	ELM	875x756x3	9	0.14	2048	1024	225	230400	138240	82944
BCI Competition DB	AlexNet	875x756x3	16	0.58	4096	2048	360	737280	442368	265421
	ResNet-18	875x756x3	174	0.28	32768	16384	360	5898240	3538944	2123366
	SqueezeNet	875x756x3	16	1.08	15360	7680	360	2764800	1658880	995328
	ELM	875x756x3	9	0.29	2048	1024	360	368640	221184	132710

Further, we have computed JTFDF using different pre-trained neural networks, including ELM, AlexNet, SqueezeNet, and ResNet. To assess the effectiveness of the MSSDM-JTFDF features, experiments were carried out employing subject-independent cross-validation analysis and a quadratic SVM classifier on both IMS-EEG datasets. For the cross-subject dataset, leave-one-out cross-validation was conducted using a 70%-30% split for training and testing data. This procedure was validated using k-fold validation schemes, and the average outcome was reported, taking into account performance metrics such as average accuracy (ACCavg) [126]. The performance study has evaluated performance across three key aspects, and the findings are discussed below. The first part of

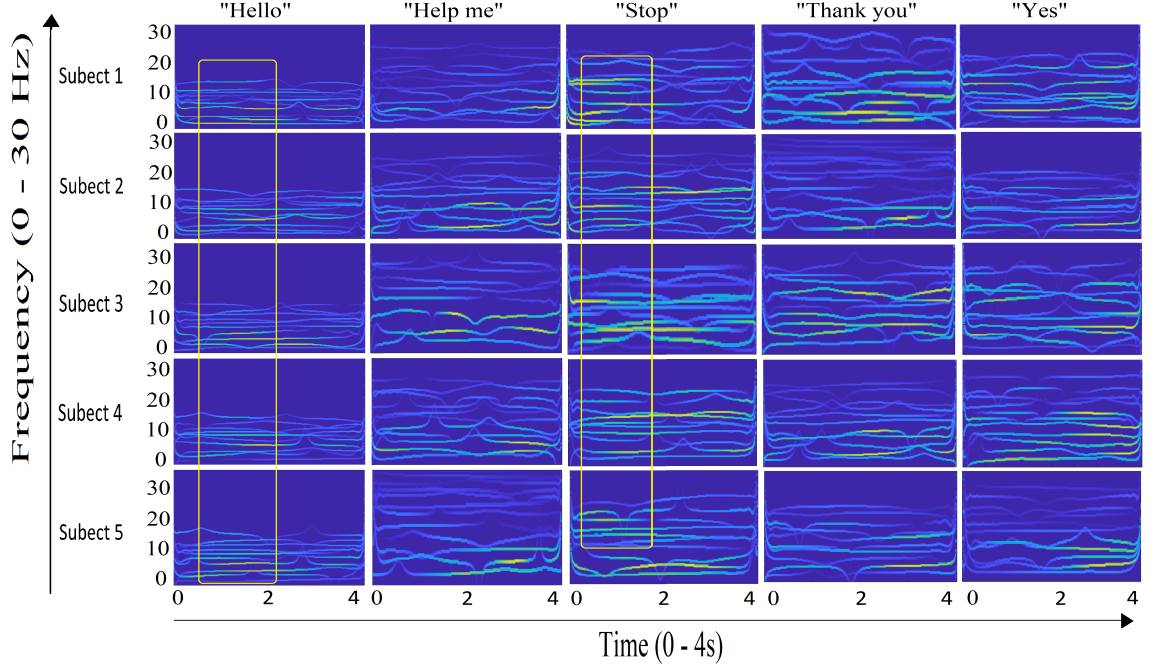


Figure 4.5: The JTF representation of the IMS word 5-class data using proposed MSSDM for 15 subjects (BCI Competition dataset). The highlighted yellow boxes in the plots demonstrate similar imagined stimulus responses with extracted  $\theta$  and  $\alpha$  bands.

the simulation study is for the subject-wise classification performance of computed JTFDF features on both datasets. In the second part of the simulation study, the performance of feature complexity with the different dense networks are compared. In the third part of the simulation study, analyze the effect of the selection of correlation factors in two different feature mapping schemes. To investigate the JTFDF feature with four pre-trained neural network and two feature correlation techniques, we have formulated the eight different MSSDM-based classification models, proposed model 1 (PM1): MSSDM-ELM-JTFDF-CCA, proposed model 2 (PM2): MSSDM-AlexNet-JTFDF-CCA, proposed model 3 (PM3): MSSDM-SqueezeNet-JTFDF-CCA, proposed model 4 (PM4): MSSDM-ResNet-JTFDF-CCA, proposed model 5 (PM5): MSSDM-ELM-JTFDF-HDC, proposed model 6 (PM6): MSSDM-AlexNet-JTFDF-HDC, proposed model 7 (PM7): MSSDM-SqueezeNet-JTFDF-HDC, and proposed model 8 (PM8): MSSDM-ResNet-JTFDF-HDC.

Tables 4.3, 4.4, 4.5, and 4.6, demonstrate the subject-wise classification performance on JTFDF features obtained from four different pre-trained neural networks (ELM, AlexNet, SqueezeNet, and ResNet) using the SVM classifier on the Coretto dataset and BCI Com-

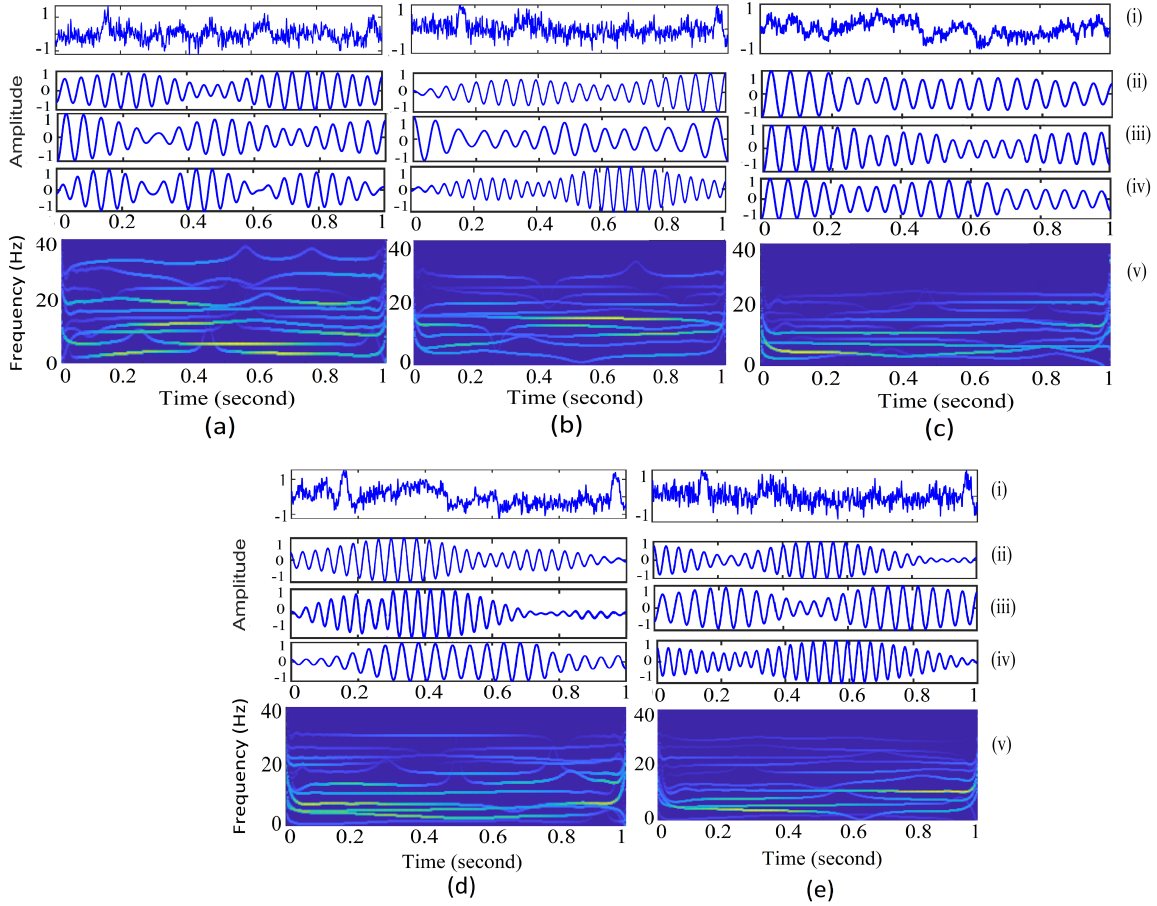


Figure 4.6: Plots of: (i)-(iii) decomposed CAOCs from 3-channel EEG signals and (iv) their JTF analysis for IMS classes (a) Hello, (b) Help me, (c) Stop, (d) Thank you, and (e) Yes, respectively.

petition dataset, respectively. Here we have verified the performance by conducting k-fold cross-validation across all subjects (see Fig. 4.10). The presented results show the obtained average accuracies in 10-fold cross-validation across fifteen subjects with correlation factor ( $CF > 0.7$ ) and exhibit the efficient selection of DFs without compromising performance. CF was calculated across the range of 0.3 to 0.8. However, the best classification results along with reduced features (Rf), were achieved when CF was set to 0.7 for all proposed models, which is shown in Fig. 4.9. In the Coretto dataset, our proposed model PM4 shows an average accuracy of  $ACC_{avg} = 60.80 \pm 1.13\%$  with reduced features ( $Rf > 45\%$ ) for the 6-class imagined word classification across all subjects. Whereas in the BCI Competition dataset, the proposed model PM7 has delivered  $ACC_{avg} = 59.10 \pm 1.68\%$  with  $Rf > 46\%$  for the 5-class problem across all subjects. We have also compared four pre-trained neural net-



# CHAPTER 4. MULTIVARIATE SSDM-BASED JOINT TIME-FREQUENCY ANALYSIS WITH IMPROVED IMAGINED SPEECH DECODING FOR INTUITIVE BCI

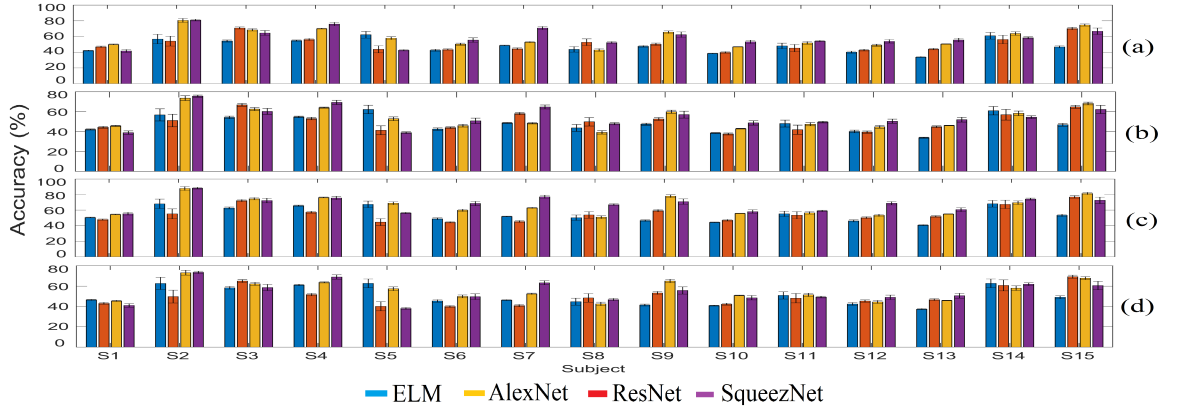


Figure 4.7: Subject-wise evaluation of 10-fold classification accuracy performed by two different feature correlation schemes, (a)-(c) HDC and (b)-(d) CCA based on Coretto dataset and BCI Competition dataset, respectively. The error bars denote the standard deviation of values  $< 1\%$ .

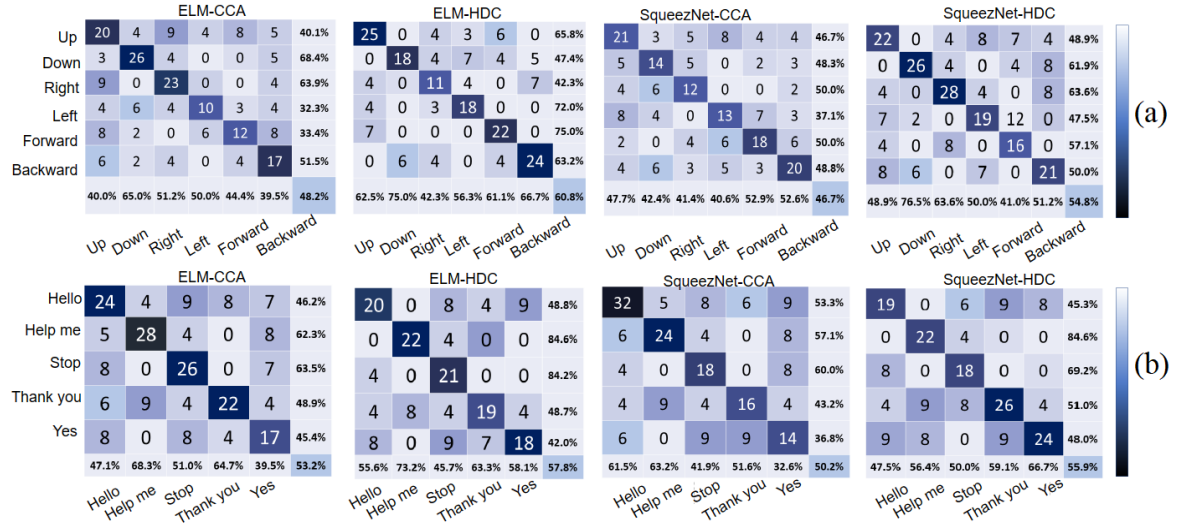


Figure 4.8: Averaged confusion matrix of the classification across all subjects in (a) the Coretto dataset and (b) the BCI Competition dataset.

work methods to determine how effective feature reduction is done with different proposed models, which is shown in Fig. 4.7. In the Coretto dataset, it is found that the highest ACCavg among all frameworks were PM1 (ACCavg = 48.07% and Rf  $> 36\%$ ), PM2 (ACCavg = 50.74% and Rf  $> 42\%$ ), PM3 (ACCavg = 56.16% and Rf  $> 39\%$ ), and PM4 (ACCavg = 60.80% and Rf  $> 45\%$ ) in the case of deep feature using ELM, AlexNet, SqueezeNet, and ResNet network, respectively. In the case of the CCA feature mapping scheme, the highest accuracy was attained by PM4 (ACCavg = 60.80% and Rf  $> 45\%$ ). Whereas in the HDC



# CHAPTER 4. MULTIVARIATE SSDM-BASED JOINT TIME-FREQUENCY ANALYSIS WITH IMPROVED IMAGINED SPEECH DECODING FOR INTUITIVE BCI

Table 4.3: Average accuracy score (in %) obtained from reduced features (in %) for four different pre-trained neural networks using CCA scheme and SVM classifier on the Correto dataset.

CCA-based feature selection: ACCavg (Rf)																
CF →	ELM				AlexNet				SqueezeNet				ResNet			
	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8
Subject 1	35.3 (18)	38.7 (27)	<b>42.1</b> (39)	41.5 (40)	39.2 (22)	41.9 (30)	<b>46.9</b> (43)	46.3 (44)	45.6 (28)	48.3 (36)	48.1 (40)	<b>52.9</b> (44)	38.0 (20)	45.8 (34)	<b>41.7</b> (33)	40.8 (35)
Subject 2	50.1 (25)	54.2 (38)	56.8 (40)	<b>63.0</b> (49)	60.6 (36)	60.0 (43)	54.2 (52)	<b>60.4</b> (54)	72.2 (47)	78.6 (58)	80.6 (43)	<b>88.6</b> (48)	68.3 (43)	77.1 (46)	<b>81.1</b> (44)	80.0 (45)
Subject 3	49.1 (25)	47.2 (37)	<b>54.3</b> (43)	53.7 (48)	58.3 (34)	60.0 (50)	<b>70.8</b> (46)	68.1 (69)	40.5 (16)	47.2 (33)	<b>50.9</b> (36)	45.8 (40)	61.1 (37)	66.7 (52)	64.5 (50)	<b>71.0</b> (52)
Subject 4	46.7 (23)	52.1 (35)	<b>54.8</b> (39)	54.1 (46)	56.1 (33)	57.4 (40)	56.3 (54)	<b>58.8</b> (58)	58.3 (35)	58.3 (37)	<b>65.1</b> (41)	63.1 (45)	71.9 (49)	70.8 (50)	<b>71.9</b> (47)	67.5 (48)
Subject 5	43.6 (26)	49.9 (39)	<b>62.2</b> (50)	61.5 (51)	43.6 (26)	39.9 (29)	43.7 (31)	<b>45.7</b> (37)	50.1 (33)	54.2 (40)	<b>57.9</b> (41)	52.1 (47)	38.9 (22)	43.3 (29)	<b>62.8</b> (45)	61.6 (47)
Subject 6	28.0 (14)	32.8 (21)	<b>42.5</b> (25)	41.9 (27)	40.0 (26)	41.8 (30)	43.5 (34)	<b>45.2</b> (38)	39.7 (26)	45.0 (29)	<b>50.3</b> (36)	46.7 (41)	44.9 (31)	50.1 (34)	<b>57.6</b> (43)	57.0 (44)
Subject 7	35.6 (18)	38.3 (27)	<b>48.7</b> (38)	46.7 (35)	53.7 (36)	47.9 (36)	44.5 (38)	<b>47.8</b> (44)	48.2 (30)	50.3 (35)	50.9 (39)	<b>51.4</b> (46)	65.2 (47)	68.3 (53)	70.8 (59)	<b>77.9</b> (61)
Subject 8	42.4 (21)	37.0 (32)	<b>43.7</b> (37)	42.9 (41)	48.1 (27)	50.0 (45)	52.8 (50)	<b>53.3</b> (60)	45.6 (24)	44.6 (35)	42.9 (35)	<b>43.3</b> (41)	56.1 (35)	59.3 (50)	<b>52.4</b> (45)	46.6 (47)
Subject 9	38.3 (19)	45.8 (29)	47.5 (35)	<b>48.4</b> (37)	45.0 (26)	46.6 (30)	50.2 (53)	<b>58.8</b> (56)	58.2 (39)	58.1 (37)	<b>60.6</b> (36)	59.9 (41)	55.8 (37)	54.2 (33)	<b>62.4</b> (38)	61.9 (40)
Subject 10	24.1 (12)	32.6 (18)	<b>38.6</b> (33)	36.8 (23)	38.6 (27)	40.5 (26)	39.8 (38)	<b>42.1</b> (37)	41.7 (30)	48.0 (30)	<b>47.0</b> (40)	46.5 (46)	48.1 (36)	31.0 (12)	<b>54.2</b> (47)	53.7 (50)
Subject 11	36.7 (18)	38.1 (28)	48.1 (33)	<b>49.7</b> (36)	38.8 (20)	41.7 (31)	<b>45.3</b> (39)	44.5 (39)	42.9 (25)	46.2 (32)	<b>51.9</b> (41)	50.8 (47)	54.8 (36)	58.3 (44)	54.2 (43)	<b>59.6</b> (46)
Subject 12	34.2 (17)	37.1 (26)	<b>40.1</b> (32)	36.2 (33)	36.1 (19)	39.0 (27)	42.7 (44)	<b>45.7</b> (44)	37.5 (20)	47.5 (32)	<b>47.0</b> (43)	46.5 (50)	40.0 (33)	47.5 (32)	<b>53.9</b> (50)	51.2 (53)
Subject 13	27.6 (14)	28.6 (21)	33.9 (25)	<b>34.8</b> (27)	39.6 (26)	43.1 (35)	44.0 (43)	<b>45.8</b> (46)	43.3 (30)	50.0 (38)	50.5 (44)	<b>51.2</b> (51)	46.3 (33)	56.3 (44)	<b>55.6</b> (49)	54.5 (52)
Subject 14	44.8 (22)	55.6 (34)	<b>61.0</b> (40)	57.8 (44)	52.8 (30)	56.8 (35)	56.2 (31)	<b>60.6</b> (54)	61.0 (39)	60.3 (34)	63.9 (34)	<b>70.2</b> (39)	63.9 (42)	66.7 (41)	<b>58.3</b> (42)	57.8 (31)
Subject 15	40.7 (20)	35.7 (31)	46.9 (36)	<b>49.7</b> (40)	58.3 (38)	63.9 (46)	70.2 (48)	<b>70.8</b> (51)	73.8 (53)	71.0 (49)	70.7 (43)	<b>77.6</b> (50)	58.3 (38)	61.9 (40)	<b>66.7</b> (41)	65.9 (42)
Average	38.5 (19)	41.6 (29)	<b>48.1</b> (36)	47.9 (38)	47.3 (28)	48.7 (36)	50.7 (42)	<b>52.9</b> (48)	50.6 (31)	53.8 (37)	56.2 (39)	<b>56.5</b> (45)	54.8 (35)	57.1 (39)	<b>60.8</b> (45)	60.3 (46)

feature mapping scheme, the PM7 model reported the best classification performance (ACCavg = 54.68% and Rf > 37%) in the case of the SqueezeNet feature. The inferior result was reported by PM5 (ACCavg = 48.07% and Rf > 36%) in the case of PM1 with the ELM feature. In the case of BCI Competition dataset, computed deep features employing ELM, AlexNet, SqueezeNet, and ResNet networks had the highest average accuracies across all proposed models, with PM1 (ACCavg = 53.08% and Rf > 43%), PM2 (ACCavg = 56.47% and Rf > 38%), PM7 (ACCavg = 59.07% and Rf > 46%), and PM4 (ACCavg = 57.75% and Rf > 41%), respectively. Overall, the proposed models PM4-ResNet (ACCavg = 60.80% and Rf > 45%) and PM7-SqueezeNet (ACCavg = 59.07% and Rf > 46%) demonstrated superior performance with reduced features for classifying 5-class and 6-class IMG-EEG

imagined signals in both datasets [17] and [119], respectively.

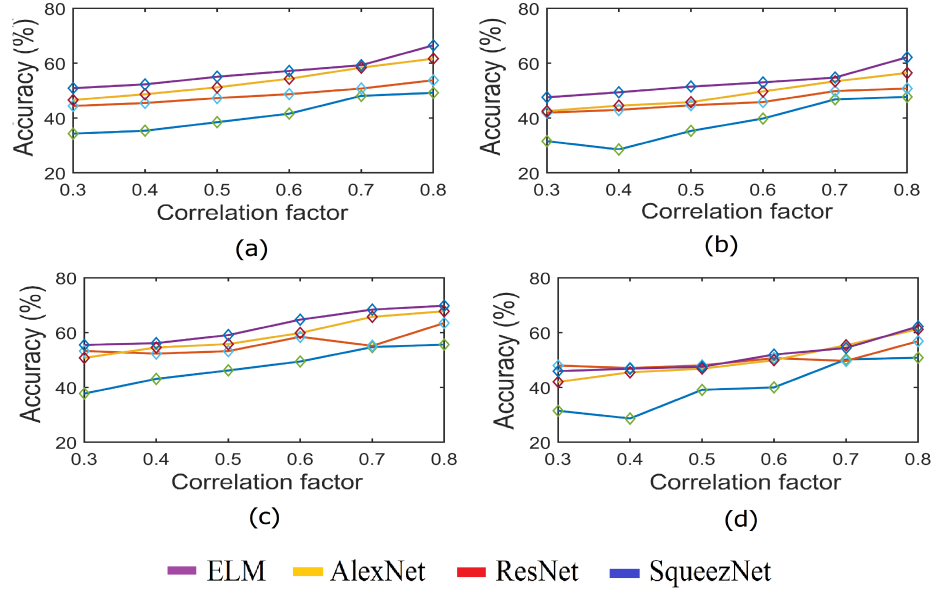


Figure 4.9: Cross-validation of average accuracy plotted against each CF value for four selected deep features in case of (a) CCA with Coretto dataset, (b) HDC with Coretto dataset, (c) CCA with BCI Competition dataset, and (d) HDC with BCI Competition dataset.

Further, we evaluate the proposed model's performance and complexity with different deep networks to determine its effectiveness with deep features. The used ELM model exhibits comparatively shallow layers with less computation time compared to the other dense classifier models, AlexNet, SqueezeNet, and ResNet. In the subsequent experiments, we configure a few ELM layers (refer Table 4.2) in the proposed models PM1 and PM5, which achieves more stable performance on both datasets according to Tables 4.3, 4.4, 4.5, and 4.6. In the Coretto dataset, even though ResNet has delivered the highest accuracy with PM4 ( $ACC_{avg} = 60.80\%$  and  $R_f > 45\%$ ), the ELM shows significant performance by delivering competent performance PM1 ( $ACC_{avg} = 48.07\%$  and  $R_f > 36\%$ ) with fewer layer configuration. Similarly, the PM1 model achieved the highest classification performance ( $ACC_{avg} = 53.10\%$  and  $R_f > 43\%$ ) using the ELM feature in the BCI Competition dataset. The experimental findings demonstrate that achieving accurate prediction of imagined word speech with reasonably high accuracy ( $ACC_{avg}$ ) is possible using fewer dense layers on the obtained JTFDF features by MSSDM, thereby reducing computational complexity. To examine the potential selection of JTFDF features in the proposed classification models,

# CHAPTER 4. MULTIVARIATE SSDM-BASED JOINT TIME-FREQUENCY ANALYSIS WITH IMPROVED IMAGINED SPEECH DECODING FOR INTUITIVE BCI

Table 4.4: Average accuracy score (in %) obtained from reduced features (in %) for four different pre-trained neural networks using HDC scheme and SVM classifier on the Correto dataset.

Hellinger distance-based feature selection: ACCavg (Rf)																
CF →	ELM				AlexNet				SqueezeNet				ResNet			
	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8
Subject 1	32.3 (16)	38.2 (24)	38.9 (38)	<b>40.8</b> (42)	37.0 (21)	39.6 (26)	<b>44.3</b> (42)	40.8 (44)	0.6 (25)	44.2 (30)	46.2 (38)	<b>48.1</b> (39)	35.5 (19)	42.8 (29)	38.9 (31)	<b>51.9</b> (34)
Subject 2	45.8 (23)	46.5 (34)	<b>47.5</b> (50)	47.1 (55)	57.2 (34)	56.7 (42)	51.2 (48)	<b>52.1</b> (60)	64.6 (42)	64.9 (50)	<b>65.9</b> (57)	64.2 (58)	63.9 (41)	72.0 (57)	<b>75.7</b> (35)	72.6 (39)
Subject 3	44.9 (22)	42.0 (34)	<b>48.3</b> (42)	46.9 (54)	55.1 (33)	56.7 (48)	<b>66.9</b> (61)	62.3 (68)	46.2 (24)	53.2 (45)	54.9 (43)	<b>55.4</b> (43)	57.1 (35)	62.3 (54)	60.3 (48)	<b>64.2</b> (53)
Subject 4	42.7 (21)	50.5 (32)	<b>58.7</b> (51)	57.6 (52)	53.0 (32)	54.2 (36)	53.1 (45)	<b>55.5</b> (48)	52.2 (31)	58.3 (40)	61.1 (47)	<b>61.7</b> (48)	71.0 (50)	66.2 (48)	69.4 (37)	<b>74.0</b> (41)
Subject 5	39.8 (20)	42.4 (30)	<b>45.8</b> (48)	42.6 (49)	41.2 (21)	37.3 (19)	41.3 (27)	<b>43.1</b> (37)	44.8 (25)	47.5 (29)	<b>48.7</b> (27)	48.2 (28)	36.3 (16)	40.5 (22)	38.9 (19)	<b>53.1</b> (20)
Subject 6	25.6 (13)	31.8 (19)	<b>39.3</b> (32)	39.2 (35)	37.8 (25)	39.0 (26)	<b>44.0</b> (36)	42.7 (44)	45.5 (33)	45.1 (33)	<b>47.5</b> (34)	45.2 (35)	41.9 (29)	45.8 (33)	<b>50.8</b> (37)	49.0 (41)
Subject 7	32.5 (16)	37.1 (24)	41.8 (37)	<b>42.7</b> (42)	50.7 (34)	44.8 (32)	<b>58.1</b> (53)	56.5 (61)	51.4 (35)	50.9 (38)	49.8 (39)	<b>50.3</b> (40)	60.9 (45)	62.5 (50)	64.7 (40)	<b>65.3</b> (44)
Subject 8	38.8 (19)	35.9 (29)	39.5 (36)	<b>43.2</b> (48)	45.5 (26)	46.7 (40)	<b>49.9</b> (46)	48.1 (60)	40.8 (21)	38.9 (32)	46.8 (37)	<b>48.6</b> (38)	52.4 (33)	54.2 (47)	47.9 (38)	<b>53.0</b> (44)
Subject 9	35.1 (18)	44.5 (26)	45.6 (44)	<b>47.3</b> (44)	42.5 (25)	43.5 (25)	52.7 (39)	<b>55.6</b> (45)	52.1 (35)	38.2 (20)	<b>47.8</b> (28)	47.2 (28)	52.1 (35)	49.5 (31)	57.0 (37)	<b>65.4</b> (42)
Subject 10	22.0 (11)	31.6 (17)	<b>34.9</b> (32)	33.7 (31)	36.5 (25)	37.8 (23)	<b>37.6</b> (34)	36.1 (46)	37.3 (26)	40.5 (25)	47.0 (38)	<b>48.9</b> (38)	45.0 (34)	28.3 (13)	48.8 (39)	<b>50.6</b> (45)
Subject 11	35.0 (17)	37.0 (26)	43.5 (37)	<b>45.4</b> (44)	36.6 (19)	39.4 (29)	<b>41.9</b> (35)	40.1 (49)	38.3 (21)	40.0 (29)	44.3 (32)	<b>46.0</b> (32)	51.2 (34)	53.3 (43)	49.5 (37)	<b>50.2</b> (42)
Subject 12	31.2 (16)	41.0 (23)	<b>43.3</b> (41)	42.5 (40)	34.1 (18)	36.8 (19)	39.5 (37)	<b>40.2</b> (49)	33.5 (18)	34.8 (17)	<b>46.4</b> (38)	43.5 (39)	46.7 (31)	44.4 (27)	<b>50.4</b> (42)	49.0 (48)
Subject 13	25.2 (13)	26.2 (19)	33.6 (26)	<b>38.8</b> (35)	37.4 (25)	39.7 (33)	<b>44.9</b> (37)	40.3 (47)	38.8 (26)	44.1 (37)	<b>44.8</b> (31)	43.0 (32)	43.3 (31)	52.6 (45)	52.0 (39)	<b>55.4</b> (44)
Subject 14	40.9 (20)	50.8 (31)	<b>54.2</b> (35)	52.9 (43)	49.9 (29)	53.7 (34)	<b>57.0</b> (45)	54.3 (46)	54.6 (34)	50.6 (31)	<b>53.1</b> (35)	50.9 (35)	59.7 (39)	62.3 (42)	<b>54.5</b> (36)	52.6 (41)
Subject 15	37.3 (19)	32.7 (28)	40.6 (33)	<b>45.4</b> (46)	55.1 (36)	60.3 (45)	<b>64.9</b> (57)	61.9 (68)	63.0 (44)	65.2 (50)	<b>68.8</b> (55)	62.2 (56)	54.5 (36)	57.9 (43)	<b>62.3</b> (48)	60.6 (55)
Average	35.3 (17)	39.8 (26)	44.7 (38)	<b>45.5</b> (43)	44.6 (27)	45.8 (31)	<b>49.8</b> (42)	48.6 (51)	46.9 (29)	47.8 (34)	<b>51.5</b> (39)	50.9 (39)	51.4 (34)	53.0 (38)	54.7 (37)	<b>58.1</b> (42)

we have compared each of the models on feature selection to different CF values. Figure 4.9 (a)-(d) compares cross-validation classification accuracy using two feature correlation schemes (CCA and HDC) with varying CFs (0.3 to 0.8) for two datasets. The figures show that accuracy increases with higher CF values but also increases feature complexity. Therefore, based on the experiment, a lower CF of 0.7 was selected to reduce feature complexity without compromising the accuracy performance. Considering the performances depicted in Tables 4.3, 4.4, 4.5, 4.6, and Fig. 4.9 (a)-(d), it is observed that among the proposed model, the PM1-ELM (ACCavg = 48.07%, Rf > 36%, CF = 0.7) with CCA and PM5-ELM (ACCavg = 50.18%, Rf > 45%, CF = 0.7) with HDC are found to be the best model with reduced features for classifying IMG-EEG signals, which can yield the highest perfor-

CHAPTER 4. MULTIVARIATE SSDM-BASED JOINT TIME-FREQUENCY ANALYSIS WITH IMPROVED IMAGINED SPEECH DECODING FOR INTUITIVE BCI

Table 4.5: Average accuracy score (in %) obtained from reduced features (in %) for four different pre-trained neural networks using CCA scheme and SVM classifier on the BCI Competition dataset.

Canonical correlation analysis-based feature selection: ACCavg (Rf)																
CF →	ELM				AlexNet				SqueezeNet				ResNet			
	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8
Subject 1	42.4 (21)	46.0 (32)	52.8 (46)	<b>53.6</b> (49)	47.1 (26)	50.3 (36)	47.9 (41)	<b>54.7</b> (46)	38.9 (34)	40.9 (27)	<b>41.1</b> (30)	40.4 (35)	39.7 (32)	43.2 (29)	44.4 (33)	<b>45.1</b> (39)
Subject 2	60.1 (30)	64.4 (45)	<b>64.7</b> (49)	63.2 (50)	52.6 (43)	60.1 (40)	<b>65.3</b> (36)	65.0 (38)	70.1 (40)	76.4 (44)	<b>76.0</b> (34)	75.9 (40)	65.1 (41)	72.6 (40)	<b>78.5</b> (40)	74.7 (46)
Subject 3	58.9 (29)	56.1 (44)	<b>65.6</b> (41)	63.0 (45)	52.1 (41)	60.5 (47)	<b>72.3</b> (34)	70.6 (48)	62.7 (33)	68.3 (42)	66.6 (33)	<b>70.4</b> (38)	64.0 (35)	62.8 (36)	<b>66.4</b> (32)	66.3 (38)
Subject 4	56.0 (28)	56.9 (42)	59.2 (42)	<b>59.5</b> (43)	51.9 (33)	54.2 (41)	<b>60.4</b> (29)	58.2 (32)	70.1 (42)	76.4 (41)	<b>73.5</b> (29)	72.4 (34)	61.2 (44)	62.1 (43)	70.8 (45)	<b>71.7</b> (53)
Subject 5	52.3 (26)	59.3 (39)	57.8 (44)	<b>60.5</b> (39)	47.5 (21)	47.9 (28)	<b>50.6</b> (23)	48.3 (29)	39.9 (14)	43.5 (23)	44.0 (34)	<b>46.9</b> (40)	40.7 (26)	40.8 (21)	<b>44.7</b> (34)	42.1 (40)
Subject 6	33.6 (17)	48.6 (25)	48.9 (34)	<b>50.1</b> (35)	43.6 (27)	50.1 (27)	<b>54.9</b> (26)	53.4 (28)	46.0 (29)	50.2 (27)	57.0 (40)	<b>61.8</b> (47)	47.0 (30)	46.2 (23)	46.1 (33)	<b>48.6</b> (39)
Subject 7	42.7 (21)	45.5 (32)	47.3 (45)	<b>48.3</b> (49)	58.5 (37)	57.5 (44)	56.3 (30)	<b>60.6</b> (42)	56.8 (45)	60.8 (55)	67.8 (39)	<b>66.8</b> (46)	56.2 (41)	60.3 (50)	62.2 (42)	<b>64.6</b> (48)
Subject 8	46.9 (23)	44.0 (35)	<b>47.7</b> (44)	46.6 (46)	42.5 (29)	50.3 (51)	<b>53.8</b> (36)	50.3 (46)	47.6 (34)	49.3 (50)	51.4 (29)	<b>59.1</b> (30)	48.7 (35)	49.6 (46)	<b>54.1</b> (32)	52.2 (36)
Subject 9	46.0 (23)	49.5 (35)	50.0 (49)	<b>50.3</b> (52)	49.1 (26)	55.9 (41)	<b>60.9</b> (46)	58.9 (51)	57.2 (34)	59.0 (44)	<b>63.0</b> (44)	61.9 (45)	54.4 (35)	57.6 (45)	<b>59.5</b> (46)	58.7 (52)
Subject 10	28.9 (14)	38.7 (22)	<b>41.3</b> (39)	40.5 (43)	42.1 (28)	48.6 (32)	<b>47.2</b> (44)	46.8 (38)	49.4 (35)	50.9 (34)	53.9 (47)	<b>57.8</b> (48)	50.4 (36)	50.5 (33)	49.1 (42)	<b>50.1</b> (47)
Subject 11	44.0 (22)	50.3 (33)	<b>54.9</b> (50)	52.6 (55)	46.5 (25)	50.0 (33)	53.5 (49)	<b>53.5</b> (54)	50.2 (28)	51.3 (34)	<b>52.7</b> (44)	52.2 (44)	50.1 (35)	53.8 (36)	52.5 (43)	<b>52.8</b> (49)
Subject 12	41.0 (20)	48.2 (31)	47.1 (48)	<b>51.8</b> (53)	43.3 (23)	46.8 (29)	<b>48.9</b> (50)	47.6 (55)	51.3 (31)	52.3 (35)	<b>56.6</b> (54)	54.8 (55)	52.3 (32)	52.9 (35)	56.0 (50)	<b>58.9</b> (57)
Subject 13	33.1 (17)	40.0 (25)	40.9 (40)	<b>41.7</b> (44)	43.1 (27)	48.7 (36)	<b>50.0</b> (51)	48.3 (56)	47.5 (31)	58.1 (43)	<b>57.4</b> (52)	56.7 (53)	48.5 (32)	50.8 (39)	<b>54.8</b> (49)	53.1 (56)
Subject 14	53.7 (27)	60.7 (40)	<b>64.5</b> (39)	63.4 (43)	48.2 (31)	62.1 (42)	<b>67.5</b> (43)	66.1 (47)	49.7 (33)	50.9 (35)	<b>57.7</b> (29)	55.0 (29)	48.6 (34)	56.3 (37)	<b>63.6</b> (45)	60.6 (51)
Subject 15	48.9 (24)	42.5 (37)	55.2 (42)	<b>57.6</b> (48)	58.2 (39)	62.1 (39)	<b>66.9</b> (40)	64.3 (42)	59.8 (35)	61.0 (36)	<b>63.6</b> (46)	62.2 (47)	49.2 (37)	58.3 (42)	63.8 (47)	<b>64.0</b> (52)
Average	45.9 (23)	48.4 (34)	49.1 (43)	<b>50.6</b> (46)	48.8 (30)	54.0 (37)	<b>56.5</b> (38)	56.1 (43)	54.5 (33)	57.8 (38)	58.6 (39)	<b>59.5</b> (42)	53.1 (35)	55.4 (37)	<b>57.8</b> (41)	57.4 (47)

mance in datasets [17] and [119], respectively. Furthermore, to extend the significance of the proposed models, we have also conducted an additional study, which involved generating a confusion matrix to classify IMG-EEG imagined word classes across all fifteen cross-subjects, as illustrated in Fig. 4.8. Our findings indicate that PM1-ELM (in dataset [17]) is more accurate at recognizing imagined word classes (Down and Right) than PM2-AlexNet and PM3-QueueNet, however, the other models (PM2-AlexNet and PM3-QueueNet) are more accurate at recognizing other classes (Up, Forward, and Backword). A similar case was demonstrated with another dataset [119], shown in Fig. 4.8. The proposed MSSDM-JTFDF models have exhibited distinct feature discriminability and significantly improved classification performance across different imagined word recognition classes. Moreover,

# CHAPTER 4. MULTIVARIATE SSDM-BASED JOINT TIME-FREQUENCY ANALYSIS WITH IMPROVED IMAGINED SPEECH DECODING FOR INTUITIVE BCI

Table 4.6: Average accuracy score (in %) obtained from reduced features (in %) for four different pre-trained neural networks using HDC scheme and SVM classifier on the BCI Competition dataset.

Hellinger distance-based feature selection: ACCavg (Rf)																
CF →	ELM				AlexNet				SqueezeNet				ResNet			
	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8	0.5	0.6	0.7	0.8
Subject 1	35.4 (18)	45.2 (27)	<b>49.1</b> (45)	48.8 (48)	42.4 (25)	45.3 (27)	43.1 (39)	<b>48.8</b> (44)	41.4 (24)	51.9 (25)	<b>57.1</b> (48)	54.9 (49)	34.5 (17)	41.7 (23)	45.3 (36)	<b>50.3</b> (37)
Subject 2	50.2 (25)	59.8 (38)	<b>63.5</b> (60)	62.1 (63)	55.5 (40)	58.8 (43)	60.2 (48)	<b>64.2</b> (51)	65.7 (41)	71.4 (49)	<b>78.0</b> (69)	77.2 (70)	62.1 (37)	66.4 (48)	69.3 (60)	<b>70.1</b> (61)
Subject 3	49.1 (25)	50.9 (37)	<b>61.2</b> (51)	60.0 (54)	58.0 (38)	62.8 (51)	<b>65.0</b> (55)	64.1 (52)	56.8 (32)	62.9 (49)	68.4 (53)	<b>70.3</b> (54)	55.6 (31)	56.5 (47)	<b>59.6</b> (44)	59.2 (45)
Subject 4	46.7 (23)	50.2 (35)	55.1 (50)	<b>56.0</b> (55)	55.4 (32)	58.0 (47)	60.2 (51)	<b>62.4</b> (52)	53.0 (30)	63.0 (48)	70.2 (65)	<b>72.1</b> (66)	45.9 (43)	54.4 (49)	<b>58.2</b> (47)	57.4 (49)
Subject 5	43.6 (22)	50.1 (33)	<b>55.7</b> (50)	55.6 (55)	42.3 (21)	43.1 (26)	<b>46.3</b> (40)	45.6 (48)	45.5 (24)	55.7 (38)	63.1 (52)	<b>64.4</b> (53)	33.7 (12)	42.4 (25)	<b>54.7</b> (43)	54.3 (44)
Subject 6	28.1 (14)	40.8 (21)	45.5 (41)	<b>46.1</b> (45)	38.8 (25)	45.1 (25)	<b>48.9</b> (44)	47.5 (53)	36.1 (22)	52.0 (32)	57.0 (47)	<b>61.2</b> (48)	38.9 (25)	45.6 (26)	50.1 (40)	<b>54.6</b> (41)
Subject 7	35.6 (18)	40.5 (27)	<b>45.2</b> (40)	44.5 (44)	49.6 (32)	51.8 (38)	52.9 (48)	<b>54.1</b> (58)	46.2 (28)	49.3 (36)	48.4 (38)	<b>50.9</b> (39)	49.6 (40)	55.3 (48)	<b>60.4</b> (50)	57.3 (51)
Subject 8	38.5 (21)	42.9 (32)	<b>44.4</b> (43)	43.9 (47)	46.7 (25)	47.4 (36)	48.5 (47)	<b>52.1</b> (56)	41.5 (20)	44.2 (33)	44.8 (38)	<b>45.3</b> (38)	44.4 (23)	53.9 (43)	53.6 (46)	<b>55.1</b> (48)
Subject 9	40.5 (20)	44.7 (30)	<b>46.5</b> (45)	46.2 (49)	43.7 (23)	44.2 (30)	53.4 (52)	<b>62.8</b> (61)	50.1 (30)	52.6 (38)	<b>53.9</b> (42)	53.5 (47)	56.2 (36)	59.2 (50)	<b>63.3</b> (53)	62.6 (54)
Subject 10	25.4 (13)	35.0 (19)	<b>38.4</b> (35)	36.9 (37)	37.9 (25)	38.4 (22)	42.3 (39)	<b>45.0</b> (46)	42.4 (30)	40.8 (25)	44.8 (33)	<b>45.7</b> (36)	39.9 (27)	45.1 (29)	46.9 (35)	<b>52.5</b> (36)
Subject 11	38.7 (19)	40.3 (29)	<b>51.0</b> (40)	47.9 (43)	41.9 (23)	39.5 (28)	<b>48.2</b> (37)	47.6 (44)	39.0 (20)	48.2 (37)	52.9 (34)	<b>53.5</b> (37)	49.4 (30)	50.2 (42)	52.6 (33)	<b>54.9</b> (34)
Subject 12	36.1 (18)	42.5 (27)	<b>46.9</b> (43)	46.1 (45)	39.9 (22)	40.9 (21)	<b>49.6</b> (45)	48.8 (54)	34.1 (16)	44.4 (29)	<b>54.3</b> (41)	53.9 (45)	41.9 (24)	43.2 (28)	<b>50.3</b> (37)	48.2 (38)
Subject 13	29.1 (15)	36.9 (22)	39.6 (37)	<b>41.1</b> (43)	39.7 (25)	46.5 (32)	46.7 (44)	<b>49.5</b> (52)	39.4 (25)	46.7 (32)	52.8 (41)	<b>53.1</b> (42)	42.1 (28)	51.1 (36)	56.0 (45)	<b>58.4</b> (46)
Subject 14	44.8 (22)	47.5 (34)	59.4 (48)	<b>61.2</b> (55)	52.9 (31)	61.4 (47)	60.8 (49)	<b>65.5</b> (58)	50.5 (33)	54.9 (41)	<b>64.9</b> (44)	62.0 (45)	58.1 (36)	60.6 (47)	62.5 (42)	<b>65.9</b> (43)
Subject 15	45.6 (23)	50.3 (34)	52.0 (50)	<b>53.0</b> (58)	61.0 (38)	67.2 (39)	<b>69.3</b> (42)	60.3 (49)	61.0 (38)	64.5 (48)	<b>72.4</b> (50)	72.2 (50)	50.9 (30)	52.3 (40)	<b>56.7</b> (44)	55.1 (46)
Average	39.4 (20)	45.2 (29)	<b>50.2</b> (45)	49.8 (49)	48.0 (28)	50.6 (34)	52.2 (45)	<b>54.6</b> (52)	47.2 (28)	53.0 (37)	59.1 (46)	<b>59.7</b> (48)	48.3 (29)	53.6 (39)	<b>55.9</b> (44)	56.7 (45)

the classification rate for other classes was also improved to a certain extent in comparison to existing work [18, 19]. Thus, these results reveal that the proposed MSSDM-based models for IMS word class exhibited significantly higher performance as compared to the existing approach [17, 18, 19, 126, 127] with reducing feature space, indicating the effectiveness of the proposed framework for a feasible solution.

## 4.5 Discussions

The effectiveness of the proposed method was verified by comparing its performance to state-of-the-art works on the same datasets. Our demonstrated results, shown in Ta-

ble 4.7, were in line with the neural representation discovered in the previous works [17, 18, 19, 126, 127] and have been tested with the same experimental paradigm, including the number of channel selection and formulation of deep features of performing the same imagined word classes in both datasets. To recognize imagined words in various IMG-EEG datasets, several feature extraction approaches were combined with several machine learning approaches. Previous studies in IMS classification on the Correto dataset EEG signals achieved the highest accuracy rates of 17.46% with DWT-based bands in 6-class (Coretto et al. [17]), 18.89% with CNN deep feature in 5-word class (Cooney et al. [126]), 19.81% with deep ConvNets and FBCSP in three-class [127], and 22.01% with the BoF approach method (García-Salinas et al. [18]) in the 5-class. For further improvement, Lee et al. [19] have achieved accuracies of 38.67% (Correto dataset) and 37.33% (BCI Competition dataset) by exploring IF and SE features with the SNN. However, compared to [19], our MSSDM-based proposed PM1 and PM7 models have demonstrated improved classification performance. The best accuracies were achieved with PM4-ResNet ( $ACC_{avg} = 60.80\%$ ,  $Rf > 45\%$ ,  $CF = 0.7$ ) and PM7-SqueezeNet ( $ACC_{avg} = 59.07\%$ ,  $Rf > 46\%$ ,  $CF = 0.7$ ) on the datasets [17] and [119], respectively. Whereas, some proposed models (PM5-ELM on

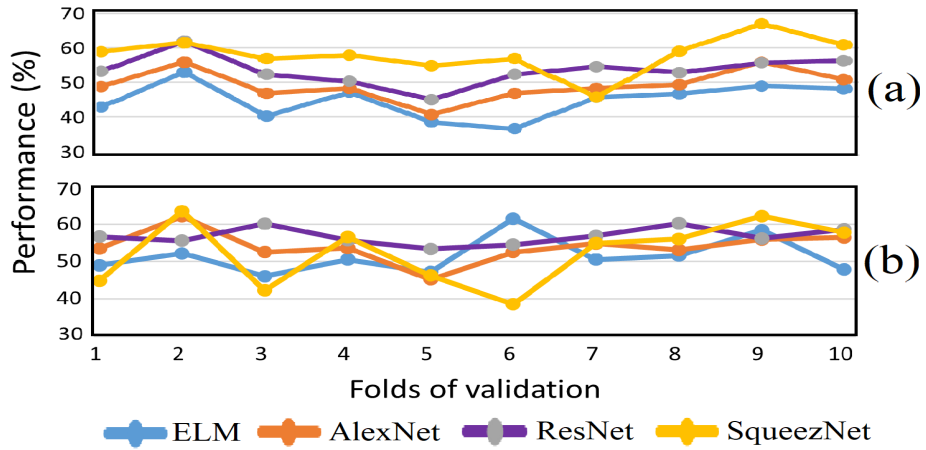


Figure 4.10: k-fold cross validation across all subjects in CCA approach on (a) the Correto dataset and (b) the BCI Competition dataset.

Correto dataset) could not discriminate between two imagery classes and delivered slightly inferior results for Lee et al. [19] in BCI Competition dataset. This can be improved by increasing the size of epochs in each trial with different word choices. Overall, our MSSDM-

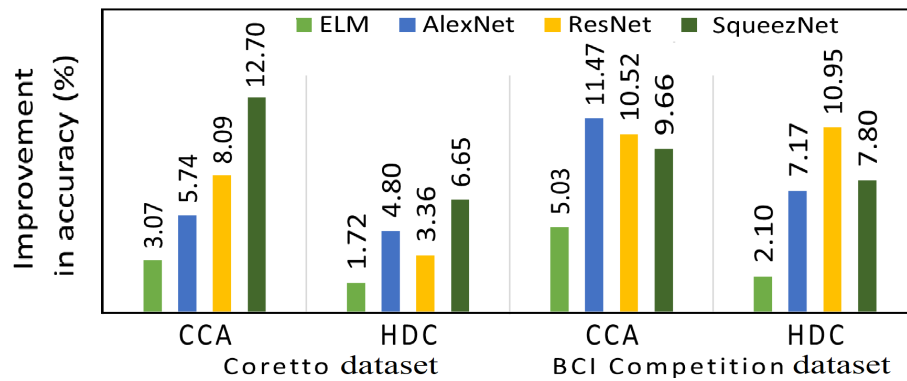


Figure 4.11: Comparison of improvement of averaged decoding accuracies (in %) during imagination across all subjects.

Table 4.7: Performance comparison with the other state-of-the-art methods.

Dataset	Authors and reference	Methodology	Accuracy (%)	<i>p</i> -value
Coretto dataset [17]	Coretto et al. [17]	DWT, RF DB	$17.46 \pm 0.75$	$< 0.01$
	Cooney et al. [126]	CNN	$18.89 \pm 1.41$	$< 0.01$
	Schirrmester et al. [127]	Deep ConvNets, FBCSP	$19.81 \pm 2.10$	$< 0.01$
	García-Salinas et al. [18]	BoF	$22.01 \pm 2.33$	$< 0.01$
	Lee et al. [19]	IF and SE features, SNN, SVM	$45.00 \pm 3.13$	$< 0.01$
	Proposed method-1	MSSDM-ELM-JTFDF-CCA (SVM)	$48.09 \pm 3.13$	$< 0.01$
	Proposed method-2	MSSDM-ResNet-JTFDF-CCA (SVM)	$60.80 \pm 2.17$	$< 0.01$
BCI Competition dataset [119]	Coretto et al. [17]	DWT, RF	$37.24 \pm 6.24$	$< 0.01$
	Cooney et al. [126]	CNN	$19.79 \pm 1.45$	$< 0.01$
	Schirrmester et al. [127]	Deep ConvNets, FBCSP	$20.06 \pm 1.05$	$< 0.01$
	García-Salinas et al. [18]	BoF	$23.26 \pm 1.66$	$< 0.01$
	Lee et al. [19]	IF and SE features, SNN, SVM	$48.10 \pm 3.68$	$< 0.01$
	Proposed method-1	MSSDM-ELM-JTFDF-CCA (SVM)	$53.14 \pm 3.68$	$< 0.01$
	Proposed method-2	MSSDM-SqueezeNet-JTFDF-HDC (SVM)	$59.07 \pm 8.26$	$< 0.01$

DWT: Discrete wavelet transform, RF: Random forest, FBCSP: Filter bank common spatial patterns, BoF: Bag of features, ConvNets: Convolutional neural networks, IF: Instantaneous frequency, SE: Spectral entropy, SNN: Siamese neural network.

JTFDF feature-based models showed a significant improvement over existing methods by 12.70% and 5.03% in the case of dense and shallow features, respectively, as validated by statistical analysis (shown in Fig. 4.11).

Therefore, it is evident from this study that the proposed MSSDM method can be useful to enhance multivariate channel-aligned common features from a few channels, leading

to improved IMG-EEG imagined word multi-class classification performance. Also, our results revealed that the MSSDM method generated multivariate JTF images that precisely represented imagined word recognition, and reduced JTFDF features extracted from these JTF images were perfectly suited for imagined multi-class classification. Thus, we can state that the proposed MSSDM-JTFDF-based models deliver robust classification performance despite a large number of classes, proving the feasibility of reducing deep features and computational complexity for multi-class classification for feasible BCI systems.

## **4.6 Summary**

In this study, we have proposed a novel MSSDM-JTFDF features-based classification model for the automatic recognition of imagined word speech using multichannel IMS-EEG signals. Our approach consists of (i) implementation of the MSSDM decomposition method to extract an ensemble of CAOCs from multichannel IMS-EEG signals using improved multivariate swarm filtering and SFT spectrum, (ii) construction of multivariate JTF images using JIA and JIF functions, and (iii) formulation of JTF-based deep features using different pre-trained NN. In addition, we have proposed a feature reduction approach using CCA and HDC to formulate reduced deep features from JTF images while reducing computational complexity for multi-class classification. The study shows the efficacy of the proposed method in accurately classifying IMS across 6-class and 5-class recognition tasks, outperforming existing state-of-the-art approaches on both datasets. Moreover, the experimental outcomes indicate that the proposed MSSDM-based models achieve more efficient IMS classification, even with fewer JTFDF features and a limited number of EEG channels, despite a large number of classes. In the future, we will extend our problem to the broader range of IMS problem-based applications, optimize channels, and ultimately deploy our solution to address practical clinical systems.





# **Chapter 5**

## **Enhanced SSDM with Multi-Class Motor Imagery-based EEG-BCI System**

This chapter presents a novel decomposition technique, ESSDM to solve uniform decomposition and hyperparameter selection issues in the existing SWD method. The technique uses SHO, a modified SwF filter bank and SFT spectrum to improve hyperparameter selection and decomposition performance. It adopts a newly designed fitness function criteria based on KLD distance from spectral kurtosis for optimal mode extraction and fast convergence. The designing of ESSDM and its effectiveness over existing methods with synthetic and real EEG signals are discussed in detail in this chapter. Additionally, this proposed method is then used to develop a framework for the automatic classification of MI tasks from MI-EEG signals for BCI applications. The classification performance of the MI-EEG recognition BCI framework is detailed in later sections.

### **5.1 Introduction**

MI has developed a new intuitive paradigm in the BCI and has become an active rehabilitation tool for the treatment of paralytic stroke patients [34, 73]. MI is a cognitive process in which the subject imagines the motory action rather than performing the actual action, e.g., right-hand or left-hand. In active rehabilitation training, MI-BCI constructs a direct information transfer pathway, which converts EEG brain activity signals into MI

commands to control prosthesis instead of traditional muscle control pathways [25, 26]. In these studies [24, 25, 26, 28, 29, 30, 31, 32, 33], EEG has been widely used because of its non-invasiveness and affordability, and it offers an improved fine-grained analysis by determining spectral, spatial, and temporal components underlying MI categories. Additionally, MI-EEG has multi-class scalability, thus showing the possibility of building an extensible BCI system. However, in an MI-BCI system, accurate recognition of MI actions from EEG signals is difficult due to its limited spatial resolution, intra-channel interference, and non-stationary and nonlinearity nature; thus, it is challenging to develop a reliable EEG-based MI-BCI recognition model.

To examine the efficacy of using EEG signals, several studies have attempted to extract informative features for MI-EEG recognition systems in recent years. It majorly focuses on two categories. The first category is based on filtering-based feature extraction, such as spatial filtering and its variants [23, 24, 29, 31, 32, 73], frequency-domain features (e.g., PSD [128], Riemannian covariance [129], and RE [130]), TF methods like MEMD [131], and WT [132]. In the second category, shallow learning and deep learning methods have been employed to estimate the MI-EEG tasks with promising results based on EEG [24, 25, 28, 30, 34]. These approaches involve, spatial TF features with optimized Deep&Wide (ORDW) [28], 3D-convolutional neural network (3D-CNN) [25], frequential deep belief network (FDBN) [30], spatial filters based deep CNN [24], and multi-scale fusion attention mechanism (MS-AMF) [26], demonstrating the high discriminability of the feature for enhancement of MI-based actions. Among these approaches, the decomposition methods became crucial for MI recognition and gaining attention. However, the used methods such as MEMD [133] or EWT [134] limit the performance of extracting spectral characteristics and identifying MI state due to mode mixing issues. Whereas most of the above-reported methods are constrained by their performance. Mostly derived handcrafted features are based on shallow and deep learning methods. It strongly depends upon the formulation of classification models to find mutual feature characteristics across multiple EEG channels with optimal multi-class scalability and introduced computational complexity for real-time deployable BCI systems.

Recent advancements in modern signal processing technologies utilizing multi-scale

adaptive decomposition have already proved the potential to analyze MI-EEG signals. By virtue of their ability to adaptively decompose an original signal into several components, these methods, such as MIF [66], FBSE [53], and VMD [52], provide a useful investigation to analyze the features, particularly the weak spectral characteristics in nonlinear and nonstationary signals. However, the majority of these methods have been demonstrated to be superior in addressing both mode mixing elimination and adaptive decomposition. Yet, the challenge of tuning multiple preset parameters across a wide range severely restricts their potential uses. Among these approaches, SWD [67] has shown its effectiveness in improving decomposition adaptability and addressing mode mixing issues when analyzing nonstationary signals. However, the performance of SWD relies heavily on two predefined tuning parameters:  $P_{th}$  and  $StD_{th}$ . To ensure optimal SWD decomposition, it is imperative to meticulously choose the accurate threshold settings for each signal prior to decomposition. A few significant attempts in selecting SWD thresholds using metaheuristic optimization algorithms have been explored for TF analysis [68, 71]. These approaches have deliberately attempted in solving various complex optimization problems such as data-adaptive, fast convergence, derivation-free mechanism, and local optima avoidance to a certain extent. Moreover, methods in this literature are limited by their performance in terms of the trade-off between optimization of threshold parameters tuning and performance. Thus, there is scope for improvement related to individual threshold optimization in MI action recognition across MI-EEG channels to improve mode aliasing issues and deliver optimal mode selection.

To address this issue, we propose a novel decomposition technique, namely, the ESSDM based on SHO [135], modified SwF filter bank, and SFT spectrum to solve the issue of choice of uniform decomposition and threshold parameters in existing SWD and, further it is applied to enhance multi-class MI-EEG classification performance. ESSDM adopts improved swarm filtering to automatically deliver optimal frequency bands in SFT spectrum with appropriate threshold parameters to extract meaningful OCs of any MI-EEG signal and improve TF localization. In this approach, new fitness function criteria is designed based on the KLD distance [136] from the spectral kurtosis of all modes to select threshold parameters that optimize decomposition effect, avoid excessive iterations, and provide fast

convergence with optimal modes. The selection of the SHO optimizer is due to a newly evolved optimization technique which has proved its capability of solving many optimization problems. With a newly designed fitness function, SHO conducts an optimization approach that mimics the natural hunting mechanism and hierarchical leadership structure observed in herd hunting. SHO offers distinct advantages with its fast convergence, computational efficiency, and the minimal risk of getting trapped in local minima. The fitness function in SHO is formulated to signify the relationship between the SWD thresholds and the overall best decomposition performance of similar signals. The experimental finding reveals that the proposed approach offers a unified optimization in SWD threshold selection for different similar signals, leading to accurate MI-EEG recognition. In order to automate mutual features across multi-channel MI-EEG signals, we have introduced a new subspace clustered-based FTFG feature from computed TF information. The multi-scale fused FTFG features are automatically extracted from the ESSDM-based TF images in a categorical way, which enhances channel-aligned common information present in the obtained OCs. Finally, the performance of ESSDM-based classification performance has been evaluated using obtained FTFG features and using three computationally efficient classifiers, namely, ELM [123], compact CNN (CNet) [137], and capsule neural network (CapsNet) [138]. The novelty of this study is to propose an efficient ESSDM for optimal mode extraction and TF analysis method, and extract optimized FTFG features in order to enhance the MI-EEG recognition model for feasible BCI application. The key contributions of the presented work are given as follows.

1. A new ESSDM is proposed to enhance the optimization efficiency and accuracy of SWD to improve mode aliasing issues and deliver optimal mode selection in specific channels for MI-EEG classification enhancement.
2. Significant SHO optimizer and new fitness function criteria have been designed based on the KLD distance from the spectral kurtosis of all modes to select threshold parameters that optimize the decomposition effect, avoid excessive iterations, and provide fast convergence with optimal modes.
3. FTFG features were derived from computed TF information to find mutual spectral

information across channels for improved MI-EEG BCI classification.

4. The proposed ESSDM-based MI-EEG classification framework not only outperforms existing works in the specific-subject classification but also seamlessly adapts to cross-subject classification on two well-known MI-EEG datasets.

## 5.2 Dataset and experimental paradigm

In our study, we have tested the performance of a proposed method on two public datasets for MI-EEG signals. Details of these datasets were described as follows.

Dataset 1: This used dataset IVa is from BCI competition III [139], involving five subjects ("aa," "al," "av," "aw," and "ay") performing MI tasks (right-hand or right-foot movement). EEG data from 118 electrodes were processed using BPF for filtering noise and artifacts with a pass-band frequency range [0.05 Hz to 200 Hz]. A sampling frequency of 250 Hz was used in the experimental setup for recording. Each subject has completed 280 trials, and the duration of an imagined stimulus segment that ranged from 500 to 2500 milliseconds was used to conduct the analysis.

Dataset 2: The BCI IV-2a dataset [140] involves 4 MI classes, including left-hand, right-hand, foot, and tongue movements, and it consists of data from 9 subjects. The data were recorded at a sampling rate of 250 Hz using 22 Ag/AgCl electrodes. Each subject contributed two sessions, each comprising 288 trials, resulting in an average of 72 trials per class. Through the MI stimulus analysis, we have considered the data segment from 2 seconds to 6 seconds into each trial. To preprocess the EEG signals, a filter bank with BPFs ranging from 8 Hz to 30 Hz was applied to all recorded channels. Additionally, a notch filter at 50 Hz was applied to the EEG signals. The trial of the experimental deployment paradigm is demonstrated in Fig. 5.2.

## 5.3 Proposed methodology

The objective of this study is to develop an automated ESSDM-based MI recognition system using EEG. Fig. 5.1 shows the proposed framework for MI-EEG recognition using

the ESSDM method and all the methods involved are briefly explained in the subsequent sub-sections.

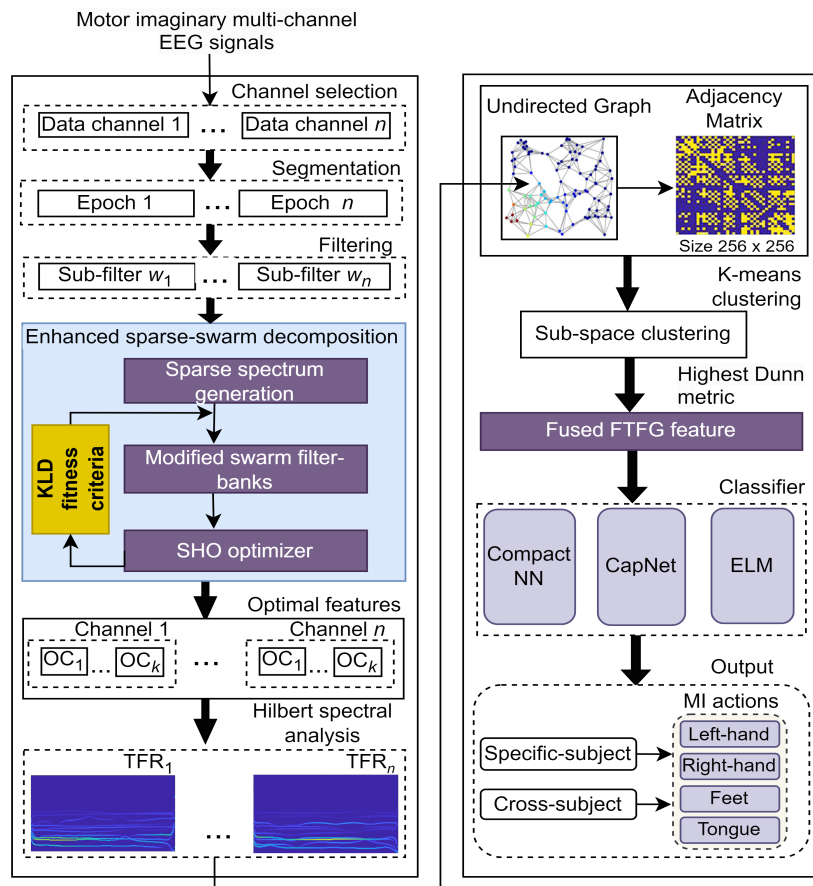


Figure 5.1: Block diagram of the proposed framework for MI-EEG recognition using ESSDM method.

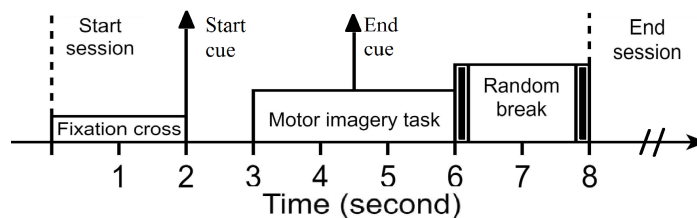


Figure 5.2: Experimental paradigm for MI-EEG recording.

### 5.3.1 Enhanced sparse swarm decomposition method

The proposed ESSDM is a novel enhanced extension of the SWD method [67] to decompose multichannel nonstationary signals into optimal OCs and improve TF analysis of nonstationary signals. The execution of the proposed ESSDM method comprises of formulation of a new SFT spectrum, design of an iterative swarm filtering method with modification of convergence criteria, and adoption of SHO with a new fitness function for optimal threshold parameter tuning that converge channel-specific optimal OCs for efficient decomposition. The motivation behind the adoption of SHO is described as follows.

SWD demonstrates superior performance in addressing mode mixing and extracting weak features from multicomponent signals. However, being a data-driven method, SWD's varied coefficients can lead to diverse decomposition outcomes, not necessarily optimized for the selection of modes related to MI-EEG separation and detection. The key issues arise from setting the Pth and StdTh, which significantly impact the decomposition performance. To understand the effect of selecting these threshold values on decomposition performance, we conducted a study on the mode reconstruction capability of existing SWD using a synthetic multicomponent AM signal. The synthetic AM signal uses five constant close frequencies with amplitude variation, represented in Eq. (2.18). Inappropriate selection of threshold parameters can result in either over-decomposition or under-decomposition, which is demonstrated in Fig. 5.3. In Fig. 5.3 (a), the threshold setting at  $P_{th} = 0.1$  and  $StdTh = 0.1$  lead to the over-decomposition results. In this,  $OC_4$  and  $OC_5$  deliver the same spectral characteristic of signal details but require extra computation to obtain these two components. Conversely, the under-decomposition result occurs in the same input signals when  $P_{th} = 0.5$  and  $StdTh = 0.5$ , which is shown in Fig. 5.3 (a). It produces a mode-mixing issue by decomposing different significant components in the single mode, thereby sacrificing significant spectral detail. In fact, setting lower thresholds in SWD yields more modes and sophisticated decomposition, but it does not guarantee better results and leads to higher computational costs. The existing recommended threshold values ( $P_{th} = 0.1$  and  $StdTh = 0.1$ ) may not be practical for all scenarios, especially in analyze the multi-spectral OCs related to MI-EEG conditions. To address this issue, the proposed ESSDM method employs



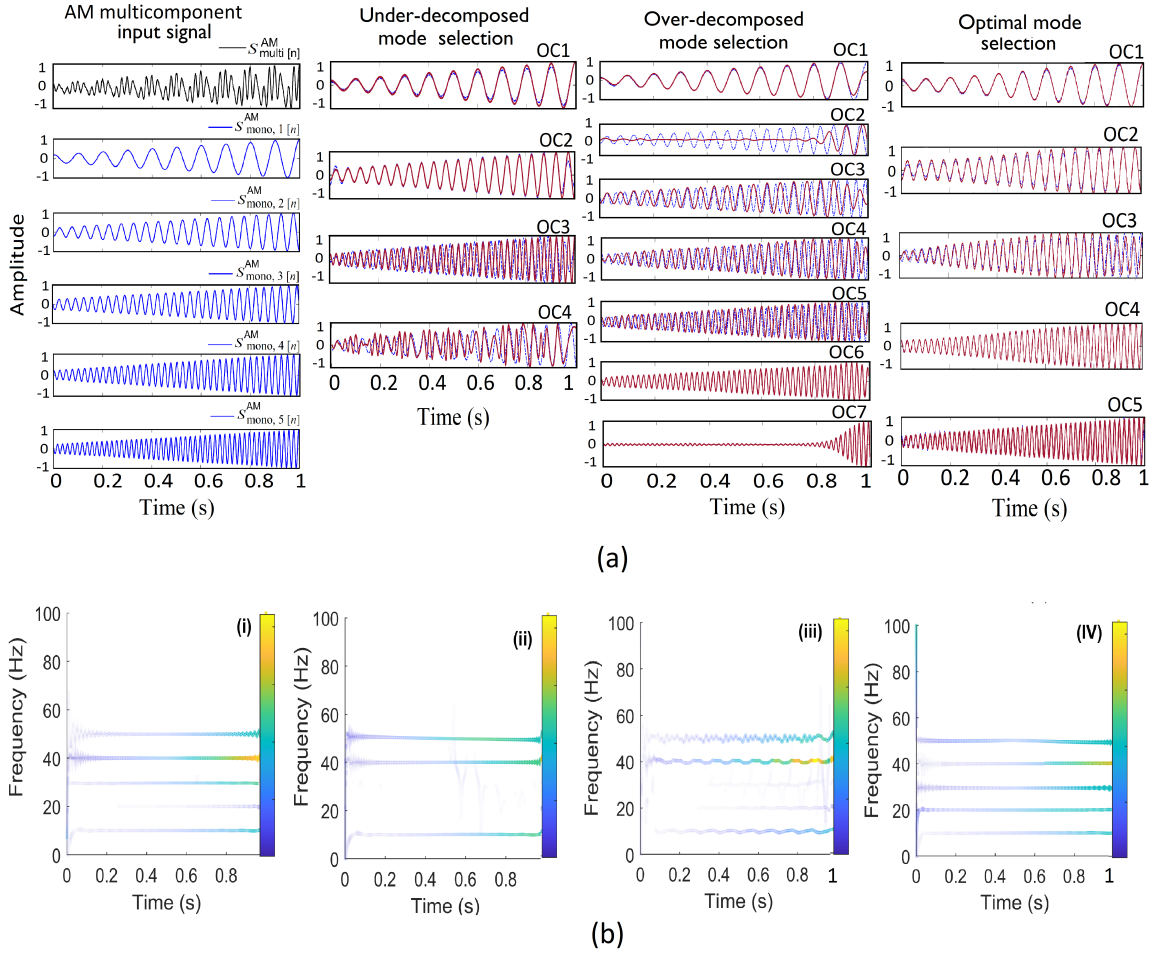


Figure 5.3: Plots of (a) SWD-based decomposed OCs with different threshold parameter selection and (b) obtained TFRs using HSA.

SHO and an appropriate optimization fitness criterion to determine suitable thresholds for separation and extraction from diverse MI-EEG multicomponent signals. The steps of the proposed SSDM to extract optimal modes from multicomponent non-stationary signal  $x(n)$  comprise three steps. Initially, SHO is designed to select the optimal threshold value from the selectable threshold range using new fitness criteria. Further, the SFT spectrum estimation model is applied to get an optimized spectrum. The detailed description of the SFT model is described in Section 2.2.1. Each obtained threshold parameter value based on the SHO is used to design a USSDM-based filter bank and extract the optimal OCs from the obtained SFT spectrum. A mathematical expression derived to obtain the decomposed OCs using SSDM is given in Section 2.2.3. The SHO with a new fitness function for optimal

threshold parameter tuning is given in detail in the next section.

### 5.3.2 Adoption of SHO for optimal threshold parameter tuning

The SHO is a biologically-inspired and population-based optimization approach with a higher capability for global optimization in comparison [135] and its parameter initialization values of SHO for Pth and StdTh are illustrated in Table 5.1. The algorithm of the proposed ESSDM method is given in Algorithm 5.1. At first, SHO is performed to calculate the optimal threshold value (Pth, StdTh) of SSDM to avoid over-decomposition or under-decomposition modes. SHO is responsible for searching this value from the selectable threshold range to determine the optimal threshold setting. In the second phase, new fitness criteria, KLD distance from spectral kurtosis (SK) of all modes, have been employed to avoid excessive iterations and provide fast convergence with optimized modes. To quantify the relationship between the decomposition performance and the threshold setting, a fitness function,  $f(\text{Pth}, \text{StdTh})$ , is designed as follows:

Table 5.1: Parameters initialization SHO optimizer for threshold parameter tuning

Parameter	Experimental value
Population size ( $N$ )	20 (max)
Size of herd ( $H$ ) and predator ( $P$ )	20
Dimensionality ( $D$ )	30
Maximum iterations	14 (maximum)
Objective function ( $f$ )	@Ackley
Search space lower bound (xd)	0.001
Search space upper bound (xu)	0.1
Exploration rate	0.1-0.5
Repulsion strength ( $\alpha$ )	0.5-2.0
Attraction strength ( $\beta$ )	0.5-2.0

$$f(\text{Pth}, \text{StdTh}) = \frac{|D_{\text{KLD}}(x_{\text{it}}, x_{\text{it}-1})| + M_{\text{it}}}{\sum_{\text{it}=1}^L |D_{\text{KLD}}(x_{\text{it}}, x_{\text{it}-1})|} \quad (5.1)$$

$$f(\text{Pth}, \text{StdTh}) = \frac{\left| D_{\text{KLD}} \left( \mathcal{P}(x_{\text{it}}) l_n \frac{\mathcal{P}(x_{\text{it}})}{\mathcal{P}(x_{\text{it}-1})} \right) \right| + \frac{\text{SK}_{M_{\text{it}}}}{\text{SK}_{R_{\text{it}}}}}{\sum_{\text{it}=1}^L \left| D_{\text{KLD}} \left( \mathcal{P}(x_{\text{it}}) l_n \frac{\mathcal{P}(x_{\text{it}})}{\mathcal{P}(x_{\text{it}-1})} \right) \right|} < 0.05 \quad (5.2)$$

$$M_{it} = \frac{1}{N} \sum_{n=1}^{N-1} \frac{|X(n, f)|^4}{(|X(n, f)|^2)^2} \begin{cases} +SK_{it} & D_{KLD} < 0.1 \\ -SK_{it} & D_{KLD} > 0.1 \end{cases} \quad (5.3)$$

$$x_j = \sum_{j=1}^{k-1} DOM_j \quad (5.4)$$

Where  $L$  denotes the number of iterations in every herd-pray optimization,  $DOM$  is the obtained dominant OCs,  $x_{it}$  is the reconstructed original signal having  $(k-1)$  OCs of the  $it_{th}$  sample length. The term  $D_{KLD}(x_{it}, x_{it-1})$  calculates the spectral distance by finding spectral entropy using criteria, KLD between adjacent OCs and converge the process when it has minimum optimal value ( $< 0.05$ ).  $M_{it}$  represents the ratio of spectral kurtosis between  $(K-1)$  modes and residual of extracted signal. In SWD, the last OC represents the residual of decomposition; thus, only the first  $(N-1)$  OCs are used for the analysis of the fitness function. In this fitness function, the selection of the lowest  $D_{KLD}$  value indicates that the extracted OCs are closer to the original signal and signifies less probability of under-decomposition condition. Whereas  $M_{it}$  controls the over-decomposition conditionally (in Eq. 5.3) by adding term  $\pm SK_{it}$  and reduces excessive iterations with fast convergence. Thus, with a newly designed fitness function, SHO is able to correlate between different threshold settings and decomposition performance, and extract optimal OCs with the selection of unique threshold values for any signal.

### 5.3.3 Comparison with existing other methods

To demonstrate the effectiveness of the ESSDM method in identifying optimal mode, synthetic multicomponent AM signals with additive uniformly distributed AWGN of SNR (-10 dB) were demonstrated and compared with decomposition results of EMD, quasi-VMD (Q-VMD), SWD, which are shown in Fig. 5.5. In Fig. 5.5 (b), it shows that the TFR obtained by HHT is overlapped and non-distinguishable. The EMD fails to decompose the signal into the original monocomponents due to the mode mixing issue. We hardly find three monocomponents ( $OC_1$ ,  $OC_2$ , and  $OC_3$ ) reconstructed properly in the obtained plot with overlapping nature. To justify the advantage in mitigating mode mixing, Q-VMD ( $k =$

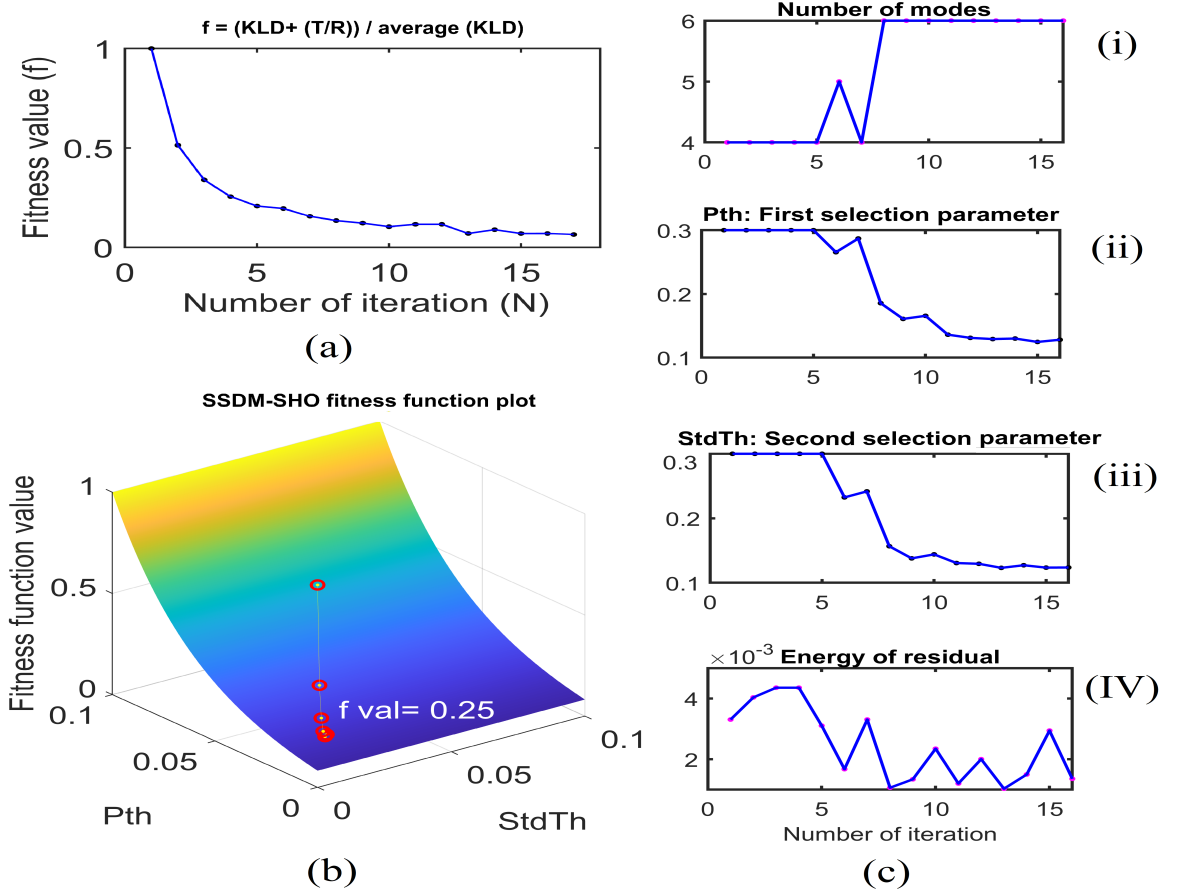


Figure 5.4: Plots of (a)-(b) convergence curves, and (c) fitness values for each iteration showing the best selection of threshold parameters ( $P_{th} = 0.18$  and  $StdTh = 0.15$ ) for optimal mode extraction using ESSDM.

5,  $\alpha = 2000$ ) is applied, and the result is presented in Fig. 5.5 (c). The first two components,  $OC_1$  and  $OC_2$ , are reconstructed properly, but noise interference distorts the  $OC_3$ ,  $OC_4$ , and  $OC_5$  components. However, even when many  $k$  and  $\alpha$ -related factors are used, it is still difficult to obtain a satisfactory result. In the case of SWD, it accurately replicates five monocomponents and shows better ability to the mode mixing against the cost of the selection of lowest threshold values ( $P_{th} = 0.01$ ,  $StdTh = 0.01$ ) with computational complexity. But it also delivers insignificant results with the selection of improper threshold parameters ( $P_{th} = 0.03$  and  $Stdth = 0.05$ ) and delivers few OCs modes accurately, as shown in Fig. 5.5 (d). In case of Q-VMD, ESR for  $OC_1$  and  $OC_3$  is greatly improved but demonstrated poor in  $OC_3$ ,  $OC_4$ , and  $OC_5$ , which is shown in Fig. 5.5 (c). On the contrary, the proposed ESSDM demonstrates all five converged components with five modes. Due to SHO-based converge

criteria and modified swarm filter bank, ESSDM optimally detects boundary frequencies in the sparse domain and delivers optimal OCs. It achieved computational efficiency by optimizing threshold value to  $P_{th} = 0.02$  and  $Std_{th} = 0.3$  with fewer average iterations (8 – 10). Table 5.2 indicates the performance of SSDM, SWD, and HHT decomposition methods. The values of ESR and MSE for the ESSDM method are the lowest as compared to the SWD and EMD methods. The value of the ESR of every extracted component for the SWD method is slightly inferior, whereas the ESR for the EMD method was the highest. The obtained ESR for the reconstructed monocomponent signals  $OC_1$ ,  $OC_2$ ,  $OC_3$ ,  $OC_4$ , and  $OC_5$  using the ESSDM is proved to be better than the non-optimized SWD method.

Table 5.2: Assessment of ESR and MSE metrics for AM reconstructed monocomponent signals 1<sup>st</sup>, 2<sup>nd</sup>, 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> using EMD, SWD, Q-VMD, and proposed ESSDM method. (Note: '-' represents missing values for non-constructed components).

Methods	Proposed ESSDM	SWD	EMD	Q-VMD
ESR <sub>1</sub> (1 <sup>st</sup> )	0.0001	1.0725	0.0216	0.0010
ESR <sub>2</sub> (2 <sup>nd</sup> )	0.0879	0.0104	0.2190	0.0558
ESR <sub>3</sub> (3 <sup>rd</sup> )	0.0266	0.0017	0.3553	0.2502
ESR <sub>4</sub> (4 <sup>th</sup> )	0.0007	0.0010	-	1.7405
ESR <sub>5</sub> (5 <sup>th</sup> )	0.0028	0.0010	-	0.0491
MSE <sub>1</sub> (1 <sup>st</sup> )	0.0043	0.0618	0.0594	0.0016
MSE <sub>2</sub> (2 <sup>nd</sup> )	0.0115	0.0031	0.0676	0.0115
MSE <sub>3</sub> (3 <sup>rd</sup> )	0.0116	0.0024	0.0808	0.1094
MSE <sub>4</sub> (4 <sup>th</sup> )	0.0006	0.0009	-	0.0836
MSE <sub>5</sub> (5 <sup>th</sup> )	0.0091	0.0012	-	0.0947

### 5.3.4 Feature formulation and classification

For an efficient and feasible MI-EEG classification model, extracting suitable features from multichannel signals is a crucial task in developing BCI applications. In our work, we have introduced a new feature approach that employs FTFG features, which were derived from computed channel-specific TF images. To extract the FTFG feature, the TF images were processed by applying a graph spectral filter to learn the feature attributes of the TF image of the size of  $875 \times 756 \times 3$  and generate the sparse adjacency matrix. In order to find the most discriminative FTFG features across cross-channel, a k-means algorithm-based sub-space clustering approach is employed to find similarity between nodes among inter-

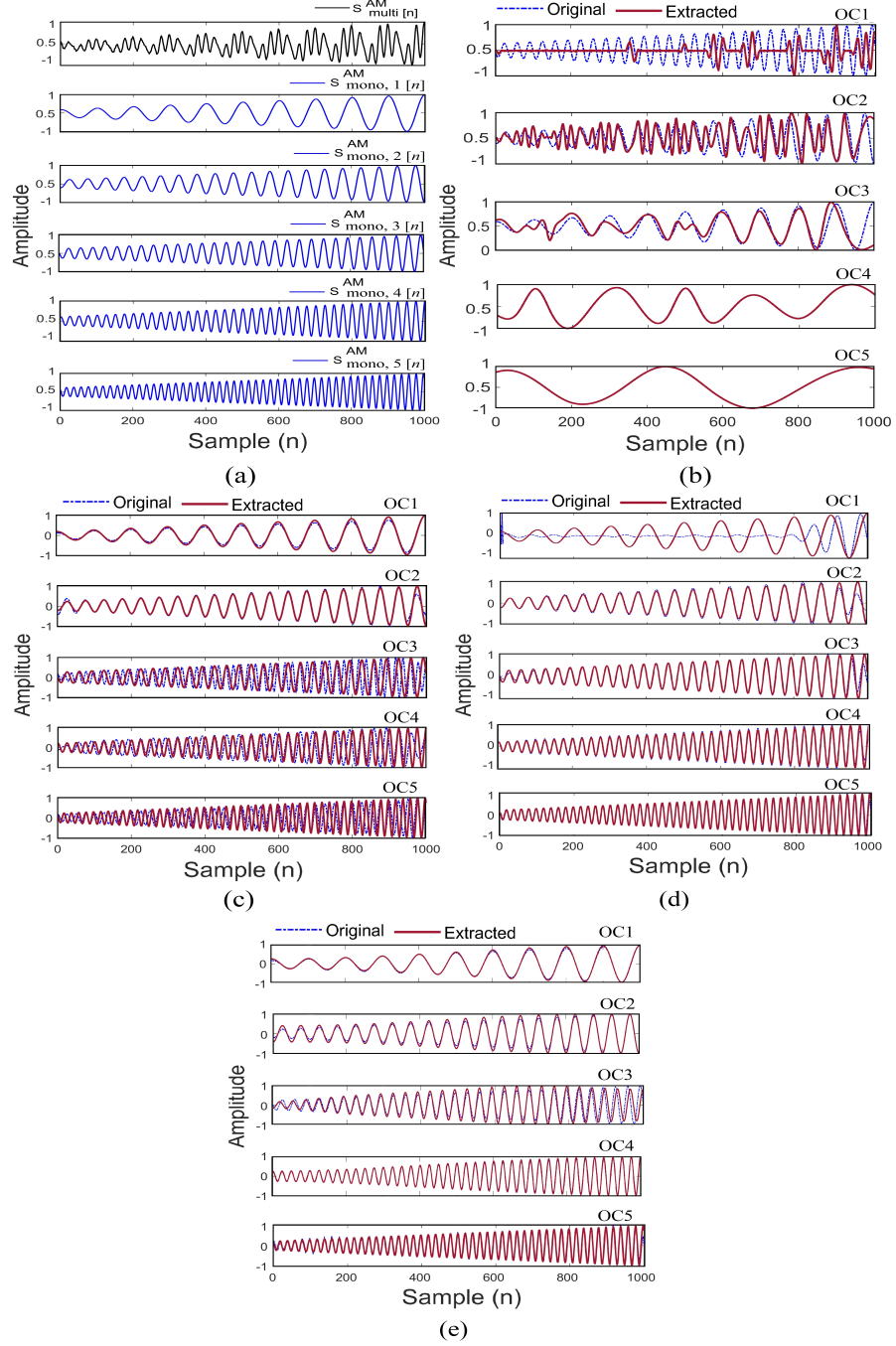


Figure 5.5: Plots of (a) nonstationary AM input signal and decomposition results using (b) EMD, (c) Q-VMD, (d) SWD, and (e) proposed ESSDM method.

channel graphs with respect to directly connected and indirectly connected nodes. In order to understand how different computed sub-space clustered features affect the performance complexity, we have tested against the performance parameter: Normalized Dunn index (NDI). By selecting the highest  $NDI > 0.5$ , the proposed FTFG feature shows significance

for the performance enhancement of MI-EEG class recognition. Fig. 5.6 illustrates the extracted FTFG features for four different MI-EEG classes. These obtained features are further fed to three different low computational classifiers: ELM [123], CNet [97], and CapsNet [138]. In this work, ELM, CNet, and CapsNet were specifically selected due to their shorter training times and fewer layer requirements. Algorithm 5.1 outlines the FTFG feature formulation flow.

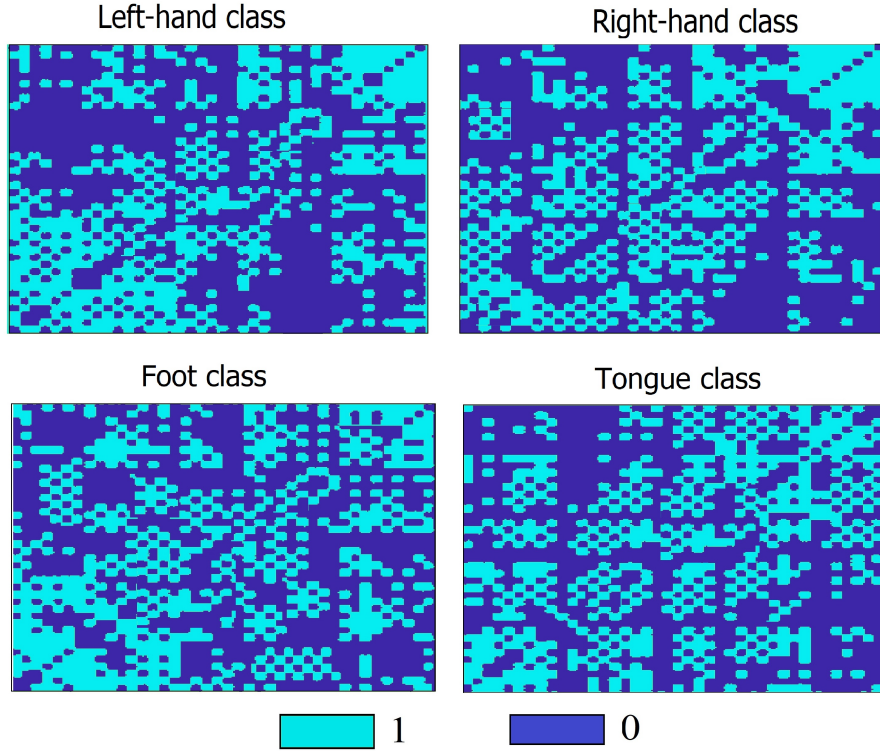


Figure 5.6: The extracted FTFG features based adjacency matrix from TF image for four MI-EEG classes.

### 5.3.5 Evaluating classifier metrics and validation scheme

In order to validate the performance of the proposed method, we have conducted specific-subject and cross-subject validation by employing leave-one-out cross-validation using 80%-20% training and testing data configuration. This process was validated with  $k$ -fold validation schemes, and the average result was reported by considering the performance metrics, average accuracy (ACCavg) and Cohen's kappa coefficient [141]. The kappa ( $k$ ) is a robust measure of classification performance by evaluating by correctness of the accuracy

---

**Algorithm 5.1** Computation of ESSDM-based FTFG features

---

Input: Nonstationary multichannel MI-EEG signal  $s(k)$

Output: FTFG features from ESSDM-based optimal OC modes

Initialization:  $it = 0, c = 0, k = 0, \text{Max}_{tr} = 0, \text{member}A = 30$

---

**Repeat**

// Apply SHO optimization

**Repeat**

Initialize threshold parameters as population  $A = \{H, P\}$

Calculate H and P movement

Compute best survival value SV of each individual ( $a_i$ )

$$SV = [P_{th}, S_{td}Th] = SV_{a_i} = \frac{f(a_i) - f_{best}}{f_{best} - f_{worst}}$$

// Apply modified swarm filter bank and extract optimal OCs

**Repeat**

**Repeat**

Compute SFT spectrum by Eq. (2.4)

Smooth spectrum by estimating highest ESD  $s_{it}(k)$  by Eq. (2.7)

Find spectral boundaries on the obtained SFT spectrum

Select appropriate swarm filter parameters  $\delta$  and  $M$  Eq. (2.9a) and Eq. 2.9b), respectively

**Repeat**

$y_{j+1} \leftarrow \text{SwF}(s_j(k), M, \delta)$ ; SwF: Swarm filter

$j \leftarrow j + 1$

$B_{it+1}(k) \leftarrow B_{it}(k)$

**Until**  $B_{it} < 0.1$

$s'_{it}(k) \leftarrow y_j(k)$

$s_{it+1}(k) \leftarrow s_{it}(k) - s'_{it}(k)$

$it \leftarrow it + 1, j = 0$

$y_0(k) \leftarrow s_{it}(k)$

**Until**  $E'_{s_{it}}(\omega) < P_{th}$

$C_{it} \leftarrow C_{it+1}$

**Until**  $N_{it} < N_c$

$Y_{it}(k) \leftarrow y_0(k)$

Calculate fitness function  $f(P_{th}, S_{td}Th)$  by (5.2)

Recalculate survival value of each individual with restoration phase

**Until**  $f_{it} < f_{it-1}$

// Estimation of FTFG features

Computation of IF and IA using HT

Compute adjacency matrix  $V$  using un-directional graph from TF image

Generate  $K$ -cluster using sub-space clustering

Compute distance function between cluster pair  $V_i$  and  $V_j$  using normalized Dunn index

$$NDI = \begin{cases} \max\left(\frac{\min_{i \neq j} \{d_{inter}(V_i, V_j)\}}{\max_{i \neq j} \{d_{intra}(V_i, V_j)\}}\right) & , \text{ if } (V_i \cong V_j) > 0.5 \\ 0 & , \text{ Otherwise} \end{cases}$$

Select  $K$ -cluster using NDI to get FTFG feature

$T_{it} \leftarrow T_{it+1}$

**Until**  $T_{it} < \text{Max}_{tr}$  ;  $\text{Max}_{tr}$  - maximum number of trails

---

Notations:  $B$ : Bhattacharya distance,  $M$ : Number of swarm members,  $\delta$ : Virtual time interval,  $P_{th}$ : Spectrum peak threshold selection,  $N_c$ : Number of channels,  $it$ : Iteration.

---



level, which is expressed as follows:

$$k = \frac{A_o - A_r}{1 - A_r} \quad (5.5)$$

$$A_r = \frac{x_1 \times y_1 + x_2 \times y_2 + \dots + x_c \times y_c}{n \times n} \quad (5.6)$$

Where  $A_o$  and  $A_r$  represent the accuracy and the model's random classification rate.  $x_1, \dots, x_c$  signify the actual samples in each MI class, while  $y_1, y_2, \dots, y_c$  represent the samples predicted by the model.  $c$  corresponds to the number of MI classes and  $n = x_1 + x_2 + \dots + x_c$  is the total sample count. We employ  $p$ -values from Wilcoxon signed-rank test to assess the statistical significance of ESSDM-based classification models against other baseline models [23, 24, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 73, 142]:  $p > 0.05$  implies no statistical difference,  $p < 0.05$  indicates statistical difference, and  $p < 0.01$  suggests high statistical significance.

## 5.4 Results and discussions

In order to study the effectiveness of the proposed ESSDM-JTFDF, channel-aligned multi-scale fused JTFDF features have been tested by designing an end-to-end classification framework for MI-EEG signals-based motory imaginary enhancement. We have investigated the multi-class classification performance on two public datasets with the 4-class in BCI IV-2a [140] and the 2-class in BCI III-4a [139] during MI imagination actions. In our experiment, MI-EEG signals from different electrode configurations were taken into consideration, including 22 electrodes and 118 electrodes, respectively. The 30 s analysing window of each of these EEG signals was selected and further segmented into 10-second epochs with considered imagined activation stimulus. The ESSDM method is used to extract the optimal OC modes from the multichannel MI-EEG signals. Then channel-specific TF coefficients are computed from these extracted OCs to transform TF images using HSA [41]. These TF images explore the TF analysis for discriminative neural activity patterns due to the MI imagination tasks. In the TF plots, the highest activation during the MI tasks was observed in the alpha ( $\alpha$ : 8 – 13 Hz) and beta ( $\beta$ : 13 – 30 Hz) rhythms with imag-

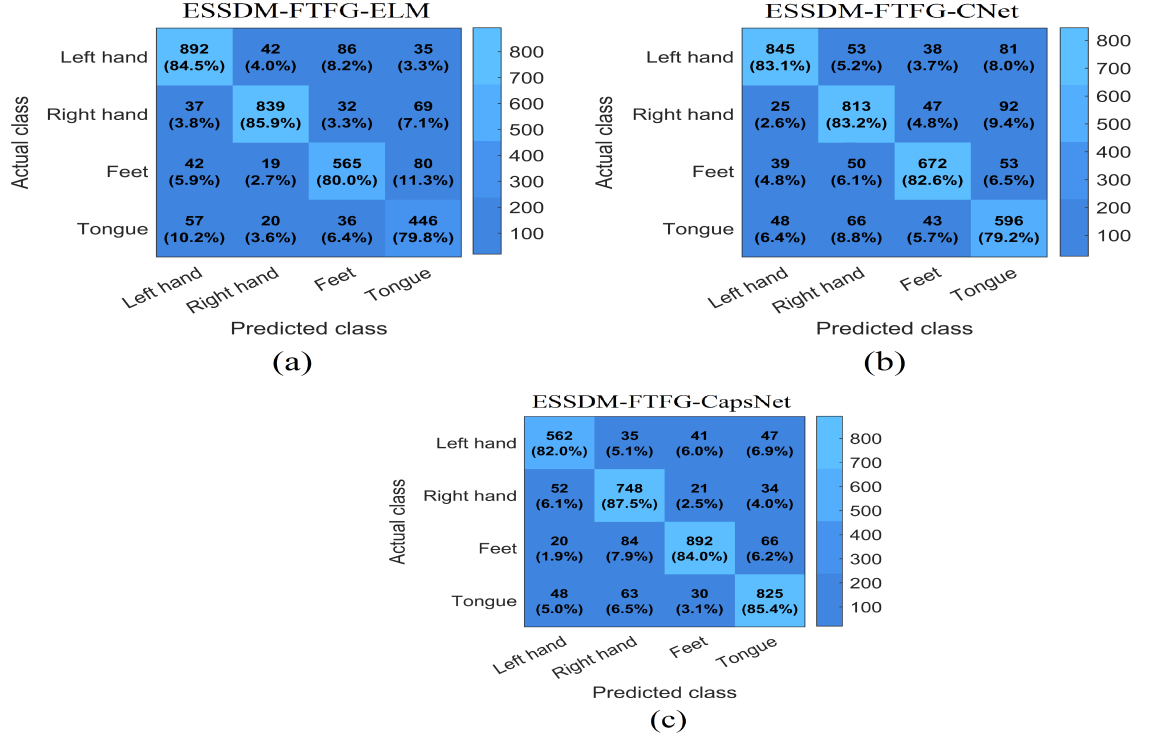


Figure 5.7: Averaged confusion matrix across all subjects using (a) ESSDM-FTFG-ELM, (b) ESSDM-FTFG-CNet, and (c) ESSDM-FTFG-CapsNet on the BCI IV-2a dataset.

ined stimulus response during the time frame (0.5 - 2.5 s) after the visual onset of imagination. The performance study has been carried out in three significant aspects, whose results are discussed in the following sections. Firstly, classification accuracy and visualization of the specific-subject benchmark performance are presented. In the second phase, cross-subject benchmarks were used to assess the classification performance of obtained ESSDM-based-fused features across the two datasets. The third phase of the simulation study determines how subspace clustering-based feature dimensionality enhances classification performance. To investigate the novel FTFG feature, we have formulated three different ESSDM-based classification frameworks, proposed model 1: SSDM-FTFG-ELM, proposed model 2: ESSDM-FTFG-CNet, and proposed model 3: ESSDM-FTFG-CapsNet. To compare with the proposed models, the classification results of non-optimised SWD are also carried out and validated with 10-fold cross-validation.

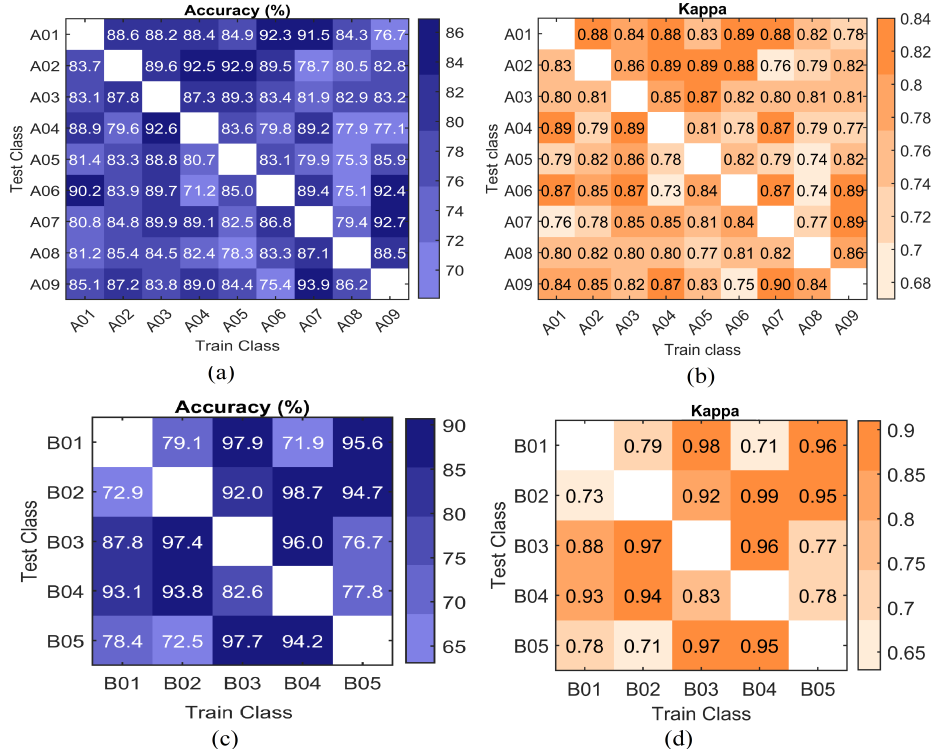


Figure 5.8: The classification performance of cross-subject benchmarks (ACCavg, kappa) on the BCI IV-2a and III-4a datasets.

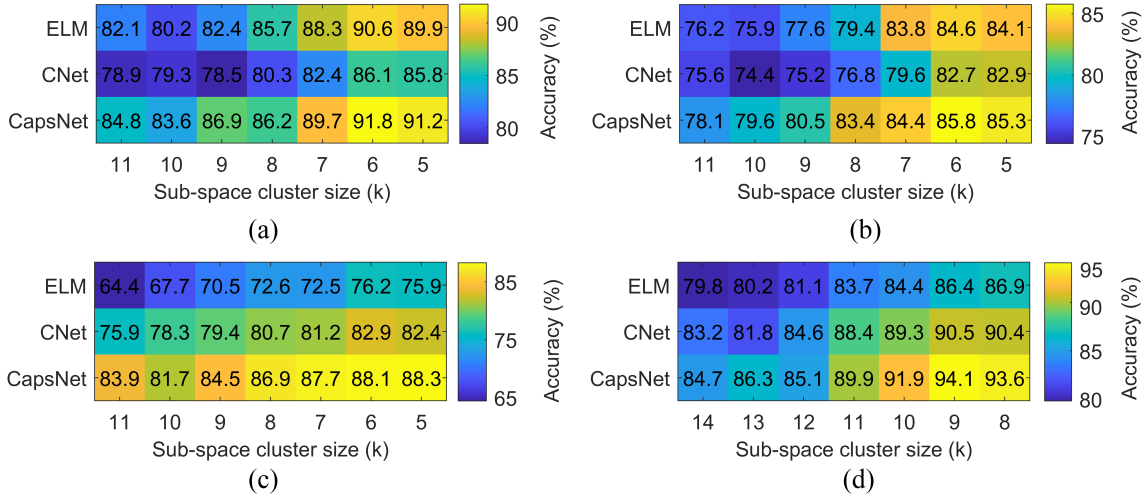


Figure 5.9: Average accuracy obtained in specific-subject (a) bi-class and (b) multi-class on IV-2a, and cross-subject (c) and (d) for IV-2a and III-4a, respectively.

#### 5.4.1 Performance comparison with specific-subject benchmarks

To verify the effectiveness of the ESSDM-based fused feature, we have carried out several comparative experiments using the proposed ESSDM and non-optimised SWD method

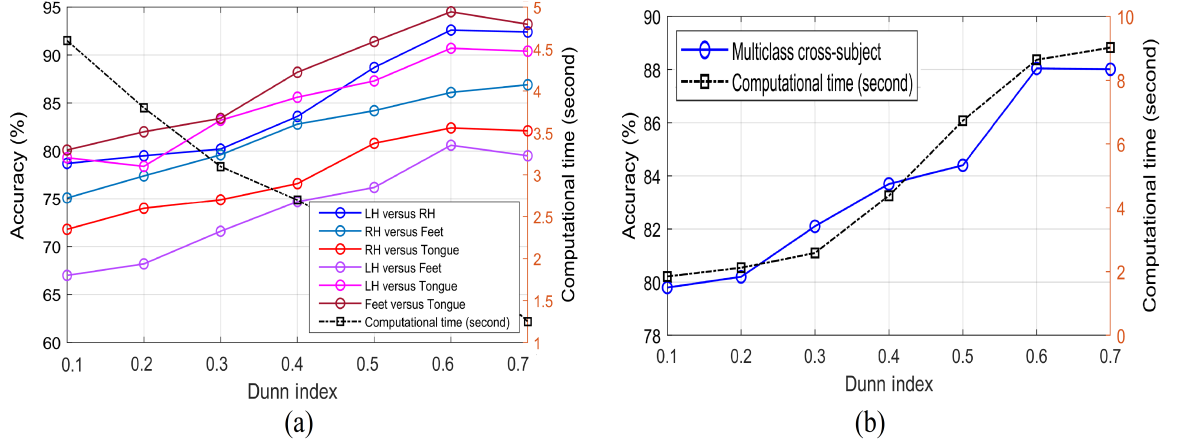


Figure 5.10: The effect of a sub-clustered factor in (a) bi-class specific-subject and (b) cross-subject performance on the IV-2a dataset.

on the BCI IV-2a and III-4a datasets and also compared with some specific-subject benchmark baseline models. In Table 5.3, the maximum average accuracy is marked in bold font. Table 5.3 demonstrates the classification performance (ACCavg, kappa, and  $p$ -values) following 10-fold cross-validation for proposed ESSDM and SWD with the state-of-the-art models on the BCI IV-2a dataset (four classes). The following baseline models were utilized for performance comparison: sparse filter band common spatial pattern (SFBCSP) [23], compact convolutional recurrent neural network (SCCRNN) [27], ORDW [28], metric-based spatial filtering transformer (MSFT) [73], CNN-long short-term memory (CNN-LSTM) [27], Global [142], and MS-AMF [26]. For a more reasonable method analysis, these four types of imagination classes are first paired for bi-classification. In the bi-class scenario, classification results between the two mental tasks are evaluated for nine subjects in six pairs: left-hand versus right-hand (LH-RH), right-hand versus feet (RH-F), right-hand versus tongue (RH-T), left-hand versus feet (LH-F), left-hand versus tongue (LH-T), and feet versus tongue (F-T). Significant differences between the results are evaluated by Wilcoxon signed-rank test. According to the classification of six pairs, results show that the proposed classification framework ESSDM-FTFG-CapsNet has the highest average accuracy with 91.8% ( $k = 0.88$ ,  $p = 0.02$ ) among all methods, and it is 4.5% higher than SWD-FTFG-CapsNet ( $p = 0.05$ ). It indicates that fused features extracted by ESSDM have a significant contribution to the classification results, and these results are consistent with

10-fold cross-validation. There is no significant difference between the results of ESSDM-FTFG-ELM ( $ACC_{avg} = 90.6$ ,  $k = 0.85$ ,  $p = 0.02$ ) and SWD-FTFG-CapsNet ( $ACC_{avg} = 87.2\%$ ,  $k = 0.82$ ,  $p = 0.04$ ). In the context of specific-subject longitudinal classification, certain subjects have exhibited enhanced individual performance. Specifically, subject A07 demonstrated an improvement from 82% to 92.8%, A08 improved from 85% to 88.6%, A09 showed an improvement from 83% to 91.9%, whereas A05 decreased from 91.8% to 83.0%, and A03 decreased from 97% to 91.6%.

For multi-class scenarios, the classification performance of four-task MI-EEG has been tested on BCI IV-2a and BCI III-4a datasets, as shown in Table 5.3. Classification results of the proposed FTFG features under different feature clustered combinations are analyzed, and the best results are reported. It is found that the highest average accuracy of ESSDM-FTFG-CapsNet and ESSDM-FTFG-ELM are 85.8% ( $k = 0.84$ ,  $p = 0.01$ ) and 84.6% ( $k = 0.82$ ,  $p = 0.02$ ), respectively. ESSDM-FTFG-CNet model reported the average performance was ( $ACC_{avg} = 82.7\%$ ,  $k = 0.79$ ,  $p = 0.01$ ). Whereas in the case of SWD based scheme, the highest accuracy was attained by SWD-FTFG-CapsNet ( $ACC_{avg} = 82.3$ ,  $k = 0.78$ ,  $p = 0.03$ ). The inferior result was reported by SWD-FTFG-CNet ( $ACC_{avg} = 79.82\%$ ,  $k = 0.75$ ,  $p = 0.05$ ), when compared to the best baseline model MSFT-MLP [73] which had the highest reported  $ACC_{avg} = 86.1\%$ . The Wilcoxon significance reveals a significant difference ( $p = 0.009$ ) between the proposed method and non-optimized SWD, which indicating that the computed FTFG features exhibit significant TF characteristics which are most relevant to brain motory actions. Compared to individual subjects, the ESSDM-FTFG-CapsNet model has the highest accuracy and kappa for subjects A01, A02, A04, A06, and A09. Especially for A06, the accuracy has been improved by at least 15.5% compared with the best baseline models [73], indicating the presence of subject-specific sensitivities. For subjects A03 and A07, the classification accuracy obtained by ESSDM-FTFG-CapsNet is significantly lower than the best baseline model at 7.8% and 6.5%, respectively, probably because their EEG signals are covered by artifacts and noises, resulting in ESSDM-FTFG-CapsNet could not extract significant spectral characteristics. In terms of average accuracy, ESSDM-FTFG-CapsNet achieves the highest accuracy of 85.8% and kappa of 0.84 with fewer layers, which is better than the baseline models. Especially, the average accuracy

is 12.1% higher than the ORDW [28] with optimal subject-specific network parameters and slightly deteriorated with the MSFT [73] model. ESSDM-FTFG-CapsNet does not change any threshold parameters when testing all subjects. As compared to baseline models [24, 27, 28], the proposed ESSDM method-based frameworks show significant classification performance due to two reasons. First, the derived fused feature across channels shows the most discriminative feature that exhibits a specific TF band involved in MI. Second, it is less computationally efficient due to the use of low-dense ELM, CNet, and CapsNet classifier-based framework in contrast to existing deep learning-based approaches used in baseline models [24, 27, 28], especially with the model MSFT-MLP [73]. Furthermore, to evaluate the performance of the proposed models, we have carried out an additional study that involved plotting a confusion matrix to classify MI-EEG classes across all nine specific subjects, which is shown in Fig. 5.7. We observed that the proposed model ESSDM-FTFG-CapsNet in the dataset [139, 140] is advantageous in recognizing MI-EEG classes (feet and tongue) more accurately than the ESSDM-FTFG-ELM and ESSDM-FTFG-CNet, whereas the other models (ESSDM-FTFG-ELM and ESSDM-FTFG-CNet) have greater accuracy in recognizing other classes (right-hand and left-hand). These findings indicate the classification rate for other classes was also improved due to the enhanced FTFG features to a certain extent in comparison to prior works [24, 27, 28].

#### 5.4.2 Performance comparison with cross-subject benchmarks

To thoroughly validate, the proposed ESSDM method has been tested for cross-subject scenarios. The cross-subject validation was utilized with every trained specific-subject feature vector in the cross with different subjects' test features employing a 10-fold cross-validation approach. Table 5.4 and Fig. 5.8 exhibit the cross-subject classification performance on the BCI IV-2a and III-4a datasets. It can be observed that the classification performance of the ESSDM model composed of multi-channel features is significantly higher than that of the non-optimized SWD-based model, including baseline models only using the novel fused ESSDM-FTFG features. The results show that all cross-subject results exceed classification performance of 88.04% for IV-2a, with the exception of cross-subjects 4 and 6 (lowest ACCavg = 71.2%) and 98.12% for III-4a with the exception of cross-subject 5 and 2

Table 5.3: Average accuracy score (in %) and kappa ( $k$ ) in both MI-EEG labeled scenarios on the BCI IV-2a dataset. The best individual values are marked in bold. Notation NR denotes non-reported values.

Method	Subjects									Average
	A01	A02	A03	A04	A05	A06	A07	A08	A09	
<b>Bi-class scenario</b>										
SFBCSP [23]	98.8(NR)	97.8(NR)	97.0(NR)	94.7(NR)	91.8(NR)	90.6(NR)	67.9(NR)	63.5(NR)	58.4(NR)	84.5(NR)
Expert-CNN [24]	91.6(NR)	90.8(NR)	92.3(NR)	86.5(NR)	87.8(NR)	88.8(NR)	78.4(NR)	80.4(NR)	68.4(NR)	85.0(NR)
SCCRNN [27]	95.0(NR)	85.0(NR)	95.0(NR)	85.0(NR)	89.0(NR)	78.0(NR)	74.0(NR)	85.0(NR)	79.0(NR)	85.0(NR)
ORDW [28]	96.0(0.92)	93.5(0.87)	97.2(0.94)	89.4(0.79)	91.5(0.83)	81.2(0.62)	82.0(0.64)	79.0(0.58)	83.0(0.66)	88.0(0.76)
SWD-FTFG-ELM	81.9(0.77)	83.8(0.78)	93.8(0.81)	74.0(0.69)	75.4(0.70)	78.0(0.73)	86.3(0.81)	85.4(0.84)	92.6(0.87)	83.6(0.79)
SWD-FTFG-CNet	82.1(0.78)	80.1(0.74)	83.5(0.78)	69.9(0.65)	77.4(0.72)	71.4(0.66)	85.5(0.80)	83.7(0.88)	89.5(0.87)	81.4(0.77)
SWD-FTFG-CapsNet	88.6(0.84)	89.9(0.85)	91.9(0.87)	82.9(0.78)	75.5(0.70)	80.6(0.75)	88.4(0.83)	82.4(0.87)	86.2(0.81)	<b>87.2(0.82)</b>
Proposed-1 (ESSDM-FTFG-ELM)	86.2(0.80)	86.6(0.81)	89.7(0.85)	91.0(0.87)	80.6(0.77)	87.4(0.81)	91.3(0.90)	85.3(0.87)	<b>93.6(0.89)</b>	<b>90.6(0.85)</b>
Proposed-2 (ESSDM-FTFG-CNet)	87.2(0.81)	84.6(0.79)	85.9(0.81)	84.6(0.80)	79.5(0.76)	87.7(0.82)	90.2(0.86)	84.8(0.79)	90.3(0.86)	86.1(0.81)
Proposed-3 (ESSDM-FTFG-CapsNet)	96.7(0.90)	91.8(0.87)	91.6(0.87)	91.8(0.87)	83.0(0.80)	89.1(0.87)	<b>92.8(0.91)</b>	<b>88.6(0.94)</b>	91.9(0.87)	<b>91.8(0.88)</b>
<b>Multi-class scenario</b>										
NCSP [29]	79.2(NR)	52.1(NR)	83.3(NR)	62.2(NR)	54.5(NR)	39.2(NR)	83.3(NR)	82.6(NR)	66.7(NR)	67.0(NR)
FDBN [30]	71.1(NR)	55.6(NR)	76.9(NR)	65.6(NR)	69.1(NR)	65.0(NR)	71.7(NR)	92.4(NR)	82.4(NR)	72.2(NR)
FB-CSP [31]	76.0(NR)	56.5(NR)	81.3(NR)	61.0(NR)	55.0(NR)	45.3(NR)	82.8(NR)	81.3(NR)	70.8(NR)	67.8(NR)
CM-CNN [32]	87.5(NR)	65.3(NR)	90.3(NR)	66.7(NR)	62.5(NR)	45.5(NR)	89.6(NR)	83.3(NR)	79.5(NR)	74.5(NR)
FB-CSP-SVM [32]	82.3(NR)	60.4(NR)	83.0(NR)	72.6(NR)	60.7(NR)	44.1(NR)	86.1(NR)	77.1(NR)	75.0(NR)	71.2(NR)
1D-AX-LSTM [33]	75.1(NR)	71.4(NR)	72.2(NR)	72.9(NR)	82.6(NR)	69.6(NR)	89.0(NR)	80.3(NR)	75.1(NR)	76.5(NR)
DFB-CSP [24]	84.9(NR)	66.4(NR)	84.7(NR)	81.4(NR)	79.2(NR)	70.7(NR)	86.1(NR)	83.8(NR)	83.0(NR)	80.0(NR)
3D-CNN multibranch [25]	77.4(NR)	60.1(NR)	82.9(NR)	72.3(NR)	75.8(NR)	69.0(NR)	76.0(NR)	76.9(NR)	84.7(NR)	75.0(NR)
MS-AMF [26]	88.3(NR)	65.7(NR)	92.0(NR)	77.7(NR)	60.9(NR)	63.7(NR)	88.2(NR)	93.2(NR)	89.5(NR)	79.9(NR)
CD-LOSS [34]	91.3(NR)	71.6(NR)	92.3(NR)	78.4(NR)	80.1(NR)	61.6(NR)	92.6(NR)	90.3(NR)	78.4(NR)	81.9(NR)
CNN-LSTM [27]	NR(0.87)	NR(0.90)	NR(0.83)	NR(0.95)	NR(0.85)	NR(0.77)	NR(0.66)	NR(0.78)	NR(0.54)	NR(0.79)
Global [142]	86.7(0.82)	77.1(0.69)	80.9(0.74)	88.4(0.84)	88.6(0.84)	66.5(0.54)	56.0(0.41)	71.0(0.61)	55.9(0.40)	74.6(0.64)
ORDW [28]	89.8(0.86)	79.8(0.73)	85.2(0.8)	83.9(0.79)	80.1(0.79)	54.1(0.39)	60.1(0.47)	60.5(0.47)	70.2(0.6)	73.7(0.66)
MSFT-MLP [73]	86.7(NR)	85.0(NR)	96.7(NR)	83.3(NR)	68.3(NR)	70.0(NR)	98.3(NR)	98.3(NR)	88.3(NR)	86.1(NR)
SWD-FTFG-ELM	78.3(0.73)	80.2(0.75)	82.1(0.77)	72.2(0.67)	73.6(0.69)	76.1(0.71)	84.2(0.79)	90.8(0.86)	87.0(0.82)	<b>80.5(0.78)</b>
SWD-FTFG-CNet	80.1(0.76)	78.1(0.73)	81.4(0.76)	68.2(0.63)	75.6(0.71)	70.3(0.65)	84.2(0.79)	91.7(0.87)	88.1(0.83)	<b>79.8(0.75)</b>
SWD-FTFG-CapsNet	84.8(0.80)	86.0(0.81)	87.9(0.83)	79.3(0.74)	72.2(0.67)	77.2(0.72)	<b>92.3(0.87)</b>	88.5(0.83)	82.5(0.77)	<b>82.3(0.78)</b>
Proposed-1 (ESSDM-FTFG-ELM)	82.5(0.78)	84.1(0.79)	87.1(0.83)	88.3(0.84)	78.2(0.74)	81.1(0.76)	94.7(0.90)	91.2(0.86)	<b>89.6(0.85)</b>	<b>84.6(0.82)</b>
Proposed-2 (ESSDM-FTFG-CNet)	83.4(0.78)	82.2(0.77)	83.4(0.79)	82.2(0.78)	77.2(0.73)	81.3(0.76)	84.0(0.80)	79.6(0.74)	85.4(0.81)	<b>82.7(0.79)</b>
Proposed-3 (ESSDM-FTFG-CapsNet)	<b>92.5(0.88)</b>	<b>89.1(0.84)</b>	88.9(0.84)	<b>89.1(0.85)</b>	80.6(0.76)	<b>85.5(0.80)</b>	91.8(0.87)	91.3(0.86)	87.9(0.84)	<b>85.8(0.84)</b>

SFBCSP: Sparse filter band common spatial pattern, SCCRNN: Compact convolutional recurrent neural network, ORDW: Optimized Deep&Wide, MSFT: Metric-based spatial filtering transformer, CNN-LSTM: Convolutional neural network-long short-term memory, MS-AMF: Multi-scale fusion attention mechanism, 3D-CNN: 3D-convolutional neural network, CSSBP: Common spatial-spectral boosting pattern, CSSSP: Common sparse spectral-spatial pattern, FB-CSP: Sparse filter band common spatial pattern, CM-CNN: Channel mixing CNN, DFB-CSP: Discriminative filter bank common spatial pattern, FDBN: Frequential deep belief network, 1D-AX: One dimension-aggregate approximation, SWD: Swarm decomposition, CD-LOSS: Central distance loss.

(lowest ACCavg = 72.5%). In comparison to baseline models tested on the IV-2a dataset, including ORDW [28], MS-AMF [26], 3D-CNN [25], common spatial-spectral boosting pattern (CSSBP) [128], SFBCSP [23], common sparse spectral-spatial pattern (CSSSP) [128], and non-optimized SWD, the average classification accuracy improved by 3.84% ( $k = 0.65$ ), 8.14% ( $p = 0.04$ ), 13.03% ( $p = 0.0034$ ), 13.24% ( $p = 0.003$ ), 15.44% ( $p > 0.05$ ), 22.04% ( $> 0.05$ ), and 8.96% ( $k = 0.785$ ,  $p = 0.03$ ), respectively (shown in Fig. 5.11).

The proposed ESSDM-FTFG-CapsNet model achieves a classification accuracy of 88.04% ( $k = 0.84$ ,  $p = 0.02$ ) and 98.12% ( $k = 0.927$ ,  $p = 0.01$ ) in IV-2a and III-4a, respectively. Compared with other baseline models, the classification accuracy is improved by 12.09% (in IV-2a) and 7.62% (in III-4a) at most and 3.84% (in IV-2a) and 5.12% (in III-4a) at least. In the kappa measure, results are improved by 4.96% at most and 2.90% at least. In the Wilcoxon results, the results of ESSDM-FTFG-CapsNet were ( $p = 0.014$ , very significant), and no significant change ( $p > 0.05$ ) in classification accuracy for baseline

Table 5.4: Cross-subject classification accuracy in cross-subject scenario on the BCI IV-2a and III-4a datasets.

Dataset	Cross-subject	Threshold parameters	Fused features	Cluster size	NDI	Enhanced SSDM						Non-optimized SWD					
						ELM		CNet		CapsNet		ELM		CNet		CapsNet	
						ACC <sub>avg</sub> (in %)	K	ACC <sub>avg</sub> (in %)	K	ACC <sub>avg</sub> (in %)	K	ACC <sub>avg</sub> (in %)	K	ACC <sub>avg</sub> (in %)	K	ACC <sub>avg</sub> (in %)	K
BCI IV-2a	A01/A06	$f(\text{Pth} = 0.391, \text{StdTh} = 0.523)$	800	7	0.587	74.51	0.739	80.34	0.808	85.34	0.832	74.45	0.706	75.82	0.758	78.40	0.780
	A02/A05	$f(\text{Pth} = 0.246, \text{StdTh} = 0.527)$	835	8	0.315	76.29	0.770	76.89	0.759	85.91	0.834	73.82	0.724	76.10	0.761	78.69	0.770
	A03/A05	$f(\text{Pth} = 0.332, \text{StdTh} = 0.397)$	811	7	0.524	77.61	0.785	78.77	0.774	82.32	0.813	73.96	0.784	76.01	0.760	79.05	0.795
	A04/A03	$f(\text{Pth} = 0.284, \text{StdTh} = 0.418)$	831	8	0.475	73.95	0.734	81.23	0.820	85.60	0.828	76.52	0.747	76.37	0.764	78.87	0.795
	A05/A03	$f(\text{Pth} = 0.365, \text{StdTh} = 0.532)$	815	6	0.624	79.03	0.798	80.11	0.810	81.81	0.804	75.67	0.770	75.45	0.755	78.91	0.775
	A06/A09	$f(\text{Pth} = 0.309, \text{StdTh} = 0.465)$	818	9	0.457	69.27	0.701	82.08	0.817	85.38	0.832	73.11	0.760	74.98	0.750	78.01	0.792
	A07/A09	$f(\text{Pth} = 0.220, \text{StdTh} = 0.545)$	841	8	0.636	74.89	0.758	78.65	0.774	85.72	0.830	75.89	0.759	75.62	0.756	78.47	0.780
	A08/A09	$f(\text{Pth} = 0.475, \text{StdTh} = 0.468)$	894	8	0.563	75.18	0.744	81.01	0.824	81.54	0.802	73.33	0.753	75.09	0.751	78.57	0.797
	A09/A07	$f(\text{Pth} = 0.344, \text{StdTh} = 0.449)$	836	6	0.336	85.31	0.862	79.92	0.781	86.86	0.843	75.99	0.768	75.55	0.756	79.14	0.785
	<b>Average</b>					<b>76.22</b>	<b>0.766</b>	<b>82.89</b>	<b>0.796</b>	<b>88.04</b>	<b>0.842</b>	<b>74.75</b>	<b>0.752</b>	<b>75.67</b>	<b>0.757</b>	<b>79.00</b>	<b>0.785</b>
BCI III-4a	B01/B04	$f(\text{Pth} = 0.327, \text{StdTh} = 0.328)$	2302	6	0.698	91.56	0.889	94.81	0.915	96.58	0.937	81.76	0.796	88.65	0.865	90.62	0.882
	B02/B03	$f(\text{Pth} = 0.415, \text{StdTh} = 0.564)$	2319	9	0.312	80.31	0.814	87.45	0.836	89.83	0.886	78.59	0.769	81.56	0.812	85.73	0.803
	B03/B05	$f(\text{Pth} = 0.487, \text{StdTh} = 0.571)$	2307	7	0.596	87.09	0.867	91.99	0.893	92.89	0.881	80.57	0.767	87.73	0.823	88.46	0.846
	B04/B02	$f(\text{Pth} = 0.482, \text{StdTh} = 0.368)$	2321	8	0.409	83.02	0.813	88.57	0.853	93.55	0.924	79.98	0.772	82.14	0.804	87.28	0.828
	B05/B01	$f(\text{Pth} = 0.258, \text{StdTh} = 0.542)$	2343	7	0.688	90.18	0.874	89.9	0.853	97.86	0.947	82.14	0.784	90.37	0.852	92.59	0.885
	<b>Average</b>					<b>86.43</b>	<b>0.850</b>	<b>90.54</b>	<b>0.87</b>	<b>94.12</b>	<b>0.915</b>	<b>80.60</b>	<b>0.776</b>	<b>86.09</b>	<b>0.831</b>	<b>88.93</b>	<b>0.848</b>

K: Cohen's kappa

models [28, 73]. These findings confirm that the proposed ESSDM method offers significant advantages over other baseline models in cross-subject scenarios.

### 5.4.3 Interpretability of the learned relevant fused features representations versus classification performance

In this section, we investigate the performance of novel FTFG features with varying dimensionality and evaluate the relationship between the number of cluster data subsets and classification accuracy and time cost in different datasets. To address the impracticality of using high-dimensional features in a BCI system, it is needed to cluster most correlated features across channels to a lower-dimensional space. To assess the results of the clustering of fused features, we averaged the normalized Dunn index (NDI) with a range [0 – 1]. Higher NDI value exhibits the efficient selection of fused FTFG features without compromising



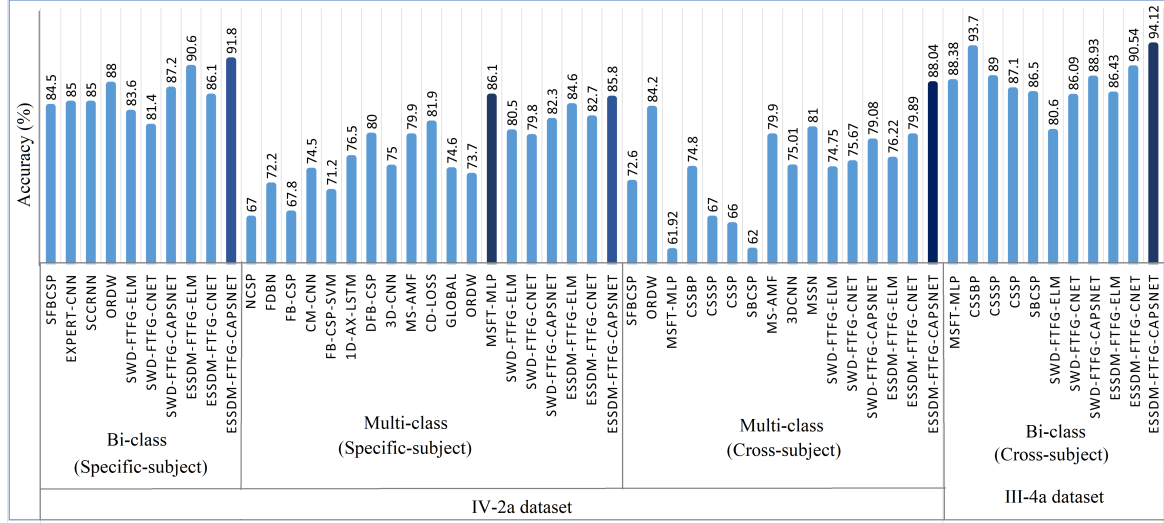


Figure 5.11: Performance comparison with the proposed method with the benchmark base-line models.

performance and produces the best classification results. Figure 5.9 shows heatmaps for classification performance on fused FTFG features obtained from different NDI values with models (ELM, CNet, and CapsNet) on the datasets (BCI IV-2a and III-4a). These heatmaps report the grid search classification results with the different subspace cluster values on the x-axis and the model type on the y-axis. Here, we showed that the feature dimensionality increased ( $NDI > 0.5$ ), and the mean classification accuracies changed significantly in all models. Obviously ESSDM model yielded relatively higher classification accuracies compared with the other competing approaches. In order to understand how clustering of features enhances BCI recognition performance, we plotted the number of data subsets ( $k$ ) corresponding to the NDI values against classification accuracy and time complexity for bi-class (specific-subject) and multi-class (cross-subject) scenarios on datasets IV-2a and III-4a, as displayed in Fig. 5.9. The computation time cost is obtained on MATLAB-R2022b on a desktop computer with a Core (TM) i7-2600 CPU and 16 GB RAM. In dataset IV-2a, when we increase the number of clusters index from 5 to 9, computation time increases, and the average accuracy for different bi-class tasks decreases slightly. For instance, MI tasks RH-LH and F-T have demonstrated average accuracy reductions to 7.4% and 8.9%, respectively (shown in Fig. 5.10 (a)). However, as we continue decreasing subsets, the accuracy improves significantly until it levels out at the 6-subspace cluster index and delivers

the highest average accuracy. Further, it deteriorates performance. In dataset IV-2a, computation time decreases from 4.6 s to 1.28 s as we decrease the cluster index from 11 to 5, and accuracy rises from 82.37% to 91.8% (NDI = 0.67). In the cross-subject case, accuracy sees a notable increase up to 6 subsets, but then it remains constant and even decreases after 6 subsets, which is shown in Fig. 5.10 (b). Hence, dataset IV-2a works best with 6 subsets in our experimentation, while dataset III-4a benefits most from 9 subsets. The further study compares the performance of optimized and non-optimized FTFG features formulation after and before the proposed ESSDM method in the specific-subject scenario, which is shown in Table 5.4. This study compares the accuracy of ESSDM and non-optimized SWD with different optimizing parameters (threshold parameter, cluster size, and NDI) and the comparison results are presented in Table 5.4. The NDI was also computed between 0.3 and 0.7, but  $NDI > 0.6$  has produced the best classification results in terms of ACCavg and kappa with optimized FTFG features for all proposed models. Note that the dimensionality of the feature vector for each model was determined according to the training performance. Table 5.3 shows that the FTFG feature obtained using the ESSDM considerably improves the classification accuracy. After enhanced features, the multi-class accuracy performance of these models in IV-2a is improved by 2.7% ( $k = 0.88$ ,  $p = 0.02$ , subject A01), 4.1% ( $k = 0.84$ ,  $p = 0.01$ , subject A02), 5.8% ( $k = 0.85$ ,  $p = 0.01$ , subject A04), and 15.6% ( $k = 0.80$ ,  $p = 0.01$ , subject A06). Similarly, in III-4a, improvement were observed by 3.2% ( $k = 0.92$ ,  $p = 0.01$ , subject B01), 4.6% ( $k = 0.94$ ,  $p = 0.01$ , subject B05). Moreover, the classification performance for subjects who performed poorly in the baseline models was significantly improved. When we looked more closely at the results in Table 5.4, we discovered that ESSDM-FTFG features, specifically subjects A06, A07, and A08, significantly improved the classification accuracy of the subjects with weak accuracies in baseline models in IV-2a. Even the statistical analysis (Wilcoxon signed rank tests) proved that ESSDM achieved significant improvement in classification performance with enhanced features at ( $p < 0.02$ ) optimum level for almost all the subjects could even exceed 94.12%. Based on the experimental outcomes, it is clearly stated that MI-EEG actions can be predicted with reasonably high ACCavg with fewer layer networks and FTFG features by ESSDM, thereby reducing the computational burden and significantly enhancing the classification perfor-

mance for different MI-EEG recognition classes. Figure 5.11 summarises the comparative analysis of the proposed ESSDM-based model with state-of-the-art baseline models. In overall, the proposed model ESSDM-FTFG-CNET achieved the highest performance and demonstrated effective models with enhanced features for classifying bi-class and multi-class MI-EEG signals in both datasets [139, 140], respectively. Moreover, the classification rate for other classes was also improved to a certain extent in comparison to existing baseline models [24, 27, 28]. The table shows that the average accuracies were higher in specific-subject tests (3.8% in bi-class MI tasks and 3.84% in multi-class MI tasks) and cross-subject ( $ACC_{avg} = 12.1\%$ ) than the benchmark baseline models, which proves that the proposed EESDM can greatly improve MI-EEG decoding performance, indicating the effectiveness for a feasible BCI solution.

## 5.5 Summary

In this study, the ESSDM and FTFG features-based classification models have been proposed for automatic recognition of MI classes from multichannel MI-EEG signals. To solve the threshold problem, ESSDM has adopted a novel nature-inspired meta-heuristic optimization algorithm-SHO, to select the optimal SWD thresholds and improve mode aliasing issues in signal specific. The newly designed fitness function enhances threshold parameter tuning that optimizes the decomposition effect, avoids excessive iterations, and offers optimal modes with fast convergence that provide more insight information from the MI-EEG signals. Thus, benefiting from the merits of SFT spectrum, modified swarm filter bank and SHO, ESSDM offers a valuable approach for feature extraction from multicomponent modulation signal, especially in the case of nonstationary multichannel MI-EEG signals. Experimentally, it is proved that this evolutionary method holds potential for the analysis and synthesis of various biomedical nonstationary signals. The results of the study demonstrate the effectiveness of the proposed methodology in classifying MI in 4-class and 2-class recognition problems and achieved higher classification accuracy than baseline models on two public datasets, namely, BCI IV-2a and BCI IV-3a. Additionally, the Wilcoxon statistical significance tests reveal that utilizing sub-clustered FTFG features ( $p < 0.01$ ) with a

limited number of EEG channels, the proposed ESSDM-based models deliver more efficient MI-EEG classification despite having a large number of channels in BCI IV-2a. These findings are significant and demonstrate the potential tool for MI-EEG recognition as an intuitive BCI for real-world use. In future, we will extend this problem to more classes of MI problems with optimizing channels and threshold improvement in cross-modalities intended to deploy for feasible clinical solutions.



## **Chapter 6**

# **Clustering SSDM with MI-EEG based Upper Limb Movement Recognition for BCI Application**

This chapter introduces the CSSDM to improve upper limb movement recognition in BCI systems using MI-EEG data. CSSDM addresses issues in processing nonhomogeneous EEG signals in existing SWD by clustering them into homogeneous sets with DBSC and CMI criteria. It integrates SFT spectrum estimation and modified swarm filters to enhance TF analysis and extract optimal OCs. The designing of the CSSDM is explained in this chapter. This chapter also presents an automated upper limb movement recognition BCI framework using novel TFGS features and classifiers (SAViT and ConvNet (AlexNet)) to validate the performance of CSSDM on the BNCI Horizon 7-class database. These frameworks enhance classification by utilizing optimal channels and rhythm-specific analysis and deliver improved accuracy and efficient feature generalization across channels. The classification performance of the proposed frameworks is illustrated in detail in later sections of this chapter.

## 6.1 Introduction

MI-assisted BCI offers an effective rehabilitation technique for people with neuromuscular diseases to restore or replace impaired communication or motor function [143, 144]. In past research, MI-EEG has been explored over the other modalities for the restoration of upper limb movement execution (ME) due to its efficacy in fine-grained analysis and exhibits stronger spectral and amplitude correlation during the execution of upper limb ME, thus leads to decode ME significantly during movement imagination [145].

### 6.1.1 Existing baseline approaches

Over the past decade, numerous signal processing and BCI algorithms-based methodologies have been proposed for automatic upper limb ME recognition using the MI-EEG signals [117, 127, 143, 144, 145, 146, 147, 148]. It primarily contributes to two major approaches. The first approach deals with feature extraction based on spatial and spectral filtering, encompassing methods such as theoretic CSP features [143], optimal neuronal components (ONC) [146], minimum distance to mean (MDM) [147], Riemannian distance (RD) [147], critical features (CF) [148], source power comodulation (SPoC) [146], and multi-class filter bank task-related component analysis (mFBTRCA) [145]. The second approach involves shallow learning and deep learning-based classification frameworks: Shallow CNN (ConvNet) [127], deep ConvNet [127], 10-layer one-dimensional ConvNet [148], EEGWaveNet [144], multiscale ConvNet, (LDA) [143], and tangent space linear discriminant analysis (TSLDA) [147]. Most studies [127, 144, 148] employ all channels and demonstrate limited performance for deployable BCI systems due to subject-specific analysis and lack of mutual feature generalization across-channel. Secondly, it ignores the significance of the cognitive neural effects of brain regions with the selection of appropriate channel localization that is highly correlated to upper limb MI tasks. Therefore, it is crucial to design BCI as an optimal solution with minimal channels and enhance mutual feature characteristics for efficient upper limb MI task classification performance.

To address these problems, the novel decomposition method namely, the CSSDM has been proposed to efficiently select channels in the most relevant cortical regions of the

MI tasks and improve homogeneous spectral characteristics across-channel with optimal mode extraction. Further, the TFGS features are computed from the decomposed modes using CSSDM and classify the 7-class upper limb MI-EEG tasks with a self-attention vision transformer (SAViT) [149] and ConvNet (AlexNet) [97] classifiers. The developed CSSDM-based classification frameworks have shown significant improvement in classifying 7-class upper limb MI tasks on the database [150].

## 6.2 Proposed methodology

The objective of this study is to develop an automated ESSDM-based MI recognition system using MI-EEG signals. The block diagram of the proposed automated CSSDM-based classification system using MI-EEG signals is shown in Fig. 6.1. The description of the proposed framework using the CSSDM method, the dataset used, and all the methods involved are briefly explained in the following sub-section.

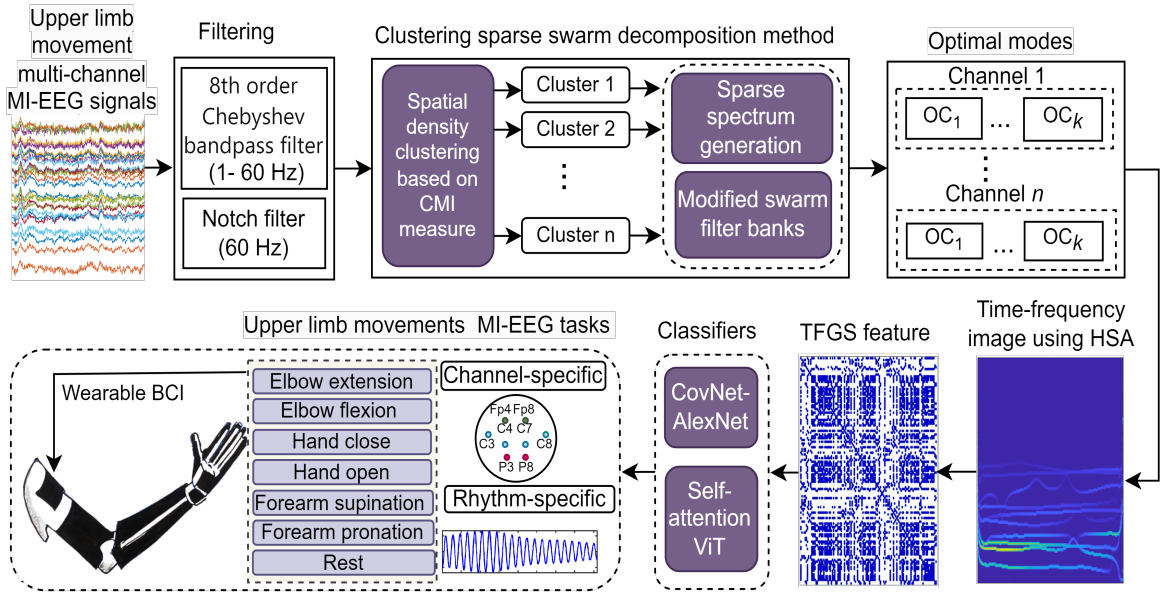


Figure 6.1: Block diagram of the proposed automated CSSDM-based upper-limb movement recognition from MI-EEG sensors data.  $OC_k$  denotes the oscillatory component.



### 6.2.1 Dataset used

In this work, we have used the upper limb movement BNCI Horizon dataset (001–2017) [150], which is openly available at <https://bnci-horizon-2020.eu/database/data-sets>. The dataset included upper limb movement EEG data from 15 subjects, who performed seven types of MI tasks: Elbow extension (EE), elbow flexion (EF), hand close (HC), hand open (HO), forearm supination (FS), forearm pronation (FP), and rest. The data were acquired using a 61-channel EEG with 72 trials per class. EEG signals were recorded from 61 channels across frontal, central, parietal, and temporal regions based on the international 10–20 system. For preprocessing, we applied an 8th-order Butterworth bandpass filter with a pass-band from 0.01 Hz to 80 Hz to remove noise and sampled at a rate of 512 Hz. A notch filter at 50 Hz was used to suppress power line interference.

### 6.2.2 Clustering sparse swarm decomposition method

The CSSDM is proposed to address the limitation of processing nonhomogeneous signals from a large number of channels and extract homogeneous modes. It consists of a DBSC-based clustering technique, optimized spectrum estimation, and the design of an iterative swarm filtering method. The unique advantage of CSSDM is to extract homogeneous spectral characteristics across nonhomogeneous multichannel EEG data with homogeneous channel selection for efficient decomposition. The proposed CSSDM is explained by the following stages:

Stage 1: For a given non-homogeneous MI-EEG  $n$ -channel signal  $s(n)$ , the DBSC-based clustering technique is applied to extract unique optimum sample points (OSPs) from homogeneous channel data. In this channel clustering technique, the CMI measure is used to obtain the set of OSPs across  $CL_k$  homogeneous cluster. A mathematical expression to obtain CMI is given in Eq. (6.1). These obtained OSPs of homogeneous EEG clusters show better stationary characteristics and contribute homogeneous spectral information across all channel clusters. The DBSC-based clustering technique along with designed CMI criteria is explained in Section 6.2.2.1.

Stage 2: The SFT spectrum estimation model is then applied to each obtained OSP to

get an optimized spectrum  $\psi(\tilde{c}_f)$ . The detailed description of the SFT model is described in Section 2.2.1 of Chapter 2.

Stage 3: The SFT spectrum of each OSP is considered to extract the decomposed optimal modes  $Y_q$  using a swarm filter bank. A mathematical expression for obtaining the decomposed modes using SSDM is provided in Section 2.2.3 of Chapter 2. The detail of clustering SSDM is given in Algorithm 6.1.

### 6.2.2.1 DBSC algorithm with CMI measure

The designed DBSC algorithm with CMI measure determines a small number of homogeneous cluster sets of unique canonical pairs from non-homogeneous EEG channels and delivers OSPs with optimal channel selection and less iterations [151]. It finds minimum distance relation between two canonical data set pairs adaptively against the threshold parameters criteria: Circle centred at each point with a radius (Epr) and minimal OSPs (OSPmin). In our work, DBSC-CMI segments the non-homogeneous EEG time series into homogenous clusters and extracts OSPs according to a CMI measure. CMI finds the degree of similarity between two homogeneous clusters ( $u$  and  $v$ ) that represent the same spectral characteristics by computing FT spectra, which is given as follows:

$$\Omega^{\text{CMI}}(U, V) = \frac{\sum_{h=1}^r \sum_{l=1}^s k_{(h,l)}^{(u,v)} \log \frac{k_{(h,l)}^{(u,v)}}{k_h^u k_l^v}}{\sqrt{(\sum_{h=1}^r k_h^u \log \frac{k_h^u}{k}) (\sum_{l=1}^s \log \frac{k_l^v}{k})}} \quad (6.1)$$

Where  $k$  is the OSP points in a single EEG channel of size  $N$ .  $U = (U_1, U_2, U_3, \dots, U_r)$  and  $V = (V_1, V_2, V_3, \dots, V_s)$  represent the two channels with different length  $r$  and  $s$ , respectively.  $k_h^x$  and  $k_l^y$  represent the number of sample points in the clusters  $h$  and  $l$ , respectively.  $k_{(h,l)}^{(u,v)}$  is the number of intersected OSPs in the channel clusters  $U_h$  and  $V_l$ . In contrast to the raw EEG signals, the OSPs of each channel cluster form the homogeneous EEG cluster, which exhibits better stationary characteristics patterns specific to the MI-EEG classes.

### 6.2.3 Feature formulation and classification

In our work, a new feature namely, the TFGS feature has been derived from computed CSSDM modes-based TF image using HSA [41]. To obtain the TFGS features, the TF images of size  $875 \times 756 \times 3$  have been processed using the graph spectral filter ( $G_f = \sum_{n=1}^Q g_n \langle f_g, \phi_n \rangle \phi_n$ ) to extract the  $Q$  feature attributes and generate the sparse adjacency matrix.  $G_f$  is derived from TF coefficients  $(g_n)_{n=1}^Q$  and unidirectional computed graph signal  $f_g$ .  $\phi_{n=1}^Q$  is an orthonormal basis. With optimal clustering channel, the proposed TFGS features show its significance by exhibiting discriminative neural activity patterns related to 7-class upper limb MI-EEG tasks, which is demonstrated in Fig. 6.2.

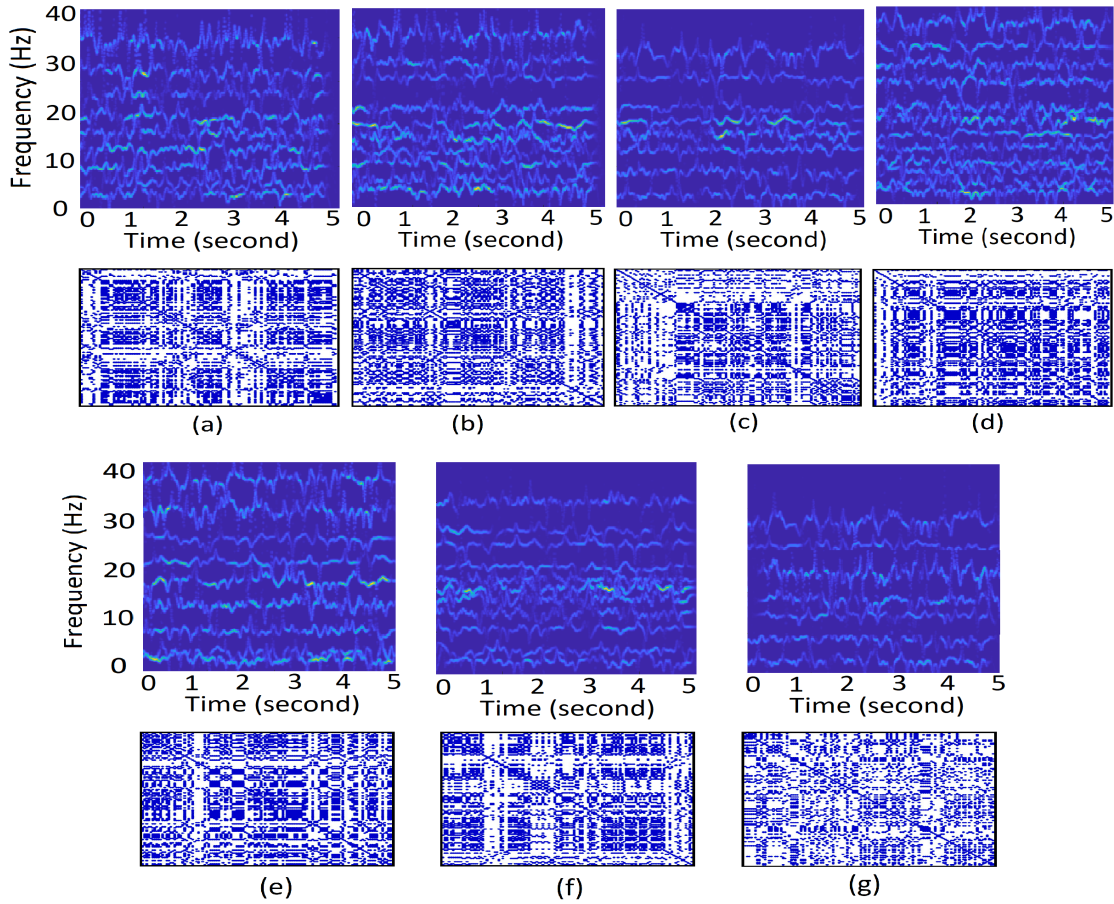


Figure 6.2: The extracted TFGS features from TF images for (a) EE, (b) EF, (c) HC, (d) HO, (e) FS, (f) FP, and (g) rest MI-EEG classes.

Further TFGS features were fed to two classifiers: SAViT and ConvNet (AlexNet). ConvNet (AlexNet) was selected due to its proven ability for spatial feature modeling. Whereas

SAViT was chosen because of its efficient sequence layer modeling and minimal training time compared to hierarchical transformer designs such as swin SViT and vision transformer Lite (ViT-Lite) [152].

## 6.3 Results and Discussions

To test the effectiveness of CSSDM-TFGS features, the experiments were conducted using the subject-independent 10-fold cross-validation analysis. To address the challenges of upper limb MI task classification, this study involves two major aspects: (i) Firstly, it analyses the effect of optimal cluster-based channel selection on classification performance. (ii) Secondly, it investigates the significance of features complexity using CSSDM bands based on (theta ( $\theta$ : 4–8 Hz), alpha ( $\alpha$ : 8–13 Hz), and beta ( $\beta$ : 13–35 Hz)) rhythms selection with distinct cortical regions. Using SAViT and ConvNet (AlexNet) classifiers, two classification models, CSSDM-TFGS-SAViT and CSSDM-TFGS-ConvNet, are developed and compared to baseline models [143, 144, 146, 147, 148] on performance metrics namely, ACC, SEN, SPE, and Cohen’s kappa coefficient ( $k_p$ ).

Table 6.1: The obtained classification performance (in %) for different optimal channel selection of cross-subject on the BNCI Horizon 7-class database.

Optimal channel	Feature samples	Cluster	CSSDM-TFGS-SAViT				CSSDM-TFGS-ConvNet			
			ACC (%)	SEN (%)	SPE (%)	Kappa	ACC (%)	SEN (%)	SPE (%)	Kappa
4	5096	3	40.61	42.71	41.64	0.417	38.12	38.65	39.02	0.382
8	10192	6	<b>49.02</b>	<b>50.62</b>	<b>50.46</b>	<b>0.503</b>	<b>47.69</b>	<b>45.94</b>	<b>47.13</b>	<b>0.481</b>
14	20384	8	39.85	44.23	43.77	0.447	43.57	42.25	40.21	0.447
21	26754	11	40.94	45.61	44.79	0.423	42.71	39.94	38.03	0.419
28	35672	13	39.89	41.13	41.18	0.408	42.04	41.35	39.51	0.424
35	44590	17	44.76	45.84	45.30	0.451	40.02	41.10	36.67	0.401
48	61152	19	39.88	46.66	46.34	0.471	41.34	39.56	40.09	0.393
54	68796	21	43.17	49.34	48.54	0.489	39.01	38.21	37.82	0.409
61	77714	23	43.53	47.38	47.42	0.478	40.70	40.62	39.35	0.404

### 6.3.1 Effect of optimal clustering channel-based features

Table 6.1 shows the cross-subject classification performance of the upper limb MI-EEG class on obtained TFGS features for optimal cluster-wise channel selection. Features are chosen based on the lowest student’s t-test-based p-value ( $p < 0.05$ ). It shows that the

---

**Algorithm 6.1** Computation of CSSDM-based TFGS features

---

**Input:** Non-homogeneous MI-EEG  $n$ -channel signal  $s(n)$

**Initialize** : Iteration  $it = 0$ ; Channel  $N = 61$ ; Cluster  $k$  with different class labels;  
Cluster sample points  $minPts = 0$ ; Radius  $Epr = 0$ .

**// Stage 1: Apply DBSC for homogeneous channel clustering to get OSPs**

**while**  $N_i \neq 0$

$p_i = N_i$ ;  $minEpr = 0$ ; Assign current channel.

**for**  $k$  in 0: ( $minPts - 1$ ) **do**

$minEpr = (p_i)_{Max}$ ; Assign maximum sample points of  $N$ .

$minClusterId = 0$ ; Assign initial cluster identifier (id).

**if**  $|Epr(p_i)| \geq MinPts$  **then**

**if**  $minEpr \geq CMI$  **then** ; Apply CMI criteria.

$minEpr = Epr(p_i)$  ; Get intersected OSPs by finding neighbourhood  $Epr$ .

$minClusterId = p_i$

**end**

$CL_k = minEpr$ ; Generate homogeneous cluster  $CL_k$  with OSPs.

**end**

**// Design modified swarm filter bank and find optimal OCs.**

**// Stage 2: Computation of SFT spectrum and spectral boundary estimation** (refer Section 2.2.1)

**while**  $k \neq 0$

Compute SFT spectrum  $c_k = \psi(\tilde{c}_f)$  on each OSP from cluster  $CL_k$ .

$\omega_{dom}^q = \arg_{\omega} \max \left( E'_{c_k}(\omega) > P_{th} \right)$ ; Where  $E(\cdot)$  is energy of dominant spectrum  $\omega$ . The  $P_{th}$  is set as 0.1.

Find spectral boundaries  $B_i$  on the obtained SFT spectrum.

$B_i = \frac{Fc_k + Fc_{k+1}}{2}$ ;  $i = 1, 2, \dots, M - 1$ . Where  $Fc_k$  and  $Fc_{k+1}$  represent two successive center  $\omega$ .

**while**  $\hat{\omega}_i \neq 0$

**// Stage 3: Designing of SwF filter bank** (refer Section 2.2.3)

Compute swarm filter parameters  $\delta(\hat{\omega})$  and  $M(\hat{\omega})$  to get the desired filter response

$\delta(\hat{\omega}) = -1.5\hat{\omega}^2 + 3.454$

$M(\hat{\omega}) = \lfloor 33.46 \hat{\omega}^{-0.735} - 29.1 \rfloor$

Compute driving force ( $F_{Dr}$ ) and cohesive ( $F_{Coh}$ ) force.

$F_{Dr,i}^n = p_{prey}(n) - p_i(n - 1)$

$F_{Coh,i}^n = \frac{1}{M-1} \sum_{j=1, j \neq i}^M f(p_i(n - 1) - p_j(n - 1))$

Find position  $P_i(n)$  and velocity  $V_i(n)$  with respect to every swarm  $M$ .  $\delta = 0.1$  decides the swarm-pray mobility.

$p_i(n) = p_i(n - 1) + \delta V_i(n)$

$v_i(n) = v_i(n - 1) + \delta(F_{Dr,i}^n + F_{Coh,i}^n)$

Apply BD-based convergence criteria to get significant OCs.

$C_{BD} = -\ln \sum_{q=1}^d \sqrt{\left( |y_{q,\delta,M}(n)| \cdot |x_q(n)| \right)^2}$ ;  $d$  is number of OCs.

Where  $|x(n)|$  and  $|y_{\delta,M}(n)|$  are amplitude of the FT of the extracted OCs signal  $c_k(n)$  and  $Y_{\delta,M}(n)$ .

Find residual  $CL_k(n) = CL_k(n) - CL'_k(n)$

Extract  $q$  modes  $Y_{q,\delta,M} = H \sum_{i=1}^n p_i(n)$ ;  $H$  is scaling factor.

**// Stage 3: Estimation of TFGS feature**

Compute IA and IF using HSA [41] from each obtained mode  $m$ .

Find TFGS feature from graph adjacency matrix of TF image.

**Output:** The extracted TFGS features from the CSSDM modes.

---

proposed CSSDM-TFGS-SAViT model has delivered highest classification accuracy  $ACC = 49.02 \pm 0.61\%$  ( $SEN = 50.6\%$ ,  $SPE = 50.46\%$ ,  $k_p = 0.503$ , and  $p < 0.05$ ) with optimal channel selection (cluster = 6 and features = 10) without compromising performance. It also shows better discriminative mutual characteristics across all MI-EEG classes for all obtained features and outperformed in all models against all maximal 61-channel selection, which is shown in confusion matrix plots (Fig. 6.4).

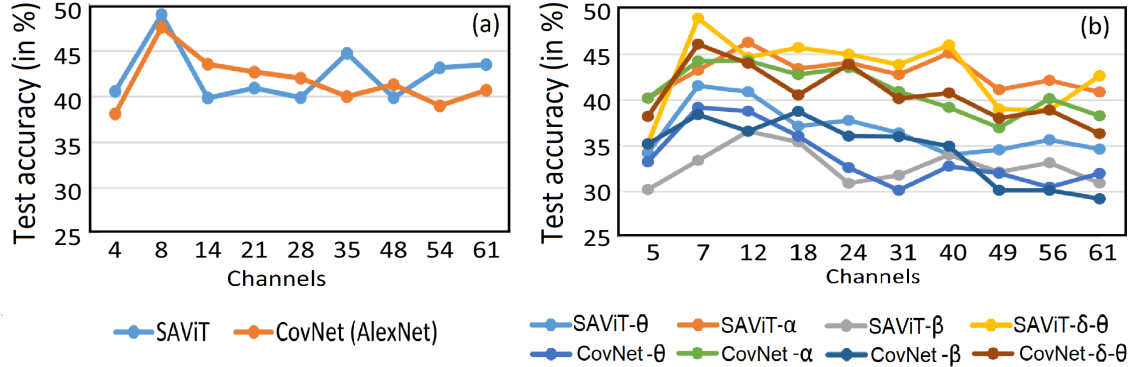


Figure 6.3: Performance comparison of proposed CSSDM-based models for optimal channel selection in (a) all rhythms and (b) rhythm-specific.

The second model, CSSDM-TFGS-ConvNet has performed slightly deteriorate performance with selected clusters are  $ACC = 47.69 \pm 0.45\%$  ( $SEN = 45.94\%$ ,  $SPE = 47.13\%$ ,  $k_p = 0.4815$ , and  $p < 0.08$ ). Figure 6.3 compares accuracy versus optimal channel selection with clusters. Even when the number of channels is reduced from 61 to 8, the proposed models have attained the highest accuracy against the baseline models.

Table 6.2: The cross-subject performance (in %) of obtained features from  $\theta$ ,  $\alpha$ , and  $\beta$  rhythms with distinct cortical regions on on the BNCI Horizon 7-class database.

Rhythm	Cortical region	Feature samples	Channel	CSSDM-TFGS-SAViT				CSSDM-TFGS-ConvNet			
				ACC (%)	SEN (%)	SPE (%)	Kappa	ACC (%)	SEN (%)	SPE (%)	Kappa
Theta	Full placement	12868	14	38.23	39.03	38.57	0.393	36.72	37.10	34.58	0.359
	FL, PL, OL lobes	11009	10	39.46	41.76	38.06	0.409	38.60	37.29	37.84	0.390
	FL, CL, OL lobes	8200	9	40.92	40.72	41.35	0.412	39.45	40.16	40.57	0.389
Alpha	Full placement	10740	11	44.15	42.45	41.69	0.451	43.23	44.02	41.08	0.430
	FL, PL, OL lobes	7254	8	44.69	44.99	46.03	0.445	44.02	41.98	42.76	0.437
	FL, CL, OL lobes	6328	6	46.31	45.61	44.11	0.463	45.19	46.64	44.03	0.454
Beta	Full placement	10090	12	32.17	32.97	30.14	0.339	29.97	28.59	30.80	0.305
	FL, PL, OL lobes	9740	10	31.54	33.84	34.34	0.306	28.56	29.97	30.96	0.294
	FL, CL, OL lobes	7902	9	36.59	38.72	36.99	0.381	38.03	36.50	35.99	0.391
Theta and Alpha	Full placement	16860	9	43.12	42.9	42.61	0.458	40.15	42.27	40.95	0.408
	FL, PL, OL lobes	12542	8	43.96	45.26	46.68	0.435	42.80	45.89	44.03	0.446
	FL, CL, OL lobes	10380	7	<b>48.68</b>	50.98	49.11	0.491	<b>47.69</b>	48.51	46.71	0.483

### 6.3.2 Effect of rhythm-based features

To investigate the significance of the feature complexity using rhythms ( $\theta$ ,  $\alpha$ , and  $\beta$ ) from CSSDM bands, statistical analysis is conducted on distinct cortical regions namely, frontal lobe (FL), central lobe (CL), parietal lobe (PL), and occipital lobe (OL). Table 6.2 shows that the CSSDM-based features computed from ( $\theta$ ,  $\alpha$ , and  $\beta$ ) rhythms with optimal clustering provide a significant contribution for the recognition of 7-class upper limb MI-EEG signals. The table indicates that the optimal cluster selection consistently retains most of the significant channels on the FL, CL, and OL lobes. Compared with other brain regions, the FL and CL lobes significantly contribute to the classification of MI tasks and delivered the highest performance (channel = 8, ACC =  $48.68 \pm 1.09\%$ ) in CSSDM-TFGS-SAViT model with cortical regions (FL, CL, and OL) in the rhythms ( $\theta$  and  $\alpha$ ). When combining FL and CL lobes channels from optimal clustering ( $k = 7$ ), the accuracy is improved from 42% to 49% and 40% to 48% in the case of CSSDM-TFGS-SAViT and CSSDM-TFGS-ConvNet, respectively. Similarly, combining the channels of the FL, PL, and OL lobes in  $\alpha$  rhythm has shown significant accuracy improvement from 38% to 44% ( $p < 0.01$ ). Whereas  $\beta$  rhythm-based features ( $p > 0.05$ ) have not performed significantly.

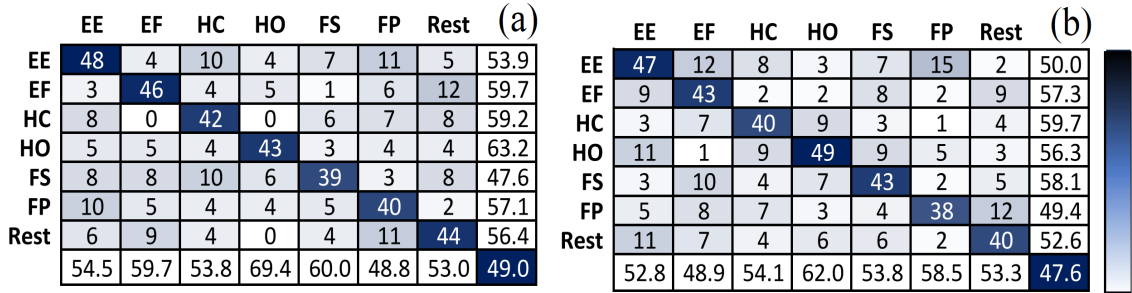


Figure 6.4: The obtained confusion matrix using (a) CSSDM-TFGS-SAViT and (b) CSSDM-TFGS-ConvNet models across all subjects.

The performance comparison in Table 6.3 reveals that the proposed CSSDM-based models deliver the highest classification performance despite a large number of classes, proving suitability for practical BCI systems. The proposed CSSDM method effectively captures homogeneous spectral characteristics in multichannel with optimal mode extraction and improves TF analysis. However, it requires precise parameter tuning in the designed DBSC scheme.

Table 6.3: Performance comparison with the existing baseline models.

Reference	Methodology	Channel used	Cross- validation	Accuracy (mean $\pm$ std)	p-value
Grosse-Wentrup et al. [143]	Theoretic CSP, LDA	60	NA	23.13 $\pm$ 5.51	< 0.0037
Dähne et al. [146]	SpoC, Ridge	58	10-fold	22.07 $\pm$ 4.97	< 0.0006
Barachant et al. [147]	MDM, RD	22	30-fold	25.50 $\pm$ 7.00	< 0.0032
Barachant et al. [147]	TSLDA	22	30-fold	27.12 $\pm$ 6.31	< 0.0053
Schirrmeister et al. [127]	Shallow ConvNet	22	NA	25.90 $\pm$ 6.45	< 0.1258
Schirrmeister et al. [127]	Deep ConvNet	22	NA	35.29 $\pm$ 7.40	< 0.1287
Thuwajit et al. [144]	Multiscale ConvNet, EEG-WaveNet	60	10-fold	20.27 $\pm$ 5.69	< 0.0005
Mattioli et al. [148]	CF, HopeFullNet	64	NA	33.77 $\pm$ 7.70	< 0.0835
Jia et al. [145]	mFBTRCA	60	10-fold	41.93 $\pm$ 0.78	NA
<b>Proposed method 1</b>	<b>CSSDM-TFGS-SAViT</b>	<b>8</b>	10-fold	<b>49.02<math>\pm</math>0.61</b>	<b>&lt; 0.05</b>
<b>Proposed method 2</b>	<b>CSSDM-TFGS-ConvNet</b>	<b>9</b>	10-fold	<b>47.69<math>\pm</math>0.14</b>	<b>&lt; 0.08</b>

Note: std: Standard deviation, NA: Not available

## 6.4 Summary

In this work, we have proposed a CSSDM-based classification framework for the improvement of automatic recognition of multi-class upper limb MI-EEG tasks. The CSSDM adopts the DBSC algorithm, modified swarm filtering and SFT spectrum to solve non-homogeneous EEG decomposition issues in the existing SWD method and improves homogeneous spectral characteristics across channels with optimal mode extraction and TF analysis. The proposed CSSDM-based classification frameworks are validated and outperformed against baseline models. The presented analysis of TFGS features complexity on the optimal channel and rhythm selection demonstrates a significant role in upper limb MI-EEG classification and proved to be an efficient upper limb BCI tool for patients with neuromuscular diseases. In the future, it would be interesting to extend CSSDM to analyze multichannel signals of diverse MI-EEG motory actions for different BCI applications.





# Chapter 7

## Conclusions and future scope

The EEG and MEG are found to be the most useful brain signals that assist in the diagnosis of neurological diseases and BCI applications. This thesis focuses on the development of a novel extension of the univariate SWD-based nonstationary signal-processing techniques to decompose multichannel signals for brain signal analysis. We have developed new signal processing techniques, including SSDM, MSSDM, ESSDM, and CSSDM for robust nonstationary brain signal analysis. These techniques adopt newly designed optimized SFT spectrum representation and filter banks for efficient extraction of components from multi-channel nonstationary signals and time-frequency improvement. The proposed SSDM-based methods are then used to design different frameworks for automatic diagnosis of neurological disease detection and BCI frameworks. The developed frameworks include cognitive visual object recognition using EEG and MEG signals, EEG-based motor imagery recognition, EEG-based imagined speech task detection, and EEG-based upper limb movements detection. The proposed frameworks have delivered the highest classification performance as compared with other state-of-the-art approaches presented for neurological disease detection and BCI frameworks. We have presented five objectives in Chapter 1, and the findings of each objective are consolidated and presented in this chapter. Additionally, this chapter includes a separate section highlights potential directions for future research for this thesis.

## 7.1 Conclusions

In Chapter 2, we proposed a novel univariate SSDM which is an extension of SWD. It adopts improved SFT spectrum estimation model and designed swarm filter banks. The proposed SSDM method delivers substantial improvements in time-frequency analysis and mode extraction from nonstationary signals compared to two existing methods, namely SWD and HHT. Using SSDM, we have also designed an automated framework for the sleep apnea detection system using newly formulated features (FF and TFR) and classifiers (CNN and BiLSTM) on a polysomnographic EEG database. The automated sleep apnea detection system outperformed state-of-the-art approaches in subject-independent cross-validation on a cross-subject dataset.

In Chapter 3, a new adaptive MSSDM has been proposed for multivariate analysis of nonstationary EEG-MEG signals, which is an extension of the univariate SSDM method. The MSSDM method effectively captures homogeneous spectral characteristics in multi-channel with channel-aligned mode extraction and improves common multivariate mutual patterns related to visual objects. In addition, we have compared the proposed method with the univariate SSDM approach and direct rhythms analysis which is computed using a BPF. The MSSDM has been then used for decoding different visual object categories for BCI applications. The proposed framework has outperformed other existing methods in the literature through subject-independent analysis.

In Chapter 4, we presented a novel MSSDM-based approach to extract multivariate joint oscillatory features from multichannel EEG signals. JTF multivariate features are generated using the JIF and JIA functions from extracted CAOCs. We developed an imagined speech task detection framework that utilizes JTFDF features from four pre-trained neural networks and an SVM classifier for cross-channel IMS-EEG signals in BCI applications. Two feature mapping strategies (CCA and HDC) were employed to generate the most discriminative features. The proposed framework has better classification performance than other existing methods using a cross-subject dataset.

In Chapter 5, a novel decomposition technique ESSDM has been proposed to solve uniform decomposition and hyperparameter selection issues in the existing SWD method.

Using SHO with a KLD-based fitness function, a modified SwF filter bank, and an SFT spectrum, the ESSDM enhances hyperparameter selection and decomposition performance. In this work, a new FTFG feature has been designed using the subspace clustering technique. Additionally, this method is then used to develop a framework for the automatic classification of motor imagery tasks from MI-EEG signals for BCI applications. The proposed framework not only outperforms existing works in the specific-subject classification but also seamlessly adapts to cross-subject classification on two well-known MI-EEG datasets.

In Chapter 6, CSSDM is introduced to improve upper limb movement recognition in BCI systems using MI-EEG data. CSSDM addresses issues in processing nonhomogeneous EEG signals in existing SWD by clustering them into homogeneous channel sets with DBSC and CCA-assisted CMI criteria and delivering optimal oscillatory components. The CSSDM has been then used for an automated upper limb movement recognition BCI framework using novel TFGS features and classifiers (SAViT and ConvNet (AlexNet)) on the BNCI Horizon 7-class database. The proposed frameworks demonstrate superior classification performance compared to all existing methods in the literature.

## **7.2 Future directions**

The accurate analysis of brain signals using EEG and MEG for diagnosing neurological diseases and BCI is an emerging research field that has the potential to transform clinical practices and offer deeper insights into these neurological conditions. Although our current solutions have shown promising results and practical applications, there is still a significant need to enhance existing methodologies to overcome certain issues. A few of the future directions are as follows:

- The proposed concept can be further explored for developing signal analysis methods for signal decomposition, time-frequency representation, and feature extraction for various applications.
- The proposed methods in this thesis have been studied on small clinical datasets. In the future, the performance of these methods can be validated on larger and diverse

EEG datasets with varying sampling rates to develop more robust, generalized, and accurate models that enhance decision-making in real-time clinical applications.

- In multichannel analysis, MSSDM has been explored primarily for EEG and MEG modalities which limit its application scope. Therefore, it is necessary to extend the MSSDM technique to other multichannel physiological signals, such as EOG and EMG for effective analysis and wider practical applications.
- In our proposed neurological disease detection and BCI frameworks, we developed classification model using EEG signals. To enhance feasibility, it is needed to investigate multimodal frameworks that combine EEG with other physiological signals to enhance BCI performance and accuracy of disease detection.
- The performance of the proposed methods such as SSDM, MSSDM, ESSDM and CSSDM solely depends on the preprocessing of EEG signals. Even though a preprocessing filter has already been applied in our work there is a need to develop new preprocessing techniques for EEG signals to further reduce noise and eliminate artifacts for more accurate performance in real-time signal classification.
- Due to the iterative nature of the SSDM approach, its application is limited. Therefore, there is scope to reduce the computational complexity of the SSDM method through other optimization approaches.
- Explore ESSDM and CSSDM methods for improved channel selection and homogeneous signal processing across a broader range of biomedical applications.
- Novel features can be proposed based on developed signal processing algorithms for efficient representation of the nonstationary signals.
- Although the hybrid feature selection method has partially addressed the dimensionality issue, further research is needed to identify the most significant features using more robust feature learning methodologies.
- In the future, deploy the developed frameworks on dedicated hardware for real-time

applications including portable medical devices to ensure accurate disease diagnosis in remote areas.



# Bibliography

- [1] A. Daamouche, L. Hamami, N. Alajlan, F. Melgani, A wavelet optimization approach for ECG signal classification, *Biomedical Signal Processing and Control* 7 (2012) 342–349.
- [2] Q. Huang, Q. Li, M. Ran, X. Liu, Y. Zhou, Threshold-optimized swarm decomposition using grey wolf optimizer for the acoustic-based internal defect detection of arc magnets, *Shock and Vibration* 2021 (2021) 21.
- [3] R. M. Cichy, D. Pantazis, A. Oliva, Resolving human object recognition in space and time, *Nature Neuroscience* 17 (2014) 455–462.
- [4] N. C. Kong, B. Kaneshiro, D. L. Yamins, A. M. Norcia, Time-resolved correspondences between deep neural network layers and EEG measurements in object processing, *Vision Research* 172 (2020) 27–45.
- [5] M. Guggenmos, P. Sterzer, R. M. Cichy, Multivariate pattern analysis for MEG: A comparison of dissimilarity measures, *Neuroimage* 173 (2018) 434–447.
- [6] Y. Zhang, P. Xu, P. Li, K. Duan, , Y. Wen, Q. Yang, T. Zhang, D. Yao, Noise-assisted multivariate empirical mode decomposition for multichannel EMG signals, *Biomedical Engineering Online* 16 (2017).
- [7] S. J. Redmond, C. Heneghan, Cardiorespiratory-based sleep staging in subjects with obstructive sleep apnea, *IEEE Transactions on Biomedical Engineering* 53 (2006) 485–496.



- [8] S. Taran, V. Bajaj, D. Sharma, Robust Hermite decomposition algorithm for classification of sleep apnea EEG signals, *Electronics Letters* 53 (2017) 1182–1184.
- [9] W. S. Almuhammadi, K. A. Aboalayon, M. Faezipour, Efficient obstructive sleep apnea classification based on EEG signals, in: *2015 Long Island Systems, Applications and Technology*, IEEE, 2015, pp. 1–6.
- [10] T. Mahmud, I. A. Khan, T. I. Mahmud, S. A. Fattah, W.-P. Zhu, M. O. Ahmad, Sleep apnea detection from variational mode decomposed EEG signal using a hybrid CNN-BILSTM, *IEEE Access* 9 (2021) 102355–102367.
- [11] C. Shahnaz, A. T. Minhaz, S. T. Ahamed, Sub-frame based apnea detection exploiting delta band power ratio extracted from EEG signals, in: *2016 IEEE Region 10 Conference (TENCON)*, 2016, pp. 190–193.
- [12] A. Bhattacharjee, S. Saha, S. A. Fattah, W.-P. Zhu, M. O. Ahmad, Sleep apnea detection based on Rician modeling of feature variation in multiband EEG signal, *IEEE Journal of Biomedical and Health Informatics* 23 (2019) 1066–1074.
- [13] T. Mahmud, I. A. Khan, T. Mahmud, S. A. Fattah, W.-P. Zhu, M. O. Ahmad, Sleep apnea event detection from sub-frame based feature variation in EEG signal using deep convolutional neural network, in: *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, 2020, pp. 5580–5583.
- [14] R. M. Cichy, D. Pantazis, Multivariate pattern analysis of MEG and EEG: A comparison of representational structure in time and space, *Neuroimage* 158 (2017) 441–454.
- [15] C. S. Dasalla, H. Kambara, M. Sato, Y. Koike, Single-trial classification of vowel speech imagery using common spatial patterns, *Neural Networks* 22 (2009) 1334–1339.
- [16] P. Saha, M. Mageed, S. Fels, Speak your mind! towards imagined speech recognition with hierarchical deep learning, *arXiv preprint arXiv:1904.05746* (2019).

- [17] G. A. P. Coretto, I. E. Gareis, H. L. Rufiner, Open access database of EEG signals recorded during imagined speech, in: 12th International Symposium on Medical Information Processing and Analysis, volume 10160, SPIE, 2017, p. 1016002.
- [18] J. S. García-Salinas, L. Villaseñor-Pineda, C. A. Reyes-García, A. A. Torres-García, Transfer learning in imagined speech EEG-based BCIs, *Biomedical Signal Processing and Control* 50 (2019) 151–157.
- [19] S. H. Lee, M. Lee, J. H. Jeong, S. W. Lee, Towards an EEG-based intuitive BCI communication system using imagined speech and visual imagery, in: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), IEEE, 2019, pp. 4409–4414.
- [20] C. Cooney, A. Korik, F. Raffaella, D. Coyle, Classification of imagined spoken word-pairs using convolutional neural networks, in: The 8th Graz BCI Conference, 2019, Verlag der Technischen Universitat Graz, 2019, pp. 338–343.
- [21] M. N. I. Qureshi, B. Min, H. J. Park, D. Cho, W. Choi, B. Lee, Multiclass classification of word imagination speech with hybrid connectivity features, *IEEE Transactions on Biomedical Engineering* 65 (2017) 2168–2177.
- [22] D. Y. Lee, M. Lee, S. W. Lee, Classification of imagined speech using siamese neural network, in: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2020, pp. 2979–2984.
- [23] Y. Zhang, G. Zhou, J. Jin, X. Wang, A. Cichocki, Optimizing spatial patterns with sparse filter bands for motor-imagery based brain–computer interface, *Journal of Neuroscience Methods* 255 (2015) 85–91.
- [24] B. E. Olivas-Padilla, M. I. Chacon-Murguia, Classification of multiple motor imagery using deep convolutional neural networks and spatial filters, *Applied Soft Computing* 75 (2019) 461–472.

- [25] X. Zhao, H. Zhang, G. Zhu, F. You, S. Kuang, L. Sun, A multi-branch 3D convolutional neural network for EEG-based motor imagery classification, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27 (2019) 2164–2177.
- [26] D. Li, J. Xu, J. Wang, X. Fang, Y. Ji, A multi-scale fusion convolutional neural network based on attention mechanism for the visualization analysis of EEG signals decoding, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 28 (2020) 2615–2626.
- [27] L. Wang, W. Huang, Z. Yang, C. Zhang, Temporal-spatial-frequency depth extraction of brain-computer interface based on mental tasks, *Biomedical Signal Processing and Control* 58 (2020) 101845.
- [28] D. F. Collazos-Huertas, A. M. Álvarez-Meza, G. Castellanos-Dominguez, Spatial interpretability of time-frequency relevance optimized in motor imagery discrimination using Deep&Wide networks, *Biomedical Signal Processing and Control* 68 (2021) 102626.
- [29] H. Yang, S. Sakhavi, K. K. Ang, C. Guan, On the use of convolutional neural networks and augmented CSP features for multi-class motor imagery of EEG signals classification, in: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015, pp. 2620–2623.
- [30] N. Lu, T. Li, X. Ren, H. Miao, A deep learning scheme for motor imagery classification based on restricted Boltzmann machines, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25 (2016) 566–576.
- [31] S.-H. Park, D. Lee, S.-G. Lee, Filter bank regularized common spatial pattern ensemble for small sample motor imagery classification, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26 (2017) 498–505.
- [32] S. Sakhavi, C. Guan, S. Yan, Learning temporal information for brain-computer interface using convolutional neural networks, *IEEE Transactions on Neural Networks and Learning Systems* 29 (2018) 5619–5629.

- [33] P. Wang, A. Jiang, X. Liu, J. Shang, L. Zhang, LSTM-based EEG classification in motor imagery tasks, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26 (2018) 2086–2095.
- [34] L. Yang, Y. Song, K. Ma, L. Xie, Motor imagery EEG decoding method based on a discriminative feature learning strategy, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29 (2021) 368–379.
- [35] L. Stanković, M. Daković, T. Thayaparan, Chapter 3 - non-stationary signal analysis time-frequency approach, *Academic Press Library in Signal Processing* 3 (2014) 27–142.
- [36] V. Gupta, R. B. Pachori, Classification of focal EEG signals using FBSE based flexible time-frequency coverage wavelet transform, *Biomedical Signal Processing and Control* 62 (2020) 102124.
- [37] F. J. Harris, Time domain signal processing with the DFT, *Handbook of Digital Signal Processing* (1987) 633–699.
- [38] T. Karp, N. J. Fliege, Modified DFT filter banks with perfect reconstruction, *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing* 46 (1999) 1404–1414.
- [39] R. L. Allen, D. Mills, *Signal analysis: time, frequency, scale, and structure*, John Wiley & Sons, 2004.
- [40] B. L. Sturm, A wavelet tour of signal processing, *Computer Music Journal* 31 (2007) 83–85.
- [41] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, H. H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, in: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, volume 454, 1998, pp. 903–995.

- [42] İ. Bayram, An analytic wavelet transform with a flexible time-frequency covering, *IEEE Transactions on Signal Processing* 61 (2012) 1131–1142.
- [43] I. W. Selesnick, Wavelet transform with tunable Q-factor, *IEEE Transactions on Signal Processing* 59 (2011) 3560–3575.
- [44] J. Gilles, Empirical wavelet transform, *IEEE Transactions on Signal Processing* 61 (2013) 3999–4010.
- [45] M. A. Cody, The wavelet packet transform: Extending the wavelet transform, *Dr. Dobb's Journal* 19 (1994) 44–46.
- [46] M. N. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, *IEEE Transactions on Image Processing* 14 (2005) 2091–2106.
- [47] I. W. Selesnick, R. G. Baraniuk, N. C. Kingsbury, The dual-tree complex wavelet transform, *IEEE Signal Processing Magazine* 22 (2005) 123–151.
- [48] A. Bhattacharyya, L. Singh, R. B. Pachori, Fourier–Bessel series expansion based empirical wavelet transform for analysis of non-stationary signals, *Digital Signal Processing* 78 (2018) 185–196.
- [49] I. Daubechies, J. Lu, H.-T. Wu, Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool, *Applied and Computational Harmonic Analysis* 30 (2011) 243–261.
- [50] I. Selesnick, Tunable-Q wavelet transform for signal analysis, *IEEE Transactions on Signal Processing* 59 (2011) 3560–3575.
- [51] J. McClellan, T. Parks, Eigenvalue and eigenvector decomposition of the discrete Fourier transform, *IEEE Transactions on Audio and Electroacoustics* 20 (1972) 66–74.
- [52] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, *IEEE Transactions on Signal Processing* 62 (2013) 531–544.

- [53] J. Schroeder, Signal processing via Fourier-Bessel series expansion, *Digital Signal Processing* 3 (1993) 112–124.
- [54] R. B. Pachori, P. Sircar, Analysis of multicomponent AM-FM signals using FB-DESA method, *Digital Signal Processing* 20 (2010) 42–62.
- [55] P. K. Chaudhary, R. B. Pachori, FBSED based automatic diagnosis of COVID-19 using X-ray and CT images, *Computers in Biology and Medicine* 134 (2021) 104454.
- [56] V. Gupta, R. B. Pachori, FBDM based time-frequency representation for sleep stages classification using EEG signals, *Biomedical Signal Processing and Control* 64 (2021) 102265.
- [57] Z. Wu, N. E. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, *Advances in Adaptive Data Analysis* 1 (2009) 1–41.
- [58] P. C. Chu, C. Fan, N. Huang, Compact empirical mode decomposition: an algorithm to reduce mode mixing, end effect, and detrend uncertainty, *Advances in Adaptive Data Analysis* 4 (2012) 1250017.
- [59] R. R. Sharma, R. B. Pachori, Time–frequency representation using IEVDHM–HT with application to classification of epileptic EEG signals, *IET Science, Measurement & Technology* 12 (2018) 72–82.
- [60] P. Jain, R. B. Pachori, An iterative approach for decomposition of multi-component non-stationary signals based on eigenvalue decomposition of the Hankel matrix, *Journal of the Franklin Institute* 352 (2015) 4017–4044.
- [61] P. Singh, S. D. Joshi, R. K. Patney, K. Saha, The Fourier decomposition method for nonlinear and non-stationary time series analysis, in: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, volume 473, 2017, p. 20160871.
- [62] J. B. Elsner, A. A. Tsonis, *Singular spectrum analysis: a new tool in time series analysis*, Springer Science & Business Media, 2013.

- [63] N. Rehman, H. Aftab, Multivariate variational mode decomposition, *IEEE Transactions on Signal Processing* 67 (2019) 6039–6052.
- [64] N. Rehman, D. P. Mandic, Multivariate empirical mode decomposition, in: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, volume 466, 2010, pp. 1291–1302.
- [65] A. Bhattacharyya, R. B. Pachori, A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform, *IEEE Transactions on Biomedical Engineering* 64 (2017) 2003–2015.
- [66] K. Das, R. B. Pachori, Schizophrenia detection technique using multivariate iterative filtering and multichannel EEG signals, *Biomedical Signal Processing and Control* 67 (2021) 102525.
- [67] E. Apostolidis, L. J. Hadjileontiadis, Swarm decomposition for nonstationary signal analysis, *IEEE Transactions on Biomedical Engineering* 64 (2017) 198–207.
- [68] Y. Miao, M. Zhao, V. Makis, J. Lin, Optimal swarm decomposition with whale optimization algorithm for weak feature extraction from multicomponent modulation signal, *Mechanical Systems and Signal Processing* 122 (2019) 673–691.
- [69] S. Alnuaimi, S. Jimaa, Y. Kimura, G. K. Apostolidis, L. J. Hadjileontiadis, A. H. Khandoker, Fetal cardiac timing events estimation from doppler ultrasound signals using swarm decomposition, *Mechanical Systems and Signal Processing* 10 (2019) 789.
- [70] K. Dragomiretskiy, D. Zosso, Variational mode decomposition, *IEEE Transactions on Signal Processing* 62 (2014) 531–544.
- [71] B. Samanta, C. Nataraj, Use of particle swarm optimization for machinery fault detection, *Engineering Applications of Artificial Intelligence* 22 (2009) 308–316.
- [72] J. Hammond, P. White, The analysis of non-stationary signals using time-frequency methods, *Journal of Sound and Vibration* 190 (1996) 419–447.

- [73] X. Jia, Y. Song, L. Xie, Excellent fine-tuning: From specific-subject classification to cross-task classification for motor imagery, *Biomedical Signal Processing and Control* 79 (2023) 104051.
- [74] S. P. Kelly, E. C. Lalor, C. Finucane, G. McDarby, R. B. Reilly, Visual spatial attention control in an independent brain-computer interface, *IEEE Transactions on Biomedical Engineering* 52 (2005) 1588–1596.
- [75] X. Wang, Y. Zhou, M. Shu, Y. Wang, A. Dong, ECG baseline wander correction and denoising based on sparsity, *IEEE Access* 7 (2019) 31573–31585.
- [76] N. Mourad, ECG denoising algorithm based on group sparsity and singular spectrum analysis, *Biomedical Signal Processing and Control* 50 (2019) 62–71.
- [77] M. Kaleem, A. Guergachi, S. Krishnan, Empirical mode decomposition based sparse dictionary learning with application to signal classification, in: *2013 IEEE Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE)*, 2013, pp. 18–23.
- [78] Y. Ding, I. W. Selesnick, Artifact-free wavelet denoising: Non-convex sparse regularization, convex optimization, *IEEE Signal Processing Letters* 22 (2015) 1364–1368.
- [79] B. P. Prathaban, R. Balasubramanian, Dynamic learning framework for epileptic seizure prediction using sparsity based EEG reconstruction with optimized CNN classifier, *Expert Systems with Applications* 170 (2021) 114533.
- [80] K. Brzostowski, Sparse reconstruction for enhancement of the empirical mode decomposition-based signal denoising, *IEEE Access* 8 (2020) 111566–111584.
- [81] L. Yang, S. Ding, F. Zhou, X. Yang, Y. Xiao, Robust EEG feature learning model based on an adaptive weight and pairwise-fused LASSO, *Biomedical Signal Processing and Control* 68 (2021) 102728.
- [82] B. Wang, Y. Zhang, W. Lu, J. Geng, A robust and efficient sparse time-invariant Radon transform in the mixed time–frequency domain, *IEEE Transactions on Geoscience and Remote Sensing* 57 (2019) 7558–7566.



- [83] F. Ahmed, P. Paromita, A. Bhattacharjee, S. Saha, S. Azad, S. Fattah, Detection of sleep apnea using sub-frame based temporal variation of energy in beta band in EEG, in: 2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), IEEE, 2016, pp. 258–261.
- [84] S. Saha, A. Bhattacharjee, M. A. A. Ansary, S. Fattah, An approach for automatic sleep apnea detection based on entropy of multi-band EEG signal, in: 2016 IEEE Region 10 Conference (TENCON) (2016) 420–423.
- [85] Y. Ichimaru, G. B. Moody, Development of the polysomnographic database on CD-ROM, *Psychiatry and Clinical Neurosciences* 53 (1999) 175–177.
- [86] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, H. E. Stanley, Physiobank, Physiokit, and Physionet: Components of a new research resource for complex physiologic signals, *Circulation*. 101 (2000) e215–e220.
- [87] M. K. Uçar, M. R. Bozkurt, C. Bilgin, K. Polat, Automatic detection of respiratory arrests in OSA patients using PPG and machine learning techniques, *Neural Computing and Applications* 28 (2017) 2931–2945.
- [88] M. Uddin, C. M. Chow, S. W. Su, Classification methods to detect sleep apnea in adults based on respiratory and oximetry signals: a systematic review, *Physiological measurement* 39 (2018) 03TR01.
- [89] X. Zhao, X. Wang, T. Yang, S. Ji, H. Wang, J. Wang, Y. Wang, Q. Wu, Classification of sleep apnea based on EEG sub-band signal characteristics, *Scientific Reports* 11 (2021) 5824.
- [90] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, K. Knight, Sparsity and smoothness via the fused lasso, *Journal of the Royal Statistical Society Series B: Statistical Methodology* 67 (2005) 91–108.
- [91] R. W. Schafer, What is a Savitzky-golay filter?[lecture notes], *IEEE Signal Processing Magazine* 28 (2011) 111–117.

- [92] J. Zhou, X.-m. Wu, W.-j. Zeng, Automatic detection of sleep apnea based on EEG detrended fluctuation analysis and support vector machine, *Journal of clinical monitoring and computing* 29 (2015) 767–772.
- [93] B. Hjorth, EEG analysis based on time domain properties, *Electroencephalography and Clinical Neurophysiology* 29 (1970) 306–310.
- [94] T. Higuchi, Approach to an irregular time series on the basis of the fractal theory, *Physica D: Nonlinear Phenomena* 31 (1988) 277–283.
- [95] M. Khateeb, S. M. Anwar, M. Alnowami, Multi-domain feature fusion for emotion classification using DEAP dataset, *IEEE Access* 9 (2021) 12134–12142.
- [96] A. Rényi, On measures of entropy and information, *Berkeley Symposium on Mathematical Statistics and Probability* 4.1 (1961) 547–561.
- [97] L. O. Chua, CNN: A vision of complexity, *International Journal of Bifurcation and Chaos* 7 (1997) 2219–2425.
- [98] M. Schuster, K. K. Paliwal, Bidirectional recurrent neural networks, *IEEE Transactions on Signal Processing* 45 (1997) 2673–2681.
- [99] J. Patterson, A. Gibson, *Deep learning: A practitioner’s approach*, O’Reilly Media, Inc (2017).
- [100] P. Kaushik, A. Gupta, P. P. Roy, D. P. Dogra, EEG-based age and gender prediction using deep BLSTM-LSTM network model, *IEEE Sensors Journal* 19 (2019) 2634–2641.
- [101] Y. Wang, Z. Xiao, S. Fang, W. Li, J. Wang, X. Zhao, Bi-Directional long short-term memory for automatic detection of sleep apnea events based on single channel EEG signal, *Computers in Biology and Medicine* 142 (2022) 105211.
- [102] H. Pitkänen, B. Duce, T. Leppänen, S. Kainulainen, A. Kulkas, S. Myllymaa, J. Töyräs, H. Korkalainen, Gamma power of electroencephalogram arousal is modu-

- lated by respiratory event type and severity in obstructive sleep apnea, *IEEE Transactions on Biomedical Engineering* 69 (2021) 1417–1423.
- [103] M. P. Bonnesen, H. B. D. Sorensen, P. Jennum, Mobile apnea screening system for at-home recording and analysis of sleep apnea severity, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018, pp. 457–460.
- [104] H. Phan, F. Andreotti, N. Cooray, O. Y. Chén, M. De Vos, DNN filter bank improves 1-max pooling CNN for single-channel EEG automatic sleep stage classification, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2018, pp. 453–456.
- [105] H. Park, D. Jeong, K. Park, Automated detection and elimination of periodic ECG artifacts in EEG using the energy interval histogram method, *IEEE Transactions on Biomedical Engineering* 49 (2002) 1526–1533.
- [106] D. Lai, M. B. B. Heyat, F. I. Khan, Y. Zhang, Prognosis of sleep bruxism using power spectral density approach applied on EEG signal of both EMG1-EMG2 and ECG1-ECG2 channels, *IEEE Access* 7 (2019) 82553–82562.
- [107] E. Choi, C. Lee, Feature extraction based on the Bhattacharyya distance, *Pattern Recognition* 36 (2003) 1703–1709.
- [108] J. Kalafatovich, M. Lee, S.-W. Lee, Decoding visual recognition of objects from EEG signals based on attention-driven convolutional neural network, in: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, 2020, pp. 2985–2990.
- [109] G. Smith, On the foundations of quantitative information flow, in: *International Conference on Foundations of Software Science and Computational Structures*, Springer, 2009, pp. 288–302.

- [110] A. Garde, L. Sörnmo, R. Jané, B. F. Giraldo, Correntropy-based spectral characterization of respiratory patterns in patients with chronic heart failure, *IEEE Transactions on Biomedical Engineering* 57 (2010) 1964–1972.
- [111] M. I. Khalid, T. Alotaiby, S. A. Aldosari, S. A. Alshebeili, M. H. Al-Hameed, F. S. Y. Almohammed, T. S. Alotaibi, Epileptic MEG spikes detection using common spatial patterns and linear discriminant analysis, *IEEE Access* 4 (2016) 4629–4634.
- [112] M. P. Do Carmo, J. Francis, *Riemannian geometry*, Springer, 1992.
- [113] F. Lotte, C. Guan, Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms, *IEEE Transactions on Biomedical Engineering* 58 (2010) 355–362.
- [114] Y. Zhang, C. S. Nam, G. Zhou, J. Jin, X. Wang, A. Cichocki, Temporally constrained sparse group spatial patterns for motor imagery BCI, *IEEE Transactions on Cybernetics* 49 (2018) 3322–3332.
- [115] H. Higashi, T. Tanaka, A. Funase, Classification of single trial EEG during imagined hand movement by rhythmic component extraction, in: *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, 2009, pp. 2482–2485.
- [116] D. Dash, A. Wisler, P. Ferrari, E. M. Davenport, J. Maldjian, J. Wang, MEG sensor selection for neural speech decoding, *IEEE Access* 8 (2020) 182320–182337.
- [117] R. B. Pachori, *Time-frequency analysis techniques and their applications*, CRC Press, 2023.
- [118] C. Cortes, V. Vapnik, Support-vector networks, *Machine learning* 20 (1995) 273–297.
- [119] 2020 International BCI, Competition., Accessed: Apr. 2020. [Online]. Available: <https://osf.io/pq7vb/>.

- [120] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
- [121] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems* 25 (2012).
- [122] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, K. Keutzer, Squeezenet: Alexnet-level accuracy with 50x fewer parameters and < 0.5 MB model size, *CoRR* (2016).
- [123] A. Iosifidis, A. Tefas, I. Pitas, On the kernel extreme learning machine classifier, *Pattern Recognition Letters* 54 (2015) 11–17.
- [124] D. Weenink, Canonical correlation analysis, in: Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam, volume 25, University of Amsterdam Amsterdam, 2003, pp. 81–99.
- [125] T. Kailath, The divergence and Bhattacharyya distance measures in signal selection, *IEEE Transactions on Communication Technology* 15 (1967) 52–60.
- [126] C. Cooney, R. Folli, D. Coyle, Optimizing layers improves CNN generalization and transfer learning for imagined speech decoding from EEG, in: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), IEEE, 2019, pp. 1311–1316.
- [127] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggenberger, M. Tangermann, F. Hutter, W. Burgard, T. Ball, Deep learning with convolutional neural networks for EEG decoding and visualization, *Human Brain Mapping* 38 (2017) 5391–5420.
- [128] Y. Liu, H. Zhang, M. Chen, L. Zhang, A boosting-based spatial-spectral model for stroke patients' EEG analysis in rehabilitation training, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 24 (2015) 169–179.

- [129] P. Yang, J. Wang, H. Zhao, R. Li, MLP with Riemannian covariance for motor imagery based EEG analysis, *IEEE Access* 8 (2020) 139974–139982.
- [130] C.-Y. Kee, S. Ponnambalam, C.-K. Loo, Binary and multi-class motor imagery using Renyi entropy for feature extraction, *Neural Computing and Applications* 28 (2017) 2051–2062.
- [131] S. K. Bashar, M. I. H. Bhuiyan, Classification of motor imagery movements using multivariate empirical mode decomposition and short time Fourier transform based hybrid method, *Engineering Science and Technology, an International Journal* 19 (2016) 1457–1464.
- [132] B. Xu, L. Zhang, A. Song, C. Wu, W. Li, D. Zhang, G. Xu, H. Li, H. Zeng, Wavelet transform time-frequency image and convolutional network-based motor imagery EEG classification, *IEEE Access* 7 (2018) 6084–6093.
- [133] C. Park, D. Looney, N. Rehman, A. Ahrabian, D. P. Mandic, Classification of motor imagery BCI using multivariate empirical mode decomposition, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 21 (2012) 10–22.
- [134] M. T. Sadiq, X. Yu, Z. Yuan, M. Z. Aziz, Motor imagery BCI classification based on novel two-dimensional modelling in empirical wavelet transform, *Electronics Letters* 56 (2020) 1367–1369.
- [135] E. Cuevas, F. Fausto, A. González, E. Cuevas, F. Fausto, A. González, The selfish herd optimizer, *New Advancements in Swarm Algorithms: Operators and Applications* (2020) 69–109.
- [136] F. Pérez-Cruz, Kullback-leibler divergence estimation of continuous distributions, in: *2008 IEEE International Symposium on Information Theory*, IEEE, 2008, pp. 1666–1670.
- [137] H. Zhou, J. M. Alvarez, F. Porikli, Less is more: Towards compact CNNs, in: *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV* 14, Springer, 2016, pp. 662–677.

- [138] S. Sabour, N. Frosst, G. E. Hinton, Dynamic routing between capsules, *Advances in Neural Information Processing Systems* 30 (2017).
- [139] G. Dornhege, B. Blankertz, G. Curio, K.-R. Müller, Boosting bit rates in noninvasive EEG single-trial classifications by feature combination and multiclass paradigms, *IEEE Transactions on Biomedical Engineering* 51 (2004) 993–1002.
- [140] C. Brunner, R. Leeb, G. Müller-Putz, A. Schlögl, G. Pfurtscheller, BCI competition 2008–graz data set a, *Institute for Knowledge Discovery (Laboratory of Brain-Computer Interfaces)*, Graz University of Technology 16 (2008) 1–6.
- [141] S. M. Vieira, U. Kaymak, J. M. Sousa, Cohen’s kappa coefficient as a performance measure for feature selection, in: *International Conference on Fuzzy Systems*, IEEE, 2010, pp. 1–8.
- [142] J. J. Liao, J. J. Luo, T. Yang, R. Q. Y. So, M. C. H. Chua, Effects of local and global spatial patterns in EEG motor-imagery classification using convolutional neural network, *Brain-Computer Interfaces* 7 (2020) 47–56.
- [143] M. Grosse-Wentrup, M. Buss, Multiclass common spatial patterns and information theoretic feature extraction, *IEEE Transactions on Biomedical Engineering* 55 (2008) 1991–2000.
- [144] P. Thuwajit, P. Rangpong, P. Sawangjai, P. Autthasan, R. Chaisaen, N. Banluesombatkul, P. Boonchit, N. Tatsaringkansakul, T. Sudhawiyangkul, T. Wilaiprasitporn, EEG-WaveNet: Multiscale CNN-based spatiotemporal feature extraction for EEG seizure detection, *IEEE Transactions on Industrial Informatics* 18 (2021) 5547–5557.
- [145] H. Jia, F. Feng, C. F. Caiafa, F. Duan, Y. Zhang, Z. Sun, J. Solé-Casals, Multi-class classification of upper limb movements with filter bank task-related component analysis, *IEEE Journal of Biomedical and Health Informatics* 27 (2023) 3867–3877.
- [146] S. Dähne, F. C. Meinecke, S. Haufe, J. Höhne, M. Tangermann, K.-R. Müller, V. V. Nikulin, SPOC: a novel framework for relating the amplitude of neuronal oscillations to behaviorally relevant parameters, *Neuroimage* 86 (2014) 111–122.

- [147] A. Barachant, S. Bonnet, M. Congedo, C. Jutten, Multiclass brain–computer interface classification by Riemannian geometry, *IEEE Transactions on Biomedical Engineering* 59 (2011) 920–928.
- [148] F. Mattioli, C. Porcaro, G. Baldassarre, A 1D CNN for high accuracy classification and transfer learning in motor imagery EEG-based brain-computer interface, *Journal of Neural Engineering* 18 (2022) 066053.
- [149] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo, Swin transformer: Hierarchical vision transformer using shifted windows, in: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 10012–10022.
- [150] P. Ofner, A. Schwarz, J. Pereira, G. R. Müller-Putz, Upper limb movements can be decoded from the time-domain of low-frequency EEG, *PloS One* 12 (2017) e0182578.
- [151] G. Mai, K. Janowicz, Y. Hu, S. Gao, ADCN: An anisotropic density-based clustering algorithm for discovering spatial point patterns with noise, *Transactions in Geographic Information Systems* 22 (2018) 348–369.
- [152] A. Vaswani, Attention is all you need, *Advances in Neural Information Processing Systems* (2017) 5998–6008.





## List of Publications

### A. Published

#### A1. In Refereed Journals

1. **S. V. Bhalerao** and R. B. Pachori, ESSDM: An enhanced sparse swarm decomposition method and its application in multi-class motor imagery-based EEG-BCI system, *IEEE Transactions on Automation Science and Engineering*, vol. 22, pp. 18509-18521, 2025.
2. **S. V. Bhalerao** and R. B. Pachori, Imagined speech-EEG detection using multivariate swarm sparse decomposition-based joint time-frequency analysis for intuitive BCI, *IEEE Transactions on Human-Machine Systems*, vol. 55, no. 3, pp. 347-357, June 2025.
3. **S. V. Bhalerao** and R. B. Pachori, Automated classification of cognitive visual objects using multivariate swarm sparse decomposition from multichannel EEG-MEG signals, *IEEE Transactions on Human-Machine Systems*, vol. 54, no. 4, pp. 455-464, 2024.
4. **S. V. Bhalerao** and R. B. Pachori, Sparse spectrum based swarm decomposition for robust nonstationary signal analysis with application to sleep apnea detection from EEG, *Biomedical Signal Processing and Control*, vol. 77, 103792, 2022.
5. **S. V. Bhalerao** and R. B. Pachori, Clustering sparse swarm decomposition for automated recognition of upper limb movements from non-homogeneous cross-channel EEG signals, *IEEE Sensors Letters*, vol. 8, issue 1, 7000804, pp. 1-4, 2024.

#### A2. In Book chapter

1. **S. V. Bhalerao** and R. B. Pachori, Automatic detection of motor imagery EEG signals using swarm decomposition for robust BCI systems, *Human-Machine Interface Technology Advancements and Applications*, CRC Press, 2023. eBook ISBN9781003326830.

## **C. Journal publications (other than thesis)**

### **C1. In Refereed Conference**

1. A. Mahato, **S. V. Bhalerao**, R. B. Pachori, V. M. Gadre and D. D. Mahapatra, "Neurological Responses to Meditation with EEG Analysis Using Novel Empirical Fourier-Bessel Decomposition Approach," in Proc. 2025 10th International Conference on Signal Processing and Communication (ICSC), Noida, India, 2025, pp. 694-699

### **C2. In Book chapter**

1. **S. V. Bhalerao**, S. S. Ainwad, and R.B. Pachori, FBSE based automated classification of motor imagery EEG signals in brain-computer interface, *Handbook of Neural Engineering, Volume 2: Brain-Computer Interfaces*, Elsevier, 2024.