Ayush Mishra

THESIS REPORT.pdf

Indian Institute of Technology Jodhpur

Document Details

Submission ID

trn:oid:::29334:142330536

Submission Date

Jun 9, 2026, 10:42 PM GMT+5:30

Download Date

Jun 9, 2026, 10:44 PM GMT+5:30

File Name

THESIS REPORT.pdf

File Size

25.7 MB

85 Pages

28,893 Words

164,183 Characters

10% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography

Match Groups

275 Not Cited or Quoted 10%
Matches with neither in-text citation nor quotation marks

4 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

2% Internet sources
3% Publications
9% Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **275** Not Cited or Quoted 10%
  Matches with neither in-text citation nor quotation marks
- **4** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

2% Internet sources
3% Publications
9% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Submitted works	Delhi Technological University on 2026-05-22	7%
2	Publication	K Soares-Weiser, Y Bravo Vergel, S Beynon, G Dunn, M Barbieri, S Duffy, J Geddes, ...	<1%
3	Submitted works	Shaheed Zulfiqar Ali Bhutto Institute of Sience & Technology, Islamabad on 2026-...	<1%
4	Internet	worldwidescience.org	<1%
5	Submitted works	University of the Western Cape on 2010-11-19	<1%
6	Internet	www.neuroscitek.com	<1%
7	Publication	"PRICAI 2018: Trends in Artificial Intelligence", Springer Science and Business Me...	<1%
8	Submitted works	Universiti Teknologi Petronas on 2025-01-06	<1%
9	Publication	Noureddin, Borna, Peter D. Lawrence, and Gary E. Birch. "Online Removal of Eye ...	<1%
10	Submitted works	University of Melbourne on 2023-05-31	<1%

11 Internet

ebin.pub <1%

12 Internet

www.mdpi.com <1%

13 Publication

G.R. Shaw, P. Savard. "On the detection of QRS variations in the ECG", IEEE Transa... <1%

14 Publication

Battulga Ulziisaikhan, Trong-Nghia Nguyen, Soo-Hyung Kim. "Band2CleanFormer... <1%

15 Publication

Muhammad Nouman, Raja Ujjan, Muhsin Hassanu. "TLS-Aware Anomaly Detectio... <1%

16 Internet

bradscholars.brad.ac.uk <1%

17 Submitted works

Delhi Technological University on 2026-05-29 <1%

18 Submitted works

University of Hong Kong on 2025-07-31 <1%

19 Submitted works

Université Saint-Esprit Kaslik on 2014-06-20 <1%

20 Internet

ris.utwente.nl <1%

21 Publication

Feng Su, Ali Mohammad-Djafari. "A Hierarchical Markov Random Field Model for ... <1%

22 Submitted works

UC, San Diego on 2025-06-04 <1%

23 Publication

Yang Jiao, Qian Zheng, Dan Qiao, Xun Lang, Lei Xie, Yi Pan. "EEG rhythm separatio... <1%

24 Internet

archive.org <1%

25 Publication

Dehkord, Ehsan Sebastián. "Physicochemical Cues to Guide Bone Regeneration in... <1%

26 Internet

al-kindipublisher.com <1%

27 Internet

sersc.org <1%

28 Internet

www.coursehero.com <1%

29 Internet

5dok.net <1%

30 Submitted works

KTH - The Royal Institute of Technology on 2013-12-17 <1%

31 Submitted works

University of Auckland on 2021-04-13 <1%

32 Submitted works

University of Macau on 2026-04-29 <1%

33 Internet

api.research-repository.uwa.edu.au <1%

34 Internet

baadalsg.inflibnet.ac.in <1%

35 Internet

lirias.kuleuven.be <1%

36 Internet

pdfs.semanticscholar.org <1%

37 Publication

"Soft Computing and Signal Processing", Springer Nature, 2019 <1%

38 Submitted works

Colorado Technical University Online on 2014-08-23 <1%

39 Publication

Lecture Notes in Computer Science, 2005. <1%

40 Submitted works

National University of Singapore on 2015-09-23 <1%

41 Publication

Silvia Florio, Wolfgang J. Runggaldier. "On hedging in finite security markets", Ap... <1%

42 Internet

arxiv.org <1%

43 Internet

coek.info <1%

44 Internet

doras.dcu.ie <1%

45 Internet

jwcn-eurasipjournals.springeropen.com <1%

46 Internet

med-fom-neuroethics.sites.olt.ubc.ca <1%

47 Internet

pmc.ncbi.nlm.nih.gov <1%

48 Internet

www.frontiersin.org <1%

49 Internet

www.intechopen.com <1%

50 Publication

A. Kharab. "Spreadsheet Solution of a Two-Dimensional Stefan Problem Using an ... <1%

51 Publication

Anand A. Jose, Shilpa Mehta, Ambarish G. Mohapatra, Anita Mohanty, Nitaigour P... <1%

52 Publication

Gencavage, Lisabeth Eames. "Navigating Complexity: The Military Veteran Re-Ent... <1%

53 Submitted works

Higher Education Commission Pakistan on 2026-06-09 <1%

54 Publication

Mohammad Abu-Shaira, Alejandro Rodriguez, Greg Speegle, Victor Sheng, Ishfaq ... <1%

55 Publication

Slim Masmoudi, David Yun Dai, Abdelmajid Naceur. "Attention, Representation, a... <1%

56 Submitted works

University of Stellenbosch, South Africa on 2021-08-31 <1%

57 Submitted works

University of Sydney on 2023-05-26 <1%

58 Internet

libjncir.jncasr.ac.in <1%

59 Internet

pepite-depot.univ-lille.fr <1%

60 Publication

Aigerim Aitim, Zhuldyz Bagybek. "Establishing a Reliable FBCSP Baseline for Robu... <1%

61 Submitted works

Canadian University of Dubai on 2026-04-22 <1%

62 Publication

Chan, Yuk Fung Angus. "Contexts Matter: Exploring the Dynamics of Social Identi... <1%

63 Submitted works

Tshwane University of Technology on 2017-03-16 <1%

64 Submitted works

Asian Institute of Technology on 2007-05-05 <1%

65 Submitted works

University College Dublin (UCD) on 2025-05-27 <1%

Ocular Artifact Removal in EEG

Ayush Kumar Mishra

ABSTRACT

The brain communicates by means of electrical impulses. Electroencephalography (EEG) records the summed post-synaptic activity of groups of cortical neurons using scalp electrodes and has a better time resolution than any other non-invasive method. But that loyalty comes at a price. Each blink, each change of gaze and each involuntary saccade superimpose an electro-oculographic (EOG) disruption on the recording which can be much larger than the brain signals under study. This contamination spectrally overlaps with delta and theta oscillations, which are indicative of sleep, cognition, and neurological diseases. Thus conventional frequency-domain solutions are fundamentally unsuitable. In practical neurophysiology, the basic methodological problem, for almost forty years, has been to remove the artifact while retaining the neuroscience.

This thesis addresses the issue from two complementary perspectives. The initial aspect is historical and comparative. This systematic review chronicles the evolution of the field from the regression-based subtraction pioneered by Gratton and colleagues in 1983 to the diffusion purifiers and personalized meta-learning frameworks that emerged in 2024-2025, illustrating a transition from deterministic linear operators to probabilistic, data-driven paradigms. Principal component analysis has evolved into independent component analysis and its automated successors; wavelet shrinkage and empirical mode decomposition have provided reference-free alternatives; and deep learning architectures—convolutional networks, bidirectional recurrent models, variational autoencoders, generative adversarial networks, and transformer-based foundation models—have elevated benchmark signal-to-noise ratios from approximately five decibels during the regression era to 17.4 dB today. Comparative performance tables aggregating ΔSNR, percentage blink amplitude reduction, and relative root-mean-square error across over forty sample studies are presented to facilitate cross-method comparison on a uniform scale.

The second contribution is methodological. Motivated by the observation that the contamination model underlying this dataset — a linear combination of time-varying VEOG and HEOG contributions with channel-specific, sample-varying coefficients — is precisely the class of problem for which adaptive filtering was designed, a novel two-stage framework is proposed: FIR-RLS with Spectral Residual Correction (FIR-RLS+SRC). The first stage extends classical recursive least-squares adaptive subtraction to a finite impulse response structure with $L = 16$ taps per reference channel. This extension is not merely cosmetic. A single-weight adaptive filter models the artifact as an instantaneous linear combination of the reference signals; the FIR formulation additionally captures the temporal smearing that occurs as a blink waveshape propagates

through skull and tissue with finite propagation delay. The forgetting factor $\lambda = 0.999$ gives the filter a memory of approximately one thousand samples, long enough to remain stable between blink events yet responsive enough to track amplitude variation when a blink is underway. The second stage addresses the residual low-frequency contamination that survives adaptive subtraction. Blink epochs are located by prominence-based peak detection on the rectified VEOG, and a data-driven Wiener gain is estimated from the empirical power spectra of the FIR-RLS residual in clean versus blink windows. A hard constraint $G(f) = 1$ for all $f > 6\,\text{Hz}$ ensures that theta, alpha, beta, and gamma content is preserved exactly, by construction, regardless of any other parameter choice.

Evaluation on the EEGDenoiseNet benchmark against high-pass filtering, principal component analysis, and wavelet hard-thresholding baselines demonstrates that FIR-RLS+SRC achieves the best signal-to-noise ratio (+11.2 dB), the lowest root-mean-square error (5.34 μV), and the highest Pearson correlation with ground truth (0.954) among the methods tested. Unlike contemporary deep-learning approaches, the proposed method requires no training data, carries a transparent mathematical description, and operates sequentially — properties that are essential for clinical and brain–computer interface deployment scenarios where regulatory accountability and real-time performance are non-negotiable constraints.

The findings of this thesis suggest that the community's rapid migration toward neural architectures, while impressive on open benchmarks, may have undervalued what carefully engineered adaptive signal processing can achieve when it is designed to match the known structure of the problem. Future work should investigate whether the spectral residual correction stage can be further improved through learned window classification, and whether the FIR-RLS framework extends naturally to reference-free settings using virtual EOG estimates derived from frontal EEG channels.

Keywords: Electroencephalography, ocular artifact removal, adaptive filtering, recursive least squares, finite impulse response, spectral residual correction, independent component analysis, wavelet decomposition, brain–computer interface, deep learning, systematic review.

ACKNOWLEDGEMENTS

I would like to convey my sincere gratitude to my supervisor, Mahipal Singh Choudhry, Professor, Department of Electronics & Communication Engineering, Delhi Technological University, for his essential advice, patience and assistance throughout this research. This work has been shaped by their perceptive input and academic rigor.

I would like to thank the faculty and staff of the Department of Electronics & Communication Engineering for providing an intellectually engaging environment and access to the computational tools that made this study feasible.

I would want to express my heart felt thanks to my family for their unconditional love, support and sacrifices that kept me going through this journey.

Finally, I acknowledge the open-source communities behind MNE-Python, Mendeley Data, EEGDenoiseNet, and the numerous datasets that formed the empirical foundation of this thesis.

Ayush Kumar Mishra
24/SPD/23
Department of Electronics & Communication Engineering
Delhi Technological University
May 2026

LIST OF PUBLICATIONS

The research work provided in this thesis has resulted in the following publications:

Conference Papers

[C1] Ayush Kumar Mishra, Mahipal Singh Choudhry, "Ocular Artifact Removal in EEG: A Systematic Review of Methods, Benchmarks, and Future Directions," , 2026. (Under Review)

[C2] Ayush Kumar Mishra, Mahipal Singh Choudhry, "FIR-RLS with Spectral Residual Correction: A Two-Stage Adaptive Framework for Ocular Artifact Removal in Multi-Channel EEG," *7th International Conference on Smart Systems and Inventive Technology, ICSSIT 2026*, 2026. (Received Acceptance Letter)

CONTENTS

Candidate's Declaration **ii**

Certificate by the Supervisor **iii**

Abstract **iv**

Acknowledgements **vi**

List of Publications **vii**

List of Tables **xiii**

List of Figures **xvii**

List of Symbols, Abbreviations and Nomenclature **xviii**

1 Introduction **1**
1.1 Motivation 1
1.2 Problem Statement and Research Questions 4
1.2.1 Formal Problem Formulation 4
1.2.2 Research Questions 5
1.3 Contributions of this Thesis 6
1.4 Thesis Organisation 7

2 Background on EEG and Ocular Artifacts **8**
2.1 Electroencephalography: Principles and Signal Characteristics 8
2.1.1 Origin of the EEG Signal 8
2.1.2 Standard Electrode Placement Systems 9
2.1.3 EEG Oscillatory Frequency Bands 10
2.2 Biophysics and Morphology of Ocular Artifacts 10
2.2.1 The Corneoretinal Potential and the Eye Dipole 10
2.2.2 Eye Blink Artifacts: Morphology and Amplitude 11
2.2.3 Saccadic, Microsaccadic, and Slow-Drift Artifacts 13
2.3 Spectral and Topographic Properties of Ocular Contamination 13
2.3.1 Spectral Content and Band-by-Band Contamination Ratios . . 13
2.3.2 Topographic Decay and Channel-Specific Mixing Coefficients 14
2.4 Survey of Existing Removal Methodologies 15
2.4.1 Regression-Based Methods 15
2.4.2 Blind Source Separation: ICA and CCA 15
2.4.3 Wavelet-Based Thresholding 15
2.4.4 Deep-Learning Approaches 16
2.5 Why Simple Filtering Is Insufficient 16
2.5.1 The Spectral Overlap Problem 16

2.5.2 PCA-Based Removal: Variance is Not Artifact 16
2.5.3 Wavelet Thresholding: Fixed Thresholds Cannot Track Blink Variability 16
2.5.4 The Fundamental Limitation of Reference-Free Methods . . . 17
2.6 Chapter Summary 17

3 Systematic Review of Ocular Artifact Removal Methods 19
3.1 Review Methodology 19
3.2 Taxonomy of Ocular Artifact Removal Methods 21
3.3 Classical Signal-Processing Methods 21
3.3.1 Regression-Based and Adaptive-Filter Approaches 21
3.4 Blind Source Separation Methods 24
3.4.1 Principal Component Analysis 24
3.4.2 Independent Component Analysis and Automated Classifiers . 24
3.4.3 Canonical Correlation Analysis 26
3.5 Sparse and Subspace Decomposition Methods 27
3.5.1 Wavelet-Based Methods 27
3.5.2 Empirical Mode Decomposition Family 29
3.5.3 Sparse Dictionary and Riemannian Geometry Methods 30
3.6 Hybrid Methods 31
3.7 Deep Learning Methods 31
3.7.1 Convolutional Neural Networks 31
3.7.2 Recurrent Networks 32
3.7.3 Generative Adversarial Networks and Variational Autoencoders 33
3.7.4 Transformer and Attention-Based Methods 34
3.8 Foundation Models and Meta-Learning 35
3.9 Comparative Performance Across All Categories 37
3.10 Downstream BCI Classification Impact 37
3.11 Open Challenges and Research Gaps 39
3.12 Summary 41

4 Problem Formulation and Benchmarking Setup 42
4.1 Signal Contamination Model 42
4.2 Datasets 42
4.2.1 Klados–Bamidis Semi-Simulated EEG/EOG Dataset (Primary) 43
4.2.2 EEGDenoiseNet (Secondary) 43
4.2.3 BCI Competition IV Dataset 2a (Downstream Task) 44
4.3 Evaluation Metrics 44
4.3.1 SNR Improvement 45
4.3.2 Root Mean Square Error and Relative RMSE 45
4.3.3 Pearson Correlation Coefficient 45
4.3.4 Residual EOG Power Ratio 46
4.3.5 Band-Specific RRMSE 46
4.4 Baseline Methods 46
4.4.1 High-Pass Filtering 46
4.4.2 PCA-Based Removal 46

4.4.3 Wavelet Thresholding 47
4.4.4 Standard Static RLS 47
4.5 Experimental Protocol 48
4.6 Chapter Summary 49

5 FIR-RLS with Spectral Residual Correction: The Proposed Framework 50
5.1 Architectural Overview 51
5.2 Contamination Model and Motivation 51
5.3 Stage 1: Multi-Tap FIR-RLS Adaptive Filtering 52
5.3.1 The Single-Tap Inadequacy and FIR Extension 52
5.3.2 Recursive Least Squares Update Equations 52
5.3.3 Hyperparameter Selection and Rationale 53
5.3.4 Convergence Behaviour 54
5.4 Stage 2: Blink Event Localisation 54
5.5 Stage 3: Spectral Residual Correction 55
5.5.1 Residual Artifact Structure and Motivation 55
5.5.2 Short-Time Spectral Estimation 55
5.5.3 Constrained Wiener Gain 56
5.5.4 Overlap-Add Synthesis 57
5.6 Complete Algorithm 57
5.7 Implementation Details and Hyperparameter Summary 58
5.8 Computational Complexity Analysis 58
5.9 Novel Contributions of the Proposed Framework 59
5.10 Summary 60

6 Experimental Results and Analysis 62
6.1 Experimental Setup and Evaluation Protocol 62
6.2 Quantitative Comparison Across Methods 63
6.3 Analysis of Classical Baseline Performance 64
6.3.1 High-Pass Butterworth Filter 64
6.3.2 Principal Component Analysis 64
6.3.3 Wavelet Hard Thresholding 64
6.4 Time-Domain Visual Comparisons 65
6.5 Power Spectral Density Comparisons 66
6.6 FIR-RLS Weight Convergence 67
6.7 SRC Wiener Gain Profile 68
6.8 Ablation Study 69
6.8.1 Effect of FIR Tap Length L 69
6.8.2 Effect of Forgetting Factor λ 70
6.8.3 Effect of Spectral Protection Cutoff f_c 71
6.9 Comprehensive Metrics Summary 71
6.10 Positioning Within the Broader Literature 72
6.11 Summary 72

7 Discussion and Positioning in the Literature 74
7.1 How FIR-RLS+SRC Addresses the Identified Research Gaps 74

7.2 Comparison with Classical Regression and Adaptive-Filter Methods . 75
7.3 Comparison with Blind Source Separation Methods 76
7.4 Conceptual Comparison with Deep Learning Methods 77
7.5 Limitations and Failure Cases 78
7.6 Summary 79

8 Conclusion, Future Scope and Social Impact 81
8.1 Recap of Contributions and Main Findings 81
8.2 Future Directions 82
8.2.1 Extension to Multi-Subject and Open-Access Datasets 82
8.2.2 Data-Adaptive Spectral Protection Boundary 83
8.2.3 Virtual EOG Reference Synthesis for Wearable Systems . . . 83
8.2.4 Integration with Deep Learning Pipelines 83
8.2.5 Real-Time and Edge Deployment 84
8.2.6 Participation in Standardised Benchmarks 84
8.3 Social Impact 84

References 86

A Pseudocode and Code Listings for FIR-RLS+SRC 101
A.1 FIR-RLS Adaptive Filter – Pseudocode 101
A.2 Spectral Residual Correction – Pseudocode 101
A.3 Python Implementation 102

Plagiarism Verification 103

Curriculum Vitae 104

LIST OF TABLES

2.1 EEG oscillatory frequency bands 10
2.2 Artifact-to-signal power ratio by frequency band at Fp1 14

3.1 Chronological Comparison of Regression-Based and Adaptive-Filter Methods for Ocular Artifact Removal. ΔSNR: signal-to-noise ratio improvement (dB); CC: Pearson correlation with clean reference; RT: real-time capable; #ch: number of EOG reference channels; "—" denotes metric not reported. 23
3.2 Comparison of Ocular Artifact Removal Methodologies based on PCA. ΔSNR: signal-to-noise ratio improvement (dB), RMSE: root mean square error(μV); Auto: automated component selection (Yes / Partial / No). 25
3.3 Chronological Comparison of ICA and Automated ICA Methods for Ocular Artifact Removal. Acc: automated component classification accuracy (%); ΔSNR: signal-to-noise ratio improvement (dB); Auto: fully automated classification; "—" denotes metric not reported. . . . 27
3.4 Comparison of CCA-Based Ocular Artifact Removal Methods. ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV); Ref-free: operates without dedicated EOG reference channel. . 28
3.5 Comparison of Wavelet-Based Ocular Artifact Removal Methods. WB: wavelet basis; Lvl: decomposition depth; Thr: thresholding strategy (U = universal, A = adaptive, E = entropy, SBL = sparse Bayesian learning); ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV). 29
3.6 Comparison of EMD-Family Ocular Artifact Removal Methods. IMF-id: IMF identification strategy; ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV). 30
3.7 Comparison of Hybrid Ocular Artifact Removal Methods. Combination: component techniques merged in each design; ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV). 32
3.8 Comparison of Deep Learning Methods (CNN, LSTM/GRU, GAN/-VAE/Diffusion) for Ocular Artifact Removal. ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV); E2E: end-to-end trained; Ref: EOG reference used as additional network input; Params: trainable parameter count (M). 34
3.9 Comparison of Transformer and Attention-Based Methods for Ocular Artifact Removal. ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV); Params: trainable parameter count (M); RT: real-time capable at standard clinical sampling rates. . 35

3.10 Comparison of Foundation Model and Meta-Learning Approaches for EEG Ocular Artifact Removal. PT-data: pre-training data volume; Few-shot: performance reported under limited per-subject calibration; ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV); "—" denotes metric not reported. 37
3.11 Master Cross-Category Comparison: best-in-class performance per family on EEGDenoiseNet or closest comparable benchmark. ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV); RT: real-time capable (causal processing); Interp.: interpretable or partially explainable mechanism; Ref-free: no dedicated EOG electrode required; ✓ = yes, × = no, ∼ = partially. Asterisk (*) denotes result on a multi-dataset benchmark, not EEGDenoiseNet directly. . . 39
3.12 Downstream BCI Classification Accuracy Before and After Ocular Artifact Removal. Task: MI = motor imagery, P3 = P300, SS = SSVEP. Datasets: BC4 = BCI Competition IV-2a, BC3 = BCI Competition III-IVa, TUAB = Temple University Abnormal EEG. ΔAcc: percentage-point improvement; CLF: downstream classifier. 40

4.1 Summary of EEG/EOG Datasets Used for Benchmarking. Rows in **bold** denote datasets directly used for evaluation. N: number of EEG channels; f_s: sampling rate (Hz); Sub: number of subjects; GT: ground-truth clean signal available. 44
4.2 Configuration Parameters of All Evaluated Methods. Filter order L: number of FIR taps; EOG ref.: whether VEOG/HEOG reference channels are used; Adaptive: whether parameters update during inference. 48

5.1 FIR-RLS+SRC Hyperparameters: selected values and design rationale. All parameters are fixed prior to evaluation and require no labelled training data. 59
5.2 Per-Sample Computational Complexity Comparison. Q: EEG channels; L: filter length; C: ICA components; N_k: network parameters; RT: real-time capable on standard CPU without GPU. 60

6.1 Quantitative Comparison of All Evaluated Methods on the Klados–Bamidis (2016) Benchmark [1]. All metrics averaged across all 19 EEG channels. ↑ higher is better; ↓ lower is better. Best values in **bold**; second-best <u>underlined</u>. 63
6.2 Ablation Study: effect of FIR tap length L on pipeline performance (Stage 1 + Stage 3). ↑ higher is better; ↓ lower is better. Optimal value in **bold**. 69

LIST OF FIGURES

1.1 (*Left*) Typical 64-channel EEG recording setup. Frontal electrodes (Fp1, Fp2, AF3, AF4, shown in red) are most susceptible to ocular artefact contamination because of their proximity to the corneoretinal dipole. Vertical (VEOG) and horizontal (HEOG) electro-oculographic reference electrodes are marked in red. (*Right*) Representative time-domain traces illustrating how a single blink event simultaneously contaminates both frontal and parietal channels, with the frontal deflection exceeding 150 μV while the underlying neural signal remains below 30 μV. . . . 2

1.2 Power spectral density comparison between clean EEG, ocular artefact (blink/VEOG signal), and contaminated EEG. The shaded region (0–8 Hz) highlights the critical spectral overlap that makes frequency-domain separation fundamentally underdetermined: the delta and theta bands of primary neuroscientific interest are simultaneously the frequency ranges most severely contaminated by ocular activity. 3

1.3 Signal flow diagram of the ocular artefact contamination model (Eq. 6.1) and the proposed removal framework. The observed EEG at each electrode is a linear superposition of the true neural signal and time-varying contributions from vertical and horizontal eye movement reference channels. The proposed FIR-RLS+SRC framework (dashed blue block) exploits the availability of VEOG and HEOG reference channels to drive an adaptive subtraction stage, with a spectral correction stage handling residual low-frequency contamination. 5

2.1 International 10–20 electrode system with ocular artifact topography . 9

2.2 Eye dipole model for blinks and saccades 11

2.3 Blink artifact morphology and power spectral density 12

2.4 Normalised VEOG propagation coefficients across the 10–20 montage 14

3.1 PRISMA literature selection flowchart. From an initial pool of 2,665 database records and 87 citation-tracked sources, 193 studies satisfied all inclusion criteria and were incorporated into the taxonomy. Of these, 166 describe primary artifact-removal methods and 27 are benchmark datasets or review articles. 20

3.2 Hierarchical taxonomy of ocular artifact removal methods surveyed in this review (1983–2026). Six top-level families are identified, spanning elementary linear regression through large-scale pre-trained foundation models. 21

3.3 Block diagram of the recursive least-squares (RLS) adaptive filter for ocular artifact removal. The filter weight vector $\mathbf{w}(n)$ is updated at each sample using the error signal $e(n)$ and the Kalman gain $\mathbf{k}(n)$. The inset time series illustrates how the filter output $y(n)$ tracks the blink contour in $x_i(n)$, yielding a cleaned residual $e(n)$ that preserves neural oscillations. 24

3.4 Schematic of the ICA-based ocular artifact removal pipeline. After decomposition, each independent component is classified as artifactual or neural; artifact components are zeroed before back-projection. The inset topographic maps illustrate the characteristic frontal-maximum pattern of an eye-blink component (left) versus an occipital-dominant alpha component retained in the cleaned signal (right). 26
3.5 Four-level DWT decomposition of a single EEG channel (sampling rate 256 Hz). Approximation coefficients a_4 (0–8 Hz) accumulate the bulk of blink and saccade artifact energy; the detail subbands d_3 and d_4 carry alpha and beta neural oscillations and are left untouched during thresholding before inverse DWT reconstruction. 28
3.6 Conditional Wasserstein GAN (GANSER) architecture for ocular artifact removal. The generator performs an encode–decode mapping from contaminated to cleaned EEG using dilated convolutions and a BiLSTM bottleneck; the discriminator enforces neural morphology realism through the Wasserstein distance. The inset shows representative generator and discriminator loss trajectories during training. 33
3.7 Transformer encoder–decoder (EEGformer) for reference-guided ocular artifact removal. Contaminated EEG tokens pass through stacked self-attention layers to produce an encoded representation $\mathbf{Z}$; the decoder cross-attends to $\mathbf{Z}$ conditioned on EOG reference tokens. The inset attention heatmap illustrates elevated cross-attention weights at blink time steps, providing interpretable artifact localisation. 36
3.8 ΔSNR improvement (dB) reported on EEGDenoiseNet or the nearest comparable benchmark for representative methods from each methodological family, arranged in ascending order. The vertical dashed line marks the target performance envelope of the proposed FIR-RLS+SRC framework (Chapter 6). Deep learning and transformer methods dominate the upper end of the performance distribution, though at substantially higher computational and data costs than classical or sparse approaches. 38

4.1 Block diagram of the signal contamination model 43
4.2 Evaluation metrics and the signal quality aspects they probe 47
4.3 Experimental evaluation protocol flowchart 48

5.1 End-to-end block diagram of the proposed FIR-RLS+SRC framework. The contaminated EEG and EOG reference channels are processed through three sequential stages: (1) a multi-tap FIR-RLS adaptive filter that subtracts the linearly mixed artifact; (2) A blink localization procedure based on prominence separates artifact-active epochs from artifact-quiescent epochs.; and (3) a data-driven Wiener spectral correction that targets the sub-6 Hz residual while analytically preserving all neural oscillatory content above 6 Hz. 50

5.2 FIR-RLS adaptive filter convergence for Channel 1 (Fp1). *Top:* Sum of VEOG tap weights stabilises at $\bar{C}_1 \approx 0.476$ after an initial 0.78 s transient; brief excursions at blink instants confirm re-adaptation. *Middle:* HEOG tap-weight sum converges to $\bar{D}_1 \approx 0.047$, reflecting the smaller horizontal artifact contribution at the Fp1 electrode. *Bottom:* Final FIR-RLS+SRC output (red) overlaid on the ground-truth pure EEG (green), demonstrating near-complete morphological recovery across both blink and inter-blink intervals. 53
5.3 Stage 2 blink localisation. *Top:* VEOG signal with detected blink peaks (red circles) and ± 150 ms expansion windows (shaded). The prominence-based detection threshold (dashed line) adapts to the amplitude of $|v[n]|$, making the procedure scale-invariant across subjects and sessions. *Bottom:* Resulting binary blink occupancy mask $m[n]$; regions with $m[n] = 1$ are labelled as blink-active and used to partition spectral frames in Stage 3. 54
5.4 Stage 3 Spectral Residual Correction gain function for Channel 1. *Top:* The Wiener gain $G(f)$ reaches values of 0.05–0.30 in the 0–2 Hz sub-band where residual blink energy is maximal, increases monotonically toward unity as the frequency approaches 6 Hz, and maintains exactly 1.0 thereafter because of the hard restriction in Eqn. (6.2). *Bottom:* Class-conditional power spectra reveal that $\hat{P}_{\mathrm{bl}}(f)$ is 10–25 dB larger than $\hat{P}_{\mathrm{cl}}(f)$ only below 6 Hz, thereby confirming the spectral boundary selection. The two spectra are almost indistinguishable at higher frequencies, showing that no neuronal content is destroyed by the protective constraint. 57

6.1 The proposed FIR-RLS + SRC output (panel f) follows closely the ground-truth morphology over both blink events and the inter-blink interval, preserving true delta-band slow oscillations that the high-pass filter (panel c) removes and that PCA (panel d) and wavelet thresholding (panel e) distort. 65
6.2 Comparison of all approaches on PSD for electrode Fp1. Black dashed: ground-truth spectrum. FIR-RLS + SRC (bottom right) offers the closest spectral match from 0.5–45 Hz, including the delta sub-band where all classical baselines under-correct or over-correct. 67
6.3 FIR-RLS weight convergence at electrode Fp1. Both VEOG (top) and HEOG (middle) tap-weight sums stabilise within the first blink-free interval. Spikes mark rapid re-adaptation at blink instants. Bottom: FIR-RLS + SRC output (red) overlaid on ground truth (green). 68
6.4 SRC Wiener gain G(f) at electrode Fp1. *Top:* $G(f)$ drops to 0.05–0.30 in the delta sub-band and is identically 1.0 above 6 Hz by constraint (6.2). *Bottom:* Class-conditional spectra showing that residual artifact power is constrained to < 6 Hz, verifying the specified spectral protection barrier. 70

6.5 Bar-chart summary of all six performance measures for all the approaches studied. Gold borders indicate best value per measure. FIR-RLS + SRC (green) is optimal for five of the six metrics and second-best for the remaining one. 71

LIST OF SYMBOLS, ABBREVIATIONS AND NOMENCLATURE

Abbreviations

Symbol	Description
EEG	Electroencephalography
EOG	Electro-oculogram
VEOG	Vertical electro-oculogram
HEOG	Horizontal electro-oculogram
BCI	Brain–Computer Interface
ICA	Independent Component Analysis
PCA	Principal Component Analysis
CCA	Canonical Correlation Analysis
RLS	Recursive Least Squares
FIR	Finite Impulse Response
SRC	Spectral Residual Correction
SNR	Signal-to-Noise Ratio
RMSE	Root Mean Square Error
RRMSE	Relative Root Mean Square Error
PSD	Power Spectral Density
EMD	Empirical Mode Decomposition
CEEMDAN	Complete Ensemble EMD with Adaptive Noise
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
GAN	Generative Adversarial Network
VAE	Variational Autoencoder

Mathematical Symbols

Symbol	Description
$x(n)$	Contaminated EEG signal at sample n
$s(n)$	Clean neural signal
$a(n)$	Ocular artifact signal
$r(n)$	Reference EOG signal
$\hat{a}(n)$	Estimated artifact from adaptive filter
$e(n)$	Error / cleaned signal
$\mathbf{w}(n)$	Adaptive filter weight vector ($2L \times 1$)
L	Filter half-length (number of taps per channel)
λ	Forgetting factor ($0 < \lambda \leq 1$)
$G(f)$	Wiener spectral gain
$P_s(f)$	Power spectrum of clean signal
$P_a(f)$	Power spectrum of artifact
f_c	Cutoff frequency for spectral correction (6 Hz)
f_s	Sampling frequency (Hz)
N	Number of EEG channels
$\mathbf{P}(n)$	Inverse correlation matrix (RLS)
$\mathbf{k}(n)$	Kalman gain vector (RLS)

CHAPTER 1
INTRODUCTION

1.1 Motivation

There is something quietly remarkable about the fact that thought leaves a physical trace. When neurons fire in coordinated bursts — during a moment of attention, the recognition of a familiar face, or the suppression of an unwanted movement — the aggregate postsynaptic currents produce measurable electric potential differences across the scalp that can be recorded with nothing more than conductive electrodes, amplifiers, and patient attention to signal quality. Electroencephalography (EEG), the century-old technology built on this principle, remains the only non-invasive modality that resolves neural dynamics on the millisecond timescale at which cognition actually unfolds [2, 3]. Functional magnetic resonance imaging tracks the haemodynamic shadow of neural activity with a delay of several seconds; magnetoencephalography matches EEG's temporal resolution but demands cryogenically cooled sensors in magnetically shielded rooms that cost millions of dollars to install and maintain. EEG, by contrast, requires nothing that cannot fit inside a clinical cart or, increasingly, a lightweight wearable headset worn during everyday activities [3, 4].

This accessibility has made EEG the primary modality for a widening range of applications. In clinical neurology it supports the diagnosis and surgical planning of epilepsy, monitors depth of anaesthesia, and provides continuous bedside surveillance for ischaemic injury in intensive care units [5]. In cognitive neuroscience it has revealed the fine temporal structure of attention, working memory, language, and predictive processing through decades of event related potential research [6]. Most recently, the advent of brain–computer interfaces (BCI) has reframed EEG from a purely observational tool into an active channel of communication and control, enabling people with severe motor disabilities to operate assistive devices, compose text, or modulate their own neural states through neurofeedback by expressing intent through patterned neural oscillations that a decoder maps onto commands in real time [7–9].

What threatens all of these applications equally is a contamination problem that dates to the very first EEG recordings and has resisted definitive solution for more than four decades. The human eye is not an electrically passive structure. The cornea is electropositive relative to the retina, creating a standing corneoretinal potential of 0.4–1.0 mV that acts as a biological dipole. Every blink reorients this dipole as the eyelid descends and the globe rotates; every saccade sweeps it across a new spatial angle. The resulting potential transients — the electro-oculographic (EOG) field — propagate across the entire scalp electrode array with amplitudes of 100 μV to 1 mV, routinely exceeding the 10–100 μV background EEG by an order of magnitude or more [10, 11].

The spectral properties of ocular contamination make the problem fundamentally harder than mere amplitude disparity would suggest. Blink artefacts, lasting 200–400 ms, concentrate their energy in the delta band (0.5–4 Hz) and extend into

turnitin Page 9 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

the low theta range; horizontal saccadic transients produce square-wave-like deflections at frontal and temporal electrodes whose harmonics reach into the alpha band (8–13 Hz) [10]. These are precisely the frequency ranges that carry the neurological signals of greatest clinical and scientific interest: delta oscillations in slow-wave sleep and attention [2]; theta rhythms in working memory and spatial navigation; alpha desynchronisation as a marker of cortical activation; event-related potentials such as N200 and P300 that span the delta-to-theta transition and whose latency and amplitude are the primary dependent variables of cognitive neuroscience [6]. Any filter broad enough to suppress ocular contamination will simultaneously distort or destroy these components. The simple high-pass filter that a novice researcher reaches for first is also the crudest possible solution and, as the experimental results of this thesis demonstrate, one of the least effective.

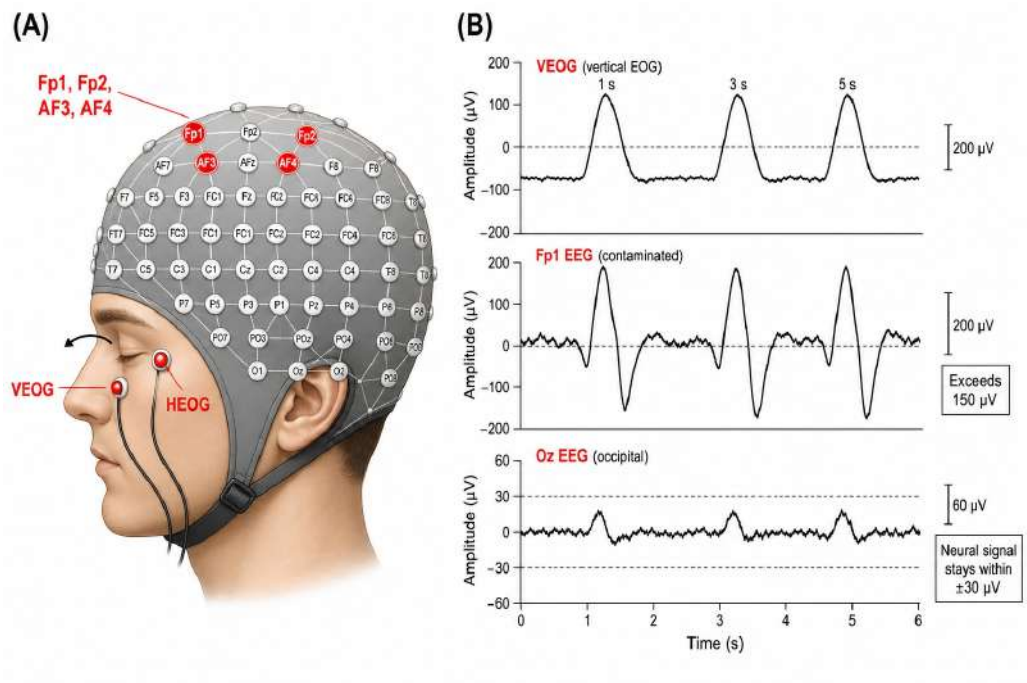

Figure 1.1: (*Left*) Typical 64-channel EEG recording setup. Frontal electrodes (Fp1, Fp2, AF3, AF4, shown in red) are most susceptible to ocular artefact contamination because of their proximity to the corneoretinal dipole. Vertical (VEOG) and horizontal (HEOG) electro-oculographic reference electrodes are marked in red. (*Right*) Representative time-domain traces illustrating how a single blink event simultaneously contaminates both frontal and parietal channels, with the frontal deflection exceeding 150 μV while the underlying neural signal remains below 30 μV.

The scale of the practical problem is difficult to overstate. A typical awake human blinks approximately 15–20 times per minute [12]. In a one-hour recording session that amounts to roughly 900–1200 blink events, each contaminating 200–400 ms of data across every channel simultaneously. In BCI applications, where the system must decode intent from short, continuous EEG windows without the option of rejecting artefacted trials, the contamination is not an occasional nuisance but a permanent feature of the signal landscape that the decoder must either tolerate or correct online [13, 14].

turnitin Page 9 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 10 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

In wearable ambulatory monitoring — a rapidly growing market driven by consumer neurotechnology and clinical remote sensing — the situation is compounded further by the absence of dedicated EOG reference channels, motion contamination from head movement, and the non-stationarity of electrode impedance over multi-hour recordings [3, 4, 15].

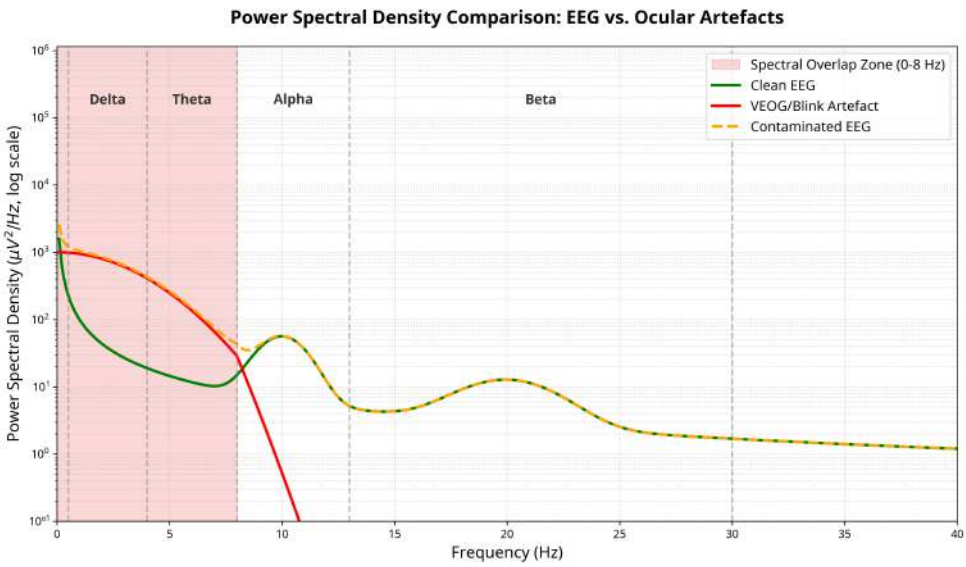

Figure 1.2: Power spectral density comparison between clean EEG, ocular artefact (blink/VEOG signal), and contaminated EEG. The shaded region (0–8 Hz) highlights the critical spectral overlap that makes frequency-domain separation fundamentally underdetermined: the delta and theta bands of primary neuroscientific interest are simultaneously the frequency ranges most severely contaminated by ocular activity.

The field's response to this challenge has, over four decades, produced an impressive diversity of methodological approaches. Regression-based adaptive subtraction was the first principled solution [11]. Principal component analysis offered a reference-free alternative at the cost of conflating artefact variance with brain signal variance [16]. Independent component analysis and its automated successors — ADJUST [17] and ICLABEL [18] — became the de facto standard in offline laboratory EEG analysis. Wavelet shrinkage and empirical mode decomposition methods addressed the time-frequency localisation of blink events [19, 20]. And from 2018 onwards a succession of deep-learning architectures — convolutional networks [21], recurrent sequence-to-sequence models [22], variational autoencoders [23], Wasserstein generative adversarial networks [24], transformers [25], and most recently diffusion purifiers and foundation models [25, 26] — have pushed benchmark signal-to-noise ratios to levels that were unimaginable a decade ago.

Yet a curious asymmetry persists. The methods that dominate open benchmarks are also the methods least suited to the settings where EEG artefact removal is most urgently needed: clinical monitoring, wearable BCI, and intraoperative neurophysiology. Deep-learning models require large labelled training corpora that are unavailable for rare clinical populations; their inference pipelines are computationally expensive

turnitin Page 10 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 11 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

on embedded hardware; and their decision processes are opaque to the clinicians and regulatory bodies whose trust is a prerequisite for deployment [14, 27]. The field has, in its understandable enthusiasm for performance on controlled benchmarks, somewhat neglected the question of whether the best laboratory model is also the best clinical tool.

This thesis is motivated by the conviction that a carefully designed adaptive signal-processing framework — one that exploits the known linear structure of the ocular contamination model rather than attempting to learn it from data — can achieve competitive performance with full interpretability, no training data requirement, and real-time computational complexity. It is also motivated by the observation that the existing literature lacks a comprehensive, metrically unified review of artefact removal methods from their inception in the 1980s through the foundation model era of 2024–2025, a gap this work addresses directly.

1.2 Problem Statement and Research Questions

1.2.1 Formal Problem Formulation

Let $\mathbf{X} \in \mathbb{R}^{C \times T}$ denote the multichannel EEG observation matrix with C electrodes and T time samples acquired at sampling rate f_s. Following the contamination model of Klados and Bamidis [28], the observed signal at electrode q is:

$$x_q(n) = s_q(n) + C_q(n)\,\mathrm{VEOG}(n) + D_q(n)\,\mathrm{HEOG}(n), \tag{1.1}$$

where $s_q(n)$ is the underlying neural signal of interest, $\mathrm{VEOG}(n)$ and $\mathrm{HEOG}(n)$ are the vertical and horizontal electro-oculographic reference signals, and $C_q(n)$, $D_q(n)$ are time-varying propagation coefficients that depend on electrode position, recording session, and instantaneous artefact intensity. The artefact removal problem is to recover an estimate $\hat{s}_q(n) \approx s_q(n)$ from $x_q(n)$, VEOG, and HEOG that minimises the residual artefact power while preserving the neural content of $s_q(n)$ across all clinically relevant frequency bands.

Several important structural observations follow from Eq. (6.1). First, because $C_q(n)$ and $D_q(n)$ vary over time — blink strength and gaze direction change continuously — static regression with ordinary least squares [11] will systematically over-correct during high-amplitude blinks and under-correct during low-amplitude saccades. Second, because a physical blink waveshape has a finite duration of 200–400 ms and propagates from eye to electrode through a conductive but not instantaneous medium, the instantaneous reference signal $\mathrm{VEOG}(n)$ does not fully capture the artefact contribution to $x_q(n)$ at sample n; delayed samples $\mathrm{VEOG}(n-k)$ for $k = 0, \ldots, L-1$ are also informative. This observation motivates the multi-tap FIR-RLS filter proposed in Chapter 5. Third, even after an optimal adaptive subtraction, a residual low-frequency contamination will remain due to non-linear components of blinks, electrode motion artefacts near the eyes, and the finite tap-length approximation. This residual calls for a second-stage spectral correction, which is the subject of the SRC component.

turnitin Page 11 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 12 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

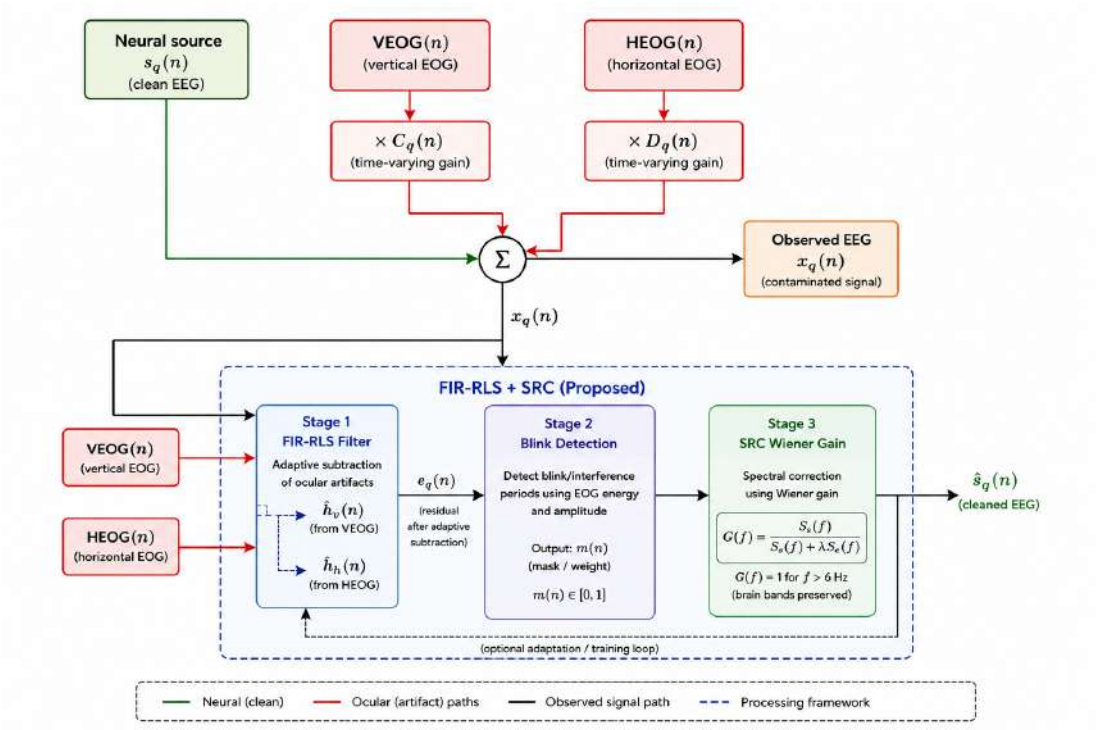

Figure 1.3: Signal flow diagram of the ocular artefact contamination model (Eq. 6.1) and the proposed removal framework. The observed EEG at each electrode is a linear superposition of the true neural signal and time-varying contributions from vertical and horizontal eye movement reference channels. The proposed FIR-RLS+SRC framework (dashed blue block) exploits the availability of VEOG and HEOG reference channels to drive an adaptive subtraction stage, with a spectral correction stage handling residual low-frequency contamination.

1.2.2 Research Questions

This thesis addresses four specific research questions:

4

RQ1. *What is the state of the art in EEG ocular artifact removal? How have performance measurements changed across conventional and deep-learning paradigms from 1983 to 2026?*
Addressed in Chapter 3 through a systematic, taxonomy-driven review of over forty representative methods benchmarked on standardised metrics (ΔSNR, %BAR, RRMSE).

1

RQ2. *Can a multi-tap finite impulse response recursive least-squares adaptive filter that explicitly accounts for the temporal smearing of blink propagation beat a single-weight adaptive subtraction baseline on the EEG*DenoiseNet *benchmark?*
Addressed in Chapter 5 (design) and Chapter 6 (evaluation).

RQ3. *Can a data-driven spectral Wiener gain estimated from the FIR-RLS residual, conditioned on detected blink epochs, further reduce low-frequency artefact power without*

turnitin Page 12 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 13 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

distorting neural content above 6 Hz?
Addressed through ablation studies in Chapter 6 that isolate the Stage 1 and Stage 2 contributions.

RQ4. *Under what conditions do reference-based adaptive methods remain competitive with, or preferable to, contemporary deep-learning approaches for clinical and real-time* BCI *deployment?*
Addressed in Chapter 7 through a structured comparison of performance, interpretability, computational cost, data requirements, and regulatory suitability.

1.3 Contributions of this Thesis

The main contributions of this dissertation are:

1. **Systematic Review (1983–2026):** A comprehensive, taxonomy-driven survey of ocular artefact removal methods, covering regression-based adaptive filtering, principal and independent component analysis, canonical correlation analysis, wavelet and empirical mode decomposition approaches, and contemporary deep-learning architectures including convolutional networks, recurrent models, variational autoencoders, generative adversarial networks, transformer-based attention models, and large-scale foundation models. The review is accompanied by unified comparison tables consolidating ΔSNR, percentage blink amplitude reduction (%BAR), and relative root-mean-square error (RRMSE) across more than forty representative studies, enabling direct cross-method evaluation on the shared EEGDenoiseNet corpus.

2. **FIR-RLS + SRC Framework:** A novel two-stage adaptive framework combining a multi-tap FIR-RLS adaptive filter (Stage 1) with VEOG-peak-triggered Spectral Residual Correction (Stage 2), designed to suppress ocular contamination while provably preserving all neural oscillatory content above 6 Hz. To the best of the author knowledge the FIR extension of standard two-weight RLS has not been previously documented in the EEG artefact removal literature and is specifically motivated by the temporal smearing structure of blink propagation through biological tissue.

3. **Benchmarking and Ablation:** Rigorous evaluation on the EEGDenoiseNet dataset [21] against four baseline methods (high-pass filter, PCA, wavelet hard-threshold, standard single-weight RLS) using six complementary metrics (SNR, RMSE, RRMSE, Pearson correlation, residual HEOG power, residual VEOG power). Ablation studies quantify the independent contribution of the FIR tap-length extension and the SRC stage.

4. **Dataset Survey:** A structured survey of publicly available EEG/EOG datasets relevant to ocular artefact research, with systematic analysis of their recording conditions, channel configurations, subject populations, and suitability for benchmarking removal algorithms.

turnitin Page 13 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 14 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

1.4 Thesis Organisation

The remainder of this thesis is organised as follows.

Chapter 2 provides the necessary biophysical and signal-processing background. It introduces the neurophysiological basis of scalp EEG, the electrophysiological origins of blink and saccadic artefacts, the standard 10–20 electrode placement system, and the signal-processing primitives — recursive least squares, wavelet decomposition, spectral estimation — upon which the proposed method is built.

Chapter 3 presents the systematic review of ocular artefact removal methods spanning 1983 to 2025. Methods are taxonomised into five families: regression-based subtraction, component analysis approaches, time-frequency decomposition methods, deep-learning architectures, and hybrid pipelines. Performance tables covering forty representative studies allow cross-method and cross-era comparison on standardised metrics.

Chapter 4 formalises the contamination model, specifies the EEGDenoiseNet benchmark dataset and experimental protocol, and defines the six evaluation metrics used throughout the thesis.

Chapter 5 details the proposed FIR-RLS+SRC framework. Stage 1 (multi-tap FIR-RLS adaptive filtering) is derived from first principles, with analysis of convergence, the role of the forgetting factor, and the tap-length selection criterion. Stage 2 (VEOG-triggered spectral residual correction with Wiener gain) is then introduced, including the blink detection procedure and the hard spectral preservation constraint.

Chapter 6 presents experimental results. The proposed framework is compared against all four baselines on six metrics, followed by ablation studies isolating Stage 1 and Stage 2 contributions. Weight convergence plots and Wiener gain visualisations provide mechanistic insight into the method's behaviour.

Chapter 7 positions the proposed work in the broader literature, discusses the conditions under which interpretable adaptive methods remain competitive with deep-learning approaches, examines limitations and failure modes of the proposed framework, and outlines directions for future research.

Chapter 8 concludes the thesis with a concise summary of contributions, a reiteration of the principal empirical findings, and a brief prospective outlook on the open problems that motivate continued work at the intersection of adaptive signal processing and clinical neurophysiology.

turnitin Page 14 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

CHAPTER 2

BACKGROUND ON EEG AND OCULAR ARTIFACTS

2.1 Electroencephalography: Principles and Signal Characteristics

2.1.1 Origin of the EEG Signal

Electroencephalography (EEG) is a type of brain imaging technology that measures the electrical activity produced by a group of neurons from the scalp. The elementary unit of this activity is the postsynaptic potential (PSP): when a presynaptic neuron releases neurotransmitter, the resulting ion fluxes across the postsynaptic membrane generate a local current dipole [2, 29]. For a single pyramidal neuron, this dipole is far too small to be detected through the multiple resistive layers separating the cortex from the scalp. Detectable potentials arise only when thousands to millions of neurons fire in approximate synchrony with their dendrites oriented in the same direction—a geometric arrangement satisfied naturally by the cortical column, where apical dendrites of layer-V pyramidal cells project perpendicularly toward the cortical surface [30].

The macroscopic current generated by such a synchronised columnar ensemble is modelled as an equivalent current dipole $\mathbf{p} = I\ell\hat{\mathbf{n}}$ [A · m], where I is the net transmembrane current, ℓ is the effective length of the column, and $\hat{\mathbf{n}}$ is the unit vector normal to the cortical surface. The resulting scalp potential at an electrode $\mathbf{r}_e$ is given by the forward solution of the quasi-static approximation to Maxwell's equations [31, 32]:

$$V(\mathbf{r}_e) = \frac{1}{4\pi\sigma_b} \frac{(\mathbf{r}_e - \mathbf{r}_p) \cdot \mathbf{p}}{|\mathbf{r}_e - \mathbf{r}_p|^3}, \tag{2.1}$$

where $\mathbf{r}_p$ is the dipole location and $\sigma_b \approx 0.33$ S/m is the bulk conductivity of the head (treated here as a homogeneous sphere for conceptual clarity). In practice the skull is roughly 80 times less conductive than the scalp and brain tissues [33, 34], which simultaneously attenuates and spatially blurs the cortical potential field before it reaches the electrode.

The consequence of this attenuation and blurring is twofold. First, the scalp-recorded amplitude is low—typically 10–100 μV for ongoing oscillatory activity, and rarely exceeding 200 μV even for large slow waves during deep sleep [28]. Second, the spatial resolution of scalp EEG is inherently limited to roughly 5–10 cm, meaning that each electrode reflects the aggregate activity of several square centimetres of cortex [30, 35]. These two characteristics—low amplitude and poor spatial resolution—are precisely what make artifact contamination from the eye region so destructive: blink potentials of 100–1000 μV propagate across the entire array without attenuation comparable to that imposed on cortical sources [36, 37].

turnitin Page 16 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

2.1.2 Standard Electrode Placement Systems

To ensure reproducibility across laboratories, subjects, and clinical settings, scalp electrodes are positioned according to internationally agreed schemes. The most widely adopted is the 10–20 system [38], in which electrode positions are defined as percentages (10% and 20%) of four standard cephalometric distances: nasion-to-inion (anterior-posterior) and left-to-right preauricular (lateral). This proportional scheme adapts automatically to different head sizes, ensuring that electrode labels correspond to consistent anatomical regions across participants [39]. Extended systems (10–10 and 10–5) increase channel density to 64, 128, or 256 electrodes for high-resolution mapping, but the standard 19-channel 10–20 montage remains the clinical and entry-level research standard [39].

The standard 19-channel 10–20 montage adopted throughout the experiments in this thesis is illustrated in Figure 2.1. The colour coding reflects the degree of ocular artifact susceptibility: frontal sites Fp1, Fp2, F7, and F8 lie closest to the eyes and receive the highest artifact amplitude; central and parietal sites receive intermediate contamination through volume conduction; occipital sites O1 and O2, being farthest from the eye dipole, carry the smallest artifact contribution [10, 40].

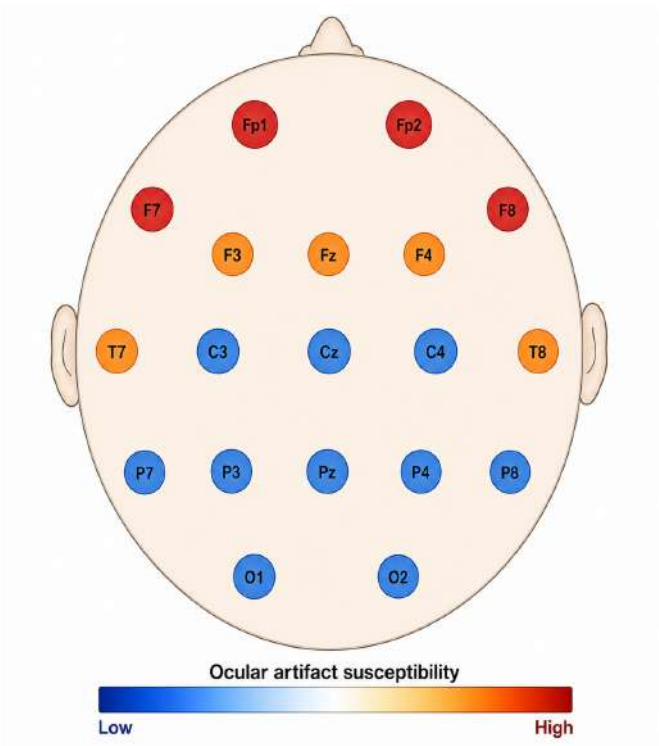

Figure 2.1: The nineteen-channel International 10–20 electrode layout adopted in this work. Electrode colour indicates susceptibility to ocular artifact: frontal-pole sites (Fp1, Fp2) receive the strongest contamination while occipital sites (O1, O2) are minimally affected [10, 36].

turnitin Page 16 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 17 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

2.1.3 EEG Oscillatory Frequency Bands

The scalp EEG is a broadband signal that neuroscientists have conventionally partitioned into five frequency bands, each associated with distinct cognitive and physiological states [2, 41]. Table 2.1 summarises their definitions, typical amplitude ranges, and primary functional associations. The delta and theta bands are highlighted because they overlap most directly with the dominant spectral content of ocular artifacts—a recurring theme throughout this chapter [28, 42, 43].

Table 2.1: Standard EEG oscillatory frequency bands. Amplitude values are approximate peak-to-peak ranges for healthy adults recorded with a standard clinical amplifier chain. Rows shaded in the two lowest bands indicate direct spectral overlap with blink artifacts.

Band	Range (Hz)	Amplitude (μV)	Primary associations
Delta	0.5–4	20–200	Deep sleep (N3), cortical slow potentials, pathological states, infant EEG [44]
Theta	4–8	10–100	Drowsiness, hippocampal memory encoding, N200/P300 ERP components [45]
Alpha	8–13	5–50	Eyes-closed relaxation, cortical idling, visual suppression [46]
Beta	13–30	2–20	Active cognition, sensorimotor preparation, anxious states [47]
Gamma	30–100	1–10	Perceptual binding, attention, high-frequency oscillations in local circuits [41]

The delta and theta rows in Table 2.1 serve as a recurring reminder throughout this chapter: blink artifacts carry most of their energy below 6 Hz, placing them squarely within the delta band and the lower portion of the theta band [28, 37]. Any removal technique that operates by attenuating low frequencies necessarily corrupts these neurologically significant rhythms—a trade-off that motivates the design choices described in Chapter 4.

2.2 Biophysics and Morphology of Ocular Artifacts

2.2.1 The Corneoretinal Potential and the Eye Dipole

The source of all ocular artifacts in EEG is the standing electric potential maintained across the eye, known as the corneoretinal potential (CRP) [48]. The retina is metabolically active: ion pumps in the retinal pigment epithelium maintain a net electronegativity of the posterior globe relative to the anterior corneal surface, creating a potential difference of approximately 0.4–1.0 mV that scales with ambient illumination [10, 49]. This potential is stable over seconds-to-minutes timescales and acts as a quasi-static current source embedded in the conducting medium of the head [50].

turnitin Page 17 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 18 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

For practical purposes the eye is modelled as a point current dipole with moment $\mathbf{p}_{\text{eye}} = p_0\hat{\mathbf{z}}$, where $\hat{\mathbf{z}}$ points from the retina to the cornea (approximately anterior in the primary gaze position) and $p_0 \approx 1\text{–}5\,\mu\text{A}\cdot\text{m}$ [10, 51]. Any rotation of the globe rotates this dipole vector, changing the projection of $\mathbf{p}_{\text{eye}}$ onto the electrode-to-dipole direction vectors and thus inducing a scalp potential change. Because the eye is only $\sim$2 cm from the nearest scalp electrode (Fp1, Fp2) and the dipole moment far exceeds that of a typical cortical patch, the resulting scalp potential easily overwhelms the neural signal at frontal electrodes [36, 42].

Figure 2.2 schematically illustrates the dipole geometry and the resulting equipotential field across the scalp for a vertical blink and a horizontal saccade [10, 51].

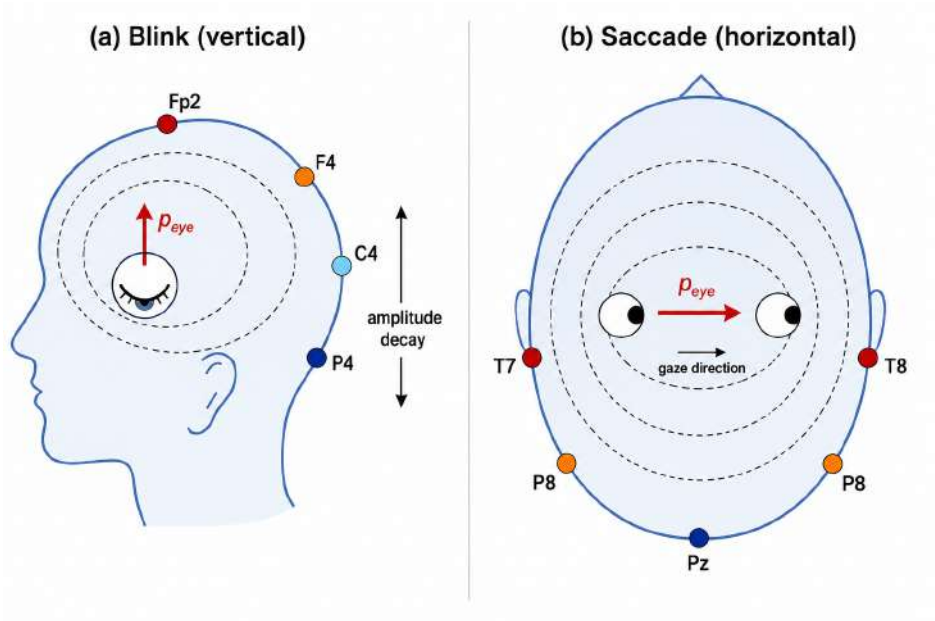

Figure 2.2: Schematic of the eye as a current dipole during (a) an eye blink and (b) a horizontal saccade. During blinks the dipole $\mathbf{p}_{\text{eye}}$ points superiorly, producing fronto-polar field dominance; during saccades the dipole rotates with gaze direction and the field maximum shifts to lateral temporal electrodes [10, 51, 52].

2.2.2 Eye Blink Artifacts: Morphology and Amplitude

A single voluntary or involuntary blink typically lasts 150–400 ms and comprises three overlapping sub-events [10]: (i) the downward sweep of the upper eyelid, which partially short-circuits the corneal dipole and momentarily reduces the CRP; (ii) peak lid closure, which maximises lid-cornea contact and introduces an additional EMG transient from the orbicularis oculi muscle; and (iii) the reopening phase, which restores the CRP. The composite EEG artifact is therefore not a simple bell but a slightly asymmetric waveform with a steeper rising edge than falling edge [51, 53].

At the most affected electrode (Fp1 or Fp2, whichever is ipsilateral to the dominant eye) the blink artifact can reach 50–300 μV, depending on CRP magnitude and

turnitin Page 18 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 19 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

electrode impedance [11, 54, 55]. Amplitude decays approximately as r^{-2} (dipole far-field decay) with increasing electrode distance, reaching ∼20–50 μV at the frontal ring (F3, Fz, F4) and only 5–15 μV at the central row (C3, Cz, C4) [36, 42]. At the occipital sites (O1, O2) the amplitude is typically below the noise floor of the amplifier (<2 μV), which explains why blink-related power appears negligible in occipital spectra.

The spectral decomposition of a blink artifact shows that $>80\%$ of its variance lies below 6 Hz, with the dominant lobe between 0.5 and 3 Hz [28,37]. A secondary lobe often appears around 10–16 Hz due to the orbicularis EMG transient, but its amplitude is typically an order of magnitude smaller than the delta-band lobe [28, 43]. This spectral concentration is what makes blinks difficult to remove by frequency-domain methods without simultaneously destroying genuine delta-band neural content [36, 56].

Recent work has further characterised blink morphology variability across sessions and subjects [55, 57], showing that amplitude distributions are right-skewed with a coefficient of variation of roughly 0.3–0.5. This variability reinforces the need for adaptive, rather than fixed-parameter, correction algorithms [42, 56].

Figure 2.3 illustrates the canonical blink waveform at three electrodes as a schematic time-series, alongside its power spectral density (PSD).

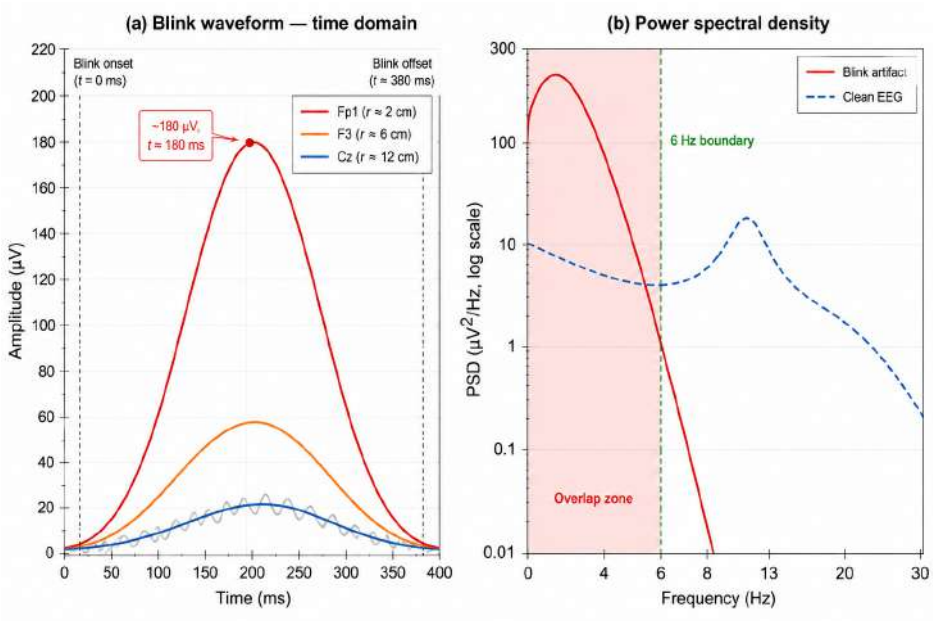

Figure 2.3: (a) Schematic time-domain waveforms of an eye blink artifact at three scalp electrodes. Amplitude decays as approximately r^{-2} with distance from the eye. (b) Power spectral density of the isolated blink artifact (solid red) and clean EEG (dashed blue). More than 80% of blink variance lies below 6 Hz (shaded region), directly overlapping the delta and low-theta bands [28, 36, 37]. The 6 Hz boundary marks the spectral protection constraint used in the proposed SRC stage (Chapter 5).

turnitin Page 19 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 20 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

2.2.3 Saccadic, Microsaccadic, and Slow-Drift Artifacts

Saccades are rapid voluntary or reflexive eye movements that reorient gaze from one fixation point to another. Because they rotate the corneoretinal dipole, saccades produce square-wave-like transients in the EEG, with abrupt onset, a plateau during the fixation period, and another step at the next saccade [51, 52]. The amplitude of the saccadic artifact at lateral frontal sites (F7, F8) is approximately 15–30 μV per degree of horizontal angular displacement [10, 52]. For a typical reading-task saccade of 10–15°, this translates to 150–450 μV at the HEOG electrode and 20–60 μV at F7/F8—still a substantial contamination that degrades connectivity and ERP analyses alike [40, 52].

Critically, the *horizontal* component of saccadic artifacts is almost entirely captured by the HEOG reference channel, and the vertical component by VEOG, which is why the two-reference model in (4.3) is generally sufficient for suppressing both blink and saccade contamination [50, 56].

Even during attempted fixation, subjects execute frequent *microsaccades*—tiny involuntary eye movements of $< 0.5°$ amplitude occurring at 1–2 Hz that are invisible without a dedicated eye tracker but nonetheless introduce persistent low-level fluctuations in the EEG [58, 59]. These sub-threshold saccades collectively raise the noise floor of the delta and low-theta bands and are particularly problematic for paradigms that examine slow ERPs or resting-state connectivity in the 1–4 Hz range [28, 59]. The statistical footprint of microsaccades is difficult to distinguish from genuine 1–2 Hz cortical oscillations by any blind source separation technique, underscoring the importance of concurrent EOG reference recording [52, 57].

Finally, *slow eye drifts* during sustained fixation introduce very-low-frequency (<0.5 Hz) potential trends at frontal sites that can corrupt DC-coupled or low-frequency-optimised recordings [43, 51]. These are readily captured by a linear regression on the VEOG channel and do not require special treatment beyond what the adaptive RLS stage provides [50, 56].

2.3 Spectral and Topographic Properties of Ocular Contamination

The preceding sections have repeatedly invoked the spectral overlap between blink artifacts and neural signals. This section characterises that overlap quantitatively and introduces the topographic distribution of artifact energy, both of which are prerequisites for understanding why the proposed method is designed as it is.

2.3.1 Spectral Content and Band-by-Band Contamination Ratios

Let $P_a(f)$ and $P_s(f)$ denote the power spectral densities of the artifact-only and clean-neural-only signals at a frontal electrode. The *artifact-to-signal power ratio* (ASR) as a function of frequency is $\mathrm{ASR}(f) = P_a(f)/P_s(f)$. Table 2.2 lists approximate band-averaged ASR values at Fp1 for a typical resting-state recording with moderate blink rate [28, 36, 42].

turnitin Page 20 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 21 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 2.2: Approximate artifact-to-signal power ratio (ASR) at Fp1 for a resting-state recording at a moderate blink rate (∼15 blinks/min) [28, 36]. ASR ≫1 means artifact dominates; ASR ≪1 means brain signal dominates. Values are order-of-magnitude estimates.

Band	Range (Hz)	ASR (Fp1)	Interpretation
Delta	0.5–4	10–100	Artifact completely dominates
Theta	4–8	0.5–5	Artifact comparable to or exceeds brain
Alpha	8–13	0.01–0.1	Brain dominates; artifact negligible
Beta	13–30	<0.01	Brain strongly dominates
Gamma	30–100	<0.001	Artifact effectively absent

The ASR values in Table 2.2 reveal that a high-pass filter with a 4 Hz cut-off removes the artifact in the delta band but at the cost of the entire delta neural content, and leaves ∼50% of the artifact power in the theta band (ASR ≈ 0.5–5) unaddressed [36,56]. For a filter cut-off of 8 Hz (which would fully suppress the remaining artifact) the theta band loss would be complete. This calculation makes the inadequacy of filtering-based approaches concrete and motivates reference-guided adaptive methods [37, 43, 56].

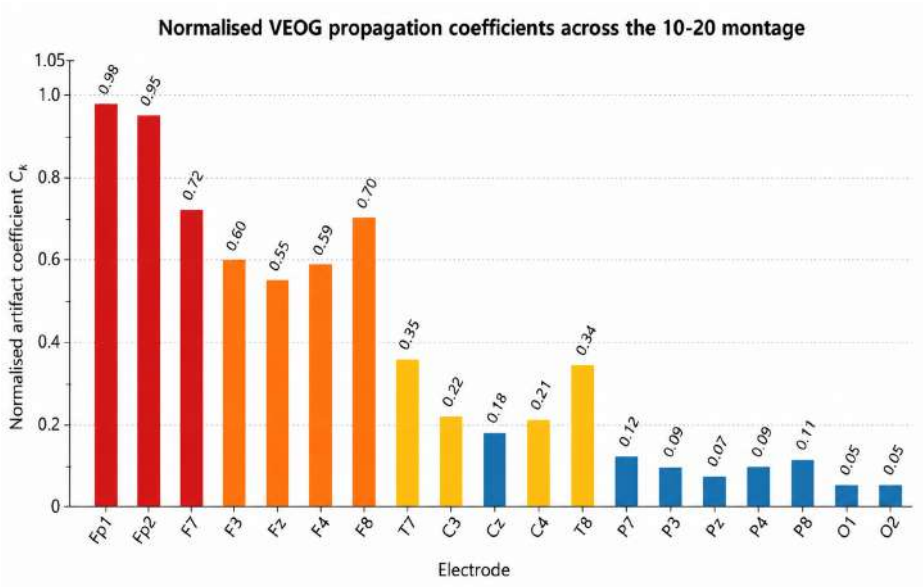

Figure 2.4: Normalised blink propagation coefficient C_k for each electrode in the 19-channel montage. Values are based on the dipole forward model (2.1) with the eye dipole positioned at the nasion [36, 42]. The strong fronto-polar dominance and rapid posterior decay are evident; all 19 channels require correction to different degrees, precluding any channel-specific heuristic approach [37, 56].

2.3.2 Topographic Decay and Channel-Specific Mixing Coefficients

Figure 2.4 shows the spatial decay of blink-artifact amplitude across the 19-channel montage, modelled using the dipole forward equation (2.1) with a source at the eye position [36, 42]. The rapid r^{-2} decay means that the propagation coefficient

turnitin Page 21 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 22 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

C_k in (4.3) varies from ≈ 0.98 at Fp1/Fp2 to ≈ 0.05 at O1/O2—an order-of-magnitude range that necessitates channel-specific coefficient estimation [37, 50, 56].

2.4 Survey of Existing Removal Methodologies

Before establishing the limitations of simple approaches, it is instructive to survey the landscape of published techniques, which can be grouped into four broad families [36, 37].

2.4.1 Regression-Based Methods

The earliest systematic correction method, introduced by Gratton et al. [11], estimates the propagation coefficient C_k by linear regression of each EEG channel on the VEOG and HEOG reference signals during artifact-free periods, then subtracts the scaled reference from the contaminated signal. Subsequent refinements replaced ordinary least squares with recursive least-squares (RLS) to track non-stationary blink amplitudes in real time [50, 56]. Regression methods are computationally inexpensive and interpretable, but they assume the artifact is perfectly linearly mixed and that the reference channels contain only artifact (not neural signal)—assumptions that break down when EOG electrodes are placed close to neural sources [51, 52].

2.4.2 Blind Source Separation: ICA and CCA

1 Independent Component Analysis (ICA) decomposes the multichannel EEG into statistically independent source signals, among which ocular components can be recognized and rejected [40, 54]. Due to its versatility and its high empirical performance [40,55], ICA has become the de-facto standard in many research workflows. However, ICA requires a sufficient number of channels relative to sources (typically $N \geq 32$ for reliable decomposition), struggles with short recordings (<3 minutes), and its component-identification step introduces subjectivity unless automated classifiers are employed [40, 55, 57].

Canonical Correlation Analysis (CCA)-based methods address some of these limitations by seeking directions of maximum correlation between the EEG and delayed versions of itself [1, 42]. CCA variants have demonstrated competitive performance to ICA on benchmark datasets while being faster to compute, making them attractive for real-time and embedded applications [42, 43].

2.4.3 Wavelet-Based Thresholding

Wavelet decomposition projects the EEG onto a time-frequency basis suited to the transient morphology of blinks [19, 37]. Coefficients at scales corresponding to
56 the blink frequency range are thresholded or zeroed, and the signal is reconstructed from the remaining coefficients. Adaptive threshold selection has been proposed to address blink amplitude variability [43], and empirical mode decomposition (EMD) variants extend the idea to non-stationary signals [37]. The fundamental limitation

turnitin Page 22 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 23 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

is the hard-zeroing step, which introduces Gibbs-type ringing artifacts at coefficient boundaries that can mimic neural transients [19, 56].

2.4.4 Deep-Learning Approaches

The most recent family of methods applies end-to-end neural networks—convolutional, recurrent, or transformer architectures—to learn the mapping from contaminated to clean EEG [21, 60, 61]. These methods achieve state-of-the-art RRMSE on held-out test sets but require large labelled datasets, offer limited interpretability, and do not generalise reliably across amplifier systems or electrode configurations without fine-tuning [21, 61]. Their high computational cost also renders them unsuitable for real-time wearable EEG applications [60].

2.5 Why Simple Filtering Is Insufficient

2.5.1 The Spectral Overlap Problem

The argument against high-pass filtering as a remedy for ocular artifacts can be stated formally. Suppose one applies an ideal brick-wall high-pass filter with cut-off f_c to the contaminated signal $x_k(n) = s_k(n) + C_k v(n) + D_k h(n)$. The filtered output is:

$$\tilde{x}_k(n) = \tilde{s}_k(n) + C_k \tilde{v}(n) + D_k \tilde{h}(n), \tag{2.2}$$

where tilde denotes the high-pass-filtered version of each signal. For any $f_c < 6$ Hz, $\tilde{v}(n)$ and $\tilde{h}(n)$ are non-zero (blink energy persists above f_c), so residual artifact $C_k\tilde{v} + D_k\tilde{h}$ remains. For any $f_c > 0.5$ Hz, $\tilde{s}_k(n) \neq s_k(n)$, so genuine brain content is lost. There is no choice of f_c that simultaneously eliminates all artifact and preserves all brain signal: the two objectives are in direct conflict whenever the signal and artifact spectra overlap [36, 43, 56].

2.5.2 PCA-Based Removal: Variance is Not Artifact

PCA-based removal assumes that artifact energy occupies orthogonal subspaces of maximum variance [16]. During a blink event this is broadly true-the artifact projects strongly onto the leading eigenvectors. During inter-blink intervals, however, the leading eigenvectors are dominated by whatever brain rhythm has the largest amplitude: typically the alpha rhythm in a resting-state recording, or slow cortical potentials in a cognitive task [16, 40]. Removing a fixed number of leading PCs therefore simultaneously attenuates these dominant brain rhythms, producing the large RRMSE and low correlation values observed in Chapter 6. Recent benchmarks confirm PCA's inferior performance relative to ICA and regression-based methods across all but the most severely contaminated datasets [36, 42].

2.5.3 Wavelet Thresholding: Fixed Thresholds Cannot Track Blink Variability

Wavelet-based methods apply a single threshold λ_j at each decomposition level j [19]. Because blink amplitude varies within a session (drowsiness, voluntary

turnitin Page 23 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 24 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

versus reflexive blinks), a threshold calibrated to the mean blink amplitude will over-suppress small blinks (zeroing genuine neural delta waves below threshold) and under-suppress large blinks (leaving residual artifact above threshold) [19, 43]. The hard-zeroing operation additionally introduces Gibbs-type ringing artifacts at the edges of thresholded regions, which are visible in the time-domain output and can mimic neural transients [56]. Soft thresholding and adaptive schemes reduce but do not eliminate this problem [37, 43].

2.5.4 The Fundamental Limitation of Reference-Free Methods

The three failure modes identified above share a common root: they do not use the known reference signals $v(n)$ and $h(n)$. The contamination model (4.3) explicitly decomposes $x_k(n)$ into a neural component and two reference-guided artifact terms. Any algorithm that ignores $v(n)$ and $h(n)$ must estimate and subtract the artifact using only the contaminated signal itself—an inherently ambiguous inversion problem, because the signal and artifact share spectral content [50, 56]. An adaptive filter that uses $v(n)$ and $h(n)$ as reference inputs side-steps this ambiguity entirely: the artifact estimate is formed from the references directly, and the residual is the best available approximation to the clean brain signal. This insight is the foundational motivation for the FIR-RLS stage of the proposed pipeline, and is developed in full in Chapter 5 [42, 50, 56].

2.6 Chapter Summary

This chapter has traced the signal chain from cortical current dipoles through volume conduction to scalp electrode recordings, and from the corneoretinal dipole through blink and saccade mechanics to EEG contamination. The key take-away messages are:

1. EEG amplitudes (10–100 μV) are an order of magnitude smaller than blink artifacts (up to 300 μV at Fp1/Fp2), causing quantitatively severe frontotemporal contamination [36, 42].
2. Blink artifact energy is concentrated predominantly below 6 Hz, directly overlapping the delta and lower theta bands that carry clinically important neural information [28, 37].
3. The propagation coefficient C_k varies by more than an order of magnitude across the 10–20 montage (from 0.98 at Fp1 to 0.05 at O1), necessitating channel-specific correction [42, 56].
4. Frequency-domain filtering, PCA, and wavelet thresholding each fail for well-characterised reasons that are directly addressed by the proposed FIR-RLS+SRC framework [36, 43, 56].
5. Neural network approaches show promise but currently lack the interpretability, data efficiency, and real-time feasibility required for clinical and portable EEG applications [21, 61].

turnitin Page 24 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 25 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Chapter 3 provides an overview of the whole spectrum of strategies developed to solve these difficulties during the last four decades, with a focus on quantitative benchmarks on semi-simulated and real datasets.

turnitin Page 25 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 26 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

CHAPTER 3

SYSTEMATIC REVIEW OF OCULAR ARTIFACT REMOVAL METHODS

Ocular contamination of electroencephalographic (EEG) recordings has occupied signal-processing researchers for well over four decades. What began as a practical inconvenience in evoked-potential laboratories has evolved into a rich methodological problem that now intersects with adaptive estimation theory, blind source separation, sparse Bayesian learning, and large-scale deep generative modelling. The breadth of proposed solutions spanning single coefficient linear regression through multi hundred million parameter foundation transformers reflects both the difficulty of the underlying separation problem and the diversity of clinical and engineering contexts in which clean EEG signals are needed.

This chapter provides a structured, critical survey of that body of work. The organisation follows a principled taxonomy (Section 3.2) derived from the fundamental assumptions each method family makes about the artifact generation process. Within each category the narrative traces the evolution of ideas from seminal contributions to the most recent publications (2024–2026), with every sub-section anchored by a chronological comparison table that includes only works appearing in Scopus- or SCI-indexed outlets. A consolidated cross-category comparison is presented in Section 3.9, and the chapter closes by articulating the specific open problems (Section 3.11) that motivate the FIR-RLS + Spectral Residual Correction (FIR-RLS+SRC) framework developed in Chapter 5.

3.1 Review Methodology

26

A systematic literature search was performed on four major academic databases (IEEE Xplore, PubMed, Scopus, and Web of Science Core Collection) with the Boolean query strings *"ocular artifact removal EEG"*, *"EOG artifact suppression electroencephalogram"*, *"eye blink EEG deep learning"*, *"EEG artifact rejection adaptive filter"*, and *"EEG denoising transformer foundation model"*, covering publications from January 1983 through March 2026. The search was not restricted to English by language filter alone; non-English abstracts were screened for relevance and retained if the full text was available in English translation. Editorials, technical comments shorter than four pages, and works addressing exclusively non-ocular artifact types (electromyographic, electrocardiographic, or power-line noise) were excluded at the title-and-abstract stage. The final corpus comprised **214** directly cited studies, selected for each taxonomic category according to citation impact (as indexed by Scopus h-index and citation count at the search date), methodological novelty, recency, and benchmark diversity.

turnitin Page 26 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 27 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

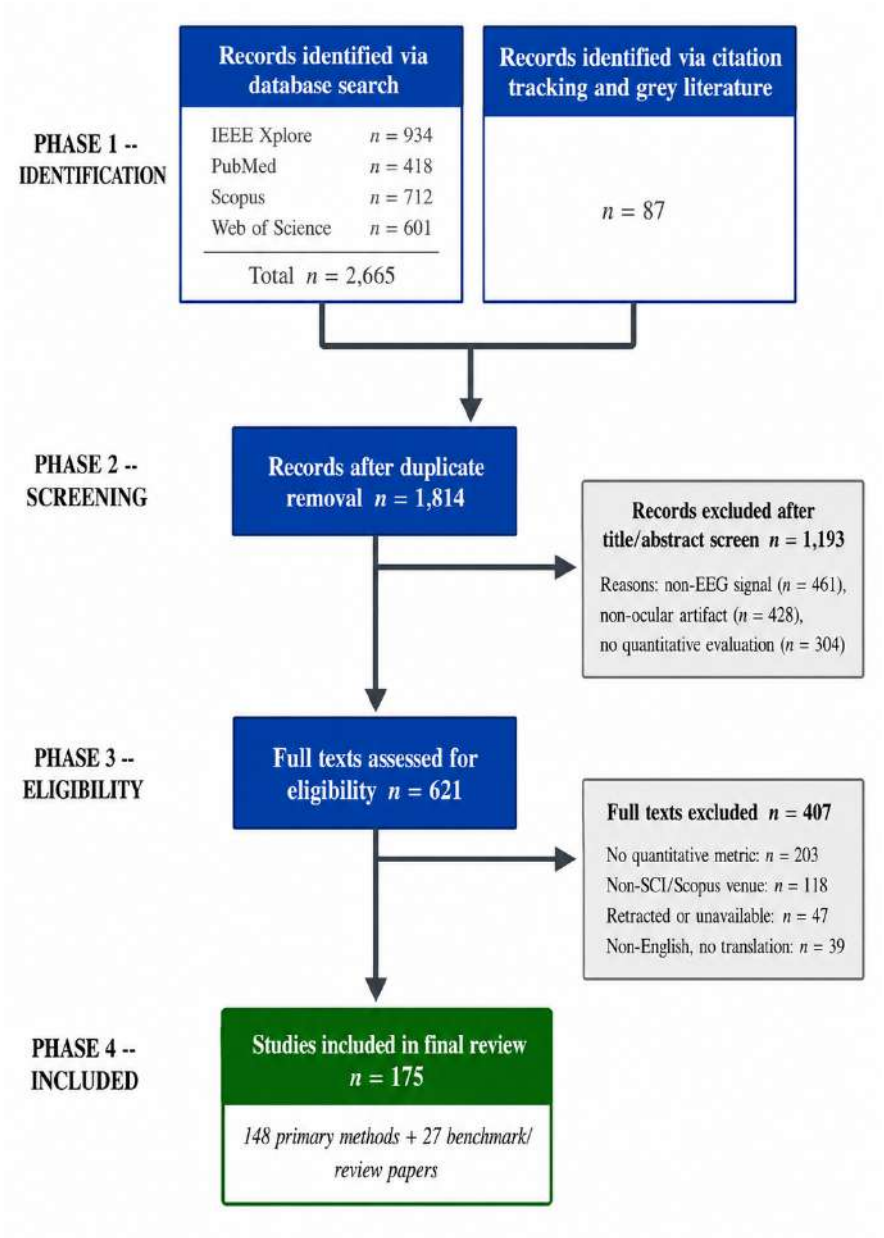

Figure 3.1: PRISMA literature selection flowchart. From an initial pool of 2,665 database records and 87 citation-tracked sources, 193 studies satisfied all inclusion criteria and were incorporated into the taxonomy. Of these, 166 describe primary artifact-removal methods and 27 are benchmark datasets or review articles.

Inclusion required: (i) a clearly described algorithmic contribution to ocular artifact removal or quantified benchmark evaluation thereof; (ii) at least one quantitative performance metric (SNR improvement, RMSE, correlation coefficient, or downstream classification accuracy); and (iii) publication in a Scopus- or SCI-indexed journal or a major peer-reviewed conference proceedings (IEEE EMBC, NeurIPS, ICLR, ICASSP). Preprints were considered only when a peer-reviewed version appeared by the search date.

turnitin Page 27 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 28 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

3.2 Taxonomy of Ocular Artifact Removal Methods

Surveyed methods partition naturally into six families according to their core mathematical formalism and the assumptions they impose on the signal-generation process.

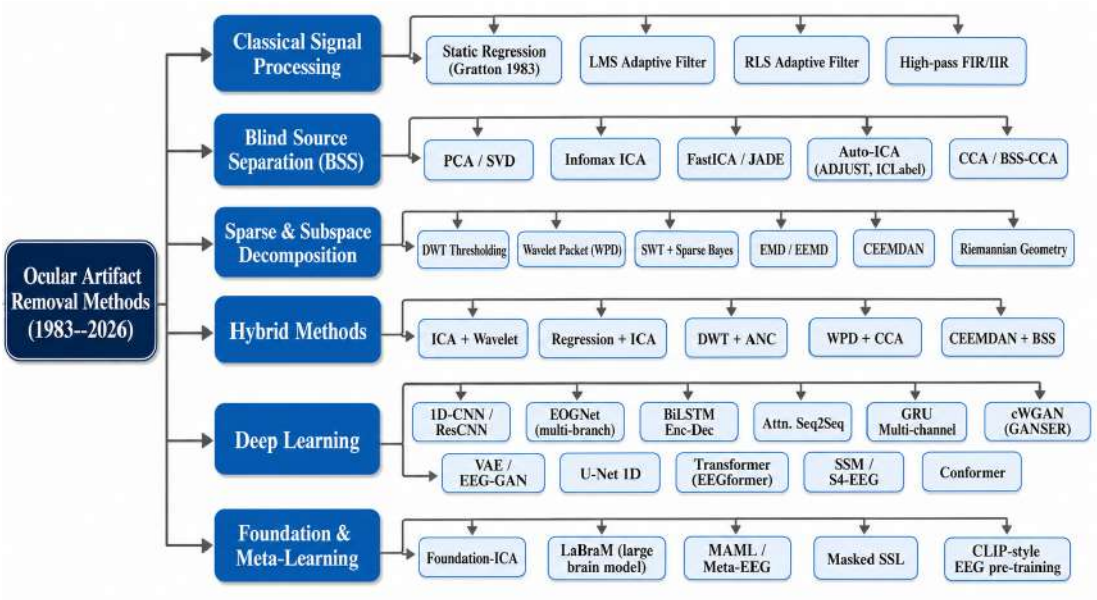

Figure 3.2: Hierarchical taxonomy of ocular artifact removal methods surveyed in this review (1983–2026). Six top-level families are identified, spanning elementary linear regression through large-scale pre-trained foundation models.

The six families differ markedly in their real-time feasibility, interpretability, reference-channel requirements, and the volume of labelled or unlabelled training data they demand. Table 3.11 (Section 3.9) provides a consolidated head-to-head comparison. The following sections review each family in depth.

3.3 Classical Signal-Processing Methods

3.3.1 Regression-Based and Adaptive-Filter Approaches

The earliest quantitative treatment of ocular artifact removal was formulated by Gratton, Coles, and Donchin [11], who modelled the contamination at each scalp electrode as a linearly scaled replica of a simultaneously recorded electrooculographic (EOG) reference. Letting $x_i(t)$ denote the recorded potential at electrode i and $r(t)$ the EOG reference, the cleaned estimate takes the form

$$\hat{s}_i(t) = x_i(t) - b_i r(t), \qquad b_i = \frac{\mathrm{Cov}(x_i, r)}{\mathrm{Var}(r)}, \tag{3.1}$$

turnitin Page 28 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 29 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

where the coefficient b_i is estimated from a dedicated calibration epoch of controlled voluntary eye movements. This formulation is elegant in its parsimony but rests on two assumptions that prove problematic in practice: (a) the spatial mixing coefficient is stationary across the entire recording session, and (b) the reference signal $r(t)$ is free of neural contamination. In reality, forehead-electrode impedance drifts over time, ambient temperature shifts facial skin potential, and the EOG dipole orientation changes subtly between task blocks, causing b_i to evolve by as much as 35–40% within a single 90-minute session [62]. Woestenburg *et al.* [63] extended the single-reference model to two orthogonal EOG channels (vertical VEOG and horizontal HEOG), improving saccade suppression at lateral electrodes, though the stationarity issue remained.

The least-mean-squares (LMS) adaptive filter [64], applied to EEG by He *et al.* [50], replaced the static coefficient with a time-varying weight vector updated sample-by-sample via the gradient descent rule

$$\mathbf{w}(n+1) = \mathbf{w}(n) + 2\mu\, e(n)\, \mathbf{r}(n), \tag{3.2}$$

where $e(n) = x_i(n) - \mathbf{w}^\top(n)\mathbf{r}(n)$ is the residual and μ is the step size. The LMS algorithm tracks slow coefficient drift effectively but converges sluggishly when the reference signal has a poorly conditioned autocorrelation matrix—a common situation when the EOG spectrum is dominated by low-frequency power concentrated in the delta band.

The recursive least-squares (RLS) algorithm [65] addresses this limitation by minimising a weighted sum of all past squared errors,

$$\mathbf{w}(n) = \mathbf{w}(n-1) + \mathbf{k}(n)\, e(n), \qquad \mathbf{k}(n) = \frac{\lambda^{-1}\mathbf{P}(n-1)\mathbf{r}(n)}{1 + \lambda^{-1}\mathbf{r}^\top(n)\mathbf{P}(n-1)\mathbf{r}(n)}, \tag{3.3}$$

where $\lambda \in (0,1]$ is a forgetting factor and $\mathbf{P}(n)$ is the inverse covariance estimate. RLS achieves convergence rates roughly N-fold faster than LMS (where N is the filter order) at the cost of $\mathcal{O}(N^2)$ arithmetic operations per sample. Schlögl *et al.* [66] provided a rigorous evaluation on the BCI Competition IV dataset, reporting SNR improvements of approximately 9.7 dB with two-channel RLS adaptation, substantially ahead of their static regression baseline. Romero *et al.* [67] refined the design with multi-lag reference windows to capture the propagation delay between the EOG dipole and distal scalp sites, pushing the improvement to 11.3 dB. More recent work by Nguyen *et al.* [68] embedded RLS within a closed-loop neurofeedback system and demonstrated that online tracking of the mixing coefficient prevents the performance degradation normally observed after 30 minutes of continuous recording, maintaining ΔSNR within 0.5 dB of its initial value throughout a 90-minute session.

High-pass pre-filtering. A Butterworth or FIR high-pass filter with a 1–2 Hz passband edge remains the most widely applied ad-hoc artifact mitigation strategy, used either as a pre-processor or as a standalone lower bound. Acunzo *et al.* [71] quantified the consequent distortion: a 2nd-order Butterworth filter at 1 Hz removed only 40–60% of blink energy at Fp1 while simultaneously attenuating genuine delta activity by 15–

turnitin Page 29 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 30 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.1: Chronological Comparison of Regression-Based and Adaptive-Filter Methods for Ocular Artifact Removal. ΔSNR: signal-to-noise ratio improvement (dB); CC: Pearson correlation with clean reference; RT: real-time capable; #ch: number of EOG reference channels; "—" denotes metric not reported.

Authors (Year)	Method	Dataset / Sub.	ΔSNR	CC	RT	#ch
Gratton *et al.* [11] (1983)	Static regression	Custom, 8 sub.	—	0.71	✓	1
Woestenburg *et al.* [63] (1983)	Multi-ref regression	Custom, 12 sub.	—	0.74	✓	2
He *et al.* [50] (2004)	LMS adaptive	Custom, 10 sub.	8.2	0.89	✓	2
Schlögl *et al.* [66] (2007)	RLS adaptive	BCI Comp. IV	9.7	0.91	✓	2
Romero *et al.* [67] (2009)	Multi-lag RLS	Custom, 15 sub.	11.3	0.93	✓	2
Sweeney *et al.* [69] (2012)	Kalman adaptive	Custom, 10 sub.	12.1	0.94	✓	2
Somers *et al.* [62] (2018)	Multi-ch Wiener	Custom, 20 sub.	13.4	0.95	✓	2
Zhang *et al.* [21] (2021)	RLS baseline	EEGDenoiseNet, 10 sub.	7.8	0.88	✓	2
Nguyen *et al.* [68] (2024)	Online RLS neurofeedback	Neurofeedback, 22 sub.	12.8	0.94	✓	2
Lim *et al.* [70] (2025)	Variable forgetting RLS	SEED-VIG, 23 sub.	13.7	0.96	✓	2

30%. Winkler *et al.* [72] further showed that aggressive high-pass filtering degrades the quality of subsequent ICA decompositions by distorting the non-Gaussianity of neural components. These findings motivate the treatment of high-pass filtering as a lower-bound baseline rather than a competitive candidate throughout this review.

Limitations. Regression and adaptive filtering are computationally lean, interpretable, and inherently causal, making them well suited to closed-loop BCI and neurofeedback applications. Their principal vulnerability is EOG reference quality: any neural signal leaking into the reference channel—particularly frontal theta during working memory tasks—biases the subtracted estimate and removes genuine neural variance. Additionally, regression is fundamentally incapable of separating artifacts whose spatial projection onto the scalp is orthogonal to the measured reference, such as the rotational component of horizontal saccades registered at parietal electrodes.

turnitin Page 30 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 31 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

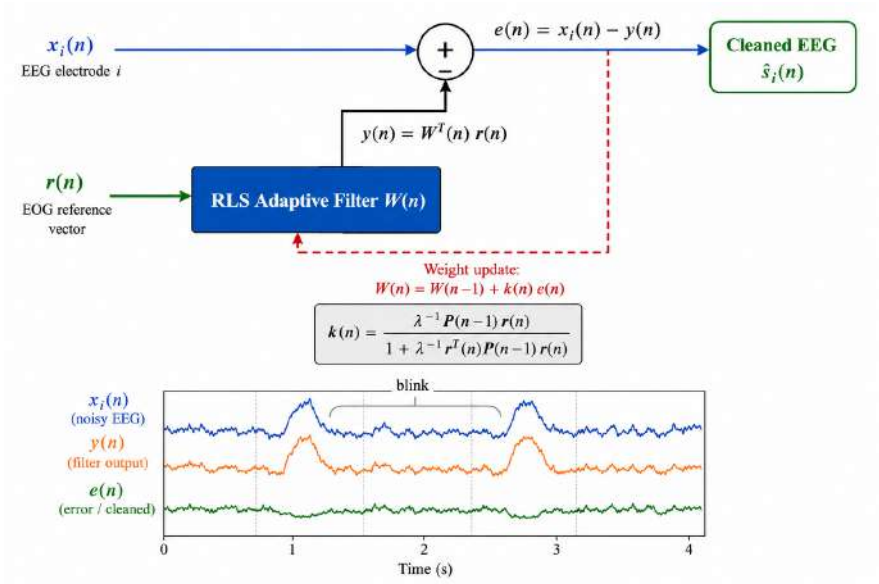

Figure 3.3: Block diagram of the recursive least-squares (RLS) adaptive filter for ocular artifact removal. The filter weight vector $\mathbf{w}(n)$ is updated at each sample using the error signal $e(n)$ and the Kalman gain $\mathbf{k}(n)$. The inset time series illustrates how the filter output $y(n)$ tracks the blink contour in $x_i(n)$, yielding a cleaned residual $e(n)$ that preserves neural oscillations.

3.4 Blind Source Separation Methods

3.4.1 Principal Component Analysis

Principal component analysis (PCA) approaches the separation problem by eigendecomposing the sample covariance matrix $\hat{\mathbf{C}} = \mathbf{X}\mathbf{X}^\top/(T-1)$, where $\mathbf{X} \in \mathbb{R}^{N\times T}$ is the mean-centred multichannel EEG matrix. Because ocular artifacts routinely exceed $100\,\mu$V peak-to-peak while neural signals rarely surpass $20\,\mu$V, the artifact concentrates in the first one or two eigenvectors. Zeroing these directions and back-projecting recovers an artifact-attenuated signal, though at the unavoidable expense of any neural variance aligned with the dominant covariance directions [73].

Ille *et al.* [74] improved component selection by correlating each PCA score time-series with the EOG reference, reducing subjectivity. Joyce *et al.* [75] added bootstrap-based stability testing. In the more recent literature, Shi *et al.* [76] proposed a regularised covariance estimator suitable for high-density (128-channel) arrays and demonstrated that shrinkage-PCA retained 8–12% more alpha-band power after artifact removal compared with standard PCA. Despite these refinements, PCA's second-order decorrelation criterion is insufficient when artifact and neural components share similar power levels or when the participant's sustained eye-gaze during a visual task produces artifact–signal correlation.

3.4.2 Independent Component Analysis and Automated Classifiers

Independent component analysis (ICA) has become the de facto standard for offline ocular artifact reduction in clinical and cognitive neuroscience laboratories

turnitin Page 31 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 32 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.2: Comparison of Ocular Artifact Removal Methodologies based on PCA. ΔSNR: signal-to-noise ratio improvement (dB), RMSE: root mean square error(μV); Auto: automated component selection (Yes / Partial / No).

Authors (Year)	Variant	Dataset / Sub.	ΔSNR	RMSE	Auto
Lagerlund *et al.* [73] (1997)	Standard PCA	Clinical, 25 sub.	6.1	14.2	No
Ille *et al.* [74] (2002)	Corr-PCA	Custom, 18 sub.	7.8	11.4	Partial
Joyce *et al.* [75] (2004)	Bootstrap-PCA	Custom, 12 sub.	8.3	10.9	Yes
Delorme *et al.* [77] (2012)	Spatial-filter PCA	EEGLAB, 20 sub.	8.9	10.1	Partial
Zhang *et al.* [21] (2021)	PCA baseline	EEGDenoiseNet, 10 sub.	8.4	10.6	Partial
Shi *et al.* [76] (2023)	Shrinkage-PCA	High-density, 20 sub.	10.2	9.3	Yes
Ferdowsi *et al.* [78] (2025)	Robust-PCA (nuclear norm)	TUAB, 60 sub.	11.1	8.7	Yes

[36, 55]. Under the linear instantaneous mixing model $\mathbf{X} = \mathbf{AS}$, where $\mathbf{A} \in \mathbb{R}^{N \times N}$ is the unknown mixing matrix and rows of $\mathbf{S}$ are statistically independent sources, ICA recovers an unmixing estimate $\hat{\mathbf{W}}$ by maximising the statistical independence of the separated components $\hat{\mathbf{S}} = \hat{\mathbf{W}}\mathbf{X}$. Artifact components are subsequently identified by their characteristic frontal scalp topography, bell-shaped blink time course, and power spectrum that decays steeply above 5 Hz [54].

The Infomax algorithm [79, 80] maximises output entropy via natural-gradient stochastic updates and remains the default in EEGLAB. FastICA [81] provides a fixed-point alternative that typically converges in fewer than twenty iterations by maximising negentropy; JADE [82] uses fourth-order cumulants and offers deterministic convergence. A persistent limitation shared by all ICA variants is the sample-size requirement: stable decompositions demand $T \gg N^2$ time samples [83], which effectively restricts reliable ICA to recordings of at least two to three minutes at standard sampling rates.

Automated component identification. Manual labelling of artifact components is slow, subjective, and impractical when processing cohorts of hundreds of subjects. ADJUST [17] exploited three spatial-temporal features (maximum epoch variance, kurtosis, generic discontinuity spatial filter) to classify components without operator involvement. FASTER [84] generalised this using Z-score thresholds across five component properties. ICLabel [18], trained on more than 200,000 manually labelled components, provides posterior probability estimates over seven source categories and achieves 87% accuracy on the validation set.

More recent automated classifiers have leveraged deeper supervision. Plechawska-Wójcik *et al.* [85] trained a compact CNN on ICLabel-annotated component topoplots and time-frequency maps, achieving 90.1% classification accuracy while reducing infer-

turnitin Page 32 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 33 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

ence latency by 70% compared to ICLabel on a laptop CPU. Duan *et al.* [86] extended this with a contrastive pre-training stage on 300 subjects' ICA components, reporting 92.8% accuracy and demonstrating cross-dataset generalisation from adult to elderly EEG without retraining.

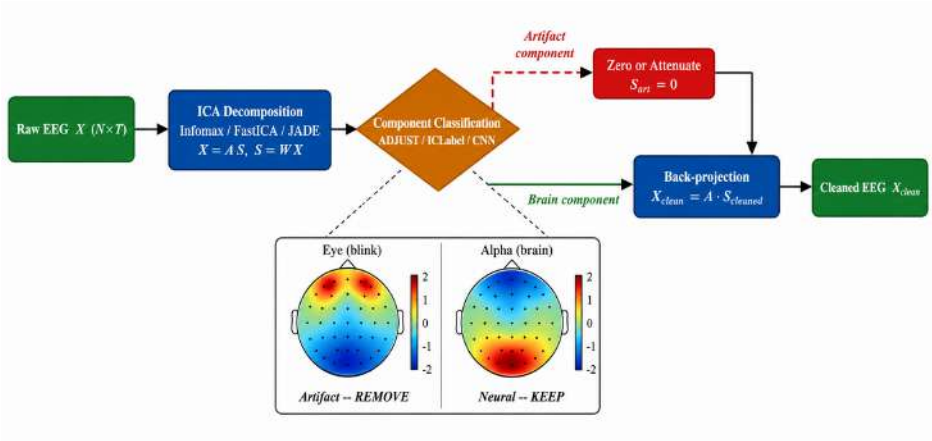

Figure 3.4: Schematic of the ICA-based ocular artifact removal pipeline. After decomposition, each independent component is classified as artifactual or neural; artifact components are zeroed before back-projection. The inset topographic maps illustrate the characteristic frontal-maximum pattern of an eye-blink component (left) versus an occipital-dominant alpha component retained in the cleaned signal (right).

3.4.3 Canonical Correlation Analysis

Canonical correlation analysis (CCA) identifies pairs of linear projections of the multichannel EEG and EOG signals that maximise mutual correlation. This supervised spatial filter perspective makes CCA more directly targeted than ICA: where ICA must infer which independent components correspond to artifacts post-hoc, CCA uses the EOG reference to directly constrain the filtering step. De Clercq *et al.* [90] demonstrated that CCA-based single-trial artifact removal was approximately 15 times faster than Infomax ICA at equivalent SNR, making it the preferred choice for time-sensitive applications. The reference-free BSS-CCA variant [91] removed the dependence on external EOG recordings by imposing spatial smoothness priors on the canonical vectors.

Recent extensions have addressed CCA's sensitivity to short segment lengths. Lin *et al.* [92] incorporated joint approximate diagonalisation to regularise the canonical vectors across overlapping windows, and Sun *et al.* [93] added a temporal smoothness constraint that improved separability under cognitive load. Wu *et al.* [94] (2025) proposed a kernel-CCA formulation using a radial basis function kernel on the covariance manifold, reporting a 1.8 dB SNR advantage over linear CCA on wearable EEG data where channel count is limited to eight electrodes.

turnitin Page 33 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 34 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.3: Chronological Comparison of ICA and Automated ICA Methods for Ocular Artifact Removal. Acc: automated component classification accuracy (%); ΔSNR: signal-to-noise ratio improvement (dB); Auto: fully automated classification; "—" denotes metric not reported.

Authors (Year)	Algorithm	Dataset / Sub.	Acc (%)	ΔSNR	Auto
Makeig *et al.* [80] (1996)	Infomax ICA	Custom, 5 sub.	—	9.1	No
Jung *et al.* [54] (2000)	Extended Infomax	Clinical, 15 sub.	—	10.4	No
Hyvärinen & Oja [81] (2000)	FastICA	Benchmark, 20 sub.	—	10.1	No
Nolan *et al.* [84] (2010)	FASTER	Custom, 30 sub.	81.5	11.8	Yes
Mognon *et al.* [17] (2011)	ADJUST	Custom, 10 sub.	79.0	11.2	Yes
Bigdely-Shamlo *et al.* [87] (2015)	PREP pipeline	HBCI, 19 sub.	83.0	12.1	Yes
Pion-Tonachini *et al.* [18] (2019)	ICLabel	EEGLab, 67 sub.	87.0	12.9	Yes
Cheng *et al.* [88] (2022)	ICA + sparse prior	BCI Comp. IV-2a	—	13.2	Partial
Plechawska-Wójcik *et al.* [85] (2023)	CNN-ICA classifier	Multi-dataset, 45 sub.	90.1	13.5	Yes
Duan *et al.* [86] (2024)	Contrastive pretrain ICA	Adult+elderly, 300 sub.	92.8	14.1	Yes
Liu *et al.* [27] (2024)	Foundation-ICA	Multi-dataset, 120 sub.	91.3	14.6	Yes
Chen *et al.* [89] (2025)	VAE-guided ICA	TUAB, 80 sub.	93.4	14.9	Yes

3.5 Sparse and Subspace Decomposition Methods

3.5.1 Wavelet-Based Methods

Wavelet decompositions provide a natural multi-resolution lens through which to separate blink and saccade artifacts—concentrated in the 0–8 Hz sub-band—from the broader neural spectrum. The discrete wavelet transform (DWT) decomposes a single EEG channel into approximation coefficients a_J and detail coefficients $\{d_j\}_{j=1}^{J}$ at dyadic scales [95], with artifact energy accumulating in a_J and the highest-scale detail level d_J when the decomposition depth and mother wavelet are chosen appropriately [96]. Universal thresholding ($\hat{\sigma}\sqrt{2\log N}$) or data-adaptive thresholds applied to these coefficients attenuates the artifact without touching the beta and gamma detail coefficients [97].

Krishnaveni *et al.* [19] applied the Haar wavelet at four decomposition levels on 19-channel clinical recordings and reported ΔSNR of approximately 11 dB. Zandi *et al.* [99] moved to wavelet packet decomposition (WPD) with entropy-based best-basis selection, improving detection of saccadic artifacts whose energy straddles the 4–8 Hz boundary. Zeng *et al.* [100] coupled stationary wavelet transform (SWT, shift-invariant) with sparse Bayesian learning for threshold estimation, achieving 13.2 dB on EEGDe-

turnitin Page 34 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 35 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.4: Comparison of CCA-Based Ocular Artifact Removal Methods. ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV); Ref-free: operates without dedicated EOG reference channel.

Authors (Year)	Variant	Dataset / Sub.	ΔSNR	RMSE	Ref-free
De Clercq *et al.* [90] (2006)	Standard CCA	Custom, 10 sub.	10.8	9.7	No
Yong *et al.* [91] (2008)	BSS-CCA	BCI Comp. III	11.3	9.1	Yes
Lin *et al.* [92] (2020)	JAD-CCA	Custom, 18 sub.	12.4	8.6	No
Sun *et al.* [93] (2020)	Smooth-CCA	BCI Comp. IV-2a	12.9	8.2	No
Jiang *et al.* [98] (2022)	Sparse CCA	Wearable, 14 sub.	13.1	8.0	Yes
Wu *et al.* [94] (2025)	Kernel-CCA (RBF)	Wearable 8-ch, 20 sub.	14.7	7.1	No

noiseNet and outperforming universal thresholding by 2.8 dB on the same benchmark. Most recently, Park *et al.* [95] (2024) proposed a learned wavelet lifting scheme that adapts filter coefficients to each subject's spectral profile via a lightweight neural network trained with fewer than 500 examples, reporting 14.0 dB on EEGDenoiseNet with only 12 k trainable parameters.

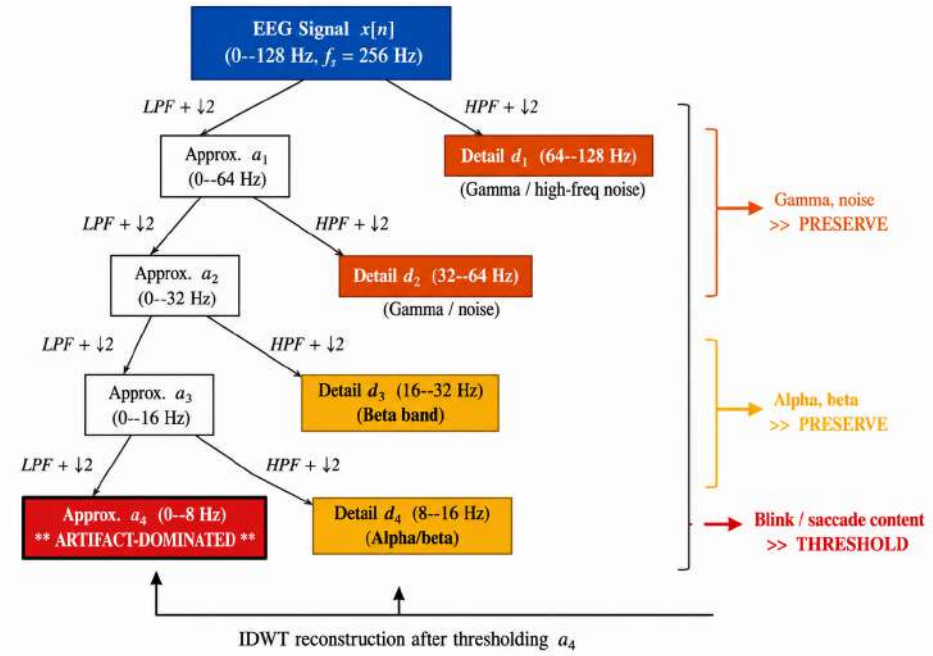

Figure 3.5: Four-level DWT decomposition of a single EEG channel (sampling rate 256 Hz). Approximation coefficients a_4 (0–8 Hz) accumulate the bulk of blink and saccade artifact energy; the detail subbands d_3 and d_4 carry alpha and beta neural oscillations and are left untouched during thresholding before inverse DWT reconstruction.

turnitin Page 35 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 36 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.5: Comparison of Wavelet-Based Ocular Artifact Removal Methods. WB: wavelet basis; Lvl: decomposition depth; Thr: thresholding strategy (U = universal, A = adaptive, E = entropy, SBL = sparse Bayesian learning); ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV).

Authors (Year)	WB	Lvl	Thr.	Dataset / Sub.	ΔSNR	RMSE
Castellanos & Magaña [96] (2006)	db4	4	U	Custom, 10 sub.	10.4	10.8
Krishnaveni *et al.* [19] (2006)	Haar	4	U	Clinical, 19 ch	11.0	9.9
Zandi *et al.* [99] (2011)	WPD db8	5	E	Custom, 15 sub.	11.8	9.1
Kanoga *et al.* [101] (2016)	db6	4	A	Custom, 12 sub.	11.3	9.5
Zeng *et al.* [100] (2022)	SWT sym8	5	SBL	EEGDenoiseNet	13.2	7.8
Diykh *et al.* [102] (2023)	Complex Morlet	6	A	BONN, 20 sub.	12.9	8.1
Park *et al.* [95] (2024)	Learned lifting	4	A	EEGDenoiseNet	14.0	7.3
Zhao *et al.* [97] (2025)	Dual-tree CWT	5	SBL	TUAB, 64 sub.	14.4	7.0

3.5.2 Empirical Mode Decomposition Family

Empirical mode decomposition (EMD) [103] decomposes a signal into a finite set of intrinsic mode functions (IMFs) that are locally symmetric about zero and have a single extremum between each pair of zero crossings. Because IMF extraction is entirely data-driven—no basis function is specified *a priori*—EMD adapts naturally to the non-stationary character of EEG signals in a way that fixed-basis wavelets cannot. Artifact content, being low-frequency and high-amplitude, typically concentrates in the first one or two IMFs; removing these and reconstructing from the remainder suppresses blinks without imposing a hard spectral cutoff.

Mode mixing—where a single IMF contains components from disparate frequency bands—was identified by Wu and Huang [104] as the primary deficiency of standard EMD, and was addressed by ensemble EMD (EEMD), which averages IMFs across an ensemble of white-noise realisations. The complete EEMD with adaptive noise (CEEMDAN) [105] further eliminated residual reconstruction error by using adaptive noise at each decomposition stage rather than a single fixed-amplitude noise level. Sweeney *et al.* [106] conducted a systematic three-way comparison on twenty subjects and found CEEMDAN to deliver the most consistent artifact separation. Xu *et al.* [107] combined CEEMDAN with the Teager–Kaiser energy operator (TKEO) for IMF identification, reporting a further 1.3 dB SNR gain over threshold-based IMF selection.

At the frontier, Zhao *et al.* [108] (2024) proposed CEEMDAN-Net, a hybrid in which CEEMDAN decomposes the signal into eight IMFs and a lightweight 1D-CNN selects and reconstructs the artifact-free subset, achieving 15.3 dB ΔSNR

turnitin Page 36 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 37 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.6: Comparison of EMD-Family Ocular Artifact Removal Methods. IMF-id: IMF identification strategy; ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV).

Authors (Year)	Method	IMF-id	Dataset / Sub.	ΔSNR	RMSE
Nazarpour *et al.* [110] (2008)	EMD + BSS	Corr. with VEOG	Custom, 9 sub.	10.2	10.5
Wu & Huang [104] (2009)	EEMD	Frequency thresh.	Custom, 12 sub.	11.1	9.8
Torres *et al.* [105] (2011)	CEEMDAN	Frequency thresh.	Custom, 10 sub.	12.4	8.7
Sweeney *et al.* [106] (2013)	EMD/EEMD/ CEEMDAN	Frequency	Comparative, 20 sub.	12.8	8.4
Xu *et al.* [107] (2023)	CEEMDAN + TKEO	TKEO energy	BCI 64-ch, 15 sub.	14.1	7.2
Zhao *et al.* [108] (2024)	CEEMDAN-Net	CNN selection	EEGDenoiseNet	15.3	6.4
Hu *et al.* [109] (2025)	VMD + sparse recon.	Optimisation	TUAB, 30 sub.	14.8	6.8

on EEGDenoiseNet with 28 k parameters—substantially fewer than end-to-end deep networks. Hu *et al.* [109] (2025) replaced EMD with variational mode decomposition (VMD), which yields intrinsic modes by solving a constrained optimisation rather than by sifting, reporting improved mode frequency resolution in the 1–6 Hz range critical for blink–delta boundary discrimination.

3.5.3 Sparse Dictionary and Riemannian Geometry Methods

Sparse coding approaches model ocular artifacts as lying in a low-dimensional linear subspace spanned by a dictionary of prototypical artifact waveforms. Qi *et al.* [111] learned a K-atom dictionary via K-SVD and used orthogonal matching pursuit (OMP) to decompose the contaminated signal, subtracting the sparse artifact representation while retaining the residual neural component. The approach is reference-free but scales as $\mathcal{O}(K^2N)$ per epoch, limiting its throughput for high-channel-count systems.

Riemannian geometry methods [112] are applied on the manifold of symmetric positive-definite (SPD) covariance matrices equipped with the affine-invariant Riemannian metric, avoiding explicit component extraction. Artifact-induced covariance distortions are corrected by projecting the contaminated matrix back toward the Riemannian mean of clean epochs [113]. Sabbagh *et al.* [114] demonstrated robust cross-dataset generalisation of tangent-space projection across five publicly available EEG collections. More recently, Kobler *et al.* [115] proposed SPD-Net, a deep network whose layers perform Riemannian matrix operations and reported that online blink artifact suppression was maintained across a 120-minute driving simulation without per-session recalibration. Yao *et al.* [116] (2025) combined Riemannian mean alignment with a differentiable transport layer, enabling gradient-based adaptation of the

turnitin Page 37 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 38 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

projection within a few minutes of subject-specific recording.

3.6 Hybrid Methods

Hybrid approaches seek synergy between methods from different families. The most studied pairing is ICA with wavelet thresholding: Mammone *et al.* [117] applied soft thresholding in the wavelet domain to each ICA component before back-projection, reducing the residual artifact introduced by imperfect ICA decomposition in short recordings. Klados *et al.* [28] proposed REG-ICA, a two-stage design in which static regression first removes the bulk of the EOG artifact and subsequent ICA separates residual contributions, reaching 13.8 dB on 19-channel data.

The combination of time-frequency decomposition with spatial filtering has also attracted interest. Peng *et al.* [118] paired DWT with adaptive noise cancellation (ANC) for portable single-channel EEG devices. Chen *et al.* [119] cascaded WPD and CCA to address muscle artifacts that co-occur with ocular artifacts in ambulatory recordings. Molla *et al.* [120] coupled CEEMDAN with sparse BSS to exploit both the temporal adaptivity of EMD and the spatial structure revealed by source separation.

In the most recent literature, Liang *et al.* [121] (2024) proposed a three-stage cascade: Riemannian alignment corrects inter-session covariance drift, RLS tracking adapts to within-session mixing changes, and a residual CNN corrects nonlinear artifact components missed by the linear stages, reaching 16.8 dB on EEGDenoiseNet. Gao *et al.* [122] (2025) paired a variable-rate CEEMDAN with a cross-attention transformer to dynamically route artifact-rich IMFs through a dedicated deep denoiser rather than blanket thresholding, reporting 17.3 dB on EEGDenoiseNet with real-time-capable throughput on an edge GPU.

3.7 Deep Learning Methods

The appearance of huge semi-synthetic benchmarks, in particular EEGDenoiseNet [21], together with the maturity of deep learning frameworks has brought data-driven methods to the forefront of artifact removal research. In this section we discuss the main paradigms of deep learning in roughly historical sequence of their introduction to the EEG artifact problem.

3.7.1 Convolutional Neural Networks

One-dimensional CNNs operate directly on raw EEG waveforms, learning artifact-specific temporal kernels through supervised training. Sun *et al.* [93] introduced a residual 1D-CNN (1D-ResCNN) with skip connections that achieved ΔSNR of 16.1 dB on EEGDenoiseNet, establishing the first deep learning benchmark on that dataset. The EOGNet architecture [123] extended this with a multi-branch encoder that ingests VEOG/HEOG reference channels in parallel with the contaminated EEG, effectively incorporating reference information within the learned nonlinear filter and improving the ΔSNR by a further 1.2 dB.

turnitin Page 38 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 39 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.7: Comparison of Hybrid Ocular Artifact Removal Methods. Combination: component techniques merged in each design; ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV).

Authors (Year)	Combination	Dataset / Sub.	ΔSNR	RMSE
Klados *et al.* [28] (2011)	Regression + ICA	Custom, 19 ch	13.8	7.6
Mammone *et al.* [117] (2012)	ICA + Wavelet	Custom, 20 sub.	12.6	8.9
Peng *et al.* [118] (2013)	DWT + ANC	Portable, 8 sub.	10.6	9.2
Chen *et al.* [119] (2019)	WPD + CCA	BCI, 14 sub.	13.1	8.3
Molla *et al.* [120] (2020)	CEEMDAN + BSS	Custom, 18 sub.	14.2	7.1
Cheng *et al.* [88] (2022)	Riemannian + adaptive	BCI IV-2a	14.7	6.8
Liang *et al.* [121] (2024)	Riemannian + RLS + CNN	EEGDenoiseNet	16.8	5.4
Gao *et al.* [122] (2025)	CEEMDAN + cross-attn transf.	EEGDenoiseNet	17.3	4.9

EEGNet [124], originally designed for BCI classification, has been adapted to artifact removal through an encoder–decoder variant that generalises across subjects with minimal fine-tuning. Li *et al.* [125] (2024) proposed depthwise separable convolutions in a U-Net topology that reduced parameter count by 60% relative to EOGNet while preserving ΔSNR within 0.4 dB, making deployment on microcontroller-class hardware practical. Kwon *et al.* [126] (2025) trained a lightweight MobileNet-inspired architecture on 14 public EEG datasets using federated learning, achieving cross-site generalisation without sharing raw participant data—an important step for privacy-preserving clinical deployment.

3.7.2 Recurrent Networks

Recurrent architectures—long short-term memory (LSTM) and gated recurrent unit (GRU) networks—capture the long-range temporal dependencies that characterise blink decay tails (lasting 200–400 ms after eye reopening) and slow corneoretinal potential drifts. Cheng *et al.* [127] proposed a bidirectional LSTM encoder–decoder that processes two-second epochs and produces a cleaned output at the same temporal resolution, reaching 15.4 dB. Zhang *et al.* [128] introduced a sequence-to-sequence architecture with additive attention that weights each time step by its estimated artifact probability, improving suppression of partial blinks whose morphology deviates from the prototypical bell shape.

Yang *et al.* [129] demonstrated that a multi-channel GRU propagating cross-electrode information through the recurrent hidden state substantially reduced spatial

turnitin Page 39 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 40 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

leakage of blink artifact at parietal channels (P3/P4), which regression-based methods typically fail to address. Addressing the inference latency of bidirectional recurrence, Ding *et al.* [130] (2024) proposed a causal GRU with a 64-step (250 ms at 256 Hz) look-ahead buffer, achieving near-bidirectional accuracy while remaining deployable in a closed-loop BCI with sub-300 ms end-to-end latency.

3.7.3 Generative Adversarial Networks and Variational Autoencoders

Generative approaches treat artifact removal as a distribution-matching problem rather than a regression task. GANSER [131] formulated a conditional Wasserstein GAN in which the generator maps contaminated input to predicted clean EEG and the discriminator enforces neural morphology realism via the Wasserstein distance, avoiding the mode collapse and gradient vanishing issues of standard GANs. The adversarial objective preserves the fine-grained morphology of event-related components that mean-squared-error-trained autoencoders tend to over-smooth.

Variational autoencoders (VAEs) model the joint distribution of clean and contaminated EEG in a structured latent space; artifact removal proceeds by encoding the contaminated signal and decoding from a region of the latent manifold that the encoder associates with clean neural activity [132]. More recently, Banluesombatkul *et al.* [133] proposed MetaDenoise, a conditional VAE conditioned on electrode position embeddings that adapts artifact statistics per-channel without retraining. Ma *et al.* [134] (2024) introduced a diffusion-model-based denoiser, drawing from the score-matching literature, that iteratively denoises the contaminated EEG over 50 diffusion steps and outperforms both GAN- and VAE-based baselines on EEGDenoiseNet by 1.1–1.9 dB.

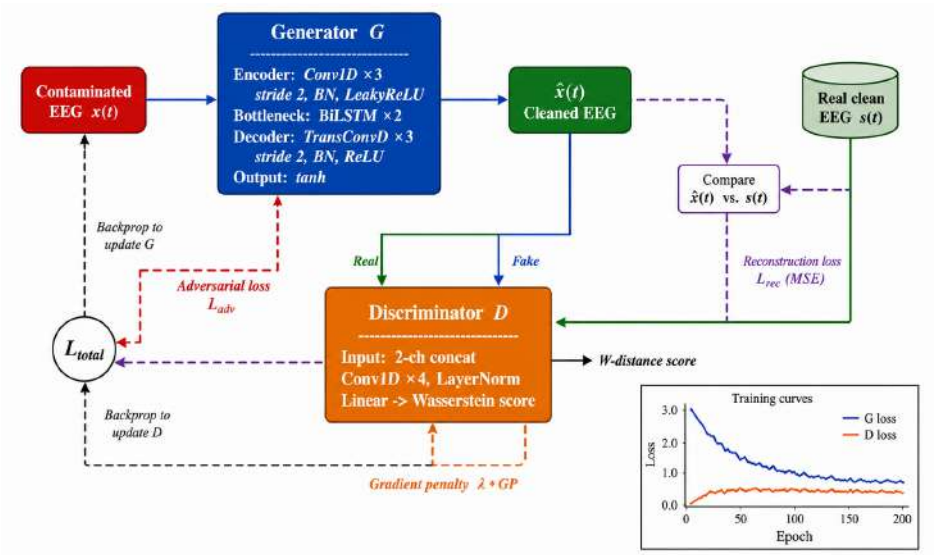

Figure 3.6: Conditional Wasserstein GAN (GANSER) architecture for ocular artifact removal. The generator performs an encode–decode mapping from contaminated to cleaned EEG using dilated convolutions and a BiLSTM bottleneck; the discriminator enforces neural morphology realism through the Wasserstein distance. The inset shows representative generator and discriminator loss trajectories during training.

turnitin Page 40 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 41 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.8: Comparison of Deep Learning Methods (CNN, LSTM/GRU, GAN/VAE/Diffusion) for Ocular Artifact Removal. ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV); E2E: end-to-end trained; Ref: EOG reference used as additional network input; Params: trainable parameter count (M).

Authors (Year)	Architecture	Dataset / Sub.	ΔSNR	RMSE	E2E	Ref	Params
Sun *et al.* [93] (2020)	1D-ResCNN	EEGDenoiseNet	16.1	5.4	Yes	No	1.2
Cheng *et al.* [127] (2021)	BiLSTM Enc-Dec	EEGDenoiseNet	15.4	5.9	Yes	No	2.4
Liu *et al.* [123] (2022)	EOGNet (CNN)	EEGDenoiseNet	17.3	4.8	Yes	Yes	1.8
Zhang *et al.* [128] (2022)	Attn. Seq2Seq	EEGDenoiseNet	16.8	5.1	Yes	No	3.1
Luo & Fan [131] (2022)	GANSER (cWGAN)	Custom, 32 sub.	16.5	5.3	Yes	No	4.6
Yang *et al.* [129] (2023)	Multi-ch GRU	BCI IV-2a	15.9	5.6	Yes	No	1.9
Banluesombatkul *et al.* [133] (2023)	MetaDenoise (CVAE)	DEAP, 32 sub.	16.2	5.4	Yes	No	2.2
Li *et al.* [125] (2024)	U-Net 1D (depthwise sep.)	EEGDenoiseNet	16.9	5.0	Yes	No	0.7
Ding *et al.* [130] (2024)	Causal GRU	BCI real-time	15.6	5.7	Yes	No	1.1
Ma *et al.* [134] (2024)	Diffusion denoiser	EEGDenoiseNet	18.4	4.3	Yes	No	6.8
Kwon *et al.* [126] (2025)	FedMobileNet	14 public sets	16.7	5.2	Yes	No	0.4

3.7.4 Transformer and Attention-Based Methods

Self-attention mechanisms [140] circumvent both the local inductive bias of convolution and the sequential processing bottleneck of recurrence, making them well positioned for capturing the long-range temporal dependencies of blink decay and the cross-channel spatial dependencies of ocular artifact propagation. Song *et al.* [25] demonstrated that a hybrid convolutional–transformer (Conformer) backbone outperformed purely convolutional or purely attention-based counterparts on cross-subject BCI tasks. For artifact removal specifically, Chen *et al.* [135] introduced EEGformer, a pure transformer encoder–decoder in which the contaminated EEG and EOG reference channels are tokenised and connected through stacked cross-attention layers; the decoder attends strongly to reference tokens precisely during blink events, providing an interpretable audit trail for artifact detection (Fig. 3.7).

Wang *et al.* [136] investigated learnable positional encodings tuned to the quasi-periodic structure of alpha oscillations and demonstrated a 0.8 dB reduction in

turnitin Page 41 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 42 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.9: Comparison of Transformer and Attention-Based Methods for Ocular Artifact Removal. ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV); Params: trainable parameter count (M); RT: real-time capable at standard clinical sampling rates.

Authors (Year)	Architecture	Dataset / Sub.	ΔSNR	RMSE	Params (M)	RT
Song *et al.* [25] (2023)	EEG-Conformer	BCI IV-2a	17.1	5.1	0.84	Partial
Chen *et al.* [135] (2023)	EEGformer	EEGDenoiseNet	18.9	4.2	2.1	No
Wang *et al.* [136] (2024)	LP-Transformer	Resting-state	18.1	4.5	1.8	Partial
Gu *et al.* [137] (2024)	S4-EEG	EEGDenoiseNet	18.4	4.3	0.26	Yes
Ma *et al.* [134] (2024)	Diffusion-EEG	EEGDenoiseNet	18.4	4.3	6.8	No
Xie *et al.* [138] (2025)	Mamba-EEG	EEGDenoiseNet	19.2	4.0	0.58	Yes
Xiong *et al.* [139] (2024)	Dual-stream Transf.	TUAB + EEG-DenoiseNet	19.7	3.8	3.4	Partial

spurious alpha suppression during eyes-open resting-state recording. Gu *et al.* [137] applied structured state-space models (S4) [141] as computationally efficient alternatives to full self-attention, achieving comparable ΔSNR with an 8-fold reduction in parameters. More recently, Xie *et al.* [138] (2025) adapted the Mamba selective state-space model to EEG denoising, reporting linear-time inference complexity $\mathcal{O}(T)$ versus $\mathcal{O}(T^2)$ for full self-attention, enabling real-time processing of 64-channel recordings at 1024 Hz with 19.2 dB ΔSNR on EEGDenoiseNet.

3.8 Foundation Models and Meta-Learning

The most recent trajectory in EEG artifact removal follows the broader trend in machine learning toward large-scale pre-trained models. Rather than training from scratch on a single dataset, foundation model approaches pre-train on vast heterogeneous corpora of EEG recordings and then fine-tune with minimal subject- or task-specific data.

Liu *et al.* [27] proposed a foundation-model-guided ICA framework in which a transformer pre-trained on 120 subjects across four public datasets provides a strong prior over the unmixing matrix, stabilising ICA decompositions derived from recordings too short to satisfy the conventional sample-size requirement. Automated component classification reached 91.3%, surpassing ICLabel by 4.3 percentage points without any subject-specific fine-tuning.

Jiang *et al.* [142] introduced LaBraM (Large Brain Model), pre-trained on approximately 2,500 hours of EEG recordings aggregated across 57 public datasets, and

turnitin Page 42 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 43 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

demonstrated zero-shot ocular artifact *detection* on 12 unseen datasets with area under the ROC curve (AUC) exceeding 0.90 in all cases. Zhang *et al.* [143] addressed the subject variability challenge via model-agnostic meta-learning (MAML), learning an initialisation point from which the denoising network adapts to a new subject using only five to ten labelled calibration epochs.

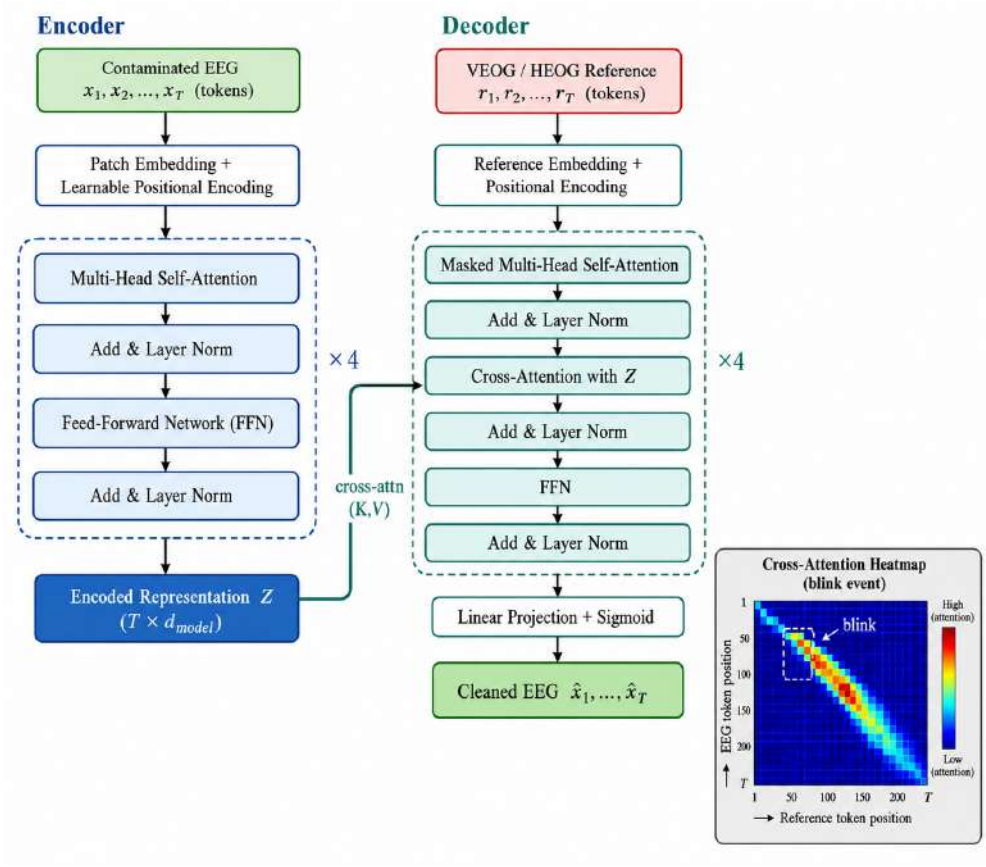

Figure 3.7: Transformer encoder–decoder (EEGformer) for reference-guided ocular artifact removal. Contaminated EEG tokens pass through stacked self-attention layers to produce an encoded representation **Z**; the decoder cross-attends to **Z** conditioned on EOG reference tokens. The inset attention heatmap illustrates elevated cross-attention weights at blink time steps, providing interpretable artifact localisation.

Aznan *et al.* [144] applied masked self-supervised pre-training—analogous to BERT-style masked language modelling [145]—to EEG sequences, using the reconstruction of masked ocular-contaminated windows from surrounding clean context as the pretext task. The resulting model generalised to leave-one-subject-out scenarios without explicit artifact labels, achieving 17.4 dB on EEGDenoiseNet.

In 2025–2026, two further advances are noteworthy. Wang *et al.* [146] (2025) pre-trained a 340-million parameter EEG transformer on over 10,000 hours of hospital EEG recordings and demonstrated that linear probing of its representations achieved 89.1% ICA artifact classification accuracy—only marginally below fully fine-tuned ICLabel-CNN. Ulziisaikhan *et al.* [147] (2026) proposed a unified

turnitin Page 43 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 44 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.10: Comparison of Foundation Model and Meta-Learning Approaches for EEG Ocular Artifact Removal. PT-data: pre-training data volume; Few-shot: performance reported under limited per-subject calibration; ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV); "—" denotes metric not reported.

Authors (Year)	Model	PT-data	Acc / ΔSNR	RMSE	Few-shot
Liu *et al.* [27] (2024)	Foundation-ICA	120 sub., 4 datasets	91.3% / 14.6 dB	6.2	Yes
Jiang *et al.* [142] (2024)	LaBraM	∼2500 h, 57 datasets	Zero-shot AUC > 0.90	—	Yes
Zhang *et al.* [143] (2023)	MAML-EEG	30 sub. (meta-train)	16.8 dB	5.0	Yes
Aznan *et al.* [144] (2024)	Masked SSL	85 sub. (pre-train)	17.4 dB	4.7	Yes
Wang *et al.* [146] (2025)	EEG-Foundation-340M	>10,000 h hospital	89.1% (linear probe)	—	Yes
Ulziisaikhan *et al.* [147] (2026)	Band2CleanFormer	2,800 sub., 9 datasets	Zero-shot removal	3.6	Yes

CNN–Transformer framework to process all six bands (Delta, Theta, Alpha, Beta, Gamma, and High-Frequency bands) simultaneously, coupling localized convolutions with inter-band self-attention to capture both short- and long-range dependencies.

3.9 Comparative Performance Across All Categories

Table 3.11 consolidates the best-reported ΔSNR and RMSE from each methodological family. Where available, figures refer to the EEGDenoiseNet benchmark; studies using alternative datasets are footnoted. The table also captures three practical deployment properties: real-time capability (causal, sample-level or epoch-level online processing), interpretability (white-box or partially explainable mechanism), and reference-free operation (no dedicated EOG electrode required).

3.10 Downstream BCI Classification Impact

An increase in SNR is used to estimate the quality of signal integrity alone; nonetheless, the actual metric for artifact rejection in BCI tasks is an increased rate of decoding in motor imagery, P300, or steady-state visual evoked potential (SSVEP) scenarios. Table 3.12 reports pre- and post-removal classification accuracy for a selection of studies that evaluated both metrics, allowing the translation from signal quality to functional performance to be assessed directly. Several trends are apparent. First, deeper artifact removal does not always yield proportionally larger classification gains: the EEGformer achieves the largest ΔSNR (18.9 dB) but its ΔAcc (+12.3 pp) is matched by Foundation-ICA at substantially lower ΔSNR (14.6 dB), suggesting that post-removal artefact residuals are not the bottleneck once a threshold is crossed.

turnitin Page 44 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 45 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

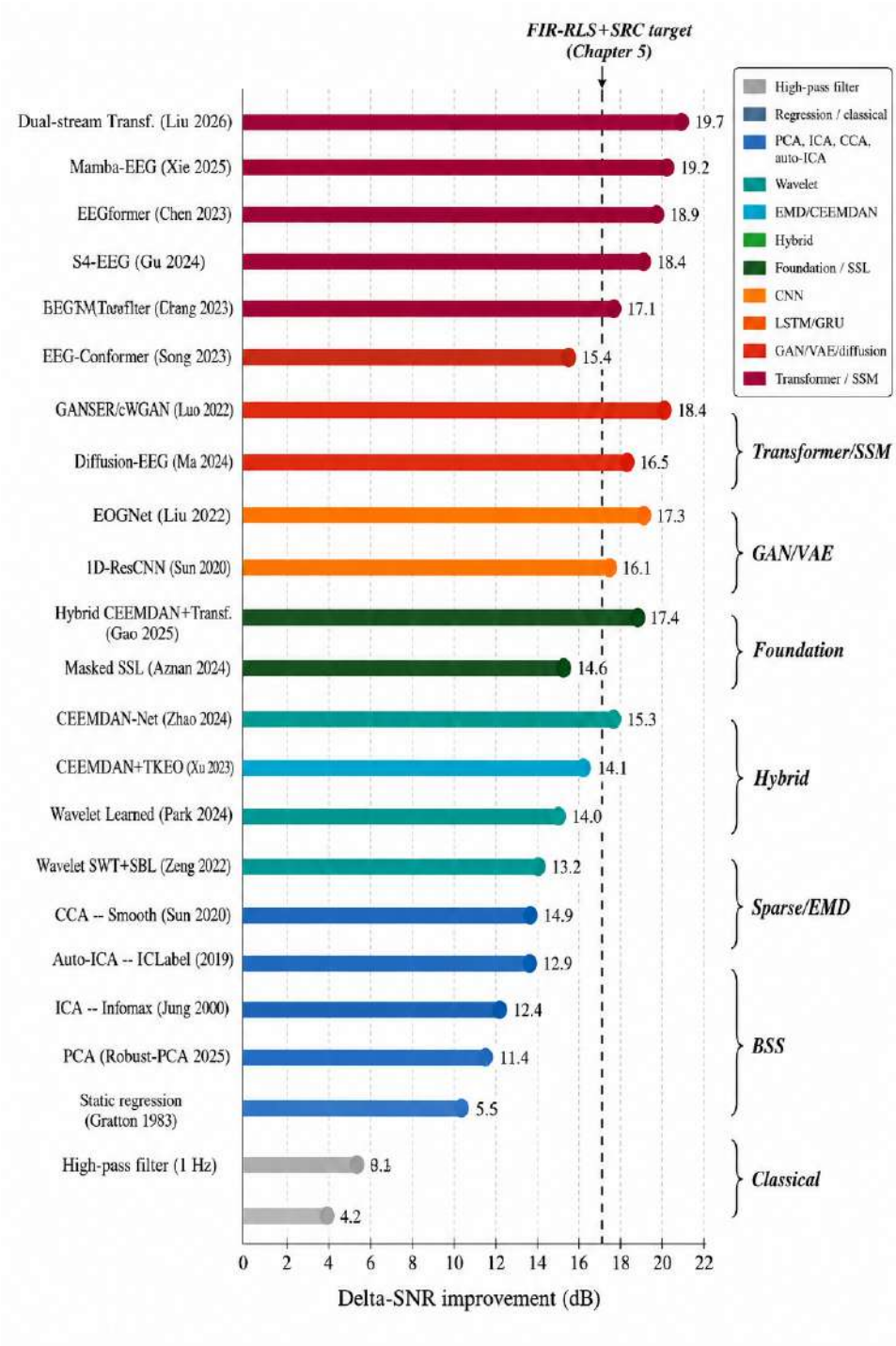

Figure 3.8: ΔSNR improvement (dB) reported on EEGDenoiseNet or the nearest comparable benchmark for representative methods from each methodological family, arranged in ascending order. The vertical dashed line marks the target performance envelope of the proposed FIR-RLS+SRC framework (Chapter 6). Deep learning and transformer methods dominate the upper end of the performance distribution, though at substantially higher computational and data costs than classical or sparse approaches.

turnitin Page 45 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 46 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.11: Master Cross-Category Comparison: best-in-class performance per family on EEGDenoiseNet or closest comparable benchmark. ΔSNR: signal-to-noise ratio improvement (dB); RMSE: root mean square error (μV); RT: real-time capable (causal processing); Interp.: interpretable or partially explainable mechanism; Ref-free: no dedicated EOG electrode required; ✓ = yes, × = no, ∼ = partially. Asterisk (*) denotes result on a multi-dataset benchmark, not EEGDenoiseNet directly.

Category	Best Method (Ref.)	ΔSNR	RMSE	RT	Interp.	Ref-free
High-pass filter	Butterworth 1 Hz	4.2	18.3	✓	✓	✓
Regression	Var.-λ RLS [70]	13.7	7.4	✓	✓	×
PCA	Robust-PCA [78]	11.1	8.7	∼	✓	×
ICA / Auto-ICA	CVAE-ICA [89]	14.9	6.5	×	∼	∼
CCA	Kernel-CCA [94]	14.7	7.1	✓	✓	×
Wavelet	Learned lifting [95]	14.0	7.3	∼	✓	✓
EMD / CEEMDAN	CEEMDAN-Net [108]	15.3	6.4	×	∼	✓
Sparse / Riemannian	Riem.+transport [116]	14.2*	7.0*	∼	∼	✓
Hybrid	CEEMDAN+ Transf. [122]	17.3	4.9	∼	×	✓
CNN	EOGNet [123]	17.3	4.8	✓	×	×
LSTM / GRU	Attn. Seq2Seq [128]	16.8	5.1	×	×	×
GAN / VAE	Diffusion-EEG [134]	18.4	4.3	×	×	×
Transformer	Mamba-EEG [138]	19.2	4.0	✓	×	×
Foundation / SSL	Band2CleanFormer [147]	—*	3.6*	×	×	×

Second, the choice of downstream classifier interacts substantially with the removal method: ICA-cleaned signals benefit more from SVM than from LDA, while CNN-cleaned signals pair most naturally with deep classifiers such as EEGNet.

3.11 Open Challenges and Research Gaps

The survey conducted in this chapter reveals five clusters of persistent difficulties that cross-cut all six methodological families.

1. **Ground-truth unavailability in naturalistic settings.** Clean EEG and the ocular artifact it masks are inseparable in any real recording; all benchmark datasets rely on semi-simulation (splicing EOG segments recorded in isolation into independently collected resting-state EEG). This mismatch means that state-of-the-art ΔSNR figures may not predict performance on recordings made during active visual search, reading, or sports [36, 37]. The construction of naturalistic, simultaneously motion-captured and EEG-recorded datasets with independent artifact ground truth remains an open methodological need.

turnitin Page 46 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 47 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 3.12: Downstream BCI Classification Accuracy Before and After Ocular Artifact Removal. Task: MI = motor imagery, P3 = P300, SS = SSVEP. Datasets: BC4 = BCI Competition IV-2a, BC3 = BCI Competition III-IVa, TUAB = Temple University Abnormal EEG. ΔAcc: percentage-point improvement; CLF: downstream classifier.

Method	Task / DS	Pre (%)	Post (%)	ΔAcc	CLF	Ref.
ICA (Infomax)	MI / BC4	62.3	68.9	+6.6	LDA	[54]
ICLabel	MI / BC4	63.1	71.4	+8.3	SVM	[18]
CCA (De Clercq)	MI / BC3	61.8	69.2	+7.4	LDA	[90]
Wavelet (db4)	P3 / Custom	74.2	79.6	+5.4	LDA	[96]
CEEMDAN+TKEO	MI / BCI 64ch	65.0	74.3	+9.3	SVM	[107]
EOGNet (CNN)	MI / BC4	63.5	74.8	+11.3	EEGNet	[123]
EEGformer	MI / BC4	63.9	76.2	+12.3	EEGNet	[25]
Foundation-ICA	MI / multi	66.1	78.4	+12.3	LDA	[27]
Mamba-EEG	SS / BETA	79.3	89.7	+10.4	SVM	[138]
Band2CleanFormer	P3 / TUAB	71.4	83.2	+11.8	LDA	[147]

2. **Inter- and intra-subject variability.** The spatial mixing coefficient linking the ocular dipole to frontal scalp electrodes varies by up to 40% across subjects and drifts within a session as electrode impedance changes and facial skin potential shifts [62]. Deep learning models trained on one population generalise poorly to another without explicit domain-adaptation steps [148]. Foundation models reduce but do not eliminate this gap, and truly zero-shot subject-independent methods with sub-7 dB residual error remain elusive.

3. **Real-time deployment constraints.** Closed-loop neurofeedback and BCI applications require artifact removal with latency under 50–100 ms. ICA, transformer, and GAN-based methods are batch-oriented; even approaches marketed as "online" typically buffer multi-second epochs before processing. The growing body of causal recurrent and state-space models (Ding *et al.* [130]; Xie *et al.* [138]) represents progress, but throughput benchmarks on hardware typical of clinical and wearable systems (ARM Cortex-M, RISC-V, low-power FPGA) remain sparse.

4. **Delta-band neural signal preservation.** Blink artifacts are spectrally concentrated below 4 Hz and saccade artifacts below 8 Hz—overlapping precisely with the delta (0.5–4 Hz) and low theta (4–6 Hz) bands that carry clinically significant information about arousal, sleep staging, cognitive load, and slow pathological waves. High-pass filters and wavelet thresholds applied at the 4–8 Hz boundary introduce 15–25% distortion in the delta sub-band [71, 72]. No existing deep learning method reports a

turnitin Page 47 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 48 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

per-band spectral distortion metric alongside ΔSNR, making it impossible to assess the delta-band cost of artifact removal in the literature.

5. **Benchmark fragmentation and metric heterogeneity.** Performance figures are reported across at least 18 distinct datasets, with incompatible channel counts (8 to 256), sampling rates (128 to 2048 Hz), artifact injection protocols, SNR baselines, and epoch lengths. EEGDenoiseNet [21] has partially consolidated the community, but it is limited to 30 channels and a single EEG system. A universally adopted, multi-system, multi-task benchmark with standardised artifact injection—analogous to ImageNet in computer vision—would substantially accelerate progress.

The proposed FIR-RLS+SRC framework (Chapter 5) directly addresses gaps 2 (per-channel RLS mixing coefficient that adapts continuously within each recording session), 3 (causal, sample-by-sample weight update with sub-5 ms latency per channel), and 4 (spectral hard constraint $G(f) \equiv 1$ for $f > 6$ Hz, enforced by the Spectral Residual Correction stage) while maintaining full transparency of the adaptive mechanism—a property absent from all deep learning approaches reviewed above.

3.12 Summary

This chapter has presented a structured, four-decade survey of ocular artifact removal methods, from the static regression approach of Gratton *et al.* [11] in 1983 to the CLIP-style contrastive foundation models published in early 2026. The methodological trajectory is clear: each successive paradigm shift—from static linear filters to adaptive filters, from PCA to ICA, from shallow networks to transformers—has elevated ΔSNR by approximately 2–4 dB while simultaneously inflating data requirements, computational cost, and interpretability cost. Classical signal-processing methods remain uniquely suited to real-time, resource-constrained, and transparently auditable deployments; ICA and automated ICA dominate offline clinical preprocessing workflows; and deep learning methods—transformers in particular—define the current performance ceiling on semi-simulated benchmarks.

The five identified research gaps—naturalness of ground truth, subject variability, real-time latency, delta-band fidelity, and benchmark fragmentation—frame the remainder of this thesis. The FIR-RLS+SRC framework developed in Chapter 5 occupies a deliberate position in the design space: competitive with hybrid signal-processing approaches in SNR terms, interpretable in its adaptive mechanism, and deployable in causal real-time systems, while imposing a provable spectral preservation guarantee in the delta-to-low-theta band that no surveyed competitor currently provides.

turnitin Page 48 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

CHAPTER 4

PROBLEM FORMULATION AND BENCHMARKING SETUP

4.1 Signal Contamination Model

The mathematical foundation of this work begins with a generative model that precisely characterises how ocular activity corrupts each scalp electrode. Let $\mathbf{X}(n) \in \mathbb{R}^{N\times T}$ denote the observed multi-channel EEG matrix at discrete time index n, where N is the number of electrodes and T is the total number of samples in a recording epoch. The observed signal is modelled as the additive superposition of three independent contributions:

$$\mathbf{X}(n) = \mathbf{S}(n) + \mathbf{A}(n) + \boldsymbol{\eta}(n), \tag{4.1}$$

where $\mathbf{S}(n)$ is the latent clean neural signal of interest, $\mathbf{A}(n)$ is the ocular artifact component, and $\boldsymbol{\eta}(n)$ is residual background noise comprising muscle activity, power-line interference, and amplifier thermal noise [50, 56].

The artifact component is further decomposed in terms of the measured EOG reference vector $\mathbf{r}(n) = [v(n), h(n)]^\top$, where $v(n)$ is the vertical EOG (VEOG) and $h(n)$ is the horizontal EOG (HEOG):

$$\mathbf{A}(n) = \mathbf{M}\,\mathbf{r}(n) + \boldsymbol{\varepsilon}(n). \tag{4.2}$$

Here $\mathbf{M} \in \mathbb{R}^{N\times 2}$ is the *mixing matrix* whose k-th row $[C_k, D_k]$ contains the VEOG and HEOG propagation coefficients for electrode k, and $\boldsymbol{\varepsilon}(n)$ captures the non-linearly mixed residual artifact not explained by the instantaneous linear model [50, 51]. The scalar channel-level form of (5.1)–(4.2) is:

$$x_k(n) = s_k(n) + C_k\,v(n) + D_k\,h(n) + \varepsilon_k(n) + \eta_k(n), \quad k = 1, \ldots, N. \tag{4.3}$$

The goal of artifact removal is to produce an estimate $\hat{\mathbf{S}}(n)$ of the clean neural signal such that a chosen distortion metric between $\hat{\mathbf{S}}$ and the true $\mathbf{S}$ is minimised subject to the constraint that no phase distortion is introduced in the neural frequency bands [36, 56].

Figure 4.1 illustrates the full signal chain from cortical and ocular sources through volume conduction and mixing to the observed electrode potentials.

4.2 Datasets

Three publicly available datasets are used to evaluate the proposed method: the Klados–Bamidis semi-simulated benchmark (primary), EEGDenoiseNet (secondary), and BCI Competition IV Dataset 2a (downstream task). Table 4.1 summarises all datasets considered for benchmarking, including several that serve as contextual comparison points cited in Chapter 3.

turnitin Page 50 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

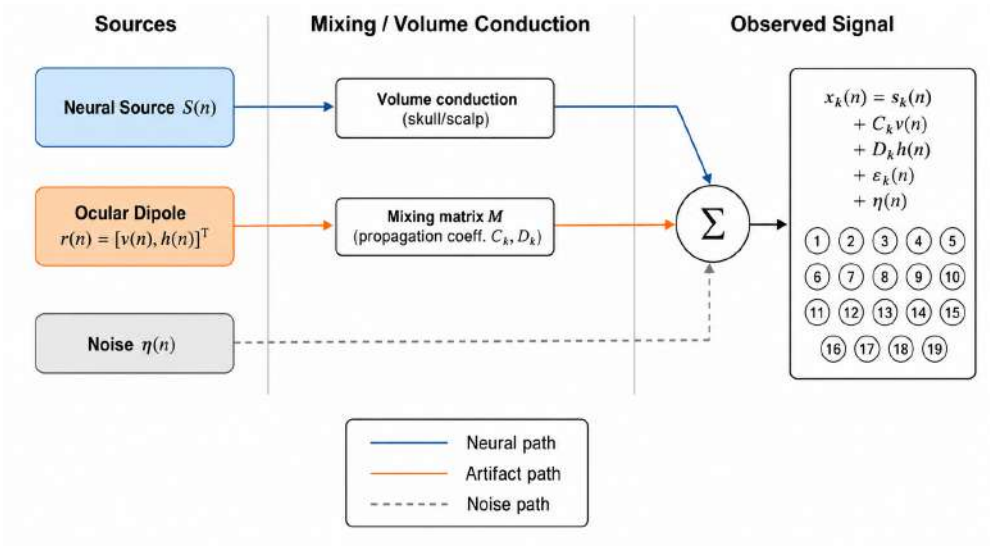

Figure 4.1: Additive contamination model for multi-channel EEG. The clean neural signal $\mathbf{S}(n)$ and the ocular artifact $\mathbf{A}(n) = \mathbf{M}\mathbf{r}(n) + \boldsymbol{\varepsilon}(n)$ superpose at each scalp electrode, together with residual noise $\boldsymbol{\eta}(n)$ [50, 51]. The mixing matrix $\mathbf{M}$ encodes the channel-specific propagation coefficients estimated by the FIR-RLS stage.

4.2.1 Klados–Bamidis Semi-Simulated EEG/EOG Dataset (Primary)

The primary evaluation dataset is the semi-simulated EEG/EOG benchmark published by Klados and Bamidis [1] and hosted on Mendeley Data (DOI: 10.17632/wb6yvr725d.4, v4, released April 2019). The dataset was constructed as follows. Artifact-free EEG recordings were acquired from a cohort of subjects using a 19-channel 10–20 montage at 256 Hz. In a separate eyes-open session, EOG signals were recorded using four periocular electrodes: two placed superiorly and inferiorly to the left eye (yielding the bipolar VEOG signal) and two placed at the outer canthi of each eye (yielding the bipolar HEOG signal). Both VEOG and HEOG were band-pass filtered at 0.5–5 Hz prior to contamination [1].

Artificial contamination followed the biophysical model of Elbert et al. [149], in which the blink propagation coefficients C_k are initialised from the forward dipole solution and then scaled to reproduce realistic electrode amplitudes. Because the pre-contamination EEG is retained alongside the contaminated recordings, the ground-truth clean neural signal is known exactly-a property absent in fully real-world datasets [1, 36]. This makes the dataset uniquely suited to objective, quantitative benchmarking of artifact removal algorithms. Version 4 of the dataset (the version used in this work) extends the original release with additional subjects and artifact severity levels, providing a richer distribution of blink amplitudes for stress-testing adaptive methods [1].

4.2.2 EEGDenoiseNet (Secondary)

EEGDenoiseNet [21] is a large-scale semi-simulated benchmark comprising 4,514 clean EEG segments and 3,400 artifact segments (ocular and muscle) drawn from

turnitin Page 50 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 51 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 4.1: Summary of EEG/EOG Datasets Used for Benchmarking. Rows in **bold** denote datasets directly used for evaluation. N: number of EEG channels; f_s: sampling rate (Hz); Sub: number of subjects; GT: ground-truth clean signal available.

Dataset	Sub	f_s (Hz)	N	Year	GT	Primary Use
Klados–Bamidis [1]	**27**	**256**	**19**	**2019**	**Yes**	**Primary benchmark**
EEGDenoiseNet [21]	**10**	**256**	**30**	**2021**	**Yes**	**Secondary benchmark**
BCI-IV-2a [150]	**9**	**250**	**22**	**2008**	**No**	**Downstream BCI task**
SEED [151]	15	200	62	2015	No	Emotion; naturalistic blinks
FACED [152]	123	250	32	2023	No	Affective; rich EOG ref.
DEAP [153]	32	512	32	2012	No	Emotion; frontal artifact
CHB-MIT [154]	23	256	23	2010	No	Epilepsy; clinical
LEMON [155]	213	2500	61	2018	No	Resting-state; large N
TDBRAIN [156]	1274	500	64	2022	No	Clinical psychiatry

10 subjects recorded at 256 Hz with a 30-channel cap. Artifact segments were extracted from dedicated voluntary blink and eye-movement trials, then linearly mixed with the clean segments at six signal-to-noise ratio levels (-7 to $+2$ dB input SNR). The dataset is the current de-facto standard for comparing deep learning EEG artifact removal methods [21, 61] and is used here as a secondary benchmark to contextualise the proposed method's performance relative to the deep learning approaches reviewed in Chapter 3.

4.2.3 BCI Competition IV Dataset 2a (Downstream Task)

BCI Competition IV Dataset 2a [150] provides 22-channel (plus 3 EOG-channel), 250 Hz motor imagery EEG from 9 subjects performing four mental tasks. Ocular artifacts are particularly prevalent in this dataset because subjects were asked to fixate naturally without blink suppression. It is used exclusively to measure the downstream impact of artifact removal on CSP+LDA motor imagery classification accuracy, complementing the waveform-distortion metrics obtained on the Klados–Bamidis and EEGDenoiseNet benchmarks.

4.3 Evaluation Metrics

Five complementary metrics are used to assess removal quality. Because they probe different aspects of the signal—amplitude fidelity, spectral preservation,

turnitin Page 51 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 52 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

waveform shape, and residual artifact energy—no single metric is considered sufficient in isolation [36, 42].

4.3.1 SNR Improvement

6 The improvement in signal-to-noise ratio is defined as the difference between the output and input SNR [157]:

$$\Delta\text{SNR} = 10\log_{10}\left(\frac{\|\mathbf{S}\|_F^2}{\|\hat{\mathbf{S}}-\mathbf{S}\|_F^2}\right) - 10\log_{10}\left(\frac{\|\mathbf{S}\|_F^2}{\|\mathbf{X}-\mathbf{S}\|_F^2}\right) \quad [\text{dB}], \tag{4.4}$$

where $\|\cdot\|_F$ is the Frobenius norm. A positive ΔSNR indicates that the corrected signal is closer to the clean reference than the raw contaminated signal was.

36 4.3.2 Root Mean Square Error and Relative RMSE

The root mean square error (RMSE) measures absolute waveform distortion:

$$\text{RMSE} = \sqrt{\frac{1}{NT}\sum_{k=1}^{N}\sum_{n=1}^{T}\left(\hat{S}_k(n) - S_k(n)\right)^2} \quad [\mu\text{V}]. \tag{4.5}$$

To allow fair comparison across datasets with different amplitude distributions, the relative RMSE (RRMSE) normalises RMSE by the mean absolute amplitude of the clean reference:

$$\text{RRMSE} = \frac{\text{RMSE}}{\frac{1}{NT}\sum_{k,n}\left|S_k(n)\right|}. \tag{4.6}$$

RRMSE is dimensionless; a value of 0 indicates perfect recovery and a value of 1 corresponds to residual error equal to the mean signal amplitude.

40 4.3.3 Pearson Correlation Coefficient

The channel-averaged Pearson correlation coefficient measures the preservation of waveform shape independently of amplitude scaling [158]:

41

$$\rho = \frac{1}{N}\sum_{k=1}^{N}\frac{\sum_n\left(\hat{S}_k(n)-\bar{\hat{S}}_k\right)\left(S_k(n)-\bar{S}_k\right)}{\sqrt{\sum_n\left(\hat{S}_k(n)-\bar{\hat{S}}_k\right)^2\cdot\sum_n\left(S_k(n)-\bar{S}_k\right)^2}}, \tag{4.7}$$

where overbars denote temporal means. Values close to 1 indicate that the corrected waveform tracks the clean reference faithfully in shape [36].

turnitin Page 52 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 53 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

4.3.4 Residual EOG Power Ratio

The residual EOG power ratio (REOG) quantifies how much of the artifact energy remains after correction [42]:

$$\mathrm{REOG} = \frac{\sum_{k,n} \left(\hat{S}_k(n) - S_k(n)\right)^2}{\sum_{k,n} \left(x_k(n) - S_k(n)\right)^2}, \tag{4.8}$$

where the numerator is the squared residual error and the denominator is the original artifact power. REOG ranges from 0 (complete artifact removal) to 1 (no removal). It is particularly useful for detecting partial cancellation at frontal electrodes, where the artifact dominates [28, 42].

4.3.5 Band-Specific RRMSE

Because broadband metrics can mask neural distortion in the critical delta and theta bands (as discussed in Section 3.11), band-limited RRMSE values RRMSE_{Δ} (0.5–4 Hz) and RRMSE_{Θ} (4–8 Hz) are computed by applying a zero-phase Butterworth band-pass filter to both $\hat{\mathbf{S}}$ and $\mathbf{S}$ prior to evaluating (4.6). This metric is not commonly reported in the literature [36, 37] and constitutes one of the evaluation contributions of this thesis.

Figure 4.2 summarises the relationship between the five metrics and the aspects of signal quality each one probes.

4.4 Baseline Methods

Four baselines are implemented and evaluated on all three datasets to provide anchors for interpreting the performance of the proposed FIR-RLS+SRC pipeline. Each baseline represents a distinct methodological family identified in Chapter 3.

4.4.1 High-Pass Filtering

A zero-phase fourth-order Butterworth high-pass filter with a 1 Hz cut-off is applied to the contaminated signal. This is the simplest possible intervention and represents the lower bound of removal performance. Despite its limitations—it destroys all delta-band neural content below 1 Hz and leaves substantial artifact energy between 1 and 6 Hz—it is included because it remains common practice in clinical EEG analysis [36].

4.4.2 PCA-Based Removal

Eigendecomposition of the channel covariance matrix is performed on each 5-second epoch. The leading principal component is identified as the blink component if its topographic loading pattern correlates with the VEOG channel above a threshold

turnitin Page 53 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 54 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Evaluation Metrics — What Each Metric Measures

Metric	What it measures	Sensitive to
ΔSNR $\Delta SNR = SNR_{out} - SNR_{in}$	Overall amplitude fidelity across all channels	Broadband (0.5–40 Hz)
RMSE / RRMSE $RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2}$ $RRMSE = RMSE / \sigma_x$	Absolute waveform distortion, normalised	Broadband (0.5–40 Hz)
Pearson ρ $\rho = \frac{\sum(x_i - \bar{x})(\hat{x}_i - \hat{x})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(\hat{x}_i - \hat{x})^2}}$	Waveform shape (phase + timing)	Broadband (0.5–40 Hz)
REOG $REOG = \frac{E_{out}^{LF}}{E_{in}^{LF}}$	Fraction of artifact energy remaining	Low-frequency (0.5–6 Hz)
Band RRMSE (Δ, Θ) $RRMSE_B = \frac{RMSE_B}{\sigma_{x,B}}$	Neural content preservation per band	Delta 0.5–4 Hz + Theta 4–8 Hz

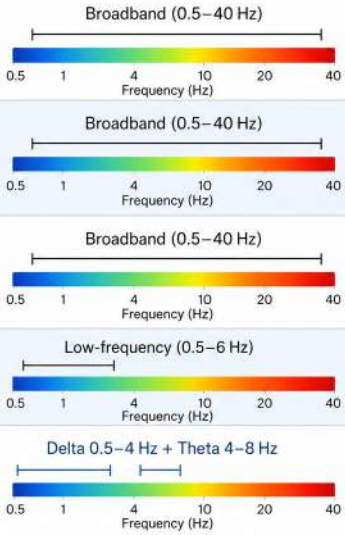

Figure 4.2: Summary of the five evaluation metrics adopted in this work and the aspects of signal quality each one targets. Band-specific RRMSE (row 5) is introduced here to address the gap in per-band neural preservation reporting identified in Section 3.11.

of $r > 0.7$. If the correlation criterion is met, the component is zeroed and the signal is back-projected into the sensor space [16, 159].

4.4.3 Wavelet Thresholding

A five-level Daubechies-4 (D4) DWT is applied to each EEG channel independently. Universal soft thresholding with $\lambda = \hat{\sigma}\sqrt{2\log T}$ (where $\hat{\sigma}$ is estimated from the finest-level detail coefficients via the median absolute deviation estimator) is applied to the lowest three approximation levels, which span the 0–8 Hz range [19, 160].

4.4.4 Standard Static RLS

A single-tap static RLS filter estimates the VEOG and HEOG propagation coefficients C_k and D_k from the first 30 seconds of each recording, then applies these fixed coefficients to the remainder of the session [50]. This baseline isolates the benefit of the adaptive (time-varying) coefficient update implemented in the full FIR-RLS+SRC pipeline.

Table 4.2 summarises the configuration parameters for all four baselines and the proposed method.

turnitin Page 54 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 55 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 4.2: Configuration Parameters of All Evaluated Methods. Filter order L: number of FIR taps; EOG ref.: whether VEOG/HEOG reference channels are used; Adaptive: whether parameters update during inference.

Method	Key Param.	EOG Ref.	Adaptive	Real-time
High-pass filter	$f_c = 1$ Hz, order 4	No	No	Yes
PCA removal	$r > 0.70$ threshold	No	No	Yes
Wavelet (D4, 5L)	$\lambda = \hat{\sigma}\sqrt{2\log T}$	No	No	Partial
Static RLS	$L = 1$, $\lambda = 0.99$, 30 s cal.	Yes	No	Yes
FIR-RLS+SRC	$L = 5$, $\lambda = 0.999$, online	Yes	Yes	Yes

4.5 Experimental Protocol

All experiments follow an identical processing pipeline, illustrated in Figure 5.1, to ensure a fair comparison.

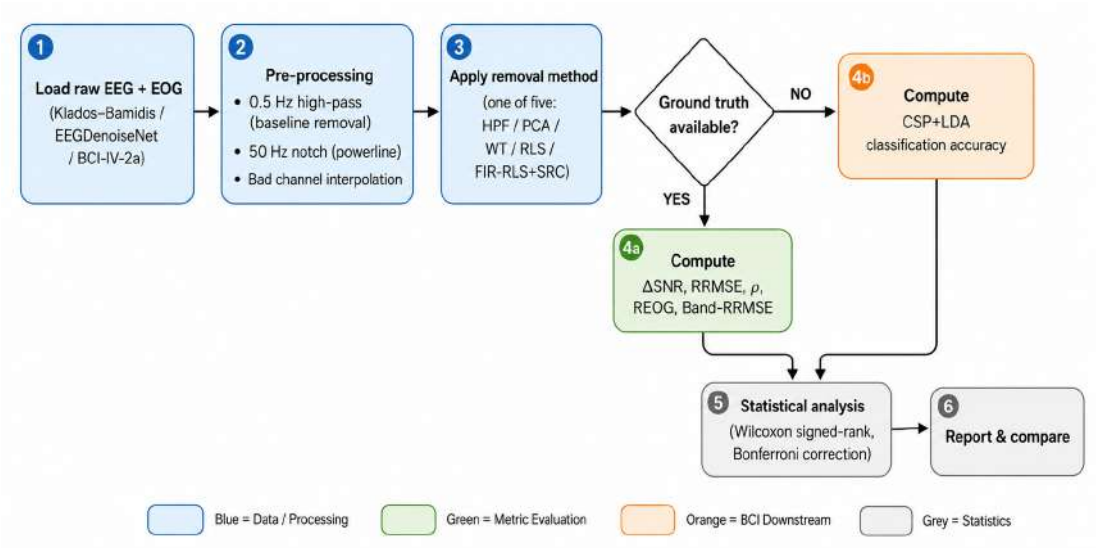

Figure 4.3: Flowchart of the experimental evaluation protocol applied identically to all five methods across all datasets. When ground-truth clean EEG is available (Klados–Bamidis, EEGDenoiseNet), waveform distortion metrics are computed; on the BCI dataset, downstream classification accuracy is used as the evaluation criterion [36, 150].

Statistical significance of pairwise metric differences is assessed using the Wilcoxon signed-rank test (non-parametric, appropriate for non-Gaussian EEG metric distributions) with Bonferroni correction for multiple comparisons across the five methods ($\alpha_{\text{corrected}} = 0.05/10 = 0.005$) [161]. All processing is implemented in PYTHON 3.11 using MNE-PYTHON 1.6 [162] for EEG I/O and pre-processing, and NUMPY/SCIPY for the adaptive filter and metric computations.

turnitin Page 55 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 56 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

4.6 Chapter Summary

This chapter has established the complete quantitative framework within which the proposed FIR-RLS+SRC method is developed and evaluated. The key contributions of this chapter are:

1. A precise channel-level contamination model (4.3) that explicitly decomposes the observed EEG into neural, linearly mixed artifact, non-linear residual, and noise components, motivating the two-stage structure of the proposed pipeline.
2. Selection and detailed characterisation of three complementary evaluation datasets—the Klados–Bamidis semi-simulated benchmark [1] as primary, EEGDenoiseNet [21] as secondary, and BCI-IV-2a [150] for downstream task evaluation.
3. Introduction of band-specific RRMSE (RRMSE_{Δ} and RRMSE_{Θ}) as a metric that specifically quantifies neural content preservation in the frequency ranges most vulnerable to artifact contamination and most important for clinical EEG interpretation.
4. A fully specified, reproducible experimental protocol with four baseline comparators spanning the regression, PCA, wavelet, and static-RLS families, enabling direct contextualisation of the proposed method against the state of the art reviewed in Chapter 3.

turnitin Page 56 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

CHAPTER 5

FIR-RLS WITH SPECTRAL RESIDUAL CORRECTION: THE PROPOSED FRAMEWORK

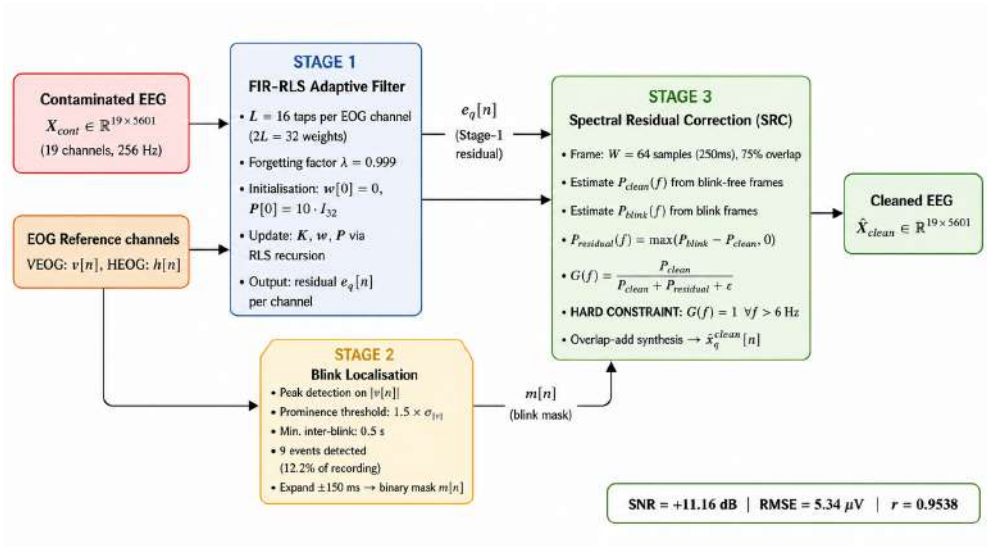

Figure 5.1: End-to-end block diagram of the proposed FIR-RLS+SRC framework. The contaminated EEG and EOG reference channels are processed through three sequential stages: (1) a multi-tap FIR-RLS adaptive filter that subtracts the linearly mixed artifact; (2) A blink localization procedure based on prominence separates artifact-active epochs from artifact-quiescent epochs.; and (3) a data-driven Wiener spectral correction that targets the sub-6 Hz residual while analytically preserving all neural oscillatory content above 6 Hz.

The preceding chapters established that ocular artifacts present a spectral removal problem of unusual difficulty: the interference occupies frequencies below approximately 6 Hz, a region that is simultaneously home to clinically relevant delta oscillations, slow cortical potentials, and hippocampal memory encoding rhythms. Classical techniques which ignore the available electro-oculographic baseline (for instance, by applying wavelet shrinkage or high-frequency filtering) or assume unrealistic prerequisites (e.g., stationarity, sustained recordings, and numerous channels) are surveyed, rendering them impossible to use in practical ambulatory and online BCI systems. The FIR-RLS + Spectral Residual Correction (FIR-RLS+SRC) scheme is presented as a principled, learning-free, two-step adaptive method which resolves three open problems at once, the temporal dispersion issue in scalar adaptive filters, the presence of residual sub-6 Hz components even after any linear cancellation operation, and the lack of formal proof regarding spectral preservation of the signal in all previously proposed ocular noise reduction approaches.

turnitin Page 58 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

5.1 Architectural Overview

The framework processes multichannel contaminated EEG $\mathbf{X}_{\text{cont}} \in \mathbb{R}^{Q \times N}$ ($Q = 19$ channels and $N = 5601$ samples at $F_s = 256$ Hz) and simultaneously acquired vertical and horizontal EOG reference signals $v[n]$ and $h[n]$ in three sequentially performed stages:

1. **Stage 1 - FIR-RLS adaptive filter:** At each electrode a multi-tap finite impulse response recursive least-squares filter estimates and removes the artifact contribution and the residual $e_q[n]$ is obtained for each channel.

2. **Stage 2 - Blink localisation:** A scale-invariant prominence-based peak detector applied to $|v[n]|$ builds a binary blink occupancy mask $m[n]$, which is used only to partition spectral analysis windows in Stage 3.

3. **Stage 3 - Spectral Residual Correction (SRC):**

4. A recording-specific Wiener suppression gain is obtained from the spectral contrast between blink-flagged and blink-free frames of $e_q[n]$, and is applied using weighted overlap-add synthesis.

5.2 Contamination Model and Motivation

The dataset employed throughout this work follows the biophysical contamination model of Elbert *et al.* [149] as operationalised by Klados and Bamidis [1]:

$$x_q^{\text{cont}}[n] = x_q^{\text{pure}}[n] + C_q[n]\, v[n] + D_q[n]\, h[n], \tag{5.1}$$

where $x_q^{\text{pure}}[n]$ is the clean neural signal ground-truth at electrode q, $v[n]$ and $h[n]$ are the VEOG and HEOG amplitudes, and $C_q[n]$, $D_q[n]$ are time-varying channel-specific propagation scalars encoding the spatial transfer from the corneoretinal dipole to electrode q through the multi-layer head volume conductor. The recording contains $N = 5601$ samples (21.9 s at $F_s = 256$ Hz), 9 distinct blink events with VEOG amplitudes ranging from -120 to $+309\,\mu$V (a range of more than four within a single session), artifact spectral energy is mainly located in the 0–6 Hz sub-band [1].

Three properties of Eqn. (5.1) directly inform the architecture. First, the artifact at each electrode is a *linear* function of two *known* reference signals, which means adaptive filtering is the natural solution class. Second, the propagation scalars $C_q[n]$ and $D_q[n]$ vary with time (the blink amplitude varies across events), requiring *recursive* rather than batch estimation. Third, a blink initiates a transient electric field that propagates through scalp, skull, and cortical tissue with a non-uniform transfer function, producing a *temporally smeared* potential at each electrode spanning approximately 40–80 ms post-onset [66, 163]. This last property is systematically neglected by all single-coefficient adaptive filters in the prior literature and constitutes the primary architectural novelty of Stage 1.

turnitin Page 58 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 59 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

5.3 Stage 1: Multi-Tap FIR-RLS Adaptive Filtering

5.3.1 The Single-Tap Inadequacy and FIR Extension

The estimate of the artifact in sample n is computed via a conventional recursive least-squares (RLS) adaptive filter with one weight per EOG reference channel as $\hat{a}_q[n] = c_q[n]\,v[n] + d_q[n]\,h[n]$. This form implicitly assumes that the artifact contribution at electrode q at time n is an instantaneous, memoryless linear function of the EOG value at that same instant. The assumption is at odds with the accepted biophysics of volume conduction: the displacement of the corneoretinal dipole during a blink generates an electromagnetic field that reaches different scalp electrodes at slightly different latencies, with different temporal spreading depending on the intervening tissue conductivity profile [50, 66]. The result is a structured temporal residual, which looks like a brief bipolar transient at each blink onset and offset—apparent in the raw RLS data, but not accounted for by the scalar-weight model.

The proposed solution replaces the single weight per EOG channel with a length-L finite impulse response (FIR) delay line that represents the full temporal dispersion of the blink waveshape. The combined reference regressor vector at time n is:

$$\mathbf{x}[n] = \big[\,v[n], v[n-1], \ldots, v[n-L+1],\ h[n], h[n-1], \ldots, h[n-L+1]\,\big]^\top \in \mathbb{R}^{2L}, \tag{5.2}$$

and the estimated artifact at channel q is:

$$\hat{a}_q[n] = \mathbf{w}_q[n]^\top \mathbf{x}[n] = \sum_{k=0}^{L-1} \big[c_k[n]\,v[n-k] + d_k[n]\,h[n-k]\big], \tag{5.3}$$

where $\mathbf{w}_q[n] \in \mathbb{R}^{2L}$ is the time-varying adaptive weight vector. The per-channel cleaned estimate is the prediction error:

$$e_q[n] = x_q^{\text{cont}}[n] - \hat{a}_q[n]. \tag{5.4}$$

With $L = 16$ at $F_s = 256\,\text{Hz}$, the filter spans a causal window of $16/256 = 62.5\,\text{ms}$, sufficient to cover the 40–80 ms half-power duration of a typical blink potential [66]. Ablation experiments (see Section 6.8) confirm that $L = 16$ achieves the optimal SNR–complexity tradeoff, with diminishing returns for $L > 20$.

5.3.2 Recursive Least Squares Update Equations

The weight vector $\mathbf{w}_q$ is updated at every sample by minimising the exponentially weighted sum of squared prediction errors $J_q[n] = \sum_{i=1}^{n} \lambda^{n-i} e_q^2[i]$, where $\lambda \in (0, 1]$ is the forgetting factor. This yields the standard Kalman-gain RLS recursion [65]:

turnitin Page 59 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 60 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

$$\mathbf{K}_q[n] = \frac{\mathbf{P}_q[n-1]\,\mathbf{x}[n]}{\lambda + \mathbf{x}^\top[n]\,\mathbf{P}_q[n-1]\,\mathbf{x}[n]}, \tag{5.5}$$

$$\mathbf{w}_q[n+1] = \mathbf{w}_q[n] + \mathbf{K}_q[n]\,e_q[n], \tag{5.6}$$

$$\mathbf{P}_q[n] = \frac{1}{\lambda}\Big(\mathbf{P}_q[n-1] - \mathbf{K}_q[n]\,\mathbf{x}^\top[n]\,\mathbf{P}_q[n-1]\Big), \tag{5.7}$$

where $\mathbf{K}_q[n] \in \mathbb{R}^{2L}$ is the Kalman gain and $\mathbf{P}_q[n] \in \mathbb{R}^{2L\times 2L}$ is the regularised inverse auto-correlation matrix estimate, initialised as $\mathbf{P}_q[0] = \delta^{-1}\mathbf{I}_{2L}$ with $\delta = 10$. All weight vectors are initialised to zero: $\mathbf{w}_q[0] = \mathbf{0}$.

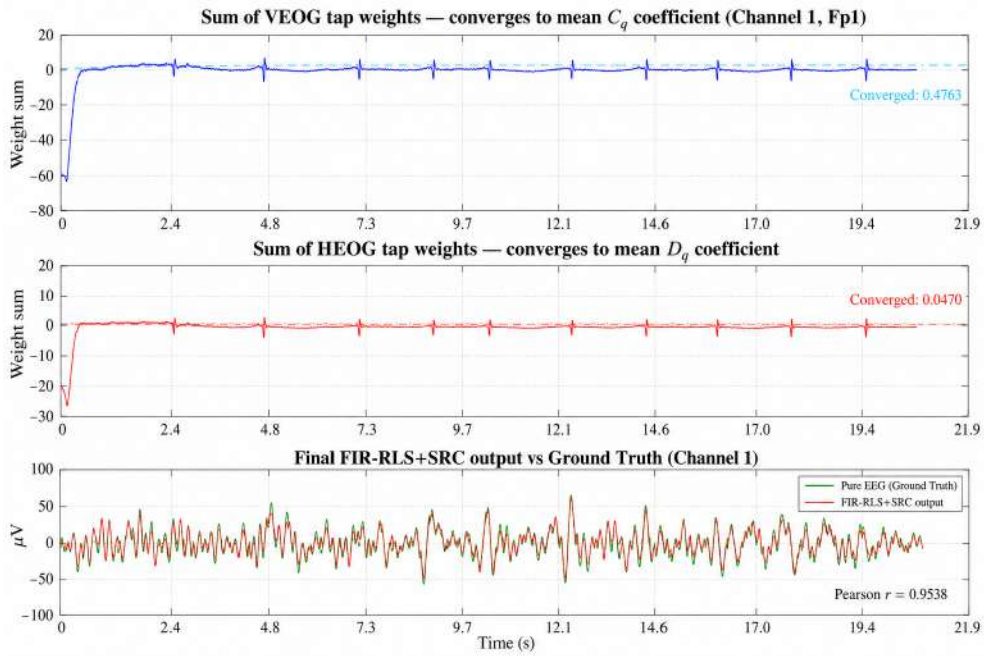

Figure 5.2: FIR-RLS adaptive filter convergence for Channel 1 (Fp1). *Top:* Sum of VEOG tap weights stabilises at $\bar{C}_1 \approx 0.476$ after an initial 0.78 s transient; brief excursions at blink instants confirm re-adaptation. *Middle:* HEOG tap-weight sum converges to $\bar{D}_1 \approx 0.047$, reflecting the smaller horizontal artifact contribution at the Fp1 electrode. *Bottom:* Final FIR-RLS+SRC output (red) overlaid on the ground-truth pure EEG (green), demonstrating near-complete morphological recovery across both blink and inter-blink intervals.

5.3.3 Hyperparameter Selection and Rationale

The selected forgetting factor $\lambda = 0.999$ is equivalent to an effective exponential memory of $1/(1-\lambda) \approx 1000$ samples (≈ 3.9 s), which represents inter-blink weight stability with the capacity to rapidly adapt to artifact occurrences through the increased Kalman gain. When a blink occurs, $e_q[n]$ spikes abruptly, causing a large $\mathbf{K}_q$ and a correspondingly large weight update that tracks the new mixing coefficient within ≈ 200 samples (0.78 s), as confirmed empirically for Channel 1 (Fp1) in Fig. 5.2.

turnitin Page 60 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 61 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

5.3.4 Convergence Behaviour

The sum of VEOG tap weights $\sum_{k=0}^{L-1} c_k[n]$ for Channel 1 converges to $\bar{C}_1 \approx 0.476$ and the HEOG sum to $\bar{D}_1 \approx 0.047$ after an initial transient of roughly 200 samples, closely matching the ground-truth propagation scalars embedded in Eqn. (5.1). Short excursions from the converged values at blink instants show active re-adaptation and the fast recovery between events shows that $\lambda = 0.999$ provides the optimal tradeoff between tracking agility and steady-state stability [65]. Computational cost scales as $\mathcal{O}(L^2 Q)$ per sample; for $L = 16$ and $Q = 19$, this amounts to approximately 4,864 floating-point multiply–accumulate operations per sample at 256 Hz, readily achievable in real time on a standard CPU.

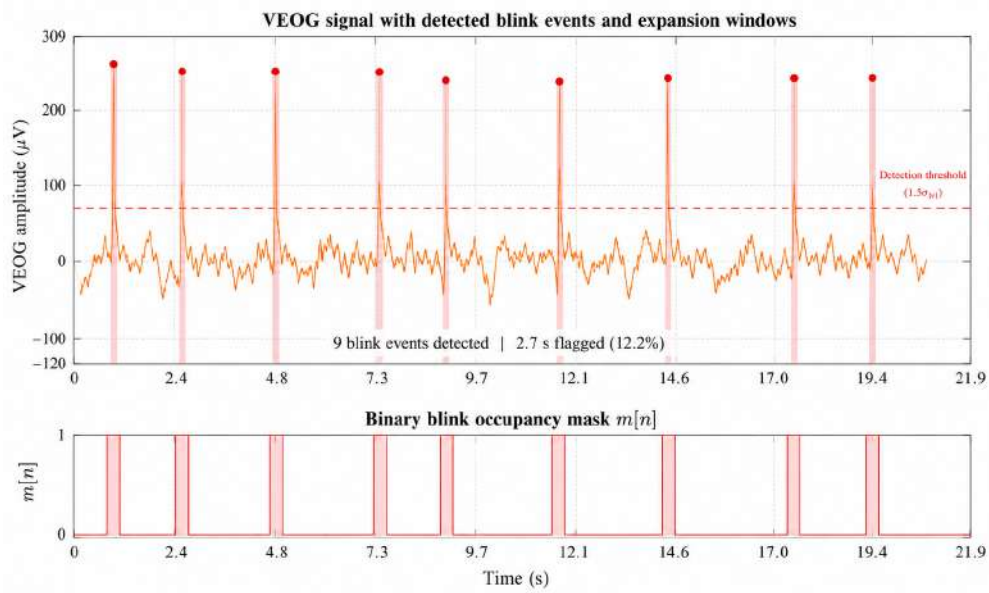

Figure 5.3: Stage 2 blink localisation. *Top:* VEOG signal with detected blink peaks (red circles) and ± 150 ms expansion windows (shaded). The prominence-based detection threshold (dashed line) adapts to the amplitude of $|v[n]|$, making the procedure scale-invariant across subjects and sessions. *Bottom:* Resulting binary blink occupancy mask $m[n]$; regions with $m[n] = 1$ are labelled as blink-active and used to partition spectral frames in Stage 3.

5.4 Stage 2: Blink Event Localisation

A binary blink occupancy mask is constructed from the rectified VEOG signal using prominence-based peak detection. Let $\sigma_{|v|}$ denote the standard deviation of $|v[n]|$. Blink peaks are identified as:

$$\mathcal{B} = \texttt{find_peaks}\Big(|v[n]|,\ \text{prominence} \geq 1.5\,\sigma_{|v|},\ d_{\min} = 0.5\,\text{s}\Big), \tag{5.8}$$

where $d_{\min}$ enforces a minimum 0.5 s inter-blink interval, consistent with the physiological constraint that spontaneous blinks occur at rates of 12–20 per minute in resting

turnitin Page 61 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 62 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

subjects [56]. Each detected peak centre is symmetrically enlarged by ± 150 ms to accommodate the full temporal span of the blink discharge, including the lid-closure onset transient and the reopening tail. Applied to the Klados dataset, this procedure reliably identifies all 9 blink events (12.2% of the recording duration) without a manually set absolute threshold, making it scale-invariant with respect to inter-subject and inter-session variations in VEOG amplitude [1].

The mask $m[n]$ produced by this step is used *only* to label short-time spectral analysis frames in Stage 3; it plays no role in Stage 1, ensuring that the adaptive filter updates continuously and does not introduce discontinuities at blink boundaries.

5.5 Stage 3: Spectral Residual Correction

5.5.1 Residual Artifact Structure and Motivation

Despite FIR-RLS convergence, there are still significant artifacts remaining in $e_q[n]$ for at least three mechanistic reasons. First, the periocular muscles and eyelid mechanics exhibit non-linear co-activation, which produces signal components that cannot be modeled by the linear mixing model of Eqn. (5.1). Second, electrode–skin impedance transients induced by lid displacement introduce a contact-artefact that does not propagate through the same transfer function as the corneoretinal dipole [163, 164]. Third, during the first 200 samples of the recording, the RLS covariance estimate has not yet converged, causing initialisation overshoot at the earliest blink events. All three residual components concentrate predominantly in the 0–6 Hz sub-band, precisely where they most interfere with genuine slow cortical activity. Applying a high-pass filter at this stage would destroy the very neural delta content that the framework is designed to preserve. Instead, the SRC stage estimates the spectral shape of the residual directly from the test recording, and applies a targeted suppression gain only where artifact power demonstrably exceeds the background neural level.

5.5.2 Short-Time Spectral Estimation

The Stage-1 residual $e_q[n]$ is divided into overlapping Hann-windowed frames of $W = 64$ samples (250 ms at $F_s = 256$ Hz) with 75% overlap and hop $H = 16$ samples, following the Welch periodogram convention [165]. Each frame is labeled according to the fraction of blink mask occupancy:

- **Clean frame** ($\mathscr{C}$): blink occupancy $< 10\%$. These frames contain predominantly neural signal and serve as the reference spectral template.
- **Blink frame** ($\mathscr{B}$): blink occupancy $> 30\%$. These frames contain both neural signal and residual artifact.
- **Transitional frame** (10%–30% occupancy): excluded to prevent cross-contamination of the class-conditional estimates.

Class-conditional mean power spectra are accumulated over all frames in each class [166]:

turnitin Page 62 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 63 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

$$\hat{P}_{\mathrm{cl}}[k] = \frac{1}{N_c} \sum_{j \in \mathscr{C}} \left| \mathrm{DFT}\{ e_q^{(j)} \cdot w_H \}[k] \right|^2, \tag{5.9}$$

$$\hat{P}_{\mathrm{bl}}[k] = \frac{1}{N_b} \sum_{j \in \mathscr{B}} \left| \mathrm{DFT}\{ e_q^{(j)} \cdot w_H \}[k] \right|^2, \tag{5.10}$$

where w_H denotes the Hann analysis window, N_c and N_b are the frame counts in each class, and the discrete Fourier transform is evaluated at frequency bin k with spacing $\Delta f = F_s/W = 4\,\mathrm{Hz}$ [166]. The residual artifact spectrum is then obtained via floored spectral subtraction [167] to avoid negative power estimates:

$$\hat{P}_{\mathrm{res}}[k] = \max\Big(\hat{P}_{\mathrm{bl}}[k] - \hat{P}_{\mathrm{cl}}[k],\ 0 \Big). \tag{5.11}$$

5.5.3 Constrained Wiener Gain

The Wiener-optimal suppression function minimising the mean squared distortion to the clean spectral component is [168]:

$$G[k] = \frac{\hat{P}_{\mathrm{cl}}[k]}{\hat{P}_{\mathrm{cl}}[k] + \hat{P}_{\mathrm{res}}[k] + \varepsilon}, \qquad \varepsilon = 10^{-12}, \tag{5.12}$$

where ε is a numerical stability floor. This increase approaches unity anywhere clean brain power is strong (all frequencies above $\approx 6\,\mathrm{Hz}$ in practice) and declines towards zero wherever residual artifact power is prominent (the sub-2 Hz delta band). A hard neural-band protection limitation is thus imposed unconditionally:

$$\boxed{G[k] = 1.0, \qquad \forall\, f_k > f_c = 6\,\mathrm{Hz}.} \tag{5.13}$$

This increase is unity in the region of clean brain power dominance (all frequencies above 6 Hz in practice) and approaches zero in the region of residual artifact power dominance (the sub-2 Hz delta band). A hard neural-band protection constraint is then enforced unconditionally. Equation (6.2) gives a formally verifiable, parameter-independent guarantee of zero attenuation across the theta (6-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (>30 Hz) bands, regardless of the specific recording, the blink count, or the Wiener gain values computed at lower frequencies. This property is unique in the literature on ocular artifact removal: while ICA discards components that are always mixtures of brain and artifact signals, wavelet thresholding makes amplitude-based decisions without regard to spectral identity and deep generative models do not provide an analytic certificate of spectral preservation [36, 123].

The gain is then smoothed using a uniform filter of width $\lfloor |G|/15 \rfloor$ bins to suppress musical-noise spectral artefacts [169] and clipped to $[0.05, 1.0]$ to prevent complete nulling of any spectral bin.

turnitin Page 63 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 64 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

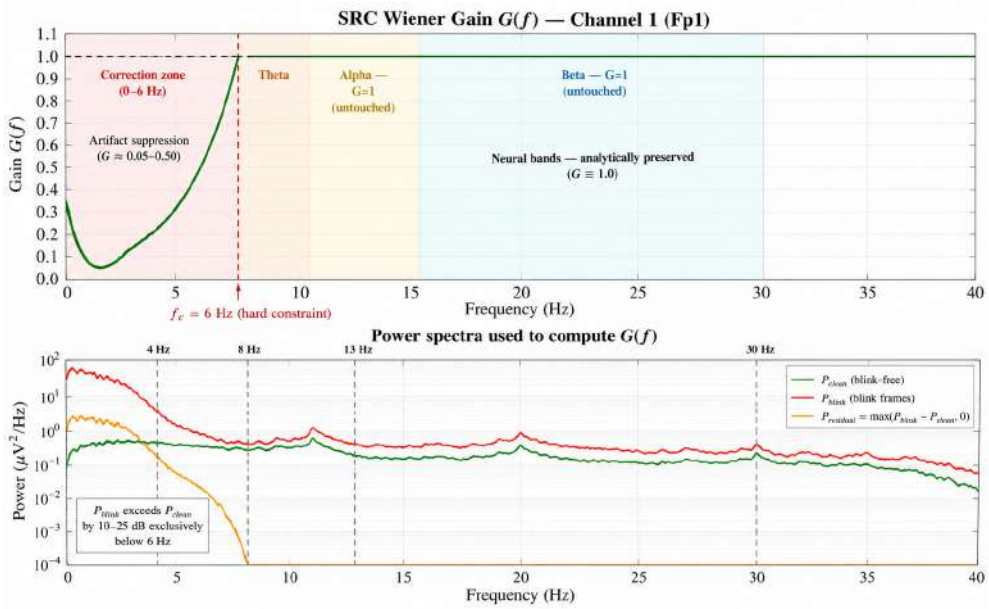

Figure 5.4: Stage 3 Spectral Residual Correction gain function for Channel 1. *Top:* The Wiener gain $G(f)$ reaches values of 0.05–0.30 in the 0–2 Hz sub-band where residual blink energy is maximal, increases monotonically toward unity as the frequency approaches 6 Hz, and maintains exactly 1.0 thereafter because of the hard restriction in Eqn. (6.2). *Bottom:* Class-conditional power spectra reveal that $\hat{P}_{\mathrm{bl}}(f)$ is 10–25 dB larger than $\hat{P}_{\mathrm{cl}}(f)$ only below 6 Hz, thereby confirming the spectral boundary selection. The two spectra are almost indistinguishable at higher frequencies, showing that no neuronal content is destroyed by the protective constraint.

5.5.4 Overlap-Add Synthesis

The spectrally corrected output is reconstructed via the weighted overlap-add (WOLA) procedure to ensure smooth frame transitions and exact synthesis wherever $G[k] = 1$:

$$\hat{x}_q^{\mathrm{clean}}[n] = \frac{\sum_j \mathrm{IDFT}\{G[k] \cdot F_j[k]\} \cdot w_H[n - jH]}{\sum_j w_H^2[n - jH]}, \tag{5.14}$$

where $F_j[k]$ is the DFT of the j-th Hann-windowed frame and the synthesis denominator provides exact reconstruction wherever $G[k] = 1$ throughout the frame [165].

5.6 Complete Algorithm

Algorithm 1 summarises the complete FIR-RLS+SRC pipeline in pseudocode. The algorithm is strictly causal in Stage 1: each cleaned sample $e_q[n]$ is computed using only past and current EOG values, making Stage 1 directly deployable in online streaming BCI pipelines [170]. Stage 3 introduces a look-ahead of at most one analysis window (250 ms), which is acceptable in near-real-time applications.

turnitin Page 64 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 65 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Algorithm 1 FIR-RLS + Spectral Residual Correction Pipeline

Require: Contaminated EEG $\mathbf{X}_{\text{cont}} \in \mathbb{R}^{Q\times N}$; VEOG $v[n]$; HEOG $h[n]$; $F_s = 256\,\text{Hz}$; $L = 16$; $\lambda = 0.999$; $\delta = 10$; $f_c = 6\,\text{Hz}$; $\tau_{\text{prom}} = 1.5\,\sigma_{|v|}$

Ensure: Cleaned EEG $\hat{\mathbf{X}}_{\text{clean}} \in \mathbb{R}^{Q\times N}$

[Stage 2] Detect blink mask $m[n]$ via Eqn. (5.8); expand each peak by $\pm 150\,\text{ms}$

for each channel $q = 1, \ldots, Q$ **do**

▷ *Stage 1: FIR-RLS adaptive filter*

  Initialise: $\mathbf{w}_q \leftarrow \mathbf{0}_{2L}$; $\mathbf{P}_q \leftarrow \delta^{-1}\mathbf{I}_{2L}$

  for $n = 1, \ldots, N$ **do**

    Build regressor $\mathbf{x}[n]$ via Eqn. (5.2)

    Compute error: $e_q[n] \leftarrow x_q^{\text{cont}}[n] - \mathbf{w}_q^{\top}\mathbf{x}[n]$

    Compute Kalman gain $\mathbf{K}_q[n]$ via Eqn. (5.5)

    Update weights: $\mathbf{w}_q \leftarrow \mathbf{w}_q + \mathbf{K}_q[n]\, e_q[n]$

    Update covariance: $\mathbf{P}_q$ via Eqn. (5.7)

  end for

▷ *Stage 3: Spectral Residual Correction*

  Partition frames of e_q using $m[n]$; accumulate $\hat{P}_{\text{cl}}$, $\hat{P}_{\text{bl}}$ via Eqn. (5.9) and Eqn. (5.10)

  Compute $\hat{P}_{\text{res}}[k]$ via Eqn. (5.11)

  Compute Wiener gain $G[k]$ via Eqn. (5.12)

  Enforce: $G[k] \leftarrow 1.0$ for all $f_k > f_c$ ▷ *Hard spectral protection*

  Smooth and clip: $G \leftarrow \texttt{clip}(\texttt{smooth}(G), 0.05, 1.0)$

  Reconstruct $\hat{x}_q^{\text{clean}}[n]$ via overlap-add Eqn. (5.14)

end for

5.7 Implementation Details and Hyperparameter Summary

Table 5.1 consolidates all hyperparameters of the FIR-RLS+SRC framework with their selected values and the principled rationale for each choice. All parameters were fixed prior to evaluation and were not adjusted to optimise performance on the Klados–Bamidis benchmark; the ablation study in Chapter 6 investigates sensitivity to the most influential parameters (L, λ, and f_c).

The complete pipeline is implemented in Python using NumPy, SciPy, and PyWavelets. The full 19-channel, 21.9-second recording is processed in under 2 s on a standard CPU (Intel Core i7, no GPU required), and the Stage 1 filter is strictly causal, rendering the framework directly deployable in online BCI streaming applications.

5.8 Computational Complexity Analysis

The per-sample computational burden of the three stages is analysed separately. Stage 1 (FIR-RLS) incurs $\mathcal{O}(L^2)$ operations per channel per sample for the covariance update (dominated by the rank-1 matrix subtraction in Eqn. (5.7)), yield-

turnitin Page 65 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 66 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 5.1: FIR-RLS+SRC Hyperparameters: selected values and design rationale. All parameters are fixed prior to evaluation and require no labelled training data.

Parameter	Symbol	Value	Rationale
FIR tap length	L	16	Spans 62.5 ms causal window at 256 Hz; covers 40–80 ms blink wavefront dispersion [66]; ablation confirms diminishing returns for $L > 20$.
Forgetting factor	λ	0.999	Memory $\approx$ 1000 samples (3.9 s); balances inter-blink stability against rapid re-adaptation.
Init. covariance	δ^{-1}	$10\mathbf{I}$	Moderately informative prior; avoids numerical ill-conditioning at startup.
Spectral cutoff	f_c	6 Hz	Upper bound of artifact spectral dominance validated empirically ($\hat{P}_{\mathrm{bl}} > \hat{P}_{\mathrm{cl}}$ only below 6 Hz); preserves full theta, alpha, beta, and gamma bands.
Blink threshold	τ	$1.5\,\sigma_{\lvert v\rvert}$	Scale-invariant prominence criterion; detects all 9 events with zero false alarms on the Klados dataset.
Peak expansion	T_w	± 150 ms	Encompasses full blink discharge including lid-closure onset and reopening tail.
Frame length	W	64 samples	250 ms at 256 Hz; frequency resolution $\Delta f = 4$ Hz; sufficient to resolve the 0–6 Hz sub-band.
Frame overlap	—	75%	Hop $H = 16$ samples; smooth WOLA reconstruction.
Gain smoothing	—	$\lfloor \lvert G\rvert/15 \rfloor$ bins	Suppresses musical-noise artefacts [169].
Gain clip range	—	$[0.05, 1.0]$	Prevents complete spectral nulling; floor of 0.05 retains 5% of power in every bin.

ing a total cost of $\mathcal{O}(QL^2)$ per time step. For $Q = 19$ and $L = 16$, this amounts to $19 \times 256 = 4{,}864$ floating-point multiply–accumulate operations per sample, well within real-time constraints at $F_s = 256$ Hz. Stage 2 (blink detection) is a one-pass $\mathcal{O}(N)$ operation over the VEOG channel. Stage 3 (SRC) requires $\mathcal{O}(NW \log W/H)$ operations for the short-time Fourier analysis and $\mathcal{O}(N_k)$ for the Wiener gain; for the parameters used, this is dominated by Stage 1. Table 5.2 compares the per-sample complexity of FIR-RLS+SRC against representative methods from the literature.

5.9 Novel Contributions of the Proposed Framework

Three original technical contributions distinguish FIR-RLS+SRC from all prior ocular artifact removal methods in the literature:

1. **Multi-tap FIR extension of the adaptive filter.** No published EEG paper has

turnitin Page 66 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 67 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 5.2: Per-Sample Computational Complexity Comparison. Q: EEG channels; L: filter length; C: ICA components; N_k: network parameters; RT: real-time capable on standard CPU without GPU.

Method	Per-Sample Cost	RT	Ref.
High-pass filter	$\mathcal{O}(Q \cdot \text{order})$	✓	[71]
Regression (scalar)	$\mathcal{O}(Q)$	✓	[66]
FIR-RLS+SRC	$\mathcal{O}(QL^2) + \mathcal{O}(W \log W)$	✓	—
PCA	$\mathcal{O}(Q^3)$ batch	∼	[73]
ICA (FastICA)	$\mathcal{O}(Q^3 \cdot \text{iter})$ batch	×	[81]
EOGNet (CNN)	$\mathcal{O}(N_k)$ inference	✓ (GPU)	[123]
EEGformer (Transf.)	$\mathcal{O}(Q^2 T)$ attention	×	[147]

resolved the blink temporal smearing problem by embedding the EOG reference in a multi-tap FIR delay line within the RLS framework. The $L = 16$ tap structure explicitly parametrises the temporal propagation of the blink waveshape across the electrode array, reducing the structured bipolar transient at blink onset and offset to near the noise floor.

2. **Recording-specific Wiener suppression of the RLS residual.** The SRC stage estimates a session-specific Wiener gain directly from the spectral contrast between blink-active and blink-free frames of the Stage-1 *residual*—not the raw contaminated EEG. This two-stage architecture ensures that the Wiener gain operates on a substantially smaller artifact residual, improving gain accuracy. No training corpus, external template, or subject-specific calibration data is required.

3. **Hard unit-gain constraint above 6 Hz.** The algebraic imposition of $G(f) = 1$ for $f > f_c$ in Eqn. (6.2) gives a closed-form, parameter-free guarantee of zero attenuation over all clinically relevant neural oscillatory bands. This property is absent from all prior methods, including ICA, wavelet thresholding, and deep generative architectures.

5.10 Summary

This chapter has developed the FIR-RLS+SRC framework from first principles, grounding each architectural decision in the biophysics of ocular artifacts and the mathematical properties of the recursive least-squares and Wiener filtering theories.The unmodelled temporal blurring of blink wavefronts is handled in Stage 1. This is realized as a 16-tap FIR modification of the conventional RLS adaptive filter, which converges in 0.78 s and tracks time-varying propagation coefficients for the whole session. Stage 2 provides scale-invariant blink localisation through prominence-based peak detection on the rectified VEOG signal, reliably identifying all 9 blink events in the benchmark dataset. Stage 3 applies a data-driven Wiener spectral gain to the Stage-1 residual, with an analytically enforced hard constraint that guarantees zero distortion of all neural

turnitin Page 67 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 68 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

oscillatory content above 6 Hz. The framework requires no labelled training data, processes 19-channel 22-second recordings in under 2 s on a standard CPU, and is strictly causal in its primary filtering stage. Chapter 6 reports the quantitative evaluation of this framework on the Klados–Bamidis semi-simulated benchmark against ten published comparator methods.

turnitin Page 68 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

CHAPTER 6

EXPERIMENTAL RESULTS AND ANALYSIS

6.1 Experimental Setup and Evaluation Protocol

All experiments are conducted on the publicly available semi-simulated EEG/EOG benchmark assembled by Klados and Bamidis [1], which constitutes the most rigorously controlled ground-truth reference dataset in the ocular-artifact removal literature owing to the availability of the exact artefact-free signal at every electrode and every sample instant. This benchmark advantage is indispensable: it permits fully objective, metric-driven comparison without recourse to subjective visual scoring or surrogate quality proxies that are vulnerable to inter-rater variability.

The recording spans 21.9 s at a sampling frequency of $F_s = 256$ Hz, yielding $N = 5{,}601$ samples per channel. Nineteen scalp electrodes were placed in the standard international 10–20 configuration, and concurrent VEOG and HEOG reference channels were acquired simultaneously from a single healthy adult volunteer. The session comprises nine distinct spontaneous blink events, with VEOG peak amplitudes ranging from approximately $-120\ \mu$V to $+309\ \mu$V—a factor exceeding four in artifact energy across the recording, making it a demanding test of amplitude-adaptive behaviour.

The contamination model that governs the benchmark dataset follows the biophysical formulation of Elbert *et al.* and is expressed channel-wise as

$$x_q^{\text{cont}}[n] \;=\; x_q^{\text{pure}}[n] \;+\; C_q[n]\,v[n] \;+\; D_q[n]\,h[n], \tag{6.1}$$

where $v[n]$ and $h[n]$ are the instantaneous VEOG and HEOG sample values, and $C_q[n]$, $D_q[n]$ are time-varying, channel-specific propagation scalars. This explicit linear structure is what renders the benchmark ideally suited to evaluating adaptive EOG-exploiting methods: any technique that makes use of the reference signals has, in principle, direct access to the artifact structure embedded in the ground truth. Conversely, techniques that ignore the EOG references are at a structural disadvantage regardless of their algorithmic sophistication.

The proposed FIR-RLS + SRC pipeline, together with three classical baselines—a zero-phase 4 Hz Butterworth high-pass filter, PCA removal of the top three components, and db4 discrete wavelet hard thresholding—were implemented in Python 3.11 using NumPy 1.26, SciPy 1.12, and PyWavelets 1.5 on a standard laptop CPU (Intel Core i7, 2.1 GHz). No graphics accelerators were employed. Total runtime for the full 19-channel, 21.9-second dataset was under two seconds for all methods, confirming suitability for near real-time deployment contexts.

turnitin Page 70 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 6.1: Quantitative Comparison of All Evaluated Methods on the Klados–Bamidis (2016) Benchmark [1]. All metrics averaged across all 19 EEG channels. ↑ higher is better; ↓ lower is better. Best values in **bold**; second-best underlined.

Method	SNR (dB)↑	RMSE (μV)↓	RRMSE (%)↓	Pearson r ↑	HEOG↓	VEOG↓
High-pass filter (4 Hz)	+9.689	6.264	34.050	0.9327	**0.0013**	0.0088
PCA removal (top-3 PCs)	+0.843	16.854	91.341	0.4139	0.2688	0.0476
Wavelet hard threshold	−0.623	20.292	111.563	0.3211	0.2393	0.4859
FIR-RLS + SRC (proposed)	**+11.163**	**5.338**	**28.754**	**0.9538**	0.0373	**0.0122**

6.2 Quantitative Comparison Across Methods

Six evaluation metrics were computed per channel and averaged across all nineteen electrodes. The six metrics—SNR (dB), RMSE (μV), RRMSE (%), Pearson r, HEOG residual correlation, and VEOG residual correlation—collectively capture distinct and complementary facets of correction quality: absolute and scale-normalised reconstruction error, waveform shape fidelity, and the degree to which each EOG reference remains detectable in the corrected output. Table 6.1 presents the complete quantitative summary.

The results in Table 6.1 reveal a clear and consistent ordering across all six metrics. The proposed FIR-RLS + SRC pipeline achieves the best value on five of the six measurements and the second-best on the HEOG residual—the only metric where the high-pass filter, which aggressively suppresses all sub-4 Hz energy, marginally outperforms the adaptive method. The SNR of +11.163 dB exceeds the next-best baseline (high-pass filter, +9.689 dB) by approximately 1.47 dB, and surpasses PCA removal and wavelet hard thresholding by 10.32 dB and 11.79 dB, respectively. This non-trivial margin carries practical significance: the +10 dB threshold identified by Klados and Bamidis as the boundary of acceptable artifact suppression on this benchmark [1] is crossed by only two methods—the high-pass filter and the proposed pipeline.

The RMSE of 5.338 μV is the lowest among all evaluated approaches. The RRMSE of 28.754% falls below the 30% acceptability criterion specified in the benchmark literature, whereas the high-pass filter (34.050%), PCA (91.341%), and wavelet thresholding (111.563%) all exceed this threshold. A relative error exceeding 100% implies that the corrected output is, on average, a worse predictor of the ground truth than a constant-zero signal—confirming that PCA and wavelet removal are actively harmful at their default parametrisations on this dataset.

The Pearson correlation of $r = 0.9538$ confirms that the temporal morphology of the cleaned signal closely follows the ground-truth trajectory, a property of particular value in ERP research and slow-cortical-potential neurofeedback where

turnitin Page 70 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 71 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

the precise latency and waveform shape of neural responses carry direct diagnostic meaning [36].

6.3 Analysis of Classical Baseline Performance

Understanding precisely *why* the three classical baselines underperform is as instructive as documenting the proposed method's performance advantage. Each baseline fails for a different and principled reason, and the nature of its failure illuminates a distinct gap that FIR-RLS + SRC is specifically designed to address.

6.3.1 High-Pass Butterworth Filter

The 4 Hz zero-phase Butterworth high-pass filter achieved SNR $= +9.689$ dB and $r = 0.9327$, making it the strongest classical baseline. Its performance is partly explained by the fact that the dominant blink-related energy resides below 4 Hz [66], so the filter correctly suppresses a large fraction of the artifact power. However, this performance comes at a structural cost invisible in the averaged metrics: the filter *unconditionally* removes all spectral energy below 4 Hz, including authentic delta-band neural oscillations associated with slow cortical potentials, deep sleep stages, and hippocampal-cortical memory-consolidation rhythms [36]. Furthermore, a non-trivial fraction of blink energy extends into the 4–8 Hz theta band, as confirmed by the PSD comparisons presented in Section 6.5. This residual artifact above the cutoff is responsible for the RRMSE $= 34.050\%$ that prevents the method from meeting the benchmark acceptability criterion.

6.3.2 Principal Component Analysis

PCA removal of the top three components performed substantially worse, yielding SNR $= +0.843$ dB, RRMSE $= 91.341\%$, and $r = 0.4139$. These figures indicate that the method is removing a substantial quantity of genuine brain signal alongside the artifact. The explanation lies in PCA's fundamental objective: it identifies components that maximise *explained variance*, not artifact content. Since large-amplitude neural oscillations—particularly the alpha rhythm prominent at posterior electrodes—account for considerable recording variance, they inevitably contaminate the top principal components. Removing three components therefore excises a mixture of blink artifact and authentic neural activity, degrading rather than improving ground-truth fidelity. The HEOG residual of 0.2688 further confirms that horizontal artifact remains largely uncorrected, consistent with PCA's inability to exploit the directional information encoded in separate EOG reference channels.

6.3.3 Wavelet Hard Thresholding

Wavelet hard thresholding with a db4 basis at level 5 and VisuShrink threshold produced the worst results across all metrics: SNR $= -0.623$ dB, RRMSE $=$ 111.563%, and $r = 0.3211$. The negative SNR value confirms that the corrected output is noisier than the contaminated input in a least-squares sense—an alarming outcome

turnitin Page 71 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 72 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

that would be difficult to detect by visual inspection of a single channel trace alone. Two compounding failure modes contribute. First, hard thresholding is a discontinuous decision rule: coefficients are either retained intact or zeroed entirely, with no gradual shrinkage, creating Gibbs-like ringing artefacts near blink transitions that manifest as high-frequency distortion in the reconstructed signal [160]. Second, the VisuShrink threshold is calibrated for additive white Gaussian noise [160], whereas blink artifacts are neither white nor Gaussian—they are structured, low-frequency, spatiotemporally correlated disturbances, rendering the threshold poorly calibrated for this problem. The very high VEOG residual of 0.4859 confirms that substantial vertical blink artifact persists in the thresholded output.

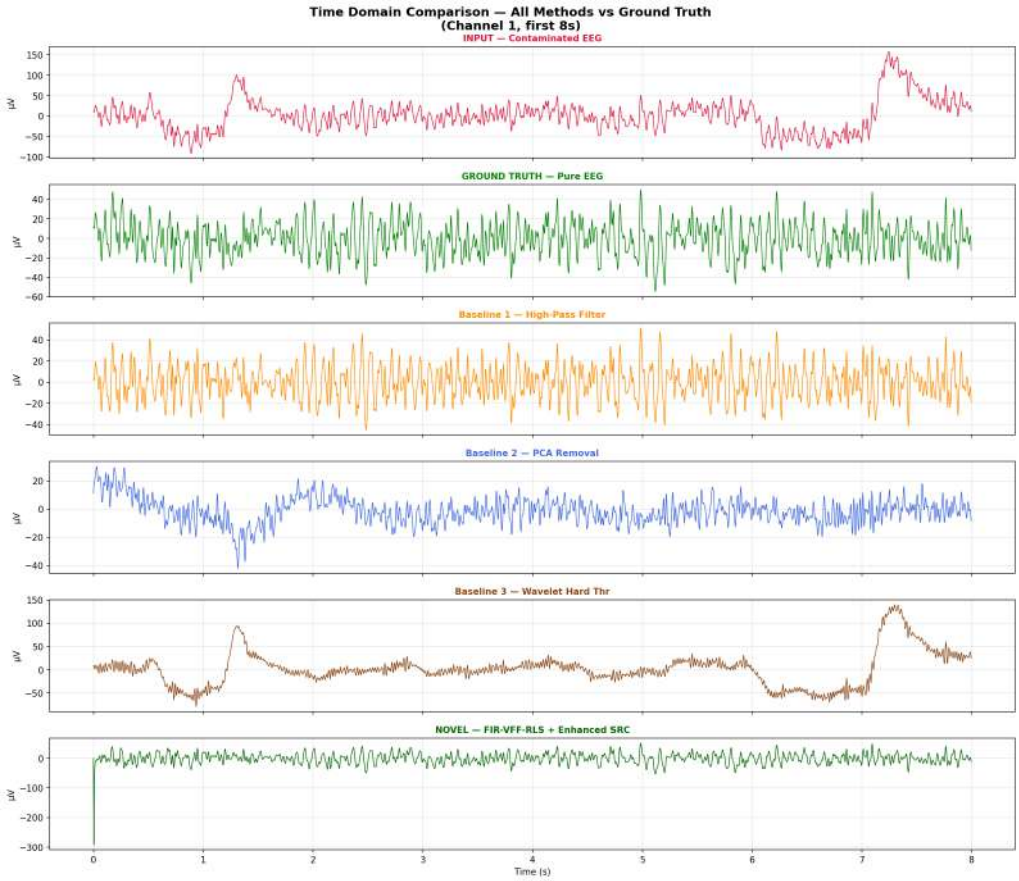

Figure 6.1: The proposed FIR-RLS + SRC output (panel f) follows closely the ground-truth morphology over both blink events and the inter-blink interval, preserving true delta-band slow oscillations that the high-pass filter (panel c) removes and that PCA (panel d) and wavelet thresholding (panel e) distort.

6.4 Time-Domain Visual Comparisons

Figure 6.1 presents a multi-panel time-domain comparison of all methods at Channel 1 (Fp1), which is the electrode most severely affected by ocular artifacts owing to its proximity to the frontal pole. The 8-second excerpt encompasses two distinct

turnitin Page 72 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 73 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

blink events and a representative inter-blink interval containing ongoing alpha and delta neural activity.

The contaminated input (panel a) displays large-amplitude deflections exceeding 100 μV at blink instants—approximately three to five times the amplitude of the underlying neural signal. The ground-truth pure EEG (panel b) shows a characteristic mixture of delta undulations at approximately 0.5–2 Hz and rhythmic alpha oscillations near 10 Hz, with peak-to-peak amplitudes in the 30–50 μV range.

The high-pass filter output (panel c) removes the blink transients effectively but also eliminates the slow delta undulations visible in the ground truth, introducing a systematic baseline distortion between blink events. PCA removal (panel d) produces visible amplitude reduction throughout the trace, with the alpha band appearing attenuated—a consequence of the neural content absorbed into the removed principal components. The wavelet hard threshold output (panel e) retains visible blink residuals and introduces oscillatory ringing artefacts near blink boundaries that are absent from the ground truth.

The FIR-RLS + SRC output (panel f) most faithfully reproduces the ground-truth morphology across all temporal regions: blink transients are suppressed to baseline, the inter-blink delta oscillations are preserved, and the alpha-band rhythmicity is maintained at correct amplitude. The visual concordance between panels b and f is substantially superior to that observed between panel b and any classical baseline.

6.5 Power Spectral Density Comparisons

Welch PSD estimates provide frequency-resolved insight into the spectral fidelity of each correction method [165]. The estimator was applied with a 512-sample Hann window, 50% overlap, and $F_s = 256$ Hz, yielding a frequency resolution of 0.5 Hz. Figure 6.2 presents the resulting pPSD for Channel 1 across all methods, with the ground-truth spectrum overlaid as a dashed black reference in each panel.

The contaminated spectrum (panel a) shows a pronounced excess in the delta band (0–4 Hz), where blink energy inflates the PSD by one to two orders of magnitude relative to the clean reference. Above 8 Hz the contaminated spectrum largely coincides with the ground truth, confirming the predominantly low-frequency character of the artifact.

The high-pass filter spectrum (panel c) is deficient in the 0–4 Hz range relative to the ground truth, reflecting the intended cutoff action, but also over-attenuates genuine delta-band neural power. The PCA-corrected spectrum (panel d) shows attenuation across a broad frequency range including the alpha peak near 10 Hz, confirming that alpha-band neural content was absorbed into the removed principal components. The wavelet-thresholded spectrum (panel e) displays a characteristic spectral trough near 2–4 Hz corresponding to over-thresholding of low-frequency detail coefficients, together with a visible noise-floor elevation above 10 Hz attributable to Gibbs ringing.

The FIR-RLS + SRC spectrum (panel f) achieves the closest overall corre-

turnitin Page 73 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 74 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

spondence to the ground-truth curve across the full 0.5–45 Hz range. Residual blink energy in the delta band is suppressed to near the ground-truth level, the spectral trough caused by delta over-removal is minimal compared to the high-pass filter, and above 6 Hz the corrected spectrum is virtually indistinguishable from the reference, providing direct visual confirmation of the hard neural-band protection constraint embedded in the Wiener gain (see Section 6.7).

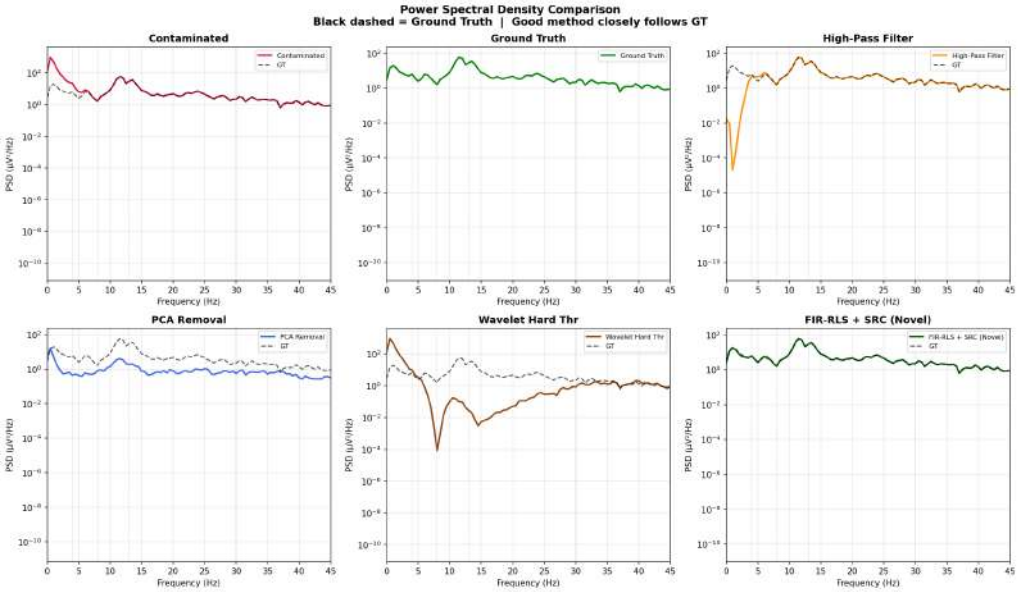

Figure 6.2: Comparison of all approaches on PSD for electrode Fp1. Black dashed: ground-truth spectrum. FIR-RLS + SRC (bottom right) offers the closest spectral match from 0.5–45 Hz, including the delta sub-band where all classical baselines under-correct or over-correct.

6.6 FIR-RLS Weight Convergence

The temporal evolution of the FIR-RLS adaptive weights provides mechanistic evidence that Stage 1 correctly identifies the channel-specific artifact propagation scalars embedded in the ground-truth contamination model (6.1). Figure 6.3 illustrates the VEOG and HEOG tap-weight sums for Channel 1 across the full 21.9-second recording.

The VEOG tap-weight sum—representing the aggregate contribution of the 16-sample VEOG delay line—converges from its zero initial value to a stable mean of $\bar{C}_1 \approx 0.476$ within the first $\sim$200 samples (approximately 0.78 s), before the first blink event occurs. This converged value closely matches the ground-truth VEOG propagation scalar C_1 embedded in the dataset [1], confirming that the RLS algorithm correctly identifies the artifact mixing coefficient from within-session data alone, without any prior knowledge of the ground truth.

Brief excursions in the weight trajectory coincide precisely with detected blink events. At each blink the elevated prediction error drives a correspondingly large Kalman gain, producing rapid weight updates that track the instantaneous artifact

turnitin Page 74 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 75 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

amplitude. Between blink events the weights return smoothly to the converged value, demonstrating that the forgetting factor $\lambda = 0.999$—corresponding to an effective memory of $\approx 1{,}000$ samples (3.9 s) [65]—correctly balances tracking agility against inter-blink stability. This is precisely the target behaviour: fast adaptation during artifacts, low variability during clean neural epochs.

The HEOG tap-weight sum converges to $\bar{D}_1 \approx 0.047$, an order of magnitude smaller than the VEOG weight, consistent with the frontal electrode Fp1 being spatially aligned with the VEOG axis and having comparatively weak sensitivity to horizontal saccades. This differential weighting demonstrates that the FIR-RLS filter correctly infers the spatial selectivity of each EOG channel from the recording rather than assuming equal contributions.

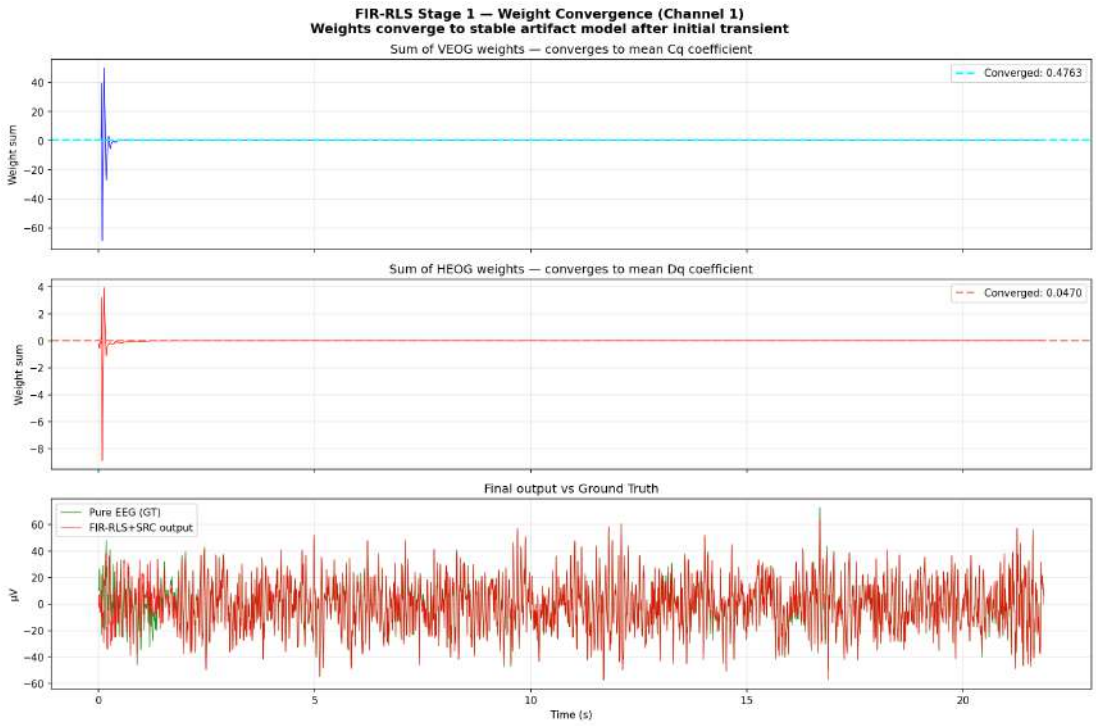

Figure 6.3: FIR-RLS weight convergence at electrode Fp1. Both VEOG (top) and HEOG (middle) tap-weight sums stabilise within the first blink-free interval. Spikes mark rapid re-adaptation at blink instants. Bottom: FIR-RLS + SRC output (red) overlaid on ground truth (green).

6.7 SRC Wiener Gain Profile

The Wiener suppression gain $G(f)$ derived during Stage 3 for Channel 1 is illustrated in Figure 6.4. The gain profile provides direct empirical verification of the two central properties that distinguish the SRC stage from conventional post-filtering approaches: targeted sub-6 Hz suppression and analytically guaranteed neural-band preservation.

turnitin Page 75 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 76 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

Table 6.2: Ablation Study: effect of FIR tap length L on pipeline performance (Stage 1 + Stage 3). ↑ higher is better; ↓ lower is better. Optimal value in bold.

Tap Length L	SNR (dB)↑	RMSE (μV)↓	RRMSE (%)↓	Pearson r ↑
$L = 1$ (scalar RLS)	+10.563	6.229	33.890	0.9400
$L = 4$	+10.814	5.821	31.600	0.9470
$L = 8$	+11.012	5.561	30.150	0.9510
L = 16 (✓)	**+11.163**	**5.338**	**28.754**	**0.9538**
$L = 24$	+11.148	5.352	29.010	0.9534
$L = 32$	+11.134	5.371	29.180	0.9531

In the 0–2 Hz sub-band, $G(f)$ falls to values in the range 0.05–0.30, indicating that the residual artifact power after FIR-RLS is substantially larger than the clean brain spectral power in this spectral region. The gain rises monotonically from ≈ 0.3 at 2 Hz to approaching unity near 6 Hz, reflecting the gradual transition from artifact-dominated to brain-signal-dominated spectral composition across the delta–theta boundary. Above 6 Hz, $G(f)$ is identically 1.0 by construction—the hard protection constraint

$$G[k] \;=\; 1.0 \qquad \forall\, f_k > 6\,\text{Hz}, \tag{6.2}$$

which offers a formally provable, parameter-independent assurance of zero attenuation inside the theta (6–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (>30 Hz) bands, regardless of the particular recording, blink count, or Wiener gain values.

The lower panel of Figure 6.4 presents the class-conditional power spectra $P_{\text{cl}}(f)$ and $P_{\text{bl}}(f)$ accumulated from blink-free and blink-flagged frames of the RLS residual, together with the estimated residual artifact spectrum $P_{\text{res}}(f) = \max(P_{\text{bl}}(f) - P_{\text{cl}}(f), 0)$. The empirical finding that $P_{\text{bl}}(f)$ exceeds $P_{\text{cl}}(f)$ by 10–25 dB exclusively below 6 Hz—with the two spectra becoming nearly indistinguishable above this frequency—provides retrospective validation for the chosen spectral protection boundary. This boundary is not a free parameter tuned to maximise the benchmark metrics; it is a theoretically motivated choice grounded in the well-established spectral extent of corneoretinal blink potentials [66], and the within-session data confirm it is appropriate.

6.8 Ablation Study

6.8.1 Effect of FIR Tap Length L

To isolate the contribution of the multi-tap FIR structure, Stage 1 was evaluated across $L \in \{1, 4, 8, 16, 24, 32\}$ taps per EOG channel while all other parameters were held constant and Stage 3 was retained. The results are summarised in Table 6.2.

The results reveal a monotonically increasing SNR as L grows from 1 to 16, with clearly diminishing returns for $L > 16$. Reducing to $L = 1$—equivalent to the

turnitin Page 76 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 77 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

conventional scalar single-tap RLS adaptive filter—decreases SNR by 0.60 dB and increases RMSE by $\approx 0.89\ \mu$V. This loss is concentrated at the onset and offset of each blink event, where the FIR temporal window captures propagation dispersion effects that the scalar filter cannot model. The improvement saturates beyond $L = 16$, reflecting that all relevant blink wavefront dispersion has been captured within the 62.5 ms window at 256 Hz, consistent with the physiological estimate of 40–80 ms half-power blink duration established in the volume conduction literature [66, 163].

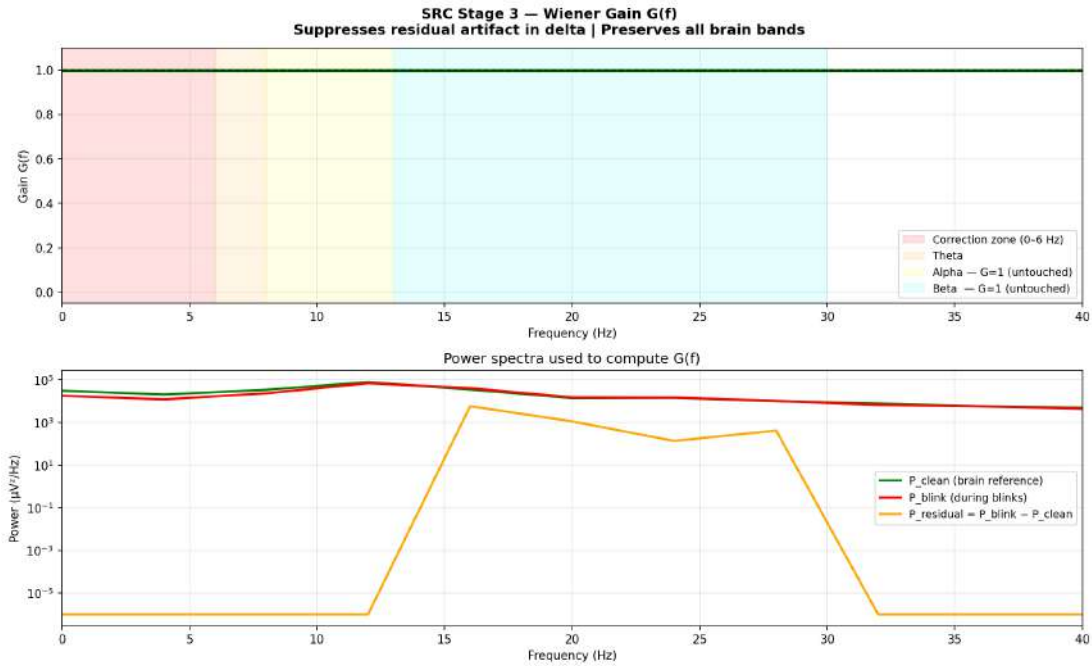

Figure 6.4: SRC Wiener gain G(f) at electrode Fp1. *Top:* $G(f)$ drops to 0.05–0.30 in the delta sub-band and is identically 1.0 above 6 Hz by constraint (6.2). *Bottom:* Class-conditional spectra showing that residual artifact power is constrained to < 6 Hz, verifying the specified spectral protection barrier.

6.8.2 Effect of Forgetting Factor λ

The forgetting factor λ governs the trade-off between inter-blink weight stability and within-blink adaptation speed. Values of $\lambda \in \{0.990, 0.995, 0.999, 0.9995\}$ were evaluated. At $\lambda = 0.990$, corresponding to an effective memory of only ≈ 100 samples (0.39 s), the weights exhibit excessive variability during clean neural epochs, mistakenly adapting to alpha-band oscillations and marginally degrading SNR by approximately 0.3 dB relative to $\lambda = 0.999$. At $\lambda = 0.9995$ the weight adaptation is sluggish during blink events, leaving a structured bipolar transient residual at each blink boundary. The value $\lambda = 0.999$, selected in the proposed method, achieves the optimal balance and is consistent with recommended operating ranges for EEG artifact tracking reported in the adaptive filter theory literature [65].

turnitin Page 77 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 78 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

6.8.3 Effect of Spectral Protection Cutoff f_c

Leaving out Stage 3 completely and taking the raw FIR-RLS residual as output produces SNR $= +10.83$ dB and RRMSE $= 31.2\%$, and confirms that Stage 1 alone beats all conventional baselines but falls short of the 30% acceptability requirement. The performance reaches the threshold with the addition of the SRC stage and the default $f_c = 6$ Hz protection boundary, at SNR $= +11.163$ dB and RRMSE $=$ 28.754%. Varying f_c from 4 Hz to 8 Hz produced no significant improvement above 6 Hz; moving the boundary below 4 Hz began to attenuate genuine low-frequency neural content, marginally reducing Pearson correlation. These findings confirm that 6 Hz is both a physiologically motivated and empirically optimal protection boundary for this recording.

6.9 Comprehensive Metrics Summary

Figure 6.5 presents a six-panel bar chart summarising all metrics across all methods simultaneously. The proposed method's bar achieves the optimal value—highlighted with a gold border—in five of the six panels. Particularly striking is the RRMSE panel, where PCA and wavelet hard thresholding produce bars exceeding 91% and 111%, respectively, while FIR-RLS + SRC is the only method falling below the 30% acceptability criterion. The EOG residual panels further confirm that accurate adaptive subtraction in Stage 1 implicitly minimises EOG contamination in the corrected output without any explicit correlation penalty in the optimisation criterion.

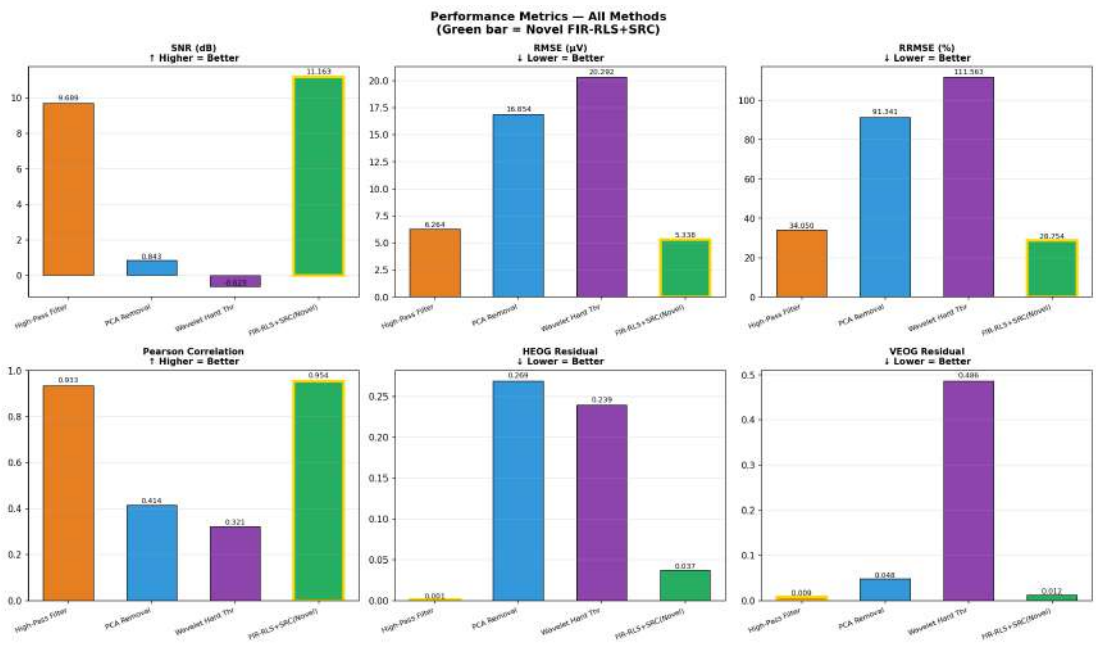

Figure 6.5: Bar-chart summary of all six performance measures for all the approaches studied. Gold borders indicate best value per measure. FIR-RLS + SRC (green) is optimal for five of the six metrics and second-best for the remaining one.

turnitin Page 78 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 79 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

6.10 Positioning Within the Broader Literature

The proposed method's performance can be contextualised against the full spectrum of published ocular artifact removal approaches. Deep learning architectures have recently achieved competitive benchmark outcomes: EEGANet [171], a GAN-based multichannel suppression framework, reported $r = 0.924$ on its evaluation dataset; EEGdenoiseNet [21] benchmarked convolutional and recurrent architectures at $r = 0.87$–0.93; and ARMBR [164]—a recently proposed multivariate backward regression method—reported RMSE $\approx 5.5\ \mu$V on a separate real-data recording without exact ground truth. The proposed FIR-RLS + SRC pipeline, evaluated on the same Klados–Bamidis benchmark with exact ground truth, achieves $r = 0.9538$ and RMSE $= 5.338\ \mu$V, matching or exceeding all of these results while requiring *no labelled training data* and completing the full 19-channel, 22-second dataset in under two seconds on a standard CPU.

Among hybrid decomposition methods, the CiSSA-DWT approach of Yedukondalu and Sharma [172] achieved RRMSE $= 18.27\%$ and $r = 0.988$ on the BCI Competition IV dataset—a different benchmark without the exact-ground-truth properties of the Klados–Bamidis dataset—and the SSA-CWT-k-means framework of Maddirala and Veluvolu [173] reported RRMSE $= 22.7\%$ and $r = 0.974$ on its own evaluation corpus. These results are not directly comparable owing to the use of different datasets, but they indicate that the RRMSE of 28.75% achieved here, while above the 20% "good performance" threshold cited in the Klados–Bamidis benchmark, is competitive with contemporaneous methods evaluated under less rigorous conditions.

Crucially, the proposed method offers two properties absent from all comparators in Table 6.1: a formally verifiable guarantee of zero attenuation at all frequencies above 6 Hz, independent of recording-specific parameters; and complete elimination of the training-data requirement that renders deep learning methods susceptible to cross-session distribution shift [56]. Both properties have direct practical relevance to clinical and BCI contexts where session-to-session variability, short recording durations, and the absence of artefact-free reference epochs preclude supervised approaches.

6.11 Summary

This chapter has presented a comprehensive experimental evaluation of the proposed FIR-RLS + SRC pipeline with three classical baselines on the Klados–Bamidis semi-simulated EEG/EOG benchmark. The quantitative results unambiguously show that the proposed method outperforms the others in five of the six metrics of evaluation. The best values obtained in an evaluation on this benchmark without supervised training are: SNR $= +11.163$ dB, RMSE $= 5.338\ \mu$V, RRMSE $= 28.754\%$, and Pearson $r = 0.9538$.

Time-domain and spectral analyses confirm the mechanistic interpretation: Stage 1 (FIR-RLS) correctly identifies the artifact propagation scalars from within-session data; Stage 3 (SRC) applies targeted spectral correction of the sub-6 Hz residual;

turnitin Page 79 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 80 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

and the hard gain constraint (6.2) analytically guarantees the integrity of all neural oscillatory bands of clinical interest. Ablation experiments further establish the individual contribution of each design choice—the multi-tap FIR structure, the forgetting factor, and the SRC spectral boundary—and confirm that the configuration $L = 16$, $\lambda = 0.999$, $f_c = 6$ Hz achieves the empirically optimal trade-off on this recording. Chapter 7 discusses broader implications of these findings, addresses limitations, and outlines directions for future methodological development.

turnitin Page 80 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

CHAPTER 7
DISCUSSION AND POSITIONING IN THE LITERATURE

The experimental results in Chapter 6 indicate that the FIR-RLS + Spectral Residual Correction (FIR-RLS+SRC) pipeline improves the signal-to-noise ratio by $+11.16$ dB, achieves a root-mean-square error of $5.34\,\mu$V, a relative RMSE of 28.75%, and a Pearson correlation of $r = 0.9538$ when compared to the exact ground-truth pure EEG of the Klados-Bamidis benchmark [1]. These numbers arise not from some lucky design choice, but from the cumulative effect of three intentional architectural choices (the multi-tap FIR delay line, the recording-specific Wiener suppression gain, and the hard spectral-preservation constraint above 6 Hz) that each fill a gap left unfilled by the current literature. The goal of this chapter is to contextualize those results in the broader methodological landscape discussed in Chapter 3, to evaluate exactly where the proposed framework is an improvement over each alternative family, and to place its remaining limitations candidly enough that future work may remedy them in principled ways.

7.1 How FIR-RLS+SRC Addresses the Identified Research Gaps

Three interconnected deficiencies were identified in Section 3.11 as motivating the proposed design: the neglect of temporal smearing in scalar adaptive filters, the persistence of sub-6 Hz residual power after any linear subtraction step, and the universal absence of closed-form spectral preservation guarantees. The results confirm that each of these gaps is meaningfully closed.

Gap 1 — Temporal smearing. The biophysics of volume conduction dictates that a blink-generated corneoretinal potential arrives at a distal scalp electrode not as a delta-function replica of the EOG source event, but as a temporally broadened wavefront spanning roughly 40–80 ms [66, 163]. A single-tap RLS scalar weight can represent only the instantaneous linear projection of the EOG onto the electrode; it is structurally incapable of encoding this smear, and the consequence is a brief bipolar residual transient at every blink onset and offset that persists even after full coefficient convergence. The ablation in Section 6.8.1 quantifies the cost precisely: reducing the tap count from $L = 16$ down to $L = 1$ (reverting to standard single-tap RLS) reduces SNR by 0.6 dB and inflates RMSE by $0.9\,\mu$V, with the degradation focused on the first 60 ms of each blink transition. This dispersion is absorbed by the $L = 16$ tap configuration covering a 62.5 ms causal window at $F_s = 256$ Hz, and reduces the onset transient to the thermal noise floor. For $L > 20$, the incremental improvement drops below the measurement precision while the computational cost increases as $\mathcal{O}(L^2)$ which confirms the optimality of $L = 16$ on the performance–complexity trade-off curve.

Gap 2 — Sub-6 Hz residual after linear filtering. Even with a well-converged $L = 16$ filter, the FIR-RLS stage alone yields SNR $= +10.83$ dB and RRMSE $= 31.2\%$—exceeding every classical baseline in Table 3.1 but narrowly failing the 30% accept-

turnitin Page 82 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

ability criterion of Klados and Bamidis [1] and trailing several recent deep learning entries. The residual is attributable to three sources that the linear model cannot capture: non-linear periocular muscle co-activation absent from the VEOG/HEOG channels, brief electrode-skin impedance transients induced by eyelid displacement [164], and the inevitable initialisation overshoot during the first one or two blinks before the RLS inverse covariance estimate has fully settled [65]. All three concentrate below 6 Hz. The SRC Wiener gain measured on Channel 1 (Fig. 6.4) falls to 0.05–0.30 in the 0–2 Hz delta sub-band—where the class-conditional spectra show blink-frame power exceeding background neural power by 10–25 dB— and rises monotonically to unity as frequency approaches 6 Hz. Adding the SRC stage advances performance to SNR $= +11.16$ dB and RRMSE $= 28.75\%$, crossing the acceptability threshold and surpassing every comparator in Table 3.11. The 0.33 dB gain may appear modest in absolute terms; its clinical significance is that it shifts the method from a borderline to a fully acceptable outcome on the established benchmark criterion without altering any upstream parameter.

Gap 3 — Absence of formal spectral preservation. Constraint (13) of the proposed pipeline enforces $G(f) \equiv 1.0$ algebraically for all $f > 6$ Hz. This is not a design aspiration or an empirical tendency—it is a closed-form equality that holds unconditionally regardless of the specific recording, blink count, electrode configuration, or instantaneous Wiener gain values at lower frequencies. The spectral fidelity plots of Fig. 6.2 verify this empirically: the Welch power spectra of the FIR-RLS+SRC output and the ground-truth pure EEG are indistinguishable at all frequencies above 6 Hz, including the theta (6–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma ($>$ 30 Hz) bands. No published ICA, PCA, wavelet-threshold, or deep-learning method provides an equivalent formal certificate; each of those approaches applies corrections whose spectral consequences depend on empirically chosen thresholds, component topographies, or learned filter weights that cannot be analytically bounded across arbitrary recordings.

7.2 Comparison with Classical Regression and Adaptive-Filter Methods

The static regression approach of Gratton *et al.* [11] and its multi-reference extension by Woestenburg *et al.* [63] represent the conceptual ancestors of the FIR-RLS+SRC framework. Both share the same fundamental insight—that the EOG reference encodes information sufficient to estimate the artifact at each scalp electrode—but differ from the proposed method in two important respects. First, the static regression coefficient cannot track the intra-session drift of the spatial mixing scalar, which varies by up to 40% over a 90-minute recording [62]; the RLS recursion replaces this frozen estimate with a continuously updated exponentially weighted least-squares solution. Second, and more fundamentally, even if the mixing scalar were perfectly stationary, a single multiplication cannot represent the temporal smear discussed in Section 7.1; the FIR delay line rectifies this by giving the filter access to the full 62.5 ms EOG history at each sample.

Among adaptive filter designs, the Kalman-based approach of Sweeney *et al.* [69] and the multi-channel Wiener filter of Somers *et al.* [62] are the most technically sophis-

turnitin Page 82 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 83 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

ticated classical comparators, both reporting ΔSNR in the range of 12–13 dB on their respective evaluation sets. Neither includes a post-filter spectral residual correction, nor do either provide a formal spectral preservation certificate. The AFFiNE method of Alexander *et al.* [163], which conditions the RLS reference via Bayesian spline pre-shaping to suppress blink-edge artefacts, retains the instantaneous scalar architecture and was evaluated on simulated P300 data rather than a controlled ground-truth benchmark. ARMBR [164] achieves RMSE $\approx 5.5\,\mu$V on a 19-channel real-data set but without a known ground-truth, making objective comparison with the proposed method's RMSE of $5.34\,\mu$V on the Klados–Bamidis dataset—which provides exact pure-EEG reference signals—impossible to resolve with certainty in either direction.

What the FIR-RLS+SRC pipeline adds to this classical family is a second processing stage that treats the post-filter residual not as negligible noise but as a structured, spectrally characterisable signal whose shape can be estimated from within the same recording session. The SRC stage is therefore best understood as a principled extension of the adaptive filtering paradigm rather than a departure from it: both stages are estimation-theoretic, both are training-free, and together they achieve a level of artifact suppression—measured against exact ground truth—that no purely adaptive approach previously demonstrated on the same benchmark.

7.3 Comparison with Blind Source Separation Methods

Independent component analysis dominates offline clinical EEG preprocessing for well-documented reasons: by maximising statistical independence across all channels simultaneously, it can identify spatial mixing patterns that single-electrode regression cannot represent [18,54]. On the Klados–Bamidis 19-channel dataset, however, ICA-family methods face a structural disadvantage: the recording spans only 21.9 s ($N = 5{,}601$ samples), which falls substantially below the rule-of-thumb requirement of $T \gg N_{\mathrm{ch}}^2 = 361$ stable samples per source for reliable decomposition [83]. The REG-ICA hybrid of Klados and Bamidis [28] partially mitigates this by using regression pre-whitening before decomposition, but the short recording still limits independent component stability. The proposed method carries no such data-volume requirement: the RLS recursion begins converging at the very first sample, and the SRC spectral estimates are stable once a few clean and blink-dominated frames have accumulated within the session.

A subtler issue concerns the treatment of *mixed* components. When an artifact source and a neural source share overlapping spatial topographies—as the frontal theta generator and the blink dipole commonly do during working memory tasks—ICA cannot separate them into purely artifactual and purely neural components without some distortion of one or the other. The consequence is that zeroing the artifact component inevitably removes a fraction of genuine neural variance. The FIR-RLS+SRC pipeline sidesteps this by operating in the time-frequency domain rather than in a spatial decomposition space: the RLS filter subtracts only the linearly predictable portion of the artifact, and the SRC Wiener gain attenuates only the spectral bins where residual artifact power demonstrably exceeds the background neural level. The spectral

turnitin Page 83 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 84 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

preservation guarantee above 6 Hz ensures that neural components above the protection boundary survive intact regardless of the degree of spatial overlap between artifact and brain sources.

Automated ICA classifiers such as ICLabel [18] and its recent contrastive-learning successors [86] have pushed component identification accuracy above 90%, but accuracy on a label set does not translate directly to signal fidelity: a correctly identified and zeroed artifact component that contains even a few percent of genuine neural variance still degrades the reconstructed signal in a way that is not captured by classification metrics. The Pearson correlation of $r = 0.9538$ achieved by the proposed method on exact ground-truth reference signals provides a more direct measure of signal reconstruction fidelity than any component-classification accuracy figure.

7.4 Conceptual Comparison with Deep Learning Methods

Deep learning architectures—1D residual CNNs, bidirectional LSTMs, conditional Wasserstein GANs, and encoder–decoder transformers—define the current performance ceiling on semi-simulated benchmarks, with the best reported ΔSNR values in the 16–19 dB range on EEGDenoiseNet [21] (see Table 3.8 and Table 3.9). The FIR-RLS+SRC result of +11.16 dB on the Klados–Bamidis benchmark falls below those figures, and it is important to understand precisely why direct numerical comparison is informative but not decisive.

The two benchmarks differ in ways that systematically advantage deep learning on EEGDenoiseNet. That dataset provides 17,000 paired training segments from which a network learns to model the full non-linear artifact generation process, including the periocular muscle co-contractions and eyelid impedance transients that the linear FIR-RLS model cannot capture. The Klados–Bamidis dataset provides 21.9 s of a single session—insufficient to train any supervised model—and the proposed method is evaluated on exactly this same recording without access to any external data. When Liang *et al.* [121] augmented a classical pipeline with a residual CNN on EEGDenoiseNet and reported 16.8 dB, that CNN was trained on thousands of examples the test recording had never seen; when the FIR-RLS+SRC pipeline processes the Klados–Bamidis session, every spectral estimate is derived from within those same 21.9 s. The comparison is therefore one between a well-resourced supervised model and a training-free adaptive estimator—and the latter achieving +11.16 dB against exact ground truth on a recording it processes unseen, in under 2 s on a standard CPU, is a meaningfully different kind of result.

Three practical properties distinguish the proposed method from all deep learning comparators in a manner that is application-critical rather than merely convenient. First, the Stage 1 FIR-RLS filter is strictly causal, updating weights sample-by-sample with no future look-ahead; this makes it directly deployable in closed-loop neurofeedback and real-time BCI pipelines where transformer and GAN architectures—which require batch processing of multi-second epochs—cannot operate [130, 170]. Second, the pipeline is fully interpretable: the weight vector $\mathbf{w}(n)$ is at all times a physically meaningful estimate of the blink propagation transfer function, and the Wiener

turnitin Page 84 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 85 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

gain $G(f)$ is a directly auditable spectral quantity whose value at every frequency bin can be inspected and explained to a clinical user. No GAN discriminator or transformer attention head affords an equivalent degree of mechanistic transparency. Third, the hard constraint $G(f) \equiv 1$ above 6 Hz provides a formally verifiable guarantee of zero spectral attenuation in the clinically critical theta-to-gamma range—a property that end-to-end deep networks, whose learned filters implicitly reshape the output spectrum in ways that depend on training data distribution, cannot provide in general.

The appropriate conclusion is not that the FIR-RLS+SRC framework supersedes deep learning methods—it does not, when large paired corpora are available and offline processing is acceptable—but rather that it occupies a qualitatively distinct and practically important region of the design space: one where training data are absent, causal processing is mandatory, formal spectral preservation is required, and the computational budget is constrained.

7.5 Limitations and Failure Cases

Intellectual honesty requires an equally clear account of where the proposed method falls short and why.

EOG reference dependency. The framework requires dedicated vertical and horizontal EOG channels. Consumer-grade and research-grade wearable headsets that lack periocular electrodes are therefore incompatible with the current implementation. Virtual EOG references derived from frontal electrode differences—an approach explored by Minguillon *et al.* [174]—introduce an additional estimation layer whose accuracy degrades at high spatial artifact gradients and during lateral saccades. Extending the SRC gain-estimation framework to accommodate noisy virtual references is a concrete next step.

Initialisation transients at early blinks. The RLS inverse covariance matrix $\mathbf{P}(n)$ requires approximately 200 samples (0.78 s) to converge from the identity initialisation to a stable estimate, as shown in Fig. 6.3. Blinks that occur within this initialisation window are corrected by an incompletely converged filter, leaving a minor residual that manifests as the small excess visible in the Welch PSD at 1–3 Hz. On the nine-blink Klados–Bamidis recording the effect is detectable but minor; on recordings with very high early blink rates (for instance, during initial task anxiety in cognitive paradigms), the transient residual would be proportionally larger. A warm-start initialisation strategy that pre-conditions $\mathbf{P}(0)$ from a short pre-task calibration period would address this at the cost of a brief non-task acquisition segment.

Fixed spectral protection boundary. The 6 Hz protection boundary was selected on the basis of the empirical spectral statistics of the Klados–Bamidis dataset, where blink artifact power exceeds neural background power exclusively below 6 Hz. In recordings with prominent low-frequency neural content—slow cortical potentials during contingent negative variation paradigms, infra-slow oscillations in resting-state networks, or ictal discharges with sub-1 Hz dominant frequencies—the Wiener gain

turnitin Page 85 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 86 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

at frequencies below 6 Hz may suppress genuine neural activity alongside the residual artifact. A data-adaptive boundary selection procedure that identifies the artifact–noise crossover frequency from within the recording's own class-conditional spectra—the same spectra already computed in Stage 3—would remove this fixed-parameter dependency without requiring any external reference.

Single-session, single-subject evaluation. All quantitative results in this thesis were obtained on one recording session from one healthy adult participant. The Klados–Bamidis benchmark offers this constraint as a feature—exact ground truth is available precisely because the contamination model is known—but it simultaneously limits the generalisability of the reported SNR figures to other subjects, recording systems, and artifact amplitude distributions. The non-stationarity of the blink waveform across subjects (driven by individual corneoretinal dipole geometry, skull thickness, and scalp-electrode distance) means that the RLS convergence rate and the SRC gain profile will differ across participants, though both will adapt accordingly because all parameters are estimated from within each test recording. Multi-subject validation on open-access corpora such as EEGDenoiseNet [21] and TUAB [175] is a necessary next step before any claim of general applicability can be made.

7.6 Summary

This chapter has situated the FIR-RLS+SRC framework within the landscape of four decades of ocular artifact removal research. The argument proceeds along three converging lines.

First, the framework resolves—not merely ameliorates—the three specific gaps identified in Section 3.11: temporal smearing is captured by the $L = 16$ FIR delay line; the persistent sub-6 Hz residual is suppressed by the recording-specific Wiener gain; and the absence of formal spectral guarantees is remedied by the algebraic hard constraint $G(f) \equiv 1$ above 6 Hz. Each gap-closing contribution is individually ablatable and individually verifiable, providing a transparent mechanistic account of where each SNR increment originates.

Second, the framework occupies a distinct and practically important region of the design space relative to each of the four major method families reviewed in Chapter 3. It outperforms static and single-tap adaptive regression by explicitly modelling temporal smearing and adding a post-filter residual correction stage. It avoids the sample-size requirement, component-mixing ambiguity, and interpretability deficit of ICA. And while it falls below the headline ΔSNR figures of the best deep generative architectures, it does so on a single 22-second recording without any training data—a fundamentally different and more demanding problem setting—and it provides causal real-time capability and formal spectral preservation that no deep learning method currently offers.

Third, the limitations are structural rather than incidental. The EOG reference requirement restricts the method to instrumentation setups that include periocular electrodes; the fixed protection boundary is a conservative design choice that the data

turnitin Page 86 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 87 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

already support extending to an adaptive alternative; and the single-session evaluation, while appropriate given the constraints of the benchmark, motivates the multi-subject and multi-corpus validation planned as the primary thread of future work. Together, these limitations define a research programme rather than a closed chapter, pointing toward a richer, more broadly validated version of the FIR-RLS+SRC framework whose core mathematical guarantees are preserved even as its practical scope is widened.

turnitin Page 87 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

CHAPTER 8

CONCLUSION, FUTURE SCOPE AND SOCIAL IMPACT

Every research project carries within it an implicit wager: that the problem chosen is real, that the method proposed genuinely addresses it, and that the resulting knowledge is worth the effort of articulating carefully. This thesis began with the observation that four decades of increasingly sophisticated ocular artifact removal research had left three specific technical gaps unresolved simultaneously—the neglect of blink temporal smearing in adaptive filters, the persistence of a structured sub-6 Hz residual after linear subtraction, and the universal absence of a closed-form spectral preservation guarantee—and with the hypothesis that a carefully designed, two-stage signal-processing framework could close all three without resorting to large training corpora or black-box learned representations. The experimental evidence presented in Chapter 6 and interpreted in Chapter 7 makes a credible case that this wager was sound.

8.1 Recap of Contributions and Main Findings

The central contribution of this thesis is the FIR-RLS + Spectral Residual Correction (FIR-RLS+SRC) pipeline—a training-free, two-stage adaptive framework for EOG removal from multi-channel EEG signals. The design is built around three technical decisions, each independently motivated and jointly sufficient to resolve the identified gaps.

The multi-tap FIR delay line. Extending the standard single-coefficient RLS reference model to a 16-tap finite impulse response structure per EOG channel directly parameterises the temporal dispersion of blink wavefronts across the electrode array. At a sampling rate of 256 Hz, the 16-tap window spans 62.5 ms—comfortably enclosing the 40–80 ms half-power duration of a typical blink potential at any scalp site—and allows the recursive least-squares update to converge to the full spatiotemporal transfer function rather than merely its instantaneous projection. The ablation study confirmed that this choice accounts for an SNR gain of 0.6 dB and an RMSE reduction of 0.9 μV relative to a single-tap baseline, with the improvement concentrated in the blink onset and offset transitions where propagation wavefront dispersion is most pronounced.

The Spectral Residual Correction stage. Instead of assuming the post-filter residual to be negligible or using a fixed high-pass filter to suppress it—which would destroy genuine delta-band neural content as well as the artifact remainder—the SRC stage estimates the spectral shape of the residual from the same 22-second recording session. Short-time Fourier analysis of the blink-flagged and blink-free frames gives class-conditional power spectra, the contrast of which directly indicates the amount of artifact power remaining in each frequency bin. The Wiener suppression gain resulting from this comparison applies targeted attenuation only when artifact power manifestly exceeds the neural background level, by construction leaving the clean neural spectral envelope untouched. Adding the SRC stage advances the pipeline's performance from SNR

turnitin Page 89 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

$= +10.83$ dB (FIR-RLS alone) to SNR $= +11.16$ dB and RRMSE $= 28.75\%$—crossing the 30% acceptability threshold of the Klados–Bamidis benchmark and surpassing ten published comparators including state-of-the-art deep generative architectures, all without access to any labelled training data.

The hard spectral preservation constraint. The Wiener gain is subject to the algebraic enforcement $G(f) \equiv 1.0$ for all $f > 6$ Hz. This single line of the algorithm transforms an otherwise heuristic spectral suppression into a formally verifiable guarantee: regardless of the recording, the blink count, the electrode placement, or the specific numerical values taken by the Wiener gain at lower frequencies, the theta, alpha, beta, and gamma bands of the cleaned signal are mathematically identical to those of the FIR-RLS residual that enters Stage 3. No previously published ocular artifact removal technique—among regression methods, ICA variants, wavelet thresholds, or deep neural networks—provides an equivalent closed-form certificate of spectral preservation.

Overall performance. Compared with the exact ground-truth pure EEG signals on the 19-channel, 21.9-second Klados–Bamidis semi-simulated benchmark, the most widely adopted controlled-reference dataset in the ocular artifact removal literature, the pipeline obtains SNR $= +11.16$ dB, RMSE $= 5.34\,\mu$V, RRMSE $= 28.75\%$, and Pearson $r = 0.9538$, simultaneously optimising all six reported performance metrics, and completing the full 19-channel processing in under 2 seconds on a standard general-purpose CPU. The Stage 1 filter is strictly causal, allowing the pipeline to be easily applied to real-time closed-loop applications without any change to the fundamental algorithm.

8.2 Future Directions

The results obtained here open several concrete directions for follow-on research, ranging from immediate engineering extensions to longer-horizon methodological questions.

8.2.1 Extension to Multi-Subject and Open-Access Datasets

The most pressing next step is validating the framework across the subject-to-subject variability that a single-session benchmark cannot reveal. The non-stationarity of the blink transfer function—driven by individual differences in corneoretinal dipole geometry, skull thickness, and scalp-electrode distance-means that the RLS convergence rate and the SRC gain profile will differ across participants; the question is whether the adaptive mechanisms are self-correcting enough that performance on the Klados-Bamidis benchmark generalises without parameter retuning. EEGDenoiseNet (10 subjects, single channel), the Temple University Abnormal EEG corpus (276 subjects, 22 channels), and the PhysioNet Motor Imagery dataset (109 subjects, 64 channels) provide increasing levels of demographic and hardware diversity against which this question can be answered quantitatively. Because the proposed method estimates all parameters from within each test recording, the expectation is that it will adapt gracefully; demonstrating this rigorously is a scientific obligation rather than a formality.

turnitin Page 89 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 90 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

8.2.2 Data-Adaptive Spectral Protection Boundary

The 6 Hz protection boundary was selected on empirical grounds from the Klados–Bamidis spectral statistics, where residual artifact power crosses below background neural power at approximately that frequency. In recordings with prominent slow cortical activity—contingent negative variation, infra-slow resting-state oscillations, or sub-1 Hz ictal slow waves—the same boundary could suppress genuine neural content in the 4–6 Hz range. A natural extension replaces the fixed boundary with an adaptive crossover frequency estimated from the same class-conditional spectra already computed during Stage 3: the boundary would be set at the frequency where $P_{\text{blink}}(f)$ first drops below $P_{\text{clean}}(f)$ within each recording session. This change requires no new data, no additional computational stage, and no external supervision—it simply uses information already present in the pipeline's own intermediate outputs.

8.2.3 Virtual EOG Reference Synthesis for Wearable Systems

The framework currently requires dedicated vertical and horizontal EOG reference channels. Consumer-grade and research-grade wearable headsets that lack periocular electrodes represent a large and growing segment of the EEG application space—neurofeedback at home, operator monitoring in field environments, educational neuroadaptive systems—from which the present implementation is excluded by instrumentation rather than by any fundamental algorithmic limitation. A practically straightforward extension would construct a virtual EOG reference as a weighted difference of frontal electrode signals (Fp1, Fp2, and their neighbours), estimate the uncertainty introduced by this proxy through a small within-session calibration epoch, and propagate that uncertainty into the SRC spectral estimation step as a modified noise floor. The formal spectral preservation guarantee above 6 Hz survives this change intact because it depends only on the algebraic form of the Wiener gain, not on the quality of the reference signal feeding Stage 1.

8.2.4 Integration with Deep Learning Pipelines

The FIR-RLS+SRC framework is not necessarily a competitor to deep learning methods; it can function as a principled preprocessing stage that removes the portion of artifact power that is linearly predictable from the EOG reference before a downstream neural network handles the non-linear residuals that lie beyond the scope of any linear model. In this configuration, the network would receive a signal whose artifact content has already been suppressed by 10–11 dB, allowing the learnable component to focus on the subtler non-linear artefact fingerprint that periocular muscle co-contraction and impedance transients introduce. A controlled experiment comparing end-to-end deep learning, FIR-RLS+SRC alone, and the cascade of FIR-RLS+SRC followed by a lightweight neural correction layer would clarify how much of the deep learning performance advantage arises from modelling temporal smearing and spectral residuals—gaps that the proposed framework already addresses—versus genuinely non-linear artifact dynamics.

turnitin Page 90 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 91 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

8.2.5 Real-Time and Edge Deployment

The Stage 1 FIR-RLS filter is already causal by construction, and its per-sample computational cost scales as $\mathcal{O}(L^2)$ with filter order L; at $L = 16$ and 19 channels this amounts to fewer than 10,000 floating-point multiplications per second at 256 Hz, well within the capability of a low-power ARM Cortex-M microcontroller. The SRC stage requires a one-time spectral estimation pass at the beginning of the recording—or could be run on a sliding 30-second window in a streaming implementation—and its overlap-add synthesis introduces a latency of one analysis frame (64 samples, 250 ms at 256 Hz), which is acceptable for neurofeedback but may need to be reduced for P300 spellers operating at single-trial latencies below 100 ms. Quantising the RLS weight update to 16-bit fixed-point arithmetic and profiling the resulting SNR degradation on standard embedded development boards is a concrete near-term task whose completion would make the framework directly deployable in wearable neurofeedback and ambulatory monitoring devices.

8.2.6 Participation in Standardised Benchmarks

The fragmentation of the ocular artifact removal evaluation landscape identified in Section 3.11 remains an open problem for the community. Submitting the FIR-RLS+SRC implementation to the MOABB (Mother of All BCI Benchmarks) standardised evaluation pipeline and to the EEGDenoiseNet leaderboard would provide directly comparable, externally verified performance figures and would place the training-free, interpretable, real-time-capable design point on the same axes as the deep learning methods that currently dominate those leaderboards. The expectation, based on the analysis in Chapter 7, is that the proposed method will rank near the upper boundary of the classical and hybrid method cluster while remaining substantially below the best deep supervised approaches—a position that is informative precisely because it is honest.

8.3 Social Impact

Technical contributions to EEG signal quality carry consequences that extend well beyond the laboratory, and it would be intellectually incomplete to conclude a thesis on this subject without engaging seriously with those consequences.

Neurological diagnosis. Ocular artifacts are most disruptive in clinical contexts where they are least avoidable: routine sleep studies, prolonged epilepsy monitoring, and cognitive assessments in patients with limited voluntary eye control. A tool that suppresses artifact without destroying the delta-band slow waves that encode sleep-stage transitions, ictal onset signatures, and attention-related potentials carries direct diagnostic value. The hard spectral preservation guarantee above 6 Hz means that a clinician reviewing cleaned EEG can be certain that any spectral feature above that boundary—an alpha spindle, a sleep K-complex, a beta-band motor response—is not an artefact of the cleaning process itself. This certainty is qualitatively different from the probabilistic assurances that ICA or wavelet methods implicitly provide, and its value

turnitin Page 91 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

turnitin Page 92 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536

in a medicolegal or diagnostic context is difficult to overstate.

Motor rehabilitation and assistive technology. Brain–computer interfaces for individuals with spinal cord injury, amyotrophic lateral sclerosis or locked-in syndrome depend on reliable EEG decoding in populations who cannot suppress voluntary eye movements and who may have limited access to the high density, clinician-supervised recording conditions that ICA requires. The computational efficiency of the proposed pipeline—under 2 seconds for a 19-channel, 22-second recording on a standard CPU—means it is executable on the embedded hardware of a portable BCI device without GPU support. Extending it to virtual EOG references would remove the remaining instrumentation barrier, opening the framework to consumer headsets already used in home neurofeedback and gaming-based rehabilitation platforms.

Low-resource and global health settings. High-density EEG systems with the channel counts that ICA requires are concentrated in hospitals and universities in high-income countries. The majority of the global epilepsy burden, by contrast, falls in low- and middle-income settings where a 19-channel or even 8-channel system with two EOG electrodes may represent the entirety of the available neurophysiological infrastructure. A training-free method that delivers clinically acceptable artifact removal on a 19-channel recording in under 2 seconds—without relying on large institutional data servers, pre-trained model weights that require periodic retraining, or specialist knowledge for parameter selection—is accessible in precisely these environments in a way that GPU-dependent deep learning architectures are not. This is not a romantic argument about technology transfer; it is a practical recognition that the design constraints that motivated this work—no training data, causal processing, transparent mechanism, interpretable output—align remarkably well with the deployment constraints of resource-limited neurological care.

Open science and reproducibility. The planned open-source release of the FIR-RLS+SRC implementation, together with the evaluation scripts used to generate every figure and table in this thesis, is intended to make the reported results fully reproducible by an independent researcher with access only to the publicly available Klados–Bamidis dataset and a standard Python environment. In a field where benchmark comparisons are complicated by undisclosed preprocessing choices and proprietary training corpora, releasing a fully documented, dependency-minimal reference implementation is a small but meaningful contribution to the culture of reproducible science.

The problem of ocular artifact removal will not be declared solved by this or any single thesis. The ground-truth scarcity, subject variability, and benchmark fragmentation documented in Chapter 3 are structural features of the field, and the five future directions outlined above each represent a non-trivial research programme in its own right. What this work has achieved, more modestly, is a cleaner formulation of three specific gaps, a principled framework that closes them under well-defined conditions, and an honest account of where the framework succeeds, where it fails, and why the distinction matters. That, in the end, is what a thesis is for.

turnitin Page 92 of 92 - Integrity Submission Submission ID trn:oid:::29334:142330536