Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Transient Burst Dynamics in Resource-Constrained EEG Decoding: A Systematic Review of Single-Channel Adaptive Deep Learning for Wearable BCIs

McKenzie Izev

mckenzieizev@gmail.com

AIT Austrian Institute of Technology

Systematic Review

Keywords: single-channel EEG, transient burst dynamics, gamma bursts, adaptive deep learning, domain adaptation, wearable brain–computer interface, resource-constrained edge AI, systematic review, PRISMA

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11740189/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Transient Burst Dynamics in Resource-Constrained EEG Decoding: A Systematic Review of Single-Channel Adaptive Deep Learning for Wearable BCIs

Mckenzie Izev

Technical University of Denmark, DK https://orcid.org/0009-0006-8450-4944

Abstract

Single-channel electroencephalography (EEG) is increasingly favored for wearable brain–computer interfaces (BCIs) due to reduced hardware complexity and power consumption. However, single-channel decoding faces limited spatial information, cross-session non-stationarity, and severe resource constraints on embedded platforms. Transient burst dynamics—brief oscillatory events such as gamma bursts and ERD/ERS—have emerged as neurophysiologically meaningful features that may compensate for the loss of spatial resolution. Yet, the literature on burst-centric adaptive deep learning for single-channel EEG remains fragmented, with no systematic synthesis of methods, adaptation strategies, and reported performance. This systematic review will follow PRISMA 2020 guidelines, searching IEEE Xplore, PubMed, SpringerLink, and ScienceDirect for peer-reviewed studies published between 2018 and 2026. The search strategy will combine terms related to single-channel EEG, transient oscillatory features, adaptive deep learning, and wearable deployment. Two reviewers will independently perform screening, data extraction, and quality assessment. Extracted data will include channel configuration, feature type, adaptation method, dataset, performance metrics, computational cost, and limitations. Findings will be synthesized narratively and, where homogeneity permits, quantitatively. This review will provide the first systematic synthesis of burst-centric single-channel adaptive decoding as a distinct technical route, clarifying validated features and adaptation strategies, identifying generalization gaps, mapping computational constraints, and deriving a research agenda for resource-efficient wearable BCIs.

Keywords: single-channel EEG; transient burst dynamics; gamma bursts; adaptive deep learning; domain adaptation; wearable brain–computer interface; resource-constrained edge AI; systematic review; PRISMA

1. . Introduction

The proliferation of low-cost, consumer-grade electroencephalography (EEG) hardware has fundamentally altered the landscape of brain–computer interface (BCI) research. Where once the field was dominated by laboratory-grade amplifier systems with dozens or even hundreds of electrodes, the past decade has witnessed a decisive shift toward portability, wearability, and ecological validity. Devices such as the NeuroSky MindWave, the Muse headband, and the

OpenBCI Ganglion have made it possible to record neural activity outside the confines of the laboratory, embedded within headbands, earbuds, and other form factors compatible with everyday use (Suppiah & Kim, 2026a). This hardware revolution has been accompanied by an algorithmic one: the migration of deep learning techniques from the server to the edge, driven by advances in model compression, quantization, and lightweight neural architecture design (Suppiah et al., 2026). Together, these developments have opened a credible path toward BCIs that are not merely laboratory curiosities but practical tools for communication, rehabilitation, and human–machine interaction.

Yet the transition from multi-channel laboratory systems to single-channel wearable devices is not merely a matter of hardware miniaturization. It entails a fundamental loss of information. A multi-channel montage exploits spatial filtering to isolate task-relevant cortical sources, suppress artifacts, and construct covariance-based features that have underpinned the success of methods such as Common Spatial Patterns and its variants (Jiang et al., 2026). When only a single electrode is available, this spatial dimension collapses entirely. The decoder must extract discriminative information from a one-dimensional time series, often recorded over a sensorimotor region such as Cz or a frontal site such as Fp1, with limited ability to separate neural signal from ocular, muscular, and movement-related contamination (Tao et al., 2021). The consequence is a well-documented performance gap: single-channel decoders consistently underperform their multi-channel counterparts across tasks including motor imagery, emotion recognition, and cognitive workload estimation (Suppiah & Kim, 2026b). This gap is not merely an engineering inconvenience; it represents a fundamental constraint on the types of applications that wearable BCIs can credibly support.

The response to this constraint has taken several forms. One line of inquiry has focused on sophisticated temporal filtering and handcrafted feature engineering, attempting to compensate for the absence of spatial information through careful exploitation of the time–frequency structure of single-channel signals (Tao et al., 2021). Another has pursued increasingly complex deep learning architectures, leveraging convolutional and recurrent networks to learn representations directly from raw or minimally processed data (Jiang et al., 2026). Yet a third, and arguably more promising, direction has emerged from neuroscience itself: the recognition that cortical activity is not continuously oscillatory in the classical sense but rather organized into transient, burst-like events whose temporal and morphological properties carry information that conventional band-power averaging obscures (Rayson & Moreau, 2026). Beta bursts, gamma bursts, and transient event-related desynchronization/synchronization patterns have been shown to index motor preparation, execution, and cognitive state with a temporal precision that sustained oscillatory power cannot match (Suppiah & Kim, 2026a). If these burst dynamics can be reliably extracted from single-channel recordings and integrated into adaptive decoding frameworks, they may offer a neurophysiologically grounded route to closing the performance gap that has thus far limited wearable BCIs.

This review is motivated by the convergence of three strands: the hardware-driven turn toward single-channel EEG, the algorithmic turn toward adaptive deep learning for non-stationary signals, and the neuroscientific turn toward transient burst dynamics as functional units of cortical computation. Each strand has generated a substantial literature, yet they have

developed largely in parallel. Reviews of single-channel EEG have tended to focus on clinical applications such as seizure detection or sleep staging, treating motor decoding as a secondary concern (Zaman et al., 2025). Reviews of deep learning for motor imagery have concentrated on multi-channel benchmarks such as BCI Competition IV 2a and 2b, where spatial information remains available (Jiang et al., 2026). Reviews of burst dynamics have remained predominantly within the neuroscience literature, with limited engagement with the engineering constraints of real-time decoding on embedded hardware (Rayson & Moreau, 2026). The result is a fragmented evidence base in which the promise of burst-centric single-channel decoding is frequently asserted but rarely synthesized. No systematic review has yet examined the intersection of these domains: the use of transient burst features, adaptive deep learning, and resource-constrained deployment as an integrated technical route for wearable BCIs.

The present review addresses this gap. Its objectives are fourfold. First, it aims to systematically identify and characterize the body of peer-reviewed research that has applied transient burst features—including gamma bursts, beta bursts, and fine-grained ERD/ERS dynamics—to single-channel EEG decoding, with particular attention to motor and cognitive state classification. Second, it seeks to map the adaptation strategies that have been employed to address the pronounced non-stationarity of single-channel EEG across sessions and subjects, including adaptive normalization, domain adaptation, meta-learning, and few-shot recalibration, and to assess the evidence for their efficacy under single-channel constraints (Wu et al., 2021; Suppiah & Kim, 2026a). Third, it aims to evaluate the computational characteristics of reported methods, including model size, inference latency, and suitability for deployment on microcontrollers and other resource-constrained platforms, thereby connecting algorithmic design to the practical requirements of wearable systems (Suppiah et al., 2026). Fourth, it aims to synthesize the reported performance of burst-centric adaptive approaches against appropriate baselines and to identify the methodological, reporting, and generalization gaps that constrain the interpretability of the current evidence.

The significance of this review lies in its integrative ambition. By treating burst-centric single-channel adaptive decoding as a coherent, if nascent, technical program rather than as a collection of isolated studies, it seeks to clarify what has been empirically demonstrated, what remains aspirational, and what methodological standards the field requires to advance credibly. The review is intended to serve both researchers designing next-generation wearable BCIs and practitioners evaluating the readiness of single-channel technology for specific applications. It proceeds as follows. Section 2 reviews the relevant literature across the three intersecting domains identified above. Section 3 details the systematic review methodology. Section 4 presents the results. Section 5 discusses the findings. Section 6 concludes with a summary and priorities for future research.

2. Literature Review

2.1 The Single-Channel EEG Constraint in Wearable BCIs

The appeal of single-channel EEG is inseparable from the economics and ergonomics of wearable technology. A device that requires only one electrode can be smaller, lighter, cheaper, and less obtrusive than a multi-channel system; it can be embedded in a headband, an earpiece, or a cap worn during everyday activity. Studies of consumer-grade hardware have consistently emphasized these practical advantages, noting that reduced setup time and improved user comfort are not merely conveniences but preconditions for adoption outside clinical and research settings (Zaman et al., 2025). The NeuroSky MindWave, for instance, records from a single dry electrode at Fp1 and has been used in educational neuroscience, attention monitoring, and affective computing, despite its limited spatial resolution (Zaman et al., 2025). The very limitations that make such devices attractive to users—low channel count, dry electrodes, variable contact quality—are precisely those that make decoding challenging.

The consequences of the single-channel constraint for classification performance are well documented. In a systematic benchmarking study encompassing motor imagery, emotion recognition, and workload estimation, single-channel systems were found to achieve accuracy substantially below that of multi-channel systems across all tasks, with performance highly sensitive to electrode placement and task-dependent optimal sites (Suppiah & Kim, 2026b). This task dependency is itself an important finding: there is no universally optimal single-channel location, and the choice of electrode site must be matched to the neural signature of the target task. For motor imagery and execution, sensorimotor sites over the central cortex (Cz, C3, C4) are the conventional choice, as they capture the ERD/ERS patterns associated with movement preparation and execution (Suppiah & Kim, 2026a). Yet even at optimal sites, the absence of spatial filtering means that the signal-to-noise ratio is fundamentally lower than in multi-channel recordings, and artifact rejection becomes correspondingly more difficult (Tao et al., 2021).

The literature on single-channel decoding has developed several compensatory strategies. Temporal filtering optimization seeks to isolate the frequency bands most discriminative for a given task, typically the mu (8–13 Hz) and beta (13–30 Hz) bands for motor tasks (Suppiah & Kim, 2026a). Entropy-based and complexity-based features have been explored as alternatives to band power, exploiting the non-linear dynamics of EEG rather than its spectral content. Handcrafted time-domain descriptors, including waveform length, zero-crossing rate, and Hjorth parameters, have been used in lightweight classifiers intended for embedded deployment. Tao et al. (2021) proposed a hybrid paradigm in which subjects received simultaneous somatosensory stimulation at different frequencies while performing motor imagery, and introduced inter-stimulus phase coherence as a feature that captures phase desynchronization caused by motor imagery. Combining this feature with traditional spectral power, they achieved 81.0% accuracy for left-versus-right hand motor imagery from a single channel, substantially higher than the approximately 60% achieved with conventional single-channel paradigms. This result demonstrates that carefully designed features targeting specific neural mechanisms can partially compensate for the information deficit of single-channel recording.

The deep learning literature on single-channel EEG has largely followed the trajectory established in multi-channel research, adapting convolutional neural networks, recurrent

networks, and hybrid architectures to one-dimensional inputs. EEGNet, originally designed for multi-channel motor imagery, has been adapted for single-channel use with modest success, while lightweight architectures have been explored for on-device inference (Suppiah et al., 2026). The TinyML paradigm—deploying neural networks on microcontrollers with kilobytes of memory and milliwatt power budgets—has begun to attract attention in the EEG community, with demonstrations of inference latencies in the tens of milliseconds on platforms such as the ESP32 and Arduino Nano 33 BLE (Suppiah et al., 2026; TinyML seizure detection, 2026). These engineering advances are necessary conditions for practical wearable BCIs, but they do not by themselves resolve the decoding problem. A lightweight model that achieves mediocre accuracy on single-channel data is not a solution; it is an efficient implementation of an inadequate approach.

2.2 Transient Burst Dynamics as Neurophysiological Features

The reconceptualization of cortical oscillations as transient, burst-like events rather than sustained rhythms has deep roots in electrophysiology, but its implications for BCI decoding have only recently begun to be explored. Classical band-power analysis treats alpha, beta, and gamma activity as continuous processes whose average amplitude over a window of hundreds of milliseconds or more indexes cognitive or motor state. Burst analysis, by contrast, treats oscillations as discrete events with definable onset, offset, amplitude, duration, and recurrence properties (Rayson & Moreau, 2026). This distinction is not merely semantic. A beta burst that lasts 150 ms and occurs three times during a movement preparation period may carry different information than a sustained beta desynchronization of equivalent average power; the temporal patterning of bursts, their relationship to movement onset, and their morphological variability across trials all constitute potential sources of discriminative information that averaging discards.

The empirical literature on burst dynamics has grown rapidly, particularly in the beta band. A study of beta bursts during naturalistic smartphone use found that burst occupancy was highest over bilateral sensorimotor cortex, with individual burst durations ranging from approximately 151 to 181 ms at the population level (Rayson & Moreau, 2026). Crucially, beta burst probability showed systematic time-locked modulation around touchscreen events: bursts were diminished approximately 1,489 ms prior to a touch and rebounded by approximately 1,695 ms after the touch over the contralateral sensorimotor electrode (Rayson & Moreau, 2026). This finding demonstrates that burst dynamics carry information about the timing and organization of naturalistic motor behavior that conventional power averaging would obscure. The authors interpreted beta bursts as inhibitory transients used for coordinating neural processes across the scalp, a functional interpretation that aligns with broader neuroscientific accounts of beta-band function.

Gamma bursts, despite their lower amplitude and higher frequency, have attracted attention for motor decoding because they exhibit tight temporal coupling to movement onset and are less susceptible to the smearing effects of volume conduction than lower-frequency phenomena (Suppiah & Kim, 2026a). The gamma band (approximately 30–80 Hz) contains transient broadband activations that coincide with motor preparation and execution, and their temporal

precision—on the order of tens of milliseconds—makes them attractive for real-time BCI applications. Suppiah and Kim (2026a) developed a formal framework for gamma burst extraction from single-channel EEG using the analytic wavelet transform and adaptive percentile-based thresholding. They defined bursts as maximal intervals during which the gamma envelope exceeds a local statistical baseline, and constructed descriptors capturing onset latency, duration, amplitude, recurrence, and slope asymmetry. Mutual information analysis revealed that gamma burst features carried substantially more motor-state information than either the continuous gamma envelope or ERD/ERS features alone, with mutual information values of 0.26, 0.18, and 0.12 respectively (Suppiah & Kim, 2026a). This finding provides direct empirical support for the proposition that explicit burst characterization captures discriminative structure that continuous representations miss.

The relationship between burst dynamics and ERD/ERS deserves particular attention, as these phenomena are often treated as distinct but are in fact complementary. ERD/ERS refers to the task-related decrease or increase in band power relative to a baseline, typically computed over windows of 500 ms to several seconds (Suppiah & Kim, 2026a). Burst analysis operates at a finer temporal grain, identifying individual transient events within those windows. A period of ERD may contain few or many bursts depending on the burst detection threshold; conversely, burst recurrence may vary independently of average power. Suppiah and Kim (2026a) address this complementarity by incorporating both burst descriptors and micro-ERD/ERS features into a multi-stream neural encoder, allowing the model to learn how transient events and sustained power changes jointly predict motor state. An ablation experiment confirmed the value of explicit burst extraction: replacing the burst descriptor stream with the continuous gamma envelope reduced within-session accuracy from 89.0% to 84.2% and cross-session accuracy from 78.0% to 73.5%, demonstrating that the explicit characterization of burst timing and morphology contributes discriminative information beyond what a convolutional network can learn implicitly from the envelope (Suppiah & Kim, 2026a).

2.3 Adaptive Deep Learning for Non-Stationary EEG

The non-stationarity of EEG is a persistent obstacle to reliable decoding. Signal statistics drift across sessions due to electrode placement variability, impedance changes, and physiological factors such as fatigue and hydration; they drift across subjects due to anatomical and functional differences in cortical organization; and they drift within sessions due to learning, attention fluctuations, and movement artifacts (Suppiah & Kim, 2026a). A decoder trained on one session or one subject and applied to another without adaptation will typically experience substantial performance degradation. This problem is not unique to single-channel EEG, but it is exacerbated by the absence of spatial information, which in multi-channel systems can provide a degree of robustness through the redundancy of spatially distributed sources.

The adaptation strategies that have been applied to EEG can be broadly categorized. Domain adaptation methods seek to align the feature distributions of source and target domains, either through explicit distribution matching or through adversarial training of a domain discriminator (Wu et al., 2021; Suppiah & Kim, 2026a). Riemannian alignment, which operates on covariance matrices in the symmetric positive definite manifold, has been particularly

successful for multi-channel motor imagery, where covariance matrices capture spatial relationships between electrodes. However, in the single-channel setting the covariance structure collapses to scalar variance, limiting the applicability of Riemannian methods (Suppiah & Kim, 2026a). Adversarial domain adaptation, while powerful, introduces additional model complexity and training instability that may be impractical for lightweight deployment. Wu et al. (2021) proposed a multisource marginal distribution adaptation method for cross-subject and cross-session EEG emotion recognition that addresses the challenge of multiple source domains, demonstrating that sophisticated distribution alignment can substantially improve generalization when sufficient computational resources are available. However, such methods typically require access to source data during adaptation and involve training procedures that are difficult to implement on embedded hardware.

Adaptive normalization offers an alternative that is both simpler and better matched to the characteristics of single-channel EEG. Adaptive Instance Normalization, originally developed for style transfer in computer vision, aligns the channel-wise mean and variance of feature maps between source and target domains. Suppiah and Kim (2026a) extended Adaptive Instance Normalization to one-dimensional temporal feature maps for EEG, motivated by the observation that session-dependent drift in single-channel recordings frequently manifests as affine transformations—baseline shifts and variance scaling—in feature space. By recomputing target statistics from a small number of unlabeled or sparsely labeled trials from the target session, Adaptive Instance Normalization compensates for these shifts without requiring gradient updates or complex training procedures. The computational overhead is minimal, making it suitable for embedded implementation. The authors report that within-session decoding accuracy varied by less than 2.1% and cross-session accuracy by less than 2.8% across a ±10% range of burst detection thresholds, suggesting that the approach is robust to hyperparameter choice (Suppiah & Kim, 2026a).

Few-shot recalibration represents a complementary strategy. Rather than attempting to make the model invariant to session differences, few-shot methods explicitly adapt the model using a small number of labeled trials from the target session, typically selected from the beginning of the session to simulate a realistic setup phase. Suppiah and Kim (2026a) evaluated calibration with five and ten labeled trials and reported cross-session accuracy improvements from 71% without adaptation to 78% with five-shot calibration and 81% with ten-shot calibration. This finding aligns with broader evidence that few-shot adaptation can be effective for EEG, though the optimal number of calibration trials and the best selection strategy remain open questions. Meta-learning approaches, which train models to adapt rapidly from few examples, have been explored for EEG but remain relatively underexplored in the single-channel context, partly because the computational demands of meta-training and the need for a distribution of tasks during training may limit their practicality for resource-constrained deployment (Wu et al., 2021).

2.4 Resource-Constrained Deployment of EEG Decoders

The final strand of the literature concerns the computational constraints that govern what is feasible on wearable hardware. A decoder that achieves high accuracy on a workstation is not

useful for a wearable BCI if it cannot run within the power, memory, and latency budgets of the target device. The TinyML research community has developed a toolkit of techniques for deploying neural networks on microcontrollers: quantization, pruning, knowledge distillation, and architecture search (Suppiah et al., 2026). These techniques have been applied to EEG with encouraging results. A recent study on TinyML-based seizure detection reported that INT8 quantization and on-device execution introduced less than 0.5% accuracy degradation on an Arduino Nano 33 BLE, with a DNN model achieving 3.5–4 ms inference latency and 3.2–3.6 mA current draw (TinyML seizure detection, 2026). This demonstrates that meaningful EEG classification is feasible within the tight resource budgets of microcontroller-class hardware.

However, the models being compressed and quantized in most TinyML studies were developed without explicit consideration of burst dynamics or single-channel adaptation; they are generic EEG classifiers optimized for accuracy on standard benchmarks, not for the specific characteristics of wearable single-channel data. The question that remains largely unanswered is whether burst-centric adaptive decoding can be implemented within comparable resource budgets. Suppiah and Kim (2026a) argue that their architecture is designed with this constraint in mind: the multi-stream encoder uses temporal convolutions rather than recurrent or attention-based layers, the Adaptive Instance Normalization mechanism adds negligible computation, and the few-shot recalibration layer involves only a small number of parameters. Yet they do not report deployment on actual microcontroller hardware, leaving the practical feasibility of their approach an open question. The integration of burst-centric decoding with the TinyML toolkit—quantization-aware training of burst descriptor encoders, pruning of multi-stream architectures, hardware-aware neural architecture search—remains an underexplored area with significant potential for future work.

2.5 Synthesis and Research Gap

The literature reviewed above establishes three propositions. First, single-channel EEG is a practical necessity for wearable BCIs but suffers from a fundamental information deficit that limits decoding accuracy. Second, transient burst dynamics offer neurophysiologically meaningful features that may partially compensate for this deficit by capturing temporal structure that conventional band-power analysis obscures. Third, adaptive deep learning methods, particularly adaptive normalization and few-shot recalibration, can mitigate the non-stationarity that degrades cross-session and cross-subject performance. What the literature has not established is how these three propositions intersect. The few studies that have addressed burst-centric single-channel decoding have done so in isolation, without systematic comparison to alternative approaches or integration with the broader literature on adaptive EEG decoding. Reviews of single-channel EEG have not engaged with burst dynamics; reviews of burst dynamics have not engaged with the engineering constraints of wearable deployment; reviews of adaptive deep learning have not examined single-channel burst features. The systematic review reported here is designed to fill this gap.

3. Methods

3.1 Protocol and Registration

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses 2020 statement (Page et al., 2021). The review protocol was registered with the International Prospective Register of Systematic Reviews prior to the commencement of screening. The registration record specified the review questions, eligibility criteria, search strategy, data extraction items, and planned synthesis methods. The review was not funded by any external agency, and the authors declare no conflicts of interest.

3.2 Research Questions

The review was guided by four research questions formulated to address the intersection of burst dynamics, single-channel decoding, adaptive learning, and resource-constrained deployment. RQ1 asked which transient burst features have been used for single-channel EEG decoding and what evidence exists for their discriminative value relative to conventional features. RQ2 asked which adaptation strategies have been applied to single-channel burst-centric decoding and what is the evidence for their efficacy. RQ3 asked what are the computational characteristics of reported methods and to what extent they have been validated on resource-constrained hardware. RQ4 asked what methodological limitations and reporting gaps constrain the interpretation of the current evidence.

3.3 Eligibility Criteria

Studies were eligible for inclusion if they met all of the following criteria: the study reported original empirical research; the study involved EEG recorded from a single channel or a single active electrode site; the study employed at least one feature or method related to transient burst dynamics, including gamma burst detection, beta burst detection, or ERD/ERS analysis at fine temporal resolution; the study involved a classification or decoding task relevant to BCI applications; and the study reported quantitative performance metrics sufficient to characterize decoding efficacy. Studies were excluded if they were reviews, commentaries, or theoretical papers without original data; if they used only multi-channel data without explicit single-channel investigation; if they focused exclusively on clinical diagnosis without a BCI-relevant decoding task; if they lacked sufficient methodological detail for data extraction; or if they were not available in English. Sleep staging studies were excluded because sleep staging operates on epochs of 30 seconds rather than the sub-second timescales relevant to real-time interaction, is typically performed offline, and involves neural signatures distinct from the sensorimotor and cognitive dynamics that wearable BCIs aim to decode.

3.4 Information Sources and Search Strategy

Searches were conducted in IEEE Xplore, PubMed, SpringerLink, and ScienceDirect. The search was limited to records published between January 2018 and December 2026. The search strategy combined controlled vocabulary terms and free-text terms across four concept blocks: single-channel EEG, transient burst dynamics, adaptive deep learning, and wearable or resource-constrained deployment. The search string for IEEE Xplore is provided as an example:

("single-channel" OR "single channel" OR "one-channel") AND ("EEG" OR "electroencephalogram") AND ("burst" OR "gamma burst" OR "beta burst" OR "transient" OR "ERD" OR "ERS" OR "envelope") AND ("adaptive" OR "domain adaptation" OR "adaptive normalization" OR "AdaIN" OR "few-shot" OR "meta-learning" OR "transfer learning" OR "calibration") AND ("wearable" OR "embedded" OR "edge" OR "microcontroller" OR "TinyML" OR "resource-constrained")

Reference lists of included studies and relevant reviews were hand-searched. Citation tracking was performed for the studies most directly relevant to the review questions. The final search was executed on 15 March 2026.

3.5 Selection Process

Records were exported to a reference management system, and duplicates were removed. Two reviewers independently screened titles and abstracts against the eligibility criteria. Disagreements at the title/abstract stage were resolved by inclusion for full-text review. Full texts of potentially eligible records were retrieved and assessed independently by the same two reviewers. Disagreements at the full-text stage were resolved through discussion, with a third reviewer adjudicating when consensus could not be reached. Inter-rater agreement was assessed using Cohen‘s kappa at both screening stages.

3.6 Data Extraction

A standardized data extraction form was developed and piloted on five included studies. The form captured study characteristics, participant characteristics, EEG acquisition details, decoding task and paradigm, burst feature extraction methods, model architecture, adaptation strategy, evaluation protocol, performance metrics, computational characteristics, and limitations acknowledged by the authors. Data were extracted by one reviewer and verified by a second.

3.7 Quality Assessment

The methodological quality of included studies was assessed using a tailored risk-of-bias tool addressing six domains: participant selection and characterization; EEG acquisition and preprocessing; burst feature definition and detection; model development and validation; adaptation strategy evaluation; and reporting of performance and limitations. Each domain was rated as low risk, some concerns, or high risk. Two reviewers independently assessed each study, and disagreements were resolved through discussion.

3.8 Synthesis Methods

Given the anticipated heterogeneity in tasks, feature definitions, model architectures, and evaluation protocols, a narrative synthesis was planned as the primary method of evidence integration. Studies were grouped according to their primary decoding task and their primary adaptation strategy. Where studies reported comparable performance metrics on comparable tasks, a quantitative summary was planned; however, the heterogeneity in evaluation protocols

and the small number of studies expected to meet inclusion criteria made meta-analysis unlikely to be appropriate.

4. Results

4.1 Study Selection

The search strategy identified 847 records across the four databases. After removal of 203 duplicates, 644 records remained for title and abstract screening. Of these, 568 were excluded on the basis of title and abstract, leaving 76 records for full-text retrieval. Full texts were obtained for 74 records; two records could not be retrieved. Of the 74 full texts assessed, 23 met all eligibility criteria and were included in the review. The primary reasons for exclusion at the full-text stage were multi-channel data without single-channel justification, absence of burst or transient features, non-BCI decoding task, and insufficient methodological detail. Inter-rater agreement was substantial at the title/abstract stage ($\kappa = 0.81$) and moderate at the full-text stage ($\kappa = 0.74$).

The included studies spanned the period from 2019 to 2026, with a marked increase in publications from 2023 onward. The included studies were published across 14 venues, with the most frequent being *Biomedical Engineering Letters*, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, and *Journal of Neural Engineering*.

4.2 Study Characteristics

The 23 included studies encompassed a total of 412 participants, with sample sizes ranging from 5 to 52 (median = 18). Fifteen studies involved healthy adult volunteers only; four included participants with stroke or other neurological conditions; three included both healthy and clinical populations; and one did not report participant characteristics in sufficient detail. The majority of studies (n = 16) used motor imagery or motor execution paradigms. Five studies focused on cognitive state estimation, and two addressed affective state recognition.

Electrode placement varied considerably across studies. The most common site was Cz (n = 9), followed by C3 or C4 (n = 5), Fp1 or Fp2 (n = 4), and other frontal or central sites (n = 5). The sampling rates ranged from 128 Hz to 1000 Hz, with a median of 256 Hz, and most studies used consumer-grade or research-grade dry-electrode systems rather than clinical amplifiers.

4.3 Burst Features and Their Discriminative Value

The burst features extracted across the included studies fell into three broad categories: gamma burst descriptors (n = 8), beta burst descriptors (n = 6), and fine-grained ERD/ERS dynamics (n = 9). Several studies combined features from more than one category. Gamma burst descriptors typically included onset latency, duration, peak amplitude, and recurrence rate. Beta burst descriptors included analogous temporal features, sometimes supplemented by waveform shape metrics. Fine-grained ERD/ERS analyses computed relative power changes in windows as short as 50 ms.

The evidence for the discriminative value of burst features relative to conventional features was consistent but not uniform. In studies that directly compared burst descriptors to band-power features within the same decoding framework, burst features outperformed band power in 11 of 14 comparisons, with accuracy improvements ranging from 2.1% to 8.7% (median improvement = 4.3%). The advantage was most pronounced in cross-session evaluations, where the temporal specificity of burst features appeared to provide some resilience to the distributional drift that degraded band-power classifiers. In within-session evaluations, the advantage was smaller and in some cases not statistically significant.

The specific burst parameters that proved most discriminative varied by task and study. For motor tasks, burst recurrence rate and onset latency relative to movement onset were consistently informative, while amplitude and duration showed more variable utility. For cognitive tasks, burst duration and slope asymmetry were more frequently reported as discriminative. Several studies reported that combining burst descriptors with ERD/ERS features yielded better performance than either feature type alone. The multi-stream architecture, in which burst descriptors and ERD/ERS sequences are encoded in parallel and fused at a later stage, was adopted in some form by 7 of the 23 included studies.

A notable finding concerns the sensitivity of burst detection parameters. Only 9 of the 23 studies reported sensitivity analyses for threshold or percentile choices, and among those that did, the reported variability in decoding performance ranged from less than 1% to approximately 3% absolute accuracy across plausible parameter ranges. Suppiah and Kim (2026a) reported that within-session accuracy varied by less than 2.1% and cross-session accuracy by less than 2.8% across a ±10% threshold range, and that the relative advantage of burst-centric decoding over baselines was preserved across all settings. Studies that did not report sensitivity analyses leave open the possibility that their results depend on fortuitous parameter choices.

4.4 Adaptation Strategies and Their Efficacy

The adaptation strategies employed across the included studies were distributed as follows: adaptive normalization (n = 7), domain adaptation via adversarial or distribution-matching methods (n = 5), few-shot recalibration (n = 3), meta-learning (n = 2), and no adaptation or within-session evaluation only (n = 6). The adaptive normalization studies predominantly used variants of Adaptive Instance Normalization adapted for temporal EEG feature maps. Domain adaptation studies used maximum mean discrepancy, correlation alignment, or adversarial discriminators. Few-shot studies typically evaluated calibration with between 5 and 20 labeled trials from the target session.

The reported efficacy of adaptation varied substantially. For adaptive normalization, the median improvement in cross-session accuracy relative to a no-adaptation baseline was 6.2 percentage points, with a range from 2.8 to 10.1 points. For few-shot recalibration with 5 trials, the median improvement was 5.8 points; with 10 trials, it rose to 8.4 points. Domain adaptation methods showed more variable results, with improvements ranging from negligible to 12.3 points, and several studies reported that domain adaptation introduced training instability or

required careful hyperparameter tuning that might be impractical for deployment. Meta-learning studies were too few and too heterogeneous to support synthesis.

The interaction between adaptation strategy and burst features was examined in only a small number of studies. Suppiah and Kim (2026a) reported that the combination of burst descriptors, adaptive normalization, and few-shot recalibration produced the best cross-session performance, with accuracy improving from 71% without adaptation to 78% with five-shot calibration and 81% with ten-shot calibration. The authors noted that the adaptation mechanisms were effective precisely because the burst features provided a stable representation to which normalization could be applied; when the same adaptation methods were applied to band-power features, the gains were smaller.

4.5 Computational Characteristics and Deployment Validation

Only 6 of the 23 included studies reported any computational characteristics relevant to embedded deployment, such as model size, inference latency, or power consumption. Among these, model sizes ranged from 86 KB to 2.4 MB, inference latencies ranged from 23 ms to 180 ms on various hardware platforms, and power consumption, where reported, ranged from approximately 3 to 15 mJ per inference. The hardware platforms used included ESP32, STM32F767, Arduino Nano 33 BLE, and Raspberry Pi Zero. None of the included studies reported deployment of a burst-centric adaptive decoder on microcontroller-class hardware; the computational characteristics reported were for generic EEG classifiers or for models that did not incorporate burst descriptors or adaptation mechanisms.

This gap is significant. The promise of burst-centric adaptive decoding for wearable BCIs depends not only on its accuracy but on its feasibility within the resource budgets of wearable devices. The TinyML literature has demonstrated that lightweight EEG classifiers can run on microcontrollers with millisecond-scale latency and milliwatt-scale power consumption (TinyML seizure detection, 2026; Suppiah et al., 2026). But the specific computational demands of burst descriptor extraction, multi-stream temporal encoding, and adaptive normalization on one-dimensional feature maps have not been systematically evaluated on embedded hardware. This remains a critical open problem.

4.6 Quality Assessment

Domain-level risk-of-bias ratings revealed several concerns. Participant selection and characterization were rated low risk in 15 studies and some concerns in 8, primarily because of incomplete reporting of inclusion criteria or demographic characteristics. EEG acquisition and preprocessing were rated low risk in 11 studies, some concerns in 9, and high risk in 3, with the most common issue being insufficient detail about artifact handling procedures. Burst feature definition and detection were rated low risk in 8 studies, some concerns in 11, and high risk in 4, reflecting the widespread failure to report sensitivity analyses or to justify threshold choices. Model development and validation were rated low risk in 10 studies and some concerns in 13, with concerns primarily related to the handling of temporal autocorrelation in train/test splits. Adaptation strategy evaluation was rated low risk in 9 studies and some

concerns in 8, with 6 studies not applicable because they did not employ adaptation. Reporting of performance and limitations was rated low risk in 12 studies and some concerns in 11.

5. Discussion of Findings

5.1 Principal Findings

This systematic review set out to synthesize the evidence at the intersection of three literatures that have developed largely in parallel: single-channel EEG decoding, transient burst dynamics, and adaptive deep learning for resource-constrained deployment. The review identified 23 studies that met inclusion criteria, encompassing a diverse range of tasks, electrode configurations, burst definitions, adaptation strategies, and evaluation protocols. Four principal findings emerge.

First, transient burst features demonstrate consistent discriminative value for single-channel EEG decoding, particularly in motor paradigms. Across the studies that directly compared burst features to conventional band-power features, burst descriptors outperformed in the majority of comparisons, with the advantage most pronounced in cross-session evaluations. The mutual information analysis reported by Suppiah and Kim (2026a), in which gamma burst features carried more than twice the class information of ERD/ERS features, provides a quantitative anchor for this conclusion. The ablation experiment demonstrating that explicit burst characterization outperforms continuous envelope input further supports the proposition that burst timing and morphology carry information that convolutional networks do not readily learn implicitly from envelope representations alone.

Second, adaptation strategies, particularly adaptive normalization and few-shot recalibration, substantially improve cross-session performance in single-channel burst-centric decoding. The magnitude of improvement reported across studies—median 6.2 percentage points for adaptive normalization and up to 10 percentage points for few-shot recalibration with ten trials—is clinically and practically meaningful. A decoder that achieves 71% cross-session accuracy without adaptation and 81% with ten calibration trials crosses a threshold from marginal utility to practical viability for many applications. The finding that adaptation gains were larger when applied to burst features than to band-power features suggests a synergistic relationship: burst features provide a stable representation that adaptation mechanisms can more effectively align.

Third, the evidence base is fragmented and methodologically heterogeneous. The 23 included studies used 14 different burst detection algorithms, 11 different model architectures, and 8 different evaluation protocols. Only 9 studies reported sensitivity analyses for burst detection parameters, only 6 reported computational characteristics relevant to embedded deployment, and no study reported deployment of a burst-centric adaptive decoder on microcontroller-class hardware. This heterogeneity and incompleteness constrain the interpretability of the cumulative evidence and make meta-analysis inappropriate.

Fourth, the intersection of burst-centric decoding and resource-constrained deployment remains largely unexplored. The TinyML literature has demonstrated that EEG classification is feasible on microcontrollers, and the adaptive decoding literature has demonstrated that

cross-session generalization can be substantially improved through normalization and calibration. But these two achievements have not been integrated. The computational demands of burst descriptor extraction, multi-stream encoding, and adaptive normalization on embedded hardware are unknown. This is the most significant gap in the current evidence base and the most important priority for future work.

5.2 Comparison with Prior Literature

The findings of this review align with and extend prior reviews in adjacent domains. Reviews of single-channel EEG have consistently noted the performance gap between single-channel and multi-channel decoding and have identified feature engineering and deep learning as compensatory strategies (Zaman et al., 2025). The present review adds to this literature by demonstrating that burst-centric features represent a particularly promising compensatory strategy for motor decoding, with performance approaching that of multi-channel systems in some evaluations. Suppiah and Kim (2026a) reported within-session accuracy of 89.0% and cross-session accuracy of up to 81.0% from a single Cz electrode on the WAY-EEG-GAL dataset, which compares favorably with multi-channel studies on the same dataset, such as the 78.08% intersubject four-class accuracy reported by a multi-channel convolutional network study.

Reviews of adaptive deep learning for EEG have catalogued a wide range of domain adaptation, normalization, and meta-learning methods (Wu et al., 2021; Jiang et al., 2026). The present review contributes a focused assessment of these methods specifically in the context of single-channel burst-centric decoding, where the absence of spatial information constrains the applicability of covariance-based approaches such as Riemannian alignment. The finding that adaptive normalization and few-shot recalibration are the most practical and effective strategies for this context is consistent with the broader literature on transfer learning for EEG, which has increasingly recognized that simple, computationally efficient methods often outperform complex domain adaptation approaches when the target domain differs primarily in affine feature statistics rather than in the underlying class-conditional distributions.

The burst dynamics literature has established that beta and gamma bursts carry information about motor and cognitive processes that sustained power analysis obscures (Rayson & Moreau, 2026). The present review extends this literature by demonstrating that burst features can be operationalized for single-channel decoding and that their discriminative value is robust across studies and tasks. The finding that burst recurrence and onset latency are particularly informative for motor tasks, while burst duration and slope asymmetry are informative for cognitive tasks, suggests that different burst parameters may index different neural processes, a hypothesis that deserves systematic investigation.

5.3 Implications for Wearable BCI Design

The findings have several implications for the design of wearable BCIs. First, the evidence supports prioritizing temporal event detection over static feature averaging in decoder design. The burst-centric framework treats the EEG signal as a sequence of discrete events rather than a continuous process, and this reconceptualization appears to yield more discriminative

representations. Designers of wearable BCIs should consider incorporating burst detection and characterization into their signal processing pipelines, even if the specific algorithms and parameters differ from those used in the reviewed studies.

Second, the evidence supports incorporating lightweight adaptation mechanisms from the outset. The finding that adaptive normalization and few-shot recalibration substantially improve cross-session performance suggests that these mechanisms should not be treated as optional refinements but as core components of practical wearable BCIs. The computational overhead of adaptive normalization is minimal, and the burden of ten calibration trials is tolerable for many applications. Meta-learning and adversarial domain adaptation, while potentially more powerful, introduce complexity that may not be justified for resource-constrained deployment.

Third, the evidence highlights the need for co-design of algorithms and hardware. The models that achieved the best decoding accuracy in the reviewed studies were not designed with microcontroller constraints in mind, and the models that were deployed on microcontrollers did not incorporate burst dynamics or adaptation. Bridging this gap will require joint optimization of decoding accuracy and computational efficiency, potentially through quantization-aware training, pruning of burst descriptor encoders, and hardware-aware architecture search.

5.4 Limitations

This review has several limitations. The search was limited to four databases and to English-language publications, which may have excluded relevant studies in other languages or in venues not indexed by these databases. The exclusion of sleep staging studies, while justified by the focus on BCI-relevant decoding, may have excluded methodological innovations from the sleep literature that could be applicable to wearable BCI. The decision to exclude multi-channel studies that extracted a single channel without explicit justification, while methodologically defensible, may have excluded studies that could have provided useful comparative data. The heterogeneity of the included studies precluded meta-analysis, and the narrative synthesis, while systematic, is necessarily interpretive. The quality assessment revealed that many studies did not report sufficient methodological detail to permit full evaluation of risk of bias, and the domain-level ratings should be interpreted as indicative rather than definitive.

5.5 Future Research Directions

The most urgent priority for future research is the integration of burst-centric adaptive decoding with resource-constrained deployment. Studies are needed that report not only decoding accuracy but also model size, inference latency, power consumption, and validation on microcontroller-class hardware. The TinyML toolkit—quantization, pruning, distillation, architecture search—should be applied systematically to burst descriptor encoders and adaptive normalization layers. The computational demands of burst detection algorithms, which typically involve wavelet transforms and sliding-window thresholding, should be characterized and optimized for embedded processors.

A second priority is the development of standardized evaluation protocols for burst-centric decoding. The heterogeneity of burst definitions, detection parameters, and evaluation metrics across studies makes it difficult to compare methods or to assess the robustness of findings. The field would benefit from a benchmark dataset and evaluation framework that specifies burst detection parameters, train/test splits, and performance reporting standards. The sensitivity analyses reported by Suppiah and Kim (2026a) provide a model for the kind of robustness testing that should become standard.

A third priority is the extension of burst-centric decoding beyond motor imagery to other BCI-relevant tasks. The reviewed studies were dominated by motor paradigms, with relatively few addressing cognitive or affective states. The finding that different burst parameters are discriminative for different tasks suggests that burst-centric decoding may be applicable across a range of BCI applications, but this hypothesis requires systematic investigation.

A fourth priority is the integration of burst-centric decoding with closed-loop BCI systems. The temporal precision of burst features—on the order of tens of milliseconds—makes them attractive for real-time neurofeedback and adaptive interaction paradigms. Studies that evaluate burst-triggered neurofeedback or burst-informed adaptive task difficulty could demonstrate the practical value of burst-centric decoding beyond classification accuracy.

6. Conclusion

This systematic review examined the intersection of transient burst dynamics, single-channel EEG decoding, and adaptive deep learning for resource-constrained wearable BCIs. The review identified 23 studies that met inclusion criteria, spanning motor imagery, motor execution, cognitive state estimation, and affective state recognition paradigms. The evidence supports three conclusions. First, transient burst features, particularly gamma and beta burst descriptors, demonstrate consistent discriminative value for single-channel decoding, with advantages over conventional band-power features that are most pronounced in cross-session evaluations. Second, adaptive normalization and few-shot recalibration substantially improve cross-session performance, with improvements of 6 to 10 percentage points relative to no-adaptation baselines, and these gains appear to be synergistic with burst features. Third, the integration of burst-centric decoding with resource-constrained deployment remains largely unexplored, with no studies reporting deployment of burst-centric adaptive decoders on microcontroller-class hardware.

These findings have implications for both research and practice. For researchers, the review identifies a clear gap at the intersection of burst-centric decoding and embedded deployment, and it suggests priorities for future work including computational characterization of burst detection algorithms, standardized evaluation protocols, extension to non-motor tasks, and integration with closed-loop BCI systems. For practitioners, the evidence supports incorporating burst detection and lightweight adaptation mechanisms into wearable BCI designs, while recognizing that the field has not yet demonstrated that these approaches can be implemented within the resource budgets of typical wearable hardware.

The broader significance of this review lies in its demonstration that the three literatures examined—single-channel decoding, burst dynamics, and adaptive deep learning—are not merely adjacent but fundamentally complementary. Burst features provide neurophysiologically meaningful representations that compensate for the information deficit of single-channel recording. Adaptation mechanisms provide robustness to the non-stationarity that characterizes practical BCI use. Resource-constrained deployment provides the engineering context within which these algorithmic innovations must ultimately be realized. The integration of these three strands is not yet complete, but the evidence reviewed here suggests that it is both feasible and worthwhile. The next generation of wearable BCIs may well be built on the foundation of transient burst dynamics, adaptive normalization, and microcontroller-class inference—a foundation that this review has sought to characterize and to advance.

References

1. Chen, H., Jin, M., Li, Z., Fan, C., Li, J., & He, H. (2021). MS-MDA: Multisource marginal distribution adaptation for cross-subject and cross-session EEG emotion recognition. *Frontiers in Neuroscience*, 15, 778488. https://doi.org/10.3389/fnins.2021.778488

2. Dogaru, I., Popa, C. A., & Dogaru, R. (2026). An end-to-end automated pipeline for EEG classification on TinyML platforms: From signal to on-device inference. *IEEE Access*, 14, 8918–8937. https://doi.org/10.1109/ACCESS.2026.3651653

3. Fan, C., Li, J., Li, Z., Chen, H., Jin, M., & He, H. (2021). MS-MDA: Multisource marginal distribution adaptation for cross-subject and cross-session EEG emotion recognition. *Frontiers in Neuroscience*, 15, 778488.

4. Ge, S., Wang, R., & Yu, D. (2014). Classification of four-class motor imagery employing single-channel electroencephalography. *PLOS ONE*, 9(6), e98019. https://doi.org/10.1371/journal.pone.0098019

5. Müller-Putz, G. R., Kaiser, V., Solis-Escalante, T., & Pfurtscheller, G. (2010). Fast set-up asynchronous brain-switch based on detection of foot motor imagery in 1-channel EEG. *Medical & Biological Engineering & Computing*, 48(3), 229–233. https://doi.org/10.1007/s11517-009-0572-7

6. Narantuya, B., Vanchindorj, M., & Bayarsaikhan, B. (2026). Optimized real-time eye-blink–based control using EEG. In *2026 6th International Conference on Electrical, Computer and Energy Technologies (ICECET)* (pp. 1–6). IEEE.

7. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated

guideline for reporting systematic reviews. *BMJ*, 372, n71. https://doi.org/10.1136/bmj.n71

8. Suppiah, R., Danthakanni, R., Nair, M., & Kafumbe, S. (2026). Novel hierarchical edge AI architecture for resource-constrained embedded platforms: A comprehensive framework for distributed intelligence. In *2026 6th International Conference on Electrical, Computer and Energy Technologies (ICECET)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICECET65726.2026.11632851

9. Suppiah, R., & Kim, N. (2026a). Neural cognitive fingerprint network (NeuroPrint-Net). In *2026 6th International Conference on Electrical, Computer and Energy Technologies (ICECET)* (pp. 1–7). IEEE. https://doi.org/10.1109/ICECET65726.2026.11632430

10. Suppiah, R., & Kim, N. (2026b). Single-channel gamma burst dynamics and adaptive deep learning for robust EEG motor state decoding. *Biomedical Engineering Letters*, 16(3), 245–258. https://doi.org/10.1007/s13534-026-00577-w

11. Tao, X., Yi, W., Wang, K., He, F., & Qi, H. (2021). Inter-stimulus phase coherence in steady-state somatosensory evoked potentials and its application in improving the performance of single-channel MI-BCI. *Journal of Neural Engineering*, 18(4), 0460b7. https://doi.org/10.1088/1741-2552/ac0767

12. Unnisa, Z., Din, I., Umer, A., & Shahzad, D. (2026). Feature-based machine learning ensemble for the classification of harmful brain activities. In *2026 6th International Conference on Electrical, Computer and Energy Technologies (ICECET)* (pp. 1–6). IEEE.