

International Journal Research Publication Analysis

Page: 01-19

NON-INVASIVE BRAIN–COMPUTER INTERFACES (BCIS): TECHNOLOGIES, SIGNAL PROCESSING, ARTIFICIAL INTELLIGENCE, APPLICATIONS, CHALLENGES, AND FUTURE DIRECTIONS A COMPREHENSIVE STATE-OF-THE-ART REVIEW

***Kuldip K. Karale, Gauri S. Mankar, Atharva R. Darvekar, Prof Narendra J. Padole**

P. G. Department of Computer Science and Technology, Degree College of Physical
Education, Shree H. V. P. Mandal, Amravati, Maharashtra, India.

Article Received: 18 August 2026

*Corresponding Author: Kuldip K. Karale

Article Revised: 06 September 2026

P. G. Department of Computer Science and Technology, Degree College of Physical
Education, Shree H. V. P. Mandal, Amravati, Maharashtra, India.

Published on: 26 September 2026

DOI: <https://doi-doi.org/101555/ijrpa.1209>

ABSTRACT

Non-invasive brain–computer interfaces (BCIs) translate scalp- or externally measured neural activity into commands, communication, or feedback without breaching the nervous system. Their accessibility and safety make them central to assistive communication, neurorehabilitation, cognitive assessment, and emerging human–computer interaction; nevertheless, weak and non-stationary signals, artifacts, calibration burden, and limited cross-subject generalization continue to restrict dependable deployment. This review integrates the technological, algorithmic, clinical, and governance dimensions of non-invasive BCI. It compares electroencephalography (EEG), functional near-infrared spectroscopy (fNIRS), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and hybrid systems; explains acquisition, preprocessing, feature extraction, decoding, and feedback; and evaluates classical machine learning, compact deep networks, transformers, self-supervised learning, and generative decoders. Particular attention is given to motor imagery, P300, steady-state visually evoked potentials, EEG-to-text translation, and fMRI-based visual reconstruction. The review argues that artificial intelligence can compensate for—but cannot repeal—the information limits imposed by non-invasive sensing. Reliable progress therefore requires leakage-resistant evaluation, uncertainty estimation, multimodal co-design, privacy-preserving learning, transparent human oversight, and longitudinal clinical validation. A layered roadmap is proposed for calibration-light, adaptive, secure, and clinically accountable closed-loop BCIs.

INDEX TERMS: Brain–Computer Interface, Eeg, Fnirs, Meg, Fmri, Motor Imagery, P300, Ssvep, Deep Learning, Transformers, Generative Ai, Neural Privacy, Neurorehabilitation.

I. INTRODUCTION

A brain–computer interface establishes a functional pathway between measured neural activity and an external system. In the canonical loop, neural signals are acquired, conditioned, converted into features or latent representations, decoded into an intention or state estimate, and returned to the user through visual, auditory, haptic, robotic, or therapeutic feedback [1], [2]. The term BCI should therefore denote more than a classifier: it describes a coupled human–machine control system whose performance depends on both algorithmic adaptation and the user’s capacity to learn a stable neural strategy.

Invasive interfaces can sample cortical activity with high spatial and temporal fidelity, but surgery, long-term tissue response, maintenance, and eligibility constrain their scale. Non-invasive systems instead observe electrical, magnetic, or hemodynamic consequences of neural activity through intact tissue. They trade signal fidelity for substantially lower physical risk and broader experimental, clinical, and consumer accessibility [3]. EEG remains the most widely deployed modality because it is comparatively inexpensive, temporally precise, and portable. fNIRS adds localized hemodynamic information; MEG provides high temporal resolution with better source localization at high cost; and fMRI offers whole-brain spatial detail but poor portability and delayed dynamics.

Three developments motivate a renewed synthesis. First, wearable sensors and dry or semi-dry electrodes are moving BCI beyond tightly controlled laboratories. Second, representation learning has shifted decoding from handcrafted features toward compact convolutional networks, graph models, transformers, and self-supervised pretraining [4]–[7]. Third, generative models are expanding the output space from a few discrete commands toward language and visual reconstruction [8]–[11]. These gains are significant, but many reported results remain task-specific and are vulnerable to non-independent train–test partitions, small cohorts, session drift, and optimistic offline evaluation.

This review makes four contributions. It provides a unified comparison of non-invasive sensing technologies; connects neurophysiological paradigms to appropriate signal-processing and AI methods; distinguishes demonstrated capability from speculative interpretation; and proposes a practical architecture and evaluation agenda for trustworthy, calibration-light closed-loop systems. Unlike a systematic meta-analysis, the paper is a

structured state-of-the-art review. Quantitative values are reported as typical ranges or study-specific results and should not be interpreted as universal head-to-head benchmarks.

II. REVIEW SCOPE AND METHOD

The review prioritizes peer-reviewed journal articles, major conference proceedings, standards, and official regulatory texts available through August 2026. Foundational papers were retained when they define widely used concepts such as event-related desynchronization, P300 spelling, common spatial patterns, independent component analysis, and information-transfer rate. Recent sources were selected to represent deep learning, foundation-model pretraining, EEG-to-language decoding, fMRI-to-image reconstruction, hybrid sensing, clinical translation, and neurodata governance. Commercial webpages and unsourced performance claims were excluded from the evidence base.

Evidence was organized into six analytic dimensions: sensing physics; elicitation paradigm; acquisition and preprocessing; representation and decoding; application readiness; and safety, ethics, and governance. Because datasets, class definitions, trial counts, and validation protocols differ substantially, accuracy values were not pooled. Greater evidentiary weight was assigned to cross-session, cross-subject, prospective, and clinically supervised evaluations than to random window-level splits from a single recording session.



Fig. 1. Structured review and synthesis workflow used in this paper.

III. NEUROPHYSIOLOGICAL FOUNDATIONS AND BCI PARADIGMS

A. Electrophysiological and Hemodynamic Sources

Scalp EEG predominantly reflects synchronized postsynaptic potentials generated by populations of similarly oriented cortical pyramidal neurons. Conductive tissue and the skull attenuate and spatially blur these fields; measured amplitudes are typically in the microvolt range. EEG consequently provides millisecond-scale temporal information but relatively coarse source localization and high sensitivity to ocular, muscular, cardiac, movement, and electromagnetic artifacts [3], [12]. MEG records femtotesla-scale magnetic fields associated with neural currents. Magnetic fields are less distorted by the skull, yet conventional superconducting systems require cryogenics and magnetically shielded rooms. Optically pumped magnetometers promise more naturalistic movement, although environmental field control remains demanding [13].

fNIRS and fMRI infer neural activity indirectly through neurovascular coupling. fNIRS measures wavelength-dependent absorption changes associated mainly with oxygenated and deoxygenated hemoglobin in superficial cortex. fMRI measures the blood-oxygen-level-dependent response across the brain. Both offer better spatial specificity than routine scalp EEG but inherit a hemodynamic delay measured in seconds. This delay is acceptable for state monitoring, assessment, and rehabilitation but limits rapid command selection. Hybrid EEG–fNIRS systems combine complementary temporal and spatial information, though improved accuracy must justify synchronization, setup, and computational complexity [14], [15].

B. Motor Imagery and Sensorimotor Rhythms

Motor imagery BCIs use changes in sensorimotor rhythms, especially mu and beta activity, during imagined movement. Event-related desynchronization indicates a task-related reduction in rhythmic power; post-movement event-related synchronization is often observed as a beta rebound [16]. Common spatial patterns (CSP), filter-bank CSP, regularized covariance methods, Riemannian geometry, and deep convolutional networks are common decoders. Motor imagery is self-paced and suitable for rehabilitation or continuous control, but performance varies markedly across users and sessions. Feedback design and user training are therefore part of the learning system rather than peripheral interface choices.

C. P300 and Other Event-Related Potentials

P300 interfaces detect a positive event-related potential following an infrequent, task-relevant stimulus. The classic row–column speller infers a symbol from responses to flashed groups [17]. Averaging improves signal-to-noise ratio but reduces speed; modern paradigms use optimized stimulus timing, language models, and adaptive stopping. P300 systems often require less volitional rhythm training than motor imagery, yet depend on attention, vision, and controlled stimulus presentation.

D. Steady-State Visually Evoked Potentials

SSVEP BCIs present targets modulated at distinct frequencies or phases. Attention to a target produces an occipital response at the stimulus frequency and harmonics. Canonical correlation analysis, task-related component analysis, and filter-bank methods provide strong performance and high information-transfer rates under controlled conditions [18]. Practical limits include visual fatigue, photosensitivity precautions, display refresh constraints, gaze dependence in some users, and performance degradation in unconstrained illumination.

E. Speech, Language, and Visual Decoding

Language decoding spans overt speech, attempted speech, covert speech, imagined speech, and stimulus-evoked reading. These are not interchangeable. Overt and attempted speech

provide stronger and more time-aligned motor signals, whereas unrestricted internal language lacks observable boundaries and reliable ground truth. DeWave introduced discrete quantization for EEG-to-text translation without eye-fixation markers, and EEG2TEXT combined masked pretraining with regional multi-view transformers [8], [9]. Both demonstrate research progress, but current non-invasive language decoding should not be described as unrestricted mind reading: datasets are small, tasks are constrained, and semantic outputs may be strongly influenced by pretrained language priors.

Visual reconstruction from fMRI has advanced through diffusion-model priors. MinD-Vis combines sparse masked brain modeling with double-conditioned latent diffusion, while related work maps fMRI representations into latent diffusion or multimodal embedding spaces [10], [11]. Reconstructions can recover recognizable semantic and structural information under controlled image-viewing experiments. They do not demonstrate arbitrary decoding of private, self-generated visual thought.

TABLE I. COMPARISON OF PRINCIPAL NON-INVASIVE SENSING MODALITIES.

Modality	Measured quantity	Temporal resolution	Spatial resolution	Mobility	Principal strengths	Principal limitations
EEG	Scalp voltage	Milliseconds	Centimeters; inverse-model dependent	High	Low cost, mature, real-time	Low SNR; artifact and volume conduction
fNIRS	Hemoglobin absorption changes	Seconds	$\approx 1-3$ cm, superficial cortex	Moderate-high	Wearable; complementary to EEG	Hemodynamic delay; extracerebral contamination
MEG	Neuromagnetic field	Milliseconds	Millimeter-centimeter, system dependent	Low; improving with OPM	High temporal fidelity; less skull distortion	Cost, shielding, motion/environment sensitivity
fMRI	BOLD contrast	Seconds	Millimeter scale	Very low	Whole-brain spatial coverage	Cost, immobility, delay, acoustic environment
EEG-fNIRS	Electrical + hemodynamic	Milliseconds to seconds	Complementary	Moderate	Multimodal robustness potential	Synchronization and wearable complexity

IV. SIGNAL ACQUISITION AND PROCESSING



Fig. 2. End-to-end closed-loop BCI processing chain. Feedback changes both user strategy and decoder state.

A. Acquisition, Referencing, and Quality Control

Reproducible decoding begins with acquisition metadata. Reports should identify the device, electrode or optode montage, reference, sampling rate, analog filters, impedance policy, synchronization method, stimulus timing, rejected channels, and calibration procedure. Electrode count alone is not a quality metric: coverage must match the expected cortical source and the downstream decoder. Wet electrodes typically offer stable low impedance but require preparation and cleaning. Dry systems reduce setup time but may increase contact noise and motion sensitivity. Quality-control dashboards should expose impedance or contact indicators, saturation, line noise, missing packets, and event-timing error before model inference.

B. Artifact Management

A typical EEG pipeline applies detrending or high-pass filtering, a task-appropriate low-pass filter, and narrowband power-line suppression when necessary. Referencing may use common average, linked mastoids, REST, or spatial Laplacians. Independent component analysis can separate stereotyped ocular or muscular sources, but automatic component rejection must be validated because neural and artifact activity can mix [19]. Artifact subspace reconstruction, wavelet denoising, regression against EOG, and robust channel interpolation are useful alternatives. Filtering must be applied consistently within each training fold to prevent information leakage from future samples.

C. Feature and Representation Families

Time-domain features include event-related potentials, amplitudes, latencies, autoregressive coefficients, and Hjorth parameters. Frequency-domain representations use band power or power spectral density (PSD); time–frequency methods include short-time Fourier transforms and wavelets. Spatial filters such as CSP maximize variance contrast between two conditions, while xDAWN improves evoked-potential signal-to-noise ratio. Covariance matrices mapped through Riemannian geometry provide robust compact representations. Deep networks learn temporal filters, spatial projections, and nonlinear decision boundaries jointly, but their apparent convenience does not remove the need for artifact control and validation discipline.

$$SNR_{dB} = 10\log_{10}(P_{signal}/P_{noise}) \quad (1)$$

Equation (1) expresses signal-to-noise ratio in decibels. In BCI, “signal” must be operationally defined—often as task-related power, evoked response, or model-relevant variance—because neural background activity is not simply noise.

$$w * \argmax_w (w^T \Sigma_1 w) / (w^T \Sigma_2 w) \quad (2)$$

Equation (2) gives the core two-class CSP objective, where Σ_1 and Σ^2 are class-specific covariance estimates. Regularization is important when channel dimension is large relative to available trials.

V. MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

A. Classical Statistical Learning

Linear discriminant analysis, logistic regression, support-vector machines, shrinkage covariance classifiers, random forests, and regularized linear models remain strong baselines. They train efficiently, behave predictably on small datasets, and can be easier to calibrate than large neural networks. For many clinical datasets, a carefully regularized linear model on physiologically meaningful features is more credible than a deeper model evaluated with an easier split. Every advanced method should therefore be compared with transparent baselines under identical partitions and preprocessing.

B. Convolutional, Recurrent, and Graph Networks

EEGNet uses temporal convolutions, depthwise spatial filtering, and separable convolutions to provide a compact architecture across several BCI paradigms [5]. Deep and shallow ConvNets reproduce frequency-selective and spatial operations at different levels of abstraction [6]. Recurrent networks can model longer temporal dependencies but are harder to optimize and may be unnecessary when convolutional receptive fields cover the relevant epoch. Graph neural networks encode electrode topology or learned functional connectivity, although conclusions can be sensitive to graph construction and montage differences.

C. Transformers and Self-Supervised Pretraining

Transformers support long-range interactions and flexible channel–time tokenization. Their data requirements motivate self-supervised objectives such as masked reconstruction, contrastive learning, temporal context prediction, and cross-modal alignment. BENDR showed that large-scale pretraining can produce transferable EEG representations [7]. Foundation-model claims nevertheless require evidence across devices, montages, populations, tasks, and institutions. A model fine-tuned on a new participant is not calibration-free merely because its encoder was pretrained.

D. Generative AI and Semantic Decoding

Generative modeling contributes in three distinct ways: synthetic signal augmentation, latent representation learning, and generation of semantic outputs. These uses should not be conflated. GANs, VAEs, and diffusion models can augment scarce data, but synthetic trials may reproduce training bias or artificially simplify class boundaries. Language and image

priors can make decoded outputs coherent even when neural evidence is weak. Evaluation must therefore report both task performance and the incremental information contributed by the brain signal—for example, through shuffled-signal, subject-permutation, and prior-only ablations.

DeWave quantizes continuous EEG embeddings into a learned codebook before language decoding [8]. EEG2TEXT uses masked EEG pretraining and multiple spatial views [9]. MinD-Vis first learns fMRI representations through sparse masked modeling and then conditions a latent diffusion model [10]. These architectures are important demonstrations of cross-modal representation learning. Their current role is best understood as controlled neural decoding research, not general-purpose thought extraction.

E. Domain Adaptation, Personalization, and Continual Learning

Inter-subject anatomy, cap placement, impedance, medication, fatigue, strategy, and disease state shift the data distribution. Domain-adversarial learning, alignment of covariance geometry, normalization, meta-learning, and parameter-efficient adaptation can reduce—but rarely eliminate—calibration. Continual learning must avoid catastrophic forgetting and should include drift detection, rollback, and safe fallback behavior. Personalization should be bounded by clinically validated operating conditions.

TABLE II. ALGORITHM FAMILIES AND APPROPRIATE EVALUATION QUESTIONS.

Family	Typical input	Advantages	Key risks	Minimum comparison
LDA / logistic / SVM	Engineered temporal, spectral, spatial features	Data-efficient; interpretable baseline	Feature dependence; session drift	Nested CV; regularized linear baseline
CNN / EEGNet	Raw or lightly processed channel $\times$ time	Learns local temporal/spatial filters	Overfitting; montage dependence	Subject-wise split; parameter-matched baseline
RNN / temporal model	Sequential embeddings	Models temporal state	Optimization and latency	Causal inference-time test
GNN	Electrode graph or connectivity	Explicit topology	Arbitrary graph assumptions	Graph and montage ablations
Transformer / SSL	Tokens from large unlabeled corpora	Transfer and long context	Scale, leakage, hidden pretraining	External dataset and prior-only ablations

Family	Typical input	Advantages	Key risks	Minimum comparison
			overlap	
Generative decoder	Neural latent + language/image prior	Open-ended output	Hallucination; prior dominance	Shuffled-signal, semantic and human evaluation

VI. PERFORMANCE EVALUATION AND REPRODUCIBILITY

A. Metrics Must Match the Task

Discrete selection should report balanced accuracy, sensitivity, specificity, precision, recall, F1 score, confusion matrices, and calibration—not accuracy alone. Continuous control requires trajectory error, success rate, time-to-target, stability, and failure recovery. Spellers and command systems often report information-transfer rate (ITR), but ITR can be misleading when assumptions about independent selections, error correction, dwell time, or asynchronous operation are violated. Clinical studies should include validated functional outcomes and adverse events in addition to decoder accuracy.

$$ITR = (60/T)[\log_2 N + P \log_2 P + (1 - P) \log_2 ((1 - P)/(N - 1))] \quad (3)$$

In (3), N is the number of alternatives, P is selection accuracy, and T is seconds per selection. Comparisons are meaningful only when trial timing, rejected selections, and correction procedures are disclosed.

B. Leakage-Resistant Validation

Randomly splitting overlapping windows from a continuous recording can place nearly identical samples in both training and test sets. The correct unit of independence is determined by the intended deployment claim. A within-session claim requires held-out trials; a cross-session claim requires complete held-out sessions; and a cross-subject claim requires participants absent from all training and hyperparameter selection. Preprocessing, normalization, channel selection, augmentation, early stopping, and threshold selection must occur inside the training fold. Test data should be touched once after model selection.

C. Reporting Checklist

- Identify participants, inclusion criteria, clinical status, demographics, attrition, ethics approval, and consent.
- Report hardware, montage, sampling, references, timing, artifact policy, preprocessing, rejected data, and software versions.

- Describe the independent unit, split construction, nested model selection, random seeds, confidence intervals, and statistical tests.
- Publish task definitions, label provenance, baselines, ablations, error cases, computational cost, latency, and code or pseudocode where permissible.
- Separate offline decoding, simulated online testing, supervised online use, and independent home or clinical deployment.

VII. APPLICATIONS AND TRANSLATION READINESS

A. Communication and Assistive Control

P300 and SSVEP spellers provide communication channels when reliable eye movement or muscle control is limited. Motor imagery can control cursors, wheelchairs, robotic devices, or environmental systems, but safe real-world control generally requires shared autonomy, obstacle avoidance, confirmation states, and an emergency stop. Language-model assistance can improve typing efficiency but must not overwrite the user's intended message. Interfaces should make uncertainty visible and allow rapid correction.

B. Neurorehabilitation

In rehabilitation, the purpose of decoding is often to deliver temporally contingent feedback or stimulation rather than to maximize command bandwidth. BCI-triggered functional electrical stimulation, robotics, virtual reality, or sensory feedback may strengthen the association between intended movement and peripheral consequences. Controlled studies have reported benefits after stroke, but heterogeneity in dose, impairment, comparator therapy, and outcome timing limits generalization [20]. BCI should complement—not replace—clinical assessment and evidence-based therapy.

C. Cognitive and Affective Monitoring

EEG and fNIRS have been investigated for workload, fatigue, vigilance, attention, and affect. These states are context-dependent constructs rather than directly observable brain labels. Models can learn task, movement, facial muscle, lighting, and demographic proxies. Workplace or educational monitoring therefore poses scientific and ethical risks even when technically feasible. High-stakes inference should require a validated use case, proportionality, consent, uncertainty reporting, and a meaningful right to refuse.

D. HCI, VR/AR, Gaming, and Creative Tools

Non-invasive BCI can supplement conventional input with hands-free selection, adaptive difficulty, neurofeedback, and accessibility features. The most robust products are likely multimodal: eye tracking, electromyography, inertial sensors, speech, and conventional

controls can compensate for neural uncertainty. Claims should distinguish passive physiological sensing from intentional BCI control and should avoid implying access to precise private thoughts.

TABLE III. APPLICATION READINESS AND PRIMARY EVIDENCE GAPS.

Application	Common paradigm	Readiness	Primary gap	Safety requirement
Laboratory spelling	P300 / SSVEP	Established research use	Home robustness and fatigue	Error correction; user confirmation
Assistive device control	MI / SSVEP / hybrid	Prototype-clinical research	Longitudinal reliability	Shared autonomy; emergency stop
Stroke rehabilitation	MI + feedback/FES/robotics	Clinical research	Large multicenter effectiveness	Therapist oversight; adverse-event tracking
Workload/fatigue	Passive EEG/fNIRS	Research-limited deployment	Construct validity and context shift	No punitive automated decisions
Open-vocabulary text	Reading/covert-task EEG	Early research	Independent semantic evidence	Prevent overclaiming and involuntary inference
Visual reconstruction	Stimulus-evoked fMRI	Early research	Generalization and privacy	Strict consent and controlled access

VIII. TECHNICAL, CLINICAL, AND HUMAN-FACTOR CHALLENGES

A. Physical Information Limits

Skull attenuation, volume conduction, hemodynamic delay, limited depth sensitivity, and environmental interference impose genuine information constraints. More parameters cannot recreate spatial or temporal detail absent from the measurement. AI can exploit structure and priors, denoise observations, and combine modalities, but uncertainty should increase when neural evidence is weak. The distinction between reconstructing a likely stimulus and directly measuring a cognitive representation is fundamental.

B. Non-Stationarity and Calibration

BCI distributions drift within a session and across days. Sources include physiology, attention, fatigue, learning, sensor displacement, impedance, medication, and environment.

Calibration burden discourages clinical and home adoption, while aggressive adaptation may silently fail when labels are unavailable. Practical systems require signal-quality gating, drift alarms, conservative adaptation, and periodic human-confirmed recalibration.

C. Human Learning, Fatigue, and Accessibility

The user and decoder co-adapt. Instructions, feedback latency, reward, fatigue, frustration, and perceived agency change neural strategy. A participant who performs poorly with one paradigm should not be labeled “BCI illiterate”; the paradigm, hardware, feedback, or training protocol may be inaccessible. Inclusive design should accommodate sensory, cognitive, motor, linguistic, and cultural differences and should report non-completers rather than excluding them from performance summaries.

D. Clinical Evidence and Operational Integration

Clinical value depends on more than statistical significance. A system must fit staffing, infection-control, setup-time, cleaning, reimbursement, maintenance, cybersecurity, and caregiver workflows. Outcomes should be meaningful to patients and sustained beyond supervised demonstrations. False activations, skin irritation, fatigue, frustration, and opportunity cost should be measured. Model updates after deployment require controlled change management and post-market monitoring.

IX. ETHICS, SECURITY, PRIVACY, AND GOVERNANCE

A. Why Neural Data Require Enhanced Protection

Neural recordings may reveal health conditions, cognitive states, identity-linked patterns, or sensitive inferences beyond the immediate control task. Risk arises not only from raw signals but also from embeddings, model parameters, inferred labels, and future reinterpretation. Consent to operate a device is not automatically consent to secondary research, advertising, employment assessment, or model training. Data minimization, purpose limitation, short retention, encryption, access logging, and deletion rights should be default engineering requirements [21], [22].

B. Security Threat Model

Threats include sensor spoofing, adversarial perturbation, model extraction, membership inference, poisoned updates, unauthorized cloud access, and manipulation of feedback. Safety-critical systems should authenticate sensor streams, isolate control channels, use signed updates, monitor distribution shift, and fail safely. Federated or distributed learning can reduce centralization but does not by itself guarantee privacy; gradients and updates can

leak information unless combined with secure aggregation and carefully calibrated privacy mechanisms.

C. Emerging Governance

The GDPR regulates personal data and provides additional safeguards for special categories of data in the European Union [23]. The EU AI Act adds risk-based duties and prohibits or constrains particular manipulative and emotion-recognition uses [24]. Chile’s constitutional and statutory initiatives brought neurorights into national debate, while Colorado explicitly added neural data to its privacy framework [25], [26]. UNESCO adopted the Recommendation on the Ethics of Neurotechnology in 2025, emphasizing dignity, mental privacy, autonomy, fairness, and responsible innovation [27]. These instruments differ in scope and enforceability, but collectively signal that technical capability does not establish social legitimacy.

TABLE IV. TRUSTWORTHY BCI CONTROL MEASURES ACROSS THE DATA LIFECYCLE.

Lifecycle stage	Principal risk	Recommended controls
Collection	Excessive or covert acquisition	Task-limited channels; consent; visible recording state; local preprocessing
Storage	Breach and indefinite retention	Encryption; separation of identifiers; retention schedule; deletion verification
Model training	Leakage, bias, unauthorized reuse	Data governance; subject-wise evaluation; audit trail; privacy testing
Inference	Unsupported or discriminatory claims	Calibrated uncertainty; abstention; human review; use-case restrictions
Feedback/control	Unsafe action or manipulation	Confirmation; rate limits; safe state; emergency stop; clinical oversight
Model update	Poisoning or silent performance drift	Signed releases; validation gates; rollback; post-deployment monitoring

X. PROPOSED ARCHITECTURE FOR NEXT-GENERATION NON-INVASIVE BCI

A trustworthy next-generation system should be layered so that signal quality, representation learning, decision uncertainty, safety policy, and feedback remain independently testable. The proposed Adaptive Multimodal Trustworthy BCI (AMT-BCI) architecture is a design synthesis rather than a claim of experimentally validated superiority.

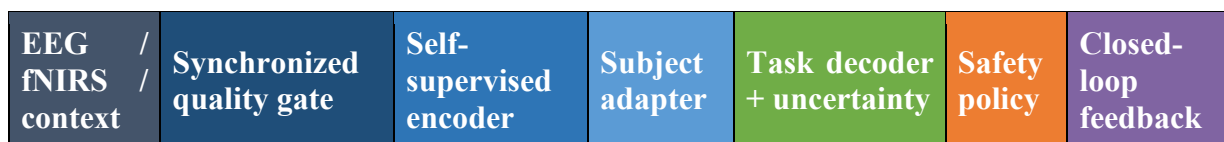


Fig. 3. Proposed AMT-BCI architecture. The safety-policy layer can abstain, request confirmation, or return the system to a safe state.

A. Layered Operation

- Multimodal acquisition: use the least intrusive sensing configuration sufficient for the task; auxiliary sensors are disclosed and separately governed.
- Quality gate: estimate contact quality, artifacts, packet loss, timing integrity, and out-of-distribution input before decoding.
- Pretrained representation: learn reusable temporal and spatial features from appropriately governed unlabeled data.
- Lightweight subject adapter: calibrate a small parameter set rather than retraining the full network; quantify residual shift.
- Task decoder with uncertainty: produce a command or semantic output together with confidence and an abstention option.
- Safety-policy engine: enforce task-specific constraints, confirmation requirements, rate limits, and emergency behavior.
- Closed-loop co-adaptation: deliver feedback designed for user learning while logging drift, fatigue, and corrections.

B. Evaluation Program

The architecture should be assessed in successive gates: offline subject-wise validation, cross-session replication, prospective online testing, supervised clinical or home pilots, and longitudinal deployment. Advancement requires predefined thresholds for task success, calibration time, abstention, latency, adverse events, fairness, privacy, and recovery from sensor or network failure. Performance should be compared with a non-BCI alternative because a technically impressive BCI may not improve the user's real task.

XI. RESEARCH GAPS AND FUTURE DIRECTIONS

A. Calibration-Light Generalization

The immediate goal is not universal zero-calibration decoding but predictable adaptation with bounded data and time. Multisite pretraining, montage-aware encoders, geometry-based alignment, and parameter-efficient adapters should be evaluated on participants, devices, and

institutions absent from development. Public leaderboards should hide complete subjects and sessions to discourage iterative leakage.

B. Causal and Multimodal Representation Learning

High-performing models may rely on ocular, muscular, stimulus, or environmental correlates. Causal interventions, modality dropout, sensor ablations, and source-informed architectures can test whether predictions depend on neural evidence. EEG–fNIRS fusion should be justified by incremental value under realistic movement, setup, and missing-modality conditions rather than by accuracy alone.

C. Generative Decoding with Evidence Attribution

For text and image generation, future systems should identify which output components are supported by neural evidence and which arise from the generative prior. Counterfactual and prior-only baselines, token-level uncertainty, semantic calibration, constrained decoding, and user confirmation are required. Evaluation must move beyond BLEU or visual similarity toward meaning preservation, factual control, error severity, and user agency.

D. Wearable and Edge Intelligence

Efficient architectures, low-power acquisition, on-device artifact suppression, and privacy-preserving inference can reduce latency and cloud exposure. Edge deployment also creates constraints: thermal comfort, battery life, memory, update security, and degraded-mode behavior. Hardware–algorithm co-design should optimize end-to-end clinical or user utility rather than isolated model accuracy.

E. Longitudinal Clinical and Participatory Research

BCI studies need larger, diverse cohorts and longer observation, but participation quality matters as much as sample size. Patients, caregivers, clinicians, disabled users, and affected workers or students should help define outcomes, acceptable trade-offs, consent processes, and failure modes. Participatory design can reveal when a proposed neural interface solves the wrong problem or creates burdens that an accuracy metric cannot capture.

XII. DISCUSSION

Non-invasive BCI is best understood as a portfolio of technologies rather than a single performance ladder. EEG is well matched to rapid cortical dynamics, fNIRS to slower localized hemodynamic changes, MEG to high-fidelity laboratory or specialized wearable measurement, and fMRI to spatially detailed scientific decoding. Hybridization is valuable only when complementary information survives the operational costs of additional sensors.

The strongest near-term progress will likely come from better systems engineering: robust acquisition, transparent quality control, physiologically informed representations, careful adaptation, uncertainty-aware decisions, and useful feedback. Large models may improve transfer and semantic output, but claims should remain proportional to the experimental task. In particular, stimulus reconstruction and language-assisted decoding do not justify claims of unconstrained thought reading.

A second conclusion is that evaluation design is part of the algorithm. Subject leakage, overlapping-window leakage, inconsistent baselines, and teacher-forced generation can create impressive but non-deployable results. The field should reward prospective, cross-session, and independent-site validation, even when headline accuracy is lower. Negative results and failure analysis are essential for patient safety and realistic translation.

Finally, privacy and agency cannot be postponed until commercialization. Neural data pipelines should be designed around purpose limitation, minimal collection, user control, security, and human oversight. Trustworthy BCI requires the capacity to abstain, explain operational confidence, resist unauthorized reuse, and preserve a user's ability to stop or correct the system.

XIII. CONCLUSION

Non-invasive BCIs have matured from laboratory demonstrations into a broad research platform spanning communication, rehabilitation, cognitive assessment, human–computer interaction, and generative neural decoding. EEG remains the practical core; fNIRS, MEG, fMRI, and hybrid configurations expand the accessible spatial, temporal, and physiological information. Artificial intelligence has improved feature learning and opened new output spaces, yet sensing physics, non-stationarity, dataset scale, and human variability continue to set hard constraints.

The next generation should be judged not solely by offline accuracy but by calibration burden, longitudinal reliability, uncertainty, clinical utility, privacy, safety, fairness, and user agency. The proposed AMT-BCI roadmap combines quality-gated multimodal sensing, self-supervised representations, lightweight personalization, uncertainty-aware decoding, explicit safety policy, and closed-loop co-adaptation. Progress along this path can make non-invasive BCI more useful without overstating what current systems can infer from the human brain.

REFERENCES

1. J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," *Clinical Neurophysiology*, vol. 113, no. 6, pp. 767–791, 2002, doi: 10.1016/S1388-2457(02)00057-3.
2. L. F. Nicolas-Alonso and J. Gomez-Gil, "Brain computer interfaces, a review," *Sensors*, vol. 12, no. 2, pp. 1211–1279, 2012, doi: 10.3390/s120201211.
3. B. J. Edelman et al., "Non-invasive brain-computer interfaces: State of the art and trends," *IEEE Reviews in Biomedical Engineering*, vol. 18, pp. 26–49, 2025, doi: 10.1109/RBME.2024.3449790.
4. F. Lotte et al., "A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update," *Journal of Neural Engineering*, vol. 15, no. 3, 031005, 2018, doi: 10.1088/1741-2552/aab2f2.
5. V. J. Lawhern et al., "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *Journal of Neural Engineering*, vol. 15, no. 5, 056013, 2018, doi: 10.1088/1741-2552/aace8c.
6. R. T. Schirrmeister et al., "Deep learning with convolutional neural networks for EEG decoding and visualization," *Human Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, 2017, doi: 10.1002/hbm.23730.
7. D. Kostas, S. Aroca-Ouellette, and F. Rudzicz, "BENDR: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data," *Frontiers in Human Neuroscience*, vol. 15, 653659, 2021, doi: 10.3389/fnhum.2021.653659.
8. Y. Duan, J. Zhou, Z. Wang, Y.-K. Wang, and C.-T. Lin, "DeWave: Discrete EEG waves encoding for brain dynamics to text translation," in *Advances in Neural Information Processing Systems*, vol. 36, 2023.
9. H. Liu, D. Hajjaligol, B. Antony, A. Han, and X. Wang, "EEG2TEXT: Open vocabulary EEG-to-text decoding with EEG pre-training and multi-view transformer," *arXiv:2405.02165*, 2024.
10. Z. Chen et al., "Seeing beyond the brain: Conditional diffusion model with sparse masked modeling for vision decoding," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, 2023, pp. 22710–22720.
11. Y. Takagi and S. Nishimoto, "High-resolution image reconstruction with latent diffusion models from human brain activity," in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, 2023, pp. 14453–14463.

12. E. Niedermeyer and F. H. Lopes da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 5th ed. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2005.
13. E. Boto et al., “Moving magnetoencephalography towards real-world applications with a wearable system,” *Nature*, vol. 555, pp. 657–661, 2018, doi: 10.1038/nature26147.
14. R. Sitaram et al., “Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain–computer interface,” *NeuroImage*, vol. 34, no. 4, pp. 1416–1427, 2007, doi: 10.1016/j.neuroimage.2006.11.005.
15. S. Fazli et al., “Enhanced performance by a hybrid NIRS–EEG brain computer interface,” *NeuroImage*, vol. 59, no. 1, pp. 519–529, 2012, doi: 10.1016/j.neuroimage.2011.07.084.
16. G. Pfurtscheller and F. H. Lopes da Silva, “Event-related EEG/MEG synchronization and desynchronization: Basic principles,” *Clinical Neurophysiology*, vol. 110, no. 11, pp. 1842–1857, 1999, doi: 10.1016/S1388-2457(99)00141-8.
17. L. A. Farwell and E. Donchin, “Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials,” *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510–523, 1988, doi: 10.1016/0013-4694(88)90149-6.
18. X. Chen, Y. Wang, S. Gao, T.-P. Jung, and X. Gao, “Filter bank canonical correlation analysis for implementing a high-speed SSVEP-based brain–computer interface,” *Journal of Neural Engineering*, vol. 12, no. 4, 046008, 2015, doi: 10.1088/1741-2560/12/4/046008.
19. S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski, “Independent component analysis of electroencephalographic data,” in *Advances in Neural Information Processing Systems*, vol. 8, 1996, pp. 145–151.
20. A. Biasucci et al., “Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke,” *Nature Communications*, vol. 9, 2421, 2018, doi: 10.1038/s41467-018-04673-z.
21. M. Ienca and R. Andorno, “Towards new human rights in the age of neuroscience and neurotechnology,” *Life Sciences, Society and Policy*, vol. 13, 5, 2017, doi: 10.1186/s40504-017-0050-1.
22. OECD, *Recommendation of the Council on Responsible Innovation in Neurotechnology*, OECD/LEGAL/0457, 2019.

23. European Parliament and Council, Regulation (EU) 2016/679 (General Data Protection Regulation), Official Journal of the European Union, 2016.
24. European Parliament and Council, Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), Official Journal of the European Union, 2024.
25. Republic of Chile, Law No. 21.383, Amendment to the Political Constitution for the Protection of Brain Activity and Information, 2021.
26. Colorado General Assembly, House Bill 24-1058: Protect Privacy of Biological Data, 2024.
27. UNESCO, Recommendation on the Ethics of Neurotechnology, Paris, France, 2025.
28. S. Han et al., “Advancing brain–computer interfaces with generative AI: A review of state-of-the-art and future outlook,” *The Innovation Life*, vol. 4, 100198, 2026, doi: 10.59717/j.xinn-life.2026.100198.
29. S. Wang et al., “Non-invasive brain–computer interfaces: Converging frontiers in neural signal decoding and flexible bioelectronics integration,” *Nano-Micro Letters*, 2026, doi: 10.1007/s40820-025-02042-2.
30. N. Sharma et al., “Recent trends in EEG-based motor imagery signal analysis and recognition: A comprehensive review,” *IEEE Access*, vol. 11, pp. 80518–80542, 2023, doi: 10.1109/ACCESS.2023.3299497.