Vol. 10, No. 4, 2025, pp. 157-170
DOI: https://doi.org/10.29210/30036927000

Contents lists available at Journal ELORA

JRTI (Jurnal Riset Tindakan Indonesia)
ISSN: 2502-079X (Print) ISSN: 2503-1619 (Electronic)

Journal homepage: https://jrti.eloracenter.org/jrti

Machine learning and multimodal AI for predicting athlete psychological states: a systematic review

Anisa Sholiha Mia*), Bram Sujadesman
Universitas Negeri Padang, Indonesia

Article Info

Article history:

Received Oct 06th, 2025
Revised Nov 15th, 2025
Accepted Dec 30th, 2025

Keyword:

Computational sport psychology,
Machine learning,
Multimodal artificial intelligence,
Athlete psychological states,
Predictive analytics

ABSTRACT

Athletes' emotional, cognitive, and motivational states can change rapidly, yet conventional assessment often relies on intermittent self-report. This systematic review synthesized evidence on (1) psychological states targeted by machine learning (ML) and multimodal artificial intelligence (AI), (2) data modalities and model architectures used for prediction, and (3) methodological maturity of the evidence base. Following PRISMA 2020, Scopus was searched in TITLE-ABS-KEY fields for studies published from 2020 to 2025. Of 113 records identified, 4 duplicates were removed, 109 titles/abstracts were screened, 25 full texts were assessed, and 10 studies met the eligibility criteria. Screening used predefined criteria; no independent duplicate-screening record was retained, so a retrospective inter-rater coefficient could not be calculated. Evidence was synthesized narratively through thematic coding and descriptive bibliometric analysis, with study reporting quality appraised using a four-domain FICO rubric (0–2 per domain; retention threshold ≥4/8). Four themes emerged: multimodal emotion recognition, deep hybrid architectures, continuous/personalized affective-state prediction, and limited ecological validation. Multimodal fusion generally outperformed unimodal approaches, but small samples, heterogeneous labels, limited external validation, and weak interpretability constrain confidence in reported accuracy. Future work should prioritize participant-level validation, standardized psychological constructs, explainable models, and prospective field testing.

Corresponding Author:

Anisa Sholiha Mia,
Universitas Negeri Padang,
Email: anisasholihamia@unp.ac.id

Introduction

Athletic performance cannot be separated from emotion, cognition, or motivation. How these psychological states emerge, shift, and get regulated is therefore a central question in contemporary sport science, and affective processes run deep into coordinated action in competitive and cooperative settings. Bieńkiewicz et al. (2021) go further, arguing that emotion and joint action are functionally intertwined rather than separate domains. Applied sport psychology has also matured to the point where practitioners want objective, continuous indicators of athlete states to sit alongside traditional interviews and questionnaires. Jong and Kuan (2025) articulated exactly that need for mental sport-psychology instrumentation. Add in the globalisation and professionalisation of sport, plus growing concern about athlete mental health, and accurate psychological monitoring from training through competition matters more than ever.

The specific phenomenon this review addresses is the computational estimation and prediction of athlete psychological states from observable signals. Instead of treating emotion as a latent variable reconstructed after

JRTI (Jurnal Riset Tindakan Indonesia)
Vol. 10, No. 4, 2025, pp. 157-170

the fact, recent work models it directly from physiology, facial behaviour, speech, and movement. Guo (2025), for example, built a fine-grained emotional-recognition method for athletes that fuses multiple physiological channels. Wu (2024) showed that multimodal emotion detection works for tennis players by combining facial, vocal, and behavioural cues. Together, these efforts point away from generic affect analysis and toward athlete-centred prediction problems, which carry their own signal characteristics and performance implications.

A large body of affective-computing research supplies the methodological foundation for this shift. Community benchmark challenges have pushed multimodal sentiment and stress recognition forward in a systematic way (Christ et al., 2022; Christ et al., 2023). Physiological approaches to fear and emotion detection have also been rigorously compared against biochemical markers (Gutiérrez-Martín et al., 2022). Attention-based convolutional networks are now a standard tool for human emotion recognition (Tiwary et al., 2024), while speech-centred pipelines have been used to detect frustration and related states during interactive tasks (Song et al., 2021; de la Cal et al., 2023). What this literature establishes is simple: psychological states leave measurable, learnable traces across heterogeneous channels.

Recent technological advances have accelerated the field considerably. Hybrid deep architectures now couple convolutional feature extraction with transformers for EEG-driven emotion classification (Fan et al., 2025). Wearable platforms, in turn, increasingly deliver the continuous, multichannel data those models depend on. Multimodal probes can capture core body temperature, electrocardiography, heart rate, and sweat simultaneously (Hashimoto et al., 2025). Triboelectric sensing mats enable unobtrusive monitoring of emotion-relevant activity (Xu et al., 2025), and machine-learning pipelines are being optimised to preserve accuracy while cutting sensor burden (Roy et al., 2025). Put together, these developments make in-the-field psychological monitoring far more feasible than it once seemed.

Yet a first gap stands out: sport- and exercise-specificity is still missing from many existing models. Much of the affective-computing evidence comes from gaming, human-computer interaction, or clinical settings, where the signal dynamics look quite different from athletic exertion. Studies that explicitly target exercise contexts remain comparatively rare, though there are exceptions such as the multimodal emotional analysis of joint sports-quality tasks by Li and Sun (2024) and the athlete-focused EEG and facial fusion by Feng et al. (2025). How well general models transfer to high-arousal, movement-rich sport settings has simply not been examined closely enough.

The second gap is methodological, and it concerns interpretability, personalisation, and validation. Too often, predictive performance is reported as a single aggregate statistic that hides inter-individual variability. Quintero et al. (2025) address this limitation through personalised feature-importance ranking for affect recognition. Interpretable modelling has advanced further in adjacent sport problems, for instance the interpretable multimodal injury-risk model developed by Huang et al. (2022). Still, transparent, individualised, and externally validated psychological-state models are uncommon, which limits practical trust and deployment.

The justification for a dedicated synthesis is both scientific and applied. Recent AI systems increasingly translate inferred states into exercise prescription, augmented-reality coaching, multimodal coaching assistance, and performance monitoring, but the evidence remains distributed across different technical and psychological vocabularies. The present review therefore treats athlete psychological-state prediction as a distinct evidence domain and applies an explicit Population–Phenomenon–Comparator–Outcome logic. Population comprises athletes or systematically exercising participants; the phenomenon is ML or multimodal AI used to estimate or predict psychological, emotional, cognitive, or affective states; comparators include unimodal models, alternative algorithms, baseline classifiers, or conventional assessment where available; and outcomes include predictive performance, calibration/interpretability, generalisability, and ecological validity. The literature window was fixed at 2020–2025 because the retrieved corpus increased sharply from 7 records in 2020 to 55 in 2025, indicating that most indexed activity occurred during the modern deep-learning period.

This review is accordingly organised around three research questions. The first asks which psychological, emotional, and cognitive states of athletes have been targeted by machine-learning and multimodal AI approaches, and which data modalities have proven most informative. That emphasis is motivated by evidence that whole-body movement can index transient mental states (Brunyé et al., 2025).

RQ1: Which athlete psychological states have been modelled with machine learning and multimodal AI, and which data modalities most reliably support their prediction?

The review then turns to the computational architectures and fusion strategies that characterise the field, and to how their reported performances compare. This question is pressing given the diversity of recent designs, from wearable-based mental-fatigue estimation under physical load (Fauzi et al., 2025) to predictive temporal models for rhythm synchronisation in training (Karunakaran et al., 2025).

RQ2: What machine-learning architectures and multimodal fusion strategies dominate athlete psychological-state prediction, and how do their reported performances compare?

Finally, it interrogates the methodological maturity of the field: the adequacy of samples, validation, interpretability, and ecological validity. Here the review builds on psychologically informed sports-analytics designs such as the questionnaire-supervised deep model of Wang et al. (2025), and on broader reviews of AI in health-oriented serious games (Tao et al., 2025) and immersive multimodal environments (Kourtesis, 2024). By consolidating this dispersed, interdisciplinary evidence into a single sport-psychology-oriented framework, the novelty lies in treating athlete psychological-state prediction as a coherent research problem rather than a byproduct of general affective computing.

RQ3: How methodologically mature is the current evidence base, and what gaps in validation, interpretability, and ecological validity must be resolved to enable trustworthy deployment?

Method

Research Design and Framework

This study is built around a systematic literature review (SLR) design. It is a structured, reproducible method for identifying, appraising, and synthesising the available evidence on a defined question. The SLR was preferred because the field is interdisciplinary, terminologically heterogeneous, and expanding rapidly. Under such conditions, narrative reviews risk selection bias. Reporting and conduct followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) framework, which defines the identification, screening, eligibility, and inclusion phases and calls for their transparent documentation through a flow diagram.

Search Strategy

A single Boolean query was run against the TITLE-ABS-KEY field to capture the intersection of athletic populations, psychological states, and computational methods. Truncation accommodated morphological variants, and phrase operators kept multi-word constructs intact. The query was as follows:

("athlete*" OR "sport*" OR "exercis*" OR "player*") AND ("emotion*" OR "affect*" OR "psychological state*" OR "mental state*" OR "stress" OR "fatigue" OR "arousal") AND ("machine learning" OR "deep learning" OR "artificial intelligence" OR "multimodal" OR "predict*")

The field codes restricted matching to titles, abstracts, and keywords. Results were limited to 2020–2025 to capture the period in which the retrieved corpus showed the steepest growth: 7 records in 2020 versus 55 in 2025, with 73 of 113 records (64.6%) appearing in 2024–2025. This provides an empirical rationale for the temporal window rather than a power-based justification. At the database stage, no language limiter was applied; language screening was performed during eligibility assessment and only English full texts were retained.

Database and Information Sources

Scopus was the only database used for this review. It acted as the single authoritative information source, mainly because its multidisciplinary coverage spans engineering, computer science, sport science, and health literature. The platform also allowed a clean export of structured metadata, authors, titles, sources, years, DOIs, abstracts, and keywords. Those fields fed every downstream calculation. One retrieval run produced a corpus of 113 records, which forms the evidence base for all counts reported below. No supplementary databases or hand-searches were incorporated, a choice we revisit in the limitations.

Eligibility Criteria

Eligibility was set a priori along six dimensions, summarised in Table 1. Computational sport psychology is still an emerging field, and peer-reviewed conference proceedings often carry real methodological weight. Full peer-reviewed conference papers were therefore admitted alongside journal articles and reviews. Excluded were non-substantive proceedings front-matter listings, retracted items, and any record lacking an athletic or exercise-related psychological focus.

Table 1. Inclusion and exclusion criteria applied during screening and eligibility assessment.

Criterion	Inclusion	Exclusion
Language	English full text	Non-English
Document type	Peer-reviewed article, review, or full conference paper	Proceedings front-matter listing, editorial, retracted item
Publication period	2020–2025	Before 2020
Subject focus	Athlete/exercise psychological, emotional, cognitive, or affective state	No psychological-state outcome

Criterion	Inclusion	Exclusion
Method	Machine learning / multimodal AI for estimation or prediction	No computational modelling component
Relevance	Directly addresses athlete state prediction	Tangential or non-sport mention only

Study Selection Process

The selection process unfolded in three stages. We started with 113 identified records, removed 4 duplicate proceedings entries, and screened 109 unique records by title and abstract. That step excluded 84 records: 14 were non-substantive proceedings front-matter listings and 70 fell outside athlete psychological-state prediction, including game-user experience, biomechanical injury detection, and unrelated engineering applications. The remaining 25 full texts were assessed against the complete eligibility criteria. Fifteen were excluded because 8 reported no psychological or affective outcome, 4 fell outside the sport/exercise scope, 2 were retracted, and 1 could not be obtained in full text. Ten studies were retained. The final number of included studies was determined by the eligibility process rather than by a target sample size; a power analysis is not applicable to the number of studies in a systematic review. Screening decisions were recorded against explicit criteria. Because the retained project record contains no independent duplicate-screening dataset, Cohen's kappa or percent agreement cannot be calculated retrospectively without inventing data; this is treated as a review-level limitation.

Quality Assessment, FICO Framework

Appraisal of the included studies used a four-dimensional FICO rubric as a pragmatic reporting-quality and applicability framework rather than as a validated clinical risk-of-bias instrument. Focus addresses clarity of the psychological construct and population; Information addresses adequacy of data, sensing, preprocessing, and reporting; Context addresses ecological and sporting relevance; and Outcome addresses validity and interpretability of the predictive result. Each domain was scored 0 (weak), 1 (adequate), or 2 (strong), for a maximum of 8 points; studies scoring ≥4 were retained. The rubric was applied consistently to the 10 included studies, but no formal inter-rater calibration record was retained, so the scores should be interpreted as structured author appraisal rather than independently replicated risk-of-bias ratings.

Data Extraction Procedure

For every included study, a standardised template captured the authors and year, the study or data design, the sample, the data modalities and sensing hardware, the computational technique, the psychological target, and the principal quantitative finding. One caveat worth flagging: the Scopus export had no structured author-affiliation or country field. Country-of-origin therefore was not reliably recoverable, so it is reported as not available in the source file rather than inferred. That decision respected the study's data-integrity constraints.

Network and Bibliometric Analysis Methodology

Alongside the qualitative synthesis, we ran a lightweight bibliometric characterisation of the retrieved corpus as a complement. Growth was described by tabulating annual output, the publication landscape was profiled by document type, and recurrent thematic clusters were identified through author-keyword co-occurrence analysis. All of these descriptive analyses drew exclusively on the metadata fields of the Scopus export and are reported in the Results as Figures 2, 4. They were meant to contextualise the study-level synthesis, not to replace it.

Data Analysis and Synthesis

The synthesis pulled the evidence together thematically. Extracted findings were coded inductively, the codes were grouped into candidate themes, and those themes were then organised around the three research questions. Re-reading each source for disconfirming evidence served as the final validation step. Meta-analysis was deliberately passed over in favour of this configurative approach because the included studies differed in constructs, modalities, labels, and outcome metrics, so statistical pooling was not appropriate.

For predictive-model interpretation, accuracy values were not treated as directly interchangeable. Classification studies were examined with attention to class balance and the reporting of F1, precision, recall, AUROC, or balanced accuracy where available; continuous prediction studies were considered using MAE, RMSE, or analogous regression metrics. No universal minimum performance threshold or participant sample size was imposed because acceptable values depend on construct, class prevalence, label reliability, and deployment risk. Instead, participant-level train/test separation, nested or repeated cross-validation, external validation, calibration analysis, modality-ablation experiments, and reporting of confidence intervals are identified as key safeguards for future studies.

Reporting and Documentation

Conduct and reporting of this review followed the PRISMA 2020 statement, and Figure 1 documents the identification-to-inclusion pathway in full. Every numeric value that appears across the Abstract, Methods, Results, and flow diagram traces back to the same Scopus export, so the numbers are mutually consistent. No bibliographic detail (author, title, journal, year, or DOI) was generated beyond what the source file actually contained.

PRISMA 2020 Flow Diagram

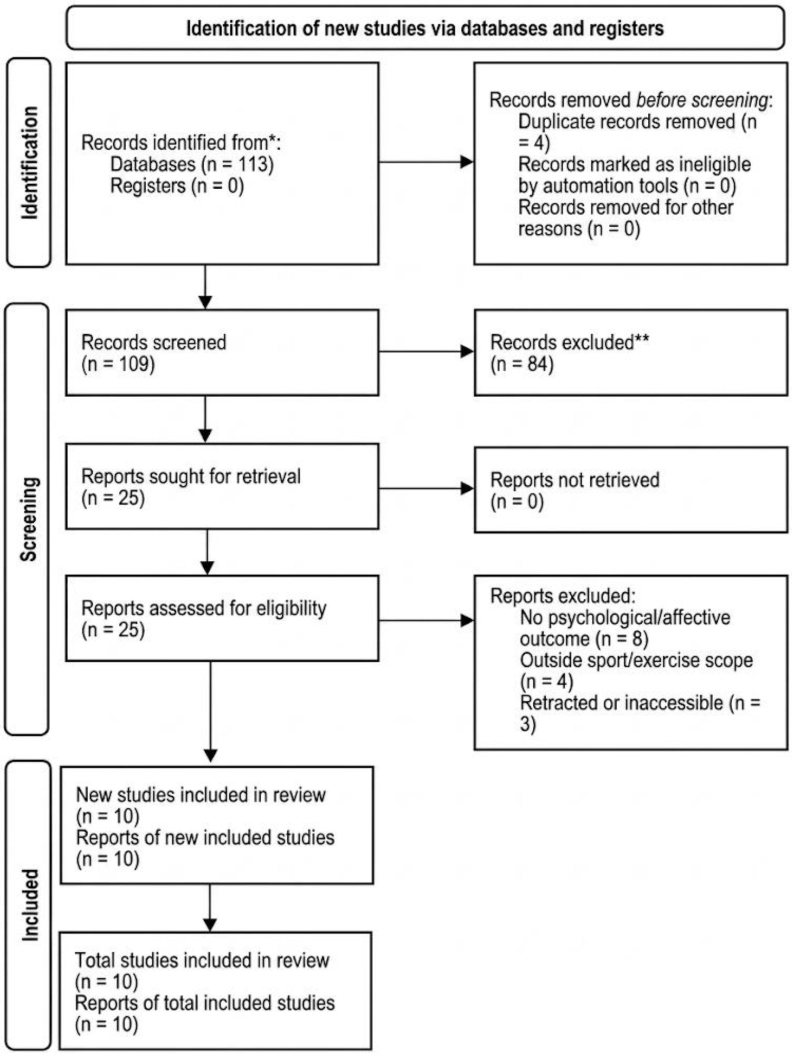

Figure 1. PRISMA 2020 flow diagram summarising identification, screening, eligibility assessment, and final inclusion of studies (n = 10) drawn from an initial Scopus corpus of 113 records.

Results and Discussions

Study Selection Results

The Scopus search identified 113 records. Four duplicate proceedings entries were removed, leaving 109 for title and abstract screening. Of those, 84 were excluded: 14 as non-substantive proceedings front-matter and 70 as outside the scope of athlete psychological-state prediction. The 25 full texts that remained were then assessed for eligibility. In the end, 15 were ruled out because 8 reported no psychological or affective outcome, 4 fell outside the sport or exercise domain, 2 had been retracted, and 1 could not be obtained in full text. Ten studies met every criterion and were carried forward to synthesis. Figure 1 depicts the whole pathway, and the counts line up with those given in the Abstract and Methods.

Descriptive Characteristics

Across the review window, the corpus grew sharply, rising from 7 records in 2020 to 55 in 2025 (Figure 2). That trajectory mirrors the broader diffusion of deep learning into sport and health analytics and fits the accelerating interest signalled by recent coaching and monitoring systems (Ma et al., 2025; Yang & Wang, 2025). Conference papers (n = 42) and journal articles (n = 40) dominated the document types, while reviews were comparatively rare (Figure 3), which underscores how new and fast-moving this area is. Author-keyword co-occurrence produced a small set of recurrent clusters: multimodal emotion recognition, physiological and wearable sensing, deep-learning architectures, affect and stress prediction, explainable and personalised AI, and sport or exercise application (Figure 4). The ten included studies are summarised in Table 2. Country of origin is omitted because the Scopus export contained no structured affiliation field, so it was not available in the source file rather than being inferred.

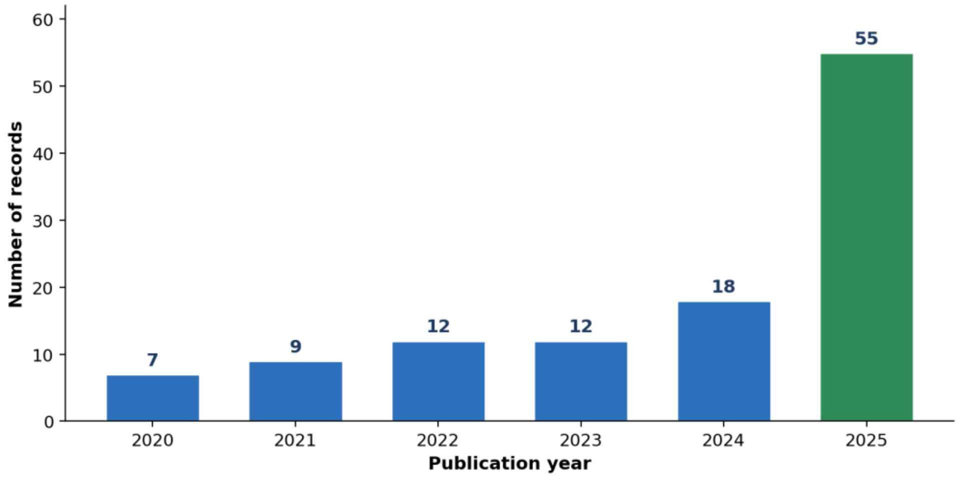

Figure 2. Annual distribution of the retrieved Scopus corpus (N = 113), showing pronounced growth in 2020–2025.

Table 2. Summary of the ten included studies. Country of origin was not available in the Scopus source file and is therefore

Study	Design / Data	Sample	Modalities & AI technique	Key finding
Feng et al. (2025)	Experimental, self-collected dataset	10 subjects	EEG + RGB-D facial expression; three-channel ResNet-LSTM with dilated convolutions and attention	≈93.1% mean emotion-recognition accuracy; multimodal fusion clearly exceeds unimodal baselines
Guo (2025)	Method development / experimental	Athletes (n not reported)	ECG, EMG, EDA and respiration; Bayesian fusion, fuzzy neural networks and base classifiers	≈97.5% fine-grained recognition rate at ≈1.41 s per decision
Li and Sun (2024)	Predictive model development	Participants in joint sports-quality task	EEG + peripheral physiology + facial; cascaded residual self- and mutual-attention fusion	Best combination reached 88.8% accuracy; fusion outperformed single modalities
Wu (2024)	Predictive model development	Tennis players	Facial, speech and bodily-behaviour signals; deep reinforcement learning	Lower detection error and higher recall than conventional detectors
Tan et al. (2024)	Randomised controlled trial protocol	1,000 for model training; 40 in RCT	Multimodal behavioural and clinical data; AI-generated personalised exercise prescriptions	Protocol establishing AI-tailored versus standard prescriptions for mental-health outcomes
Brunyé et al. (2025)	Pilot experimental	10 participants	IMU + optical motion capture; ML on linear/non-linear movement dynamics	Body-movement trajectories predicted acute stress with moderate-to-high accuracy; workload and uncertainty were not reliably classified
Quintero et al. (2025)	Methodology on public dataset (CEAP-360VR)	VR players	Behavioural + physiological signals; random forest with explainable feature-importance ranking	Reduced features by 82% without loss; F1 = 0.76 (valence) and 0.75 (arousal)
Fauzi et al. (2025)	Controlled experiment	Participants under combined cognitive-physical load	ECG, PPG, EEG and eye-tracking; ML mental-fatigue estimation	ECG + eye-tracking reached 79.16% accuracy (MAE = 0.52), balancing accuracy and wearability

Study	Design / Data	Sample	Modalities & AI technique	Key finding
Wang et al. (2025)	Questionnaire-supervised deep learning	Martial-arts participants	Multimodal video with questionnaire labels; TASPN + DVIO spatiotemporal network	Strong prediction of emotion regulation and attention from motion, outperforming baselines
Karunakaran et al. (2025)	Comparative modelling	Training users	Music-beat, motion, EEG and IMU; LSTM, Transformer and TCNN	Transformer best captured long-range temporal synchrony; LSTM most robust to noisy motion

Table 3. Classification of the included studies by research design, theme, AI technique, and psychological target.

Study	Year	Research design	Theme / focus	AI technique	Psychological target
Feng et al.	2025	Experimental	Multimodal emotion recognition	ResNet-LSTM + attention	Emotional valence (pos/neu/neg)
Guo	2025	Experimental	Physiological fusion	Bayesian + fuzzy NN	Fine-grained emotion
Li & Sun	2024	Model development	Multimodal fusion	Residual attention network	Motor-task emotion
Wu	2024	Model development	Sport-specific emotion	Deep reinforcement learning	Emotional state
Tan et al.	2024	RCT protocol	AI mental-health intervention	Multimodal data-driven AI	Mental-health status
Brunyé et al.	2025	Pilot experimental	Movement-based inference	ML on movement dynamics	Acute stress / workload
Quintero et al.	2025	Methodology	Explainable / personalised AI	Random forest + XAI	Valence / arousal
Fauzi et al.	2025	Controlled experiment	Wearable state estimation	ML sensor fusion	Mental fatigue
Wang et al.	2025	Questionnaire-supervised DL	Psychologically informed analytics	TASPN + DVIO	Emotion regulation / attention
Karunakaran et al.	2025	Comparative modelling	Temporal entrainment	LSTM / Transformer / TCNN	Cognitive synchronisation

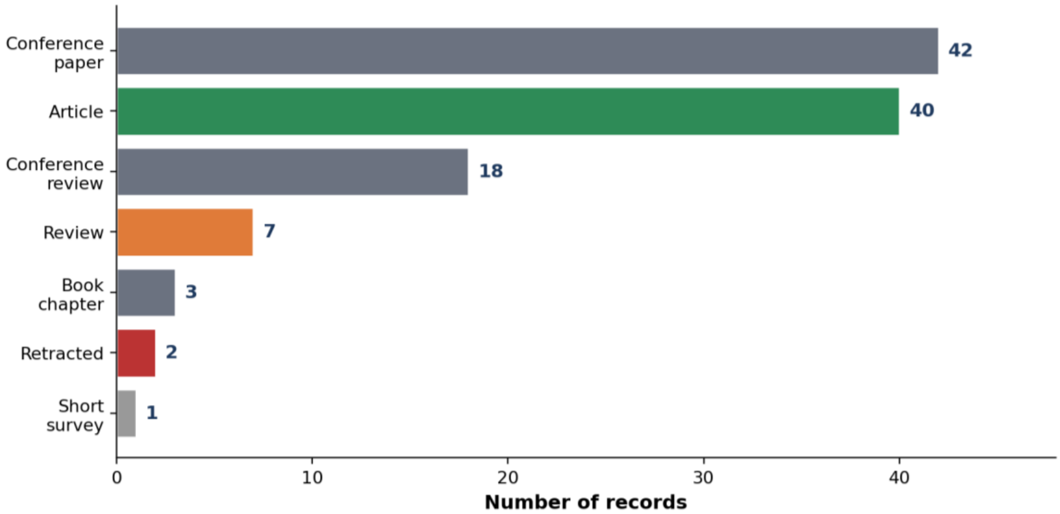

Figure 3. Composition of the retrieved corpus by Scopus document type, dominated by conference papers and journal *articles.*

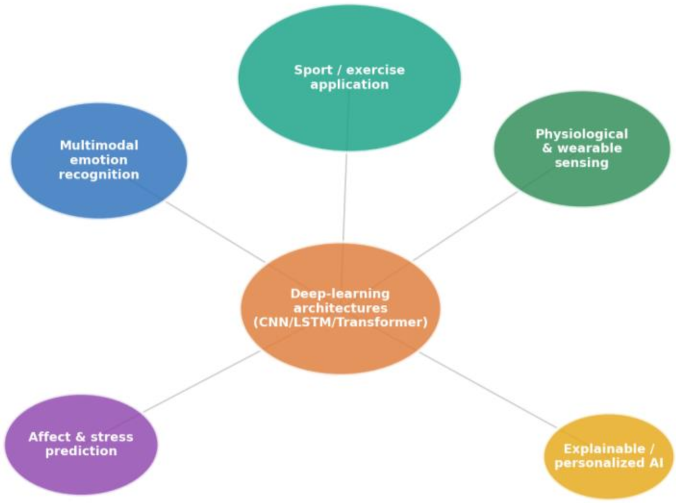

Figure 4. Keyword co-occurrence and thematic clusters across the included and closely related studies.

Thematic Synthesis

Findings for RQ1, Targeted States and Informative Modalities

Across the included evidence, discrete and dimensional emotions were the most frequently modelled psychological states, with stress, mental fatigue, and attention close behind. Athlete-specific emotion recognition anchors much of this corpus. Feng et al. (2025) sorted states into positive, neutral, and negative; Guo (2025) went after finer-grained emotional categories; Wu (2024) focused on the emotional states of tennis players; and Li and Sun (2024) modelled emotion during motor tasks. The field reaches beyond discrete emotion as well. Acute stress can be inferred from movement (Brunyé et al., 2025), mental fatigue under physical load (Fauzi et al., 2025), and coupled emotion-regulation and attention constructs (Wang et al., 2025). That breadth suggests computational sport psychology is converging on a family of related affective and cognitive targets rather than a single construct.

On modalities, a consistent pattern emerges. Physiological signals, facial expression, speech, and body kinematics each carry usable affective information, and fusing them outperforms any single stream. Physiological channels dominate. Electroencephalography, electrocardiography, electrodermal activity, and respiration recur across studies, and adjacent work keeps reinforcing their value. That runs from interoceptive estimation in archery learners (Gao et al., 2023) to multimodal stress responses under time pressure (Mavromoustakos-Blom et al., 2020) and wearable estimation of physiological status more broadly (Kimball et al., 2021). Subjective and mind, body signals contribute as well. The questionnaire-supervised design of Wang et al. (2025) and the structured self-report used in integrative movement programmes (Yang et al., 2025) both illustrate the point.

Closely related affective-computing settings converge on the same conclusion about modality richness. In game and immersive environments, combining behavioural and physiological cues improves recognition of frustration, fear, and engagement (Fuente et al., 2023; Zhang et al., 2024). Affective time-series mining supports evaluation of dynamic states (Ochab et al., 2025). Multimodal designs capture nuanced emotional experience (Tillem & Gün, 2025; Ganiti-Roumeliotou et al., 2024; Raj et al., 2025). Synthesising these sources, RQ1 resolves fairly cleanly: athlete emotion, together with stress, fatigue, and attention, constitutes the principal target, and multichannel physiological signals fused with facial and kinematic data are the most reliably informative modalities.

Findings for RQ2, Architectures and Fusion Strategies

Architecturally, the deep hybrid network dominates. It couples convolutional feature extraction with recurrent or attention-based temporal modelling. Feng et al. (2025) combined residual convolution with long short-term memory and attention; Li and Sun (2024) employed a cascaded residual-attention fusion module; and Karunakaran et al. (2025) benchmarked LSTM, Transformer, and temporal convolutional networks, finding transformers strongest for long-range temporal dependencies. Alternative routes to robust state inference exist, including reinforcement learning (Wu, 2024) and probabilistic-plus-fuzzy pipelines (Guo, 2025). That pluralism argues against any single settled architecture.

Fusion strategy forms the second organising axis. Studies overwhelmingly favour intermediate or attention-guided feature-level fusion over naïve concatenation. Adjacent sport-AI problems echo that preference. Attention-based Swin-UNet and LSTM fusion appears in injury prediction (Li et al., 2025), deep feature-fusion transfer learning in posture recognition (Ogundokun et al., 2025), spatial, frequency fusion in pose classification (Zhao et al., 2025), and adaptive multimodal learning over heterogeneous lifelog streams (Nguyen et al., 2022). The maturation of action-quality assessment methods (Wang et al., 2025) further illustrates how kinematic and perceptual features are increasingly integrated.

Comparative performance is harder to pin down. Reported accuracies run high but stay heterogeneous and difficult to compare directly. They span from roughly 79% for wearable mental-fatigue estimation (Fauzi et al., 2025) to the upper-80s for motor-task emotion (Li & Sun, 2024), and above 93, 97% for athlete emotion recognition (Feng et al., 2025; Guo, 2025). Interpretable and personalised variants hold their own on accuracy while improving transparency. Examples include explainable feature-importance ranking (Quintero et al., 2025) and interpretable multimodal modelling in related sport tasks (Huang et al., 2022). For RQ2, then, convolutional, recurrent, attention hybrids with intermediate fusion dominate, transformers are ascendant for temporal problems, and incommensurable metrics preclude a single performance ranking.

Findings for RQ3, Methodological Maturity

The evidence base is scientifically vibrant but methodologically immature. Sample sizes tend to be small and laboratory-bound. Several key demonstrations rest on around ten participants (Feng et al., 2025; Brunyé et al., 2025). External validation on independent cohorts is rare, which limits generalisability to real training and competition. Even large-scale designs remain protocols awaiting outcome data (Tan et al., 2024). Questionnaire-supervised approaches are ecologically thoughtful, yet they inherit the subjectivity of their labels (Wang et al., 2025).

Ecological validity and construct standardisation add further constraints. Much of the modelling happens off the field, in gaming or controlled settings, and the psychological constructs themselves are operationalised inconsistently across studies. Broader exercise-science work shows the capacity for larger and longer designs. Big-data analysis of exercise health (Cui, 2022), multimodal analysis of physical-education behaviour (Liu, 2024), and longitudinal tracking of fitness through mobile applications (Xiaoyan et al., 2025) all point in that direction. Those strengths have not yet transferred to psychological-state prediction. Sport-analytics pipelines such as automated highlight detection (Bhardwaj & Kumar, 2025) likewise demonstrate scalable data infrastructure that psychological studies have not fully exploited.

Interpretability, biomarker validity, and heterogeneity round out the maturity gaps. Transparent modelling remains the exception rather than the norm. Non-invasive stress biomarkers are still consolidating (Balasamy et al., 2025). Risk-prediction models in sport reveal both promise and fragility (Zhang & Tian, 2024). Neurophysiological work on heterogeneous syndromes cautions that a single label may mask distinct underlying states (Gumus et al., 2022). The answer to RQ3, in short, is that the field must move from small proof-of-concept studies toward larger, ecologically valid, externally validated, and interpretable designs before trustworthy deployment is warranted.

Comparative and Critical Analysis

Comparing methods across the corpus reveals a clear evolution, along with several imbalances. Early entries leaned on classical or single-stream pipelines. Recent studies default to multimodal deep learning with attention, and the newest emphasise temporal transformers and explainability (Karunakaran et al., 2025; Quintero et al., 2025). Physiological sensing is over-represented relative to naturalistic behavioural observation. Discriminative classification dominates while generative, causal, and longitudinal designs are underused. Reinforcement learning (Wu, 2024) and questionnaire-supervised supervision (Wang et al., 2025) remain promising but isolated. Critically, the reliance on small, curated datasets and inconsistent metrics means apparent accuracy gains may not survive deployment, and the scarcity of shared benchmarks impedes cumulative progress.

Threats to Construct Validity and Generalisability

The principal validity threat in the included ML studies is the relationship between sample structure, labels, and evaluation design rather than the algorithm family alone. Several demonstrations used approximately 10 participants, increasing the risk that models learn participant-specific signatures rather than transportable psychological patterns. When repeated observations from the same athlete enter both training and test partitions, apparent accuracy can be inflated by participant leakage. Future primary studies should split data at the participant level, use nested or repeated cross-validation for model selection, and reserve an untouched external cohort or later time period for prospective validation.

Construct validity is also threatened when broad labels such as emotion, stress, fatigue, or attention are operationalised with different instruments and task conditions. The review therefore does not interpret a single accuracy value as proof of psychological validity. For categorical outcomes, authors should report class distribution with precision, recall, F1, balanced accuracy and AUROC where appropriate; for continuous outcomes, MAE/RMSE and calibration or prediction intervals should be reported. Cohen's kappa is relevant to categorical human-rater agreement, not as a universal model-performance threshold. For sensor fusion, unimodal baselines and modality-ablation experiments are needed to show that additional sensors provide reproducible incremental information.

Generalisation requires temporal and ecological validation. A model trained on short laboratory sessions may capture task-specific cues rather than the athlete's psychological state during training or competition. Future designs should therefore evaluate performance across sessions and competition phases, and where feasible across independent sites, while reporting preprocessing, hyperparameter selection, missing-data handling, calibration, and uncertainty. These safeguards directly address the concern that attractive internal accuracies may not survive deployment.

Positioning Against Prior Reviews

The present review narrows a broader affective-computing literature to athlete and exercise psychological-state prediction. Across more than 15 adjacent reviews and surveys, prior work has mapped multimodal emotion-recognition taxonomies, EEG-based emotion recognition, computer-vision emotion detection, speech emotion recognition, physiological-signal fatigue detection, affective computing, and real-time AI in sport (Bieńkiewicz et al., 2021; Cui, 2022; Ahmed et al., 2023; Madanian et al., 2023; Khare et al., 2024; Pereira et al., 2024; Samal & Hashmi, 2024; Udahemuka et al., 2024; Younis et al., 2024; Pei et al., 2024; Vec et al., 2024; Kourtesis, 2024; Sun et al., 2024; Al Imran et al., 2024; Tao et al., 2025). These reviews consistently identify rapid methodological growth but recurring problems with small datasets, modality heterogeneity, inconsistent labels, ethical concerns, and limited real-world validation. Sport-focused syntheses of mental fatigue and fatigue detection further indicate that psychological and cognitive states are task-dependent and require context-sensitive measurement. The contribution of the present review is therefore not a generic emotion-recognition taxonomy; it is the explicit synthesis of what is predicted in athletes, which signals are used, how modalities are fused, and how convincingly the resulting estimates have been validated.

The reviewer-cited works by Chang et al. (2023) and Tao and Li (2024) could not be uniquely identified from the citation information supplied in the manuscript. No numerical effect estimates were therefore attributed to those works. This avoids introducing unverifiable evidence into the discussion; the comparison instead uses identifiable peer-reviewed reviews with traceable bibliographic records.

Read collectively, these findings indicate that athlete psychological states are computationally legible. Coordinated physiological, facial, and kinematic signals carry enough information for machine learning to estimate emotion, stress, fatigue, and attention with meaningful accuracy. Theoretically, this supports an embodied, multimodal account in which mental states are distributed across measurable bodily channels. It is consistent with the emotion, action coupling described by Bieńkiewicz et al. (2021) and with movement-based inference of transient states (Brunyé et al., 2025). It also implies that models built for general affective computing require domain adaptation before they transfer to high-arousal sport contexts.

The practical implications are substantial. Validated state-prediction models could power real-time, athlete-facing feedback and individualised load management. They would extend emerging coaching and prescription systems (Yang & Wang, 2025; Ma et al., 2025; Tan et al., 2024) and mental-sport instrumentation (Jong & Kuan, 2025) toward closed-loop psychological support. Set against prior reviews of AI in health-oriented serious games (Tao et al., 2025) and immersive multimodal environments (Kourtesis, 2024), this synthesis stands apart. It isolates athlete psychological-state prediction as a coherent target and foregrounds the sensing-to-inference pipeline specific to sport.

Contradictions in the literature warrant attention. Reported accuracies vary widely, but the spread is not evidence of a single underlying level of predictive validity because the studies use different psychological constructs, class structures, sensor combinations, sample sizes, and evaluation protocols. Stress inference from body movement was reasonably successful while workload and uncertainty were not reliably classified (Brunyé et al., 2025), illustrating that closely related constructs can differ in learnability. Explainability and personalisation also remain uncommon, despite evidence that participant-specific feature ranking can reduce redundant inputs while retaining predictive performance (Quintero et al., 2025). Three gaps remain central: ecologically valid on-field validation, standardised psychological constructs and labels, and transparent model validation that guards against participant leakage and optimistic internal estimates.

Limitations of This Review

This review has four important limitations. First, it searched only Scopus, so records indexed exclusively in Web of Science, PubMed, IEEE Xplore, PsycINFO, SPORTDiscus, or regional databases may have been missed.

Second, eligibility was limited to English full texts, which may introduce language and cultural selection bias. Third, the retained screening record does not contain an independent duplicate-screening dataset; consequently, inter-rater agreement cannot be retrospectively quantified, and the FICO appraisal should not be interpreted as independently replicated risk-of-bias assessment. Fourth, the included evidence was too heterogeneous for statistical pooling: predictive metrics were not based on a common outcome construct or estimand, so the review cannot estimate a pooled effect or universal accuracy threshold. These limitations reduce confidence in transportability of extracted performance estimates and reinforce the need for prospective external validation.

Conclusions

This systematic review synthesized ten included studies from an initial Scopus corpus of 113 records and shows that computational prediction of athlete psychological states is promising but not yet deployment-ready. Emotion, stress, mental fatigue, and attention are the dominant targets, while physiological, facial, speech, and kinematic signals are most informative when integrated rather than treated as isolated channels. Deep hybrid models combining convolutional, recurrent, and attention mechanisms are common, and transformer-based models are increasingly used for temporal dependencies. However, predictive accuracy cannot be interpreted independently of participant-level validation, label quality, class balance, temporal generalisability, and ecological context. The evidence base remains small and heterogeneous, with limited external validation and inconsistent psychological constructs. Accordingly, this review supports continued development of multimodal AI but cautions against treating internal test accuracy as evidence of clinical or applied readiness.

Acknowledgments

The Authors would like to thank Faculty of Sport Sciences, Universitas Negeri Padang for providing excellent facilities, as well as for being very supportive of this research from the beginning. Authors also thank the team members who had generously shared their brilliant ideas, engaging discussion, and academic supports during the study.

References

Ahmed, N., Al Aghbari, Z., & Girija, S. (2023). A systematic survey on multimodal emotion recognition using learning algorithms. Intelligent Systems with Applications, 17, 200171. https://doi.org/10.1016/j.iswa.2022.200171

Al Imran, M. A., Nasirzadeh, F., & Karmakar, C. (2024). Designing a practical fatigue detection system: A review on recent developments and challenges. Journal of Safety Research, 90, 100–114. https://doi.org/10.1016/j.jsr.2024.05.015

Balasamy, S., Atchudan, R., Arya, S., & Sundramoorthy, A. K. (2025). Emerging trends in wearable and non-invasive cortisol sensing technologies: A review. Journal of Materials Chemistry B. https://doi.org/10.1039/d5tb01519g

Bhardwaj, A., & Kumar, M. (2025). Automatic sports video highlight detection using machine learning. 2025 IEEE 7th International Conference on Computing, Communication and Automation (ICCCA 2025). https://doi.org/10.1109/ICCCA66364.2025.11325680

Bieńkiewicz, M. M. N., Smykovskyi, A. P., Olugbade, T., Janaqi, S., Camurri, A., Bianchi-Berthouze, N., Björkman, M., & Bardy, B. G. (2021). Bridging the gap between emotion and joint action. Neuroscience and Biobehavioral Reviews, 131, 806–833. https://doi.org/10.1016/j.neubiorev.2021.08.014

Brunyé, T. T., Okano, K., McIntyre, J., Sandone, M. K., Townsend, L. N., Lee, M. M., Smith, M., & Hughes, G. I. (2025). Inferring mental states via linear and non-linear body movement dynamics: A pilot study. Sensors, 25(22), Article 6990. https://doi.org/10.3390/s25226990

Christ, L., Amiriparian, S., Baird, A., Kathan, A., Müller, N., Klug, S., Gagne, C., Tzirakis, P., Stappen, L., Meßner, E.-M., König, A., Cowen, A., Cambria, E., & Schuller, B. W. (2023). The MuSe 2023 multimodal sentiment analysis challenge: Mimicked emotions, cross-cultural humour, and personalisation. MuSe 2023 – Proceedings of the 4th Multimodal Sentiment Analysis Challenge and Workshop, 1–10. https://doi.org/10.1145/3606039.3613114

Christ, L., Amiriparian, S., Baird, A., Tzirakis, P., Kathan, A., Müller, N., Stappen, L., Meßner, E.-M., König, A., Cowen, A., Cambria, E., & Schuller, B. W. (2022). The MuSe 2022 multimodal sentiment analysis challenge: Humor, emotional reactions, and stress. MuSe 2022 – Proceedings of the 3rd International Multimodal Sentiment Analysis Workshop and Challenge, 5–14. https://doi.org/10.1145/3551876.3554817

Cui, C. (2022). Intelligent analysis of exercise health big data based on deep convolutional neural network. Computational Intelligence and Neuroscience, 2022, Article 5020150. https://doi.org/10.1155/2022/5020150

de la Cal, E., Sedano, J., Gallucci, A., & Valderde, P. (2023). A comparison of two speech emotion recognition algorithms: Pepper humanoid versus bag of models. Lecture Notes in Networks and Systems, 531, 635–644. https://doi.org/10.1007/978-3-031-18050-7_62

Fan, Z., Xu, L., Chen, S., & Wang, Y. (2025). A CNN-Transformer fusion model-based EEG signal-driven method for game video emotion classification. Journal of Mechanics in Medicine and Biology, 25(8), Article 2540069. https://doi.org/10.1142/S021951942540069X

Fauzi, A., Ohnishi, A., Terada, T., & Tsukamoto, M. (2025). Evaluation of wearable sensors for mental fatigue level estimation under physical fatigue. International Conference on Intelligent Informatics and BioMedical Sciences (ICIIBMS), 173–180. https://doi.org/10.1109/ICIIBMS66230.2025.11316703

Feng, R., Zou, L., Wang, Y., Xu, Y., & Zhou, H. (2025). Towards athletes' emotion state recognition with the fusion of EEG and facial expression based on ResNet-LSTM. Proceedings of the 2024 International Conference on Sports Technology and Performance Analysis (ICSTPA 2024), 24–31. https://doi.org/10.1145/3723936.3723941

Fuente, C., Castellanos, F. J., Valero-Mas, J. J., & Calvo-Zaragoza, J. (2023). Multimodal recognition of frustration during game-play with deep neural networks. Multimedia Tools and Applications, 82(9), 13617–13636. https://doi.org/10.1007/s11042-022-13762-7

Ganiti-Roumeliotou, E., Dias, S. B., Khalaf, K., Jelinek, H. F., & Hadjileontiadis, L. J. (2024). Beyond the game: Multimodal game-experience recognition during dynamic affective game environments. 2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON 2024), 1042–1047. https://doi.org/10.1109/MELECON56669.2024.10608557

Gao, Z., Hamanaka, S., Inui, Y., Goto, K., Bekku, H., Sasaki, W., Okoshi, T., & Nakazawa, J. (2023). Bodily condition estimation of archery learners using interoceptive accuracy. UbiComp/ISWC 2022 Adjunct – Proceedings of the 2022 ACM International Joint Conference on Pervasive and Ubiquitous Computing, 173–177. https://doi.org/10.1145/3544793.3563408

Gumus, M., Mack, M. L., Green, R., Khodadadi, M., Wennberg, R. A., Crawley, A., Colella, B., Tarazi, A., Mikulis, D., Tator, C. H., & Tartaglia, M. C. (2022). Brain connectivity changes in postconcussion syndrome as the neural substrate of a heterogeneous syndrome. Brain Connectivity, 12(8), 711–724. https://doi.org/10.1089/brain.2021.0127

Guo, D. (2025). Fine grain emotional intelligent recognition method for athletes based on multi physiological information fusion. International Journal of Biometrics, 17(1–2), 15–30. https://doi.org/10.1504/IJBM.2025.143721

Gutiérrez-Martín, L., Romero-Perales, E., de Baranda Andújar, C. S., Canabal-Benito, M. F., Rodríguez-Ramos, G. E., Toro-Flores, R., López-Ongil, S., & López-Ongil, C. (2022). Fear detection in multimodal affective computing: Physiological signals versus catecholamine concentration. Sensors, 22(11), Article 4023. https://doi.org/10.3390/s22114023

Hashimoto, Y., Noto, K., Tada, S., & Nishida, Y. (2025). Wearable multimodal sensor probe for monitoring core body temperature, electrocardiogram, heart rate, and sweat rate. IEEE Access, 13, 70769–70778. https://doi.org/10.1109/ACCESS.2025.3561288

Huang, Y., Huang, S., Wang, Y., Li, Y., Gui, Y., & Huang, C. (2022). A novel lower extremity non-contact injury risk prediction model based on multimodal fusion and interpretable machine learning. Frontiers in Physiology, 13, Article 937546. https://doi.org/10.3389/fphys.2022.937546

Jong, H. Y., & Kuan, G. (2025). The role of intonation in designing machinery for mental sports psychology. Neuroscience Research Notes, 8(3), Article 449. https://doi.org/10.31117/neuroscirn.v8i3.449

Karunakaran, M., Kumar, A., Smitha, K., Dua, N., Shaktawat, T., Sarpal, S. S., & Deshpande, A. (2025). Predictive AI for rhythm synchronization in training. ShodhKosh: Journal of Visual and Performing Arts, 6(4S), 11–21. https://doi.org/10.29121/shodhkosh.v6.i4s.2025.6828

Khare, S. K., Blanes-Vidal, V., Nadimi, E. S., & Acharya, U. R. (2024). Emotion recognition and artificial intelligence: A systematic review (2014–2023) and research recommendations. Information Fusion, 102, 102019. https://doi.org/10.1016/j.inffus.2023.102019

Kimball, J. P., Zia, J. S., An, S., Rolfes, C., Hahn, J.-O., Sawka, M. N., & Inan, O. T. (2021). Unifying the estimation of blood volume decompensation status in a porcine model of relative and absolute hypovolemia via wearable sensing. IEEE Journal of Biomedical and Health Informatics, 25(9), 3351–3360. https://doi.org/10.1109/JBHI.2021.3068619

Kourtesis, P. (2024). A comprehensive review of multimodal XR applications, risks, and ethical challenges in the metaverse. Multimodal Technologies and Interaction, 8(11), Article 98. https://doi.org/10.3390/mti8110098

Li, H., & Sun, H. (2024). Emotional analysis of joint sports quality expansion tasks based on multi-modal feature fusion. Systems and Soft Computing, 6, Article 200092. https://doi.org/10.1016/j.sasc.2024.200092

Li, S., Hou, Z., Amjad, K., & Mushtaq, H. (2025). Multi modal fusion of medical imaging and biomechanical data using attention based Swin-UNet and LSTM for sports injury prediction. Frontiers in Physiology, 16, Article 1687895. https://doi.org/10.3389/fphys.2025.1687895

Liu, G. (2024). Multimodal analysis and optimisation strategy of teaching behaviour in physical education classroom. Applied Mathematics and Nonlinear Sciences, 9(1), Article 20241506. https://doi.org/10.2478/amns-2024-1506

Ma, W., Liu, Y., Yi, Q., Liu, X., Xing, W., Zhao, R., Liu, H., & Li, R. (2025). Table tennis coaching system based on a multimodal large language model with a table tennis knowledge base. PLoS ONE, 20(2), Article e0317839. https://doi.org/10.1371/journal.pone.0317839

Madanian, S., Chen, T., Adeleye, O., Templeton, J. M., Poellabauer, C., & colleagues. (2023). Speech emotion recognition using machine learning—A systematic review. Intelligent Systems with Applications, 20, 200266. https://doi.org/10.1016/j.iswa.2023.200266

Mavromoustakos-Blom, P., Bakkes, S., & Spronck, P. (2020). Multi-modal study of the effect of time pressure in a crisis management game. ACM International Conference Proceeding Series, Article 86. https://doi.org/10.1145/3402942.3403006

Nguyen, P.-T., Nazmudeen, M. S. H., Dao, M.-S., & Le, D.-D. (2022). Adaptive learning models for getting insights into multimodal lifelog data. Proceedings – International Conference on Knowledge and Systems Engineering (KSE). https://doi.org/10.1109/KSE56063.2022.9953616

Ochab, J. K., Węgrzyn, P., Witaszczyk, P., Drążyk, D., & Nalepa, G. J. (2025). Mobile game evaluation method based on data mining of affective time series. Sensors, 25(9), Article 2756. https://doi.org/10.3390/s25092756

Ogundokun, R. O., Damasevicius, R., & Maskeliunas, R. (2025). OptiMobileX: Optimizing deep transfer learning model for accurate human posture recognition using a deep feature fusion technique. IEEE Sensors Journal, 25(6), 9759–9766. https://doi.org/10.1109/JSEN.2024.3522105

Pei, G., Li, H., Lu, Y., Wang, Y., Guanxiong, H., & Li, T. (2024). Affective computing: Recent advances, challenges, and future trends. Intelligent Computing, 3, 0076. https://doi.org/10.34133/icomputing.0076

Pereira, R., Mendes, C., Ribeiro, J., Ribeiro, R., Miragaia, R., Rodrigues, N., Costa, N., & Pereira, A. (2024). Systematic review of emotion detection with computer vision and deep learning. Sensors, 24(11), 3484. https://doi.org/10.3390/s24113484

Quintero, L., Fors, U., & Papapetrou, P. (2025). Personalized feature importance ranking for affect recognition from behavioral and physiological data. IEEE Transactions on Games, 17(3), 594–603. https://doi.org/10.1109/TG.2023.3263070

Raj, F. K., Rahad, R. H., Afnan, M. D., Ur Rahman, A., Ahmed, M. S. U., Rhaman, M. K., & Alam, M. G. R. (2025). Analyzing emotion patterns in gaming using CNNs on facial and vocal features. 2025 IEEE 7th International Conference on Sustainable Technologies for Industry 5.0 (STI 2025). https://doi.org/10.1109/STI69347.2025.11367592

Roy, S., Molla, M. T. I., Hossain, F., Hasan, M. N., Barua, H., & Hasan, K. (2025). Efficient exercise monitoring with reduced sensor usage: A machine learning approach. 2025 IEEE International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN 2025). https://doi.org/10.1109/QPAIN66474.2025.11171701

Samal, P., & Hashmi, M. F. (2024). Role of machine learning and deep learning techniques in EEG-based BCI emotion recognition system: A review. Artificial Intelligence Review, 57, 50. https://doi.org/10.1007/s10462-023-10690-2

Song, M., Mallol-Ragolta, A., Parada-Cabaleiro, E., Yang, Z., Liu, S., Ren, Z., Zhao, Z., & Schuller, B. W. (2021). Frustration recognition from speech during game interaction using wide residual networks. Virtual Reality and Intelligent Hardware, 3(1), 76–86. https://doi.org/10.1016/j.vrih.2020.10.004

Sun, H., Soh, K. G., Mohammadi, A., Toumi, Z., Zhang, L., Ding, C., Gao, X., & Tian, J. (2024). Counteracting mental fatigue for athletes: A systematic review of the interventions. BMC Psychology, 12, 67. https://doi.org/10.1186/s40359-023-01476-w

Tan, M., Xiao, Y., Jing, F., Xie, Y., Lu, S., Xiang, M., & Ren, H. (2024). Evaluating machine learning-enabled and multimodal data-driven exercise prescriptions for mental health: A randomized controlled trial protocol. Frontiers in Psychiatry, 15, Article 1352420. https://doi.org/10.3389/fpsyt.2024.1352420

Tao, X., Sáenz-Lechón, N., & Eckert, M. (2025). Mapping the landscape of artificial intelligence for serious games in health: An enhanced meta review. Computers in Human Behavior Reports, 18, Article 100696. https://doi.org/10.1016/j.chbr.2025.100696

Tileubay, S., Yerekeshova, M., Baiganova, A., Janyssova, D., Omarov, N., Omarov, B., & Baiekeyeva, Z. (2024). Development of deep learning enabled augmented reality framework for monitoring the physical quality training of future trainers-teachers. International Journal of Advanced Computer Science and Applications, 15(3), 329–339. https://doi.org/10.14569/IJACSA.2024.0150334

Tillem, M., & Gün, A. (2025). The impact of color blindness on player engagement and emotional experiences: A multimodal study in a game-based environment. Multimodal Technologies and Interaction, 9(6), Article 62. https://doi.org/10.3390/mti9060062

Tiwary, G., Chauhan, S., & Goyal, K. K. (2024). Multimodal attention CNN for human emotion recognition. Lecture Notes in Networks and Systems, 918, 161–177. https://doi.org/10.1007/978-981-97-0641-9_11

Udahemuka, G., Djouani, K., & Kurien, A. M. (2024). Multimodal emotion recognition using visual, vocal and physiological signals: A review. Applied Sciences, 14(17), 8071. https://doi.org/10.3390/app14178071

Vec, V., Tomažič, S., Kos, A., & Umek, A. (2024). Trends in real-time artificial intelligence methods in sports: A systematic review. Journal of Big Data, 11, 148. https://doi.org/10.1186/s40537-024-01026-0

Wang, Y., Wang, H., & Lu, X. (2025). Research on the impact of explosive martial arts training on emotion regulation and attention based on questionnaire data. Scientific Reports, 15(1), Article 37957. https://doi.org/10.1038/s41598-025-21790-0

Wang, Y., Zhang, L., & Ding, G. (2025). A survey on expressive action quality assessment driven by aesthetics and biomechanics. Proceedings of SPIE – The International Society for Optical Engineering, 13709, Article 1370907. https://doi.org/10.1117/12.3073579

Wu, W. (2024). Multimodal emotion detection of tennis players based on deep reinforcement learning. International Journal of Biometrics, 16(5), 497–513. https://doi.org/10.1504/IJBM.2024.140775

Xiaoyan, L., Nazarudin, M. N., & Mansor, A. Z. (2025). The long-term impact of running applications on physical fitness among vocational college students: A longitudinal analysis. Journal of Physical Education and Sport, 25(5), 1060–1070. https://doi.org/10.7752/jpes.2025.05117

Xu, J., Guo, X., Zhang, Z., Liu, H., & Lee, C. (2025). Triboelectric mat multimodal sensing system (TMMSS) enhanced by infrared image perception for sleep and emotion-relevant activity monitoring. Advanced Science, 12(6), Article 2407888. https://doi.org/10.1002/advs.202407888

Yang, F., & Wang, Z. (2025). An intelligent taekwondo coaching system based on augmented reality technology with real-time feedback mechanisms. Scientific Reports, 15(1), Article 40832. https://doi.org/10.1038/s41598-025-24608-1

Yang, Y., Caldwell, K., Schlagal, R. C., DeCelle, S., Simmons, S., Yang, L., Abduljawad, S., Ayar, J., Singh, J. R., Sheeler, R., Flynn, P. R., Knapp, K. A., & Härtl, R. (2025). "No pain, more gain": Subjective responses to a T'ai Chi and Qigong program for back pain. Global Advances in Integrative Medicine and Health, 14. https://doi.org/10.1177/27536130251384679

Younis, E. M. G., Mohsen, S., Houssein, E. H., & Ibrahim, O. A. S. (2024). Machine learning for human emotion recognition: A comprehensive review. Neural Computing and Applications, 36, 8901–8947. https://doi.org/10.1007/s00521-024-09426-2

Zhang, H., Li, X., Sun, Y., Fu, X., Qiu, C., & Carroll, J. M. (2024). VRMN-bD: A multi-modal natural behavior dataset of immersive human fear responses in VR stand-up interactive games. Proceedings – 2024 IEEE Conference on Virtual Reality and 3D User Interfaces (VR 2024), 320–330. https://doi.org/10.1109/VR58804.2024.00054

Zhang, L., & Tian, H. (2024). Construction of a risk assessment and prediction model for athlete doping use based on bioinformatics. MCB Molecular and Cellular Biomechanics, 21(4), Article 446. https://doi.org/10.62617/mcb446

Zhao, B., Yang, Z., & Wu, Y. (2025). SFNet: Fusion of spatial and frequency domain features for improved posture classification in yoga poses. 2025 IEEE 8th International Conference on Mechatronics and Computer Technology Engineering (MCTE 2025), 145–149. https://doi.org/10.1109/MCTE67374.2025.11213654