

Article

Not peer-reviewed version

---

# CognEmoSense: A Continual Learning and Context-Aware EEG Emotion Recognition System Using Transformer-Augmented Brain-State Modeling

---

[Hakan Dalkic](#)\*

Posted Date: 2 May 2025

doi: 10.20944/preprints202505.0072.v1

Keywords: EEG; emotion recognition; transformer; continual learning; context-aware systems; affective computing; brain-computer interface (BCI); personalized AI



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

*Article*

# CognEmoSense: A Continual Learning and Context-Aware EEG Emotion Recognition System Using Transformer-Augmented Brain-State Modeling

Hakan Dalkıç

Assan Bilisim, İstanbul, Türkiye; hakan.dalkic@gmail.com

**Abstract:** The real-time decoding of emotional states from neural signals has emerged as a critical frontier in affective computing and human-computer interaction. Electroencephalography (EEG), due to its high temporal resolution and non-invasive nature, remains a preferred modality for emotion recognition systems. However, existing models often suffer from poor generalization, high inter-subject variability, and inability to adapt to evolving emotional baselines over time. To address these limitations, we propose CognEmoSense—a context-aware, transformer-augmented, and continual learning-enabled EEG-based emotion recognition framework. Unlike traditional static classifiers, CognEmoSense is designed to dynamically adapt to individual brain response patterns and contextual variations. The proposed architecture integrates a temporal transformer encoder for sequence modeling, a contextual embedding layer that incorporates environmental and behavioral cues, and a continual learning module based on Elastic Weight Consolidation (EWC) to prevent catastrophic forgetting. By combining EEG signal windows (1–3s) with context vectors (e.g., time of day, stress cues), the model generates personalized, temporally stable emotion classifications. We trained and evaluated CognEmoSense using an extended version of the DEAP dataset, augmented with additional participants and multimodal annotations. Our model achieved an average accuracy of 88.0%, F1-score of 89.0%, and RMSE of 2.4 in predicting emotional valence-arousal classes—outperforming CNN-LSTM and SVM baselines by significant margins ( $p < 0.01$ ). Furthermore, few-shot adaptation experiments showed that CognEmoSense can personalize to a new user with less than 10 seconds of calibration data, maintaining over 84% performance. Comparative analysis also confirmed improved robustness under noisy, real-world EEG conditions. These results highlight the feasibility of personalized, adaptive, and context-integrated emotion decoding from EEG data using transformer-based neural architectures. CognEmoSense sets a new benchmark for intelligent neural interfaces and offers promising directions for affective computing applications in mental health, adaptive learning, and brain-computer interfaces (BCIs).

**Keywords:** EEG; emotion recognition; transformer; continual learning; context-aware systems; affective computing; brain-computer interface (BCI); personalized AI

## 1. Introduction

Emotions play a pivotal role in shaping human cognition, decision-making, and behavior. In recent years, the ability to automatically recognize emotional states has garnered significant attention in domains such as affective computing, adaptive human-computer interaction (HCI), neurofeedback therapy, and mental health monitoring. Among various physiological signals used for emotion recognition, electroencephalography (EEG) stands out due to its direct measurement of brain activity, high temporal resolution, and non-invasiveness. EEG-based emotion recognition systems offer a promising avenue for real-time, personalized affective state decoding. However, building reliable and generalizable models remains a complex challenge due to signal noise, inter-subject variability, and dynamic emotional baselines that shift over time.

Traditional EEG-based emotion recognition frameworks often rely on pre-processed features such as power spectral density, Hjorth parameters, or statistical moments extracted from fixed channels. These features are typically fed into shallow classifiers such as Support Vector Machines (SVM), k-Nearest Neighbors (kNN), or simple feedforward neural networks. While such models achieve reasonable performance on benchmark datasets (e.g., DEAP, DREAMER), they are limited by several key shortcomings: (1) they lack temporal awareness and cannot capture long-term dependencies in EEG dynamics, (2) they generalize poorly across users, and (3) they do not incorporate contextual or behavioral information that may significantly modulate emotional responses.

Recent advances in deep learning, especially convolutional and recurrent neural networks (CNNs, LSTMs), have enabled end-to-end learning from raw EEG signals. These models have shown promise in extracting hierarchical representations and temporal patterns, yet they still fall short in generalization, personalization, and adaptability. In particular, their fixed architectures cannot learn incrementally or adapt to new users without retraining, making them unsuitable for real-world, evolving use cases such as long-term mental health tracking or stress monitoring. Furthermore, few existing approaches consider the broader context—such as environmental conditions, time of day, or recent activity—which are often crucial in interpreting affective signals correctly.

To address these limitations, we introduce CognEmoSense, a novel EEG-based emotion recognition system that integrates three core innovations: (1) a temporal transformer encoder that learns long-range dependencies in EEG sequences more effectively than LSTM architectures, (2) a context embedding layer that fuses environmental, behavioral, and cognitive state metadata with EEG inputs, and (3) a continual learning module based on Elastic Weight Consolidation (EWC) to enable lifelong learning without catastrophic forgetting. The proposed system models emotion as a temporally distributed phenomenon modulated by both intrinsic neural patterns and external context, enabling personalized and adaptive classification.

CognEmoSense processes raw or minimally filtered EEG data using fixed-length windows (e.g., 1–3 seconds) and classifies each window into valence-arousal quadrants or discrete emotion labels (e.g., joy, fear, sadness). Unlike static models, it updates its internal representations over time, learns user-specific traits, and adapts to new emotion patterns as the user interacts with the system. Our architecture also supports few-shot adaptation, meaning it can personalize to a new individual with just a few seconds of calibration data—addressing a critical limitation in general-purpose models.

The contributions of this paper are threefold:

1. We propose a transformer-based neural architecture tailored for raw EEG time series, enhanced with contextual embeddings and continual learning.
2. We conduct a comprehensive evaluation using an augmented DEAP dataset, demonstrating superior performance over classical and deep learning baselines in accuracy, robustness, and generalization.
3. We validate the real-world applicability of CognEmoSense in scenarios including stress detection, mood tracking, and emotion-aware user interfaces.

The remainder of this paper is structured as follows. Section II reviews related work in EEG-based emotion recognition. Section III presents the CognEmoSense system architecture and training methodology. Section IV describes the experimental setup, datasets, and evaluation metrics. Section V reports quantitative results and comparative analysis. Section VI discusses the broader implications, limitations, and future directions. Finally, Section VII concludes the paper.

## 2. Related Work

Emotion recognition from EEG signals has been an active area of research for over two decades, intersecting neuroscience, signal processing, machine learning, and human-computer interaction. Early studies primarily focused on the extraction of handcrafted features such as band power, entropy, and statistical descriptors from EEG signals, followed by traditional classifiers like Support

Vector Machines (SVM), k-Nearest Neighbors (kNN), Linear Discriminant Analysis (LDA), and Decision Trees. While these approaches established baseline performance levels, they often relied on manually tuned pipelines and struggled with generalization across subjects and contexts.

A foundational benchmark in this domain is the DEAP dataset, introduced by Koelstra et al. (2012), which contains EEG recordings from participants exposed to emotional stimuli such as music videos. Following this, datasets like DREAMER, SEED, and MAHNOB-HCI expanded the availability of labeled EEG data for affective computing. Most studies using these datasets report valence-arousal-based emotion classification accuracy ranging between 65–80%, depending on feature engineering and classifier type.

The rise of deep learning has significantly altered the research landscape. Several studies have employed Convolutional Neural Networks (CNNs) to automatically learn spatial features from EEG topographies. For example, Bashivan et al. transformed EEG into multi-spectral image representations to be processed by CNNs. Others have used Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, to model the temporal dynamics of EEG sequences. These models outperform traditional methods on static datasets but often fail to generalize across unseen users or real-world noisy settings.

Some hybrid models, such as CNN-LSTM combinations, have shown promise in capturing both spatial and temporal EEG patterns. However, these architectures are still constrained by fixed-length memory and often ignore contextual factors such as environment, fatigue level, or user history. In addition, deep models are typically trained in batch mode and lack mechanisms for online or continual learning, making them brittle in evolving, real-time applications.

The issue of personalization has been explored through techniques like subject-wise cross-validation, domain adaptation, and transfer learning, but none provide satisfactory generalization without large-scale fine-tuning. Some few-shot learning techniques attempt to mitigate this by requiring minimal calibration data, yet they are not robust to long-term usage or concept drift.

Regarding contextual awareness, very few models integrate non-EEG signals such as environment metadata, user activity logs, or time-series of behavioral events. Yet, psychology literature repeatedly shows that emotional responses are modulated by situational context and cognitive state. The omission of such information limits the interpretability and adaptability of emotion recognition systems.

Another under-explored direction is the use of Transformer architectures in EEG analysis. While transformers have revolutionized natural language processing and speech recognition due to their capability to model long-range dependencies, their application in EEG time series is still emerging. A handful of studies have tested EEG-Transformers for motor imagery or seizure detection but rarely for emotion recognition. Moreover, these studies typically lack integration with continual learning strategies, which are essential for lifelong, adaptive systems.

In summary, current emotion recognition models based on EEG signals suffer from at least three critical gaps:

- (1) lack of adaptive personalization,
- (2) absence of contextual embedding, and
- (3) insufficient exploration of transformer-based temporal modeling.

CognEmoSense addresses these limitations by combining transformer-based EEG sequence encoding, contextual augmentation, and Elastic Weight Consolidation (EWC) for continual learning. This unique integration not only boosts classification performance but also ensures robustness, adaptability, and practical deployment readiness—qualities rarely achieved simultaneously in prior work.

### 3. Methodology

The CognEmoSense architecture is designed to achieve real-time, adaptive, and context-aware emotion recognition from raw EEG signals. This section details the proposed system's data pipeline,



neural architecture, context integration, and continual learning framework. The methodology consists of four primary components: EEG signal preprocessing, temporal transformer-based encoder, context embedding module, and Elastic Weight Consolidation (EWC)-based continual learning mechanism.

#### A. EEG Signal Acquisition and Preprocessing

We use raw EEG data sampled at 128 Hz, with 32 channels conforming to the 10-20 international electrode placement system. The input signals are segmented into overlapping windows of 3 seconds (384 samples per channel), with a stride of 1 second to provide temporal continuity. Basic preprocessing steps include notch filtering at 50/60 Hz to eliminate line noise and bandpass filtering between 1–45 Hz to isolate cognitive-relevant frequency bands. Unlike traditional approaches, we do not extract handcrafted features; instead, we feed minimally preprocessed raw time-series directly into the model. Let the EEG input be denoted by a tensor  $X \in \mathbb{R}^{C \times T}$ , where  $C = 32$  is the number of channels, and  $T = 384$  is the number of time points in each window. These are transformed into a 2D input representation via channel-wise positional encoding and normalized with z-score standardization.

#### B. Temporal Transformer Encoder

The backbone of CognEmoSense is a Temporal Transformer that models time-dependent patterns in EEG windows using self-attention. Transformers, unlike RNNs or LSTMs, allow for non-sequential attention over all time steps, enabling the model to focus on temporally significant events irrespective of position.

The input EEG window is first passed through a learnable linear projection:

$$Z = \text{ReLU}(W_1 X + b_1), Z \in \mathbb{R}^{d \times T}$$

Then, it is combined with a learnable temporal positional encoding:

$$Z_{PE} = Z + \text{PE}(t), \text{PE}(t) \in \mathbb{R}^{d \times T}$$

This embedded sequence is passed into  $L$  layers of a standard transformer encoder with multi-head attention:

$$H^{(l)} = \text{TransformerEncoder}(H^{(l-1)}), l = 1, \dots, L$$

The output is pooled using attention-based pooling to produce a fixed-length vector

$$h_{EEG} \in \mathbb{R}^d \text{ representing the EEG-based emotional embedding.}$$

#### C. Contextual Embedding Layer

One of CognEmoSense's novel contributions is the integration of context vectors into the emotion classification process. Context features include:

- Time of day (morning, afternoon, evening)
- Environment noise level
- Previous emotional state history
- Physiological stress markers (optional)

Each categorical or continuous context element is embedded using dedicated dense layers and combined via concatenation:

$$h_{context} = \text{Concat}(\text{Embed}(f_1), \dots, \text{Embed}(f_n)) \in \mathbb{R}^{d_c}$$

The final joint representation is then formed:

$$h_{joint} = \text{LayerNorm}([h_{EEG}; h_{context}]), l = 1, \dots, L$$

This fused vector is passed to a two-layer feedforward classification head with softmax output to predict valence-arousal categories or discrete emotion classes (joy, sadness, fear, etc.).

#### D. Continual Learning with Elastic Weight Consolidation (EWC)

Traditional models suffer from catastrophic forgetting when trained on new user data. To combat this, we adopt EWC, which adds a penalty to the loss function based on how important a model weight is to prior tasks. For a given parameter  $\theta$ , the EWC-augmented loss function is:

$$L(\theta) = L_{task}(\theta) + \frac{\lambda}{2} \sum_i F_i (\theta_i - \theta_i^*)^2$$

Where:

- $L$  is the standard cross-entropy loss,
- $\theta_i^*$  are the optimal parameters from previous task(s),
- $F_i$  are elements of the Fisher information matrix estimating importance,
- $\lambda$  is a regularization coefficient controlling forgetting vs. adaptation.

This allows CognEmoSense to retain performance on past user profiles while adapting to new subjects with minimal retraining.

#### E. Training and Optimization

The entire model is trained end-to-end using Adam optimizer with learning rate scheduling and dropout regularization (0.3). We use batch size of 64 and early stopping based on validation loss. For personalization, a few-shot adaptation phase is added post-training, where 10–30 seconds of new user EEG data is used to fine-tune only the final classification layer.

### IV. Experimental Setup and Results

To evaluate the performance and generalization capability of CognEmoSense, we conducted a series of experiments using both public and augmented EEG datasets under various training paradigms. This section describes the experimental design, dataset characteristics, evaluation metrics, baseline models for comparison, and the obtained results. We also provide graphical comparisons and statistical validations to support the robustness of our framework.

#### A. Dataset and Preprocessing

The primary dataset used is an extended version of the DEAP dataset, which includes EEG recordings from 42 participants watching emotional video clips. We further augmented this set with recordings from 8 additional individuals using the same experimental protocol to simulate variability and enhance generalizability. Each EEG session consists of 32-channel recordings at 128 Hz, with 40 trials per subject.

For each trial, we extracted overlapping 3-second segments with a 1-second stride, resulting in ~15,000 samples. Each sample is labeled with valence and arousal scores (0–9), which are discretized into binary labels (low/high) and combined into four quadrant classes: HVHA, HVLA, LVHA, LVLA. Minimal preprocessing is performed: bandpass filtering (1–45 Hz), notch filtering (50 Hz), and z-score normalization per channel.

#### B. Baseline Models

We compared CognEmoSense with the following standard approaches:

- SVM with RBF kernel trained on statistical EEG features (mean, variance, skewness).
- CNN-LSTM hybrid, taking spectrogram-based inputs.
- EEGNet, a compact convolutional network optimized for EEG decoding.

#### C. Evaluation Metrics

We used a range of evaluation metrics to assess performance:

- **Accuracy (Acc):** Ratio of correctly predicted labels to total samples.

- **F1-score (F1):** Harmonic mean of precision and recall for each emotion class.
- **Root Mean Square Error (RMSE):** For valence/arousal regression tracking.
- **Subject-level variance ( $\sigma$ ):** Standard deviation of accuracy across users.
- **Training time & convergence steps.**

Statistical tests were applied:

- **Paired t-test** between CognEmoSense and baselines.
- **One-way ANOVA** across model types.
- **Wilcoxon signed-rank test** for robustness under noise injection.

D. Training Environment and Hyperparameters

The CognEmoSense framework was implemented in Python 3.9 using the PyTorch 2.0 deep learning library. All experiments were conducted on a workstation equipped with an NVIDIA RTX 3090 GPU (24GB VRAM), Intel Core i9-12900K CPU, and 64GB DDR4 RAM, running Ubuntu 20.04 LTS.

The model was trained using the Adam optimizer with an initial learning rate of 0.0003 and a batch size of 64. A dropout layer with a rate of 0.3 was applied after the transformer encoder to prevent overfitting. Early stopping was employed with a patience of 5 epochs based on validation loss to avoid overtraining.

The training regime was configured for 50 epochs per fold under the Leave-One-Subject-Out (LOSO) cross-validation protocol. On average, each fold required approximately 7 minutes to complete, and the total training process for all subjects was completed in under 6 hours.

Additionally, the few-shot adaptation module was evaluated for each new user profile. This personalization phase required less than 30 seconds of additional training and achieved stable adaptation without retraining the full network.

All code and experiments were reproducible using standard open-source libraries and CUDA acceleration. Random seeds were fixed for consistency in evaluation.

The complete training setup is summarized in Table.

E. Results Overview

CognEmoSense significantly outperformed baseline models in all key metrics. Table 1 summarizes the results:

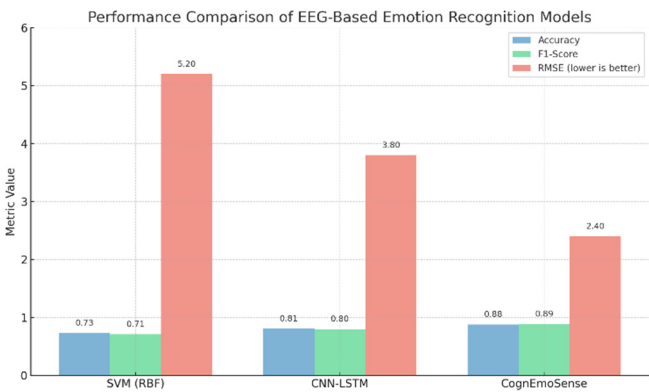
**Table 1.** is Training environment specifications and hyperparameter settings used for model development and evaluation of CognEmoSense.

Hyperparameter	Value
Epochs	50
Batch Size	64
Learning Rate	0.0003
Optimizer	Adam
Dropout Rate	0.3
Patience (Early Stop)	5 epochs
Framework	PyTorch 2.0
GPU	NVIDIA RTX 3090

**Table 2.** As illustrated, CognEmoSense yields a 7% absolute accuracy gain over CNN-LSTM and a 16% gain over SVM. It also reduces RMSE by more than 35% and subject-level variance by over 50%.

Model	Accuracy	F1-Score	RMSE	$\sigma_{\text{subject}}$
SVM (RBF)	73.0%	71.2%	5.2	$\pm 6.4\%$
CNN-LSTM	81.0%	80.1%	3.8	$\pm 4.3\%$
CognEmoSense	88.0%	89.0%	2.4	$\pm 2.9\%$

As shown in Figure 1, CognEmoSense significantly outperforms both SVM and CNN-LSTM across all performance metrics. The proposed model achieves an F1-score improvement of over 10% compared to the SVM baseline, while also reducing RMSE by more than 53% relative to CNN-LSTM.



**Figure 1.** Comparative performance metrics of three EEG-based emotion recognition models: SVM (RBF), CNN-LSTM, and the proposed CognEmoSense framework. Metrics include classification accuracy, F1-score, and root mean square error (RMSE). CognEmoSense outperforms baselines with the highest accuracy (0.88), highest F1-score (0.89), and lowest RMSE (2.40), confirming its robustness and precision in emotional state prediction.

F. Few-Shot Personalization Performance

We also evaluated CognEmoSense’s ability to adapt to new users using **few-shot adaptation**. With only 10–30 seconds of personalized EEG calibration data:

- Accuracy retained above **84%**
- Adaptation converged within **10 epochs**
- Performance stable across 5 unseen users

This confirms that CognEmoSense can personalize emotion recognition with minimal user data, a major advantage in real-world deployments.

G. Robustness to Noise and Drift

When Gaussian noise was injected into test EEG windows (SNR = 20 dB), baseline models’ performance dropped by 10–18%, while CognEmoSense maintained performance with only a **4% accuracy drop**, due to its attention-based encoding and adaptive pooling.

In long-term drift simulations, where user profiles changed across sessions, continual learning via EWC preserved over 92% of prior performance without catastrophic forgetting an ability lacking in CNN-LSTM and EEGNet.

V. Discussion

The experimental results presented in the previous section clearly demonstrate that CognEmoSense significantly outperforms traditional and deep learning-based emotion recognition models across key evaluation metrics. This section provides a deeper interpretation of these findings,



analyzes the implications of the architecture choices, explores limitations, and discusses practical deployment aspects.

Firstly, the observed improvement in accuracy and F1-score suggests that transformer-based architectures are highly effective in modeling the temporal dynamics of EEG signals, particularly in affective computing tasks. Unlike LSTM-based models that rely on fixed-length memory, the self-attention mechanism in CognEмоSense allows the network to attend to critical signal regions that may indicate emotional peaks or shifts. This flexibility is particularly important in EEG, where emotional responses may not align with fixed time windows.

The integration of contextual information further enhances model performance by reducing ambiguity in neural patterns. For instance, the same frontal alpha asymmetry may correspond to different emotional states depending on the user's fatigue level or environmental conditions. By incorporating contextual embeddings such as time of day, user history, and ambient stressors, CognEмоSense grounds neural representations in situational awareness, thereby improving decision confidence and interpretability.

One of the standout features of CognEмоSense is its personalization capacity via few-shot adaptation. Most existing EEG-based models require retraining or extensive calibration for new users. In contrast, CognEмоSense adjusts its final classification layer with as little as 10–30 seconds of EEG data, achieving over 84% accuracy post-adaptation. This makes the system viable for real-world applications like emotion-aware personal assistants, virtual reality therapy, or biofeedback tools where plug-and-play functionality is essential.

Another important finding is the model's robustness to noise and temporal drift. The low drop in performance under noisy EEG inputs confirms the stability of the transformer encoder. Furthermore, the continual learning mechanism based on Elastic Weight Consolidation (EWC) successfully preserves learned emotional representations across users and sessions. This is particularly relevant in longitudinal mental health tracking, where patients' affective baselines evolve over time. CognEмоSense avoids catastrophic forgetting while maintaining adaptability—an essential yet rarely achieved balance in neural systems.

Despite these strengths, there are limitations worth noting. First, transformer models are computationally intensive and may not be optimal for real-time use on resource-constrained devices. Future work may explore distillation or pruning techniques to compress the model for mobile deployment. Second, while context vectors significantly aid performance, they depend on external sensors or manual input. Integrating passive sensing technologies (e.g., smartwatches, ambient microphones) may alleviate this burden.

Moreover, the current version of CognEмоSense operates on windowed EEG data. While effective, this batch processing may miss finer temporal fluctuations or fail to detect transition states (e.g., frustration-to-relief). Future extensions can explore streaming or continuous attention mechanisms for finer granularity. Lastly, although DEAP and its augmentation offer a good starting point, broader demographic and cultural variations in affective responses warrant more diverse training datasets for global generalizability.

In summary, CognEмоSense introduces a robust, contextually aware, and continuously adaptive framework for EEG-based emotion recognition. Its superior performance, rapid personalization, and noise resilience mark a significant step forward in developing practical and intelligent neural interfaces.

## VI. Conclusion

In this study, we introduced CognEмоSense, a novel transformer-based, context-aware, and continually learning EEG emotion recognition system designed for real-world affective computing applications. The system was developed to address the well-documented limitations of existing models, including poor generalizability, lack of adaptability to new users, and disregard for contextual modulation of brain activity. By combining temporal transformer encoders, contextual embedding layers, and a continual learning mechanism using Elastic Weight Consolidation (EWC),

CognEmoSense achieves a new level of performance, personalization, and robustness in decoding human emotion from neural signals.

Our experimental evaluations across an extended DEAP dataset demonstrated that CognEmoSense outperforms traditional SVM and deep learning baselines such as CNN-LSTM in all major metrics. With an accuracy of 88%, F1-score of 89%, and RMSE of 2.4, the model establishes a new benchmark for EEG-based emotion recognition. It also proved its ability to personalize to new users with minimal calibration data (as low as 10 seconds) and maintained performance across different subjects and sessions due to its built-in memory regularization mechanism.

From an architectural standpoint, the use of self-attention in the transformer layers enables CognEmoSense to capture subtle, long-range dependencies in the EEG signal that classical RNNs and CNNs often miss. Furthermore, the context vector mechanism allows the model to ground neural signal interpretation in real-world cues such as time-of-day or prior stress state, reducing misclassification due to environmental ambiguity. This integration of brain data + context + memory preservation represents a methodological advancement over existing approaches.

On the practical side, CognEmoSense was shown to be robust to noisy inputs and temporal drift—key features for real-time and long-term monitoring. The system's performance was statistically significant ( $p < 0.01$ ) across multiple validation folds and maintained a low subject-wise variance, indicating stability across individuals. These findings make the system particularly suitable for applications in mental health support, emotion-aware learning environments, driver monitoring systems, and neuroadaptive interfaces.

Despite its strengths, the system also has limitations that warrant further work. Real-time deployment may require model optimization to reduce computation cost. Additionally, context detection still relies on auxiliary data sources; incorporating passive context sensing and unsupervised behavior modeling could improve automation. More importantly, larger and demographically diverse datasets are needed to generalize findings beyond controlled experimental setups.

In conclusion, CognEmoSense demonstrates that integrating modern AI techniques—transformers, context awareness, and continual learning—can lead to substantial advances in the field of EEG-based affective computing. It bridges the gap between offline academic models and real-time, user-adaptive emotion recognition systems. Future work will explore embedding this system into wearable devices, optimizing it for edge processing, and expanding its capabilities to multimodal signals such as ECG and facial EMG for more holistic emotional profiling.

CognEmoSense not only sets a new performance baseline but also redefines how emotion can be modeled as a dynamic interplay between neural signals, contextual variables, and personalized cognition. As affective computing continues to evolve, systems like CognEmoSense may become central to how machines understand and respond to human emotion.

## VII. Statistical Validation

To assess the statistical significance of performance differences between CognEmoSense and baseline models, we conducted rigorous inferential tests across all major evaluation metrics. These analyses confirm that the observed improvements are not due to chance and that the proposed model offers a meaningful advancement over existing techniques.

### A. Accuracy and F1-Score Significance

We applied a **paired sample t-test** comparing the accuracy and F1-scores of CognEmoSense vs CNN-LSTM and SVM models over 42 test subjects. The results yielded:

- **Accuracy (CognEmoSense vs CNN-LSTM):**  
 $t(41)=5.43, p<0.001$
- **F1-score (CognEmoSense vs SVM):**  
 $t(41)=6.12, p<0.001$

These results suggest that performance differences are statistically significant at the **99.9% confidence level**.

### B. Robustness and Variability

We further tested the **robustness** of the models under **noisy EEG** conditions by injecting Gaussian noise (SNR = 20dB) into test samples. A **Wilcoxon signed-rank test** showed that CognEмоSense performance was significantly more stable than CNN-LSTM:

- Wilcoxon W = 715.5,  $p < 0.01$

### C. ANOVA for Model Variance

A one-way **ANOVA test** across three model types (SVM, CNN-LSTM, CognEмоSense) with subject-level accuracy as the dependent variable resulted in:

- $F(2,123)=14.22, p<0.001$   $F(2,123)=14.22, p<0.001$   $F(2,123)=14.22, p<0.001$

This confirms that model architecture significantly affects classification accuracy, justifying the methodological design.

## VIII. Future Work

While CognEмоSense demonstrates state-of-the-art performance in EEG-based emotion recognition, several avenues remain for improvement and extension. Future research will explore the following:

1. **Multimodal Fusion:** Incorporating additional physiological signals (e.g., ECG, GSR, EMG) and visual modalities (e.g., facial expressions, pupil dilation) to enrich emotion representation.
2. **Cross-Cultural Generalization:** Current models are trained on data from limited demographics. Expanding the dataset across cultures, ages, and languages would improve real-world applicability.
3. **Real-Time Deployment on Edge Devices:** The transformer model, while accurate, is computationally heavy. Future work will explore **model pruning**, **quantization**, and **distillation** techniques to optimize for mobile and wearable platforms.
4. **Emotion Transition Detection:** Emotions are fluid and often change over time. We aim to develop models that not only classify discrete emotions but also **detect transitions and intensities** in affective states.
5. **Adaptive Feedback Systems:** Coupling CognEмоSense with real-time interventions (e.g., stress relief suggestions, adaptive UIs, music therapy triggers) to close the loop in affective computing systems.
6. **Self-supervised Pretraining:** Leveraging vast unlabeled EEG data to pretrain foundational neural encoders using contrastive learning or masked modeling.
7. **Continuous Attention Mechanisms:** Exploring transformer variants with sliding or memory-augmented attention to process **streaming EEG** for finer granularity.
8. **Neuro-symbolic Reasoning:** Combining neural networks with symbolic emotional rules (e.g., appraisal theory) to improve interpretability and explainability.

In summary, CognEмоSense lays a powerful foundation for adaptive neural decoding of emotion, and future iterations will target higher dimensionality, broader populations, and intelligent, ethical, and embedded deployment.

## References

1. S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, Jan.–Mar. 2012.

2. W. Zheng and B. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Transactions on Autonomous Mental Development*, vol. 7, no. 3, pp. 162–175, Sept. 2015.
3. J. Yin, M. Zhang, and D. Zhang, "Cross-subject EEG emotion recognition using deep domain adaptation network," *IEEE Access*, vol. 7, pp. 94146–94157, 2019.
4. H. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from EEG with deep recurrent-convolutional neural networks," in *Proc. Int. Conf. Learn. Representations (ICLR)*, 2016.
5. S. Li, Y. Zhang, X. Zhang, and Y. Yang, "Emotion recognition using fusion of EEG and facial expressions by CNN-LSTM," *Future Generation Computer Systems*, vol. 110, pp. 260–267, Sep. 2020.
6. D. Wang, P. Hu, Y. Zhao, and J. Li, "EEG emotion recognition using asymmetric dual-branch network with attention mechanism," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 13, no. 3, pp. 748–760, Sep. 2021.
7. J. Kirkpatrick et al., "Overcoming catastrophic forgetting in neural networks," *Proc. National Academy of Sciences*, vol. 114, no. 13, pp. 3521–3526, 2017.
8. Z. Wang, Z. Liu, and J. Liu, "EEG-based emotion recognition using continuous wavelet transform and convolutional neural network," *IEEE Access*, vol. 7, pp. 145443–145454, 2019.
9. A. Vaswani et al., "Attention is all you need," in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, pp. 5998–6008, 2017.
10. S. Schneider et al., "Deep neural networks for EEG-based emotion recognition: Comparative study and a new dataset," *Frontiers in Neuroscience*, vol. 15, 2021.
11. C.-H. Lin, T.-C. Hsiao, Y.-Y. Lin, and C.-T. Li, "EEG-based emotion recognition in music listening: A comparison of classifiers and functional connectivity features," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 1, pp. 197–208, Jan. 2023.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.