Generalizable Psychiatric Screening from Wearable: Causal Disentanglement of Multi-channel Biosignals with Interpretability Analysis

Wenting Kuang*, Yufei Zhang*, Yuqi Cheng*, Kaishun Wu†, Changhe Fan‡, Yongpan Zou*,
*College of Computer Science and Software Engineering, Shenzhen University
†Information Hub, Hong Kong University of Science and Technology (Guangzhou)
‡Department of Psychiatry, Affiliated Guangdong Second Provincial General Hospital of Jinan University
kuangwenting2023@email.szu.edu.cn, yfyaya@foxmail.com, cyuqi2090@gmail.com,
wuks@hkust-gz.edu.cn, changhefan@yahoo.com, yongpan@szu.edu.cn

2025 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 979-8-3315-1557-7/25/$31.00 ©2025 IEEE | DOI: 10.1109/BIBM66473.2025.11356596

***Abstract*—Psychiatric disorders represent a growing global public health concern. Current screening methods primarily rely on subjective self-reports, which prone to bias and often impede timely intervention. In this paper, we propose G-PsychSW (Generalizable Psychiatric Screening from Wearable), a continuous and objective mental health monitoring based on multi-channel biosignals, including photoplethysmography (PPG), galvanic skin response (GSR), skin temperature (SKT), and blood oxygen (OXY). To address challenges such as individual variability and limited interpretability, we introduce causal disentanglement to learn subject-invariant representations to enable domain generalization (DG). We further incorporate multi-scale attention convolution (MAConv) and functional graph construction (FGC) to model temporal dynamic and inter-channel interaction, uncovering meaningful interpretability analysis corresponding to psychial biomarkers. Experiments on a customized dataset under leave-one-subject-out cross-validation show that G-PsychSW achieves 80.16% accuracy and 86.52% F1-score across four-class psychiatric screening, including healthy, depression, anxiety and bipolar disorders, outperforming state-of-the-art DG methods and validating its effectiveness.**

***Index Terms*—psychiatric screening, wearable system, biosignals, domain generalization, interpretability analysis.**

I. Introduction

Psychiatric disorders, particularly depression and anxiety, have become a global health concern, with prevalence rising by 25% annually. [1]. These disorders manifest through emotional instability and behavioral maladaptation, burdening both individuals and healthcare systems. Their diagnosis still relies heavily on subjective self-reports, making assessments prone to bias and delayed intervention. This underscores the need for objective, data-driven methods for early and reliable screening.

Biosignals provide promising physiological correlates of mental states. Frontal alpha asymmetry and elevated beta power in Electroencephalography (EEG) are associated with depression, anxiety, and bipolar disorders [2]. Similarly, disrupted skin temperature (SKT) rhythms and increased galvanic skin response (GSR) are associated with depressive and manic states [3], [4]. These biomarkers provide objective insights into the somatic correlates of mental health. The rapid growth of wearable technologies further enables long-term, non-intrusive monitoring of multi-channel biosignals, including SKT, GSR, Oxygen (OXY), Photoplethysmography (PPG). This paves the way for scalable and continuous psychiatric screening, particularly for individuals with limited self-awareness. Despite these advances, two major challenges remain:

(1) Individual variability in physiological responses introduces a domain gap that hinders model generalization. Domain adaptation methods [5], [6] tackle this by aligning feature distributions between the source domain (*i.e.*, known subject) and the target domain (*i.e.*, unknown subject). Such approaches rely on target-domain data, which is often unavailable in real-world screening. In contrast, domain generalization (DG) aims to learn invariant representations without access to target domain, though sometimes at the cost of accuracy. Liu et al. [7] employ regional decomposition with gating networks to extract invariant features, while Jia et al. [8] obtain invariant and subject-specific components through disentanglement. However, the subject-specific encoder scales with the number of subjects, potentially increasing model complexity.

(2) Limited interpretability constrains the clinical applicability of psychiatric screening models. Recent models often act as black boxes with limited transparency. Existing studies mainly focus on channel-wise attribution, such as applying GNNExplainer for EEG relevance analysis in depression detection [9] or using saliency maps for spatial attention visualization in emotion classification [10]. Beyond individual channel analysis, several works investigate inter-channel interactions. Yang et al. [11] use mutual information and adaptive class activation mapping to identify emotion-related EEG channels, while Wang et al. [12] apply dynamic attention to visualize abnormal connectivity. However, these methods infer channel importance from outputs without distinguishing class-discriminative features from domain noise, and overlook temporal dynamics within channels, which are crucial for detecting high-risk periods and enabling timely intervention.

In this paper, we develop a hardware–software system called **G-PsychSW** that functions independently of commercial APIs. The wearable device collects multi-channel biosignals, including PPG, GSR, OXY, and SKT, while the software

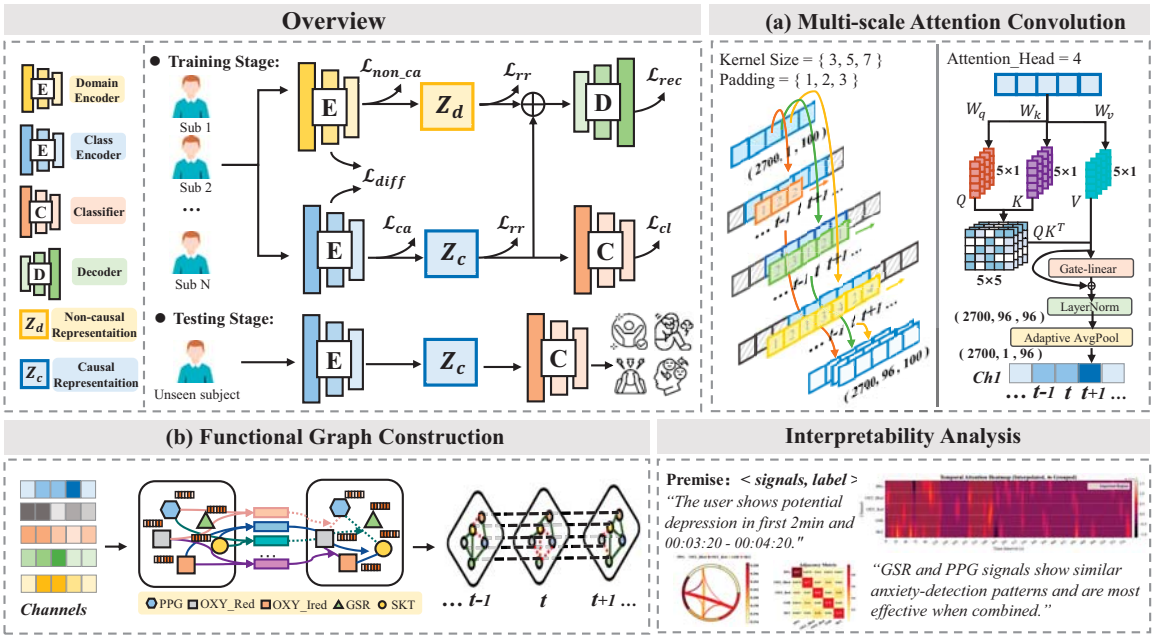


Fig. 1: Overview of G-PsychSW.

terminal classifies users into four-class mental health states such as healthy, depression, anxiety, and bipolar disorder, and provides interpretable visual feedback. To address challenges of individual variability and limited interpretability, we design an end-to-end framework that integrates mobile health sensing, causal representation learning, and interpretable modeling for reliable psychiatric screening in real-world settings. The overall architecture is illustrated in Fig. 1. Our main contributions are summarized as follows:

- We develop a complete hardware–software system designed for objective and ubiquitous psychiatric screening in daily life, driven by biosignal analysis.
- We employ causal disentanglement learning that separates class-related factor and domain-related factor, extracting invariant representation to mitigate individual variability.
- We incorporate multi-scale attention convolution (MAConv) and functional graph construction (FGC) into the encoder, enhancing interpretability by revealing which biosignals are important, when they matter, and how they interact to support model predictions.

II. Methodology

We define the source set $X_s = \left\{\left(X_s^i, Y_s^i, D_s^i\right)\right\}_{i=1}^{n}$, where X_s, Y_s and D_s denote the source data, class label, and domain label for n subjects. The model trained on this set is then applied to predict the class of the target subject X_t. The signals are segmented using a 1-second sliding window with T time steps, filtered to retain key components and suppress noise, and normalized via min–max scaling for intra-channel consistency. Each preprocessed sample is represented as $X = (x_1, x_2, \ldots, x_T) \in \mathbb{R}^{C\times T}$, where C is the number of biochannels, and T is the total number of time steps.

A. Multi-scale Attention-based Convolution

Different signals reflect distinct physiological mechanisms related to mental conditions. Early fusion may obscure their specific patterns and degrade discriminability [13]. To avoid this, we design MAConv in a channel-independent manner, which is formulated as:

$$V_{\text{ATT}} = \text{MHA}\big(\text{Concat}[\text{Conv1D}_k(x)]\big), \quad k \in 3, 5, 7, \quad (1)$$

where Conv1D_k denotes a temporal convolution with kernel size k applied to each channel $x \in \mathbb{R}^{1\times T}$. Smaller kernels capture rapid autonomic changes, while larger ones encode long-term trends. The concatenated multi-scale features are fed into a multi-head attention (MHA) module that adaptively assigns weights across time steps, uncovering temporal and channel-level relevance and improving the interpretability of physiological dynamics in psychiatric screening.

B. Functional Graph Construction

Prior work such as Jia et al. [14] and Yang et al. [11] show that certain biosignals exhibit functional similarities that can be modeled via graph construction. We propose an FGC module that adaptively captures inter-channel interactions in a data-driven manner. Given channel embeddings $v_i \in V$, pairwise similarity is computed via scaled dot-product attention (SDP-ATT):

$$\mathbf{A}_{i,j} = \text{SDP-ATT}(v_i, v_j) = \frac{\exp\left(\frac{(W_q v_i)^\top (W_k v_j)}{\sqrt{d}}\right)}{\sum_{j'=1}^{C} \exp\left(\frac{(W_q v_i)^\top (W_k v_{j'})}{\sqrt{d}}\right)} \quad (2)$$

where $\mathbf{A}_{i,j}$ is the (i, j) -th entry of the adjacency matrix $\mathbf{A} \in \mathbb{R}^{C\times C}$. Based on symmetrically normalized adjacency

matrix $\hat{A}$, we further apply a graph convolution layer (GCN) to construct inter-channel functional connections:

$$Z = \hat{A}VW, \quad \text{where } \hat{A} = D^{-1/2}\mathbf{A}D^{-1/2} \tag{3}$$

where $\hat{A} = D^{-1/2}AD^{-1/2}$ is the symmetrically normalized adjacency matrix, and $W \in \mathbb{R}^{C\times H}$ is a learnable linear projection to H dimensions. The resulting representation $Z \in \mathbb{R}^{C\times H}$ captures inter-channel semantics, alleviating the need for prior assumptions about channel roles and facilitating discovery of informative interactions in screening.

C. Causal Disentanglement For Domain Generalization

To mitigate spurious correlations and domain-specific biases, we adopt causal disentanglement motivated by Jia et al. [15]. The framework comprises a class encoder E_c, a domain encoder E_d, a reconstruction decoder D, and a classifier C. E_c and E_d extract causal representations Z_c and non-causal representations Z_d, both built upon a backbone integrating the **MAConv** and **FGC** modules to capture temporal and channel dependencies. D regularizes the encoders via latent reconstruction, and C predicts domain-invariant labels from the causal representation Z_c through a multi-layer perceptron.

1) representation learning: To obtain discriminative and generalizable representations, we introduce a prototype-based contrastive alignment that jointly exploits class and domain supervision. Learnable prototypes P_c, P_d serve as semantic anchors, attracting representations toward their respective class or domain centers while repelling mismatched ones. The unified alignment loss aggregates both objectives as follows:

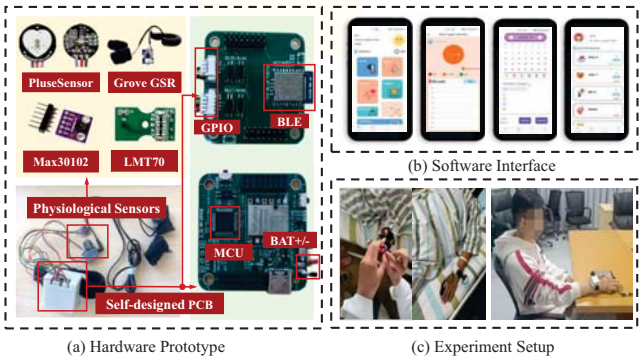
$$\mathcal{L}_{\text{align}} = -\sum_{t\in c,d} \lambda_t \log \frac{\exp(sim(z_i^t, p_{y_i}^t)/\tau)}{\sum_{k\in s_t}\exp(sim(z_i^t, p_k^t)/\tau)} \tag{4}$$

where $sim(\cdot)$ denotes cosine similarity, and τ is a temperature parameter controlling the distribution sharpness. s_c and s_d denote the sets of class and domain prototypes learned from the source domain, with $\lambda_c : \lambda_d = 2 : 1$ balancing their contributions. Causal alignment is prioritized since class factors are more informative for prediction, whereas domain factors serve auxiliary supervision for shift mitigation.

2) confounder mitigation: Although the class and domain encoders are designed to extract distinct subspaces, residual entanglement may remain due to latent confounders such as environmental noise. To promote statistical independence, we apply zero-mean row-wise normalization and enforce an orthogonality constraint on the learned representations. A dynamically weighted difference loss further encourages disentanglement:

$$\mathcal{L}_{\text{diff}} = \frac{\lambda}{H^2}\left\| \frac{\hat{Z}_c^{\top}\hat{Z}_d}{\|\hat{Z}_c\|\|\hat{Z}_d\|}\right\|_F \tag{5}$$

where λ increases from 0 to 1 during training to avoid early over-constraint and $\|\cdot\|_F$ denotes the Frobenius norm. This constraint enforces $\hat{Z}_c^{\top}\hat{Z}_d \approx 0$, encouraging orthogonality between subspaces and suppressing confounding effects.



(a) Hardware Prototype (b) Software Interface (c) Experiment Setup

Fig. 2: The system prototype and experiment setup.

3) collaborative reconstruction: To prevent any single subspace from dominating the model optimization, we introduce a collaborative reconstruction mechanism that jointly decodes causal and non-causal representations via a shared decoder D.

$$\mathcal{L}_{\text{rec}} = \frac{1}{B}\sum_{i=1}^{B}\left\| D([Z_c^{(i)}, Z_d^{(i)}]) - x^{(i)}\right\|_2^2, \tag{6}$$

where B is batch size and $[\cdot,\cdot]$ denotes feature concatenation.

4) redundancy reduction: Furthermore, we introduce a redundancy reduction loss to improve subspace distinctiveness. It enforces decorrelation by minimizing the deviation between the self-correlation matrix and the identity matrix, suppressing redundant dependencies among feature dimensions. Formally,

$$\mathcal{L}_{\text{rr}} = \left\|Z_cZ_c^{\top} - \mathbf{I}\right\|_F + \left\|Z_dZ_d^{\top} - \mathbf{I}\right\|_F \tag{7}$$

where $\mathbf{I}$ is identity matrix, ensuring unit variance and independence across representations.

5) classification and overall objective: The psychiatric state is predicted from the causal representation Z_c, optimized using a standard cross-entropy loss $\mathcal{L}_{\text{cl}}$. The overall objective integrates all loss components to guide the model toward learning disentangled, informative, and generalizable representations:

$$\mathcal{L}_{\text{total}} = \alpha\mathcal{L}_{\text{align}} + \beta\mathcal{L}_{\text{diff}} + \gamma\mathcal{L}_{\text{rec}} + \delta\mathcal{L}_{\text{rr}} + \mathcal{L}_{\text{cl}} \tag{8}$$

where α, β, γ, and δ are hyperparameters controlling the contribution of each term for robust psychiatric screening.

III. Experiment

A. Hardware–software System

The wearable device integrates a self-designed circuit board, microcontroller, Bluetooth module, and finger-mounted biosensors such as PulseSensor, MAX30102, Grove GSR, and LMT70 for PPG, OXY (red and infrared), GSR, and SKT acquisition. Signals are processed by the microcontroller via GPIO interfaces, and transmitted to smartphone through Bluetooth. The device measures 52 mm × 45 mm × 30 mm, weighs 111.8 grams, and supports over 5 h of continuous monitoring, making it suitable for wearable use. The software terminal receives real-time data, performs biosignal analysis and psychiatric screening using our model, and visualizes results

TABLE I: Performance on Subject-Independent Psychiatric Screening: Baselines vs. Ablated Variants

Model	Normal		Depression		Anxiety		Bipolar		Overall	
	mACC(%)	mF1(%)	mACC(%)	mF1(%)	mACC(%)	mF1(%)	mACC(%)	mF1(%)	mACC(%)	mF1(%)
CNN [16]	46.78	60.81	64.05	76.43	69.85	79.82	18.44	26.51	51.78	62.87
MSTGCN [14]	55.08	66.23	58.80	66.68	56.58	64.73	72.66	80.65	60.30	69.05
ManyDG [17]	70.01	81.27	78.97	86.32	80.14	87.06	**80.44**	**88.26**	77.56	85.79
MoGE [7]	64.63	77.22	78.99	87.51	73.86	83.60	59.45	71.77	69.82	80.48
MDNet [8]	52.59	67.49	**79.01**	**87.52**	79.33	87.69	44.96	56.56	65.49	67.49
CDDG [15]	72.40	82.13	74.19	82.56	82.23	88.53	73.47	83.21	76.04	84.40
Ours	**81.04**	**87.58**	75.28	82.60	**84.26**	**89.07**	79.08	86.29	**80.16**	**86.52**
w/o align	83.74	89.84	72.25	81.04	82.48	87.43	73.20	81.33	78.22	85.08
w/o diff	80.56	87.23	76.06	83.25	82.81	88.20	79.23	85.71	79.84	86.21
w/o rec	79.95	87.32	73.37	81.65	83.25	88.45	75.79	83.43	78.41	85.41
w/o rr	79.58	86.18	73.31	81.55	78.88	84.53	79.58	86.03	77.82	84.51

Note: The table reports mean accuracy (mACC) and mean F1-score (mF1) over healthy controls and psychiatric disorder groups. In the baseline comparison, the top results for each metric are highlighted in **bold** and are underlined, respectively.

on the smartphone. Fig. 2 illustrates the hardware–software system and experimental setup.

B. Dateset Detail

The self-collected dataset comprises 74 participants (40 males, 34 females), aged 14–60, all recruited under an approved Institutional Review Board (IRB) protocol with informed consent. The patient group comprises 18 with depression, 22 with anxiety, and 16 with bipolar disorder, diagnosed through standardized scales and psychiatrist evaluations, ensuring the reliability of the labels. 18 healthy controls from a university population are screened using the SDS, SCL-90, and SAS to confirm mental stability. Each participant wears the custom wristband in a resting state while five-channel biosignals are continuously recorded for 45 minutes at varying sampling rates. After preprocessing, 10-minute segments per subject were resampled to 100 Hz for consistency, yielding 44,400 total samples for subsequent physiological analysis.

C. Baseline

We compare our method with a baseline CNN [16] and several state-of-the-art DG models for time-series analysis. The CNN serves as a classical convolutional neural network architecture that extracts features through 1D convolution without domain-level processing. MSTGCN [14] applies adversarial learning to confuse domain discrimination, while ManyDG [17] leverages mutual reconstruction and orthogonal projection to suppress domain-specific features. MoGE [7] introduces graph experts to capture cross-channel heterogeneity. MDNet [8] performs subject-wise disentanglement to extract domain-invariant features, and CDDG [15] separates causal and domain representations via dual encoders.

D. Experiment Setting And Implementation

We evaluate subject-independent performance using leave-one-subject-out cross-validation (LOSO-CV). In each fold, 1 subject is reserved for testing, and the remaining 73 subjects are split into training and validation sets in an 8:2 ratio. Final results are reported as the average accuracy and F1-score across all test folds. Regarding hyperparameters, the temperature parameter τ is set to 0.1. The overall objective comprises multiple terms weighted as follows: $\alpha = 1$, $\beta = 0.001$, $\gamma = 1$, and $\delta = 0.5$, corresponding to the $\mathcal{L}_{\text{align}}$, $\mathcal{L}_{\text{diff}}$, $\mathcal{L}_{\text{rec}}$, and $\mathcal{L}_{\text{rr}}$. The model is optimized with Adam with learning rate $1\text{e}{-3}$ for up to 300 epochs, applying early stopping with a patience of 50 based on validation results. All baselines share the same settings for fair comparison. Models are implemented in PyTorch and trained on an NVIDIA RTX A6000 GPU.

IV. Result

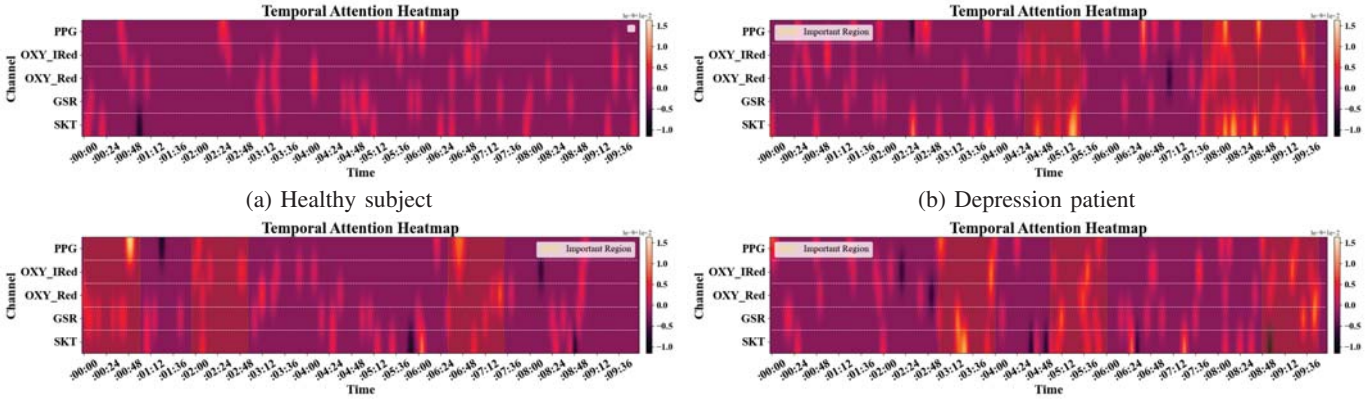
A. Subject-independent Performance

We conduct subject-independent four-class psychiatric screening on our customized dataset, with the results summarized in Tab. I. Our method attains an average accuracy of 80.16% and F1-score of 86.52%, surpassing all baselines. Compared with the best-performing ManyDG, it improves accuracy and F1-score by 2.60% and 0.79% and outperforms CNN by 28.38% in accuracy and 23.65% in F1-score. Normal and Anxiety classes achieve the highest accuracies. We further perform paired t-tests to assess the statistical significance of performance differences, a common practice in time-series forecasting [18]. Statistically significant improvements are observed over CNN ($p < 1e{-}9$), MSTGCN ($p < 1e{-}5$), MDNet ($p < 1e-4$), MoGE ($p < 0.01$), and CDDG ($p < 0.05$) in either or both metrics. Although the gain over ManyDG is not significant ($p = 0.14$), our approach still achieves the best overall performance. These results demonstrate the effectiveness and generalization capability of our approach for psychiatric screening without subject-specific calibration.

B. Ablation Study

To assess the contribution of each loss term, we conduct ablation studies by individually removing $\mathcal{L}_{\text{align}}$, $\mathcal{L}_{\text{diff}}$, $\mathcal{L}_{\text{rec}}$, and $\mathcal{L}_{\text{rr}}$, yielding four ablated variants: **w/o align**, **w/o diff**, **w/o rec**, and **w/o rr**. The results are shown in Tab. I.

The exclusion of any single objective leads to a consistent drop in performance, with overall accuracy reduced by 1.94%, 0.32%, 1.75% and 2.34% for **w/o align**, **w/o diff**, **w/o rec**,



(a) Healthy subject (b) Depression patient

(c) Anxiety patient (d) Bipolar patient

Fig. 3: Temporal Attention Heatmap Across Subjects.

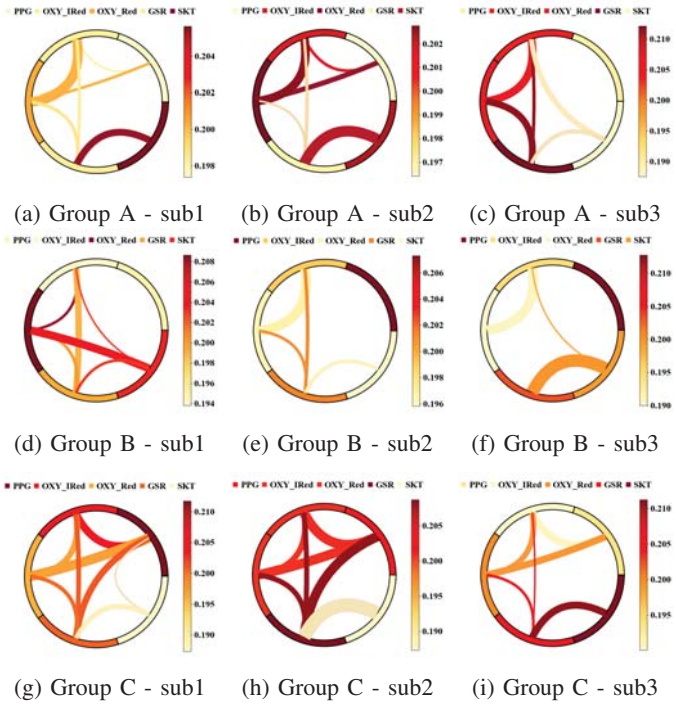
and **w/o rr**, respectively. Among them, removing $\mathcal{L}_{\text{rr}}$ results in the pronounced decline, suggesting that both causal and non-causal representations become more redundant and correlated, likely due to multi-channel biosignal co-variation, environmental noise, and inter-subject amplitude differences. In contrast, removing $\mathcal{L}_{\text{diff}}$ causes the smallest decline, suggesting effective representation learning even without explicit subspace separation. These results demonstrate that each objective complements the others, collectively enhancing causal feature learning, suppressing spurious correlations, and improving subject-independent psychiatric screening robustness.

C. Model Interpretability

1) Interpretable Patterns of Temporal Attention: To elucidate the temporal dynamics of biosignals, we visualize attention patterns derived from causally disentangled representations. By removing domain factors, the attention maps more precisely capture class-relevant physiological variations. As shown in Fig. 3, heatmaps highlight critical periods and channels influencing model predictions, where warmer colors (e.g., red) represent higher importance for classification, while cooler colors (e.g., purple) indicate lower relevance. The top three informative temporal segments, identified by cross-channel attention weights ranking, are marked in yellow.

Although healthy controls exhibit weaker activations, clear boundaries between healthy and patient groups are not visually apparent. Nevertheless, when aligned with psychiatric labels, condition-specific temporal physiological information can be captured. For instance, the depression case shows significant activations around 00:04:30–00:05:30 and exhibits autonomic dysregulation and blunted emotional reactivity during the final two minutes. In addition, the color-based comparison among channels highlights that SKT and PPG show the strongest activations under depressive conditions. Tracking these high-importance segments enables detection of symptom-related episodes and supports early clinical intervention.

2) Interpretable Patterns of Channel Connection: We employ chord diagrams to visualize inter-channel connection derived from causal representations, to further explore how their interactions contribute to psychiatric screening. As illustrated in Fig. 4, each arc segment corresponds to a specific biochannel, with color intensity representing its importance in the model's decision process. The thickness of connecting chord indicates the strength of inter-channel dependencies, where thicker chords suggest stronger synergistic contributions. For comparison, we focus on the top three high-performing subjects from depression, anxiety, and bipolar groups.



(a) Group A - sub1 (b) Group A - sub2 (c) Group A - sub3

(d) Group B - sub1 (e) Group B - sub2 (f) Group B - sub3

(g) Group C - sub1 (h) Group C - sub2 (i) Group C - sub3

Fig. 4: Channel Importance and Inter-Channel Connection Across Patient Groups. Group A: Depressive Disorder; Group B: Anxiety Disorder; Group C: Bipolar Disorder.

Within the depressive cohort, strong interactions are observed between OXY_Red and OXY_IRed, as well as between

SKT and GSR across subjects. These patterns likely reflect the sensitivity of SKT and OXY to autonomic physiological regulation, often disrupted in depression [3], [19]. Our observations align with findings by Zhang et al. [20], who report that combination of SKT and GSR yields optimal results during the monitoring of depression episode; For anxiety, GSR and PPG exhibit strong connectivity, reflecting their shared sympathetic origin in sympathetic nervous system activation, which is a hallmark of anxiety-related stress responses. Prior studies [21], [22] further confirm both GSR and PPG as robust biomarkers of anxiety, lending physiological credence to the observed patterns; In the case of bipolar disorder, more diverse and complex channel interactions emerge. This complexity likely reflects the oscillation between depressive and manic or anxious episodes, necessitating a comprehensive multi-channel representation. Among them, interactions involving PPG with OXY_IRed, PPG with OXY_Red, and GSR with SKT are particularly salient. These findings corroborate prior research linking bipolar disorder to autonomic dysregulation and altered PPG responses [23], with additional evidence emphasizing the relevance of GSR and SKT in bipolar symptomatology [4].

V. Conclusion

In this paper, we present G-PsychSW, a generalizable and interpretable wearable system designed for psychiatric screening with physiological insights. Causal disentanglement framework with multi-loss optimization isolates class factor from subject-specific variations, enhancing robustness to domain shifts. Integrated dual-attention mechanism captures temporal dynamics and functional connectivity, adaptively identifying when, which, and how biosignals influence predictions. Visualization of these attention patterns enhances interpretability and aligns with clinical knowledge, supporting practical application. Evaluated on a customized dataset under LOSO-CV, our G-PsychSW achieves 80.16% accuracy and 86.52% F1-score, surpassing existing DG methods.

VI. Acknowledgement

This work was supported in part by China NSFC (No. 62172286, 62472366); in part by key program of DEGP (No. 2024ZDZX1011, 2024GCZX003, 2023KCXTD042); in part by Guangdong Provincial Key Lab of Integrated Communication, Sensing and Computation for Ubiquitous Internet of Things(No.2023B1212010007); in part by Shenzhen Science and Technology Program (No. JCYJ20250604181429038); in part by Shenzhen Science and Technology Foundation (No. ZDSYS20190902092853047); in part by 111 Center (No. D25008); in part by Guangdong Yiyang Healthcare Charity Foundation (No. JZ2022001-3); in part by Talents' plan Foundation of Guangdong Second Provincial General Hospital (No. 2024C003). Both Changhe Fan and Yongpan Zou are the corresponding authors.

References

[1] W. H. Organization, *World mental health report: Transforming mental health for all*. World Health Organization, 2022.

[2] M. İ. Atagün, "Brain oscillations in bipolar disorder and lithium-induced changes," *Neuropsychiatric disease and treatment*, pp. 589–601, 2016.

[3] N. Lorenz, J. Spada, C. Sander, S. G. Riedel-Heller, and U. Hegerl, "Circadian skin temperature rhythms, circadian activity rhythms and sleep in individuals with self-reported depressive symptoms," *Journal of psychiatric research*, vol. 117, pp. 38–44, 2019.

[4] M. Sharma, S. Kacker, and M. Sharma, "A brief introduction and review on galvanic skin response," *Int. J. Med. Res. Prof*, vol. 2, no. 6, pp. 13–17, 2016.

[5] W. Ye, Z. Zhang, F. Teng, M. Zhang, J. Wang, D. Ni, F. Li, P. Xu, and Z. Liang, "Semi-supervised dual-stream self-attentive adversarial graph contrastive learning for cross-subject eeg-based emotion recognition," *TAFFC*, 2024.

[6] W. Kuang, S. Jin, D. Wang, Y. Zheng, Y. Zou, and K. Wu, "Emotracer: A user-independent wearable emotion tracer with multi-source physiological sensors based on few-shot learning," in *BIBM*. IEEE, 2024, pp. 2680–2687.

[7] X.-H. Liu, W.-B. Jiang, W.-L. Zheng, and B.-L. Lu, "Moge: Mixture of graph experts for cross-subject emotion recognition via decomposing eeg," in *BIBM*. IEEE, 2024, pp. 3515–3520.

[8] Z. Jia, F. Zhao, Y. Guo, H. Chen, T. Jiang, and B. Center, "Multi-level disentangling network for cross-subject emotion recognition based on multimodal physiological signals," in *IJCAI*, 2024, pp. 3069–3077.

[9] J. Shen, J. Chen, Y. Ma, Z. Cao, Y. Zhang, and B. Hu, "Explainable depression recognition from eeg signals via graph convolutional network," in *BIBM*. IEEE, 2023, pp. 1406–1412.

[10] Y. Ding, N. Robinson, S. Zhang, Q. Zeng, and C. Guan, "Tsception: Capturing temporal dynamics and spatial asymmetry from eeg for emotion recognition," *TAFFC*, vol. 14, no. 3, pp. 2238–2250, 2022.

[11] H. Yang, C. P. Chen, B. Chen, and T. Zhang, "Improving the interpretability through maximizing mutual information for eeg emotion recognition," *TAFFC*, 2024.

[12] Y. Wang, J. Guo, Z. Jia, G. Cao, Y. Yang, G. Kang, and J. Huang, "Dynseizuregat: Multi-band dynamic graph attention network for interpretable seizure detection and analysis of drug-resistant epilepsy using seeg," *JBHI*, 2025.

[13] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Information fusion*, vol. 37, pp. 98–125, 2017.

[14] Z. Jia, Y. Lin, J. Wang, X. Ning, Y. He, R. Zhou, Y. Zhou, and L.-w. H. Lehman, "Multi-view spatial-temporal graph convolutional networks with domain generalization for sleep stage classification," *TNSRE*, vol. 29, pp. 1977–1986, 2021.

[15] L. Jia, T. W. Chow, and Y. Yuan, "Causal disentanglement domain generalization for time-series signal fault diagnosis," *Neural Networks*, vol. 172, p. 106099, 2024.

[16] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 2002.

[17] C. Yang, M. B. Westover, and J. Sun, "Manydg: Many-domain generalization for healthcare applications," in *ICLR*, 2023.

[18] F. Chen, Y. Zhang, L. Chen, X. Meng, Y. Qi, and J. Wang, "Dynamic traveling time forecasting based on spatial-temporal graph convolutional networks," *FCS*, vol. 17, no. 6, p. 176615, 2023.

[19] R. Takizawa, M. Fukuda, S. Kawasaki, K. Kasai, M. Mimura, S. Pu, T. Noda, S.-i. Niwa, Y. Okazaki, J. P. for Psychiatric Application of Near-Infrared Spectroscopy (JPSY-NIRS) Group *et al.*, "Neuroimaging-aided differential diagnosis of the depressive state," *Neuroimage*, vol. 85, pp. 498–507, 2014.

[20] Y. Zhang, S. Jin, W. Kuang, Y. Zheng, Q. Song, C. Fan, Y. Zou, V. C. Leung, and K. Wu, "Depguard: Depression recognition and episode monitoring system with a ubiquitous wrist-worn device," *TMC*, 2025.

[21] E. Zhou, M. Soleymani, and M. J. Matarić, "Investigating the generalizability of physiological characteristics of anxiety," in *BIBM*. IEEE, 2023, pp. 4848–4855.

[22] J. Šalkevicius, R. Damaševičius, R. Maskeliunas, and I. Laukienė, "Anxiety level recognition for virtual reality therapy system using physiological signals," *Electronics*, vol. 8, no. 9, p. 1039, 2019.

[23] H.-A. Chang, C.-C. Chang, N.-S. Tzeng, T. B. Kuo, R.-B. Lu, and S.-Y. Huang, "Heart rate variability in unmedicated patients with bipolar disorder in the manic phase," *Psychiatry and clinical neurosciences*, vol. 68, no. 9, pp. 674–682, 2014.