RESEARCH ARTICLE

Real-time detection and early warning system of psychological crisis based on meta-learning and small sample learning

Xumei Chen[1, *], Chuan Wang[2]

[1]Mental Health Education Center, Sichuan University of Media and Communications, Chengdu, Sichuan, China. [2]R&D Center, KSW Technologies Co, Ltd., Chengdu, Sichuan, China.

Received: March 3, 2026; **accepted:** June 8, 2026.

Psychological crisis has the characteristics of great individual differences, rapid evolution, and scarce labeled samples, which bring great challenges to real-time detection and early warning. This study proposed a real-time psychological crisis detection and early warning framework leveraging meta-learning and few-shot learning paradigms and introduced a novel Domain-Adaptive Meta-Transfer Learning (DAMTL) methodology. The proposed model underwent pre-training on a publicly available mental health dataset within the source domain, where it acquired a cross-individual, multi-modal feature representation that generalized across subjects. It then combined a meta-learning strategy and domain-adaptive mechanism in the target domain to achieve rapid personalized modeling of new subjects. The studies were carried out on the Computational Linguistics and Clinical Psychology Workshop (CLPsych) and Reddit Suicide Watch and Mental Health Collection (SWMH) datasets, which were set as 3-way and 25-shot tasks. The results showed that the DAMTL method achieved an accuracy of 76.4%, an F1-score of 75.1%, and an Area Under the Curve (AUC) of 0.84. In cross-subject experiments with different datasets, the accuracy rate was 71.6%, the F1-score was 70.3%, and the AUC was 0.81. The overall performance of proposed method was better than that of the model-independent meta-learning algorithm, various transfer learning, and monomodal methods. Ablation experiments showed that domain adaptation, multi-modal feature fusion, and a meta-learning strategy all contributed significantly to the model's performance. The results demonstrated that the proposed method exhibited robust stability and strong generalization capabilities under the constraints of few-shot learning and cross-distribution scenarios, thereby establishing a viable technical framework for the early identification and precise forecasting of psychological crises.

Keywords: psychological crisis detection; meta-learning; small sample learning; domain adaptation; multimodal feature.

[*]**Corresponding author:** Xumei Chen, Mental Health Education Center, Sichuan University of Media and Communications, Chengdu, Sichuan 611745, China. Email: cxm112022@163.com.

Introduction

The accelerating tempo of contemporary society has led to a notable escalation in psychological distress among the general population. Mental health has become an important topic of social concern. Psychological crises are often sudden and hidden, and individuals may not take the initiative to seek help in their daily lives. Traditional mental health intervention methods mostly rely on questionnaire survey or face-to-face evaluations, which are time-consuming and laborious, and are unable to realize real-time monitoring [1]. The advancement of information and communication technologies has led social media platforms and mobile devices to generate vast volumes of user behavioral data, thereby opening novel avenues for the analysis and forecasting of psychological states. Text, behavior, and social interaction data contain rich information about emotions, cognition, and social patterns. The challenge of identifying early indicators of individual psychological crises from such data has emerged as a central research focus within the domains of psychological computing and artificial intelligence [2].

Currently, research on psychological crisis detection has established a technical framework centered on natural language processing, machine learning, and multimodal analysis. Some studies have attempted to extract negative emotions and risk indicators from user text to identify potential psychological abnormalities. Swaminathan *et al*. developed a two-stage model based on keyword filtering and logistic regression in a telemedicine setting to screen for psychological crisis information in patient conversations [3]. Wu developed a psychological crisis identification system for college students, conducting a joint analysis of psychological profiles and questionnaire data to validate the application value of multi-source data fusion in early detection [4]. Nikitina *et al*. analyzed parameters from two text corpora with one comprising clinical depression patients and the other individuals with high depressive tendencies to identify common and unique markers in depressive texts and proposed three comprehensive depression indices, providing a methodological reference for AI-based online psychological diagnosis [5]. In addition to traditional supervised learning, multimodal modeling has gradually garnered attention. Text content, behavioral characteristics, and social relationships are collectively incorporated into the analytical framework to characterize more complex processes of psychological state changes [6]. However, the identification of mental health crises in real-world scenarios still faces significant challenges. There are marked differences among individuals in terms of expression habits, emotional responses, and social behaviors, thus, the same negative language may indicate entirely different levels of risk across different users [7]. Many models rely on large amounts of labeled data for training, which can lead to performance degradation when applied across users or platforms. Mental health data itself is characterized by privacy sensitivity, high annotation costs, and imbalanced class distributions, meaning that high-quality data available for training is often limited. In recent years, few sample learning and meta-learning methods have shown unique advantages in quickly adapting to new tasks. They can use limited data to establish personalized forecasting ability [8, 9], and transfer learning provides shared knowledge for the model from a large amount of historical data [10]. Meanwhile, the difference in the distribution of different individual data will lead to migration deviation, and domain adaptive technology can build a bridge between different data distributions, making the model more stable when facing new users [11]. Compared to traditional supervised learning, meta-learning and few-shot learning place greater emphasis on a model's ability to adapt rapidly under sample-constrained conditions [12]. These approaches have already yielded some results in tasks such as depression detection, medical diagnosis, and intelligent question-answering. Nasim *et al*. analyzed 1,503 questionnaire data points from medical institutions, integrating decision trees, K-nearest neighbor classifiers, random forest models, and the meta-learning multi-layer perceptron to propose a meta-learning model for predicting postpartum depression, thereby providing an effective tool for assessing mental health risks in pregnant women and new mothers [13]. Liu *et al*. introduced the Model-Agnostic Meta-Learning (MAML) algorithm into low-sample-size disease diagnosis. By jointly optimizing knowledge graph model parameters and learning rate settings, they improved the diagnostic accuracy for rare diseases in primary care settings [14]. Jia *et al*. addressed the limited generalization capability of MAML in few-shot classification by proposing the Mix-MAML method. By combining data augmentation, initialization decay, and resolution enhancement strategies, they achieved classification performance that surpassed or matched existing methods on multiple public datasets while maintaining the simplicity of the model architecture [15]. These studies demonstrate that meta-learning and transfer mechanisms hold significant potential in low-data-sample settings. However, research on real-time psychological crisis detection remains limited. Existing psychological crisis detection studies largely rely on ample labeled data, and models often exhibit performance degradation when applied across different subjects or scenarios with limited adaptability to individual differences. Although meta-learning and few-shot learning methods have performed well in other fields, they are relatively underdeveloped in the context of psychological crisis detection.

This research proposed a Domain-Adaptive Meta-Transfer Learning (DAMTL) model to enable real-time detection and early warning of personalized psychological crises through cross-subject and cross-dataset validation using the Computational Linguistics and Clinical Psychology Workshop (CLPsych) and Reddit Suicide Watch and Mental Health Collection (SWMH) datasets. Through pre-training feature encoders on extensive user datasets from the source domain, the model acquired cross-subject shared knowledge, subsequently mitigating the distribution divergence between the source domain and novel user data *via* a domain adaptation module. The target domain phase combined meta-learning strategies to rapidly establish personalized psychological baselines with limited data and identify abnormal patterns of change. The system ultimately output psychological crisis risk probabilities and tiered early warning results to support real-time intervention and dynamic monitoring.

Materials and methods

Overall framework of the DAMTL model

DAMTL model consisted of multimodal feature coding, source domain pre-training, domain adaptive adjustment, and target domain element learning. The model encoded the user's text, behavior, and social data into a unified feature representation and learned the general mental model on the large-scale data in the source domain. The domain adaptation module fine-tuned a subset of the pre-trained model parameters within the target domain, thereby mitigating the distribution shift between the source domain and novel user data. Meta-learning strategy used support set and query set to adapt quickly, so that the model could establish personalized psychological baseline and identify cooperative signal patterns that deviated from the baseline. The model output risk probability and graded early warning, which provided a basis for early identification of individual psychological state. The whole framework maintained a clear division of labor among modules, and features extraction, adaptation, and meta-learning cooperated with each other, to realize rapid response to new users' small sample data (Figure 1).

Source domain task learning

The primary objective of source domain task learning within the DAMTL framework was to derive shared feature representations from source domain datasets, which then served as the initialization point for subsequent target domain meta-learning. Multi-modal feature encoders were used to pre-train different modes to generate feature representations. The text modality was initialized *via* pre-training on Bidirectional Encoder Representations from Transformers (BERT) (https://doi.org/10.48550/arXiv.1810.04805), which projected source domain textual inputs into a high-dimensional semantic feature space [16]. Behavior mode used a Multi-Layer Perceptron (MLP) (https://en.wikipedia.org/wiki/Multilayer_perceptron) to extract user behavior characteristics such as clicking, browsing and interaction times [17]. Social patterns were pre-trained by Graph Neural Network (GNN) (https://doi.org/10.1038/s41598-021-95102-7), and social relations and interaction patterns were mapped into node-level features [18]. The features of the three

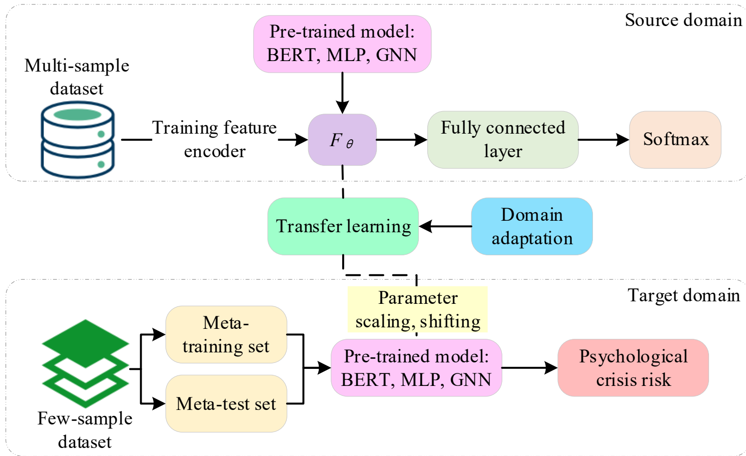

Figure 1. Overall structure diagram of DAMTL.

modes were fused at the encoder output, and a unified multi-modal representation $F_\theta(X)$ was obtained. In the source domain training stage, the influence of other individuals was not considered for the time being, and only the preprocessed multimodal data was used to iteratively optimize the parameter F_θ of the feature encoder until the performance reached the preset goal or the upper limit of training rounds. Let the source domain sample be X, the corresponding label be Y, the classifier be C_ϕ, and the model prediction $\hat{Y}$ be the follows.

$$\hat{Y} = C_\phi(F_\theta(X)) \tag{1}$$

The empirical loss was formally defined using a cross-entropy loss function below.

$$\mathcal{L}_{exp} = -\sum_i Y_i \log(\hat{Y}_i) \tag{2}$$

The feature encoder parameters θ and classifier parameters ϕ were then updated using the optimizer below.

$$\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L}_{exp} \tag{3}$$

$$\phi \leftarrow \phi - \eta \nabla_\phi \mathcal{L}_{exp} \tag{4}$$

where η was the learning rate. After completing source domain training, the feature encoder parameters θ were frozen for use in the target domain task, while the classifier parameters ϕ were discarded for the few-shot target task. To minimize parameter updates during transfer, scaling factors S_W and translation factors S_b were introduced to fine-tune the neural network weights W and biases b as follows [19].

$$W' = W \odot S_W \tag{5}$$

$$b' = b \odot S_b \tag{6}$$

where $\odot$ was element multiplication. For the training subset T_{train} associated with the target task T, the corresponding loss function was defined as below.

$$\mathcal{L}_T = \mathcal{L}_{exp}(C_\phi(F_\theta(X_T)), Y_T) \tag{7}$$

The scaling and translation factors were initially $S_W = 1$ and $S_b = 0$ and optimized by gradient descent in the training process as follows.

$$S_W \leftarrow S_W - \beta \nabla_{S_W} \mathcal{L}_T \tag{8}$$

$$S_b \leftarrow S_b - \beta \nabla_{S_b} \mathcal{L}_T \tag{9}$$

where β was the factor learning rate. After updating, the feature encoder output the following results.

$$F'_\theta(X) = S_W \odot F_\theta(X) + S_b \tag{10}$$

This mechanism was deployed to facilitate support set and query set operations during the target domain meta-learning phase. On the premise of ensuring knowledge sharing in the source domain, this method avoided parameter forgetting in the migration process and provided reliable initialization for small sample adaptation of new tasks.

Domain adaptation module

The domain adaptation component integrated into the DAMTL framework served to mitigate distribution divergence between the source and target domains, while preserving the meta-learning paradigm's adaptability to novel tasks and minimizing feature discrepancy across source and target task distributions. In the process of target domain migration, directly migrating the pre-trained multi-modal feature encoder to the meta-learning stage would update all parameters, which would easily lead to the loss of learned shared features [20]. The domain adaptive module allowed most of the pre-training parameters to be frozen, and only a few parameters were updated, which ensured the retention of source domain knowledge and improved the adaptability in small sample learning in the target domain [21]. Maximum Mean Discrepancy (MMD) is a widely adopted metric for addressing domain adaptation challenges [22]. Let the sample set in the source domain be X_s, the sample set in the target domain be X_t, and the sample mean values be μ_s and μ_t, respectively. MMD quantified the distribution divergence between the source and target domains by evaluating discrepancies in their respective empirical means with the formal definition given as follows.

$$MMD(X_s, X_t) = ||\mu_s - \mu_t||^2 \quad (11)$$

Domain adaptive loss $\mathcal{L}_{DA}$ incorporated an MMD term into the training objective, enabling the feature encoder to minimize distribution divergence between the two domains while preserving source domain knowledge as below.

$$\mathcal{L}_{DA} = MMD(F_\theta(X_s), F_\theta(X_t)) \quad (12)$$

In the meta-learning stage, the optimizer only updated some trainable parameters θ' and frozen the rest of the parameters. The parameter updating equation was shown below.

$$\theta' \leftarrow \theta' - \eta \nabla_{\theta'} \mathcal{L}_{DA} \quad (13)$$

The refined feature encoder was capable of producing feature representations that discriminated between the source and target domains, and meanwhile, it considered the fast adaptability of a few samples, providing stable input for the meta-learning task in the target domain.

Target domain meta-learning

Target domain meta-learning was mainly used to quickly establish a personalized psychological baseline under the conditions of new users and few samples. In the proposed model, it leveraged the domain-adapted multi-modal feature encoder F'_θ to extract multi-modal features from the target task and the parameters of the basic learner L_ϕ being trained by the support set. The support set comprised a limited number of annotated samples, which were utilized to train base learners for capturing task-specific individual characteristics. Following support set training, the base learner's performance was assessed on the query set, and the meta-learner parameters were optimized using the query set loss to achieve rapid cross-task adaptability. Let the target task set be $\{T_i\}$, and each task included a support set S_i and a query set Q_i. The loss function of the basic learner training on the support set was expressed below.

$$\mathcal{L}^i_{support} = CE(L_\phi(F'_\theta(S_i)), Y_{S_i}) \quad (14)$$

The base learner L'_{ϕ_i} computed the test loss on the query set *via* parameter updates, which was then used to refine the meta-learner parameter ϕ as follows.

$$\mathcal{L}^i_{query} = CE(L'_{\phi_i}(F'_\theta(Q_i)), Y_{Q_i}) \quad (15)$$

Upon completion of all meta-training tasks, the meta-learner was capable of rapidly adapting to limited data inputs when confronted with unseen tasks and realized personalized prediction of the psychological state of new users. The support set of the unknown task $\{T_j^{\text{new}}\}$ was employed to fine-tune the base learner, while the query set was utilized to assess its performance. The model's generalization performance on novel tasks was quantified by using the average of evaluation results. This approach preserved the shared knowledge embedded in the source domain while effectively leveraging limited target domain data, giving consideration to stability and rapid adaptability.

Psychological crisis risk output

After the completion of the learning stage of target domain elements, the DAMTL model could evaluate the psychological state of new users in real time. The characteristics of prediction meant that the risk scores were obtained through the basic learner first, and these scores reflected the possibility of individual psychological crisis in the current time period. To facilitate understanding and decision-making, the model mapped the risk score to three levels of hierarchical early warning including low, medium, and high. Each level corresponded to different concerns and intervention suggestions. Risk output not only considered a single negative emotion but also integrated the collaborative patterns of multi-dimensional signals such as emotion, cognition and social behavior like depression accompanied by social withdrawal or cognitive distress plus abnormal behavior patterns. The model could quickly adjust the prediction results with a small amount of data, and timely capture the abnormal state that deviated from the personal psychological baseline. The outputs generated were applicable to a diverse range of scenarios such as mental health monitoring, personalized intervention recommendations, and research analytics. Under the premise of protecting users' privacy, the system could realize the early identification of high-risk psychological crisis by using only a small amount of historical data and provide a basis for further intervention or decision-making. The whole output process was closely connected with the previous feature extraction, domain adaptation, and meta-learning, so that the model could maintain stable prediction ability in a dynamic environment, considering individual differences and small sample adaptation needs.

Experimental data set and experimental setup

In this research, two public mental health-related datasets were used to verify the performance of the DAMTL model under the same dataset and cross-dataset conditions, which included the Computational Linguistics and Clinical Psychology Workshop (CLPsych) 2016 dataset [23] (https://github.com/IKBLab/CLPsych2016) and SWMH dataset [24] (https://zenodo.org/records/6476179). CLPsych dataset contained emotional, cognitive and social behavior labels of multiple subjects, which was suitable for cross-subject experiments with the same dataset, while SWMH dataset contained suicide-related text and sentiment/behavior labels, which could be used for domain transfer experiments. Both datasets were preprocessed including text vectorization, behavioral data standardization, and time series feature generation to be input into DAMTL model. All experiments were conducted on a workstation equipped with a 64-bit Windows 10 operating system, an Intel Core i7 processor, 8 GB of RAM, and an NVIDIA RTX 3060 graphics card. It was implemented in Python 3.9 (https://www.python.org) using PyCharm (https://www.jetbrains.com/pycharm/) and the GPU-Keras (https://keras.io) framework. For the DAMTL model, the number of epochs was set to 10 during the pre-training phase and 50 during the meta-learning phase. The learning rate for the base learner was set to 5×10^{-3}, and the learning rate for the meta-learner was set to 5×10^{-4}. The Adam optimizer was used with a dropout rate of 0.2 and a batch size of 64. During training, the learning rate was halved twice every 100 steps, and an early stopping strategy was employed to prevent overfitting. All experiments were evaluated using 10-fold cross-validation. Three types of tasks were conducted in the study, which included cross-subject experiments with the same dataset, cross-subject experiments with different datasets, and ablation experiments. The cross-subject experiments with the same dataset divided the CLPsych dataset into a training set and a test set, in which the training set was used for task learning and meta-training in the source domain, and the test set was used for small sample adaptation in the target domain. Psychological risk levels were categorized into three groups of low risk, medium risk, and high risk with N-way being set to 3. To examine the impact of different K-shot support set sizes on model performance, the support set sizes were set sequentially to 5, 10, 15, 20, 25, 30, 35, and 40, and the accuracy, F1-score, and Area Under the Curve (AUC) were recorded. The cross-subject experiments with different datasets defined CLPsych as the source domain and SWMH as the target domain, enabling evaluation of the DAMTL model's adaptability and generalization capability under conditions of domain distribution shift. In experiments on the same dataset, DAMTL adopted a 3-way, 25-shot task setup, where each task consisted of three classes with 25 training samples provided for each class. To reduce the variability caused by random partitioning, the models in both experiments were run independently three times under the same settings with the training and test sets repartitioned each time. The results were analyzed based on the statistical performance across these three runs. MAML (https://doi.org/10.48550/arXiv.1703.03400), Prototypical Network (ProtoNet) (https://doi.org/10.48550/arXiv.1703.05175), Fine-tuned Multimodal Bidirectional Encoder Representations from Transformers (FT-MBERT) (https://doi.org/10.1007/s10660-022-09560-w), Multimodal Convolutional Neural Network–Long Short-Term Memory (CNN-LSTM) (https://doi.org/10.1016/j.compmedimag. 2022. 102128), BERT, MLP, and GNN were used as baseline models. DAMTL was compared with these baseline models under consistent settings. Paired t-tests were used to analyze the results of three independent experiments, verifying the statistical significance of performance differences between DAMTL and each baseline model with a significance level set at $P < 0.05$. Statistical analysis was performed in a Python 3.9 environment using the scipy.stats module from SciPy library (https://scipy.org/). All models were compared under the same data partitioning and experimental settings to ensure the comparability of statistical results. Ablation experiments were designed to sequentially remove the domain adaptation module, multimodal feature fusion, and meta-learning strategy under the same dataset and cross-subject conditions, which was done to evaluate the contribution of each component module of the DAMTL model to performance. The task settings remained the same as 3-way and 25-shot.

Results and discussion

Support set partition analysis

The results demonstrated that expanding the support set size from 5-shot to 25-shot yielded significant improvements in the model's accuracy, F1-score, and AUC with the increases of 7.15%, 7.59%, and 7.69%, respectively. When the K-shot exceeded 25, the performance improvement tended to be stable. The accuracies of 30-shot, 35-shot, and 40-shot were all about 76.3%, while F1-scores and AUCs were basically the same (Figure 2). If the support set was too small, the rapid adaptability of the model would be limited, and a small number of samples could not fully characterize the task characteristics, which made the accuracy and F1-score fluctuate greatly. With the increase of support set samples, the model could establish an individual psychological baseline more stably and improve the prediction accuracy. After more than 30 samples, the performance gain was not obvious, which showed that the meta-learning strategy of DAMTL had fully played its role in the scene with few samples. Based on comprehensive accuracy, F1-score, and AUC performance, 25-shot as a support set size could give consideration to both performance and training cost and was a reasonable parameter selection.

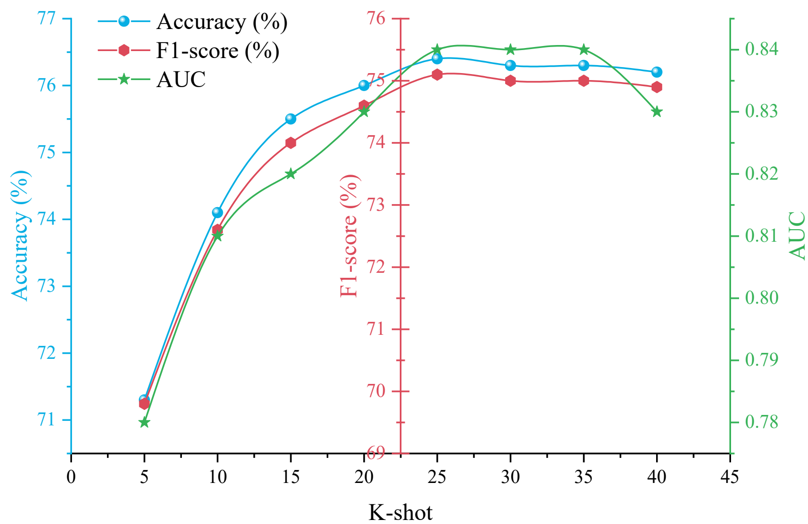

Figure 2. Impact of K-shot Support Set Size on Performance.

Table 1. Results from three cross-subject experiments of the DAMTL model with the same dataset.

Number of experiments	Accuracy (%)	F1-score (%)	AUC
Experiment 1	76.3	74.9	0.83
Experiment 2	76.5	75.2	0.84
Experiment 3	76.3	75.1	0.84
Mean ± SD	76.4 ± 0.12	75.1 ± 0.15	0.84 ± 0.006

Model performance verification

The accuracy, F1-score, and AUC values across the three experiments exhibited narrow ranges and low standard deviations (SD) (Table 1), indicating the DAMTL model's high stability and reliability when applied to cross-subject analysis on the same dataset. The model could quickly establish an individual psychological baseline with a small number of support sets while maintaining adaptability to different subjects. The comparison results between DAMTL and the baseline models MAML, ProtoNet, FT-MBERT, multimodal CNN-LSTM, monomodal BERT, MLP, and GNN showed that DAMTL was superior to other methods in accuracy, F1-score, and AUC. Compared with MAML, the accuracy, F1 score, and AUC of DAMTL increased by 3.4%, 3.6%, and 0.04%, respectively, while compared with ProtoNet, the improvement was more obvious. Compared with the best performing FT-MBERT, the accuracy, F1-score, and AUC of DAMTL were improved by 2.96%, 3.3%, and 3.7%, respectively, which showed that the domain adaptation and multimodal element-transfer learning strategy of DAMTL were effective in stably capturing individual psychological characteristics. The performance of the single-mode model was generally low, which verified the necessity and advantages of multi-mode feature fusion (Figure 3). The results confirmed that DAMTL was stable and had high prediction accuracy in cross-subject experiments with the same dataset. Compared with traditional meta-learning or transfer learning methods, DAMTL was more effective in small sample adaptation and individualized risk identification. The results of the performance metrics of the DAMTL model showed that the accuracy, F1-score, and AUC values across the three experimental replicates remained within narrow ranges of 71.5 - 71.8% for accuracy, 70.2 - 70.5% for F1-score, and 0.81 - 0.82 for AUC, demonstrating minimal performance fluctuation (Table 2). The model maintained adaptability in cross-domain tasks, effectively extracting

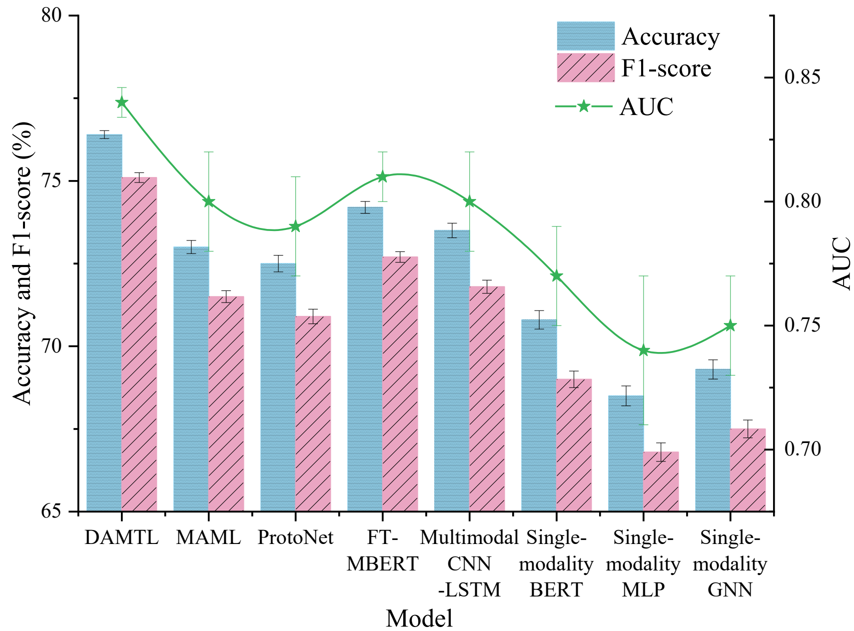

Figure 3. Comparison of cross-subject experimental results using the same dataset.

Table 2. Results from three cross-subject experiments using different datasets for the DAMTL model.

Number of experiments	Accuracy (%)	F1-score (%)	AUC
Experiment 1	71.8	70.5	0.82
Experiment 2	71.5	70.3	0.81
Experiment 3	71.6	70.3	0.81
Mean ± SD	71.6 ± 0.15	70.3 ± 0.12	0.81 ± 0.006

individual psychological characteristics for risk prediction even when source and target domains exhibited distributional differences. The performance of DAMTL alongside MAML, ProtoNet, FT-MBERT, Multimodal CNN-LSTM, and unimodal models in cross-dataset tasks demonstrated that DAMTL still maintained high performance in cross-dataset experiments with accuracy, F1-score, and AUC of 71.6%, 70.3%, and 0.81, respectively. Compared with MAML, DAMTL increased by 5.29%, 5.71%, and 3.85%, respectively. The unimodal model displayed an obvious decline in the migration from the source domain to the target domain, which showed that multimodal feature fusion and domain adaptive strategy played an important role in dealing with the distribution difference between the source domain and the target domain (Figure 4). DAMTL not only maintained stability in cross-subject experiments with different datasets but also was superior to traditional meta-learning, transfer learning, and monomodal models, and could quickly adapt to new tasks and make accurate psychological crisis risk prediction in the face of distribution differences and few samples. The results of paired t-test analyzing the outcomes of three independent experiments indicated that, in cross-subject experiments on the same dataset, comparisons between DAMTL and MAML, ProtoNet, FT-MBERT, and other baseline models showed that DAMTL achieved statistically significant improvements in accuracy, F1-score, and AUC (P < 0.05) with particularly notable differences compared to FT-MBERT and MAML. In cross-subject experiments across different datasets, DAMTL also demonstrated a significant advantage over all baseline models (P < 0.05), indicating that it retained stable generalization

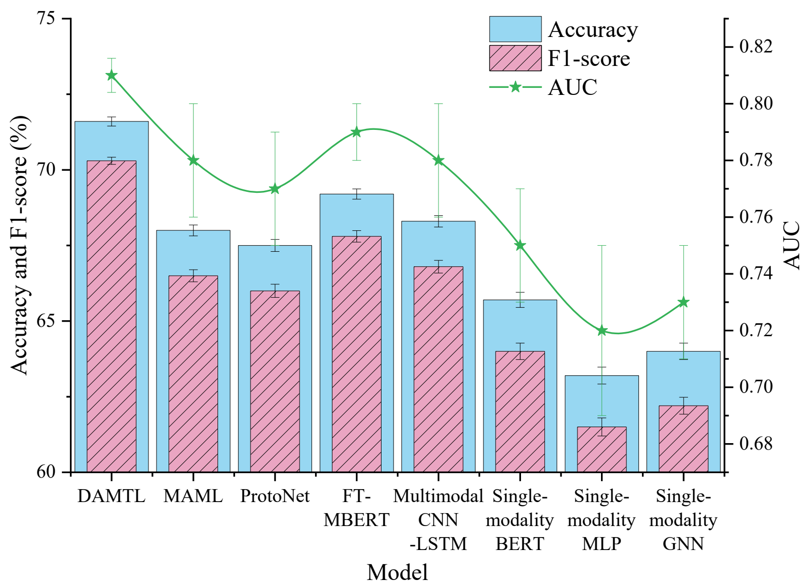

Figure 4. Comparison of cross-subject experimental results across different datasets.

Table 3. Results of the paired t-test for model performance across the two experiments.

Experimental setup	Comparison models	t-value (Accuracy)	p-value (Accuracy)	t-value (F1-score)	p-value (F1-score)	t-value (AUC)	p-value (AUC)
Cross-subject experiments using the same dataset	MAML	4.12	0.014	4.35	0.011	3.98	0.016
	ProtoNet	4.68	0.009	4.91	0.007	4.22	0.013
	FT-MBERT	3.75	0.018	3.89	0.015	3.41	0.022
	CNN-LSTM	4.01	0.015	4.20	0.012	3.88	0.017
	BERT	5.02	0.006	5.21	0.004	4.76	0.008
	MLP	5.89	0.003	6.10	0.002	5.44	0.005
	GNN	5.36	0.004	5.58	0.003	4.92	0.007
Cross-subject experiments using different datasets	MAML	5.03	0.006	5.18	0.005	4.62	0.009
	ProtoNet	5.41	0.004	5.63	0.003	4.95	0.006
	FT-MBERT	4.36	0.012	4.52	0.010	4.01	0.015
	CNN-LSTM	4.89	0.007	5.05	0.006	4.38	0.011
	BERT	6.02	0.002	6.18	0.002	5.55	0.004
	MLP	6.74	0.001	6.91	0.001	6.08	0.003
	GNN	6.11	0.002	6.29	0.002	5.67	0.004

capabilities in distribution transfer scenarios (Table 3). The statistical results confirmed that the performance improvement of DAMTL was not due to random fluctuations but maintained a consistent advantage across multiple repeated experiments, further validating the effectiveness of domain adaptation and meta-transfer learning strategies.

Ablation experiment

The results of the ablation study showed that, after the domain adaptation module was removed, the accuracy rate decreased by 2.4%, F1-score decreased by 2.5%, and AUC decreased by 0.03, which confirmed that domain adaptation played an obvious role in alleviating the distribution difference between the source domain and the target domain and improving cross-subject adaptability. The removal of multi-modal feature fusion had the greatest impact on

Table 4. Ablation experiment results.

Model Variants	Accuracy (%)	F1-score (%)	AUC
Full DAMTL Model	76.4	75.1	0.84
Without Domain Adaptation Module	74.0	72.6	0.81
Without Multimodal Feature Fusion	72.2	70.5	0.79
Without Meta-Learning Strategy	73.0	71.2	0.80

performance with the accuracy reduced to 72.2%, F1-score reduced to 70.5%, and AUC reduced to 0.79, indicating the core position of multi-modal feature fusion in capturing individual psychological state. After removing the meta-learning strategy, the performance also dropped, but it was slightly better than only removing the multi-modal features, which showed that meta-learning provided a stable optimization framework for rapid adaptation with few samples. From the results of ablation experiments, it could be concluded that each module contributed significantly to the overall performance of the DAMTL model. The domain adaptive module ensured the stability of cross-task migration, multimodal feature fusion improved feature expression ability, and the meta-learning strategy enhanced the adaptation speed of few samples. The combination of the three formed a complete DAMTL model, which could maintain accuracy and stability in tasks with a few samples, cross-subjects, and cross-data sets.

Conclusion

This study addressed the challenge of real-time psychological crisis detection under low-sample-size conditions by developing the DAMTL model, which was based on meta-learning and domain adaptation to enable psychological risk identification across subjects and datasets. The model demonstrated stable performance in experiments on CLPsych and SWMH datasets, achieving superior accuracy, F1-score, and AUC under 3-way and 25-shot settings, indicating strong generalization capabilities under conditions of limited samples and varying distributions. The results indicated that the model exhibited some sensitivity to the K-shot scale with more pronounced performance improvements at lower sample sizes, gradually stabilizing as the support set increases, reflecting the role of meta-learning mechanisms in rapidly adapting to new tasks. Ablation experiments further validated the combined contributions of domain adaptation, multimodal feature fusion, and meta-learning strategies to performance enhancement, highlighting the critical role of multimodal information in psychological risk modeling. However, there were still some limitations in this study. The experiment was mainly based on text-related mental health datasets, and the modal types and scenarios were relatively limited. The real-time and long-term stability of the model in the real online environment was not fully verified. In the future, more modal information such as physiological signals and behavior logs should be introduced to further improve the comprehensiveness of risk identification. Meanwhile, the lightweight model structure should be explored to meet the actual deployment needs, and continuous time series modeling may be combined to promote the psychological crisis early warning from offline analysis to continuous monitoring and dynamic intervention.

References

1. Li Z, An Z, Cheng W, Zhou J, Zheng F, Hu B. 2023. MHA: A multimodal hierarchical attention model for depression detection in social media. Health Inf Sci Syst. 11(1):6.
2. Khoo LS, Lim MK, Chong CY, McNaney R. 2024. Machine learning for multimodal mental health detection: A systematic review of passive sensing approaches. Sensors. 24(2):348.
3. Swaminathan A, López I, Mar RAG, Heist T, McClintock T, Caoili K, *et al*. 2023. Natural language processing system for rapid

detection and intervention of mental health crisis chat messages. NPJ Digit Med. 6(1):213.
4. Wu Y. 2025. Data fusion model for psychological crisis early warning system using data mining techniques. Informatica. 49(23):179-196.
5. Nikitina EN, Stankevich MA, Chudova NV. 2025. The problem of identifying text markers of depression and depressiveness in automatic text analysis. RUDN J Psychol Pedagogics. 22(1):123-143.
6. Rezk M, Elmadany N, Hamad RK, Badran EF. 2023. Categorizing crises from social media feeds *via* multimodal channel attention. IEEE Access. 11:72037-72049.
7. Ding Z, Zhou Y, Dai AJ, Qian C, Zhong BL, Liu CL, *et al*. 2025. Speech based suicide risk recognition for crisis intervention hotlines using explainable multi-task learning. J Affect Disord. 370:392-400.
8. Gharoun H, Momenifar F, Chen F, Gandomi AH. 2024. Meta-learning approaches for few-shot learning: A survey of recent advances. ACM Comput Surv. 56(12):1-41.
9. Tian P, Yu H. 2023. Can we improve meta-learning model in few-shot learning by aligning data distributions? Knowl Based Syst. 277:110800.
10. Mansoor MA, Ansari KH. 2024. Early detection of mental health crises through artificial-intelligence-powered social media analysis: A prospective observational study. J Pers Med. 14(9):958.
11. Greco CM, Simeri A, Tagarelli A, Zumpano E. 2023. Transformer-based language models for mental health issues: A survey. Pattern Recognit Lett. 167:204-211.
12. Tang Y, Wu Y, He Y, Liu J, Zhang W. 2025. Dual-task contrastive meta-learning for few-shot cross-domain emotion recognition. Comput Mater Contin. 82(2):2331.
13. Nasim S, Al-Shamayleh AS, Thalji N, Raza A, Abualigah L, Alzahrani AI, *et al*. 2024. Novel meta learning approach for detecting postpartum depression disorder using questionnaire data. IEEE Access. 12:101247-101259.
14. Liu Q, Tian Y, Zhou T, Lyu K, Xin R, Shang Y, *et al*. 2024. A few-shot disease diagnosis decision making model based on meta-learning for general practice. Artif Intell Med. 147:102718.
15. Jia J, Feng X, Yu H. 2024. Few-shot classification *via* efficient meta-learning with hybrid optimization. Eng Appl Artif Intell. 127:107296.
16. Beniwal R, Saraswat P. 2024. A hybrid BERT-CNN approach for depression detection on social media using multimodal data. Comput J. 67(7):2453-2472.
17. Wang M, Xu A, Fan C, Sun X. 2025. Machine learning for predicting personality and psychological symptoms from behavioral dynamics. Electronics. 14(3):583.
18. Xia Y, Liu L, Dong T, Chen J, Cheng Y, Tang L. 2024. A depression detection model based on multimodal graph neural network. Multimed Tools Appl. 83(23):63379-63395.
19. Zhang X, Zhang X, Chen W, Li C, Yu C. 2024. Improving speech depression detection using transfer learning with wav2vec 2.0 in low-resource environments. Sci Rep. 14(1):9543.
20. Sharma A, Sharma K, Kumar A. 2023. Real-time emotional health detection using fine-tuned transfer networks with multimodal fusion. Neural Comput Appl. 35(31):22935-22948.
21. Khoei TT, Singh A. 2024. A survey of emotional artificial intelligence and crimes: Detection, prediction, challenges and future direction. J Comput Soc Sci. 7(3):2359-2402.
22. Chen T, Guo Y, Hao S, Hong R. 2023. Semi-supervised domain adaptation for major depressive disorder detection. IEEE Trans Multimed. 26:3567-3579.
23. Garg M. 2023. Mental health analysis in social media posts: A survey. Arch Comput Methods Eng. 30(3):1819.
24. Bauer B, Norel R, Leow A, Rached ZA, Wen B, Cecchi G. 2024. Using large language models to understand suicidality in a social media–based taxonomy of mental health disorders: Linguistic analysis of reddit posts. JMIR Ment Health. 11:e57234.