

RESEARCH ARTICLE

Data-Segmentation Verification and a Target Generative Adversarial Network: EEG-Based Emotion Recognition

Lufeng Yin¹  and Yong Li^{1,*} ¹*School of Electrical Engineering, Shenyang University of Technology, China*

Abstract: Emotion recognition is a crucial component of artificial intelligence. As one of the main factors in emotion recognition, data-driven affective computing heavily relies on high-quality training data, which may not always be readily available due to various reasons. Addressing the challenge of data augmentation with inter-class and intra-class imbalances in emotion-evoking data is a critical issue in affective computing. Currently, many researches have addressed the problem of inter-class imbalance, in which data segmentation processing methods are widely used. However, the rationality of data segmentation methods needs to be verified; meanwhile, the solution to the intra-class imbalance problem remains to be solved. In this paper, we validate the rationality of data segmentation methods through experiments and propose a targeted data generation mechanism. This mechanism intentionally generates pseudo samples in proximity to the often-overlooked samples, aimed at mitigating intra-class imbalance. Combined with Wasserstein generative adversarial network-gradient penalty and generative adversarial network-based self-supervised, we have got T-WGAN-GP and T-GANSER. We then apply these approaches to the DEAP dataset for emotion recognition with different data segmentation methods, including segmenting the data after (MI) or before (MII) the division of the training and testing sets, as well as no data segmentation at all (MIII). Results show that MII and MIII in the segmentation method are theoretically sound, and T-WGAN-GP obtains the best accuracy in the reasonable segmentation method due to its targeted data generation mechanism. This mechanism effectively mitigates the intra-class imbalance to some extent.

Keywords: emotions, generative adversarial networks (GANs), intra- and inter-class imbalance, physiological signals, data segmentation method

1. Introduction

With the swift advancement of population aging, the demand for elderly care has surged, placing substantial strain on the provision of emotional well-being support. Recognizing the emotional states of the elderly, whether they are content, despondent, or discontent, has become a formidable challenge, hindering the delivery of effective services and care. This issue is particularly pronounced in places such as nursing homes and hospitals, where staffing is limited and the growing number of elderly individuals poses challenges for caregivers, making comprehensive care a complex task. As population aging continues to progress, these problems have become more evident.

With the advancement of robot technology, emotion recognition technology based on physiological signals has been proposed as a potential solution to this problem. Emotion, defined as a conscious mental reaction (such as anger or fear) subjectively experienced as a strong feeling, typically accompanied by physiological and behavioral changes in the body [1], can be objectively measured using physiological signals such as EEG signals, ECG signals, and skin signals. These physiological signals are spontaneously generated by the human body and are not easily influenced by subjective

thoughts [2], making emotion recognition based on physiological signals a promising approach to obtain more objective and reliable results. In recent years, many researchers have considered physiological signals as reliable indicators for emotion recognition [3]. This technology has wide application prospects in fields such as distance education [4], medical care [5, 6], human-computer interaction [7], and others, including care for the elderly.

By utilizing emotion recognition based on physiological signals, it is possible to identify the emotions of the elderly and provide tailored services based on their emotional state, leading to improved care outcomes. This technology has the potential to achieve a better care effect by addressing the challenges associated with providing care for the elderly in settings such as nursing homes and hospitals, where staff resources may be limited.

EEG has certain limitations that need to be addressed in the context of emotion recognition. Firstly, EEG signals have a low signal-to-noise ratio as they represent the combined activity of millions of neurons, making it challenging to extract meaningful information from the data. Secondly, the original feature dimension of EEG is typically high, and the limited availability of data samples in typical cognitive neuroscience datasets further exacerbates the issue of a low sample-to-feature ratio. This can pose challenges for machine learning algorithms. Additionally, EEG signals are non-stationary, meaning that they can change over time, further complicating the analysis and interpretation of the data.

*Corresponding author: Yong Li, School of Electrical Engineering, Shenyang University of Technology, China. Email: liyong@sut.edu.cn

To address the issue of limited data, data augmentation techniques, such as data segmentation, have been used in emotion recognition. Data segmentation involves dividing the continuous EEG data into shorter segments for analysis. However, there are challenges associated with data segmentation. In emotion recognition verification, it is important to avoid data crossover between the training and testing sets, meaning that the segmented data, obtained from the same original data, should not be used in both sets. Data crossover can significantly impact the accuracy and validity of the verification process. For instance, when applying a trained network for emotion recognition in real-world scenarios, the data must be collected from individuals or situations that are outside of the dataset used for training the network. This means we must avoid data crossover during the network training phase.

Moreover, the manifestation of emotional expression in EEG signals necessitates a specific time interval, and the exact duration of emotional expression cannot be precisely determined. Therefore, data segmentation may result in short segments that do not capture the complete emotional expression, leading to issues such as lack of emotional features, increased noise, and inadequate representation of labeled emotions. In contrast, the use of complete data may be more appropriate in certain cases to ensure a more comprehensive representation of the emotional state.

Considering these limitations and challenges, it is important to carefully design and implement data augmentation techniques, such as data segmentation, while considering issues such as data crossover and the duration of emotional expression in EEG signals. This is crucial to ensure accurate and reliable emotion recognition outcomes.

Additionally, in emotion recognition, individual differences among subjects and the accompanying emotional arousal often result in imbalanced sample data, both in terms of inter-class and intra-class imbalances [8]. Inter-class imbalance refers to the uneven distribution of samples among different classes in the dataset, while intra-class imbalance refers to the imbalance in the density distribution within a single class. Both types of imbalances can lead to low classification accuracy, particularly in terms of *F1* score.

Deep learning is currently widely used in multiple fields. Jo et al. [9] studied the application of deep learning in Alzheimer's disease for diagnostic classification and prognostic prediction using neuroimaging data. They performed a systematic review of publications that utilized deep learning approaches and neuroimaging data for the diagnostic classification of Alzheimer's disease. Hua et al. [10] reveal that the latest research in deep learning and reinforcement learning has paved the way for highly complex tasks to be performed by robots. Deep reinforcement learning, imitation learning, and transfer learning in robot control are discussed in detail. Zhou et al. [11] describe the structure of some popular architectures of deep neural networks and the approaches for training a model. The encouraging results in classification and regression problems achieved by deep learning will attract more research efforts to apply deep learning into the field of emotion in the future.

Although existing generative adversarial networks (GANs), including WGAN-GP, can address inter-class imbalance by generating different numbers of samples for each class, they do not consider intra-class imbalance, as noted in the state-of-the-art research. In fact, data augmentation with GANs can exacerbate intra-class imbalance if not properly managed. This is because GANs may mistakenly perceive sparse samples within a class, which are distant from the majority samples or have low sample density, as noise. As a result, GANs may focus on generating data that resembles the majority subclass, leading to a further imbalanced distribution of the class. Moreover, classes with fewer samples are more prone to intra-class imbalance compared to classes with larger sample sizes.

Unfortunately, to mitigate inter-class imbalance, generating more data for the minority class is necessary, which requires addressing intra-class imbalance during data augmentation using GANs.

To tackle the above-mentioned problem, this study proposes the following contributions:

- 1) A targeted data generation method is proposed, based on which a new loss function is established. It can measure the intra-class distribution of newly generated data (by GANs) relative to the original data.
- 2) Developing the framework of T-WGAN-GP and T-GANSER, which enables targeted generation of data within the minority class while considering the overall distribution, thus addressing intra-class imbalance, and improving classification accuracy.
- 3) Verifying the rationality of enhancing data for emotion recognition by data segmentation with different manners.

The rest of this paper is organized as follows: Section 2 provides a summary of related studies. Section 3 presents the proposed targeting method and its details, including the training procedure of T-WGAN-GP and T-GANSER. The section also proposes a crossover rate of the dataset to verify the degree of crossover in different data segmentation methods. Section 4 presents simulations, results, and analysis. Finally, conclusions are drawn in Section 5.

2. Related Research

In recent years, a growing body of research has focused on developing methods for enhancing small sample data, with various approaches proposed to address this challenge in emotion recognition. Emotion recognition involves combining emotion induction methods, feature extraction techniques, and classification methods to analyze physiological signals, providing a foundation for the development of emotion analysis and human-computer interaction. Commonly used classification methods for emotion recognition fall into two categories: (1) methods that extract features from different datasets, manually select features from among many, and then classify emotions based on the selected features; and (2) methods that directly classify the data or automatically extract features. However, both modeling methods require an adequate amount of data to ensure the development of a sound emotion classification model. Training with a small number of samples can lead to overfitting, causing the model to have poor generalization ability and lower accuracy on the testing set. Since physiological signals obtained in experiments can be subject to individual differences and signal instability, they are usually small sample signals, prompting researchers to propose various methods to improve accuracy in such cases.

Several methods have been proposed to address the challenge of small sample size. For example, Piho and Tjahjadi [12] introduced a window selection method based on mutual information to select an appropriate signal window to reduce the length of the signals. Nandi et al. [13] proposed a real-time emotion classification system based on logistic regression trained in an online fashion using the Stochastic Gradient Descent algorithm. Demir et al. [14] proposed five deep convolutional neural network models: AlexNet, VGG16, ResNet50, SqueezeNet, and MobilNetv2. Chao et al. [15] proposed a deep learning framework based on a multiband feature matrix (MFM) and a capsule network to combine the frequency domain, spatial characteristics, and frequency band characteristics of the multi-channel EEG signals to construct the MFM. Pereira et al. [16] proposed an experimental evaluation of three different EEG datasets (DEAP, MAHNOB, and STEED) characterized by short, intermediate, and long signal (or stimulus) durations. Chao and Liu [17] proposed a deep belief-conditional random

field framework that integrates the improved deep belief networks with glia chains and conditional random field. Tao et al. [18] proposed an attention-based convolutional recurrent neural network to extract more discriminative features from EEG signals and improve the accuracy of emotion recognition. Liu et al. [19] proposed an electroencephalogram (EEG) emotion recognition method based on a hybrid feature extraction method in empirical mode decomposition domain combining with optimal feature selection based on sequence backward selection. Mert and Akan [20] investigated the feasibility of using time-frequency representation of EEG signals for emotional state recognition. Chen et al. [21] conducted research on emotion recognition, with a focus on the use of classifiers, achieving 67.89% and 69.06% accuracy. Piho and Tjahjadi [12] paid more attention to the feature extraction stage. Du et al. [22] proposed different experimental ideas, such as using one person's experimental data for training and another person's data for testing. Hu et al. [23] took the experimental data of 22 people from 32 people in the DEAP dataset as samples. Gupta et al. [24] aimed to investigate the channel-specific nature of EEG signals comprehensively and provide an effective method based on flexible analytic wavelet transform for the recognition of emotion.

The studies cited in the above references focus on different aspects of emotion recognition, including feature extraction, classification methods, and the use of various signal types. However, they do not address the problem of small sample size, which can result in overfitting and poor generalization ability.

Insufficient data collection is a widespread issue that researchers often face. To overcome this challenge, data augmentation methods have been proposed. In image classification, data augmentation is commonly used to enhance accuracy, especially when the dataset is limited. Geometric transformation techniques such as inversion, rotation, shearing, scaling, reduction, translation, and jitter are widely employed to increase dataset diversity. Additionally, pixel conversion techniques such as adding salt and pepper noise, Gaussian noise, applying Gaussian blur, and adjusting the Hue, Saturation, Value contrast, brightness, saturation, histogram equalization, and white balance are commonly used. Another method involves generating new samples by using multiple existing samples to address the issue of sample imbalance.

In the field of emotion recognition, geometric data segmentation methods are often used to increase the number of samples. For example, Li et al. [25] segmented data from the DEAP dataset in 2018, dividing 4-second segments into individual samples while paying attention to the relationship between frequency bands and channels. Yin et al. [26] divided their data into one 20-second sample and achieved 67.97% and 65.10% recognition accuracy in two classifications. Similarly, in a 2020 study, Zhong et al. [27] segmented their data into 1-second samples and achieved accuracy rates of 83.09% and 79.77%. In another study, Sharma et al. [28] used a different approach by taking the last 30 s of one-minute data as the sample, which also led to good emotion recognition accuracy. Islam et al. [29] adopted a 3-second segment treatment while focusing on the connection between channels. He et al. [30] used a method in their 2022 study of dividing every 9 s into a section and overlapping every 8-second window. Zheng et al. [31] focused on the data processing method of data segmentation when using the DEAP dataset for four-category classification of emotions. They segmented the original sample into 1-second samples and achieved 69.67% recognition accuracy. Shen et al. [32] used the 1-second segment processing method and studied different combinations of four frequency bands in their 2021 study, achieving 74.22% accuracy.

Paper by Li et al. [25], Yin et al. [26], Zhong et al. [27], Sharma et al. [28], Islam et al. [29], and He et al. [30] applied different segmentation processes to the DEAP dataset to increase the amount of data, which assist their research methods to improve the accuracy of emotion recognition. Meanwhile, literature by Zheng et al. [31] and Shen et al. [32] focuses on the effect of data augmentation on the four categories of the DEAP dataset under different segmentation conditions, as well as some auxiliary research. While these studies use data segmentation to increase the amount of data, the segmentation-based enhancement method and data volume increase are still limited by the original data, without much control. Additionally, segmentation may exacerbate the data type imbalance.

The increasing demand for high-quality data has presented various challenges, prompting the development of sophisticated neural network-based data enhancement methods. Notably, the advent of deep neural networks and GANs has revolutionized the field of data enhancement. GANs have demonstrated impressive capabilities in generating near-real images, videos, and sounds through a game-like training approach between a generation model and a discriminant model. The successful application of GANs in generating high-quality data has sparked rapid development and numerous improved methods in this area.

Inspired by GANs, several researchers have proposed new methods for enhancing data quality and addressing issues such as data imbalance. For instance, Luo and Lu [33] introduced the Conditional Wasserstein GAN (CWGAN) to enhance EEG data and improve the accuracy of emotion recognition based on EEG. Shao et al. [34] and Waheed et al. [35] developed a framework based on the auxiliary classifier GAN (ACGAN) to learn the characteristics of sample data and address the lack of data types. Yan et al. [36] proposed a CWGAN with gradient penalty based on a variational self-encoder to address data imbalance. Jin et al. [37] proposed a new GAN discriminator network structure based on WGAN-GP and ACGAN for data enhancement. In Luo and Lu [33], Shao et al. [34], Waheed et al. [35], Yan et al. [36], and Jin et al. [37], the use of GANs enables the arbitrary control of the number of generated data, but the lack of original data may limit the performance of the network model.

Dong and Ren [38] explored various segmentation methods to improve the generative network's ability to produce data samples with PSD features in their four-category classification study. Luo et al. [39] employed a 1 s segmentation approach to preprocess the data and then trained a GAN to generate high-quality data. Zhang et al. [40] also used a 1 s segmentation method to train the GANSER network model to generate high-quality data and enhance the classifier's performance.

In Dong and Ren [38], Luo et al. [39], and Zhang et al. [40], the two enhancement methods were combined by first segmenting the data and then enhancing it using a GAN. This improved the sample's basic data volume to meet data quantity and quality requirements and enhanced emotion recognition accuracy. However, this combined method did not separate the training set and the test set during data segmentation. As previously discussed, this approach is logically flawed (as will be demonstrated in the fourth part of this article) and does not address data imbalance issues.

Due to the high cost of data collection, most EEG datasets contain only a limited amount of data, leading to unbalanced sample categories. To address this issue, new and efficient oversampling methods for unbalanced data synthesis have been proposed. For example, Zheng et al. [41] proposed the Conditional Wasserstein GAN-Gradient Penalty method, which adds auxiliary condition information to the model. Additionally, Pan and Zheng [42] proposed a new sample generation method using GANs to address the problem of EEG sample shortage and sample category

imbalance. These studies explored the performance of emotion recognition with frequency band correlation and frequency band separation computational models before and after data augmentation on standard EEG-based emotion datasets, thereby solving the problem of imbalance between categories. However, these studies did not consider the imbalance within categories.

To address this issue, we propose a new data enhancement framework that focuses on intra-class data imbalance. By expanding the data within each category, we can achieve a balance between categories and improve the accuracy of emotion recognition. Among various data enhancement methods, the suitability of data segmentation methods for emotion recognition remains unexplored. The data of the training set and the testing set should avoid crossover, and we plan to verify the rationality of different data segmentation methods through experiments. This will provide experimental reference for subsequent related studies.

3. Research Content

To address the issue of intra-class imbalance, we propose a targeted data generation method. And we propose new data augmentation frameworks, T-WGAN-GP and T-GANSER, by combining targeted data generation with WGAN-GP [43] and GANSER [40], respectively. Three distinct experimental modes for data segmentation have been designed, along with the introduction of a dataset cross-rate metric. The rationality of different data segmentation methods is analyzed through the experimental results of each mode.

3.1. Targeting method

Imbalances may occur between emotion categories, where the amount of data differs, and within categories, where the distribution of data varies greatly. Such imbalances can lead to biased emotion recognition results and decreased accuracy. This approach involves training the generator in the GAN to generate high-quality data around specific targets, effectively expanding the unbalanced data within each category to achieve balance and improve the accuracy of emotion recognition. Target points refer to sample points with a lower distribution of surrounding data.

3.2. T-WGAN-GP and T-GANSER

The frameworks aim to address the problem of data imbalance in EEG datasets. T-WGAN-GP and T-GANSER generate high-quality data around the calculated relative position area or target to address this issue to achieve a balance in the data.

T-WGAN-GP and T-GANSER introduce a targeted loss function, Target-Loss, in addition to the Original-Loss which is calculated from Equation (1). The loss functions of WGAN-GP and GANSER share similar components, with the Equation (1) given as:

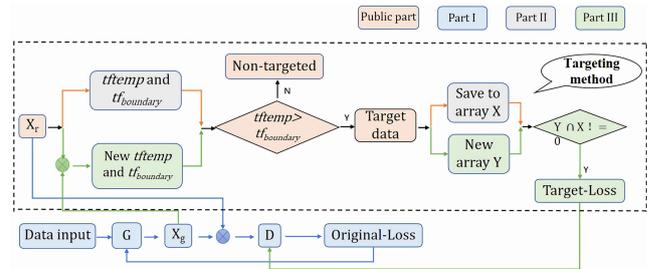
$$L = E_{e \sim P_e} [D(G(\delta(e, \tau)))] - E_{e \sim P_e} [D(\delta(e, \tau))] + \lambda_p E_{\hat{e} \sim P_{\hat{e}}} [||\nabla D(\hat{e})||_2 - 1]^2 \tag{1}$$

where $P_{\hat{e}}$ is defined sampling uniformly along straight lines between pairs of points sampled from the data distribution P_e and the generator distribution among $G(\delta(e, \tau))$. The gradient of the discriminator D is denoted by $\nabla D(\hat{e})$, and λ_p is a hyperparameter presenting the weight of the penalty term.

The training process for T-WGAN-GP and T-GANSER is illustrated in Figure 1. First, as shown in part I of the diagram, generated data are obtained from the countermeasure network, and the Original-Loss function is calculated through the network

model to update the discriminator's parameters and then update the generator's parameters based on the output of the discriminator. The training process in part I follows the training methods used by the WGAN-GP and GANSER models.

Figure 1
Flow chart of T-WGAN-GP and T-GANSER enhancement framework



Then in part II, firstly, the target factor ($tftemp$) and critical value ($tfboundary$) are calculated for real data, then the target factor of each real data is compared with the critical value, and the data whose target factor is greater than the critical value are saved in a new array (Xa) for subsequent calculation.

Finally, in part III, the generative model is employed initially to generate pseudo-data, which is then merged with real data to construct a new dataset for training. Then, the target factor is calculated again for each sample in the dataset at this point, and a new critical value is selected. Select the data whose target factor is greater than the critical value by comparison, compare the selected data with the data saved in the part II, and select the duplicated data, at this time, the selected data are the final target point, and use the target point data to calculate the Target-Loss for the update of the model parameters. The part II and part III are the training process of the targeting method in the new model.

Among them, the specific calculation flow of target detection method and target loss function is shown in Table 1. Targets are selected by calculating target factors and thresholds for real and generated samples in different computational orders followed by calculating the target loss optimization model.

The calculation of the target factor ($TFtemp$) in the process of Table 1 is shown in Figure 2, where X_i belongs to c_x , $dist(a, b)$ is the Euclidean Distance between a and b and $Center_{c_i}$ is the center of subclass c_i , where a and b represent the location of the sample. In Figure 2, the data are first clustered into k clusters and then sorted based on the number of elements in each cluster. When dividing into small and large clusters, a threshold is set, with a threshold value of 0.9 in this study. After the division is completed, the target factors are calculated sequentially according to the algorithm shown in the figure.

3.3. Model framework and parameter design

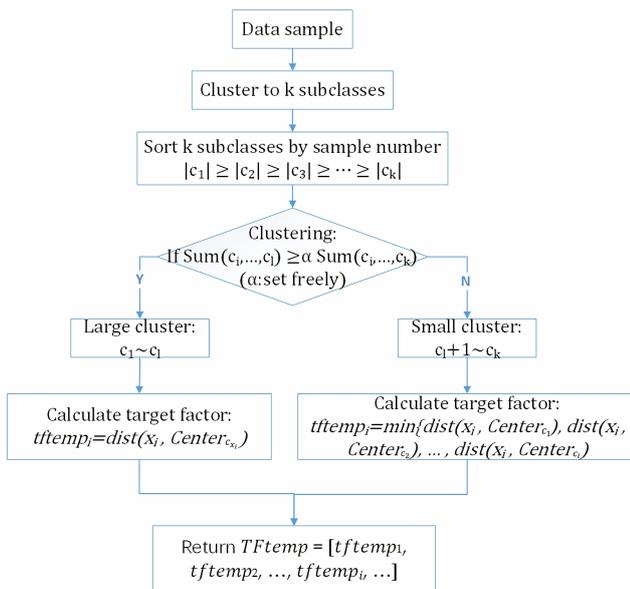
T-WGAN-GP framework is shown in Figure 3. The input of the generator consists of 128 random noise data, which is then passed through a fully connected layer with 1024 neurons to expand the data to 1024, and then reorganized into a three-dimensional tensor (256,2,2). It utilizes the LeakyReLU activation function and finally generates a tensor of size (32,4,1) through 3 transposed convolutional layers. The first transpose convolutional layer has 256 convolutional kernels with a stride of (4,2), the second layer has 128 convolutional kernels with a stride of (2,1), and the third

Table 1
Target detection and target loss calculation process

Input: $Xr = [xr1, xr2, \dots, xrn]$, the original data.

Let $X = Xr$, $X = [x1, x2, \dots, xi, \dots]$
 ===== Target factor computation, TFC =====
 Compute target factor of $X = [x1, x2, \dots, xi, \dots]$, which is denoted as $TfTemp = [tftemp1, tftemp2, \dots, tftempi, \dots]$, where $tftempi$ is the target factor of xi (Figure 2)
 Sort $tftempi$ from min to max store it in a new variable $Tf = [tf1, tf2, \dots, tfp]$, $tf1 \leq tf2 \leq \dots \leq tfp$. Where p is the dimension of X (it equals to n currently).
 Set boundary = Round $(0.8 * p)$ where Round () is a rounding function.
 =====
 For all the xi in X , if $tftempi \geq tfboundary$, put xi into Xa . ($tfboundary$: criticality factor, $Xa = []$)
 Generate new samples with generator using random noise z , $Xg \leftarrow G\theta(z)$
 Let $X = Xr \cup Xg$. Do Target factor computation (Figure 2), which is the same as the code between the “====” lines. Then we have $TfTemp = [tftemp1, tftemp2, \dots, tftempi, \dots, tftempn+1]$ and a new $tfboundary$.
 Set Target_Losstemp = 0, For all the items in X
 If $xi \in Xa$ and $tftempi > tfboundary$
 Target_Losstemp = Target_Losstemp + $(tftempi - tfboundary) * 2$
 End if
 Target_Loss = $\sqrt{\text{Target_Losstemp} / n}$

Figure 2
Target factor calculation process



layer has 64 convolutional kernels with a stride of (2,1). The size of the convolutional kernels is all (3,3), and LeakyReLU is used as the activation function for the first two layers, while Tanh is used as the activation function for the third layer.

In the discriminator D , there are a total of 3 convolutional layers with kernel sizes of (3,3), and strides of (1,1), with 64, 128, and 256 convolutional kernels respectively. LeakyReLU is used as the activation function for all layers, and the last layer is a fully connected layer with 1 node and no activation function. During training, Adam optimizer is used with learning rates of 0.0001 for the discriminator and 0.0002 for the generator.

The classifier adopts support vector machine (SVM) with a Gaussian kernel function, where the hyperparameters $g = 1$ and $c = 1$.

T-GANSER framework and usage parameters are the same as GANSER [40].

In T-WGAN-GP, random noise is input into a network model constructed using multiple layers of convolutional layers. EEG samples are obtained by feeding the multi-layer convolution with a predefined dimensional vector; In T-GANSER, the data are first subjected to a transformation function to remove some signal values. Then, it is input into the model where multiple layers of convolution are employed to fill in the missing signals, resulting in complete data and obtaining EEG samples.

3.4. Experimental design of different data segmentation methods

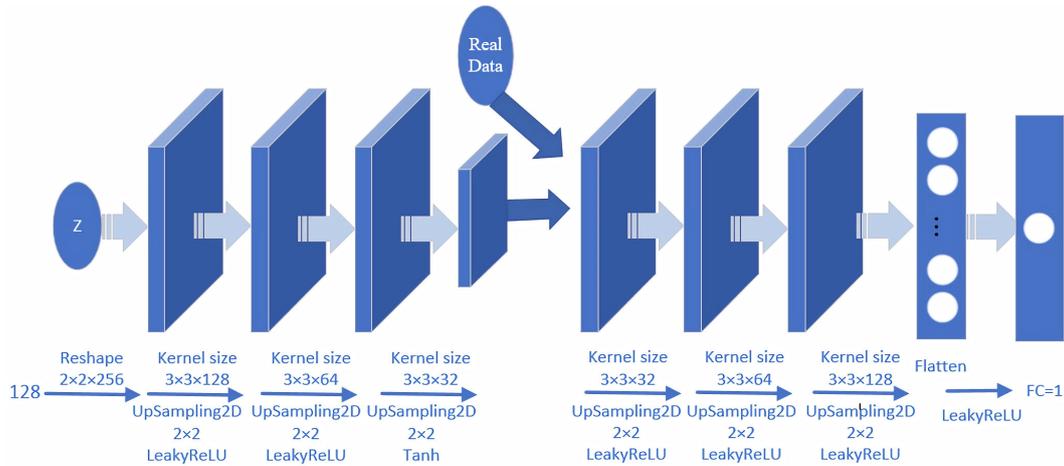
One of the commonly used methods for emotion recognition is to use data segmentation for data augmentation. Data segmentation divides the original sample into multiple samples, and the segmented data from the same sample often have similar features. Therefore, when the segmented data from the same sample exist in the training and testing sets, it is easier to identify cross-sample data during testing, which can improve recognition accuracy. However, in real-world scenarios, when applying the trained network for emotion recognition, the EEG physiological signals of one or more individuals collected outside the dataset should be identified. These signals and the data used in our training network should not crossover. Both theory and practical considerations show that the rationality of data crossover between the training and testing sets should be low.

Additionally, in signal data capturing emotional expression, there is a necessity for a certain duration for emotions to manifest fully, and this timeframe is not always determinable. Consequently, after segmenting the data, short-duration segments may lack emotional characteristics, contain more noise, and fail to accurately represent the corresponding emotional labels. In contrast, utilizing the complete dataset is more justifiable.

Based on the theoretical analysis above, this article designs three experimental modes:

- 1) MI: segmenting the data before the division of the training and testing sets

Figure 3
T-WGAN-GP model framework



Each sample has an experimental duration of 60 s, and it is segmented based on a 1-second interval. Following segmentation, the data are divided into training and testing sets.

2) MII: segmenting the data after the division of the training and testing sets

The entire dataset is firstly divided into training and testing sets. Subsequently, the data samples in the training and testing sets are further partitioned separately.

3) MIII: no data segmentation

Perform experiments utilizing the complete dataset, where involves treating the entire 60-second duration as a single sample for analysis and model training.

3.5. Crossover rate of dataset

To assess the influence of data crossover on the accuracy of emotion recognition, we introduce the concept of the dataset crossover rate.

This rate is calculated to determine the degree of crossover between data in the training and testing sets after the data is segmented. The crossover rate allows us to accurately measure the degree of crossover and evaluate its impact on recognition accuracy.

3.5.1. Calculation of crossover rate for single sample

To calculate the crossover rate for a single sample, we use the following steps:

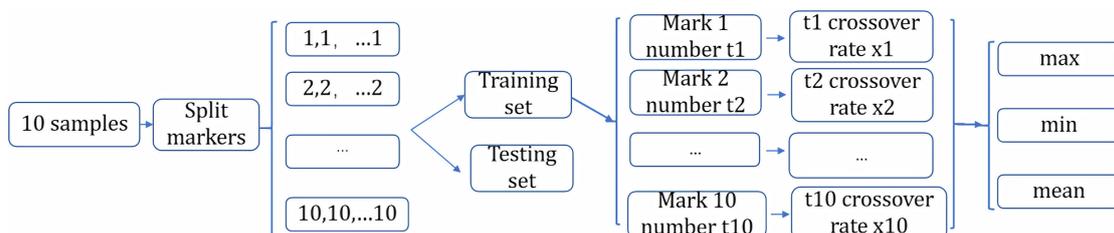
- 1) Count the number of data that are divided into the training set after the same sample is segmented into multiple data points. We denote this as 'n'.
- 2) Count the total number of data points that the sample is segmented. We denote this as 'm'.
- 3) The crossover rate can be calculated as $C = n/m$.

We believe that having more labeled data in the training set leads to better model training and makes it easier to identify data segmented from the same sample in the testing set. Therefore, a higher value of n/m indicates a greater degree of crossover.

3.5.2. Determination of crossover rate for multiple samples

When multiple samples are segmented for labeling, each sample has a calculated crossover rate. For example, in Figure 4, 10 samples are divided into the training and testing sets using different labels (e.g., labels 1, 2, ..., 10), and the crossover rate is calculated for each sample. Each sample thus has its own crossover rate value. To comprehensively represent the crossover situation of multiple sets of data experiments, we select the minimum, maximum, and average values of all sample values as the crossover rate. The maximum value indicates the deepest degree of data crossover, while the minimum value indicates the smallest degree of crossover. By selecting these values, we can obtain a comprehensive view of the crossover situation across multiple sets of data experiments. Formula (2) shows the calculation of DC, which is the crossover rate of multiple sample data sets. Here, j represents the number of

Figure 4
Multiple sample crossover rate selection example diagram



selected samples, while C_{Max} , C_{Min} , C_{Mean} represent the maximum, minimum, and average values calculated across multiple samples, respectively.

$$DC = \sum_{i=1}^j (C_{Max}, C_{Min}, C_{Mean}) \quad (2)$$

4. Simulation and Results Analysis

To validate the effectiveness of T-WGAN-GP, T-GANSER, and the rationality of data segmentation methods, we conducted experiments using the DEAP dataset [44]. Furthermore, in the process of validating the effectiveness of the targeted approach, we conducted ablation experiments to compare the results between models with and without a targeted framework. In the experiments, Python and Pytorch were used for T-WGAN-GP and T-GANSER, respectively.

4.1. DEAP

In this article, we conducted a study on four categories based on different valence and arousal levels: High Valence High Arousal, High Valence Low Arousal, Low Valence High Arousal, and Low Valence Low Arousal (LVLA).

In our experiments, we exclusively used EEG data from the DEAP dataset for two primary reasons:

- 1) The DEAP dataset is widely used by researchers for emotion recognition tasks involving four classification, making it a commonly employed and comprehensive dataset.
- 2) The authors who proposed GANSER, which we used, have previously employed dataset in their literature. This facilitates the comparison of our results with theirs, thereby demonstrating the effectiveness of our method.

4.2. The experimental results of T-WGAN-GP and T-GANSER

4.2.1. DEAP dataset by T-WGAN-GP

1) Data preprocessing

In this section, experiments are conducted using complete data for experimental data processing. To capture pertinent information for affective computing, the research by Zheng and Lu [45] suggests that the differential entropy (DE) of EEG signals is an effective feature. Therefore, instead of using the original signal, the extracted DE features are employed for data augmentation after segmentation. We will conduct 60% of the samples as the training set, 20% of the samples were used as the validating set, and another 20% as the testing set.

2) Experimental results

Figures 5 and 6 illustrate the density distribution gradient plots before and after data enhancement, respectively. Figure 7 illustrates the fitting performance among different models. The results of the evaluations are presented in Table 2.

3) Results analysis

Figure 5 depicts the original sample distribution and density distribution gradient after clustering in LVLA, which has the smallest number among four classes in training set. The sample distribution density is highly uneven and imbalanced within the class. The density distribution gradient is obtained when calculating

Figure 5
The density distribution gradient diagram

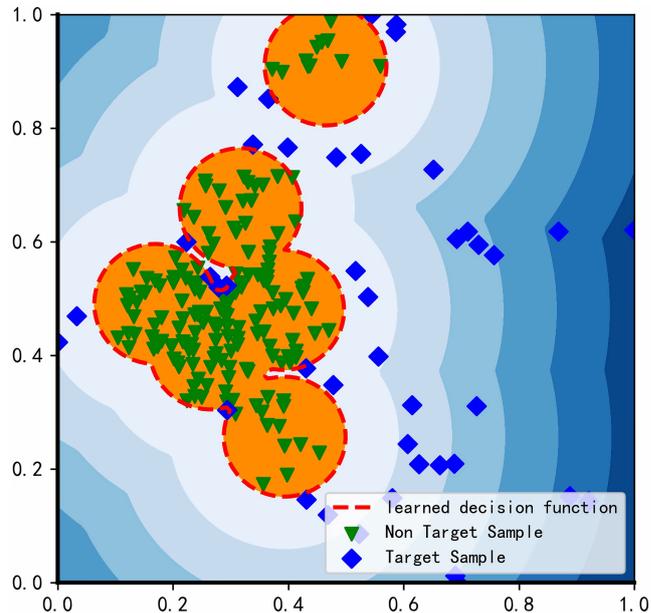
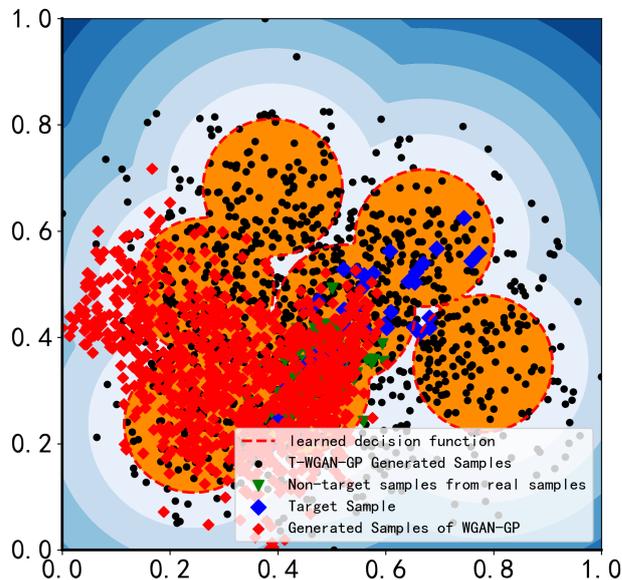


Figure 6
The density distribution gradient diagram after data augmentation



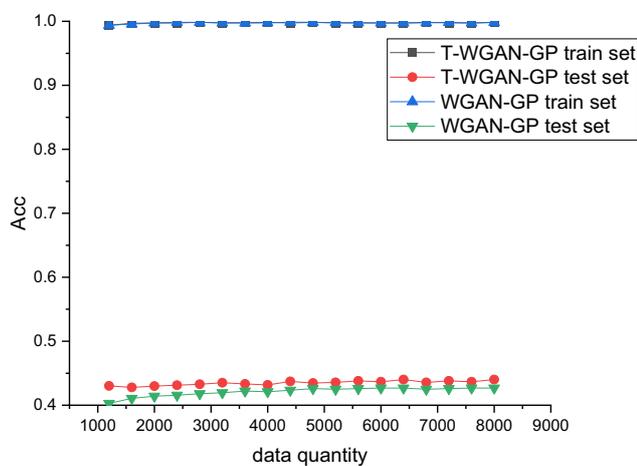
target factor, where blue dots represent detected target samples, and green dots represent remaining samples except for target samples.

In Figure 6, the blue points represent the detected target samples of the original data, the green points represent the remaining samples of the original data, the black points represent the samples generated by T-WGAN-GP, and the red points represent the samples generated by WGAN-GP. It can be clearly seen that T-WGAN-GP generates more data points around the blue points, which correspond to the target samples of the original data. In contrast, WGAN-GP generates fewer data points around the blue points and focuses more on generating data in other areas. This implies that

Table 2
The experimental results of WGAN-GP and T-WGAN-GP with SVM

Sample quantity		1200	1600	2000	2400	2800	3200	3600	4000	4400
WGAN-GP	Acc	0.4029	0.4107	0.4141	0.4156	0.4180	0.4195	0.4220	0.4210	0.4234
	F1	0.3360	0.3484	0.3532	0.3561	0.3588	0.3619	0.3652	0.3642	0.3675
T-WGAN-GP	Acc	0.4302	0.4278	0.4298	0.4312	0.4327	0.4351	0.4332	0.4317	0.4371
	F1	0.3975	0.3951	0.3968	0.3979	0.4002	0.4046	0.4010	0.3994	0.4057
WGAN-GP	Acc	0.4259	0.4249	0.4259	0.4268	0.4268	0.4249	0.4259	0.4268	0.4268
	F1	0.3712	0.3706	0.3712	0.3727	0.3739	0.3781	0.3782	0.3806	0.3775
T-WGAN-GP	Acc	0.4346	0.4356	0.4380	0.4366	0.4400	0.4356	0.4380	0.4366	0.4400
	F1	0.4033	0.4026	0.4079	0.4058	0.4088	0.4066	0.4101	0.4087	0.4077

Figure 7
The fitting performance of the WGAN-GP and T-WGAN-GP



T-WGAN-GP is more effective in generating data that closely resembles the target samples in the original dataset.

In Figure 7, the identification accuracy of the training and testing sets with different amounts of data increase is depicted for two types of models. The black and red lines represent the T-WGAN-GP model, while the blue and green lines represent the WGAN-GP model. The accuracy of the training sets for both models is quite similar, with some overlap between them. However, the analysis of the results reveals that there is a significant difference between the accuracy of the training and testing sets for both models, indicating overfitting, which requires further investigation. Notably, the performance of the T-WGAN-GP model appears to be superior to that of the WGAN-GP model in this regard.

Table 2 displays the results of the ablation experiment. The accuracy and F1 scores are used as evaluation metrics. The F1 is selected because it is an important indicator for imbalanced data classification. ‘‘Sample quantity’’ refers to the number of samples in the training set for each experiment. In the initial step, the number of samples from each class in the training set was adjusted to 300, ensuring that there was no significant inter-class imbalance. This resulted in a total of 1200 samples. Subsequently, $N = 100k$ ($k = 1, 2, 3, \dots$) additional samples were added to each class, gradually increasing the size of the training set, as shown in Table 2 (1600, 2000, ..., 8000).

Table 2 displays the accuracy (Acc) and F1 values acquired through the utilization of various data augmentation frameworks, namely WGAN-GP and T-WGAN-GP with SVM. The table

records the experimental results for each trial. From the table, it can be observed that the experimental results of the T-WGAN-GP model consistently outperform those of the WGAN-GP model. Specifically, at data increments of 6400 and 7200, the Acc and F1 score reach their peak values, at 44% and 41.01%, respectively.

In summary, T-WGAN-GP is effective in generating new samples with a better distribution for emotion classification. It performs better than WGAN-GP in terms of accuracy and the F1 value on the DEAP dataset.

4.2.2. DEAP dataset by T-GANSER

1) Data handling

In this experimental section, the data processing methods are the same as those used in the GASNER model. In this experiment, to segment the data, a non-overlapping sliding window of size 128 is used to separate the trial data into one-second-long chunks, resulting in a total of 2400 EEG samples from 40 trials. 80% of the samples are randomly selected as the training set, and 20% are used as the testing set.

2) Experimental result

Figure 8 depicts the density distribution gradient plot after data augmentation. Figure 9 illustrates the fitting performance between the two models. Table 3 presents the results of the experiment conducted on the preprocessed and segmented DEAP dataset.

3) Result analysis

After applying data augmentation models such as T-GANSER and GANSER, all samples, including the original data and the newly generated samples, were plotted together in a density distribution gradient diagram, as depicted in Figure 8.

In the diagram, the blue points represent the detected target samples from the original data, while the green points correspond to the remaining samples from the original data excluding the target samples. The black points indicate the samples generated by T-GANSER, whereas the red points represent the samples generated by GANSER. By observing the diagram, it can be noted that T-GANSER generated a larger number of data points around the blue points, which correspond to the target samples from the original data.

In Figure 9, the identification accuracy of the training and testing sets after 300 iterations is depicted for two models, T-GANSER and GANSER. The black and red lines represent the T-GANSER model, while the blue and green lines represent the GANSER model. The accuracy of the training sets for both models is quite similar, with some overlap between them. From

Figure 8
The density distribution gradient diagram after data augmentation

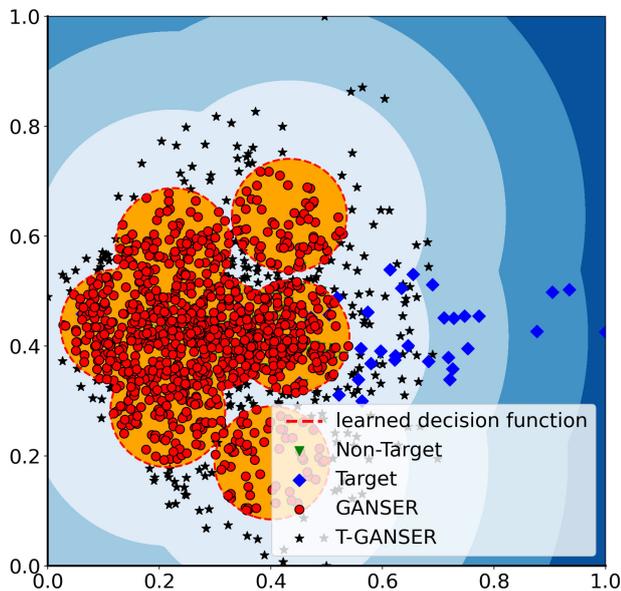
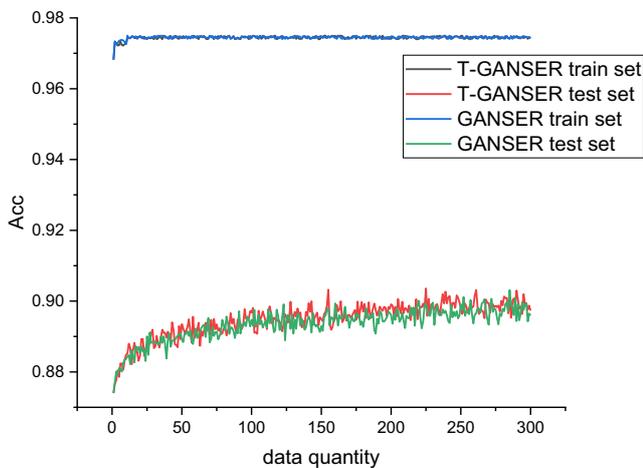


Figure 9
The fitting performance of the GANSER and T-GANSER



the analysis of the results, it can be observed that there is not a significant difference between the accuracy of the training and testing sets for both models, indicating a relatively small degree of overfitting. Particularly, the performance of the T-GANSER model appears to be superior to that of the GANSER model.

Table 3 represents the results of the ablation experiment comparing GANSER and T-GANSER. The table compares the results of three different methods: GANSER, T-GANSER 1, and T-GANSER 2. GANSER represents the results obtained from re-examining GANSER with 300 iterations of GAN and classifier. T-GANSER 1 represents the results obtained from T-GANSER with 300 iterations of GAN and classifier. T-GANSER 2 represents the results obtained from introducing model training after extracting DE features from the data and performing target calculation with 200 iterations of GAN and classifier. T-GANSER 1 performs target

calculation on segmented data, while T-GANSER 2 segments data after target calculation.

The data are initially segmented into 1-second segments and subsequently employed for both the training and testing sets. After data partitioning, the size of the training set increases to 61,440. Therefore, assuming Q iterations, a total of $Q \times 61,440$ new data points are generated for training. It is important to note that the number of iterations remains consistent across all experiments in the five-fold cross-validation.

We reproduced the GANSER model with 300 iterations and achieved the same accuracy as reported in the paper. Experiments using the T-GANSER model 1 and 2 in Table 3 show the improved results obtained by the targeted data generation method in the T-GANSER model. It can be observed that the improved T-GANSER model achieved an accuracy of 89.97% and F1 score of 89.88% with 200 iterations, which is slightly higher than the accuracy obtained by the GANSER model. In the cross-validation experiments, some results even exceeded 90%, indicating the effectiveness of the T-GANSER model.

Table 3
The experimental results of GANSER and T-GANSER

Method	Means	Sample quantity	Average
GANSER	re-examination	300*61440	Acc 89.74
			F1 89.65
T-GANSER	1	300*61440	Acc 89.74
			F1 89.65
	2	200*61440	Acc 89.97
			F1 89.88

4.2.3. Summary and analysis of targeted approach experiments

Table 4 compares the results of the current highest accuracy models in different modes and the results of the models before and after combining the target detection methods. Tables 5 and 6 record the training efficiency of the model before and after optimization, respectively.

Table 4
Results comparison

	MI		MIII	
	Acc	F1	Acc	F1
The highest	89.74	89.65	–	–
GANSER	89.74	89.65	–	–
WGAN-GP	–	–	42.68	38.06
Our	89.97	89.88	44	41.01

The results are shown in Table 4, it can be observed that our model has achieved better results under different modes. It is worth noting that no studies of emotion recognition conducted with unsegmented data (explicitly mentioned) can be found in the available literature. Consequently, T-WGAN-GP achieved the best accuracy of four-classification recognition in the related experiments. In segmentation experiments utilizing this dataset, GANSER previously attained the highest accuracy to date. However, our proposed T-GANSER outperformed the performance of GANSER.

The training time of the optimized and unoptimized models with the same number of trainings is recorded in Tables 5 and 6, respectively. From the results, due to the increase of the algorithm, the training time of its model increases accordingly. Subsequent studies can continue to optimize the models to reduce the training time and improve the usefulness.

Table 5
The training time for WGAN-GP and T-WGAN-GP

Models	WGAN-GP	T-WGAN-GP
Epochs	5000	5000
Train times	3663.15 s	3897.26 s

Table 6
The training time for GANSER and T-GANSER

Models	GANSER	T-GANSER
Epochs	200	200
Train times	8 h 35 m 29 s	11 h 48 m 45 s

4.3. Reasonability analysis of different data segmentation methods

In this section, experiments are conducted on the currently most accurate GANSER model using three experimental modes (MI, MII, and MIII) based on different data segmentation methods designed. The experimental effects of other enhancement models are generally like GANSER model and will not be analyzed repeatedly.

1) Data handling

In this experiment, 80% of the samples are randomly selected as the training set, 20% are used as the testing set.

2) Experimental result

Tables 7 and 8 show the dataset crossover rate calculation results for the MI, MII, and MIII processing method, and Table 9 summarizes the GANSER experimental results under different modes.

3) Results analysis

Table 7 shows the results of the intersection rate calculation for the dataset in MI mode. To avoid extreme and contingency results, five experiments were conducted with different amounts of data, and the maximum, minimum, and average values were recorded for each

experiment. The mean value represents the representative value for the crossover rate of the dataset, while the minimum and maximum values indicate different levels of cross-validation. The experimental results indicate that the MI method can lead to data crossover. The higher the value, the higher the degree of dataset crossover.

Table 8 displays the crossover rate calculation results in the MII and MIII experiments. It is noteworthy that the method of first dividing the data into the training and testing sets and then segmenting them and no data segmentation does not involve data crossover.

Table 9 summarizes the recognition accuracy and dataset crossover rate of the GANSER model under three modes. In the MII and MIII modes, the segmentation methods have been adjusted, with one involving segmenting the data after the division of the training and testing sets, and the other involving no data segmentation. In the MIII, the data dimension was down-sampled from 7680 to 1280 after removing the baseline, and the input and output of the model neural network were adjusted accordingly.

Based on the data presented in the table, it is evident that the GANSER model attained a recognition accuracy of 89.74% when the mean dataset crossover rate was 80%. And the recognition rate of GANSER decreases to 31.08% and 25.86% when the crossover rate of dataset is 0.

Based on the analysis of recognition accuracy and crossover rate, it can be concluded that a higher crossover rate generally leads to higher recognition accuracy. However, it is also important to note that a crossover rate of 0 results in a significant drop in recognition accuracy. This highlights the importance of carefully considering the crossover rate when performing recognition tasks. Therefore, it can be concluded that it is the higher crossover rate that leads to higher recognition accuracy, other than GANSER itself.

4.4. Summarize

The experiments, results, and analysis which are presented in this paper can illustrate the effectiveness of the targeted method in handling data intra-class imbalance situations. And for emotional expression and emotion recognition, utilizing more complete physiological signals and dividing the training and the testing sets in advance seems to be a more reasonable approach.

According to the experimental results, T-WGAN-GP achieves the highest accuracy in four-category emotion recognition by using the complete 1,280 samples in the DEAP dataset.

In this paper, we conducted experiments using the DEAP dataset. When using other datasets, a simple processing according to our model and algorithm can be used for direct experiments. Different EEG datasets have insufficient data volume due to acquisition difficulties, which leads to data imbalance. Our method can effectively improve the imbalance.

Table 7
Calculation results of dataset crossover rate (MI)

Sample quantity Sequence number	100			500			1280		
	Max	Min	Mean	Max	Min	Mean	Max	Min	Mean
1	0.68	0.9	0.8	0.65	0.933	0.8	0.62	0.933	0.8
2	0.6	0.917	0.8	0.65	0.95	0.8	0.62	0.95	0.8
3	0.68	0.917	0.8	0.63	0.933	0.8	0.62	0.95	0.8
4	0.62	0.917	0.8	0.65	0.95	0.8	0.63	0.95	0.8
5	0.65	0.933	0.8	0.67	0.933	0.8	0.63	0.95	0.8
Average	0.658	0.9168	0.8	0.65	0.9398	0.8	0.624	0.9466	0.8

Table 8
Calculation results of dataset crossover rate (MII and MIII)

Sample quantity	100/500/1280		
	Max	Min	Mean
Sequence number			
Average	0	0	0

Table 9
The experimental results for the GANSER model
(MI, MII, and MIII)

Modes	Method	Acc	F1	Crossover rate of dataset
MI	GANSER	89.74	89.65	0.8
MII	GANSER	31.08	29.42	0
MIII	GANSER	25.86	23.59	0

5. Conclusions and Future Expectations

5.1. Conclusions

Based on the experimental results, it can be seen that the targeted method can improve the intra-class imbalance of the data. T-WGAN-GP and T-GANSER models are proposed in combination with the targeted method, and the effectiveness of the targeted method is verified using the two models, and the results are better than the model without the inclusion of the targeted method. Additionally, different data segmentation methods using GANSER showed varying recognition accuracies, highlighting the importance of the cross-validation rate. However, avoiding data cross-validation is deemed more reasonable despite the more accurate segmentation method of segmenting before dividing. In T-WGAN-GP, complete data are used to avoid noise, inaccurate expression, and incomplete information. T-WGAN-GP achieves the best accuracy of four-classification recognition using complete data in the current DEAP dataset in our experiment.

5.2. Potential limitations

In this study, we employed appropriate data processing methods and target method to facilitate the model in learning the distribution characteristics of data more comprehensively and reasonably. This approach enables the model to handle highly variable real-world data effectively. Specific applications to real-world methods need to be considered.

Research on targeting methods needs to continue to be optimized. Currently, research on targeting a specific physiological signal faces some challenges when using multiple signals simultaneously for target computation. And using low-quality data as the target to train the model will greatly affect the model performance. In addition, this study utilizes EEG signals, so it is crucial to study the fusion of multiple signals. And the effect of the model on different datasets needs to be further investigated.

5.3. Future work

In response to the limitations that the research in this paper may have encountered, there are several areas that should be focused on in the future.

For application in real-life scenarios, EEG signals and corresponding labels can be obtained from a certain number of subjects through EEG signal acquisition experiments. The use

of our trained model is tested to verify its use in real-world scenarios.

Further investigation is needed to pinpoint the areas where generating fake data is necessary. The target screening methods need to be investigated for low-quality targets that affect model performance. Future research should aim to fuse various physiological signals, such as ECG signals, to improve accuracy. The degree of applicability in real-life scenarios and generalizability across different datasets also needs to be investigated. And studying how to utilize complete data to achieve better identification accuracy is crucial.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available at <https://doi.org/10.1109/T-AFFC.2011.15>, reference number [44]. The DEAP datasets that support the findings of this study are openly available at <http://www.eecs.qmul.ac.uk/mmv/datasets/deap/download.html>.

Author Contribution Statement

Lufeng Yin: Software, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Visualization.
Yong Li: Conceptualization, Methodology, Validation, Writing – review & editing, Supervision, Project administration.

References

- [1] Ezzameli, K., & Mahersia, H. (2023). Emotion recognition from unimodal to multimodal analysis: A review. *Information Fusion*, 99, 101847. <https://doi.org/10.1016/j.inffus.2023.101847>
- [2] Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, 1(1), 18–37. <https://doi.org/10.1109/T-AFFC.2010.1>
- [3] Alarcão, S. M., & Fonseca, M. J. (2019). Emotions recognition using EEG signals: A survey. *IEEE Transactions on Affective Computing*, 10(3), 374–393. <https://doi.org/10.1109/TAFFC.2017.2714671>
- [4] Wang, S. H., Phillips, P., Dong, Z. C., & Zhang, Y. D. (2018). Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm. *Neurocomputing*, 272, 668–676. <https://doi.org/10.1016/j.neucom.2017.08.015>
- [5] Seo, J., Laine, T. H., Oh, G., & Sohn, K. A. (2020). EEG-based emotion classification for Alzheimer's disease patients using conventional machine learning and recurrent neural network models. *Sensors*, 20(24), 7212. <https://doi.org/10.3390/s20247212>
- [6] Belkacem, A. N., Jamil, N., Palmer, J. A., Ouhbi, S., & Chen, C. (2020). Brain computer inter-faces for improving the quality of life of older adults and elderly patients. *Frontiers in Neuroscience*, 14, 692. <https://doi.org/10.3389/fnins.2020.00692>
- [7] Chaturvedi, I., Satapathy, R., Cavallari, S., & Cambria, E. (2019). Fuzzy commonsense reasoning for multimodal sentiment analysis. *Pattern Recognition Letters*, 125, 264–270. <https://doi.org/10.1016/j.patrec.2019.04.024>
- [8] He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263–1284. <https://doi.org/10.1109/TKDE.2008.239>

- [9] Jo, T., Nho, K., & Saykin, A. J. (2019). Deep learning in Alzheimer's disease: Diagnostic classification and prognostic prediction using neuroimaging data. *Frontiers in Aging Neuroscience*, *11*, 220. <https://doi.org/10.3389/fnagi.2019.00220>
- [10] Hua, J., Zeng, L., Li, G., & Ju, Z. (2021). Learning for a robot: Deep reinforcement learning, imitation learning, transfer learning. *Sensors*, *21*(4), 1278. <https://doi.org/10.3390/s21041278>
- [11] Zhou, L., Zhang, C., Liu, F., Qiu, Z., & He, Y. (2019). Application of deep learning in food: A review. *Comprehensive Reviews in Food Science and Food Safety*, *18*(6), 1793–1811. <https://doi.org/10.1111/1541-4337.12492>
- [12] Pihö, L., & Tjahjadi, T. (2020). A mutual information based adaptive windowing of informative EEG for emotion recognition. *IEEE Transactions on Affective Computing*, *11*(4), 722–735. <https://doi.org/10.1109/TAFFC.2018.2840973>
- [13] Nandi, A., Xhafa, F., Subirats, L., & Fort, S. (2021). Real-time emotion classification using EEG data stream in e-learning contexts. *Sensors*, *21*(5), 1589. <https://doi.org/10.3390/s21051589>
- [14] Demir, F., Sobahi, N., Siuly, S., & Sengur, A. (2021). Exploring deep learning features for automatic classification of human emotion using EEG rhythms. *IEEE Sensors Journal*, *21*(13), 14923–14930. <https://doi.org/10.1109/JSEN.2021.3070373>
- [15] Chao, H., Dong, L., Liu, Y., & Lu, B. (2019). Emotion recognition from multiband EEG signals using CapsNet. *Sensors*, *19*(9), 2212. <https://doi.org/10.3390/s19092212>
- [16] Pereira, E. T., Gomes, H. M., Veloso, L. R., & Mota, M. R. A. (2021). Empirical evidence relating EEG signal duration to emotion classification performance. *IEEE Transactions on Affective Computing*, *12*(1), 154–164. <https://doi.org/10.1109/TAFFC.2018.2854168>
- [17] Chao, H., & Liu, Y. (2020). Emotion recognition from multi-channel EEG signals by exploiting the deep belief-conditional random field framework. *IEEE Access*, *8*, 33002–33012. <https://doi.org/10.1109/ACCESS.2020.2974009>
- [18] Tao, W., Li, C., Song, R., Cheng, J., Liu, Y., Wan, F., & Chen, X. (2023). EEG-based emotion recognition via channel-wise attention and self attention. *IEEE Transactions on Affective Computing*, *14*(1), 382–393. <https://doi.org/10.1109/TAFFC.2020.3025777>
- [19] Liu, Z. T., Xie, Q., Wu, M., Cao, W. H., Li, D. Y., & Li, S. H. (2019). Electroencephalogram emotion recognition based on empirical mode decomposition and optimal feature selection. *IEEE Transactions on Cognitive and Developmental Systems*, *11*(4), 517–526. <https://doi.org/10.1109/TCDS.2018.2868121>
- [20] Mert, A., & Akan, A. (2018). Emotion recognition based on time–frequency distribution of EEG signals using multivariate synchrosqueezing transform. *Digital Signal Processing*, *81*, 106–115. <https://doi.org/10.1016/j.dsp.2018.07.003>
- [21] Chen, J., Hu, B., Moore, P., Zhang, X., & Ma, X. (2015). Electroencephalogram-based emotion assessment system using ontology and data mining techniques. *Applied Soft Computing*, *30*, 663–674. <https://doi.org/10.1016/j.asoc.2015.01.007>
- [22] Du, X., Ma, C., Zhang, G., Li, J., Lai, Y. K., Zhao, G., . . . , & Wang, H. (2022). An efficient LSTM network for emotion recognition from multichannel EEG signals. *IEEE Transactions on Affective Computing*, *13*(3), 1528–1540. <https://doi.org/10.1109/TAFFC.2020.3013711>
- [23] Hu, J., Wang, C., Jia, Q., Bu, Q., Sutcliffe, R., & Feng, J. (2021). ScalingNet: Extracting features from raw EEG data for emotion recognition. *Neurocomputing*, *463*, 177–184. <https://doi.org/10.1016/j.neucom.2021.08.018>
- [24] Gupta, V., Chopda, M. D., & Pachori, R. B. (2019). Cross-subject emotion recognition using flexible analytic wavelet transform from EEG signals. *IEEE Sensors Journal*, *19*(6), 2266–2274. <https://doi.org/10.1109/JSEN.2018.2883497>
- [25] Li, M., Xu, H., Liu, X., & Lu, S. (2018). Emotion recognition from multichannel EEG signals using K-nearest neighbor classification. *Technology and Health Care*, *26*, 509–519. <https://doi.org/10.3233/THC-174836>
- [26] Yin, Z., Liu, L., Chen, J., Zhao, B., & Wang, Y. (2020). Locally robust EEG feature selection for individual-independent emotion recognition. *Expert Systems with Applications*, *162*, 113768. <https://doi.org/10.1016/j.eswa.2020.113768>
- [27] Zhong, Q., Zhu, Y., Cai, D., Xiao, L., & Zhang, H. (2020). Electroencephalogram access for emotion recognition based on a deep hybrid network. *Frontiers in Human Neuroscience*, *14*, 589001. <https://doi.org/10.3389/fnhum.2020.589001>
- [28] Sharma, R., Pachori, R. B., & Sircar, P. (2020). Automated emotion recognition based on higher order statistics and deep learning algorithm. *Biomedical Signal Processing and Control*, *58*, 101867. <https://doi.org/10.1016/j.bspc.2020.101867>
- [29] Islam, M. R., Islam, M. M., Rahman, M. M., Mondal, C., Singha, S. K., Ahmad, M., . . . , & Moni, M. A. (2021). EEG channel correlation based model for emotion recognition. *Computers in Biology and Medicine*, *136*, 104757. <https://doi.org/10.1016/j.compbiomed.2021.104757>
- [30] He, Z., Zhong, Y., & Pan, J. (2022). An adversarial discriminative temporal convolutional network for EEG-based cross-domain emotion recognition. *Computers in Biology and Medicine*, *141*, 105048. <https://doi.org/10.1016/j.compbiomed.2021.105048>
- [31] Zheng, W. L., Zhu, J. Y., & Lu, B. L. (2019). Identifying stable patterns over time for emotion recognition from EEG. *IEEE Transactions on Affective Computing*, *10*(3), 417–429. <https://doi.org/10.1109/TAFFC.2017.2712143>
- [32] Shen, F., Peng, Y., Kong, W., & Dai, G. (2021). Multi-scale frequency bands ensemble learning for EEG-based emotion recognition. *Sensors*, *21*(4), 1262. <https://doi.org/10.3390/s21041262>
- [33] Luo, Y., & Lu, B. L. (2018). EEG data augmentation for emotion recognition using a conditional Wasserstein GAN. In *40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2535–2538. <https://doi.org/10.1109/EMBC.2018.8512865>
- [34] Shao, S., Wang, P., & Yan, R. (2019). Generative adversarial networks for data augmentation in machine fault diagnosis. *Computers in Industry*, *106*, 85–93. <https://doi.org/10.1016/j.compind.2019.01.001>
- [35] Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., & Pinheiro, P. R. (2020). CovidGAN: Data augmentation using auxiliary classifier GAN for improved Covid-19 detection. *IEEE Access*, *8*, 91916–91923. <https://doi.org/10.1109/ACCESS.2020.2994762>
- [36] Yan, K., Su, J., Huang, J., & Mo, Y. (2022). Chiller fault diagnosis based on VAE-enabled generative adversarial networks. *IEEE Transactions on Automation Science and Engineering*, *19*(1), 387–395. <https://doi.org/10.1109/TASE.2020.3035620>
- [37] Jin, Q., Lin, R., & Yang, F. (2020). E-WACGAN: Enhanced generative model of signaling data based on WGAN-GP and ACGAN. *IEEE Systems Journal*, *14*(3), 3289–3300. <https://doi.org/10.1109/JSYST.2019.2935457>
- [38] Dong, Y., & Ren, F. (2020). Multi-reservoirs EEG signal feature sensing and recognition method based on generative

- adversarial networks. *Computer Communications*, 164, 177–184. <https://doi.org/10.1016/j.comcom.2020.10.004>
- [39] Luo, Y., Zhu, L. Z., Wan, Z. Y., & Lu, B. L. (2020). Data augmentation for enhancing EEG-based emotion recognition with deep generative models. *Journal of Neural Engineering*, 17(5), 056021. <https://doi.org/10.1088/1741-2552/abb580>
- [40] Zhang, Z., Liu, Y., & Zhong, S. H. (2023). GANSER: A self-supervised data augmentation framework for EEG-based emotion recognition. *IEEE Transactions on Affective Computing*, 14(3), 2048–2063. <https://doi.org/10.1109/TAFFC.2022.3170369>
- [41] Zheng, M., Li, T., Zhu, R., Tang, Y., Tang, M., Lin, L., & Ma, Z. (2020). Conditional Wasserstein generative adversarial network-gradient penalty-based approach to alleviating imbalanced data classification. *Information Sciences*, 512, 1009–1023. <https://doi.org/10.1016/j.ins.2019.10.014>
- [42] Pan, B., & Zheng, W. (2021). Emotion recognition based on EEG using generative adversarial nets and convolutional neural network. *Computational and Mathematical Methods in Medicine*, 2021(1), 2520394. <https://doi.org/10.1155/2021/2520394>
- [43] Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein generative adversarial networks. In *Proceedings of the 34th International Conference on Machine Learning*, 214–223.
- [44] Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., . . . , & Patras, I. (2012). DEAP: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 3(1), 18–31. <https://doi.org/10.1109/T-AFFC.2011.15>
- [45] Zheng, W. L., & Lu, B. L. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development*, 7(3), 162–175. <https://doi.org/10.1109/TAMD.2015.2431497>

How to Cite: Yin, L., & Li, Y. (2024). Data-Segmentation Verification and a Target Generative Adversarial Network: EEG-Based Emotion Recognition. *Journal of Computational and Cognitive Engineering*, 3(4), 421–433. <https://doi.org/10.47852/bonviewJCCE42022571>