



OPEN ACCESS

EDITED BY

Patrick K. A. Neff,
University of Zurich, Switzerland

REVIEWED BY

Venkata Ramana Murthy Oruganti,
Amrita Vishwa Vidyapeetham University, India
Abdul Rehman Javed,
Air University, Pakistan

*CORRESPONDENCE

Tingshao Zhu
✉ tszhu@psych.ac.cn
Shaoshuai Gao
✉ ssgao@ucas.ac.cn

SPECIALTY SECTION

This article was submitted to
Public Mental Health,
a section of the journal
Frontiers in Public Health

RECEIVED 27 October 2022

ACCEPTED 27 February 2023

PUBLISHED 17 March 2023

CITATION

Wen Y, Li B, Liu X, Chen D, Gao S and Zhu T
(2023) Using gait videos to automatically assess
anxiety. *Front. Public Health* 11:1082139.
doi: 10.3389/fpubh.2023.1082139

COPYRIGHT

© 2023 Wen, Li, Liu, Chen, Gao and Zhu. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Using gait videos to automatically assess anxiety

Yeye Wen^{1,2}, Baobin Li³, Xiaoqian Liu², Deyuan Chen¹,
Shaoshuai Gao^{1*} and Tingshao Zhu^{2,4*}

¹School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing, China, ²Institute of Psychology, Chinese Academy of Sciences, Beijing, China, ³School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing, China, ⁴Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

Background: In recent years, the number of people with anxiety disorders has increased worldwide. Methods for identifying anxiety through objective clues are not yet mature, and the reliability and validity of existing modeling methods have not been tested. The objective of this paper is to propose an automatic anxiety assessment model with good reliability and validity.

Methods: This study collected 2D gait videos and Generalized Anxiety Disorder (GAD-7) scale data from 150 participants. We extracted static and dynamic time-domain features and frequency-domain features from the gait videos and used various machine learning approaches to build anxiety assessment models. We evaluated the reliability and validity of the models by comparing the influence of factors such as the frequency-domain feature construction method, training data size, time-frequency features, gender, and odd and even frame data on the model.

Results: The results show that the number of wavelet decomposition layers has a significant impact on the frequency-domain feature modeling, while the size of the gait training data has little impact on the modeling effect. In this study, the time-frequency features contributed to the modeling, with the dynamic features contributing more than the static features. Our model predicts anxiety significantly better in women than in men ($r_{Male} = 0.666$, $r_{Female} = 0.763$, $p < 0.001$). The best correlation coefficient between the model prediction scores and scale scores for all participants is 0.725 ($p < 0.001$). The correlation coefficient between the model prediction scores for odd and even frame data is 0.801~0.883 ($p < 0.001$).

Conclusion: This study shows that anxiety assessment based on 2D gait video modeling is reliable and effective. Moreover, we provide a basis for the development of a real-time, convenient and non-invasive automatic anxiety assessment method.

KEYWORDS

anxiety assessment, mental health, gait video, machine learning, reliability and validity

1. Introduction

The increasing pressure of modern life has led to a decline in global mental health and an increase in anxiety and depression (1). Anxiety disorders are the most common mental health problems worldwide and may cause physiological reactions such as irritability, fatigue, and increased heart rate. A long-term intense anxious state not only affects an individual's social, life, and work responsibilities but also has a serious impact on their physical health (2). Therefore, to improve the mental health of different groups, the demand for mental health services has increased worldwide (3, 4). Fortunately, in recent years, researchers have made new progress in the treatment of mental diseases such as anxiety and depression (5, 6). At the same time, we urgently need to develop a convenient and timely method for assessing anxiety states.

In psychology, the anxiety scale has been carefully designed, revised and tested, and various scale-based assessment methods have been developed (7). Self-reports rely on individuals reporting their symptoms, behaviors, and attitudes (8). At present, self-reports remain the most commonly used and most effective anxiety assessment method (9). However, scale-based assessments have some limitations and are not applicable in some scenarios (10). For example, in scenarios that require multiple measurements, participants completing the same questionnaire multiple times can lead to practice effects (11). In scenarios such as job interviews, scale results may be inaccurate due to social desirability (12). In addition, the self-report method is not suitable for certain populations, such as illiterate or dyslexic individuals. Therefore, we hope to develop more objective indicators to assess anxiety.

Anxiety can affect an individual's physiological responses. Anxious individuals may experience shortness of breath and accelerated heartbeat (2). In addition, fear is a typical symptom of anxiety disorders, and patients may experience muscle tension (13), sweating, trembling (14), and skin conductance and heart rate changes (15). Anxiety-induced fear can also be reflected through facial expressions (16). Giannakakis et al. showed that some specific facial cues, such as eye and mouth movements, are suitable as discriminative indicators of anxiety (17). Anxiety may also be reflected in voice changes. In anxious states, individuals tend to speak quickly at a loud volume (18), showing fewer voice changes and more pauses (19). Gait and anxiety are also related. Gait posture and movement characteristics can indicate a variety of emotions (20, 21). For example, individuals with anxiety tend to pace back and forth (22). Feldman et al. found that compared with healthy people, anxious patients have shorter stride distances and take fewer steps per minute, displaying movement disorders to some extent (23). Other researchers have noted similar characteristics, such as slow gait (24, 25) and balance dysfunction (26, 27). In addition, arm swings, vertical head movements, and lateral upper body swings have also been associated with anxiety (28). Among the various physiological and behavioral characteristics related to anxiety, gait has several advantages, including large variations, non-invasiveness and ease of observation. Thus, gait can serve as an objective indicator for assessing anxiety.

To acquire gait data, some researchers have used body-worn sensors (29), human motion capture systems (30, 31), Kinects (Xbox One Kinect Sensor) (32) and other devices. However, these devices are expensive and complex to operate, which is not conducive to improving the applicability of anxiety assessment methods. In this study, we recorded 2D gait videos using a common camera that is simple to operate, increasing the ease of obtaining data.

In recent years, with the development of machine learning technology, various researchers have used gait to assess anxiety. Jing et al. found that a prediction model based on gait features performed better than a prediction model based on speech features (33). Miao et al. and Zhao et al. established anxiety assessment models, and the correlation coefficients between the anxiety prediction score and the scale score reached 0.4 (34) and 0.51 (35), respectively. Both studies considered the basic statistics of the gait time series data and the amplitude in the frequency domain after a Fourier transform as features. These features are relatively simple, which may increase the make it difficult to express the

rich movement characteristics of gait. In addition, these features lack biological or kinematic interpretations. Stark et al. considered five main gait parameters to identify anxiety, namely, the turning angle, neck variance, lumbar rotation, lumbar movement in the sagittal plane, and arm movement (36). Although the above studies established different anxiety assessment models, they did not comprehensively evaluate the model reliability and validity, and did not adequately validate the performance of their models.

In this study, we used 2D gait videos to construct static and dynamic time-domain features and frequency-domain features and established anxiety prediction models through machine learning algorithms. To validate the proposed models, we examined the effects of different frequency-domain feature construction methods, training data sizes and gender on model performance and compared the contributions of different time-frequency features to the modeling results. In addition, we tested the odd-even split-half reliability of the proposed anxiety assessment model. The goal of this study is to provide a convenient auxiliary anxiety assessment method.

The contributions of this study are as follows:

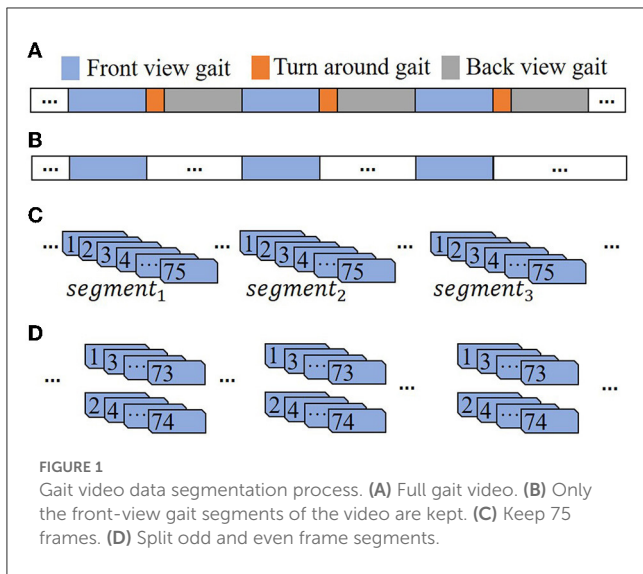
- Build anxiety assessment models using easily accessible 2D gait videos, reducing cost and increasing convenience of anxiety assessment. It was verified that a good anxiety assessment model can be built without using longer gait videos.
- We constructed static and dynamic time-domain features and frequency-domain features with biological kinematic significance, and proved the rationality and necessity of constructing features.
- This study carefully evaluated the performance (validity and reliability) of the anxiety assessment model through experiments. We validated differences in anxiety assessment between men and women, and verified the robustness of our model in a video odd-even split-half test.

The rest of this paper is organized as follows. First, we introduce the research methods and experiments in Section Methods, including the collection and preprocessing of gait data, feature engineering and modeling, and experimental procedures. Then, the results of several comparative experiments are reported in Section Results. A general discussion of the results is given in Section Discussion, explaining the findings of the study and illustrating further work. Finally, concluding remarks is presented in Section Conclusion.

2. Methods

In this study, we used a camera to capture participant gait videos (walking back and forth) indoors. The specific gait video collection method is similar to the method described in Wen et al. (37).

After the gait videos were collected, the participants immediately completed a 7-item Generalized Anxiety Disorder (GAD-7) scale assessment. The GAD-7 assessment is a valid and efficient tool for identifying GAD and assessing its severity in clinical practice and research (9). It evaluates anxiety states in the previous 2 weeks and divides anxiety into four levels according



to the scale scores, namely, minimal anxiety (0–4), mild anxiety (5–9), moderate anxiety (10–14), and severe anxiety (15–21). The GAD-7 assessment shows good internal consistency (Cronbach $\alpha = 0.92$) and test-retest reliability (intraclass correlation = 0.83) (9).

Permission for the above protocol was obtained from the Institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences (Approval number: H15010).

We obtained ~2-min gait videos for each participant, including front and back gaits. Since the front-view gait skeleton evaluation is more accurate than that the back-view evaluation (38), we analyzed skeletons only from the front view to obtain more precise features. Previous studies have shown that good models can be built using a small number of gait frames (35). We kept three consecutive front-view gait segments for each participant, and each segment included 75 frames. To assess the odd-even split-half reliability of the model, we divided the first 74 frames in the gait data into two sets by considering odd and even frames. The gait data segmentation process is shown in Figure 1.

The preprocessing method is similar to the approach proposed in Wen et al. (37). We used OpenPose (39) (a multiperson 2D pose recognition system) to extract the 2D coordinates of 25 body key points from the gait videos and performed coordinate translation (with the MidHip key point as the coordinate origin) and smoothing on the coordinate sequence. Figure 2 shows the 25 human body key points in OpenPose.

The gait coordinate sequence obtained after preprocessing includes only isolated coordinate points and thus does not reflect changes between frames and variations between different key points. We call the features obtained from such data static time-domain features. To reflect the changing gait characteristics (40), we calculate the interframe difference and construct the distances between joints (see Supplementary Table A) and angles between joints (see Supplementary Table B) to express dynamic information. We term these features dynamic time-domain features. The method for obtaining the static and dynamic time-domain features is similar to Wen et al. (37). Figure 3 shows a diagram of the interframe difference between f_{j-1} , f_j , and f_{j+1} in

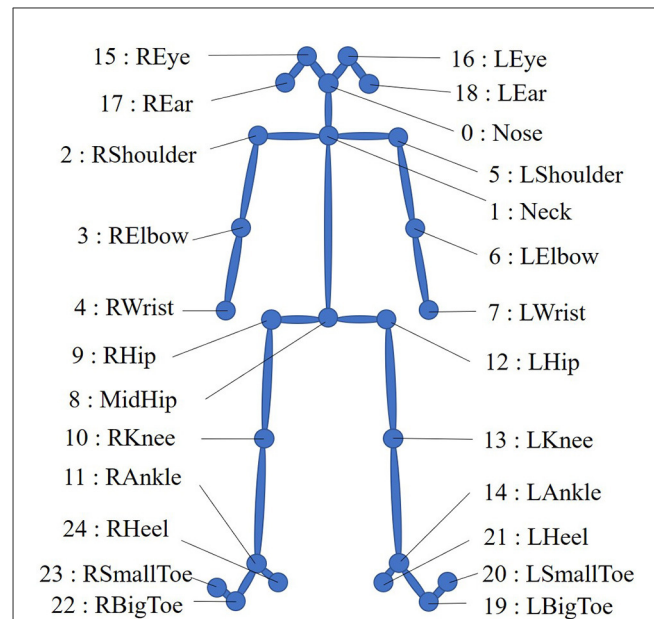


FIGURE 2
Twenty-five human body key points in OpenPose.

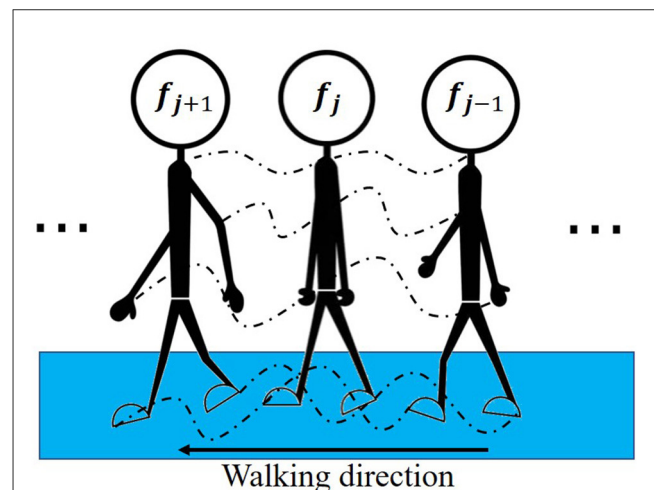
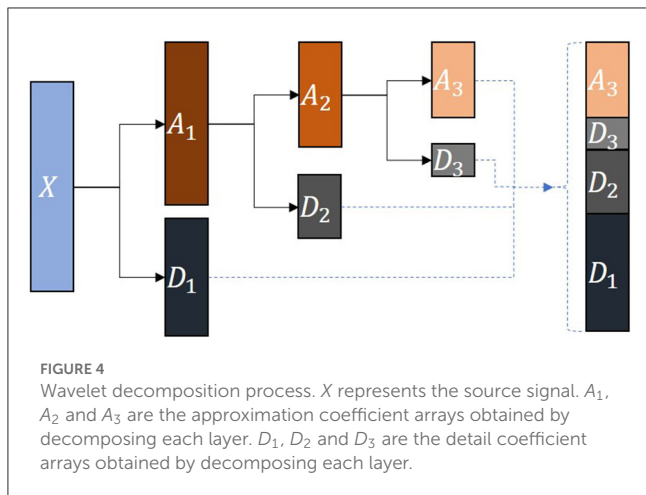


FIGURE 3
Diagram of the interframe difference. f_{j-1} , f_j and f_{j+1} represent three adjacent gait images in the gait video. The dotted line represents the movement trajectory of the key point.

a gait video. The motion track of the key points between each frame contains the interframe difference information.

In gait, some movement patterns are more easily reflected in the frequency domain (41). Relevant studies have extracted frequency-domain gait features through Fourier transforms (34, 35). However, Fourier transforms (42) cannot be applied in multiresolution analyses in the frequency domain. Thus, we use wavelet transforms (43) to analyze the frequency variation characteristics of the joint distances in the frequency domain.

We use the *db1* wavelet base to decompose the distance between joints into an approximation coefficient array A_3 representing low-frequency signals and detail coefficient arrays D_1 , D_2 , and D_3



representing high-frequency signals. Figure 4 shows the three-layer wavelet decomposition process.

We used 10 feature extraction functions to extract the above time-domain and frequency-domain features. These functions include the maximum, minimum, mean, median, variance, root mean square, skewness, kurtosis, absolute energy, and coefficient of variation in the sequence data. The specific feature extraction functions are shown in Supplementary Table C.

We used z-score standardization (44) to eliminate differences in the values and dimensions of features. The z-score standardization is defined as:

$$x' = \frac{x - \bar{x}}{\sigma_x}$$

Where x is the sample mean and σ_x is the sample standard deviation. Then, we used principal component analysis (PCA) (45) to remove redundant features and sequential forward selection (SFS) (46) to automatically identify feature combinations that resulted in optimal model performance. SFS is a greedy search algorithm. At each stage, according to the evaluation rules, the SFS algorithm continuously selects the optimal feature from the remaining features to determine the optimal feature subset. The SFS pseudocode is shown in Algorithm 1.

We selected 3 typical machine learning regression algorithms for modeling, namely, Gaussian process regression (GPR), linear regression (LR), and support vector regression (SVR), where the SVR models included the “linear,” “poly,” “rbf,” and “sigmoid” kernel functions. We trained and tested the models with 10 rounds of 10-fold cross validation. The complete modeling process is shown in Figure 5.

In computer science, the root mean square error (RMSE) is often used to evaluate regression model performance (47) and is defined as:

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (Model_n - Scale_n)^2}$$

Where $Model_n$ and $Scale_n$ represent the anxiety model prediction score and anxiety scale score of the n th participant, respectively.

Algorithm: Sequential Forward Selection.

Input:

X : The whole feature set

J : The model evaluation rules (Using RMSE)

Output:

S : The best subset of features

Method:

- (1) Create an empty subset $Z = \{\emptyset\}$
- (2) **repeat**
- (3) Select best remaining feature:
 $x = \arg \min_{x' \in Z_k} [J(Z + x)]$
- (4) Update $Z = Z + x$
- (5) $S = Z$
- (6) **until** not decreased in J OR $Z = X$

Algorithm 1. Pseudocode for the Sequential Forward Selection algorithm.

To comprehensively evaluate the performance of the proposed anxiety assessment models, we considered reliability and validity assessment methods used in psychology. We used the Pearson correlation between the anxiety assessment model prediction scores and the anxiety scale scores as the model criterion validity. In addition, we fed different data segments into the model to obtain prediction scores and used the Pearson correlation between these different model prediction scores to evaluate model reliability.

To explore the influence of the number of wavelet decomposition layers during the construction of the frequency-domain features on the prediction results, we set the wavelet decomposition *level* parameter from 1 to 4 (the *level* parameter controls the number of wavelet decomposition layers). Figure 6 shows the effect of decomposing the original time series signal according to different numbers of wavelet layers. The signals in each column can be restored to the original signal X after they are superimposed on each other.

To explore the influence of the gait video training data size on the model, we used gait segments with different numbers of frames to build various models and compared the model performance. In gait data segmentation, each participant has three segments of gait data, as shown in Figure 1. First, we used $segment_1$, $segment_2$ and $segment_3$ to establish three single-segment models. Then, two of the three segments were combined to establish three double-segment fusion models. Finally, the three segments were combined to establish a three-segment fusion model. The gait segments were combined as follows:

$$segment_{12} = segment_1 + segment_2$$

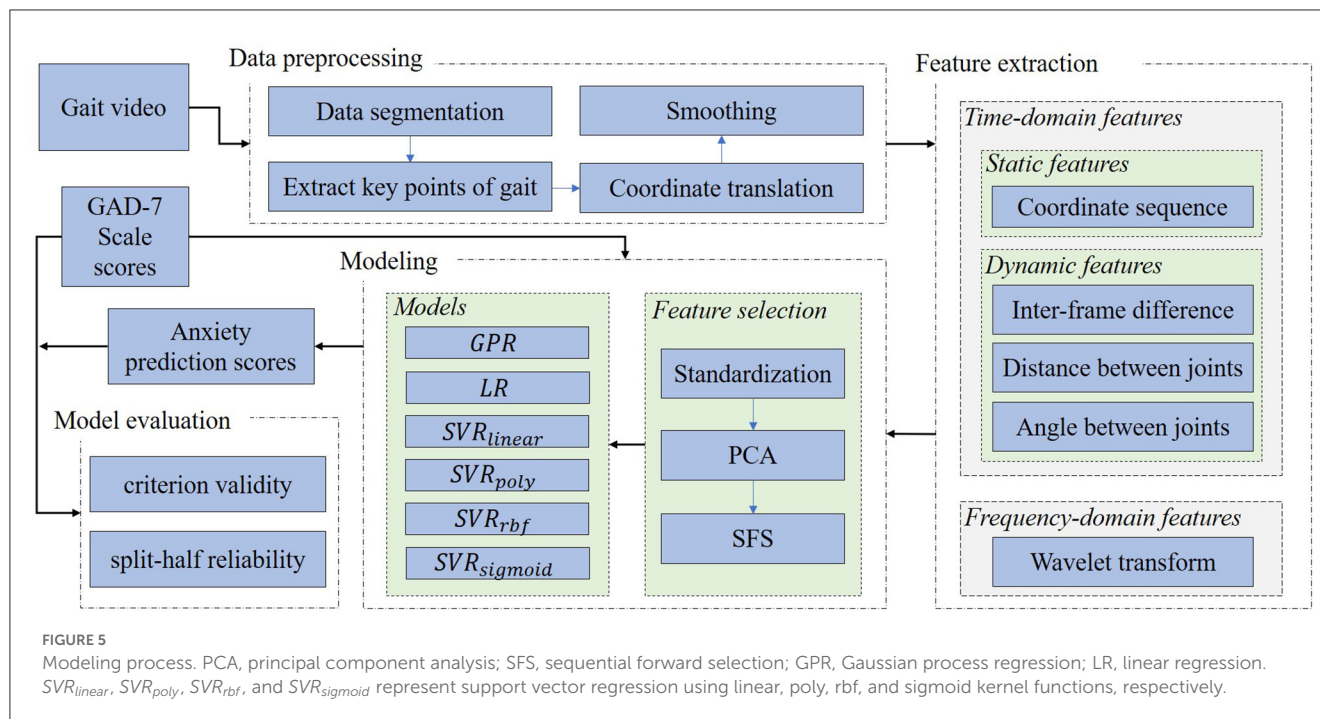
$$segment_{13} = segment_1 + segment_3$$

$$segment_{23} = segment_2 + segment_3$$

$$segment_{123} = segment_1 + segment_2 + segment_3$$

The Pearson correlation coefficients between the model prediction scores and the anxiety scale scores were calculated to evaluate the influence of the number of gait segment frames on the performance of the models.

In machine learning, some neural network components can be removed to understand their impact on the network (48). In this study, we explored the impact of different features on model



performance through feature ablation studies to determine whether the constructed features are effective. We used the static time-domain features, dynamic time-domain features, all time-domain features (including dynamic and static features), frequency-domain features, and all features (including all time-domain and frequency-domain features) to build 5 anxiety assessment models. The Pearson correlation coefficients between the model prediction scores and the scale scores were used to evaluate the contribution of different features to the model.

We also explored whether gender has an effect on anxiety prediction models. To accomplish this, we input the male and female gait data into the anxiety assessment model. Then, we calculated the Pearson correlation coefficients between the anxiety prediction scores of males and females and the corresponding scale scores to evaluate whether gender impacts the anxiety prediction model.

In psychology, odd-even split-half reliability is often used to characterize the degree of internal consistency of scales (49). We input the odd and even frame gait data into the anxiety assessment model to obtain the corresponding model prediction scores and used the Pearson correlation coefficient between the two prediction scores to evaluate the robustness and reliability of the model.

3. Results

We recruited 152 participants. According to the experimental processing requirements, 150 valid data remained after screening, including 79 males (52.67%) and 71 females (47.33%). The proportion of males and females was essentially balanced. The ages of the participants ranged from 21 to 28 years (mean = 22.99, SD = 1.07). The mean and standard deviation of the participant GAD-7 scores were 4.31 and 4.45, respectively. As shown in Table 1, the

participants mainly showed minimal and mild anxiety, with 132 participants at this anxiety level (88%). There were 5 participants with severe anxiety, and all were women.

Table 2 show that in terms of the different algorithms, the GPR and LR models had the best effect, regardless of the number of wavelet decomposition layers. In terms of the number of wavelet decomposition layers, except for the SVR_{poly} model (the SVR_{poly} model had the best effect when $level = 2$), the performance of the other models continuously improved as the number of layers increased from $level = 1$ to $level = 3$ (the mean values of r_{L1} , r_{L2} and r_{L3} were 0.401, 0.504, and 0.565, respectively). When $level = 4$, the model performance declined (the mean value of r_{L4} was 0.464). In summary, the GPR and LR models showed optimal performance when $level = 3$ ($r_{L3_GPR} = 0.677$, $r_{L3_LR} = 0.677$, $p < 0.001$, and their RMSE values were less than those of the other algorithms). We determined the optimal number of wavelet decomposition layers by iteratively searching parameters.

As shown in Table 3, among the 7 data combinations, the GPR and LR models had the best results. In the GPR and LR models, the modeling effects of the $segment_1$, $segment_{12}$, $segment_{13}$ and $segment_{123}$ gait segments (which all contained $segment_1$ and had mean r_{s1} , r_{s12} , r_{s13} and r_{s123} values of 0.559, 0.495, 0.495, and 0.516, respectively) were better than those of the other segments (the mean values of r_{s2} , r_{s3} and r_{s23} were 0.425, 0.435, and 0.447, respectively). Similar trends were found for the SVR_{rbf} and $SVR_{sigmoid}$ models. In conclusion, the GPR and LR models had the best performance when modeled on $segment_1$ ($r_{s1_GPR} = 0.731$, $r_{s1_LR} = 0.702$, $p < 0.001$). We found that there are some differences in the modeling effect of gait segments in different periods. Moreover, the increase in the number of gait segments did not significantly improve the model effect.

As shown in Table 4, the modeling effects of the GPR and LR models on different features were significantly better than those of

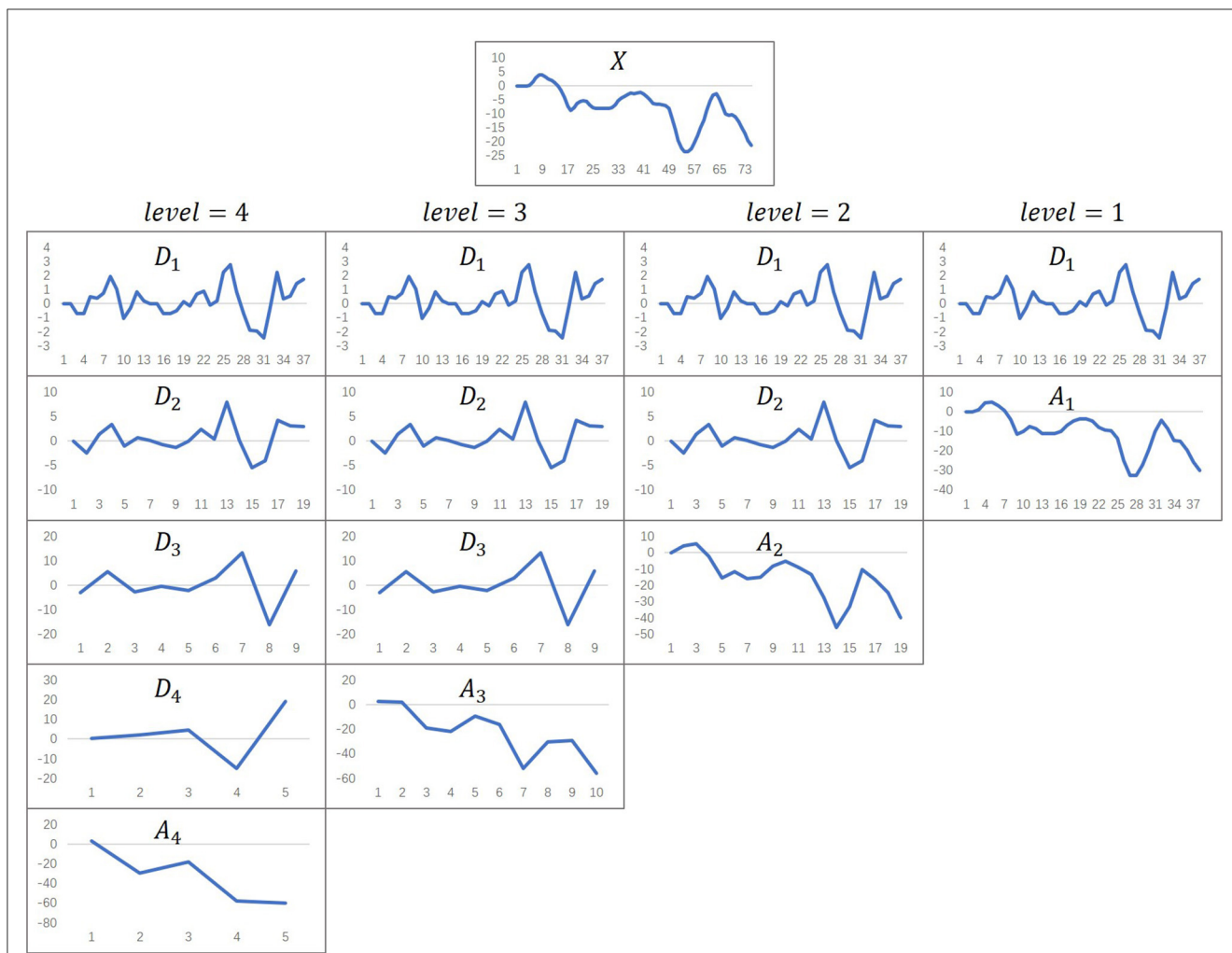


FIGURE 6
The effect of wavelet decomposition. The level parameter represents the number of wavelet decomposition layers. X represents the original time series data. D_i ($i \in \{1, 2, 3, 4\}$) represent detail coefficient arrays. A_j ($j \in \{1, 2, 3, 4\}$) represent approximation coefficient arrays.

TABLE 1 Population distribution of GAD-7 scale scores.

	GAD-7 scale score range				Total
	0~4	5~9	10~14	15~21	
Male	55	18	6	0	79
Female	41	18	7	5	71
Total	96	36	13	5	150

GAD-7, the 7-item Generalized Anxiety Disorder scale; 0~4, minimal anxiety; 5~9, mild anxiety; 10~14, moderate anxiety; 15~21, severe anxiety.

the other models. The GPR model achieved the best modeling effect on all features, including the time-domain and frequency-domain features ($r_{5_GPR} = 0.725, p < 0.001$). The mean values of $r_1, r_2, r_3, r_4,$ and r_5 were 0.399, 0.446, 0.536, 0.565, and 0.560, respectively, showing a slow increasing trend. These trends were particularly noticeable in the GPR and LR models, with $r_{5_GPR} > r_{4_GPR}$ and $r_{5_LR} > r_{4_LR}$ ($p < 0.001$). We found that the anxiety assessment models are sensitive to different gait features. And gait

features with kinematic characteristics can significantly improve the performance of the model.

As shown in Table 5, the GPR model performed significantly better than the other models ($r_{All_GPR} = 0.725, r_{Male_GPR} = 0.666, r_{Female_GPR} = 0.763, p < 0.001$, and its RMSE value was lower than those of the other algorithms). The anxiety prediction effect was better for women than for men (the mean values of r_{Male} and r_{Female} were 0.547 and 0.566, respectively). Except for the SVR_{linear} and SVR_{poly} models, all other models reflected this characteristic. We found that the prediction performance of anxiety assessment model for different groups is different.

As shown in Table 6, except for SVR_{poly}, all models showed good reliability, and their odd-even split-half reliability was > 0.8 . This proved the stability of the model to a certain extent. In conclusion, the GPR model obtained the best criterion validity and split-half reliability performance.

Gait-based anxiety assessment methods have not been fully established. Here we migrated our method to a similar dataset (34). The results showed that the GPR model had the best effect. The Pearson correlation coefficient between the predicted scores of the

TABLE 2 Criterion validity of frequency-domain feature modeling using different numbers of wavelet decomposition layers.

	$RMSE_{L_1}$	r_{L_1}	$RMSE_{L_2}$	r_{L_2}	$RMSE_{L_3}$	r_{L_3}	$RMSE_{L_4}$	r_{L_4}
GPR	4.027	0.475	3.568	0.594	3.273	0.677	3.830	0.564
LR	4.092	0.471	3.593	0.594	3.291	0.677	3.859	0.565
SVR _{linear}	4.024	0.408	3.967	0.430	3.619	0.562	3.946	0.441
SVR _{poly}	4.223	0.269	3.772	0.520	3.915	0.437	4.085	0.409
SVR _{rbf}	4.105	0.405	4.008	0.434	3.967	0.496	4.071	0.390
SVR _{sigmoid}	4.045	0.375	3.952	0.451	3.773	0.542	3.988	0.416

The subscripts L_1 , L_2 , L_3 and L_4 indicate that the numbers of wavelet decomposition layers are 1, 2, 3, and 4 (the *level* parameter ranges from 1 to 4), respectively, when constructing the frequency-domain features. $RMSE$ and r represent the root mean square error and criterion validity of the model established using the frequency-domain features, respectively. All correlation coefficients are highly significant ($p < 0.001$).

TABLE 3 Criterion validity of modeling with different training data sizes.

	r_{s_1}	r_{s_2}	r_{s_3}	$r_{s_{12}}$	$r_{s_{13}}$	$r_{s_{23}}$	$r_{s_{123}}$
GPR	0.731	0.543	0.578	0.633	0.592	0.545	0.634
LR	0.702	0.547	0.578	0.630	0.583	0.545	0.637
SVR _{linear}	0.542	0.276	0.320	0.362	0.540	0.426	0.494
SVR _{poly}	0.403	0.386	0.372	0.392	0.314	0.425	0.354
SVR _{rbf}	0.460	0.454	0.403	0.526	0.487	0.425	0.490
SVR _{sigmoid}	0.518	0.346	0.359	0.424	0.454	0.314	0.488

r_{s_1} , r_{s_2} and r_{s_3} represent the criterion validity of the models established using gait segments *segment*₁, *segment*₂ and *segment*₃, respectively. $r_{s_{12}}$, $r_{s_{13}}$ and $r_{s_{23}}$ represent the criterion validity of the models established after combining any two of the three gait segments. $r_{s_{123}}$ represents the criterion validity of the model established after combining all three gait segments. All correlation coefficients are highly significant ($p < 0.001$).

TABLE 4 Ablation studies with different modeling features.

	r_1	r_2	r_3	r_4	r_5
GPR	0.462	0.602	0.681	0.677	0.725
LR	0.461	0.595	0.680	0.677	0.704
SVR _{linear}	0.349	0.274	0.498	0.562	0.540
SVR _{poly}	0.410	0.368	0.467	0.437	0.404
SVR _{rbf}	0.378	0.428	0.459	0.496	0.457
SVR _{sigmoid}	0.336	0.407	0.432	0.542	0.528

r_1 , r_2 , r_3 , r_4 and r_5 represent the criterion validity of the models developing using static time-domain features, dynamic time-domain features, all time-domain features (including dynamic and static features), frequency-domain features, and all features (including all time-domain and frequency-domain features), respectively. All correlation coefficients are highly significant ($p < 0.001$).

anxiety assessment model and the scale scores reached 0.6, which was higher than the 0.4 reported by Miao et al. (34). In addition, we also tested the odd-even split-half reliability of the model on this dataset to 0.8. This shows that our anxiety assessment model has good robustness.

4. Discussion

We demonstrated that automated anxiety assessment using 2D gait videos is feasible. Based on 2D gait videos, we constructed

TABLE 5 Criterion validity of the anxiety assessment model for males and females.

	$RMSE$	r_{All}	r_{Male}	r_{Female}
GPR	3.185	0.725	0.666	0.763
LR	3.430	0.704	0.639	0.722
SVR _{linear}	3.698	0.540	0.632	0.446
SVR _{poly}	4.018	0.404	0.404	0.361
SVR _{rbf}	3.948	0.457	0.469	0.512
SVR _{sigmoid}	3.823	0.528	0.474	0.590

$RMSE$, root mean square error. r_{All} , r_{Male} and r_{Female} represent the criterion validity of the model for all participants, male participants, and female participants, respectively. All correlation coefficients are highly significant ($p < 0.001$).

TABLE 6 The odd-even split-half reliability of anxiety assessment models.

	GPR	LR	SVR _{linear}	SVR _{poly}	SVR _{rbf}	SVR _{sigmoid}
$r_{split-half}$	0.803	0.801	0.808	-0.696	0.876	0.883

$r_{split-half}$ represents the odd-even split-half reliability. All correlation coefficients are highly significant ($p < 0.001$).

and fused static and dynamic time-domain features and frequency-domain features and used machine learning methods to establish anxiety assessment models. Moreover, we evaluated the criterion validity and split-half reliability of the proposed anxiety prediction models. We also assessed the effects of different frequency-domain feature construction methods, gait training data sizes, and gender differences on the modeling results, verifying the contributions of various time-domain and frequency-domain features. Our results showed that the proposed gait video-based anxiety assessment method had good reliability and validity.

People with anxiety disorders tend to be between 15 and 35 years old (50). Higher education levels appear to have a protective effect on anxiety and depression (51). In our study, the participants ranged from 21 to 28 years old, their educational backgrounds were mainly involved postgraduate education, and their anxiety levels were concentrated between minimal and mild anxiety. This showed that our sample had a certain representativeness in the higher education student groups.

We used the $RMSE$ to evaluate the relative performance of different models. Smaller $RMSE$ and larger r values indicate better

model performance. In Tables 2, 4, the RMSE and r values showed inverse trends. This result showed that it was reasonable to use the criterion validity to evaluate the performance of the models.

As the number of wavelet decomposition layers increases, we can obtain more detail coefficient arrays representing high-frequency information and more approximate coefficient arrays representing low-frequency information. Since our sequence length was 75, the coefficient arrays that cannot be divided into half are filled with zeros in each wavelet decomposition. When the wavelet decomposition level was too high, the length of the coefficient array was too short, and the zero-padding operation introduced more errors, which led to inaccurate frequency-domain features. This was why the mean value of r_{L_4} was smaller than that of r_{L_3} . Therefore, in wavelet decomposition, as the number of decomposition layers increases, we can more easily distinguish between low-frequency and high-frequency signals. However, the interference errors caused by the continuous subdivision also increase.

In general, in machine learning, more training data leads to better model effects (52). In our experiments, the model performance did not improve and even decreased as the number of gait training segment frames increased. For example, as shown in Table 3, the modeling effect after fusing two or three gait segments was worse than that of single gait segment modeling. On the one hand, gait is a periodic process (53). More gait segments lead to redundant information that does not contribute to modeling. Therefore, it is sufficient to model with fewer gait frames, which is similar to previous research results (34, 35, 37). On the other hand, different gait segments are discontinuous, and directly merging these sequences may cause mutations that reduce model performance to some extent. We also observed that the modeling effect of gait data including $segment_1$ was better than that of data including other segments, which may be due to the fatigue of participants walking back and forth in a narrow space, which led to inaccuracies in the subsequent gait videos.

Feature ablation studies were performed to examine how different features contribute to modeling. Taking the GPR model with good reliability and validity as an example, $r_{3_GPR} > r_{2_GPR} > r_{1_GPR}$ verified that gait contains both dynamic and static information and that dynamic information expresses gait characteristics better than static information. Moreover, $r_{5_GPR} > r_{4_GPR}$ and $r_{5_GPR} > r_{3_GPR}$ verified that time-domain and frequency-domain information both contribute to modeling. The results of the feature ablation studies showed that the various constructed features were effective and necessary.

Previous studies have shown that the muscular strength of anxious women is significantly lower than that of healthy women and that these two groups show differences in gait, while these differences are not obvious among males (23). In addition, anxiety differs between the genders, and females are more likely to be anxious than males (54). This may be the reason why the anxiety prediction results are better for women than for men. This fact also supports the finding that participants with severe anxiety in Table 1 were all women.

Cronbach's alpha for the GAD-7 scale was 0.92 (9). In general, an alpha value >0.7 is considered to indicate acceptable reliability. In this study, except for the SVR_{poly} model, the split-half reliability of the models was >0.8 . This result indicates that the odd-even split-half reliability can be applied to evaluate model performance.

This study is a continuation and extension of our previous work (37). We have optimized the methods of data segmentation, frequency-domain feature construction, and feature selection in experiments. Compared with previous studies, we explored in detail the impact of various factors (different features, gait dataset size, gender) on the model through comparative experiments with various parameters. In this study, the modeling method is more objective and reasonable, and the robustness and predictive performance of the anxiety assessment model are improved. Our research has some limitations. During data collection, a single camera was used to capture gait videos of the participants walking back and forth. Thus, the data contained some gait segments (such as turning and back gaits) that were not suitable for modeling. During preprocessing, the segmentation and recombination of different gait segments might introduce data breakpoints that can impact the model effects. In the future, we set the gait data collection scene as participants walking normally on the treadmill, ensuring that only the participants' front-view gait videos are recorded. We will try to avoid damaging the continuity of gait videos in preprocessing. In addition, although we verified the feasibility of assessing anxiety state based on gait videos, the participants were mainly college graduate students. Since this model was trained on only one social group, the generalizability may be insufficient. Thus, we will recruit participants from different groups according to the differences in age, gender, region, culture and economic background to increase the diversity of training data.

Due to the convenience, real-time, and non-invasive properties of our model, our approach can be applied in various scenarios. For example, the model can be applied for personal daily anxiety assessment. Moreover, companies can learn the employee anxiety levels through video data to provide psychological counseling in a timely manner and improve work efficiency. Using this method to assess the anxiety level of social groups in a timely manner can help to improve community mental health and public health. In future work, our proposed method still has some room for improvement. First, our current research uses traditional machine learning models and artificially constructed features. Although we have demonstrated the rationality and effectiveness of the constructed features in experiments, we still rely on a lot of subjective experience in the early stage. In recent years, many studies have made breakthroughs using deep learning (55). So next we will apply deep neural network to automatically extract gait features and train anxiety assessment models with better predictive performance. Second, our current research needs to convert gait video frame by frame into human body key point coordinates, and then calculate and analyze based on these 2D coordinates. In the process of extracting key points, some gait information will be lost, which will affect the model's learning of gait information. In the future work, we will use image streams for modeling directly based on gait video, so that the neural network can capture more detailed information in the gait.

5. Conclusion

In this study, we developed a convenient and timely anxiety assessment method that may contribute to improving mental health services. Our experiments show that gait can be used

as an objective cue to measure anxiety, the gait video-based anxiety assessment model has good criterion validity and split-half reliability, and the model has a better prediction effect on females than males. In addition, due to the periodicity of gait, increasing the number of gait training segment frames has little effect on the performance of the anxiety assessment model. The results of comparative experiments showed that the static and dynamic time-domain features and frequency-domain features improved model performance. Our preliminary study provides ideas for developing a convenient real-time anxiety assessment method.

Data availability statement

To protect the privacy of the participants, the original datasets in the article cannot be made public. If necessary, feature datasets of gait are available from the corresponding author on reasonable request. Requests to access the datasets should be directed to TZ, tszhu@psych.ac.cn.

Ethics statement

The studies involving human participants were reviewed and approved by Institutional Review Board of the Institute of Psychology, Chinese Academy of Sciences. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YW, BL, XL, and TZ proposed the idea of the research and designed the research method. DC and SG put forward constructive suggestions. TZ and XL provided research data. YW completed

the data analysis and modeling and completed the first draft of the manuscript. TZ and SG guided the research process. All authors participated in the editing and reviewing of manuscripts and contributed to the article and approved the submitted version.

Funding

This research was funded by the Scientific Foundation of Institute of Psychology, Chinese Academy of Sciences, No. E2CX4735YZ.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2023.1082139/full#supplementary-material>

References

1. Vindegaard N, Benros ME. Covid-19 pandemic and mental health consequences: systematic review of the current evidence. *Brain Behav Immun.* (2020) 89:531–42. doi: 10.1016/j.bbi.2020.05.048
2. Association AP. *Diagnostic and Statistical Manual of Mental Disorders : Dsm-5*. 5th Ed Arlington VA Washington DC: American Psychiatric Association (2013).
3. Salari N, Hosseini-Far A, Jalali R, Vaisi-Raygani A, Rasoulpoor S, Mohammadi M, et al. Prevalence of stress, anxiety, depression among the general population during the covid-19 pandemic: a systematic review and meta-analysis. *Global Health.* (2020) 16:11. doi: 10.1186/s12992-020-00589-w
4. World Health Organization. *World Health Statistics 2022: Monitoring Health for the Sdgs, Sustainable Development Goals* Geneva: World Health Organization (2022).
5. Chalah MA, Ayache SS. Noninvasive brain stimulation and psychotherapy in anxiety and depressive disorders: a viewpoint. *Brain Sci.* (2019) 9:82. doi: 10.3390/brainsci9040082
6. Chalah MA, Ayache SS. Disentangling the neural basis of cognitive behavioral therapy in psychiatric disorders: a focus on depression. *Brain Sci.* (2018) 8:150. doi: 10.3390/brainsci8080150
7. Rossi PH, Wright JD, Anderson AB. *Handbook of Survey Research*. London: Academic press (2013).
8. Jupp V. *The Sage Dictionary of Social Research Methods*. London: SAGE Publications, Ltd (2006).
9. Spitzer RL, Kroenke K, Williams JB, Löwe B, A. Brief measure for assessing generalized anxiety disorder: the Gad-7. *Arch Intern Med.* (2006) 166:1092–7. doi: 10.1001/archinte.166.10.1092
10. Althubaiti A. Information bias in health research: definition, pitfalls, and adjustment methods. *J Multidiscip Healthc.* (2016) 9:211. doi: 10.2147/JMDH.S104807
11. Duff K, Beglinger LJ, Schultz SK, Moser DJ, McCaffrey RJ, Haase RE, et al. Practice effects in the prediction of long-term cognitive outcome in three patient samples: a novel prognostic index. *Arch Clin Neuropsychol.* (2007) 22:15–24. doi: 10.1016/j.acn.2006.08.013
12. Podsakoff PM, MacKenzie SB, Podsakoff NP. Sources of method bias in social science research and recommendations on how to control it. *Annu Rev Psychol.* (2012) 63:539–69. doi: 10.1146/annurev-psych-120710-100452
13. Hemmings C, Bouras N. *Psychiatric and Behavioural Disorders in Intellectual and Developmental Disabilities*, 3rd Edn. Cambridge: Cambridge University Press (2016).
14. Stein MB, Sareen J. Generalized anxiety disorder. *N Engl J Med.* (2015) 373:2059–68. doi: 10.1056/NEJMc1502514
15. Marks I, Marset P, Boulougouris J, Huson J. Physiological accompaniments of neutral and phobic imagery. *Psychol Med.* (1971) 1:299–307. doi: 10.1017/S0033291700042264

16. Surcinelli P, Codispoti M, Montebanacci O, Rossi N, Baldaro B. Facial emotion recognition in trait anxiety. *J Anxiety Disord.* (2006) 20:110–7. doi: 10.1016/j.janxdis.2004.11.010
17. Giannakakis G, Padiaditis M, Manousos D, Kazantzaki E, Chiarugi F, Simos PG, et al. Stress and anxiety detection using facial cues from videos. *Biomed Signal Process Control.* (2017) 31:89–101. doi: 10.1016/j.bspc.2016.06.020
18. Siegman AW, Boyle S. Voices of fear and anxiety and sadness and depression: the effects of speech rate and loudness on fear and anxiety and sadness and depression. *J Abnorm Psychol.* (1993) 102:430–7. doi: 10.1037/0021-843X.102.3.430
19. Wortwein T, Morency LP, Scherer S, Ieee, editors. Automatic assessment and analysis of public speaking anxiety: a virtual audience case study. In: *6th AAAC Affective Computing and Intelligent Interaction International Conference (ACII); 2015 Sep 21–24; Xian, Peoples Republic China.* New York: IEEE (2015). doi: 10.1109/ACII.2015.7344570
20. Roether CL, Omlor L, Christensen A, Giese MA. Critical features for the perception of emotion from gait. *J Vis.* (2009) 9:15. doi: 10.1167/9.6.15
21. Montepare JM, Goldstein SB, Clausen A. The identification of emotions from gait information. *J Nonverbal Behav.* (1987) 11:33–42. doi: 10.1007/BF00999605
22. Seligman ME, Walker EF, Rosenhan DL. *Abnormal Psychology.* Norton (2001).
23. Feldman R, Schreiber S, Pick CG, Been E. Gait, balance, mobility and muscle strength in people with anxiety compared to healthy individuals. *Hum Mov Sci.* (2019) 67:10. doi: 10.1016/j.humov.2019.102513
24. Reelick MF, van Iersel MB, Kessels RP, Rikkert MGO. The influence of fear of falling on gait and balance in older people. *Age Ageing.* (2009) 38:435–40. doi: 10.1093/ageing/afp066
25. Staab JP, Balaban CD, Furman JM, editors. Threat assessment and locomotion: clinical applications of an integrated model of anxiety and postural control. *Semin Neurol.* (2013) 33:297–306. doi: 10.1055/s-0033-1356462
26. Hainaut J-P, Caillet G, Lestienne FG, Bolmont B. The role of trait anxiety on static balance performance in control and anxiogenic situations. *Gait Posture.* (2011) 33:604–8. doi: 10.1016/j.gaitpost.2011.01.017
27. Bolmont BT, Gangloff P, Vouriot A, Perrin PP. Mood states and anxiety influence abilities to maintain balance control in healthy human subjects. *Neurosci Lett.* (2002) 329:96–100. doi: 10.1016/S0304-3940(02)00578-5
28. Michalak J, Troje NF, Fischer J, Vollmar P, Heidenreich T, Schulte D. Embodiment of sadness and depression—gait patterns associated with dysphoric mood. *Psychosom Med.* (2009) 71:580–7. doi: 10.1097/PSY.0b013e3181a2515c
29. Greene BR, Foran TG, McGrath D, Doheny EP, Burns A, Caulfield B, et al. Comparison of algorithms for body-worn sensor-based spatiotemporal gait parameters to the gaitrite electronic walkway. *J Appl Biomech.* (2012) 28:349–55. doi: 10.1123/jab.28.3.349
30. Moeslund TB, Hilton A, Kruger V. A survey of advances in vision-based human motion capture and analysis. *Comput Vis Image Underst.* (2006) 104:90–126. doi: 10.1016/j.cviu.2006.08.002
31. Cloete T, Scheffer C, Ieee, editors. Benchmarking of a full-body inertial motion capture system for clinical gait analysis. In: *30th Annual International Conference of the IEEE-Engineering-in-Medicine-and-Biology-Society; 2008 Aug 20–24; Vancouver, CANADA.* New York: IEEE (2008).
32. Li QN, Wang YF, Sharf A, Cao Y, Tu CH, Chen BQ, et al. Classification of gait anomalies from kinect. *Visual Comput.* (2018) 34:229–41. doi: 10.1007/s00371-016-1330-0
33. Jing C, Liu X, Zhao N, Zhu T, editors. Different performances of speech and natural gait in identifying anxiety and depression. In: *International Conference on Human Centered Computing.* Cham: Springer (2019). doi: 10.1007/978-3-030-37429-7_20
34. Miao B, Liu X, Zhu T. Automatic mental health identification method based on natural gait pattern. *PsyCh J.* (2021) 10:453–64. doi: 10.1002/pchj.434
35. Zhao N, Zhang Z, Wang Y, Wang J, Li B, Zhu T, et al. See your mental state from your walk: recognizing anxiety and depression through kinect-recorded gait data. *PLoS ONE.* (2019) 14:e0216591. doi: 10.1371/journal.pone.0216591
36. Stark M, Huang H, Yu L-F, Martin R, McCarthy R, Locke E, et al. Identifying individuals who currently report feelings of anxiety using walking gait and quiet balance: an exploratory study using machine learning. *Sensors.* (2022) 22:3163. doi: 10.3390/s22093163
37. Wen Y, Li B, Chen D, Zhu T. Reliability and validity analysis of personality assessment model based on gait video. *Front Behav Neurosci.* (2022) 16:901568. doi: 10.3389/fnbeh.2022.901568
38. Fang J, Wang T, Li C, Hu X, Ngai E, Seet B-C, et al. Depression prevalence in postgraduate students and its association with gait abnormality. *IEEE Access.* (2019) 7:174425–37. doi: 10.1109/ACCESS.2019.2957179
39. Cao Z, Hidalgo G, Simon T, Wei SE, Sheikh Y. Openpose: realtime multi-person 2d pose estimation using part affinity fields. *IEEE Trans Pattern Anal Mach Intell.* (2021) 43:172–86. doi: 10.1109/TPAMI.2019.2929257
40. Murray MP. Gait as a total pattern of movement: including a bibliography on gait. *Am J Phys Med Rehabil.* (1967) 46:290–333.
41. Orović I, Stanković S, Amin M. A new approach for classification of human gait based on time-frequency feature representations. *Signal Process.* (2011) 91:1448–56. doi: 10.1016/j.sigpro.2010.08.013
42. Nussbaumer HJ. The fast fourier transform. In: Nussbaumer HJ, editor. *Fast Fourier Transform and Convolution Algorithms.* Berlin, Heidelberg: Springer Berlin Heidelberg (1981). p. 80–111. doi: 10.1007/978-3-662-00551-4_4
43. Daubechies I. *The Wavelet Transform, Time-Frequency Localization and Signal Analysis.* Princeton, NJ: Princeton University Press (2009). doi: 10.1515/9781400827268.442
44. Zill DG. *Advanced Engineering Mathematics.* Burlington, MA: Jones & Bartlett Publishers (2020).
45. Bishop CM. *Pattern Recognition and Machine Learning.* New York, NY: Springer (2006).
46. Reeves SJ, Zhe Z. Sequential algorithms for observation selection. *IEEE Trans Signal Process.* (1999) 47:123–32. doi: 10.1109/78.738245
47. Zhou Z-H. *Machine Learning.* Singapore: Springer (2021). doi: 10.1007/978-981-15-1967-3
48. Meyers R, Lu M, de Puiseau CW, Meisen T. Ablation studies in artificial neural networks. *arXiv [Preprint].* (2019). *arXiv:* 1901.08644. doi: 10.48550/arXiv.1901.08644
49. Bartko JJ, Carpenter WT. On the methods and theory of reliability. *J Nerv Ment Dis.* (1976) 163:307–17. doi: 10.1097/00005053-19761000-00003
50. Craske MG, Stein MB. Anxiety. *Lancet.* (2016) 388:3048–59. doi: 10.1016/S0140-6736(16)30381-6
51. Bjelland I, Krokstad S, Mykletun A, Dahl AA, Tell GS, Tambs K. Does a higher educational level protect against anxiety and depression? The Hunt study. *Soc Sci Med.* (2008) 66:1334–45. doi: 10.1016/j.socscimed.2007.12.019
52. Luyckx K, Daelemans W. The effect of author set size and data size in authorship attribution. *Lit Linguist Comput.* (2011) 26:35–55. doi: 10.1093/llc/fqq013
53. Baker R, Hart HM. *Measuring Walking: A Handbook of Clinical Gait Analysis.* London: Mac Keith Press (2013).
54. Lewinsohn PM, Gotlib IH, Lewinsohn M, Seeley JR, Allen NB. Gender differences in anxiety disorders and anxiety symptoms in adolescents. *J Abnorm Psychol.* (1998) 107:109. doi: 10.1037/0021-843X.107.1.109
55. Amanat A, Rizwan M, Javed AR, Abdelhaq M, Alsaqour R, Pandya S, et al. Deep learning for depression detection from textual data. *Electronics.* (2022) 11:676. doi: 10.3390/electronics11050676