How to estimate the emotions hidden behind spatio-temporal

Abstract

Emotion estimation of online spatiotemporal behavior is a technique for studying mental health and its changing laws based on spatiotemporal trajectory data of objects. According to WHO data, the proportion of patients with depression worldwide is as high as 3.7%, and mental health detection technology has become a new hotspot in current international research. Traditional technologies mainly collect physiological data such as heart rate, blood pressure, blood oxygen and sleep through wearable devices (such as wristbands) to achieve online analysis of mental health levels. However, the low measurement accuracy of wearable devices makes it difficult to meet the quality requirements for emotion estimation. More importantly, emotional changes are not only affected by physiological factors, but social factors are more important. This paper studies the relationship between the object's spatiotemporal behavior and emotional state, focusing on the mechanism of the object's social behavior pattern and its changes on emotional changes. A social activity pattern extraction method based on spatio-temporal trajectory data is proposed, a social activity sequence expression model of the subject's daily behavior is established, and the mapping relationship between the social activity sequence and the emotional index under multi-resolution is explored. The experimental results show that the object's social and social activity patterns are closely related to its emotional index. The proposed **SADS** emotion estimation model is better than the baseline paper on both **SAPD22111510** and **SAPD23031530** datasets, with an average increase in accuracy of 3.9% and 8.1% respectively. For the first time, the paper expands the research object of online emotion estimation from traditional spatiotemporal behavior to social behavior pattern research, which provides new research ideas and technical approaches for online emotion estimation research.

1 Introduction

In recent years, smartphones have become an indispensable part of people's daily lives. With the widespread use of smartphones and the rapid development of mobile applications and sensor technologies, we can collect a large amount of spatiotemporal trajectory data through smartphones. Currently, there is a large body of literature that uses spatiotemporal trajectory data for trajectory classification（Yu xuan Liang et. al. 2022), abnormal trajectory detection (Han x et.al. 2022), trajectory encryption (Li D et.al. 2020), and provides rich information for studying human behavior and emotions. This paper aims to explore how to use spatiotemporal trajectory data collected by smartphones for emotion analysis, in order to provide new perspectives and methods for psychological health research and applications.

Spatiotemporal trajectory data includes the location information of users at specific times and latitude and longitude

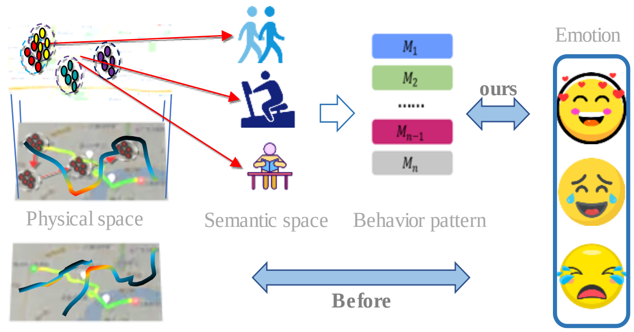


Figure 1: Extraction of social attributes from spatio-temporal trajectories.

locations. This information can be further mined to analyze users' activity patterns, behavioral habits, and social interactions. By analyzing this data, we can uncover users' emotional states and psychological characteristics. For example, by analyzing users' individual spatiotemporal activities and patterns during a certain period of time, we can infer their emotional fluctuations, such as anxiety, depression, or happiness. In addition, by analyzing users' social behaviors, we can also understand the relationship between their emotional states and social situations, as well as further analyze their emotional support networks and psychological needs.

Emotions are described as impulsive responses to significant internal or external events, with subjects always reacting to the same event. They involve a range of coordinated linguistic, physiological, behavioral, and neural mechanisms of response. (Fox et.al. 2008)Emotions generate different physiological, behavioral, and cognitive changes, and motivate behavior. (Schacter et.al. 2011) Emotional responses are a step that provides drive and direction. When a person has emotional or psychological issues, their behavior often differs greatly from that of a normal person. Emotional state is one of the key factors that affect behavior.

Due to the potential impact of emotions, extensive research has been conducted on emotion detection. We hope to detect changes in emotions in advance to avoid dangerous events caused by abnormal behavior. In addition, given the increasing importance of mental health to individuals and society, researchers are now looking for accurate and rapid methods to detect human emotions. Therefore, many existing methods are machine learning techniques that attempt to automatically detect emotions based on physiological data using sensor datasets (Kanjo et.al. 2018) (Kanjo et.al. 2017) (Jerritta et.al. 2011) (Busso et.al. 2004) So far, a large amount of automatic emotion detection research has used visual, audio, and motion data (such as facial expressions, postures, and speech) (Jerritta et.al. 2011) (Busso et.al. 2004)(Kanjo et.al. 2015). With the increasing availability of low-cost wearable sensors (such as Fitbit, Microsoft wristbands), research using human physiological data (such as skin conductance), heart rate (HR), electroencephalogram (EEG), etc. has become interesting for emotion monitoring. However, using physiological data for emotion detection may be challenging. Firstly, many available physiological signals may not be distinguishable in emotion detection (Agrafioti et.al. 2012). Secondly, emotion detection based on physiological feature data is subject to strict experimental environment control, which may have a significant impact on the daily life of the subjects and is relatively weak in complex scenarios of daily life and outdoor personnel, except for medical diagnosis (Agrafioti et.al. 2012).

In this article, we attempt to detect the corresponding emotions of our subjects by utilizing the spatiotemporal trajectories of their mobile phones. We focus on the relationship between the subjects' behavioral patterns and their emotions, and analyze their emotional states based on their behavior. In 2013, R Li Kam Wa et al. (LiKamWa et.al. 2013) first used mobile phone data to analyze human emotional states. By analyzing the history and usage patterns of social communication software, we can calculate the user's average daily emotions. The initial accuracy rate was 66%. After two months of personalized training, the accuracy gradually increased to 93%. Eiman-Kanjo (Exler et.al. 2016) and others first used location data for emotion detection and classification. They trained physiological, environmental, and location data on three models, MLP, CNN, and CNN-LSTM, respectively, and then combined these three data for training. Finally, they evaluated the performance of each model based on the combined model. Compared with environmental data, location data has higher accuracy. It also indicates the feasibility of using the spatiotemporal trajectory of the target of interest for emotion classification. However, using the spatiotemporal features of the target's behavior for emotion classification requires fixed sampling locations with specified behaviors and significant changes in spatiotemporal features to be effective. Additionally, modeling based on raw spatiotemporal trajectories requires high data sampling frequency, high computational complexity, and can suffer from the problem of vanishing gradients during model training. Furthermore, it cannot be extended to different regions or applied to emotion classification tasks for key individuals in complex real-life scenarios. The interpretability of the classification results is insufficient, as it cannot provide the basis for classification.

Specifically, our research on emotion perception based on spatiotemporal trajectories mainly focuses on the spatiotemporal trajectories of location-based objects. Our method first detects stay points based on the spatiotemporal trajectory data of the target of interest to discover their behavioral activities. We extract multiple semantic information from the stay points, including but not limited to stay point attributes, stay duration, stay time, and participants. As shown in Figure 1,our approach differs from existing methods in the literature that analyze the behavior of the target of interest in physical space or semantic space using stay points. Instead, we focus on the impact mechanism of social behavior patterns and their changes on emotion changes, exploring the influence of social factors on individual emotions. In this study, we analyze stay point data containing multiple semantic information to discover the significance and regularity of social attributes behind the behavioral activities of the target of interest. We then classify and predict their emotions based on the changes in social attribute-level behavior patterns. Our contributions are as follows:

- This study explores the correlation between spatiotemporal behavior and emotional states, focusing on the impact mechanism of social behavior patterns and their changes on emotion changes. We propose and verify the hypothesis of behavior pattern differentiation in the case of individual emotional abnormalities.
- We propose a method for extracting social activity patterns based on spatiotemporal trajectory data and establish a social activity sequence representation model for the target's daily behaviors. Based on the target's activity intentions, we adaptively provide activity types, which can more accurately express the target's behavior patterns.
- The paper explores a non-intrusive, sustainable, and real-time method for detecting abnormal emotions using behavior pattern analysis. For the first time, the research object of online emotion estimation is expanded from traditional spatiotemporal behavior to social behavior patterns, providing new research ideas and technical approaches for online emotion estimation.

2 Related Work

It contains said We summarize related work from three aspects: Track mining, activity pattern modeling and representation learning.

2.1 Behavioral pattern representation

Behavioral pattern representation is an important branch of spatiotemporal behavioral data analysis, which transforms difficult spatiotemporal behavioral locus data into easy to understand semantic data. Spatiotemporal behavior trajectory data are a typical time series data, which is generally composed of a series of position points that vary with time. The spatiotemporal trajectory data series record the time and location of the social activities of the concerned object, and reflect the complex relationship between the social activities and the environment. In 2015, L Guo (Guo L et.al. 2015) and others, for the first time considering duration as a single dimension in the expression of behavioral patterns for objects of concern, proposed a new way to find common behavior by considering the length of time spent at each location in the trajectory (**DoSTra**). The method can be used to detect groups with similar lifestyles, habits or behavior patterns. In 2020, Zhangqing Shan et al. (Shan Z et.al. 2020) pointed out the problems of semantic back, semantic error and semantic complexity in the expression of human behavior patterns. Second, they elaborate semantic cleansing functions to detect semantic complexity from human mobility. Third, objective data sources are used to avoid semantic bias. Compared with the existing methods, the urban semantic map shows satisfactory validity and accuracy in finding fine-grained semantic patterns. Hancheng Cao, Fengli Xu and others from Tsinghua University in 2020 (Cao H et.al.) propose a habit2vec approach based on representational learning to address the issue of how to accurately construct user preference patterns. This method chooses the POI type most frequently visited by the object of interest in each time slice as a typical habit, and obtains the vector representation of its living habits through habit2vec representation.

To sum up, there are two kinds of ways to express the behavior of the object of concern: one is to express the behavior of the object of concern efficiently by using the space-time relationship. Zhang Qingshan, as the representative, uses the attribute "time, region attribute, staying time, etc." to express the behavior of the object of concern. But this method is not good for the calculation of the distance measure of the behavior pattern. On the other hand, NLP, represented by Hancheng Cao, trains a word embedding model to express the semantic information of user trajectory in vector space, and constructs the user behavior preference pattern.

2.2 Distance measure of behavior pattern

LM Petry, CA Ferrero, LO Alvares, et al., Federal University of Santa Catalina, San Catalina, 2019 (Petry et.al.2019) In the Transactions in GIS Journal (JCR zone 2), an improved MUITAS approach was proposed to address the neglect of attribute relationships in the multidimensional trajectory (MAT) similarity measurement. This method introduces semantic relations between attributes, including independent attributes and dependent attributes, and sets up different distance functions for each attribute. In 2020, Caihong Liu, Chonghui Guo and others of Dalian University of Technology put forward a clustering algorithm of network semantic sequence based on community detection in order to solve the problem that the existing method does not analyze the similarity of trajectories from a global perspective. This algorithm can better measure the semantic similarity of trajectories from the network angle, analyze the global relationship between trajectories, and obtain better clustering effect. In 2018, Xiucheng Li et al. (Li x et.al. 2018) proposed the first deep learning method to learn the robustness of trajectory representation to low data quality, so as to support accurate and efficient trajectory similarity calculation and search. Experimental results show that the proposed method has higher accuracy and is at least an order of magnitude faster than the most advanced k-nearest trajectory search method. In 2021, Ziquan Fang et al. (Fang Z et.al. 2021) Will propose a self-trained end-to-end depth trajectory clustering framework called E2DTC at ICDE, inspired by the data-driven capabilities of deep neural networks. E2DTC does not need any additional manual feature extraction operations, and can easily adapt to any track on the track data cluster analysis, easy to classify different tracks processing. In 2022, an Liu et al. (Liu A et.al. 2020) put forward a time-space trajectory similarity measurement model AT2VEC on TKDE, which integrates NPL and attention mechanism. Firstly, it trains the activity attributes in the dataset to get the vector representation of the activity words, and then trains and tests the autonomous intention AT2VEC by synthesizing the dataset generated by time, space and activity characteristics. The trajectory clustering result of this method is better than that of the existing methods.

To sum up, there are two kinds of methods in the literature, one of which is based on machine learning to analyze the local or global similarity of behavior patterns, such as time, space, type of social activity, etc. On the other hand, the fusion of NLP and attention mechanism, represented by Liu An, has achieved some results in computing the semantic similarity between different types of interest points.Therefore, how to realize the semantic analysis of long-distance spatiotemporal behavior has become a key technical bottleneck in social security behavior analysis.

2.3 Emotional state perception

Early in the research of emotion detection, there are (Padgett et.al. 1996) and (Walter et.al. 2011) emotion detection based on facial data and physiological data. In 2013, R Li Kam Wa et al. (LiKamWa et.al. 2013) used mobile phone data for the first time to analyze human emotional state. By analyzing the history of social software communication and application usage patterns, they can calculate the average daily emotion of users. The initial accuracy rate is 66%. After two months of personalized training, the accuracy rate gradually increases to 93%. In 2016, Anja Exler and others collected data and heart rates from smartphones. Three emotional dimensions — valence, vigor arousal and calmness — were assessed, but valence was considered only in the context of this paper. The correlation between time characteristics and heart rate was determined by analysing information gain. The decision tree classifier achieves recognition accuracy up to 0.91. For the first time in 2020, Eiman Kanjo(Kanjo et.al. 2019) and others used location data to detect and categorize emotions. They trained physiological, environmental, and location data separately on three models, MLP, CNN, and CNN-LSTM, and then combined the data to evaluate the performance of each model. Compared with the environmental data, the location data have higher accuracy. In 2021, Kangning Yang (Yang K et.al. 2021) et al proposed a novel attention -based LSTM system using sensors from smartphones (front-facing cameras, microphones, touch screens) and wristbands (photo plethysmographs, skin electrical activity and infrared thermopile sensors) to accurately determine a user's emotional state, yielding an average positive and negative emotional classification rate of 89.2 per cent.

To sum up, the research on emotion detection is more and more inclined to use the data acquired by portable sensor devices to study emotion. In the past, the research on emotion detection based on accurate physiological signals and facial features is more accurate, but it should not be popularized. Current research on mood detection and mood classification has been developed using data collected by portable sensors (e.g. mobile phones, bracelets, etc.) as input. Using location information collected by mobile phone to detect and classify emotion is also a popular and novel research direction.

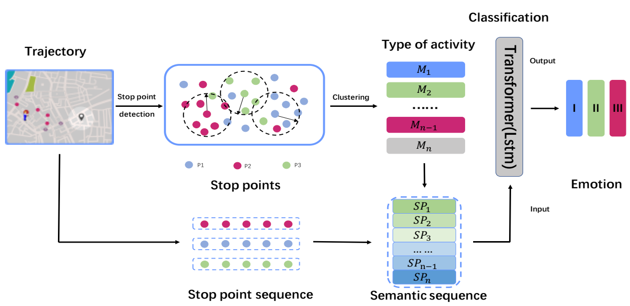


Figure 2: Overall architecture for our proposed RILA model

3 Approach

In this section, we introduce the SADS model, which is an emotional state classification model based on the analysis of behavioral patterns. Specifically, our proposed method is introduced as follows, 3.1, which introduces how to discover the stay point from the original trajectory, 3.2, Introduced how to analyze the activity behavior from the stay point and how to understand the meaning of the social attribute level behind the activity, 3.3, introduced how to express the behavior pattern of the object of concern at the social attribute layer, 3.4, described how to use the space-time trajectory to classify emotions. The overall architecture is illustrated in Figure 2.

3.1 Spatiotemporal behavior stop point detection

Individual object spatio-temporal behavior stop point detection. The social activities studied in this project refer to social activities that last for a certain period of time in a certain area. It is assumed that the individual space-time trajectory point can be defined as $p = (l_o, l_a, t)$, and $l_o$and l_aa are longitude and latitude. The kth individual trajectory point set can be formalized as:

$$P_k = \{p_m, p_{m+1}, \ldots, p_n\}$$

At this time, the adjacent trajectory points in P_k satisfy the constraints $D(p_i, p_j) \le D_t$and$T(p_i, p_j) \le T_t$, using clustering can find the center point $C_k$ofP_k , then the i_{th}individual stay area can be expressed as the center point Ci, the activity radius r_i activity area. In addition, the maximum time interval in Pk is taken as the individual stay time$TL_k = (t_m, t_n)$, and the person id of the individual stay point is taken as e. Then, at this time, thei_{th}individual stay region can be expressed formally as: $S_i = (C_i, r_i, t_m, t_n, e)$

Multi-object single spatio-temporal social activity detection. The above definition of social activity. After determining the individual stay area S, all two-object single spatio-temporal social activities in the set S can be derived by judging two-by-two in the set S. The k_{th} twoobject single spatio-temporal social activity can be formally expressed as: $U_k = (S_i, S_j)$ At this time, the two individual stay are asS_i, S_j in U_k satisfy the constraints $D(S_i, S_j) \le D_{th}$ and $\mathrm{T(S_i)} \cap \mathrm{T(S_j)} \ge \mathrm{T_{th}}$. Take the center point of the line connecting C_i and C_j in S_i and S_j as the center point of U_k CE_k, and the radius of activity is r_k. Take the concatenation of the two individual stay times in U_k as the time space The social activity time is written as $TU_k = TL_i \cup TL_j = (t_{k_{begin}}, t_{k_{end}})$, then, at this time, the kth two-object single spatiotemporal social activity can be expressed formally as:

$$W_k = (CE_k, r_k, t_{k_{begin}}, t_{k_{end}}, \mathrm{CC})$$

CC is the set of two-object ids, and subsequently all the two-object single social activities in W are merged to obtain the multi-object single social activity, which is represented formally as:$Y_k = \{W_\mathrm{m}, W_{\mathrm{m}+1}, \ldots, W_n\}$.

At this time, any two single social activities W_i and W_j of two objects in Y_k satisfy the constraints $\mathrm{D(W_i, W_j)} \le \mathrm{D_{the}}$ and $\mathrm{T(W_i)} \cap \mathrm{T(W_j)} \ge \mathrm{T_{the}}$. At this time, take the concatenation of people in Y_k as EC, take the central set of all individual stay points in Y_k for clustering, then we can get the central point of Y_k CR_k, the radius of Y_k The activity radius of Y_k is R_k. Take the concatenation of all individuals'

stay time in U_k as the spatio-temporal social activity time as $TP_k = TU_m \cup TU_{m+1} \cup \ldots \cup TU_n = (t_{k_{be}}, t_{k_{en}})$. Then, at this point, the kth multi-object spatiotemporal social activity can be expressed formally as:
$WE_k = (CR_k, R_k, t_{k_{be}}, t_{k_{en}}, \text{EC})$

3.2 Social activity types and their tag library creation

Social activity types and their labeling database are created by clustering techniques to classify multiple single social activity samples into different types. The "social activity" here is mainly composed of three elements (activity location/region, activity time, and activity person). Different types of activities are closely related to the area where they are located, and the main focus here is on the social attributes of the location area, rather than on specific latitude and longitude points. So it is necessary to characterize the activity area according to the semantics of area attributes. This part provides a sample that can divide the location of regional activities in schools into six dimensions, such as rest area, living area, learning area, sports area, recreation area, and experimental meeting area, etc. Each different application scenario can make necessary adjustments to the above division according to the different characteristics of their own application areas. Social activities located at different time periods may imply differences in the nature of the activities. In this project, activity time is divided into two dimensions, activity duration Q_{t_1} and activity time Q_{t_2},then activity duration Q_{t_1} can be defined as:$Q_{t_1} = t_{k_{end}} - t_{k_{begin}}$。

We only collect the trajectory data of the tested subjects from 8:00 a.m. to 23:00 p.m. The activity time is represented by Q_{t_2},then the vector of activity time dimension can be expressed as $Q_t =$ （Q_{t_1}， Q_{t_2}） The vector of hidden social activities generated from the above vectorization of social activities is represented as Q. This section clusters the social activities using a graphical model clustering scheme. Assuming that the relationship between samples within Q is modeled as a graph $G = (V, E)$, where V is the set of points and E denotes the relationship between samples, dividing the graph into disjoint individual subgraphs yields the classification result, i.e., minimizing the loss function $Cut(G_1, G_2, \ldots, G_r) = \sum_{i \in G_1, \ldots\ldots k \in G_r} w_{ij\ldots\ldots k}$, where r is the category of classification and w is the similarity between individual sample points.

Assuming that g categories are co-classified from the social activity clustering analysis, the social activities can be sorted based on the number of social activities in each category, which are defined as 1,2,3 … … g type social activities in order of higher number, noted as $M_1, M_2, M_3, \cdots, M_g$.

Table 1: Dataset Analysis Situation Analysis

Dataset	Available	trajectory	Emotional labels
Geolife	√	√	×
Grab-posisi	√	√	×
EnvBodySens	×	√	√
SAPD22111510	√	√	√
SAPD23031530	√	√	√

3.3 Expression of the social activity of the individual object

After establishing a library of social activity types, a sequence of spatio-temporal trajectories for a specific object, which spatio-temporal behavior can be expressed in a semantic sequence of social activity types as follows:

$$Z =< M_1, M_2, M_3, \cdots, M_b >$$

where M_i denotes the social activity type of the ith activity of the object in that observation period, and b denotes the number of activity data generated by the current individual object.

3.4 Emotion classification

In this study, we transformed the emotional state detection problem based on the spatio-temporal trajectory data of the concerned object into a classification problem based on the social activity semantic sequence

$$Z =< M_1, M_2, M_3, \cdots, M_b >.$$

Base on Transformer.The model consists of an embedding layer, a position encoding layer, a transformer encoder, and a fully connected layer. The model parameters are given in 4.3. First, the input short-term social activity type semantic sequence is encoded through the embedding layer and the position encoding layer to obtain a (32, n, 64), and then use the encoder of the transformer to extract the behavioral features in the sequence, and encode it into a fixed-length vector representation, and then we average the encoded vectors of all its positions to get A tensor of shape (32, 64), and finally we classify the average vector through a fully connected layer to obtain a tensor of shape (32, 3), where each element represents the sequence belongs to each category Probability, to realize the classification of emotions corresponding to the behavior data of the measured object in a certain period of time.

4 Experiments

4.1 Datasets

There are 5 datasets involved in this paper. The conditions for using the space-time trajectory of the object of interest to carry out emotion classification research are met as shown in Table 1. The geolife and Grab-positi datasets are open source and are spatio-temporal trajectory data, but there are no emotional labels in the dataset. It cannot be used as a dataset for emotional classification research using the temporal and spatial trajectories of the objects of interest. EnvBodySens

Table 2: Introduction to Dataset

Dataset	SAPD22111510	SAPD23031530
Period	10day	1month
number of people	15	15
Data size	346152 items	1016376 items
Sex ratio	10:5	9:6
Positive:neutral:passive	21.2:60.4:18.4	18.6:51.4:29.0
collection time	8-23 o'clock	8-23 o'clock

The dataset is not open source and cannot be used as a dataset for emotional classification research using the temporal and spatial trajectories of the objects of interest. The current public spatiotemporal trajectory datasets at home and abroad cannot satisfy our method to verify the effectiveness of the emotion classification method based on behavioral pattern analysis. Therefore, We plan to build our own pedestrian trajectory datasets SAPD22111510 and SAPD23031530 containing emotional labels. All experiments in this paper are carried out on the basis of datasets SAPD22111510 and SAPD23031530.

SAPD22111510, SAPD23031530: This project recruited 15 volunteers to participate in the collection of datasets SAPD22111510 and SAPD23031530. Volunteers' trajectories and emotional indexes ("emotional indexes are divided into: positive, normal, and negative") were uploaded to the Locate app-location&emotion for collection. We did not intervene in any data volunteers. Volunteers are distinguished by unique ids to allow them to carry out daily activities in their daily state. The dataset SAPD22111510 will be collected in November 2022 for a period of 10 There are 168 days of GPS trajectories and a total of 346,152 pieces of data The dataset collected SAPD23031530 in March 2023, a period of about 30 days of daily campus activity trajectories, a total of 488 days of GPS trajectories, and a total of 1,016,376 pieces of data. Volunteers use the location&emotion Android program to perform background positioning every 10s, and store the longitude and latitude, current time, and POI information of the current location of the volunteers into the database.

The detailed indicators of the SAPD22111510 and SAPD23031530 datasets are shown in Table 2

4.2 BaseLines

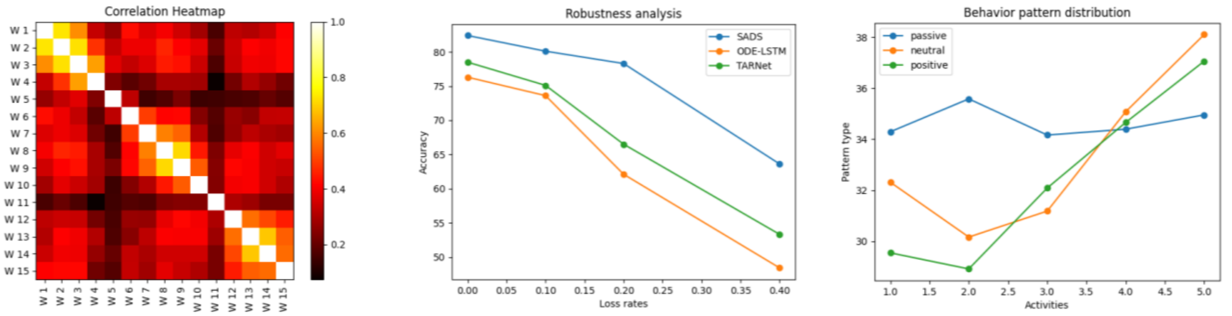
We compare our proposed model with the following two baselines. The first baseline proposes TARNet [32], a new model that uses Transformers to learn task-aware data reconstruction. The authors design a data-driven masking strategy using the distribution of self-attention scores from terminal task training. The timestamps considered important by the terminal task are sampled. The data at these timestamps are then masked and reconstructed, making the reconstruction task-aware, allowing the reconstruction of learned representations to improve the performance of the final task, and enabling classification of time-series data.The second baseline is that we use the open source literature dealing with trajectory classification as our second type of comparison baseline, which takes time series data as input. ODE-LSTM, this article designs a new algorithm based on long-term short-term memory (LSTM), which separates its memory from the time-continuous state, so that the continuous-time dynamic flow can be encoded in the RNN, enabling it to arrive at responds to inputs with arbitrary time-lags while ensuring constant error propagation through the storage path and enabling classification of time-series data.

4.3 Implementation Details

We use Adam as the optimizer. The learning rate of LSTM is set to 1e-4, the maximum number of epochs and batch size are set to 200 and 32 respectively, for the super parameters of transformer, learning_rate is 1e-4, batch_size is 32, dropout is 0.1, epoch is 200, The number of encoder layers is 4, the number of heads is 8, and the hidden layer size is 128.
We implement our model using PyTorch. We train our models on NVIDIA RTX3090 GPUs.

4.4 Experimental Results

We conducted comparative experiments on two datasets to compare the results of the two baselines on the two datasets. We found that our SADS model achieved better results compared to all baseline models, and we will It is attributed to the further information extraction and social attribute level analysis of the stay point data containing multi-semantic information. We explored the influence of social factors on the emotions of the object of concern, and thus found the social attribute level behind the behavior of the object of interest. meaning and rules. Finally, we classify and predict their emotions by analyzing the changes in their behavior patterns at the level of social attributes. Compared with the current literature method for sentiment analysis based on spatiotemporal trajectories, better sentiment classification results are



(a) (b) (c)

Figure 3:Correlation analysis of behavior and emotion

obtained, as shown in Table 3, which also proves the effectiveness and advancement of our method.

Table 3: Results on the dataset

Dataset	SAPD22111510		SAPD23031530	
	Acc	F1	Acc	F1
SADS	82.4%	0.736	65.35%	0.512
ODE-LSTM	76.3%	0.661	55.65%	0.421
TARNet	78.6%	0.673	61.32%	0.483

4.5 Additional Results

Short-term behavior pattern correlation analysis. In order to study the correlation between each time window in the short-term behavior pattern sequence, we conducted a correlation analysis on the short-term behavior patterns between the time windows within the collection period of 8-23 o'clock, and the correlation analysis results are shown in Figure 3a, The results show that w1-5, w6-10, and w11-15 are significantly correlated. Therefore, in the follow-up experiments, the length of the track to be tested is 5 hours, and the short-term behavior of the attention object in different emotional states Whether there is a difference in the pattern was investigated.

Significant difference analysis. Verify the difference in the distribution of short-term behavior pattern sequences of attention objects under different emotional conditions, under the assumption that: H0: the distribution of short-term behavior pattern sequences of attention objects under different emotional conditions is consistent; H1: attention objects are under different emotional conditions The distribution of short-term behavior pattern sequences is not consistent; this paper extracts 20 short-term behavior pattern sequences of objects of interest in different emotional states for comparison, and the distribution function image of short-term behavior patterns is shown in the figure 3c. Through calculation, the P values of the ANOVA analysis of variance obtained in the two datasets are SAPD22111510: 0.0067, SAPD23031530: 0.0052, respectively, which are less than the significance level of 0.05, which proves that the original hypothesis is established and confirms that in the dataset used in this paper, The sequence distribution of the short-term behavior pattern of the attention object under different emotional conditions is not consistent, and there are significant differences.

Robustness analysis. We analyzed the robustness of the model on the SAPD22111510dataset. We randomly sampled 10% of the trajectory points in each trajectory, let the original trajectory be defective, and then looked at the impact of data defects on the performance of the model to study their robustness. To ensure the rigor of the data, we have done this experiment. We extracted 10%, 20%, 40%, and 60% of the data to conduct the experiment to check the impact of missing data on the performance of the model. The structure is shown in Figure 3b. In the case of 20% loss, ode-lstm and TARNet showed a significant drop in accuracy, while sdas only showed a significant drop at 40%, but the drop was smaller than the benchmark literature. This shows that sdas exhibits good robustness in the presence of missing trajectory data. In addition, sdas also shows better performance in noisy environments such as positioning offsets.

5 Conclusion

This paper designs a model that includes the functions of object-of-interest activity detection, activity classification and understanding, behavior pattern expression, and emotion perception to complete the modeling of the object-of-interest trajectory and emotional state mapping relationship. It has the ability to discover the hidden activities of the objects of interest, understand and classify the hidden activities, and then more accurately express the behavior of the objects of interest at the model layer, thereby discovering the behavior pattern mechanism behind the trajectory of the objects of interest and exploring the emotional index of the objects of interest are closely related to multidimensional patterns of behavior. In the future, we plan to explore the influence of different characteristics of characters on emotions, and further explore the factors behind emotions, such as social status, education, income, etc.

References

Fox, Elaine. Emotion Science: *An Integration of Cognitive and Neuroscientific Approaches*. New York: Palgrave MacMillan. 2008: 16-17. ISBN 0230005179.

Schacter, D.L., Gilbert, D.T., Wegner, D.M., & Hood, B.M. (2011). *Psychology* (European ed.). Basingstoke: Palgrave Macmillan.

E. Kanjo, E.M. Younis, N. Sherkat.*Towards unravelling the relationship between on-body, environmental and emotion data using sensor information fusion approach*.Inf. Fusion, 40 (2018), pp. 18-31

E. Kanjo, D.J.DJ. Kuss, C.S. Ang *NotiMind: utilizing responses to smart phone notifications as affective sensorsNotiMind*: IEEEAccess (2017), 10.1109/ACCESS.2017.2755661

S. Jerritta, M. Murugappan, R. Nagarajan, K. Wan, *Physiological signals based human emotion recognition*: a review, in: Proceedings of the IEEE Seventh International Colloquium on Signal Processing and its Applications (CSPA), IEEE, 2011, pp. 410–415. IEEE，2011,410-415

C. Busso, Deng Z., S. Yildirim, M. Bulut, Lee C.M., A. Kazemzadeh, Lee S., U. Neuman, S. Narayanan *Analysis of emotion recognition using facial expressions*, speech and multimodal information Proceedings of the Sixth International Conference on Multimodal Interfaces, ACM (2004), pp. 205-211

E. Kanjo, A. Chamberlain.*Emotions in context: examining pervasive affective sensing systems*, applications, and analyses.Pers. Ubiquitous Comput. (2015), pp. 1-16

F. Agrafioti, D. Hatzinakos, A.K. Anderson.Ecg *pattern analysis for emotion detection*.IEEE Trans. Affect. Comput., 3 (1) (2012), pp. 102-115

N. D. Lane, E. Miluzzo, H. Lu, D. Peebles, T. Choudhury and A. T. Campbell, *"A survey of mobile phone sensing"*, IEEE Commun. Mag., vol. 48, no. 9, pp. 140-150, Sep. 2010.

LiKamWa R, Liu Y, Lane N D, et al. Moodscope: Building a mood sensor from smartphone usage patterns[C]//Proceeding of the 11th annual international conference on Mobile systems, applications, and *services*. 2013: 389-402.

Exler A, Schankin A, Klebsattel C, et al. A wearable system for mood assessment considering smartphone features and data from mobile ECGs[C]//Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: Adjunct. 2016: 1153-1161.

Guo L, Huang G, Gao X, et al. DoSTra: Discovering common behaviors of objects using the duration of staying on each location of trajectories[C]//Workshops at the twenty-ninth AAAI conference on artificial intelligence. 2015.

Shan Z, Sun W, Zheng B. Extract human mobility patterns powered by city semantic diagram[J]. IEEE Transactions on Knowledge and Data Engineering, 2020.

Cao H, Xu F, Sankaranarayanan J, et al. Habit2vec: Trajectory semantic embedding for living pattern recognition in population[J]. IEEE Transactions on Mobile Computing, 2019, 19(5): 1096-1108.

Petry L M, Ferrero C A, Alvares L O, et al. Towards semantic-aware multiple-aspect trajectory similarity measuring[J]. Transactions in GIS, 2019, 23(5): 960-975.

Liu C, Guo C. STCCD: Semantic trajectory clustering based on community detection in networks[J]. Expert Systems with Applications, 2020, 162: 113689.

Li X, Zhao K, Cong G, et al. Deep representation learning for trajectory similarity computation[C]//2018 IEEE 34th international conference on data engineering (ICDE). IEEE, 2018: 617-628.

Fang Z, Du Y, Chen L, et al. E 2 dtc: An end to end deep trajectory clustering framework via self-training[C]//2021 IEEE 37th International Conference on Data Engineering (ICDE). IEEE, 2021: 696-707.

Liu, A., Zhang, Y., Zhang, X., Liu, G., Zhang, Y., Li, Z., ... & Zhou, X. (2020). Representation learning with multi-level attention for activity trajectory similarity computation. IEEE Transactions on Knowledge and Data Engineering.

Padgett C, Cottrell G. Representing face images for emotion classification[J]. Advances in neural information processing systems, 1996, 9.

Walter S, Scherer S, Schels M, et al. Multimodal emotion classification in naturalistic user behavior[C]//International Conference on Human-Computer Interaction. Springer, Berlin, Heidelberg, 2011: 603-611.

LiKamWa R, Liu Y, Lane N D, et al. Moodscope: Building a mood sensor from smartphone usage patterns[C]//Proceeding of the 11th annual international conference on Mobile systems, applications, and services. 2013: 389-402.

Exler A, Schankin A, Klebsattel C, et al. A wearable system for mood assessment considering smartphone features and data from mobile ECGs[C]//Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: Adjunct. 2016: 1153-1161.

Kanjo E, Younis E M G, Ang C S. Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection[J]. Information Fusion, 2019, 49: 46-56.

Yang K, Wang C, Gu Y, et al. Behavioral and physiological signals-based deep multimodal approach for mobile emotion recognition[J]. IEEE Transactions on Affective Computing, 2021.

Yuxuan Liang, Kun Ouyang, Yiwei Wang, Xu Liu, Hongyang Chen, Junbo Zhang, Yu Zheng, and Roger Zimmermann. 2022. TrajFormer: EfficientTrajectory Classification with Transformers. In Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM'22), October 17–21, 2022, Atlanta, GA, USA. ACM, New York, NY, USA,9 pages. https://doi.org/10.1145/3511808.3557481

Han X, Cheng R, Ma C, et al. DeepTEA: effective and efficient online time-dependent trajectory outlier detection[J]. Proceedings of the VLDB Endowment, 2022, 15(7): 1493-1505.

Li D, Li W, Wang X, et al. App trajectory recognition over encrypted internet traffic based on deep neural network[J]. Computer Networks, 2020, 179: 107372.

Y. Zheng, H. Fu, X. Xie, W.-Y. Ma, and Q. Li, Geolife GPS trajectory dataset, Geolife GPS trajectories 1.3 ed., August 2012. [Online]. Available: https://www.microsoft.com/enus/research/publication/geolife-gps-trajectory-dataset-user-guide/

Huang X, Yin Y, Lim S, et al. Grab-posisi: An extensive real-life gps trajectory dataset in southeast asia[C]//Proceedings of the 3rd ACM SIGSPATIAL international workshop on prediction of human mobility. 2019: 1-10.

Mathias Lechner and Ramin Hasani. Learning long-term dependencies in irregularlysampled time series. Advances in neural information processing systems, 2020.

Chowdhury R R, Zhang X, Shang J, et al. Tarnet: Task-aware reconstruction for time-series transformer[C]//Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. 2022: 212-220.