

Received 25 May 2022, accepted 23 June 2022, date of publication 30 June 2022, date of current version 8 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3187502

RESEARCH ARTICLE

Multimodal Educational Data Fusion for Students' Mental Health Detection

TENG GUO¹, WENHONG ZHAO², (Member, IEEE), MUBARAK ALRASHOUD³,
AMR TOLBA^{4,5}, (Senior Member, IEEE), SELENA FIRMIN⁶,
AND FENG XIA⁶, (Senior Member, IEEE)

¹School of Software, Dalian University of Technology, Dalian 116620, China

²Ultrprecision Machining Center, Zhejiang University of Technology, Hangzhou 310014, China

³Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh 11437, Saudi Arabia

⁴Department of Computer Science, Community College, King Saud University, Riyadh 11437, Saudi Arabia

⁵Department of Mathematics and Computer Science, Faculty of Science, Menoufia University, Shibin El Kom 32511, Egypt

⁶Institute of Innovation, Science and Sustainability, Federation University Australia, Ballarat, VIC 3353, Australia

Corresponding author: Wenhong Zhao (whzhao6666@outlook.com)

This work was supported by the International Scientific Partnership Program (ISPP) at King Saud University under Grant ISPP-78.

ABSTRACT Mental health issues can lead to serious consequences like depression, self-mutilation, and worse, especially for university students who are not physically and mentally mature. Not all students with poor mental health are aware of their situation and actively seek help. Proactive detection of mental problems is a critical step in addressing this issue. However, accurate detections are hard to achieve due to the inherent complexity and heterogeneity of unstructured multi-modal data generated by campus life. Against this background, we propose a detection framework for detecting students' mental health, named CASTLE (educational data fusion for mental health detection). Three parts are involved in this framework. First, we utilize representation learning to fuse data on social life, academic performance, and physical appearance. An algorithm, named MOON (multi-view social network embedding), is proposed to represent students' social life in a comprehensive way by fusing students' heterogeneous social relations effectively. Second, a synthetic minority oversampling technique algorithm (SMOTE) is applied to the label imbalance issue. Finally, a DNN (deep neural network) model is utilized for the final detection. The extensive results demonstrate the promising performance of the proposed methods in comparison to an extensive range of state-of-the-art baselines.

INDEX TERMS Educational data mining, mental health, network embedding, decision making.

I. INTRODUCTION

Mental health is an important component of overall health and well-being. Mental health issues can lead to serious consequences, such as self-mutilation and suicide, particularly among university students who are not yet physically and mentally mature [1]–[5]. Currently, youth mental health is deteriorating. According to The State Of Mental Health In America 2021, 9.7 percent of youth in the United States have severe major depression, up from 9.2 percent in the previous year's data [6]. The situation is even worse in developing countries [2]. Research also demonstrated that young people

are more likely than any other age group to experience moderate to severe anxiety and depression during the COVID-19 pandemic [7]–[9]. However, not all students who have mental health issues are aware of their situation and actively seek help. According to research, roughly three-quarters of college students are hesitant to seek help when they have a mental health problem [10]. In this context, a proactive detection system for students with mental health problems is the key to addressing this issue.

However, detecting these students proactively is a tremendous challenge because mental health is influenced by a variety of complex factors. Previous studies have demonstrated that social life [11], [12], academic performance [13], physical appearance [11], and demographic features [4] can all

The associate editor coordinating the review of this manuscript and approving it for publication was Laura Celentano¹.

have an impact on students' mental health, and these features are recorded by the unstructured multi-modal data generated from various systems. Social life records, for example, belong to graph or network data, whereas physical appearance belongs to image data. It is difficult to accurately represent and fuse these features. The most commonly used methods in the current study are manual scoring and calculating summary statistics (e.g., the mean or variance). For example, researchers use Grade-Point Average (GPA) as a proxy for students' academic performance or they manually assign a score to people's physical appearance [11], [14]. It is undeniable that these approaches are capable of effectively overcoming data complexity and heterogeneity. However, it is not only subject to human bias, but it also risks losing information when data distribution varies greatly. For example, two students with the same GPA may take completely different courses. In this case, an effective solution to represent and fuse these features is a crucial challenge for detecting students with mental health problems.

The rapid development of technologies such as network science and representation learning [5], [15]–[17] enables us to profile campus life accurately and effectively, which brings us an unprecedented opportunity for the detection of students' mental health [18]–[21]. Opportunities and challenges, however, coexist. Several significant issues continue to impede researchers in related fields. First, previous studies have utilised the friendship network to represent students' social life [22], [23]. However, in addition to friendship, social life includes a variety of scenarios such as information dissemination and seeking assistance [12]. Representing students' social life comprehensively remains a significant challenge. Secondly, GPA does not accurately represent a student's academic performance, as mentioned in the previous paragraph. Accurate representation of academic performance remains an open question. In addition, because the number of students with mental health issues is much smaller compared to those who are healthy mentally, the related dataset is highly imbalanced, posing another critical challenge for mental health detection.

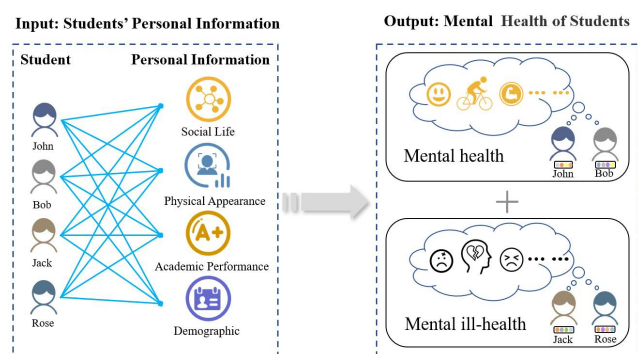


FIGURE 1. The illustration of our research problem.

In this paper, we aim to detect students with mental health problems by overcoming the challenges mentioned above.

We collect data through a combination of questionnaires and the learning management system (LMS), and select students' social life, physical appearance, academic performance, and demographic features to detect their mental health status (shown in Figure 1). As mentioned above, these four features have been demonstrated to be related to the mental state of students [4], [11]–[13]. To profile students' social life comprehensively, we collect social relationships for multiple scenarios, including friendship, life advice, academic advice, support, cooperation, intelligence, and good/bad news sharing [11], [12]. The experimental process of this research is shown in Figure 2. A framework, named **CASTLE** (eduCational dAta fuSion for menTaL hEalth detection), is proposed to achieve an accurate and effective detection, which includes three parts. First, representation learning is used here for the effective fusion of students' multi-modal information, including multi-view social network embedding, physical appearance representation, and academic performance representation. In this part, we propose **MOON** (Multi-view SOcial NetwORk EmbedDiNg), a multi-view social network embedding algorithm, to embed students' heterogeneous social relations effectively. Moreover, a convolutional neural network (CNN)-based auto-encoder is used to embed students' ID photos to obtain an accurate representation of their physical appearance. In addition, we introduce the method of combining the variant of one-hot encoding and autoencoder to overcome the heterogeneity of students' academic performance [24]. Second, a synthetic minority oversampling technique (SMOTE) algorithm is used to mitigate the effects of label imbalance. Finally, a deep neural network (DNN) model is used for the final detection.

Our contributions can be summarized as follows:

- We propose a detection framework **CASTLE** for students' mental health detection through fusing multi-modal information generated from campus life.
- We design a multi-view social network embedding algorithm **MOON**, eliminating the information redundancy across different views through simplifying the strategy of federated embedding.
- We conduct comprehensive experiments on a real-world educational dataset, and the extensive results demonstrate the promising performance of the proposed methods in comparison with an extensive range of state-of-the-art baselines.

This paper is organized as follows. In Section II, related work is reviewed. The problem formulation is presented in Section III. In Section IV, the **CASTLE** detection framework is introduced in detail. In Section V, all the data used in this research and its collection process are introduced. In Section VI, we analyze the results of our experiment. We present the discussion and conclusion of our work in Section VII.

II. RELATED WORK

The mental health of students attracts tremendous attention [25]. Scholars explore the laws behind mental health

from different dimensions through statistical analysis theory. Rossin-Slater *et al.* [26] carried out experiments to explore the link between depression and school shootings through youth antidepressant use. They found that exposure to fatal school shootings increases youth antidepressant use. Duckworth and Seligman [27] investigated the self-regulation of eighth-grade students through longitudinal research, thus completing the detection of student achievement. The results indicated that students with poor self-regulation also have poor intelligence. Usher and Curran [28] predicted the influences on the mental health status of Australian university students using a cross-sectional study including an online survey. The results showed that there is a significant positive correlation between mental health and gender, age, health status, physical activity levels, sporting club participation, and social-emotional well-being. Morelli *et al.* [12] carried out an experiment to explore the association between mental traits and social centrality, and found that different mental traits can have different effects on a student's social centrality.

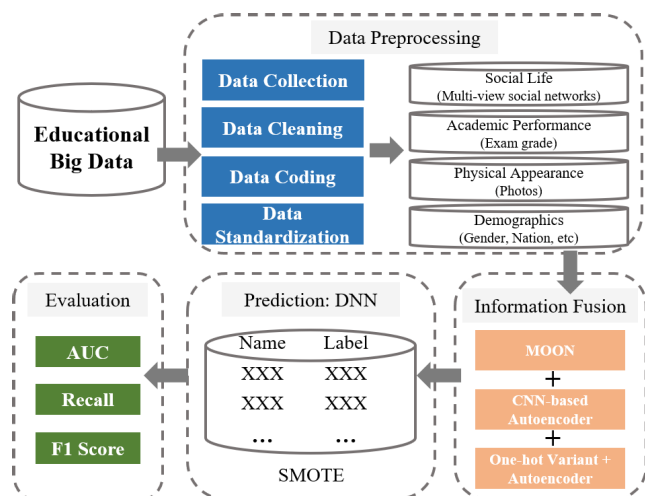


FIGURE 2. The experimental process of this research.

Recently, social behavior has received increasing interest. Gong *et al.* [1] analyzed social anxiety disorders in university students using the research sensor of smartphones. They explored the relationship between social interaction and location, and the changes in social anxiety disorder with changes in location. The results showed that, depending on social anxiety, different students demonstrate large differences in personal behavior. Wongkoblapp *et al.* [29] also concentrated on social behavior. They proposed a detection model based on social networks that can identify participants with poor mental health. The experimental results showed that social network data can effectively detect mental health problems. Meanwhile, Vanlalawmpuia and Lalhmingliana [30] analyzed data from social networking sites. They utilized data mining techniques to identify depression in Facebook users. Moreover, the study identified the number of

depression indicator words, which are significant to related works.

Thanks to the development of big data technology, machine learning algorithms are applied to detect mental health issues [31]–[33]. Unlike previous studies that concentrate on the relationship between a single feature and mental health, machine learning-based research attempts to combine multiple features for prediction or classification. Brathwaite *et al.* [34] validated an existing model for the prediction of mental health in Nigeria. They selected 11 predictors and collected data through questionnaires, such as biological sex, childhood maltreatment, school failure, social isolation, fights, running away from home, and drug use. Tate *et al.* [35] designed a model to predict mental health problems in mid-adolescence and extract features from parental reports and register data such as the National Patient Register and the Multi-Generation Register. Walsh *et al.* [36] aimed to explore the law behind the nonfatal suicide attempts of adolescents and applied the random forest to make predictions. The features they used include diagnostic, demographic, medication, and socioeconomic factors. Ge *et al.* [37] carried out experiments to predict student mental status during the COVID-19 epidemic. They used early psychometric test results to predict future test results through the Xgboost algorithm and achieved good performance. Rubaiyat *et al.* [38] analyzed the predictability of major disorders based on a series of psychological tests, such as internet addiction, depression, and low self-esteem. They used machine learning methods to create models for detection. The experimental results showed that different disorders are interconnected.

Generally, these studies are mainly based on structured data such as demographic information. However, some important pieces of information, such as social patterns and appearance, are stored in unstructured data that is difficult to quantify and represent. With the development of representation learning and deep learning, some scholars are trying to introduce the information hidden in these unstructured data into mental health prediction. Oyebode *et al.* [32] introduced processing techniques for unstructured data to capture more complex features. They made predictions for mental health by performing sentiment analysis on 88125 user reviews. Mathur *et al.* [39] designed an experiment for suicidal tweet detection through natural language processing technology. Gaur *et al.* [40] incorporated domain-specific knowledge and proposed a detection framework to determine the severity of suicide risk-based data from Reddit. Cai *et al.* [41] proposed a multi-modal fusion-based depression recognition system based on EEG (Electroencephalogram) data, and the results showed that their system provides a highly flexible technique for depression identification. The data used for all relevant studies are summarized in Table 1. It can be seen that some scholars are trying to mine information from unstructured data to improve the performance of prediction. In a word, mental health prediction based on unstructured data is becoming a hot topic in this field.

TABLE 1. Summary of the data used in the current study.

Research	Data used for prediction	Data Type
Brathwaite et al. [34]	They collected the biological sex, childhood, maltreatment, school failure, social isolation, fights, and drug use of the experimental participants utilizing questionnaires.	Structured data
Tate et al. [35]	They used participants' parental reports and register data as the features for prediction.	Structured data
Walsh et al. [36]	They made predictions based on the routinely collected clinical data: diagnostic, demographic, medication, and socioeconomic factors.	Structured data
Ge et al. [37]	They integrated information attained at the early stage to predict probable anxiety and probable insomnia in the future.	Structured data
Rubaiyat et al. [38]	They collected the results of a series of participants' psychological tests to predict their mental status in the future.	Structured data
Oyebode et al. [32]	They performed sentiment analysis of 88125 user reviews and based on this, they predicted users' mental health status.	Unstructured data
Mathur et al. [39]	They designed an experiment for suicidal tweet detection through natural language processing technology.	Unstructured data
Gaur et al. [40]	They collected Reddit posts to predict the severity of the suicide risk of an individual.	Unstructured data
Cai et al. [41]	They collected EEG data of participants for depression recognition.	Unstructured data

III. PROBLEM FORMULATION

In this section, we introduce notations and formally define the research problem of this work. In our research, we use a multi-view network to profile students' heterogeneous social relations. A multi-view network consists of a set of nodes U and a set of views V , where each view $v \in V$ contains a set of edges $E^{(v)}$ between nodes involved in U . Specifically, an edge $e_{ij}^{(v)} \in E^{(v)}$ links nodes $i, j \in U$. In this case, the multi-view network is denoted as $G = (U, V, \{E^{(v)} : v \in V\})$. Given the multi-view network, the social life of student i is represented as a low-dimension vector $\mathbf{f}_i \in \mathbb{R}^D$ through embedding the corresponding node i , where D is the dimension of the embedding space. Besides, for student i academic performance, physical appearance, and demographic features are used as features and are represented by $\mathbf{a}_i \in \mathbb{R}^P$, $\mathbf{c}_i \in \mathbb{R}^K$ and \mathbf{o}_i , respectively, where P and K are the dimensions of the embedding space. Moreover, the mental status of student i denoted as $y_i \in \{0, 1\}$ and according to the results of the mental status test, we divide students into two groups: mentally healthy students and mentally unhealthy students.

Mental Status Detection: For student i , given the features mentioned above, we detect its mental health status y_i .

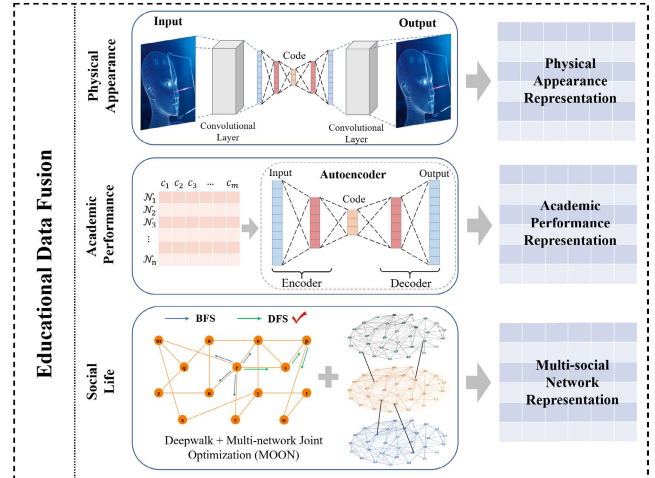
TABLE 2. Summary of main notations.

Notation	Definition
U	the set of nodes U
V	the set of views V
v_0	the source view for social network (friendship network)
V'	the set of targeted views of social network
$E^{(v)}$	the set of edges $E^{(v)}$ in view v
$i^{(v)}, j^{(v)}$	the root node i and its context j in the view v
$\Omega^{(v)}$	the set of intra-view pairs in view v
$\mathbf{f}_i^{(v)}, \tilde{\mathbf{f}}_i^{(v)}$	center and context embeddings of node i for view v
α, β	hyperparameters to balance loss components

IV. DESIGN OF CASTLE FRAMEWORK

The CASTLE framework, including educational data fusion, data augmentation, and the detection model, is introduced in this section. First of all, we introduce educational data fusion, consisting of three subparts (shown in Figure 3): multi-view

network embedding, physical appearance representation, and academic performance representation. First, the MOON algorithm proposed in this paper is introduced for multi-view social network embedding. Second, we represent students' physical appearance and academic performance accurately and effectively through a CNN-based auto-encoder and an auto-encoder based algorithm [24]. Moreover, in the data augmentation part, we use the SMOTE algorithm to generate the data of students with mental health problems to balance the dataset. Finally, a DNN with the dropout mechanism is utilized for the final detection.

**FIGURE 3.** The illustration of educational data fusion.

A. EDUCATIONAL DATA FUSION

1) SOCIAL LIFE REPRESENTATION

As mentioned before, students' heterogeneous social relations are represented by a multi-view network in this paper. In other words, eight networks are used to represent the social relationships of experimental participants in eight different scenarios. Eight social scenarios include friendship, life advice, academic advice, support, cooperation, intelligence, and good/bad news sharing (details are shown in Section V-D). Inspired by [42], we propose a multi-view

social network embedding algorithm, named MOON, to embed social information of students in multiple social scenarios into low-dimensional dense vectors. The details are shown as follows.

Ata *et al.* [42] proposed concepts of first-order and second-order collaboration and experiments demonstrate that its embedding performance is better than other state-of-the-art algorithms. In this case, we introduce these concepts into our experiments and divide the node pairs into three categories:

- **Intra-view Pairs:** Two nodes are linked in the single-view network. The representation can be generated through random walks in each single-view network to retain the diversity of different views.
- **Cross-view, intra-node Pairs:** The same node (i.e., intra-node) across two views (i.e., cross-view) forms pairs. Through the alignment process of node representations in such a pair, we aim to capture the first-order social relations.
- **Cross-view, cross-node Pairs:** A node in one view forms pairs with various nodes (i.e., cross-node) in another different view (i.e., cross-view) based on their associations in each view, to obtain the second-order social relationship.

Details of these pairs are shown as follows.

First, each node has a view-specific representation to retain the diversity of each view. The embedding process follows the Deepwalk model [43], i.e., it generates topologically associated node pairs from random walks [44]. For a certain view $v \in V$, a set of intra-views pairs is denoted as $\Omega^{(v)} \subset U \times U$. According to the mechanism of the skip-gram model, a pair $i^{(v)}, j^{(v)} \in \Omega^{(v)}$ consists of a root node $i^{(v)}$ and a context node $j^{(v)}$. The prediction task is that uses the root node $i^{(v)}$ to predict the context node $j^{(v)}$, i.e., to maximize $P(j^{(v)} | i^{(v)})$. The loss function of this step about parameters Θ is:

$$L_{\text{Div}}(\Theta) = - \sum_{v \in V} \sum_{(i^{(v)}, j^{(v)}) \in \Omega^{(v)}} \log P(j^{(v)} | i^{(v)}; \Theta) \quad (1)$$

$P(j^{(v)} | i^{(v)}; \Theta)$ is defined based on a softmax function:

$$P(j^{(v)} | i^{(v)}; \Theta) = \frac{\exp(\tilde{\mathbf{f}}_j^{(v)} \cdot \mathbf{f}_i^{(v)})}{\sum_{u \in U} \exp(\tilde{\mathbf{f}}_u^{(v)} \cdot \mathbf{f}_i^{(v)})} \quad (2)$$

where $\mathbf{f}_i^{(v)}, \tilde{\mathbf{f}}_i^{(v)} \in \mathbb{R}^{|D|/|V|}$ represent the embedding vector of the center node i and its context, respectively, for view v . Θ represents the model parameter.

However, for cross-view joint embedding, the current research strategy is to unite all networks indiscriminately without capturing the essential characteristics of different networks [42]. This strategy may lead to high computational overhead, especially when there are too many networks involved (for example, our dataset contains eight networks representing different social scenarios). In this case, to achieve more effective embedding, we design the following embedding strategy for multi-view social networks. Social

relations have an intrinsic nature that **most social relations are based on friendships**. Taking life advice as an example, students only seek help from their friends when they are in trouble. To capture this law, we divide all views of the social network into two categories:

- **Source View:** Based on the assumption mentioned above, the source view is friendship.
- **Target View:** Except for friendship, the rest of social relations are target views.

Based on these definitions, we define the **First-order social relation** and the **Second-order social relation**.

a: FIRST-ORDER SOCIAL RELATION

Cross-view, intra-node Pairs. The intuition here is that the same node across various views represents the same entity, so its view-specific representation should collaborate with one another. In other words, for cross-view intra-node pairs, the two view-specific embeddings of the same node should become similar. For the research question in this paper, each target view is a derivative of the source-view. In this case, the source-view representation should help the target-view representation, i.e., the friendship should impact other views by optimizing the following:

$$\begin{aligned} L_{S1}(\Theta) &= - \sum_{v' \in V'} \sum_{(i^{(v')}, \cdot) \in \Omega^{(v_0)}} \log P(i^{(v')} | i^{(v_0)}; \Theta) \\ &= - \sum_{v' \in V'} \sum_{(i^{(v_0)}, \cdot) \in \Omega^{(v_0)}} \log \frac{\exp(\mathbf{f}_i^{(v')} \cdot \mathbf{f}_i^{(v_0)})}{\sum_{u \in U} \exp(\mathbf{f}_u^{(v')} \cdot \mathbf{f}_i^{(v_0)})} \end{aligned} \quad (3)$$

where v_0 is the source view for social network and v' is the target view for social network. V' denotes the set of target views of social networks.

b: SECOND-ORDER SOCIAL RELATION

Cross-view, cross-node Pairs. For the second-order social relation, the two nodes without a link in the target views may entail a latent relation when they are linked in the source view. It is not appropriate to directly assume a link between them. Instead, we introduce an implicit mechanism to enable such across-view relations: for a node in each target view, its context's context in the source view may contain information that contributes to its embedding. For example, one of your friend's friends is yourself.

For the multi-view social network, the context of nodes in the friendship network should be used to guide their embedding in other views. Formally, $(i^{(v_0)}, j^{(v_0)}) \in \Omega^{(v_0)}$ represents node i and its context j in view v_0 . Given such a relation, we form a cross-view, cross-node pair $(i^{(v')}, j^{(v_0)})$, to close the distance between the representations of $i^{(v')}$ and $j^{(v_0)}$. The loss is shown as follows:

$$L_{S2}(\Theta) = - \sum_{v' \in V'} \sum_{(i^{(v')}, j^{(v_0)}) \in \Omega^{(v_0)}} \log P(i^{(v')} | j^{(v_0)}; \Theta)$$

$$= - \sum_{v' \in V'} \sum_{(i^{(v_0)}, j^{(v_0)}) \in \Omega^{(v_0)}} \log \frac{\exp(\tilde{\mathbf{f}}_j^{(v_0)} \cdot \mathbf{f}_i^{(v')})}{\sum_{u \in U} \exp(\tilde{\mathbf{f}}_u^{(v_0)} \cdot \mathbf{f}_i^{(v')})} \quad (4)$$

The complete loss function of the MOON algorithm is:

$$L = L_{\text{Div}} + \alpha \cdot L_{S1} + \beta \cdot L_{S2} \quad (5)$$

where $\alpha, \beta \geq 0$ are hyperparameters to control the importance among the three components.

2) PHYSICAL APPEARANCE REPRESENTATION

The photos of students are processed by a convolutional auto-encoder in this research to capture the spatial structure characteristics. The classical auto-encoder can be defined as: in each hidden layer, we adopted the following nonlinear transformation function:

$$\begin{aligned} \mathbf{h}_{(2)} &= f(\mathbf{W}_{(2)}\mathbf{h}_{(1)} + \mathbf{b}_{(2)}) \\ \mathbf{h}_{(3)} &= f(\mathbf{W}_{(3)}\mathbf{h}_{(2)} + \mathbf{b}_{(3)}) \\ &\dots \\ \mathbf{h}_{(i)} &= f(\mathbf{W}_{(i)}\mathbf{h}_{(i-1)} + \mathbf{b}_{(i)}), \quad i = 4, 5, \dots, k \end{aligned} \quad (6)$$

where f is the activation function and $\mathbf{W}_{(i)}, \mathbf{b}_{(i)}$ are the transformation matrix and the bias vector respectively. Based on the auto-encoder, we added convolutional layers to extract valid image features. The extraction of convolutional features is defined as follows:

$$\mathbf{h}_{(m)} = f(\mathcal{W}_{(m)} * \mathbf{h}_{(m-1)} + \mathbf{b}_{(m)}) \quad (7)$$

where $*$ represents the convolution operator and \mathcal{W} represents filter weight. The corresponding loss function is as follows:

$$\mathcal{L}(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \quad (8)$$

where \mathbf{x} represents the input and $\hat{\mathbf{x}}$ represents the output (i.e., $\hat{\mathbf{x}} = f(\mathbf{W}\mathbf{x} + \mathbf{b})$, where \mathbf{W} and \mathbf{b} represent the transformation matrix and bias vector of the model, respectively).

3) ACADEMIC PERFORMANCE REPRESENTATION

The heterogeneity of academic performance, caused by the diversity of the student curriculum, is always a challenge in the education field. For example, one student chooses courses A, B, and C and another student chooses courses C and D. The content and number of dimensions vary when using their exam grades as features. The current popular method in this field is to calculate summarizing statistics (e.g., the mean or GPA) as the agent to represent academic performance. It can effectively overcome the heterogeneity of the curriculum, but information loss may be significant when data distribution varies widely. To preserve the completed information of academic performance while overcoming the heterogeneity of grade data, we use the method of combining the variant of one-hot encoding and autoencoder for homogenization [24]. Firstly, we embed their course through one-hot encoding, replacing the 1 with the corresponding exam grade. In this

way, we create the matrix $\mathbf{C} \in \mathbb{R}^{n \times m}$ (shown as follows) where n and m represent the number of students and the number of courses, respectively.

$$\begin{Bmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{Bmatrix}$$

However, the number of courses taken by each student is much less than the total number of courses offered by the university. This issue leads to the severe sparsity of \mathbf{C} . To overcome this problem, we use an auto-encoder (Eq. 6) to reduce the high dimension caused by the previous step and obtain an accurate representation of students' intelligence features.

B. DATA AUGMENTATION FOR LABEL IMBALANCE

The number of students with mental health problems is smaller generally (shown in Section V-B), so the label imbalance problem exists in our experiments. Thus, we utilize the SMOTE algorithm [45] to augment data in order to improve generalization performance. Normal oversampling methods take a simple strategy that copies the sample of the target category, resulting in the law captured by learning models on modified data being too specific. By contrast, the basic idea of SMOTE is to analyze the categories with fewer data and to generate data by the following equation:

$$x_{\text{new}} = x + \text{rand}(0, 1) \times (\tilde{x} - x) \quad (9)$$

where x and \tilde{x} represent two different data in the same category. So, the learning model can capture more general patterns on modified data by the SMOTE algorithm.

C. DETECTION MODEL

In this study, the students' mental health detection task is regarded as a binary classification experiment. A three-layer DNN model, including the input, hidden, and output layers, is used for the final binary detection (shown in Figure 4). The input and output layers serve as nodes to buffer input and output, respectively, and the function of the hidden layer is to fit the relationship between input and output. Before any data has been run through the network, the weights for the three-layer model are random. Through the back-propagation algorithm, all weights are updated according to the laws hidden in the data. A dropout mechanism is applied to overcome the overfitting caused by a relatively small dataset. Dropout is a technique for addressing the overfitting problem of neural network-based models, and the mechanism is to drop units (along with their connections) of neural networks randomly during model training. Meanwhile, we also introduce batch normalization to optimize the training process. (Note that our experiments are carried out in the second semester of the student's university life, so only the first grades are used as the feature. A model for time-series data, like temporal

convolutional network (TCN) [46] or long short-term memory network (LSTM) [47] should be the alternative if more semester grades are involved.)

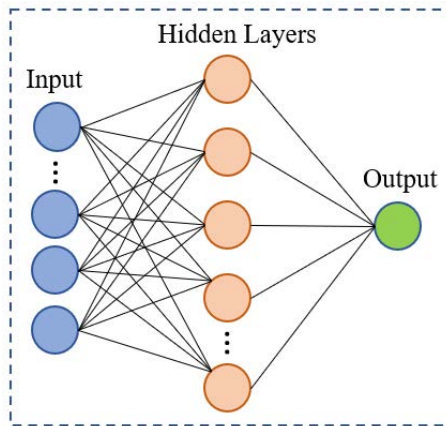


FIGURE 4. The illustration of three-layer model.

V. DATASET

In this section, we detail the data and its collection process. The dataset used in this research includes 509 university students in the same school from a Chinese university and they are freshmen who have just finished their first semester exams. They are required to be more than 18-year-old freshmen (aged 18-20, mean = 19.03, SD = 0.21), who live in several specific residential buildings (next to each other) in the same area. Removing the error data, 485 students are involved in this experiment. First, ethical consideration is introduced. Second, we present the data collected through the LMS (academic performance and demographic features) and the questionnaire (physical appearance and social networks).

A. ETHICAL CONSIDERATIONS

This research has been given ethical approval through the university's ethical approval process. Participants consent to release all related data for the study. Participants are given the option to freely withdraw at any time during the study or omit any particular answers without providing a reason. Participants' pictures and information are kept coded and confidential, and the questionnaire data is kept with separate IDs (e.g., letter code for faces, number code for other data). The key relating codes are kept separately in a password-protected file. The study is not anticipated to cause any distress, but if, for any reason, participants are distressed, they are encouraged to contact the student support service.

B. MENTAL TEST

In this research, we use the Symptom Checklist 90 (SCL-90), a widely used self-report psychometric instrument, to assess mental distress and symptoms of psychopathology [48]. The primary symptom dimensions of SCL-90 consist of total scores of psychological health, somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety,

hostility, phobic anxiety, paranoid ideation, psychoticism, and a category of "additional items" which help clinicians assess other aspects of the clients' symptoms. According to the university student norm, all students are divided into two categories: students with mental health problems and students with healthy mental status. The ratio of students with healthy mental status to students with mental health problems is 7 to 1.

C. ACADEMIC PERFORMANCE AND DEMOGRAPHIC FEATURE

The LMS is the infrastructure of the university, which records information about students' learning and daily life. Students' academic performance can generally be recorded as the exam grade of each course stored in the LMS. The academic performance data used in this research includes 13234 records, and 1455 records of the demographic data are involved in our experiment, including gender, age, and nationality.

D. PHYSICAL APPEARANCE AND SOCIAL NETWORKS

To collect physical appearance accurately, all participants are assigned to a lab and instructed to sit down and look at the center of the camera lens with neutral expressions, hair pulled back, and no adornments. Students are photographed under consistent lighting conditions with a fixed camera distance. We used a Fujifilm FinePix S5 Pro digital SLR camera (60 mm fixed length lens) and a photo booth painted white with calibrated D65 white lighting. These facial photographs are aligned according to interpupillary distance. We resize and crop the photographs to ensure the display of equal proportions of neck and hair. This process results in a set of 485 images of 485 identities.

After the photos, participants are asked to nominate members of their social networks regarding important dimensions of friendship, life advice, academic advice, support, cooperation, intelligence, and good/bad news sharing. They are instructed to write down 5-8 names of freshmen living in their dorm area (The 485 reference student names are given as a reference list.) with questions including: please choose 5-8 freshman names from the list who are your friends; who are intelligent; whom you would go to for academic advice/life advice, sharing good/bad news, support, or cooperation. For the sake of aesthetics in the format, we selected six of these networks for visualization shown in Figure 5.

VI. EXPERIMENTS AND RESULTS

In this section, we present the experimental results in detail to not only demonstrate the performance of the proposed approaches, including the CASTLE framework and MOON algorithm, but also to explore the detectability of students with mental health problems. All experiments are implemented in Python 3.6. Packages including Pandas and Scikit-learn are utilized for data analysis and detection. Origin 2018 and Graph are utilized for the visualization of data and experimental results. We first introduce the representation results of academic performance and physical

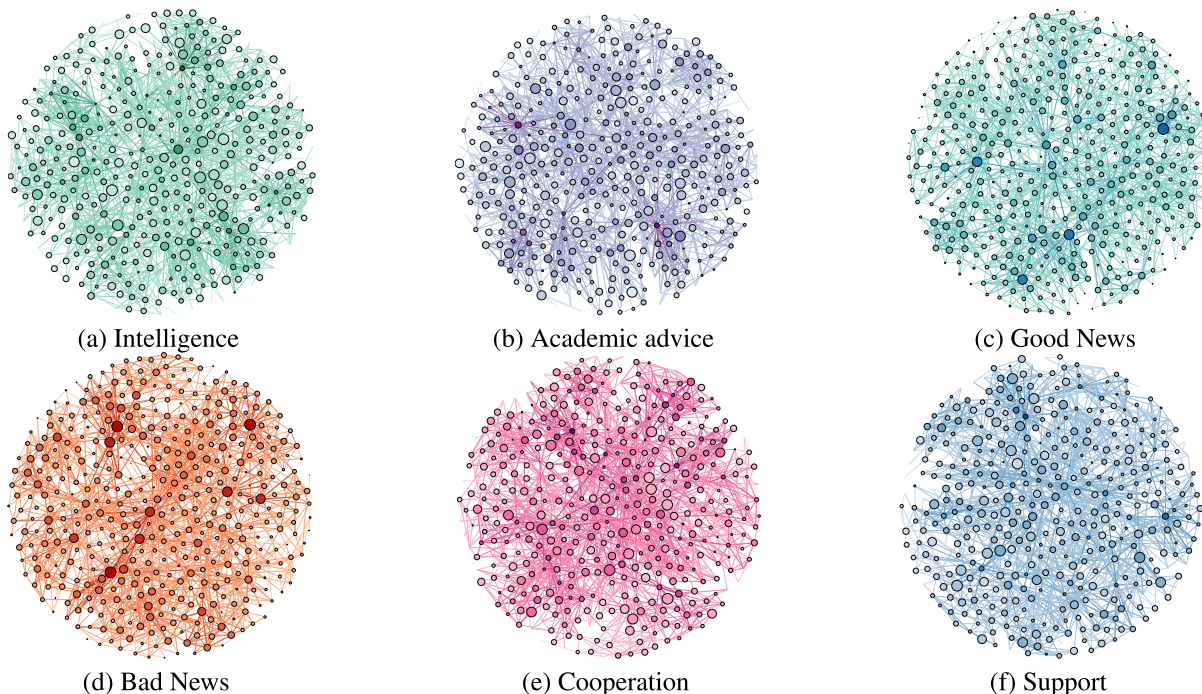


FIGURE 5. The sketch of social networks.

appearance. Then, we introduce the experimental settings of mental status detection and its results.

A. REPRESENTATION OF ACADEMIC PERFORMANCE AND PHYSICAL APPEARANCE

To deal with the heterogeneity of academic performance data, we apply the representation approach of combining the variant of one-hot encoding and autoencoder, as mentioned before. Since this experiment was conducted in the second semester, only grades from the first semester are used as a feature for detection. We test the different dimensions and the performance is shown in Figure 6. The value of the loss function fluctuates slightly, representing that even vectors with low dimensions can still accurately represent the academic performance of each student. Thus, we choose 6 as the dimension of representation for computational efficiency. Moreover, we use a CNN-based auto-encoder to process students' photos and the representations are shown in Figure 6. We choose 6 as the dimension of representation as well.

B. DETECTION RESULTS

1) RESULTS ANALYSIS

We design a series of experiments to explore the detectability of students with mental health problems and to validate the methods proposed in this paper. There are a total of 485 samples in our dataset, and as we mentioned before, the ratio of students with healthy mental status to students with mental health problems is 7 to 1. Four features are used for detection, including social life, appearance, academic performance, and demographic information.

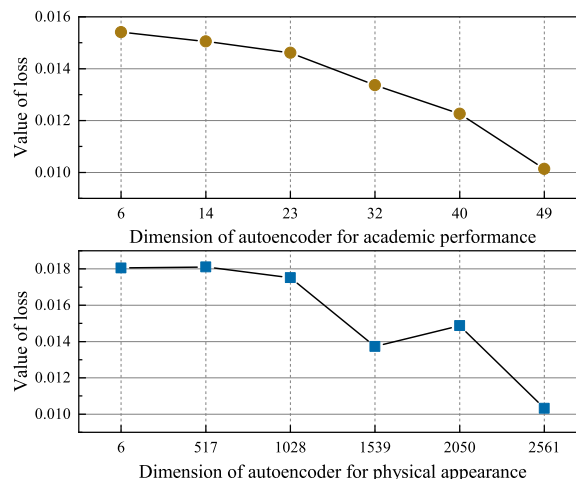


FIGURE 6. The results of feature representation for academic performance and physical appearance.

The embedding dimension of social life is 8 (details are shown in Section VI-B2). The embedding dimension of both appearance and academic performance is 6. Moreover, three pieces of demographic information were included in this experiment, including gender, age, and nationality. In this case, the final feature used for prediction is a 23-dimensional vector.

Due to privacy concerns, there is a lack of publicly available datasets in the field of student mental health. It is difficult for all scholars in the field to test the performance of the algorithm in the same data environment. In this case, we design

the following comparative experiments based on algorithms commonly used in related fields. Firstly, we replace the specified parts of the proposed framework with some popular algorithms, which include two parts. In the first part, we replace our network embedding algorithm with the following algorithms:

- **Deepwalk** [43]: Deepwalk is a classic embedding algorithm that is widely used in the field of network representation.
- **MANE** [42]: MANE is a multi-view network algorithm, which inspired us to propose the MOON algorithm.

In the second step, we replace our final detection model with current popular algorithms shown as follows:

- **Support Vector Machine (SVM)** [49]: SVM is a classic algorithm and is widely used in the field of data mining.
- **Random Forest (RF)** [50]: is a classic ensemble algorithm that achieves good performance in various applications.
- **XGBoost** [51]: XGBoost is a boosting-tree-based method and is widely used in various data mining scenarios with good performance.

We divide the training and test sets into ratios of 9:1, 8:2, 7:3, 6:4, and 5:5, respectively. For each train and test repartition, we use SMOTE to alleviate the class imbalance problem, and the data generation process is shown as follows:

- First, raw data is divided into two categories: the training set a and the testing set b by stratified sampling.
- Second, we use SMOTE on the training set a to generate samples of the minority class. Then in the new training set a' , the number of students in the two classes is equal.

We test the performance of these algorithms from two aspects for a comprehensive evaluation. On one hand, we fit algorithms based on the raw training set a and test them on the testing set b . The results are shown in Figure 7. In this paper, we quantify the experimental results by Accuracy, Recall, and F1 Score, because Precision is a constant value given that Recall and F1 Score are known. In this case, Accuracy, Recall and F1 Score can provide a comprehensive and effective evaluation of the experimental results. Especially, we are more concerned with how many students with mental problems are identified, so Recall and F1 Score are more effective indicators. It is shown that the detection is not accurate and the fluctuation is quite large due to the label imbalance of raw data. To overcome this problem, we fit algorithms based on the balanced training set a' and test them on testing set b . The results are shown in Figure 8. The performance is improved significantly.

Note that from the performance of MOON+DNN and MANE+DNN in Figure 8, the proposed MOON algorithm is only slightly better than the MANE algorithm, but the proposed algorithm has a clear advantage in computational overhead. In the original mode MANE, the complexities for cross-view-intra-node and cross-view-cross-node consistencies are both $O(|E| \cdot D/|V| \cdot K \cdot |V|)$ and the overall complexity is $O(|E| \cdot D \cdot K)$, where K is the number of negative samples. For the MOON proposed in this paper, the complexities for

cross-view-intra-node and cross-view-cross-node consistencies are both $O(|E| \cdot D/|V| \cdot K)$ through simplifying collaboration between networks (Eq. 3 and Eq. 4). Finally, the overall complexity of the MOON is $O(|E| \cdot D/|V| \cdot K)$.

Finally, to better understand the performance of our framework, we compare the performance of current popular algorithms trained on the raw training set a (with label imbalance) and the performance of CASTLE trained on the balanced training set a' (processed by SMOTE). Note that in this part, since SMOTE is part of the CASTLE framework, only CASTLE is based on the SMOTE-processed balanced dataset. Other algorithms are based on the raw dataset. The results are shown in Table 3, which demonstrates the performance of the framework.

TABLE 3. Comparison of the CASTLE framework and popular algorithms.

	Accuracy	Recall	F1 score
CASTLE	0.84468	0.71468	0.74647
Deepwalk+DNN	0.78031	0.60862	0.61431
MANE+DNN	0.74431	0.59607	0.60986
MOON+SVM	0.63636	0.60352	0.61095
MOON+RF	0.77272	0.57058	0.57528
MOON+GBDT	0.79545	0.58529	0.59384

2) INPUT ANALYSIS

First, the detection performance with different embedding dimensions of the MOON algorithm is analyzed based on the CASTLE framework, and the results are shown in Table 4. For computational efficiency, the embedding dimension of the MOON algorithm is set at 8. Second, the contribution of each type of feature is analyzed in this paper. The results are shown in Table 5. All features contribute to the detection, which is consistent with our assumption. Note that the contribution of physical appearance is relatively small because social life and physical appearance may contain redundant information [11].

TABLE 4. Detection performance with different embedding dimension of the MOON algorithm.

Dimension	Accuracy	Recall	F1 score
4	0.82177	0.69781	0.72613
8	0.84468	0.71468	0.74647
16	0.85391	0.71789	0.74998
32	0.85433	0.72431	0.75513
64	0.84681	0.70168	0.74001

TABLE 5. Performance with different input. For concise presentation in the table, we used shorthand to represent each part: F (Friendship network), S (Social behavior (Multi-view social network)), A (Academic performance), P (Physical appearance), D (Demographic features).

Inputs	Accuracy	Recall	F1 score
F	0.76969	0.61622	0.63642
$S + P$	0.78939	0.63374	0.65294
$S + P + A$	0.80242	0.68627	0.70132
$S + P + A + D$	0.84468	0.71468	0.74647
$F + P + A + D$	0.82179	0.69844	0.71771

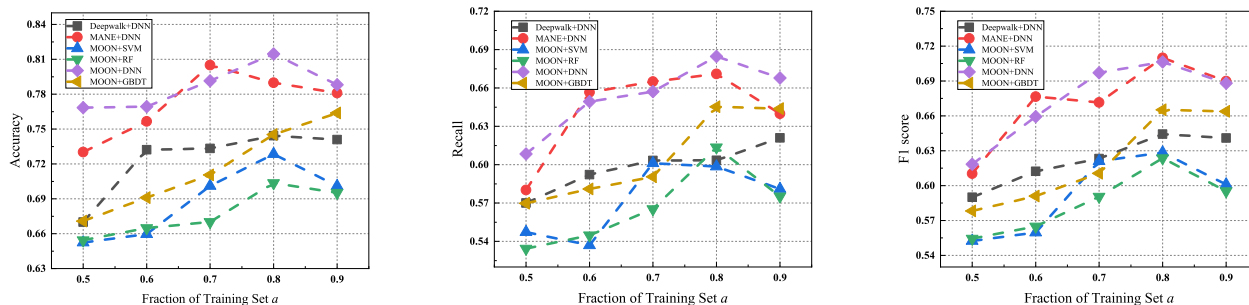


FIGURE 7. Performance of mental health detection on raw training dataset a .

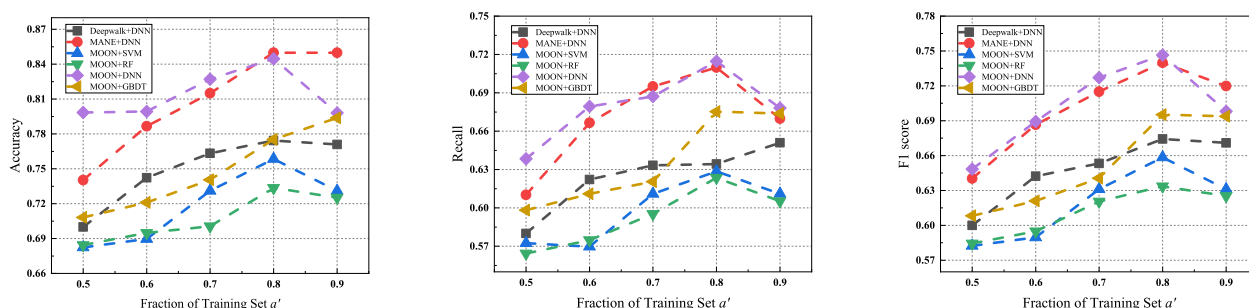


FIGURE 8. Performance of mental health detection on balanced training dataset a' .

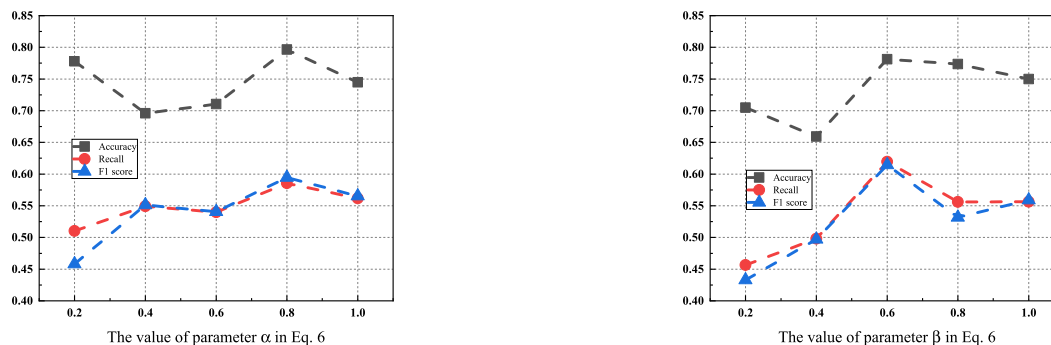


FIGURE 9. Parameters analysis of α and β in Eq. 5.

Moreover, the questionnaire is applied to collect the multi-view social network of students, and this method is time- and cost-consuming and hardly applicable to large-scale. In this case, we explore the performance of the situation without manual data collection methods such as questionnaires. We replace the multi-view network collected through questionnaires with a data-generated friendship network through their canteen co-occurrence frequency [52]. The structure of the friendship network is embedded through Deepwalk and the performance is shown in Table 5 as $F + P + A + D$. Compared with $S + P + A + D$, its performance declined slightly, but this approach contains all data from LMS, which is more practical for the education management division (More discussion in Section VII).

In addition, we design experiments to analyze the parameters α and β in Eq. 5 separately. To avoid the influence of other characteristics, this part of the experiment is based on the original data (without data augmentation) and uses only social networks to detect the targeted students. We fix α and β to 0.5 respectively to test the effect of another parameter on the detection results. The results are shown in Figure 9. It can be seen that both parameters affect the detection results in different degrees.

Finally, as mentioned above, a dropout mechanism is utilized to improve the generalization performance, and we test the sensitivity of the proposed framework on dropout rate (Figure 10). The change of dropout proportions impacts the performance slightly, and 0.3 is the best.

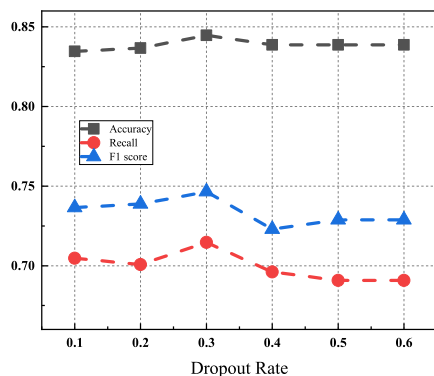


FIGURE 10. Performance of the proposed framework with different dropout proportions.

VII. CONCLUSION AND DISCUSSION

In this paper, we investigate an important problem of the detection of students' mental health. We propose an educational data fusion detection framework CASTLE for achieving an effective and accurate detection through fusing multi-modal data generated from campus life. We tackle the various challenges that exist in multi-modal data by using representation learning theories. Specifically, for the representation of students' social life, we divide social networks into the source view and the target view according to the inherent nature of social behaviors and propose the MOON algorithm for multi-view social network embedding. In essence, the idea we provided here is a new embedding strategy for the multi-view network that different networks need to be treated differently depending on the characteristics of the specific application scenario, and this strategy could be extended to other fields. Moreover, we use a SMOTE model to overcome the label imbalance problem. We conduct comprehensive experiments on a real-world educational dataset and the extensive results demonstrate the performance of the proposed methods.

Although we demonstrate the performance of the proposed algorithm through rich experiments, there is still some room for improvement in this study. First, we consider campus social networks as static weightless multi-view networks. However, social networks are dynamic and each link is weighted differently in real life. How to accurately characterize the social network of students is still an open question. In addition to campus social relationships, other social networks also have a significant impact on mental health, such as intimate relationships, family relationships, and teacher-student relationships. Subsequent research needs to consider social networks more comprehensively. Moreover, as mentioned above, the questionnaire is applied to collect multi-view social network of students, and this method is time- and cost-consuming and hardly applicable to large-scale. Although we try to automatically capture friendship relationships among students based on cafeteria co-occurrence, such methods are too crude and the accuracy cannot be guaranteed. How to capture students' social networks in diverse scenarios based on a data-driven

approach is still an open question. Finally, the deep learning model is adopted in this paper, which means losing the interpretability of the experimental results while obtaining better prediction performance. This can easily cause educators to mistrust the prediction.

According to the drawbacks mentioned above, there are multiple directions for future work as follows:

(1) First, we will develop subsequent versions of the CASTLE framework to fuse more features, like students' Internet access patterns and life orderliness, to achieve better detection performance.

(2) Second, we will attempt to develop data-driven methods for capturing students' social patterns based on group work records, or discussion records on LMS, to replace the questionnaire-based data collection.

(3) Third, in the next step, we try to introduce causal learning related techniques to analyze the experimental results.

(4) Last but not least, we also intend to integrate the CASTLE framework into the modern educational management system to assist with educational decision making.

ACKNOWLEDGMENT

The authors thank Dongyu Zhang, Chuanhui Yuan, and Qing Qing for their help with the experiments as well as all students who participated in the experiments.

REFERENCES

- [1] J. Gong, Y. Huang, P. I. Chow, K. Fua, M. S. Gerber, B. A. Teachman, and L. E. Barnes, "Understanding behavioral dynamics of social anxiety among college students through smartphone sensors," *Inf. Fusion*, vol. 49, pp. 57–68, Sep. 2019.
- [2] P. Y. Collins and S. Saxena, "Action on mental health needs global cooperation," *Nature*, vol. 532, no. 7597, pp. 25–27, Apr. 2016.
- [3] A. D. Bergin, E. P. Vallejos, E. B. Davies, D. Daley, T. Ford, G. Harold, S. Hetrick, M. Kidner, Y. Long, S. Merry, R. Morriss, K. Sayal, E. Sonuga-Barke, J. Robinson, J. Torous, and C. Hollis, "Preventive digital mental health interventions for children and young people: A review of the design and reporting of research," *NPJ Digit. Med.*, vol. 3, no. 1, pp. 1–9, Dec. 2020.
- [4] T. M. Evans, L. Bira, J. B. Gastelum, L. T. Weiss, and N. L. Vanderford, "Evidence for a mental health crisis in graduate education," *Nature Biotechnol.*, vol. 36, no. 3, p. 282, 2018.
- [5] D. Zhang, N. Shi, C. Peng, A. Aziz, W. Zhao, and F. Xia, "MAM: A metaphor-based approach for mental illness detection," in *Proc. Int. Conf. Comput. Sci.* Cham, Switzerland: Springer, 2021, pp. 570–583.
- [6] Mental Health America (MHA). (2020). *The State of Mental Health in America 2021*. [Online]. Available: <https://www.mhanational.org/research-reports/2021-state-mental-health-america>
- [7] E. A. Holman, R. R. Thompson, D. R. Garfin, and R. C. Silver, "The unfolding COVID-19 pandemic: A probability-based, nationally representative study of mental health in the United States," *Sci. Adv.*, vol. 6, no. 42, p. 5390, Oct. 2020.
- [8] D. Witteveen and E. Velthorst, "Economic hardship and mental health complaints during COVID-19," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 44, pp. 27277–27284, Nov. 2020.
- [9] S. Yu, Q. Qing, C. Zhang, A. Shehzad, G. Oatley, and F. Xia, "Data-driven decision-making in COVID-19 response: A survey," *IEEE Trans. Computat. Social Syst.*, vol. 8, no. 4, pp. 1016–1029, Aug. 2021.
- [10] D. D. Ebert, P. Mortier, F. Kaehlke, R. Bruffaerts, H. Baumeister, R. P. Auerbach, J. Alonso, G. Vilagut, K. U. Martínez, C. Lochner, P. Cuijpers, A. Kuechler, J. Green, P. Hasking, C. Lapsley, N. A. Sampson, and R. C. Kessler, "Barriers of mental health treatment utilization among first-year college students: First cross-national results from the WHO world mental health international college Student initiative," *Int. J. Methods Psychiatric Res.*, vol. 28, no. 2, Jun. 2019, Art. no. e1782.

- [11] D. Zhang, T. Guo, H. Pan, J. Hou, Z. Feng, L. Yang, H. Lin, and F. Xia, "Judging a book by its cover: The effect of facial perception on centrality in social networks," in *Proc. World Wide Web Conf. (WWW)*, 2019, pp. 2290–2300.
- [12] S. A. Morelli, D. C. Ong, R. Makati, M. O. Jackson, and J. Zaki, "Empathy and well-being correlate with centrality in different social networks," *Proc. Nat. Acad. Sci. USA*, vol. 114, no. 37, pp. 9843–9847, Sep. 2017.
- [13] M. Li, W.-Q. Li, and L. M. W. Li, "Sensitive periods of moving on mental health and academic performance among university students," *Frontiers Psychol.*, vol. 10, p. 1289, Jun. 2019.
- [14] H. Yao, D. Lian, Y. Cao, Y. Wu, and T. Zhou, "Predicting academic performance for college students: A campus behavior perspective," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 3, pp. 1–21, May 2019.
- [15] T. Guo, X. Bai, X. Tian, S. Firmin, and F. Xia, "Educational anomaly analytics: Features, methods, and challenges," *Frontiers Big Data*, vol. 4, Jan. 2022, Art. no. 811840.
- [16] R. Xiao and X. Liu, "Analysis of the architecture of the mental health education system for college students based on the Internet of Things and privacy security," *IEEE Access*, vol. 9, pp. 81089–81096, 2021.
- [17] F. Amin, A. Ahmad, and G. S. Choi, "Towards trust and friendliness approaches in the social Internet of Things," *Appl. Sci.*, vol. 9, no. 1, p. 166, 2019.
- [18] J. Liu, X. Kong, F. Xia, X. Bai, L. Wang, Q. Qing, and I. Lee, "Artificial intelligence in the 21st century," *IEEE Access*, vol. 6, pp. 34403–34421, 2018.
- [19] M. Hou, J. Ren, D. Zhang, X. Kong, D. Zhang, and F. Xia, "Network embedding: Taxonomies, frameworks and applications," *Comput. Sci. Rev.*, vol. 38, Nov. 2020, Art. no. 100296.
- [20] F. Xia, A. M. Ahmed, L. T. Yang, and Z. Luo, "Community-based event dissemination with optimal load balancing," *IEEE Trans. Comput.*, vol. 64, no. 7, pp. 1857–1869, Jul. 2015.
- [21] W. Yuan, K. He, C. Shi, D. Guan, Y. Tian, A. Al-Dhelaan, and M. Al-Dhelaan, "Multi-view network embedding with node similarity ensemble," *World Wide Web*, vol. 1, no. 2, pp. 1–16, 2020.
- [22] B. Zhou, J. Pei, and W. Luk, "A brief survey on anonymization techniques for privacy preserving publishing of social network data," *ACM SIGKDD Explor. Newslett.*, vol. 10, no. 2, pp. 12–22, Dec. 2008.
- [23] T. Guo, T. Tang, D. Zhang, J. Li, and F. Xia, "Web of students: Class-level friendship network discovery from educational big data," in *Proc. Int. Conf. Web Inf. Syst. Eng.* Cham, Switzerland: Springer, 2021, pp. 497–511.
- [24] T. Guo, F. Xia, S. Zhen, X. Bai, D. Zhang, Z. Liu, and J. Tang, "Graduate employment prediction with bias," in *Proc. 32th AAAI Conf. Artif. Intell.* Palo Alto, CA, USA: AAAI Press, 2020, pp. 670–677.
- [25] J. Akullian, A. Blank, L. Bricker, L. DuHadway, and C. Murphy, "Supporting mental health in computer science students and professionals," in *Proc. 51st ACM Tech. Symp. Comput. Sci. Educ.*, Feb. 2020, pp. 958–959.
- [26] M. Rossin-Slater, M. Schnell, H. Schwandt, S. Trejo, and L. Uniat, "Local exposure to school shootings and youth antidepressant use," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 38, pp. 23484–23489, Sep. 2020.
- [27] A. L. Duckworth and M. E. P. Seligman, "Self-discipline outdoes IQ in predicting academic performance of adolescents," *Psychol. Sci.*, vol. 16, no. 12, pp. 939–944, Dec. 2005.
- [28] W. Usher and C. Curran, "Predicting Australia's university students' mental health status," *Health Promotion Int.*, vol. 34, no. 2, pp. 312–322, Apr. 2019.
- [29] A. Wongkoblap, M. A. Vadillo, and V. Curcin, "Detecting and treating mental illness on social networks," in *Proc. IEEE Int. Conf. Healthcare Informat. (ICHI)*, Aug. 2017, p. 330.
- [30] R. Vanlalawmpuia and M. Lalhmimgliana, "Prediction of depression in social network sites using data mining," in *Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2020, pp. 489–495.
- [31] K. E. B. Ooi, M. Lech, and N. B. Allen, "Multichannel weighted speech classification system for prediction of major depression in adolescents," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 2, pp. 497–506, Feb. 2013.
- [32] O. Oyeboode, F. Alqahtani, and R. Orji, "Using machine learning and thematic analysis methods to evaluate mental health apps based on user reviews," *IEEE Access*, vol. 8, pp. 111141–111158, 2020.
- [33] T. Tran, D. Phung, W. Luo, R. Harvey, M. Berk, and S. Venkatesh, "An integrated framework for suicide risk prediction," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2013, pp. 1410–1418.
- [34] R. Brathwaite, T. B.-M. Rocha, C. Kieling, B. A. Kohrt, V. Mondelli, A. O. Adewuya, and H. L. Fisher, "Predicting the risk of future depression among school-attending adolescents in Nigeria using a model developed in Brazil," *Psychiatry Res.*, vol. 294, Dec. 2020, Art. no. 113511.
- [35] A. E. Tate, R. C. McCabe, H. Larsson, S. Lundström, P. Lichtenstein, and R. Kuja-Halkola, "Predicting mental health problems in adolescence using machine learning techniques," *PLoS ONE*, vol. 15, no. 4, Apr. 2020, Art. no. e0230389.
- [36] C. G. Walsh, J. D. Ribeiro, and J. C. Franklin, "Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning," *J. Child Psychol. Psychiatry*, vol. 59, no. 12, pp. 1261–1270, Dec. 2018.
- [37] F. Ge, L. W. D. Zhang, and H. Mu, "Predicting psychological state among Chinese undergraduate students in the COVID-19 epidemic: A longitudinal study using a machine learning," *Neuropsychiatric Disease Treatment*, vol. 16, p. 2111, Jan. 2020.
- [38] N. Rubaiyat, A. I. Apsara, D. Chaki, H. Arif, L. Israt, L. Kabir, and M. G. R. Alam, "Classification of depression, internet addiction and prediction of self-esteem among university students," in *Proc. 22nd Int. Conf. Comput. Inf. Technol. (ICIT)*, Dec. 2019, pp. 1–6.
- [39] P. Mathur, R. Sawhney, and R. R. Shah, "Suicide risk assessment via temporal psycholinguistic modeling," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 10. Palo Alto, CA, USA: AAAI Press, 2020, pp. 13873–13874.
- [40] M. Gaur, A. Alambo, J. P. Sain, U. Kursuncu, K. Thirunakaran, R. Kavuluru, A. Sheth, R. Welton, and J. Pathak, "Knowledge-aware assessment of severity of suicide risk for early intervention," in *Proc. World Wide Web Conf. (WWW)*, 2019, pp. 514–525.
- [41] H. Cai, Z. Qu, Z. Li, Y. Zhang, X. Hu, and B. Hu, "Feature-level fusion approaches based on multimodal EEG data for depression recognition," *Inf. Fusion*, vol. 59, pp. 127–138, Jul. 2020.
- [42] S. K. Ata, Y. Fang, M. Wu, J. Shi, C. K. Kwok, and X. Li, "Multi-view collaborative network embedding," *ACM Trans. Knowl. Discovery From Data*, vol. 15, no. 3, pp. 1–18, Apr. 2021.
- [43] B. Perozzi, R. Al-Rfou, and S. Skiena, "DeepWalk: Online learning of social representations," in *Proc. 20th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2014, pp. 701–710.
- [44] F. Xia, J. Liu, H. Nie, Y. Fu, L. Wan, and X. Kong, "Random walks: A review of algorithms and applications," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 4, no. 2, pp. 95–107, Apr. 2020.
- [45] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, no. 1, pp. 321–357, Jan. 2002.
- [46] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," 2018, *arXiv:1803.01271*.
- [47] Y. Yu, X. Si, C. Hu, and Z. Jianxun, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019.
- [48] Y. Wei, H. Li, H. Wang, S. Zhang, and Y. Sun, "Psychological status of volunteers in a phase I clinical trial assessed by symptom checklist 90 (SCL-90) and Eysenck personality questionnaire (EPQ)," *Med. Sci. Monitor: Int. Med. J. Experim. Clin. Res.*, vol. 24, p. 4968, Jan. 2018.
- [49] B. Scholkopf and A. J. Smola, *Learning With Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA, USA: MIT Press, 2001.
- [50] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [51] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.
- [52] H. Yao, M. Nie, H. Su, H. Xia, and D. Lian, "Predicting academic performance via semi-supervised learning with constructed campus social network," in *Proc. Int. Conf. Database Syst. Adv. Appl.* Cham, Switzerland: Springer, 2017, pp. 597–609.



TENG GUO received the bachelor's degree in automation from Henan University, Kaifeng, China, in 2013, the master's degree in design and management of sustainable built environment from the University of Reading, U.K., in 2015, and the master's degree in control engineering from the Xi'an University of Architecture and Technology, in 2016. He is currently pursuing the Ph.D. degree with the Alpha Laboratory, School of Software, Dalian University of Technology, China. His research interests include educational data mining, human behavior analysis, and human profiling.



WENHONG ZHAO (Member, IEEE) received the Ph.D. degree from Zhejiang University, Hangzhou, China, in 2002. Since 1991, he has been working at the Zhejiang University of Technology, Hangzhou, where he is currently a Full Professor with the Ultraprecision Machining Center. His research interests include big data, embedded systems, intelligent systems, and precision machining.



SELENA FIRMIN received the Ph.D. degree from Monash University, Australia. She is currently a Lecturer with the Institute of Innovation, Science and Sustainability, Federation University Australia. She is interested in active learning and enhancing students' learning experiences using technology. Her research interests include computing education, technology in health, and information systems.



MUBARAK ALRASHOUD received the Ph.D. degree in computer science from Ryerson University, Toronto, ON, Canada, in 2015. He is currently an Associate Professor and the Head of the Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia. His research interests include social network analysis, vehicular ad hoc networks, data science, and software engineering.



AMR TOLBA (Senior Member, IEEE) received the M.Sc. and Ph.D. degrees from the Department of Mathematics and Computer Science, Faculty of Science, Menoufia University, Egypt, in 2002 and 2006, respectively. He is currently a Full Professor of computer science at King Saud University (KSU), Saudi Arabia. He has authored/coauthored over 100 scientific papers in top ranked (ISI) international journals and conference proceedings. His main research interests include artificial intelligence (AI), the Internet of Things (IoT), data science, and cloud computing. He serves as a technical program committee (TPC) member for several conferences. He served as an associate editor/guest editor for several ISI journals.



FENG XIA (Senior Member, IEEE) received the B.Sc. and Ph.D. degrees from Zhejiang University, Hangzhou, China. He was a Full Professor and an Associate Dean (Research) with the School of Software, Dalian University of Technology, China. He is currently an Associate Professor and a former Discipline Leader (IT) with the Institute of Innovation, Science and Sustainability, Federation University Australia. He has published two books and over 300 scientific papers in international journals and conferences. His research interests include data science, artificial intelligence, graph learning, and systems engineering. He is a Senior Member of ACM.

...