GAT-LS: A Graph Attention Network for Learning Style Detection in Online Learning Environment

Wanying Suo[1], Bello Ahmad Muhammad[2], Zhichao Zhang[3], and Bo Liu[1]
[1]School of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi, 710062, China
[2]School of Computer Science and Technology, Henan Institute of Science and Technology, Henan, 453003, China
[3]School of Information Science and Technology, Hainan Normal University, Haikou, 571158, China

In the context of the information age, the rapid growth and increasing diversity of learning resources underscore the urgency of personalized learning, while learning style is the most crucial factor to consider in personalized learning as it significantly influences students' academic achievements and learning experiences. Traditional methods of assessing learning styles, such as completing questionnaires, have many drawbacks, including subjectivity and time costs. Therefore, in recent years, researchers have been exploring automatic methods to identify learning styles by analyzing students' interactive behaviors. Motivated by these limitations, we propose a learning style detection method using a graph attention network (GAT), named GAT-LS. We originally constructed a bipartite graph between learners and learning materials, utilizing node features to represent the students' behavior. Subsequently, we employ GAT to obtain hidden vectors for the graph nodes. These hidden vectors encapsulate both the overall graph information and the importance of neighboring nodes. We employ a multi-head attention network to process student nodes and combine a dropout mechanism with a single-layer attention network to process learning material nodes. Finally, we map the obtained hidden node features to the Felder-Silverman learning style model (FSLSM) and use K-means clustering to detect learning styles. The proposed method can be integrated into various types of educational systems or online learning platforms, providing a better educational experience and learning resource recommendations for both teachers and students. Experiments on the real-world dataset, KDD CUP 2015, demonstrated the superiority of our method. Our proposed approach achieved outstanding results with average values of 0.9647 accuracy, 0.9478 precision, 0.9171 recall, and 0.9346 F1 score.

***Index Terms*—Learning style, Graph learning, Graph attention network, Interactive learning environment, FSLSM.**

I. INTRODUCTION

IN the face of information age, the field of education has undergone profound transformations. Future education is marked by a series of emerging trends, including but not limited to online learning, personalized education, and career development. Presently, numerous online learning platforms and institutions as well as online learning environments(OLEs) such as Bilibili, China MOOC, SuperStar Learning, among others, have enriched the supply of educational resources. Particularly in the post-pandemic era, online learning has become a significant trend, providing learners with more flexible learning methods and achieving a more equitable distribution of educational resources. However, while online learning brings convenience to learners, it is also accompanied by challenges such as information overload and difficulties in learner navigation and course selection due to a "one-size-fits-all" [1]approach.

In this context, personalized learning, aimed at providing tailored educational services through precise learner profiles, emerges as a key solution to address the aforementioned issues. Learning style (LS), as one of the crucial characteristics for delivering personalized learning, becomes a core factor in realizing personalized education. LS can help learners better identify their learning preferences and help them understand what learning methods and tools can help them better grasp knowledge. And for teachers, better personalized teaching can be carried out by identifying LS of students [2]. Therefore, an efficient and appropriate learning style detection method that analyzes the interaction data between students and online platforms is urgently needed.

Manuscript received April 20, 2024; revised July 22, 2024. Corresponding author: Wanying Suo (email: wysuo@snnu.edu.cn). Email Addresses: Bello Ahmad Muhammad (mbahmad@hist.edu.cn), Zhichao Zhang (zhangzhichao@hainnu.edu.cn), Bo Liu(liubojsj@snnu.edu.cn).

The research on learning style detection is either collaborative or automatic methods in the literature [3]. The collaborative learning style detection method is through a large number of questionnaire surveys [4]. Obviously, this is easy to implement and has good credibility, however, this traditional methods are static and time-consuming in the process of filling out questionnaires, also, lacking of flexibility makes it difficult to conduct in-depth qualitative investigations. In order to detect changes in student behavior data in real-time and improve the accuracy and efficiency of learning style detection, research has focused on an automatic approach to identifying the learner's learning style based on their behavior in online educational systems [1], [5], [6]. Automatic tenhniques analyzes the behavior of learners during the learning process and uses data mining, machine learning, or deep learning techniques to detect their learning styles. Plus, the automatic approach can be broken down into categories: A literature-based (LB) approach and a data-driven approach. The core idea of LB approach is to use the behavior of learners during learning to infer their learning style and then apply a simple rule-based method to calculate the learner's learning style. LB was first proposed by [7], then many scholars applied this method [8], [9], [10]. The advantages of LB are its strong universality and applicability to data collected from any learning system, while the disadvantages are that the technology is static and requires pre-set rules. On the other hand, in the data-driven approach, researchers use a variety of machine learning and deep learning methods [11], [12], [13], [14], [15] to automatically detect learning styles. These automatic methods can automatically extract data features with high accuracy for large-scale data, furthermore, the approachs can dynamically monitor changes in student behavior. However, the interaction data between students and online platforms is large-scale, which sets obstacles for capturing changes in learner behavior, another problem is that the quality of large-scale data cannot be guaranteed, and there may be issues such as missing, duplicate, and incorrect data. Also, most existing methods heavily rely on specific learning environments or educational systems and have poor transferability, but we need to make real-time adjustments and optimizations based on the needs of students.

To overcome these limitations, and inspired by the recent flourishing development of graph representation learning and attention mechanisms, we propose a novel Graph Attention Network for learning style detection (GAT-LS) in online learning environment. In our approach, the interaction data between students and online platforms is constructed into a bipartite graph and node features are used to enhance the representation of the data, then we build different GAT models to extract hidden features of nodes with different attributes and scales, the hidden features are used to cluster learners with different learning styles according to a suitable learning style model. We utilized the Felder-Silverman learning style model (FSLSM) since research has confirmed that it is suitable for online environments and has stability and reliability [16], [4], [3], [17].The contributions of the paper are as follows:

1) We proposed a method that utilizes graph attention networks (GAT) to determine learning types. This tackles the issue of efficiently collecting student behavior data on a large scale. First, we create an S-M bipartite graph and use node features to capture both the overall graph information and specific node attribute details. Node features are more informative and scalable compared to using edge weights to represent behavioral records.
2) Subsequently, we apply an attention mechanism to learning style detection by processing two types of nodes separately using different GAT models, with the goal of extracting deep hidden features. The attention mechanism allows us to identify which neighboring nodes are more crucial. Clearly, not every interaction behavior with materials equally reflects a student's learning style. Finally, using the FSLSM model, we employ the K-means clustering method to cluster learners and obtain labels, thereby achieving learning style detection.
3) We finally evaluate the effectiveness of the proposed method on real-world large-scale datasets (KDD CUP 2015) and compare its performance with other research results. Our method demonstrates superior performance in the comparison, and notably, we introduce the concept of the "middle-of-the-road classification" problem, offering a novel perspective in academic research.

The remainder of this paper is organized as follows. Section 2 reviews the current research. Section 3 presents problem definition and formulation about our approach, Section 4 provides a detailed introduction of our proposed GAT-LS method. Section 5 introduces the evaluation indicators and gives experiment details. Section 6 summarizes the experiment and propose future research directions.

II. Related Work

This section provides a detailed explanation of learning style as well as the existing automatic approach for learning style detection.

A. Learning Style

Learning styles indicate the way each learner prefers to receive learning materials and interacts with the learning environment [1], indicating the habitual learning preferences or tendencies of individuals. This habit is developed over the long term and possesses stable and unique characteristics, playing a crucial role in personalized learning. The formation of learning styles is influenced by various factors such as society, school, and family [5].

For individual learners, understanding their own learning style can enhance the learning experience, fostering increased confidence and active engagement in the learning process. For educators, comprehending students' learning styles contributes to accurately formulating personalized teaching strategies, providing more targeted support and guidance, thereby improving teaching effectiveness. Throughout the entire learning process, researchers showed [18] that consideration of learning styles enables the both online education system and conventional educational system to better adapt to diverse learning needs, driving education towards a more personalized and flexible direction.

The detection of learning styles begins with selecting an appropriate Learning Style Model (LSM)[16]. In the field of education, several well-known learning style models include Kolb's model, Honey and Mumford's model, Gardner's multiple intelligence theory, Dunn and Dunn's model, Neil Fleming's VAR/VARK model, Myers-Briggs type indicator, Felder-Silverman model[16], [1], [4], etc. The design framework for learning style detection questionnaires is rooted in these models, similarly, automatic learning style detection methods rely on these models for detection and classification.

In this paper, we use FSLSM as the basis for learning style detection with the following reasons. First, FSLSM draws on the advantages of previous famous models (e.g. Kolb's and Myers and Briggs). FSLSM model advocates self-awareness to enhance the engagement, understanding, and retention of educators and students in the classroom [19]. Secondly, It can provide a more detailed description of the learner's behavior. FSLSM divides learners' preferences into 4 dimensions with a total of 16 combinations of each dimension's two poles, and learners can choose one from the two poles. These four dimensions include input (Visual/Verbal), perception (Sensing/Intuitive), processing (Active/Reflective), and understanding (Sequential/Global), as is shown in fig. 1. Thirdly, above all, there are experimental experiences [1], [11], [17] that prove the effectiveness and reliability of the model.

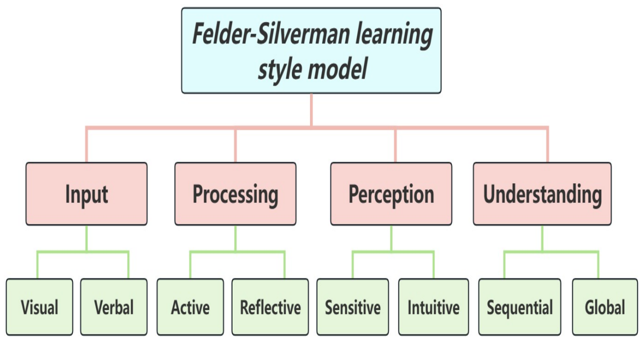

Figure 1. Dimension and Categories of FSLSM

Compared with other models, FSLSM is the most suitable model for online learning environments.

B. Existing Automatic Learning Style Identification Approaches

An automatic approach is suggested as a superior substitute for the conventional approach of learning style detection, which often involve the administration of questionnaire or survey to determine and classify learner's preferred style of learning. The literature [20] and [21] suggest that automatic learning style recognition employ two kinds of methods: literature-based and data-driven approaches.

The literature-based approach utilizes the learner's behavior data to gather insights into their learning style. Subsequently, a straightforward rule-based method is applied to determine the learner's style by counting the number of matching hints. The subsequent paragraphs present an overview of existing automatic approaches that utilize the FSLSM. [13] presents a literature-based approach that utilizes a straightforward rule-based procedure within a web-based learning management system to determine a student's learning style. The methodology autonomously identifies the learning styles and affective states of students by analyzing their preferences and behavior throughout a course. [22] utilizes a literature-based approach to propose a MOOC framework that facilitates learner engagement through adaptive provision of learning resources tailored to their individual learning styles. The author employs a set of attributes to forecast learning style, as described by [20]. In 2017, [23] utilized a literature-based approach and proposed a MOOC course by discerning the learning style of the learners. The authors observe the pattern of learner interaction by utilizing predetermined threshold values and employ a straightforward rule-based approach to effectively identify the learner's learning style. The evaluation was conducted by assigning several values to the threshold. Authors recognized that the accuracy of predictions could be impacted.

In contrast, the data-driven approach leverages data mining, AI, or machine learning algorithms to build a model based on the learner's behavior data. The model is then employed as clustering or classification input to identify their learning styles. This approach is characterized by high accuracy and the utilization of real data for learner categorization. The literature presents many approaches that utilize data-driven methods and base their studies on the FSLSM. These approaches employ either clustering algorithms or classification algorithms to ascertain the learning style of the learner[24], [25], [26], [27]. Among them, [21] employed four computational intelligence algorithms (Artificial neural network, genetic algorithm, ant colony systems, and particle swarm optimization) to detect learning style based on the extracted learners' behavior sequence. The aim was to enhance precision compared to existing methods, potentially improving adaptive learning systems and student learning outcomes. In the same vein, [28] propose a novel hybrid multi-step architecture based on ant colony system and artificial neural networks to increase the precision of learning styles identification.

Multiple machine learning models are trained by [29] to predict students in quarterly courses, then use the most accurate prediction model for each quarter to find the learning style features that maximize student pass rates. [30] and [31] utilize a decision tree to detect learning styles by initially constructing a tree and subsequently pruning it. [15], [22], [32], [33], [34] automated the extraction of learning styles by constructing a Bayesian network and employing data mining approach. Various machine learning techniques, including decision trees, random forests, neural networks, and K-nearest neighbors, have been employed by researchers such as [35], [21], [12], [36], [37], they mean to extract repersentation from students' behavior and construct classifiers. [38] and [39] advocated for using a deep belief network (DBN) to classify learning styles based on FSLSM, evaluating learners' behavior in MOOC environments.

A reinforcement learning approach is proposed by [40] to identify learning styles in E-learning. The method suggests appropriate learning objects (LOs) based on expert knowledge and associated learning styles. It employs three reinforcement strategies tailored for high-performance, low-performance, or both scenarios. For instance, when students encounter LOs, their high performance reinforcement from assessments aligns with the associated learning style. and [41] introduced an automated method for detecting learners' styles based on the four dimensions of FSLSM, aiming to provide adaptive courses on the Moodle platform. The majority of these methods employ either human feature engineering or web mining techniques to extract behavioral features, which are subsequently utilized in conjunction with a machine learning algorithm to ascertain their learning style. Ant colony systems and artificial neural networks are combined in a loose coupling in paper [42],[43] designed an automatic and reliable learning style recognition mechanism. Firstly, a learning style label framework based on multi label fusion (LSDFA) was proposed. In addition, a two-layer ensemble model based on learner online learning behavior data (SRGSML) was proposed to identify learner learning styles.

Recently, [44], [45], and [2] have introduced the use of graph representation learning (GRL) stechniques to automatically acquire features and utilize them to detect students' learning styles. For instance, [45] employed graph-based analysis to capture the activities and behaviors of learners during the learning process. The author utilized a graph autoencoder to transform the graph structure of the constructed bipartite graph into a low-dimensional representation for various downstream machine-learning tasks. Next, the author employs a K-means clustering algorithm to classify and group learners who demonstrate similarities according to the selected FSLSM categories. Furthermore, [2] employs the K-means clustering algorithm to obtain labels according to the FSLSM categories and then applies a graph neural network to accurately classify and forecast the learning style of learners in real-world scenarios. [46] used an Long Short-Term Memory (LSTM) autoencoder to transform past and present data about learners and learning resources into a low-dimensional representation. They also used a clustering algorithm to quickly find and group learners who had similar ways of learning. [47] presents an approach that compares the performance of the K-means clustering algorithm with the fuzzy C-mean (FCM) to detect learning styles. It is possible and useful to find learning styles with the FCM algorithm, as shown by its high precision and recall compared to the k-means clustering algorithm. While the approaches yield satisfactory outcomes, there is still potential for enhancing precision and accuracy.

Learning style based educational systems demonstrates the importance of learning style and its guiding significance for students, researchers design learning style-based adaptive educational systems to improve the effiectiveness of personalized learning. [48] uses artificial neural networks combined with the weighted sum model (WSM) to detect learning style in its individualized tutoring model, and the Honey and Mumford model is utilized to map. [28] propose a novel hybrid multistep architecture based on ant colony system and artificial neural networks to increase the precision of learning styles identification.

III. Problem Definition and Formulation

This section provides an introduction to the relevant ideas used in the study and the notations used in the suggested strategy.

In general, the problem we are studying can be abstracted into one sentence, which is to transform a large amount of unsupervised student interaction data with online platforms into clustered and classified labeled data. The relevant problem definitions and formulas will be described in detail in the following subsections based on the technology we use.

A. Definition of S-M Bipartite Graph and Node Features:

The S-M bipartite graph $G_{sm} = (S, M, E_{sm})$, is a bipartite graph where S and M represent the student set and set of materials. E_{sm} represents the observed student-material interaction relations. Students node features F_s represent the student's frequency access to learning materials when they interact with OLEs. Material nodes features F_m is the mapping relationship with relevant learning style models.

B. Definition of Graph Attention Network :

After constructing the interaction behavior between students and the platform into $G_{sm} = (S, M, E_{sm})$, the goal of the GAT is to encode all the nodes features F_s and F_m into a latent representation (Low dimensional representation) using the different mapping function

$$F_s \to {F_s}'$$
$$F_m \to {F_m}'$$

The mapping function should maintain the inherent structural and weight information between the set of students and the set of materials. In the low dimensional spaces, nodes that contains similar neighbor attention information should have a shorter distance from each other.

C. Definition of Learning Style identification:

Given the encoded latent representation ${F_s}'$ and ${F_m}'$ containing the low dimensional representation of the set of learners $S = \{s_1, s_2, \dots, s_X\}$, set of learning materials $M = \{m_1, m_2, \dots, m_Y\}$, and the importance of material nodes to student nodes. The mapping feature between the encoded learning materials and the selected learning style model of FSLSM. The goal of our method is to use the mapping resulting feature as an input to the clustering algorithm to identify and cluster similar learners with similar learning styles.

IV. GAT-based Learning Style Detection Model

Here, we provide a comprehensive description of the proposed GAT-LS method for detecting learning styles. The objective of this technique is to explore a graph representation learning algorithm that can enhance the precision and accuracy of learning style detection. Figure 2 illustrates that the GAT-LS technique involves three primary steps to accomplish the detection procedure. The first step involves constructing a bipartite graph, known as the S-M graph, which is designed to represent the student's behavior in a graph format, specifically focusing on their actions and interactions with learning resources. These refer to the set of students, the set of learning materials, and their connections based on the data regarding the students' actions and interactions with the learning materials. In the second phase, we utilize the Graph Attention Network (GAT) to establish the connections between nodes and assess the varying significance of neighboring nodes. This process encodes the graph into a compact representation, which may be used for further machine-learning tasks. The final step is detecting the learning style by utilizing the latent representation and mapping feature of each FSLSM dimension. This is done through a clustering algorithm, which helps identify and group learners who have similar learning styles. Below is a comprehensive description of each part in further depth.

A. S-M Bipartite Graph Construction

We utilize the interactions between students and materials to generate a bipartite graph. This grants us two independent collections of nodes, comprising the set of students $S = \{s_1, s_2, \dots, s_X\}$ and the set of learning materials

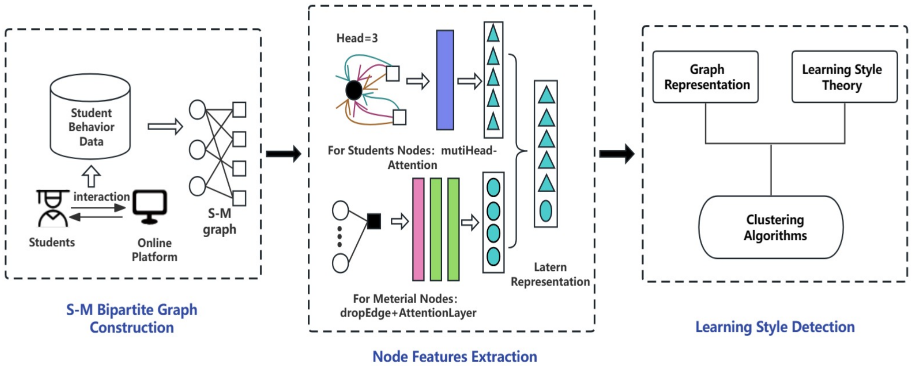

Figure 2. Overview of GAT-LS model for learning styles detection

$M = \{m_1, m_2, \ldots, m_Y\}$. The interaction data provides the connection between the independent sets. A 0-1 matrix is used to represent the connection relationship between two sets of nodes, so if a student interacts with a learning material and no matter how many times they do, an edge $e_k = 1$ is designed between them, otherwise none.

Based on the characteristics of the dataset and the requirements of the selected GRL model, we add additional node features for all nodes. Adding node features to bipartite graphs can provide contextual or full graph information as additional information, helping to better understand the meaning and relationships of nodes.

To determine the feature vector of each student node, we measure the student's frequency access to learning materials when they interact with OLEs, so we obtain a fix-length vertor that represents interaction information. Each dimension of the vector indicates how many times did the student visit each neighbor. For material nodes, the mapping relationship with relevant learning style models is their node features.

The constructed S-M bipartite network is denoted as $G_{sm} = (S, M, E_{sm})$, where S and M represent the set student set and learning material set, respectively. Node feature vectors denoted the observed student-material recorded data, which depict the student's actions and interactions with learning materials while learning. Fig.3 provides an illustration of a sample S-M graph. The graph embedding techniques would extract both structural and node features information about the graph, enabling the conversion of the graph into low dimensional space for the purpose of learning style detection.

B. Node Features Extraction

Node feature extraction refers to the process in graph-based machine learning where relevant features are extracted or learned from the properties linked to nodes in the graph. These have become the most popular research direction in recent years [49], [50]. In many real-world applications, nodes in a graph are often linked with various features or attributes that contain significant information. Extracting meaningful information from these attributes is very essential for constructing efficient models for subsequent machine-learning tasks such as node classification, link prediction, and clustering.

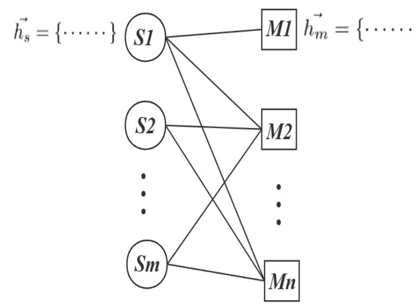

Figure 3. An example of S-M bipartite graph

Several researchers have written a lot of works and shown a lot of interest in the field of graph embedding for feature extraction and classification since the quantity of data is increasing at an exponential rate. Graph embedding techniques have been used for various real-world problems like speech recognition, picture detection, NLP, social networks, knowledge graphs, and recommender systems after numerous algorithms, theories, and large-scale training systems were created. In order to accomplish the extraction procedure, GRL approaches were recently proposed by [45], which makes use of graph autoencoder. While this method outperforms its predecessor, there is still an opportunity for enhancement in terms of accuracy and precision.

We constructed a bipartite graph with distinct characteristics: it comprises two types of nodes, student nodes and material nodes. The number of student nodes is considerably higher than that of material nodes, and the information of their node features also varies significantly. Therefore, it is imperative to consider these two parts of the node feature extraction task separately due to the substantial differences in both structure and the number of nodes. Also, apparently, not every material equally reflects students' learning preferences. Hence, we must take into account the varying importance levels of neighboring nodes. For instance, if a student frequently accesses both material A and material B, these two materials are of heightened importance to that student compared to others. Recognizing the need for a more nuanced approach, we try to introduce an attention mechanism to the learning styles detection task.

Based on the above explanation, we employ the use of Graph Attention Network (GAT) [51] in this paper, a kind of GRL, to achieve feature extraction process. We proposed a comprehensive and efficient feature extraction model based on GAT, which learns non-linear embedding of the S-M bipartite graph. The model is composed of two parts: one uses a multi-head attention network to extract hidden node features of student nodes, and the other uses an drop-edge mechanism and single-layer attention network for materials nodes. The workflow of the feature extraction model is illustrated in fig. 4.

In the following paragraphs, We will provide a detailed introduction to GAT and explain our proposed method.

1) Overview of the GAT

Graphs have aroused great interest among researchers in tackling problems such as knowledge graphs, social relationships, and network security. Also, Graph representation learning has become the most powerful tool for automatically extracting graph features. Attention Mechanism in deep learning is a method that mimics the human visual and cognitive systems. By introducing attention mechanisms, neural networks can automatically learn and selectively focus on important information in the input, improving the performance and generalization ability of the model. Combining graph representation learning with attention mechanisms constitutes GAT. GAT networks have been successfully applied in many tasks, such as social network recommendations [52], traffic flow prediction [53], short text classification [54], etc. However, to our knowledge, the application of GAT techniques has not been adequately evaluated yet in an educational context especially in the field of learning style detection. This article aims to make some preliminary attempts.

By stacking layers in which nodes are able to attend over their neighborhoods' features, GAT enable specifying different weights to different nodes in a neighborhood implicitly [51]. A single-layer structure, graph attention layer, will be introduced first as it is the cornerstone of all subsequent GAT structures in our experiments.

The input of the graph attention layer is a set of node feature vectors, $\mathbf{h} = \left\{\vec{h_1}, \vec{h_2}, \cdots, \vec{h_N}\right\}, \vec{h_i} \in \mathbb{R}^F$, where N is the number of nodes and F is the number of node features. The matrix h with size $N \times F$ represents the characteristics of all nodes, and $\mathbb{R}$ with size $F \times 1$ only represents the characteristics of a certain node. The output of each layer is a new set of node feature vectors, $\mathbf{h} = \left\{\vec{h_1}', \vec{h_2}', \cdots, \vec{h_N}'\right\}, \vec{h_i}' \in \mathbb{R}^{F'}$, where F' represents the new node feature vector dimension, which may not be equal to F.

Then we need to train a weight matrix $\mathbf{W} \in \mathbb{R}^{F' \times F}$ for all nodes in order to obtain the transformation between input and output of a single-layer network, then we implement a self attention attention mechanism for each node with attention coefficients:

$$e_{ij} = a\left(\mathbf{W}\vec{h_i}, \mathbf{W}\vec{h_j}\right) \tag{1}$$

Equation 1 represents the importance of node j to node i without considering the information of the graph structure. There are many ways to choose a, the author chose a single-layer feedforward neural network with parameter $\vec{\mathbf{a}} \in \mathbb{R}^{2F'}$ in the paper. GAT introduces attention mechanism into the graph structure through masked attention, which means that attention is only allocated to the neighboring node set $N_i\,(j \in N_i)$ of node i. In order to make the attention coefficients easier to calculate and easier to compare, we introduce $softmax$ to regularize the adjacent nodes j of all i:

$$\alpha_{ij} = softmax_j\left(e_{ij}\right) = \frac{exp\left(e_{ij}\right)}{\sum_{k \in N_i} exp\left(e_{ik}\right)} \tag{2}$$

e_{ij} and α_{ij} are both called attention coefficients, just α_{ij} is normalized based on e_{ij}. LeakyReLU nonlinearity was also added to the output layer of the feedforward neural network a in the paper. Therefore, by integrating equation 1 and 2, the complete attention mechanism is obtained as follows:

$$\alpha_{ij} = \frac{exp\left(LeakyReLU\left(\vec{\mathbf{a}}^T\left[\mathbf{W}\vec{h_i}' \mid\mid \mathbf{W}\vec{h_j}'\right]\right)\right)}{\sum_{k \in N_i} exp\left(LeakyReLU\left(\vec{\mathbf{a}}^T\left[\mathbf{W}\vec{h_i}' \mid\mid \mathbf{W}\vec{h_j}'\right]\right)\right)} \tag{3}$$

The attention coefficients between different regularized nodes were obtained through the above operation, which can be used to predict the output features of each node:

$$\vec{h_i}' = \sigma\left(\sum_{j \in N_i} a_{ij}\mathbf{W}\vec{h_j}\right) \tag{4}$$

W is the weight matrix multiplied by the features, a is the attention correlation coefficient calculated earlier, σ is a nonlinear activation function, The j traversed in $j \in N_i$ represents the neighboring nodes adjacent to all i.

In order to improve the generalization ability of the attention mechanism, GAT chooses to use a multi-head attention layer, which uses K sets of independent single-layer structure, and then concatenates their results together:

$$\vec{h_i}' = \|_{k=1}^{k}\sigma\left(\sum_{j \in N_i} a_{ij}^k\mathbf{W}^k\vec{h_j}\right) \tag{5}$$

2) Multi-head Attention for Students Node

We train muli-head attention layers on the students nodes of the obtained S-M bipartite graph with node features, which includes information about the complex interaction between students and materials. Obviously, GAT is a powerful feature extraction tool that can not only reduce dimensionality, but its biggest advantage is that it can use attention coefficients to represent the importance level between first-order neighboring nodes.

The three-head Attention layers is utilized when we want the output node features to include information about which neighbor is more important. First, the bipartite graph will be represented as a 0-1 matrix $\mathbf{A}_{ij}$ which is $n \times n$, it is worth noting that the diagonal elements of this matrix are all

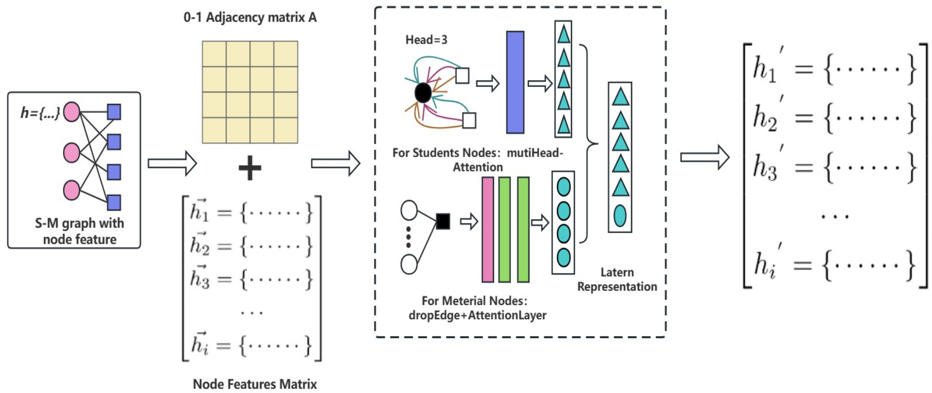

Figure 4. The architecture of feature extraction model

zero, also, each edge is bidirectional. The following equation specifically describes the 0-1 matrix:

$$\mathbf{A}_{ij} = \begin{cases} W_{ij} = 1, & s_i, m_j \in E \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

Then, we aggregate the feature vectors of the nodes into a feature matrix $\mathbf{F}_s$, the example of $\mathbf{F}_s$is as follows:

$$\mathbf{F}_s = \begin{bmatrix} \vec{h_1} = \{\cdots\cdots\} \\ \vec{h_2} = \{\cdots\cdots\} \\ \vec{h_3} = \{\cdots\cdots\} \\ \cdots \\ \vec{h_i} = \{\cdots\cdots\} \end{bmatrix}$$

Finally, inputting $\mathbf{A}_{ij}$ and $\mathbf{F}_s$ into three-head Attention layers model yields a new set of node feature vectors that we want:

$$\mathbf{F}_s{}' = \begin{bmatrix} \vec{h_1}' = \{\cdots\cdots\} \\ \vec{h_2}' = \{\cdots\cdots\} \\ \vec{h_3}' = \{\cdots\cdots\} \\ \cdots \\ \vec{h_i}' = \{\cdots\cdots\} \end{bmatrix}$$

3) Drop-edge Mechanism and Single-layer Attention for Materials Node

As discussed before, different feature extraction models for different sets of nodes and this is the core of the entire work. However, for learning materials nodes, the difference in quantity between them and student nodes requires us to consider an oversmooth problem, which is a unique problem of the GCN model, that is, as the depth of model deepens, the node features gradually become similar, so as the network layers deepen, the node representation tends to converge and the distinguishability of the node representation will deteriorate. Finally, the representation of all nodes will converge to a fixed point, and the resulting node representation will be independent of the input features, and it will also cause the gradient to disappear. As a result, we have lost the characteristics of materials nodes and become similar to many student nodes which is a problem that we need to consider separately.

In order to solve the problem of oversmooth, we introduced the DropEdge mechanism which is a very simple idea that randomly removes some edges during feature extraction. Previous articles have proven the effectiveness of this mechanism for oversmooth problems [55]. Formally, it randomly removes some non-zero elements with probability p, the entire process can be expressed as the following equation:

$$\mathbf{A}_{drop} = \mathbf{A} - \mathbf{A}^{'} \tag{7}$$

Where $A^{'}$ is a sparse matrix of a random subset removed by p.

Based on the above explaination, our model for extracting features of materials node is a DropEdge layer and two single-layer graph attention layers, detailed structure can be described in fig. 4. The input can be represented by $\mathbf{A}_{ij}$ and $\mathbf{F}_m$,

$$\mathbf{F}_m = \begin{bmatrix} \vec{h_1} = \{\cdots\cdots\} \\ \vec{h_2} = \{\cdots\cdots\} \\ \vec{h_3} = \{\cdots\cdots\} \\ \cdots \\ \vec{h_j} = \{\cdots\cdots\} \end{bmatrix}$$

and output is a new set of node feature vectors $\mathbf{F}_{m'}$.

$$\mathbf{F}_m{}' = \begin{bmatrix} \vec{h_1}' = \{\cdots\cdots\} \\ \vec{h_2}' = \{\cdots\cdots\} \\ \vec{h_3}' = \{\cdots\cdots\} \\ \cdots \\ \vec{h_j}' = \{\cdots\cdots\} \end{bmatrix}$$

Now, the feature extraction task for S-M bipartite graphs has been completed.

C. Learning Style Detetion

After we uses GAT model to automatically obtain a fixed-length vectors we want, the next primary task is to select a

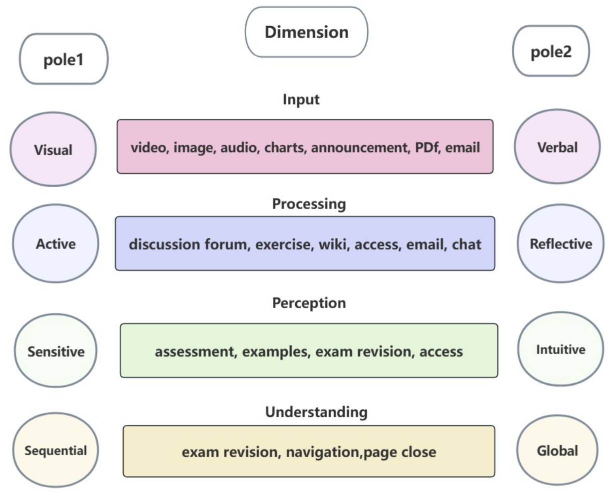

Figure 5. Mapping the encoded learning resources with FSLSM dimensions

suitable learning style model so we can map the vectors of learning materials to per model dimensions. By doing so, the latent representation vector of student nodes can be clustered because they have information about the entire graph and interaction with learning materials. The clustering algorithm uses learning material mapping as feature values to label the behavior of learners.

In this part, we use FSLSM as the basis for learning style detection. The specific reasons for choosing this model have been stated in the related work section. A detailed description about the FSLSM model is shown in fig. 1.

According to the mapping relationship provided by literature [56] [32] as shown in Fig. 5 , we can determine the model dimension to which learners belong based on their behavioral characteristics.

The resulting material nodes latent representation associated with each FSLSM dimension are provided as inputs to the K-means clustering algorithm.

K-means divides a given sample set into K clusters based on the distance between samples. Make the points within the cluster as closely connected as possible, while keeping the distance between clusters as large as possible. Assuming the cluster is divided into $(\mathbf{C}_1, \mathbf{C}_2, \cdots, \mathbf{C}_k)$, our goal is to minimize the squared error Er as well as the following equation,

$$Er = \sum_{i=1}^{k} \sum_{x \in C_i} ||x - \mu_i||_2^2 \tag{8}$$

where μ_i is the mean vector of C_i, sometimes also known as the center of mass, expressed as:

$$\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x \tag{9}$$

It is obvious that we take $k = 2$ because the dimensions of each of the four categories in the FSLSM model are divided into two poles. After being processed by mapping fig. 5, the materials feature related to each FSLSM dimension are fed into K-means algorithm. Its results can be explained as each cluster representing a group of students that they are very close in vector distance.

V. Experimental Results and Performance Analysis

In this section, we use real datasets to validate the effectiveness of the method we propose. First, introduce the dataset we use in the experiment and preprocess it, then conduct the experiment. Second, we introduced the evaluation of the experimental as well as the analysis and discussion of the results. Finally, we compare the performance of our approach with other existing studies.

A. *Datasets Description and Preprocessing*

To assess the efficacy of the proposed GAT-LS, we utilized the identical dataset employed in the evaluation of GRL-LS approach by Muhammad in 2022. The files comprise behavioral records from the 2015 KDD Cup. The dataset is sourced from XuetangX, a Chinese MOOC learning platform established by Tsinghua University. This study documents the online learning patterns exhibited by students across 40 courses from autumn 2013 to spring 2014. The entirety of the user's behavioral data was stored in the enrollment system. The CSV file comprises seven behavior records: access, conversation, navigation, page-close, problem, video, and wiki. Table I presents a comprehensive breakdown of the dataset information of the dataset.

Table I. The information about the 2015 kdd cup dataset

Dataset	KDD Cup 2015
No. of students	5069
No. of materials	07
event records	27,163

B. *Experiment*

Interaction information of the dataset was utilized for us to construct an S-M bipartite graph and each student node has a node feature vector which represents the frequency of student access to materials, opposite to it each material node has a vector shows mapping relationship with FSLSM. The bipartite graph we obtained has 5069 student nodes and 7 material nodes, complex interaction information is comprised inside in high dimensional. For that, we first develop muti-head GAT to obtain full graph information and most importantly to get information on the importance of neighbors for each student node, then drop-edge mechanism and single-layer attention are used for materials nodes to extract its hidden representations.

Libraries numpy, pandas, pytorch etc. are implemented to our GAT-LS. A three-head GAT was applied with a dropout rate of 0.2 to prevent overfitting, also the parameter alpha of leakyRelu is set to 0.2 to ensure that all nodes have non-zero values during gradient updates. As for material nodes, we randomly remove some edges with a probability $p = 0.7$

when extracting features, and we set alpha of leakyRelu to 0.2, dropout rate to 0.5 to start experiment.

The latent represention of all nodes have been determined and the mapping with FSLSM classification makes material nodes equivalent to getting label. We set $k = 2$, calculating the center value of each cluster based on the mapping of material resources to per FSLSM pole. Then K-means clustering algorithm group similar vertors closely and put dissimilar ones in a different group. The clustering results as well as the results of learning style detection are shown in Table II.

Table II. The clustering results

FSLSM Dimension	FSLSM Categories	No. of Learner
Input	Visual	3,104
	Verbal	1,968
Processing	Active	1,869
	Reflective	3,204
Perception	Sensitive	4,742
	Intuitive	330
Understanding	Sequential	1889
	Global	3,183

C. Evaluation of GAT-LS

To evaluate the effectiveness of our proposed GAT-LS method, the accuracy, percision, recall, and F1 score are used on the obtained clustering results. Machine learning models require quantitative evaluation metrics to evaluate which models perform better, the above four indicators are commonly used. Next is a brief introduction to them.

1) Accuracy

Accuracy is the most straightforward indicator for measuring a classification model, it represents the proportion of correctly classified samples to the total number of samples. A correctly classified sample consists of two parts, a situation where the prediction is positive and the reality is positive, namely TP; there is also a situation where the prediction is negative and the reality is also negative, namely TN. The calculation equation for Accuracy is as follows

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

2) Precision

Precision represents the proportion of samples with positive predicted results that are actually positive. The calculation equation for precision is as follows:

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

3) Recall

Recall represents the proportion of the actual number of positive samples in the predicted result to the total number of positive samples in the entire sample. The calculation equation for recallis as follows:

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

4) F1 score

F1 score takes into account both precision and recall factors and is a weighted average of them. The calculation equation for F1 score is as follows:

$$F1\ score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{13}$$

D. Results and Discussion

This section provides a comprehensive evaluation of the effectiveness of the proposed GAT-LS in relation to the clustering outcomes, utilizing four key metrics: Accuracy, Precision, F1 Score, and Recall. These evaluation metrics will provide us with a multifaceted understanding of the model's performance and enable a thorough analysis and assessment of the clustering results.

Table III displays the results for accuracy, precision, recall, and F1 score. The average values obtained were 0.9647, 0.9478, 0.9171, and 0.9346, respectively, for all the FSLSM dimensions. Based on the results obtained, it is evident that the proposed GAT-LS approach outperforms in terms of accuracy, precision, recall, and F1 score when applied to the final clustering result across all four dimensions of FSLSM. The experimental results demonstrate that the GAT-LS approach has strong performance and efficiently accomplishes the task of learning style detection.

In the next stage of our study, we will compare the performance of the proposed method against widely adopted methods in the current domain. This comparison aims to reveal the superiority of our method or improvements in specific aspects.

Table III. Evaluation of GAT-LS approach for each classification of FSLSM

FSLSM / Metrics	Vis/Ver.	Act/Ref.	Sen/Int.	Seq/Glo.	Avg.
Accuracy	0.9864	0.9603	0.9553	0.9568	**0.9647**
Precision	0.9783	0.9384	0.9309	0.9437	**0.9478**
Recall	0.9680	0.9054	0.8925	0.9027	**0.9171**
F1_score	0.9731	0.9216	0.9113	0.9227	**0.9346**

E. Performance Comparison

In this section, we compare the performance of the GAT-LS approach with the existing learning style detection and prediction methods. We have selected a wide range of representative methods that uses the same dataset and closely related as a point of the comparison, however, due to the different evaluation indicators of different studies, we divided the results comparison into two parts. The first part is the comparison of Accuracy and the second part is the comparison of Precision.

1) Comparison of the GAT-LS approach with existing approaches based on Accuracy

We compare our approach, GAT-LS, with the existing approaches in terms of accuracy. We consider existing approaches that use the same dataset to perform the comparison [21], [29], [41], [12], [2], [2], and some other related works. The performance comparison is shown in Table IV.

Table IV. Comparison of accuracy results

FSLSM dimension	Approach	Accuracy
Vis/Ver	LSID-ANN	0.8400
	Deles	0.7880
	KNN	0.881
	LGBM	0.8325
	GRL-LS	0.9100
	GNN-LS	0.9657
	GAT-LS	**0.9864**
Act/Ref	LSID-ANN	0.8020
	Deles	0.7990
	KNN	0.809
	LGBM	0.8325
	GRL-LS	0.9300
	GNN-LS	0.9665
	GAT-LS	**0.9603**
Sen/Int	LSID-ANN	0.7900
	Deles	0.7900
	KNN	-
	LGBM	0.8325
	GRL-LS	0.8500
	GNN-LS	0.9632
	GAT-LS	**0.9553**
Seq/Glo	LSID-ANN	0.7970
	Deles	0.7020
	KNN	0.881
	LGBM	0.8325
	GRL-LS	0.8400
	GNN-LS	0.9696
	GAT-LS	**0.9568**

As shown in Table IV, our proposed GAT-LS method has demonstrated competitive performance in terms of the accuracy evaluation indicator. Notably, our method exhibits substantial advantages in accuracy when compared to the approaches presented in article [21], [41], [12]. In comparison to GRL-LS, our proposed method surpasses it across all dimensions of FSLSM. However, despite excelling in the input dimension, we achieved the highest value in the comparison with our method reaching 0.9864 and GNN-LS method attained 0.9657, our algorithm exhibits slightly lower performance than GNN-LS in the remaining three dimensions and average values. The observed outcome is attributed to the fact that GNN-LS underwent 200 rounds of high computational power iterations, thereby achieving superior results. Moreover, GNN-LS uses a classification algorithm to predict the learner's learning style in the real world, while our GAT approach employs the use of a clustering algorithm to detect learning style. Using classification algorithms in the future to predict learning styles in the real world can provide a more effective performance of our GAT-LS approach. Consequently, in comparison to the results reported in other studies, our GAT-LS demonstrates competitive performance in terms of accuracy.

2) Comparison of the GAT-LS approach with existing approaches based on Precision

We further compare the performance of our approach, GAT-LS, with the existing approaches based on precision. The precision results were compared with the existing approach proposed by [56], [22], [41], [13], [45], [57]. Table V displays the performance comparison.

Table V. Comparison of precision results

FSLSM dimension	Approach	Precision
Vis/Ver	Graf	0.7667
	Liyanage	0.7625
	Karagiannis	0.9216
	Khan	0.6356
	Bernard	0.8020
	GRL-LS	0.8300
	GAT-LS	**0.9783**
Act/Ref	Graf	0.7933
	Liyanage	0.6500
	Karagiannis	0.7500
	Khan	0.6547
	Bernard	0.7410
	GRL-LS	0.8700
	GAT-LS	**0.9384**
Sen/Int	Graf	0.7733
	Liyanage	0.7500
	Karagiannis	0.6600
	Khan	0.6496
	Bernard	0.7270
	GRL-LS	0.7300
	GAT-LS	**0.9309**
Seq/Glo	Graf	0.7333
	Liyanage	0.7750
	Karagiannis	0.8000
	Khan	0.7121
	Bernard	0.8250
	GRL-LS	0.7100
	GAT-LS	**0.9437**

For each FSLSM dimension, we can see that our GAT-LS approach is successful and achieves higher precision compared to the existing approaches with 0.9783 for visual/verbal, 0.9384 for active/reflective, 0.9309 for sensitive/ intuitive, and 0.9437 for sequential/global precision respectively. Based on the average, it is evident that the average precision of our proposed method stands at 0.9478, significantly surpassing the precision values of other methods listed in Table V. This suggests the proposed approach exhibits high performance, the integration of graph-based analysis and attention mechanism in our method for feature vector creation and label formation has markedly enhanced the effectiveness of learning style detection.

We further support our proposed GAT-LS approach by conducting further comparisons with the existing methods to show the effectiveness of the proposed approach. For that, we consider using recall and F1 score to compare the performance of our approach to that of the existing methods presented by [12], [29], [45], these methods employ the use of a deep neural network algorithm (DNN), Light Gradient Boosting Machine (LGBM), and Graph representation techniques for learning style detection. Table VI shows the comparison results.

Table VI. Comparison of recall and F1 score results

FSLSM dimension	Approach	Recall	F1 score
Vis/Ver	KNN	0.8800	0.8800
	LGBM	0.8457	0.8071
	GRL-LS	0.9216	0.852
	GAT-LS	**0.9680**	**0.9731**
Act/Ref	KNN	0.8100	0.8700
	LGBM	0.8457	0.8071
	GRL-LS	0.9216	0.852
	GAT-LS	**0.9054**	**0.9216**
Sen/Int	KNN	-	-
	LGBM	0.8457	0.8071
	GRL-LS	0.9216	0.852
	GAT-LS	**0.8925**	**0.9113**
Seq/Glo	KNN	0.8800	0.8800
	LGBM	0.8457	0.8071
	GRL-LS	0.9216	0.852
	GAT-LS	**0.9027**	**0.9227**

As can be seen in Table VI, the recall and F1-score of our GAT-LS outperform the existing approaches with an average of 0.9171, and 0.9346 respectively. This shows that our GAT-LS has superior results and efficiency compared with other methods. However, it is worth noting that the number of recalls is relatively lower than the other three indicators, only 0.9171. This is because the dataset we are using does not have enough features for our model to fully extract. Finding a suitable student behavior dataset is also a major challenge for learning styles.

The results of the experiment and comparisons demonstrate the superiority and robustness of our proposed GAT-LS method in efficiently achieving style recognition and detection in learning. The introduction of an attention mechanism allows the feature extraction process to focus more on learners' preferences for learning materials. Our work addresses the issues of accurately conducting learner behavior analysis in the era of big data and effectively implementing learner style detection. Applying our method to online learning systems can achieve more accurate recommendations of learning resources, addressing the problem of knowledge navigation. Furthermore, our approach serves as a pivotal aid for educators and students, enhancing the overall quality of the learning experience.

Besides, another significant revelation is a puzzle that we guess the direct application of authoritative learning style models from the field of education may no longer be suitable for the current era of big data. The motivation behind this observation stems from the comparison of different models. It was observed that, in comparison to GRL-LS, the GAT model shows minimal differences in accuracy, but substantial disparities in clustering results. The complexity of student behavior, after undergoing embedding, results in vectors that can be categorized into both Visual and Verbal clusters. This phenomenon is tentatively termed as "Middle-of-the-road Classification" by us. Investigating the root causes, it is attributed to the relatively coarse granularity of learning style models in the field of education. When transposed into the computer science domain, these models either involve binary or quaternary classifications. Additionally, the lack of efficiency in the models prevents precise vector embedding, leading to inherent fuzziness. A number of papers have already done some work on refining the learning style model, [34] proposed a classification model constructed with an algorithm based on Object-Oriented Bayesian networks, the 16 classifications of the FALSM model are each divided into balanced, moderated and strong. [12] divided each dimension of the FSLSM model into five types instead of two, for instance, (strong verbal, moderate verbal, balanced, moderate visual, strong visual) for input dimensions.

VI. Conclusion

In this paper, our study abstracts the interaction information between students and the learning platform into a bipartite graph with node features. Subsequently, Graph Attention Network (GAT) is employed to extract features from student and material nodes. Multi-head attention is used for student nodes, while material nodes employ a dropEdge mechanism and single-layer attention network to prevent oversmoothing. Finally, K-means clustering is applied for learning style detection based on the FSLSM model.Our approach provides an effective means to understand students' learning styles, offering potential support for personalized learning in the field of education. By delving into the interactions between students and learning resources, we provide educators with a more comprehensive insight, enabling them to better adjust teaching strategies to meet the individual needs arising from student differences.

However, "Middle-of-the-road Classification" problem inspires our future research will focus on further refining the model to better accommodate a broader range of learning environments and disciplines. We plan to explore how our approach can be utilized to offer customized learning experiences for various student populations, while also enhancing performance to improve the accuracy and robustness of learning style classification. Then it is worth noting that the refinement of dimensions in learning style models in the field of education and the paradigm shift of learning style models into the computer science domain pose challenges involving interdisciplinary collaboration and issues of reasonable evaluation. Addressing the "middle-of-the-road classification problem" further will involve the collaboration of scholars from different domains.

VII. ACKNOWLEDGEMENT

This work is supported by the Fundamental Research Funds for the Central Universities (GK202405003), the Shaanxi Provincial Natural Science Foundation project (2024JC-YBMS-543) and the Digital Education Research and Practice Project of Shaanxi Normal University (JYSZH201301).

REFERENCES

[1] M. Raleiras, A. H. Nabizadeh, and F. A. Costa, "Automatic learning styles prediction: A survey of the state-of-the-art (2006–2021)," *Journal of Computers in Education*, vol. 9, no. 4, pp. 587–679, 2022.

[2] B. A. Muhammad, B. Liu, H. K. Ahmad, M. Umar, and K. Qin, "Gnn-ls: A learning style prediction in online environments using graph neural networks," *Journal of Networking and Network Applications*, vol. 2, no. 4, pp. 172–182, 2022.

[3] H. M. Truong, "Integrating learning styles and adaptive e-learning system: Current developments, problems and opportunities," *Computers in human behavior*, vol. 55, pp. 1185–1193, 2016.

[4] J. Feldman, A. Monteserin, and A. Amandi, "Automatic detection of learning styles: state of the art," *Artificial Intelligence Review*, vol. 44, no. 2, pp. 157–186, 2015.

[5] N. R. Ngatirin and Z. Zainol, "Automatic detection of learning styles: A decade review on data-driven approaches," in *Journal of Physics: Conference Series*, vol. 1997, p. 012001, IOP Publishing, 2021.

[6] S. Wibirama, A. P. Sidhawara, G. L. Pritalia, and T. B. Adji, "A survey of learning style detection method using eye-tracking and machine learning in multimedia learning," in *2020 International Symposium on Community-centric Systems (CcS)*, pp. 1–6, IEEE, 2020.

[7] S. Graf and P. Kinshuk, "An approach for detecting learning styles in learning management systems," in *Sixth IEEE International Conference on Advanced Learning Technologies (ICALT'06)*, pp. 161–163, IEEE, 2006.

[8] P. Q. Dung and A. M. Florea, "An approach for detecting learning styles in learning management systems based on learners' behaviours," in *International Conference on Education and Management Innovation*, vol. 30, pp. 171–177, 2012.

[9] M. Hasibuan, L. E. Nugroho, P. I. Santosa, and S. Kusumawardani, "A proposed model for detecting learning styles based on agent learning.," *International Journal of emerging technologies in learning*, vol. 11, no. 10, 2016.

[10] S. V. Kolekar, R. M. Pai, and M. P. MM, "Rule based adaptive user interface for adaptive e-learning system," *Education and Information Technologies*, vol. 24, pp. 613–641, 2019.

[11] O. El Aissaoui, Y. El Alami El Madani, L. Oughdir, and Y. El Allioui, "A fuzzy classification approach for learning style prediction based on web mining technique in e-learning environments," *Education and Information Technologies*, vol. 24, pp. 1943–1959, 2019.

[12] B. Hmedna, A. El Mezouary, and O. Baz, "A predictive model for the identification of learning styles in mooc environments," *Cluster Computing*, vol. 23, no. 2, pp. 1303–1328, 2020.

[13] F. A. Khan, A. Akbar, M. Altaf, S. A. K. Tanoli, and A. Ahmad, "Automatic student modelling for detection of learning styles and affective states in web based learning management systems," *IEEE Access*, vol. 7, pp. 128242–128262, 2019.

[14] J. Yang, Z. X. Huang, Y. X. Gao, and H. T. Liu, "Dynamic learning style prediction method based on a pattern recognition technique," *IEEE Transactions on Learning Technologies*, vol. 7, no. 2, pp. 165–177, 2014.

[15] S. Fatahi, H. Moradi, and E. Farmad, "Behavioral feature extraction to determine learning styles in e-learning environments.," *International Association for Development of the Information Society*, 2015.

[16] R. M. Felder, L. K. Silverman, *et al.*, "Learning and teaching styles in engineering education," *Engineering education*, vol. 78, no. 7, pp. 674–681, 1988.

[17] Z. Othmane, A. DEROUICH, and A. TALBI, "A comparative study of the most influential learning styles used in adaptive educational environments," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 11, 2019.

[18] M. Hasibuan and R. A. Aziz, "Systematic literature review detection learning style," in *2022 International Conference on Platform Technology and Service (PlatCon)*, pp. 29–33, IEEE, 2022.

[19] B. Xing, "A review of research on learning style," *Open Journal of Modern Linguistics*, vol. 13, no. 2, pp. 263–275, 2023.

[20] S. Graf *et al.*, *Adaptivity in learning management systems focussing on learning styles*. PhD thesis, Technische Universität Wien, 2007.

[21] J. Bernard, T.-W. Chang, E. Popescu, and S. Graf, "Learning style identifier: Improving the precision of learning style identification through computational intelligence algorithms," *Expert Systems with Applications*, vol. 75, pp. 94–108, 2017.

[22] M. Prabhani, P. Liyanage, K. S. L. Gunawardena, and M. Hirakawa, "Detecting learning styles in learning management systems using data mining," *J. Inf. Process.*, vol. 24, pp. 740–749, 2016.

[23] H. A. Fasihuddin, G. Skinner, and R. I. Athauda, "Towards adaptive open learning environments: Evaluating the precision of identifying learning styles by tracking learners' behaviours," *Education and Information Technologies*, vol. 22, pp. 807–825, 2017.

[24] M. Jebbari, B. Cherradi, W. Moutaouakil, O. El Gannour, S. Hamida, and A. Raihani, "Prediction and classification of learning styles using machine learning approach," in *2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, pp. 1–7, IEEE, 2023.

[25] H. Y. Ayyoub and O. S. Al-Kadi, "Learning style identification using semi-supervised self-taught labeling," *IEEE Transactions on Learning Technologies*, 2024.

[26] H.-H. Nguyen, K. Do Trung, L. N. Duc, L. D. Hoang, P. T. Ba, and V. A. Nguyen, "A model to create a personalized online course based on the student's learning styles," *Education and Information Technologies*, vol. 29, no. 1, pp. 571–593, 2024.

[27] H.-H. Nguyen, L. Nguyen Duc, K. Do Trung, L. Dang Hoang, T. Vu, T. V. Vu, and V. A. Nguyen, "Applying machine learning techniques to detect student's learning styles," in *Proceedings of the 14th International Conference on Education Technology and Computers*, pp. 456–462, 2022.

[28] J. Bernard, E. Popescu, and S. Graf, "Improving online education through automatic learning style identification using a multi-step architecture with ant colony system and artificial neural networks," *Applied Soft Computing*, vol. 131, p. 109779, 2022.

[29] R. Nazempour and H. Darabi, "Personalized learning in virtual learning environments using students' behavior analysis," *Education Sciences*, vol. 13, no. 5, p. 457, 2023.

[30] H. J. Cha, Y. S. Kim, S. H. Park, T. B. Yoon, Y. M. Jung, and J.-H. Lee, "Learning styles diagnosis based on user interface behaviors for the customization of learning interfaces in an intelligent tutoring system," in *Intelligent Tutoring Systems: 8th International Conference, ITS 2006, Jhongli, Taiwan, June 26-30, 2006. Proceedings 8*, pp. 513–524, Springer, 2006.

[31] E. Özpolat and G. B. Akar, "Automatic detection of learning styles for an e-learning system," *Computers & Education*, vol. 53, no. 2, pp. 355–367, 2009.

[32] P. García, A. Amandi, S. Schiaffino, and M. Campo, "Evaluating bayesian networks' precision for detecting students' learning styles," *Computers & Education*, vol. 49, no. 3, pp. 794–808, 2007.

[33] N. Hidayat, R. Wardoyo, A. Sn, and H. D. Surjono, "Enhanced performance of the automatic learning style detection model using a combination of modified k-means algorithm and naive bayesian," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 3, pp. 638–648, 2020.

[34] Y. P. Valencia Usme, M. Normann, I. Sapsai, J. Abke, A. Madsen, and G. Weidl, "Learning style classification by using bayesian networks based on the index of learning style," in *Proceedings of the 5th European Conference on Software Engineering Education*, pp. 73–82, 2023.

[35] I. Azzi, A. Jeghal, A. Radouane, A. Yahyaouy, and H. Tairi, "A robust classification to predict learning styles in adaptive e-learning systems," *Education and Information Technologies*, vol. 25, pp. 437–448, 2020.

[36] F. Rasheed and A. Wahid, "Learning style detection in e-learning systems using machine learning techniques," *Expert Systems with Applications*, vol. 174, p. 114774, 2021.

[37] H. Y. Ayyoub and O. S. Al-Kadi, "Learning style identification using semisupervised self-taught labeling," *IEEE Transactions on Learning Technologies*, vol. 17, pp. 1093–1106, 2024.

[38] M. Ndognkon Manga and M. Fouda Ndjodo, "An approach for non-deterministic and automatic detection of learning styles with deep belief net," in *Intelligent Computing: Proceedings of the 2021 Computing Conference, Volume 2*, pp. 427–452, Springer, 2021.

[39] M. Gao, X. He, L. Chen, T. Liu, J. Zhang, and A. Zhou, "Learning vertex representations for bipartite networks," *IEEE transactions on knowledge and data engineering*, vol. 34, no. 1, pp. 379–393, 2020.

[40] F. A. Dorça, L. V. Lima, M. A. Fernandes, and C. R. Lopes, "Comparing strategies for modeling students learning styles through reinforcement learning in adaptive and intelligent educational systems: An experimental analysis," *Expert Systems with Applications*, vol. 40, no. 6, pp. 2092–2101, 2013.

[41] I. Karagiannis and M. Satratzemi, "An adaptive mechanism for moodle based on automatic detection of learning styles," *Education and Information Technologies*, vol. 23, pp. 1331–1357, 2018.

[42] "Improving online education through automatic learning style identification using a multi-step architecture with ant colony system and artificial neural networks," *Applied Soft Computing*, vol. 131, p. 109779, 2022.

[43] Q. Ni, Y. Mi, Y. Wu, L. He, Y. Xu, and B. Zhang, "Design and implementation of the reliable learning style recognition mechanism based on fusion labels and ensemble classification," *IEEE Transactions on Learning Technologies*, vol. 17, pp. 241–257, 2024.

[44] B. A. Muhammad, Z. Wu, and H. K. Ahmad, "A conceptual framework for detecting learning style in an online education using graph representation learning," in *2020 International Conference on Networking and Network Applications (NaNA)*, pp. 136–140, IEEE, 2020.

[45] B. A. Muhammad, C. Qi, Z. Wu, and H. K. Ahmad, "Grl-ls: A learning style detection in online education using graph representation learning," *Expert Systems with Applications*, vol. 201, p. 117138, 2022.

[46] B. A. Muhammad, C. Qi, Z. Wu, and H. K. Ahmad, "An evolving learning style detection approach for online education using bipartite graph embedding," *Applied Soft Computing*, vol. 152, p. 111230, 2024.

[47] B. A. Muhammad, W. Jianping, G. Gao, C. Zhao, Q. Li, X. Zuo, and Y. Lv, "A fuzzy c-means algorithm to detect learning styles in online learning environment," *Journal of Networking and Network Applications*, vol. 4, no. 1, pp. 39–47, 2024.

[48] C. Troussas, A. Krouska, and M. Virvou, "A multilayer inference engine for individualized tutoring model: adapting learning material and its granularity," *Neural Computing and Applications*, vol. 35, no. 1, pp. 61–75, 2023.

[49] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," *Advances in neural information processing systems*, vol. 30, 2017.

[50] I. Chami, S. Abu-El-Haija, B. Perozzi, C. Ré, and K. Murphy, "Machine learning on graphs: A model and comprehensive taxonomy," *The Journal of Machine Learning Research*, vol. 23, no. 1, pp. 3840–3903, 2022.

[51] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.

[52] Y. Jiang, H. Ma, Y. Liu, Z. Li, and L. Chang, "Enhancing social recommendation via two-level graph attentional networks," *Neurocomputing*, vol. 449, pp. 71–84, 2021.

[53] C. Zhang, J. James, and Y. Liu, "Spatial-temporal graph attention networks: A deep learning approach for traffic forecasting," *IEEE Access*, vol. 7, pp. 166246–166256, 2019.

[54] H. Linmei, T. Yang, C. Shi, H. Ji, and X. Li, "Heterogeneous graph attention networks for semi-supervised short text classification," in *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*, pp. 4821–4830, 2019.

[55] Y. Rong, W. Huang, T. Xu, and J. Huang, "Dropedge: Towards deep graph convolutional networks on node classification," *arXiv preprint arXiv:1907.10903*, 2019.

[56] S. Graf, Kinshuk, and T.-C. Liu, "Supporting teachers in identifying students' learning styles in learning management systems: An automatic student modelling approach," *Journal of Educational Technology & Society*, vol. 12, no. 4, pp. 3–14, 2009.

[57] J. Bernard, E. Popescu, and S. Graf, "Improving online education through automatic learning style identification using a multi-step architecture with ant colony system and artificial neural networks," *Appl. Soft Comput.*, vol. 131, p. 109779, 2022.

Wanying Suo received her B.Sc degree in Computer Science from Shaanxi Normal University, Xian,China, in 2020. She is currently pursuing the Master degree in Computer Science with Shaanxi Normal University, Xian, China. Her research interests include learning style dection and graph learning.

Bello Ahmad Muhammad received his B.Sc. and M.Sc. degrees in computer science from Bayero University, Kano, Nigeria, in 2010 and 2015, respectively, and received his Ph.D. degree in Computer Science from Shaanxi Normal University, China. He is currently a postdoctoral researcher in the Department of Information Engineering, Henan Institute of Science and Technology, China. His research interests include Learning style detection, graph representation learning, Learning resources recommendation, and Underwater acoustic image communication.

Zhichao Zhang received his M.S degrees from Harbin Institute of Technology, Heilongjiang China in 2011. Now he is presently working as a lecturer at the Hainan Normal University, Hainan, China. He is currently working toward the D.S. degree with the College of Computer Science, Shaanxi Normal University, Shaanxi, China. His research focuses on artificial intelligence in education.

Bo Liu received his M.S. degree from Xi'an Jiaotong University, Xi'an, China and Ph.D. degree in the School of Systems Information Science, Future University Hakodate, Hokkaido, Japan and is also a faculty member in the School of Computer Science, Shaanxi Normal University, Xi'an, China. His current research interests include Nanocommunications and Smart-education.