RESEARCH ARTICLE

The Prediction of Flow in Railway Station Based on RRC-STGCN

XIAOSHU WANG [1], **WEI BAI** [2], **ZHIKANG MENG** [3], **BINBIN XIN** [3], **RUIFENG GAO** [3], **AND XIAOJUN LV** [2]

[1]Postgraduate Department, China Academy of Railway Sciences, Beijing 100081, China
[2]Institute of Computing Technology, China Academy of Railway Sciences Company Ltd., Beijing 100081, China
[3]School of Automation and Software Engineering, Shanxi University, Taiyuan 030006, China

Corresponding author: Wei Bai (baiweisxu@163.com)

ABSTRACT Predicting passenger flow is crucial for effective management and safety in railway stations. Accurate prediction of passenger flow facilitates effective allocation of staff tasks, enhances the efficient utilization of waiting areas, ensures passenger safety, and promotes a smooth travel experience for passengers. However, accurately establishing the spatial-temporal relationship of passenger flow within the station and predicting the passenger flow in each region and time period is challenging due to the varying waiting habits and train preferences of each individual passenger. In this paper, we propose a Residual-RNN-Channel Spatial-Temporal Graph Convolutional Network (RRC-STGCN), which utilizes the channel attention mechanism and residual structure. The model divides the data with a time dimension into multiple periods, capturing spatial-temporal correlations through the channel attention mechanism, and extracting spatial-temporal dependencies from the feature maps using the spatial-temporal convolution module. The model uses a residual structure to fuse features in order to enhance the accuracy of prediction results. In addition, we conduct a comprehensive experimental evaluation using a real dataset of railway station passenger, demonstrating that the RRC-STGCN model outperforms five well-known baselines. Moreover, we provide visualizations of the prediction results, effectively showcasing the dynamic changes in passenger flow in each waiting area.

INDEX TERMS Spatial-temporal graph neural network, deep learning, passenger flow prediction, railway station, residual structure, channel attention.

I. INTRODUCTION

The primary objective of daily operations at railway stations is to ensure the safe and precise boarding and alighting of passengers. To achieve this, the station formulates operational plans by taking into account the total number and distribution of passengers present within the station. This information is crucial in effectively scheduling the tasks of station personnel and avoiding inefficient or insufficient allocation of labor resources. By having a comprehensive understanding of the quantity and the distribution of passengers, the station can efficiently coordinate the duties of its staff, thereby

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu [ID].

preventing any inefficiencies or shortages in the allocation of labor resources. Additionally, dynamically adjusting equipment status based on the quantity and the distribution of passengers within the station contributes to improving the quality of passenger service. In situations where there is a high volume of passenger flow within the station, advance notice of the distribution of this flow and other detailed data, which will help rapid emergency response, efficient emergency disposal, and maximize the guarantee of passenger safety. However, detailed data on the quantity and the distribution of passengers within the station is closely related to individual passenger habits. Each passenger has different preferences for train tickets and waiting habits, which results in varying durations of passenger stays in different areas

within the station. Therefore, it becomes extremely challenging to establish the spatiotemporal relationship efficiently and accurately between the quantity and the distribution of passengers within the station, and to convert the quantity and the distribution of passengers into the passenger flow in each area and time period for accurate prediction.

Numerous scholars have conducted in-depth research on railway passenger flow prediction. From a macro perspective, using the railway passenger transport big data platform to forecast passenger flow across the entire railway network [1]. For the control regions, ticket data from trains passing through a specific section are used to predict future passenger flow in that area [2]. From a micro perspective, the prediction of station passenger flow is utilized to calculate passenger arrival patterns and identify the peak number of passengers congregating [3], [4]. Regarding passenger distribution within stations, due to the difficulty in accurately collecting passenger flow data in different areas, Feng [5] manually counted the number of passengers using video streams from surveillance cameras as a benchmark for prediction. While intelligent video analysis techniques employing deep learning can facilitate passenger flow statistics [6], the placement of surveillance cameras in challenging positions at certain stations hinders the precise collection of passenger flow data in those areas.

The Spatial-Temporal Graph Convolutional Network (STGCN) [7], [8] is a deep learning model specifically designed to handle spatiotemporal data. It is capable of representing spatiotemporal data as a graph structure and utilizing techniques such as graph convolution and attention mechanisms for feature learning and prediction. Given its favorable characteristics, STGNN has been extensively researched in urban traffic flow prediction [7], [8]. STGNN can partition a fixed scene into multiple regions and establish connections between adjacent regions, thereby transforming the problem of traffic flow prediction into a graph modeling and prediction problem. This approach has advantages in capturing the dynamic traffic patterns and spatiotemporal variations between different regions within a city.

Based on the aforementioned features, in this paper, we propose a Residual-RNN-Channel Spatial-Temporal Graph Convolutional Network (RRC-STGCN) based on the channel attention mechanism and residual structure. The model transforms the flow lines within the passenger station into a graph to predict railway station passenger flow. In the model, we divide the data into three different periods by a Multi-Period Partition Module. Then, each period utilizes a Spatio-Temporal Attention Module to capture spatiotemporal correlations. Specially, a Spatio-Temporal Convolution Module is used to extract spatial and temporal features for further processing. Finally, the residual structure is employed to fuse the features extracted from each period and predict passenger flow. The main contributions of this paper can be summarized as follows. (1) Enhancing the efficiency of capturing spatiotemporal correlations by introducing the channel attention mechanism based on the existing attention mechanism.

(2) Mitigating the problem of high model complexity and low computational efficiency by incorporating the residual structure into the prediction of passenger flow. (3) Considering the limitations of conventional methods, the extraction of temporal features is improved by introducing long short-term memory (LSTM). (4) Validating the effectiveness of the proposed model through experiments conducted on a real railway passenger station dataset.

II. RELATED WORK
A. RAILWAY PASSENGER FLOW PREDICTION METHODS

The prediction of railway passenger flow has always been a popular research topic [1], [2], [3], [4], [5]. From a spatial perspective, researchers have focused on predicting passenger flow within the national railway network, narrowing down to local railway lines, and ultimately individual stations. From a temporal perspective, the prediction has been divided into long-term, medium-term, and short-term. The main challenge in passenger flow prediction is how to extract features from existing passenger flow data. Spatial characteristics encompass the attributes of passenger flows across diverse regions, whereas temporal features encapsulate the cyclical patterns and evolving trends of data over time. Prediction methods can be classified into model-driven and data-driven approaches.

1) MODEL-DRIVEN PREDICTION METHODS

Model-driven methods, also known as parameter methods, are widely used in railway passenger flow prediction. These meth-ods include statistical-based model prediction [9], Autoregres-sive Integrated Moving Average (ARIMA) model [10], [11] and so on. In the early stage, Xu et al. [9] used statistical prin-ciples to forecast short-term passenger flow during the spring transportation period. Subsequently, the ARIMA model, which is suitable for predicting strong regular and stationary sequences, demonstrated excellent performance. Xing and Dong [10] proposed that the seasonal ARIMA model better reflects the current patterns of railway passenger flow. Wu [11] conducted passenger flow prediction based on the ARIMA model for Shanghai Hongqiao Station during the COVID-19 pandemic. Liu [12] combined wavelet decomposition and the ARIMA model to construct a weak model for passen-ger flow prediction, and then used the AdaBoost ensemble al-gorithm to aggregate multiple weak models into a strong model for short-term passenger flow prediction. Shi et al. [13] applied wavelet decomposition and the ARIMA model to fore-cast the high-frequency and low-frequency components of passenger flow data for Guangzhou-Zhuhai intercity stations. However, the ARIMA model performs well only under strict assumptions and fixed algorithm structures. Passenger flow at stations, being susceptible to a range of external influences, exhibits considerable uncertainty. There-fore, this method can-not effectively capture the nonlinear characteristics of pedes-trian flow and fails to achieve the desired accuracy.

2) DATA-DRIVEN PREDICTION METHODS

In recent years, due to the development of railway station information systems, there has been a significant accumulation of data related to passenger flow. By analyzing and mining passenger flow data at railway stations, we can extract and identify patterns in passenger flow changes, providing new insights for predicting passenger flow. The data-driven approach primarily consists of traditional machine learning methods and deep learning methods.

Traditional machine learning methods, such as back propagation neural network (BP) [14] and K-nearest neighbors (KNN) [15], can extract nonlinear temporal relationships. The aforementioned approaches can conduct feature modeling on railway passenger flow data, uncovering nonlinear relationships within the data, and thereby achieving a substantial improvement in prediction accuracy. Wang et al. [14] utilized BP neural network to normalize and predict railway passenger volume from 1980 to 1998. Cao [15] used the KNN to accurately forecast short-term railway passenger flow and considered specific factors such as departure date, train number, tickets and origin-destination. However, these methods heavily rely on manual feature engineering, which is powerless when handling large-scale pedestrian data and cannot meet practical needs.

To further improve the accuracy of model predictions, deep learning models driven by big data have become a new re-search focus for predicting passenger flow. Li et al. [16] pro-posed a passenger flow prediction model for stations based on LSTM and validated it using data from the Beijing-Guang-zhou high-speed railway. Liu and Ni [17] utilized Shanghai station data and constructed a BP- Gate Recurrent Unit (GRU)model for short-term passenger flow prediction. Deep learning models have strong nonlinear capabilities, which can better capture the diverse transformation patterns inherent in pedestrian flow, and greatly enhance the prediction accuracy. However, the aforementioned studies only utilized temporal features for prediction and rarely incorporated spatiotemporal features. Feng [5] proposed a short-term prediction of spatiotemporal distribution of passenger by manually counting passenger numbers from videos within high-speed railway sta-tions based on graph embedding and LSTM.

B. PREDICTION RESEARCH IN OTHER FIELDS

City traffic flow and pedestrian flow prediction share similarities with passenger flow prediction at transit stations. Both city traffic flow and pedestrian flow prediction involve calculating the volume of vehicles or pedestrians on a specific road segment from the origin to the destination. This is fundamentally similar to calculating the passenger flow from the station entrance to the platform within transit stations. Therefore, we draw upon the experiences from predicting city traffic flow and pedestrian flow to study the issue of passenger flow prediction at transit stations.

1) CITY TRAFFIC PREDICTION

Li et al. [18] and Wu et al. [19] proposed the Diffusion Convolutional Recurrent Neural Network (DCRNN) [18] and Graph WaveNet for Deep Spatial-Temporal Graph [19], respectively, as deep learning frameworks for traffic prediction. These models combine spatial and temporal dependencies in traffic flow. Xu et al. [20] introduced the Spatial-Temporal Transformer Network, which utilizes the multi-head attention mechanism in Transformer to model both temporal and spatial correlations.

Due to the irregularity of graph structures, traditional deep learning models are unable to handle this type of data. The emergence of Graph Neural Networks (GNN) [21] has attracted extensive attention from scholars. GNN and Graph Convolutional Network (GCN) [22] demonstrate remarkable capabilities in dealing with unstructured data, making various spatiotemporal learning models based on GNN and GCN become the trend and focus of research in the field of urban road traffic flow and pedestrian flow prediction. He et al. [21] proposed Spatial-Temporal Transformer Networks, which can utilize the global-dynamic information for long-term prediction. Zhao et al. [22] proposed a combined prediction method called Temporal Graph Convolutional Network, which integrates GCN and GRU [23]. Chen et al. [24] proposed an enhanced GCN model that incorporates node embedding [25] to aggregate weighted spatial features and extract temporal features by using the spatial features as input to the GRU model. Song et al. [26] proposed the Spatial-Temporal Synchronous Graph Convolutional Networks (STSGCN), stacking multiple Spatial-Temporal Synchronous Graph Convolutional Layers to capture localized spatial-temporal correlations. Li and Zhu [27] introduced the Spatial-Temporal Fusion Graph Neural Networks (STFGNN), incorporating a Gated Dilated CNN module to simultaneously capture local and global correlations. Guo et al. [8] proposed the Attention Based Spatial-Temporal Graph Convolutional Networks (ASTGCN), which can directly process traffic data on the original graph-based traffic network. This model introduces attention mechanisms to learn the dynamic temporal and spatial correlations in traffic flow data, effectively capturing the dynamic spatiotemporal features of traffic data. These methods effectively address the problem of urban road traffic flow prediction and provide new insights about predicting passenger flow at railway stations for this study.

2) CITY PEDESTRIAN FLOW PREDICTION

Yu et al. [7] proposed the Spatio-Temporal Graph Convolutional Networks (STGCN), which are capable of handling pedestrian flow prediction. By combining graph convolution and causal convolution, STGCN can extract both temporal and spatial features. The network consists of a convolutional structure that allows for parallel computing, enabling efficient learning the spatial and temporal correlations in pedestrian flow through the spatiotemporal modules composed of graph

convolution and temporal convolution. Cardia et al. [28] presented a GCN-based solution for predicting crowd flow, specifically on predicting pedestrian flow in urban public spaces. Chen et al. [29] proposed a toolbox for spatiotemporal crowd flow prediction, named Urban Computing Toolbox, which integrates multiple GCN models. The toolbox enables spatiotemporal prediction of crowd flow in smart cities. It is evident that considering both temporal and spatial correlations in the data is an inevitable trend in research. The use of spatiotemporal graphs for predicting traffic flow or pedestrian flow is currently a mainstream approach. Building upon this, the RRC-STGCN method is proposed and applied to predict passenger flow at railway stations.

III. METHODS

A. PROBLEM FORMULATION

To address the issue of predicting the distribution of passengers within a railway station, we draw inspiration from traffic flow prediction that transforming the passengers distribution of the station into a problem of predicting the flow at key nodes. The passenger route of the station consists of the following steps: arrival at the entry, real-name verification, baggage screening, waiting, ticket check-in, and boarding. After entering the station, the general purpose of passengers is to board a train and leave the station to reach their destination. Passengers choose a waiting area near the ticket inspection point for the train which they have purchased a ticket for. Therefore, we consider the starting point of the passenger's route as the entrance and the endpoint as the platform. This process parallels the progression of a vehicle from its origin to its designated endpoint in urban traffic. Based on this, we transform the station's internal areas into nodes of a personnel flow network, defined as a graph. By predicting the flow at the nodes of this graph, we can achieve the prediction of passengers distribution within the station.

A graph can be represented as $G = (V, E, A)$, where V is the set of nodes in the graph, E is the set of edges in the graph, and A represents the adjacency matrix, which is a two-dimensional matrix that represents the connection relationship between nodes in the graph. Specifically, the V in the network represent the locations of the collectors, such as real-name verification gates and check-in gates. E represent the flow lines of passengers connecting collectors locations. The readings of the collectors are represented as attributes that change over time on the nodes. The Laplacian matrix of the graph can fully reflect the structural properties of the graph, describing the differences between central nodes and surrounding nodes, and is often used to represent graphs. Moreover, according to spectral graph theory, in order to improve computational efficiency, the Laplacian matrix is often transformed into the normalized form. The calculation formula is as Eq.(1), (2), and (3), where L_{sym} represents the normalized form of the Laplacian matrix, L represents the Laplacian matrix, D is a diagonal matrix that represents the degree matrix, and I represents the identity matrix. Its essential meaning is to replace the diagonal of the
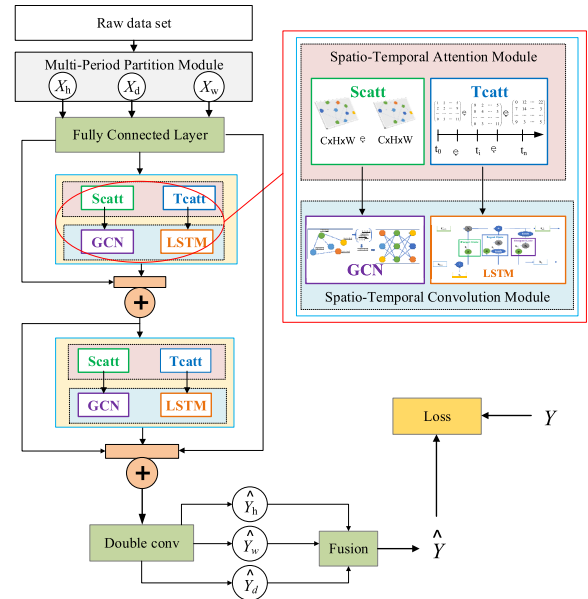

FIGURE 1. RRC-STGCN architecture. Assume that the input flow data shape is (32, 10, 1, 12), where 32 is batch size, 10 is the vertice count, 1 is the feature count, and 12 is timestep count. When it passes through the spatio-temporal attention module and the spatio-temporal convolution module, the data is first processed by Tcatt and the shape becomes (32, 12, 12). Then the data is fed into the Scatt module and the shape turns into (32, 10, 10). The shape of the data processed by GCN is (32, 10, 64, 12), and the final data passes through LSTM and the shape becomes (32, 12, 64, 10).

adjacency matrix with 1, and the other 1s representing the edge are replaced by the opposite of the -1/2 power of the degree product of the two vertices connected by the edge.

$$L_{sym} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}} = I - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} \quad (1)$$

$$L = D - A \quad (2)$$

$$D_{ii} = \sum_{j=1}^{N} A_{ij} \quad (3)$$

Any node in the graph G represents the location of a data collector. We set the data collector to record the cumulative value of passengers passing through it in a certain time period. This cumulative value exhibits a certain time series pattern. Therefore, the problem of predicting the flow at the nodes of the graph can be transformed into predicting the future flow sequence for a certain period by using the various historical measurements of all nodes on the graph G over the past time intervals.

B. RRC-STGCN ARCHITECTURE

In this paper, we propose a Residual-RNN-Channel Spatial-Temporal Graph Convolutional Network (RRC-STGCN) to predict the flow at railway station. The overall framework of RRC-STGCN is shown in Figure 1. The model consists of three modules: Multi-Period Partition Module, Spatio-Temporal Attention Module, Spatio-Temporal Convolution Module, and Multiple Components Fusion. The Multi-Period Partition Module maps the information to a high-dimensional space using a Fully Connected Layer (FC) to enhance
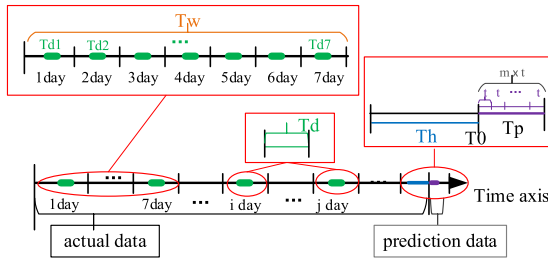
FIGURE 2. The meaning of T_w, T_d, T_h, T_p.

FIGURE 3. Channel attention.

the model's performance. This module divides the original data into three components and performs feature extraction on each component separately. The Spatio-Temporal Attention Module enables the network to automatically pay more attention to valuable information. The Spatio-Temporal Convolution Module extracts features from the information that the attention module focuses on. The extracted feature information from each component is fused through multiple consecutive convolutions to predict the future passenger flow in the waiting area. Residual connections are added in the model to enhance the representation of spatio-temporal features and improve the efficiency of the model.

C. MULTI-PERIOD PARTITION MODULE

Inspired by the paper [8] and [30], in real-world scenarios of waiting rooms at train stations, pedestrian flow data exhibits different periodic characteristics during different time intervals. Assuming that the original pedestrian flow data at the station has a sampling period of t minutes, with the current time point as the starting time T_0, the future prediction period is T_p, satisfying $T_p = m \times t$, where m is a positive integer. Within the past time interval, original data of lengths T_d, T_h, and T_w are respectively extracted as input for the cycle component: recent X_h, daily X_d, and weekly X_w, as shown in Figure 1. Since the near-cycle component is close to the predicted cycle, the flow of people is directly affected by it. Since the daily frequency of trains at the station is basically the same, it is intuitively manifested as a certain repeatability of the daily cycle component, such as during the afternoon peak period every day. Due to the regularity of travel planning, it is manifested in the certain regularity of the weekly component, such as the train plan is divided into daily lines and weekend lines. Additionally, T_d, T_h, and T_w are all multiples of T_p, as illustrated in Figure 2. For example, T_p represents a time period from 19:30 to 20:00 on May 21, 2021. T_h represents a time period from 17:30 to 19:00 on May 21, 2021. T_d represents time periods from 19:30 to 20:00 on May 19 and 20, 2021. Finally, T_w represents a time period from 19:30 to 20:00 between May 7 and May 13, 2021.

The recent cycle component T_h represents the segment of original data that is closest to the starting time T_0, and it represented by as $X_h = \{x_{t_0-t_h+1}, x_{t_0-t_h+2}, \cdots, x_{t_0}\}$. The variation of passenger flow at the station is a gradual process, and the changes in the recent cycle component directly affect the magnitude of T_P. Similar to the recent cycle component,
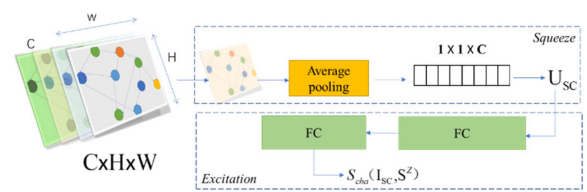
the daily cycle component T_d and the weekly cycle component T_w are specifically expressed as follows:

$$X_d = \{x_{t_0-(T_d \div T_p)*q+1}, \ldots, x_{t_0-(T_d \div T_p)*q+T_p},$$
$$\ldots, x_{t_0-(T_d \div T_p-1)*q+1}, \ldots, x_{t_0-(T_d \div T_p-1)*q+T_p},$$
$$\ldots, x_{t_0-q+1}, x_{t_0-q+T_p}\} \quad (4)$$
$$X_w = \{x_{t_0-7*(T_d \div T_p)*q+1}, \cdots, x_{t_0-7*(T_d \div T_p)*q+T_p},$$
$$\cdots, x_{t_0-7*(T_d \div T_p-1)*q+1}, \cdots, x_{t_0-7*(T_d \div T_p-1)*q+T_p},$$
$$\cdots, x_{t_0-7*q+1}, \cdots, x_{t_0-7*q+T_p}\} \quad (5)$$

where q represents the step length.

D. SPATIO-TEMPORAL ATTENTION MODULE

When describing the spatio-temporal correlation of station passenger flow, it is closely related to the individual habits of each passenger, such as the train they purchased, the waiting time, and their waiting habits, which can vary greatly. In order to reduce the impact of these factors on the prediction results, the Spatio-Temporal Attention Module is selected to focus on the key information of the input data. The Spatio-Temporal Attention Module consists of two parts: the Spatial Channel Attention (SCAtt) and the Temporal Channel Attention (TCAtt), which capture spatial and temporal correlations respectively.

1) SPATIAL CHANNEL ATTENTION (SCATT)

At the spatial scale, each waiting area is closely related to the train purchased by the passengers. In order to model this attribute, the spatial channel attention mechanism is used to adaptively learn the dynamic connection between passengers and waiting areas, known as spatial dynamic correlation. The Spatial Channel Attention module (SCAtt) consists of a Channel Attention module and a Spatial Attention module. The Channel Attention module is used to extract the importance of each channel, while the Spatial Attention module is used to extract the importance of different regions in the image. To improve the ability to capture features, the Channel Attention module is first applied to focus on important channel information, followed by the Spatial Attention module. The weighted feature maps using attention weights are then used to emphasize important locations, as some in [31].

a: CHANNEL ATTENTION

The Channel Attention module is utilized to extract the importance of each channel, enabling more accurately focus on and utilize of the key information provided by each channel. The Channel Attention, as shown in Figure 3,

consists of two parts: the *Squeeze* operation and the *Excitation* operation [32].

Assuming that the spatial feature map S^Z has dimensions $C*H*W$, where C the number of channels, H represents the height, W represents the width. S_C^Z represents the C-th two-dimensional matrix in the time series feature map S, where C corresponds to the channel. The *Squeeze* operation is a feature compression technique applied to the input feature map in the spatial dimension. Its principle involves using global average pooling to transform the feature map S^Z into a feature vector of size $1 \times 1 \times C$, as shown in Eq. (6). Here, Sq($\bullet$) represents the *Squeeze* operation, and U_{SC} represents the feature vector obtained after feature compression.

$$U_{SC} = Sq(S_C^Z) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} S_C^Z(i,j) \qquad (6)$$

The *Excitation* operation models the interdependencies between the C channels using two fully connected layers. After each fully connected layer, an activation function is applied for stimulation, as shown in Eq. (7). Here, I_{SC} represents the weighted coefficients after activation, $Ec(\bullet)$ represents the *Excitation* operation, ω_1 represents the weight values of the first fully connected layer, $\delta(\bullet)$ represents the *ReLU* activation function of the first fully connected layer. ω_2 represents the weight values of the second fully connected layer, and $\sigma(\bullet)$ represents the *sigmoid* activation function of the second fully connected layer.

$$I_{SC} = Ec(U_{SC}) = \sigma(\omega_2 \delta(\omega_1 U_{SC})) \qquad (7)$$

We recalibrate the features by combining the weighted coefficients I_{SC} after activation with the input feature map and S^Z, obtaining the weighted feature map after channel attention processing, as shown in Eq. (8).

$$S_{cha}(I_{SC}, S^Z) = I_{SC} \times S^Z \qquad (8)$$

b: SPATIAL ATTENTION

To simplify the equation, we assume that the weighted feature map $S_{cha}(I_{SC}, S^Z)$ as X_Z, obtained from the channel attention module, serves as both the output of the Channel Attention and the input for the Spatial Attention. The calculation method is shown in Eq. (9). Here, $X_Z \in \{X_h, X_d, X_w\}$ represents the input feature map for the spatial attention mechanism, Y_1, Y_2, Y_3, b_s denotes the learnable dynamic parameters, and $\sigma(\bullet)$ is the activation function. In this case, the *ReLU* activation function is chosen to process the data and generate the spatial similarity score matrix S_{att}. The *softmax* function is then applied to normalize the spatial similarity score matrix, ensuring that the sum of the correlation strength coefficients is 1. This yields the Spatial Attention strength coefficients. Multiplying these coefficients with the original feature map results in the feature-weighted feature map, which serves as the output of the spatial channel attention module, as Eq. (10). Here, $S_{i,j}^Z$ represents the spatial weighted feature map $Z \in \{h, d, w\}$ between node i and node j in the Z

time series.

$$S_{att} = \sigma((X_Z Y_1) Y_2 (X_Z Y_3)^T + b_s) \qquad (9)$$

$$S_{i,j}^Z = X_Z \bullet softmax(S_{att}) = X_i \bullet \frac{\exp(S_{att}^{i,j})}{\sum_{j=1}^{N} \exp(S_{att}^{i,j})} \qquad (10)$$

2) TEMPORAL CHANNEL ATTENTION (TCATT)

At the temporal scale, the passenger flow in each waiting room at the station is mutually influenced at different time periods. For example, a large number of passengers entering the wait-ing room in the previous moment will inevitably lead to a sharp increase in the passenger flow. When the train arrives at the station, it will then result in a decrease in the passenger flow in the waiting room. Therefore, captur-ing the character-izes the temporal correlation is necessary. In this section, the Time Channel Attention mechanism is selected to dynami-cally assign different levels of impor-tance to the temporal maps. The Temporal Channel Atten-tion consists of the Chan-nel Attention and the Temporal Attention.

a: CHANNEL ATTENTION

The processing principles of the Chan-nel Attention are described in Eq. (11), (12), (13), and (14). Here, $T_C^Z(i,j)$ represents the element at position (i,j) in the C-th matrix of the time map T in the Z time series. U_{TC} denotes the compressed time map feature vector. ω_3 and ω_4 represent the weights of each fully connected layer. $\delta(\cdot)$ and $\sigma(\cdot)$ correspond to the activation functions of ω_3 and ω_4, respectively. $S_{cha}(U_{TC}, J_C)$ represents the channel attention fea-ture map after weighted processing.

$$J_C = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} T_C^Z(i,j) \qquad (11)$$

$$U_{TC} = \sigma(\omega_4 \delta(\omega_3 J_C)) \qquad (12)$$

$$S_{cha}(U_{TC}, J_C) = U_{TC} \times J_C \qquad (13)$$

$$Y^Z = S_{cha}(U_{TC}, J_C) \qquad (14)$$

b: TEMPORAL ATTENTION

The processing principles of the Temporal Attention are described in Eq. (15) and (16). T_{att} represents the temporal attention weights. V_t, U_1, U_2, U_3, and b_t are the parameters that need to be trained. Y^Z represents the input of the temporal attention mechanism. $T_{att}^{i,j}$ represents the correlation strength between time i and time j, which is normalized by the softmax function. Multiplying this normalized value with the input feature map yields the atten-tion-weighted annotated feature map $T_Z^{i,j}$ of the attention mechanism.

$$T_{att} = V_t \cdot \sigma(((x_{i,j})^T U_1) U_2 (U_3 x_{i,j}) + b_t) \qquad (15)$$

$$T_Z^{i,j} = Y^Z \cdot \frac{\exp(T_{att}^{i,j})}{\sum_{j=1}^{N} \exp(T_{att}^{i,j})} \qquad (16)$$
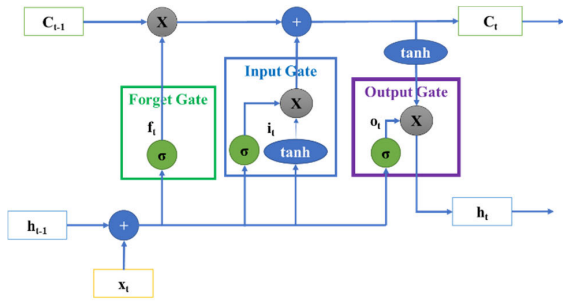
FIGURE 4. The structure of LSTM [35], [36].

E. SPATIO-TEMPORAL CONVOLUTION MODULE

After being processed by the Spatio-Temporal Attention Module, a new spatio-temporal feature map is obtained, which consists of feature maps with different attention weights assigned to important information. The output of the spatio-temporal attention module is then fed into the Spatio-Temporal Convolution Module for further processing, extracting spatial dependencies from the spatial feature map and temporal dependencies from the temporal feature map. In this section, GCN and LSTM are combined to extract spatiotemporal dependencies.

1) GRAPH CONVOLUTIONAL NETWORKS (GCN)

As analyzed in Section III-A, the problem in this paper is transformed into a graph $G = (V, E, A)$. To address this, we adopt GCN [8]. GCN is a type of convolutional neural network that directly operates on graphs and utilizes their structural information, making it suitable for handling irregular graph-structured data. GCN performs convolutional operations on the graph structure to extract information from the features of nodes and their neighboring nodes. It updates the representation of each node based on the features of its neighboring nodes, thereby leveraging the global information of the graph structure. Each convolutional layer only processes the first-order neighborhood information. By stacking multiple convolutional layers, high-order neighborhood information can be propagated [33]. The specific information propagation rule for each layer is given by Eq. (17), where H_{l+1} represents the feature information matrix of the $l + 1$-th layer, $\sigma(\bullet)$ denotes the non-linear activation function, and W_l represents the weight coefficients of the l-th layer.

$$H_{l+1} = \sigma(L_{sym}H_lW_l) \tag{17}$$

Thus, for an N-layer graph convolutional network model, it can be represented by Eq. (18). Here, X represents the feature matrix, which serves as the input to the GCN network model. $\widehat{A}$ is the adjacency matrix obtained by applying Laplacian regularization to the adjacency matrix A. $V_1, \cdots, V_{N-1}, V_N$ and so on represent the weight coefficients of each convolutional layer, while $\sigma(\bullet)$ and $ReLU(\bullet)$ are the activation functions of the GCN network model.

$$F(X, A) = \sigma(\widehat{A}\text{Re}LU(\widehat{A}\text{Re}LU(\cdots, \text{Re}LU(\widehat{A}XV_1))V_{N-1})V_N) \tag{18}$$

2) LONG SHORT-TERM MEMORY (LSTM)

The passenger flow exhibits strong temporal correlations, and some of the flow data also display significant similarities. LSTM [5], [6] is variants of RNN. It demonstrates outstanding performance in processing time series data and effectively ad-dress the issues of gradient vanishing and exploding that are prevalent in traditional RNNs [34]. In this study, LSTM is em-ployed to extract temporal dependencies. LSTM consists of three gates: the input gate, the forget gate, and the output gate. The structure diagram of LSTM [35], [36] is shown in Figure 4.

The specific definitions of each part in the LSTM cell are given as follows.

$$i_t = \sigma(W_{xi} \cdot x_t + W_{hi} \cdot h_{t-1} + b_i) \tag{19}$$
$$f_t = \sigma(W_{xf} \cdot x_t + W_{hf} \cdot h_{t-1} + b_f) \tag{20}$$
$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh(W_{xc} \cdot x_t + W_{hc} \cdot h_{t-1} + b_c) \tag{21}$$

where i_t, f_t and o_t represent the input gate, forget gate and output gate respectively. W_* represents the weight matrices, b_* is the biases, $\sigma(\bullet)$ denotes the activation function, and C_t represents the state of the memory unit at time t. Based on the output gate o_t and the memory state C_t, the new hidden state h_t is computed.

$$o_t = \sigma(W_{xo} \cdot x_t + W_{ho} \cdot h_{t-1} + b_o) \tag{22}$$
$$h_t = o_t \cdot \tanh(C_t) \tag{23}$$

F. MULTIPLE COMPONENTS FUSION

After capturing the spatio-temporal correlations through the Spatio-Temporal Attention Module and extracting the spatio-temporal dependencies with the Spatio-Temporal Convolution Module, the passenger flow data in the waiting area produces outputs for each component. However, each component has a different level of influence on the final prediction result. In or-der to improve the accuracy of the prediction, this paper incor-porates residual connections into the model to enhance the re-liability of the outputs. Additionally, suitable parameter matri-ces are selected based on historical data to effectively blend the outputs.

Assuming there are m attention modules and convolution modules, residual connections can be represented by Eq. (24) and (25), respectively. In the equations, $\widehat{Y_i}$ represents the out-put of any component, $a(m)$ represents the fusion of data ag-gregation by the previous m modules, x_m represents the output of the m-th module, and $f(x_m)$ represents the output af-ter processing by the m-th module.

$$\widehat{Y_i} = comv(a(m) + x_m) \tag{24}$$
$$x_m = x_{m-1} + f(x_m) \tag{25}$$

The fusion of component outputs can be achieved using Eq. (26). In the equation, $\odot$ represents the Hadamard product, W_h, W_d, W_w are the dynamically learned parameters that rep-resent the strengths of different components in the prediction result. These parameters need to be continuously
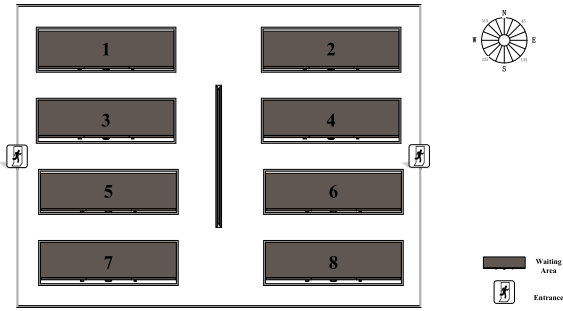
FIGURE 5. The simplified topological structure of the station.

updated based on historical pedestrian flow data to ensure the accuracy of the prediction result.

$$\widehat{Y} = W_h \odot \widehat{Y}_h + W_d \odot \widehat{Y}_d + W_w \odot \widehat{Y}_w \qquad (26)$$

IV. EXPERIMENTS

To validate the effectiveness of the proposed RRC-STGCN for predicting passenger flow in railway stations, comprehensive experiments were conducted using data collected from a real information system of a railway station. The station is located in the northern of China. This section provides a detailed introduction to four aspects: the dataset and evaluation metrics, experimental settings, experimental results and analysis, and visualization.

A. DATASET AND METRICS

1) DATASET

The data set utilized in this paper is derived from actual system data collected at a railway station. The station comprises two entrances oriented in the east-west direction, along with a waiting room. The two entrances are situated on the first floor and each side is equipped with four real-name verification gates and an additional artificial real-name verification port. Access to the waiting room is possible via an escalator or staircase. There are 8 ticket gates in the waiting room, each corresponding to a designated waiting area. Each gate is equipped with 5 sets of ticket gates and an artificial ticket gate. For ease of description, a simplified topological structure of the station is illustrated in Figure 5.

The data on passenger flow at the entrance can be directly obtained from the real-name verification gates. The equipment collects raw data once per minute. However, due to the train departure interval of approximately half an hour at this station and the variation in the number of people due to train arrivals, we aggregate the data every thirty minutes for ease of analysis. A total of 40,204 pieces of data on passenger flow were collected from the east and west entrances of the station from March 30, 2021, to May 23, 2022. After handling outliers and missing values, 40,128 pieces of data remained, with 20,064 data points for each entrance. Taking the east entrance as an example, the distribution of the data is shown in Figure 6.

Passengers choose the waiting area corresponding to their own train's ticket gate for waiting. Therefore, we use the
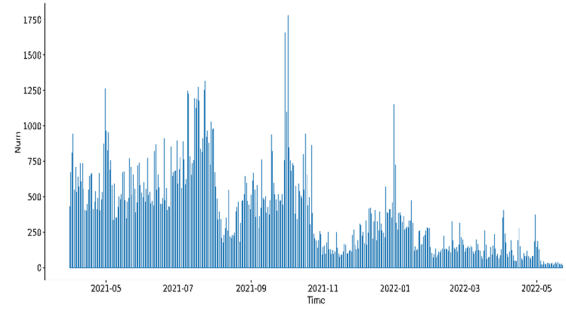

FIGURE 6. Distribution of pedestrian flow data at the east entrance.
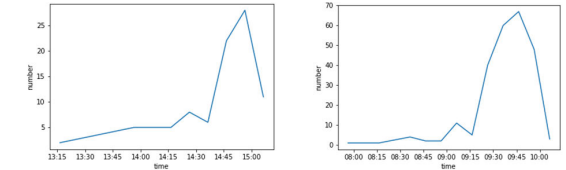

FIGURE 7. Distribution of passenger real-name verification time.

data from the ticket gates as the passenger flow data for the respective waiting areas. However, the ticket gates only collect data when trains are being checked. We aggregate the data in 30-minute intervals, summing up the number of passengers for all ticketed trains within that time frame. At the same time, there are manual ticket gates in the ticket gate area, which may open for ticket checking in advance based on passenger volume and have no data records. To address this issue, we make adjustments using the total number of tickets sold for the corresponding train.

The specific correction method is as follows: we calculate the total number of manual real-name verifications by subtracting the total number of real-name verifications from the total number of ticketed passengers for a specific train. For example, if the number of passengers boarding a certain train at this station is 432 and the number of real-name verifications is 365, the remaining 67 passengers are manually checked. We use the distribution characteristics of passengers at the real-name verification gates to fit their distribution at the entrance and waiting area. After fitting, it is found that passengers typically arrive at the station approximately 2 hours before the departure time, as shown in Figure 7.

Given that for each moment, the number of passengers entering the waiting area through manual real-name verification gate at the entrance represented as $\mathcal{P}_e = \{p_n\}_{n=1}^N$, where N denotes the data volume, and the proportion of passengers entering the waiting area is r_n, the number of passengers entering waiting area at each moment $\tilde{\mathcal{P}}_w = \{p_n * r_n\}_{n=1}^N$. The probability of passengers arriving at the waiting area in different time intervals before the train departure can be calculated using Eq. (27), where t_p represents the time in hours that passengers arrive in advance to the waiting area.

$$\mathbb{P}_1 = \{prob_1^i\}_{i=1}^4 = \begin{cases} prob_1^1, & 0 < t_p \leq 0.5 \\ prob_1^2, & 0.5 < t_p \leq 1 \\ prob_1^3, & 1 < t_p \leq 1.5 \\ prob_1^4, & 1.5 < t_p \leq 2 \end{cases} \qquad (27)$$
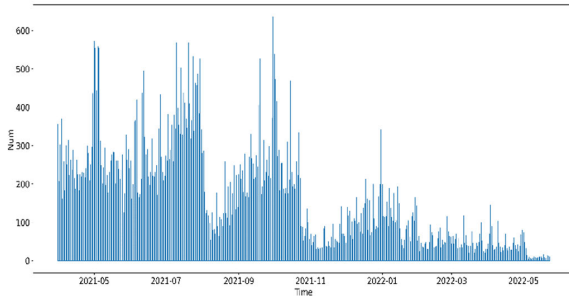
FIGURE 8. Data distribution of people in waiting area 1.

After arriving at the waiting area, the probability of the remaining passengers in the waiting area at each time interval can be represented by Eq. (28), where t_d represents the delayed time in hours for passengers entering the waiting area. The number of passengers in the waiting area at each moment can be calculated using Eq. (29).

$$\mathbb{P}_2 = \{prob_2^j\}_{j=0}^3 = \begin{cases} 1, & t_d = 0 \\ 1 - prob_1^1, & t_d = 0.5 \\ 1 - prob_1^1 - prob_1^2, & t_d = 1 \\ 1 - prob_1^1 - prob_1^2 - prob_1^3, & t_d = 1.5 \end{cases} \tag{28}$$

$$\mathcal{P}_w = \{p_n r_n + p_{n-1} r_{n-1} prob_2^1 + p_{n-2} r_{n-2} prob_2^2 \\ + p_{n-3} r_{n-3} prob_2^3\}_{n=1}^N \tag{29}$$

By aggregating the aforementioned data, we obtain the dataset detailing the distribution of individuals in each waiting area. This allows us to obtain the final dataset. Taking Waiting Area 1 as an example, the dataset illustrates the distribution of individuals, as depicted in Figure 8.

2) METRICS
In order to facilitate better comparison with other relevant works, we adopt the Mean Absolute Error (MAE) as Eq.(30), Root Mean Square Error (RMSE) as Eq.(31), and Mean Absolute Percentage Error (MAPE) as Eq.(32) as experimental evaluation metrics. Smaller values indicate better model performance, and combining the results of these three metrics comprehensively reflects the model's performance. The specific formulas are shown in Eq. (30), Eq. (31) and Eq. (32), where N N represents the sample size, y_i indicates the true value of the passenger flow, and $\hat{y}_i$ denotes the predicted value.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \tag{30}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \tag{31}$$

$$MAPE = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \tag{32}$$

TABLE 1. Ablation experiment results (Bold indicates the best result).

No.	Model	MAE ↓	RMSE ↓	MAPE(%) ↓
①	B(Baseline)	7.95	31.38	9.03
②	B+RC	7.09	29.31	8.59
③	B+LSTM	7.83	31.31	8.68
④	B+SE	7.29	29.83	8.16
⑤	B+RC+LSTM	6.85	28.40	8.12
⑥	B+RC+LSTM+SE	**6.83**	**28.20**	**7.69**

B. EXPERIMENT SETTINGS
The proposed model was built using the PyTorch framework on an Ubuntu 20.04.6 LTS system. Two NVIDIA Quadro RTX 8000 GPUs were utilized for accelerated computations. The dataset was divided into training, validation, and testing sets in a ratio of 6:2:2. A batch size of 32 was used, with a learning rate of 0.001. The training process consisted of 80 epochs. The MAE between the actual values and predicted values was employed as the loss function during training, and the Adam optimization algorithm was used.

C. EXPERIMENT RESULTS AND ANALYSIS
1) ABLATION STUDY
To comprehensively demonstrate the efficacy of the proposed model components, pertinent ablation experiments were conducted. The experimental findings are presented in Table 1. The experimental modules primarily encompass the incorporated residual connection (RC), the LSTM module for capturing temporal dependencies, and the Channel Attention (SE) for capturing channel-specific dependencies.

(1) Comparing modules ② , ③ and ④ to ① , each module demonstrates varying degrees of improvement in the model. The incorporation of the Residual module reduces MAE from 7.95 to 7.09, resulting in a decrease of 0.86. Similarly, RMSE decreases from 31.38 to 29.31, indicating a decrease of 2.07, and MAPE decreases from 9.03% to 8.59%, demonstrating a decrease of 0.44%. This indicates that the added residual connections effectively enhance the model. Likewise, the inclusion of the LSTM module leads to reductions in MAE, RMSE, and MAPE by 0.12%, 0.07%, and 0.35%, respectively. Moreover, the addition of the SE channel attention mechanism also contributes to decreases in MAE, RMSE, and MAPE by 0.66%, 1.55%, and 0.87%, respectively. These findings demonstrate that LSTM and SE also contribute to the enhancement of the model's predictive performance. Regarding the MAE and RMSE metrics, the Residual module exhibits the most significant impact, while for the MAPE metric, the SE module yields the best performance, reducing it by 0.87% compared to the baseline.

(2) Comparing modules ②, ③ and ⑤, the addition of LSTM to module B results in a decrease in MAE from 7.09 to 6.85, RMSE from 29.31 to 28.40, and MAPE from 8.59% to 8.12%. Furthermore, adding the Residual module to module ③ leads to a decrease in MAE from 7.83 to 6.85, RMSE from 31.31 to 28.40, and MAPE from 8.68% to 8.12%. This once again confirms the effectiveness of both the Residual and

TABLE 2. Comparison with other models (Bold indicates the best result).

Model	MAE ↓	RMSE ↓	MAPE(%) ↓
DCRNN	8.10	33.59	9.42
STGCN	7.61	31.16	8.35
ASTGCN	7.95	31.38	9.03
STSGCN	7.57	30.66	8.24
STFGNN	6.98	28.78	7.91
Ours	**6.83**	**28.20**	**7.69**
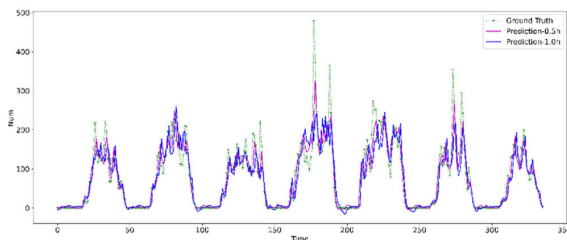
FIGURE 9. Predication value compared them with the real-time data.

FIGURE 10. UI of the prediction of flow in railway station system.

FIGURE 11. The current moment's passenger flow at the station.

LSTM modules and demonstrates that these two modules do not conflict with each other in terms of improving predictive performance.

(3) Comparing modules ① , ⑤ and ⑥, the addition of the SE module to module ⑤ shows minimal improvements in MAE and RMSE, with a decrease of only 0.02 from 6.85 to 6.83 and a decrease of 0.20 from 28.40 to 28.20, respectively. However, a more significant enhancement is observed in MAPE, which decreases from 8.12% to 7.69%, a reduction of 0.43%. The inclusion of the SE channel attention mechanism may contribute to a better understanding of crucial features and trends in the pedestrian flow sequence, thereby improving the MAPE and other percentage error metrics. Ultimately, compared to the baseline, the final model ⑥ achieves noticeable improvements, with reductions of 1.12 in MAE, 3.18 in RMSE, and 1.34% in MAPE, reaching values of 6.83, 28.20, and 7.69%, respectively.

2) OVERALL COMPARISON

In order to further assess the predictive performance of the proposed model, the established dataset was trained and evaluated using the following models under the same environmental configuration, including DCRNN [18], STGCN [7], ASTGCN [8], STSGCN [26], and STFGNN [27]. The specific comparison results are presented in Table 2.

It can be observed that the proposed model outperformed the alternative models, achieving the best experimental outcomes with MAE, RMSE, and MAPE values of 6.83, 28.20, and 7.69%, respectively. These metrics were reduced by 0.15%, 0.58%, and 0.22%, respectively, compared to the best results obtained by STFGNN (MAE: 6.98, RMSE: 28.78, MAPE: 7.91%). Our model has achieved good performance and has a particularly strong advantage in describing the
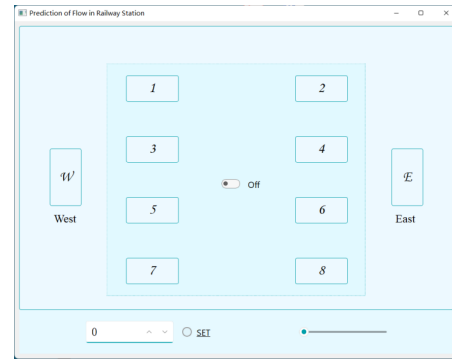
spatiotemporal characteristics of pedestrian flow data. The inclusion of residual connections helps alleviate the issue of gradient vanishing and enhances information propagation efficiency. Moreover, the LSTM module introduces a gate mechanism to effectively capture long-term information in sequences, while the incorporation of SE allows the model to prioritize important feature channels, thereby improving both the performance and generalization capability of the model.

D. VISUALIZATION OF THE REAL-TIME RESULTS

In final stage of the experiment, we visualized the prediction effect of the proposed RRC-STGCN model using Waiting Area 1 as an example. We visualized a total of 336 data points for a week in this waiting area. We observed the predicted flow of passenger 0.5 hours ahead and 1.0 hour ahead, respectively, and compared them with the real values. The comparison results are shown in Figure 9. From the figure, we can see that the proposed method can fit the real values better, which once again verifies the effectiveness of RRC-STGCN. Additionally, it can be noticed that the predictions made 0.5 hours in advance fit better at extreme values and accurately predict the trend of passenger flow in the railway station. This highlights the necessity of the Recent Cycle Component setting in RRC-STGCN.

In order to better apply our theoretical methods and guide engineering applications, we have developed a visualization software. Specifically, we used PyQt to design the user interface (UI) for the passenger flow prediction in railway stations, taking Figure 5 as a template. The resulting UI design is shown in Figure 10.

FIGURE 12. The predicted value of passenger flow in the next half hour at the station.

The switch button in the middle controls the start and stop. The spin box and the "SET" button are responsible for selecting the target time node. The sliding bar in the lower right corner controls the selection of the prediction time period, including the next half hour and one hour.

By randomly selecting a target time node, Figures 11 and 12 display the current moment and the predicted passenger flow at the entrances and waiting areas in the next half hour, respectively. The Numbers indicate the magnitude of the pedestrian flow at each location, and the colors are used to visually represent the flow intensity (blue: small flow, green: moderate flow, red: large flow, brown: very large flow). Through this UI design, railway passenger station managers can have a more intuitive understanding of the current and future pedestrian flow situation at the entrances and waiting areas. This enables them to better carry out passenger flow prediction and scheduling, optimize the operation and management of the station.

V. CONCLUSION AND FUTURE WORK

This paper addresses the problem of predicting the total number of passengers and their distribution in railway stations using a spatiotemporal graph neural network. By transforming the flow lines within the station into a graph, the problem is converted into a passenger flow prediction task. We propose a railway station passenger flow prediction model called RRC-STGCN. The model divides input data into three distinct periods and utilizes spatio-temporal channel attention mechanisms to capture spatio-temporal correlations. Furthermore, a Spatio-Temporal Convolution Module is employed to extract spatio-temporal dependencies from spatial and temporal feature maps, thereby enhancing the predictive capability of the model. Extensive experiments validate the effectiveness and accuracy of the proposed model in predicting passenger flow in railway stations. Comparative experiments with other models demonstrate that the RRC-STGCN model achieves the best results in terms of MAE, RMSE, and MAPE metrics, surpassing other models by at least 0.15%, 0.58%, and 0.22%, respectively. Additionally, the predicted results are visualized to intuitively demonstrate the model's predictions of passenger flow. The research results of this paper indicate that the RRC-STGCN model has significant

potential for predicting passenger flow in railway stations. The model not only explains the passenger arrival patterns, but also completes the distribution prediction of passengers in the station. It supports the improvement of station operational efficiency and safety. Future work will focus on optimizing the model and extending its application to the entire railway network for passenger flow prediction in control areas.

REFERENCES

[1] Z. Wei, X. Shan, H. Wang, X. Lv, and J. Zhang, "Railway passenger flow forecast system based on passenger transport big data platform," *Railway Comput. Appl.*, vol. 31, no. 1, pp. 37–42, Jan. 2022.

[2] Y. Xiao, B. L and H. Yang, "Railway passenger transport demand analysis and short-term passenger flow forecast," *Sci. Technol. Eng.*, vol. 22, no. 9, pp. 3727–3734, Mar. 2022.

[3] R. Li, H. Wu, and W. Yang, "Design of passenger flow monitoring and early warning system for railway stations based on ticketing data," *Railway Transp. Economy*, vol. 38, no. 12, pp. 54–58 and 73, Dec. 2016.

[4] Y. Ye and W. Li, "Calculating method of maximum aggregate number at high-speed rail terminal," *J. East China Jiaotong Univ.*, vol. 35, no. 4, pp. 76–82, Aug. 2018.

[5] S. Feng, "Temporal and spatial distribution short-term prediction of passenger flow in high-speed railway stations based on graph embedding LSTM," M.S. thesis, School Traffic Transp., Beijing Jiaotong Univ., Beijing, China, 2021.

[6] R. Li, "Research on key technologies of intelligent recognition of passenger flow in high speed railway station based on deep learning," Ph.D. dissertation, China Acad. Railway Sci., Beijing, China, 2022.

[7] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," in *Proc. 27th Int. Joint Conf. Artif. Intell.*, Jul. 2018, pp. 3634–3640, doi: 10.24963/ijcai.2018/505.

[8] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," in *Proc. AAAI*, Jul. 2019, vol. 33, no. 1, pp. 922–929, doi: 10.1609/aaai.v33i01.3301922.

[9] W. Xu, H. Huang and Y. Qin, "Study of railway passenger flow forecasting method based on spatio-temporal data mining," *J. Beijing Jiaotong Univ.*, vol. 28, no. 05, pp. 16–19, Oct. 2004.

[10] L. Xing and J. Dong, "Prediction of the demand of railway station transport capacity based on seasonal," *Shanxi Sci. Technol.*, vol. 34, no. 4, pp. 113–116, Apr. 2019.

[11] Z. Wu, "Analysis and prediction of passenger flow data of Shanghai Hongqiao hub based on ARIMA model," *Traffic Transp.*, vol. 35, no. S1, pp. 315–320, Mar. 2022.

[12] J. Liu, "Short-term forecast model of railway passenger flow based on ensemble algorithm," *J. Chongqing Jiaotong Univ., Natural Sci.*, vol. 41, no. 5, pp. 20–25, May 2022.

[13] Y. Shi, L. Chen, Y. Liang, K. Chen and X. Li, "Research on prediction of passenger flow of intercity railway based on wavelet decomposition and ARIMA model," *Jiangsu Sci. Technol. Inf.*, vol. 36, no. 29, pp. 30–34, Oct. 2019.

[14] Z. Wang, Y. Wang, L. Jia, and P. Li, "The application of improved bp neural network in the prediction of railway passenger volume time serial," *China Railway Sci.*, vol. 26, no. 2, pp. 127–131, Mar. 2005.

[15] W. Cao, "Study on optimization of railway short-term passenger flow forecastillg method based on machine learning," M.S. thesis, Beijing Jiaotong Univ., Beijing, China, 2018.

[16] J. Li, Q. Peng, and C. Wen, "Short term passenger flow prediction of high speed railway based on LSTM deep neural network," *Syst. Eng. Theory Pract.*, vol. 41, no. 10, pp. 2669–2682, Oct. 2021.

[17] C. Liu and S. Ni, "Short-time passenger flow prediction of high-speed railway based on BP-GRU combined model," *China Transp. Rev.*, vol. 45, no. 03, pp. 104–109, Mar. 2023.

[18] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," Feb. 2018, *arXiv:1707.01926*. Accessed: Aug. 27, 2023.

[19] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph WaveNet for deep spatial-temporal graph modeling," May 2019, *arXiv:1906.00121*. Accessed: Aug. 27, 2023.

[20] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph WaveNet for deep spatial-temporal graph modeling," May 2019, *arXiv:1906.00121*. Accessed: Aug. 27, 2023.

[21] S. He, Q. Luo, R. Du, L. Zhao, G. He, H. Fu, and H. Li, "STGC-GNNs: A GNN-based traffic prediction framework with a spatial–temporal Granger causality graph," *Phys. A, Stat. Mech. Appl.*, vol. 623, Aug. 2023, Art. no. 128913, doi: 10.1016/j.physa.2023.128913.

[22] L. Zhao, Y. Song, C. Zhang, C. Liu, P. Wang, T. Lin, M. Deng, and H. Li, "T-GCN: A temporal graph convolutional network for traffic prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 9, pp. 3848–3858, Sep. 2020, doi: 10.1109/TITS.2019.2935152.

[23] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," Dec. 2014, *arXiv:1412.3555*. Accessed: Aug. 27, 2023. http://arxiv.org/abs/1412.3555

[24] D. Chen, H. Chen, and A. Ren, "Short-time traffic flow prediction of graph convolutional network considering influence of space and time," *Comput. Eng. Appl.*, vol. 57, no. 13, pp. 269–275, Jul. 2021.

[25] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," Sep. 2018, *arXiv:1706.02216*. Accessed: Aug. 27, 2023.

[26] C. Song, Y. Lin, S. Guo, and H. Wan, "Spatial–temporal synchronous graph convolutional networks: A new framework for spatial–temporal network data forecasting," in *Proc. AAAI*, Apr. 2020, vol. 34, no. 1, pp. 914–921, doi: 10.1609/aaai.v34i01.5438.

[27] M. Li and Z. Zhu, "Spatial–temporal fusion graph neural networks for traffic flow forecasting," *arXiv:2012.09641*, Mar. 6, 2021. Accessed: Aug. 27, 2023.

[28] M. Cardia, M. Luca, and L. Pappalardo, "Enhancing crowd flow prediction in various spatial and temporal granularities," Mar. 2022, *arXiv:2203.07372*. Accessed: Aug. 27, 2023.

[29] L. Chen, D. Chai, and L. Wang, "UCTB: An urban computing tool box for spatiotemporal crowd flow prediction," Jun. 2023, *arXiv:2306.04144*. Accessed: Aug. 27, 2023.

[30] S. Zhang, H. Zheng, H. Su, B. Yan, J. Liu, and S. Yang, "GACAN: Graph attention-convolution-attention networks for traffic forecasting based on multi-granularity time series," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2021, pp. 1–8, doi: 10.1109/IJCNN52387.2021.9534064.

[31] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Comput. Vision—ECCV 2018* (Lecture Notes in Computer Science), vol. 11211, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham, Switzerland: Springer, 2018, pp. 3–19.

[32] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," May 2019, *arXiv:1709.01507*. Accessed: Aug. 27, 2023.

[33] J. Li, Y. Liu, and L. Zou, "A dynamic graph convolutional network based on spatial-temporal modeling," *Acta Scientiarum Naturalium Universitatis Pekinensis*, vol. 57, no. 4, pp. 605–613, Jul. 2021.

[34] R. Shi and L. Du, "Multi-section traffic flow prediction based on MLR-LSTM neural network," *Sensors*, vol. 22, no. 19, p. 7517, Oct. 2022, doi: 10.3390/s22197517.

[35] R. Bai, R. Shi, M. Yue, and X. Du, "Hybrid model based on K-means++ algorithm, optimal similar day approach, and long short-term memory neural network for short-term photovoltaic power prediction," *Global Energy Interconnection*, vol. 6, no. 2, pp. 184–196, Apr. 2023, doi: 10.1016/j.gloei.2023.04.006.

[36] J. Mulerikkal, S. Thandassery, V. Rejathalal, and D. M. D. Kunnamkody, "Performance improvement for metro passenger flow forecast using spatio-temporal deep neural network," *Neural Comput. Appl.*, vol. 34, no. 2, pp. 983–994, Jan. 2022.

WEI BAI was born in 1989. She is currently an Assistant Researcher with the Institute of Computing Technology, China Academy of Railway Sciences Company Ltd. Her research interests include railway passenger station informatization, and passenger management and prediction.

ZHIKANG MENG was born in 1998. He is currently pursuing the master's degree with the School of Automation and Software Engineering, Shanxi University. His research interests include graph neural networks and deep learning.

BINBIN XIN was born in 1998. He is currently pursuing the master's degree with Shanxi University. His research interests include deep learning and pedestrian re-identification.

RUIFENG GAO was born in 2000. He is currently pursuing the master's degree with Shanxi University. His research interests include deep learning and pedestrian recognition.

XIAOSHU WANG was born in 1988. She is currently pursuing the Ph.D. degree with the China Academy of Railway Sciences. Her research interests include intelligent railway passenger stations and prediction of passenger flow.

XIAOJUN LV was born in 1968. He is currently an Assistant Researcher with the Institute of Computing Technology, China Academy of Railway Sciences Company Ltd. His research interests include railway passenger station informatization and railway station safety.

● ● ●