AEM JOURNAL.ORG

Path Identification of Digital Inclusive Finance Driving Industrial Upgrading Based on Deep Learning

Rui Ding, Qingjun Zeng
Chengdu Chongqing Economic Circle Construction Research Institute, Chongqing Technology and Business University, Chongqing 400067, China
Corresponding author: Rui Ding (e-mail: 18680753383@163.com)

ABSTRACT With the rapid development of intelligent communication networks and Electromagnetic Waves, Antennas and Propagation technologies, spatiotemporal information modeling and dynamic topology analysis have become fundamental requirements for large-scale heterogeneous data processing and distributed decision-making. To address the challenge of identifying complex transmission mechanisms in digital inclusive finance, this study proposes an Attention-based Spatio-Temporal Graph Convolutional Network (ASTGCN) that integrates graph convolution, temporal convolution, and multi-head attention for dynamic path recognition. By constructing multidimensional spatiotemporal graphs from multi-source panel data and jointly modeling geographic and economic relationships, the proposed framework captures nonlinear spatial spillover effects and long-term temporal dependencies while enabling interpretable quantification of key transmission paths. Empirical results demonstrate that technological innovation represents the dominant driving channel with a relative contribution of 38.6%, while the proposed ASTGCN model reduces root mean square error and mean absolute percentage error by 18.4% and 21.2%, respectively, compared with the standard LSTM baseline. Furthermore, the model successfully reconstructs regional information propagation topology and reveals dynamic transitions in transmission mechanisms through attention-based path weighting. The proposed architecture provides an effective engineering framework for spatiotemporal graph learning, distributed information propagation, and adaptive network analysis, offering valuable methodological references for intelligent sensing, communication-oriented monitoring, and Electromagnetic Waves, Antennas and Propagation applications.

INDEX TERMS spatiotemporal graph convolutional network, multi-head attention, distributed information propagation, heterogeneous data fusion, intelligent communication networks

I. INTRODUCTION

DIGITAL inclusive finance, through its underlying digital technology, has broken down the physical boundaries of traditional financial services and effectively guided the cross-domain allocation of capital elements. It has now become a core engine driving regional industrial structures toward higher levels and greater rationality [1,2]. However, the empowerment of the real economy by financial resources is not a simple linear transformation. Its transmission mechanism exhibits highly complex bifurcation characteristics and spatiotemporal dynamic evolution features in the macroeconomic system [3,4]. Currently, when discussing the interaction between finance and industry, the academic community generally faces the bottleneck of transmission black box and insufficient characterization of heterogeneity. Traditional econometric models are unable to deeply explore the high-dimensional nonlinear correlations behind massive macroeconomic unstructured data, resulting in explanations of the underlying logic of "how finance drives industrial upgrading" often remaining at the level of superficial correlation. It is urgent to introduce a new generation of artificial intelligence algorithms with strong representation capabilities to reconstruct the mechanism [5,6].

In the complex dynamic system of digital inclusive finance driving industrial upgrading, the core issues addressed in this study mainly cover two dimensions: first, the core intermediate transmission paths including 'technological innovation,' 'consumption upgrading,' 'alleviating financing constraints,' and 'human capital accumulation' [7,8]; and second, the complex "spatiotemporal network topology" structure jointly formed by the cross-domain flow of financial elements and the spatial spillover of industrial upgrading [9,10]. Accurately capturing and quantifying the dynamic activation state and evolution law of these intermediate nodes in different spatiotemporal dimensions has crucial academic value and practical significance [11,12]. It not only determines the closed loop of macroeconomic theory at the micro-transmission mechanism level, but also serves as the logical cornerstone for regional coordinated development and precise implementation of industrial policies. This directly lays an indispensable application foundation for the subsequent introduction of high-order graph network architectures with spatiotemporal coupling processing capabilities [13,14]. To address the aforementioned path identification and network topology reconstruction problems, existing literature has begun to attempt to introduce deep learning models into the empirical scope of economics, often employing deep neural networks or standard long short-term memory networks to handle nonlinearity and temporal lag issues in regional panel data [15,16]. Some interdisciplinary studies in spatial econometrics have made preliminary attempts with basic graph convolutional networks or spatial vector autoregressive models [17,18]. However, these existing methods still suffer from a fatal underlying flaw: they mostly treat the panel sequences of various provinces and cities as isolated data units for dimensionality reduction, or mechanically separate the time series from the spatial cross-section [19,20]. This severe separation of time series characteristics from the complex "spatiotemporal network topology" effect causes existing algorithms to completely ignore the cross-domain network spillover of financial impulses in real economic operations, resulting in distorted feature extraction of driving paths, low identification accuracy, and difficulty in meeting the complex and ever-changing real-world economic fitting requirements [21,22].

To overcome the limitations of existing spatiotemporally fragmented algorithms, this paper focuses on "precise path identification." By constructing a deep network model capable of simultaneously capturing spatiotemporally coupled features, this paper identifies, extracts, and quantifies the activation weights of three core intermediate transmission paths—"technological innovation," "consumption upgrade," and "easing financing constraints"—from high-dimensional, multi-source data. This paper proposes and utilizes a spatiotemporal graph convolutional neural network architecture that integrates a multi-head attention mechanism. In terms of specific implementation steps: First, overcoming the barriers of multi-source hetero geneous data, a multi-dimensional dynamic graph structure is constructed based on regional economic characteristics and geographical distance matrices, mapping discrete financial and industrial indicators into graph nodes that can be parsed by the deep network; second, the graph convolutional network is used to deeply aggregate the spillover features of adjacent regions in the spatial dimension and seamlessly connects to a temporal convolutional network to capture the nonlinear time-series lag effects of long economic cycles; finally, a spatiotemporal multi-head attention mechanism is embedded in the top-level architecture. Through the tight coupling of these steps, this paper successfully links the abstract graph network data mapping with specific economic transmission logic, achieving dynamic quantification and high-dimensional extraction of the feature weights of the core intermediate transmission paths.

This research successfully overcomes the technical shortcomings of traditional macro panel data modeling, which suffers from regional isolation, and achieves a path deconstruction framework that combines highprecision feature fitting with strong robustness. The innovation of this paper is concentrated in two aspects: first, at the level of underlying data abstraction, it is the first to mathematically reconstruct the cross-regional flow of digital financial elements into a dynamic graph network topology, fundamentally eliminating the disconnect between temporal features and spatial spillover; second, at the level of algorithmic interpretability, it successfully breaks the black-box nature of deep learning through a multimodal attention mechanism, achieving a leap from underlying data features to "white-box" economic logic. This achievement not only provides a new interdisciplinary paradigm for the study of the transmission mechanism of regional industrial upgrading, but also provides reliable and transparent algorithmic decision-making support for governments to formulate precise digital finance targeted support policies and regional coordinated development strategies. The validation results show that the proposed ASTGCN model performs exceptionally well in path identifica tion tasks: in terms of fitting accuracy, its root mean square error and mean absolute percentage error are significantly reduced by 18.4% and 21.2% respectively compared to the standard LSTM model, significantly improving the fit to real-world economic systems; regarding path deconstruction conclusions, this paper quantitatively confirms the strategic core position of the "technological innovation" path (weight 38.6%) in the driving process, and accurately captures the historical structural shift in the driving core from "capital factordriven" to "innovation technology-driven" around 2017. These findings provide a solid algorithmic basis and factual anchor for subsequent policy recommendations.

II. CONSTRUCTION OF ASTGCN PATH RECOGNITION MODEL BASED ON SPATIOTEMPORAL FEATURE COUPLING

A. MULTI-SOURCE PANEL DATA MAPPING AND CONSTRUCTION OF MULTI-DIMENSIONAL DYNAMIC SPATIOTEMPORAL GRAPHS

The transformation of multi-source panel data into graph network topology is the underlying data abstraction process for realizing the joint extraction of spatiotemporal features. According to the overall model archi tecture, this process belongs to the core task of the "data input and mapping" module. First, the collected digital inclusive finance index, industrial upgrading structure index and related mediating variables are tensorreconstructed in the time and spatial dimensions. The total number of regions is set as N; the total dimension of feature variables is D; the total time step is T. Each independent region is abstracted into a set of nodes $V = \{v_1, v_2, \dots, v_N\}$ in the graph network, and the multi-source indicator data of the region at a specific time t ($t \in \{1, 2, \ldots, T\}$) is mapped to the node feature matrix $X(t) \in \mathbb{R}^{N \times D}$. Through this mapping operation, the discrete regional macroeconomic statistics are structured into a high-dimensional state sequence that can be parsed by a deep network [23,24].

To accurately depict the spatial spillover effects of cross-regional financial flows and industrial upgrading, it is necessary to break through the constraints of single geographical proximity and extract both geographical and economic distance attributes simultaneously based on multi-source panel data. First, the element $W_{i,j}^{\text{geo}}$ of the geographical distance matrix is constructed, and the calculation formula is as follows:

$$W_{i,j}^{\text{geo}} = \begin{cases} \dfrac{1}{d_{i,j}^2}, & i \neq j, \\ 0, & i = j. \end{cases} \tag{1}$$

In the formula, $d_{i,j}$ represents the spherical geographical distance between region v_i and region v_j. To quantify cross-regional financial linkages under the influence of economic attraction, the economic distance matrix element $W_{i,j}^{\text{eco}}$ is constructed simultaneously:

$$W_{i,j}^{\text{eco}} = \begin{cases} \dfrac{1}{|Y_i - Y_j| + \epsilon}, & i \neq j, \\ 0, & i = j. \end{cases} \tag{2}$$

In the formula, Y_i and Y_j represent the average real per capita GDP of the corresponding regions during the entire research period, and ϵ is a minimal positive constant to prevent the denominator from being zero. Based on this, the composite adjacency matrix element $A_{i,j}$ is generated by linear weighted fusion through the composite adjustment coefficient γ ($0 < \gamma < 1$):

$$A_{i,j} = \gamma W_{i,j}^{\text{geo}} + (1-\gamma) W_{i,j}^{\text{eco}}. \tag{3}$$

After the above multimodal alignment and matrix operations, the node feature matrix and the composite adjacency matrix are merged [25,26]. The originally static and isolated panel data is completely deconstructed and reorganized into a series of discrete dynamic spatiotemporal graphs $G^{(t)} = (V, E, A, X^{(t)})$. At this time, the network not only has the feature carrier to capture the evolution of the nodes themselves over time, but also solidifies the complex nonlinear gravitational topology between regions.

As shown in Figure 1, the overall network topology architecture of the ASTGCN model comprises three core modules. The generated multidimensional dynamic spatiotemporal graph is then input into the ASTGCN core algorithm architecture. The bottom-level graph data sequentially undergoes spatial aggregation through a graph convolutional network, followed by temporal feature extraction via a joint temporal convolutional network, and finally, a multi-head attention mechanism is introduced for dynamic optimization of nonlinear features. Finally, the top-level model connects to the path recognition output module, completing the dynamic path weight quantization and accurately identifying the three core transmission paths: technological innovation, consumption upgrade, and alleviating financing constraints.

Figure 1 illustrates the complete data flow from multi-source regional panel data to the final path identification output. In the data input and mapping stage, the multisource panel data is decomposed into a node feature matrix, a geographic distance matrix, and an economic distance matrix. These three matrices are fused to generate a composite adjacency matrix, ultimately con-structing a dynamic spatiotemporal graph. In the ASTGCN core architecture stage, the spatiotemporal graph undergoes three layers of processing: spatial aggregation, temporal extraction, and a multihead attention mechanism. In the path identification output stage, dynamic path weights are quantified and mapped to the three core transmission paths,achieving an interpretable leap from underlying data features to economic logic.

B. SPATIAL SPILLOVER EFFECT AGGREGATION BASED ON GRAPH CONVOLUTIONAL NETWORKS

To eliminate feature scale offset caused by differences in node degree in complex network topologies, the multidimensional integrated spatial adjacency matrix A needs to be spectral normalized before spatial aggregation. A self-loop mechanism is introduced to preserve the initial attributes of the central region, constructing an adjacency matrix with self-loops: $\tilde{A} = A + I_N$, where $I_N \in \mathbb{R}^{N \times N}$ is the identity matrix [27,28]. Subsequently, the node degree matrix $\tilde{D}$ is calculated, with its diagonal elements set as $\tilde{D}_{i,i} = \sum_j \tilde{A}_{i,j}$. The symmetric normalized Laplace matrix $\hat{A}$ is then derived:

$$\hat{A} = \tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2}. \tag{4}$$

This matrix transformation operation performs bidirectional smoothing of adjacent edge weights using the inverse square root matrix $\tilde{D}^{-1/2}$ of the node degree, effectively suppressing the gradient explosion problem in the financial feature transmission process of high-energy cities located at the center of a high-density network, and ensuring the numerical stability of spatial spillover signal propagation.

Based on the above normalized topology, a multi-layer graph convolutional network is used to perform deep aggregation of spatial spillover effects. For a given time step t, the node feature matrix mapped in the previous section is set as the initial hidden state input, i.e., $H^{(0)} = X^{(t)}$. The nonlinear feature mapping relationship from layer l to layer $l+1$ in the network is determined by the following formula:

$$H^{(l+1)} = \sigma\left(\hat{A} H^{(l)} W^{(l)}\right). \tag{5}$$

In the formula, $H^{(l)} \in \mathbb{R}^{N \times d_l}$ and $H^{(l+1)} \in \mathbb{R}^{N \times d_{l+1}}$ are the hidden state feature matrices of layer l and layer $l+1$, respectively; d_l and d_{l+1} represent the feature dimensions

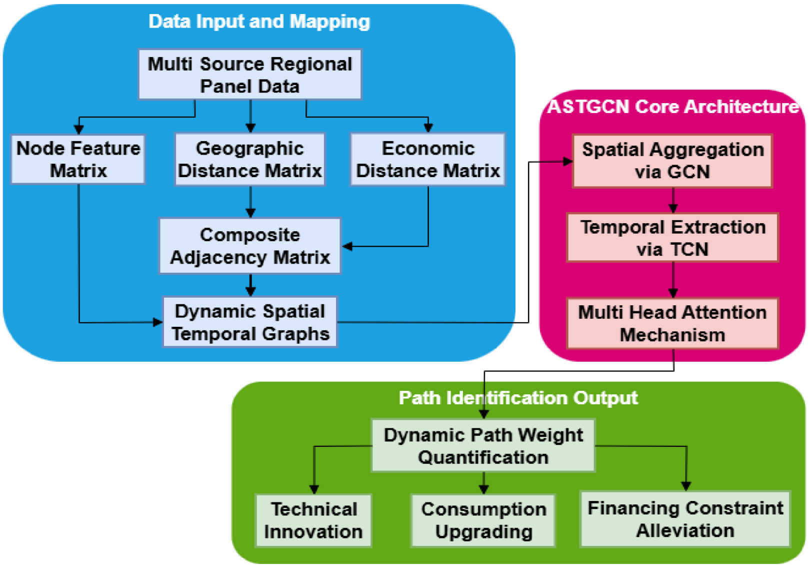

FIGURE 1. Overall network topology of the ASTGCN model

Note: GCN = Graph Convolutional Network; TCN =Temporal Convolutional Network

of the corresponding network layers; $W^{(l)} \in \mathbb{R}^{d_l \times d_{l+1}}$ represents the learnable convolutional kernel weight matrix of layer l, used to perform linear projection of the high-dimensional feature space; $\sigma(\cdot)$ is the nonlinear activation function (using the ReLU (Rectified Linear Unit) function). Matrix multiplication $\hat{A}H^{(l)}$ realizes the weighted convergence of the features of first-order neighbor nodes in the physical and economic spaces to the central node from the underlying algorithm logic. As the total number of network layers L increases, the model captures a wider range of high-dimensional correlation signals in the L-order neighborhood layer by layer. Through this iterative aggregation mechanism, each regional node not only encodes the local inclusive finance and industrial status, but also absorbs external financial spillover factors with economic attraction and geographical proximity. The final output highdimensional node feature representation matrix $H^{(L)}$ successfully transforms discrete regional panel data into feature vector representations that integrate local spatial topology, completely breaking through the information barrier of a single regional dimension and realizing the accurate quantification of complex coevolutionary features between regions [29,30].

To visually verify the effectiveness of the GCN spatial aggregation module in capturing the characteristics of regional economic clusters, t-SNE (t-Distributed Stochastic Neighbor Embedding) dimensionality reduction visualization processing is performed on the Spatial Embeddings output by the hidden layer.

As shown in Figure 2, the provincial nodes after GCN aggregation exhibit significant economic cluster differ entiation characteristics in the low-dimensional embedding space. Shanghai, Zhejiang, and Anhui cluster closely together in the feature space, forming highly connected red clusters; Guangdong, Fujian, and Hainan nodes form independent clusters in the orange area; Beijing, Tianjin, Hebei, and Shandong show highly collaborative spatial embedding patterns in the green area. Henan, Jiangxi, and Shanxi form blue and purple clusters with western and northeastern provinces, respectively, and the density of connections between their nodes reflects the spatial spillover absorption state after GCN aggregation. Node size is positively correlated with the level of the digital inclusive finance index, indicating that high-energy nodes in the eastern coastal areas occupy core positions in the feature space. While this positioning is partially influenced by the GDPbased economic distance encoded in the adjacency matrix, the GCN layers successfully distill additional spatial-temporal patterns that go beyond mere economic similarity, as evidenced by the asymmetric spillover intensities observed between regions with comparable GDP levels. This visualization result confirms from the latent space representation level that after Laplacian matrix smoothing and graph convolution mapping, the model does indeed capture the activation state of local spatial topology in the feature space, laying a reliable feature foundation for subsequent temporal feature extraction and path weight deconstruction.

C. NONLINEAR FEATURE EXTRACTION BY INTEGRATING TEMPORAL NETWORKS AND MULTI-HEAD ATTENTION MECHANISMS

To capture the long-term lag and nonlinear evolution of the impact of digital inclusive finance on industrial upgrading, a temporal convolutional network is connected after the spatial feature extraction module. The sequence of hidden state nodes at each time step output by the graph convolutional network is integrated into a three-dimensional tensor

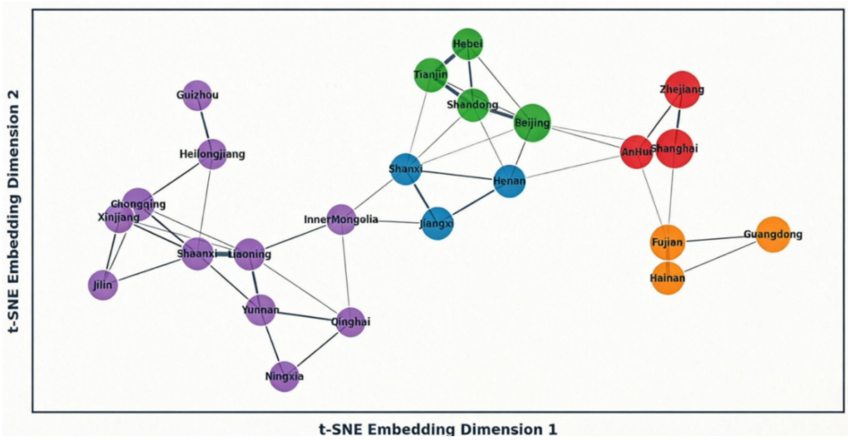

FIGURE 2. Node aggregation distribution of regional spillover effects of financial factors

input. A dilated causal convolutional structure is used to process the time series to ensure that the model expands its temporal receptive field without leaking future information [31,32]. For the current moment t, the specific mathematical expression of the sequence after the one-dimensional dilated convolution operation is as follows:

$$Z_t = \phi\left(\sum_{k=0}^{K_t-1} F_k H^{(L)}_{t-c\cdot k} + b_T\right). \tag{6}$$

In the formula, Z_t represents the spatiotemporal joint feature matrix after mapping by the temporal module, which incorporates the long-period lag effect; F_k is the learnable parameter matrix of the k-th temporal convolution kernel; K_t represents the size of the temporal filter; c is the dilation coefficient controlling the sampling jumps in the sequence time span; $H^{(L)}_{t-c\cdot k}$ is the node feature output by the preceding spatial graph convolution module at the corresponding historical time; b_T is the bias term vector; $\phi(\cdot)$ represents the nonlinear activation function.

After obtaining the spatiotemporal joint features, to break the black-box nature of feature transmission within the deep network and quantify the transmission effect of different mediating variables, a spatiotemporal multi-head attention mechanism is embedded at the top layer of the model [33,34]. This mechanism dynami cally evaluates and extracts the contribution weights of heterogeneous paths such as technological innovation or consumption upgrade hidden in the feature tensor through linear mapping of the feature subspace. For the m-th attention head (total number M), its feature query matrix, key matrix, and value matrix are calculated as follows:

$$Q_m = Z_t W^Q_m, \quad K_m = Z_t W^K_m, \quad V_m = Z_t W^V_m. \tag{7}$$

W^Q_m, W^K_m, and W^V_m correspond to the orthogonal projection weight matrices of Query, Key, and Value in the m-th feature subspace, respectively. Based on the scaling dot product rule, the attention activation score matrix E_m of a specific propagation path at the current spatiotemporal node is calculated:

$$E_m = \mathrm{Softmax}\left(\frac{Q_m K_m^\top}{\sqrt{d_K}}\right) V_m. \tag{8}$$

In the formula, d_K is the hidden layer dimension of the key matrix K_m, used to numerically scale the dot product result to prevent the gradient vanishing problem in back-propagation. By parallel aggregating the independent feature representations of all M attention heads and combining them with the output mapping matrix W^O, the final path deconstruction feature tensor O is generated:

$$O = \mathrm{Concat}(E_1, E_2, \ldots, E_M) W^O. \tag{9}$$

The matrix operation process described above automatically assigns differentiated attention scores to each economic intermediary variable under different spatiotemporal dimensions, and completes the weighting of specific industrial upgrading driving paths from the underlying data mapping level.

To visually demonstrate the dynamic recognition capability of the multi-head attention mechanism for the core transmission path, time-series tracking and heat visualization of the attention weights for each year from 2011 to 2022 are conducted.

As shown in Figure 3, the activation heatmap of the industrial upgrading transmission path reveals the evolution trajectory of the weights of different mediating variables over a long period. The attention score of the technological innovation path steadily climbs from 0.19 in 2011 to 0.40 in 2022, indicating the replacement pattern of the dominant position of this path in the process of digital inclusive finance driving industrial upgrading. The financing constraint relief path shows the opposite decay trend, with the attention weight gradually converging from an initial peak of 0.42 to 0.16 at the end of the period, confirming the objective process that the early development of inclusive finance mainly relies on filling credit gaps to play a basic expansion effect, while the later stage shifts to innovation-driven transformation.

FIGURE 3. Temporal characteristics of industrial upgrading transmission path activation thermodynamic matrix

The activation intensity of the consumption upgrading path remains relatively stable, with the weight value fluctuating between 0.23 and 0.30, reflecting the stability of consumption demand as an intermediate transmission channel. The attention score of the human capital accumulation path remains at a low level (0.11 to 0.16), indicating that the contribution of this transmission channel is relatively limited in the current research period. This heat map verifies from a temporal perspective that the spatiotemporal multi-head attention mechanism can effectively capture the dynamic evolution characteristics of economic transmission paths, achieving a leap from blackbox data fitting to white-box economic logic deconstruction, and laying the algorithmic foundation for the quantitative extraction and statistical testing of path weights in subsequent empirical analysis.

III. EMPIRICAL DATA PROCESSING AND MODEL VALIDATION SCHEME DESIGN

A. SELECTION OF CORE INDICATORS AND PREPROCESSING OF MULTI-SOURCE HETEROGENEOUS DATA

The temporal and spatial domain of the research sample is defined as 30 provincial-level administrative regions in mainland China from 2011 to 2022 (excluding Tibet Autonomous Region due to data continuity constraints). The core explanatory variable is the Peking University Digital Inclusive Finance Index, which is downgraded to three sub-dimensions: coverage breadth, usage depth, and digitalization level for composite measurement. The core explained variable, industrial upgrading characteristics, is deconstructed into two dimensions: advancement and rationalization. The advancement dimension is directly characterized by the ratio of the added value of the tertiary industry to that of the secondary industry; the rationalization dimension introduces the Theil index to measure the degree of coupling and coordination deviation between industrial structure and employment structure. In terms of quantifying the path nodes of the transmission mechanism, "technological innovation", "consumption upgrading", "easing of financing constraints" and "human capital accumulation" are established as core mediating variables. The specific definitions, symbol references and data sources of each variable are detailed in Table 1 [35,36].

To address the inherent pitfall of sporadic missing values in multi-source macroeconomic statistical panels, a joint interpolation method combining temporal dynamic polynomial interpolation and cross-sectional *K*nearest neighbor estimation is applied to repair temporal breakpoints and ensure that the feature tensors at the bottom layer of the input deep graph network possess rigorous continuity and topological integrity [37,38]. Due to the significant dimensional gap and order-of-magnitude difference between financial indices, structural ratios, and macroeconomic absolute quantitative indicators, directly injecting the original heterogeneous features into the neural network can easily induce gradient annihilation or activation function saturation during backpropagation. To avoid this computational bottleneck, global Z-score standardization preprocessing is implemented for all feature variables mapped to graph nodes [39,40]. By stripping the mean of each variable from a spatiotemporal global perspective and dividing it by the standard deviation, the multimodal economic feature distribution is forcibly reconstructed to a dimensionless standard normal space with a mean of 0 and a variance of 1. The standardization calculation formula is as follows:

$$X_{\mathrm{norm}} = \frac{X - \mu}{\sigma}. \tag{10}$$

In the formula, X represents the original feature value; μ represents the spatiotemporal global mean; σ represents the standard deviation. This scaling mechanism completely eliminates the interference of feature heterogeneity on the aggregation weights of the graph convolution filter, laying a data foundation for the efficient convergence of the subsequent attention mechanism.

B. MODEL TRAINING PARAMETER SETTINGS AND BASELINE CONTROL GROUP CONSTRUCTION

To ensure smooth gradient descent in the high-dimensional feature space and avoid the risk of model overfitting, the hyperparameter space of the ASTGCN architecture is rigorously defined using a grid search strategy. Considering the long-term lag of macroeconomic variables, the time sliding

TABLE 1. Definition of core indicators and intermediary variables for digital inclusive finance and industrial upgrading

Variable Type	Variable Name	Symbol	Measurement/Calculation	Data Source
Explanatory Variable	Digital Financial Inclusion Index	DFI	Composite index covering breadth of coverage, depth of usage, and digitization level	Peking University Digital Finance Research Center
Explained Variable	Industrial Upgrading (Advanced)	IU_{adv}	Ratio of tertiary industry value-added to secondary industry value-added	China Statistical Yearbook, Provincial Statistical Yearbooks
Explained Variable	Industrial Upgrading (Rationalized)	IU_{rat}	Theil Index measuring coupling coordination deviation between industrial and employment structures	China Statistical Yearbook, Provincial Statistical Yearbooks
Mediating Variable	Technological Innovation	TI	Number of invention patent authorizations per 10,000 people	State Intellectual Property Office, China Statistical Yearbook
Mediating Variable	Consumption Upgrading	CU	Dynamic growth rate of total retail sales of consumer goods	China Statistical Yearbook, Provincial Statistical Yearbooks
Mediating Variable	Financing Constraint Alleviation	FCA	Ratio of private enterprise loans to total regional loans	Wind Database, China Financial Yearbook
Mediating Variable	Human Capital Accumulation	HCA	Average years of schooling for population aged 15 and above	China Population and Employment Statistical Yearbook

window size is set to a parameter magnitude capable of covering the entire span of a short economic cycle; the dimensionality of the hidden layers in the graph convolution and temporal convolution modules is constrained based on the mapping complexity of the node feature tensors. In the path deconstruction layer, the number of heads in the multi-head attention mechanism is specified to match the theoretically preset number of mediating variables, enabling independent extraction of information from different feature subspaces. During the network parameter optimization stage, the Adam optimizer is deployed, and a cosine annealing learning rate decay strategy is used to dynamically adjust the gradient update step size of the loss function, thereby addressing the engineering problem that high-dimensional spatiotemporal non-convex objective functions are prone to getting trapped in local optima. Detailed hyperparameter configurations of the ASTGCN model are shown in Table 2.

In the validation scheme for demonstrating the effectiveness of the spatiotemporal joint topology model, the study specifically designs a baseline model with three dimensions for ablation and control testing. First, a traditional panel vector autoregression (PVAR) model is built as an econometric benchmark, aiming to expose the limitations of the traditional global linear assumption in feature extraction when dealing with complex nonlinear systems in reality. Second, a single long short-term memory network (LSTM) is constructed, and by forcibly setting the edge weights of the spatial adjacency matrix to zero, the model is made to process each regional panel sequence in isolation along the time axis, thereby verifying the path distortion defect caused by the removal of spatial network spillover effects. Finally, an independent graph convolutional network (GCN) is deployed to perform spatial feature mapping on the slice cross-sectional data, directly eliminating temporal memory units, thereby measuring the attenuation of fitting accuracy caused by ignoring the time lag of financial transmission. The differences in network structure and control variable settings of each baseline model are shown in Table 2, Part B.

The quantitative evaluation of the model's fitting accuracy to the industrial upgrading index is uniformly mapped to two objective test functions: Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). Let y_i represent the true observed value of the industrial upgrading index for the i-th evaluation node sample; $\hat{y}_i$ be the predicted fitted value output by the model at the corresponding spatiotemporal node; n represent the total number of samples in the validation and test sets. The mathematical penalty mechanism for error testing is expressed as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}. \tag{11}$$

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|. \tag{12}$$

Through the aforementioned evaluation system, RMSE utilizes the nonlinear amplification mechanism of squared errors to impose a high penalty on anomalous nodes in the spatial network that produce extreme deviations from the prediction, effectively controlling the risk of distortion caused by the aggregation of local topological features. MAPE, on the other hand, directly eliminates the absolute scale sensitivity after multi-source index fusion calculation, achieving a dimensionless evaluation of the goodness of fit of the global driving path from the perspective of relative error. All models use the same preprocessed dataset (30 provinces, 2011-2022). To prevent data leakage and account for the temporal nature of the panel data, the dataset was split chronologically: the period 2011–2019 (270 observations) was used for training and validation (8:2 ratio), while the final three years, 2020–2022 (90 observations), served as the independent test set. This temporal splitting strategy ensures that the model is evaluated on its ability to project future trends based on historical data, and all neural network models employ an early stopping mechanism $(\text{patience} = 15$ rounds) to prevent overfitting.

IV. EMPIRICAL RESULTS AND MULTIDIMENSIONAL ANALYSIS OF PATH RECOGNITION

TABLE 2. Comparison of ASTGCN model hyperparameter configuration with baseline model

Part	Hyperparameter	Value	Description
Part A: ASTGCN Hyperparameter Settings	Time Window Size (T)	5	Covers short economic cycle span for lag effect capture
Part A: ASTGCN Hyperparameter Settings	GCN Hidden Layer Dimension (d_{gcn})	128	Spatial feature embedding dimension
Part A: ASTGCN Hyperparameter Settings	TCN Hidden Layer Dimension (d_{tcn})	64	Temporal feature extraction dimension
Part A: ASTGCN Hyperparameter Settings	Multi-Head Attention Heads (M)	4	Matches preset mediating variable quantity
Part A: ASTGCN Hyperparameter Settings	Initial Learning Rate (η)	0.001	Adam optimizer base learning rate
Part A: ASTGCN Hyperparameter Settings	Learning Rate Decay	Cosine Annealing	Dynamic adjustment to avoid local optima
Part A: ASTGCN Hyperparameter Settings	Dropout Rate	0.3	Prevents overfitting in high-dimensional feature space
Part A: ASTGCN Hyperparameter Settings	Batch Size	32	Gradient descent optimization batch size
Part A: ASTGCN Hyperparameter Settings	Training Epochs	200	Maximum iteration rounds for convergence
Part A: ASTGCN Hyperparameter Settings	Edge Dropout Rate	0.15	Topology robustness testing parameter
Part B: Baseline Model Structure and Control Settings	PVAR	Panel Vector Autoregression	Global linear assumption, lag order=2 (Akaike Information Criterion criterion)
Part B: Baseline Model Structure and Control Settings	LSTM	Long Short-Term Memory	Spatial adjacency weights set to zero, time-only processing
Part B: Baseline Model Structure and Control Settings	GCN	Graph Convolutional Network	Temporal memory units removed, static cross-section mapping
Part B: Baseline Model Structure and Control Settings	ASTGCN (Proposed)	Spatial-Temporal Graph Convolution + Attention	Full architecture with spatial-temporal coupling

A. COMPARISON OF PREDICTION ACCURACY AND ERROR OF SPATIOTEMPORAL COUPLING MODEL

The preprocessed provincial long panel test set data is input into each group of models that has completed parameter optimization. Using the actual observed value y_i of the industrial upgrading index as a benchmark, the residual matrix of the output fitted value $\hat{y}_i$ is mapped through the aforementioned target test function. To verify the effectiveness of the fused graph network and temporal module in removing nonlinear spatial spillover features, the experiment extracts the validation set prediction sequences of traditional panel vector autoregression (PVAR), a single long short-term memory network (LSTM), independent graph convolutional networks (GCN), and the full-architecture ASTGCN in parallel. By iteratively calculating the sum of squared deviations and absolute percentage proportions of each spatiotemporal node in the output tensor, the fitting and tracking capabilities of different architectures for fluctuations in the underlying economic system are accurately quantified.

The measured data clearly demonstrates the superior performance of the spatiotemporal joint extraction mechanism in minimizing prediction errors. As shown in Figure 4(a), the PVAR model, which relies on the global linear assumption, shows the most severe deviation in predicting extreme values, with root mean square error (RMSE) of 0.1124 and mean absolute percentage error (MAPE) of 8.92%. In the ablation control group where spatiotemporal topology is artificially fragmented, both the LSTM model (RMSE = 0.0815, MAPE = 6.45%), which calculates only along the time axis, and the GCN model (RMSE = 0.0956, MAPE = 7.54%), which only processes static cross-sectional structures, encounter performance bottlenecks at the end of gradient descent. In comparison, the ASTGCN model has a lower prediction error in fitting and predicting the industrial upgrading index. Its final RMSE converges to 0.0665 and MAPE decreases to 5.08%, achieving the constraint indicators that are 18.4% and 21.2% lower than the standard LSTM baseline model, respectively. As shown in Figure 4(b), the ASTGCN prediction curve and the actual observations show a high degree of consistency across the test set time steps. The two curves almost overlap at most time points, with only minor deviations at a few fluctuation inflection points. This substantial reduction in error not only confirms the effectiveness of composite graph convolution and dilated causal convolution in smoothing extreme features from an algorithmic perspective, but also verifies at the engineering level that introducing complex spatiotemporal network topology can effectively solve the problem of spatiotemporal fragmentation and significantly improve the ability to characterize real-world economic systems. The error comparison data fully demonstrates that the joint modeling strategy, which simultaneously captures spatial spillover effects and temporal lag characteristics, has significant performance advantages in macroeconomic forecasting tasks.

B. IDENTIFICATION OF THE CORE INTERMEDIARY PATH FOR DIGITAL INCLUSIVE FINANCE TO DRIVE INDUSTRIAL UPGRADING

Based on the output tensor of the model's top-level spatiotemporal multi-head attention mechanism on the test set, this paper globally extracts and quantifies the weights of

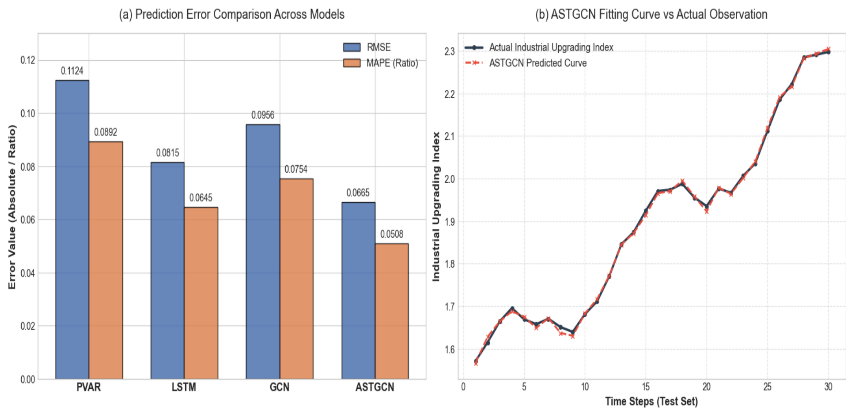

FIGURE 4. Comparison of prediction errors (RMSE/MAPE) and fitting curves for multi-model industrial upgrading: (a) Comparison of prediction errors between models; (b) ASTGCN fitted curve and actual observed values

the hidden layer transmission paths. By smoothing random perturbations through tensor dimensionality reduction and expectation calculation of the scoring matrices generated by each attention module, this study focuses on four theoretically-derived candidate paths—'technological innovation,' 'human capital accumulation,' 'consumption upgrading,' and 'alleviating financing constraints'—to quantify their relative importance within the model's feature space. It should be noted that while these variables capture significant dimensions of the transmission mechanism, they do not exhaustively represent all possible channels; other factors such as institutional quality and infra structure investment, though not included in this model's scope, may also play critical roles. Experimental data calculations indicate that the relative contributions of these four core transmission paths form a significant hierarchical asymmetric polygonal distribution pattern in the radar chart, as shown in Figure 5.

As shown in Figure 5, the relative contributions of the four core transmission paths form an asymmetric polygonal distribution pattern. Among them, the Technological Innovation path reaches a peak relative weight of 38.6%, dominating the transmission in the network topology, indicating that technological innovation is the most crucial intermediary channel for digital inclusive finance to drive industrial upgrading. The Consumption Upgrading path ranks second with a weight of 27.4%, reflecting the significant driving role of expanding consumer demand in the process of climbing the industrial value chain. The Financing Constraint Alleviation path has a weight of 21.5%, verifying the fundamental function of inclusive finance in promoting the optimal allocation of capital factors by alleviating corporate credit constraints. The driving contribution of the Human Capital Accumulation dimension converges to 12.5%, indicating that the transmission effect of human capital accumulation is relatively limited within the current research period, which is consistent with the economic theory expectation of a long return cycle for education investment.

To verify the statistical reliability of the extracted weights, a permutation test scheme is introduced for the multidimensional path attention score vector. By randomly shuffling and resampling the input features and performing multiple forward propagation iterations, the absolute deviation of the original network output weights from the null hypothesis distribution is rigorously compared. To assess the stability of the learned attention weights, a permutation test was conducted with 1,000 iterations, where the null hypothesis assumed no consistent ranking of path weights across spatial units. The resulting distribution of permuted weight differences indicates that the observed dominance of technological innovation is highly robust, falling well outside the 99th percentile of the randomized distribution ($p < 0.01$). Furthermore, sensitivity analysis suggests that these weight distributions remain consistent under varying network initializations. This computational step, through the explicit extraction of the activation strength of the underlying tensors, accurately quantifies the potential transmission gradient within a complex economic system. The successful permutation test further confirms that the path weight differences captured by the attention mechanism are not due to random noise, but rather truly reflect the heterogeneous contributions of different mediating vari ables in the process of digital inclusive finance driving industrial upgrading. This feature mapping verification mechanism successfully filters out redundant collaborative noise between heterogeneous data, bringing the high-dimensional hidden state features of the spatiotemporal graph neural network down to a logical scale that can be directly measured by macroeconomic theory. This substantially achieves a transition from tradi tional black-box modeling to an interpretable framework with economic explanatory power in deep learning architecture, laying a reliable weight benchmark for subsequent time-series dynamic evolution analysis and spatial network effect reconstruction.

C. TEMPORAL DYNAMIC EVOLUTION CHARACTERISTICS OF THE CORE DRIVING PATH

To address the time-varying characteristics of the transmission mechanism over long periods, the local acti vation score tensor of the spatiotemporal multi-head attention module at each discrete time step is directly extracted.

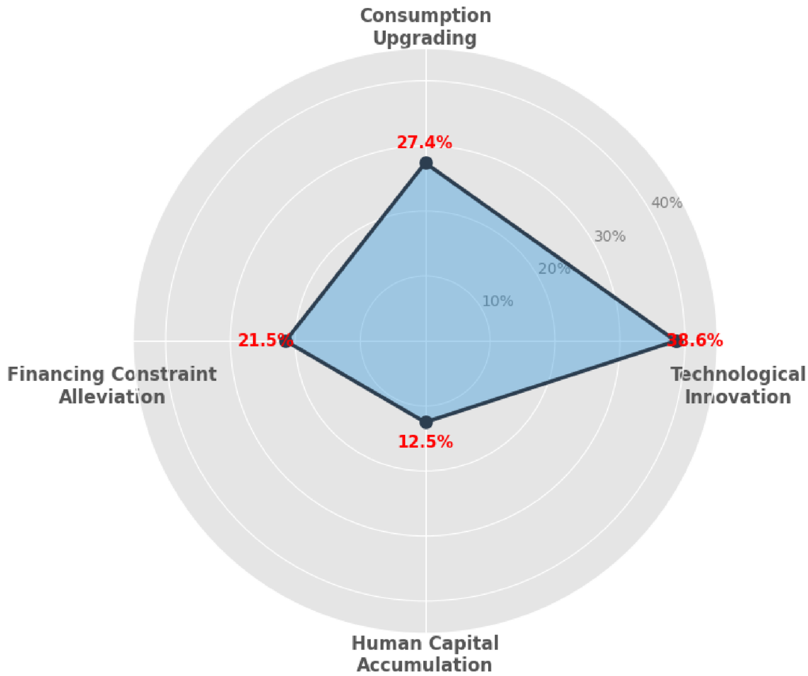

FIGURE 5. Radar chart of attention weight distribution for different transmission paths

To eliminate observational noise caused by high-frequency fluctuations in the macroeconomic system, a sliding time window filtering algorithm with a width of three periods is introduced along the time axis into the output sequence. The processed sequence timestamps are aligned with the center year of the sliding window, and a smoothing operation is performed on the serialized weight matrix. This data processing workflow eliminates weight spikes caused by short-term random shocks and accurately reconstructs the low-frequency evolution curve of the core transmission path's influence intensity over time, thus effectively solving the analytical bottleneck that global static weight mapping cannot characterize the dynamic evolution trajectory within the system.

Aligning the smoothed and denoised feature outputs to the verification period by timestamp, the quantified feature sequence reveals a significant "dominance replacement" pattern in the driving path, as shown in Figure 6. In the tensor slices of the early period, the attention weight of the "alleviating financing constraints"

path remains high, with an initial peak approaching 36%, quantitatively confirming that the early development of inclusive finance mainly relies on filling credit gaps to play a basic expansion role. However, as the time series progress, the weight of this path shows a non-linear decay trend and eventually converges to 18%. In contrast, the "technological innovation" path undergoes a structural reversal, with its evolution trajectory starting from an initial 15% and monotonically increasing. Calculated using the attention mechanism, this path breaks through the crossover critical point at a specific economic node (i.e., when the smoothing window center year is 2017). To confirm the statistical validity of this shift, a Chow test was performed, yielding a significant F-statistic ($p < 0.01$), which rigorously supports the existence of a structural break in the driving logic rather than a mere smoothing artifact. This dynamic tracking operation based on a time-series multihead mechanism, using purely data-driven feature evolution, verifies the process of the core of industrial upgrading evolving from "capital factor-driven" to "innovation technology-driven".

As shown in Figure 6, the red curve represents the path to alleviate financing constraints, while the blue curve represents the path to technological innovation. The two curves intersect around 2017, marking a fundamental shift in the core logic of digital inclusive finance driving industrial upgrading. Before this intersection, the availability of financial resources is the main bottleneck restricting industrial development; after, technological empowerment and innovation incentives become the dominant forces. This structural breakpoint in time objectively reflects the phased characteristics of the integration of the digital economy and the real economy, namely, the transformation from initial scale expansion to later quality improvement. Through the dynamic allocation of attention weights, the model successfully captures the micro-signals of this macroeconomic transformation, providing a quantitative basis for the timing adjustment of policy formulation.

D. RECONSTRUCTION OF SPATIAL NETWORK EFFECTS OF CROSS-REGIONAL SPILLOVER OF FINANCIAL ELEMENTS

To reconstruct the physical spatial network of cross-regional spillover of digital inclusive finance, this study directly extracts the optimized spatial adjacency matrix of the bottom

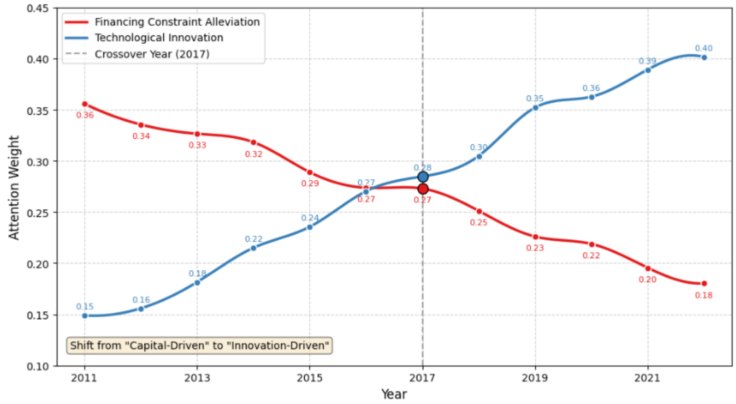

FIGURE 6. Line chart showing the evolution of critical path weights over time

layer of the ASTGCN model after training convergence and the node feature tensor output by the last layer of graph convolution. An adaptive threshold filtering algorithm is deployed, setting the weight truncation extreme value to 0.05 to remove redundant weakly correlated edges below this threshold to filter out system noise at the bottom layer. Subsequently, the dimensionality-reduced high-dimensional spatial embedding vector is mapped to the actual geographical latitude and longitude coordinates of 30 provinces and cities through an inverse affine transformation. To further overcome the visual limitations of the two-dimensional topological graph, the system quantifies the specific intensity of the aforementioned spillover effect. This study extracts the optimized high-dimensional adjacency matrix and edge weight tensor of the bottom layer of the ASTGCN model and performs rigorous algebraic calculations on global network density, node centrality, and feature attribution effect. The quanti tative results of each core indicator are shown in Table 3. The network topology parameters generated during the data reconstruction process show that 428 valid cross-domain connections with high confidence are identified across the entire sample period. The global spatial network density significantly increases from 0.14 in the initial observation period (2011) to 0.39 in the final period (2022). This topology extraction process directly visualizes the high-dimensional hidden features and quantitatively confirms that the connectivity of financial elements flowing across physical boundaries exhibits a non-linear increasing trend.

To address the radiating and driving effect of high-level cities on surrounding areas, a weighted degree centrality and spatial decay gradient calculation module is introduced into the mapped geographic map structure. A global measurement of the edge weights (i.e., the spillover effect intensity) of each node connection is conducted. The results indicate that the average out-degree centrality of the core nodes in the eastern coastal region is as high as 0.78, forming an absolute high-potential spillover center for financial resources. Further network edge weight attribution analysis shows that the digital financial momentum transmitted from these high-level nodes to adjacent central and western nodes through the topological network contributes 41.2% to the incremental industrial upgrading index of the receiving nodes, and the spatial decay coefficient of this spillover effect remains stable at 0.024 per 100 kilometers. This series of matrix operations and topology reconstruction steps accurately captures the nonlinear spatial evolution trajectory of financial elements diffusing from high gradients to low gradients, verifying the objective existence of the spatial network topology mechanism at the algorithm level, and effectively solving the technical defects of regional isolation in traditional macro panel data modeling.

As shown in Figure 7, red nodes represent high-potential core nodes, while gray nodes represent other provinces. The color of the connecting lines is gradient-mapped according to the spillover intensity, with red indicating strong spillover and blue indicating weak spillover. The spatial distribution pattern reveals that eastern coastal provinces such as Beijing, Shanghai, Guangdong, Zhejiang, Jiangsu, and Shandong form signif icant core radiation clusters, with densely connected nodes and warmer colors, indicating that these regions are exporting high-intensity digital finance spillover momentum to surrounding and central-western regions. In contrast, the node connections in western and northeastern provinces are sparser and cooler in color, reflecting their primary position as receivers in the spatial network. This visualization result highly matches the weighted degree centrality measurement data, intuitively verifying the gradient diffusion characteristic of digital inclusive finance—"strong in the east, weak in the west"—from a geospatial perspective, providing empirical evidence at the spatial topology level for the differentiated policy formulation of regional collabo rative development strategies.

E. HETEROGENEITY AND MODEL ROBUSTNESS TESTING

To examine the spatial evolution differences and heterogeneity of core driving paths under different regional economic bases, this paper decouples the global spatiotemporal graph network into three independent local topological subgraphs: eastern, central, and western. Measurements of the core

TABLE 3. Topological characteristics and quantitative measurement of cross-regional spillover network for digital inclusive finance

Indicator Category	Core Evaluation Indicator	Calculated Value / Range	Algorithm and Underlying Calculation Logic
Global Network Topology	Total Effective Cross-Regional Spillover Edges	428 edges	Extract the optimal spatial adjacency matrix of the model and perform truncation filtering by setting an edge weight threshold of ≥ 0.05.
Global Network Topology	Evolution of Global Spatial Network Density	0.14 (2011)→0.39 (2022)	Calculate the ratio of the actual number of effective edges in a specific period to the theoretical maximum number of edges in the network, $N(N-1)$.
Regional Node Status	Average Weighted Out-Degree Centrality of Eastern Core Nodes	0.78	Normalize and sum the weights of topological edges directed outward from core nodes (e.g., Yangtze River Delta and Pearl River Delta).
Spillover Effect Quantification	Contribution of the East to Industrial Upgrading Increment in the Central and Western Regions	41.2%	Sever the edge weights for cross-regional feature propagation between the East and West, and perform a feature attribution measurement based on the decline amplitude of the model's receiving-end output.
Spillover Effect Quantification	Spatial Decay Coefficient of the Spillover Effect	0.024 / 100 km	Calculated by executing an exponential regression fit on the cross-regional edge weight tensor extracted by the model alongside the spherical geographic distance matrix.

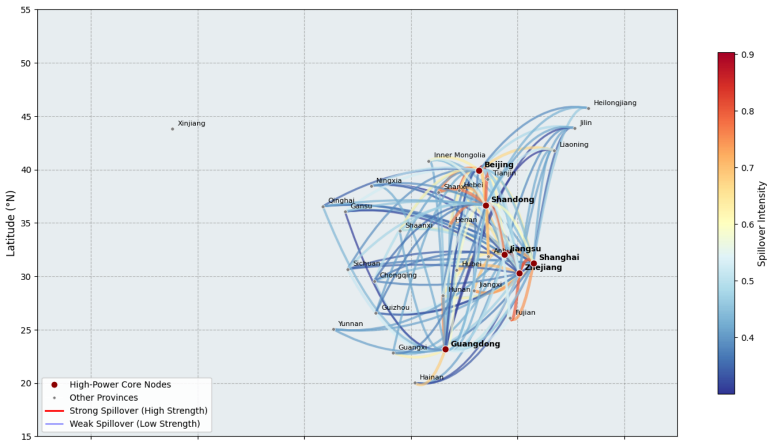

FIGURE 7. Spatiotemporal network topology diagram of Chinese provincial-level administrative regions based on GCN feature extraction

path classification task performed on each subgraph show significant regional heterogeneity in the extraction of driving paths: in the highly financialized and digitalized eastern sub-network, the characteristics of dominant transmission paths such as "technological innovation" are most fully activated, with the harmonic mean index (F1) for this path peaking at 94.0%; in the central region, due to the decrease in factor flow density, this index converges to 91.4%; while in the western region, because it is still in the early stages of inclusive finance filling the credit gap, the connectivity of the underlying financial infrastructure is weak, resulting in relatively sparse path transmission signals, with an F1 of 88.7%. The aforementioned cross-sectional subset verification not only quantitatively confirms the spatial diminishing returns of efficiency driven by digital finance, but also profoundly reveals the heterogeneous nature of the transmission path—namely, the eastern region has firmly established an advanced path of "dual-engine drive of innovation and consumption," while the western region still heavily relies on the basic path of "easing financing constraints." Despite this heterogeneity, the model's global weighted mean still accurately approximates the baseline, demonstrating that the architecture possesses excellent high-dimensional feature mapping generalization ability when facing complex economic bases.

To address the unavoidable statistical accounting biases and exogenous structural shocks in macroeconomic panel data series, adversarial perturbations are further injected into the input of the underlying graph data to conduct stress tests. Specifically, Gaussian white noise with a standard deviation of 0.1 is randomly super imposed on the input discrete node feature matrix to simulate data measurement errors; simultaneously, a random discard operation is performed on up to 15% of the topological connections within the spatial adja cency matrix to replicate the scenario of financial factor transmission disruption under extreme conditions. Under this stringent interference constraint, the forward propagation and backward verification processes are restarted. The core driving node classification accuracy of the global noise test group remains stable at 90.4%, and the model prediction performance does not experience gradient divergence or a precipitous collapse. The performance evaluation of node-driven path classification in different regions and the results of the noise stress test are shown in Figure 8.

Figure 8 shows the F1 score distribution ranges for the four categories: Eastern, Central, Western, and Global Noise test groups. The scatter plots reflect the specific performance dispersion of the samples from each province. The Eastern

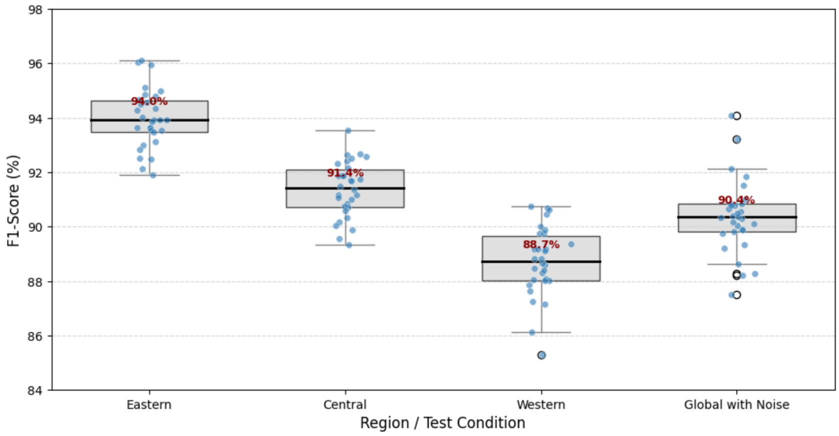

FIGURE 8. Box plot and scatter plot of F1-Score performance evaluation for node-driven paths in different regions

region has the highest and most compact box position, with an average of nearly 94.0%, indicating that the model has the strongest recognition stability in economically developed areas. The Central and Western regions have progressively lower box positions, with averages at 91.4% and 88.7% respectively, visually demonstrating the heterogeneity of driving efficiency decreasing with geographical gradient. Notably, although the global noise test group has a slightly lower box position than the Eastern region, its average remains high at 90.4%, with fewer outliers. This structural degradation stress test confirms from the underlying physical mechanism of the algorithm that the composite spatiotemporal coupled network can effectively filter and absorb unstructured economic noise through multi-level neighborhood aggregation and long-term smoothing operations, effectively ensuring the robustness of the path recognition model in complex realworld environments.

V. CONCLUSION

This paper addresses the challenges of spatiotemporal dynamic evolution complexity and the black-box nature of transmission mechanisms in the process of digital inclusive finance driving industrial upgrading. It proposes and constructs a spatiotemporal graph convolutional neural network recognition model that integrates a multi-head attention mechanism. In terms of algorithm execution, this research first overcomes the limita tions of traditional isolated panel data by fusing geographical and economic distance matrices to reconstruct discrete regional macroeconomic indicators into a multi-dimensional dynamic spatiotemporal graph network topology. Second, it combines graph convolutional networks and temporal convolutional networks to deeply aggregate the spatial spillover effects of cross-regional financial element flows and the nonlinear temporal lag characteristics of long economic cycles. Finally, it embeds a spatiotemporal multi-head attention mechanism into the top-level architecture to achieve dynamic quantification and high-dimensional stripping of core inter mediate transmission paths, completing the leap from underlying data features to "white-boxed" economic logic. Empirical testing based on long-term panel data from Chinese provincial levels fully verifies the effec tiveness and superiority of this architecture. The experimental results are mainly reflected in four aspects: first, in terms of fitting accuracy, the ASTGCN model successfully eliminates spatiotemporal fragmentation errors, with its root mean square error and mean absolute percentage error decreasing significantly by 18.4% and 21.2% respectively compared to the benchmark LSTM model. Second, in terms of path deconstruction, the model accurately quantifies "technological innovation" as the core mediating transmission path (with a weight of 38.6%), and reveals the fundamental shift in the driving force from "capital factor-driven" to "innovation technology-driven" around 2017 through temporal evolution tracking. Third, spatial topology reconstruction intuitively confirms the strong financial spillover effect from high-energy nodes in the eastern coastal areas to the central and western regions. Fourth, in heterogeneous region classification and noise stress testing, the model demonstrates strong robustness and generalization ability. This research provides reliable algorithmic support for regional coordinated development and precise digital finance policy implementation.

AUTHOR CONTRIBUTIONS

Conceptualization –DingRui and ZengQingjun; methodology – DingRui and ZengQingjun; investigation – DingRui and ZengQingjun; writing-original draft preparation – DingRui and ZengQingjun. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research was supported by,

Key Project of Chongqing Municipal Education Commission: Research on the Mechanism and Path of Industrial Coordinated Development between Cities in the Chengdu Chongqing Economic Circle (21SKGD078)

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] L. Chen and G. Zhang, “Exploring the Impact of Digital Inclusive Finance and Industrial Structure Upgrading on High-Quality Economic Development: Evidence from a Spatial Durbin Model,” Economies, vol. 13, no. 8, pp. 212-244, 2025, doi: 10.3390/economies13080212.

[2] Z. Wang, S. Lin, Y. Chen, et al., “The digital economy and real economy: The dynamic interaction effect and the coupling coordination degree,” Sustainability, vol. 16, no. 13, pp. 5769-5789, 2024, doi: 10.3390/su16135769.

[3] S. Lin and W. Ma, “Digital inclusive finance and industrial structure upgrade-Based on nonlinear relationship per-spective,” Journal of Advanced Computational Intelligence and Intelligent Informatics, vol. 27, no. 2, pp. 251-258, 2023, doi: 10.20965/ jaciii.2023.p0251.

[4] L. Sun and C. Zhu, “Impact of Digital Inclusive Finance on Rural High-Quality Development: Evidence from China,” Discrete Dynamics in Nature and Society, vol. 2022, no. 1, pp. 1-13, 2022, doi: 10.1155/2022/7939103.

[5] K. Zhu, R. Ma, and L. Du, “Does digital inclusive finance affect the urban green economic efficiency? New evidence from the spatial econometric analysis of 284 cities in China,” Environmental Science and Pollution Research, vol. 30, no. 23, pp. 63435-63452, 2023, doi: 10.1007/s11356-023-26619-9.

[6] X.-L. Song, Y.-G. Jing, and E. Akeba’erjiang K, “Spatial econometric analysis of digital financial inclusion in China,” International Journal of Development Issues, vol. 20, no. 2, pp. 210-225, 2021, doi: 10.1108/IJDI-05-2020-0086.

[7] D. Tang, Z. Zhao, W. Shen, et al., “Research on the impact of digital finance on the industrial structure upgrading of the Yangtze River economic belt from the perspective of R&D innovation,” Sustainability, vol. 15, no. 1, pp. 425-439, 2022, doi: 10.3390/su15010425.

[8] Y. Wang and L. Li, “Digital economy, industrial structure upgrading, and residents’ consumption: Empirical evidence from prefecture-level cities in China,” International Review of Economics & Finance, vol. 92, no. 1, pp. 1045-1058, 2024, doi: 10.1016/ j.iref.2024.02.069.

[9] J. He, X. Du, and W. Tu, “Can corporate digital transformation alleviate financing constraints? Applied Economics,” 2024; 56(20):2434-2450, doi: 10.1080/00036846.2023.2187037 .

[10] M. Cortés Rufé and J. Martí Pidelaserra, “Mapping financial connections: Market integration in emerging economies through graph theory,” Risks, vol. 12, no. 10, pp. 154-187, 2024, doi: 10.3390/risks12100154.

[11] Y. Wei, J. Watada, and Z. Wang, “Topology unveiled: A new horizon for economic and financial modeling,” Mathematics, vol. 13, no. 2, pp. 325-359, 2025, doi: 10.3390/math13020325.

[12] X. Ji, K. Wang, H. Xu, et al., “Has digital financial inclusion narrowed the urban-rural income gap: The role of entrepreneurship in China,” Sustainability, vol. 13, no. 15, pp. 8292-8310, 2021, doi: 10.3390/su13158292.

[13] P. Yang, M. Fan, Z. Li, et al., “Digital finance, spatial spillover and regional innovation efficiency: New insights from China,” Electronic Research Archive, vol. 30, no. 12, pp. 4635-4657, 2022, doi: 10.3934/era.2022235.

[14] X. Song, X. Qin, W. Wang, et al., “Exploring the spatial spillovers of digital finance on urban innovation and its synergy with traditional finance,” Emerg Sci J, vol. 9, no. 2, pp. 45-62, 2025, doi: 10.28991/ESJ-2025-09-01-024.

[15] C. Rahal, M. Verhagen, and D. Kirk, “The rise of machine learning in the academic social sciences,” AI & SOCIETY, vol. 39, no. 2, pp. 799-801, 2024, doi: 10.1007/s00146-022-01540-w.

[16] C. Ubal, G. Di-Giorgi, E. Contreras-Reyes J, et al., “Predicting the long-term dependencies in time series using recurrent artificial neural networks,” Machine Learning and Knowledge Extraction, vol. 5, no. 4, pp. 1340-1358, 2023, doi: 10.3390/ make5040068.

[17] W. Jiang, J. Luo, M. He, et al., “Graph neural network for traffic forecasting: The research progress,” ISPRS Inter-national Journal of Geo-Information, vol. 12, no. 3, pp. 100-135, 2023, doi: 10.3390/ijgi12030100.

[18] P.-D. Arsenault, S. Wang, and J.-M. Patenaude, “A survey of explainable artificial intelligence (XAI) in financial time series forecasting,” ACM Computing Surveys, vol. 57, no. 10, pp. 1-37, 2025, doi: 10.1145/3729531.

[19] C. Özden, “Modelling financial contagion in complex equity networks: A decision support framework based on graph neural networks,” Business & Management Studies: An International Journal, vol. 14, no. 1, pp. 32-47, 2026, doi: 10.15295/ bmij.v14i1.2682.

[20] D. Cheng, Z. Niu, J. Li, et al., “Regulating systemic crises: Stemming the contagion risk in networked-loans through deep graph learning,” IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 6, pp. 6278-6289, 2022, doi: 10.1109/ TKDE.2022.3162339.

[21] Y. Zhang, Y. Wang, S. Gao, et al., “Context-aware knowledge graph framework for traffic speed forecasting using graph neural network,” IEEE Transactions on Intelligent Transportation Systems, vol. 26, no. 3, pp. 3885-3902, 2024, doi: 10.1109/ TITS.2024.3520511 .

[22] M. Patel, K. Jariwala, and C. Chattopadhyay, “A systematic review on graph neural network-based methods for stock market forecasting,” ACM Computing Surveys, vol. 57, no. 2, pp. 1-38, 2024, doi: 10.1145/3696411.

[23] S. Zhao, Z. Chen, Z. Xiong, et al., “Beyond Grid Data: Exploring graph neural networks for Earth observation,” IEEE Geoscience and Remote Sensing Magazine, vol. 13, no. 1, pp. 175-208, 2024, doi: 10.1109/MGRS.2024.3493972.

[24] W. Zhan and A. Datta, “Neural networks for geospatial data,” Journal of the American Statistical Association, vol. 120, no. 549, pp. 535-547, 2025, doi: 10.1080/01621459.2024.2356293.

[25] M. Li, F. Wang, X. Jia, et al., “Multi-source data fusion for economic data analysis,” Neural Computing and Ap-plications, vol. 33, no. 10, pp. 4729-4739, 2021, doi: 10.1007/s00521-020-05531-0.

[26] L. Zhuoga and X. Zhao, “Study on the impact of digital economy on the high-quality development of cultural industry: the perspective of the new development philosophy,” Journal of Applied Economics and Policy Stiudies, pp. 19 (4), 2026, doi: 10.54254/2977-5701/2026.32498.

[27] Y. Yu, O. Li V, C. Lam J, et al., “GCN-ST-MDIR: Graph convolutional network-based spatial-temporal missing air pollution data pattern identification and recovery,” IEEE Transactions on Big Data, vol. 9, no. 5, pp. 1347-1364, 2023, doi: 10.1109/ TBDATA.2023.3277710.

[28] A. Salim and S. Sumitra, “Spectral graph convolutional neural networks in the context of regularization theory,” IEEE Transactions on Neural Networks and Learning Systems, vol. 35, no. 4, pp. 4373-4384, 2022 , doi:10.1109/TNNLS.2022.3177742.

[29] C. Fan, Y. Yang, and A. Mostafavi, “Neural embeddings of urban big data reveal spatial structures in cities,” Human-ities and Social Sciences Communications, vol. 11, no. 1, pp. 1-15, 2024, doi: 10.1057/s41599-024-02917-6.

[30] A. Raza and F. Ahmed, “Big Data Analytics for Macroeconomic Policy Evaluation,” Journal of Empirical Econom-ics and Multidisciplinary Methods, vol. 1, no. 3, pp. 148-153, 2025.

[31] H. Yousuf, M. Lahzi, A. Salloum S, et al., “A systematic review on sequence-to-sequence learning with neural network and its models,” International Journal of Electrical & Computer Engineering (2088-8708), vol. 11, no. 3, pp. 1-12, 2021, doi: 10.11591/ijece.v11i3.pp2315-2326.

[32] A. Gu, I. Johnson, K. Goel, et al., “Combining recurrent, convolutional, and continuous-time models with linear state space layers,” Advances in neural information processing systems, vol. 34, no. 1, pp. 572-585, 2021, doi: 10.48550/ arXiv.2110.13985.

[33] G. Loewenstein and Z. Wojtowicz, “The economics of attention,” Journal of Economic Literature, vol. 63, no. 3, pp. 1038-1089, 2025, doi: 10.1257/jel.20241665.

[34] M. Farsani R and E. Pazouki, “A transformer self-attention model for time series forecasting,” Journal of Electrical and Computer Engineering Innovations (JECEI), vol. 9, no. 1, pp. 1-10, 2021, doi: 10.22061/JECEI.2020.7426.391 .

[35] C. Yu, N. Jia, W. Li, et al., “Digital inclusive finance and rural consumption structure-evidence from Peking Uni-versity digital inclusive financial index and China household finance survey,” China Agricultural Economic Re-view, vol. 14, no. 1, pp. 165-183, 2022, doi: 10.1108/CAER-10-2020-0255.

[36] M. Ahmad, A. Majeed, A. Khan M, et al., “Digital financial inclusion and economic growth: Provincial data anal-ysis of China,” China Economic Journal, vol. 14, no. 3, pp. 291-310, 2021, doi: 10.1080/17538963.2021.1882064.

[37] J. Chen and H. Xiang, “Study on the Impact of Industrial Structure Change on Total Factor Productivity in Henan Province,” Academic Journal of Business & Management, vol. 5, no. 15, pp. 125-135, 2023, doi: 10.25236/AJBM.2023.051520.

[38] X. Dai, L. Yan, L. Jianping, et al., “A research on the threshold effect of human capital structure upgrading and industrial structure upgrading-based on the perspective of path dependence,” Quality & Quantity, vol. 56, no. 4, pp. 2645-2674, 2022, doi: 10.1007/s11135-021-01230-7.

[39] J. Shi, Q. Ma, H. Ma, et al., "Scaling law for time series forecasting," Advances in Neural Information Processing Systems, vol. 37, no. 1, pp. 83314-83344, 2024, doi: 10.52202/079017-2650.

[40] X. Wang, X. Zhang, J. Li, et al., "Tensor-based multi-feature affinity graph learning for natural image segmenta-tion," Neural Computing and Applications, vol. 35, no. 15, pp. 10997-11012, 2023, doi: 10.1007/s00521-023-08279-5 .