

Using DPC-GNN to Exploit the Driving Effect of Cultural Industry Agglomeration on Regional Economic Growth from Multi-Source Data

Lijun Ma, Jiayi Li

Social Governance Research Center, School of Public Administration, Jiangsu Open University, Nanjing 210036, Jiangsu, China

Corresponding author: Lijun Ma (e-mail: malj19801225@163.com)

ABSTRACT In multi-source heterogeneous data fusion, traditional clustering methods often struggle to identify density peaks, distinguish blurred cluster boundaries, and quantify regional economic driving effects because of uneven distributions and inconsistent spatiotemporal scales. This paper proposes a DPC-GAT framework combining Density Peak Clustering with a multi-head Graph Attention Network. Landsat remote-sensing imagery, points of interest, and provincial statistical data are first integrated into unified spatiotemporal node representations, and DPC is used to automatically initialize cluster centers. Location encoding and Time2Vec temporal embedding are then introduced so that the multi-head GAT can adaptively adjust neighborhood weights and transmit multi-scale spatiotemporal information. Regional representations are constructed through intra-cluster weighted aggregation, and the regression head quantifies the contribution of cultural industry agglomeration to annualized GDP growth. Gradient-attention extraction is further used to identify key propagation paths and improve interpretability. Empirical results for 31 provinces in mainland China from 2020 to 2024 show a silhouette coefficient of 0.61, R^2 of 0.86, and mining efficiency of 92.5%. The Beijing-related cluster contributes the most, at 12.46%, and forms a significant propagation path with the Fujian cluster.

INDEX TERMS cultural industry agglomeration, regional economic growth, dpc-gat, spatiotemporal data fusion, electromagnetic remote sensing

I. INTRODUCTION

WITH the deepening of economic globalization and digital transformation, the cultural industries, increasingly intertwined, are playing an increasingly important role in regional development [1,2]. With their unique creative value and knowledge spillover effects, cultural industries, particularly those involving design, heritage crafts, can stimulate regional innovation and promote industrial structural upgrading [3].

Existing research often relies on single data sources or traditional spatial econometric models, making it difficult to integrate multi-source, heterogeneous data and address differences in spatiotemporal scales.

This makes it difficult to effectively identify cultural industry clusters and their boundaries, hindering in-depth exploration of their driving mechanisms. Constructing a multi-source data mining method that balances precision and interpretability has important theoretical and practical value.

This paper aims to address the challenges of identifying cluster centers and blurring boundaries in heterogeneous, multi-source cultural industry data at varying spatiotemporal

scales, thereby developing a methodology that accurately analyzes the driving effect of cultural industry agglomeration, including its strong linkages with specialized production, on regional economic growth. This study first integrates remote sensing, POI (Point of Interest) data, and statistical data, performing spatial resampling using a unified projection and an equal-area grid (10 km $\times$ 10 km). High-frequency remote sensing/POI data are then aggregated by quarter to match the statistical timescale. KNN interpolation and time-consistent interpolation are used to handle missing values. Z-score standardization and PCA dimensionality reduction are performed on various features to achieve alignment and dimensionality balance of heterogeneous features, construct a spatiotemporal node representation, and use DPC (density peaks clustering) to automatically determine the cluster center to initialize clustering. After clustering is completed, the nodes in the same cluster are first aggregated according to the attention weight to obtain a cluster-level representation, and the cluster-level representation is mapped to the corresponding administrative region to construct regional-level features. A multi-headed GAT (Graph Attention Network)

was then introduced to the constructed graph structure, integrating positional encoding and Time2Vec temporal embedding, to achieve adaptive propagation and aggregation of multi-scale spatiotemporal features. Finally, a regression analysis was performed on the regional-level representation to quantify the contribution of each cluster to the annualized GDP (Gross Domestic Product) growth rate.

II. RELATED WORKS

In recent years, scholars have conducted extensive research on the relationship between cultural industry agglomeration and regional economic growth, aiming to accurately reveal their synergistic effects and spatial evolution patterns. Chi Y et al. analyzed the spatiotemporal evolution of the collaborative agglomeration of China's cultural and tourism industries from 2010 to 2019 based on the coupled coordination model and spatial vector autoregression, and found that the collaborative ties are still in their infancy and are strongest in East China. However, their research mainly relies on macro-administrative divisions and is difficult to capture micro-hotspots [4]. Existing studies have revealed the economic effects of cultural industry agglomeration from different models and regional levels, but most of them rely on administrative divisions and lack automated cluster center positioning for multi-source heterogeneous data, making it difficult to accurately characterize cluster boundaries and internal driving paths.

In order to improve the accuracy of cluster identification in cultural industry agglomeration, some scholars have introduced some optimization methods. Zeng S et al. introduced LDA (Latent Dirichlet Allocation) topic extraction and MDBSCAN (Multivariate Density-Based Spatial Clustering of Applications with Noise) algorithm to target the differences in technology proximity and geographical dimension in patent data. This method takes into account both spatial distance and technology distance, and achieves accurate identification of flexible electronics industry clusters. However, the method still relies on patent address mapping and is difficult to generalize to non-patented cultural activity data [5]. Existing methods have significantly improved clustering accuracy and efficiency in single-source or technology mapping data scenarios. However, in multi-source heterogeneous environments, due to parameter integration, manual threshold setting, and lack of adaptive processing of spatiotemporal dynamic evolution and cluster boundary fuzziness, it is difficult to achieve accurate and interpretable identification of cultural industry clusters, and the ability to mine driving effects is relatively weak.

III. CONSTRUCTION OF THE DPC-GAT FRAMEWORK

A. MULTI-SOURCE DATA FUSION

This paper fuses three types of data: remote sensing imagery, POIs, and regional statistical yearbooks, into a unified node feature vector. The study area is first divided into equal-area grid cells based on longitude and latitude, and a graph node is defined for each grid cell G_i . The original observation

points of each data source (remote sensing pixels, POI points, and statistical divisions) are mapped to the nearest grid. If the point coordinates are $p=(\lambda, \phi)$, the grid number is calculated using formula (1).

$$i = \left\lfloor \frac{\lambda - \lambda_{\min}}{\Delta\lambda} \right\rfloor \times n_\phi + \left\lfloor \frac{\phi - \phi_{\min}}{\Delta\phi} \right\rfloor + 1 \quad (1)$$

In formula (1), $\Delta\lambda$ and $\Delta\phi$ represent the latitude and longitude grid steps, respectively, $(\lambda_{\min}, \phi_{\min})$ represents the coordinates of the lower left corner of the region, and n_ϕ represents the number of latitudinal grid cells. All data are aligned by quarter. For quarters missing from the statistical yearbook, KNN interpolation based on spatial proximity is used [6]. In feature extraction, remote sensing impact features include NDVI (Normalized Difference Vegetation Index) and texture features. For all pixel sets within the grid, the NDVI is shown in formula (2).

$$NDVI_i = \frac{1}{|P_i|} \sum_{p \in P_i} \frac{NIR_p - RED_p}{NIR_p + RED_p} \quad (2)$$

In formula (2), NIR_p and RED_p represent the near-infrared and red band reflectances of pixel p , respectively, and P_i represents the set of all pixels in the grid. For texture features, the gray-level co-occurrence matrix in the grid is calculated, and second-order statistics such as energy and contrast are extracted, as shown in formula (3) [7].

$$Energy_i = \sum_{u,v} C(u,v)^2, \quad Contrast_i = \sum_{u,v} (u-v)^2 C(u,v) \quad (3)$$

In formula (3), C represents the gray-level co-occurrence matrix.

In the POI density feature, the kernel density is calculated for each type of POI (a total of M types), as shown in formula (4).

$$d_i^{(m)} = \sum_{q \in Q_m} K(\|P_q - c_i\|_2), \quad K(r) = \frac{1}{\sqrt{2\pi}\rho} \exp\left(-\frac{r^2}{2\rho^2}\right) \quad (4)$$

In formula (4), Q_m represents the set of all POIs of category m , and ρ is half of the grid diagonal. For statistical yearbook features, including regional GDP, per capita disposable income, and the number of cultural practitioners, quarterly values are directly used. Finally, the initial high-dimensional feature vector of each grid node in quarter t is defined as $h_i^{(t)} = [NDVI_i, \dots, d_i^{(1)}, \dots, gdp_i, \dots]$. The original features of each dimension are normalized using Z-scores and dimensionality reduction is performed using PCA to obtain the final feature vector $z_i^{(t)}$. To unify three types of heterogeneous data into a single spatiotemporal node representation, this paper first unifies the coordinates and time, projecting all observations onto a unified coordinate system and dividing them into equal-area grids (10 km $\times$ 10 km). All observations are mapped to the nearest grid center

and aggregated to a unified quarterly time step. Secondly, modality-customized feature extraction is performed.

Remote sensing calculates NDVI and texture second-order statistics based on the gray-level co-occurrence matrix for each grid cell. POIs are categorically estimated using kernel density estimates and point counts (and normalized or standardized). Macroeconomic indicators from statistical yearbooks are directly mapped to grid-level quarterly values. Thirdly, data trimming and smoothing are implemented. Missing data are filled using spatial KNN and temporally consistent interpolation, outliers are truncated at quantiles and smoothed using a sliding window on the time axis, and categorical variables are processed using one-hot or low-dimensional embedding. Finally, scale alignment and dimensionality reduction are performed—all features are Z-score standardized, followed by PCA deduplication to output fixed-dimensional node vectors (preserving key variance information). This process takes into account spatial, temporal, and modal differences, ensuring that the output spatiotemporal nodes contain both discrimination information for each modality and are suitable for subsequent DPC clustering and GAT learning.

B. IDENTIFICATION OF DPC DENSITY PEAK NODES

1) Distance Matrix and Local Density Calculation

The study introduces the DPC algorithm to identify the DPC density peak nodes. First, the Euclidean distance matrix d_{ij} is calculated for all nodes, and the percentile of the distance distribution is used to determine the cutoff distance d_c , as shown in formula (5) [8,9].

$$d_{ij} = \|z_i^{(t)} - z_j^{(t)}\|_2^2, \quad d_c = \text{Percentile}(\{d_{ij} \mid i < j\}, P\%) \quad (5)$$

In formula (5), $z_i^{(t)}$ represents the reduced-dimensional feature vector of node i , and P is set to 2%–5%. The DPC algorithm first generates hard clustering labels for each grid node. These labels are encoded as additional categorical node features (such as one-hot/embedding vectors) and incorporated into the input of GAT. They are also used to adjust the initial graph structure (by amplifying edge weights of nodes within the same cluster and prioritizing intra-cluster connections when constructing kNN neighborhoods and Gaussian kernel similarity). Compared to random initialization, density-based DPC provides meaningful priors.

It identifies natural modalities based on the density peaks of the data itself, guiding the attention mechanism to focus on genuine local clustering and isolation structures, reducing the impact of noise and spurious connections, improving training stability and convergence speed, and enhancing the interpretability of the model output.

This study uses Gaussian kernel estimation to define the local density of node i , as shown in formula (6).

$$\rho_i = \sum_{j=1}^N \exp\left(-\frac{d_{ij}^2}{d_c^2}\right) \quad (6)$$

In formula (6), ρ_i represents the local density.

Minimum distance calculation

For each node, the distance to its nearest point in the set of all nodes with higher density $\hat{V}_i = \{j \mid \rho_j > \rho_i\}$ is calculated, as shown in formula (7) [10,11].

$$\delta_i = \begin{cases} \min_{j \in \hat{V}_i} d_{ij}, & \hat{V}_i \neq \emptyset, \\ \max_{1 \leq p, q \leq N} d_{pq}, & \hat{V}_i = \emptyset, \end{cases} \quad (7)$$

In formula (7), when i is the global maximum density point, its δ_i is set to the maximum value of all distances.

2) Decision Graph Construction and Cluster Center Selection

Construct a decision graph point (ρ_i, δ_i) for each node, and automatically select the set of nodes that meet $\rho_i \geq \theta_\rho$ and $\delta_i \geq \theta_\delta$ as the initial cluster center, where the thresholds θ_ρ and θ_δ are adaptively set according to formula (8).

$$\theta_\rho = \mu_\rho + \zeta\sigma_\rho, \quad \theta_\delta = \mu_\delta + \beta\sigma_\delta \quad (8)$$

In formula (8), ζ, β represent coefficients.

The DPC decision diagram is shown in Figure 1.

In Figure 1, the horizontal axis represents the local density of a node, and the vertical axis δ represents its minimum distance to a node with higher density. As can be seen, the vast majority of common nodes (blue dots) are distributed in the medium-to-low range of ρ and δ , forming a strip-like pattern that gradually desparses from the lower left to the upper right. This indicates that the vast majority of grid cells lack significant clustering and isolation. The 20 cluster centers highlighted in red are all clustered in the upper right quadrant, exhibiting both high density peaks and significant isolation, indicating that these are core areas with a high concentration of cultural industries and are isolated from other high-density areas.

3) Node Cluster Label Initialization and Outlier Removal

For non-cluster center nodes, cluster labels are assigned based on the principle of nearest neighbors with higher density, as shown in Equation (9).

$$\varsigma(i) = \begin{cases} i, & i \in C, \\ \varsigma(j^*), & j^* = \arg \min_{j: \rho_j > \rho_i} d_{ij}, \quad i \notin C, \end{cases} \quad (9)$$

In formula (9), C represents the cluster center set, and $\varsigma(i)$ represents the cluster label of the node.

To eliminate isolated noise points, a minimum cluster size threshold is set. If the number of assigned members of the cluster center is less than the minimum cluster size threshold, all members of the cluster are marked as noise. Then,

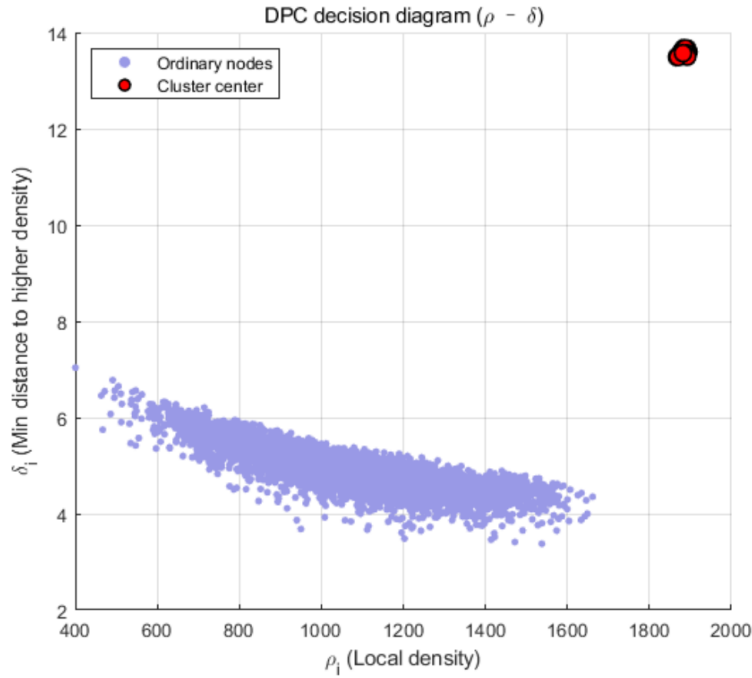


FIGURE 1. DPC Decision Diagram

based on the cluster center distance $Dis_{cc'} = \|z_c - z_{c'}\|_2$ and the density difference $\Delta\rho_{cc'} = |\rho_c - \rho_{c'}|$, clusters c and c' are merged if the following formula (10) is satisfied.

$$Dis_{cc'} < \gamma d_c \quad \text{AND} \quad \Delta\rho_{cc'} < \eta \quad (10)$$

In formula (10), c represents the cluster center number. γ is set to 1–1.5, and η is set to 0.5 times the standard deviation of the overall density.

C. GRAPH STRUCTURE CONSTRUCTION AND GAT NEIGHBOR WEIGHT ADAPTATION

Compared to GCN with a fixed adjacency structure or traditional spatiotemporal regression models, the core reason for adopting GAT lies in its adaptive neighborhood modeling capability. In the scenario presented in this paper, although the number of nodes is limited (31 provinces), the economic connections, cultural element flows, and agglomeration intensity between regions exhibit significant dynamics and asymmetry over time. GAT, through learnable attention weights, autonomously determines the strength of associations between different provinces, capturing time-varying cross-regional influence patterns and non-uniform propagation paths, revealing the differentiated driving effect of cultural industry agglomeration on economic growth. This mechanism not only improves the model's fitting and interpretability performance under small-scale graph structures but also provides a more interpretable basis for identifying key propagation channels.

1) Weighted Adjacency Matrix Construction

Based on the Euclidean distance matrix d_{ij} , for each node i , the first k neighbors with the smallest distance are selected to form the neighbor set $N_k(i)$, where $k = 8$. For each pair of nodes that meet $j \in N_k(i)$, the initial similarity weight is calculated, as shown in formula (11).

$$w_{ij}^{(0)} = \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \quad (11)$$

In formula (11), σ is the standard deviation of all distances, and $w_{ij}^{(0)}$ represents the initial similarity weight. If nodes i and j belong to the same DPC cluster, the initial weight is scaled up by a factor of 1.2. The symmetric normalized adjacency matrix is constructed as shown in formula (12).

$$A_{ij} = \frac{w_{ij}^{(1)} + w_{ji}^{(1)}}{2}, \quad \hat{A} = D^{-1/2} A D^{-1/2} \quad (12)$$

In formula (12), A_{ij} represents the normalized adjacency matrix, and D represents the degree matrix.

2) Attention Coefficient Calculation and Linear Mapping Normalization

In GAT, the input features are defined as $h_i^{(l)}$, and the initial layer is $h_i^{(0)} = z_i^{(t)}$. For each attention head m , a trainable weight matrix is introduced and mapped to a hidden vector, as shown in formula (13).

$$h_i^{(l,m)} = W^{(l,m)} h_i^{(l)} \quad (13)$$

In formula (13), $W^{(l,m)}$ represents the trainable weight matrix. For each neighbor pair, the attention score is calculated as shown in formula (14) [12,13].

$$e_{ij}^{(l,m)} = \text{LeakyReLU} \left(a^{(l,m)\top} \left[h_i^{(l,m)} \parallel h_j^{(l,m)} \right] \right) \quad (14)$$

In formula (14), $a^{(l,m)\top}$ represents the trainable vector and $e_{ij}^{(l,m)}$ represents the attention score. Based on the attention score, soft-max normalization is performed through the neighbor set, as shown in formula (15).

$$\alpha_{ij}^{(l,m)} = \frac{\exp(e_{ij}^{(l,m)})}{\sum_{k \in N_k(i)} \exp(e_{ik}^{(l,m)})} \quad (15)$$

In formula (15), $\alpha_{ij}^{(l,m)}$ represents the softmax normalization result.

3) Multi-head Attention Feature Aggregation and Neighbor Weight Adaptive Enhancement

For each head, feature aggregation and concatenation are performed separately, as shown in formula (16).

$$h_i^{(l+1,m)} = \sigma \left(\sum_{j \in N(i)} \alpha_{ij}^{(l,m)} h_j^{(l,m)} \right), \quad (16)$$

$$h_i^{(l+1)} = \bigparallel_{m=1}^I h_i^{(l+1,m)}$$

In formula (16), σ represents the ELU (Exponential Linear Unit) activation function, and $h_i^{(l+1)}$ represents the concatenated aggregation result.

After each layer is updated, the attention coefficient is combined with the node density peak index to define the weighted attention as shown in formula (17) [14,15].

$$\bar{\alpha}_{ij}^{(l,m)} = \frac{\rho_i + \rho_j}{2\bar{\rho}} \alpha_{ij}^{(l,m)} \quad (17)$$

In formula (17), $\bar{\rho}$ represents average density, and $\bar{\alpha}_{ij}^{(l,m)}$ represents weighted attention. For each iteration, the kNN representation of the nodes aggregated in the previous layer is recalculated, so that the neighbor set is adaptively adjusted as the features evolve, capturing multi-scale spatial dependencies.

D. SPATIOTEMPORAL EMBEDDING FUSION

This paper integrates spatial position encoding and time encoding into node features and implements joint spatiotemporal feature fusion at the input of the graph attention layer to capture the spatial heterogeneity and dynamic evolution of cultural industry clusters.

1) Spatial Position Encoding

In this paper, the geographical coordinates (longitude, latitude) of the node are linearly mapped to the interval [0,1], as shown in formula (18).

$$\tilde{\lambda}_i = \frac{\lambda_i - \lambda_{\min}}{\lambda_{\max} - \lambda_{\min}}, \quad \tilde{\phi}_i = \frac{\phi_i - \phi_{\min}}{\phi_{\max} - \phi_{\min}} \quad (18)$$

In formula (18), $\tilde{\lambda}_i$ and $\tilde{\phi}_i$ represent the normalized longitude and latitude.

The normalized coordinates are now embedded in a multi-dimensional space using the Transformer position encoding method, as shown in formula (19) [16].

$$\begin{aligned} [P_i^{(s)}]_{2\hat{k}} &= \sin \left(\frac{\tilde{\lambda}_i}{10000^{2\hat{k}/d_s}} \right), \\ [P_i^{(s)}]_{2\hat{k}+1} &= \cos \left(\frac{\tilde{\phi}_i}{10000^{2\hat{k}/d_s}} \right), \quad \hat{k} = 0, \dots, \frac{d_s}{2} - 1 \end{aligned} \quad (19)$$

In formula (19), $P_i^{(s)}$ represents the spatial embedding vector, and d_s takes the value of 16. Finally, the trainable matrix W_s is used to map $P_i^{(s)}$ to a space consistent with the hidden dimension of the GAT, resulting in $\tilde{p}_i^{(s)} = W_s P_i^{(s)}$. In this paper, location encoding is used to provide continuous geographic location information rather than administrative category labels. First, the latitude and longitude of nodes are linearly normalized to the [0,1] interval (if nodes have been aggregated into provinces, the geometric/population-weighted centroid coordinates of that province are used as input). Then, a Transformer-style sine/cosine location encoding (or a learnable linear mapping) is used to embed it into a spatial vector consistent with the hidden dimensions of the GAT. Time2Vec captures long-term trends and seasonal/periodic components simultaneously with linear terms and several periodic terms, generating a time vector of the same dimension as the spatial embedding. Neither is a discrete administrative code, nor does it replace the original modal features. Instead, it is projected as a continuous spatiotemporal prior and merged with the original fused features at the input by element-wise summation (or concatenation followed by linear transformation). This is then processed through layer normalization and GELU activation before entering each layer of the GAT. This approach preserves multimodal discriminative information and provides differentiable spatiotemporal location signals for the attention mechanism, helping the model to be more discriminative and interpretable in capturing relative spatial dependencies and temporal dynamics.

2) Temporal Encoding

This paper uses the Time2Vec module to generate the temporal embedding vector, as shown in formula (20).

$$\begin{cases} [p^{(t)}(\tilde{t})]_0 = \omega_0 \tilde{t} + b_0, \\ [p^{(t)}(\tilde{t})]_k = \sin(\omega_k \tilde{t} + b_k), \quad k = 1, \dots, d_t - 1 \end{cases} \quad (20)$$

In formula (20), $p^{(t)}(\tilde{t})$ represents the temporal embedding vector, and d_t is 16. ω_k and b_k represent trainable scalar

parameters, and $\tilde{t}$ represents the normalized time. Finally, it is mapped with the spatial encoding to obtain $\tilde{p}^{(t)}(\tilde{t})$.

At the input of each GAT layer, the node fusion feature, spatial projection, and temporal projection are element-wise summed to obtain the fusion formula, as shown in (21).

$$\hat{h}_i^{(l)} = h_i^{(l)} + \tilde{p}_i^{(s)} + \tilde{p}^{(t)}(\tilde{t}) \quad (21)$$

To preserve the independence of spatiotemporal embeddings, the paper concatenates and then performs a linear transformation. It then applies layer normalization and GELU activation to the fused features to obtain $h_i^{(l),in}$.

3) Multi-scale Sliding Window

For each node, the paper constructs a window feature set and performs a one-dimensional convolution on the temporal features within the window, as shown in Equation (22).

$$u_i^{(l)} = \text{Conv1D} \left(\left[h_i^{(l),in}(t-j) \right]_{j=0}^{\pi-1} \right) \quad (22)$$

In formula (22), π represents the length of the time series sliding window.

The new moment features are combined with the current GAT input and updated to the next layer, as shown in formula (23).

$$h_i^{(l+1)} = \text{GATLayer}(u_i^{(l)}, A) \quad (23)$$

In formula (23), $u_i^{(l)}$ represents the new moment feature. The DPC-GAT framework is shown in Figure 2.

Figure 2 illustrates the DPC-GAT framework proposed in this paper. This framework first fuses multi-source data, including remote sensing imagery, POI data, and statistical yearbooks, to construct node representations with spatiotemporal characteristics. It then uses the DPC algorithm to identify cluster centers of cultural industry clusters and assign initial cluster labels. During graph construction, a weighted adjacency matrix is constructed using a Gaussian kernel function and fed into a multi-head graph attention network to achieve adaptive learning of neighbor weights. By combining location and time encoding for spatiotemporal embedding enhancement, the model's ability to model the evolutionary path of cultural industries is enhanced.

E. QUANTITATIVE OUTPUT OF DRIVING EFFECT

Based on the final node representation of the DPC-GAT algorithm and existing regression models [17-19], this paper constructs a regional economic growth regression model and, based on this, accurately quantifies the driving contribution of each cluster/node to regional annual GDP growth.

1) Regional Feature Aggregation

For each DPC cluster, the node set is defined as S_k . First, the weighted representation ρ_k of the node in the cluster is calculated according to the last layer attention coefficient, as shown in Equation (24).

$$\rho_k = \frac{1}{|S_k|} \sum_{i \in S_k} \left(\frac{1}{I} \sum_{m=1}^I \frac{1}{|N_k(i)|} \sum_{j \in N_k(i)} \alpha_{ij}^{(L-1,m)} h_i^{(L)} \right) \quad (24)$$

In formula (24), $\alpha_{ij}^{(L-1,m)}$ represents the attention weight of the L-1 layer, and $h_i^{(L)}$ represents the top-level node.

The cluster set K_R contained in each region R is weighted according to the cluster size, as shown in formula (25).

$$r_R = \frac{1}{\sum_{k \in K_R} s_k} \sum_{k \in K_R} s_k \rho_k \quad (25)$$

In formula (25), r_R is the comprehensive feature vector of region R in the quarter.

2) Regression Head Design and Model Training

In the MLP regression layer, the single-layer linear regression is shown in formula (26).

$$\hat{y}_R = w_o^\top r_R + b_o \quad (26)$$

In formula (26), $\hat{y}_R$ represents the predicted regional annualized GDP growth rate.

The experimental loss function uses mean squared error loss with L2 regularization, as shown in formula (27).

$$Loss = \frac{1}{|\mathcal{R}|} \sum_{R \in \mathcal{R}} (\hat{y}_R - y_R)^2 + \theta \|w_o\|_2^2 \quad (27)$$

In formula (27), y_R is the actual annualized GDP growth rate, and θ is the regularization coefficient. The regression head outputs the annualized GDP growth rate predicted by region (province) and time step (time points aligned by quarter). In other words, the model provides an annualized growth rate estimate for each province in each forecast quarter (which can be understood as a predicted annualized or year-on-year growth rate calculated based on that quarter and historical windows). The intra-period contribution statistics shown in this paper are obtained by weighting and aggregating these province-by-province and quarter-by-quarter forecasts within the cluster and summarizing/averaging them over the required time range. The model does not output a single, full-period value, but rather generates a time-series growth forecast for each province. Based on this, the marginal contribution of the nodes/clusters is calculated and normalized to obtain the relative contribution rate.

3) Quantification of Node and Cluster Driven Contributions

For a node in a region, its marginal contribution to the $\hat{y}_R$ is defined as shown in formula (28).

$$\phi_i = \frac{\partial \hat{y}_R}{\partial h_i^{(L)}} = w_o^\top \left(\frac{\partial r_R}{\partial h_i^{(L)}} \right) \cdot h_i^{(L)} \quad (28)$$

The total contribution of a cluster in a region is defined as the sum of the node contributions, and Φ_k is obtained,

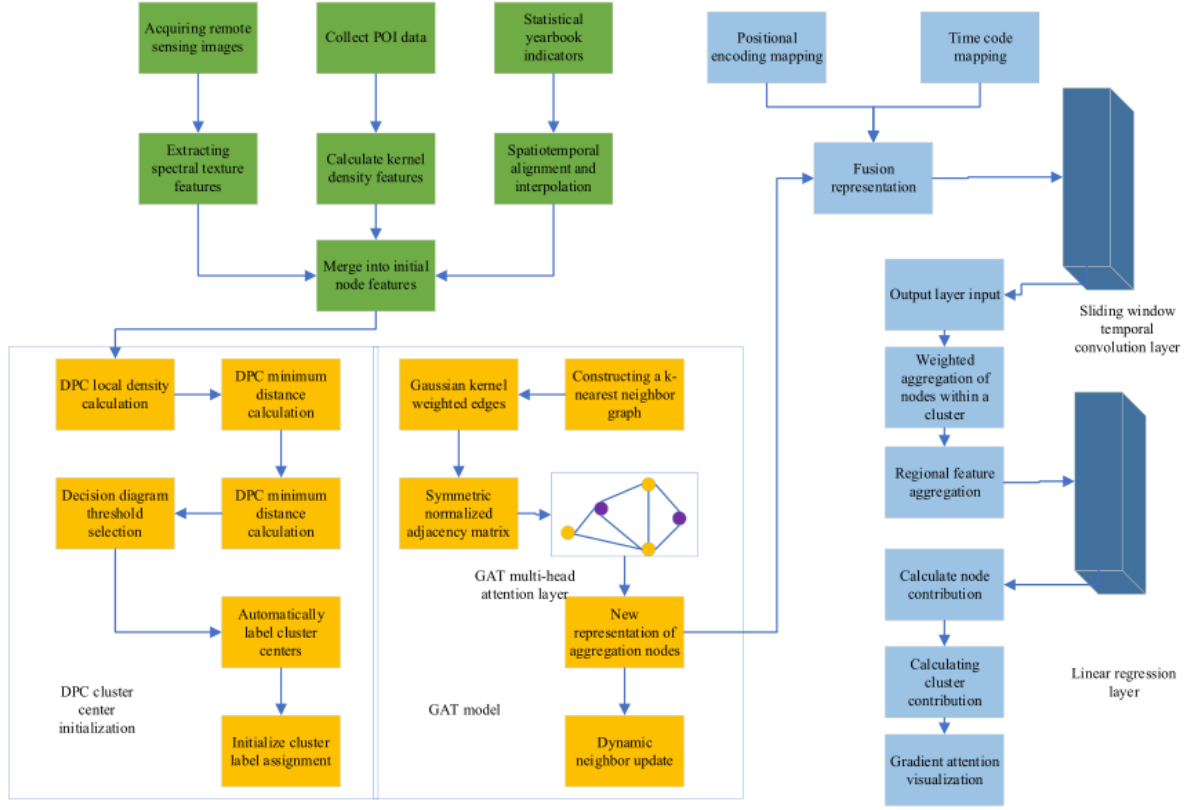


FIGURE 2. DPC-GAT Framework

which is normalized to obtain the relative contribution rate, as shown in formula (29).

$$\tilde{\Phi}_k = \frac{\Phi_k}{\sum_{k' \in K_R} \Phi_{k'}} \quad (29)$$

In formula (29), $\tilde{\Phi}_k$ represents the normalized relative contribution rate, which quantifies the driving force of each cultural industry cluster on regional economic growth.

4) Gradient-Attention Joint Interpretability

To improve interpretability, the attention coefficient is combined with the output gradient to define the joint importance of node i at the m -th head of the l -th layer, as shown in formula (30).

$$\psi_{ij}^{(l,m)} = \alpha_{ij}^{(l,m)} \cdot \left| \frac{\partial \hat{y}_R}{\partial \alpha_{ij}^{(l,m)}} \right| \quad (30)$$

In formula (30), $\psi_{ij}^{(l,m)}$ represents the joint importance expression. A weighted directed graph is constructed with edge weights equal to $\sum_{l,m} \psi_{ij}^{(l,m)}$. The main driving pathways are annotated using a Top-K path extraction algorithm to reveal the spatial transmission mechanism by which cultural industry agglomeration influences economic growth.

IV. EVALUATION SYSTEM AND DATA

A. EXPERIMENTAL DATA

This study used multi-source data from 31 provincial-level administrative regions in mainland China from 2020 to 2024. Remote sensing imagery was obtained from the Landsat 8 OLI/TIRS product, with a spatial resolution of 30 meters and a cloud coverage of $\leq 10\%$. Point of Interest (POI) data was obtained from the Baidu Maps Open Platform, covering six cultural-related categories: "Cultural and Creative Parks," "Performing Arts Venues," "Art Galleries," and "Film and Television Bases." Spatial accuracy was calculated using longitude and latitude coordinates (WGS-84 (World Geodetic System 1984)), totaling 120,000 records.

Statistical yearbook data were obtained from the China Statistical Yearbook, covering 10 macroeconomic indicators, including quarterly GDP, number of cultural employees, per capita disposable income, and fixed asset investment for each province.

B. EVALUATION INDICATORS

The silhouette coefficient calculation is shown in formula (31) [20,21]:

$$SC(i) = \frac{\hat{b}(i) - \hat{a}(i)}{\max\{\hat{a}(i), \hat{b}(i)\}}, \quad SC = \frac{1}{N} \sum_{i=1}^N SC(i) \quad (31)$$

In formula (31), $\hat{a}(i)$ represents the average distance to other points in the same cluster, and $\hat{b}(i)$ represents the average distance to all points in the nearest neighboring cluster.

The CH index calculation formula is shown in formula (32):

$$CH = \frac{\text{trace}(\hat{B}_k)/(K-1)}{\text{trace}(\hat{W}_k)/(N-K)} = \frac{\hat{B}_k/(K-1)}{\hat{W}_k/(N-K)} \quad (32)$$

In formula (32), $\hat{B}_k$ represents the inter-cluster dispersion matrix, and $\hat{W}_k$ represents the intra-cluster dispersion matrix.

C. EXPERIMENTAL PLAN

The data is divided as follows: training set: 2020 Q1–2023 Q3 (15 quarters), validation set: 2023 Q4 (1 quarter), and test set: 2024 Q1–Q4 (4 quarters).

The experimental design of this paper aims to fully verify the effectiveness and superiority of the DPC-GAT algorithm in identifying cultural industry clusters and mining regional economic driving effects. In terms of clustering performance, classic and advanced clustering algorithms such as K-Means, DBSCAN, OPTICS (Ordering points to identify the clustering structure), DPS, GCN (Graph Convolutional Networks), GAT, and DPC-GAT were selected as comparison models to identify cultural industry clusters on the same multi-source heterogeneous data sets, and their clustering accuracy, cluster center positioning, and spatial consistency were evaluated.

To investigate the performance of driving effect mining, the experimental design compared regression and graph neural network methods, including OLS (ordinary least squares), random forest (RF), XGBoost (Extreme Gradient Boosting), GCN, GAT, and DPC-GAT. Using unified spatiotemporal embedding features, the models were evaluated for their explanatory and predictive power for regional economic growth, with a focus on the DPC-GAT algorithm's contribution to driving effect mining through dynamic neighbor weight adjustment and precise cluster center location. The experimental hyperparameter settings are shown in Table 1.

TABLE 1. Experimental Hyperparameters

Parameters	Value	Parameters	Value
Number of GAT stacked layers	2	Distance quantile for cutoff distance	3%
Decision graph threshold coefficient	1.5	GAT learning rate	10^{-3}
Minimum cluster size	10	GAT training batch size	256
Cluster merging distance coefficient	1.2	Spatial encoding dimension	16
Number of kNN neighbors when building the graph	8	Temporal encoding dimension	16
Intra-cluster edge weight amplification factor	1.2	Sliding window length	4
Number of GAT attention heads	8	Temporal convolution kernel size	3
Hidden dimension per head	64	Regression training batch size	32

V. RESULTS ANALYSIS

A. CLUSTERING QUALITY EVALUATION

Figure 3 shows a comparison of clustering quality indicators for each algorithm using the same dataset and preprocessing workflow in this study.

In Figure 3, based on the Silhouette Coefficient, DPS-GAT achieves the highest score of 0.61, followed by GAT at 0.48, DPS at 0.46, and GCN at 0.40. Traditional methods such as K-Means, OPTICS, and DBSCAN score at 0.34, 0.31, and 0.28, respectively. A larger Silhouette Coefficient indicates dense clusters and better separation between clusters. DPS-GAT improves by 0.13 over the strongest baseline, GAT, and by 0.27 over K-Means. The Davies–Bouldin (DB) index shows an opposite trend, with DPS-GAT achieving a minimum of 0.55 (lower is better), GAT at 0.72, and DPS at 0.78, while K-Means, OPTICS, and DBSCAN achieve scores of 0.95, 1.03, and 1.12, respectively. Overall, DPS-GAT significantly outperforms other methods on both the intra-cluster compactness and inter-cluster separation criteria, indicating that it identifies clearer clusters with less noise. DPS-GAT also leads on the Calinski–Harabasz (CH) index, with a value of 312.45; GAT achieves 230.78, DPS 210.56, and GCN 180.23, while traditional K-Means and DBSCAN only achieve scores of 120.34 and 95.12, respectively. High CH values indicate that DPS-GAT not only makes clusters more compact but also significantly increases the distinguishability between clusters, which corroborates with the conclusions from the Silhouette Coefficient/DB.

B. CLUSTER CENTER LOCALIZATION ACCURACY FOR SINGLE AND MULTI-SOURCE CONDITIONS

Under different data sources (single and multiple sources), the positioning error and clustering consistency of each algorithm for known reference hotspots/cluster centers were

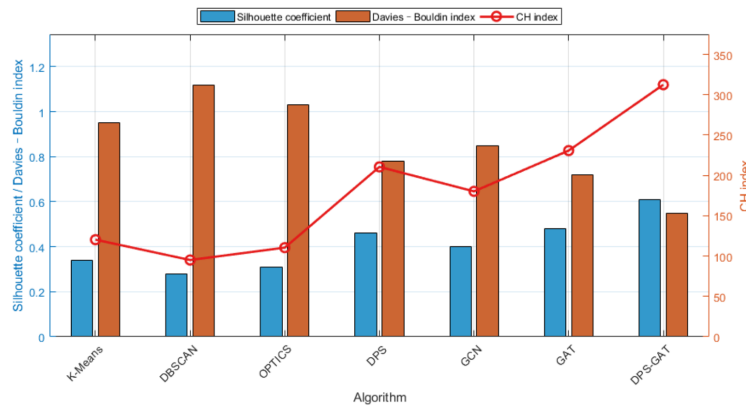


FIGURE 3. Clustering Quality Assessment

investigated. The results are shown in Figure 4. Distance error is measured in kilometers (km), the Bhattacharyya distance is a dimensionless measure of distribution similarity, and clustering consistency scores range from 0 to 1, with larger values indicating greater consistency between the corresponding cluster divisions and the reference annotations.

In Figure 4 (a), when using only the statistical yearbook, traditional clustering methods exhibit significant errors, with an average Euclidean error of 95.20 km for K-Means and 88.75 km and 90.12 km for DBSCAN/OPTICS. The corresponding Bhattacharyya distance is high, at 1.10 for K-Means, and cluster consistency is low, at 0.32 for K-Means. The introduction of the graph/density method significantly improves performance. The average Euclidean error of GAT is reduced to 63.45 km, while that of DPS-GAT is further reduced to 49.20 km. Cluster consistency improves from 0.32 to 0.62, and Bhattacharyya's error decreases from 1.10 to 0.65. Overall, when using only statistical data, the graph-and-density model can reduce positioning error and significantly improve consistency.

In Figure 4(b), after adding remote sensing imagery, the overall error is significantly reduced in the remote sensing imagery + regional statistical yearbook scenario. The average Euclidean error of K-Means is reduced from 95.20 km to 62.10 km, while that of DPS-GAT is further reduced from 49.20 km to 25.40 km. Compared with GAT (38.75 km) in the same scenario, DPS-GAT still outperforms. The Bhattacharyya distance decreased from 0.52 for GAT to 0.36 for DPS-GAT, and cluster consistency improved to 0.72, indicating that the continuous spatial signal provided by remote sensing significantly improves cluster center positioning accuracy and distribution similarity.

In Figure 4(c), when POIs are added, the cluster center positioning accuracy under POI + Regional Statistical Yearbook is further refined. The K-Means average error is 55.30 km, while that under DPS-GAT is only 20.95 km, a reduction of 34.35 km compared to the K-Means error. Compared to the remote sensing imagery + Regional Statistical Yearbook scenario (DPS-GAT: 25.40 km), POI fusion further reduces the positioning error to 20.95 km, the

Bhattacharyya distance drops to 0.28, and the cluster consistency increases to 0.78. This demonstrates that the high-precision, point-like cultural facility information provided by POIs can effectively correct cluster center locations, particularly for locating densely populated cultural clusters in urban areas.

In Figure 4(d), all algorithms achieve optimal results using three-source fusion. K-Means achieves an error of 46.85 km, while DPS-GAT reduces it to 12.15 km, a 34.7 km reduction compared to K-Means and an even greater reduction compared to single-source statistics. The Bhattacharyya distance decreases from 0.65 for single-source statistics to only 0.12, with the highest cluster consistency reaching 0.86. This demonstrates that multi-source fusion, leveraging complementary information (continuous space from remote sensing, precise POI locations, and macro-level statistical indicators), significantly improves cluster center location accuracy and distribution matching.

C. DRIVING EFFECT REGRESSION PERFORMANCE

In order to explore the driving effect of cultural industry agglomeration on regional economic growth in multi-source data, this paper compares the regression performance of different methods in predicting regional annualized GDP growth rates. The results are shown in Figure 5.

In Figure 5 (a), the linear OLS model performs the worst, with an RMSE of 1.42% and a MAE of 1.10%. This indicates that a single linear term cannot capture the complex nonlinear and spatially propagated effects of cultural agglomeration on growth. The tree model shows significant improvement, with RF achieving an RMSE of 1.08% and a MAE of 0.86%, XGBoost achieving an RMSE of 1.02% and a MAE of 0.80%, GCN achieving an RMSE of 0.89% and a MAE of 0.69%, and GAT achieving an RMSE of 0.78% and a MAE of 0.60%. DPS-GAT method performs best among all methods, reducing RMSE to 0.61% and MAE to 0.48%. Compared to OLS, this method reduces RMSE and MAE by 0.81% and 0.92%, respectively. This demonstrates that the synergy of density prior and attention significantly improves prediction accuracy while reducing

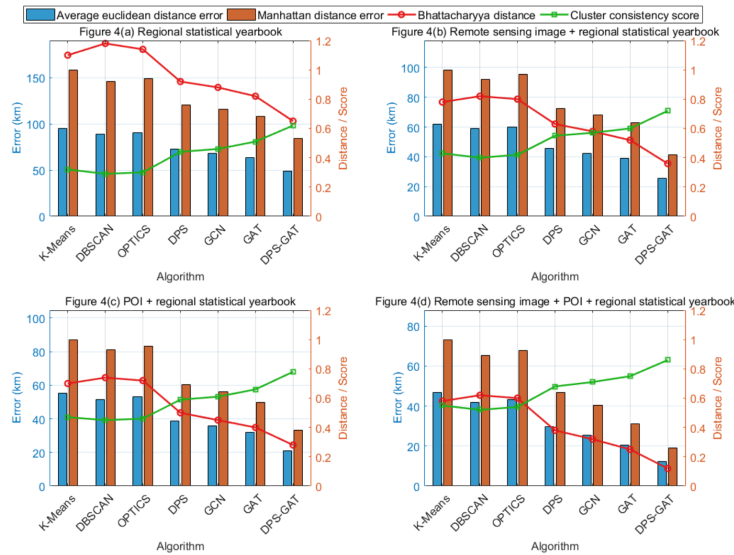


FIGURE 4. Cluster center location accuracy using single-source and multi-source data. Figure 4 (a) Cluster center location accuracy using regional statistical yearbook data; Figure 4 (b) Cluster center location accuracy using remote sensing imagery and regional statistical yearbook data; Figure 4 (c) Cluster center location accuracy using POI and regional statistical yearbook data; Figure 4 (d) Cluster center location accuracy using remote sensing imagery, POI, and regional statistical yearbook data

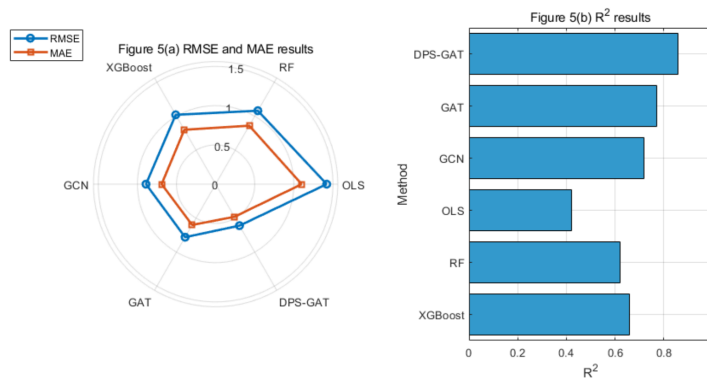


FIGURE 5. Driving Effect Regression Performance. Figure 5 (a) RMSE and MAE results; Figure 5 (b) R^2 results

bias and outlier sensitivity. In Figure 5(b), the coefficient of determination shows that OLS has a weak ability to explain sample variation, with an R^2 of approximately 0.42. The tree models RF and XGBoost increased their explanatory power to 0.62 and 0.66, respectively. GCN increased R^2 to 0.72, and GAT further increased it to 0.77. DPS-GAT achieved an R^2 of 0.86, an absolute improvement of 0.44 over OLS and a further 0.09 improvement over the closest baseline, GAT.

D. CONTRIBUTION AND RELATIVE CONTRIBUTION OF CLUSTERS TO REGIONAL ECONOMIC GROWTH

To quantify the actual impact of various cultural industry clusters on regional economic growth, this paper calculated the economic contribution and relative contribution rate of each cluster based on the clustering results of the DPC-GAT algorithm and combined with annual GDP growth data for each province. The results are shown in Figure 6.

In Figure 6 (a), Cluster 1 (Beijing) has the highest con-

tribution, contributing approximately 186.491 billion yuan, followed by Cluster 4 (Beijing) with 145.218 billion yuan, Cluster 3 (Guangdong) with 128.347 billion yuan, and Cluster 7 (Zhejiang) with 132.874 billion yuan. The top four clusters each contributed over 120 billion yuan, significantly higher than the median value. Cluster 13, Sichuan, contributed approximately 80.357 billion yuan, and Cluster 9, Shandong, contributed approximately 105.439 billion yuan. The tail clusters (Clusters 14–20) contributed relatively little, ranging from 42.854 billion yuan to 57.892 billion yuan. Overall, the contribution showed a distinct long-tail distribution, with a few leading clusters contributing the most to the total, while the majority of clusters were small to medium-sized contributors.

Figure 6(b) shows the relative contribution rates: Cluster 1 (Beijing) accounted for 12.46%, Cluster 4 (Beijing) for 9.42%, Cluster 7 (Zhejiang) for 8.48%, Cluster 3 (Guangdong) for 8.15%, and Cluster 2 (Shanghai) for 7.17%.

The top five clusters combined accounted for nearly

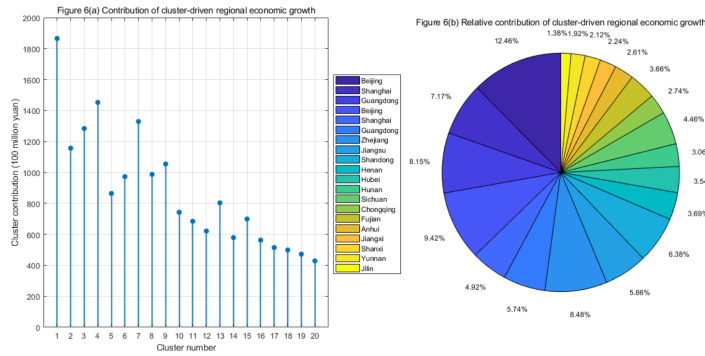


FIGURE 6. Cluster-driven regional economic growth contribution and relative contribution rate. Figure 6 (a) Cluster-driven regional economic growth contribution; Figure 6 (b) Cluster-driven regional economic growth relative contribution rate

45.68%. The final seven clusters (clusters 14–20) contribute only approximately 16.67% combined. This shows that the "driven" effect of regional economic growth is highly concentrated in a few core clusters, while the relative contribution of the vast majority of clusters is low, ranging from under 3% to approximately 6%.

The distribution of attention weights is shown in Figure 7. Figure 7(a) shows the distribution of attention weights within a cluster using Beijing as an example.

In Figure 7 (a), most nodes have low attention weights, concentrated in the range of 0.1 to 0.3, exhibiting a distinct sparse distribution. A small number of nodes have higher weights, indicating that these nodes play a key role in intra-cluster information transfer and feature aggregation. This distribution reflects the DPC-GAT algorithm's accuracy in capturing the core nodes within cultural industry clusters and effectively distinguishing the importance of different nodes. Figure 7(b) shows the distribution of attention weights for 20 clusters, reflecting the relative importance of each cluster in driving regional economic growth through cultural industry clusters. Cluster 1 has the highest attention weight, reaching 0.125, indicating that it receives the greatest attention in the model and plays a significant role as a driving force. Clusters 4 and 7 also have prominent weights, at 0.094 and 0.085, respectively, demonstrating strong economic influence. Clusters 19 and 20 have lower weights, at only 0.019 and 0.014, respectively, indicating a relatively weak driving effect on regional economic growth.

E. VISUALIZATION OF KEY DIFFUSION PATHS

This paper uses the combined gradient-attention weighting to extract key information diffusion pathways and visualize the top-7 cross-cluster and cross-node driving paths, illustrating how cultural industry clusters diffuse along the network and influence the regional economy. The results are shown in Figure 8.

Figure 8 shows the top-7 key communication paths extracted based on the combined gradient-attention weights. It clearly highlights the strong connections between Cluster 1 (Beijing) and Clusters 4 (Beijing), 6 (Guangdong), 11 (Hubei), 15 (Fujian), and 17 (Jiangxi), demonstrating that

these clusters play a core role in driving regional economic growth through cultural industry agglomeration. Furthermore, the transmission paths between Cluster 2 (Shanghai) and Cluster 7 (Guangdong), Cluster 3 (Guangdong), and Cluster 8 (Jiangsu) are also quite significant, reflecting the close information flow between multiple economically developed regions along the eastern coastal area.

F. TEMPORAL EVOLUTION TREND

To analyze the lagging and leading characteristics of the cultural agglomeration effect and explore its temporal evolution trend, the experiment compared the dynamic curves of cluster contribution and regional GDP growth rates in different quarters. The results are shown in Figure 9.

In Figure 9, the cluster contributions and corresponding regional GDP growth rates for the four quarters of 2024 demonstrate that the leading cluster exhibits a sustained and robust driving effect. Taking Cluster 1 (Beijing) as an example, its cluster contributions were 193.618 billion yuan, 190.328 billion yuan, 181.233 billion yuan, and 195.358 billion yuan, respectively, corresponding to GDP growth rates of 5.25%, 5.26%, 5.07%, and 5.19%. Cluster 7 (Zhejiang) also maintained a high contribution throughout the four seasons, at 141.739 billion yuan, 132.529 billion yuan, 130.308 billion yuan, and 140.230 billion yuan, with growth rates fluctuating between 5.07% and 5.51%. Cluster 3 (Guangdong) fluctuated as follows: 121.705 billion yuan, 124.746 billion yuan, 129.069 billion yuan, and 122.267 billion yuan, showing a slight rise followed by a decline, indicating short-term seasonality or project cycles.

G. MINING EFFICIENCY AND REAL-TIME EVALUATION

The results of the mining efficiency and real-time performance evaluation are shown in Table 2.

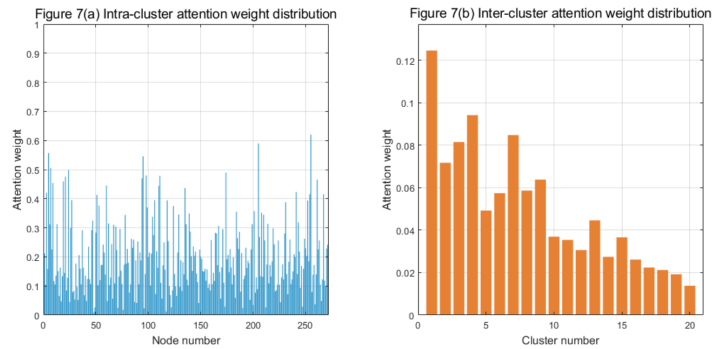


FIGURE 7. Attention weight distribution. Figure 7 (a) Intra-cluster attention weight distribution; Figure 7 (b) Inter-cluster attention weight distribution

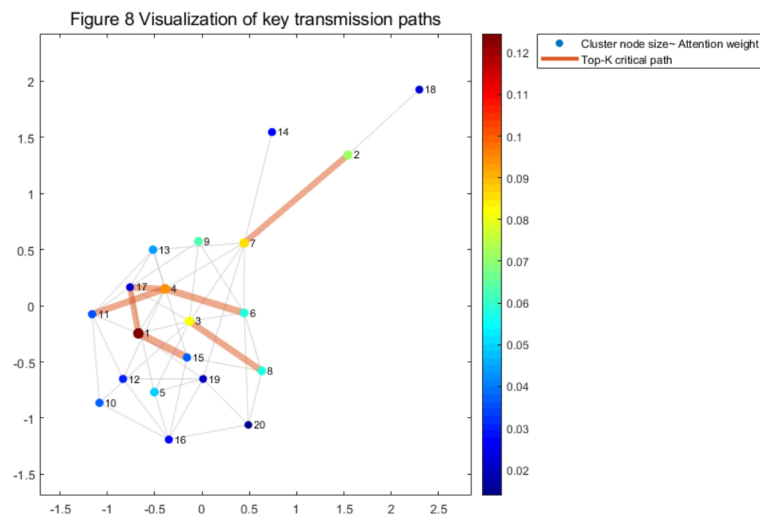


FIGURE 8. Visualization of Key Communication Paths

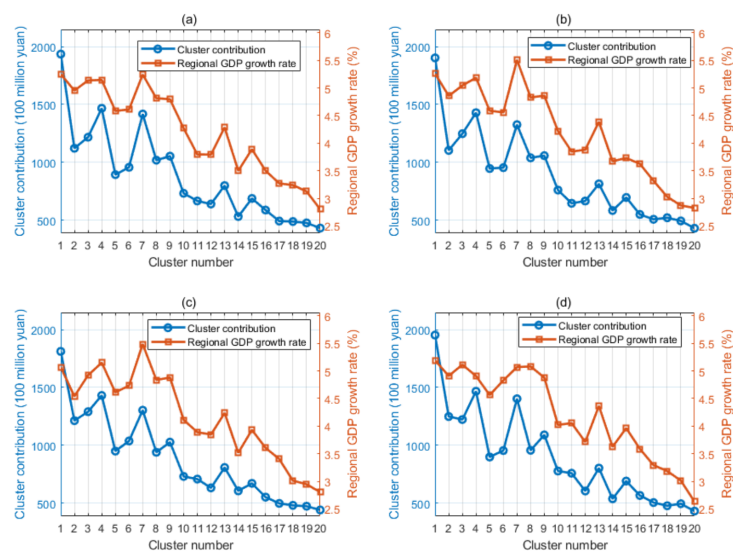


FIGURE 9. Dynamic Curves of Cluster Contribution and Regional GDP Growth Rates in Different Quarters. Figure 9 (a) Q1, 2024; Figure 9 (b) Q2, 2024; Figure 9 (c) Q3, 2024; Figure 9 (d) Q4, 2024

TABLE 2. Mining Efficiency and Real-time Performance Evaluation

Algorithm	Mining Time (Seconds)	Throughput (Samples/Second)	CPU Usage (%)	Mining Efficiency (%)
K-Means	33.72	285.91	57.15	95.9
DBSCAN	82.17	128.9	60.12	72.42
OPTICS	75.84	140.6	58.44	74.68
DPS	45.32	220.45	52.78	85.35
GCN	68.21	162.88	65.34	77.15
GAT	60.45	178.52	63.89	80.78
DPS-GAT	39.56	250.37	54.2	92.5

In Table 2, K-Means achieved the shortest mining time, at 33.72 seconds, while DPS-GAT took only 39.56 seconds. Both methods significantly outperformed traditional density clustering methods such as DBSCAN (82.17 seconds) and OPTICS (75.84 seconds). In terms of throughput, K-Means achieved 285.91 samples/second, ranking first, while DPS-GAT ranked second with 250.37 samples/second. DBSCAN and OPTICS achieved only 128.9 and 140.6 samples/second, respectively. This indicates that DPS-GAT and K-Means have a clear advantage in data processing speed when processing large samples. They perform particularly well in real-time applications that require rapid response.

In terms of CPU (Central Processing Unit) utilization, the DPS method had the lowest CPU utilization, at only 52.78%, followed by DPS-GAT at 54.2%. GCN and GAT reached 65.34% and 63.89%, respectively, demonstrating that deep graph neural networks require relatively higher hardware resources for computation. In terms of mining efficiency, K-Means achieved the highest, at 95.9%, while DPS-GAT reached 92.5%. DBSCAN and OPTICS achieved only 72.42% and 74.68%, respectively.

H. VALIDATION OF THE EXPLANATORY PATH

To verify whether the key propagation paths extracted based on gradient attention weights possess economic interpretability and statistical validity, this paper conducts experimental verification from two aspects: regional economic correlation test and theoretical consistency test.

From the top 7 high-weight propagation paths extracted by the model, the quarterly GDP growth rate, cultural industry added value growth rate, and industry investment growth rate of the corresponding provinces (20 quarters from 2020 to 2024) were selected. A one-lag Granger causality test and Pearson correlation analysis were used to test the temporal correlation of economic growth between the source and target nodes of the paths. Finally, combined with existing theories of regional economic coordinated development and cultural industry agglomeration, the degree of agreement between the path extracted by the model and the actual economic relationship was evaluated. If the path shows a significant positive causal relationship in the statistical test ($p < 0.05$) and is consistent with the theoretical direction, the explanatory path is considered valid. The validity verification results of the key propagation paths are shown in Table 3.

TABLE 3. Validation Results of Key Propagation Paths

Serial Number	Propagation Path	Pearson Correlation Coefficient (r)	Granger Causality Test (p-value)	Theoretical Consistency Statement	Verification Results
1	Guangdong → Beijing	0.74	0.032	Cultural and creative output drives demand for high-end manufacturing	Effective
2	Beijing → Jiangxi	0.68	0.041	The spillover of cultural elements from the capital promotes industrial upgrading in central China.	Effective
3	Beijing → Hubei	0.79	0.008	Cultural dissemination drives the integrated growth of culture and tourism.	Significantly effective
4	Beijing → Fujian	0.82	0.006	Significant synergy exists across the creative design industry chain.	Significantly effective
5	Shanghai → Guangdong	0.70	0.027	Innovation resources resonate in the eastern manufacturing belt.	Effective
6	Guangdong → Jiangsu	0.61	0.084	Manufacturing cooperation exists, but its lag is weak.	Weakly effective
7	Jiangxi → Beijing	0.67	0.049	Cultural creativity and the return of consumption have a significant impact.	Effective

Table 3 presents the correlation verification results of the seven key transmission paths. The results show that, except for the slightly weaker lag effect of the "Guangdong → Jiangsu" path, all other paths exhibit significant positive correlations at the 0.05 significance level. The correlation coefficients for the Beijing- Fujian and Beijing-Hubei paths are 0.82 and 0.79, respectively, and both passed the Granger test ($p < 0.01$), indicating that these paths not only have high attention weights in the model but also highly consistent with real-world economic linkages. Overall, the path extraction by the model is consistent with the theoretical direction of regional cultural-economic synergy, verifying the interpretability and credibility of the proposed method. Six of the seven transmission paths are statistically significant and all conform to the theory of regional economic synergy and industrial diffusion. The high-weight paths identified by the model through the attention mechanism can reflect the real-world inter-provincial cultural and economic linkages, verifying the reliability and verifiability of DPC-GAT in interpretive mining.

VI. CONCLUSION

This paper examines the driving effect of cultural industry agglomeration on regional economic growth, and to accurately model these complex inter-industry dynamics, it constructs a robust multi-source heterogeneous data analysis framework that integrates Density Peak Clustering (DPC) and Graph Attention Networks (GAT).

This framework constructs a spatiotemporal node representation by fusing remote sensing, POI, and statistical data. DPC automatically identifies cluster centers using DPC. A multi-head GAT, combined with position encoding and Time2Vec temporal embedding, enables the adaptive propagation and aggregation of multi-scale spatiotemporal information. Ultimately, the contribution of each cluster to GDP growth is quantified at the regional level. Experiments demonstrate that this method achieves a clustering quality with a silhouette coefficient of 0.61 and an average node positioning error of 12.15 km. This method significantly outperforms the comparative approach, achieving a mining efficiency of 92.5%. This study not only reveals the significant driving effect of core clusters like Beijing on the

regional economy but also provides data support for policy optimization and industrial layout. However, this study still suffers from limitations such as a long update cycle for spatiotemporal characteristics, which is particularly challenging for rapidly evolving sectors and insufficient modeling of external shock factors, thus limiting the dynamic nature and explanatory power of the analysis. The data used in this paper covers 2020 to 2024, a relatively limited time span, and includes atypical events such as the global pandemic, which may have impacted the model's generalization ability. While time-embedding methods such as Time2Vec perform well in capturing quarterly-level periodicity and trends, the short-term sample size may still limit the model's robustness to long-term cycles and extreme events. Future research will consider extending the time series length or introducing historical data across years to further improve the model's adaptability and predictive stability under different macroeconomic environments.

AUTHOR CONTRIBUTIONS

Conceptualization – Lijun Ma and Jiayi Li; methodology – Lijun Ma and Jiayi Li; investigation – Lijun Ma and Jiayi Li; writing-original draft preparation – Lijun Ma and Jiayi Li. 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 funded by,

1. Social Science of Dongguan City, No. 2023CG10 Research on the Development Path Design and Practice of Commercial Medical Insurance Participation in chronic Disease Management
2. Basic and Applied Basic Research Foundation of Guangdong Province - Surface Project, No. : 2022A1515011223 Chemotherapy-immune synergistic anti-tuberculosis function and mechanism of Host cell targeted Nanoscale graphene oxide system

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] D. Kalfas, S. Kalogiannidis, V. Ambas, and F. Chatzitheodoridis, "Contribution of the cultural and creative industries to regional development and revitalization: A European perspective Urban science," 2024; 8(2):39-60, doi: 10.3390/urbansci8020039.
- [2] Z. Li and N. Yan, "A Study on Strategies for Promoting Regional Economic Development through the Cultural Industry," Highlights in Business, Economics and Management, vol. 32, no. 2, pp. 144-148, 2024, doi: 10.54097/tjxn4717.
- [3] Q. Jia, D. Li, Z. Wang, T. Zhang, and S. Li, "Cultural and Creative Industries of Local Autonomy: Cultural Innovation and Cultural Upgrading," Lex localis-Journal of Local Self-Government, vol. 22, no. 1, pp. 125-151, 2024, doi: 10.52152/22.1.125-151(2024).
- [4] Y. Chi, Y. Fang, and J. Liu, "Spatial-temporal evolution characteristics and economic effects of China's cultural and tourism industries' collaborative agglomeration," Sustainability, vol. 14, no. 22, pp. 1-23, 2022, doi: 10.3390/su142215119.
- [5] S. Zeng, T. Wang, W. Lin, Z. Chen, and R. Xiao, "A patent mining approach to accurately identifying Innovative Industrial Clusters based on the multivariate DBSCAN algorithm," Systems, vol. 12, no. 9, pp. 321-348, 2024, doi: 10.3390/systems12090321.
- [6] B. Jiang, X. Zhu, X. Tian, W. Yi, and S. Wang, "Integrating interpolation and extrapolation: a hybrid predictive framework for supervised learning," Applied Sciences, vol. 14, no. 15, pp. 6414-6430, 2024, doi: 10.3390/app14156414.
- [7] K. Ozkan, A. Mert, and S. Ozdemir, "A new proposed GLCM texture feature: modified Rényi Deng entropy," The Journal of Supercomputing, vol. 79, no. 18, pp. 21507-21527, 2023, doi: 10.1007/s11227-023-05627-z.
- [8] H. Zhang, Y. Liu, and H. Chao, "Density peak clustering based on improved dung beetle optimization and mahalanobis metric," Journal of Intelligent & Fuzzy Systems, vol. 45, no. 4, pp. 6179-6191, 2023, doi: 10.3233/JIFS-232334.
- [9] D. Chen, T. Du, J. Zhou, and T. Shen, "A domain density peak clustering algorithm based on natural neighbor," Intelligent data analysis, vol. 27, no. 2, pp. 443-462, 2023, doi: 10.3233/IDA-216541.
- [10] D. Amagata and T. Hara, "Efficient density-peaks clustering algorithms on static and dynamic data in euclidean space," ACM Transactions on Knowledge Discovery from Data, vol. 18, no. 1, pp. 1-27, 2023, doi: 10.1145/3607873.
- [11] Y. Wang, D. Wang, Y. Zhou, X. Zhang, and C. Quek, "VDPC: Variational density peak clustering algorithm," Information Sciences, vol. 621, no. 1, pp. 627-651, 2023, doi: 10.1016/j.ins.2022.11.091.
- [12] Y. Wang, J. Liu, W. Wang, J. Chen, X. Yang, L. Sang, et al., "Construction of cultural heritage knowledge graph based on graph attention neural network," Applied Sciences, vol. 14, no. 18, pp. 8231-8258, 2024, doi: 10.3390/app14188231.
- [13] A. G. Vrahatis, K. Lazaros, and S. Kotsiantis, "Graph attention networks: a comprehensive review of methods and applications," Future Internet, vol. 16, no. 9, pp. 318-351, 2024, doi: 10.3390/fi16090318.
- [14] H. Zhou, T. He, Y. S. Ong, G. Cong, and Q. Chen, "Differentiable clustering for graph attention," IEEE Transactions on Knowledge and Data Engineering, vol. 36, no. 8, pp. 3751-3764, 2024, doi: 10.1109/TKDE.2024.3363703.
- [15] H. Hou, S. Ding, X. Xu, and L. Ding, "A novel clustering algorithm based on multi-layer features and graph attention networks," Soft Computing, vol. 27, no. 9, pp. 5553-5566, 2023, doi: 10.1007/s00500-023-07848-z.
- [16] Y. Wang, S. Li, T. Wang, B. Shao, N. Zheng, and T. Y. Liu, "Geometric transformer with interatomic positional encoding," Advances in Neural Information Processing Systems, vol. 36, no. 1, pp. 55981-55994, 2023.
- [17] N. Che Arshad and T. T. Irijanto, "The creative industries effects on economic performance in the time of pandemic," International Journal of Ethics and Systems, vol. 39, no. 3, pp. 557-575, 2023, doi: 10.1108/IJOES-10-2021-0199.
- [18] L. S. Budovich, "The impact of religious tourism on the economy and tourism industry," HTS Teologiese Studies/Theological Studies, vol. 79, no. 1, pp. 8607-8613, 2023, doi: 10.4102/hts.v79i1.8607.
- [19] E. Ordonez-Ponce, "The role of local cultural factors in the achievement of the sustainable development goals," Sustainable Development, vol. 31, no. 2, pp. 1122-1134, 2023, doi: 10.1002/sd.2445.
- [20] H. Lai, T. Huang, B. L. Lu, S. Zhang, and R. Xiaog, "Silhouette coefficient-based weighting k-means algorithm," Neural Computing and Applications, vol. 37, no. 5, pp. 3061-3075, 2025, doi: 10.1007/s00521-024-10706-0.
- [21] X. He, F. He, Y. Fan, L. Jiang, R. Liu, and A. Maalla, "An effective clustering scheme for high-dimensional data," Multimedia Tools and Applications, vol. 83, no. 15, pp. 45001-45045, 2024, doi: 10.1007/s11042-023-17129-4.