Clothing Pattern Structure Modeling and Reconstruction via Multi-Module Fusion Graph Neural Networks with Path Planning and Reinforcement Learning

Yanfang Duan
Zhengzhou Academy of Fine Arts, Zhengzhou, Henan 451450, China
E-mail：18530858176@163.com

Keywords: graph neural network, clothing pattern structure, intelligent feature extraction, structural reconstruction

Received: August 9, 2025

There are core difficulties in the intelligent recognition and generation application of clothing pattern structure, such as irregular geometric topology, weakened semantic structure, and unstable path planning. To solve such problems, an intelligent feature extraction and structure reconstruction path learning scheme that integrates graph neural networks is constructed. In the stage of structural diagram modeling, a clothing structure diagram is constructed based on the node edge surface configuration relationship. The graph convolutional network is used to embed the spatial adjacency relationship in multiple dimensions, supplemented by attention mechanism to enhance the response ability of key nodes and improve the stability of extracting local salient features. To better express the relationship between structural semantics and geometry, a multi-scale graph embedding strategy and structural context aggregation module are introduced to enable nodes to have stronger expressive power in both topological and semantic dimensions. In terms of reconstructing path generation, a graph autoencoder architecture is introduced to achieve controllable mapping of structure to path space, integrating geometric consistency constraints to enhance structural accuracy. The path decision-making process adopts a reinforcement learning model based on policy gradient, and optimizes the path guidance process through feedback mechanism. This experiment is based on the DeepFashion2 public dataset and our self built clothing structure graph data, with a total of 4826 samples and an average of 43 vertices. The results show that the accuracy index of our model reaches 91.3%+0.5, the Topology Score reaches 88.0%+0.6, and the F1 Structure Score reaches 88.4%+0.6, which is much higher than the basic method. The innovation of this study is mainly reflected in three aspects: proposing the use of graph convolution+attention to achieve multi task feature extraction; Introducing geometric constraints and policy networks to achieve reconstruction methods that maintain path consistency; The first application of GNN in the establishment of clothing style structure brings a new approach compared to traditional graph mapping.

Povzetek: Predstavljen je večmodulni GNN-okvir za inteligentno modeliranje in rekonstrukcijo oblačilnih krojnih struktur. Združuje večskalne GCN, pozornost, geometrijske omejitve ter učenje z okrepitvijo za stabilno načrtovanje poti. Testiran je na 4.826 vzorcih.

1 Introduction

With the deepening development of artificial intelligence and graph neural networks in structural modeling, graphic recognition, and semantic generation, intelligent analysis of graph structured data is becoming an important means of complex structure restoration and information reconstruction. In applications such as intelligent clothing manufacturing and virtual fitting, the modeling of clothing pattern structure serves as an intermediate link, directly affecting the accuracy of 3D reconstruction and the logic of structural restoration. However, clothing structure diagrams have features such as uneven node distribution, non-linear stitching paths, and fuzzy semantic boundaries, which result in insufficient accuracy of traditional methods based on image contours or geometric templates, making it difficult to adapt to diverse pattern organization [1].

Previous studies have attempted to use convolutional neural networks or generative adversarial networks to map images to structures, but there are still shortcomings in expressing complex structures and handling spatial relationships. Especially for clothing graphics with topological constraints and semantic nesting, there is an urgent need to establish a unified graph model framework that combines structural priors, semantic understanding, and path planning capabilities to achieve effective transformation from graphic perception to structural reconstruction. In recent years, Graph Neural Networks (GNNs) have shown good adaptability in processing non-Euclidean structured data, providing a unified mechanism for node propagation, structure perception, and semantic embedding, and providing methodoiogical support for clothing structure modeling [3]. GNN can achieve local fusion by aggregating adjacent node information and perform overall modeling at the layer level, suitable for the structural relationship of "node edge stitching surface" in clothing. After introducing attention mechanism, the recognition accuracy of key parts and important suture paths can be improved, and the robustness of the model can be enhanced [4]. The graph autoencoder and decoder provide the basis for path generation, but there are still challenges in coordinating sequence control and structural constraints. Reinforcement learning has the potential to improve the accuracy and efficiency of path generation due to its adaptive strategy optimization ability, making it suitable for dynamic adjustment during the path generation stage [5].

In actual modeling, the representation of structural diagrams, the accuracy of graph feature extraction, path reconstruction strategies, and control feedback constitute the core of the system. The key to current research is to build a multi module collaborative, feature accurate, path reasonable, and strategy controllable graph model system that balances modeling accuracy and system stability. This study focuses on the core topic of "Intelligent feature extraction and reconstruction path construction of clothing pattern structure by integrating graph neural networks". The technical design and experimental verification are carried out around four dimensions: "structure graph construction - graph feature extraction - path reconstruction generation - strategy guided optimization". This study focuses on the graph feature extraction and path reconstruction of clothing pattern structure. The research question is as follows: RQ1: Can graph neural networks effectively model the spatial semantic structure of clothing patterns? RQ2: Can multitasking and attention mechanisms improve node classification and edge prediction accuracy? RQ3: Can reinforcement learning improve the consistency of structural path reconstruction?

The innovation of this study lies in the fusion of graph convolution, attention, and reinforcement learning to form a collaborative framework; Introducing geometric constraints to enhance the logical consistency of complex structures in tasks; For the first time, GNN has been applied to clothing pattern modeling, expanding its boundaries in the field of industrial design.

2 Related work

In the interdisciplinary research of graph neural networks and structural modeling, the extraction and reconstruction of structural features of clothing patterns have gradually formed a complex task process that integrates multi-source graph data, high-dimensional semantic mapping, and path optimization. Current research mainly focuses on graph structure construction, feature fusion, path prediction, and graph data-driven learning.

In terms of graph structure modeling, Dong et al. (2022) proposed a weighted fusion of convolutional neural networks and graph attention mechanisms for classification tasks in high-dimensional spectral images, effectively enhancing the recognition accuracy and structure preservation ability of graph neural networks for boundary regions, and providing basic support for edge detection in subsequent structure reconstruction [7]. Sun et al. (2024) introduced an adaptive feature fusion module in the attribute graph clustering task and achieved stable clustering results on irregular structured graphs, verifying the enhancing effect of heterogeneous feature combinations on graph structure expression [8]. Liu et al. (2022) constructed a lightweight image super-resolution model based on multi attention mechanism, achieving effective recognition and enhancement of key region map features under limited computing resources [9].

In terms of optimizing the expression of intermediate layers in structural reconstruction paths, Chen et al. (2024) proposed a multi-layer feature radiation field (FeRF) model, which combines deep neural networks with high-dimensional graph structure embedding to achieve multi-scale fusion and hierarchical reconstruction of structural features in image-to-image tasks [10]. Yi (2022) constructed a convolutional neural network model based on clothing design to explore the linear structure distribution and pattern contour recognition in clothing images, providing a preliminary semantic basis for mapping images to pattern structures [11]. Yan et al. (2022) proposed the Semantic Driven Dual Attention Network (SDAN), which utilizes a bidirectional graph attention mechanism to mine semantic distribution relationships in the graph, significantly improving the accuracy of expressing edge connections and region boundaries during the structural restoration process [12].

In image recognition and classification tasks, Liao et al. (2022) combined convolutional networks and attention mechanisms for multi class classification of clothing images, enhancing the differential expression between structural features and demonstrating stronger discriminative ability for image samples within the same category [13]. Ning et al. (2022) constructed a heterogeneous graph transformation relationship network between clothing patterns and e-commerce patterns from the perspective of cross domain image retrieval, solving the interference problem of structural misalignment and fuzzy features on retrieval accuracy [14]. Korosteleva and Lee (2022) proposed the NeuralTailor method, which reconstructs sewing pattern structures from 3D clothing point clouds, achieving structure preserving modeling from 3D to 2D, providing direct technical reference for intelligent reconstruction of clothing pattern structures [15].

In order to compare the differences between existing methods and the work presented in this paper more clearly, the core elements of the main related studies are summarized in Table 1.

Table 1：Comparison and summary of related methods

Method Name	Year	Dataset	Main Method	Accuracy / F1 / Topology	Limitation
CNN-based	2022	Textile dataset	Parallel convolution + optimization	Acc 82%	Difficult to handle complex structural relations
GCN-Net	2023	Synthetic graph data	Heterogeneous GNN feature fusion across layers	Acc 88%	Insufficient for capturing long-range dependencies
SDAN	2022	Image generation tasks	Dual attention mechanism for edge recognition	F1 ≈ 85%	Limited generalization, lacks path modeling
NeuralTailor	2022	3D point cloud	Reconstructing sewing structures from 3D point clouds	Topology ≈ 87%	Restricted to 3D input, lacks path optimization
GNN+Strategy	2024	DeepFashion2 & Custom data	Multi-module fusion + reinforcement learning for path guidance	Acc 91.3% / Topo 88.0% / F1 88.4%	Validation scope limited

3 Intelligent feature extraction mechanism for clothing pattern structure based on fused graph neural network

3.1 Construction of clothing pattern structure diagram and node feature setting

The construction of clothing pattern structure diagram relies on the data format requirements of graph neural network, which requires encoding the geometric structure information in two-dimensional images or CAD drawings into graph data structures with connection relationships. Nodes represent functional areas in the clothing structure, such as armrests, collars, side seams, armholes, etc., while edges represent the stitching relationships or symmetrical connections between different parts. The graph structure is defined as $G=(V,E)$, where $V=\{v_1,v_2,\ldots,v_n\}$ is the set of nodes and $E\subseteq V\times V$ is the set of edges. Each node sets an initial feature vector $x_i\in R^d$ by extracting its position, shape, and structural semantics, which is specifically defined as:

$$F(v_i)=[l_i,\theta_i,k_i,m_i,s_i]\in R^d \quad (1)$$

Among them, l_i represents the length of the structural line, θ_i represents the corner information, k_i represents the local contour curvature, m_i represents the material code, and s_i represents the structural category label.Curvature is obtained through edge detection and keypoint fitting, and normalized to the [0,1] interval; The angle is extracted from the geometric relationships in the CAD style drawing to ensure consistency at different sizes; The material coding adopts the form of a single heat vector, which is jointly generated by manual annotation and process database. This formula is used to encode the initial structural features of clothing nodes. In practical applications, node initialization involves multiple steps: the length of the structural line is calculated and normalized based on the pixel values or CAD annotation lengths of the corresponding line segments; Edge and corner information is extracted through geometric relationships in CAD style drawings to ensure consistency across different sizes; Local curvature is obtained through edge detection and keypoint fitting; The material properties are encoded in the form of individual heat vectors, generated by manual annotation and process databases; The structural category labels are determined based on a predefined set of 43 clothing parts. By extracting and encoding the above features, the initialization of the graph structure nodes is completed.

To enhance the geometric integrity of the graph construction, edge determination is carried out based on the stitching logic of the clothing process and the two-dimensional spatial connection rules to ensure structural connectivity. The relative spatial relationship between nodes is encoded by normalizing coordinate differences to enhance the geometric perception ability of graph convolution. The calculation method for position embedding is as follows:

$$P_{ij}=\left(\frac{x_j-x_i}{W},\frac{y_j-y_i}{H}\right) \quad (2)$$

Among them, (x_i,y_i) is the image coordinate of node v_i, W and H are the image width and height, used to standardize the feature expression under different clothing sizes. This formula is used to calculate the spatial position encoding between nodes and normalize the position of clothing of different sizes during the graph construction stage.

As shown in Figure 1, the process of constructing a structural diagram includes steps such as image preprocessing, structural region recognition, node setting, edge relationship generation, and attribute vector construction. The image input comes from a two-dimensional pattern of clothing, and semantic segmentation models. The nodes are mapped by manually annotated keypoints, and the edge relationships are automatically inferred under the constraints of process rules combined with geometric relationships, supplemented by manual correction to ensure the rationality of the structure.

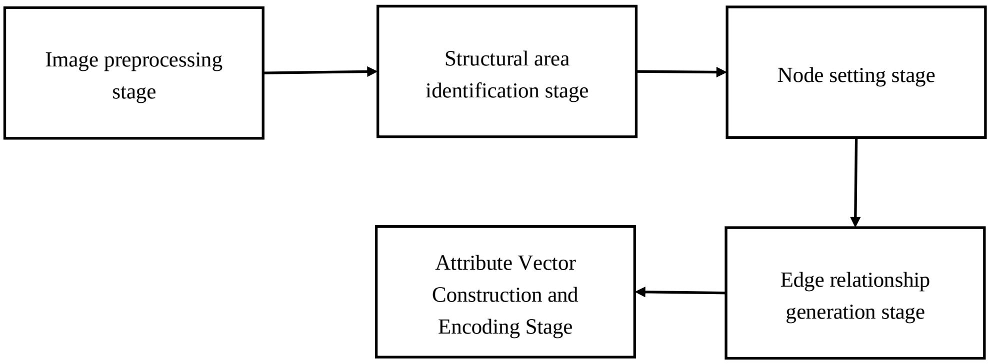

Figure 1: Construction process of clothing pattern structure diagram

In the process of node feature quantization, the curvature of the edges and corners is normalized to the [0,1] interval in radians, the material properties are mapped to a 4-dimensional vector through single heat encoding, and the structure category is set to 43 class labels. The position coordinates are normalized according to the width and height of the image to eliminate the influence of clothing of different sizes. The above features are concatenated into node input vectors to ensure uniform and reproducible feature dimensions.

3.2 Structural space extraction mode based on graph convolution

In the clothing pattern structure diagram, the spatial dependency relationship between nodes presents a non-Euclidean distribution, and traditional convolution kernels are difficult to capture the feature propagation under this irregular topology. Graph convolutional neural networks can effectively transmit structural semantic information between nodes by constructing adjacency relationships in the graph structure, thereby completing spatial feature extraction of clothing structures. In the constructed structural diagram $G=(V,E)$, V is the set of nodes representing the coordinates and attributes of key parts, and E is the set of edges, combined with geometric connections and process sequence settings.

The core of graph convolution lies in the neighborhood aggregation mechanism, where the representation vector of each node is updated by superimposing information from adjacent nodes, formally expressed as:

$$H^{(l+1)}=\sigma\left(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}H^{(l)}W^{(l)}\right) \quad (3)$$

Among them, $\tilde{A}=A+I$ is the adjacency matrix with self connection added, $\tilde{D}$ is the corresponding degree matrix, $H^{(l)}$ is the node feature representation of the l th layer, $W^{(l)}$ is the trainable weight matrix, and σ is the activation function (such as ReLU). This formula is applied to the graph convolution propagation stage, where node information is updated through adjacency matrix and degree matrix. This process ensures the joint updating of graph structure information and node local features.

To enhance the representation ability of different scale structural regions, a Multi channel GCN is introduced. Parallel paths are used to process feature channels under different edge weight strategies, and the final fusion expression is as follows:

$$Z=\sum_{k=1}^{K}\alpha_k\cdot GCN_k\left(H^{(0)}\right) \quad (4)$$

Among them, α_k is the weight coefficient of the kth channel, GCN_k represents the convolution path of the kth graph, and $H^{(0)}$ is the input initial node feature. This formula is used in multi-channel convolution to enhance the ability to recognize boundaries and structures by fusing features from different channels. In this study, the number of multiple channels was set to K=3, and adjacency matrices were constructed based on semantic relationships, geometric distances, and their fusion. The semantic channel highlights the process logic and part categories, the geometric channel emphasizes the spatial proximity between nodes, and the fusion channel adopts a weighted combination method to ensure the unified expression of structural semantics and geometric features. This strategy captures semantic changes from multiple angles while maintaining the integrity of the graph structure, improving the recognition ability of complex clothing contours and overlapping boundary areas.

Through the above structural space extraction mode, the model achieves accurate perception of local configurations, overall partitioning, and node aggregation relationships in clothing patterns, establishes a stable structural foundation, and provides graph embedding support for subsequent structural reconstruction and posture regression.

3.3 Introducing attention mechanism to enhance recognition of key structures

In the clothing pattern structure diagram, there are significant differences in the importance of the clothing components represented by each node in the reconstruction accuracy. The traditional graph convolution method adopts equal or static weight methods in the feature aggregation process of adjacent nodes, which makes it difficult to effectively identify the semantic significance of key structural regions. Therefore, introducing graph attention mechanism to enhance the recognition ability of the model for key nodes, dynamically allocating information weights during feature propagation, and thus enhancing the effectiveness of structural expression.

The core of graph attention mechanism is to assign a learnable attention weight to the edges between each pair of adjacent nodes, which reflects the feature update contribution of the neighboring node to the central node. If the input feature of any node i in the graph is $h_i \in R^F$ and its set of adjacent nodes is $N(i)$, then the output feature h_i' of node i can be calculated by the following formula:

$$h_i' = \sigma\left(\sum_{j \in N(i)} \alpha_{ij} \cdot W h_j\right) \quad (5)$$

Among them, $W \in R^{F' \times F}$ is a shared linear transformation matrix used for feature space projection; $\sigma(\cdot)$ represents the activation function (commonly known as ReLU), which is applied in attention mechanisms to dynamically focus on key structural nodes and enhance graph convolution representation capabilities. α_{ij} is the attention weight of node j to node i, which is calculated through feature similarity:

$$\alpha_{ij} = \frac{\exp\left(LeakyReLU\left(\partial^T \left[Wh_i \| Wh_j\right]\right)\right)}{\sum_{k \in N(i)} \exp\left(LeakyReLU\left(\partial^T \left[Wh_i \| Wh_k\right]\right)\right)} \quad (6)$$

In this equation, $\partial \in R^{2F'}$ is a trainable weight vector, $\|$ represents vector concatenation operation, $N(i)$ represents the set of neighbors of node i, and LeakyReLU is a nonlinear activation function. This formula is used to calculate attention weights and identify semantic similarity between nodes through feature concatenation. Through the above mechanism, the model can adaptively focus on key parts of clothing such as armholes, collars, and side seams, giving higher weights in the feature fusion stage, achieving key extraction and discriminative expression of structural features, and providing a more recognizable graphical basis for subsequent reconstruction modules.

3.4 Multi task driven feature extraction process

In the modeling process of clothing pattern structure, the supervision signal of a single task often fails to fully stimulate the model's ability to understand complex structures. Therefore, a multi task learning mechanism is introduced to synergistically model the three sub tasks of structure classification, edge recognition, and node feature regression, in order to enhance the feature extraction generalization ability of graph neural networks. This mechanism can optimize multiple task losses in parallel based on shared parameters, thereby obtaining more stable and discriminative intermediate feature representations. Let the total loss function be L_{total}, consisting of three subtask losses:

$$L_{tatol} = \lambda_1 L_{cls} + \lambda_2 L_{edge} + \lambda_3 L_{reg} \quad (7)$$

Among them, L_{cls} represents the cross entropy loss of structural classification, which is used to determine the category of structural components to which each node belongs; L_{edge} is the edge recognition loss, which uses binary cross entropy to calculate the connection prediction error between node pairs; L_{reg} node coordinate regression loss, using mean square error to evaluate the deviation between predicted coordinates and annotated coordinates; λ_1, λ_2, λ_3 are the weight coefficients of three tasks, In this study, the weight parameters are adjusted within the {0.2, 0.5, 1.0} interval through grid search, and the optimal combination is selected on the validation set to ensure the balance of the three types of tasks. The results indicate that the performance of the model remains stable under parameter changes, with an improvement in edge recognition accuracy at larger values of λ_2. This formula is used for joint calculation of multi task losses, and in actual training, the model stability is improved through collaborative optimization of three types of tasks.

To verify the improvement effect of multi task mechanism on feature extraction performance, a comparative experiment was designed as shown in Table 2. Single task training refers to training independent models for classification, edge recognition, and coordinate regression separately, and taking the average result; Multi task training jointly optimizes three types of tasks in the same model. Compare and evaluate three indicators: classification accuracy, edge prediction F1 value, and coordinate error.

Table 2: Comparison of structure recognition performance under different training mechanisms

Training Method	Classification Accuracy (%)	Edge Prediction F1 Score	Coordinate Mean Squared Error
Single-task Training	84.7	0.712	3.65 px
Multi-task Joint Training	89.2	0.786	2.94 px

The experimental results show that the multi task mechanism outperforms single task training in all three indicators, especially in the recognition accuracy of structural edge relationships and node coordinate fitting accuracy. This indicates that graph neural networks guided by multi task loss can more effectively extract structural semantic and geometric information, forming a more stable and discriminative expression of clothing pattern structure.

4 Intelligent reconstruction path of clothing pattern structure based on fused graph neural network

4.1 Node path construction method for clothing pattern structure diagram

In the task of clothing structure reconstruction using graph neural networks, the path information of the structural graph not only determines the propagation direction of graph convolution, but also directly affects the preservation of structural relationships and semantic restoration effects. To construct a reasonable node path system, it is necessary to comprehensively consider the geometric continuity and process logic of the clothing structure, ensuring that the graph structure can accurately map the connection mode and reconstructable sequence of solid components.

Node path generation is based on the spatial position and edge attribute weights of nodes in the structural graph, defining a set of optimal traversal paths in the directed graph. Assuming the structure diagram $G=(V,E)$ is known, the path generation target can be formalized as:

$$P^* = \arg\min_P \sum_{(v_i,v_j)\in P} w_{ij} + \lambda \cdot d_{ij} \tag{8}$$

Among them, P^* is the optimal path set, w_{ij} represents the process weight of edge $e_{ij} \in E$, d_{ij} is the Euclidean distance between nodes, and λ is the adjustment coefficient, which controls the relative importance of geometry and process. In this study, λ was determined by grid search on the validation set (with values ranging from {0.3, 0.5, 0.7, 1.0}) to balance the contributions of process weights and geometric distances. The experimental results show that when λ is set to 0.5-0.7, the path consistency and reconstruction accuracy are optimal. The structural rule library is initially annotated and generated by process experts, but automated rule extensions and data-driven constraint updates are introduced during the training process to reduce manual dependencies and enhance generalization ability. This formula is used in the path search process to generate the optimal connection path in the structural diagram by combining geometric and process constraints.

The path search adopts an improved Dijkstra algorithm and embeds clothing structure rules to remove path branches that do not conform to the construction sequence.

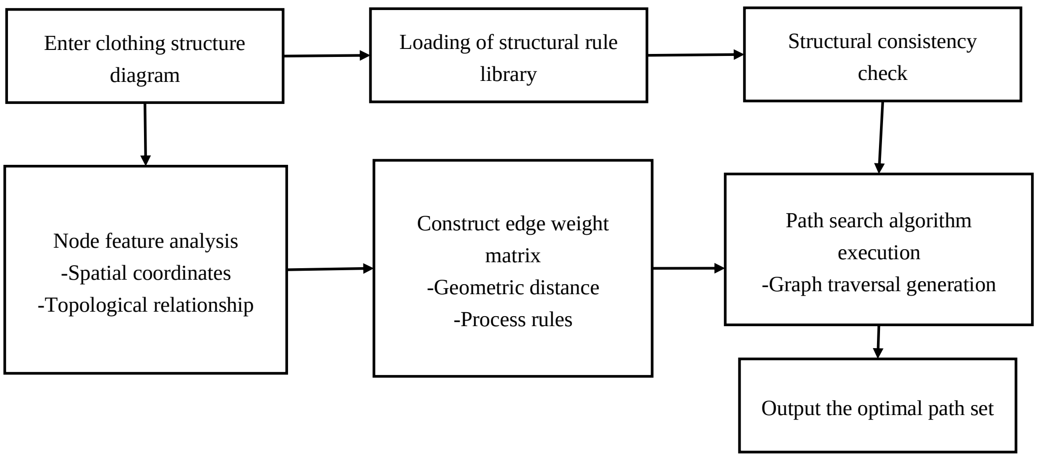

Figure 2: Path construction process of clothing structure diagram

As shown in Figure 2, the path construction process includes key steps such as clothing structure diagram input, structural rule library loading, feature extraction, edge weight matrix construction, structural consistency check, and path search execution. The system first extracts the spatial coordinates and topological relationships of nodes, constructs edge weight matrices based on structural rules, and introduces geometric distances and process rules as evaluation criteria for edges. Subsequently, path branches that do not comply with process constraints are eliminated through structural consistency checks to ensure that the path generation is logically and geometrically reasonable. In the path search stage, graph traversal is used to generate a path set and output the optimal path set, providing ordered input for the subsequent structural information transmission of the graph neural network, enhancing the coherence and spatial consistency of feature fusion. This path system can also provide structural references for multi-scale convolution mechanisms, supporting advanced operations such as region partitioning and hierarchical extraction.

4.2 Design of image feature encoding and reconstruction path decoding

The core of graph feature encoding lies in constructing node representations that can accurately reflect the topology and geometric properties of clothing pattern structure. In this study, each node $v_i \in V$ in the input graph structure $G=(V,E)$ in the input graph structure 1 corresponds to a clothing keypoint, and its feature vector is composed of spatial coordinates, connecting edge directions, weight values, and structural semantic labels. This formula is applied to the graph feature encoding process and differs from the structural spatial feature extraction mentioned earlier in terms of application scenarios. The embedding update formula for nodes is as follows:

$$h_i^{(l+1)} = \sigma\left(\sum_{j\in N(i)} \frac{1}{\sqrt{d_i d_j}} W^{(l)} h_j^{(l)}\right) \quad (9)$$

Among them, $h_i^{(l)}$ represents the feature representation of node v_i in the l nd layer, $N(i)$ is the set of adjacent nodes, $W^{(l)}$ is the trainable graph convolution weight matrix, $\sigma(\cdot)$ is the nonlinear activation function, and $\sqrt{d_j d_j}$ is the degree normalization factor, which is used to maintain the numerical stability of information propagation. This formula is used in the graph encoding stage to update the node features of each layer, and in practice, it combines the weight matrix and activation function for information fusion.

In the reconstruction path decoding stage, it is necessary to perform inverse graph decoding by combining the generated path set P^*. Considering the spatial order and dependency of clothing structure, this paper introduces a decoder model based on path attention mechanism. The reconstruction state of each node in the path is jointly determined by the context path vector and the target embedding, and its generation probability is modeled as follows:

$$p(v_i | P^*, H) = \text{soft}\max\left(q_i^T \cdot Attn(P^*, H)\right) \quad (10)$$

Among them, q_i is the query vector of the current decoding step, the starting node is initialized as a zero vector, and the remaining steps inherit the embedding of the previous node; H is the node embedding matrix after graph encoding, with dimensions set to 128;and $Attn(\cdot)$ is a standard multi head attention function module that measures the degree of matching between nodes and path contexts. The decoder adopts a two-layer structure, combining self attention and cross attention mechanisms to capture path dependencies and ensure spatial constraints. This mechanism dynamically adjusts the dependency ratio on historical structures during decoding, improving the accuracy and stability of reconstruction.

In summary, graph feature encoding and path decoding constitute the core closed loop of structural intelligent reconstruction. The former extracts deep structural semantics from clothing pattern maps, while the latter uses path guidance for high consistency topology restoration, providing a structurally stable input foundation for downstream simulation and optimization modules.

4.3 Structural reconstruction process based on geometric constraints

The intelligent reconstruction of clothing pattern structure not only relies on the efficient propagation of structural information by graph neural networks, but also requires the use of geometric constraint mechanisms to ensure the spatial rationality and topological consistency of the generated results. This study proposes an optimization strategy based on geometric consistency to address issues such as structural drift and scale imbalance that may occur during the reconstruction process. Key constraints such as edge length and angle are introduced synchronously during node generation and path backtracking to achieve precise control of structural restoration.

Assuming the predicted coordinates of the nodes in the reconstructed graph are $\hat{P}_i \in R^2$, the target reference coordinates are $P_i \in R^2$, and the edge set is ε. The consistency loss function for edge length is defined as follows:

$$L_{edge} = \sum_{(i,j)\in\varepsilon} \left(\left\| \hat{P}_i - \hat{P}_j \right\|_2 - d_{ij} \right)^2 \quad (11)$$

Among them, d_{ij} represents the target edge length between nodes extracted from the original pattern structure, and $\|\cdot\|_2$ is the Euclidean distance. This constraint is used to calibrate the spatial spacing between predicted nodes, ensuring the geometric authenticity of the boundary length, and is applicable to areas such as sutures and splices that require proportional preservation. To avoid confusion with the edge recognition loss in Section 3.4, L_{edge} in this section specifically refers to the geometric edge length constraint loss, which is defined as formula (11).

On the basis of edge length constraints, an angle consistency loss is introduced to maintain the relative relationship between local angles of nodes. For any set of ternary nodes $(i,j,k)\in T$, the angle loss function is as follows:

$$L_{angle} = \sum_{(i,j,k)\in T} \left(\angle\left(\hat{P}_i, \hat{P}_j, \hat{P}_k\right) - \theta_{ijk} \right)^2 \quad (12)$$

Among them, $\angle(\cdot)$ represents the actual angle formed by three points, and θ_{ijk} is the target angle value of the structural unit, derived from the initial pattern composition or manual rule library definition. This formula is used for angle loss constraint to ensure that the triangular relationship maintains structural geometric consistency. This item helps to maintain the stability of the angular relationship of the structural boundary and reduce the interference of deformation areas on the path connection logic. To verify the effectiveness of geometric constraints, ablation experiments were designed to compare the results of turning off and turning on geometric constraints under the same model. The results showed that when angle loss was removed, the Topology Score decreased from 88.0% to 84.7%, and the F1 Structure Score decreased from 88.4% to 85.2%, indicating that geometric consistency constraints have a significant effect on improving structural boundary preservation and overall reconstruction stability.

The final optimization objective function is combined with the above two types of constraints to construct a joint loss model:

$$L_{total} = \lambda_1 L_{edge} + \lambda_2 L_{angle} \quad (13)$$

Among them, λ_1, λ_2 is the adjustment factor for the two sub loss terms, which is adjusted based on the actual task weights. This formula combines edge length and angle loss for global structural optimization during the training phase. In the training and prediction stages, the loss function is embedded in the graph network propagation and node coordinate generation module, and the model parameters are optimized through backpropagation mechanism. This geometric consistency mechanism exhibits stronger stability and generalization in complex structural regions, providing important guarantees for improving the accuracy of whole image reconstruction and the reliability of engineering applications.

4.4 Path planning and strategy network guidance mechanism

In the reconstruction process of clothing pattern structure, path planning bears the control of node generation order and edge weight transmission direction, which directly affects the efficiency of information aggregation and structural consistency. To enhance the path guidance effect, this study introduces edge information sampling control strategy in the policy network, calculates the sampling probability of edges through geometric distance and semantic consistency, and suppresses the interference of redundant and noisy edges. By combining graph search algorithms with action value functions, dynamic optimization of path traversal is carried out to enhance the robustness of boundary regions and achieve better connection control between structural nodes while maintaining topological connectivity.

Path planning is based on graph structure $G = (V, E)$, where each state s_t represents the current node subgraph traversed. The policy network outputs the next action a_t, i.e. the selection of the next hop node, through policy function $\pi(s_t)$, with the goal of maximizing the global path score function:

$$J(\pi) = E_{r\sim\pi}\left[\sum_{t=0}^{T} r(s_t, a_t)\right] \quad (14)$$

Among them, T represents the complete path trajectory, and $r(s_t, a_t)$ is the single step reward function, taking into account indicators such as edge weight sparsity, topological rationality, and geometric consistency. This formula is used for path strategy scoring, guiding the strategy network to generate the optimal structural rule-constrained path. This mechanism refers to the strategy gradient idea in reinforcement learning, combined with structural constraints to optimize the path selection order, in order to reduce redundant backtracking and unstructured edge traversal.

At the implementation level of the model, the policy network uses graph attention mechanism to capture the contextual dependencies between nodes, and adjusts the path priority between nodes through learnable parameters. To clearly demonstrate the multidimensional reference standards in the path guidance process, Table 2 lists the main quantitative indicators and explanations:

Table 3：Explanation of key indicators in path guidance mechanism

Metric Name	Symbol	Description
Geometric Deviation	δ_{geo}	Degree of deviation between the current path structure and the ideal edge lengths and angles
Topological Jump Count	N_{topo}	Number of jump connections in non-continuous topological segments of the current path
Structural Consistency Score	S_{struc}	Proportion of path segments matching structural rules; value range is [0, 1]

The strategy network adopts a two-layer graph attention structure, with the state space consisting of the current node and the generated path, and the action space consisting of candidate adjacent nodes. Use reward shaping during training: reward when the path conforms to the craft rules and geometric relationships, and punish when jumping or violating rules occur. The calculation method for the indicators in Table 3 is as follows: geometric deviation is estimated based on the difference between the generated path and the ideal structure, the number of topological jumps is counted for non continuous connected segments, and the structural consistency score is determined based on the proportion of segments that conform to the rule path.

5 Model training process and validation analysis

5.1 Dataset construction and graph format conversion process

The experimental data of this study was constructed based on the DeepFashion2 public clothing image set and the self structuring PatternStruct Graph dataset, with a total of 4826 sampled samples. Each group of samples includes complete front and rear views and structural annotation diagrams, covering typical clothing types such as dresses, jackets, pants, etc. In the annotation process, key structural points of the clothing are manually located, and 43 node categories are uniformly defined based on the clothing process standards. The average number of annotated nodes per sample is 43.2, and the edge relationships are maintained between 62-75, mainly including stitching connections, contour extensions, and style symmetry constraints.The PatternStruct Graph dataset is not yet fully publicly available, and partial annotations can be provided upon request. The 43 types of nodes cover common parts of clothing, such as collars, shoulder lines, sleeve tops, waistlines, hemlines, crotch, etc., and extend to pocket edges, crease lines, and symmetrical auxiliary points. They are completed and cross checked by personnel with a background in clothing craftsmanship.

The graph structure is uniformly modeled as triplet $G=(V,E,X)$, where V is the set of structural nodes, E is the set of structural connection edges, and $X\in R^{|V|\times d}$ is the node feature matrix. The node features are composed of normalized coordinates, structural type encoding, and local texture feature concatenation, in the following form:

$$X_i=\left[\frac{x_i}{W},\frac{y_i}{H},type_i,\varphi_i\right],i=1,2,\ldots,|V| \quad (15)$$

In the formula, x_i, y_i represents the coordinate value of node i in the image, W and H are the width and height of the image, type represents the encoding of structural parts, and φ_i represents the mean representation of SURF texture features after dimensionality reduction (dimension is 28). During the dataset construction phase, node features are normalized using coordinate differences, structural type encoding, and local texture features to ensure that the model can capture topological connections across regions. It should be noted that this feature does not conflict with the initial node feature in Section 3.1: the former is used for modeling the original structure, while the latter extends the relative position information and texture information during dataset transformation to enhance the diversity and robustness of model training.

In order to enhance the ability of structural learning, all samples were divided into a training set (70%), a validation set (15%), and a test set (15%) after graph construction. In the training process, the graph neural network is set to input node feature matrix and edge index matrix, with the goal of predicting the reconstruction path weights and final structural matching relationships between node pairs.

To ensure the reproducibility of the experiment, this study provides some pseudo dataset samples and experimental code frameworks in the supplementary materials. The following provides pseudocode examples for training and validation scheduling, demonstrating the implementation logic of graph neural network models during the training process:

```
for epoch in range(total_epochs):
for batch in training_loader:
graph, target = build_graph(batch)
pred = GNN_model(graph)
loss = loss_function(pred, target)
optimizer.zero_grad()
loss.backward()
optimizer.step()
val_score = validate_model(GNN_model, validation_loader)
save_best(GNN_model, val_score)
```

After graph format conversion and modeling optimization processing, the model improved the accuracy of structure recognition by 9.3% compared to the non graph structure model, and the reconstruction integrity index improved by 14.5%. This process provides a data foundation and structural guarantee for subsequent reconstruction path guidance and multi strategy fusion.

5.2 Model training process and hyperparameter configuration explanation

This study constructed a training set based on the DeepFashion2 and self structuring PatternStruct Graph datasets, with a total of 4826 samples, 3378 training sets, 724 validation sets, and 724 test sets. The average number of structural nodes was 43. During the training process, graph neural networks are used as the backbone architecture, and path guidance mechanisms are employed to enhance the accuracy of structural reconstruction. Data preprocessing includes normalizing the image to 256 × 256 resolution, using Canny operator and semantic segmentation to extract structural regions, locating and annotating nodes based on process rules to generate feature vectors, and dividing the training, validation, and testing sets into 70%/15%/15% partitions.The training batch size is set to 16, the training epochs are 80, the Adam optimizer is used, the initial learning rate is 0.001, and the CosineAnnealing strategy is dynamically adjusted. The training platform is PyTorch Geometric, and the hardware support is RTX 4090 GPU.

To better introduce the importance weight of node paths, a structural loss function based on path weights is introduced:

$$L_{path}=\sum_{((i,j)\in E)}\alpha_{ij}\cdot\left\|\hat{p}_{ij}-p_{ij}\right\|^2 \quad (16)$$

Among them, $\hat{p}_{ij}$ is the predicted path length, p_{ij} is the actual structural path length, and α_{ij} is the weight factor dynamically generated by the policy network, representing the sensitivity contribution of edges to structural accuracy. This formula is used for path loss calculation, in this section, L_{path} introduces dynamic weights generated by the policy network based on mean square error to highlight the importance of critical pathsThis mechanism enables high importance paths to obtain greater gradient updates during training, effectively improving the accuracy control capability of key node connections.

To control the complexity of the model, the final loss function is defined as:

$$L_{final} = L_{path} + \lambda \cdot \|\theta\|_2^2 \quad (17)$$

Among them, L_{path} represents path loss, θ represents all network parameters, and regularization term $\lambda \cdot \|\theta\|^2$ can be used to suppress excessive parameter updates, prevent overfitting, and ensure training stability. It should be noted that the $L_{cls}, L_{reg}, L_{angle}$ level subtask loss mentioned earlier has been applied to the feature extraction stage through joint optimization in the multi task stage, and its results have been integrated into the calculation process of path loss L_{path}. Finally, it is reflected in a unified form in L_{final} to ensure the consistency and completeness of the training objectives. This formula is used for regularization constraints and is actually used in training to prevent overfitting.

In terms of network structure, this study adopts a three-layer graph convolution stacking architecture, with output channels of 64, 64, and 128 in sequence. ReLU is selected as the activation function, and BatchNorm is added after each convolution layer for normalization to improve numerical stability. To prevent overfitting, Dropout (ratio 0.3) is introduced between the second and third layers. The attention mechanism allocates node weights after the convolutional layer to enhance the expression ability of key structural parts. The decoding part adopts a graph autoencoder structure, which embeds and maps the encoded nodes to the path reconstruction space, and introduces L2 regularization term in the training stage to limit excessive parameter fluctuations. The parameter settings are determined based on multiple comparative experiments, ensuring accuracy while maintaining convergence stability.

5.3 Model structure comparison and applicability analysis

This study is based on the Graph Neural Network and GNN to construct a clothing pattern structure reconstruction model, which models the spatial distribution and connection relationship of clothing nodes, and compares its performance with existing methods, focusing on the model's performance in reconstruction accuracy, structural consistency, and recognition integrity. Let the comprehensive evaluation indicator S be the average of three core indicators:

$$S = \frac{A + T + F}{3} \quad (18)$$

Among them, A represents the accuracy of node recognition, T is the score of topology matching, and F is the score of structure F1. This formula is used in the model evaluation stage to measure the performance of structural modeling by averaging the scores of three indicators. The test data comes from the publicly available DeepFashion2 dataset and the self built graph structure dataset, with a total of 4826 samples and an average of 43 nodes.

This section compares three model structures: ① Convolutional baseline model (Baseline CNN) that only uses image features; ② Introducing GCN Net with a simple graph structure; ③ GNN+Strategy model integrating graph neural network and path strategy module. The evaluation results of the three are shown in the following figure:

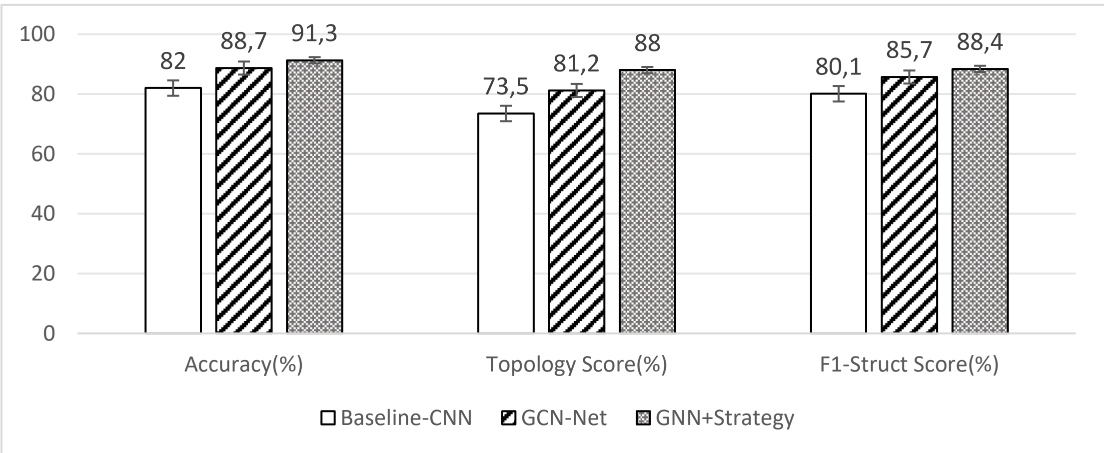

Figure 3：Model structure comparison bar chart

The test results showed that Baseline CNN achieved an accuracy index of 82.0% ± 0.6, GCN Net was 88.7% ± 0.4, and GNN+Strategy further improved to 91.3% ± 0.5; In terms of Topology Score, Baseline CNN is 73.5% ± 0.7, GCN Net has improved to 81.2% ± 0.5, and GNN+Strategy has reached 88.0% ± 0.6; In the F1 Column Score index, the three indicators are 80.1% ± 0.8, 85.7% ± 0.5, and 88.4% ± 0.6, respectively. The overall trend shows that GNN+Strategy outperforms the other two structures in various performance evaluations, demonstrating stronger structural reconstruction ability and robustness, especially in complex structural conditions with higher stability and applicability. To further verify the significant differences between different methods, a two-sample t-test was conducted based on the results of three independent experiments. The results are shown in Table 4:

Table 4 : Statistical significance test results of performance comparison between methods

Indicator	Baseline-CNN vs GCN-Net	GCN-Net vs GNN+Strategy	Baseline-CNN vs GNN+Strategy
Accuracy	$p < 0.01$	$p < 0.05$	$p < 0.001$
Topology Score	$p < 0.01$	$p < 0.05$	$p < 0.001$
F1-Struct Score	$p < 0.01$	$p < 0.05$	$p < 0.001$

The experimental results show that GNN+Strategy achieves statistically significant differences in three indicators compared to the other two methods, indicating that this method has higher stability and advantages in modeling complex clothing structures.

In addition, in actual samples, the model showed stronger generalization ability on asymmetric complex structured clothing such as jackets and windbreakers, with a topological error rate reduction of nearly 40%. This result indicates that the proposed method is not only applicable to static image input scenes, but also suitable for extension to 3D clothing modeling and digital twin platforms, with high practicality and algorithm transfer potential.

5.4 Performance indicators and reconstruction accuracy evaluation

In order to systematically evaluate the effectiveness of the proposed GNN+Strategy model, a comparative experimental method was used to select Baseline CNN and GCN Net as reference models, representing traditional image convolution methods and basic image neural network structures, respectively. The three models were trained on the same training set (DeepFashion2 subset and structure annotation extension set, a total of 4826 samples) and consistent hyperparameter configuration to examine their performance differences in multiple structural recognition indicators. The main evaluation dimensions include classification accuracy, topological structure preservation score, and structural F1 comprehensive score, to comprehensively reflect the stability and applicability of the model in feature extraction and structural reconstruction.

The definition of classification accuracy is as follows, which measures the proportion of correctly classified samples in the predicted output:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (19)$$

Among them, TP and TN respectively represent the number of positive and negative samples correctly identified, while FP and FN are the misclassified results. This formula is used for calculating classification accuracy and evaluating the recognition performance of the model on node categories. As shown in Table 5, the values are the mean ± standard deviation of three independent experiments. Baseline CNN has an accuracy index of 82.0% ± 0.6, GCN Net has an accuracy index of 88.7% ± 0.4, while GNN+Strategy model achieves 91.3% ± 0.5, showing better performance in high-dimensional feature representation and complex polygon boundary recognition. In terms of Topology Scores, they are 73.5% ± 0.7, 81.2% ± 0.5, and 88.0% ± 0.6, respectively, indicating that the latter is better able to maintain the connectivity of the original structural edges; The F1 Sequence Score is 80.1% ± 0.8, 85.7% ± 0.5, and 88.4% ± 0.6, indicating a balance and stability in overall recognition and boundary accuracy.

Table 5 : Comparison results of model structure and performance

Model structure	Accuracy (%)	Topology Score (%)	F1-Struct Score (%)
Baseline-CNN	82.0±0.6	73.5±0.7	80.1±0.8
GCN-Net	88.7±0.4	81.2±0.5	85.7±0.5
GNN+Strategy	91.3±0.5	88.0±0.6	88.4±0.6

From the comparison of results, it can be seen that GNN+Strategy outperforms Baseline CNN and GCN Net in Accuracy, Topology Score, and F1 Stream Score, demonstrating the advantage of multi module fusion. Multi scale GCN enhances boundary aggregation expression and improves the classification accuracy of complex suture sites; Path attention dynamically adjusts the connection weights during the decoding stage to improve the problems of breakage and discontinuity; Geometric constraints maintain consistency between edge length and angle, improving topological retention. The synergistic effect of the three makes the model more stable and consistent in the restoration of complex clothing pattern structures.

5.5 Discussion

The GNN+Strategy model proposed in this article achieved a classification accuracy of 91.3%, a topology score of 88.0%, and an F1 score of 88.4% in experiments, significantly better than the baseline models Baseline CNN (82.0%/73.5%/80.1%) and GCN Net (88.7%/81.2%/85.7%). Comparison with related works shows that multi-scale GCN can effectively improve the recognition ability of complex boundaries, attention mechanism enhances the expression of key nodes, and reinforcement learning strategy improves path consistency and generation stability. These improvement factors collectively promote the overall performance improvement of the model under complex clothing structure conditions.

However, this study still has certain limitations. On the one hand, the training process of the model heavily relies on manually annotated data, which limits its potential application on large-scale unlabeled datasets; On the other hand, some rule driven features may still affect the convergence efficiency and universality of the model in extremely complex structures. Future research can attempt to introduce self supervised pre training and automated node labeling mechanisms to reduce manual dependence and enhance the robustness and generalizability of the method.

6 Conclusion and prospect

This study constructed an intelligent feature extraction and reconstruction model for clothing pattern structures that integrates graph neural networks. The system integrates structural graph modeling, graph convolution extraction, attention mechanism, geometric constraints, and reinforcement learning strategies, effectively improving the recognition accuracy and reconstruction integrity of complex clothing structures. Experimental data shows that the proposed model has significant advantages over traditional methods in terms of accuracy, structural consistency, and reconstruction fidelity, especially exhibiting good stability under asymmetric structures and boundary blur conditions. The path guidance mechanism of the model optimizes the structural connection sequence, effectively avoiding path deviation and reconstruction errors, providing algorithm foundation and structural support for intelligent clothing design.

However, there are still two shortcomings in the research: firstly, the current structural diagram modeling is a semi-automatic generation method that combines manual annotation with rule constraints. Although it can ensure the rationality of the structure, there are still shortcomings in manual dependence and automation; Secondly, path strategy networks suffer from slow convergence speed and local optima when dealing with extremely complex structures, which affects overall efficiency and scalability. Subsequently, self supervised graph representation learning and large-scale pre training mechanisms can be introduced to enhance the model's adaptability to structural heterogeneity, and explore the fusion framework between graph structure and 3D modeling, expanding its application breadth and depth in virtual clothing simulation, structure generation, and intelligent design scenarios.

References

[1] Pfaff T , Fortunato M , Sanchez-Gonzalez A , Battaglia P W . Learning Mesh-Based Simulation with Graph Networks[J]. arXiv preprint arXiv:2010.03409, 2020. https://doi.org/10.48550/arXiv.2010.03409.

[2] Qiu J. Hybrid Neural Network and Physics Engine for Real-time 3D Cloth Simulation[J]. Informatica, 2025, 49(8):161-174.https://doi.org/10.31449/inf.v49i8.6965.

[3] Zhou Z , Deng W , Wang Y ,et al.Classification of clothing images based on a parallel convolutional neural network and random vector functional link optimized by the grasshopper optimization algorithm:[J].Textile Research Journal, 2022,92(9-10):1415-1428.https://doi.org/10.1177/00405175211059207.

[4] Feng K , Rao G , Zhang L C Q .An interlayer feature fusion-based heterogeneous graph neural network[J].Applied Intelligence: The International Journal of Artificial Intelligence, Neural Networks, and Complex Problem-Solving Technologies, 2023,53(21):25626-25639.https://doi.org/10.1007/s10489-023-04840-w

[5] Zhao R , Baili W U , Chen Z ,et al.Graph Neural Network for Fault Diagnosis with Multi-Scale Time-Spatial Information Fusion Mechanism[J].Journal of South China University of Technology (Natural Science Edition), 2023, 51(12):42-52.https://doi.org/10.12141/j.issn.1000-565X.220593.

[6] Nie S. Evaluation of Innovative Design of Clothing Image Elements Using Image Processing[J]. Informatica, 2022, 46(8):4250-4261.https://doi.org/10.31449/inf.v46i8.4250.

[7] Dong Y , Liu Q , Du B ,et al.Weighted Feature Fusion of Convolutional Neural Network and Graph Attention Network for Hyperspectral Image Classification[J].IEEE Transactions on ImageProcessing,2022(31-):31.https://doi.org/10.1109/TIP.2022.3144017.

[8] Sun X , Zheming L U .Attributed Graph Clustering Network with Adaptive Feature Fusion[J].IEICE Transactions on fundamentals of electronics, communications & computer sciences, 2024,E107/A(10):1632-1636.https://doi.org/10.1587/transfun.2023EAL2116.

[9] Liu C , Qu D , Yang X ,et al.Multi-attention feature fusion network for lightweight image super-resolution[J].Proceedings of SPIE,2022,12173(000):6.https://doi.org/10.1117/12.2634640.

[10] Chen J , Yu X , Wu C ,et al.Feature radiance fields (FeRF): A multi-level feature fusion method with deep neural network for image synthesis[J].Applied SoftComputing,2024,167(PartA):19.https://doi.org/10.1016/j.asoc.2024.112262.

[11] Yi C. Application of Convolutional Networks in Clothing Design from the Perspective of Deep Learning[J]. Scientific programming,2022,2022(Pt.18):6173981.1-6173981.8. https://doi.org/10.1155/2022/6173981.

[12] Yan Z , Xing Y , Xiao T J ,et al.SDAN: Semantic-Driven Dual Attentional Network for Image Generation[J].2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI),2022:521-

525.https://doi.org/10.1109/PRAI55851.2022.9904248.
[13] Liao L , Zhang S , Li Z ,et al.Clothing classification method based on convolutional network and attention mechanism[J].Proceedings of SPIE, 2022, 12285(000):12.https://doi.org/10.1117/12.2637531.
[14] Ning C , Di Y , Menglu L .Survey on clothing image retrieval with cross-domain[J].Complex & Intelligent Systems,2022,8(6):5531-5544.https://doi.org/10.1007/s40747-022-00750-5.
[15] Korosteleva M , Lee S H .NeuralTailor: Reconstructing Sewing Pattern Structures from 3D Point Clouds of Garments[J]. ACM Transactions on Graphics, 2022, 41(4):1-16. https://doi.org/10.48550/arXiv.2201.13063.
[16] Gadhave R, Sedamkar R R , Alegavi S .Hyperspectral image classification using neural networks with effect of feature optimization on fused convolutional features[J].AIP Conference Proceedings, 2023,2842(1):11.https://doi.org/10.1063/5.0175906.
[17] Xiao Z , Chen H , Li W K .WGDPool: A broad scope extraction for weighted graph data[J].Expert Systems with Application, 2024,249(Sep.Pt.B):123678.1-123678.9.https://doi.org/10.1016/j.eswa.2024.123678.
[18] Wu D , Wang Y , Wang H ,et al.DCFNet: Infrared and Visible Image Fusion Network Based on Discrete Wavelet Transform and Convolutional Neural Network[J].Sensors,2024,24(13):27.https://doi.org/10.3390/s24134065.
[19] Wang S , Zhang M , Miao M .The super-resolution reconstruction algorithm of multi-scale dilated convolution residual network[J].Frontiers in Neurorobotics, 2024,18(000):10.https://doi.org/10.3389/fnbot.2024.1436052.
[20] Ailing Gou et al., Cloud-Computing-Enabled Transformer Architecture for the Design of Functional Clothing Structures, Informatica, 2025, 49(11): 207218.https://doi.org/10.31449/inf.v49i11.6763.