

Density-Aware Graph Generation with Learnable Edge Prediction

Seyedeh Ava Razi Razavi[†], James Sargant[†], Sheridan Houghten[†], Renata Dividino^{†, *}

[†] Brock University, St Catharines, Ontario, Canada.

Abstract

Generating realistic graph-structured data is challenging due to discrete structures, variable sizes, and class-specific connectivity patterns that resist conventional generative modeling. While recent graph generation methods employ generative adversarial network (GAN) frameworks to handle permutation invariance and irregular topologies, they typically rely on random edge sampling with fixed probabilities, limiting their capacity to capture complex structural dependencies between nodes. We propose a density-aware conditional graph generation framework using Wasserstein GANs (WGAN) that replaces random sampling with a learnable distance-based edge predictor. Our approach embeds nodes into a latent space where proximity correlates with edge likelihood, which enables the generator to learn meaningful connectivity patterns. A differentiable edge predictor determines pairwise relationships directly from node embeddings, while a density-aware selection mechanism adaptively controls edge density to match class-specific sparsity distributions observed in real graphs. We train the model using a WGAN with gradient penalty, employing a GCN-based critic to ensure generated graphs exhibit realistic topology and align with target class distributions. Experiments on benchmark datasets demonstrate that our method produces graphs with superior structural coherence and class-consistent connectivity compared to existing baselines. The learned edge predictor captures complex relational patterns beyond simple heuristics and generates graphs whose density and topology closely match real structural distributions. Our results show improved training stability and controllable synthesis, making the framework effective for realistic graph generation and data augmentation. The source code is publicly available at https://github.com/ava-12/Density_Aware_WGAN.git.

Keywords: Graph Neural Networks, Generative Adversarial Networks, Conditional Generation, Wasserstein Generative Adversarial Networks, Edge Prediction

1. Introduction

Graph-structured data appears in numerous domains, from social networks and molecular structures to biological systems and communication networks, where entities and their relationships encode critical information about underlying processes and properties. The ability to generate realistic synthetic graphs has become increasingly important for applications including privacy-preserving data sharing, and data augmentation for graph-based machine learning models. However, graph generation presents fundamental challenges that distinguish it from traditional data synthesis in Euclidean spaces. Unlike images or text that have regular grid structures and fixed dimensionality, graphs exhibit irregular topologies, variable sizes, and lack canonical node ordering. Successful graph generators must capture complex dependencies between topology and node attributes while maintaining the structural properties that characterize different graph types.

Early deep learning approaches to graph generation have explored various strategies to address these challenges. Methods such as DeepGMG [1] and GraphRNN [2] model graphs through sequential generation processes, while GraphVAE [3] leverages variational autoencoders to learn latent graph representations. Generative Adversarial Networks (GANs) have shown particular promise, with GraphGAN [4], EGraphGAN [5], and MolGAN [6]

*rdividino@brocku.ca

demonstrating the potential of adversarial training for graph synthesis. LGGAN [7] further advanced the field by introducing labeled graph generation with comprehensive evaluation metrics including Maximum Mean Discrepancy (MMD) over degree, clustering, and orbit statistics. However, these GAN-based approaches often suffer from training instability, mode collapse, and difficulties in generating graphs with variable sizes.

The introduction of Wasserstein GANs (WGANs) [8] marked a significant advancement in addressing the instability issues inherent in standard GANs. WGAN-GP [9] further improved upon this by enforcing the Lipschitz constraint through gradient penalty rather than weight clipping, enabling more effective training and better sample diversity. Recent work has successfully adapted the WGAN framework to graph generation by combining it with Graph Neural Networks (GNNs) as critics. In [10], a generator produces both node features and connectivity patterns, while a GNN-based critic evaluates graph realism and class consistency. This method demonstrated superior performance in generating class-conditional graphs with improved training stability. However, despite these advances, the approach relies on simplistic edge generation using fixed, class-dependent probabilities and random sampling. This limitation results in graphs with unrealistic connectivity patterns that fail to capture the complex structural dependencies between nodes and do not reflect the learned node feature distributions.

We address these core limitations by developing a new WGAN-based framework featuring a graph generator that employs a learnable, distance-driven edge predictor, enabling the model to infer connectivity patterns and capture complex structural dependencies. We further introduce a density-aware edge selection mechanism that computes probabilities for all possible node pairs and selects edges based on class-specific density statistics. This ensures that generated graphs maintain realistic sparsity patterns. Additionally, our architecture employs node-specific latent vectors rather than shared graph-level noise, promoting diversity in node features within each graph. Experimental results show improved graph generation that matches the statistical properties of the training distribution. The generated graphs exhibit coherent structural patterns, where connectivity meaningfully reflects learned relationships between node features.

2. Foundations

We provide the formal definitions and terminology used in this work. Let a graph be represented as $G = (V, E, X, A, y)$ with a set of nodes $V = \{v_1, \dots, v_N\}$ and edges $E \subseteq V \times V$. $X \in \mathbb{R}^{N \times d}$ is the node feature matrix where $X_i \in \mathbb{R}^d$ corresponds to the d -dimensional feature vector of node $v_i \in V$. The adjacency matrix $A \in \{0, 1\}^{N \times N}$ encodes the graph structure, where $A_{ij} = 1$ if edge $(v_i, v_j) \in E$ and 0 otherwise. For graph-level tasks, $y \in \{1, \dots, C\}$ denotes the graph class label from C distinct classes.

2.1. Graph Neural Networks

Graph Neural Networks (GNNs) have enabled the training of deep learning models on graph-structured data, aiming to represent graphs as low-dimensional embeddings that can be leveraged for downstream tasks such as node classification, link prediction, and graph classification. A number of architectures have been proposed [11–14], following the message-passing scheme. GNNs learn a multi-step mapping: $f : G \rightarrow Z \in \mathbb{R}^{N \times d'}$, where Z is the updated feature matrix for a graph G , with d' being the output embedding dimension, and each row of Z represents a low-dimensional embedding of a node in G .

Each GNN layer typically consists of two phases: *Aggregation*, where information from the neighborhood of each node is gathered using a permutation-invariant function, and *Update*, where each node’s feature vector is updated based on its current state and the aggregated

neighborhood information. These steps are applied recursively over k layers in the GNN. The hidden representation $\mathbf{h}_v^{(l)}$ of a node v at layer l is computed as:

$$\mathbf{h}_v^{(l)} = \sigma \left(\mathbf{B}_l \mathbf{h}_v^{(l-1)} + \mathbf{W}_l \sum_{u \in \mathcal{N}_g(v)} g(\mathbf{h}_u^{(l-1)}) \right), \quad (2.1)$$

where $\mathbf{h}_v^{(0)} = X_v$ is the initial feature vector of node v , $\mathcal{N}_g(v)$ is the set of neighbors of v , \mathbf{B}_l and \mathbf{W}_l are learnable weight matrices at layer l , g is a differentiable, permutation-invariant aggregation function (e.g., mean, sum, or max), and σ is a nonlinear activation function.

The features of node v and its neighbors are combined by an update function, such as multi-layer perceptron (MLP), to produce new node embeddings. The output of the final layer, denoted as $Z^{(k)}$, can then be used for various downstream tasks.

2.2. Generative Adversarial Networks

Generative Adversarial Networks (GANs) constitute a powerful framework for learning complex data distributions through an adversarial training process [15]. The core principle involves two neural networks competing in a minimax game: a generator network G that learns to produce synthetic data samples, and a discriminator network D that learns to distinguish between real and generated samples. The generator maps random noise vectors $z \sim p_z(z)$ from a prior distribution to the data space, while the discriminator outputs a probability indicating whether its input is real or synthetic.

The training objective is formulated as a minimax optimization problem:

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim \mathbb{P}_r} [\log D(x)] + \mathbb{E}_{z \sim p_z} [\log(1 - D(G(z)))]. \quad (2.2)$$

During training, the discriminator maximizes its ability to distinguish real and fake samples, while the generator minimizes this ability.

2.2.1. Wasserstein GANs

Standard GANs often suffer from training instability, mode collapse, and vanishing gradients. Wasserstein GANs (WGANs) [8] address these issues by replacing the Jensen–Shannon divergence with the Wasserstein-1 (Earth-Mover) distance, which provides smoother gradients and a more meaningful loss metric. The Wasserstein distance between the real and generated distributions \mathbb{P}_r and \mathbb{P}_g is defined as:

$$W(\mathbb{P}_r, \mathbb{P}_g) = \inf_{\gamma \in \Pi(\mathbb{P}_r, \mathbb{P}_g)} \mathbb{E}_{(x, x') \sim \gamma} [\|x - x'\|], \quad (2.3)$$

where $\Pi(\mathbb{P}_r, \mathbb{P}_g)$ denotes the set of all couplings with marginals \mathbb{P}_r and \mathbb{P}_g . Using the Kantorovich–Rubinstein dual formulation, the WGAN objective becomes:

$$\min_G \max_{D \in \mathcal{D}} \mathbb{E}_{x \sim \mathbb{P}_r} [D(x)] - \mathbb{E}_{z \sim p_z} [D(G(z))], \quad (2.4)$$

where \mathcal{D} is the set of 1-Lipschitz critics. To enforce this constraint, WGAN-GP [9] introduces a gradient penalty:

$$\mathcal{L}_{\text{GP}} = \lambda \mathbb{E}_{\tilde{x}} [(\|\nabla_{\tilde{x}} D(\tilde{x})\|_2 - 1)^2], \quad (2.5)$$

where $\tilde{x} = \epsilon x_{\text{real}} + (1 - \epsilon)x_{\text{fake}}$ with $\epsilon \sim U[0, 1]$ representing interpolations between real and generated samples.

3. Related Work

Graph generation models have played a significant role in biology, social sciences, and applications such as data augmentation, drug discovery, and anomaly detection since their

introduction by Erdős and Rényi [16]. Probabilistic block models represent an important class of graph generators. The Stochastic Block Model (SBM) [17] assigns nodes to groups, with connection probability depending on group membership. The Mixed Membership Stochastic Blockmodel (MMSB) extended SBM by allowing nodes to belong to multiple groups simultaneously [18]. To address MMSB’s limitations including scalability issues, fixed community numbers, and poor performance on sparse graphs, dMMSB added a state-space model for evolving mixed memberships over time [19].

Deep learning approaches have revolutionized sequential graph generation. DeepGMG [1] formulates graph generation as a sequential decision process, adding nodes and edges one at a time using neural network probabilities. It employs GNNs with message-passing to maintain hidden representations and encode local structural information. The model supports conditional generation and handles multiple graph construction orderings through fixed or random node orderings during training. GraphRNN [2] generates graphs sequentially by modeling the graph as a sequence, and it is evaluated by comparing the generated and real graphs on properties such as degree distribution and clustering coefficients. Using two RNN components, a graph-level RNN for node addition and an edge-level RNN for connectivity prediction, GraphRNN captures both global structure and local patterns. GraphRNN uses BFS ordering to standardize graph representation and offers GraphRNN-S, a simplified variant assuming conditional edge independence for faster training. GraphVAE [3] pioneered variational autoencoders for graph generation, addressing the challenges of generating discrete graph structures. It learns a continuous latent space to encode and decode entire graphs, particularly for small molecular structures.

Closely related to our framework are adversarial methods for graph synthesis. GraphGAN [4] combines generative and discriminative models through adversarial training, with the generator modeling neighbor distributions using graph-aware softmax over BFS trees. Despite strong downstream performance, it suffered from training instability. EGraphGAN [5] addressed this using evolutionary algorithms with mutation strategies for sampling. For direct graph generation, MolGAN [6] generates molecular graphs using reinforcement learning but lacks flexibility for varying structures. LGGAN [7] overcomes this limitation, generating diverse graph types using MLPs to produce node label matrices and adjacency matrices. These approaches, however, typically generate edges through random sampling or direct matrix prediction without modeling structural dependencies. Our method differs by learning edge formation through distance-based predictions in a latent space and incorporating explicit density-aware selection to match class-specific sparsity patterns, enabling more structured and realistic graph synthesis.

4. Methodology

Our objective is to generate class-conditional synthetic graphs that mirror the structural characteristics of real data. Let $\mathcal{G} = \{(G_i, y_i)\}_{i=1}^N$ denote a set of labeled training graphs, where each $G_i = (V_i, E_i, X_i)$ contains nodes, edges, and node features, and $y_i \in \{1, \dots, C\}$ is a graph label. We learn a generator $G_\theta : \mathcal{Z} \times \mathcal{Y} \rightarrow \mathcal{G}$ that maps a latent input and a class condition to a graph whose distribution approximates \mathbb{P}_r . Unlike prior GAN-based approaches that rely on fixed-probability random edge sampling [7, 10], our method learns connectivity patterns directly from data. This change gives the model the ability to infer which nodes are likely to connect, rather than deciding edges from fixed probabilities. As shown in Figure 1, our architecture consists of a graph generator with a distance-based edge predictor and density-aware edge selection, and a discriminator for adversarial training.

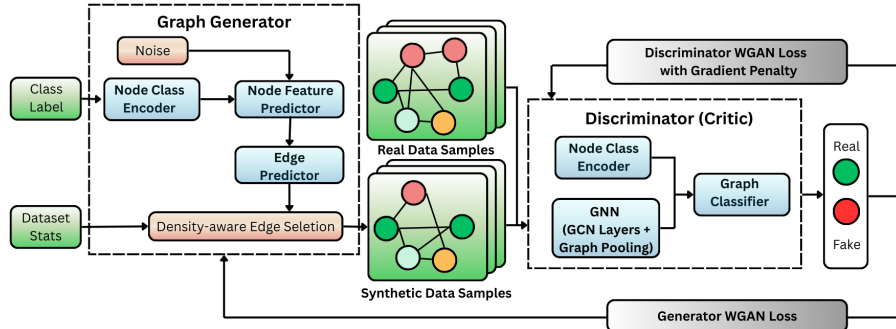


Figure 1. Model overview: the generator produces diverse node features from individual noise vectors combined with class embeddings. The distance-based edge predictor learns to map nodes into a latent space where proximity determines connectivity. Edge selection follows class-specific density statistics extracted from training data. The GNN-based critic processes graphs through multiple convolution layers, pools node features, and combines them with class embeddings to produce Wasserstein scores that ensure both distributional similarity and class alignment. Note that blue represents learning modules. In this approach, the probability distribution remains fully differentiable, allowing the predictor to learn directly from the WGAN objective.

4.1. Graph Generator

The generator creates graphs based on class-specific patterns through three stages: (1) class conditioning provides structural guidance for the target graph type; (2) node feature synthesis generates diverse node attributes appropriate for the class; and (3) edge construction learns connectivity patterns by representing structural relationships in a latent space.

Class Encoder: Class information is encoded through an embedding layer that maps class labels to dense vectors e_y . This embedding captures class-specific structural properties and is shared across all nodes in a generated graph to ensure consistent structure.

Node Feature Predictor: To generate diverse graphs within the same class, each node must have both shared class characteristics and individual variation. We provide each node with an independent noise vector $z_i \sim \mathcal{N}(0, I)$ to introduce necessary randomness for generating different structures. The generator combines this noise with the class embedding to produce node features $x_i = \text{MLP}_{\text{node}}([z_i; e_y])$, creating a feature matrix $X \in \mathbb{R}^{n \times d}$ that balances class consistency with structural diversity.

Different graph classes have different typical sizes. For example, molecular graphs usually contain 10-50 atoms while social networks may have thousands of nodes. We preserve this variation by estimating the size distribution for each class and sampling $n \sim \text{Clip}(\mathcal{N}(\mu_c, cf\sigma_c), n_{\min}^c, n_{\max}^c)$, where μ_c and σ_c represent the typical size and variation for class c . The factor cf prevents unrealistic extreme sizes while maintaining natural variation, and clipping ensures generated sizes stay within observed ranges.

Edge Predictor: Random edge sampling with fixed probabilities cannot capture the structured patterns in real graphs. In real networks, nodes with similar roles or attributes tend to connect more often. We introduce a learnable edge predictor that estimates connection likelihood based on learned structural similarity (i.e., graph homophily). This design follows the principle that similar nodes are more likely to connect, while remaining flexible enough to capture other patterns such as hub structures. The predictor maps nodes into a latent space where closeness reflects connection likelihood. For each node pair (i, j) , we compute the edge probability as:

$$p_{ij} = \sigma \left(\frac{-\|h_i - h_j\|_2 + \theta}{T} \right), \quad (4.1)$$

where the distance $\|h_i - h_j\|_2$ measures how different the node representations are. The learnable threshold θ determines the critical distance for edge formation and adapts to each class during training. The temperature T controls how sharp the decision boundary is, allowing smooth gradients during training. This approach enables the model to learn meaningful connection patterns rather than using fixed random rules.

Density-Aware Edge Selection: Real graphs have very different density patterns depending on their domain. Molecular graphs are sparse and locally structured due to chemical constraints, while social networks may be globally sparse but contain dense local groups. Capturing these class-specific density patterns is essential for generating realistic graphs.

We compute the expected edge density for each class:

$$\rho_c = \frac{2\bar{m}_c}{\bar{n}_c(\bar{n}_c - 1)} \quad (4.2)$$

where \bar{m}_c and \bar{n}_c are the average numbers of edges and nodes for class c . This density ρ_c captures the typical sparsity level for each graph type.

During generation, after computing edge probabilities for all $\binom{n}{2}$ possible pairs, we select the top $k = \lfloor \rho_c \cdot \binom{n}{2} \rfloor$ edges with highest probabilities. This selection serves three purposes: (1) matching the typical density of the target class, (2) using the model’s most confident predictions, and (3) guiding the edge predictor to learn realistic patterns. For undirected graphs, each selected edge is added in both directions to form the final edge set E .

4.2. Discriminator (Critic)

The discriminator evaluates whether generated graphs have the structural patterns found in real graphs from the target class.

Graph Encoder: The discriminator must evaluate structure at multiple levels: local patterns (triangles, stars), intermediate community structure, and global properties (connectivity, diameter). We use a GNN that builds representations through neighborhood aggregation (Equation 2.1). Over L layers, each node incorporates information from increasingly distant neighbors. This allows the discriminator to check whether generated graphs have consistent structural patterns at all scales. The resulting node embeddings $Z^{(L)} \in \mathbb{R}^{n \times d'}$ encode structural information. Global pooling creates a graph-level representation that is independent of node ordering, allowing the discriminator to evaluate graphs of different sizes consistently.

Class Encoder: To evaluate class-specific realism, the discriminator must compare generated structures against the expected patterns for that class. We combine the graph representation with the class embedding e_y to form $f = [g; e_y] \in \mathbb{R}^{d' + d_{\text{class}}}$, allowing the critic to assess whether the structure matches what is typical for the target class.

Graph Classifier: The combined representation is processed by an MLP critic that outputs a Wasserstein score $D_\phi(G, y) = \text{MLP}_\phi(f)$. This critic produces continuous scores that reflect structural quality: higher values indicate graphs whose patterns closely match real samples from class y , while lower values indicate structural problems. This scoring provides informative gradients that guide the generator toward more realistic synthesis.

4.3. Training Procedure

Training follows the WGAN-GP framework described above. In our implementation, we use a *conditional critic* $D_\phi(x, y)$, where y denotes the conditioning information.

The critic is trained to minimize: $\mathcal{L}_D = -\mathbb{E}_{x \sim \mathbb{P}_r}[D_\phi(x, y)] + \mathbb{E}_{\hat{x} \sim \mathbb{P}_g}[D_\phi(\hat{x}, y)] + \lambda \mathcal{L}_{\text{GP}}$, where $\hat{x} = G(z, y)$ and \mathbb{P}_g is the model distribution induced by the generator. The gradient penalty term reuses the definition in Equation 2.5, adapted to the conditional critic:

$\mathcal{L}_{\text{GP}} = \mathbb{E}_{\tilde{x}} \left[(\|\nabla_{\tilde{x}} D_{\phi}(\tilde{x}, y)\|_2 - 1)^2 \right]$, where \tilde{x} denotes interpolations between real and generated samples. To stabilize edge formation during training, we apply temperature annealing to the edge predictor: $T(t) = \max(T_{\text{end}}, T_{\text{start}} - \alpha \cdot t)$

The temperature parameter T governs the sharpness of the edge probability function. High initial temperature produces soft probability outputs, encouraging exploration of diverse graph structures and preventing collapse into limited edge patterns. As training progresses, T is gradually reduced, increasing contrast between close and distant node pairs and causing the edge predictor to shift from probabilistic sampling to confident, near-binary decisions. This annealing schedule provides smooth transition from stochastic graph formation early in training to sharper, more realistic structures later. This helps the generator converge to stable and expressive graph distributions while maintaining diversity during learning.

5. Evaluation

In this section, we empirically evaluate our graph generation framework’s ability to learn and reproduce complex graph distributions across diverse domains. We assess generation quality using a multi-faceted protocol that measures distributional fidelity, diversity, and structural preservation. Our experiments address three key questions: (i) how accurately the model captures the statistical properties of real graphs, (ii) whether it can generate diverse and novel graphs while preserving class-specific characteristics, and (iii) how it compares to existing graph generation methods. Source code is publicly available at https://github.com/ava-12/Density_Aware_WGAN.git.

5.1. Experimental Setup

Datasets: We evaluate our method on three benchmark graph classification datasets [20] that span different domains and structural complexities as shown in the table below. MUTAG(188 graphs) represents chemical compounds for mutagenicity prediction, with small graphs averaging 18 nodes. ENZYMES (600 graphs) contains protein structures across 6 enzyme classes, with moderate-sized graphs of approximately 33 nodes. PROTEINS (1,113 graphs) provides the largest and most structurally diverse dataset for enzyme/non-enzyme classification, with graphs ranging from 4 to 620 nodes. These datasets vary significantly in their feature representations: MUTAG includes both 7-dimensional node and 4-dimensional edge features, while ENZYMES and PROTEINS utilize only 3-dimensional node features, providing a comprehensive testbed for evaluating graph generation performance across different scales and complexities.

Property	MUTAG	ENZYMES	PROTEINS
Domain	Chemistry	Biochemistry	Biochemistry
Task Type	Binary	Multi-class (6)	Binary
Number of Graphs	188	600	1,113
Number of Classes	2	6	2
Avg. Nodes	17.93	32.63	39.06
Avg. Edges	19.79	62.14	72.82
Node Features	7	3	3

Baseline Models: We compare our approach against state-of-the-art graph generation models: DeepGMG [1], GraphRNN [2], LGGAN [7], and WPGAN [10].

Implementation Details: Our discriminator implements a Graph Convolutional Network (GCN) [12] with global mean pooling, $g = \frac{1}{n} \sum_{i=1}^n h_i^{(L)}$, which summarizes the graph by averaging node representations after multi-hop neighborhood aggregation. While we use mean pooling in our experiments, alternative operators such as sum or max pooling can be

employed. The framework is compatible with other GNN architectures, including GAT [13] and GraphSAGE [11], which we leave for future exploration. The table below presents the architecture parameters and training configuration optimized for stable adversarial learning. The hyperparameter choices reflect established best practices for WGAN training: the 5:1 critic-to-generator training ratio ensures accurate Wasserstein distance estimation, while the 2.5:1 learning rate ratio maintains adversarial balance. The reduced momentum ($\beta_1 = 0.5$) is standard for GAN training to improve convergence stability. We employ temperature annealing for the edge predictor, gradually reducing T from 2.0 to 0.5 over approximately 50 epochs to transition from exploratory edge formation to confident connectivity decisions.

Component	Parameter	MUTAG	ENZYMES	PROTEINS
Architecture	Noise dimension		16	
	Class embedding		8	
	Hidden dimension		32	
Training	Critic iterations (n_{critic})	5	5	5
	Gradient penalty (λ_{GP})	10	10	10
	Temperature schedule (T)	2.0 \rightarrow 0.5 (50 epochs)		
	Learning rate (Generator)	1×10^{-4}	3×10^{-4}	2×10^{-4}
	Learning rate (Discriminator)	5×10^{-4}	5×10^{-4}	5×10^{-4}
	Adam betas	(0.5, 0.9)	(0.5, 0.9)	(0.5, 0.9)
	Batch size	64	64	64
Evaluation	α		0.4	
	β		0.4	
	γ		0.2	

Evaluation Metrics: We evaluate generated graphs using Maximum Mean Discrepancy (MMD) [21], which measures distribution similarity in a reproducing kernel Hilbert space:

$$\text{MMD}^2(P, Q) = \mathbb{E}_{x, x' \sim P}[k(x, x')] + \mathbb{E}_{y, y' \sim Q}[k(y, y')] - 2\mathbb{E}_{x \sim P, y \sim Q}[k(x, y)], \quad (5.1)$$

where $k(\cdot, \cdot)$ is a positive-definite kernel, P is the real distribution, and Q is the generated distribution. We compute MMD over three complementary graph statistics. *Degree distribution* captures local connectivity by comparing empirical degree histograms, where $h_G[i]$ denotes the fraction of nodes with degree i . *Clustering coefficient* measures triangle formation density through the local clustering coefficient:

$$C_i = \frac{2 \cdot |\{e_{jk} : v_j, v_k \in N(v_i), e_{jk} \in E\}|}{k_i(k_i - 1)}, \quad (5.2)$$

aggregated into histograms for nodes with degree $k_i > 1$. *Spectral features* assess global topology using eigenvalues of the normalized Laplacian $L = I - D^{-1/2}AD^{-1/2}$. These three metrics provide complementary views: degree distribution captures local connectivity, clustering coefficients reveal community structure, and spectral features encode global topology. We combine them as:

$$\text{MMD}_{\text{combined}} = \alpha \cdot \text{MMD}_{\text{degree}} + \beta \cdot \text{MMD}_{\text{clustering}} + \gamma \cdot \text{MMD}_{\text{spectral}}, \quad (5.3)$$

where $\alpha + \beta + \gamma = 1$ and weights are determined empirically. Additionally, we measure *uniqueness* (proportion of distinct generated graphs) and *novelty* (proportion not in the training set) to assess diversity and generalization beyond memorization.

Results and Analysis: Table 1 presents a detailed breakdown of performance across all datasets and classes. For the ENZYMES dataset, the model achieves strong and consistent performance across all six classes, with combined MMD values ranging from 0.106 to 0.182. The model demonstrates particularly strong spectral alignment (MMD values between 0.042-0.133) and clustering preservation (0.114-0.200), indicating effective capture of both local motifs and global topology. Degree distribution alignment is also strong,

Table 1. Detailed experimental results across all datasets and classes

Dataset	Class	MMD Degree	MMD Clustering	MMD Spectral	Avg Nodes (Real→Gen)	Avg Edges (Real→Gen)	MMD Combined	Uniqueness	Novelty
MUTAG	Class 0	0.361	1.082	0.12	13.4→13.4	14.0→13.7	0.577	0.958	0.912
	Class 1	0.249	1.04	0.105	20.7→19.7	23.5→22.1	0.522	0.974	0.975
ENZYMES	Class 0	0.158	0.142	0.095	31.6→41.0	63.3→86.0	0.133	0.992	0.986
	Class 1	0.153	0.143	0.08	32.1→31.6	62.1→68.9	0.127	0.992	0.931
	Class 2	0.15	0.142	0.072	30.1→30.5	60.3→62.2	0.124	0.988	0.917
	Class 3	0.170	0.149	0.081	37.4→37.1	73.5→71.2	0.135	1.000	0.962
	Class 4	0.141	0.129	0.098	32.1→28.4	60.4→53.0	0.122	1.000	0.982
	Class 5	0.214	0.173	0.08	38.8→27.7	77.3→54.0	0.158	0.979	0.953
PROTEINS	Class 0	0.063	0.042	0.146	54.5→49.5	103.0→124.9	0.08	0.955	0.994
	Class 1	0.082	0.055	0.056	19.6→25.1	36.5→57.2	0.063	0.913	0.838

Table 2. Comparison with state-of-the-art methods using MMD metrics (lower is better) as presented in [7, 10]. Best results are bolded.

Model	PROTEINS			ENZYMES		
	Degree	Clustering	Spectral	Degree	Clustering	Spectral
DeepGMG [1]	0.96	0.63	–	0.43	0.38	–
GraphRNN [2]	0.04	0.18	–	0.06	0.20	–
LGGAN [7]	0.18	0.15	–	0.09	0.17	–
WPGAN [10]	0.03	0.31	–	0.02	0.28	–
Our Approach	0.09	0.07	0.07	0.10	0.08	0.05

with MMD values between 0.137-0.208. The model achieves near-perfect uniqueness scores (0.941-1.000) and high novelty (0.833-1.000) across all classes, confirming diverse generation without memorization.

On the PROTEINS dataset, the model achieves the lowest combined MMD scores overall (0.066-0.089), demonstrating excellent structural fidelity. Performance is particularly strong on clustering coefficients (0.039-0.050) and degree distributions (0.054-0.091), with Class 1 showing the best overall alignment (combined MMD of 0.066). Spectral MMD values are slightly higher for Class 0 (0.192) compared to Class 1 (0.062), reflecting differences in graph sizes between classes. High uniqueness (0.921-0.951) and novelty (0.859-1.000) scores confirm the model generates diverse, non-memorized samples.

On the MUTAG dataset, the model accurately reproduces graph size statistics (average nodes and edges closely match real distributions) while maintaining high diversity (uniqueness 0.857-0.933, novelty 0.933-1.000). However, clustering coefficient MMD values are notably higher (1.053–1.059) compared to other datasets. This is attributable to MUTAG’s small graph sizes (13-21 nodes) and limited structural diversity, which makes distributional matching more challenging. Despite this, combined MMD scores (0.527-0.536) indicate reasonable overall structural alignment, and low degree (0.249–0.266) and spectral (0.102-0.107) MMD values demonstrate the model still captures fundamental connectivity patterns.

Table 2 compares our approach against state-of-the-art methods on PROTEINS and ENZYMES datasets. On PROTEINS, our model achieves the best clustering coefficient preservation (MMD of 0.07) among all methods, significantly outperforming LGGAN (0.15), GraphRNN (0.18), and WPGAN (0.31), demonstrating superior modeling of local connectivity and community structure. For degree distribution, our approach achieves competitive performance (0.09) compared to WPGAN’s best result (0.03). We also report spectral MMD (0.07), which captures global topology, a metric not evaluated by prior work.

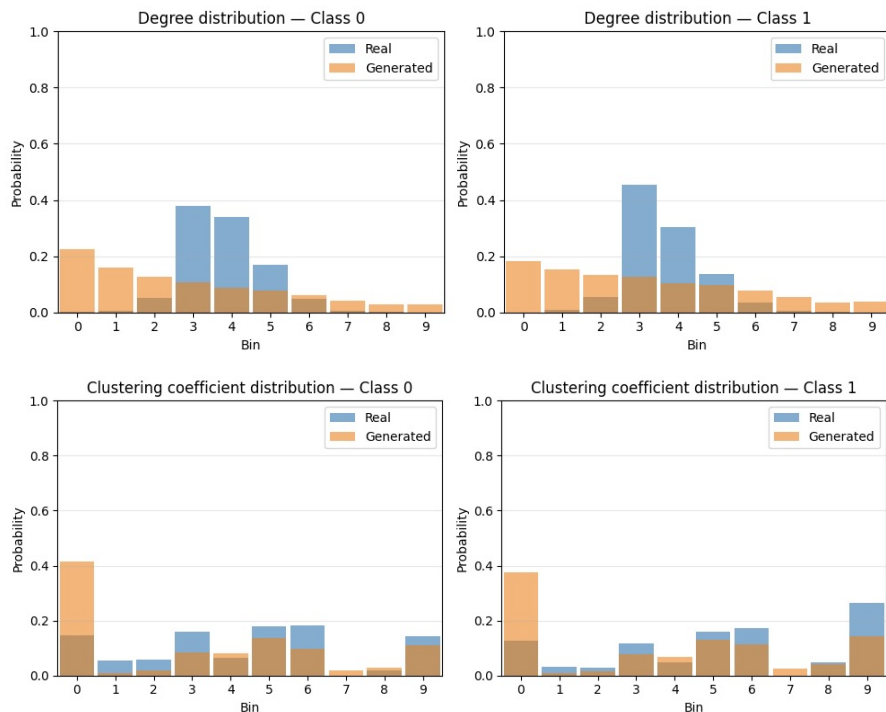


Figure 2. Distribution comparison between real and generated graphs on the PROTEINS dataset. Top: The degree distribution shows that the real graphs are concentrated around mid-range values (bins 2–6), while the generated graphs exhibit a wider spread across bins. This behavior is influenced by the density-aware edge selection mechanism, which enforces class-average sparsity but can still lead to deviations from the true degree patterns. Bottom: Clustering coefficient distribution demonstrates strong alignment with substantial overlap, confirming the model’s superior preservation of local connectivity patterns and community structure, consistent with the best clustering MMD (0.07) achieved among compared methods.

On ENZYMES, our method achieves the best clustering coefficient alignment (0.08) across all compared approaches, indicating strong capture of higher-order structural patterns. While WPGAN achieves the lowest degree MMD (0.02), our approach remains competitive (0.10) and demonstrates superior spectral performance (0.05), suggesting balanced modeling of both local and global structural properties.

Figures 2 and 3 provide distributional and qualitative analysis on the PROTEINS dataset. The clustering coefficient histogram (Figure 2, bottom) shows strong overlap between real and generated distributions, confirming superior preservation of local connectivity patterns. The degree distribution (Figure 2, top), however, reveals more conservative assignment with concentration in bins 2–6 compared to the broader real distribution. Qualitatively (Figure 3), generated graphs successfully preserve key structural characteristics including triangular motifs and community-like clusters, with our distance-based edge predictor learning meaningful node relationships that produce similar topological organization. However, isolated or low-degree nodes appear more frequently in generated samples. This difference comes from our fixed top- k edge selection method, which forces graphs to match class-average density but limits the natural variety in node degrees. While this ensures graphs have appropriate overall connectivity, it biases the range of individual node degrees. Despite

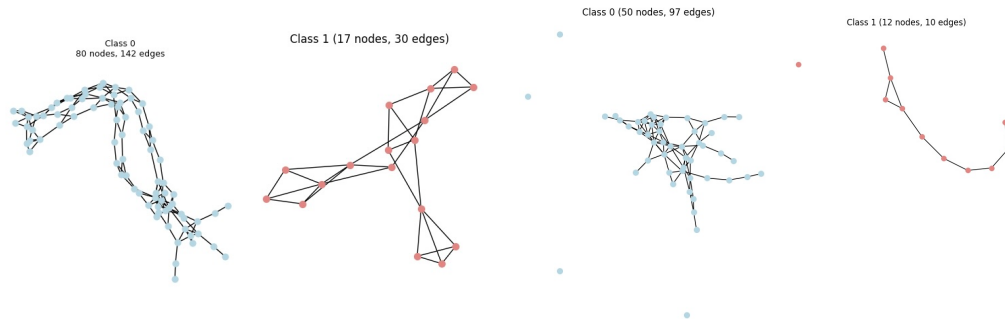


Figure 3. Visual comparison on PROTEINS dataset (left: real graphs from Classes 0 and 1; right: generated graphs). Our model preserves structural characteristics including local motifs, triangular patterns, and community clusters through learned distance-based edge prediction. While isolated nodes appear due to deterministic density-aware selection enforcing class-average sparsity, the overall topology, clustering patterns, and class-specific structural signatures are maintained.

this limitation, the model maintains overall structural consistency, clustering patterns, and class-specific characteristics while generating diverse, non-memorized instances.

A critical distinction of our evaluation is the explicit measurement of uniqueness and novelty alongside distributional metrics. Prior works [1, 2, 7, 10] focus solely on MMD scores, which can be minimized through memorization or near-replication of training samples. Our results demonstrate competitive MMD performance while maintaining high uniqueness and novelty, confirming genuine generative capability rather than overfitting.

6. Conclusion

We propose a density-aware, class-conditional graph generation framework addressing fundamental challenges in realistic graph synthesis: discrete topology, variable graph sizes, and class-specific structural patterns. The approach learns graph connectivity in latent space through a distance-based edge predictor instead of random edge sampling with fixed probabilities, capturing meaningful relational dependencies between nodes while explicitly controlling class-specific graph density. Integrated within a Wasserstein GAN framework with gradient penalty and a GCN-based critic, the proposed model ensures stable training and promotes graphs with coherent topology and realistic connectivity patterns.

Experiments on benchmark datasets (MUTAG, ENZYMES, and PROTEINS) demonstrate the model accurately captures statistical properties of real graphs, achieving competitive or superior performance on MMD-based distributional metrics, particularly for clustering and spectral characteristics reflecting higher-order connectivity. The framework consistently generates diverse and novel graphs, with high uniqueness and novelty scores, while preserving class-specific structural signatures and sparsity patterns. Compared to existing methods, our approach shows improved balance between structural fidelity and generative diversity, producing graphs whose local neighborhoods and global organization more closely match real-world distributions.

However, our top- k edge selection, while enforcing class-specific sparsity, can bias degree distributions by reducing natural variation within classes. This suggests fixed class-average density is too restrictive for capturing degree variety in real graphs. Future work includes using probabilistic sampling from learned distributions to add variation while maintaining expected density, and adding degree distribution objectives to the training loss.

References

- [1] Y. Li, O. Vinyals, C. Dyer, R. Pascanu, and P. W. Battaglia. “Learning Deep Generative Models of Graphs”. In: *arXiv preprint arXiv:1803.03324* (2018).
- [2] J. You, B. Liu, Z. Ying, V. Pande, and J. Leskovec. “Graph Convolutional Policy Network for Goal-Directed Molecular Graph Generation”. In: *Advances in Neural Information Processing Systems*. 2018.
- [3] M. Simonovsky and N. Komodakis. “GraphVAE: Towards Generation of Small Graphs Using Variational Autoencoders”. In: *International Conference on Artificial Neural Networks*. 2018.
- [4] H. Wang, J. Wang, J. Wang, M. Zhao, W. Zhang, F. Zhang, X. Xie, and M. Guo. “GraphGAN: Graph Representation Learning with Generative Adversarial Nets”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 32. 2018.
- [5] P. Wang, Z. Liu, Z. Wang, Z. Zhao, D. Yang, and W. Yan. “Graph Generative Adversarial Networks with Evolutionary Algorithm”. In: *Applied Soft Computing* 164 (2024), p. 111981.
- [6] N. De Cao and T. Kipf. “MolGAN: An Implicit Generative Model for Small Molecular Graphs”. In: *arXiv preprint arXiv:1805.11973* (2018).
- [7] S. Fan and B. Huang. “Labeled Graph Generative Adversarial Networks”. In: *arXiv preprint arXiv:1906.03220* (2019).
- [8] M. Arjovsky, S. Chintala, and L. Bottou. “Wasserstein Generative Adversarial Networks”. In: *Proceedings of the 34th International Conference on Machine Learning*. 2017, pp. 214–223.
- [9] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville. “Improved Training of Wasserstein GANs”. In: *Advances in Neural Information Processing Systems*. 2017.
- [10] S. A. Razi Razavi, J. Sargant, S. Houghten, and R. Dividino. “Generating Labeled Graphs Using Conditional Wasserstein GANs”. In: *35th International Conference on Collaborative Advances in Software and Computing*. IEEE Computer Society, 2025.
- [11] W. L. Hamilton, Z. Ying, and J. Leskovec. “Inductive Representation Learning on Large Graphs”. In: *Advances in Neural Information Processing Systems*. 2017.
- [12] T. N. Kipf and M. Welling. “Semi-Supervised Classification with Graph Convolutional Networks”. In: *5th International Conference on Learning Representations*. 2017.
- [13] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio. “Graph Attention Networks”. In: *6th International Conference on Learning Representations*. 2018.
- [14] K. Xu, W. Hu, J. Leskovec, and S. Jegelka. “How Powerful are Graph Neural Networks?” In: *7th International Conference on Learning Representations*. 2019.
- [15] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. “Generative Adversarial Nets”. In: *Advances in Neural Information Processing Systems* 27 (2014).
- [16] P. Erdős and A. Rényi. “On Random Graphs I”. In: *Publicationes Mathematicae Debrecen* 6 (1959), pp. 290–297.
- [17] I. Goodfellow. “NIPS 2016 Tutorial: Generative Adversarial Networks”. In: *arXiv preprint arXiv:1701.00160* (2016).
- [18] E. M. Airoldi, D. M. Blei, S. E. Fienberg, and E. P. Xing. “Mixed Membership Stochastic Blockmodels”. In: *Journal of Machine Learning Research* 9 (2008), pp. 1981–2014.
- [19] E. P. Xing, W. Fu, and L. Song. “A State-Space Mixed Membership Blockmodel for Dynamic Network Tomography”. In: *The Annals of Applied Statistics* 4.2 (2010), pp. 535–566.
- [20] C. Morris, N. M. Kriege, F. Bause, K. Kersting, P. Mutzel, and M. Neumann. “TUDataset: A Collection of Benchmark Datasets for Learning with Graphs”. In: *arXiv preprint arXiv:2007.08663* (2020).
- [21] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola. “A Kernel Two-Sample Test”. In: *Journal of Machine Learning Research* 13 (2012), pp. 723–773.