ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Exploration of Image Narrative Innovation of Party History Themes in Jilin: Enhancing Content Production and Dissemination Efficiency Driven by AIGC Technology and Intelligent Workflow

Yizhong Zhang[1,*]

[1] Comic Academy, Jilin Animation Academy, Changchun, Jilin, 130000, China

SUMMARY: *The increasing development and maturity of Artificial Intelligence Generated Content (AIGC) provides a solid technical support for the application of image narrative model in Jilin Party History. This paper constructs an image generation model with StyleGAN2 image generation model and ChineseCLIP model as the main components. The model is divided into a training phase and a testing phase, in which the ChineseCLIP image encoder is used to extract image features, and the ChineseCLIP text encoder is used to extract the input Chinese text features in the testing phase, thus realizing an accurate understanding of the input Chinese semantics. The spectral normalization technique is added to the model discriminator to constrain the model gradient value, which stabilizes the model training while maintaining the original expression ability of the model. In addition, the social network analysis method is adopted to carry out the quantitative analysis of the content of Jilin Party history topics by selecting the overall structure of the network, the characteristics of subgroups, key nodes and other corresponding indicators. Based on the image generation model of this paper, the image narrative communication network structure of Jilin Party history, with a graph density of 1.42, modularity coefficient of 0.931, and an average path length of 1.81, not only has a significant network community structure and a high efficiency of information dissemination, but also is an effective path for AIGC technology to assist in the innovation and development of the spirit of party history in Jilin, and to promote the inheritance of the spirit of party history.*

KEYWORDS: *image generation model; AIGC technology; Jilin party history; image narrative; social network analysis*

1 Introduction

Jilin is the base of the Northeast Resistance League, an important battlefield of the Liberation War, an important old industrial base and a commercial grain base [1, 2]. For nearly 90 years, the Party has always struggled valiantly in this hot land, and the Party has led the people of Jilin in revolution, construction and reform, and made great and brilliant historical achievements [3, 4]. It can be said that without the party, there is no new Jilin. The review and contemplation of Jilin party history topics based on image narrative is the pursuit and inheritance of the noble spirit of the party people in Jilin, and it is the rich nutrition and strong motivation to revitalize Jilin and develop Jilin [5, 6].

The history of image narrative has a long history, from cave petroglyphs to artifact decorations, the ancestors opened the initial narrative activities by way of image records [7, 8]. The image narrative of party history in Jilin is to reintegrate the created spatial images into the

*13069164188@163.com
https://doi.org/10.65102/is2026039
Published: 30 April, 2026

time sequence to achieve the purpose of reconstructing the historical context, so as to show the significant events that happened in Jilin in the party's history of more than a hundred years [9-12]. The image narrative of the party history theme has its independent self-sufficiency, and should be regarded as a kind of image text with unique historical narrative value, and its unique historical narrative ability provides an important support for the artist's artistic creation process to visualize the artistic reality as the historical reality [13-16].

With the development of artificial intelligence, Jilin party history theme image narrative ushered in the development of innovation, especially the application of artificial intelligence-generated content (AIGC) technology, which improves the production efficiency and reduces the production cost of the image narrative content [17-19]. AIGC refers to the technology based on algorithms, models, and rules for generating content such as text, pictures, sounds, videos, and codes, which provides generative AI products or services should comply with the requirements of laws and regulations and respect social morality, public order and morality [20-23]. In the image narrative of Jilin Party history topics, it is able to generate narrative content according to the characteristics of the given Party history topics and the requirements of the narrative, and disseminate them based on social media, online education platforms and other channels, so as to expand the audience scope of Jilin Party topics [24-27].

In this paper, we first summarize the operation steps of the model based on the design idea of the image generation model. It focuses on analyzing the working principle and content of the dual-channel feature fusion module, and designing the loss function in combination with the network training requirements. The spectral normalization technique is introduced into the model discriminator to optimize the model training, and the StyleGAN2-based image generation model is formed. Secondly, social network analysis is used as the research and analysis method to demonstrate the mathematical principles and calculation of the corresponding indicators in the three aspects of the overall structure of the network, the characteristics of subgroups and key nodes. Again based on the subject of Jilin Party History, collect and screen text data. Through quantitative analysis experiments and design visual effect evaluation, the output performance of the image generation model in this paper is verified. Finally, the communication network is constructed based on the generated image narratives of Jilin Party history, summarizing the overall performance of the network, analyzing the node centrality, intermediary centrality and network structure.

2 StyleGAN2 based image generation model

The StyleGAN2-based image generation model structure is capable of mapping Chinese text to the potential space of G and generating the target image, which relies on two pre-training components: the StyleGAN2 image generation model G and the image encoder and text encoder of Chinese CLIP, denoted as CN-I and CN-T, respectively.

2.1 Overview of the model

In the training phase, given an input image x, image reconstruction is implemented by the model to obtain y. An encoder-decoder structure model is used to convert the input image into potential spatial embedding using an encoder, and a decoder is used to convert the embedding vectors back to the image form, keeping the similarity between the reconstructed image and the original input image maximized. The process is shown in equation (1):

$$y = Decoder_{image}\left(Encoder_{image}\left(x\right)\right) \tag{1}$$

$Decoder_{image}$ is the image generation model G, which is responsible for decoding and generating the final image from the latent feature vector; and $Encoder_{image}$ is the image encoder. The image encoder uses a two-branch structure to obtain the latent representation of the image, which is used to extract the image features and to complement the feature extraction, respectively. The first branch uses the CN-I image encoder to encode the input image x into a 512-dimensional potential feature vector $c \in \mathbb{R}^{1\times 512}$ as in equation (2):

$$c = CN - I\left(x\right) \tag{2}$$

The CN-I encoder is used to extract the global features of the image, which cannot fully capture the detailed information of the image due to structural limitations. Therefore the second branch uses the encoder part of the conditional variational auto-encoder CVAE, taking the image x as input and encoding the feature vector c obtained from the first branch as a conditional variable, and the output potentially represents the probability distribution $q\left(z \mid x,c\right)$ of z as in Eqs. (3)-(4):

$$q\left(z \mid x,c\right) = \mathcal{N}\left(f_{enc}^{\mu}\left(x,c\right), f_{enc}^{\sigma^2}\left(x,c\right)\right) \tag{3}$$

$$t = CN - T\left(T\right) \tag{4}$$

A nonparametric sampling method, NpSamp, is used to bridge the distribution gap that exists between the ChineseCLIP text/image embeddings and to introduce a nonparametric sampling method that does not require separate training and uses only the domain images for testing.

NpSamp computes a random convex combination as in Eq. (5) by retrieving K image embeddings c^1 to c^K, selecting the $M < K$ image embeddings c^{i_1} to c^{i_K} that are closest to the text embedding t in cosine distance:

$$t' = \sum_{j=1}^{M} \alpha_j c_j^{i_j} \tag{5}$$

α_j is sampled from the Dirichlet distribution, and the computed result of the randomized convex combination t' contains some of the image embedding information to shorten the gap of the ChineseCLIP graphic embedding. Afterwards t' is jointly input into DualIC with a potential representation randomly sampled from the standard Gaussian distribution $\mathcal{N}\left(0,1\right)$.

2.2 Dual-channel feature fusion module

In this paper, a two-channel feature fusion module is designed see Fig. 1 for improving the semantic consistency of the generated image with the input text description.

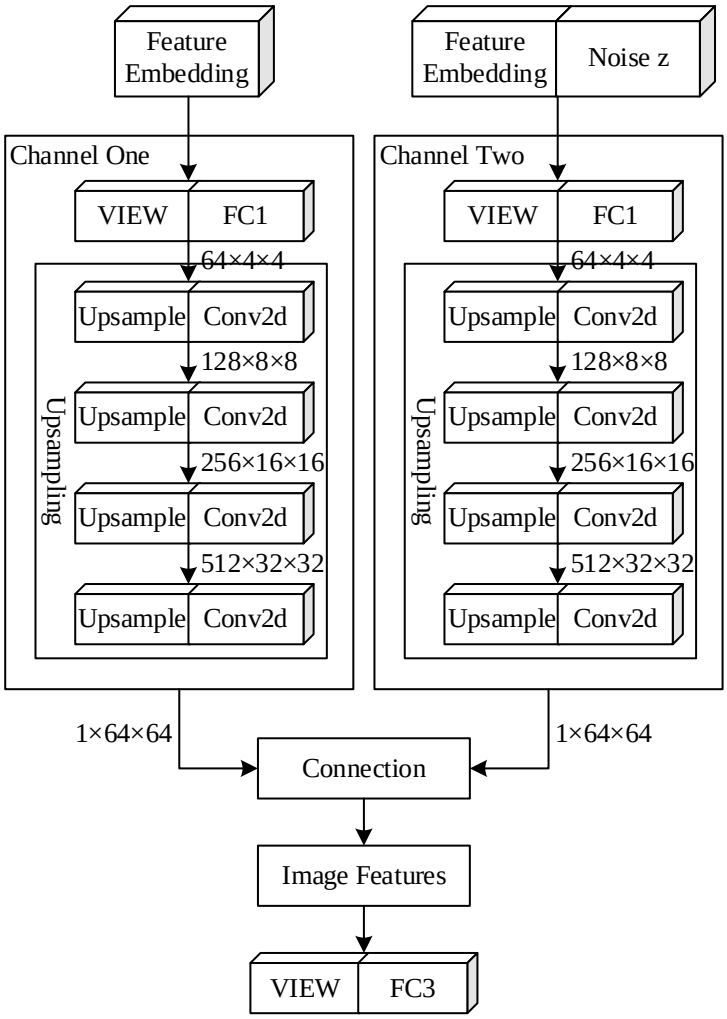

Figure 1: Dual-channel feature fusion module

Channel 1 is a pure feature vector channel with the goal of generating a high-dimensional feature map based on the feature transformation of the input text or image, preserving the semantic information in the original features. The input feature embedding image feature c or text feature t' is first mapped to a 1024-dimensional feature vector through a fully connected layer FC1. The aim is to transform the input feature embedding into a higher dimensional space with equation (6):

$$h_1 = FC1\left(c \| t'\right) \tag{6}$$

$h_1 \in \mathbb{R}^{1024}$ is the output of the fully-connected layer. 1024-dimensional feature vector h_1 is reshaped into a $64 \times 4 \times 4$ 3D feature map through the VIEW layer, which converts the one-dimensional feature vector into a 3D feature map suitable for convolution operation as in Eq. (7):

$$H_1 = VIEW\left(h_1\right) \tag{7}$$

$H_1 \in \mathbb{R}^{64\times4\times4}$ is the reshaped feature map. The H_1 is upsampled by four up-sampling and

3×3 2D convolutional layers to extract features and expand the spatial size of the feature map. The upsampling operation expands the spatial size of the feature map from 4×4 to 64×64, while the convolution operation is used to extract local features as in equation (8):

$$F_1 = upsample\_conv\left(H_1\right) \tag{8}$$

$F_1 \in \mathbb{R}^{1\times64\times64}$ is the final output feature map.

Channel 2 is the channel where feature vectors are combined with noise, and the goal is to fuse the input feature embeddings with the latent vectors to generate a joint feature map that introduces more diversity and robustness. The input feature embeddings c or t' are connected with the latent vectors z in the feature dimension to form a 640-dimensional joint feature vector. The aim is to combine the original features with noise to enhance the generative power of the model as in equation (9):

$$h_2 = Concat\left(c \| t', z\right) \tag{9}$$

The $h_2 \in \mathbb{R}^{640}$ is the connected feature vector. The joint feature vector h_2 is further mapped to the 640-dimensional feature space through a fully connected layer FC2, thus realizing the nonlinear transformation of the joint features and extracting more complex semantic information. There is equation (10):

$$h_2' = FC2\left(h_2\right) \tag{10}$$

$h_2' \in \mathbb{R}^{640}$ is the output of the fully-connected layer, and the 640-dimensional feature vector $h_{2'}$ is reshaped by the VIEW layer into a $64\times4\times4$ three-dimensional feature map, with equation (11):

$$H_2 = VIEW\left(h_2'\right) \tag{11}$$

The $H_2 \in \mathbb{R}^{64\times4\times4}$ is the reshaped feature map. The H_2 is upsampled and convolved to generate a $1\times64\times64$ feature map.

The feature maps generated by the two channels are fused into a 640-dimensional feature vector by connecting operation and linear transformation. The feature maps F_1 and F_2 generated by channel one and channel two are connected in the first dimension by connecting the layer $Concat$ to form a $2\times64\times64$ feature map as in equation (12):

$$F = Concat\left(F_1, F_2\right) \tag{12}$$

The $F \in \mathbb{R}^{2\times64\times64}$ is the connected feature map. The feature map is reshaped into a one-dimensional vector that is mapped to the target dimension by a linear transformation (fully connected layer) that transforms the $2\times64\times64$ feature map into a 640-dimensional feature vector as in equation (13):

$$s = \mathrm{Re}\,shape\left(F\right)\cdot W + b \tag{13}$$

$\mathrm{Re}\,shape(F) \in \mathbb{R}^{8192}$ is the one-dimensional vector after spreading the $2 \times 64 \times 64$ feature map, and $W \in \mathbb{R}^{8192 \times 640}$ is the weights Matrix. $b \in \mathbb{R}^{640}$ is the bias vector. $s \in \mathbb{R}^{1 \times 640}$ is the final fused feature vector.

2.3 Loss function

In this paper, we propose a multi-task loss function that combines perceptual loss, potential spatial constraints, semantic consistency loss and distributional similarity loss to fully optimize the quality and semantic consistency of the generated images.

The overall objective loss function ℓ is shown in equation (14):

$$\ell = \ell_{LPIPS} + \lambda_{\omega\text{-}norm}\ell_{\omega\text{-}norm} + \ell_{CN\text{-}CLIP} + \lambda_{KL}\ell_{KL} \tag{14}$$

λ is the corresponding weight for each loss term. By combining multiple loss terms, the model is able to perform comprehensive optimization in terms of perceptual quality, latent spatial distribution, semantic consistency and distributional similarity. The pixel-level ℓ_2 loss is replaced with the learned perceptual image patch similarity LPIPS loss ℓ_{LPIPS} to ensure a high degree of perceptual similarity between the input and output image depth representations. Comparing image features at multiple levels and computing their differences realizes ℓ_{LPIPS} loss as in equation (15):

$$\ell_{LPIPS} = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{H} Lin\left(diff_{ij}\right) \tag{15}$$

N is the number of input samples, H is the number of channels of the feature, and $Lin\left(diff_{ij}\right)$ is the linear layer's treatment of the squared $diff_{ij}$ of the feature difference.

The ω normalization loss ℓ_{ω_norm} is introduced to prevent the predicted StyleGAN2 latent vectors ω from deviating too much from the distribution as in Equation (16):

$$\ell_{\omega\_norm} = \left\| latent - latent\_avg \right\|^2 \tag{16}$$

$latent$ is the potential vector of the sample and $latent\_avg$ is the average potential vector. By constraining the distribution of potential vectors, the stability and diversity of the generated images are improved. Introducing a $\ell_{CN\text{-}CLIP}$ loss encourages the model to learn similar input and output images in the ChineseCLIP space as in equation (17):

$$\ell_{CN\text{-}CLIP} = \frac{1}{N}\sum_{i=1}^{N}\left(1 - similarity_i\right) \tag{17}$$

$similarity_i$ is the cosine similarity between the original image features and the reconstructed image features of the i th sample.

Calculating the KL scatter measures the similarity between the generated distribution and the true distribution as in equation (18), which helps the model learn to generate samples closer to the true distribution:

$$\ell_{KL} = D_{KL}\left(P_{gen} \| P_{real}\right) \tag{18}$$

P_{gen} is the distribution of generated images and P_{real} is the distribution of real images.

2.4 Spectral Normalization

Generative Adversarial Networks often face the problem of pattern collapse and pattern collapse during training, which mainly stems from the instability of the discriminator performance. When the real data distribution is unknown, the prediction of the data distribution by the discriminator often has large fluctuations. Moreover, if the predicted distribution does not intersect with the real distribution, it will be difficult for the generator to obtain effective gradient information, which will lead to problems such as training stagnation, model collapse, and missing details in the generated image.

To address the above problems, this study introduces the spectral normalization technique in the discriminator. This method restricts the Lipschitz constant of the network by constraining the spectral paradigm of the weights of the convolution kernel, thus effectively controlling the range of values of the network parameters. Specifically, spectral normalization performs singular value decomposition of the weight matrix of each convolutional layer and normalizes its maximum singular value, thus ensuring that the network satisfies the Lipschitz continuity condition. This regularization method significantly improves the stability of the training process while maintaining the expressive power of the model.

2.4.1 Discriminator structure

In order to process the multi-scale images (64×64, 128×128 and 256×256) output from the generator, this model designs three discriminators with similar structure but progressive scale as shown in Fig. 2. The base discriminator (64×64) adopts a modular design, and each Down Block consists of convolutional layers and batch normalization layers connected sequentially. The end of the discriminator is divided into two functional branches: the unconditional discriminator branch focuses on distinguishing the generated image from the real image, while the conditional discriminator branch is responsible for evaluating the semantic consistency between the image and the input text.

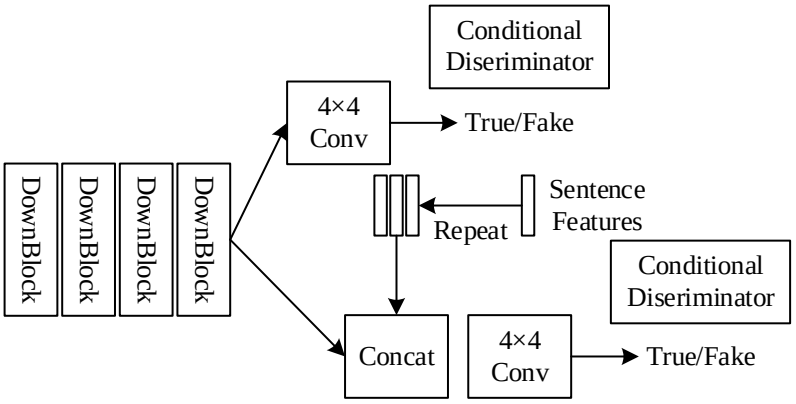

Figure 2: The construction of the discriminator

For higher resolution inputs, the 128×128 discriminator adds one DownBlock module to the front end of the base structure, and the 256×256 discriminator extends two DownBlock modules accordingly. This hierarchical discriminator design not only maintains the relevance

of feature extraction at each scale, but also ensures the parameter efficiency through module reuse, while retaining the complete multi-scale discriminant function.

2.4.2 Spectral normalization

The spectral normalization technique is introduced in the convolutional layer of the discriminator to precisely control the Lipschitz constant of the discriminator f by constraining the spectral paradigm of the weights of each layer, thus enhancing the training stability of the network.

For the n th layer network of the model, the input-output relationship is shown in equation (19):

$$x_n = a_n\left(W_n x_{n-1} + \beta_n\right) \tag{19}$$

x_{n-1}, x_n denotes the inputs and outputs of the n th layer, W_n denotes the network weight parameter of the n th layer in the model, and A_n denotes the diagonal matrix, the input-output relationship of the multilayer neural network can be expressed as equation (20):

$$f\left(x\right) = A_n W_n A_{n-1} W_{n-1} \cdots A_1 W_1 x \tag{20}$$

When the Lipschitz continuity condition is satisfied, the gradient of $f(x)$ is required as in equation (21):

$$\left\|\nabla_x\left(f\left(x\right)\right)\right\|_2 = \left\|A_n W_n A_{n-1} W_{n-1} \cdots A_1 W_1\right\|_2 \le \left\|A_n\right\|_2 \left\|W_n\right\|_2 \cdots \left\|A_1\right\|_2 \left\|W_1\right\|_2 \tag{21}$$

where $\left\|W\right\|_2$ denotes the spectral paradigm of the matrix W, defined as equation (22):

$$\sigma\left(W\right) = \max_{h\neq 0} \frac{\left\|Wh\right\|_2}{\left\|h\right\|_2} = \max_{\|h\|=1} \left\|W\right\|_2 \tag{22}$$

where h denotes the perturbation vector to the neural network input; $\left\|\cdot\right\|_2$ denotes the L2 paradigm. $\sigma\left(W\right)$ is the maximum eigenvalue of matrix W, and for the diagonal matrix A, there is $\sigma(A) = \max\left(d_1, d_2, \cdots, d_n\right)$, and d is the diagonal element. Thus, the above equation can be expressed as Eq. (23):

$$\left\|\nabla_x\left(f\left(x\right)\right)\right\|_2 \le \prod_{i=1}^{n} \sigma\left(W_i\right) \tag{23}$$

In order to make $f\left(x\right)$ satisfy the Lipschitz constraints, a normalization operation is performed on the formula as in Eq. (24):

$$\begin{aligned}\left\|\nabla_x\left(f\left(x\right)\right)\right\|_2 &= \left\|A_n\frac{W_n}{\sigma\left(W_n\right)}A_{n-1}\frac{W_{n-1}}{\sigma\left(W_{n-1}\right)}\cdots A_1\frac{W_1}{\sigma\left(W_1\right)}\right\|_2 \\ &\le \prod_{i=1}^{n}\frac{\sigma\left(W_i\right)}{\sigma\left(W_i\right)}=1\end{aligned} \tag{24}$$

Spectral normalization ensures Lipschitz continuity by constraining the spectral paradigm of the weight matrix W of each layer of the discriminator. The maximum singular value of the weight matrix is computed and normalized during parameter update. When all layers of the network satisfy the Lipschitz condition, the whole discriminator function naturally maintains Lipschitz continuity, which significantly improves the stability and convergence of the adversarial training.

3 Social network analysis

The social network analysis method provides a series of indicators to quantitatively analyze the overall structure of the network, the characteristics of the subgroups, and the key nodes according to the idea of going from the whole to the local and then to the individual, which can clarify the communication characteristics of the event in a more scientific and comprehensive way. The framework of the social network analysis method is shown in Figure 3.

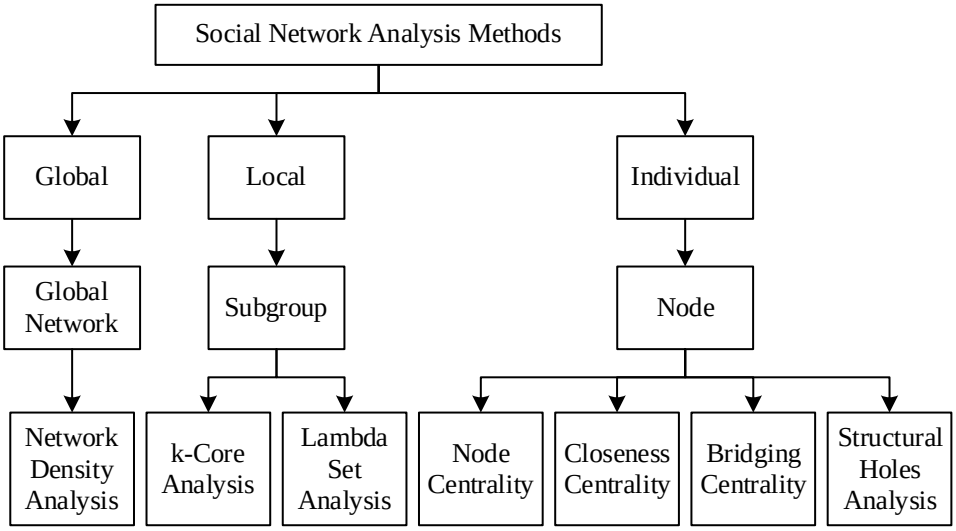

Figure 3: The analytical framework of social network analysis method

3.1 Network density analysis

Network density is the ratio of the “actual number of relationships existing in the network” to the “theoretical maximum number of possible relationships”, calculated as equation (25):

$$\Delta = \frac{L}{N\left(N-1\right)} \tag{25}$$

where N denotes the total number of nodes in the propagation network and L denotes the number of actually existing affiliations. Density is a reflection of the closeness of social network

relationships, and the larger its value, the higher the closeness between members.

3.2 k-nucleus analysis

k-kernel analysis: it means that each point in the subgroup is directly connected to at least k other points, i.e., all points in the subgroup have a node centrality of at least k, in order to measure the degree of regional clustering and to study the compositional structure. The larger the value of k, the higher the degree of clustering and the stronger the connection of the subgroup.

3.3 Lambda set analysis

The Lambda set is used to characterize the robustness of a network and is usually measured using the edge associativity within a subgroup. An edge correlation degree is $\sigma(i, j)$ if at least $\sigma(i, j)$ edges i, j have to be removed between two nodes i, j before there are no connected paths between them. The larger the edge correlation, the more robust the network is.

3.4 Node Centrality

Node centrality is to find out the points in the network graph that are directly connected to many other points. In directed network graph, the calculation of node centrality is divided into two aspects, one is the inner node centrality $d_{in}(n_i)$ determined corresponding to the inner node degree and the other is the outer node centrality $d_{out}(n_i)$ determined corresponding to the outer node degree, and then the two are summed up as in Equation (26):

$$C'_D(n_i) = d_{in}(n_i) + d_{out}(n_i) \tag{26}$$

The higher the node centrality, the more central the point is, indicating that the actor corresponding to the point is a key subject (opinion leader) in this network.

3.5 Proximity to center

Unlike node centrality, which only considers direct relationships, proximity centrality considers indirect relationships between nodes. Proximity centrality measures the closeness or distance between nodes in a network. The proximity centrality C_c of a node n_i is calculated by summing the short-range line distance $d(n_i, n_j)$ between the node and each other node n_j, which is given by equation (27):

$$C_c = \left[\sum_{j=1}^{N} d(n_i, n_j)\right]^{-1} \tag{27}$$

If the proximity centrality is lower, it indicates that the actor corresponding to that node is more closely related to other actors. Similarly, proximity centrality for directed network graphs is the sum of inner proximity centrality and outer proximity centrality.

3.6 Degree of intermediary centrality

Intermediary centrality is a measure of a node's control over other nodes, and is used to identify the key players in a node that act as bridges. Mediacy centrality measures the extent to which a point is a “mediator” for other points in the network graph by calculating the degree of

separation between a node and other points in the network graph. Let g_{jk} denote the number of short-range lines between j and k, and $g_{jk}(n_i)$ denote the number of short-range lines between two actors containing actor n_i, then the mediator centrality is calculated as in Equation (28):

$$C_B = \sum_{j<k} g_{jk}(n_i) / g_{jk} \tag{28}$$

Higher values of intermediary centrality indicate greater control over other actors. Such nodes act as “brokers” or “gatekeepers”.

3.7 Structural Hole Analysis

Structural hole theory is used to describe a structural feature formed by an intermediary connecting several actors who are not very close. For example, for three actors A, B and C, if A is connected to B, A is connected to C, and B is not connected to C, there is a hole between B and C, and A is the “intermediary” between the two, and this structure between the three is called a structural hole. The intermediary in the structural hole has a great opportunity to obtain benefits and control benefits, and has a strong competitive advantage. In this paper, the two most important indicators in structural hole analysis are effective size and restriction system to measure the situation of structural holes. The larger the effective size is, the greater the possibility of structural holes in the network, and the lower the restriction degree is, the greater the number of structural holes.

4 Image Narrative Path of Jilin Party History Based on Image Generation

As can be seen from the above, this paper establishes an image generation model based on StyleGAN2, which realizes the image narrative of Jilin Party history by extracting the text features of Jilin Party history and the generated image features using ChineseCLIP, and designing a dual-channel feature fusion module to weaken the effect of noise on the model's understanding of the text's semantic meaning, and realizes the image narrative of Jilin Party history. The communication network effectiveness of the image narrative of Jilin Party history is analyzed in detail using social network analysis in this chapter.

4.1 Text data collection and screening

First, according to the research objectives, the type of text to be collected in this paper is mainly party history text, while in order to better access the developmental vein of the objectives, the collected text should contain enough time information and be systematic as much as possible, while the volume of the text should not be too large because of the limitation of arithmetic power.

Second, obtaining text data requires finding a suitable data source. The modern Internet itself is a treasure trove of textual data, and most of the data sources for textual data acquisition in this study come from the Chinese Internet, including scattered websites, thematic pages, libraries, literature databases, and social media.

In addition, the data sources were screened to ensure that the textual data were consistent with the objectives of the study and at the same time more easily accessible.

Finally, data were acquired from selected data sources, which in this study were mainly web

pages and literature databases, and the initial version of text data was acquired through API access and crawling.

4.2 Output Performance Evaluation of Image Generation Models

4.2.1 Quantitative analysis experiments

Six baseline models in the same domain, (A1) GAN-INT-CLS, (A2) GAWWN, (A3) StackGAN, (A4) AunGAN, (A5) MirrorGAN, and (A6) Obj-GAN are selected as a control, and compared with (A7) this paper's model on the publicly available datasets CUB and MS-COCO. The performance of Inception and FID values are shown in Table 1. Where higher Inception values are better and lower FID values are better.

Table 1: Comparison of indicators of the seven models

Model	Image size	CUB	MS-COCO		Parameter quantity (M)
		Inception	Inception	FID	
A1	64*64	2.86	7.92	-	-
A2	128*128	2.99	-	-	-
A3	256*256	3.37	8.51	-	108.26
A4	256*256	3.98	25.32	29.03	170.02
A5	256*256	3.98	26.34	-	170.16
A6	256*256	-	27.16	26.01	195.08
A7	256*256	3.82	25.69	16.94	25.91

Overall, the model in this paper outperforms most of the baseline models in terms of Inception value (3.82) on the CUB dataset, and ranks third among the seven models in terms of Inception value (25.69) and best in terms of FID value (16.94) on the MS-COCO dataset, which is comparable to the baseline model. In addition, the model scale (training parameter) of this paper is 25.91, the smallest among the seven models, indicating that it is able to generate high-quality images while maintaining a lower order of magnitude of model parameter scale, which is more practical.

4.2.2 Designing visual presentations

Still using the baseline model selected in the previous subsection as a control, the performance of the abstraction understanding accuracy of the image compared with the model of this paper is shown in Fig. 4. In this paper, the image abstraction understanding accuracy of the model is assessed by the statistical model's understanding accuracy of the image elements and the number of erroneous and omitted elements of the image, and the selected images are from the dataset MS-COCO, which contains a total of 38 elements. (A1) GAN-INT-CLS, (A2) GAWWN, and (A3) StackGAN, totaling three baseline models, have low image abstraction comprehension accuracies (<80.00%), with 10 and more miss elements. (A5) MirrorGAN model performed the best with 97.37% accuracy for image element understanding with only 1 wrongly missed element. (A4) AunGAN, (A6) Obj-GAN and (A7) the models in this paper perform better with 89.47%, 92.11% and 94.74% accuracy for image abstraction comprehension, in that order, with the number of wrong and missed elements <5.

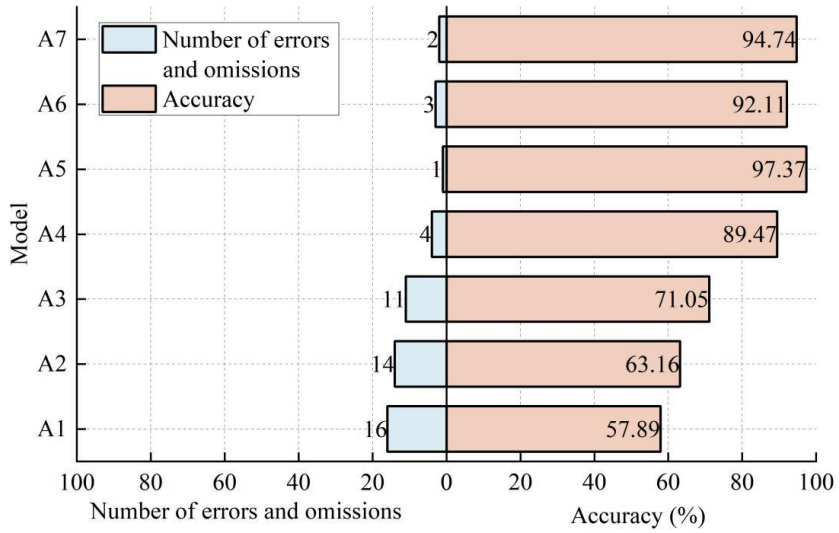

Figure 4: The accuracy of abstract understanding of the model

Color richness is used to show the color diversity of the model design works, in this paper there are 7 levels of color richness (1-7), the higher the level, the stronger the color diversity of the model design works. Based on the experiments on the accuracy of image abstract comprehension, the color diversity of the seven models based on image elements is compared in Fig. 5, in which the color richness of A1-A4 models is 5 or below, which is not in line with the actual demand, and the color richness of A5-A7 models is 7, 6 and 6 in order, which is almost the same level, with excellent color diversity performance.

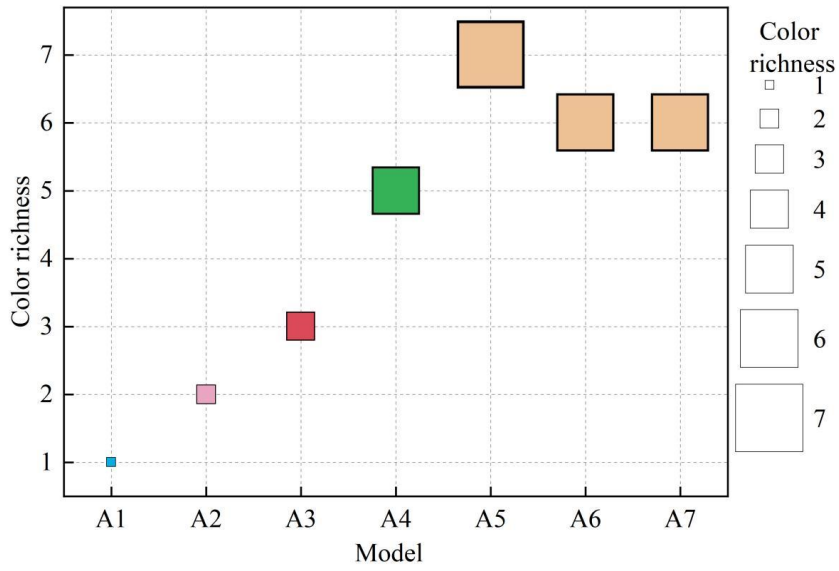

Figure 5: The color diversity contrast of the model

4.3 Dissemination Networks of Jilin Party History Image Narratives

4.3.1 Construction of communication networks

Based on the image generation model of this paper, the image narrative of Jilin Party history is carried out on the screened and collected text data. Taking Jilin Party history as the center node of the network and image narrative as the external node of the network, the communication network structure of Jilin Party history image narrative is shown in Fig. 6.The network shows

a cluster structure centered on Jilin Party history, with an average centrality of 450 for Jilin Party history, an average centrality of 1.052 for image narrative, and an average centrality of 1.056 for the whole node.The diameter of the network, i.e. the maximum value of the distance between any two nodes of the network, is 8, indicating that the network is more tightly connected. The maximum value of the distance between any two nodes in the network is 8, indicating that the network has a relatively tight connection status. The graph density is 1.42, the modularity coefficient is 0.931, and it contains 102 communities, indicating that the number of edges in the group exceeds the expected number based on chance, and the network modularity tends to be sharp, with a clear structure of network associations centered on the history of the Jilin Party. The number of strongly connected components is 65,280 and the number of weakly connected components is only 25, and the nodes in this network are tightly connected. The average clustering coefficient of the network is 1.32, and the average eigenvector centrality is 2.24. The average path length, i.e., the average value of the distance between any two nodes, is 1.81, which indicates that the nodes send information to each other with high efficiency, and that the dissemination of Jilin Party history, which takes image narrative as the main dissemination method, is more efficient.

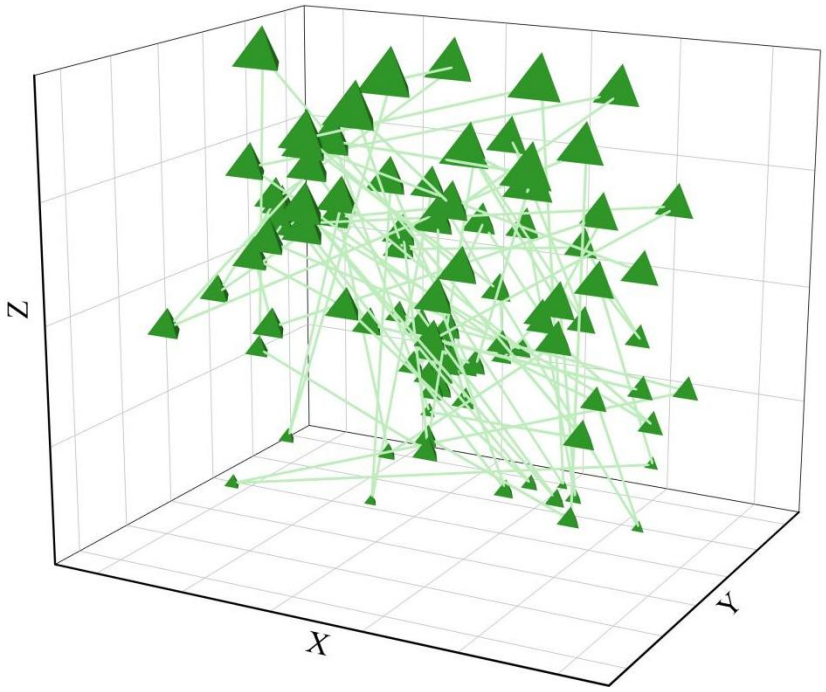

Figure 6: Disseminate the social network structure

4.3.2 Node centrality

In the node centrality degree, the in-degree indicates the degree of nodes in the network being visited by other nodes, reflecting their popularity in the overall network, while the out-degree indicates the access to other nodes in the nodes of the network, reflecting the degree of the node's initiative in obtaining the resources it needs. The top 10 nodes in the communication network structure of Jilin Party history image narrative are shown in Table 2, and the top 10 nodes in the out degree are shown in Table 3.

There are nodes “Yang Jingyu”, “Wei Zhengmin” and “Sibao Linjiang” with an in-degree of 10 and above, indicating that they are the most popular nodes in the network, and are in an important position in the network of Jilin Party history image narratives. It is in an important position in the Jilin Party history image narrative network. The three content nodes of “Red

Jilin", "Red Resources" and "Red Heritage" have an out-degree of 900 and above, indicating that they can actively obtain the required resources from other nodes in the network.

Table 2: Disseminate the top 10 nodes in terms of in-degree in the social network

Serial number	Content	In-degree
1	Yang Jingyu	23
2	Wei Zhengmin	18
3	Linjiang Campaign	15
4	Sun Guizhi	9
5	Shen Liangliang	8
6	Chen Yun	6
7	Ma Jun	5
8	Chen Ge	4
9	Wang Qing	3
10	Anti-japanese United Army Route	1

Table 3: Disseminate the top 10 nodes in terms of out-degree in the social network

Serial number	Content	Out-degree
1	Red Jilin	1116
2	Red resources	1013
3	Red Legacy in China	909
4	Red Revolution	507
5	Soldiers of the anti-Japanese United Army	400
6	Resistance base	299
7	Revolutionary relics	293
8	The Liaoning-Shenyang Campaign	186
9	Revolutionary martyr	135
10	Red spirit	82

4.3.3 Degree of intermediary centrality

As can be seen from the content of the third chapter section, the mediating centrality degree is mainly used to find the key subjects that play the role of communication bridge in the network nodes. That is, the more positions a node has in the network, the higher its degree of association with the image narrative of Jilin Party history and the stronger its mediational centrality. The top 10 nodes of mediational centrality in the communication network of Jilin Party history image narrative are summarized in Table 4. The mediational centrality of the three nodes "red resources", "anti-Japanese base" and "Liao-Shen Battle" is as high as 25,000 and above, which indicates that the dissemination of the image narratives of Jilin Party history is mainly transmitted through these three nodes directly or indirectly. This indicates that the whole Jilin Party history image narrative dissemination network mainly passes the narrative content directly or indirectly through these three nodes.

Table 4: The top 10 nodes in terms of network intermediary centrality

Serial number	Content	Intermediary centrality
1	Red resources	76683
2	Resistance base	36388
3	The Liaoning-Shenyang Campaign	25883
4	Yang Jingyu	15239
5	Red spirit	15042
6	Linjiang Campaign	14731
7	Patriotism	9699
8	Red Revolution	4967
9	The secret camp of the Northeast anti-Japanese United Army	2926
10	The Qidaojiang Conference	1073

4.3.4 Centrality index extraction and network structure analysis

Analyzing the communication network structure of the constructed image narrative of Jilin Party history, the main centrality indexes of the nodes in the network are obtained, and the degree centrality, proximity centrality, intermediary centrality, and feature centrality of each index are summarized in Table 5.There are 15 centrality indexes as follows: Red Jilin, Red Resources, Red Heritage, Yang Jingyu, Resistance Generals, Anti-Japanese Bases, Revolutionary Cultural Relics, Liao-Shen Battle, and Revolutionary Martyrs, Red Spirit, Sibao Linjiang, Resistance Route, Patriotism, Wei Zhengmin, and People's Aviation, all 15 centrality indicators have degree centrality >15, proximity centrality within the interval (17,18), mediator centrality >0.00, and feature centrality 15.00 and above. The metrics of the node lexical items in this stage of the indicators side by side reveal that the core content of Jilin party history communication lies in its red spirit.

Table 5: Multiple centrality indicators

Sort	Content	Degree	Closen	Betwee	Eigenv
1	Red Jilin	37.981	19.46	27.011	36.201
2	Red resources	88.381	19.438	25.8	35.998
3	Red Legacy in China	73.061	19.178	7.97	32.962
4	Yang Jingyu	38.352	19.068	7.652	30.788
5	Soldiers of the anti-Japanese United Army	64.114	18.842	4.256	30.85
6	Resistance base	36.265	18.643	2.377	24.82
7	Revolutionary relics	80.181	18.426	2.416	28.108
8	The Liaoning-Shenyang Campaign	43.849	18.335	1.659	23.404
9	Revolutionary martyr	82.478	18.278	1.399	24.957
10	Red spirit	60.961	17.818	1.476	23.672
11	Linjiang Campaign	42.136	17.573	1.732	24.624
12	Anti-japanese United Army Route	48.288	17.455	0.287	17.475
13	Patriotism	84.216	17.264	0.298	17.775
14	Wei Zhengmin	15.834	17.237	0.403	19.554
15	People's aviation industry	20.244	17.182	0.25	15.426

5 Conclusion

In this paper, the cross-modal alignment advantages of the StyleGAN2 image generation model and the ChineseCLIP model are fused to establish a StyleGAN2-based image generation model, which successfully realizes text-guided image generation. The special feature of this model is that it adopts a two-channel feature fusion module to weaken the influence of noise on the model's understanding of text meaning, and guides the generation of a robust representation of these features, which improves the semantic similarity between the text description and the generated image.

The overall performance of the StyleGAN2-based image generation model is not inferior to the better baseline models in the same field, with an Inception value and FID value of up to 3.82 and 16.94, respectively; the accuracy of the image abstraction understanding is 94.74%, with only two wrongly omitted image elements, and the color richness of the generated image is 6, which is an excellent performance in terms of color diversity.

After the image generation model based on StyleGAN2 is used to realize the image narrative of Jilin party history, its propagation network is constructed. The diameter of the propagation network is 8, the graph density is 1.42, the modularity coefficient is 0.931, the network modularity trend is distinct, and the network association structure centered on Jilin Party history is very obvious. The average path length is 1.81, and the overall dissemination efficiency is high. Through the social network analysis method to screen the indicators with high node centrality and intermediary centrality in the network, 15 main centrality indicators are obtained, with indicator degree centrality >15, proximity centrality in the (17,18) interval, intermediary centrality >0.00, and feature centrality >15.00, which excavate the core content of the Jilin Party history topic in the image narrative.

The AIGC technology empowers the image narrative content production of Jilin Party history topics, promotes the enhancement of its overall dissemination breadth and depth, and creatively explores the development path of Jilin Party history under the intelligent workflow.

Funding

This work was supported by the Humanities and Social Sciences Research Project of Jilin Provincial Department of Education, titled: "Research on the Creation and Dissemination of Image Narratives on Jilin's Party History Themes Using AIGC Technology" (JJKH20251514SK).

About the Author

Yizhong Zhang was born in Jilin, Jilin, P.R. China, in 1990. I obtained a bachelor's degree from Beihua University and a master's degree from the Venice Academy of Fine Arts in Italy. I am currently pursuing a PhD in Philosophy at the Faculty of Creative Technology and Heritage, Universiti Malaysia Kelantan. My main research direction is in the fields of technology and art. I work at Jilin Animation College in Changchun.

References

[1] Li, K., & Huang, Y. (2025). Research on the Construction of Northeast Anti-Japanese United Front National Cultural Park: Cultural Features, Resource Integration, and Spatial

Construction. Academic Journal of Humanities & Social Sciences, 8(8), 46-53.

[2] Zhao, J., & Geng, S. (2021). A Centennial Journey of Putting People First-" The Communist Party of China and the Progress on Human Rights in China" International Conference Summary. J. Hum. Rts., 20, 515.

[3] Chen, T., & Kung, J. K. S. (2024). The rise of the Chinese communist party. Available at SSRN 3748521.

[4] CHE, E., & ZHANG, P. (2022). Research on the hundred-year struggle process and historical experience of the Communist Party of China. RESEARCHES IN HIGHER EDUCATION OF PHARMACY, 40(3), 1

[5] Hongxian, L., Tahir, A., Bakar, S. A. S. A., & Jianan, L. (2025). Expression of Spiritual Connotation of Chinese Revolutionary Historical-Themed Oil Paintings in the 1970s. Environment-Behaviour Proceedings Journal, 10(SI29), 45-51.

[6] Weiping, Q. (2022). A Review of the Historical Narrative of the Century-old Party: Based on the History of the Communist Party of China. Teaching and Research, 56(6), 9.

[7] Fasnacht, H. (2023). The narrative characteristics of images. British Journal of Aesthetics, 63(1), 1-23.

[8] Shen, Y., & Biberman, E. (2010). A story told by a picture. Image & Narrative, 11(2), 177-197.

[9] Leslie, R. (2018). The use of victor-victim historical narratives in Chinese nationalist discourse. Cornell International Affairs Review, 11(2).

[10] Hellmann, O. (2021). The dictator's screenplay: collective memory narratives and the legitimacy of communist rule in East Asia. Democratization, 28(4), 659-683.

[11] Wang, L., & Zhao, X. (2025). Literature Review on the Image of the Communist Party of China in the Past 30 Years: Based on the Bibliometric Analysis of CiteSpace Software. China Report, 61(2), 270-291.

[12] Brown, K. (2018). The communist party and ideology. InThe SAGE Handbook of Contemporary China, edited byWeiping Wu and Mark W. Frazier, 287-301.

[13] Frenkiel, E., & Shpakovskaya, A. (2019). The evolution of representative claim-making by the Chinese communist party: From Mao to Xi (1949-2019). Politics and Governance, 7(3), 208-219.

[14] Jacob, J. T., & Subba, B. B. (2022). Towards exceptionalism: The communist party of China and its uses of history. China Report, 58(1), 7-27.

[15] Zhang, X., Brown, M. S., & O'Brien, D. (2018). "No CCP, no new China": Pastoral power in official narratives in China. The China Quarterly, 235, 784-803.

[16] Brødsgaard, K. E., & Chen, G. (2018). The Chinese Communist Party since 1949: Organization, ideology, and prospect for change. Brill Research Perspectives in

Governance and Public Policy in China, 3(1-2), 1-60.

[17] Lv, L. (2017, December). The Opportunity of Carrying Forward and Cultivating Northeast Anti-Japanese United Army Spirit under the New Situation. In 4th International Conference on Education, Language, Art and Intercultural Communication (ICELAIC 2017) (pp. 191-194). Atlantis Press.

[18] Xu, C., Guo, J., Zeng, J., Meng, S., Chu, X., Cao, J., & Wang, T. (2024, July). Enhancing AI-Generated Content Efficiency Through Adaptive Multi-Edge Collaboration. In 2024 IEEE 44th International Conference on Distributed Computing Systems (ICDCS) (pp. 960-970). IEEE.

[19] Hu, Z. (2025). A method for generating personalized learning content based on AIGC. Sustainable Futures, 10, 101331.

[20] Gao, A. (2024). From PGC to UGC to AIGC: Change of content paradigm. In SHS Web of Conferences (Vol. 199, p. 03017). EDP Sciences.

[21] Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P., & Sun, L. (2025). A survey of ai-generated content (aigc). ACM Computing Surveys, 57(5), 1-38.

[22] Zi-yang, H. U. (2024). AIGC related context: A new communication culture for human. Journal of Literature and Art Studies, 14(10), 921-931.

[23] Fan, N., Li, X., Liu, C., & Fan, Z. P. (2025). The Power of AI-Generated Content: Evidence From the Peer-to-Peer Accommodation Market. Journal of Travel Research, 00472875251332951.

[24] Xu, T., & Li, W. (2024, June). Research on the Collection of Historical Materials of Northeast Counter-Japanese United Army under the Background of Big Data. In 2024 3rd International Conference on Social Sciences and Humanities and Arts (SSHA 2024) (pp. 209-217). Atlantis Press.

[25] Zeng, H. (2025). Research on the Dissemination Effect of AIGC-Generated Content on New Media Platforms—Taking Gen Z as an Example. International Journal of Asian Social Science Research, 2(4), 24-29.

[26] Mu, Y. (2024). Research on the Dissemination Effect of AIGC-generated Content on New Media Platforms: A Case Study of Xiaohongshu. Media Studies, 1(2), 1-5.

[27] Li, X. (2025). Application and Strategy of AIGC in Omnimedia Knowledge Dissemination Under the Background of Digital Empowerment. International Journal of Knowledge Management (IJKM), 21(1), 1-16.