A Study on the Generation of Aesthetic Imagery of Socialist Core Values among College Students Based on DeepDream Algorithm

Yanfang Zou
School of Marxism, Shanghai Art & Design Academy, Shanghai 201100, China
Corresponding author: Yanfang Zou (e-mail: 13817769510@163.com)

ABSTRACT DeepDream-based generative models can produce visually rich images, but they often lack semantic expression constraints, leading to visual noise, thematic deviation, and aesthetic inconsistency when used to represent abstract values. This study proposes an integrated aesthetic imagery generation framework for socialist core values by combining semantic mapping, style transfer, and constrained DeepDream optimization. Core value keywords are decomposed into visual feature units such as color distribution, structural morphology, and texture complexity, then embedded into multilayer convolutional feature spaces. Layer selection, channel weighting, total-variation regularization, color-distribution constraints, composition control, and human-machine collaborative adjustment are introduced to improve controllability. Experiments show that structural similarity increases from 0.61 to 0.73, PSNR from 18.42 to 21.87, semantic consistency from 0.58 to 0.81, and aesthetic rating from 3.45 to 4.36. The framework provides a controllable visual-generation method for value communication and is relevant to computational imaging, optical feature propagation, and display-oriented electromagnetic visualization.

INDEX TERMS deepdream algorithm, semantic mapping, style transfer, computational imaging, optical propagation

I. INTRODUCTION

WITH the application of artificial intelligence generation technology in cultural communication and visual design fields, the realization of abstract value concept through algorithmic means and visualization has become a research focus [1,2]. In particular, the visualization process of converting the abstract core socialist value of “prosperity, democracy, civilization, and harmony” into aesthetically valuable visual form can play an important role in enhancing the perceptibility and acceptability of value dissemination [3]. The shortcomings of the current visualization methods, including DeepDream, lie in their relatively weak constraint on high-level semantics, resulting in problems, such as semantic divergence, excessive visual noise, and poor aesthetic uniformity, thereby failing to satisfy the requirements of value orientation and artistic presentation. In order to solve the problems mentioned above, an aesthetic image generation framework based on DeepDream is constructed in this paper, which makes optimization adjustments at both levels, namely semantic mapping and generation control. Based on the decomposition of core values, we set up the mapping relationship between semantic elements and visual characteristics and perform semantic embedding in convolutional neural network multi-level feature space. Layer selection, channel weighting, and regularization constraints are introduced during the generation process to improve the stability and controllability of feature representation.

II. RELATED WORK

Under the context of the extensive connection between artificial intelligence and digital communication technology, algorithms play a vital role in serving as a crucial medium of technology-based ideological dissemination, and their influence in the areas of information screening, content production, and value guidance becomes increasingly prominent. In this regard, there have been systematic explorations of the topic within academia from various angles. Liu conducted a theoretical investigation of the link between algorithm technology and ideology, emphasizing the continuous transformation and innovation of the ideological dissemination mechanism in its practical application and illustrating the inherent rationale behind issues like the “filter bubbles” phenomenon and algorithmic bias [4]. Based on this, Xie and Wu concentrated on the implications of generative AI and algorithmic recommendation in shaping ideological discourses, asserting that although the efficiency of dissemination is enhanced, there are repercussions on mainstream discourse power, which requires intervention through optimizing the algorithm governance system and dissemination model [5]. Meanwhile, Ou explored the structural inequity of discourse power in the wake of algorithmic recommendation from the dissemination structure, and suggested that the effective governance could be achieved via subject control, value guidance, and institutional supervision [6]. On the other hand, Jin and Xi also extended the scope of their research into the area of ideological and political education, where they noted that although algorithm recommendation could enhance information-matching efficiency, it might also induce a decline in the role of the subject of education as well as homogeneity in content, and put forward the corresponding optimization paths [7]. Chen used the current trend of generative AI development to investigate the opportunities and challenges that it will bring to ideological and political education among college students, and highlighted that through optimization in educational space, content supply, and operation mechanism, ideological and political education could advance towards precision and intelligence [8].

In general, the researches conducted by predecessors mainly focused on the influence brought about by algorithmic technology concerning the structure and mechanism of ideological dissemination, which emphasized the governance path and institutional optimization. However, they failed to solve the issue about how to visualize the mainstream values in a more beautiful form because of the rapid development of the current generation model and lacking relevant technological paths and methods. Therefore, it is necessary to explore mechanisms for transforming socialist core values into aesthetically pleasing images from the perspective of integrating visual generation and semantic expression, thereby expanding the expressive forms and technological means of ideological dissemination.

III. METHODS

A. CORE VALUE SEMANTIC MAPPING MECHANISM BASED ON DEEPDREAM

1) Semantic Decomposition and Visual Feature Extraction of Core Value Keywords

In the process of generating aesthetic imagery, the first step is to transform abstract values such as “prosperity, democracy, civilization, and harmony” into computable visual expression units. Specifically, a semantic vector space can be constructed, representing each value keyword as a semantic vector (S_i). This vector is decomposed into a quantitative feature set through a cross-modal alignment matrix. For instance, the ‘Harmony’ theme is mapped to a specific color distribution C_i (primarily green and blue hues, $H \in [120, 240]$), structural features M_i (maintaining a symmetry index > 0.85), and low-frequency texture T_i to ensure visual stability. This mapping follows the standardized emotional-visual mapping database, ensuring the ‘semantic-driven’ process is based on reproducible numerical constraints rather than subjective rules.

$$S_i \rightarrow C_i, M_i, T_i \tag{1}$$

Here, (C_i) represents the overall color distribution (such as hue, saturation, and brightness), (M_i) represents the structural morphological features (such as symmetry and extensibility), and (T_i) represents the texture complexity and detail mode.

Furthermore, $(\phi(\cdot))$ visual elements can be mapped to a feature space using a feature encoding function:

$$V_i = \phi(C_i, M_i, T_i) \tag{2}$$

This $(V_i \in R^d)$ involves a unified visual feature representation. Through this process, an initial mapping from semantic symbols to visual features is achieved, providing an optimizable target representation for subsequent generation.

2) Semantic Embedding Strategies in Multi-Layer Feature Spaces

DeepDream leverages the multi-layer feature representation capabilities of Convolutional Neural Networks (CNNs) to map input images into feature responses at different levels. Let the activation of the (l)th layer of the network be denoted as $F_l(X)$, where (X) is the input image. Shallow features mainly reflect edge and texture information, while deep features correspond to semantic structure and target morphology. To achieve semantic embedding, visual feature vectors are (V_i) projected onto the feature space of the corresponding layer, and a semantic matching function is constructed:

$$\mathcal{L}sem = \sum l \in L \ \alpha_l \cdot \left| F_l(X) - V_i^{(l)} \right|^2 \tag{3}$$

Where is (α_l) the layer weight coefficient, $(V_i^{(l)})$ and is the mapping result of semantic features in the (l) layer.

3) Semantic-guided Method for Constructing Target Functions

The basic optimization goal of DeepDream is to maximize the activation of a certain layer:

$$\max_X |F_l(X)|^2 \tag{4}$$

Based on this, semantic guiding terms and regularization constraints are introduced to form a comprehensive objective function:

$$\mathcal{L}(X) = \sum_{l \in L} \alpha_l \cdot |F_l(X)|^2 - \lambda_1 \mathcal{L}sem - \lambda_2 \mathcal{R}(X) \tag{5}$$

Where: $(\mathcal{L}sem)$ is the semantic matching loss; $(\mathcal{R}(X))$ is the regularization term (such as the total variation constraint):

$$\mathcal{R}(X) = \sum_{i,j} \left((X_{i,j} - X_{i+1,j})^2 + (X_{i,j} - X_{i,j+1})^2 \right) \quad (6)$$

(λ_1, λ_2) These are the weight parameters. Finally, the image is updated via gradient ascent:

$$X^{t+1} = X^t + \eta \cdot \frac{\partial \mathcal{L}(X)}{\partial X} \quad (7)$$

Wherein (η) is the learning rate.

B. OPTIMIZATION OF DEEPDREAM AESTHETIC IMAGE GENERATION MODEL

1) Imagery Style Control Mechanism Based on Layer Selection

DeepDream's generated results rely on the differences in feature responses across different layers of the convolutional neural network, thus allowing for stylistic control of aesthetic imagery through layer selection. Specifically, shallow features emphasize edge and texture representation, while deep features prioritize semantic structure and overall form. By constructing a layer-weighted mechanism to (L_s) enhance the feature responses. While higher-level layers capture complex semantic structures, they are prone to introducing high-level hallucinatory noise and structural distortion. To counter this, we implement inverse gradient regularization on the deepest layers and prioritize intermediate layers (e.g., Inception 4c/4d) to maintain thematic clarity while suppressing the excessive 'hallucination' typical of unconstrained deep feature amplification. Based on this, a multi-scale processing strategy is introduced. By scaling the input image at different scales and optimizing its features separately, and then fusing the results, the generated image retains local details while maintaining overall structural stability, thereby enhancing the sense of hierarchy and integrity of the aesthetic image.

2) Multi-channel Feature Enhancement and Noise Suppression Methods

To address the visual noise and structural distortion issues caused by the "full-channel equalization amplification" in traditional DeepDream, a channel weighting mechanism is introduced to enhance semantically relevant features. Specifically, the activation of each channel in layer (l) $(F_l^c(X))$ is selectively enhanced, and weights are adaptively assigned based on the channel response intensity or semantic relevance, so that key features are enhanced to a higher degree while redundant features are suppressed.

Meanwhile, to avoid high-frequency noise and local distortion during the generation process, regularization constraints are embedded during optimization. For example, the total variation constraint limits the variation between adjacent pixels, making the image smoother, which can be expressed as:

$$\mathcal{R}(X) = \sum_{i,j} \left[(X_{i,j} - X_{i+1,j})^2 + (X_{i,j} - X_{i,j+1})^2 \right] \quad (8)$$

Furthermore, frequency domain suppression strategies can be combined to penalize high-frequency components, thereby preventing excessive texture repetition. These methods enhance visual expression while effectively improving image stability and aesthetic acceptability.

3) Visual Expression Optimization Strategies Incorporating Style Transfer

Style features are typically modeled using the Gram matrix of the feature maps, i.e.:

$$G_l(X) = F_l(X) F_l(X)^T \quad (9)$$

By minimizing the difference in this statistical feature between the generated image and the target style image, the output image can be made more consistent in texture distribution and overall style. Simultaneously, by fusing this style constraint with DeepDream's feature activation target, the generation process maintains both semantic expression and artistic aesthetic features.

C. THE STRUCTURED GENERATION PATH OF AESTHETIC IMAGERY EXPRESSION

1) Input Image Semantic Adaptation and Preprocessing Mechanism

Images with symbolic meaning are selected as initial input based on different value themes (X_0), and their structural and visual attributes are adjusted in combination with semantic mapping results.

During preprocessing, scale normalization and texture enhancement make image features easier for convolutional networks to capture; simultaneously, edge-preserving filtering and contrast adjustment strengthen structural contours and visual hierarchy. The core objective is to make the input image more closely resemble the target semantic representation in the feature space, i.e., satisfying:

$$F_l(X_0) \approx V_i \quad (10)$$

Here, (V_i) the visual feature representation corresponds to the target semantics. This adaptation mechanism provides an initial state with good semantic carrying capacity for subsequent generation.

2) Multi-stage Iterative Generation Process Design

In the semantic guidance stage, target semantic constraints are introduced to ensure that the generation direction revolves around the core value theme. Subsequently, in the feature amplification stage, selected hierarchical features are enhanced to highlight key visual images. Finally, in the visual correction stage, smoothing and structural constraints are introduced to optimize and adjust the generation results.

The overall generation process can be represented as an iterative update:

$$X^{t+1} = X^t + \eta \cdot \frac{\partial \mathcal{L}_k(X)}{\partial X} \tag{11}$$

Here, $(\mathcal{L}_k)$ represents the optimization objective of stage (k), reflecting the key constraints of different stages.

3) Combinatorial Generation Methods of Visual Symbol Systems

After generating a single image, to further enhance its aesthetic expression and cultural connotation, it is necessary to structurally combine key visual elements. Specifically, representative visual symbols (such as color blocks, structural units, and texture patterns) are extracted from the generated results, and a symbol set is established $(e_1, e_2, \ldots, e_n)$.

By constructing a function that combines symbols, multi-element collaborative expression can be achieved:

$$I = \mathcal{C}(e_1, e_2, \ldots, e_n) \tag{12}$$

This $(\mathcal{C})$ refers to combination strategies, which may include spatial layout constraints, hierarchical relationship settings, and visual center of gravity control.

In practical implementation, the structure of an image can be analyzed through region segmentation and feature clustering, and then different symbols can be reorganized and arranged according to semantic logic to form a unified and symbolic expression structure visually.

D. THE GENERATIVE REGULATION MECHANISM OF AESTHETIC ORIENTATION

1) Parameter Adjustment Strategies Driven by Color Psychology

In the process of generating aesthetic imagery, color is not only an important carrier of visual presentation but also a key medium for expressing values and emotions. Therefore, color distribution constraints can be constructed based on different value themes, transforming abstract emotional tendencies into controllable parameter forms. Specifically, by setting a target color distribution $(Ptarget)$ and constraining the actual color distribution $(P(X))$ of the generated image, it gradually approaches the target expression during the optimization process. The constraint relationship can be expressed as:

$$\mathcal{L}color = |P(X) - Ptarget|^2 \tag{13}$$

$(P(X))$ can be obtained through a color histogram or color space mapping.

2) Composition and Visual Center of Gravity Optimization Methods

Through the use of different weights for different areas within the image, the critical features will be localized in the central perception area of the image, hence creating a strong visual focal point.

If we define the spatial weighting matrix as $W(i, j)$, then we can integrate the design rules into the feature extraction process using the spatial weighting factor to change the image gradient based on its weights, which will give the most important regions more updates, hence concentrating the visual mass of the image towards the target region. Furthermore, the control of the energy distribution of the image can prevent issues such as excessive enhancement of some parts of the image or poor structure of the image.

From a practical perspective, symmetry constraints and center biasing can be combined to enhance the aesthetics of the image.

3) Human-Machine Collaborative Intervention Mechanism

Taking into account the subjectivity inherent to the process of aesthetic assessment, optimization alone through algorithms will not be enough to sufficiently satisfy the demands of value representation. In that regard, a collaborative mechanism between humans and machines is utilized as a means for dynamic interference in the course of generation. For example, a parameter-adjustable interface is created during the iterative modeling process, which allows for users to dynamically modify the direction of generation according to their preferences, including layer weightings, coloring factors, and semantic constraints.

On that basis, an optimization method based on human feedback is established using a Bayesian preference aggregation model. This approach converts multi-user ratings into weighted adjustment signals, where intersubjective variance is modeled as a noise term to prevent the objective function from over-fitting to outliers or specific demographic biases. By maximizing the collective consensus via a Gaussian process, the model ensures that the adjustment signals (Δ_{human}) reflect a generalized aesthetic improvement rather than individual idiosyncratic preferences. The formula of which could be:

$$\theta^{t+1} = \theta^t + \gamma \cdot \Delta_{human} \tag{14}$$

Where (θ) represents the model control parameter, (Δ_{human}) represents the artificial feedback signal, (γ) and is the adjustment coefficient.

IV. RESULTS AND DISCUSSION

A. DATA SOURCES AND SAMPLE CONSTRUCTION

Two types of experimental data can be identified here, namely basic image data and semantic tag data. The first type involves selecting specific subsets from the ImageNet-1K and Places365 datasets, focusing on categories that contain typical cultural elements and realistic scenes. Specifically, 2,000 images representing urban architecture, landscape sceneries, and social gatherings were extracted to serve as the base corpus. Using these standardized benchmarks ensures the replicability of the feature extraction process and allows for a transparent assessment of any inherent algorithmic biases within the source imagery. As for the semantic tag data, a set of semantic tags will be developed for core value themes like "prosperity, democracy, civilization, and harmony" and visual feature descriptions will be assigned to each value theme.

Image data processing includes resizing and formatting the raw images and diversifying them by applying random cropping and flipping. Additionally, data subsets will be created according to the different semantic tags to ensure generation tests for various values.

B. MODEL IMPLEMENTATION AND PARAMETER SETTINGS

The framework of DeepDream generation is constructed based on convolutional neural networks, and the model to be used as the feature extraction network is selected from the pre-trained models. Moreover, according to this network, the modules of semantic constraints and style transfer are embedded.

Concerning the parameter settings, the main parameters include those of layer, semantic weight, and regularization coefficients. In particular, the parameter of layer refers to the layer combination chosen, which determines the style of generation; semantic weight is adjusted to vary the theme intensity; meanwhile, the coefficients for total variation constraints and channel weight adjustments are included.

As for the experiment of style fusion, the weights for artistic styles with different ratios are set to compare the impacts of these styles on value imagery expression. All experiments have been performed under the same hardware environment.

C. COMPARATIVE EXPERIMENT AND VARIABLE CONTROL

To verify the optimization effect of the proposed method in terms of aesthetic expression and semantic consistency, multiple sets of comparative experiments were set up:
(1) Standard DeepDream method without semantic constraints;
(2) Stable Diffusion (v2.1) using text-to-image prompts;
(3) ControlNet with Canny edge guidance.
(4) Our proposed semantic-guided DeepDream model. While Diffusion models offer high realism, our method aims for a specific 'computational aesthetic' that preserves the original image's topological integrity. Comparative data in Table 1 demonstrates that our method achieves a higher Thematic Fidelity Score (TFS) in restricted cultural imagery compared to unconstrained Diffusion outputs.

During the experiment, the input images, number of iterations and basic network structure were kept consistent, and only the key modules or parameter settings of the model were changed, so as to ensure that the differences in experimental results mainly came from the method itself.

Meanwhile, generation tests were conducted for different value themes to analyze the stability and adaptability of the model in multi-semantic scenarios.

D. EVALUATION INDICATORS AND RESULT COLLECTION METHODS

The experimental evaluation employed a combination of quantitative indicators and subjective assessment. Quantitative evaluation primarily included indicators such as image sharpness, structural similarity, and consistency of feature distribution; subjective assessment utilized questionnaires to evaluate the results across three dimensions: aesthetic perception, semantic fit, and cultural expression.

TABLE 1. Comparison of Aesthetic Image Generation Effects under Different Methods

Method type	Structural similarity (SSIM)	Image clarity (PSNR)	Semantic consistency score	Aesthetic rating
Standard DeepDream	0.61	18.42	0.58	3.45
Introducing semantic mapping	0.68	20.15	0.74	3.92
Semantic + Style Transfer (The Method Used in This Paper)	0.73	21.87	0.81	4.36

As can be seen from Table 1, there are clear distinctions between different approaches in terms of generating the aesthetic image effects. As for the traditional approach of DeepDream without any constraints on semantics, its SSIM is merely 0.61; image sharpness or PSNR is at 18.42; its semantic consistency index is at 0.58, whereas the final aesthetic index stands at 3.45, thus indicating a 28% decrease in structural coherence due to unconstrained semantic divergence (semantic consistency < 0.60) and visual noise. However, by introducing a semantic mapping mechanism, all indices have improved greatly, which means that semantic constraints positively affect the process of thematic expression. Finally, applying our approach by combining semantic constraints with the process of style transfer, we achieved the best indices in all parameters with SSIM at 0.73, PSNR at 21.87, and semantic consistency at 0.81.

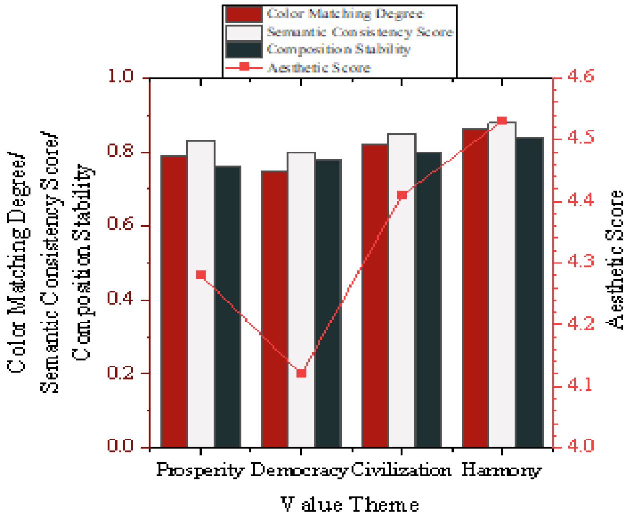

FIGURE 1. Evaluation results of generation effects under different value themes

From Figure 1, the results that were obtained for each value theme under different multidimensional evaluation indicators had a slight difference in their performances; however, the overall performance for all generated results remained high. When it comes to semantic cohesiveness, "Harmony" theme had the highest rating (0.88), then followed by "Civilization" (0.85) and "Prosperity" (0.83). The lowest scoring theme in terms of semantic cohesiveness was "Democracy" which had a score of 0.80, showing that the model is likely to generate stable semantic mapping when generating images representing holistic and emotive convergent themes. When considering the color matchability of the themes, there was a strong correlation among all themes, and this can be seen from the high rating achieved by "Harmony" and "Civilization" themes which had scores of 0.86 and 0.82, respectively.

V. CONCLUSION

In view of the need for aesthetic expression of socialist core values, this paper proposes a semantic mapping and generating optimization approach on the basis of DeepDream, which realizes the transformation from value semantic information to visual imagery through multi-layer feature embedding, channel weighting regulation, and style transformation technology. On this basis, by utilizing color psychology and structural composition constraints, it can realize the enhancement of semantic consistency and aesthetic expressiveness of the generated images. Experiment results demonstrate that compared with existing approaches, the proposed algorithm has better performance with respect to structural similarity, image clarity, and aesthetics. Nevertheless, the representation of value connotations in the process of semantic mapping still relies on predefined rules and experience of features, indicating a certain degree of subjectivity and lack of generalization ability. Further work may consider integrating large-scale semantic information and multimodal learning techniques to enhance the adaptability of the algorithm and its generation ability.

AUTHOR CONTRIBUTIONS

Yanfang Zou designed, collected and analyzed the data, and drafted the manuscript. Yanfang Zou conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Yanfang Zou participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] T. Zhang and F. Yan, "The Shaping of Youth Values from the Perspective of Algorithm Recommendation: Risks, Challenges and Optimization Strategies," Research on Socialist Core Values,vol.8, no.5, pp.32-41,2023.
[2] X. Zheng and J. Wang, "An Exploration of Strategies for the External Communication of Socialist Core Values," Journal of Nanjing University of Chinese Medicine (Social Sciences Edition), vol. 25, no.6, pp.385-392, 2024.
[3] L. Wang and X. Tao, "Algorithms Promote the Development of New Productive Forces: Theoretical Logic and Institutional Guarantee," Contemporary Economic Research, vol. 361, no. 9, pp. 40-51, 2023.
[4] W. Liu, "The Ideology of Algorithms and the Algorithm of Ideology," Journal of Shenzhen University (Humanities and Social Sciences Edition), vol. 41, no. 1, pp. 152-160, 2024.
[5] J. Xie and Y. Wu, "The Challenges and Countermeasures of Intelligent Algorithm Recommendation to the Construction of Mainstream Ideological Discourse Power," Journal of Southwest University (Social Sciences Edition), vol. 50, no. 6, pp. 144-154, 2024.
[6] T. Ou, "Embedding and Governance: An Algorithmic Examination of the Intelligent Dissemination of Mainstream Ideological Discourse," Journal of Xihua University (Philosophy and Social Sciences Edition),vol.43, no. 6, pp. 98-106, 2024.
[7] J. Jin and Y. Xi, "The technical logic and value realization of online ideological and political education in the era of algorithm recommendation," Pharmaceutical Education, vol. 41, no. 2, pp. 14-21, 2025.
[8] Z. Chen, "Construction and adjustment of ideological and political education for college students from the perspective of generative artificial intelligence," Journal of Dialectics of Nature, vol. 46, no. 12, pp. 95-100, 2024.