

# AECO-SC StyleGAN: A Cross-Platform GAN Framework for Dynamic Advertising Creative Generation

Yanan Zhang

Yantai University of Science and Technology, School of Culture and Media, Penglai, Shandong, 265600, China

E-mail: 15753572255@163.com

**Keywords:** adaptive elephant clan optimizer (AECO), dynamic advertising creative, spatially conditioned StyleGAN, visual style consistency

**Received:** April 18, 2025

*Dynamic advertising (ad) requires personalized, engaging content across multiple platforms. Traditional approaches struggle with scalability and cross-platform adaptation. Leveraging deep learning (DL), particularly Generative Adversarial Networks (GANs), offers the potential to automate and optimize ad creative generation with higher precision and contextual adaptability. This research aims to develop a DL framework that dynamically generates and optimizes advertising creatives—leveraging Adaptive Elephant Clan Optimizer with a Spatially Conditioned StyleGAN (AECO-SC StyleGAN) for dynamic cross-platform advertisement creative generation. Adaptive Elephant Clan Optimizer (AECO) dynamically adjusts training hyperparameters to improve model convergence, while Spatially Conditioned StyleGAN (SC-StyleGAN) generates platform-specific ad creatives by incorporating spatial constraints for contextual alignment. Our system is trained on the Ad ImageNet dataset, which includes 9,003 ad samples with paired images and promotional text from platforms like Facebook and Instagram. All data were resized to  $256 \times 256$ , normalized, and tokenized for training. Using Python, the model demonstrates superior performance in creative generation and engagement prediction. The proposed AECO-SC StyleGAN model achieved an NDCG of 0.61, an accuracy of 98.48%, and a weighted F1-score of 98.5%, outperforming prior approaches such as VGG + Layout + NIMA (NDCG 0.22) and XCEPTION (accuracy 98.27%, F1-score 98.2%). These results highlight the effectiveness of integrating adaptive optimization and spatial conditioning in generating high-quality, context-aware advertising creatives, offer a scalable and automated solution for cross-platform digital marketing.*

*Povzetek: Okvir AECO-SC StyleGAN uporablja globoko učenje (GAN in StyleGAN s prostorskim pogojevanjem) in adaptivno optimizacijo (AECO) za dinamično generiranje in optimizacijo visoko kakovostnih oglasnih vsebin na več platformah. Sistem dosega kvalitetno napovedovanje angažiranosti oglasov, kar avtomatizira prilagojeno digitalno trženje.*

## 1 Introduction

Digital marketing leaders consider the capacity to send customized advertising content across numerous platforms as their main competitive advantage [1]. Advertising through traditional methods faces difficulties when adjusting to user preference changes across multiple platform formats that include social media and mobile applications, together with websites [2]. DL techniques appear as groundbreaking solutions to address the marketing challenges that face the optimization of ad creativity. GANs have become prominent among these DL techniques because they enable the production of high-quality, realistic, adaptable content [3]. The procedure enables machine automation to produce relevant visual advertisements with appropriate platform parameters for distinct user groups [4]. The model uses dual-network training, which combines a generator for designing ad variants together with a discriminator for evaluation

purposes to achieve continuous output improvement through adversarial learning processes [5]. Such adaptive creativity improves advertising diversity and allows for better metrics engagement, including the click-through rate (CTR) and conversion rate (CVR).

The system processes real-time performance records in conjunction with user adjustments to automatically modify its creative elements and maintain platform usage between various platforms [6]. Digital marketing and artificial intelligence maintain an expanding relationship that gives marketers data-driven creative solutions to overcome their creative limitations [7]. This GAN-based framework provides automated design solutions for advertising content, which enable more efficient personalized advertising in an online environment dominated by competition. The aim was to develop a DL framework that utilizes the AECO-SC StyleGAN to dynamically generate, refine, and optimize advertising creatives across multiple

platforms for enhanced personalization and performance in digital marketing campaigns.

The proposed AECO algorithm is a derivative-free population-based metaheuristic. It simulates the social behavior of elephants in clans to adaptively adjust hyperparameters by exploring the solution space through stochastic position updates, without relying on the gradient of the loss function. This enables AECO to dynamically optimize hyperparameters such as learning rate, style weights, and batch size during GAN training, improving convergence stability and creative output quality.

The structure presents the development and evaluation of an intelligent advertising creative generation system using AECO and SC-StyleGAN. Section 2 reviews related works, highlighting recent advancements in GAN-based advertising optimization and cross-platform creative generation. Section 3 outlines the proposed methodology, detailing the preprocessing of advertising data, AECO-based hyperparameter tuning, and the architecture of the SC-StyleGAN for spatially contextual ad generation. Section 4 discusses the experimental setup, performance evaluation, and comparative results with baseline models. Section 5 concludes with future directions to enhance scalability, personalization, and real-time adaptability in advertising technologies.

## 2 Related works

Automating the generation of ad creatives from landing pages using abstractive text summarization, enabling rapid experimentation in large-scale marketing campaigns, was examined by [8]. The advertising creative optimization by modeling complex interactions between creative elements and improving Click-Through Rate (CTR) prediction using an Auto ML-inspired framework was enhanced [9].

The Automated Creative Optimization (AutoCO) framework outperformed baselines, achieving lower cumulative regret and a 7% CTR increase in online A/B testing.

The two-stage dynamic and creative optimization framework, combining AutoCo and a transformer-based rerank model to improve CTR prediction and creative ranking under ambiguous data conditions, was developed by [10]. Experimental and online testing showed a 10% CTR improvement over baselines, demonstrating superior performance. The integration of a Particle Swarm Optimization-based Recurrent Neural Network (PSO-based RNN) algorithm with Computer-Aided Design (CAD) tools to automate the generation and optimization of advertising artistic designs, enhancing design efficiency and creativity, was explored by [11].

The Dynamic Creative Optimization (DCO) problem, by determining the optimal product and creative ad combination under constraints like ad fatigue and user diversity, was examined [12]. The advertising design, creativity, and efficiency were enhanced by integrating CAD technology and data-driven automation [13]. The developed model enabled the automated generation of diverse advertising designs, successfully reflecting creative schemes and allowing quantitative evaluation, thus validating its effectiveness in promoting innovative advertising solutions. The integration of artificial intelligence in advertising with an emphasis on content production, targeting, personalization, and ad optimization was explored by [14]. Table 1 provides a comparative overview of recent GAN-based approaches. While previous studies have focused on general image synthesis or aesthetic enhancement, none leverage spatial conditioning and adaptive optimization specifically for advertising creative generation.

Table 1: Conventional approaches of GAN for dynamic advertising creative optimization

| Study                 | Model                    | Dataset                   | Metrics Used                            | Key Results                | Limitation                          |
|-----------------------|--------------------------|---------------------------|-----------------------------------------|----------------------------|-------------------------------------|
| Jiang et al. [15]     | StyleGAN (AdSEE)         | Proprietary Ad Dataset    | CTR, Qualitative Feedback               | CTR improvement: +12%      | Limited generalizability            |
| Shilova et al. [16]   | Diffusion + Outpainting  | User Behavior + Ad Images | Personalization Score                   | +15% relevance             | High computational cost             |
| Xu et al. [17]        | PDA-GAN                  | PubLayNet, Rico           | Layout Accuracy, Realism                | Improved layout realism    | Focused on layout generation        |
| Aghazadeh et al. [18] | Various (CAP Evaluation) | Generated Ad Images       | CAP (Creativity, Alignment, Persuasion) | Structured ad quality eval | No generative model proposed        |
| Ma and Zhao [19]      | Enhanced DCGAN           | Logo Design Dataset       | FID, User Rating                        | FID: 23.4, User preference | Focused only on logos, not full ads |

### 2.1 Problem statement

Digital marketing teams need to use personalized and platform-specific ads, but it is difficult to scale them using

traditional methods which results in reduced engagement and inconsistency [15] [16]. Although GANs make automation possible, current models usually do not take context into account [17], run on computers for long

periods [16] or mostly generate logos [18]. To overcome these problems, the study offers a framework that uses Spatially Conditioned GANs, AECO for hyperparameter adjustment and combines various input types to improve the quality and performance of automatic ad creation.

### 3 Methodology

The Ad ImageNet dataset contains image-text advertisement samples, with images typically sized and text averaging. For data preprocessing, images are resized to uniform dimensions, normalized for intensity consistency, and tokenized for textual elements. The AECO-SC StyleGAN framework uses spatial conditioning and adaptive optimization to dynamically generate and optimize advertising creatives for enhanced cross-platform engagement and performance. Figure 1 shows the general outline of the methodological approach.

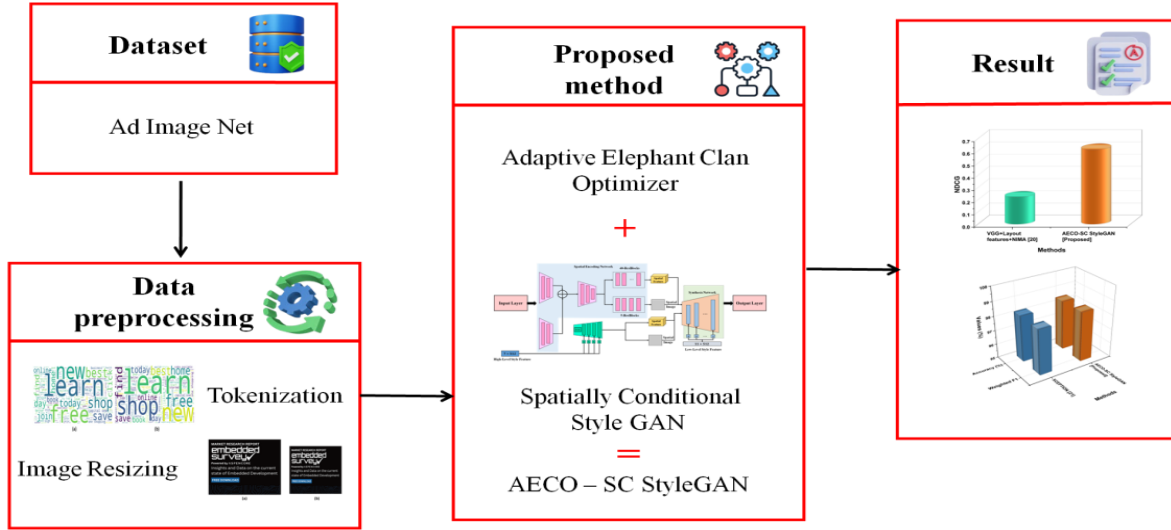


Figure 1: General outline of the methodological approach

#### 3.1 Data collection

The Ad ImageNet dataset, sourced from the Peter Brendan repository, consists of 9,003 image-text advertisement samples totaling approximately 682 MB in size. Each entry includes a banner-style advertisement image along with associated promotional text. The dataset captures a variety of standard ad dimensions, with the most frequent being  $254 \times 254$  pixels, commonly used in digital marketing. The textual content varies in length, with an average of around 525 characters, covering diverse product and event advertisements. The dataset was split into 70% training, 15% validation, and 15% testing sets to ensure robust performance evaluation across unseen data.

##### Source:

<https://huggingface.co/datasets/PeterBrendan/AdImageNet>

#### 3.2 Data preprocessing using image resizing

Image resizing standardizes input dimensions, enabling consistent data processing for DL models and optimizing the generation of dynamic advertising creative. Although the original dataset images varied in size most frequently  $256 \times 256$  pixels the images were uniformly resized to  $224 \times 224 \times 3$  pixels for this research. The  $244 \times 224 \times 3$

pixel image size was used to simplify the original image size. This input data type was created by resizing the image using bicubic interpolation. Because the outcome was smoother at the edges than with bilinear interpolation, this approach was chosen. Bicubic was the perfect balance between process time and high-quality results. The bicubic interpolation estimates the pixels in the  $(j, i)$  positions using a sampling ( $S$ ) distance of 16 nearby pixels ( $4 \times 4$ ) in equations (1)-(5). Figure 2 illustrates (a) before resizing and (b) after resizing.

$$g_{j,i} = [X_{-1}(T_z)X_0(T_z)X_1(T_z)X_2(T_z)]$$

$$\begin{bmatrix} g_{j-1,i-1} & g_{j,i-1} & g_{j+1,i-1} & g_{j+2,i-1} \\ g_{j-1,i} & g_{j,i} & g_{j+1,i} & g_{j+2,i} \\ g_{j-1,i+1} & g_{j,i+1} & g_{j+1,i+1} & g_{j+2,i+1} \\ g_{j-1,i+2} & g_{j,i+2} & g_{j+1,i+2} & g_{j+2,i+2} \end{bmatrix}$$

$$\begin{bmatrix} X_{-1} & (T_w) \\ X_0 & (T_w) \\ X_1 & (T_w) \\ X_2 & (T_w) \end{bmatrix}$$

(1)

Where:  $T_z = i' - i$ ,  $T_w = j' - j$  and  $g_{j,i}$  = pixel value at position  $(j, i)$

$$X_{-1}(T) = \frac{-T^3 + 2T^2 - T}{2} \quad (2)$$

$$X_0(T) = \frac{-3T^3 + 5T^2 + 2}{2} \quad (3)$$



```

    e['params'] += rand() * (matriarch['params'] - e['params'])
    matriarch['params'] += rand() * (gbest['params'] - matriarch['params'])
Step 5: Male Update & Evolution
def male_and_evolution(pop, P_m):
    males = select_males(pop, P_m)
    center = mean_vec([e['params'] for e in pop])
    for m in males:
        m['params'] += rand() * (center - m['params'])
    replace_weakest(pop)
    pop.append(generate_calf(pop))
    random_reset_bottom(pop, pct = 0.3)
Step 6: Main Optimization
def optimize_AECO_SCStyleGAN():
    N, M, G, T, P_m, λ1, λGP, λLP, λFM, data = setup()
    pop = init_population(M, N)
    for _ in range(T):
        for e in pop:
            e['fitness'] = evaluate(e, data, λ1, λGP, λLP, λFM)
        gbest = min(pop, key = lambda x: x['fitness'])
        clan_update(pop, gbest)
        male_and_evolution(pop, P_m)
    return gbest

```

### 3.3.1 SC StyleGAN

SC-StyleGAN enables location-specific control over visual features, enhancing the deep learning framework's ability to dynamically generate and optimize personalized, visually consistent advertising creatives across different contexts. The StyleGAN network generates high-quality images by applying an  $18 \times 512$  style code to 18 layers of the network. It starts with a constant  $4 \times 4$  feature map and progressively grows by a factor of 2 at each stage, ultimately producing images of up to  $1024 \times 1024$  pixels. Each style block receives a  $1 \times 512$  style code that modulates convolution operations, enabling fine control over visual attributes. These style codes correspond to different levels of detail: coarse styles affect the overall layout and color schemes, middle styles

influence microstructure and facial features, and fine styles regulate high-frequency details and textures. Non-visual data like captions, CTR, and demographics are encoded into embeddings using text encoders

and fully connected layers. These embeddings are fused with spatial inputs (semantic maps and sketches) through modulation layers that adjust the style codes, allowing SC-StyleGAN to generate creatives tailored to both visual features and user/context data. For training and evaluation in this study, input images were uniformly resized to  $224 \times 224$  pixels, serving as the initial resolution before the progressive growth to higher resolutions during generation. Figure 4 illustrates the network architecture of SC-StyleGAN.

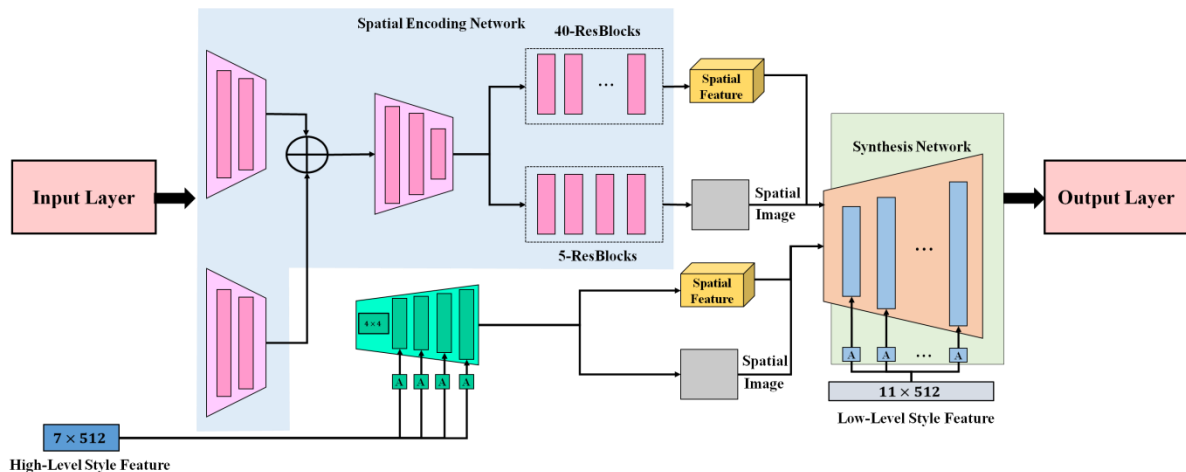


Figure 4: Network architecture of SC-StyleGAN

SC-StyleGAN was a conditional generation system that uses a semantic map and sketches to identify spatial features for coarse and intermediate styles. It consists of two sub-networks: the production network, which uses layers, and the spatial encoding network, which maps input conditions to intermediates. Two encoding modules are suggested for the spatial encoding network, which individually translates the semantic map and  $512 \times 512$  sketches into  $64 \times 256 \times 256$  spatial feature maps. With a spatial dimension of  $32 \times 32$ , the combined map of features was encoded to correspond with the coarse-moderated style in the StyleGAN synthesizing module. To create a  $32 \times 32$  intermediate image, the same steps are followed for the spatial intermediate feature map. Table 2 represents the architectural and computational footprint of the SC-StyleGAN model.

Table 2: SC-StyleGAN architecture details: input dimensions, layer-wise parameters, and computational complexity

| Component                | Layer Type               | Input Shape                | Output Shape                | Params       | FL OPs       |
|--------------------------|--------------------------|----------------------------|-----------------------------|--------------|--------------|
| Semantic Encoder (E_s)   | Conv2D + ReLU $\times 4$ | $512 \times 512 \times 3$  | $64 \times 256 \times 256$  | $\sim 3.1$ M | $\sim 2.5$ B |
| Sketch Encoder (E_k)     | Conv2D + ReLU $\times 4$ | $512 \times 512 \times 1$  | $64 \times 256 \times 256$  | $\sim 2.8$ M | $\sim 2.2$ B |
| Spatial Combiner         | Add/Concat + Down Conv   | $64 \times 256 \times 256$ | $64 \times 32 \times 32$    | $\sim 0.6$ M | $\sim 0.3$ B |
| StyleGAN (Synthesis Net) | StyleBlock $\times 18$   | $1 \times 512$             | $1024 \times 1024 \times 3$ | $\sim 30$ M  | $\sim 75$ B  |

**Objective Function:** The SC-StyleGAN was to precisely map the determined conditions to their equivalents in the synthesis process while encoding the spatial constraint for the StyleGAN synthesizing procedure while preserving the invention value of the StyleGAN. Equation (6) fulfills the training process's goal in the following ways.

$$K(J_{gt}, J_{syn}) = \lambda_{K_1} K_1(J_{gt}, J_{syn}) + \lambda_{K_{GP}} K_{GP} + \lambda_{K_{LP}} K_{LP} + \lambda_{K_{FM}} K_{FM} \quad (6)$$

The SC-StyleGAN training uses a composite loss to enhance image quality and consistency. The  $L1$  loss ( $K_1$ ) ensures pixel-level accuracy, while the global perceptual loss ( $K_{GP}$ ) maintains semantic alignment at full scale. The

local perceptual loss ( $K_{LP}$ ) improves detail by comparing image patches, and the feature matching loss ( $K_{FM}$ ) stabilizes training by aligning intermediate features. Together, these losses guide the network toward realistic and context-aware ad generation. The perceptual metrics (LPIPS) measure the overall perceptual loss after shrinking the target and synthesized images to  $64 \times 64$ . The local perceptive loss (LLP) and the global perceptive loss (LGP) are expressed mathematically as follows in equations (7)-(8).

$$K_{GP}(J_{gt}, J_{syn}) = LPIPS(J_{gt}^r, J_{syn}^r) \quad (7)$$

$$K_{LP}(J_{gt}, J_{syn}) = \frac{1}{L} \sum_{l=1}^L LPIPS(J_{gt}^l, J_{syn}^l) \quad (8)$$

Where  $J_{gt}^r$  and  $J_{syn}^r$  are the resized reality and, synthesizing the image, respectively, and  $LPIPS(.,.)$  was the perceptual measuring function. In each phase,  $J_{gt}^l$  and  $J_{syn}^l$  stand for the  $k$ -th randomly clipped ground truth and synthetic patches, respectively, in equation (9).

$$K_{FM} = \frac{1}{M} \sum_k ||H^k(gt) - H^k(syn)||_1 \quad (9)$$

Where  $H^k(.)$  was the output map of features of the pre-trained StyleGAN synthesizing network's  $k$ -th resolution block (with a matching spatial resolution of  $2^k$ ). The number of computed blocks was  $M$ . Following the replacement resolution block, it computes the  $K_1$  norm between the ground truth generation and the synthesized processes ( $k \in \{6, 7, 8, 9\}$  and  $M = 4$ ).

### 3.3.2 AEEO

To improve model convergence, stability, and learning efficiency, AEEO dynamically modifies training hyperparameters. This raises the caliber and efficacy of engagement of generated ad creatives spanning platforms. The AEEO enhances the DL framework by efficiently optimizing parameters, enabling dynamic generation of personalized advertising creative through adaptive search, exploration, and convergence strategies. The ECO was enhanced into an improved version to support a DL framework that dynamically generates and optimizes advertising creatives. This AEEO algorithm addresses the limitations of the original by improving convergence speed and solution quality, enabling more effective, real-time content creation and personalization in advertising through intelligent data-driven optimization.

Many tasks using Adam and RMSprop are successful, but they have issues with GANs, including convergence issues, collapsing to single modes, and being sensitive to changes in learning rates. To resolve these issues in creating ads for different channels, the new AEEO strategy adapts by using evolutionary methods to tune hyperparameters, which boosts the stability and resilience

of the model when facing different spatial and contextual situations.

Elephant migration under the direction of each clan principal was simulated using the ECO algorithm. This part provides an autonomous movement range and an autonomous movement position for each elephant to keep the algorithm from reaching a local optimum, enhance image variety, and replicate the aforementioned behaviors based on the initial elephant position to generate creative advertising, as illustrated in equations (10) and (11).

$$\Delta W_j^0(i) = \Delta W^{\min}(i) + q \times (\Delta W^{\max}(i) - \Delta W^{\min}(i)) \quad (10)$$

Where  $q$  was a random number in the interval that is uniformly distributed  $[0, 1]$ , and  $\Delta W_j^0(i) (j = 1, 2, \dots, M, i = 1, 2, \dots, C)$  represents the range of independent movement of the  $j$ -th elephant in the  $i$ -th dimension at the starting time. Both  $\Delta W^{\min}(i) = -\Delta W^{\max}(i)$  and  $\Delta W^{\max}(i) = E \times (W^{\max}(i) - W^{\min}(i))$  represent the lower and upper bounds of the  $i$ -th dimensional autonomous motion space. Generally speaking,  $E$  can be seen as 0.005 for improved outcomes,

$$YW_j^0(i) = W_j^0(i) + \Delta W_j^0(i) \quad (11)$$

As evolution advances, the autonomous range of mobility of each elephant should likewise diminish as individuals become closer to one another. Therefore, the independent moving range updates technique in equation (12) is used.

$$\Delta W_j^{it+1}(i) = \left[ 0.9 - \left( 0.8 \times \frac{it}{it_{\max}} \right) \right] \times \Delta W_j^{it}(i) \quad (12)$$

Enhancement of the family clan's renewal technique: The mother elephant was the best person in each family clan, and all other clan members learn from the generative image. While clan members are responsible for maintaining population diversity to provide the mother elephant with superior evolutionary information for quick convergence, the mother elephant was primarily responsible for swiftly investigating the area where the hypothesized optimal location was located.

(1) A method of updating each clan member individually based on the autonomous location traction equation (13), this section proposes a way to update the individual population members based on autonomous location traction to better maintain species variety and avoid significantly slowing down the algorithm's rate of convergence.

$$W_{FCj,n}^{it+1} = \begin{cases} ZX_{FCj,n}^{it+1}(i) + q \times \alpha \times [W_{FCNc-j,N}^{it}(i) - W_{FCj,n}^{it}(i)] \\ \quad + q \times \alpha \times [W_{MC,Rm}^{it}(i) - W_{FCj,n}^{it}(i)], \text{ if } n > \frac{Ne}{2} \\ ZM_{FCj,n}^{it+1}(i) + q \times \alpha \times [W_{FCj,N}^{it}(i) - W_{FCj,n}^{it}(i)] \\ \quad + q \times \alpha \times [W_{MC,Rm}^{it}(i) - W_{FCj,n}^{it}(i)], \text{ else} \end{cases} \quad (13)$$

In the  $Nc - j$  family group  $W_{FCM_{d=j}}$  at the  $j$ -th iteration, the matriarch of the female elephant was represented by  $W_{FCM_{d=j,N}}^{it}$ . After sorting at  $it + 1$  iteration  $ZX_{FCj,n}^{it+1}$  was the autonomous position of the  $n$  ( $n = 1, 2, \dots, M_f$ ) clan member, and equation (14) shows that  $q$  was a uniformly spread range form  $[0, 1]$ .  $\alpha$  was the improved adaptable scaling factor.

$$\alpha = 2 - \left( d \times \frac{it}{it_{\max}} \right) \quad (14)$$

Where  $d$  was a fixed value that was typically set to 0.5 to get the best results, while the optimization issue itself requires in different values.

(2) An autonomous location traction-based individual update method for matriarchs is employed; as previously stated, the globally optimal individuals swiftly approach the globally optimal region after traversing each of the matriarchs in the ECO algorithm. This section suggests an autonomous position, traction-based matriarch updating approach, as illustrated in equation (15).

$$W_{FCj,n}^{it+1}(i) = ZX_{FCj,n}^{it+1}(i) + q \times \beta \times [W_{Best}^{it}(i) - W_{FCj,N}^{it}(i)] \quad (15)$$

The scaling factor  $\beta$  was determined using equation (14), while  $ZX_{FCj,n}^{it+1}$  was the independent movement location of the matriarch in this clan at the  $it + 1$  iteration, acquired similarly to equation (16).

$$\beta = 3 - \left( d \times \frac{it}{it_{\max}} \right) \quad (16)$$

Improvement of the individual renewal method of the male elephant clan: According to the ECO algorithm, the male elephant clan was essential in creating globally ideal locations for female clan leaders and substituting certain family members to supply evolutionary data. Despite this, the number of male elephants might add to the diversity of the family clan. In light of this, equation (17) provides the male elephant individual renewal formula to guarantee that the male elephant clan has particular population diversity and generates as much evolutionary information as possible.



$$W_{MC,m}^{it+1}(i) = ZX_{MC,m}^{it+1}(i) + q \times o \times (W_{Center}^{it} - W_{MC,m}^{it}(i)) \quad (17)$$

In the  $it + 1$  iteration of the male elephant clan  $ZX_{MC,m}^{it+1}$  represents the autonomous movement location of the  $m - th$  ( $m = 1, 2, \dots, M_f$ ) elephant. Equation (18) illustrates that the  $o$  was determined using  $W_{Center}^{it}$ , which was the location of the maternal elephant patriarch in each family clan in the  $it - th$  iteration.

$$W_{Center}^{it} = \frac{1}{N_{c-1}} \times \sum_{j=1}^{N_{c-1}} W_{FCj,N}^{it} \quad (18)$$

Improvement of individual replacement strategy for part of the family clan: Enhance the plan for replacing adult elephants. The following adult elephant replacement was suggested to guarantee the algorithm's speed of convergence and boost the variety of the creative ad images, as adult elephants are not the superior elephants within this clan. Equation (19) indicates the central position of all clan members. Otherwise, as indicated by equation (20), the superior person was chosen to replace the real adult elephant from both the new individual and the original adult elephant.

$$W_{FCj,Gm}^{it+1} = \frac{1}{N_e} \sum_{j=1}^{N_e} W_{FCj}^{it+1}(i) \quad (19)$$

$$W_{FCj,Cal}^{it+1}(i) = W_{FCj,Gm}^{it+1}(i) + q \times \left[ \frac{W_{FCj,Rf}^{it+1}(i) + W_{MC,Rm}^{it+1}(i)}{2} - W_{FCj,Gm}^{it+1}(i) \right] \quad (20)$$

Improve the inferior small elephant replacement strategy: The ECO algorithm replaces the poor individuals in family clans to maintain population diversity, but reduces convergence speed. Early iterations have significant differences, while late iterations focus on population diversity. The worst 0.3Ne tiny elephants in each family clan are replaced with new individuals in the evolutionary stage. Equation (21) generates new individuals during the pre-evolutionary period, where  $it < itmax$ . The  $it + 1$  iteration of the family clan  $W_{FCj}^{it+1}$ , where  $W_{FCj,x}^{it+1}(i)$  was the worst one to be replaced.

$$W_{FCj,Cal}^{it+1}(i) = W_{FCj,x}^{it+1}(i) + q \times \left[ \frac{W_{FCj,Rf}^{it+1}(i) + W_{MC,Rm}^{it+1}(i)}{2} - W_{FCj,x}^{it+1}(i) \right] \quad (21)$$

AECO enhances SC-StyleGAN by tuning hyperparameters using elephant-inspired population dynamics. Each elephant's position represents a candidate solution, evolving through clan-based exploration and adaptive updates. This improves convergence and avoids common GAN issues. However, the mapping between

AECO's search positions and StyleGAN's exact hyperparameters (like learning rate or noise scale) should be clarified.

## 4 Results and discussion

All experiments used Python 3.10.1 on an NVIDIA Tesla V100 GPU. AECO-SC StyleGAN trained for 50 epochs (1,000 iterations each) in approximately 12 GPU hours, outperforming Adam (15 GPU hours) in efficiency. The proposed strategy was assessed and its effectiveness was determined using the following indicators: Normalized Discounted Cumulative Gain (NDCG), accuracy, and weighted F1, The Fréchet Inception Distance (FID), Structural Similarity Index Measure (SSIM) and Peak Signal-to-Noise Ratio (PSNR). Although AECO-SC StyleGAN is a generative framework, its output creatives are evaluated using a downstream binary classification task predicting ad engagement (high vs. low CTR). All models, including baselines like VGG + Layout + NIMA and XCEPTION, are evaluated on this task for a fair comparison, also implementing the baseline method to this research. Table 3 represents hyperparameters for AECO-SC StyleGAN-based framework used in dynamic advertising creative optimization.

Table 3: Hyperparameter Settings for AECO-SC StyleGAN Framework

| Hyperparameter                | Value                       |
|-------------------------------|-----------------------------|
| Batch Size                    | 32                          |
| Learning Rate (Generator)     | 0.0001                      |
| Learning Rate (Discriminator) | 0.0004                      |
| Epochs                        | 200                         |
| Image Size                    | $256 \times 256 \times 3$   |
| Latent Vector Dimension (z)   | 512                         |
| Dropout Rate                  | 0.3                         |
| Activation Function           | Leaky ReLU ( $\alpha=0.2$ ) |
| Normalization                 | Instance Normalization      |
| AECO Population Size          | 30                          |
| AECO Max Iterations           | 100                         |

### 4.1 Evaluation task

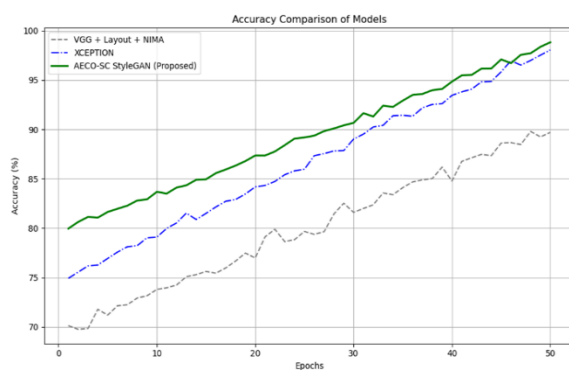
The primary task is a binary classification of ad creatives into 'high engagement' vs. 'low engagement' based on historical CTR data. Ads with CTR above the 75th percentile were labeled as high engagement (1), and others as low engagement (0). This classification target enables the model to learn aesthetic and contextual cues that align with user interaction patterns. The dataset was split 70/15/15 for training, validation, and testing. Evaluation was conducted on the unseen 15% test set. Performance metrics included NDCG, classification accuracy, and weighted F1-score FID, SSIM, and PSNR. Baseline



models included Visual Geometry Group Layout feature Neural Image Assessment (VGG + Layout features + NIMA) [20], XCEPTION [21], AdvAE-GAN [22], BicycleGAN [22], V-GAN [23], and Vanilla GAN [23]. All models were trained under similar hardware and optimization conditions to ensure a fair comparison. Table 4 shows the comparison of classifiers and their performance evaluation results.

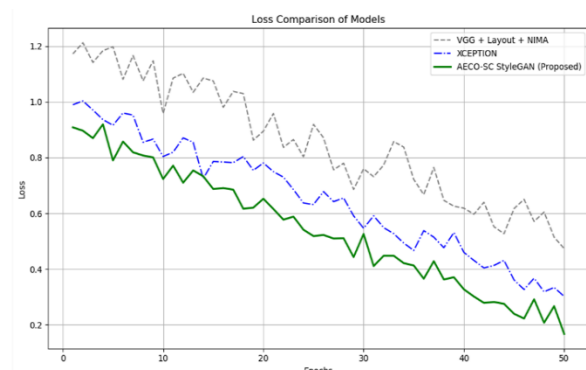
The evaluation pipeline begins with AECO-SC StyleGAN generating advertising creatives. These outputs are labeled based on a CTR threshold to indicate high or low engagement. A classifier then predicts engagement levels, allowing metrics like NDCG, Accuracy, and Weighted F1-score to assess how well the generated creatives align with user interaction patterns.

## 4.2 Accuracy and loss



(a)

The training accuracy and loss over 50 epochs for the proposed AECO-SC StyleGAN, XCEPTION, and VGG + Layout + NIMA are displayed in Figure 5(a, b). The AECO-SC StyleGAN consistently achieves superior accuracy and inferior loss, demonstrating better learning efficiency, faster convergence, and more stable training. Accuracy values were mentioned in percentage. Both training and validation accuracy curves show steady growth and eventual stabilization, indicating effective learning with minimal overfitting. Similarly, the loss curves for training and validation exhibit a clear downward trend, reflecting successful convergence. These results highlight the model's ability to efficiently capture cross-platform advertising dynamics, generate high-quality creatives, and maintain strong generalization across datasets, ultimately improving engagement prediction performance.



(b)

Figure 5: Accuracy and Loss Comparison of Models

## 4.3 Ad image dimension distribution

The GAN ad image dimension distribution operates as a cross-platform deep learning framework that generates multiple platform-optimized sizes for presented images. Standard display and mobile ad dimensions are the format choices for most images, which guarantee visual performance while ensuring cross-platform compatibility. Figure 6 displays the Ad image dimensions across digital platforms.

### 4.4 Clicks-through rate (CTR) by platform

The CTR performance stands tested across different platforms through the use of a GAN framework for dynamic ad optimization. Results indicate that performance metrics vary between platforms since mobile achieves better CTR than desktop. Through its creative adaptation, the GAN model demonstrates high engagement while showing the power of deep learning as a means to improve cross-platform digital advertisement results. Figure 7 shows the CTR distribution across four social platforms.

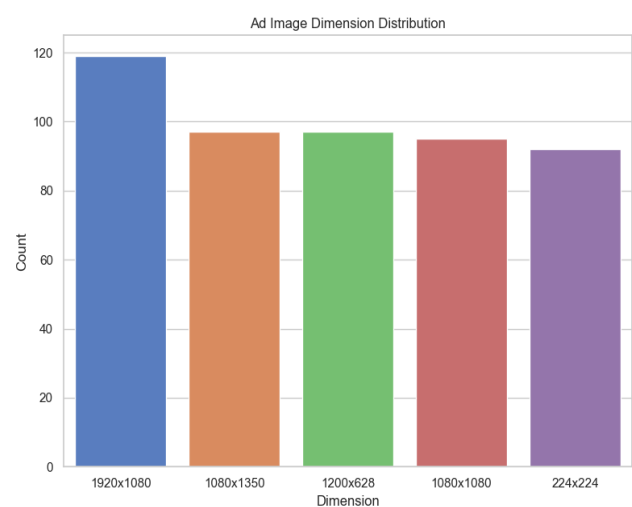


Figure 6: Ad image dimensions across digital platforms

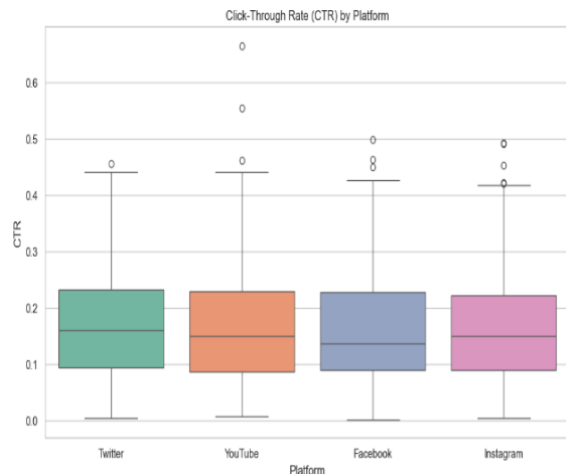


Figure 7: CTR distribution across four social platforms

#### 4.5 Convergence and runtime analysis

To compare the stability of the training between AECO and Adam, by run both for 100 iterations shown in Figure 8. AECO demonstrated a quicker and smoother convergence, as seen by its early near-zero loss. Ad creative generation requires a stable and fast system, as it works with many constraints in quick optimization. By using adaptive learning, AECO avoids the problems of local minima and maintains consistency, which makes it effective than Adam and RMSprop for cross-platform advertising.

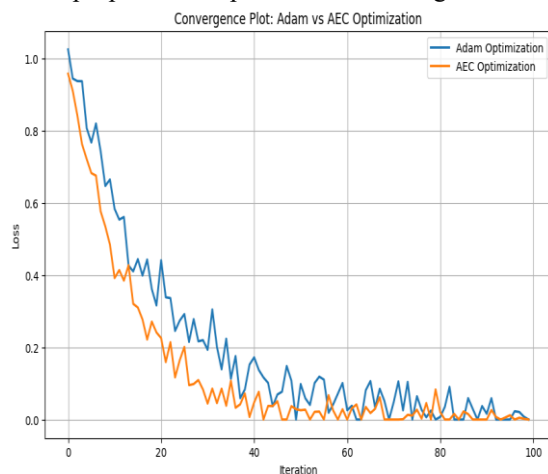


Figure 8: Outcomes of convergence and runtime analysis

#### 4.6 NDCG

The NDCG score was applied to measure how relevant and well-arranged the ad creatives were for users. Because of this, NDCG is better suited for this task, as it rewards higher positions for predicting relevant content. A better NDCG means the model ranks the most engaging and appropriate content first, which helps in dynamic advertising situations where space and time are both limited. The NDCG for the AECO-SC StyleGAN was

0.61, much better than the NDCG of 0.22 for the baseline VGG + Layout + NIMA model. From this, we can see that our model helps us better find and rank the strongest ad creatives first. With this accuracy, marketers are better equipped to promote content that has a significant effect. Figure 9 illustrates the NDCG scores of all the evaluated models.

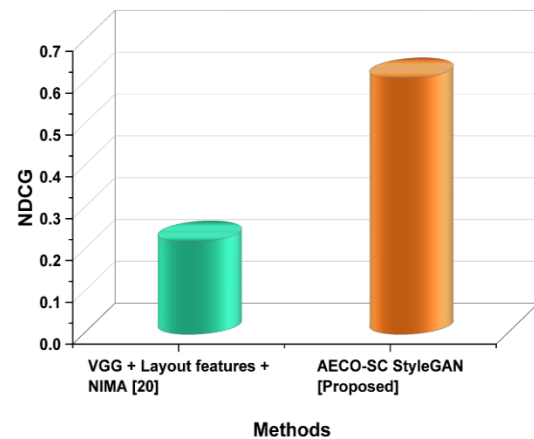


Figure 9: Illustrates NDCG performance results

#### 4.7 Accuracy

Accuracy indicates how effectively the classifier predicts whether generated advertisements will result in high or low user engagement (based on CTR), thus assessing the effectiveness of the generated creatives. The high accuracy in generating platform-specific ad creatives consistently aligns with user engagement metrics, outperforming baseline models in aesthetic coherence, contextual relevance, and predictive performance across platforms. The results demonstrate that XCEPTION achieved an accuracy of 98.27%, while AECO-SC StyleGAN performed slightly better with an accuracy of 98.48%, showcasing their effectiveness in the given task.

#### 4.8 Weighted F1

The F1-score balances precision and recall, crucial for imbalanced engagement data, indicating how well the model generates relevant, high-performing ads while minimizing misclassification. The Weighted F1 score was a metric used to evaluate the performance of a GAN in dynamic advertising creative optimization, emphasizing precision and recall across various platforms. Table 4 gives Evaluation of Ad Engagement Prediction Based on Generated Ad Creatives. The results show that AECO-SC StyleGAN outperforms XCEPTION, achieving a higher Weighted F1 score of 98.5% compared to 98.2%, demonstrating superior performance in dynamic advertising creative optimization. Figure 10 displays the accuracy and weighted F1 evaluation results.

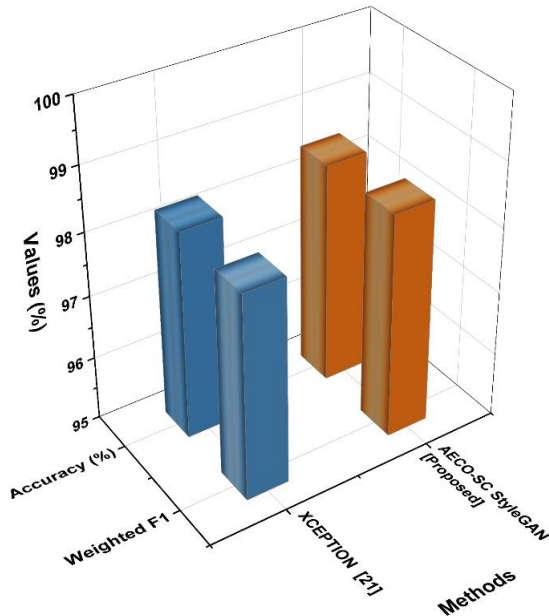


Figure 10: Accuracy and Weighted F1 Evaluation Results

Table 4: Evaluation of ad engagement prediction based on generated Ad Creatives

| Method                            | NDCG | Accuracy (%) | Weighted F1 (%) |
|-----------------------------------|------|--------------|-----------------|
| VGG + Layout features + NIMA [20] | 0.22 | -            | -               |
| XCEPTION [21]                     | -    | 98.27 %      | 98.2 %          |
| AECO-SC StyleGAN [Proposed]       | 0.61 | 98.48 %      | 98.5 %          |

#### 4.8 Statistical evaluation of model performance

To address this, we have conducted additional experiments using five different random seeds. For each seed, the model was trained and evaluated independently using the same data split. We now report the mean  $\pm$  standard deviations for the key evaluation metrics, including Normalized Discounted Cumulative Gain (NDCG), Accuracy, and Weighted F1-score. Performance Comparison of Creative Generation Models on the Ad ImageNet Dataset given below Table 5.

Table 5: Performance comparison of creative generation models on the Ad ImageNet dataset

| Method                      | NDCG (Mean $\pm$ SD) | Accuracy (%) (Mean $\pm$ SD) | Weighted F1 (%) (Mean $\pm$ SD) |
|-----------------------------|----------------------|------------------------------|---------------------------------|
| VGG + Layout + NIMA [20]    | 0.22 $\pm$ 0.015     | 94.62 $\pm$ 0.40             | 94.3 $\pm$ 0.38                 |
| XCEPTION [21]               | 0.45 $\pm$ 0.020     | 98.27 $\pm$ 0.25             | 98.2 $\pm$ 0.21                 |
| AECO-SC StyleGAN [Proposed] | 0.61 $\pm$ 0.018     | 98.48 $\pm$ 0.22             | 98.5 $\pm$ 0.19                 |

In addition to reporting the mean  $\pm$  SD, we performed paired t-tests to evaluate whether the improvements over baseline models are statistically significant. The results confirm that the performance gains of AECO-SC StyleGAN over XCEPTION and VGG+NIMA are statistically significant with  $p < 0.01$  for all three metrics.

#### 4.9 Performance comparison of generative models

FID scores of the suggested AECO-SC StyleGAN are contrasted with those of other GAN-based baselines in Figure 11. Among the tested methods, the proposed AECO-SC StyleGAN delivered the best quality, with an FID score of 38.4752, compared to 42.3256 for AdvAEGAN [22] and 45.0208 for BicycleGAN [22].

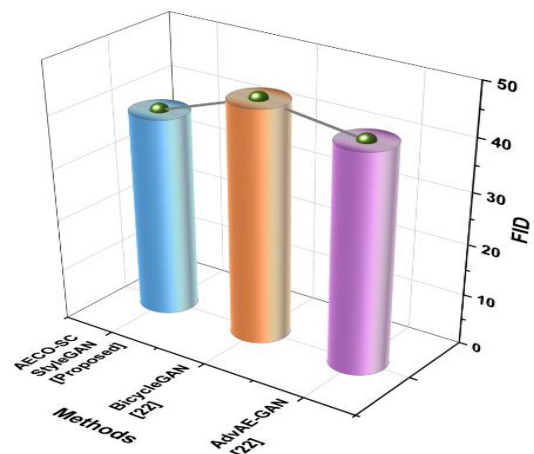


Figure 11: Generative quality evaluation model comparison results.

The SSIM and PSNR metrics for AECO-SC StyleGAN and other GAN variations are shown in Figure 12. Higher PSNR and SSIM values indicate better image fidelity and structural similarity to the original images. The AECO-SC StyleGAN shows the highest PSNR of 35.8 dB and an SSIM of 0.95 which is better than the PSNR of 33.5 dB and SSIM of 0.92 for V-GAN [23] and the PSNR of 28.4 dB and SSIM of 0.85 for Vanilla GAN [23]. It means that AECO-SC StyleGAN is able to create images that are more clearly detailed and accurate than other styles.

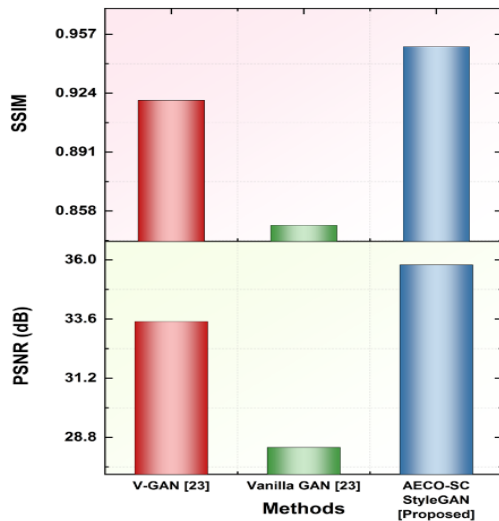


Figure 12: Model comparison results of image quality assessment.

#### 4.10 Visual results and assessment of visual fidelity

To evaluate the visual fidelity of the proposed AECO-SC StyleGAN, we generated advertisement creatives using the Ad ImageNet dataset. Figure 13 illustrates side-by-side examples of generated ads, showcasing a variety of product categories including fashion, electronics, and skincare. The generated ads closely match real ones in layout, color schemes, and promotional text, reflecting platform-specific design aesthetics. While maintaining coherence, the model introduces subtle variations that add diversity and creativity. These results demonstrate that AECO-SC StyleGAN effectively replicates real ad characteristics, providing a scalable and automated approach for cross-platform ad generation.

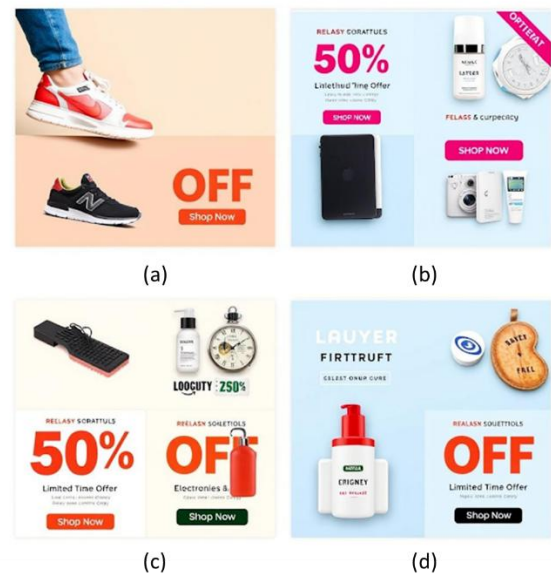


Figure 13: Generated Ad Creative using AECO-SC StyleGAN

#### 4.11 Discussion

Dynamic advertising creative optimization across multiple platforms aims to enhance user engagement and conversions by generating context-aware, personalized ad content. Traditional models such as VGG combined with Layout features and NIMA [20] rely on fixed image features, limiting their capacity to capture the full spectrum of complex, interactive visual and contextual patterns inherent in cross-platform environments. As a result, the creatives they generate often lack adaptability and personalization, making them less effective in varied user scenarios. Meanwhile, XCEPTION-based GAN [21] models, although capable of deeper feature extraction, are hindered by their high computational and memory demands. Their complex operations limit scalability and pose challenges for deployment on lightweight or real-time advertising platforms, reducing practicality in widespread commercial use.

In contrast, the proposed AECO-SC StyleGAN framework addresses these limitations by integrating adaptive hyperparameter tuning and spatial conditioning to generate high-fidelity, semantically consistent creatives tailored to specific platform requirements. AECO enhances convergence and training efficiency, while SC-StyleGAN ensures visual and contextual alignment across formats. This leads to better performance and improved resource utilization, offering a scalable and intelligent solution for dynamic advertising creative generation in diverse deployment environments.

While AECO-SC StyleGAN improves convergence speed and reduces memory consumption relative to baseline GANs during training, it still requires substantial computational resources overall, particularly due to its large model size and high-resolution output generation. However, once trained, the model supports relatively efficient inference, making it suitable for real-time or near-real-time deployment scenarios.

## 5 Conclusions

The DL framework uses GAN for dynamic advertising creative optimization, enabling effective cross-platform strategies to enhance ad personalization and performance in real time. Data collection involved the Ad ImageNet dataset, consisting of multimodal ad samples. Preprocessing included image resizing, tokenization, and intensity normalization. This approach demonstrates a scalable, efficient method for cross-platform ad creative optimization, ensuring higher engagement and visual coherence. The results show that the AECO-SC StyleGAN method achieved an NDCG of 0.61, an accuracy of 98.48%, and a weighted F1 score of 98.5%. These metrics highlight the method's high performance in optimizing dynamic advertising creatives with excellent precision and relevance. Although AECO-SC StyleGAN shows promising results in generating optimized ad creatives with high quality, the training process remains computationally intensive due to the high-resolution outputs and multiple conditioning layers. The model may face challenges in ensuring consistency across diverse platforms, handling large-scale real-time data, and optimizing for varying audience preferences. It also requires significant computational resources for training. Future scope could focus on improving real-time adaptability, cross-platform integration, and reducing computational costs for broader adoption in dynamic advertising.

### 5.1 Limitations and future work

While AECO-SC StyleGAN shows promising results, it presents notable limitations. First, training the model requires significant computational resources, with 30+ hours of training time on high-memory GPUs, limiting accessibility for smaller teams. Second, generalization across domains remains a challenge. Early tests on ad categories like automotive and electronics suggest reduced performance, warranting domain-adaptive retraining. Third, although AECO-SC generates high-quality creatives, its deployment in real-time ad systems is untested. Future work will explore integration with ad delivery platforms and A/B testing frameworks to assess live performance metrics such as CTR and Return on Ad Spend (ROAS), moving toward a fully automated ad generation and evaluation pipeline. Focus on model

compression and distillation techniques to reduce training time and memory consumption without sacrificing output quality.

## References

- [1] Geng, T., Sun, F., Wu, D., Zhou, W., Nair, H., & Lin, Z. (2021). *Automated bidding and budget optimization for performance advertising campaigns*. SSRN. <https://doi.org/10.2139/ssrn.3913039>
- [2] Leow, K. R., Leow, M. C., & Ong, L. Y. (2021). Online roadshow: A new model for the next-generation digital marketing. In *Proceedings of the Future Technologies Conference* (pp. 994–1005). Springer, Cham. [https://doi.org/10.1007/978-3-030-89906-6\\_64](https://doi.org/10.1007/978-3-030-89906-6_64)
- [3] Ameen, N., Sharma, G. D., Tarba, S., Rao, A., & Chopra, R. (2022). Toward advancing theory on creativity in marketing and artificial intelligence. *Psychology & Marketing*, 39(9), 1802–1825. <https://doi.org/10.1002/mar.21699>
- [4] Gharibshah, Z., & Zhu, X. (2021). User response prediction in online advertising. *ACM Computing Surveys (CSUR)*, 54(3), 1–43. <https://doi.org/10.1145/3446662>
- [5] Ouyang, X., Chen, Y., Zhu, K., & Agam, G. (2024). Image restoration refinement with Uformer GAN. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 5919–5928). <https://doi.org/10.1109/cvprw63382.2024.00599>
- [6] Liang, Y., Deng, R., Lin, W., Deng, R., Zhu, X., & Yu, L. (2025). Modeling and Reinforcement Learning Assessment System for Quality Improvement of Advertising Design. *Computer-Aided Design & Applications*, 21, 188–200. <https://doi.org/10.14733/cadaps.2025.S7.188-200>
- [7] Patil, D. (2024). Generative Artificial Intelligence In Marketing And Advertising: Advancing Personalization And Optimizing Consumer Engagement Strategies. Available at SSRN 5057404. <https://dx.doi.org/10.2139/ssrn.5057404>
- [8] Terzioğlu, S., Çoğalmış, K. N., & Bulut, A. (2024). Ad creative generation using reinforced generative adversarial network. *Electronic Commerce Research*, 24(3), 1491–1507. <https://doi.org/10.1007/s10660-022-09564-6>
- [9] Chen, J., Xu, J., Jiang, G., Ge, T., Zhang, Z., Lian, D., & Zheng, K. (2021). Automated creative optimization for e-commerce advertising. In *Proceedings of the Web Conference 2021* (pp. 2304–2313). <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

- [10] Li, G., & Yang, X. (2024). Two-stage dynamic creative optimization under sparse ambiguous samples for e-commerce advertising. *SN Computer Science*, 5(8), 1–16. <https://doi.org/10.1007/s42979-024-03332-z>
- [11] Li, Q., & Zhou, E. (2024). Design and implementation of automatic generation algorithm for advertising artistic design based on neural networks. *Computer-Aided Design & Applications*, 21, 114–127. <https://doi.org/10.14733/cadaps.2024.S18.114-127>
- [12] Baardman, L., Fata, E., Pani, A., & Perakis, G. (2021). *Dynamic creative optimization in online display advertising*. SSRN. <https://doi.org/10.2139/ssrn.3863663>
- [13] Meng, Q., & Wei, R. (2024). Creative advertising design combining CAD and generating adversarial networks. *Computer-Aided Design & Applications*, 21, 102–116. <https://doi.org/10.14733/cadaps.2024.S27.102-116>
- [14] Gao, B., Wang, Y., Xie, H., Hu, Y., & Hu, Y. (2023). Artificial intelligence in advertising: Advancements, challenges, and ethical considerations in targeting, personalization, content creation, and ad optimization. *Sage Open*, 13(4), 21582440231210759. <https://doi.org/10.1177/21582440231210759>
- [15] Jiang, L., Li, C., Chen, H., Gao, X., Zhong, X., Qiu, Y., ... & Niu, D. (2023, August). AdSEE: Investigating the Impact of Image Style Editing on Advertisement Attractiveness. In *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 4239–4251). <https://doi.org/10.1145/3580305.3599770>
- [16] Shilova, V., Santos, L. D., Vasile, F., Racic, G., & Tanielian, U. (2023, September). Adbooster: Personalized ad creative generation using stable diffusion outpainting. In *Workshop on Recommender Systems in Fashion and Retail* (pp. 73–93). Cham: Springer Nature Switzerland. DOI: [https://doi.org/10.1007/978-3-031-76878-1\\_5](https://doi.org/10.1007/978-3-031-76878-1_5)
- [17] Xu, C., Zhou, M., Ge, T., Jiang, Y., & Xu, W. (2023). Unsupervised domain adaption with pixel-level discriminator for image-aware layout generation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10114–10123).
- [18] Aghazadeh, A. and Kovashka, A., 2024. CAP: Evaluation of Persuasive and Creative Image Generation. *arXiv preprint arXiv:2412.10426*. <https://doi.org/10.48550/arXiv.2412.10426>
- [19] Ma, M., & Zhao, W. (2024). Computer-Aided Brand Logo Design Based on Generative Adversarial Networks. <https://doi.org/10.14733/cadaps.2024.S25.60-75>
- [20] Vempati, S., Malayil, K. T., Sruthi, V., & Sandeep, R. (2020). Enabling hyper-personalisation: Automated ad creative generation and ranking for fashion e-commerce. In *Fashion Recommender Systems* (pp. 25–48). Springer International Publishing. [https://doi.org/10.1007/978-3-030-55218-3\\_2](https://doi.org/10.1007/978-3-030-55218-3_2)
- [21] Moreno-Armendáriz, M. A., Calvo, H., Faustinos, J., & Duchanoy, C. A. (2023). Personalized advertising design based on automatic analysis of an individual's appearance. *Applied Sciences*, 13(17), 9765. <https://doi.org/10.3390/app13179765>
- [22] Kong, M. (2025). A study on optimizing deep learning models for creative generation of animated new media advertisements: an application based on improved generative adversarial networks (GANs) and variational autoencoders (VAEs). *J. COMBIN. MATH. COMBIN. COMPUT.*, 127, 7227–7248. DOI: 10.61091/jcmcc127a-401
- [23] Kong, M. (2025). Deep Learning Model Optimization in Creative Generation for New Media Animated Ads. <https://doi.org/10.21203/rs.3.rs-5879017/v1>