

Design Your Ad: Personalized Advertising Image and Text Generation with Unified Autoregressive Models

Appendix

This appendix details the design for human evaluation metrics in Sec. 1. Section 3 provides comprehensive visual comparisons with baseline methods under both personalized and non-personalized settings. To justify our model configuration, Section 4 systematically analyzes how historical length and sampling strategies affect performance. In Sec. 5, we present the instruction tuning templates used in this study. Finally, Section 6 and 7 discuss technical limitations and social impact respectively, ensuring our research advances technology while maintaining ethical and social responsibility.

1. Details of Human Evaluation Metric

To evaluate whether the generated contents (advertising images and text) comply with e-commerce standards, we established strict human evaluation criteria shown in Sec. 1.1 and Sec. 1.2.

1.1. Advertising Image Generation

The human evaluation of generated advertising images is aligned with [2] and assesses four key aspects: product size, product appearance, spatial positioning, and visual perception. Any violation of these criteria results in image rejection. The metric reports the percentage of qualified images among all generated images.

Size mismatch. As shown in Fig. 1, generated advertising images must maintain a reasonable proportion between products and indoor reference objects (e.g., ceilings, tables, baseboards). Disproportionate scaling leads to rejection. For outdoor scenes, images satisfying perspective principles (e.g., nearby objects appearing larger) are considered acceptable.



Figure 1. The examples of size mismatch.

Shape hallucination. Images are rejected if extraneous objects attached to product edges could cause user misunderstanding. For example, the wireless speaker appears as a wired speaker in Fig. 2. However, clearly distinguishable

supporting surfaces (e.g. products on stands) are acceptable. For hollow areas: minor artifacts are tolerated, but significant mismatches between hollow areas and background colors result in rejection.

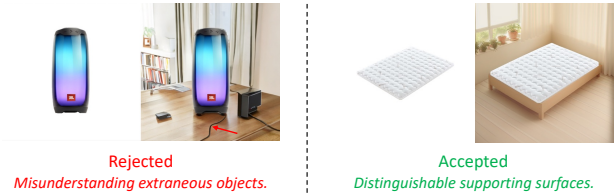


Figure 2. The examples of shape hallucination.

Space mismatch. As illustrated in Fig. 3, images are rejected if products appear spatially disconnected from backgrounds, creating unrealistic impressions, including floating products or inconsistent angles between products and backgrounds. Images with atmospheric or blurred backgrounds are exempt from spatial consistency requirements.



Figure 3. The examples of space mismatch.

Visual perception. This criterion evaluates product distinctiveness and potential discomfort. Images containing distorted human body parts (e.g., faces, hands) that cause discomfort or inappropriate associations are rejected, unless bodies naturally interact with products shown in Fig. 4. Backgrounds that obscure products due to color similarity or excessive complexity also lead to rejection.



Figure 4. The examples of visual perception.

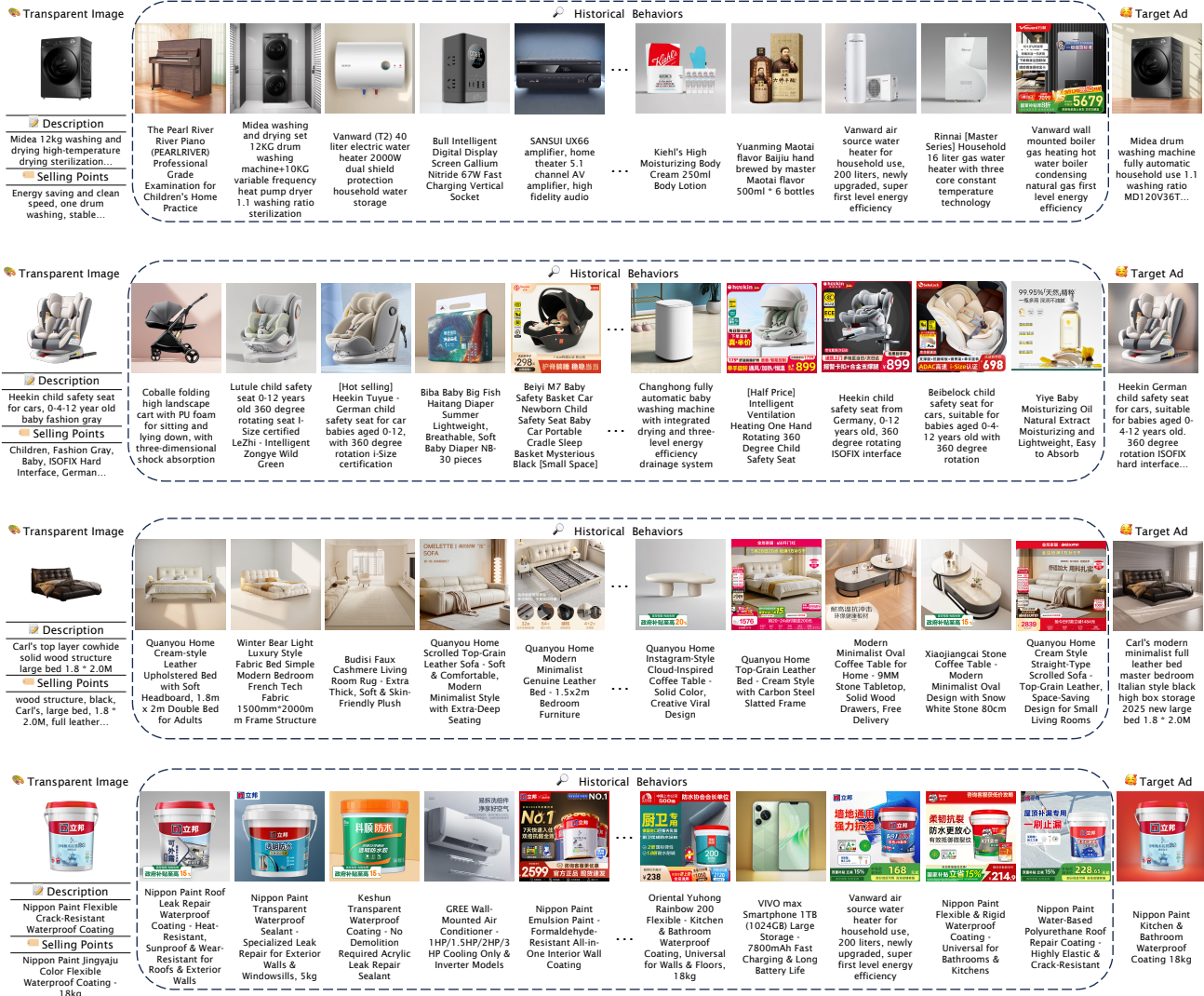


Figure 5. More visualization examples in the PADIM dataset.

Predefined Selling Points	Generated Advertising Texts	Human Evaluation Results
["course synchronization", "Smart Eye Care Mode", "BBK", "Elementary to High School"]	BBK Course Synchronization <i>Blue Light Protection</i> Elementary to High School	Reject, violate factuality.
["Original Flavor", "Engnice", "age 6 month and up", "teeth grinding function", "Infant and Toddler Standards", "48g", "Baby Rice Crackers"]	Original Flavor Engnice age 6 month and up teeth grinding function Infant and Toddler Standards 48g Baby Rice Crackers	Reject, violate readability.
["Resists Discoloration", "useful sealant", "Yuhong", "Mold-Resistant Sealant", "Kitchen & Bathroom Toilet", "multiple color specification"]	Yuhong Mold-Resistant Sealant <i>Permanently</i> Resists Discoloration Kitchen & Bathroom Toilet Sealant	Reject, exaggerated claims.

Figure 6. The examples of human evaluation for advertising texts.

1.2. Advertising Text Generation

Generated advertising texts are evaluated based on factuality and readability as shown in Fig. 6. Violation of either criterion results in rejection. The metric reports the percentage of qualified texts.

Factuality. Texts are rejected if they contain selling points

absent from the predefined list. Inclusion of sensitive content (e.g., pornography, violence) or exaggerated/ false claims also leads to rejection.

Readability. Texts fail if they contain incoherent sentences, direct copying of selling points, or severe repetition of selling points.

1.3. Annotator Details

To ensure the reliability of our annotated data, we implemented a rigorous three-stage manual review process. The annotations were performed by three independent teams of domain experts, each possessing over five years of relevant industry and regulatory experience. In the first stage, annotators performed initial labeling based on predefined guidelines. The second stage involved a verification step, where

a separate group of reviewers checked the labels for consistency and correctness. Finally, a random recheck was conducted on a subset of the data to identify any remaining discrepancies. This multi-tiered quality control guarantees an annotation accuracy of at least 98%.

2. Baseline Implementation Details

To ensure fair comparison, prompts for the text generation task are randomly sampled from a diverse template pool shown in Fig. 7, while image prompts are generated by QwenVL/GPT-4o using various templates in Fig. 8 to ensure stylistic variety.

3. Additional Visualizations

In this section, we first present additional examples from our Personalized Advertising image-text dataset (PAd1M) to illustrate its composition. We then provide more qualitative results across methods for both general and personalized advertisement generation, demonstrating the advantages of our method in jointly generating advertising images and texts.

3.1. Personalized Advertising Image-text Dataset

Figure 5 presents additional examples from the PAd1M dataset, with each row showing a complete user record. Every sample includes the target product’s transparent image, textual description, and selling points, together with its corresponding ground-truth advertising image and title. Additionally, we preserve the complete sequence of each user’s historical click behaviors, where every historically clicked product is accompanied by its original advertising image and text.

3.2. General Ad Generation

Figure 10 presents a visual comparison between our method and current state-of-the-art approaches. It can be observed that Flux-Fill [5] and PosterMaker [3] generate advertising images with richer color and higher visual realism in background design, which explains their advantages in aesthetic metrics. However, these images often contain misleading extraneous objects or unexplainable characters on the main products, compromising the practical usability of the generated results. In contrast, ReliableAd [2] achieves a higher available rate by employing structurally simple backgrounds with minimal color composition at the cost of limited visual appeal, resulting in inferior aesthetic scores. With the effective control approaches, our method maintains a high available rate while generating more realistic background content, achieving an optimal balance between aesthetic quality and practical value.

We present a comparison of advertising text generation results in Fig. 12. It reveals that existing methods [4, 9, 14]

tend to extensively copy selling points, which contributes to their strong performance on metrics like BLEU and ROUGE, yet leads to lower human evaluation scores due to poor semantic organization. In contrast, our approach employs instruction tuning and semantic constraints to ensure generated texts not only faithfully preserve selling points but also achieve natural fluency and professional organization comparable to ground-truth advertising texts. Additional examples in Fig. 14 further validate our method’s effectiveness.

3.3. Personalized Ad Generation

Figure 11 presents comparative results of different methods on personalized image generation. It can be observed that Flux-Kontext [6] struggles to maintain proper product proportions in generated images (e.g. row 2 in Fig.11) and tends to directly replicate historical content rather than extracting user preferences, resulting in significant deviations from actual clicks. While Pigeon [13] shows improved generated images through its preference extraction module, the single-modal architecture remains susceptible to sample-level noise. For instance, in the refrigerator case (row 4 in Fig. 11), although both Pigeon and ground-truth feature indoor backgrounds, noticeable discrepancies persist in color selection. In contrast, our coarse-to-fine preference understanding module effectively reduces the negative effect from noise in historical behaviors, enabling generated images to better align with user preferences in both scene and color style.

We compare text generation outputs across different methods in Fig. 13. It can be observed that Qwen3 tends to directly copy selling points while DeepSeek-R1 shows moderate improvement in this aspect. Our method shows not only more comprehensive coverage of selling points but also closer sentence structures and expressive styles to real clicks. These results validate the advantage of our approach in modeling user preferences. More generated texts are shown in Fig. 15.

4. Additional Results

History length. We examine how historical behavior sequence length affects model performance, and the results are shown in Table 1. Without historical data, the model generates content that significantly diverges from ground-truth clicks and results in poor performance across all three evaluation metrics due to the absence of user preference guidance. As the history length increases, the model progressively extracts more complete user preferences, leading to steady improvements on metrics, where the best performance for image and text generation is reached at history lengths of 5 and 10, respectively. However, further extending the sequence introduces more noise than useful information, causing performance degradation. Based on these

General Text Generation Template			
1	**Product Information** [Description]: {}, [Selling Points]: {}. **Task** Generate an advertising text based on [Description] and [Selling Points], where the text uses ONLY words from the provided [Selling Points].	6	**Product Information Provided**:[Description]: {}, [Selling Points]: {}. **Action** Generate advertising text based on these. The text must consist only of words from the Selling Points.
2	Generate a complete and attractive advertising text using EXCLUSIVELY words from Selling Point {} and Description {}.	7	Based on the Selling Point: {}, Description: {}, write a product title for a specific user with historical clicks: {}. Make sure the product title using only the words form Selling Point.
3	**Input** Product description: {}, and selling points: {}. **Task** Generate an advertising. Text uses only selling points' words.	8	Generate exactly one e-commerce product title based on the Selling Point: {}, Description: {}, ensuring the title reflects the factual information from the selling.
4	Generate a product title using the Selling Point: {}, Description: {}, with the constraint that the generated text uses only selling points words.	9	Given a product with description {} and selling points {}, generate an advertising text and image. The text must exclusively use terms from the selling points.
5	Using the provided product details - Description: {}, Selling Points: {} - generate an advertising text. The text should strictly utilize vocabulary from the Selling Points.	10	From the product description {} and selling points {}, develop an advertising text. Ensure the text uses exclusively the Selling Points' words.

Personalized Text Generation Template			
1	**Product Information** [Description]: {}, [Selling Points]: {}. **Task** Generate an advertising text based on [Description] and [Selling Points], where the text uses ONLY words from the provided [Selling Points]. Generation content following the style of the historical clicks {}.	6	**Product Information Provided**:[Description]: {}, [Selling Points]: {}. **Action** Generate advertising text based on these. The text must consist only of words from the Selling Points. Conform to text style aligned with [historical clicks]: {}.
2	Write a customized product title using the Selling Point: {}, Description: {}, and the user's Historical Clicks: {}, by analyzing the historical clicks to determine the user's preferred structure and text style, and incorporating ONLY words from the provided [Selling Points].	7	Based on the Selling Point: {}, Description: {}, write a product title for a specific user with historical clicks: {}. Make sure the product title should align with the user's historical preference and use only the words form Selling Point.
3	**Input** Product description: {}, and selling points: {}. **Task** Generate an advertising. Text uses only selling points' words. Follow the styles of historical clicks: {}	8	Generate exactly one personalized e-commerce product title based on the Selling Point: {}, Description: {}, and Historical Clicks: {}, ensuring the title reflects the factual information from the selling point and is stylistically adapted to match the text patterns found in the historical clicks.
4	Generate a customized product title using the Selling Point: {}, Description: {}, with the constraint that the generated text should follow the user's preference from the Historical Clicks: {} and uses only selling points words.	9	Given a product with description {} and selling points {}, generate an advertising text and image. The text must exclusively use terms from the selling points. The output should conform to text style from historical clicks {}.
5	Using the provided product details - Description: {}, Selling Points: {} - generate an advertising text. The text should strictly utilize vocabulary from the Selling Points, according to text style from historical click: {}.	10	From the product description {} and selling points {}, develop an advertising text. Ensure the text uses exclusively the Selling Points' words. Style the text as historical clicks {}.

Figure 7. The template pools for advertising text generation task.

Image Prompt Generation Template (General Ad)			
1	Generate a text-to-image prompt for a product background based the Selling Points {}, Description {}.	6	Generate an image prompt for a product background by analyzing the Selling Point {} and Description {}.
2	Formulate a specific text-to-image prompt for a product environment using the Selling Points {}, Description {}.	7	Develop an image generation prompt focusing on a product background which matches the Selling Point {} and Description.
3	Please generate a text-to-image prompt for a product background ensuring the generated environment directly reflects the specific atmosphere, materials, and overall vibe explicitly mentioned or logically implied by the Selling Point {} and Description {}.	8	Create a descriptive prompt to generate an attractive background for the product image in text-to-image generation task, enhancing its appeal. Base your prompt on the Description {} and Selling Points {}. Directly output the background prompt without additional explanation.
4	Write a text-to-image prompt for a product background using the Selling Point {} and Description {}, with the directive to derive the appropriate lighting setup and visual mood from the tone conveyed in the provided product details.	9	Develop an image generation prompt for a product background using the Selling Point {} and Description {}, ensuring that background elements or environmental details are inspired by the keywords and features outlined in the product information.
5	Write a background generation prompt utilizing the Selling Points {} and Description {}, instructing the generator to establish a specific mood and atmosphere that organically supports the narrative of the target product.	10	Write a text-to-image background prompt by using the Selling Point {} and Description {} to create a background that creatively translates the product's unique features into striking visual elements.

Figure 8. The template pools for image prompt generation.

Table 1. The result of different history length setting.

History Length	PBS \uparrow	BLEU \uparrow	ROUGE \uparrow
0	0.617	0.225	0.525
3	0.631	0.375	0.615
5	0.639	0.400	0.637
8	0.623	0.414	0.650
10	0.634	0.435	0.662
12	0.624	0.417	0.644

Table 2. The results of different sampling strategies.

Setting	PBS \uparrow	BLEU \uparrow	ROUGE \uparrow
Random Sampling	0.610	0.431	0.656
Most-similar Sampling	0.618	0.439	0.663
Product Similarity Sampling	0.634	0.435	0.662

findings, we select a history length of 10 as the final configuration for balanced performance in image and text generation.

Sampling strategy. In Table 2, we compare the impact of different historical behavior sampling strategies on model

performance. Specifically, we test three sampling methods on our full model: “Random Sampling” chooses N historical behaviors randomly, “Most-Similar Sampling” directly picks the top-N behaviors with the highest product similarity, and “Product Similarity Sampling” refers to our proposed approach. Results show that random sampling obtains inferior performance due to noise in historical behaviors. Although most-similar sampling achieves the best text generation quality by focusing on highly relevant samples,

Instruction Template			
1	**Product Information** [Description]: {}, [Selling Points]: {}. **Task** Generate an advertising text and image based on [Description] and [Selling Points], where the text uses ONLY words from the provided [Selling Points]. Generation content following the text style <text_ph> and the image style <image_ph>.	6	Based on the product description {} and selling points {}, compose an ad with text and image. Text must be derived solely from the Selling Points. Apply text style <text_ph> and image style <image_ph> to the generation.
2	**Input** Product description: {}, and selling points: {}. **Task** Generate an advertising text and image. Text uses only selling points words. Follow styles: text <text_ph>, image <image_ph>.	7	From the product description {} and selling points {}, develop an ad featuring text and an image. Ensure the text uses exclusively the Selling Points' words. Style the text as <text_ph> and the image as <image_ph>.
3	Given a product with description {} and selling points {}, generate an advertising text and image. The text must exclusively use terms from the selling points. The output should conform to text style <text_ph> and image style <image_ph>.	8	**Product Information Provided** : [Description]: {}, [Selling Points]: {}. **Action** : Generate advertising text and image based on these. The text must consist only of words from the Selling Points. Conform to text style <text_ph> and image style <image_ph>.
4	**Task** Generate advertising content: text and image. **Input** Product [Description]: {}, [Selling Points]: {}. Text must use ONLY words from [Selling Points]. Follow text style <text_ph> and image style <image_ph>.	9	**Input** Product Description: {}, Key Selling Points: {}. **Output Requirement** Generate an advertisement comprising text and visual. The text should exclusively contain terms from the Selling Points. Ensure the text aligns with <text_ph> style and the image with <image_ph> style.
5	**Product Details** [Description]: {}, [Key Features]: {}. **Task** Create an advertising text and image using only words from [Key Features], with text style <text_ph> and image style <image_ph>.	10	Using the provided product details - Description: {}, Selling Points: {} - generate an ad with text and imagery. The text should strictly utilize vocabulary from the Selling Points. Generate according to text style <text_ph> and image style <image_ph>.

Figure 9. The instruction template used for instruction tuning.

it overlooks instances with low product similarity but potential reference value, thus offering only limited gains in PBS. In contrast, our sampling strategy maintains both relevance and diversity, achieving the best overall balance between image and text generation performances.

5. Instruction Templates

Figure 9 presents the instruction templates used in our framework. The “{}” indicates the placeholder for textual input, while the special tokens <image_ph> and <text_ph> mark the insertion location for image and text preference token embeddings from multimodal preference extraction. These templates are randomly employed during training to enhance the model’s ability to mitigate hallucinations, effectively constraining text generation to the pre-defined selling points and ensuring accurate, controllable output.

6. Limitation

While this study introduces the first unified framework for personalized advertising image-text generation, several limitations remain. Existing autoregressive frameworks struggle with embedded visual text generation, typically producing poor results. Our work therefore focuses on establishing a reliable baseline for personalized advertisements without visual text, leaving text rendering for future work. Besides, the inference efficiency of our method cannot meet the requirements of real-time inference. Future efforts may incorporate visual text control modules and optimize model architecture to enable generation for industrial applications [1, 7, 8, 10–12, 15].

7. Social Impact

Practical impact. The proposed method could serve as a heuristic tool to collaborate with human designers, helping them reduce the routine workload while complementing their expertise. By automatically producing image-text

content aligned with user preferences, it significantly reduces manual design efforts. This technology enables e-commerce platforms to achieve more precise advertisement targeting, improving product exposure efficiency and increasing user click-through rates, thereby advancing intelligent digital marketing.

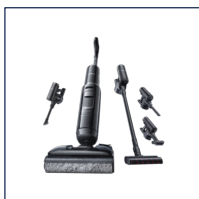
Ethical impact. The PAd1M dataset is constructed in strict compliance with legal requirements and platform regulations. All advertising content underwent a thorough review by the e-commerce platform to ensure no ethical concerns or sensitive information were raised. The dataset is solely for academic research and has passed a comprehensive legal review regarding copyright and usage scope, guaranteeing its ethical application within the research area without commercial infringement risks.

References

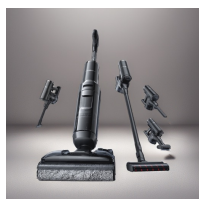
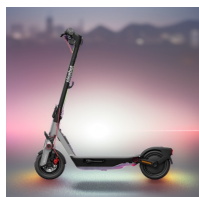
- [1] Ke Cao, Jing Wang, Ao Ma, Jiasong Feng, Zhanjie Zhang, Xuanhua He, Shanyuan Liu, Bo Cheng, Dawei Leng, Yuhui Yin, et al. Relactrl: Relevance-guided efficient control for diffusion transformers. *arXiv preprint arXiv:2502.14377*, 2025. 5
- [2] Zhenbang Du, Wei Feng, Haohan Wang, Yaoyu Li, Jingsen Wang, Jian Li, Zheng Zhang, Jingjing Lv, Xin Zhu, Junsheng Jin, et al. Towards reliable advertising image generation using human feedback. In *European Conference on Computer Vision*, pages 399–415. Springer, 2024. 1, 3
- [3] Yifan Gao, Zihang Lin, Chuanbin Liu, Min Zhou, Tiezheng Ge, Bo Zheng, and Hongtao Xie. Postermaker: Towards high-quality product poster generation with accurate text rendering. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 8083–8093, 2025. 3
- [4] Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025. 3
- [5] Black Forest Labs. Flux. <https://github.com/black-forest-labs/flux>, 2024. 3
- [6] Black Forest Labs, Stephen Batifol, Andreas Blattmann,

- Frederic Boesel, Saksham Consul, Cyril Diagne, Tim Dockhorn, Jack English, Zion English, Patrick Esser, et al. Flux. 1 kontekst: Flow matching for in-context image generation and editing in latent space. *arXiv preprint arXiv:2506.15742*, 2025. 3
- [7] Run Ling, Wenji Wang, Yuting Liu, Guibing Guo, Haowei Liu, Jian Lu, Quanwei Zhang, Yexing Xu, Shuo Lu, Yun Wang, et al. Ragar: retrieval augmented personalized image generation guided by recommendation. *arXiv preprint arXiv:2505.01657*, 2025. 5
- [8] Ao Ma, Jiasong Feng, Ke Cao, Jing Wang, Yun Wang, Quanwei Zhang, and Zhanjie Zhang. Lay2story: extending diffusion transformers for layout-toggable story generation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 16102–16111, 2025. 5
- [9] Qwen, :, An Yang, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chengyuan Li, Dayiheng Liu, Fei Huang, Haoran Wei, Huan Lin, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jingren Zhou, Junyang Lin, Kai Dang, Keming Lu, Keqin Bao, Kexin Yang, Le Yu, Mei Li, Mingfeng Xue, Pei Zhang, Qin Zhu, Rui Men, Runji Lin, Tianhao Li, Tianyi Tang, Tingyu Xia, Xingzhang Ren, Xuancheng Ren, Yang Fan, Yang Su, Yichang Zhang, Yu Wan, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, and Zihan Qiu. Qwen2.5 technical report, 2025. 3
- [10] Jing Wang, Ao Ma, Ke Cao, Jun Zheng, Zhanjie Zhang, Jiasong Feng, Shanyuan Liu, Yuhang Ma, Bo Cheng, Dawei Leng, et al. Wisa: World simulator assistant for physics-aware text-to-video generation. *arXiv preprint arXiv:2503.08153*, 2025. 5
- [11] Yun Wang, Kunhong Li, Longguang Wang, Junjie Hu, Dapeng Oliver Wu, and Yulan Guo. Adstereo: Efficient stereo matching with adaptive downsampling and disparity alignment. *IEEE Transactions on Image Processing*, 2025.
- [12] Yun Wang, Jiahao Zheng, Chenghao Zhang, Zhanjie Zhang, Kunhong Li, Yongjian Zhang, and Junjie Hu. Dualnet: Robust self-supervised stereo matching with pseudo-label supervision. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 8178–8186, 2025. 5
- [13] Yiyan Xu, Wenjie Wang, Yang Zhang, Biao Tang, Peng Yan, Fuli Feng, and Xiangnan He. Personalized image generation with large multimodal models. In *Proceedings of the ACM on Web Conference 2025*, pages 264–274, 2025. 3
- [14] An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, et al. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025. 3
- [15] Zhanjie Zhang, Ao Ma, Ke Cao, Jing Wang, Shanyuan Liu, Yuhang Ma, Bo Cheng, Dawei Leng, and Yuhui Yin. U-stydit: Ultra-high quality artistic style transfer using diffusion transformers. *arXiv preprint arXiv:2503.08157*, 2025. 5

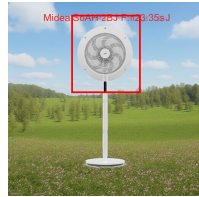
Transparent Image



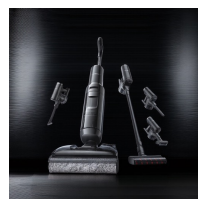
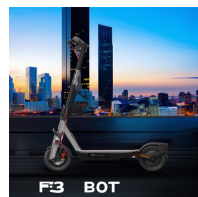
ReliableAd



Flux-Fill



PosterMaker



Ours

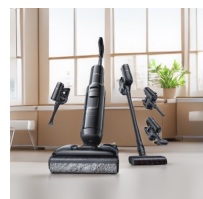


Figure 10. The generated advertising images of different methods on general advertisement generation tasks. The red boxes mark misleading extraneous objects, and the yellow boxes mark uncomfortable generated results (such as incomplete body parts).



Figure 11. The generated advertising images of different methods on personalized advertisement generation tasks.

Selling Points	Qwen2.5	Qwen3	DeepSeek-R1	Ours	Ground Truth
Men's Shoes, INITIATOR, Sports Shoes, Size 41, Nike, Versatile, IB4595-001, Fashion, NIKE, Casual Shoes	Nike INITIATOR Men's Shoes IB4595-001 Size 41 Versatile Fashion Sports Casual Shoes	Nike Men's Shoes Fashion Sports Shoes INITIATOR Casual Shoes IB4595-001 Versatile Size 41	Nike INITIATOR Versatile Fashion Men's Shoes Casual Shoes Sports Shoes IB4595-001 Size 41	NIKE Men's Shoes Fashion Sports Shoes IB4595-001 Size 41	<ol style="list-style-type: none"> NIKE Men's Shoes (Versatile Fashion) Sports Shoes NIKE Men's Shoes Casual Shoes Versatile Sports Shoes IB4595-001 Size 41 Nike Men's Shoes Casual Shoes INITIATOR Versatile Sports Shoes IB4595-001 Size 41 NIKE Men's Shoes Casual Shoes Fashion Sports Shoes IB4595-001 Size 41
Ultra-thin, TCL 520L, R520T15-UQ5, Flat-embedded, Cross four-door, 594mm, Grade 1 Energy Efficiency, Dual system, Refrigerator, Boundless, 20% National Subsidy, Frost-free, Dual purification	TCL Boundless R520T15-UQ5 520L Cross four-door 594mm Ultra-thin Flat-embedded Refrigerator Dual system Frost-free Dual purification Grade 1 Energy Efficiency 20% National Subsidy	TCL 520L Boundless Dual purification Dual system Frost-free Refrigerator 594mm Flat-embedded Cross four-door R520T15-UQ5 Grade 1 Energy Efficiency 20% National Subsidy	Boundless TCL 520L r520t15-uq5 594mm Ultra-thin Flat-embedded Cross four-door Refrigerator Dual system Dual purification Frost-free Grade 1 Energy Efficiency 20% National Subsidy	TCL 20% National Subsidy 520L Frost-free Refrigerator	<ol style="list-style-type: none"> TCL 20% National Subsidy 520L Ultra-thin Flat-embedded Refrigerator TCL 20% National Subsidy Ultra-thin Flat-embedded Refrigerator R520T15-UQ5 TCL 20% National Subsidy 520L Ultra-thin Dual system Refrigerator TCL 520L Dual system Dual purification National Subsidy Refrigerator TCL 520L Ultra-thin Flat-embedded Refrigerator (National Subsidy) R520T15-UQ5
20% National Subsidy, Top Load Washer, Haier, Direct Drive Inverter, B220E0, Fully Automatic, 10kg, Household, Bacteria and Mite Removal, High Cleaning Ratio, Shentong, Grade 1 Energy Efficiency, Large Capacity, Top Lid, Haier, 2025 New Product	Haier 2025 New Product Top Load Washer 10kg Large Capacity Top Lid Grade 1 Energy Efficiency Direct Drive Inverter Shentong Bacteria and Mite Removal Fully Automatic 20% National Subsidy B220E0	Haier 2025 New Product Top Load Washer Fully Automatic Household Large Capacity Top Lid Grade 1 Energy Efficiency High Cleaning Ratio Direct Drive Inverter Shentong Bacteria and Mite Removal B220E0 20% National Subsidy 10kg	Haier 2025 New Product Top Load Washer Fully Automatic Household Large Capacity 10kg Top Lid Grade 1 Energy Efficiency High Cleaning Ratio Direct Drive Inverter Shentong Bacteria and Mite Removal B220E0 20% National Subsidy	Haier 20% National Subsidy Top Load Washer 10kg Grade 1 Energy Efficiency	<ol style="list-style-type: none"> Haier Grade 1 Energy Efficiency Direct Drive Inverter Top Load Washer 10kg B220E0 Haier National Subsidy 10kg Fully Automatic Washer Haier National Subsidy 10kg Fully Automatic Top Load Washer Haier Grade 1 Energy Efficiency High Cleaning Ratio Top Load Washer 10kg
SAS77, Exam Preparation, Students, Exam Pen, Practice Pen, Click-type, Deli, Quick-dry, Black, 0.5mm 4pcs/box, Exam Specific, Gel Pen, 0.5mm ST Tip, deli	Deli Quick-dry Click-type Exam Pen 0.5mm ST Tip Gel Pen Student Practice Pen Exam Preparation Specific Black 0.5mm 4pcs/box SAS77	Deli Quick-dry Click-type Gel Pen 0.5mm ST Tip Exam Specific Exam Pen Student Practice Pen Black 0.5mm 4pcs/box SAS77	Deli Quick-dry Click-type Exam Specific Gel Pen 0.5mm ST Tip Student Exam Preparation Practice Pen Exam Pen Black 0.5mm 4pcs/box SAS77	Deli Exam Specific Gel Pen 4pcs/box	<ol style="list-style-type: none"> Deli Exam Specific Quick-dry Gel Pen 0.5mm 4pcs/box Deli Exam Specific Quick-dry Gel Pen 0.5mm 4pcs/box Deli Exam Specific Gel Pen Quick-dry 4pcs/box Deli Quick-dry Click-type Gel Pen 0.5mm ST Tip Practice Pen Exam Pen Exam Preparation Exam Specific Black 4pcs/box SAS77 Deli Quick-dry Click-type Gel Pen 0.5mm ST Tip Student Practice Pen Exam Pen Exam Preparation Exam Specific Black 4pcs/box SAS77 Deli Quick-dry Gel Pen 0.5mm ST Tip Student Practice Pen Exam Pen Exam Preparation Exam Specific Black 4pcs/box SAS77

Figure 12. The generated advertising texts of different methods on general advertisement generation tasks. Some of the covered selling points are marked in green.

Selling Points	Qwen3	DeepSeek-R1	Ours	Ground Truth
Glowing, MINISO, Bedroom, Creative, Night Light, Home Decor, Kuromi Crescent Sweet Dream, LED, Ornament, Bedside Lamp, Birthday Gift, MINISO	MINISO Creative Night Light LED Bedside Lamp Bedroom Ornament Birthday Gift Kuromi Crescent Sweet Dream Home Decor	MINISO Kuromi Crescent Sweet Dream Creative LED Glowing Night Light Bedroom Bedside Lamp Ornament Birthday Gift Home Decor	MINISO Kuromi Crescent Sweet Dream Glowing Night Light LED Bedside Lamp Bedroom Ornament Home Decor Birthday Gift	MINISO Kuromi Crescent Sweet Dream LED Glowing Ornament Night Light Bedroom Bedside Lamp Creative Home Decor Birthday Gift
Ultra-thin, Refined Storage Compartments, White, Dual Sterilization 2.0, Household, Long-lasting Dual Purification, Flush-built, Dual System, Ronshen, Sugar Cube 515 National Subsidy, Refrigerator, Grade 1 Energy Efficiency, BCD-515P60FZMAD, Active Dual Purification, Ronshen, French Door, Four-door, Flat-embedded, Odor Separation, Multi-door	Ronshen Dual System Four-door Refrigerator Household Grade 1 Energy Efficiency BCD-515P60FZMAD Dual Sterilization 2.0 Long-lasting Dual Purification Odor Separation	Ronshen Sugar Cube 515 National Subsidy Ultra-thin Flush-built Dual System Active Dual Purification Odor Separation Four-door Household Grade 1 Energy Efficiency Refrigerator BCD-515P60FZMAD	Ronshen Household 437L French Door Multi-door Four-door Ultra-thin Built-in Refrigerator Household 20% National Subsidy Grade 1 Energy Efficiency Frost-Free BCD-437WD3MPQ Gray	Ronshen 437L French Door Multi-door Four-door Ultra-thin Built-in Refrigerator Household Grade 1 Energy Efficiency Frost-Free Energy Saving BCD-437WD3MPQ Gray 20% National Subsidy
2025 Ryzen 9 Gaming Laptop, Strix G16 2025 R9 8940HX 16G 1T RTX5060, R9 8940HX 16G 1T, 16-inch, Gaming Screen, 2.5K 165Hz RTX5060, Ryzen 9 RTX5060, ROG, Strix G16 Gaming Laptop, 2.5K 165Hz High Refresh Rate, High Refresh Rate Screen, Cooling, Graphics Card, Powerful, High-Performance Configuration, 2025 Gaming Laptop	2025 Ryzen 9 Gaming Laptop Strix G16 2025 R9 8940HX 16G 1T RTX5060 R9 8940HX 16G 1T 16-inch Gaming Screen 2.5K 165Hz RTX5060 Ryzen 9 RTX5060 ROG Strix G16 Gaming Laptop 2.5K 165Hz High Refresh Rate High Refresh Rate Screen Cooling Graphics Card Unprecedentedly Powerful High-Performance Configuration 2025 Gaming Laptop (copy)	ROG Strix G16 2025 Ryzen 9 RTX5060 Gaming Laptop 16-inch 2.5K 165Hz High Refresh Rate Gaming Screen	ROG Strix G16 2025 R9 Ryzen 9 16-inch Gaming Laptop High-Performance RTX5060 Graphics Card 2.5K 165Hz 16G 1T	ROG Strix G16 2025 Ryzen 9 16-inch Gaming Laptop (R9 8940HX 16G 1T RTX5060 2.5K 165Hz)
Yuwell, Household Use, 14-Day Finger-prick Free, Finger-prick Free Continuous Monitoring, Anytime5 HSE 2-Box Set, Yuwell, Anytime 5th Generation, 14-Day Monitoring, Mid-Autumn Festival Gift, Finger Blood Glucose Monitor, 5th Generation Anytime5 HSE 14-Day Continuous Monitoring	Yuwell 5th Generation Anytime5 HSE 14-Day Continuous Monitoring Finger-prick Free Continuous Glucose Monitor	Yuwell 5th Generation Anytime5 HSE Finger-prick Free Continuous Monitoring 2-Box Set	Yuwell Anytime 5th Generation Continuous Glucose Monitor Finger-prick Free 14-Day Continuous Monitoring Household Use Anytime5 HSE 2-Box Set	Yuwell Anytime 5th Generation Continuous Glucose Monitor Anytime5 HSE Finger-prick Free Household Use 14-Day Continuous Monitoring 2-Box Set

Figure 13. The generated advertising texts of different methods on personalized advertisement generation tasks. Some of the covered selling points are marked in green.



Member's First Store [Long-lasting Fragrance] Amino Acid Oud Rose Shampoo 500ml



Sinabo Mid-century Style Solid Wood Large Bed [Oak]



[Valentine's Day Gift] Emporio Armani Small Sugar Cube Silver Crystal-set Mother-of-Pearl Dial Watch



Rongshida Free On-site Installation Electric Water Heater 50 Liter Grade 1 Energy Efficiency



Laotou Soft Leather Breathable Men's Shoes Black Smooth Leather



Xiaomi Hu 4K Ultra HD Flagship Screen 144Hz High Refresh Rate 5G Full Netcom Pad Pro Sky Blue 16GB+512GB



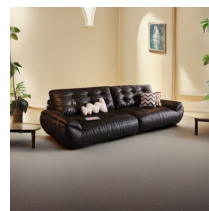
Chigo Stainless Steel Electric Kettle Wellness Pot 1L Mini Constant Temperature



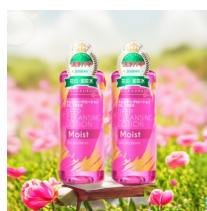
HUILETOYS Children Walker [First Birthday Gift]



Midea Omni-directional Wheels Top Cover Water Refill Air Cooler



CARZZI Top Grain Genuine Leather Sofa Small Apartment Lazy Study Room 2.8-meter Straight Row Mid-Century Sofa Space Cloud 2025 New Model Luxury Edition



Hana Print Deep Cleansing Refreshing Moisturizing Makeup Remover Water 380ml*2 Gift Box Set



Puxin Mumu Safe and High-Quality Bonded Leather High-Density Foam Genuine Leather Office Business Coffee Table Set Genuine Leather 3-Seater

Figure 14. The generated advertising images and texts of our method on general advertisement generation tasks. Some of the covered selling points are marked in green.

















Ours	Ground Truth
	
<p>Friso Royal Organic Series [Hong Kong Version] Follow-on Infant Formula Stage 2 (Follow-on Infant Formula) 800g*6 cans/box</p>	<p>Friso Royal Organic Series Hong Kong Version Stage 2 Follow-on Infant Formula 800g*6 cans/box</p>
	
<p>Western Digital (WD) SN7100 2TB PCIe 4.0 SSD Solid State Drive Large Capacity Gaming eSports Computer Components Desktop Laptop Expansion Hard Drive M.2 Interface Original Factory NAND</p>	<p>Western Digital (WD) SSD Solid State Drive M.2 Interface SN7100 PCIe 4.0 Gaming eSports Computer Expansion Hard Drive SN770 Upgrade Version SN7100 2TB</p>
	
<p>Mizone × Genshin Impact Lime Flavor Electrolyte Water Whole Case 600ML*15 Mizone+ Yae Miko No Peripheral</p>	<p>Mizone × Genshin Impact Whole Case Electrolyte Water 600ML*15 Lime Flavor Sports Drink - Yae Miko No Peripheral</p>
	
<p>Dudu Meijia Genuine Leather Sofa Modern Minimalist Living Room 2025 New Model Minimalist Leather Craft Sofa Storage Armrest A80# Genuine Leather Sofa 1.7-meter</p>	<p>Dudu Meijia Genuine Leather Sofa Living Room 2025 New Model Modern Minimalist Armrest Storage Leather Craft Sofa A80# 1.7-meter</p>
	
<p>JOMOO Bathroom Vanity Ceramic Integrated Basin 90cm Large Storage Aluminum Alloy Sink and Cabinet Set A2750-75AT-1</p>	<p>JOMOO Bathroom Vanity Ceramic Integrated Basin Aluminum Alloy Sink and Cabinet Set Large Storage 90cm A2750-75AT-1</p>
	
<p>Anke Xin Hygienic Wet Wipes 75% Alcohol 40 Sheets * 3 Packs 99% Sterilization Rate</p>	<p>Anke Xin 75% Alcohol Wipes 40 Sheets * 3 Packs Disinfection and Sterilization Hygienic Wet Wipes No-Rinse Hand Sanitizing Household Use Office Use</p>
	
<p>Youlaicoco Cream Style Bathroom Vanity Set Ceramic Integrated Basin Bathroom Vanity Set Wash Basin Cream Color 80cm Regular Wall-mounted Cabinet + Mirror Cabinet Solid Wood Washbasin</p>	<p>Youlaicoco Solid Wood Cream Color Ceramic Integrated Basin Bathroom Vanity Set Vanity Washbasin Bathroom Wash Basin 80cm Cream Style Wall-mounted Cabinet Regular Mirror Cabinet</p>
	
<p>German Continental Car Tires 235/55R18 100V UCJ Fits Tiguan L Haval H2 Tanyue Tiggo 8</p>	<p>German Continental Car Tires 235/55R18 100V UCJ Fits Tiguan L Haval H2 Tanyue Tiggo 8</p>

Figure 15. The generated advertising images and texts of our method on personalized advertisement generation tasks. Some of the covered selling points are marked in green.