ReCon: Training-Free Acceleration for Text-to-Image Synthesis with Retrieval of Concept Prompt Trajectories

Retrieval-based AI method cutting diffusion model runtime 40% while preserving image fidelity.

Researchers at Purdue University have developed RECON (Retrieving Concepts), a novel retrieval-based diffusion acceleration method designed to enhance the efficiency and fidelity of text-to-image (T2I) generation in diffusion models. Traditional diffusion models, while excelling in photorealistic image generation, are hindered by slow processing speeds due to the high number of Neural Function Evaluations (NFEs) required. Existing training-free methods, such as text-based or noise-based retrieval approaches, often sacrifice image quality and diversity, introducing extraneous details or failing to capture the essence of input prompts. RECON addresses these challenges by extracting visual "concepts" from prompts to form a knowledge base, enabling the generation of adaptable "flexible trajectories." This approach incorporates essential details from retrieved prompts while maintaining fidelity to the input text. By leveraging pre-trained language models to decompose prompts into visually meaningful components, RECON consistently produces high-fidelity images, reducing NFEs by up to 40%.

Technology Validation:

Extensive testing on datasets like MS-COCO, Pick-a-Pic, and DiffusionDB demonstrated RECON's superiority across metrics such as Pick Score, CLIP Score, and Aesthetics Score. A user study further confirmed RECON's effectiveness, with 76% of participants rating its images as the highest fidelity compared to competing methods.

Advantages:

-Improved Efficiency

-Enhaced Diversity

-Adaptable

-Robust Testing

-Broad Compatibility

Applications:

-Text-to-image Generation

-Digital Art and Design

-Entertainment and Media

-Research and Development

TRL: 4

Intellectual Property:

Provisional-Gov. Funding, 2024-09-24, United States

Keywords: Diffusion model acceleration,Text-to-image generation,AI image synthesis,Retrieval-based image generation,High-fidelity generative AI,Efficient neural rendering,Language model integration,Visual concept extraction,Creative AI tools,Digital content creation,Media and entertainment AI,Generative design technology,AI-driven image enhancement,Computational efficiency in AI