

Diffusion-based Latent Image Generation with Label Conditioning: A Score-Based Approach

ABSTRACT

We represent a score-based strategy for creating images incorporating label conditioning, diffusion models, and latent space modeling. Our method produces a wide range of high-quality images that match particular labels by utilizing diffusion models and taking label information into account. We iteratively update the latent space via score estimation to obtain precise label representation. Experiment results show how well our method produces realistic images with fine-grained label control. Our score-based methodology provides a flexible framework for precise label-guided image production, opening up possibilities in computer vision and artistic design.

Method with System Diagram/Design Complexity

Image generation is score-based in our process. Label conditioning, diffusion models, and latent space modeling make it up. By repeatedly perturbing a base picture, diffusion models capture image production processes. For efficient image attribute manipulation, latent space modeling uses a variational encoder to translate input pictures into a latent space. Label conditioning guarantees that created pictures match specific labels, giving fine-grained control over generated material. Our technique generates accurate and diversified images with controlled label guidance by integrating these components in a cohesive and scalable way.

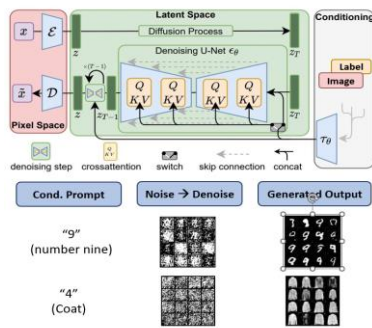


Figure 1: Diagram of latent diffusion model with label conditioning

Novelty of Project

Combining score-based modeling, diffusion models, latent space, and label conditioning, our approach to image generation is distinctive. It provides fine-grained control, efficient attribute manipulation, and targeted content generation compared to existing methods. It offers an effective solution for diverse and controlled image synthesis with precise label guidance.

Impact on society/environment

The approach profoundly impacts society, enabling personalized and targeted image generation for various applications. By reducing reliance on resource-intensive processes like photoshoots, it minimizes waste and the environmental footprint, promoting sustainable content creation.

Business Model/Feasibility/Financial Scalability plan

Our project's business plan focuses on deploying score-based image generation technology to various industries, including advertising, entertainment, and e-learning. We aim to offer the technology as a service, providing customizable image generation solutions to businesses. Revenue will be generated through subscription-based pricing models, where clients pay for access to the platform and usage-based fees for generated images. We will also explore partnerships with content creation agencies and digital platforms to reach a wider audience. Scalability will be achieved through cloud-based infrastructure, allowing efficient and cost-effective deployment to meet the growing demand for high-quality, personalized visual content.

Conclusion

The project is in progress with initial model training on a small dataset. Further work is needed to train the model on a larger dataset and incorporate advanced conditional approaches for more realistic image generation.