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

Supervisor: Dr. Mohammad Ashrafuzzaman khan

Jiaul Haque Saboj, Sadia Sifati Shammee, Tawsif Mahmud

Department of Electrical and Computer Engineering, North South University

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.

PROCESS DIAGRAM

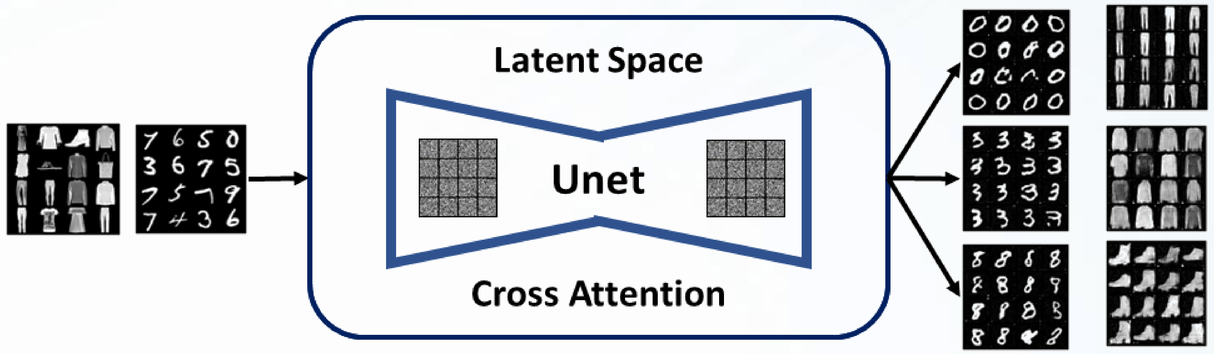

Fig: Process Diagram

METHODOLOGY

Our project involves a systematic methodology to generate label-conditioned images using a diffusion model. We begin by preprocessing the MNIST and Fashion MNIST datasets, ensuring proper normalization and incorporating label conditioning. Through rigorous training, we deploy a U-Net architecture to effectively denoise the images and enable accurate generation. Leveraging the power of latent space embeddings, we capture the underlying distribution of the data, facilitating conditional image generation. This approach allows us to seamlessly integrate label conditioning, enabling us to produce images with specific attributes or characteristics associated with the provided labels.

NOVELTY

Our project offers a unique and cutting-edge solution that unlocks the power of label-conditioned image generation, revolutionizing visual content creation. By providing precise control and unleashing creativity, our platform empowers users to effortlessly craft captivating and customized visuals, setting them apart from the crowd and enabling them to leave a lasting impression.

FUTURE WORK

As we look to the future, our research endeavors revolve around broadening the scope of our investigation to encompass various datasets of diverse complexities. Despite the challenges posed by limited computational resources, we conducted preliminary experiments on distinct flower datasets, albeit without employing conditioning techniques, yielding promising outcomes. Moving forward, our primary objective is to enhance the precision and quality of our results by incorporating label conditional approaches. Furthermore, we are committed to exploring novel conditioning techniques, such as image conditioning, in order to push the boundaries of generative models for advanced image generation and synthesis.

Fig: Generated flower image

BUSINESS MODEL

Our innovative product has the potential to revolutionize the creative industry by providing a powerful tool for generating visually stunning content, catering to professionals in design, advertising, and marketing. Additionally, our solution can extend its impact to the education sector, empowering learners of all ages with a unique and engaging learning experience through the generative capabilities of our platform. By bridging the gap between creativity and education, we aim to inspire and cultivate the creative minds of tomorrow, making a lasting impact in the field of education.

CONCLUSION

The project has successfully demonstrated the potential of label conditioned image generation using diffusion models. We have achieved promising results with the Mnist and Fashion Mnist datasets, and our future work aims to expand our research to larger and more complex datasets, exploring novel conditioning approaches to further enhance the quality and diversity of generated images.