Adrian Kukla

CS7643 HW4

Question 4.1.3

Key contributions

This paper introduces denoising diffusion probabilistic models, which is a class of generative models inspired by nonequilibrium thermodynamics. This model has a reverse denoising process which learns to reconstruct the original input image from noise by improving the generated sample using iterations. Moreover, the model has a forward noising process where Gaussian noise is added over time to an image on each iteration, increasing the signal to noise ratio. The paper's major contribution was to show that diffusion models have ability to generate high quality sample and, in some cases, outperform other types of generative models.

Strengths

The authors have shown and derived the mathematical equations underlying the approach. Moreover, the approach is based on properly structured statistical and probabilistic framework. Authors have shown that this approach can generate high quality samples with more stability in training phase as compared to GANs. Diffusion models allow for a more diverse and higher coverage of the distribution of data samples.

Weaknesses

Diffusion models require a lot of data to be tuned appropriately, low inference speed. Also, the computational cost is high as the ability to generate samples is quite slow due to the reverse sampling process, which required many iterations to generate a single image. Potentially a drawback to real world applications when low latency is key.

Adrian Kukla

CS7643 HW4

Question 4.1.4

Personal takeaways

That diffusion models have the potential to generate diverse and better-quality samples than other generative models. I was also happy to realise that researchers are aware of the ethical implications and potential misuse of their work by third parties. I had the idea that researchers don't pay attention to such issues and are mostly concerned with producing good quality papers and research breakthroughs. Glad to see some thought was put into this.

I further understood that a deep understanding of probabilistic frameworks is crucial in being able to produce better results and more suitable models given the context of problem design. Lastly, a key takeaway here is that knowledge of other domains (such as nonequilibrium thermodynamics) can help one be more creative and apply key concepts from other fields to one's own area of research.

Adrian Kukla

CS7643 HW4

Question 4.2.3

Key contributions

This paper introduces latent diffusion models which improve the efficiency of pixel-based models by operating in a latent space. This approach leverages a pretrained autoencoder and the model maps images into a lower dimensional latent space where diffusion occurs. After generation happens the decoder reconstructs the final image. It significantly reduces computational costs without reducing the quality of images. Authors release pretrained models for further research and applications.

Strengths

The paper is very thorough in explaining all the contributions, limitations, methodology, societal impact (I like how the paper does mention the potential issues with the approach). The appendices provide ample amount of information for interested parties. A thorough experimental design was conducted.

LDMs have shown to reduce computational and inference costs but enable high resolution image production. LDMs also achieve competitive performance across a range of tasks, and they require precise weighting between reconstruction and generation.

Weaknesses

The sequential sampling process is slower than GANs. The reconstruction may be an issue for high precision tasks. The model is dependent on the quality of the autoencoder (used before the diffusion latent space), where errors in encoding or decoding can worsen image quality. The paper doesn't provide suggestions on how to prevent malicious third parties from using the models for spreading misinformation or spam.

Adrian Kukla

CS7643 HW4

Question 4.2.4

Personal takeaways

It is interesting to know how latent space diffusion can improve generative models and make them more practical and reduce complexity. Once again, it’s interesting how authors of the paper mentioned societal impact, whereas I’m not used to reading papers which mentioned this often. It’s incredible how researchers can build on other bodies of work and come up with innovative solutions to reduce the space or time complexity of problems.

Paper reinforced the idea that combining techniques is worthwhile to investigate so one can overcome limitations of each approach, for example here autoencoders and diffusion models.

Further research can be done to try and reduce inference time to use fewer steps in the sequential sampling process. Could also explore latent spaces targeted at specific domains like video creation to further improve performance for those tasks.