MONAI Generative Models - Roadmap

November - December

New networks

- AutoencoderKL (AEKL) and Latent Diffusion Models (LDM) [1,2]
- Denoising Diffusion Probabilistic Models (DDPM) [3]
- Vector Quantised Variational Autoencoder (VQ-VAE) [4,5]

Adding Schedulers (Required for Diffusion models): They define the methodology for iteratively adding noise to an image or for updating a sample based on Diffusion model outputs.

- DDPM scheduler [3]
- DDIM scheduler [6]

New Layers:

- Vector Quantisation layer (Required for VQ-VAEs)

New losses

- 2D and 2.5D Perceptual loss (based on Imagenet[7] and RadImageNet [8])
- 3D Perceptual loss
- Discriminator Loss (LS-GAN loss)
- Spectral Domain Loss

New metrics

- Mean Structural Similarity (MSSIM)
- Fréchet inception distance (FID)

New engines

- Improved Trainer with Adversarial Components (Compatible with MONAI FL)

Tutorials (Time Consuming)

- Training autoencoderKL, VQVAE, and VQGAN (Torch code)
- Training unconditioned and conditioned Latent Diffusion Models in 2D and 3D data (Torch code)
- Generative Models with MONAI Bundle

New Pretrained Models for Model Zoo

- Latent Diffusion Model Trained on brain data

January-February

New networks

- Transformers [4]
- GANs
- Anomaly Detection Networks [9]

Add support to train VQVAEs (or VQGANs) with Transformers

New losses

- New discriminator loss (example: WGAN loss)

New metrics

- Maximum Mean Discrepancy (MMD)

Tutorials (Time Consuming)

- Training autoencoderKL, VQVAE, and VQGAN (Ignite code)
- Training unconditioned and conditioned Latent Diffusion Models in 2D and 3D data (Ignite code)
- Training Transformers (Torch and Ignite code)
- Training and sampling Diffusion Models with classifier-free guidance
- Inpainting with Diffusion Models
- Evaluation of Generative Models regarding Realism (using FID and MMD) and Diversity (Using SSIM based metrics)
- Privacy evaluation in Synthetic Datasets (how to obtain measures of privacy and resemblance)
- Generative Models with MONAI Federated Learning
- Segmentation on Synthetic data

New applications

- Anomaly detection

New Pretrained Models

- Networks trained on OCT and Chest XR (TBD)

After 4 months

- Super Resolution in Space and Time
- Generative Models conditioned on text data
- Generative Models on tabular data
- Domain Adaptation

- MRI reconstruction methods

References

1. Pinaya, W. H. L. *et al.* Brain Imaging Generation with Latent Diffusion Models. in *MICCAI Workshop on Deep Generative Models* 117–126 (Springer, 2022).
2. Rombach, R., Blattmann, A., Lorenz, D., Esser, P. & Ommer, B. High-Resolution Image Synthesis with Latent Diffusion Models. *arXiv Prepr. arXiv2112.10752* (2021).
3. Dhariwal, P. & Nichol, A. Diffusion models beat gans on image synthesis. *arXiv Prepr. arXiv2105.05233* (2021).
4. Esser, P., Rombach, R. & Ommer, B. Taming Transformers for High-Resolution Image Synthesis. *arXiv Prepr. arXiv2012.09841* (2020).
5. Tudosiu, P.-D. *et al.* Morphology-Preserving Autoregressive 3D Generative Modelling of the Brain. in *International Workshop on Simulation and Synthesis in Medical Imaging* 66–78 (Springer, 2022).
6. Song, J., Meng, C. & Ermon, S. Denoising diffusion implicit models. *arXiv Prepr. arXiv2010.02502* (2020).
7. Zhang, R., Isola, P., Efros, A. A., Shechtman, E. & Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. in *Proceedings of the IEEE conference on computer vision and pattern recognition* 586–595 (2018).
8. Mei, X. *et al.* Radimagenet: an open radiologic deep learning research dataset for effective transfer learning. *Radiol. Artif. Intell.* **4**, e210315 (2022).
9. Baur, C., Denner, S., Wiestler, B., Albarqouni, S. & Navab, N. Autoencoders for Unsupervised Anomaly Segmentation in Brain MR Images: A Comparative Study. *arXiv Prepr. arXiv2004.03271* (2020).