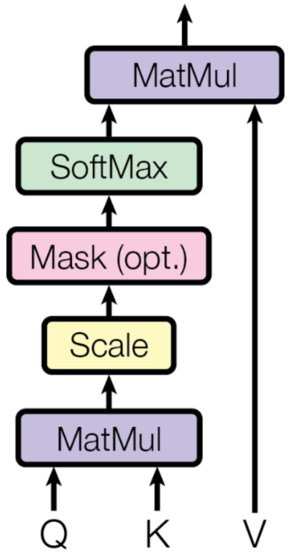
Applying attention mechanism in fast ZDC simulations

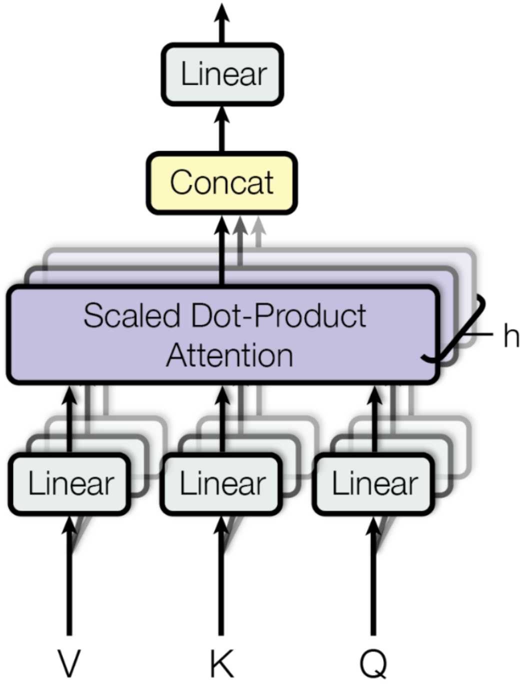
Maksymilian Wojnar

Attention mechanism

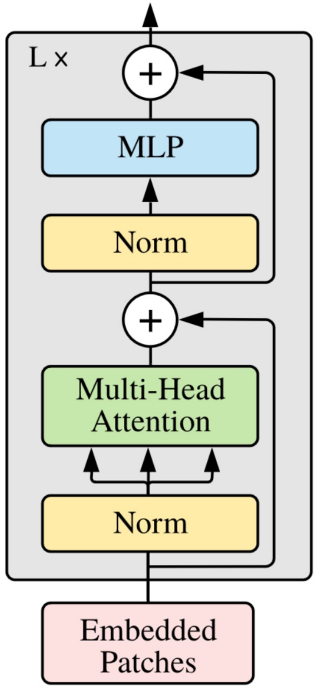
Scaled Dot-Product Attention



Multi-Head Attention

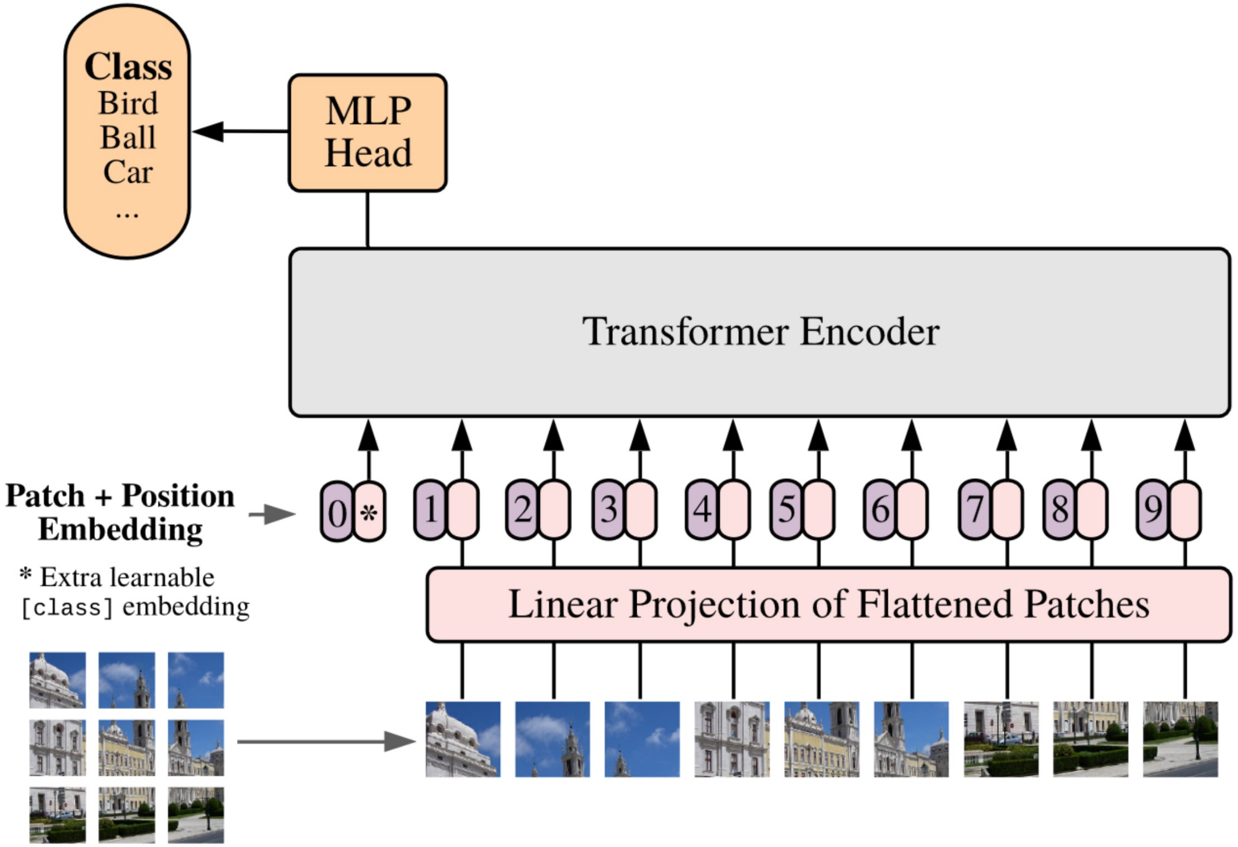


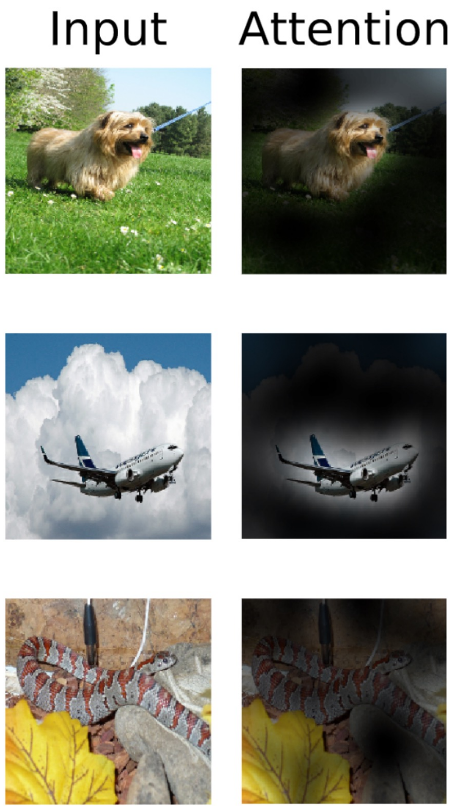
Transformer Encoder



Ashish Vaswani et al. 2017. *Attention Is All You Need*. NeurIPS, Vol. 30.
Ruibin Xiong et al. 2020. *On Layer Normalization in the Transformer Architecture*. ICML'20.

Vision Transformer (ViT)

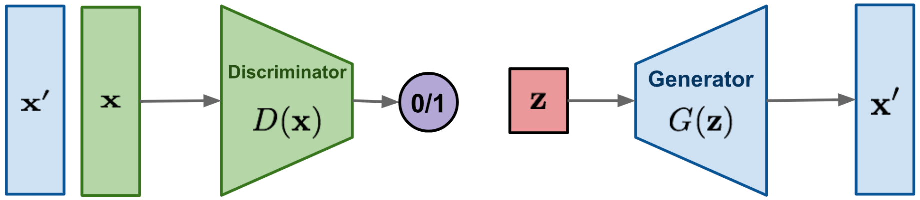




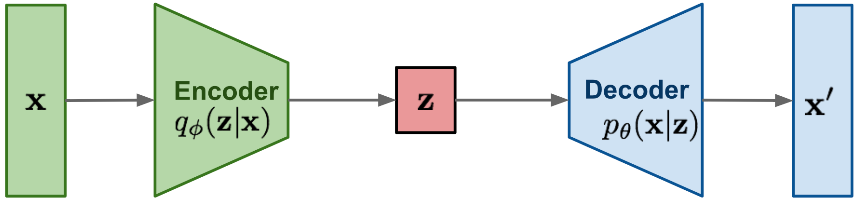
Alexey Dosovitskiy et al. 2021. *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. ICLR 2021.

Used architectures

Generative Adversarial Networks

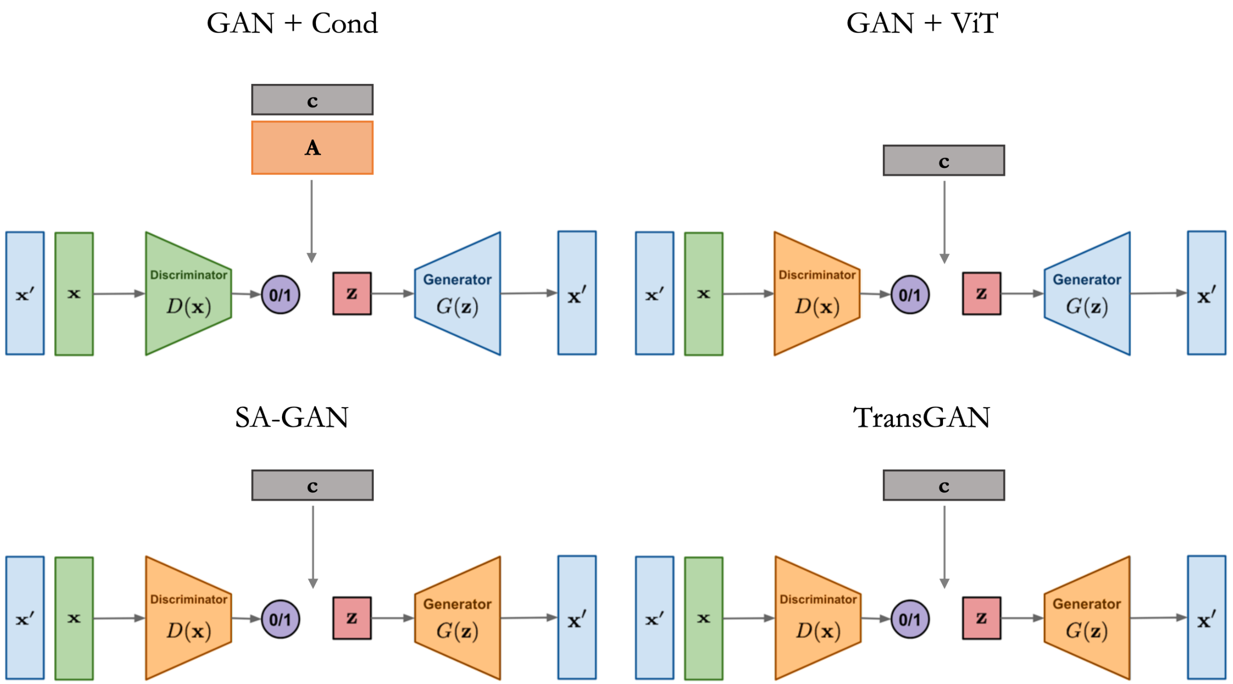


Variational Autoencoders



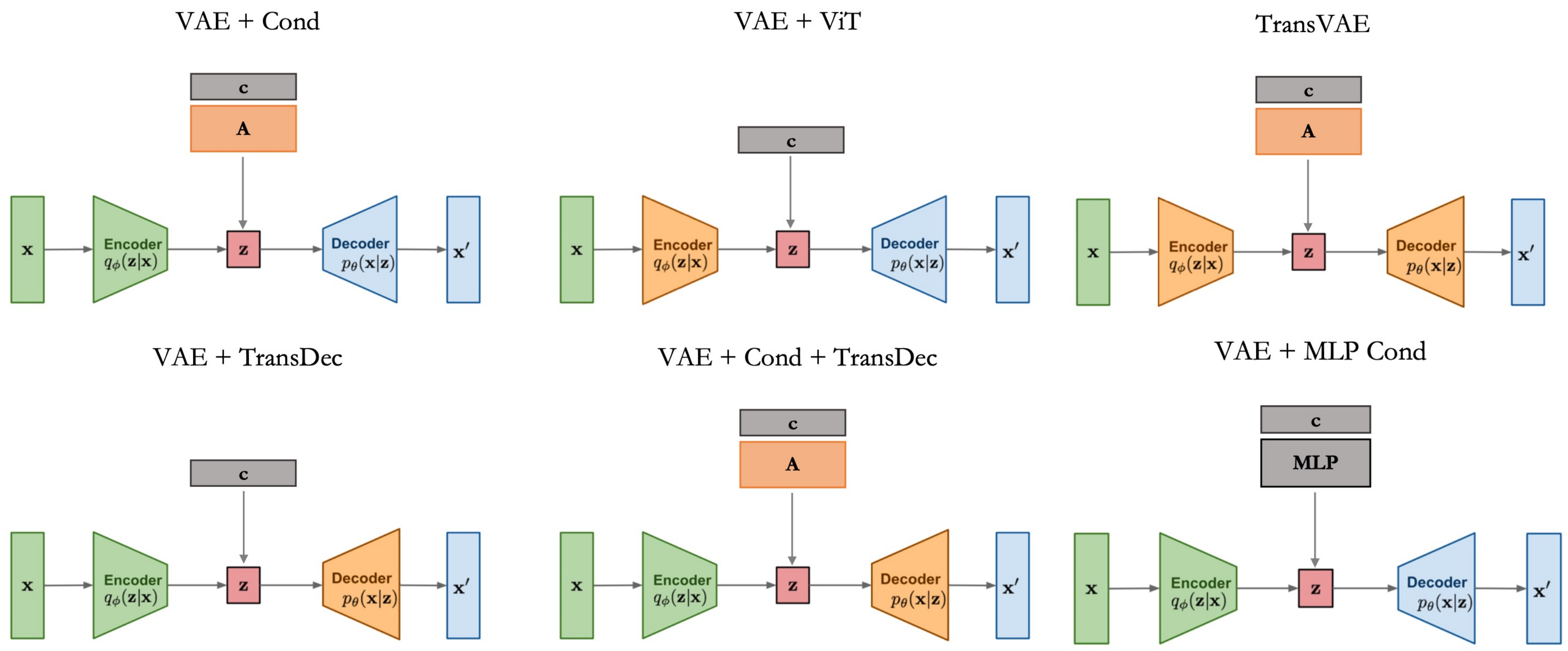
Kamil Deja et al. 2020. *End-to-End Sinkhorn Autoencoder With Noise Generator*. IEEE Access, Vol. 9. [https://gitlab.cern.ch/swenzel/zdcfastsim]
Phillip Lippe. 2022. *UvA Deep Learning Tutorials*. https://uvadlc-notebooks.readthedocs.io

Proposed models



Han Zhang et al. 2018. *Self-Attention Generative Adversarial Networks*. Remote Sensing 14, no. 10: 2425.
Yifan Jiang et al. 2021. *TransGAN: Two Pure Transformers Can Make One Strong GAN, and That Can Scale Up*. CoRR abs/2102.07074.

Proposed models



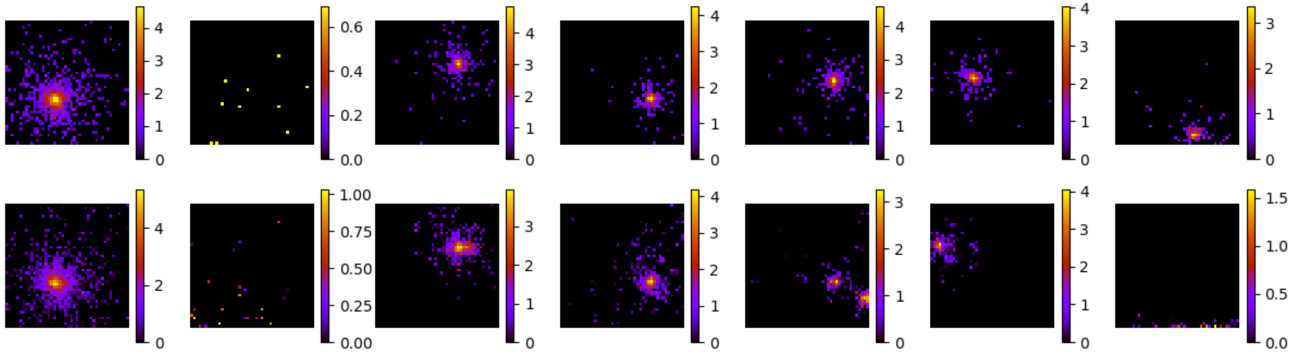
Evaluation results

Model	My results[1]		Article	
	Wasserstein	MAE	Wasserstein	MAE
Original data	–	–	2.89	6.59
GAN[2,4]	13.72	86.20	6.95	68.27
VAE	16.68	18.91	14.92	23.13
SAE	24.09	27.24	7.91	13.50
GAN + Cond[2,4]	14.96	69.40	–	–
GAN + ViT[3]	×	×	–	–
SA-GAN[4]	52.31	86.81	–	–
TransGAN	**3.95**	25.42	–	–
VAE + Cond	6.53	14.78	–	–
VAE + ViT	16.97	19.35	–	–
VAE + TransDec	8.35	15.88	–	–
VAE + Cond + TransDec	8.33	15.16	–	–
TransVAE	9.19	15.25	–	–
VAE + MLP Cond	5.59	**14.61**	–	–

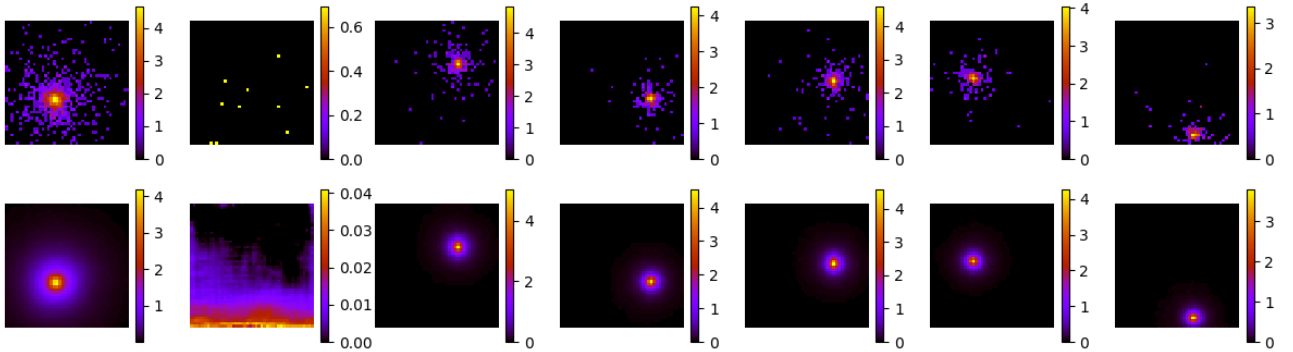
[1] Best result from models saved after each training epoch
[2] Unstable results, metrics vary significantly in successive epochs
[3] Impossible to train, it collapses in the first epochs
[4] The generated images do not visually resemble the original data

Evaluation results

TransGAN
(best Wasserstein)



VAE + MLP Cond
(best MAE)



Conclusions and future work

- GANs are difficult to train and sometimes have poor quality results.
- VAE with MLP network for conditional variables achieves the best results .

- Reproducing the same results as in the article.
- Searching for new models - normalizing flows?
- New metric?