Training Generative Reversible Networks

Robin Tibor Schirrmeister[1,2], Patryk Chrabąszcz[2], Frank Hutter[2] and Tonio Ball[1]

1 - Translational Neurotechnology Lab, University Medical Center Freiburg | 2 – Machine Learning Lab, Albert-Ludwigs-University Freiburg

Code available: github.com/robintibor/generative-reversible

BrainLinks BrainTools acting thoughts

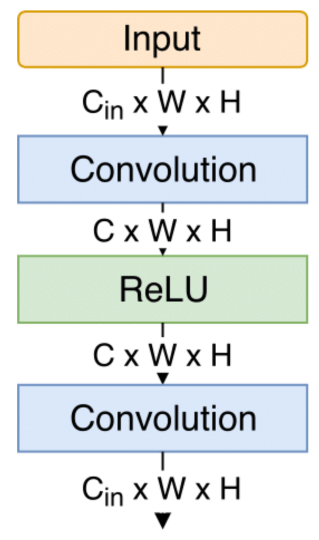
UNIVERSITÄTS KLINIKUM FREIBURG

Training Generative Reversible Networks

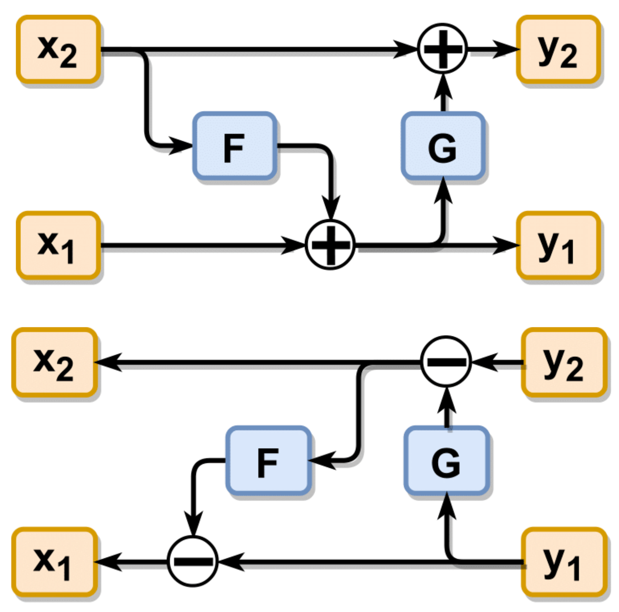
Reversible networks are by-design invertible neural networks, that had previously been trained by maximum likelihood and adversarial approaches on the generated data [1,2,3,4]. Here, we try:

1. Optimal transport on minibatches for MNIST
2. Adversary on latent space (as in adversarial autoencoders) for CelebA

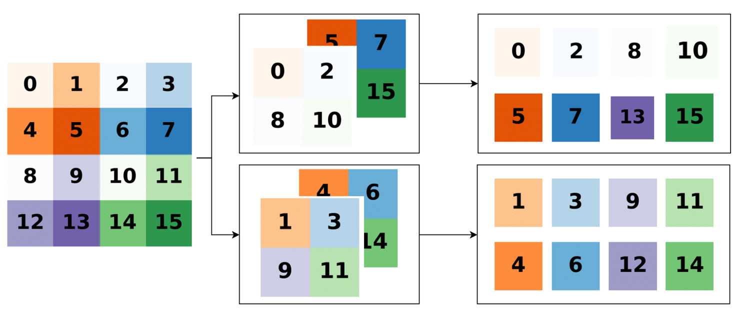
Reversible Network Architecture



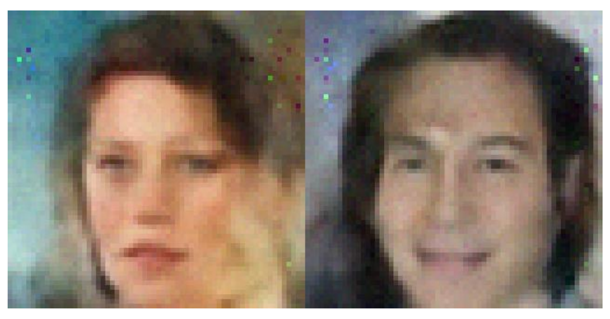
Our F/G functions (for the reversible Block)



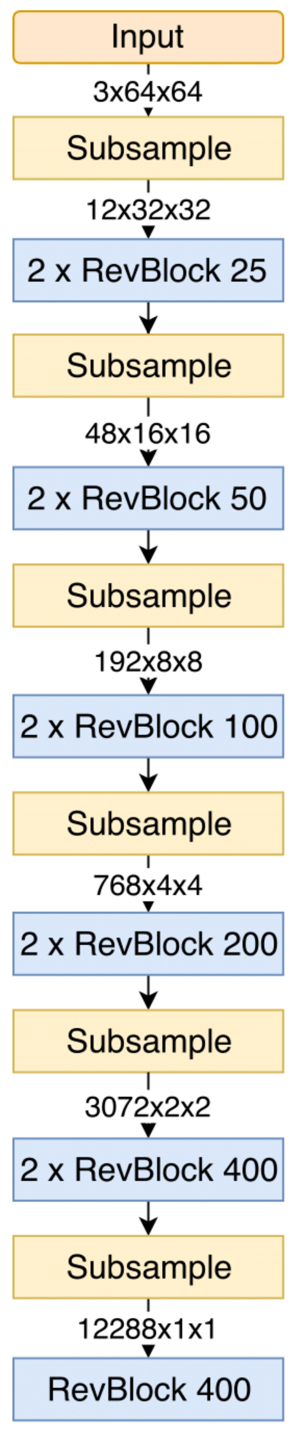
Reversible block (as in prior work)
Top: forward
Bottom: inverse



Our subsampling step
Applied twice to a 4x4 input. On the right, individual squares represent individual channels, so each channel has a single value at the end. At the end, both streams have access to pixels that cover the entire 4x4 input.



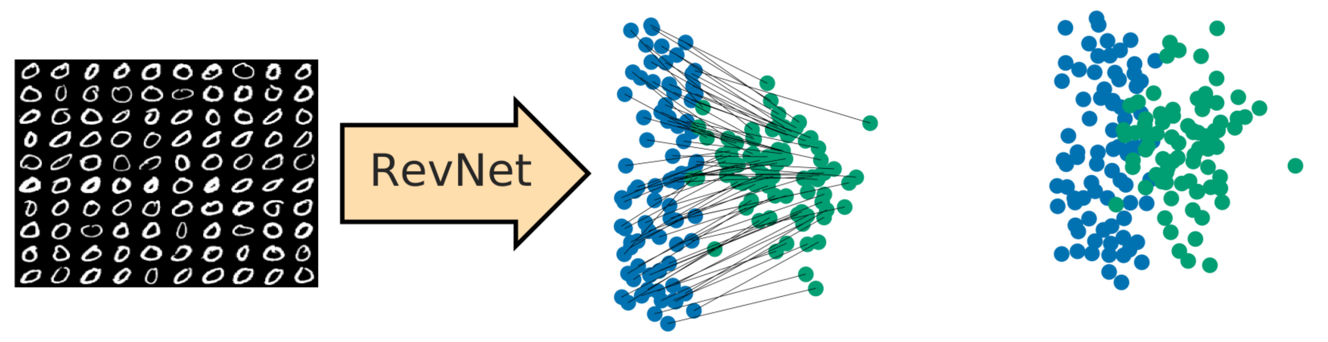
Artifacts in earlier training phases
Generated samples of unconverged RevNet with green and purple artifacts, caused by some latent dimensions strongly influencing latent dimensions they correspond to (details in paper).



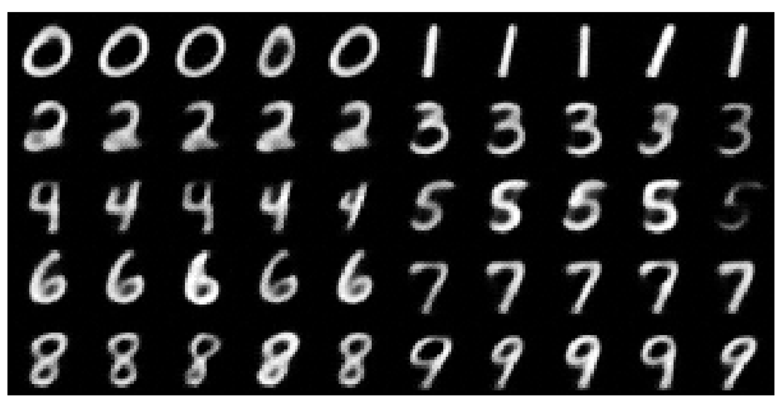
Network architecture
Numbers behind RevBlock indicate number of filters.

Optimal transport on MNIST

Optimize RevNet parameters **and prior distribution parameters** (with one distribution per class) using optimal transport on minibatches:

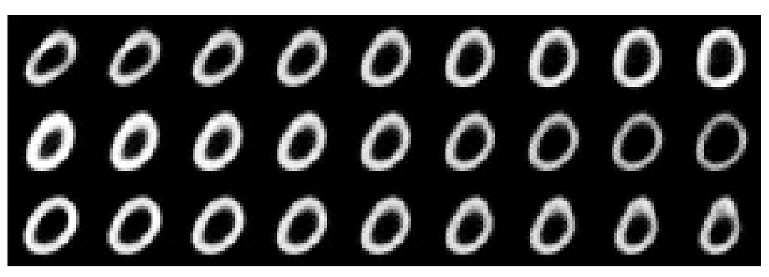


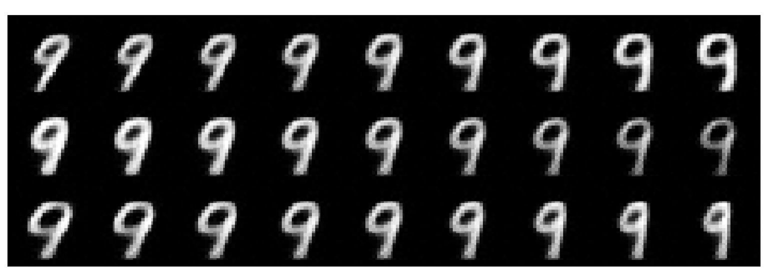
Despite biased gradients, still reasonable results on MNIST:



Samples on MNIST
Realistic digits, somewhat blurry, lack diversity

Varying three dimensions of latent prior
Three dimensions with largest standard deviations seem to roughly correspond to tilt, thickness and size, respectively. Different classes use same dimensions even though all dimensions usable by model.

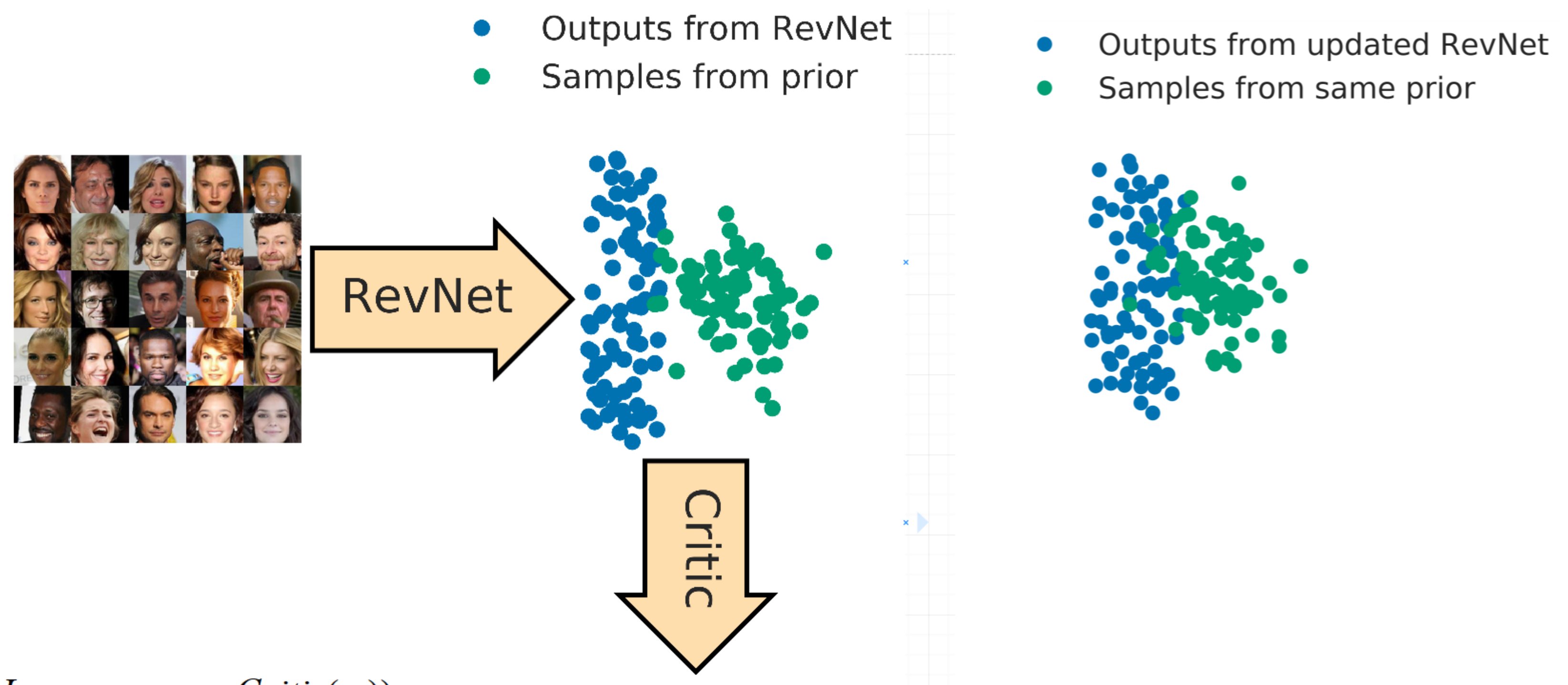




This work was partly supported by the BrainLinks-BrainTools Cluster of Excellence (DFG grant EXC 1086) and by the Federal Ministry of Education and Research (BMBF, grant Motor-BIC 13GW0053D).

Adversarial approach on CelebA

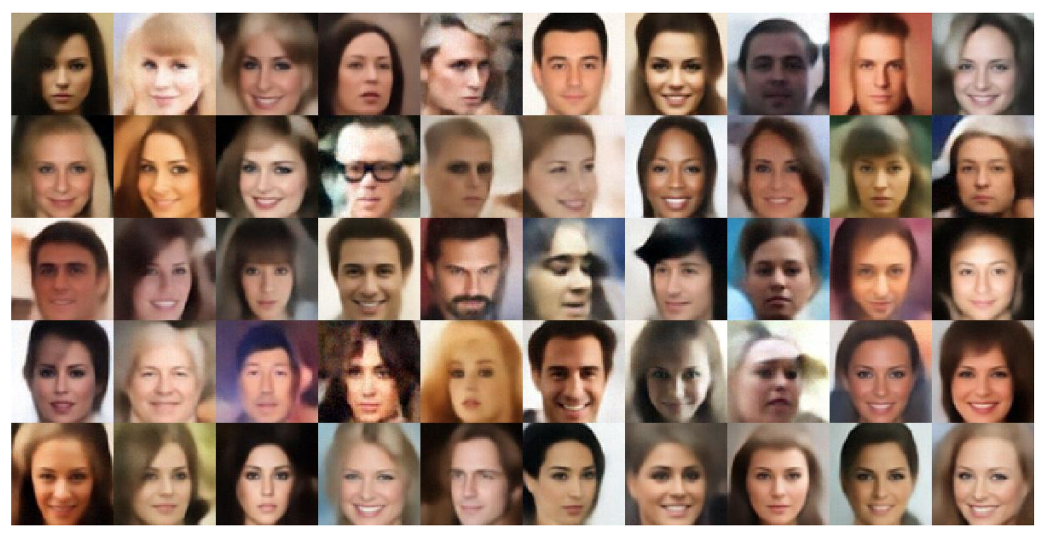
Optimize RevNet parameters using adversary in latent space:



$$Loss_{RevNet} = -Critic(\bullet)$$

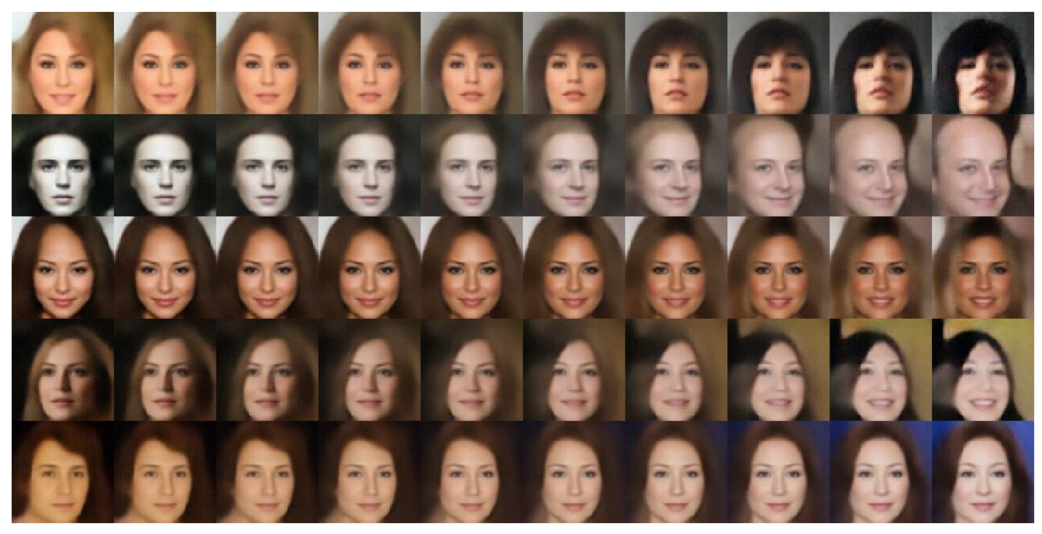
$$Loss_{Critic} = -\min(0, -1 + Critic(\bullet)) - \min(0, -1 - Critic(\bullet))$$

<- Hinge loss



Generated Samples

Reconstructions
Top row: original, middle row: reconstruction from latent space restricted the prior distribution. bottom row: reconstruction from full latent space



<- Interpolations restricted
Interpolations in the encoding space restricted to nozero dimensions of the prior. Intermediate iamges clearly resemble human faces.

Model	Frechet Inception Distance
Variational Autoencoder	63
WAE-MMD	55
WAE-GAN	42
RevNet	65

Frechet Inception Distance
Lower is better. Revnets still slightly underperform compared to WAE-MMD and WAE-GAN.

Interpolations full
Interpolations in the full encoding space. More details, but also more artifacts.

Conclusion

- Reversible Networks can be trained inside adversarial autoencoder framework
- More comparison to prior work on invertible generative models needed
- Optimal transport approach interestingly leads to same dimensions encoding same concepts for different classes
- Optimal transport approach needs further ideas how to scale to larger and more diverse datasets (semi-dual approach?)
- Further improvement of existing hierarchical invertible architectures [2] might be promising

Contact and References

Robin Tibor Schirrmeister, robin.schirrmeister@uniklinik-freiburg.de
Translational Neurotechnology Lab, *www.tnt.uni-freiburg.de*
University Medical Center Freiburg, Engelbergerstr. 21, 79106 Freiburg, Germany

Schirrmeister, R. T., Chrabąszcz, P., Hutter, F., Ball, T. (2018). Training Generative Reversible Networks. ICML 2018 workshop on Theoretical Foundations and Applications of Deep Generative Models

[1] Dinh, L., Krueger, D., and Bengio, Y. NICE: Nonlinear Independent Components Estimation. In International Conference on Learning Representations (ICLR), 2015.
[2] Dinh, L., Sohl-Dickstein, J., and Bengio, S. Density estimation using Real NVP. In International Conference on Learning Representations (ICLR), 2017.
[3] Grover, A., Dhar, M., and Ermon, S. Flow-GAN: Combining Maximum Likelihood and Adversarial Learning in Generative Models. In Thirty-Second AAAI Conference on Artificial Intelligence, 2018.
[4] Danihelka, I., Lakshminarayanan, B., Uria, B., Wierstra, D., and Dayan, P. Comparison of Maximum Likelihood and GAN-based training of Real NVPs. arXiv:1705.05263 [cs], May 2017.

UNI FREIBURG