Andrew Carr
PhD Student

Fashion Proposal

Research Vision: My research focuses on few-shot learning for GANs. Research in GANs has accelerated in recent years, and the state of the art models can produce high fidelity outputs. These generative models, like much of modern deep learning, require vast oceans of data to properly learn the underlying data manifold. As a result, many of the best results in GAN research (1; 2) come from industrial labs with large compute resources and huge data sets. However, this compute cost and data need is out of reach for many working with commodity hardware or in domains where data is scarce. For example, in the fashion world, there are many unlabeled examples, but potentially small curated data sets. A few-shot approach, inspired by recent advances in meta learning and optimal transport theory, would enable anyone with a GPU to produce novel, high quality, data given a small ground truth data set. It would also allow for quality fashion item generation. Additionally, few-shot generation would enable applications in a wide variety of fields where data is scarce (e.g., autonomous vehicles, finance, linguistics).

Lake et. al (3) was one of the first works to introduce a successful few-shot generative model. Their method uses stroke and image information to train a Bayesian model on their novel Omniglot data set. MetaGAN (4) is meta learning approach to few-shot learning. They use adversarial learning with meta learning (5; 6) inspired discriminators. The results are impressive, but they focus primarily on downstream tasks; generation quality is not studied.

To make progress on few-shot GANs I propose a three pronged research approach. These prongs are not three separate directions, but three areas that have results that can be leveraged in parallel for few-shot generative modeling. Specifically, the task can be represented as learning a distribution $p(x)$ where $x \in \mathbb{R}^{n \times m}$ is a natural image of size $n \times m$. The GAN is trained to produce novel samples from $p(x)$ with far fewer examples than traditional training approaches.

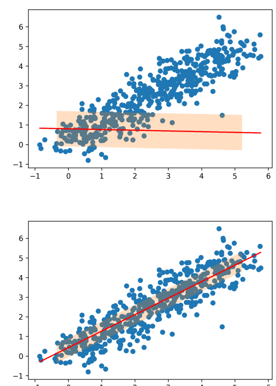
Optimal Transport / Geometry: Optimal transport theory and information geometry provide theoretically sound tools using information from the data (7). These tools can operate on arbitrary aspects of a variety of data driven problems. I plan to investigate the use of optimal transport, and differential geometry (8), on the intrinsic GAN manifold. The purpose of this investigation is to determine if transformations, acting on a trained GAN, can inform fine tuning on a new task. There have been recent, and promising, advances in unbalanced optimal transport (9) that will be key in this work. In the case where there are more samples from a certain task/class than another, unbalanced transport will allow the network to generalize quickly by transferring mass from the currently mastered task to the new task. If the goal is to generate natural images from small class $g(x)$, and we have a GAN trained on a separate task $p(x)$ (of potentially different dimension), we can transport mass from $p(x)$ to $g(x)$. By leveraging this transportation, GANs can likely be trained with far fewer examples.

Meta Learning: By learning to learn, systems designed from a meta learning perspective are especially efficient when encountering new tasks. For example, the Model Agnostic Meta-Learning (5) framework achieves impressive results on a number of supervised and semi-supervised tasks. By learning a good initialization of parameters, good performance is achieved after relatively few standard training steps. I plan to explore the MetaGAN encoder-decoder formulation with a MAML (or other) decoder and many generation techniques used in modern GANs (10). By coupling these meta learning approaches with optimal transport, the amount of training data and compute required for generative models can be greatly decreased.

Metrics and Evaluation: To ensure my methods have high quality generation my research plan is to rigorously test the model outputs with a number of modern metrics. These metrics (e.g., inception score, FID (11), geometry score (12)) are grounded in theory and give a quantitative measure of generation performance. By focusing on the generative performance of these models, we can optimize the algorithmic methods to achieve high fidelity, few-shot, generative models.

Previous Work: My work, thus far, has primarily been focused on applications of optimal transport to a variety of deep learning tasks. Most recently, I have worked with Neural Processes (13). These Neural Processes are a class of generative model that maps a set of context points to a distribution over functions. I have two papers under review.

The first of these works extends Conditional Neural Processes (14) to work on arbitrary graph structured input. The method is structured to learn a distribution over graph edge values using a set of informative input features. It scales linearly with context points and achieves competitive performance.



For the second work, it is important to note that NPs are traditionally trained using Maximum Likelihood with a KL divergence regularization term. However, there are many distributions that have analytically or computationally intractable likelihoods. For example, in the figure to the right (top), we see a distribution with a uniform noise model. The line represents the best fit, and the tube represents the model uncertainty. The model is misspecified and so traditional NPs fail to learn a model for the data. We extend NPs by using Wasserstein Distance to learn such distributions right (bottom). By using this distance, our method has signal regardless of if the likelihood exists or not. Therefore, it can learn an additional class of problems that are out of reach for traditional Neural Processes. In the generative modeling problem, Wasserstein distance and other optimal transport theory results applied to learn transport plans between tasks can enable few-shot learning. By leveraging this work in few-shot learning, I am confident that I can generate interesting, novel, and attractive fashion articles.

References

[1] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation," *CoRR*, vol. abs/1710.10196, 2017.
[2] A. Brock, J. Donahue, and K. Simonyan, "Large scale GAN training for high fidelity natural image synthesis," *CoRR*, vol. abs/1809.11096, 2018.
[3] B. Lake, R. Salakhutdinov, J. Gross, and J. Tenenbaum, "One shot learning of simple visual concepts," in *Proceedings of the annual meeting of the cognitive science society*, vol. 33, 2011.
[4] R. Zhang, T. Che, Z. Ghahramani, Y. Bengio, and Y. Song, "MetaGAN: An adversarial approach to few-shot learning," in *NIPS*, 2018.
[5] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *ICML*, 2017.
[6] F. Sung, Y. Yang, L. Zhang, T. Xiang, P. H. Torr, and T. M. Hospedales, "Learning to compare: Relation network for few-shot learning," in *CVPR*, 2018.
[7] T. Salimans, H. Zhang, A. Radford, and D. Metaxas, "Improving GANs using optimal transport," *arXiv:1803.05573*, 2018.
[8] M. Hauser and A. Ray, "Principles of Riemannian geometry in neural networks," in *NIPS*, 2017.
[9] L. Chizat, G. Peyré, B. Schmitzer, and F.-X. Vialard, "Scaling algorithms for unbalanced optimal transport problems," *Mathematics of Computation*, vol. 87, no. 314, pp. 2563–2609, 2018.
[10] J. Donahue and K. Simonyan, "Large scale adversarial representation learning," *CoRR*, vol. abs/1907.02544, 2019.
[11] M. Lucic, K. Kurach, M. Michalski, S. Gelly, and O. Bousquet, "Are GANs created equal? a large-scale study," in *NIPS*, 2018.
[12] V. Khrulkov and I. Oseledets, "Geometry score: A method for comparing generative adversarial networks," *arXiv:1802.02664*, 2018.
[13] M. Garnelo, J. Schwarz, D. Rosenbaum, F. Viola, D. J. Rezende, S. M. A. Eslami, and Y. W. Teh, "Neural processes," *CoRR*, vol. abs/1807.01622, 2018.

[14] M. Garnelo, D. Rosenbaum, C. J. Maddison, T. Ramalho, D. Saxton, M. Shanahan, Y. W. Teh, D. J. Rezende, and S. Eslami, "Conditional neural processes," *arXiv:1807.01613*, 2018.