Title: Evaluation of Generative Networks for Data Augmentation
Presenter: Spencer Churchill, Irvine Valley College
Mentor: Seth Hochwald

Data collection and preparation is a time and resource consuming process. With an increasing reliance upon statistical models to analyze data, either huge datasets are collected, or models are designed to prevent overfitting. In many cases, collecting supplementary data or designing new models is unfeasible for researchers. Recent applications of game theory to machine learning (Goodfellow, Pouget-Abadie and Mirza) produced a new classification of machine learning models called Generative Adversarial Networks (GANs).

To address occasional limitations of datasets, this research applies a GAN (Radford, Metz and Chintala) to generate varying types of data. By observing errors of the model, slopes of learning rates, and generated data, the ability of generative models to augment datasets is assessed. The results this research presents aim to reduce the costly collection of additional data.

It is found that generative networks create simple data well. Datasets of images without high-level abstractions, such as handwritten numbers formed from lines, is easily generated with high accuracy. Datasets with high-level abstractions (Chao and Churchill), such as faces made of features formed by lines, however, are much harder for GANs to generate due to increased complexity. These machine learning models are relatively new; and, as statistical methods advance, the ability for generative models to augment datasets will increase.

Works Cited

Chao, Brian and Spencer Churchill. Anime Face Dataset. 3. Kaggle, 4 August 2019. Dataset.

Goodfellow, Ian, et al. "Generative Adversarial Nets." Advances in Neural Information Processing Systems 27. Montreal: NIPS, 2014. 1-9.

Inkawhich, Nathan. "DCGAN Tutorial." Vers. 969e467. 26 August 2019. PyTorch. 12 October 2019. <https://pytorch.org/tutorials/beginner/dcgan_faces_tutorial.html>.

Radford, Alec, Luke Metz and Soumith Chintala. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks." arXiv (2016): 1-16.