Synthesizing Time-Series with Auxiliary Classifier Generative Adversarial Networks

Aaqib Saeed, Tanir Ozcelebi
a.saeed@tue.nl, t.ozcelebi@tue.nl

Keywords

deep learning, generative adversarial networks, data augmentation, synthetic data generation, temporal convolutional neural networks

Motivation and Task Description

In recent years, deep learning has shown state-of-the-art performance on an array of problems such as speech recognition, drug discovery, image segmentation and machine translation. This success is mainly attributed to hand curated datasets either by domain experts or through crowdsourcing. However, other fields (e.g. medicine and assisted living) are lagging behind, where, data curation and sharing are limited due to several factors, for instance, privacy and expensive process of collecting labeled datasets. The motivation of this work is, therefore, to leverage Generative Adversarial Networks (GANs) to develop a framework for generating synthetic time series. This data then could be shared, use to resolve class imbalance and provide better insights into the modeling process.

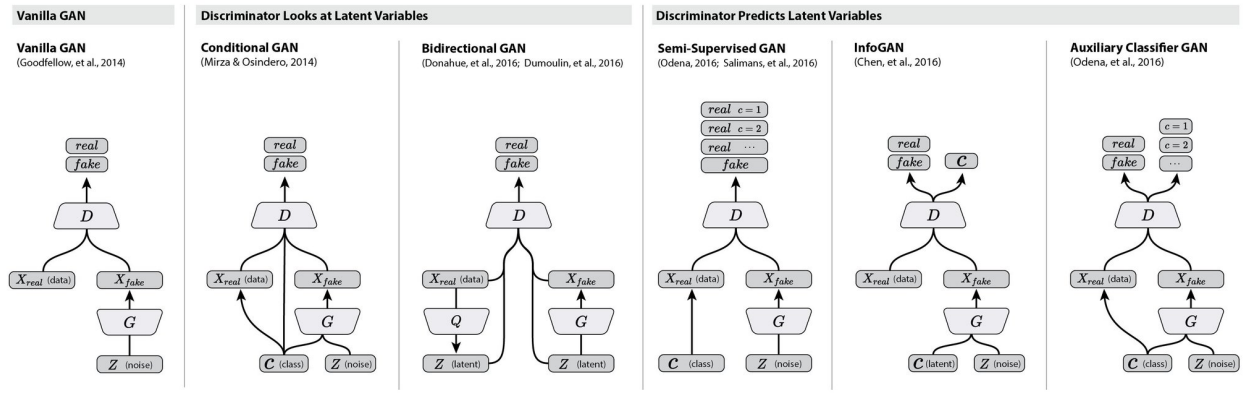

Figure by Chris Olah

In this thesis, the recently proposed variants of *Generative Adversarial Networks (GAN)* [1] particularly *Conditional GAN* [2], *InfoGAN* [3] and *AC GAN* [4] should be evaluated regarding its application of generating (class-conditional) synthetic time series. If necessary, modifications of these standard approaches are required to improve performance in the specific application context. For examining the samples' quality produced by the learned model, it will be evaluated on an end-task, in addition to visual fidelity of samples or other methods [5]. Moreover, to demonstrate the feasibility of the proposed method, diverse data sources from activity recognition domain will be used.

Objectives

- Conduct literature review of existing techniques for augmenting time-series.
- Implement GAN with a temporal convolutional neural network as generator and discriminator to process multi-dimensional time series.
- Compare different variants of GANs and perform hyper-parameter optimization.
- Evaluate performance of GANs through:
    - Training a classifier on a synthetic labeled dataset and evaluating it on a real test set, and vice-versa.
    - Maximum Mean Discrepancy [5, 6].
    - Other suitable techniques.
- Highlight potential limitations and failure modes.

Prerequisites for the student:

- Hands-on experience with machine learning.
- Understanding of (or interested to learn) how Convolutional Neural Networks can be used to process temporal data.
- Nice to have: experience implementing deep neural networks using Tensorflow (and Keras).
- Nice to have: experience with time series datasets and training deep generative models.

Literature

[1] Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.
[2] Mirza, Mehdi, and Simon Osindero. "Conditional generative adversarial nets." arXiv preprint arXiv:1411.1784 (2014).
[3] Chen, Xi, et al. "Infogan: Interpretable representation learning by information maximizing generative adversarial nets." Advances in Neural Information Processing Systems. 2016.
[4] Odena, Augustus, Christopher Olah, and Jonathon Shlens. "Conditional image synthesis with auxiliary classifier GANs." arXiv preprint arXiv:1610.09585 (2016).
[5] Theis, Lucas, Aäron van den Oord, and Matthias Bethge. "A note on the evaluation of generative models." arXiv preprint arXiv:1511.01844 (2015).
[6] Arthur Gretton, Karsten M Borgwardt, Malte Rasch, Bernhard Schölkopf, and Alex J Smola. A kernel method for the two-sample-problem. In Advances in neural information processing systems, pp. 513–520, 2007.