Signal Denoising with CycleGAN

PG2023 Collaboration Meeting

Tianai Ye

tianai.ye@queensu.ca

June 2, 2023

Queen's UNIVERSITY

Introduction

Noise reduction in detector signals can help:

- Improve measurements of pulse amplitudes → better energy resolution
- Identify low-energy signal events that are masked by electronic noise
- Improve background rejection techniques that are based on signal characteristics

Introduction

Motivation on deep learning:

- Frequently used in noise removal in other fields
- Often outperforms traditional denoising methods, e.g. moving average, Savitzky-Golay filtering
- Effective at denoising 1-dimensional electronic signals from HPGe detector
    - Our previous work using a convolutional autoencoder: *Anderson et. al. Eur. Phys. J. C 82, 1084 (2022)*

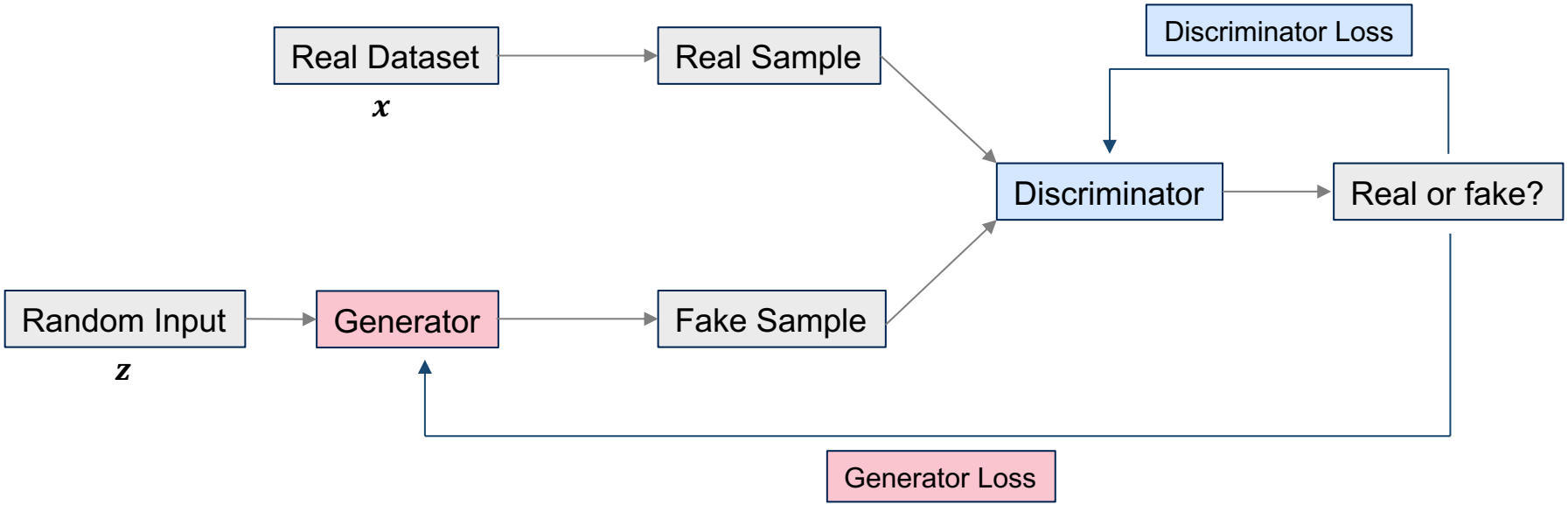
Current work:

- Denoising using *unpaired* data -> a noisy sample pulse does not need a corresponding target clean pulse for training

Ultimate goal:

- Denoising without the need of simulation

GAN - Generative Adversarial Network



[arXiv:1406.2661]

GAN - Generative Adversarial Network

Discriminator Loss

$$\arg\max_{D} \mathrm{E}_{z,x}[\log D\big(G(z)\big) + \log(1 - D(x))]$$

generated data estimated as fake

real data estimated as real

Generator Loss

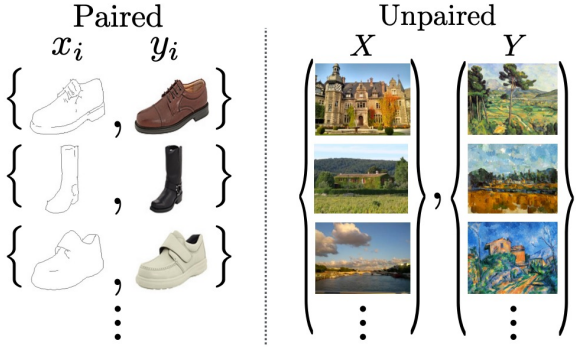
$$\arg\min_{G} \mathrm{E}_{z,x}[\log D\big(G(z)\big) + \log(1 - D(x))]$$

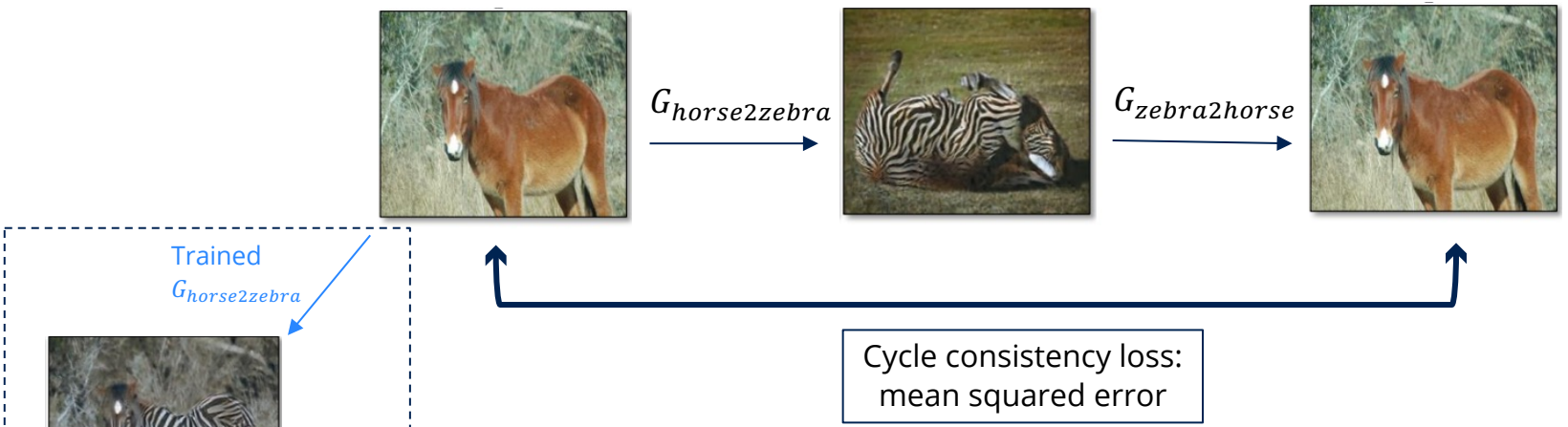
$$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$$

[arXiv:1406.2661]

CycleGAN

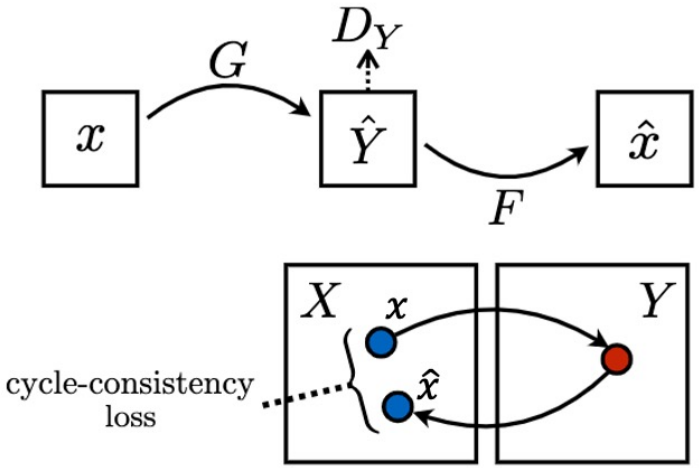
- Learns transformations across domains using **unpaired** data
- A cycle consistency is enforced between the X and Y mappings





[arXiv:1703.10593]

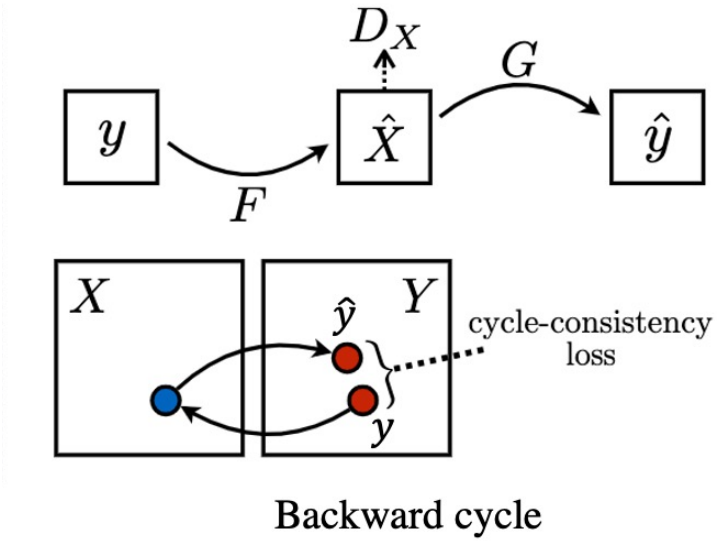
CycleGAN



Forward cycle

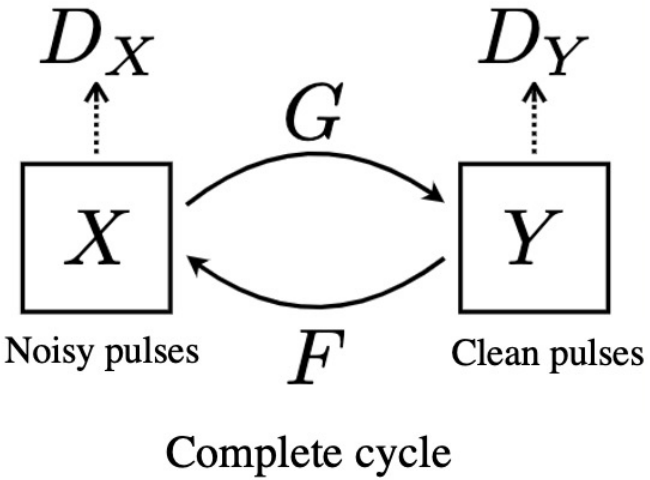
[arXiv:1703.10593]

CycleGAN



[arXiv:1703.10593]

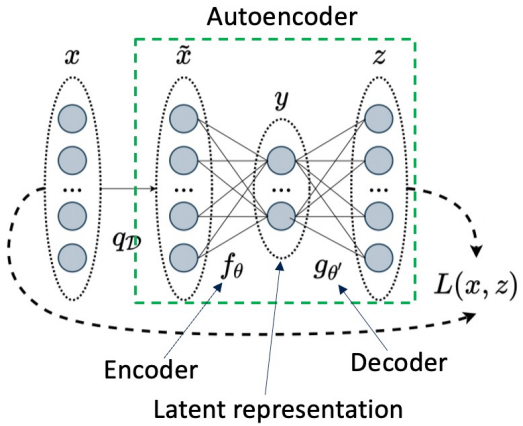
CycleGAN



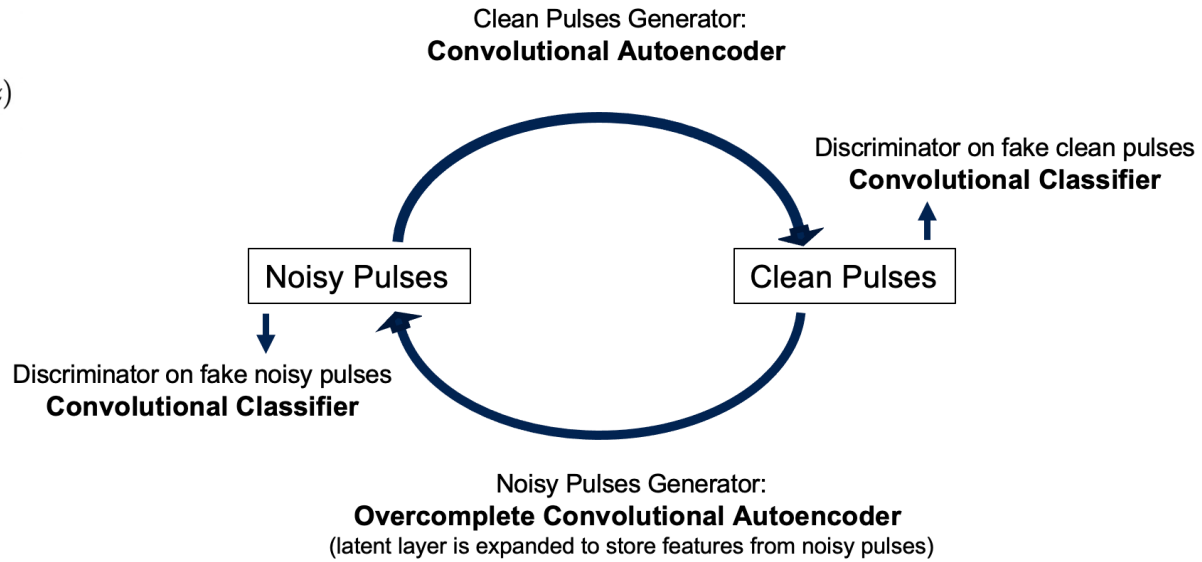
Complete cycle

[arXiv:1703.10593]

CycleGAN – Denoising Network Architecture



[arXiv:2204.06655]

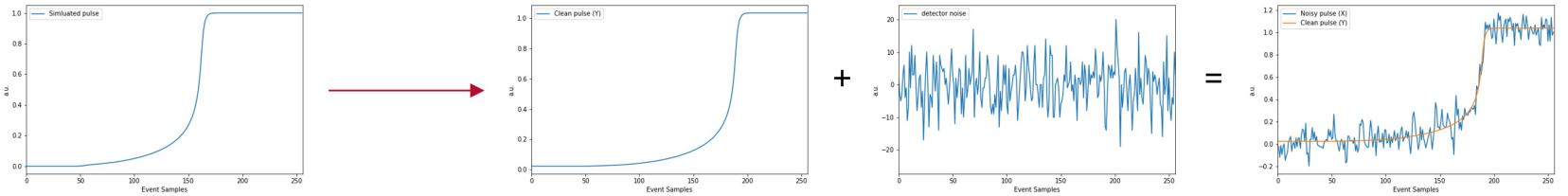


Data Preprocessing and Augmentation

scaled and shifted horizontally and vertically

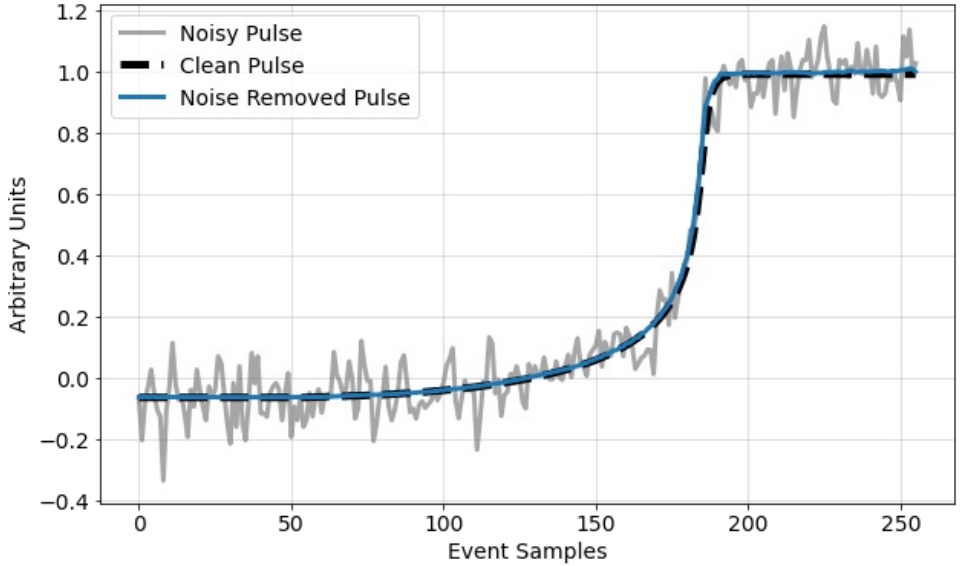
scaled to the same amplitude as the noise in 60keV pulses

Simulated pulse → Clean pulse (Y) + Real detector noise = Artificial noisy pulse (X)



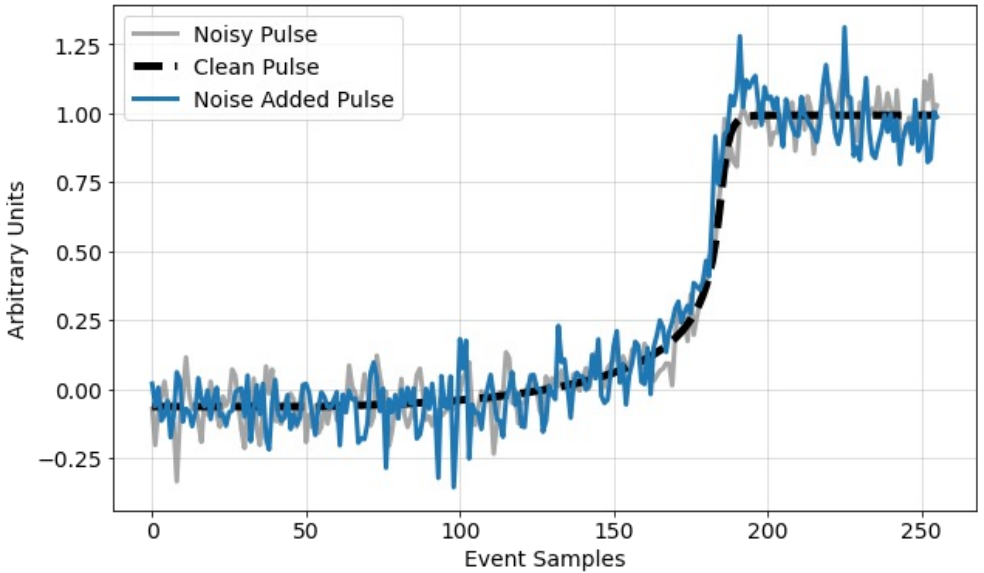
- Training dataset: ~1.3 million pulse pairs (75k multi-site events)
- Testing dataset: ~ 215k pulse samples (12.5k multi-site events)

Preliminary Results – noise removal generator



- Mean squared error of $G_{noisy2clean}$, evaluated by denoising the artificial noisy test dataset, is on the order of 10^{-4}
- Mean squared error achieved in our previous work using an autoencoder is on the order of 10^{-5}

Preliminary Results – noise addition generator



Next Steps

- Stabilize and optimize the training of GAN and CycleGAN
  - Wasserstein loss function with a gradient penalty

- Denoising without the need of simulated pulses (i.e. no target label/ground truth)
  - CycleGAN
    - Trained using pairs of detector noise traces and real pulse signals
    - Learns the underlying true waveform and the additive electronic noise
  - DualGAN : arXiv:2007.01575