Reducing Systematic Differences between Data and Simulation with Generative Models

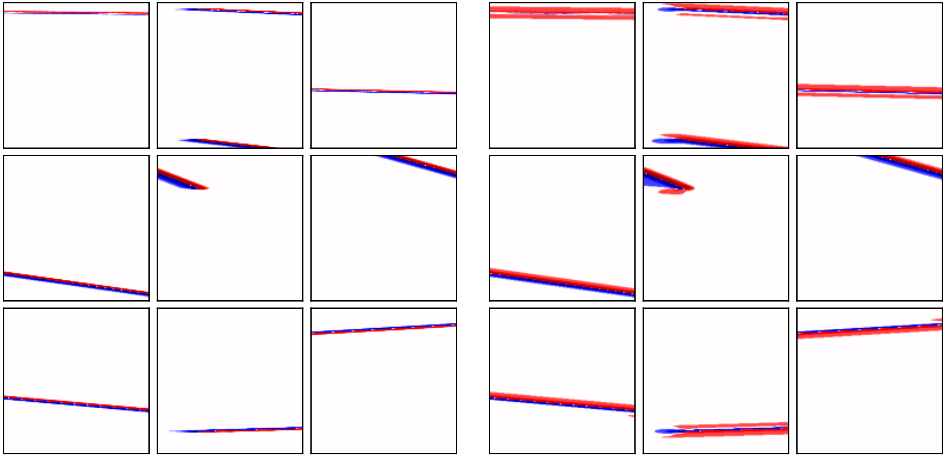
Dmitrii Torbunov,
Brookhaven National Laboratory

March 12, 2024

Introduction

- We use simulation to develop reconstruction algorithms for Liquid Argon Time Projection Chamber particle detectors (DUNE experiment).
- Simulation is not exact – has multiple systematic differences with experimental data.
- Since reconstruction algorithms trained on the simulation, sysmatic differences will degrade their performance on the real data – **domain shift problem**.
- In this talk, I will discuss our attempts to develop a generative algorithm to reduce systematic differences between experimental data and simulation.

Simplified Track Dataset (SLATS) aka "Toy" Dataset



(a) **Domain A** (low-quality simulation) (b) **Domain B** (high-quality simulation)

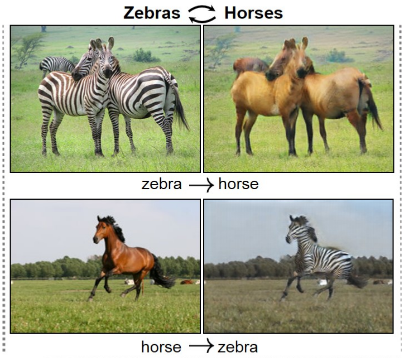
Unpaired Image-to-Image Translation

- We tried to apply Generative Models to translate make images from **Domain A** look like **Domain B** and visa versa.
- **Key Feature** of the problem: translation must be learned in an **unsupervised (unpaired)** way.
- Of the possible Unpaired Image-to-Image (UI2I) translation models there is one model that has drawn our attention due to its inherent properties – the **CycleGAN** model [1].

[1] `arXiv:1703.10593`

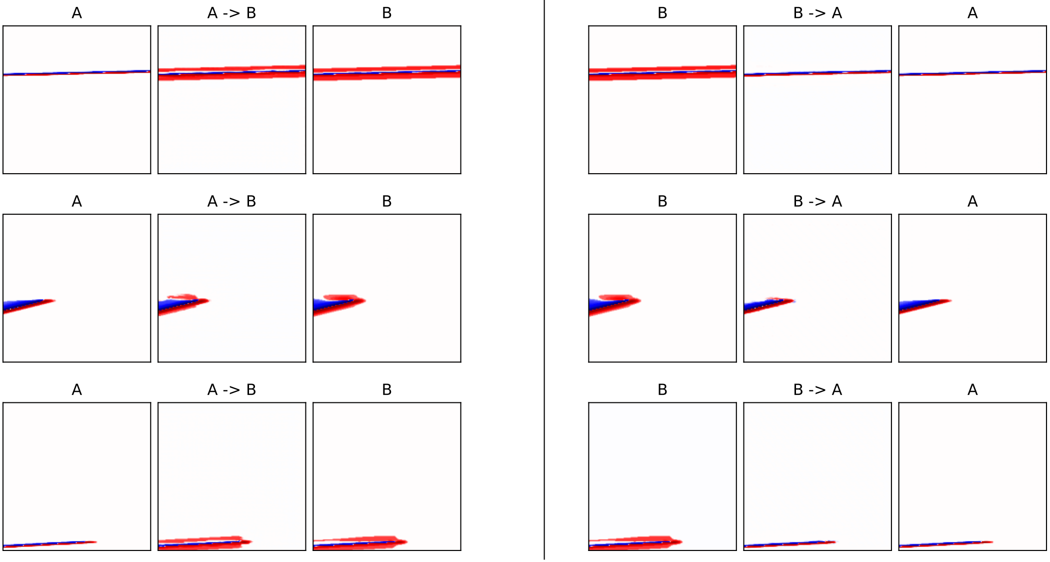
CycleGAN for UI2I

- **CycleGAN** is one of the earliest successful Unpaired Image-to-Image transfer architectures. It solves the mode collapse problem by imposing the **cycle-consisntency constraint**.
- CycleGAN **translaltes** only systematically different parts of the images (**attributes**).
- CycleGAN **preserves** preserves features commom between the two domains (**content**).
- Will CycleGAN work on scientific data?

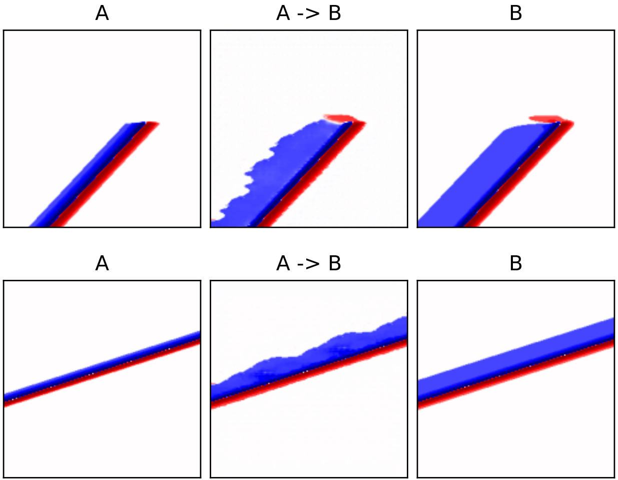


CycleGAN Teaser

Initial CycleGAN Usage Attempt



Initial CycleGAN Usage Attempt – Chunky Tracks



The default CycleGAN experiences difficulties with a translation of wide tracks

CycleGAN Translation Issues

- The default CycleGAN gives good translation on average, yet produces noticeable artifacts for some cases.
- We developed a modified CycleGAN model – UVCGAN, that fixes the failure cases.
- The UVCGAN model includes a new generator architecture and better training procedures.

UNet-ViT Fixing Bad Translations

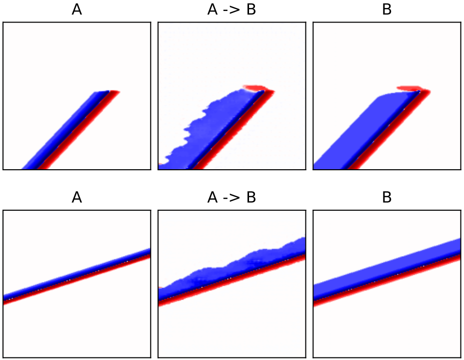


Figure: CycleGAN

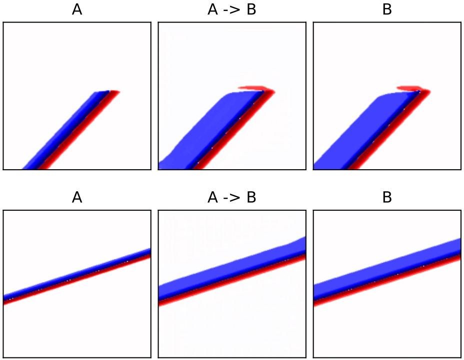


Figure: UVCGAN

Domain Shift Reduction

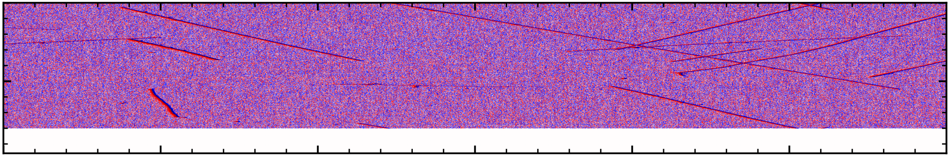
Algorithm	arXiv	A to B ℓ_1	A to B ℓ_2	B to A ℓ_1	B to A ℓ_2	Domain Shift Error MAP [%]
Baseline	-	-	-	-	-	1.904
CycleGAN	1703.10593	0.074	0.180	0.061	0.159	0.735
ACL-GAN	2003.04858	0.083	0.566	0.039	0.121	0.582
U-GAT-IT	1907.10830	0.078	1.187	0.073	1.161	0.713
UVCGAN (ours)	2304.12858	**0.030**	**0.033**	**0.025**	**0.027**	**0.391**

Publication: "Unpaired Image Translation to Mitigate Domain Shift in Liquid Argon Time Projection Chamber Detector Responses"

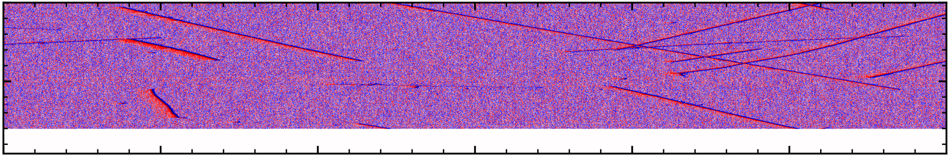
Moving to Realistic DUNE Data

Toy B 

A



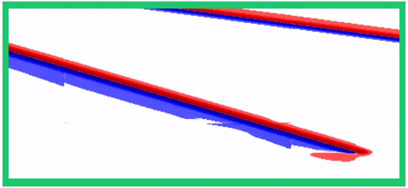
B



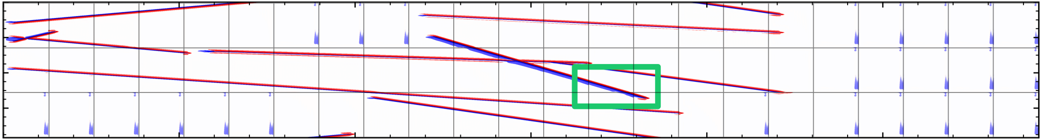
Realistic data samples are about 90 times larger than the toy data

Naive Chunking

- Naive chunked evaluation produces boundary artifacts.
- These articacts are too significant to ignore.
- Solution – use adversarial boundary merging.

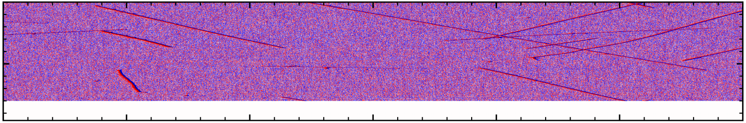


Zoom-In Chunking Artifacts

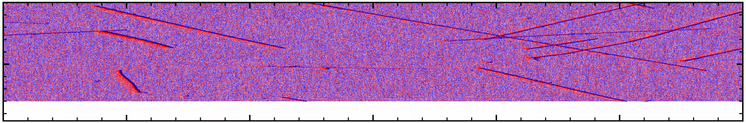


Example of Chunked Translation

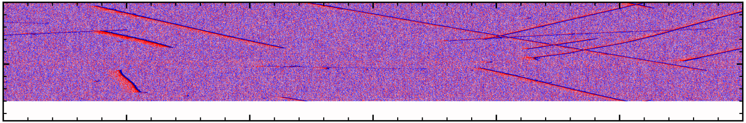
A



A -> B



B



Adversarial boundary merging fixes border artifacts

Summary

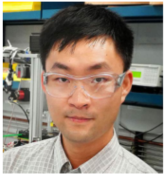
1. CycleGAN-like models seem to work reasonably well for reducing domain shift effects.
2. They can be modified to translate (3, 960, 6000) DUNE simulation images under limited hardware resources.
3. We are working to extend this algorithms to translate between the actual simulation and real data.

Our Team


Dmitrii Turbunov


Yi Huang

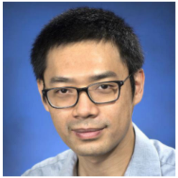

Haiwang Yu


Jin Huang


Shinjae Yoo

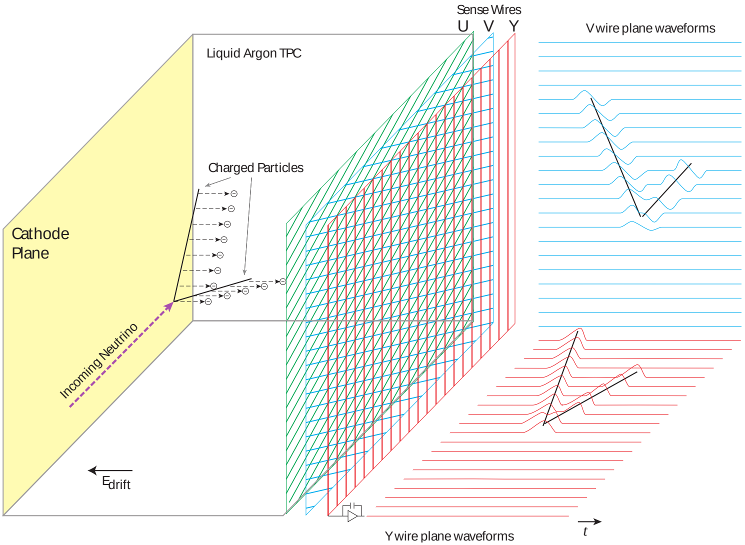

Meifeng Lin

Brett Viren


Yihui Ren

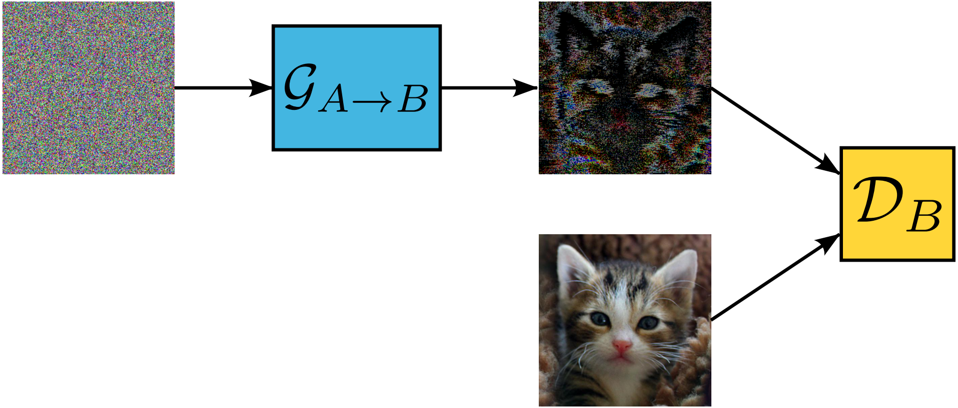
Backups

Liquid Argon Detector



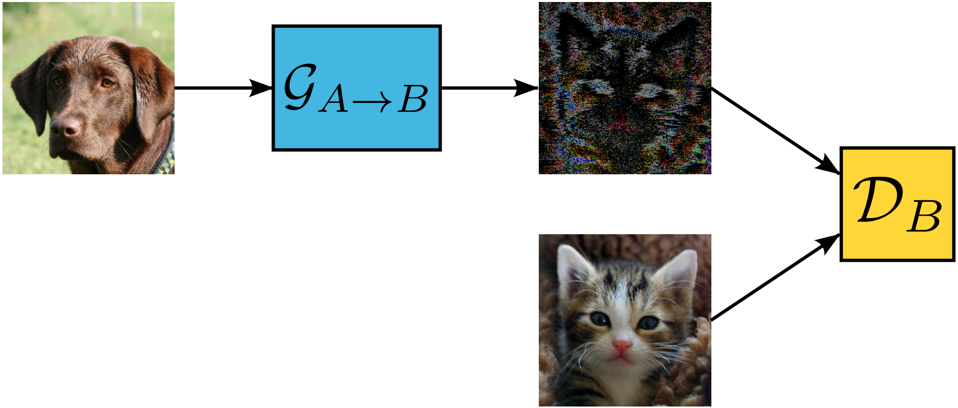
Source: DOI:10.1088/1748-0221/12/02/P02017

CycleGAN is based on GAN



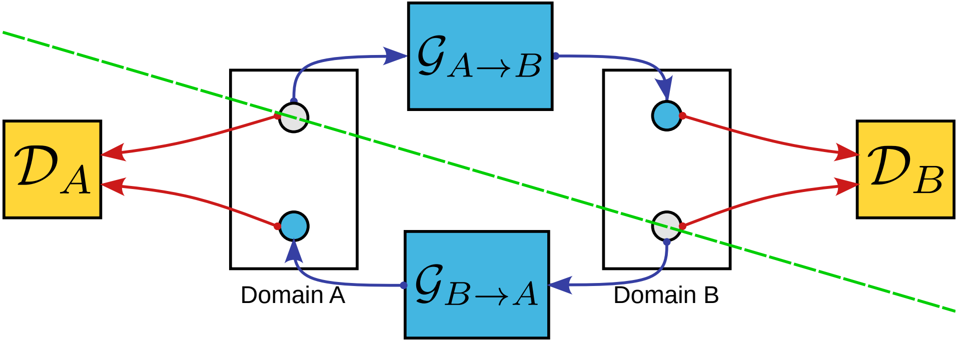
Generator $\mathcal{G}_{A\to B}$ translates random noise into images.
Discriminator $\mathcal{D}_B$ is trained to find differences between real and generated images.
$\mathcal{G}_{A\to B}$ is trained to make generated images indistinguishable from the real ones.

Old Idea – GANs for Unpaired Image-to-Image Translation (UI2I)



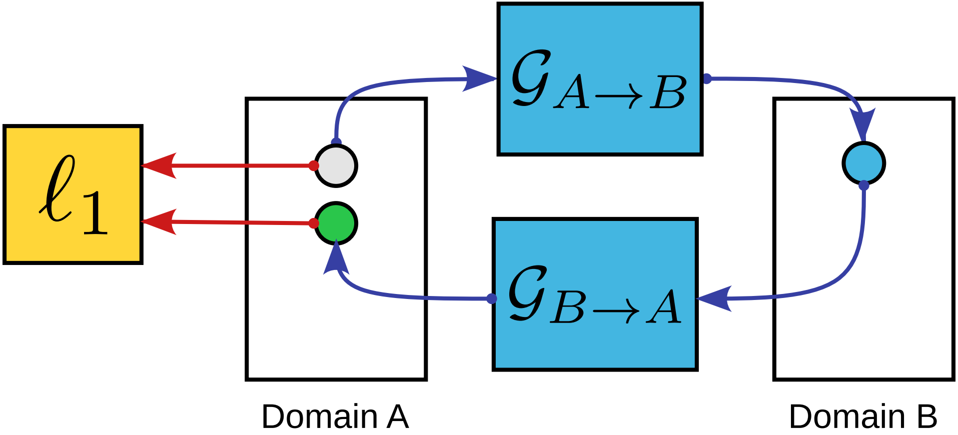
In theory, GAN can be trained to translate images between different domains (e.g. **Dog** to **Cat**)

CycleGAN Architecture



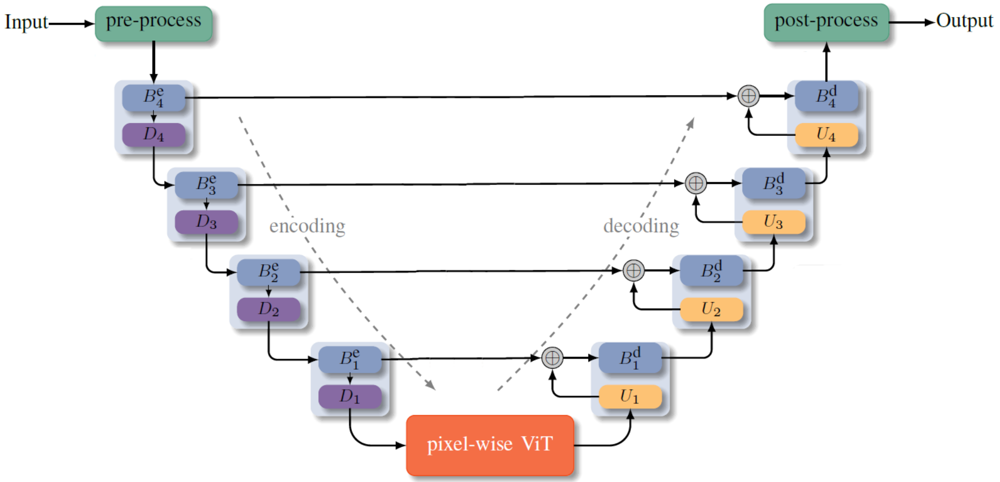
CycleGAN combines two GANs $(\mathcal{G}_{A\to B}, \mathcal{D}_B)$ and $(\mathcal{G}_{B\to A}, \mathcal{D}_A)$

CycleGAN Cycle Constraint



To solve the Mode Collapse problem, CycleGAN enforces a cycle-consistency constraint: $\mathcal{G}_{B\to A}(\mathcal{G}_{A\to B}(X_A)) == X_A$, and symmetrically in for X_B.

UNet-ViT Generator

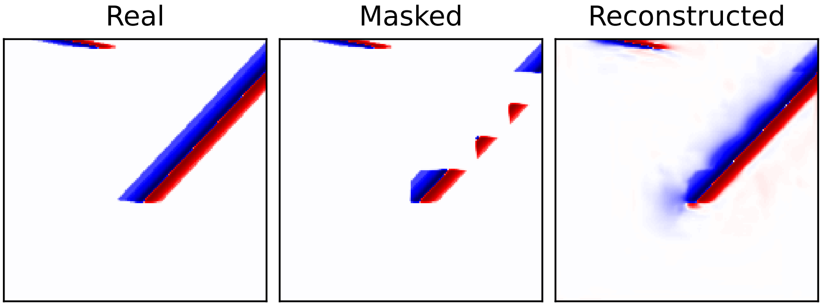


We augmented UNet by a Vision Transformer bottleneck to handle long-range dependencies.

CycleGAN Modifications, 1

- To further improve the CycleGAN performance we have tried to pre-train its generators.
- To pre-train the generators we focused on a task of a **self-supervised image inpainting**.

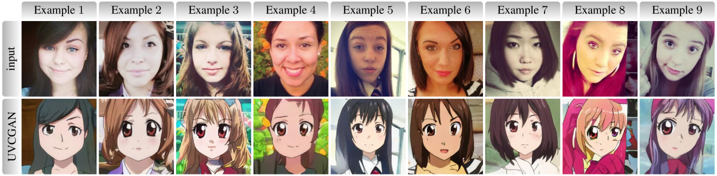
Pre-Training Generators by Image Inpainting



For the image inpainting task, some parts of the image are masked. The network is tasked with recovering the masked parts of the image from their surroundings

UVCGANv1

- We tested our architecture modifications on the standard benchmarking datasets.
- Our model achieves State of the Art performance on 4 of 6 standard benchmarks.
- Publication: "UVCGAN: UNet Vision Transformer cycle-consistent GAN for unpaired image-to-image translation" (`arXiv:2203.02557`).
- Code: `https://github.com/LS4GAN/uvcgan`
- Can we do even better?

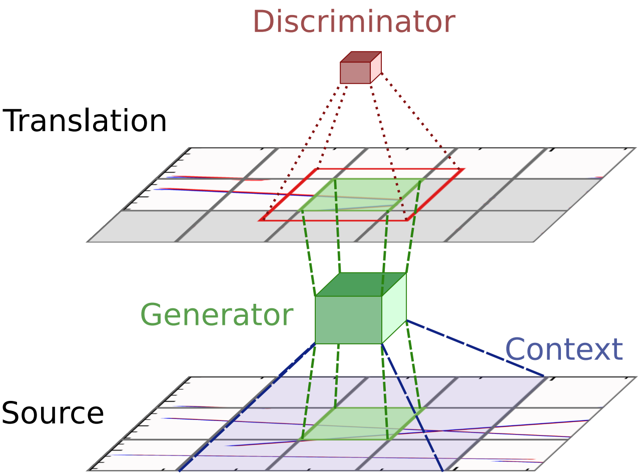


	Selfie to Anime	
	FID	KID ($\times 100$)
ACL-GAN	99.3	3.22 ± 0.26
Council-GAN	91.9	2.74 ± 0.26
CycleGAN	92.1	2.72 ± 0.29
U-GAT-IT	95.8	2.74 ± 0.31
UVCGAN	**79.0**	$\mathbf{1.35 \pm 0.20}$
	Male to Female	
	FID	KID ($\times 100$)
ACL-GAN	**9.4**	$\mathbf{0.58 \pm 0.06}$
Council-GAN	10.4	0.74 ± 0.08
CycleGAN	15.2	1.29 ± 0.11
U-GAT-IT	24.1	2.20 ± 0.12
UVCGAN	9.6	0.68 ± 0.07
	Remove Glasses	
	FID	KID ($\times 100$)
ACL-GAN	16.7	0.70 ± 0.06
Council-GAN	37.2	3.67 ± 0.22
CycleGAN	24.2	1.87 ± 0.17
U-GAT-IT	23.3	1.69 ± 0.14
UVCGAN	**14.4**	$\mathbf{0.68 \pm 0.10}$

Moving to Realistic DUNE Data

- So far, we have been developing algorithms on toy image crops of size $(256, 256)$ pixels.
- Real simulation is more complicated:
  - Real physics, instead of simplified line tracks.
  - Noise
  - Image shape: $(1, 256, 256) \rightarrow (3, 960, 6000)$.
- Large image shape was a problem due to GPU memory limits.

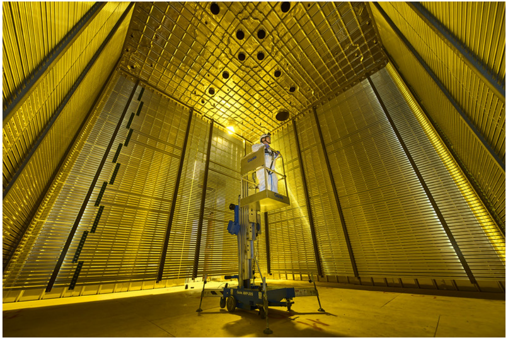
Detecting Boundary Artifacts with Discriminator



Generator translates small Chunk of the image, paying attention to a larger Context.
Discriminator sees a large area around the Chunk detecting boundary artifacts.

DUNE Experiment

- Deep Underground Neutrino Experiment (DUNE) – next-generation neutrino oscillation experiment.
- Uses state-of-the-art Liquid Argon Time Projection Chamber (LArTPC) techology.
- Our simulation of the detector is not accurate and subject to many systematic uncertainties.



LArTPC Detector