Bridge Simulation to Experiments: Unsupervised Unpaired Data Translation between Domains

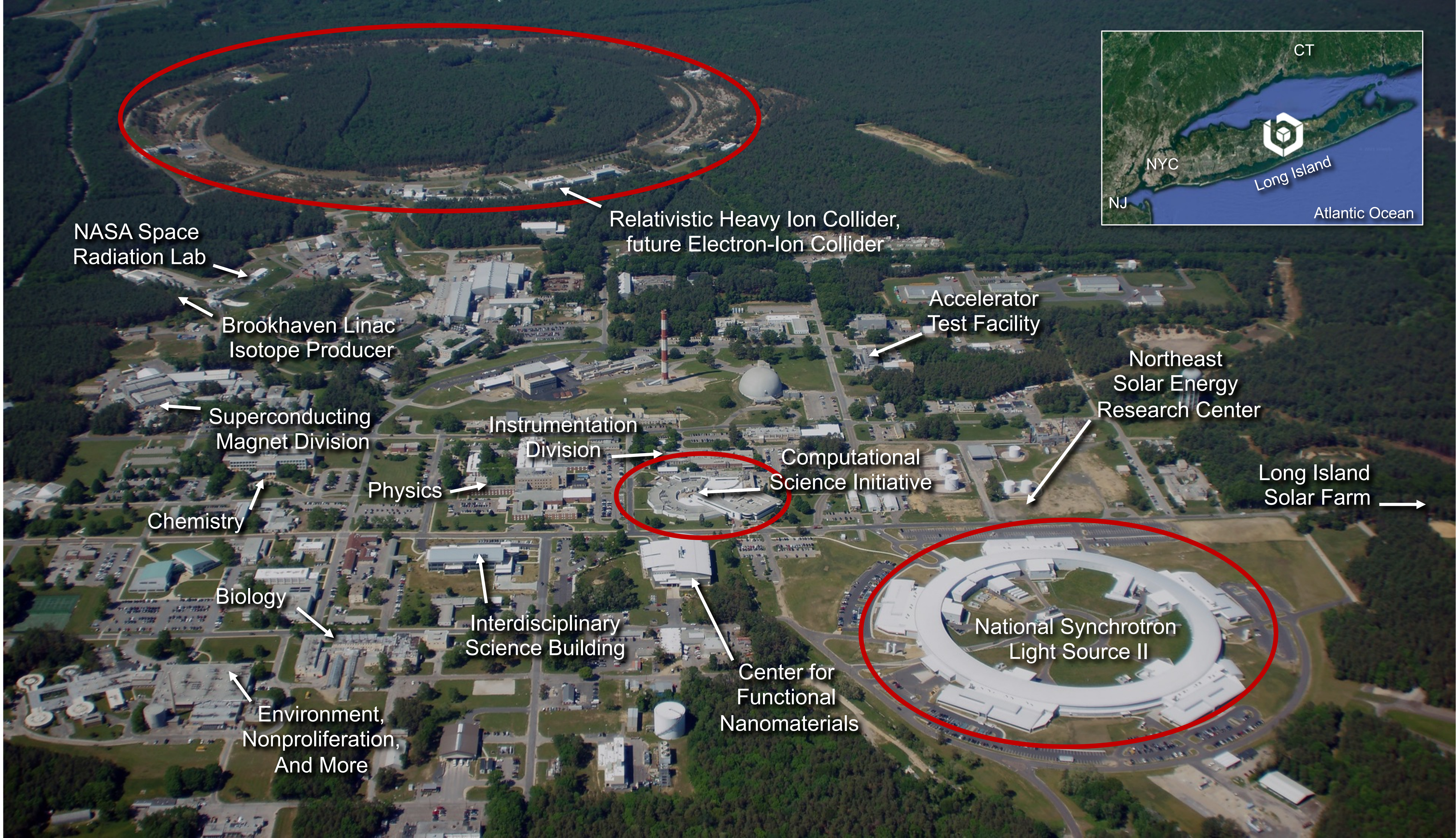
Yihui "Ray" Ren (yren@bnl.gov)

Computational Science Initiative (CSI), BNL

Dmitrii Torbunov, Yi Huang, Haiwang Yu, Jin Huang, Shinjae Yoo, Meifeng Lin, Brett Viren

10th workshop of the APS Topical Group on Hadronic Physics

Apr 12 – 14, 2023


CT
NYC
NJ
Long Island
Atlantic Ocean
Relativistic Heavy Ion Collider,
future Electron-Ion Collider
NASA Space
Radiation Lab
Brookhaven Linac
Isotope Producer
Accelerator
Test Facility
Northeast
Solar Energy
Research Center
Superconducting
Magnet Division
Instrumentation
Division
Computational
Science Initiative
Long Island
Solar Farm
Physics
Chemistry
Biology
Interdisciplinary
Science Building
Center for
Functional
Nanomaterials
National Synchrotron
Light Source II
Environment,
Nonproliferation,
And More

Computer Science Initiative (CSI) at BNL

Departments:

- Scientific Data and Computing Center (SDCC)
- Computation and Data-Driven Discovery (C3D)
- Computing for National Security
- Computational Science Laboratory (HPC, Quantum)
- Computer Science and Applied Mathematics (**ML**, Math)

Brookhaven National Laboratory

https://www.bnl.gov/compsci/

Machine Learning Group

Personnel:

- Shinjae Yoo (Group Leader)
- 8 staff scientists
- 7 postdocs (+2 onboarding)
- 3 software engineers

Project range:

- ASCR, BER, OE, NP, HEP
- NNSA, SciDAC
- LDRD

MACHINE LEARNING
Shinjae Yoo
Group Leader

Dakota Blair	Chuntian Cao
Matthew Carbone	Xin Dai
Thomas Flynn	Yi Huang
Patrick Johnstone	Ai Kagawa
Shubha Kharel	Yuewei Lin
Xihaier Luo	Yihui (Ray) Ren
Carlos Soto	Huan-Hsin Tseng
Ziming Yang	Xi Yu

Motivation

We aim to reduce Simulation & Experiment Data discrepancies.

"All models are wrong, but some are useful". George E. P. Box

Simulations:
- Can get the fundamentals correct,
- Inexpensive to run,
- Freedom of choosing parameters.

Experiments:
- Evidence for scientific advancement,
- Very expensive to run,
- "Ground truth" unknown

Gap: there are multiple discrepancies between the simulated and the real data.

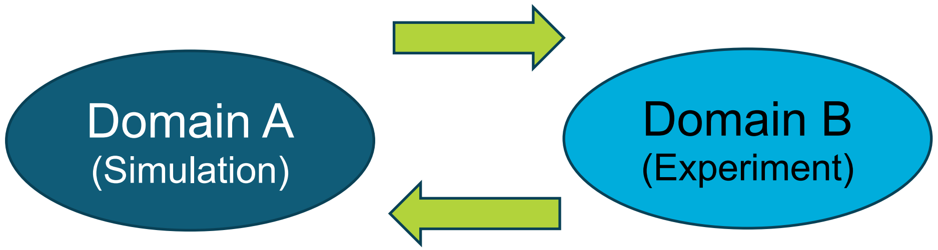
Motivation

Domain A
(Simulation)

Domain B
(Experiment)

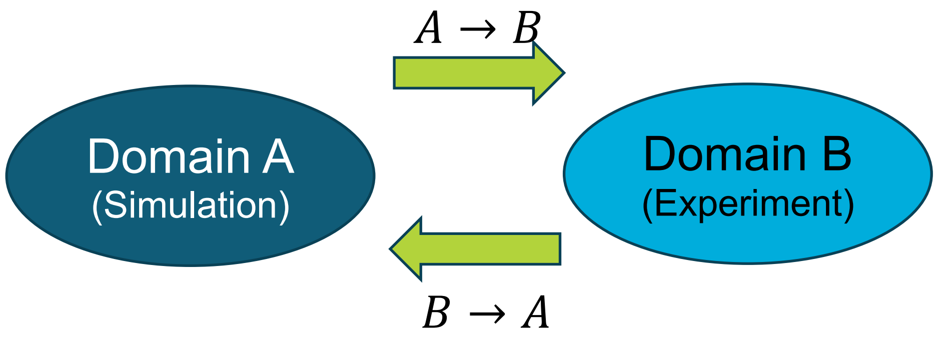
- Cause: difference between two data distributions
- Existing remedies:
  - Data Augmentation. (Heuristics, domain-agnostic, use case-dependent.)
  - Domain Adaptation. (Task-specific, required trained model, require data annotation.)
  - Transfer Learning. (Require data annotation.)

- Cause: Difference between two data distributions
- Existing Remedies:
  - Data Augmentation.
  - Domain Adaptation.
  - Transfer Learning.
- **<u>Unpaired Unsupervised Data Domain Translation</u>.**



	Task Agnostic	Model Agnostic	Domain Aware	Source Domain Label Free	Target Domain Label Free
Heuristic Data Augmentation	✅	✅	❌	✅	N/A
Domain Adaptation (DA)	❌	❌	✅	❌	❌
Unsupervised DA	❌	❌	✅	❌	✅
Transfer Learning	❌	❌	✅	❌	❌
Data Domain Translation (Ours)	✅	✅	✅	✅	✅

Impact

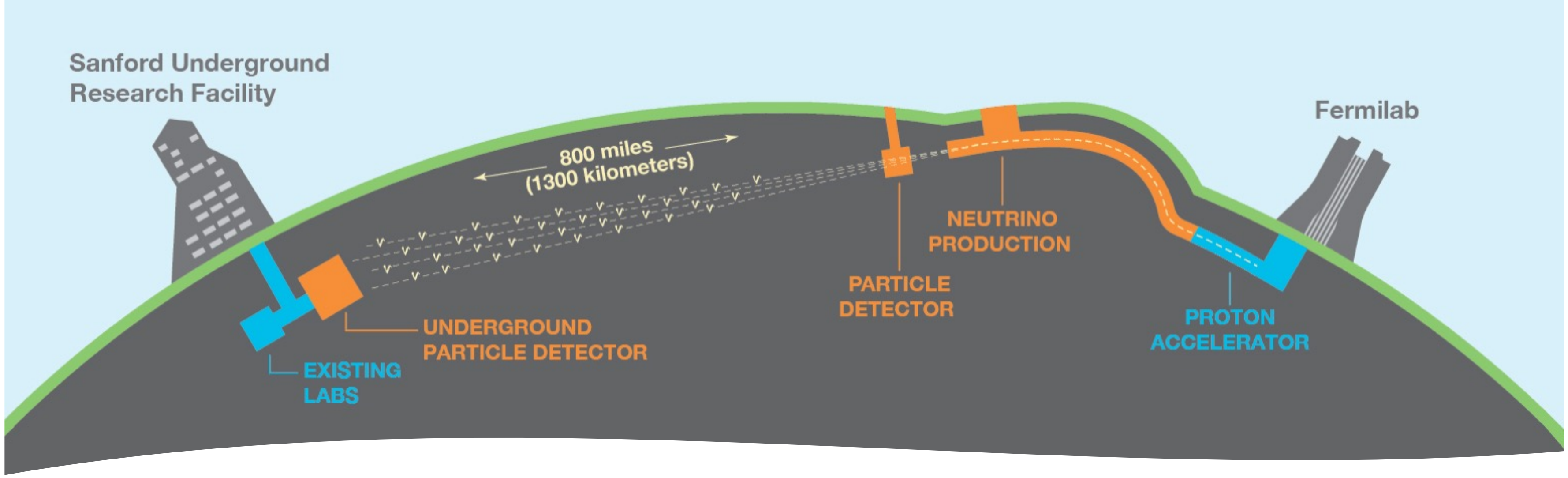


- $A \rightarrow B$: “Augmented High-Fidelity Simulation” that can produce “labeled” data.
- $B \rightarrow A$: “Data Cleaning” that can remove noise of experiment data.
- Analysis tools (w/ human-intelligence) can have better and more data to work with.
- ML models have labeled data to train and are easier to transfer to the real data.

DUNE and LArTPC

Deep Underground Neutrino Experiment (DUNE) is an experiment funded by US DOE, CERN and other international partners.

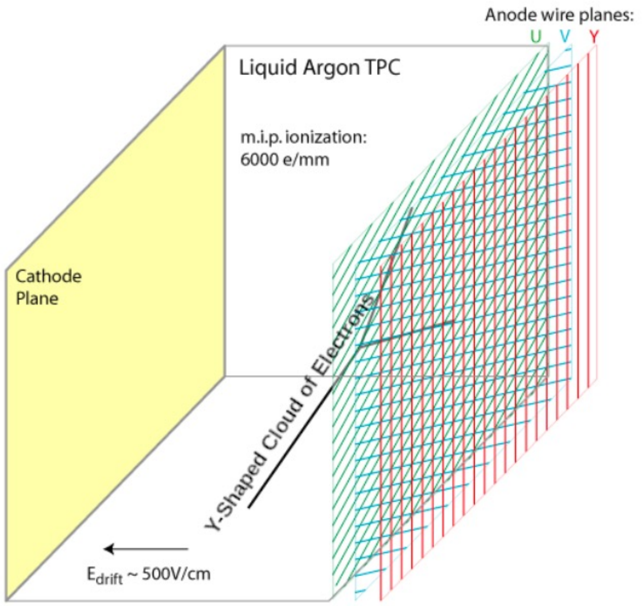
A far detector based on **liquid argon time projection chamber (LArTPC)** technology will reside in Homestake mine, South Dakota.

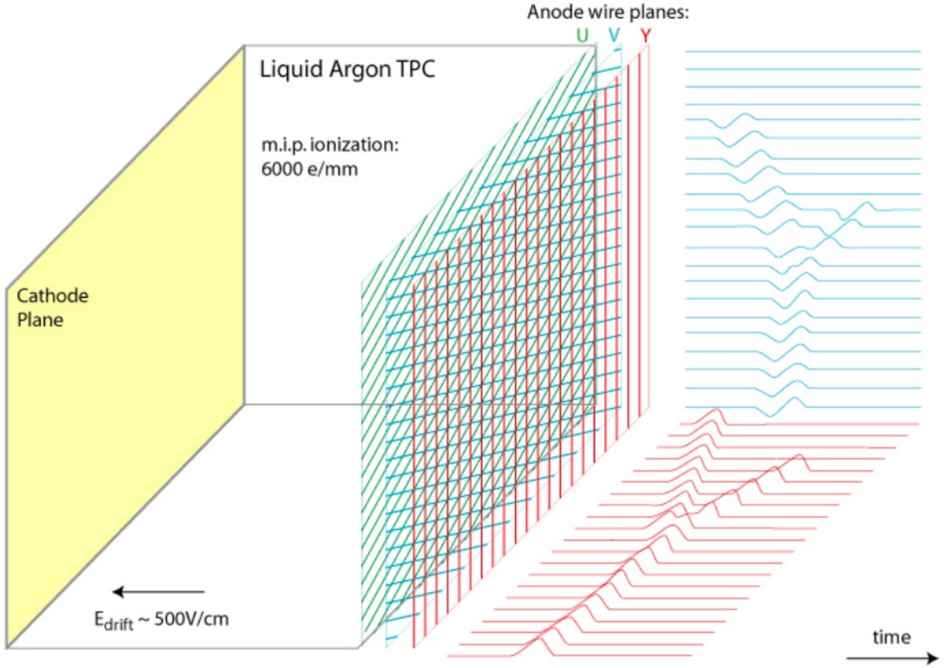


Brookhaven National Laboratory

DUNE and LArTPC

- Charged particles leave clouds of electrons in the detector.
- The clouds of electrons slowly drift towards the wire-planes under the influence of the strong electric field.
- Metallic wires record electric excitations caused by the bypassing electrons.



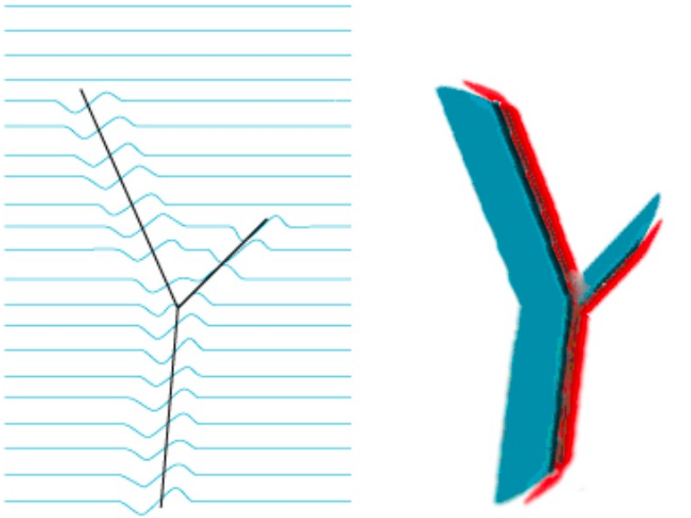


Brookhaven National Laboratory

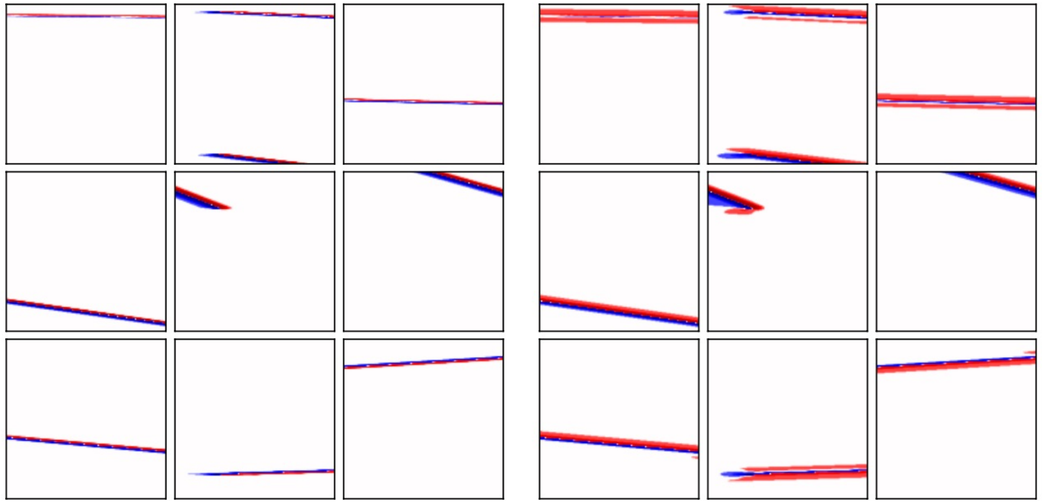
DUNE and LArTPC

Before doing this on real data, we would like to study a task under well-understood settings:

1. **Domain A** – simplified detector response, where a cloud of electrons is read only by the nearest wire.
2. **Domain B** – realistic detector response, where a cloud of electrons can produce excitations in multiple wires.



(a) Schematic Waveforms (b) Actual Data Sample



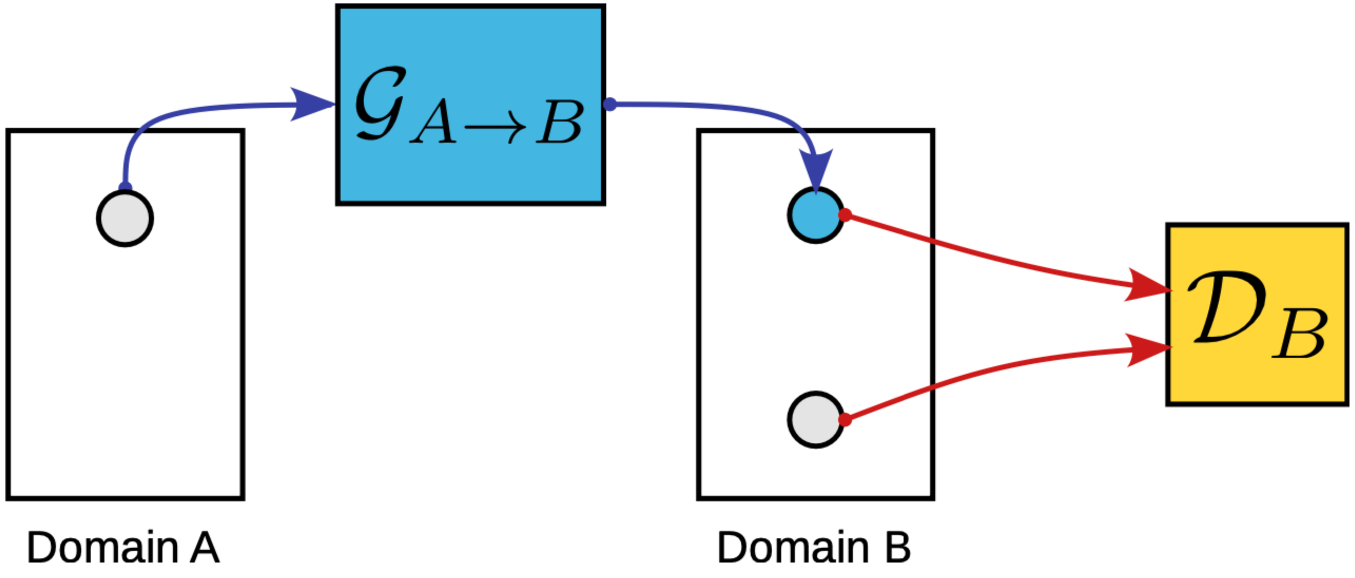
(a) **Domain A** (b) **Domain B**

Brookhaven National Laboratory

Unpaired Image-to-image translation

Since the ground truth of the experimental data is unknown, it's impossible to generate matched simulation images.

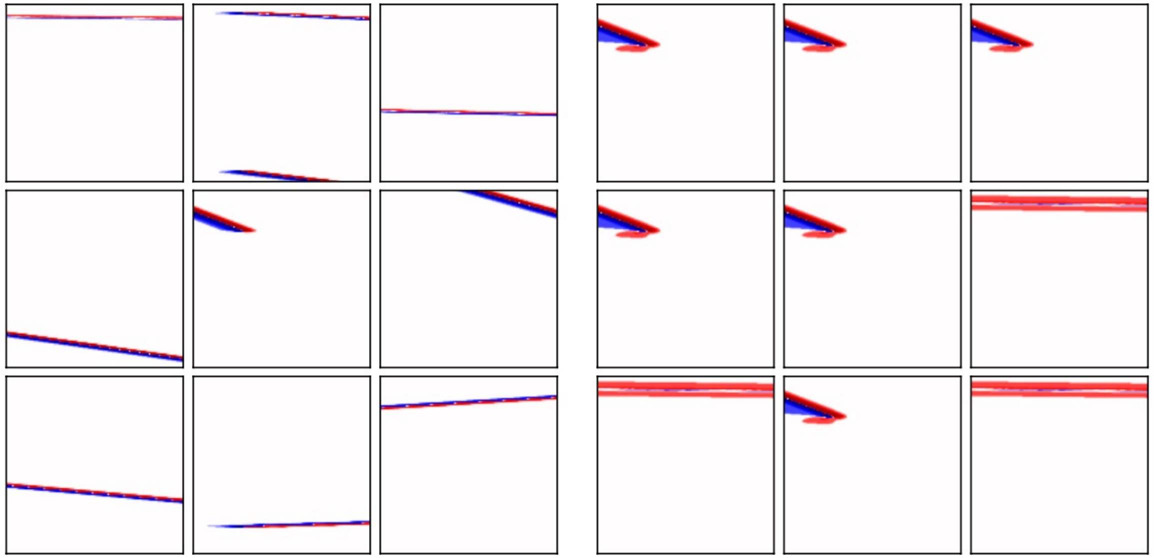
A popular way for generative tasks is GAN (generative adversarial networks).



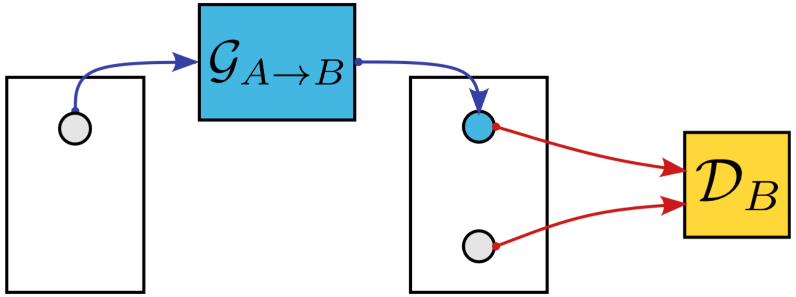
<u>Unpaired</u> Image-to-image translation

However, GAN is prone to "mode collapse".

The generator figured out some "loophole" modes that can always fool the discriminator. So, it translates all input into one of these modes disregarding the input image.



(a) Domain A (b) Translated to Domain B

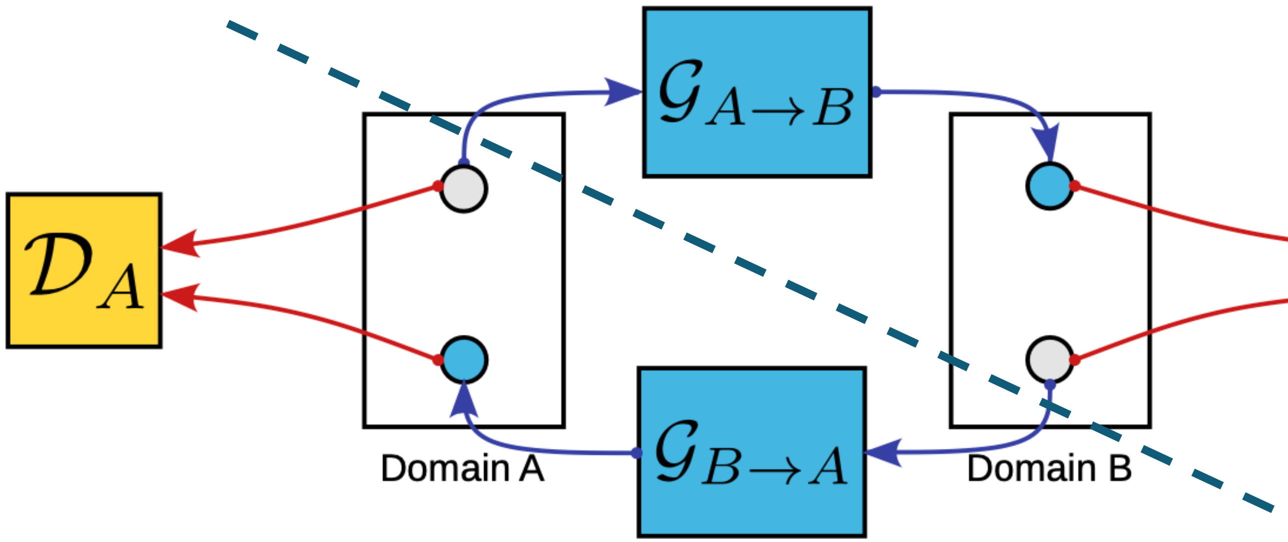


Brookhaven National Laboratory

<u>Unpaired</u> Image-to-image translation

CycleGAN connects two sets of Generator and Discriminator.

* "*Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks*" Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros, Proceedings of the ICCV 2017

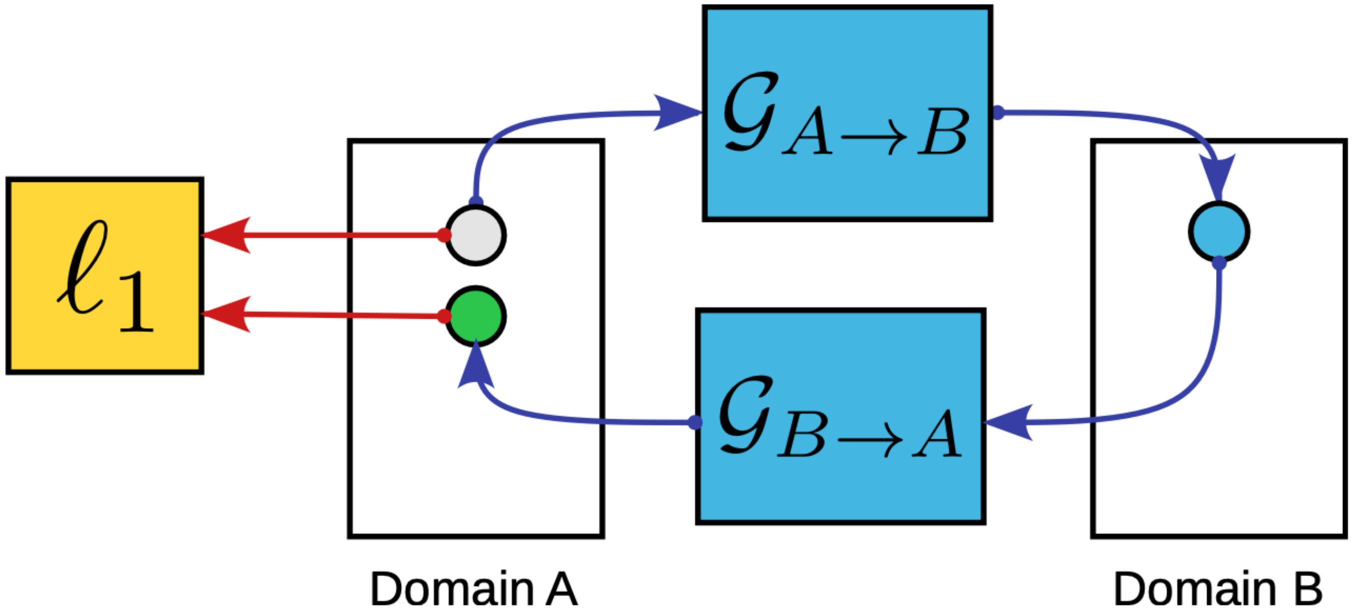


<u>Unpaired</u> Image-to-image translation

CycleGAN connects two sets of Generator and Discriminator.

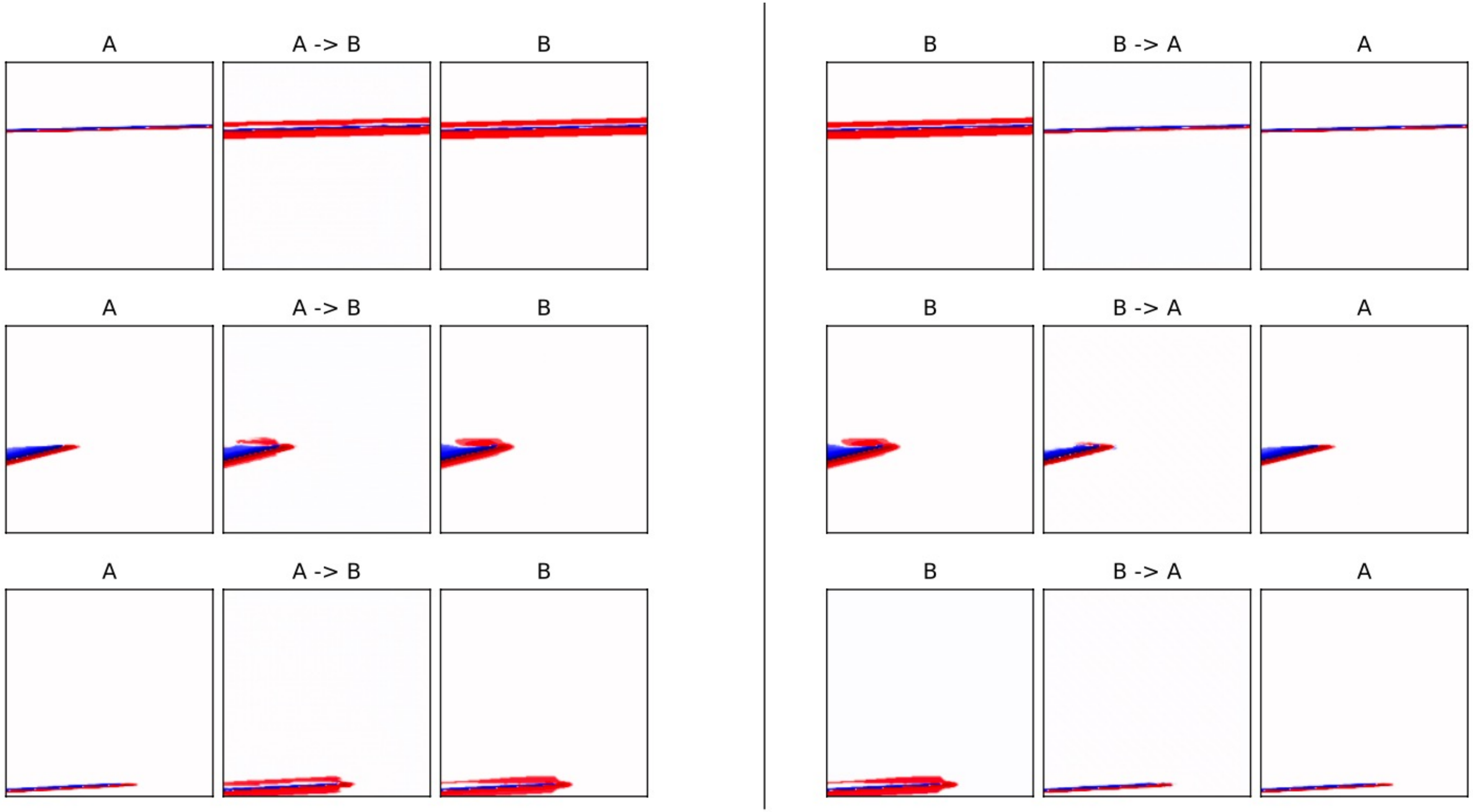
And requires a "Cycle-consistency". $\mathcal{G}_{B\rightarrow A}(\mathcal{G}_{A\rightarrow B}(X_A)) == X_A$

Which solves the mode collapse problem.



Brookhaven National Laboratory

Initial Results of CycleGAN

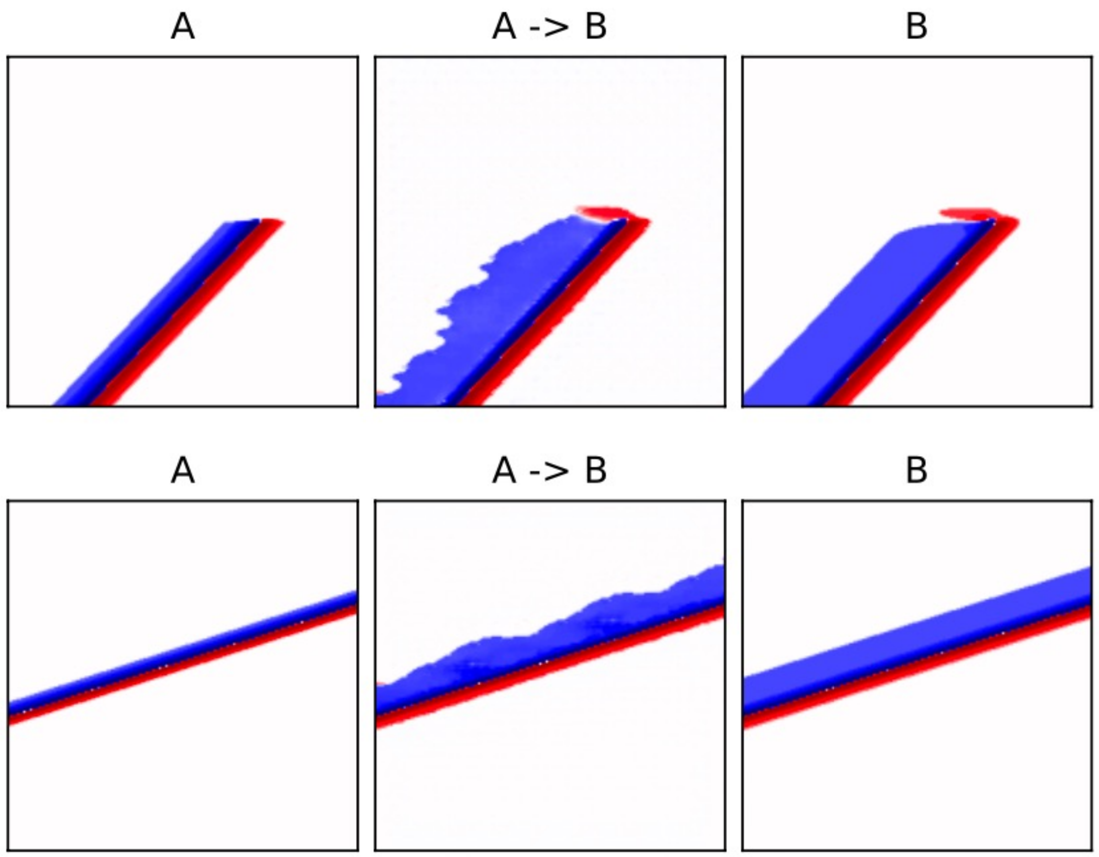


Initial Results of CycleGAN

CycleGAN solves the mode collapse problem.

Overall it is good. But it cannot create a smooth edge.

Keeping smoothness is kind-of “long-range” feature that the CNN-based generator in CycleGAN cannot reconstruct.



Brookhaven National Laboratory

Vision Transforme

We can use the transformer architecture to capture such long-range patterns.

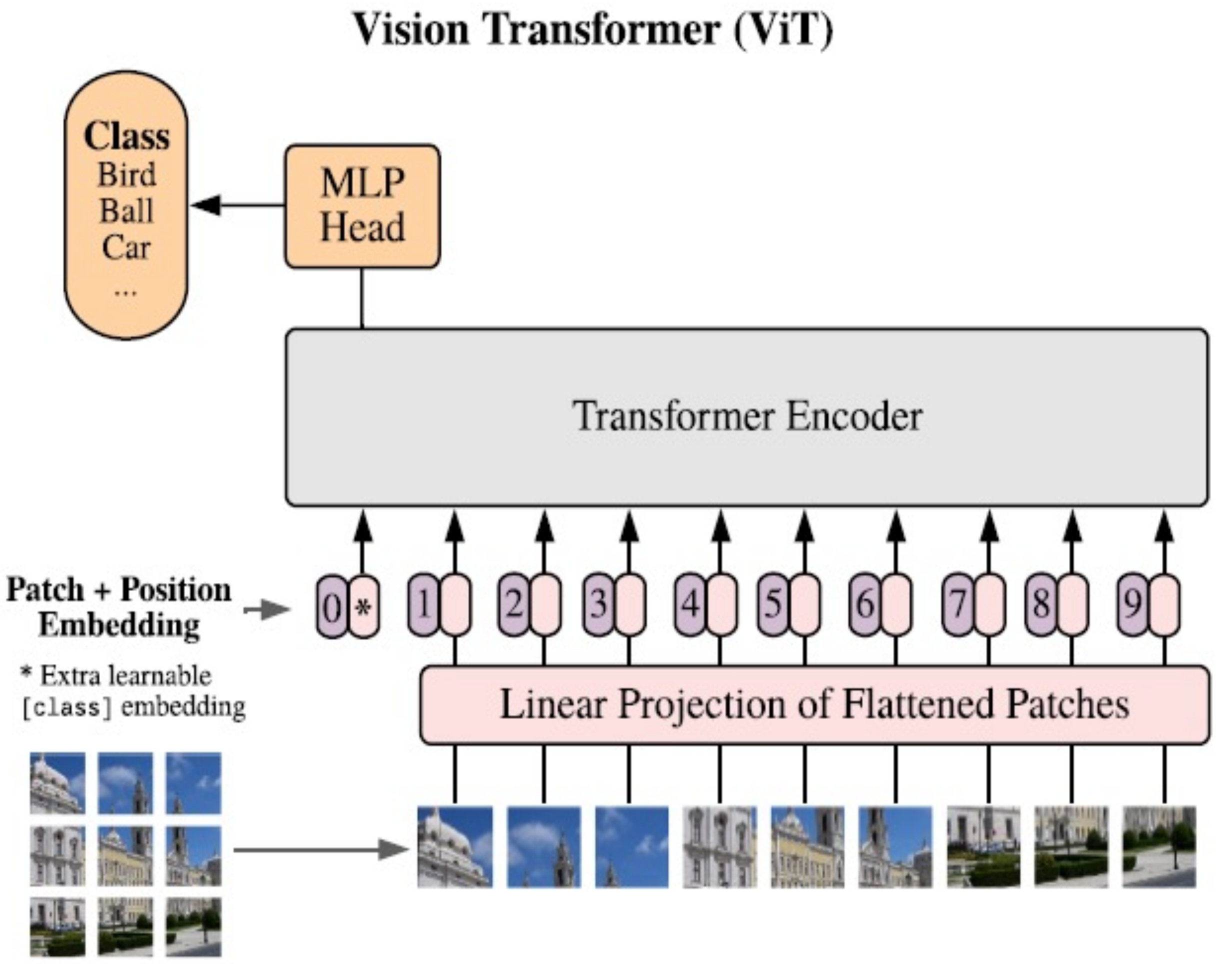
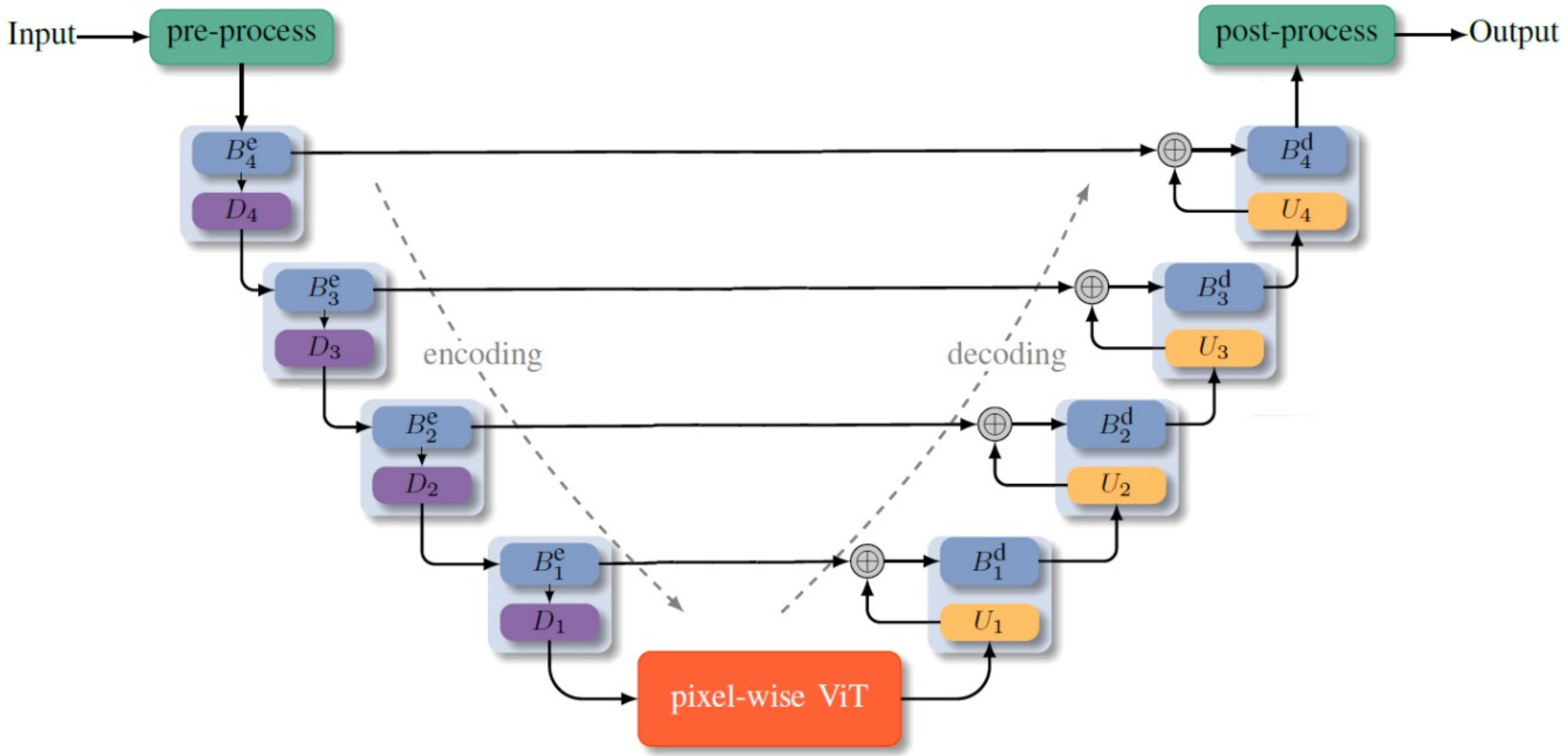


Image source: "[An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale](#)" Alexey Dosovitskiy, Lucas Beyer, et. al., ICLR2021

Unet-ViT-CycleGAN (UVCGAN)

Adding a ViT block at the bottleneck of the Unet improves long-range pattern learning.



UVCGAN fixes rough edges

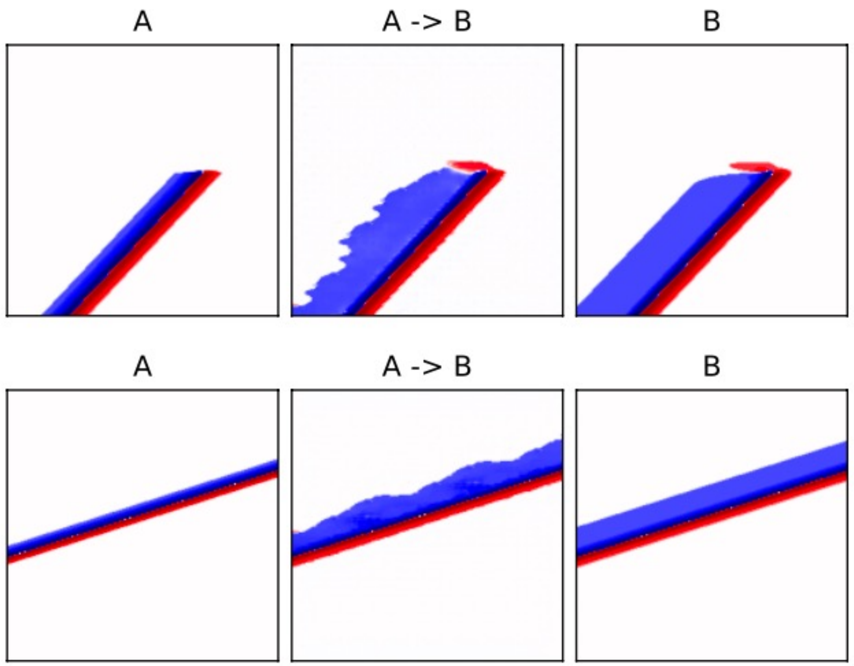


Figure: Default CycleGAN Generator

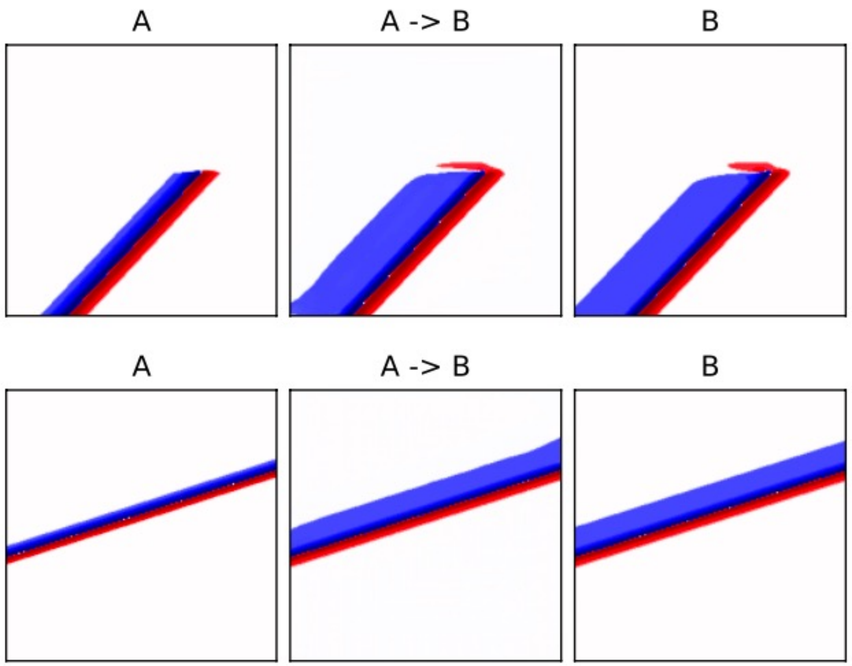


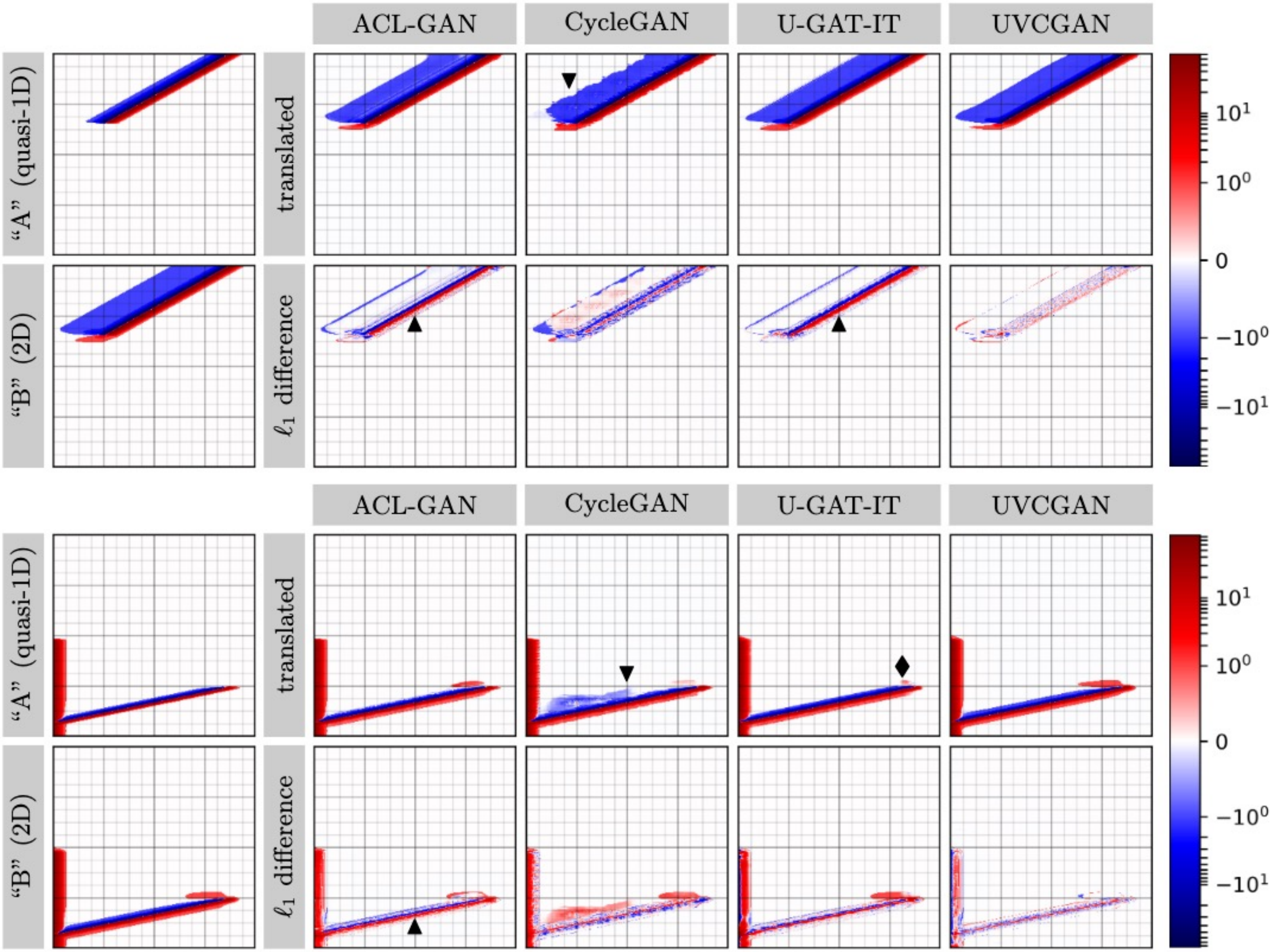
Figure: New UNet-ViT Generator

Results

We have compared our model (UVCGAN) vs advanced models:

1. ACL-GAN arXiv:2003.04858
2. Council-GAN arXiv:1911.10538
3. U-GAT-IT arXiv:1907.10830

algorithm	"A" to "B" ℓ_1	"A" to "B" ℓ_2	"B" to "A" ℓ_1	"B" to "A" ℓ_2
ACL-GAN	0.083	0.566	0.039	0.121
CycleGAN	0.074	0.180	0.061	0.159
U-GAT-IT	0.078	1.187	0.073	1.161
UVCGAN	**0.030**	**0.033**	**0.025**	**0.027**

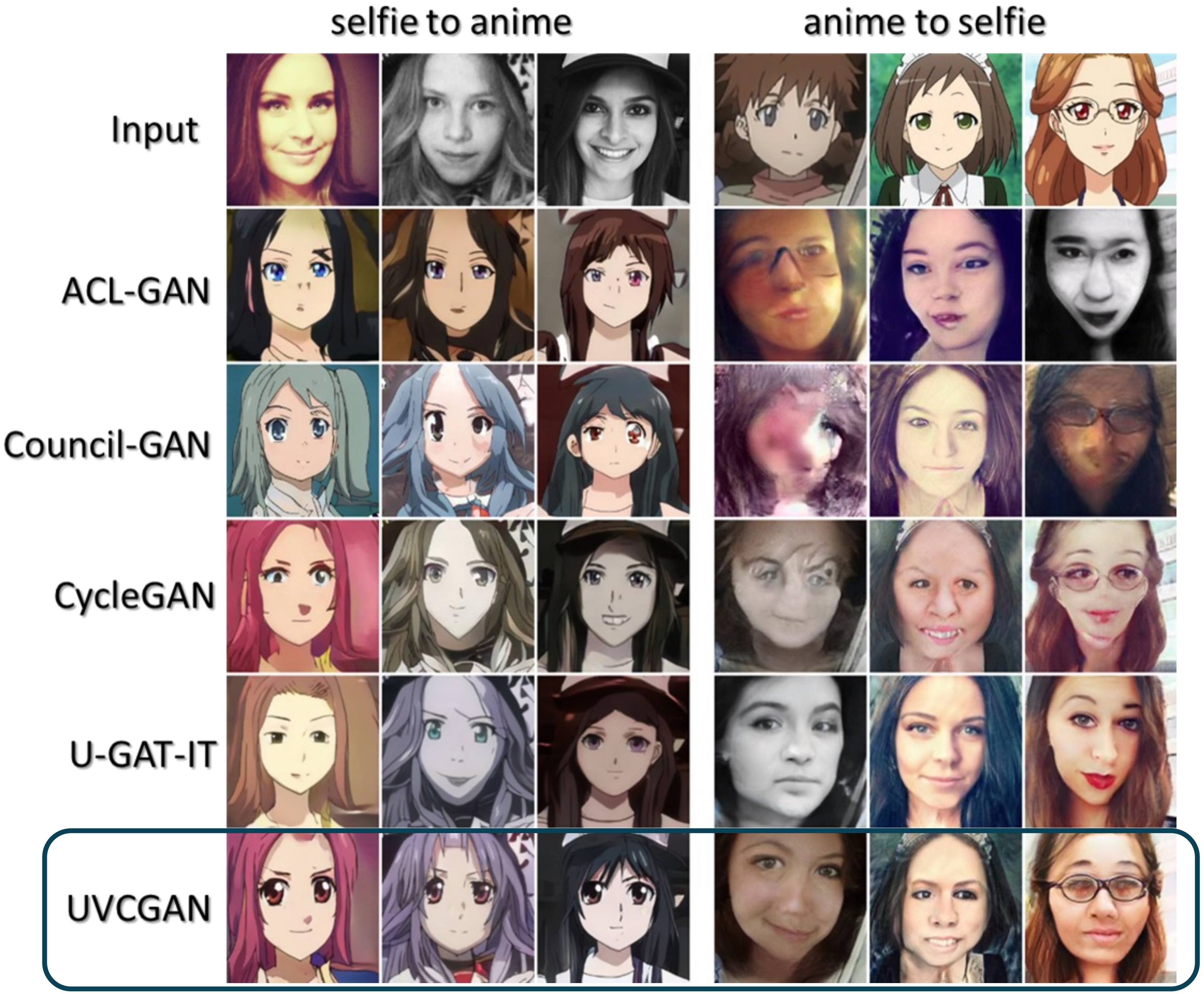


Paper submitted, should be available on arxiv soon.
(available upon email request.)
Data released on https://zenodo.org/record/7809108#.ZDV0B-zMKvB

Testing UVCGAN on open data sets

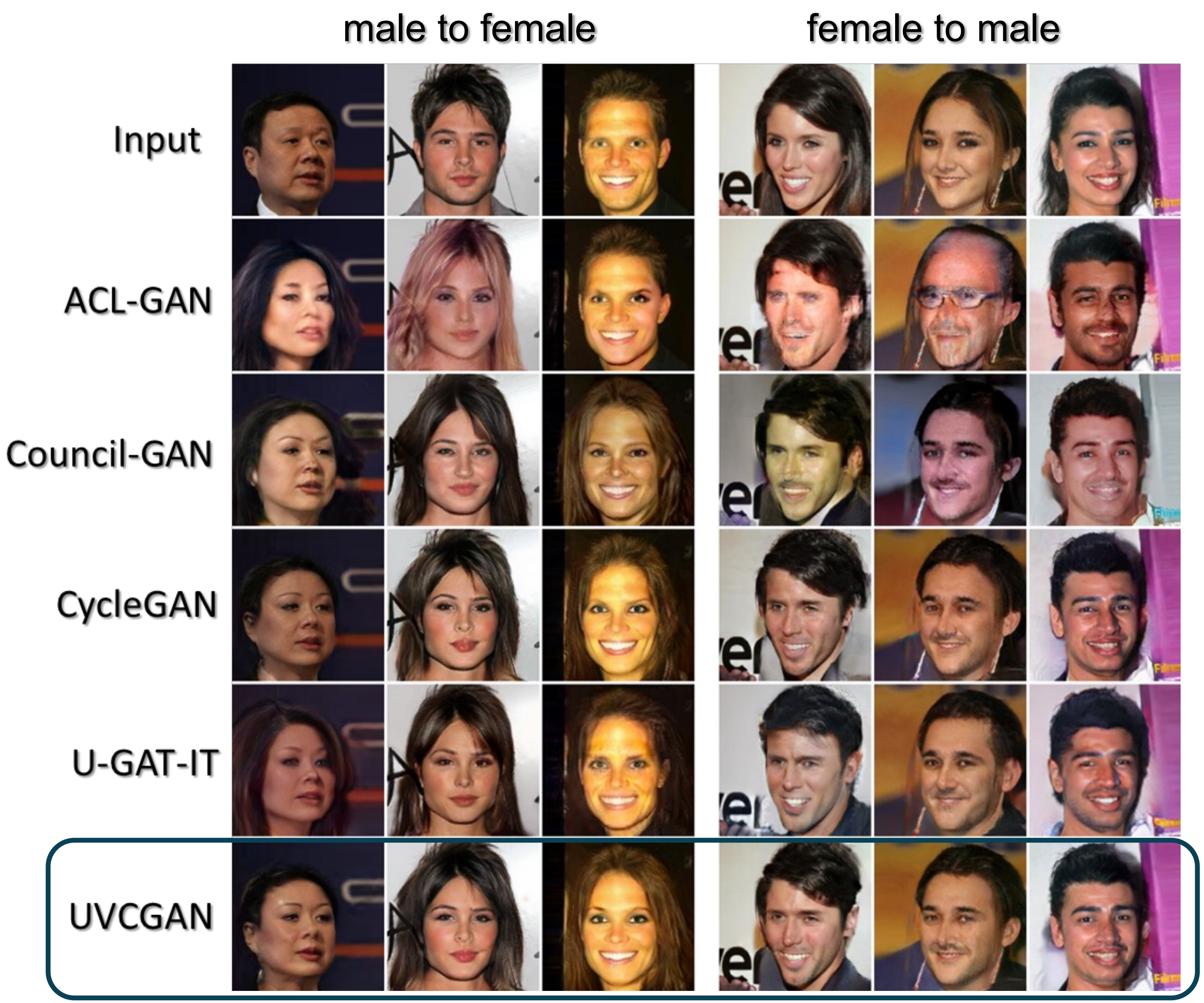
Selfie⇔Anime

- Domain A: Selfie, Domain B: Anime
- 3.4k training images, 100 test images
- Download here
- More information of the dataset can be found here.



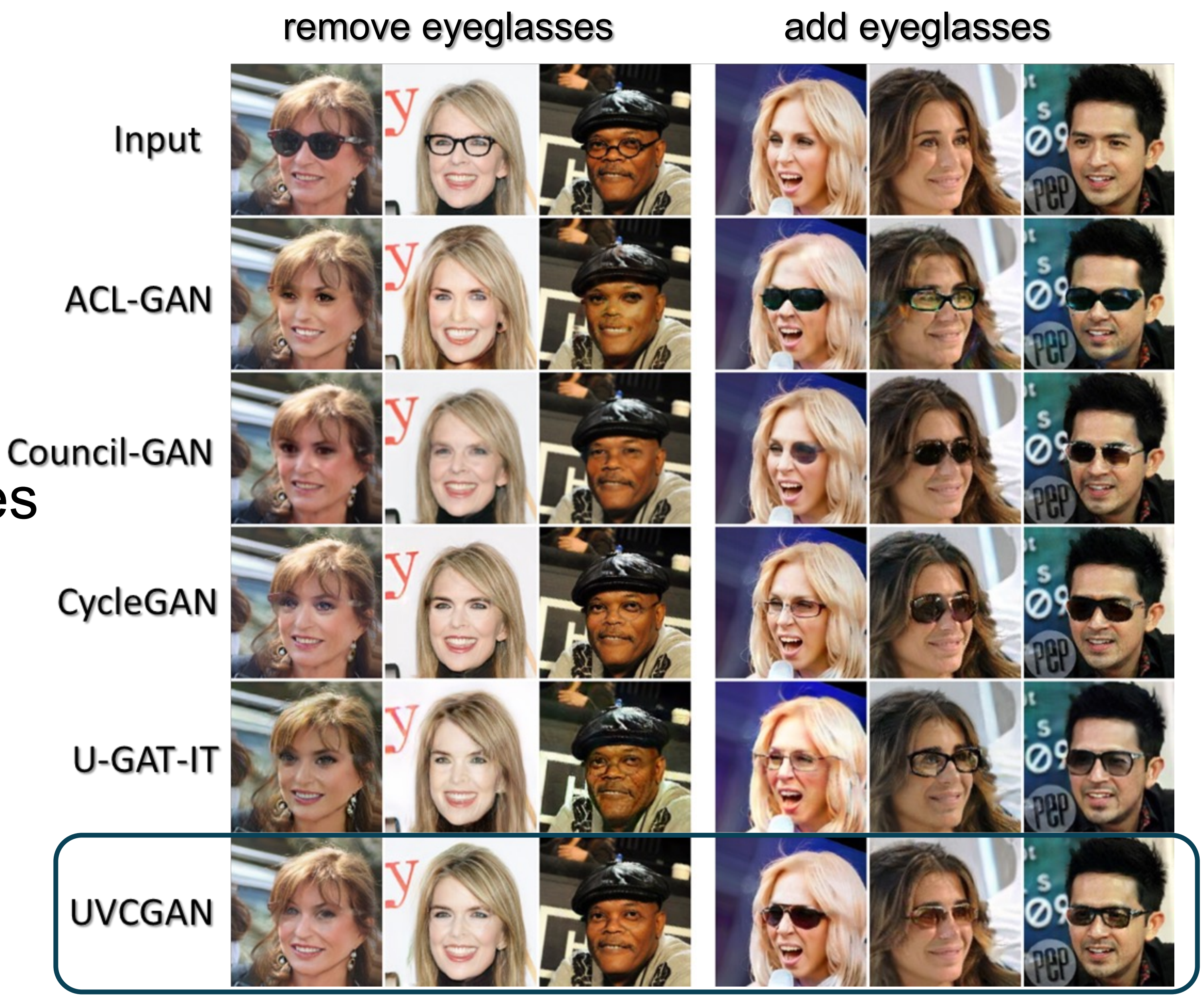
Male⇔Female

- Derived from the CelebA dataset
- Domain A: male, Domain B: female
- Train: male 68k, female 95k
- Test: male 16k, female 24k
- Download [here]



Remove and Add Eyeglasses

- Derived from the CelebA dataset
- Domain A: with glasses, Domain B: without glasses
- Train: with 10k, without 151k
- Test: with 2.6k, without 37k
- Download [here]



Quantitative Results

- Frechet Inception Distance (FID)
- Kernel Inception Distance (KID)

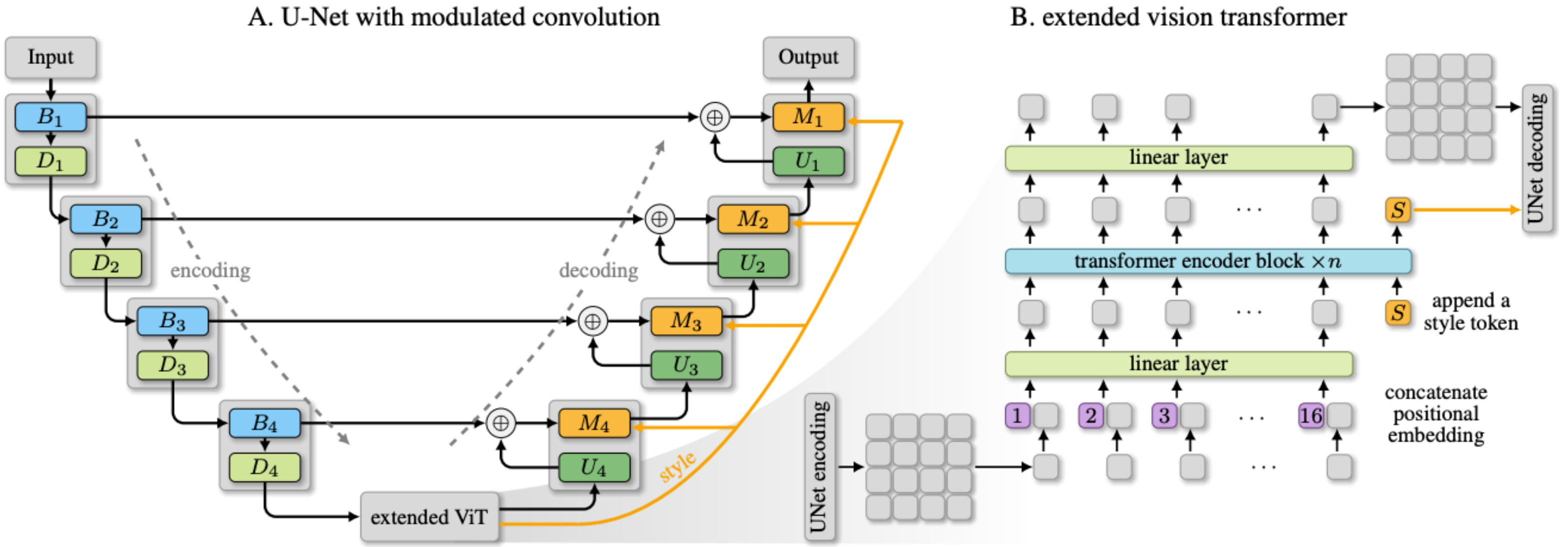
- Code can be found here https://github.com/LS4GAN/uvcgan
- Paper published in WACV2023, draft can be found here https://arxiv.org/abs/2203.02557

D. Torbunov et al., "UVCGAN: UNet Vision Transformer cycle-consistent GAN for unpaired image-to-image translation," 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2023, pp. 702-712, doi: 10.1109/WACV56688.2023.00077.

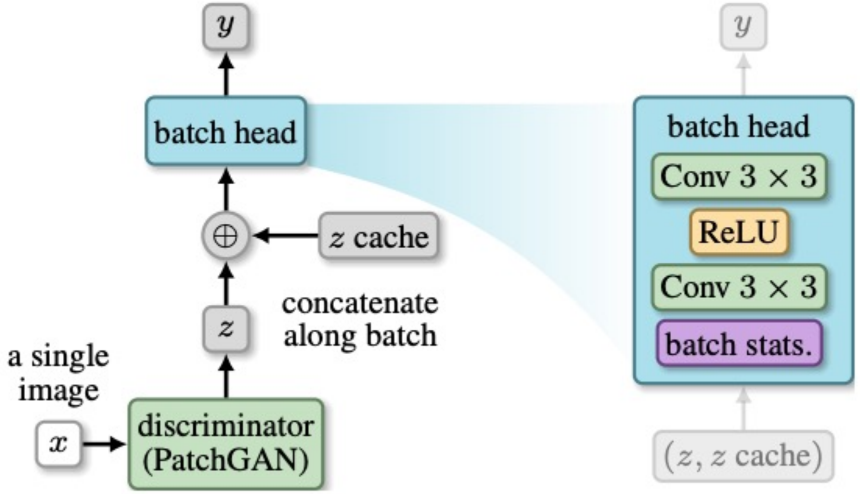
Table 2. **FID and KID scores.** Lower is better.

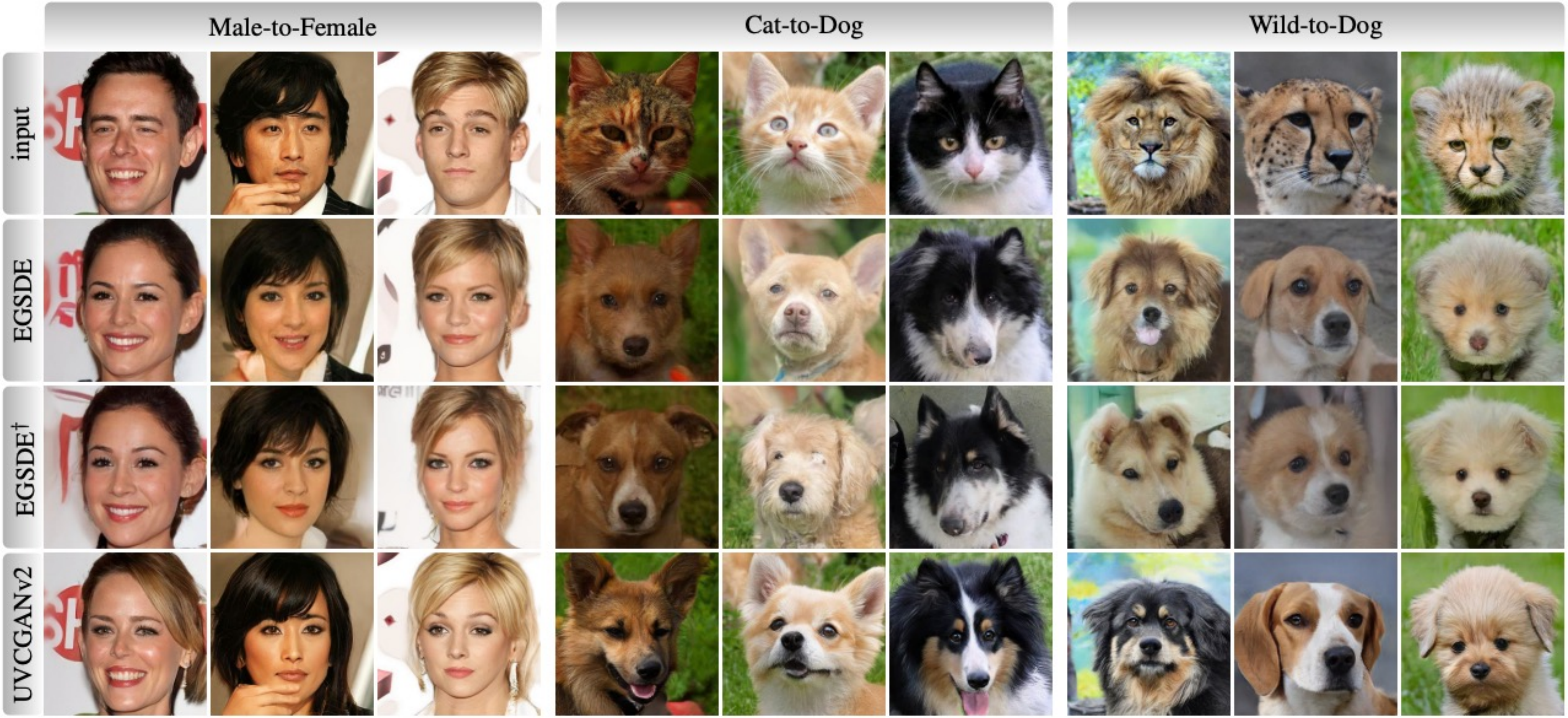
	Selfie to Anime		Anime to Selfie	
	FID	KID ($\times$100)	FID	KID ($\times$100)
ACL-GAN	99.3	3.22 ± 0.26	128.6	3.49 ± 0.33
Council-GAN	91.9	2.74 ± 0.26	126.0	2.57 ± 0.32
CycleGAN	92.1	2.72 ± 0.29	127.5	2.52 ± 0.34
U-GAT-IT	95.8	2.74 ± 0.31	**108.8**	$\mathbf{1.48 \pm 0.34}$
UVCGAN	**79.0**	$\mathbf{1.35 \pm 0.20}$	122.8	2.33 ± 0.38
	Male to Female		Female to Male	
	FID	KID ($\times$100)	FID	KID ($\times$100)
ACL-GAN	**9.4**	$\mathbf{0.58 \pm 0.06}$	19.1	1.38 ± 0.09
Council-GAN	10.4	0.74 ± 0.08	24.1	1.79 ± 0.10
CycleGAN	15.2	1.29 ± 0.11	22.2	1.74 ± 0.11
U-GAT-IT	24.1	2.20 ± 0.12	15.5	0.94 ± 0.07
UVCGAN	9.6	0.68 ± 0.07	**13.9**	$\mathbf{0.91 \pm 0.08}$
	Remove Glasses		Add Glasses	
	FID	KID ($\times$100)	FID	KID ($\times$100)
ACL-GAN	16.7	0.70 ± 0.06	20.1	1.35 ± 0.14
Council-GAN	37.2	3.67 ± 0.22	19.5	1.33 ± 0.13
CycleGAN	24.2	1.87 ± 0.17	19.8	1.36 ± 0.12
U-GAT-IT	23.3	1.69 ± 0.14	19.0	1.08 ± 0.10
UVCGAN	**14.4**	$\mathbf{0.68 \pm 0.10}$	**13.6**	$\mathbf{0.60 \pm 0.08}$

UVCGAN-v2



- Improved UVCGAN with "style" modulation and batch head.
- Compete against the SOTA diffusion-based model, EGSDE. arXiv:2207.06635
- Validated on high quality data sets: CelebA-HQ and AFHQ.




Male-to-Female
Cat-to-Dog
Wild-to-Dog
input
EGSDE
EGSDE†
UVCGANv2

UVCGAN-v2

- Paper submitted and pre-print is available https://arxiv.org/abs/2303.16280
- Source code will be released soon. https://github.com/LS4GAN/uvcgan2 (currently under internal testing.)

Table 2. **FID, PSNR, and SSIM scores.**

	Male to Female		
	FID↓	PSNR↑	SSIM↑
CUT	46.61	19.87	**0.74**
ILVR	46.12	18.59	0.510
SDEdit	49.43	20.03	0.572
EGSDE	41.93	20.35	0.574
EGSDE†	30.61	18.32	0.510
UVCGANv2	17.65	19.44	0.681
UVCGANv2-C	**17.34**	**21.18**	**0.738**
	Cat to Dog		
	FID↓	PSNR↑	SSIM↑
CUT	76.21	17.48	0.601
ILVR	74.37	17.77	0.363
SDEdit	74.17	19.19	0.423
EGSDE	65.82	**19.31**	0.415
EGSDE†	51.04	17.17	0.361
UVCGANv2	**44.76**	15.55	0.562
UVCGANv2-C	52.48	18.30	**0.638**
	Wild to Dog		
	FID↓	PSNR↑	SSIM↑
CUT	92.94	17.2	0.592
ILVR	75.33	16.85	0.287
SDEdit	68.51	17.98	0.343
EGSDE	59.75	18.14	0.343
EGSDE†	50.43	16.40	0.300
UVCGANv2	**45.56**	15.59	0.551
UVCGANv2-C	55.61	**18.65**	**0.631**

Our Team

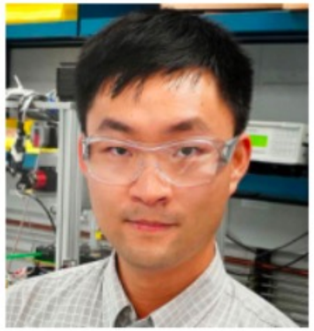
Dmitrii Turbunov

Yi Huang

Acknowledgement:
The LDRD Program at Brookhaven National Laboratory, sponsored by DOE’s Office of Science under Contract DE-SC0012704, supported this work.

(yren@bnl.gov)

Haiwang Yu

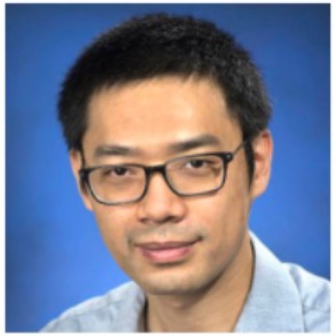


Jin Huang

Shinjae Yoo

Meifeng Lin

Brett Viren



Yihui Ren

We are hiring!

https://jobs.bnl.gov/job/upton/postdoctoral-researcher-machine-learning-in-nuclear-physics/3437/46267319136

Apply Now

Postdoctoral Researcher - Machine Learning in Nuclear Physics

Job ID: 3679

Date posted: 03/20/2023

EMPTY

Comparing with other models

We have compared our model (UVCGAN) vs advanced models:

1. ACL-GAN arXiv:2003.04858
2. Council-GAN arXiv:1911.10538
3. U-GAT-IT arXiv:1907.10830

Algorithm	Time (hrs)	Jointly Trained	# Para.
ACL-GAN	~ 86		~ 55M
Council-GAN	~ 600		~ 116M
CycleGAN	~ 40	✓	~ 28M
U-GAT-IT	~ 140	✓	~ 671M
UVCGAN	~ 60	✓	~ 68M