

Unpaired image translation to mitigate domain shift in liquid argon time projection chamber detector responses

Y. Huang

To be published in "Machine Learning: Science and Technology"

December 2024

Computational Science Initiative
Brookhaven National Laboratory

U.S. Department of Energy
Laboratory-Directed Research and Development (LDRD)

Notice: This manuscript has been authored by employees of Brookhaven Science Associates, LLC under Contract No. DE-SC0012704 with the U.S. Department of Energy. The publisher by accepting the manuscript for publication acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes.

This work was supported by the Laboratory Directed Research and Development Program of Brookhaven National Laboratory under U.S. Department of Energy Contract No. DE-SC0012704.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third party's use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PAPER • OPEN ACCESS

Unpaired image translation to mitigate domain shift in liquid argon time projection chamber detector responses

To cite this article: Yi Huang *et al* 2024 *Mach. Learn.: Sci. Technol.* **5** 045021

View the [article online](#) for updates and enhancements.

You may also like

- [Neutrino–nucleus cross sections for oscillation experiments](#)
Teppei Katori and Marco Martini
- [Exploiting data diversity in multi-domain federated learning](#)
Hussain Ahmad Madni, Rao Muhammad Umer and Gian Luca Foresti
- [New opportunities at the next-generation neutrino experiments I: BSM neutrino physics and dark matter](#)
C A Argüelles, A J Aurisano, B Batell *et al.*



PAPER

OPEN ACCESS

RECEIVED
10 April 2023REVISED
19 August 2024ACCEPTED FOR PUBLICATION
8 October 2024PUBLISHED
23 October 2024

Original Content from
this work may be used
under the terms of the
[Creative Commons
Attribution 4.0 licence](#).

Any further distribution
of this work must
maintain attribution to
the author(s) and the title
of the work, journal
citation and DOI.



Unpaired image translation to mitigate domain shift in liquid argon time projection chamber detector responses

Yi Huang , Dmitrii Torbunov , Brett Viren, Haiwang Yu , Jin Huang , Meifeng Lin
and Yihui Ren*

Brookhaven National Laboratory, Upton, NY, United States of America

* Author to whom any correspondence should be addressed.

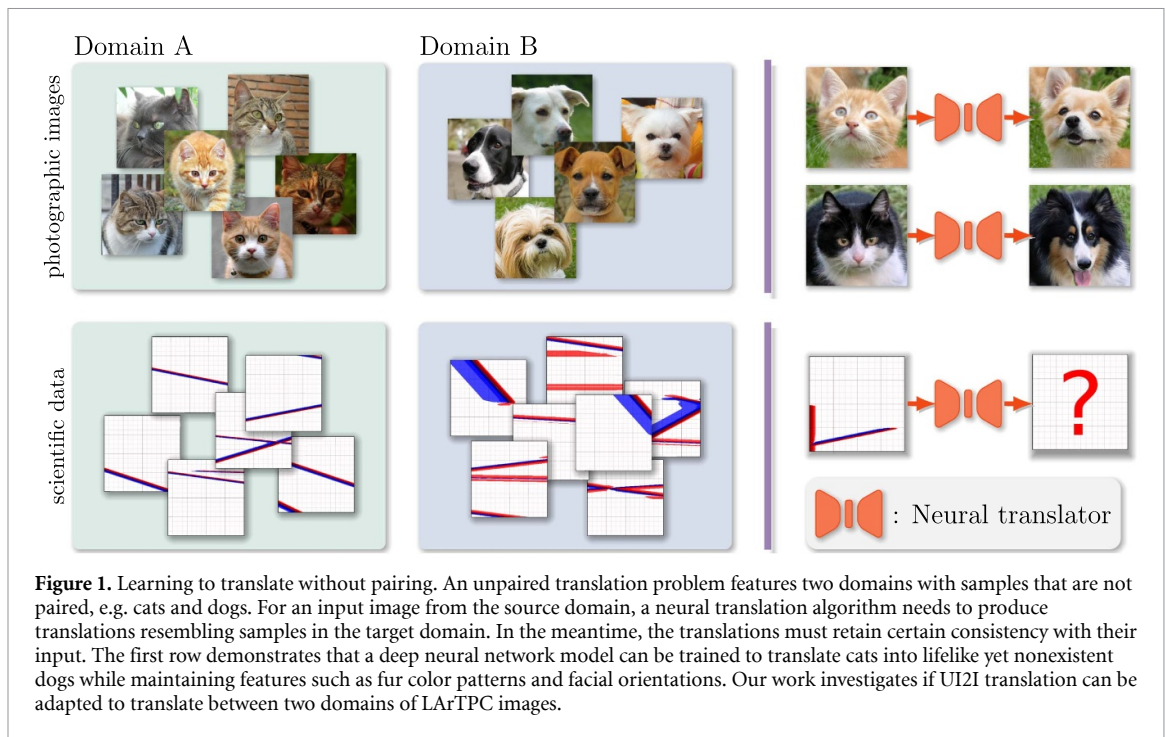
E-mail: yren@bnl.gov**Keywords:** unpaired image translation, neutrino experiment, liquid argon time projection chamber, detector response, deep neural network, domain adaptationSupplementary material for this article is available [online](#)

Abstract

Deep learning algorithms often are developed and trained on a training dataset and deployed on test datasets. Any systematic difference between the training and a test dataset may severely degrade the final algorithm performance on the test dataset—what is known as the *domain shift problem*. This issue is prevalent in many scientific domains where algorithms are trained on simulated data but applied to real-world datasets. Typically, the domain shift problem is solved through various domain adaptation (DA) methods. However, these methods are often tailored for a specific downstream task, such as classification or semantic segmentation, and may not easily generalize to different tasks. This work explores the feasibility of using an alternative way to solve the domain shift problem that is not specific to any downstream algorithm. The proposed approach relies on modern Unpaired Image-to-Image (UI2I) translation techniques, designed to find translations between different image domains in a fully unsupervised fashion. In this study, the approach is applied to a domain shift problem commonly encountered in Liquid Argon Time Projection Chamber (LArTPC) detector research when seeking a way to translate samples between two differently distributed LArTPC detector datasets deterministically. This translation allows for mapping real-world data into the simulated data domain where the downstream algorithms can be run with much less domain-shift-related performance degradation. Conversely, using the translation from the simulated data to a real-world domain can increase the realism of the simulated dataset and reduce the magnitude of any systematic uncertainties. To evaluate the quality of the translations, we use both pixel-wise metrics and a downstream task to measure the effectiveness of UI2I methods for mitigating the domain shift problem. We adapted several popular UI2I translation algorithms to work on scientific data and demonstrated the viability of these techniques for solving the domain shift problem with LArTPC detector data. To facilitate further development of DA techniques for scientific datasets, the ‘Simple Liquid-Argon Track Samples’ dataset used in this study is also published.

1. Introduction

Deep Learning (DL) methods are finding widespread and unprecedented applications in multiple areas of science and technology. Constructing supervised DL models requires access to large volumes of properly *labeled*, high-quality real-world data. However, labeling real-world scientific data is often difficult, costly, or otherwise impossible [1–4]. To workaround the issue, many scientific domains resort to using simulation as a means of obtaining large quantities of labeled data. Although this approach solves the lack of labeled data problem, it introduces another challenge. As there often exists systematic differences between the simulated



and real-world data, a DL algorithm trained on a simulated dataset can exhibit degraded performance when it is applied to real-world data. This issue is known as the *domain shift problem* [1, 5–7].

In this work, we consider tackling the typical domain shift problem in Liquid Argon Time Projection Chamber (LArTPC) detector research. LArTPC is a particle tracking and calorimetry detector technology [8–10] that forms the basis for detectors used by experiments such as MicroBooNE [11], ProtoDUNE [12], and the next-generation DUNE [13]. As with other detector technologies, obtaining human labels for real-world detector data is prohibitively costly. Therefore, physicists rely on scientific detector simulations to generate labeled data and develop analysis algorithms. While the manually designed analysis algorithms are continuously tested to ensure they are minimally affected by the domain shift problem, domain shift remains a serious concern for DL algorithms. It has sparked numerous debates, significantly slowed adoption of DL methods to LArTPC detector analysis, and driven the search for alternative solutions.

Typically, the domain shift problem is solved using various domain adaptation (DA) algorithms [1, 7, 14, 15]. DA techniques are created to enable DL algorithms to perform effectively on novel domains distinct from those where they are trained. However, LArTPC data analysis workflows make it difficult to apply traditional DA techniques. The primary obstacle is that state-of-the-art DA methods [16–18] are tightly coupled to a specific downstream task affected by the domain shift problem. LArTPC data analysis chains can employ dozens of different reconstruction algorithms possibly affected by the domain shift. This requires developing and testing dozens more DA methods, i.e. one for each downstream algorithm. Moreover, new LArTPC data analysis algorithms are constantly being developed, which requires designing even more new DA methods. Thus, it is not feasible to directly apply the traditional DA approaches to LArTPC data analysis.

Here, we consider the viability of using Unpaired Image-to-Image (UI2I) translation methods to address the domain shift problem on LArTPC data. UI2I translation methods are developed for finding translations between different domains of images in a fully unsupervised way [19–23]. For instance, the top row of figure 1 illustrates the operation of a UI2I translation algorithm for the cat-to-dog translation. In the training phase, a UI2I translation algorithm receives random images from the two domains: cats and dogs. Notably, the UI2I translation algorithm is not given what exactly the correct translation of a particular cat should look like. Thus, the algorithm is called ‘unpaired.’ Instead, the algorithm attempts to find some common ‘content’ between the two domains and learns to perform a cat-to-dog translation while preserving the ‘content.’ Once the algorithm is trained, it can transform an arbitrary image of a cat into an image of a dog where the original cat and generated dog are related on some fundamental level (share the same ‘content’).

The key question we try to answer is whether the UI2I translation methods are capable of learning the proper ‘content’ and finding the ‘correct’ translation between two domains of LArTPC data: domain *A* representing simulation and domain *B* denoting the real-world data. This question is nontrivial as there is an infinite number of possible translations between the two domains mathematically, while only a small

fraction of them is correct. The UI2I translation literature frequently overlooks the question of whether or not the image content is preserved during the translation or relies on subjective measures of the content [24]. Applying the UI2I translation methods to scientific data requires much stronger guarantees of the translation's correctness. In this work, we investigate the ability of several UI2I translation models to learn the LArTPC translations and compare their performance.

If the UI2I translation methods are capable of finding the correct translations between the data domains, then the $B \rightarrow A$ translation can be used for DA purposes. For instance, if an algorithm ϕ is trained on the A domain and applied to an image b from the B domain, the domain shift problem would manifest. However, if image b is first translated toward the domain A with the help of a UI2I translation algorithm, ϕ can be applied on the translated image, mitigating domain shift effects.

On the other hand, the correct $A \rightarrow B$ translation can be used to enhance the realism of the simulated data. Performing such a translation on a simulated A dataset will produce a more realistic B' dataset, which has several potential applications:

- It can be used to increase the fidelity of the simulation for the subsequent data analysis.
- LArTPC analysis algorithms can be developed on the translated B' data instead of the original simulated A data. This has the potential to make the algorithms less affected by the domain shift effects.
- The B' dataset is produced from A by a UI2I translation algorithm. Therefore, for each b' in B' , we know its source image a in A . Comparing b' to its source a will allow for directly observing systematic differences between the simulation and real-world data on a sample-by-sample basis. Without such an $A \rightarrow B$ correspondence, scientists can only observe systematic differences by comparing averages over the entire dataset.
- Likewise, an $A \rightarrow B$ pairing can be used to estimate the sensitivity of various downstream algorithms ϕ to the systematic differences between the simulation and real-world data by computing $\phi(a) - \phi(b')$.

Unfortunately, it is difficult to evaluate UI2I translation methods on real data. For this to be possible, for each real detector data image, a matching simulation image should be generated with the same physics. Afterwards, the simulated image would be translated to see if the outcome matches the matching real image. However, extracting the physics ground truth, such as particle momentum, from real-world LArTPC data requires meticulous and time-consuming analysis from a large scientific collaboration. Thus, we consider a **surrogate problem**, where both A and B domains are populated by the simulated data with controllable differences between the domains. Using the simulation will allow for making accurate judgments about the translation quality.

For this study, we created the Simple Liquid-Argon Track Samples (SLATS) dataset featuring two domains: A and B . The A domain is populated by a LArTPC detector simulation with a simplified version of the detector response to the particle activity within it. The B domain is populated by a LArTPC detector simulation with a more realistic version of the detector response. Incorrect simulation of the detector response is a known source of systematic errors in the LArTPC detector analysis. Thus, the SLATS dataset illustrates a common source of the domain shift problem encountered in LArTPC detector research. To facilitate the accurate evaluation of the translation's accuracy, the test portion of the SLATS dataset has an explicit pairing between the A and B domains. That is, for each test image in the A domain, we know exactly how its translation should appear in the B domain and vice versa.

We evaluate the correctness of the resulting UI2I translations via three methods. First, we use the explicit pairing of the test part of the SLATS dataset to perform pixel-wise comparisons of the translated images to their ground truth. Second, we rely on a downstream production-grade signal processing algorithm to extract the physical content of the images. This algorithm is especially sensitive to the domain shift problem. Finally, we study whether the UI2I methods can improve the performance of a supervised DL algorithm affected by the domain shift on the SLATS dataset.

The remainder of this paper is organized as follows. In section 2, we briefly describe how a LArTPC detector works and the construction of the SLATS dataset. In section 3, we review a selection of UI2I translation algorithms suitable for LArTPC data. In section 4, we evaluate the quality of translated images. Finally, in the discussion section, we summarize our findings and suggest future directions of research.

Main Contributions

- We show the feasibility of using UI2I translation techniques to perform domain translation on LArTPC data. The four UI2I translation algorithms studied in this work manage to correctly capture the 'content' of the data and preserve it during the domain translation.

- We demonstrate that UI2I translation techniques can be used to reliably mitigate the domain shift effects on LArTPC data and can provide up to 80% reduction in domain shift error of a downstream signal processing algorithm.
- Likewise, we show that UI2I translation methods can be used to improve the realism of the LArTPC simulation, suggesting the viability of using UI2I translation methods as a post-processing step to obtain a more realistic simulation.
- We release a SLATS dataset demonstrating the common source of the LArTPC detector domain shift in a controllable manner. The dataset also displays unique features of scientific datasets not commonly shared by natural images, such as signal sparsity, lack of upper/lower limits on pixel values, and exact knowledge of the correct translations. This dataset is expected to help with the future development of additional scientifically sound domain translation algorithms.
- Finally, we compare the translation performance of four UI2I translation methods (CycleGAN [20], ACLGAN [22], U-GAT-IT [23], UVCGAN [21]) on the SLATS dataset. Our results demonstrate that the UVCGAN algorithm significantly outperforms other methods across a wide range of metrics, and introduces the least amount of artifacts into the translated data. These results suggest that UVCGAN may serve as an effective basis for further development of UI2I translation methods in scientific data processing.

2. The two-domain SLATS dataset

Released with this study, the SLATS dataset has two domains, each populated by a variant of a LArTPC detector simulation used in the ProtoDUNE-SP experiment [25, 26]. The two domains differ in precisely one feature—the *response function*. This section discusses how the SLATS dataset is generated and preprocessed.

2.1. LArTPC overview

LArTPC detectors enclose a volume of liquid argon (figure 2). Energetic charged particles, like those produced from the interaction between a neutrino and an argon nucleus, pass through the volume. As they move, these particles ionize electrons from nearby argon atoms. An external electric field causes these electrons to drift through the liquid argon toward the detector's readout side. The readout of the detector comprises three parallel planes (U, V, and Y) of parallel sense wires, oriented in complementary directions. Each wire plane generates a readout, called an *ADC waveform*, that can be interpreted as one tomographic view of the particle's tracks. A tomographic view is a two-dimensional (2D) image with one dimension in space and the other in time. When the three tomographic views are combined, the three-dimensional (3D) tracks of the energized charged particles can be reconstructed.

In this work, the images used to construct the SLATS dataset are the readout from one wire plane (the U plane). The pixel value of the images is the digitized measure (or ADC value) of the current induced by the ionized electrons. The measure is the result of a convolution between the electron distribution and a detector response [27]. The real detector response is a complex function of the electrostatic fields of all electrodes in the detector's readout. However, in simulation and signal processing, this response is approximated by a simplified model. We call such an approximation a *response function*.

2.2. Two simulated domains

The two SLATS dataset domains, *A* and *B*, are generated by applying two different response functions. More precisely, for a set of simulated simple particle tracks, a low-fidelity quasi-one dimensional (1D) response function is applied to produce a domain *A* waveform and a high-fidelity 2D response function to produce a domain *B* waveform. The 2D model is state of the art in the LArTPC community. The quasi-1D model is an artificially simplified model obtained by masking all contributions from neighboring wires in the 2D response (figure 3). Additional information on response functions and track generation can be found in appendix A.1.

The SLATS dataset's design was motivated by three considerations. First, in the absence of real detector data, the contrast between the 2D model and the quasi-1D model is a reasonable proxy for the systematic difference between the real detector response and a simulated one. Second, using identical simulation conditions except for response functions allows for generating *paired test images*. With the paired test images, we can evaluate a UI2I translation algorithm by directly comparing a translated image with its known target. Lastly, the restricted source of difference in the two domains facilitates understanding of the capability (and/or potential limitations) of unpaired neural translation. The experience gained via such a constructed scenario affords a proper foundation for applying UI2I translation between domains with complex sources of difference.

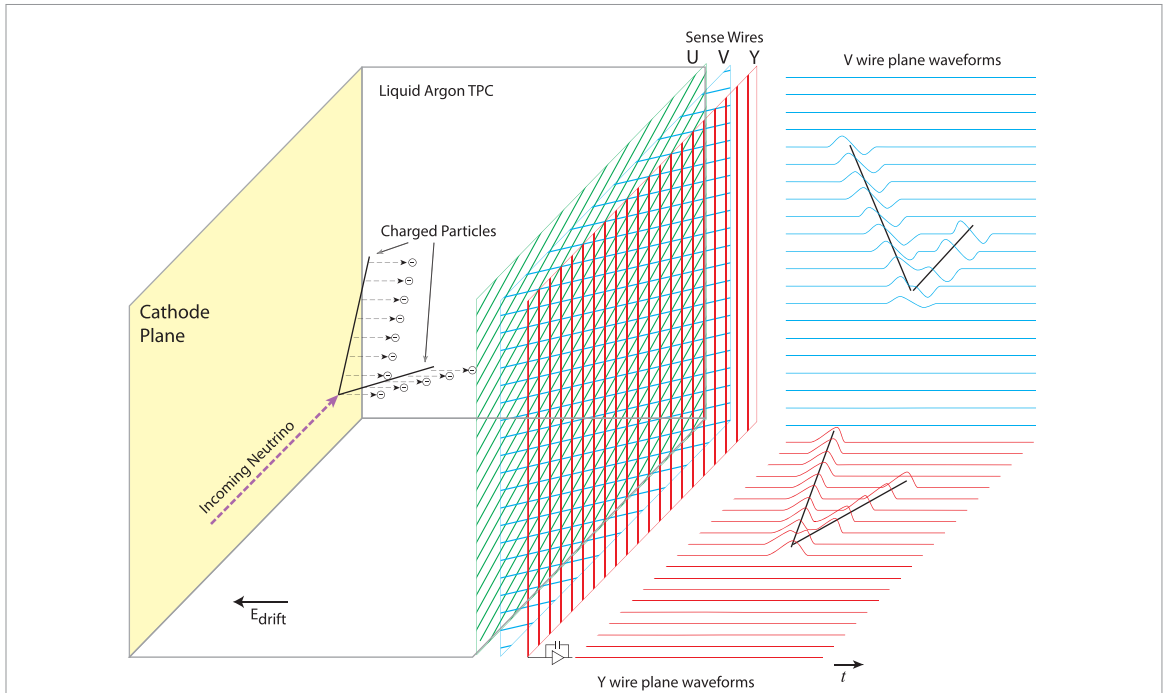


Figure 2. Signal formation in a three-wire plane LArTPC, an illustration from [11]. LArTPC detectors enclose a volume of liquid argon. Energetic charged particles ionize electrons from nearby argon atoms as they pass through the volume. An external electric field causes the electrons to drift toward the detector’s readout. The readout consists of three parallel planes of sense wires. Each wire plane generates one tomographic view of the tracks. The 3D particle tracks can then be reconstructed by combining the three tomographic views.

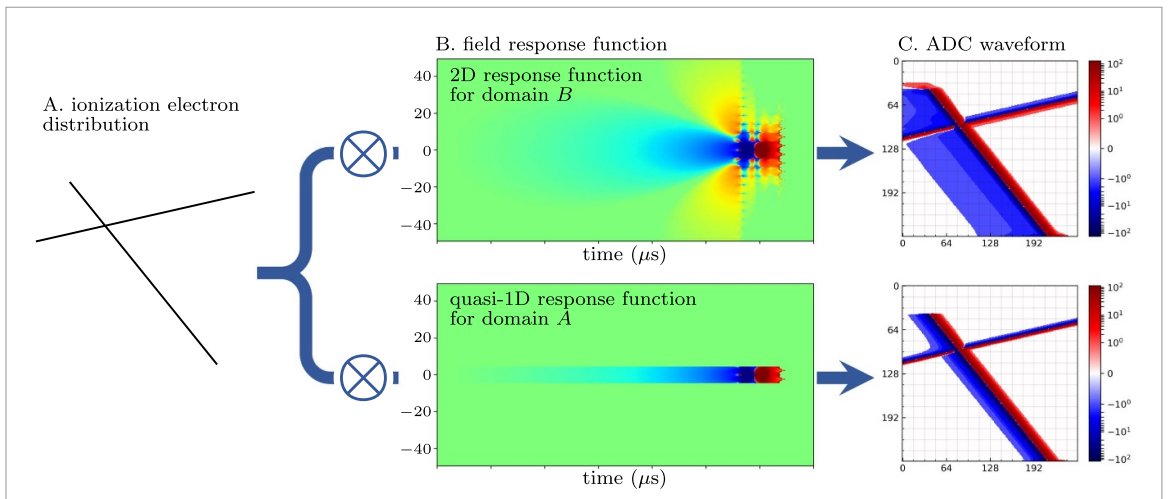


Figure 3. Response functions and ADC waveforms. The ionization electron distribution (Panel A) is convolved with two types of response functions to produce the SLATS dataset’s two domains. The 2D response (Panel B top) is used to produce domain B samples, while the quasi-1D response (Panel B bottom) is used to create domain A samples. The quasi-1D data are constructed by masking the 2D response so all contributions from neighboring wires are removed. Panel C shows examples of the ADC waveforms used as input to the translation algorithm.

2.3. Dataset preparation

This study focuses on one of the three sense wire planes, namely the U plane. Figure 4(A) depicts a full U plane readout of dimension (channel, time) = (800, 6000). Because a majority of existing neural translation algorithms take an input size of (256, 256), we use a center (768, 5888) crop (red box) in the U plane image, and then divide it into 3 × 23 non-overlapping tiles of size (256, 256).

Figures 4(B1) and (B2) show a pair of tiles from domains A and B. Two major differences appear between them. First, the domain B track exhibits long-range induction effects in both the time and space (channel) dimensions, while the A track shows variation only in time. This leads to domain B tracks being less compact than domain A. In particular, larger lobe structures can be observed at the end of domain B tracks, while domain A tracks end more abruptly. This can also lead to features in a domain B tile missing from the

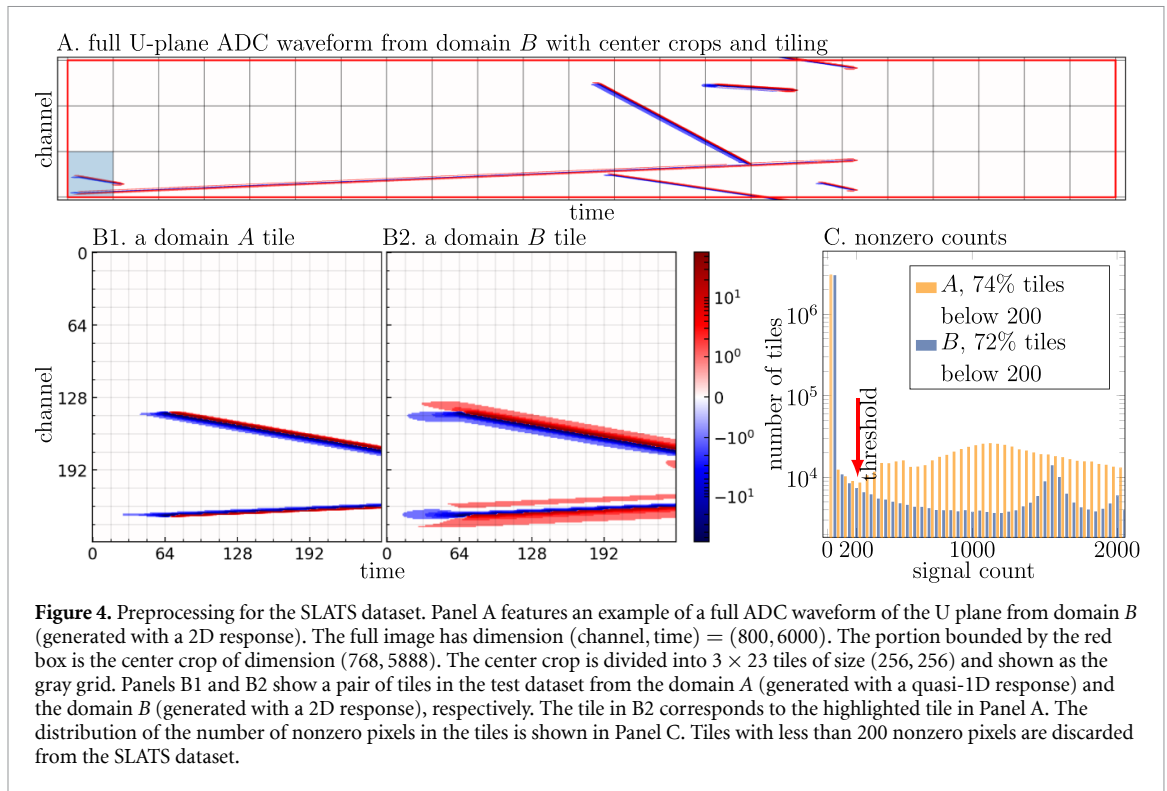


Figure 4. Preprocessing for the SLATS dataset. Panel A features an example of a full ADC waveform of the U plane from domain B (generated with a 2D response). The full image has dimension (channel, time) = (800, 6000). The portion bounded by the red box is the center crop of dimension (768, 5888). The center crop is divided into 3×23 tiles of size (256, 256) and shown as the gray grid. Panels B1 and B2 show a pair of tiles in the test dataset from the domain A (generated with a quasi-1D response) and the domain B (generated with a 2D response), respectively. The tile in B2 corresponds to the highlighted tile in Panel A. The distribution of the number of nonzero pixels in the tiles is shown in Panel C. Tiles with less than 200 nonzero pixels are discarded from the SLATS dataset.

corresponding domain A tile, such as the small red lobe between the two tracks. Second, domain B has a larger neighborhood where the electron distribution can lead to interference patterns as evidenced by the red lobe above the bottom track.

Because of the sparseness of events in the generation of SLATS, a majority of (256, 256) tiles are fully or nearly empty. According to the distribution of the number of nonzero pixels in the tiles (figure 4(C)), we choose a threshold of 200 pixels (around the first local minimum for domain A) and reject tiles below the threshold. To keep the tiles paired for testing, we retain a pair in the test dataset if only both the A and B tiles pass the threshold. More details about the preprocessing of LArTPC simulation data for neural translator training can be found in appendix A.2.

The SLATS dataset can be downloaded from <https://zenodo.org/record/7809108>. The dataset contains both center crops and tiles. The dataset's test part is paired for pairwise translation quality evaluation.

3. Deep generative models for unpaired image translation

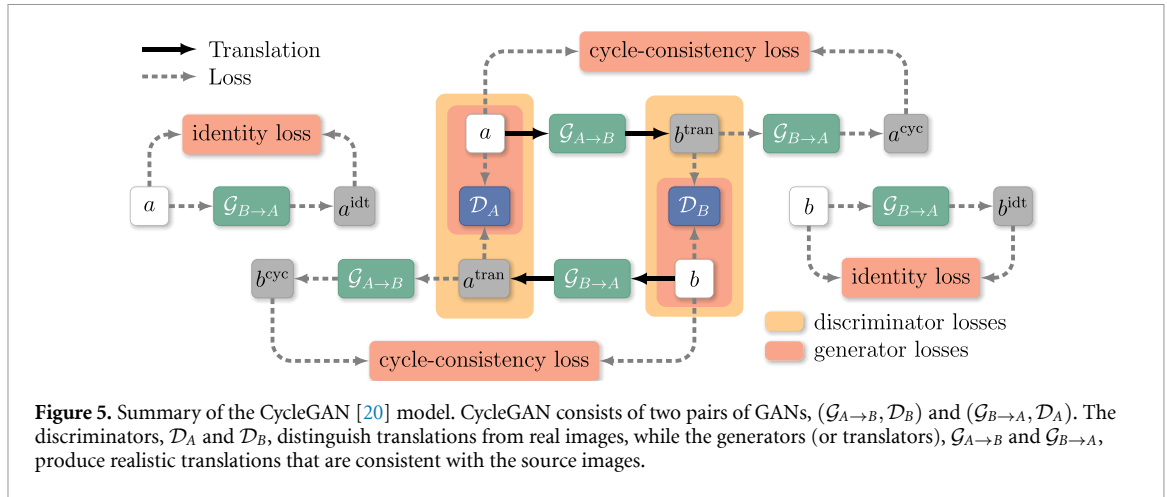
As previously outlined, our goal is to apply modern UI2I translation methods to mitigate the domain shift problem. This section addresses the challenges in designing UI2I translation algorithms and describes a family of UI2I translation methods suitable for this task. The algorithms discussed herein are based on a Generative Adversarial Network (GAN) architecture (section 3.1) and rely on a cycle-consistency constraint (section 3.2) to ensure preservation of the important features during the translation.

3.1. GAN for UI2I

The first successful models for UI2I translation were built on top of the GAN architecture [28]. GAN models are able to learn the particular data distribution and synthesize new samples indistinguishable from the real data. Their main component is a generator network \mathcal{G} that produces realistic-looking data from random noise. To train the generator network \mathcal{G} , GANs employ another neural network known as a discriminator \mathcal{D} . In each GAN training iteration, the discriminator \mathcal{D} network learns to differentiate samples produced by the generator from the real-world data. Then, using the discriminator as a guide, the generator \mathcal{G} network is trained to produce samples that are indistinguishable from the real-world data. In other words, the generator and discriminator engage in a game, throughout which the generator progressively improves the quality of the generated data.

3.2. Cycle-consistent GAN

As noted, GANs can be used to learn data distributions and produce realistic-looking samples from random noise. This means, in principle, a GAN can be trained to generate real detector data from simulated data.



However, when a GAN generates a real detector data sample from a simulated one, it is not guaranteed to preserve any information from the input. The GAN can completely discard the simulated sample and produce a random and unrelated output that looks like real detector data. This creates a challenge for our LArTPC detector example as we need to generate not only realistic-looking *real* data samples, but also ensure the generated samples preserve information from the simulated ones. The same discussion also applies to the translation in the opposite direction.

One approach to address this is provided by CycleGAN [20], which employs *two* GANs that work in the opposite directions as illustrated in figure 5. By using a pair of GANs, a CycleGAN-like model creates translation loops, so information loss during translation can be properly measured.

Specifically, we denote the two domains by A and B and the corresponding GANs as $(\mathcal{G}_{A \rightarrow B}, \mathcal{D}_B)$ and $(\mathcal{G}_{B \rightarrow A}, \mathcal{D}_A)$, respectively. Consider a source image $a \in A$. A CycleGAN-like model can translate this sample to look like those from domain B by using its generator $\mathcal{G}_{A \rightarrow B}$. To ensure the generator $\mathcal{G}_{A \rightarrow B}$ preserves information about the source image a , CycleGAN imposes a cycle-consistency constraint, requiring that a cyclically translated image $a^{\text{cyc}} \equiv \mathcal{G}_{B \rightarrow A}(\mathcal{G}_{A \rightarrow B}(a))$ matches the original image a . In practice, this constraint is enforced by the *cycle-consistency loss* $\|a - a^{\text{cyc}}\|$ that encourages the generators $\mathcal{G}_{A \rightarrow B}$ and $\mathcal{G}_{B \rightarrow A}$ to preserve the information. A similar loss function is applied for the cyclic translation starting from $b \in B$.

In addition to the cycle-consistency loss, *identity loss* may be used to encourage the generator to retain features from the source that are also present in the target domain. For an image $a \in A$, the identity loss is defined as $\|a - a^{\text{idt}}\|$, where $a^{\text{idt}} \equiv \mathcal{G}_{B \rightarrow A}(a)$. A parallel formulation applies to domain B .

3.3. CycleGAN-like UI2I translation models

Here, we introduce three CycleGAN-like UI2I translation algorithms, ACL-GAN [22], U-GAT-IT [23], and UVCGAN [21], with special emphasis on the UVCGAN, or U-Net Vision-transformer (ViT) Cycle-consistent GAN, because of its outstanding performance on the SLATS dataset (see section 4).

The motivation behind ACL-GAN is that the stringent pixel-wise cycle-consistency loss may be a hurdle for generators to produce drastic changes such as large shape changes or removing/adding large objects. To solve the problem, ACL-GAN replaces strong cycle-consistency loss with a weaker adversarial consistency that does not require the cyclically translated image to match the source exactly, merely to match the distribution of the source images.

The authors of U-GAT-IT attack the problem of effective translation from another angle, keeping the cycle/identity-consistency loss functions in their original form as those in CycleGAN but renovating the generator and discriminator network structure. They use the class attention map to guide the generators and discriminators to focus on regions distinguishing between source and target domains. U-GAT-IT has achieved outstanding performance in translation between selfie photos and anime characters, which is a tough image translation task. One downside of U-GAT-IT is its model is bulky and slow, which may limit its application to LArTPC research should throughput and computing resources become pressing considerations.

Based on this work, the UVCGAN model performs the best for translations between the two SLATS dataset domains. UVCGAN improves CycleGAN by renovating its generator network and the training procedure. The UVCGAN generator is a hybrid architecture of a U-Net backbone [29] with a ViT bottleneck [30]. U-Net is known for its outstanding accuracy in modeling local or short-range patterns and its application in the segmentation of medical images. However, it may be less effective at capturing

long-range dependencies. Conversely, based on its impressive performance in image classification [31], ViT excels in capturing long-range dependencies and semantic relationships within an image. Nevertheless, relying solely on ViT may be insufficient for addressing the complexity of an image translation task, a regression problem in nature, as it may struggle to model details. Hence, the hybrid generator architecture of UVCGAN amalgamates the strengths of convolution-based networks and ViT, striking a balance between local and long-range pattern recognition.

3.4. Alternative models for unpaired image translation

Numerous models have been developed for UI2I translation, primarily on non-scientific image datasets. The models can be categorized based on two perspectives: the DL paradigm the algorithm is based upon and the way that consistency is enforced. For example, based on the paradigm, CycleGAN [20], ACL-GAN [22], U-GAT-IT [23], Council-GAN [32], and UVCGAN [21] are GAN-based methods. CUT [33] adopts the contrastive learning paradigm. LETIT [34] utilizes the energy transport on the latent feature space, while EGSDE [24] and ILVR [35] are diffusion-based models. In terms of consistency enforcement, CycleGAN, ACL-GAN, U-GAT-IT, UVCGAN, and CUT impose explicit consistency constraints via loss functions, while the other methods do so implicitly.

Another key feature of UI2I translation algorithms is the use of artificial randomness in the image generation process. Among the aforementioned models, CycleGAN, U-GAT-IT, UVCGAN, and CUT are deterministic, while Council-GAN, EGSDE, and ILVR inject randomness into image generation. Although randomness helps boost diversity in natural image translation tasks as there tends to be no single correct translation corresponding to an input, its application to SLATS is unnecessary. Specifically, for this study of the idealized SLATS dataset, the map between the two domains is *one-to-one* in nature, which makes a deterministic model the more appropriate choice.

Given the limitation on time and computing resources, we focus on four models that enforce cycle consistency explicitly because models without explicit cycle consistency place virtually no constraints on the output and may generate images unrelated to the input.

4. Evaluation

As part of this work, the performance of the neural translation algorithms CycleGAN, ACL-GAN, U-GAT-IT, and UVCGAN is evaluated on the paired test set of SLATS.

First, we perform a direct pixel-wise comparison of the translated detector readouts (ADC waveform images) with their targets. This comparison will indicate the quality of translation on the raw detector readout level. Second, a signal processing algorithm (see appendix D, and [36]) is applied to estimate physically meaningful counts of ionized electrons (see section 2) from the raw detector readouts. The signal processing algorithm is designed to perform accurately on domain A , and it exhibits domain-shift-related performance degradation when applied to the data from domain B . Using the signal processing algorithm allows for estimating the degree to which the UI2I translation algorithms alleviate the domain-shift effects on physically meaningful quantities.

Of note, making CycleGAN, ACL-GAN, U-GAT-IT, and UVCGAN work on the SLATS dataset required several modifications. To explore their potential, we conducted a small-scale hyperparameter (HP) tuning on each of the algorithms. For simplicity, all results in this section are produced by the best-performing HP settings of each algorithm. The details regarding model modification, HP tuning, and training are available in appendix B for UVCGAN and in appendix C for the other three CycleGAN-like models.

4.1. Translation quality evaluated on ADC waveforms

To quantitatively estimate the quality of the ADC waveform translations, we calculate ℓ_1 (mean absolute error) and ℓ_2 (mean squared error) between the translated and ground truth images. Table 1 summarizes the best-performing results, while the complete results for all HP settings can be found in appendix table E1.

Two samples from the $A \rightarrow B$ translation in figure 6 and another two from the $B \rightarrow A$ translation in figure 7 are presented for a qualitative comparison, which shows all algorithms manage to reproduce the key features of the target domain in the translations to some extent. The features include more extended tracks and 'lobe' structures at the track tips in the $A \rightarrow B$ translation and compactified tracks and more abrupt track tips in the $B \rightarrow A$ translation (see section 2.3). However, there are several noticeable defects in the translations, such as rugged track edges, large errors in the track center, missing 'lobes' near track tips in $A \rightarrow B$ translation, and incompletely reduced track edges in $B \rightarrow A$ translation. That said, all of the translation algorithms perform reasonably well in maintaining a strong consistency with the input images.

Table 1. Translation performance comparison with ℓ_1 and ℓ_2 differences on the ADC waveform. The differences are produced with the best performer of each algorithm. Full results with all HP settings can be found in appendix table E1. The best results are highlighted in bold.

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

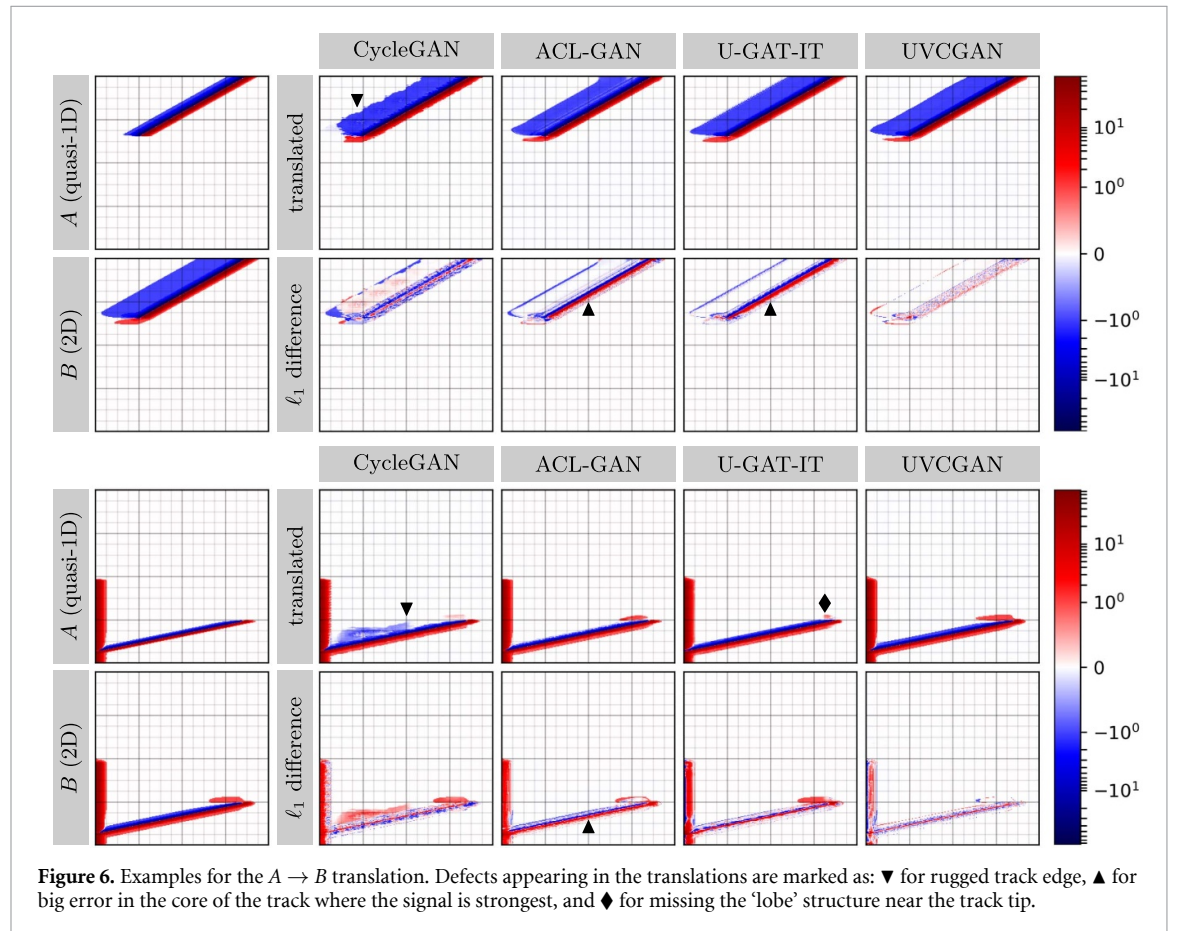


Figure 6. Examples for the $A \rightarrow B$ translation. Defects appearing in the translations are marked as: ▼ for rugged track edge, ▲ for big error in the core of the track where the signal is strongest, and ◆ for missing the ‘lobe’ structure near the track tip.

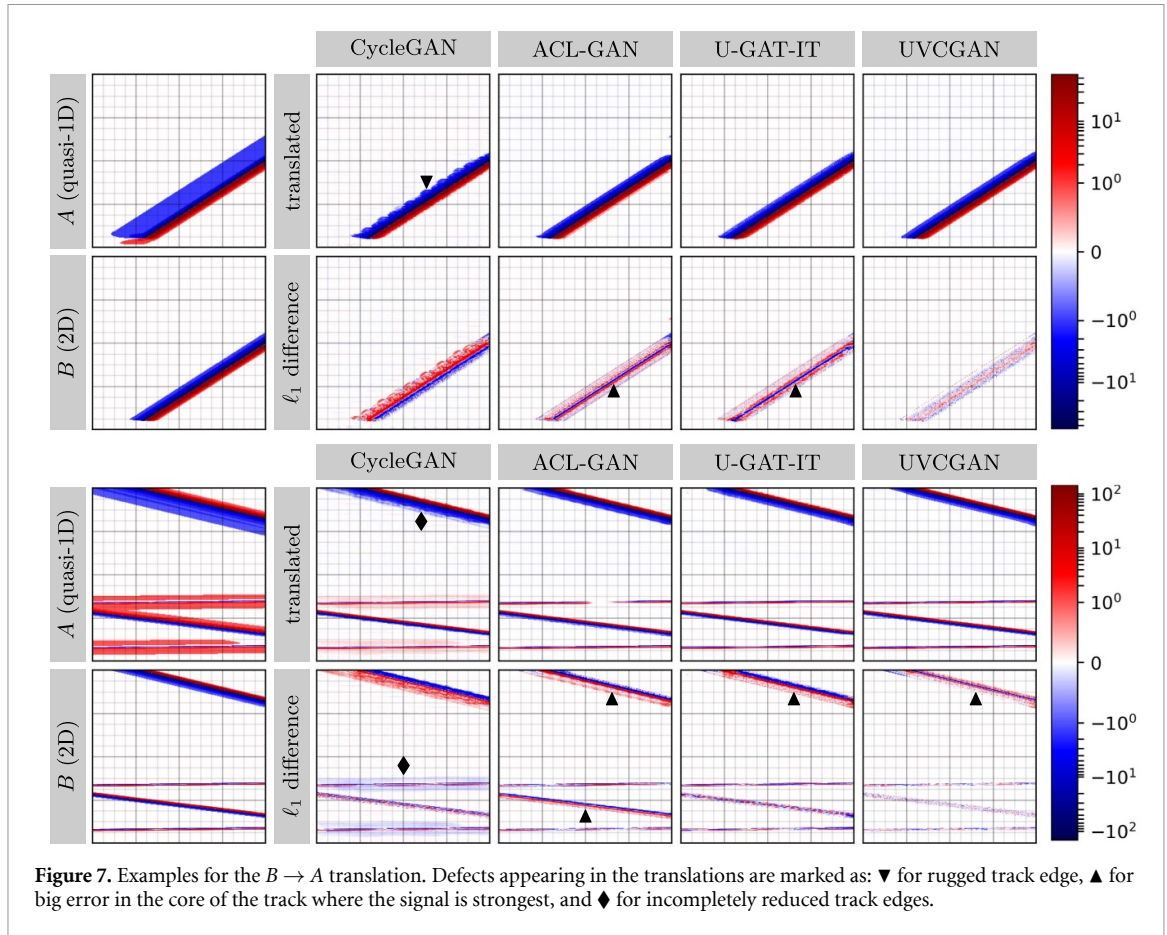
4.2. Translation quality evaluated on signal processing results

The raw LArTPC detector readouts are represented as ADC waveforms. These waveforms are difficult to interpret and have no direct relation to the physical properties of the particles that created them. Therefore, instead of building detector reconstruction pipelines directly on such waveforms, a signal processing algorithm is run first. The signal processing algorithm is designed to infer the original, physically meaningful distribution of ionized electrons that induced a particular waveform (see section 2). The recovered distributions of ionized electrons serve as a basis for the downstream reconstruction algorithms.

The signal processing algorithm involves two main stages: deconvolution of an ADC readout and high-pass filtering. The deconvolution operation is designed to act as an inverse of the simulated detector response function. Therefore, it is affected by the domain shift, as the simulated detector response may differ from the real detector response. Since signal processing is the first stage of the detector reconstruction pipelines, its domain shift error is then propagated to downstream algorithms.

The second stage of the signal processing algorithm is high-pass filtering. It is required as the bipolar nature of the deconvolution operator tends to amplify low-frequency noise. An adaptive high-pass filter, referred to as the *signal region-of-interest selection*, is subsequently applied to mitigate the impact of this amplification. Further details of this algorithm can be found in appendix D.

In this study, where domain B is used as a proxy to real detector data, we naturally use the quasi-1D response function to design the signal process procedure and denote the signal processing as ϕ_A . Because



domain A has a matching simulation and signal processing, the electron counts $\phi_A(a)$ reconstructed from $a \ll \text{disp ID disp1} > B$ will be less accurate than its counterpart $\phi_A(a)$. The difference between $\phi_A(a)$ and $\phi_A(b)$ is an indicator of the severity of the domain shift problem. Consequently, the reduction in the difference resulting from the translation can be viewed as the extent to which the domain shift problem is mitigated.

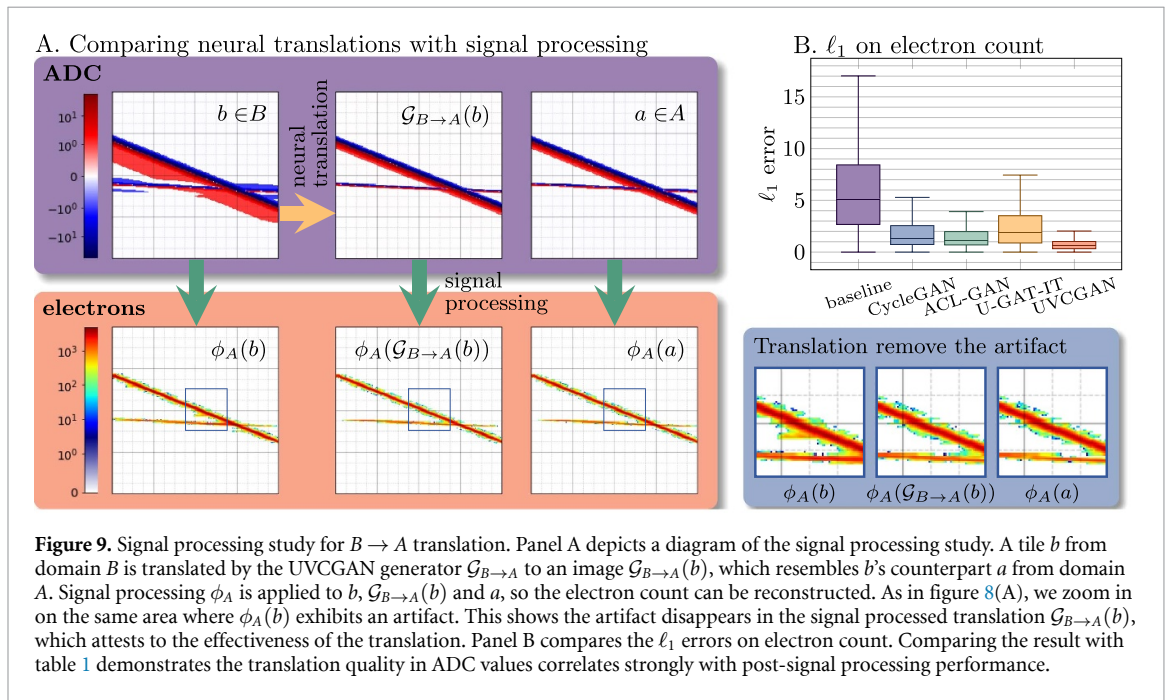
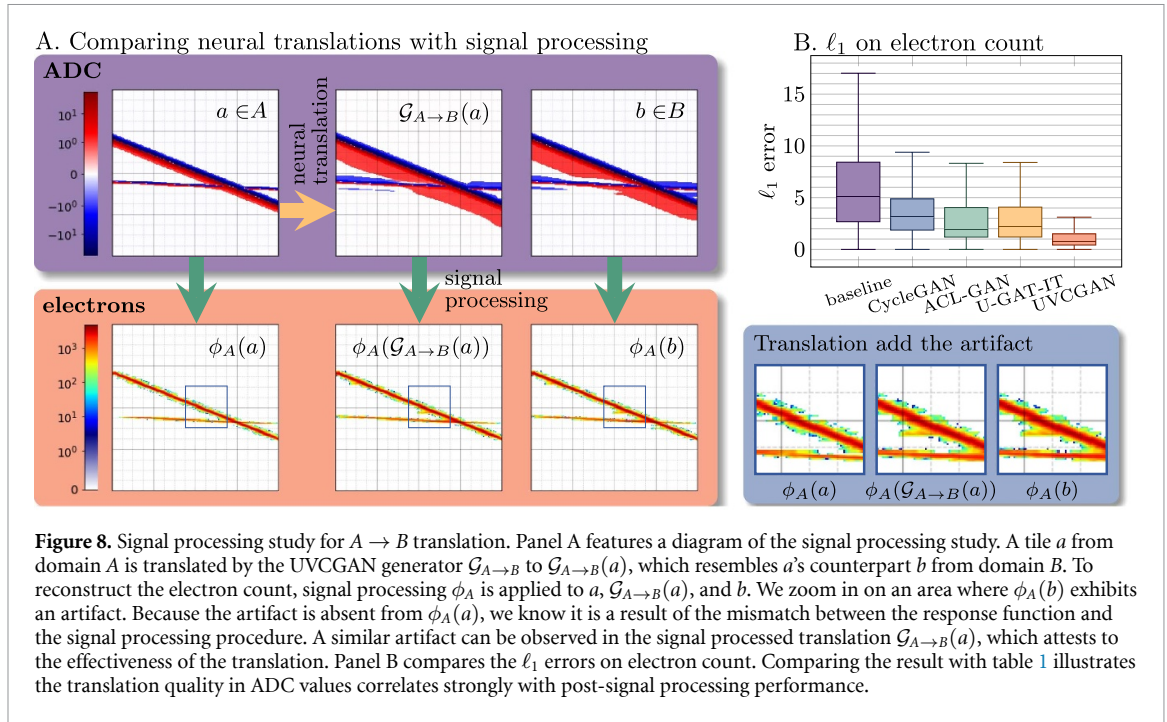
To carry out a quantitative study, we randomly sample 1000 pairs of test SLATS tiles. We calculate the baseline ℓ_1 error $\|\phi_A(a) - \phi_A(b)\|_1$ and translation ℓ_1 error $\|\phi_A(\mathcal{G}_{A \rightarrow B}(a)) - \phi_A(b)\|_1$ and $\|\phi_A(\mathcal{G}_{B \rightarrow A}(b)) - \phi_A(a)\|_1$, where $\mathcal{G}_{A \rightarrow B}$ and $\mathcal{G}_{B \rightarrow A}$ are neural translators. Figures 8(B) and 9(B) reveal the statistics.

It is worth noting that because ϕ_A is designed based on the detector response function of domain A , only the comparison between $\|\phi_A(\mathcal{G}_{B \rightarrow A}(b)) - \phi_A(a)\|_1$ and the baseline $\|\phi_A(a) - \phi_A(b)\|_1$ can be used to infer the extent to which a translation can reduce the domain-shift effect. However, $\|\phi_A(\mathcal{G}_{A \rightarrow B}(a)) - \phi_A(b)\|_1$ also is a valid indicator of the translation efficacy as it measures the translation's sensitivity in capturing the inaccuracy resulting from the domain shift.

These comparisons show that translations produced by all algorithms do improve upon the baseline with UVCGAN being the best performer in both translation directions. Notably, UVCGAN achieves a greater than 80% reduction in ℓ_1 error over the baseline on average for the $B \rightarrow A$ translation. Comparing the result in Table 1 shows that the translation quality measured in electron counts correlates strongly with those featuring ADC values. Additional evaluation of post-signal processing performance is provided in appendix E.

For qualitative comparison, the signal-processed results for one sample SLATS tile are shown in figure 8(A) for the $A \rightarrow B$ translation and figure 9(A) for the $B \rightarrow A$ translation. The translated images in both figures are produced by the best performer, UVCGAN. Due to the mismatch in response functions, the signal-processed result for a domain B sample may exhibit artifacts along the periphery of the tracks as exemplified by the area marked with the blue box. We anticipate an effective $A \rightarrow B$ translation to replicate the artifact, while a $B \rightarrow A$ translation should eliminate it. This expectation aligns with the translations generated by UVCGAN, where the artifact is introduced in $\phi_A(\mathcal{G}_{A \rightarrow B}(a))$ and is removed in $\phi_A(\mathcal{G}_{B \rightarrow A}(b))$.

Considering the neural translation algorithm is trained in a purely data-driven fashion, i.e. solely based on the ADC waveform images without any input or constraints from physics or downstream applications, these are promising results.



4.3. Domain shift mitigation for a supervised learning algorithm

In this section, we investigate the effectiveness of UI2I translation techniques in mitigating the domain shift problem in a supervised learning context. Specifically, we design a supervised DL regression model to predict the number of ionized electrons from an ADC waveform. This model exhibits decreased performance when trained on domain A and applied to domain B . Then, we test whether the UI2I translations can alleviate this degradation of the model performance.

The experiment proceeds as follows: we train predictive models— E_A , E_B , and E_G —to estimate the total count of ionized electrons e in a waveform. Here, \mathcal{G} is a neural translator that translates ADC waveform images from domain A to domain B , such as CycleGAN, ACL-GAN, U-GAT-IT, and UVCGAN. The models E_A , E_B , and E_G are trained using waveforms $a \in A$, waveforms $b \in B$, and translated waveforms $\mathcal{G}_{A \rightarrow B}(a)$ for $a \in A$, respectively.

After training, we evaluate all the models on waveforms from domain B . Due to the domain shift, we expect that E_A will perform worse on B than E_B . However, since E_G is trained on translated waveforms that

Table 2. Models trained on translated images mitigate the domain shift problem, as measured by MARE (lower is better). E_A , trained on domain A and applied to domain B , represents the worst-case scenario for domain shift. E_B , trained and tested on domain B , serves as the performance benchmark. The E_G models, trained on waveforms translated by a neural translator \mathcal{G} and tested on domain B , demonstrate varying degrees of effectiveness in mitigating domain shift through UI2I translation methods.

	E_A	E_B	E_{CycleGAN}	$E_{\text{ACL-GAN}}$	$E_{\text{U-GAT-IT}}$	E_{UVCGAN}
B	0.390	0.211	0.222	0.223	0.257	0.216

closely resemble those in domain B , we expect that it will outperform E_A when tested on B . We use an AlexNet-like [37] architecture for the regression model. Further details on the model are provided in appendix F.

Table 2 summarizes the results of the evaluation of the regression models on domain B . The performance is evaluated with mean absolute relative error (MARE), calculated as $n^{-1} \sum_{i=1}^n |(\hat{e}_i - e_i)/e_i|$, where \hat{e}_i represents the predicted total electron count and n is the number of test examples.

As expected, E_A exhibits the worst performance due to the domain shift. E_B performs the best, as it was trained and tested on data from the same domain. The four E_G models, trained on translated waveforms, show varying degrees of effectiveness in mitigating the domain shift problem. Notably, E_{UVCGAN} outperforms others, achieving comparable MARE to E_B . This aligns with our earlier findings from the pixel-wise difference analysis in section 4.1.

4.3.1. Discussion and future research direction

Findings from this work highlight the potential of UI2I translation algorithms in addressing the challenges of domain shift in LArTPC data. However, several issues require attention before these algorithms can be effectively used to translate between simulated and real detector data.

4.3.1.1. Scaling UI2I algorithms to work on large images.

Existing UI2I translation algorithms have been developed and tested on images of size (256, 256). Thus, the same-sized tiles are used in this study. However, full LArTPC images of size (800, 6000) are needed for downstream analyses. In applying the model to tiles and assembling them to form the full translated image, mismatches did occur along the tile boundaries. Therefore, as part of our future work, we need to develop network models and computational pipelines capable of handling full images.

4.3.1.2. Performing and Evaluating non-deterministic translations.

Another crucial aspect is the one-to-one nature of the translation. In this work, we addressed a problem where the translation between the two domains is fully deterministic and one-to-one. However, in real detectors, multiple stochastic processes are present. These stochastic processes will render the domain map non-deterministic, resulting in either one-to-many or even many-to-many relationships. The non-determinism of the translation presents two challenges: (1) how to adapt UI2I translation methods to handle non-deterministic mappings, and (2) how to evaluate the quality of non-deterministic translations.

There are multiple ways to make a UI2I translation non-deterministic. ACL-GAN [22] presents one such approach, replacing a strong cycle-consistency constraint with a weaker adversarial consistency. The DRIT family of models [38] demonstrates another method, separating the content (core part of the image that should be preserved) and the attributes (part of the image that changes during the translation). This separation allows for substituting multiple attributes for a single translation, resulting in a variety of output images. BiCycleGAN [39] shows another interesting way to construct a one-to-many mapping ($A \rightarrow B$) by adding an extra latent dimension L to the A domain. Then, it constructs a one-to-one map ($A \times L \rightarrow B$). This method allows us to obtain a one-to-many translation ($A \rightarrow B$) by varying points in the latent dimension L . The approaches presented by ACL-GAN, DRIT, and BiCycleGAN demonstrate that developing one-to-many and many-to-many translations is possible, indicating a promising direction for future research.

The shift to non-deterministic translations raises a question of how to evaluate the quality of the translation in a non-deterministic case. In this work, we were able to construct a paired ground truth evaluation dataset due to the one-to-one nature of the problem. This exact pairing allowed us to estimate the translation quality directly by comparing translated images to their ground truths. However, in the case of one-to-many translation, constructing such a paired evaluation dataset becomes impossible. Therefore, more sophisticated metrics are required to judge the quality of these non-deterministic translations.

We believe a robust evaluation protocol should focus on two aspects of the translation: (1) realism, a neural translator's ability to replicate the distinctive features of the target domain during translation, and (2) consistency, its capacity to translate without altering the underlying physical properties. In non-scientific UI2I translation tasks, established metrics such as Fréchet inception distance [40] and Kernel inception

distance [41] are commonly used to assess realism. However, these metrics are based on the InceptionV3 network [42] pretrained on the ImageNet dataset [43], raising doubts about their applicability to scientific datasets. On the other hand, translation consistency remains a relatively unexplored aspect of UI2I translation research. The exact definition of consistency in UI2I translation is likely dataset- and application-dependent. As far as we know, no established metrics or protocols exist for verifying such consistency, making this a critical area for future research in UI2I translation for scientific applications.

4.3.1.3. (Re)evaluation of systematic uncertainties in the presence of UI2I translation.

High-energy physics (HEP) experiments developed complex methods to estimate various systematic uncertainties affecting the final results (e.g. [44]). Incorporating UI2I translation into the standard simulation chain may present unique challenges specific to HEP experiments. These challenges are twofold: (1) estimating the systematic uncertainty (if any) stemming from the UI2I translation, and (2) understanding how the UI2I translation affects the already established systematic uncertainties.

While UI2I translation algorithms aim to bring the simulated (A) and experimental (B) domains closer, they may introduce artifacts into the translation process. These artifacts could be the source of additional systematic uncertainty, which may need to be quantified. One possible approach to establishing the magnitude of such uncertainty is to train an ensemble of UI2I models and analyze the amount of variance introduced in the experimental results by the ensemble. Alternatively, the UI2I translation could be treated as a ‘detector calibration’ step, without assigning specific systematic uncertainties to it. In this case, the uncertainties associated with UI2I would be incorporated into the uncertainties of other detector simulation parameters.

While UI2I translation techniques show promise in reducing the magnitude of systematic uncertainties, these reduced uncertainties still require careful evaluation. This evaluation process may involve a substantial effort to understand how UI2I translation algorithms interact with established methods of estimating systematic uncertainties in HEP experiments. Further research is necessary to determine the most effective approaches for handling UI2I-related uncertainties in HEP experiments.

Conclusion

In this work, we studied the potential of the UI2I translation algorithms to address the domain shift problem between simulation (domain A) and real data (domain B) in the LArTPC research. We constructed a surrogate LArTPC problem consisting of two simulated domains with a systematic difference in the detector response function. This surrogate problem illustrates the typical source of the systematic uncertainty between the simulation and real data. The deterministic nature of the detector response function allowed us to create a paired test dataset with the known ground truths for translations.

We tested four UI2I models (CycleGAN, ACL-GAN, U-GAT-IT, UVCGAN) on the surrogate LArTPC problem. Our results show that the UI2I methods can successfully perform the translation of LArTPC events as judged by pixel-wise metrics between the translation and the corresponding ground truth. Notably, UI2I methods can identify and preserve the content of each event while translating its appearance. This indicates the feasibility of the application of the UI2I methods to translate LArTPC data and improve the realism of the LArTPC simulation.

Furthermore, we tested whether the obtained UI2I translations allow us to reduce the domain shift error of detector reconstruction algorithms, which are developed on simulation but applied to real data. For this purpose, we employed a production-grade signal processing algorithm designed on simulation (domain A). This algorithm experiences domain shift error when applied directly on domain B . However, we found that its domain error can be reduced by up to 80% if we perform a UI2I translation ($B \rightarrow A$) before the application of the signal processing algorithm. These results indicate that UI2I methods can be used for domain shift reduction in LArTPC analysis.

Among the four tested UI2I models (CycleGAN, ACL-GAN, U-GAT-IT, UVCGAN), the UVCGAN model achieves the best translation quality and introduces the fewest artifacts in the translated images. This finding indicates that the UVCGAN model shows promise as a basis for more complex UI2I algorithms on scientific data. To promote the reproducibility of our research, we publicly released the SLATS dataset (<https://zenodo.org/record/7809108>) and the code used in this study (<https://github.com/LS4GAN/uvrgan4slats>).

While UI2I methods show promise in reducing the systematic differences between distinct domains of LArTPC data and help to alleviate the domain shift error of the signal processing algorithm, there are several issues that remain to be addressed before their application becomes fully feasible. First, the UI2I methods, currently developed on images up to 256 pixels in size, need to be scaled to work with larger images of up to 10 000 pixels. Second, our work investigated a problem where the relationship between two domains is one-to-one. The actual relationship between the LArTPC detector simulation and real data is many-to-many.

The performance of UI2I methods needs to be studied under many-to-many relationships. Moreover, proper translation quality metrics need to be developed for the many-to-many case. Finally, while UI2I methods may reduce the systematic difference between simulated and real data, one still needs to estimate potential systematic uncertainties introduced by these methods. Likewise, work needs to be done to ascertain how the inclusion of the UI2I translation into the detector simulation pipeline may affect other systematic uncertainties. Exploring these directions will be essential to fully leverage the potential of UI2I methods in LArTPC research and broader scientific applications.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: Source Code: <https://github.com/LS4GAN/uvcgan4slats>. Training Data: <https://zenodo.org/records/7809108>.

Acknowledgments

This work was supported by the Laboratory Directed Research and Development Program of Brookhaven National Laboratory, which is operated and managed for the U.S. Department of Energy Office of Science by Brookhaven Science Associates under contract No. DE-SC0012704.

ORCID iDs

Yi Huang  <https://orcid.org/0000-0002-7024-7469>

Dmitrii Torbunov  <https://orcid.org/0000-0003-0132-5344>

Haiwang Yu  <https://orcid.org/0000-0002-2973-4580>

Jin Huang  <https://orcid.org/0000-0002-6261-6172>

Meifeng Lin  <https://orcid.org/0000-0003-3010-0857>

Yihui Ren  <https://orcid.org/0000-0002-5750-6964>

References

- [1] Wang M and Deng W 2018 Deep visual domain adaptation: a survey *Neurocomputing* **312** 135–53
- [2] Miotto R, Wang F, Wang S, Jiang X and Dudley J T 2017 Deep learning for healthcare: review, opportunities and challenges *Brief. Bioinform.* **19** 1236–46
- [3] Min S, Lee B and Yoon S 2017 Deep learning in bioinformatics *Brief. bioinform.* **18** 851–69
- [4] Sapoval N *et al* 2022 Current progress and open challenges for applying deep learning across the biosciences *Nat. Commun.* **13** 1728
- [5] Quionero-Candela J, Sugiyama M, Schwaighofer A and Lawrence N D 2009 *Dataset Shift in Machine Learning* (The MIT Press)
- [6] Takahashi C C and Braga A P 2020 A review of off-line mode dataset shifts *IEEE Comput. Intell. Mag.* **15** 16–27
- [7] Fang Y, Yap P-T, Lin W, Zhu H and Liu M 2022 Source-free unsupervised domain adaptation: a survey (arXiv:2301.00265)
- [8] Rubbia C 1977 The liquid-argon time projection chamber: a new concept for neutrino detectors *Technical report* (European Organization for Nuclear Research)
- [9] Willis W J and Radeka V 1974 Liquid-argon ionization chambers as total-absorption detectors *Nucl. Instrum. Methods* **120** 221–36
- [10] Nygren D R 1974 The time projection chamber: a new 4 pi detector for charged particles *eConf* vol C740805 p 58
- [11] Acciarri R *et al* 2017 Design and construction of the microboone detector *J. Instrum.* **12** 02017
- [12] Abi B *et al* 2017 The single-phase protodune technical design report (arXiv:1706.07081)
- [13] Abi B *et al* 2018 The dune far detector interim design report volume 1: physics, technology and strategies (arXiv:1807.10334)
- [14] Csurka G 2017 Domain adaptation for visual applications: a comprehensive survey (arXiv:1702.05374)
- [15] Wilson G and Cook D J 2020 A survey of unsupervised deep domain adaptation *ACM Trans. Intell. Syst. Technol.* **11** 1–46
- [16] Guan L and Yuan X 2023 Iterative loop learning combining self-training and active learning for domain adaptive semantic segmentation (arXiv:2301.13361)
- [17] Hoyer L, Dai D, Wang H and Van Gool L 2023 Mic: masked image consistency for context-enhanced domain adaptation *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition* pp 11721–32
- [18] French G, Mackiewicz M and Fisher M 2017 Self-ensembling for visual domain adaptation (arXiv:1706.05208)
- [19] Liu M-Y, Breuel T and Kautz J 2017 Unsupervised image-to-image translation networks *Advances in Neural Information Processing Systems* vol 30
- [20] Zhu J-Y, Park T, Isola P and Efros A A 2017 Unpaired image-to-image translation using cycle-consistent adversarial networks *Proc. IEEE Int. Conf. on Computer Vision* pp 2223–32
- [21] Torbunov D, Huang Y, Haiwang Y, Huang J, Yoo S, Lin M, Viren B and Ren Y 2023 Uvcgan: unet vision transformer cycle-consistent gan for unpaired image-to-image translation *2023 IEEE/CVF Winter Conf. on Applications of Computer Vision (WACV) (Waikoloa, HI, USA)* (<https://doi.org/10.1109/WACV56688.2023.00077>)
- [22] Zhao Y, Ruihai W and Dong H 2020 Unpaired image-to-image translation using adversarial consistency loss *European Conf. on Computer Vision* (Springer) pp 800–15
- [23] Kim J, Kim M, Kang H and Lee K 2019 U-gat-it: unsupervised generative attentional networks with adaptive layer-instance normalization for image-to-image translation (arXiv:1907.10830)
- [24] Zhao M, Bao F, Chongxuan Li and Zhu J 2022 Egsde: unpaired image-to-image translation via energy-guided stochastic differential equations (arXiv:2207.06635)
- [25] Abi B *et al* 2017 The single-phase protodune technical design report

- [26] Abi B *et al* 2020 First results on ProtoDUNE-SP liquid argon time projection chamber performance from a beam test at the CERN Neutrino Platform *J. Instrum.* **15** 12004
- [27] Ramo S 1939 Currents induced by electron motion *Proc. IRE* vol 27 pp 584–5
- [28] Goodfellow I, Pouget-Abadie J, Mirza M, Bing X, Warde-Farley D, Ozair S, Courville A and Bengio Y 2020 Generative adversarial networks *Commun. ACM* **63** 139–44
- [29] Ronneberger O, Fischer P and Brox T 2015 U-net: convolutional networks for biomedical image segmentation *Int. Conf. on Medical Image Computing and Computer-Assisted Intervention* (Springer) pp 234–41
- [30] Dosovitskiy A *et al* 2020 An image is worth 16x16 words: transformers for image recognition at scale (arXiv:2010.11929)
- [31] Anokhin I, Demochkin K, Khakhulin T, Sterkin G, Lempitsky V and Korzhnikov D 2021 Image generators with conditionally-independent pixel synthesis *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition* pp 14278–87
- [32] Nizan O and Tal A 2020 Breaking the cycle-colleagues are all you need *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition* pp 7860–9
- [33] Park T, Efros A A, Zhang R and Zhu J-Y 2020 Contrastive learning for unpaired image-to-image translation *European Conf. on Computer Vision* (Springer) pp 319–45
- [34] Zhao Y and Chen C 2021 Unpaired image-to-image translation via latent energy transport *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition* pp 16418–27
- [35] Choi J, Kim S, Jeong Y, Gwon Y and Yoon S 2021 Ilvr: conditioning method for denoising diffusion probabilistic models (arXiv:2108.0293)
- [36] Adams C *et al* 2018 Ionization electron signal processing in single phase LArTPCs. Part I. Algorithm description and quantitative evaluation with MicroBooNE simulation *J. Instrum.* **13** 07006–07006
- [37] Krizhevsky A, Sutskever I and Hinton G E 2012 Imagenet classification with deep convolutional neural networks *Advances in Neural Information Processing Systems* vol 25, ed F Pereira, C J Burges, L Bottou and K Q Weinberger (Curran Associates, Inc)
- [38] Lee H-Y, Tseng H-Y, Huang J-B, Singh M and Yang M-H 2018 Diverse image-to-image translation via disentangled representations *Proc. European Conf. on Computer Vision (ECCV)* pp 35–51
- [39] Zhu J-Y, Zhang R, Pathak D, Darrell T, Efros A A, Wang O and Shechtman E 2017 Toward multimodal image-to-image translation *Advances in Neural Information Processing Systems* vol 30
- [40] Heusel M, Ramsauer H, Unterthiner T, Nessler B and Hochreiter S 2017 Gans trained by a two time-scale update rule converge to a local nash equilibrium *Advances in Neural Information Processing Systems* vol 30
- [41] Bińkowski M, Sutherland D J, Arbel M and Gretton A 2018 Demystifying mmd gans (arXiv:1801.01401)
- [42] Szegedy C, Vanhoucke V, Ioffe S, Shlens J and Wojna Z 2015 Rethinking the inception architecture for computer vision (arXiv:1512.00567)
- [43] Russakovsky O *et al* 2015 Imagenet large scale visual recognition challenge *Int. J. Comput. Vis.* **115** 211–52
- [44] Abratenko P *et al* 2022 Search for an anomalous excess of inclusive charged-current νe interactions in the microboone experiment using wire-cell reconstruction *Phys. Rev. D* **105** 112005