Deep Learning Approaches for Day-to-Night Image Translation: A Review with Implications for Oil and Gas Pipeline Inspection

Abdulkarim Dahiru.[1*], Sulaimon Adebayo Bashir[2], Enesi Femi Aminu[3], Eyitayo Amos Afolabi[4]

[1,2,3]Department of Computer Science, School of ICT, Federal University of Technology, Minna

[4]Department of Chemical Engineering, Federal University of Technology, Minna

*Correspondenc: eldaheer@gmail.com

ABSTRACT

Pipeline infrastructure is the most reliable, efficient, and cost-effective means of transporting crude oil and natural gas globally. Protecting this critical infrastructure is imperative because pipeline failures have significant economic consequences, affecting not only the oil and gas sector but also the Gross Domestic Product (GDP) of many oil-producing nations. Inadequate and inaccessible datasets, particularly the lack of nighttime pipeline leak images, reduce the robustness of deep learning models for pipeline leak detection under varying illumination conditions. This research investigates deep learning approaches for daytime-to-nighttime cross-domain pipeline infrastructure image-to-image translation to support early warning systems for pipeline leak monitoring. The original CycleGAN framework, without architectural modifications, was used for daytime-to-nighttime pipeline image translation due to its unpaired learning capability. The model's performance was evaluated using Fréchet Inception Distance (FID), Structural Similarity Index Measure (SSIM), and Visual Evaluation (VE). Although CycleGAN improves visual realism in cross-domain pipeline image-to-image translation from daytime to nighttime, the research quantitative results indicate that it suppresses critical anomaly features essential to early warning systems for pipeline leak detection. The study further reveals limitations in unpaired translation pertinent to pipeline infrastructure inspection, establishing a benchmark for cross-domain and anomaly-preserving image translation research for pipeline infrastructure inspection early-warning systems aimed at mitigating economic crises.

Keywords: CycleGAN, Image-to-Image Translation, Pipeline Leak Detection, Early Warning Systems, Domain Adaptation

1. Introduction

Pipeline infrastructure is critical for global crude oil and natural gas transportation, making its security imperative rather than optional. Pipelines transport most of the world's oil and gas, crossing both land and sea, which makes them vulnerable to vandalism, unintentional damage, corrosion, structural damage, theft, hot spots, and punctures (Wang *et al.*, 2023). Similarly, the transportation of petroleum and natural gas requires substantial investments because it has always been a significant concern for both the oil and gas industries and the government of many nations (Binod & Ashish, 2023).

Furthermore, despite various efforts to address threats to oil and gas pipelines, existing literature indicates that further research remains necessary. These threats affect not only the energy sector but also national GDPs of oil-producing nations.

A recurring challenge identified in the literature is the scarcity and limited accessibility of datasets, which limits research progress in this field and, in the same vein, weakens the performance of pipeline leak detection models. Ukaegbu *et al.* (2021), Jabbar*et al.* (2022), and Ali *et al.*(2023) mentioned that their models require larger and more diverse datasets to enhance the models' performance. Similarly, several studies recommend that nighttime pipeline leak images are necessary for training robust models capable of detecting leaks across varying lighting conditions. Literature such as Jiao *et al.* (2019), Thomas *et al.* (2020), and Muhammad and Ali (2023), their models were able to detect oil spills in images during the day but were unable to detect spills at night times. Therefore, this study aims to address this gap by tackling the nighttime pipeline leak images dataset challenge and also contribute meaningfully to the field of pipeline infrastructure inspection early warning systems.

1.1. Image-to-Image Translation Techniques

Image-to-Image translation, also referred to as I2I translation, is an emerging technique in deep learning and computer vision tasks that leverages existing work on Generative Adversarial Networks (GANs) in translating images from one domain to another domain or one style to another style. According to a study by Shatnawi and Bani (2024), GANs can be used to overcome challenges related to limited or rare image datasets. Furthermore, Generative Adversarial Networks have been used to generate chest images in the medical field (Agrawal & Choudhary, 2023). Similarly, GAN networks have been used to generate artificial images that not only enhance the dataset but also simulate realism in various difficult circumstances while maintaining the quality of the image dataset, which ultimately leads to more accurate detection of emergency vehicles (Shantnawe *et al.*, 2024).

Meanwhile, the oil and gas sector is yet to tap into this recent innovative technology, GAN, to enhance pipeline infrastructure inspection, especially in the I2I translation domain.
Some of the GAN notable techniques used for I2I translation include Pix2Pix and CycleGAN. Before the advent of GANs, some traditional techniques were utilised for I2I translations, such as Histogram Equalisation (HE) and Retinex. Traditional techniques, such as global transforms, can introduce artefacts or colour shift (Qian *et al.*, 2024). Similarly, Zhao *et al.* (2022) also reiterated that traditional I2I translation techniques are often associated with poor generalisation. In essence, the traditional techniques are not content-aware because they have a high chance of introducing artefacts, which are unwanted objects in an image during translation and in the case of the pipeline infrastructure inspection domain, traditional I2I translation techniques could even hide or wash out crucial pipeline

anomaly features, such as a leak in an image during translation, making the translated image lose realism. Nowadays, GAN variants such as Pix2Pix and CycleGAN have received global research attention in I2I translation and were developed as improvements and also automation of the traditional I2I translation in order to make the translation process more efficient and also maintain realism in the translated images. Furthermore, the later sections of this research highlight a broad discussion on the I2I translation techniques.

1.1.1. Traditional I2I Translation

Traditional I2I translation techniques can be utilised to enhance image illumination by manipulating pixel intensities or decomposing illumination and reflectance using techniques such as histogram equalisation (HE) and Retinex.

Retinex I2I translation technique decomposes an image into illumination and reflectance layers, then amplifies the illumination layer and reconstruct them afterwards (Lee *et al.*, 2025). Similarly, a Retinex-based I2I translation technique was used to brighten low-light weld X-ray images (Qian *et al.*, 2024).

However, despite the success recorded by the traditional techniques, Retinex and Histogram Equalisation lack generalisation because they introduce artefacts or unwanted objects and may even wash out important parts of the translated images. The introduction of unwanted artefacts and colour shift in a translated image poses a very serious economic concern.
More especially, oil and gas sector pipeline anomaly detection. Hence, there is a need for an I2I translation technique that is pipeline content-aware and ensures realistic translated images without introducing artefacts or washing out important pipeline anomaly features such as leaks during translation.

1.1.2. GAN-Based I2I Translation

Pix2Pix is a pioneering supervised I2I translation technique; it uses one generator and one discriminator and requires paired images from the same scene, such as a sketch and its corresponding photograph, to train the generator, with the discriminator learning to distinguish real pairs from generated ones (Jason, 2021; Eason, 2023). Pix2Pix generator utilised a Convolutional Neural Network (CNN) based on U-Net, while its discriminator utilised a CNN based on PatchGAN. According to Eason (2023), Pix2Pix has been applied to several tasks, such as converting black-and-white images to colored photos and translating sketches into photos. However, Pix2Pix's dependence on a paired dataset limits its application in domains such as daytime-to-nighttime I2I translation where corresponding paired dataset scenes are difficult or impossible to obtain (Lee *et al.*, 2023).

CycleGAN was proposed as an unpaired alternative, using two generators (G, F) and two discriminators (D_X, D_Y) to translate images between domains without requiring paired scenes (Zhu *et al.*, 2017). It introduces adversarial loss, which enables the generator to translate images indistinguishable from real images by the discriminator, and cycle-consistency loss, which ensures that an image translated to the target domain is reconstructed back to the original input image. The adversarial loss is defined as:

$$L\mathrm{GAN} = L\mathrm{GAN}(G, DY, X, Y) + L\mathrm{GAN}(F, DX, Y, X) \tag{1}$$

where

$$L\mathrm{GAN}(G, DY, X, Y) = Ey{\sim}p\mathrm{data}(y)\,[\log DY(y)] + Ex{\sim}p\mathrm{data}(x)\,[\log(1 - DY(G(x)))] \tag{2}$$

$$LGAN(F, DX, Y, X) = Ex \sim p\mathrm{data}(x)\ [\log DX\ (x)] + Ey \sim p\mathrm{data}(y)\ [\log(1 - DX\ (F(y)))] \tag{3}$$

Cycle-consistency loss preserves the structural content of the input image such that forward and backward mappings (G: X→Y, F: Y→X) satisfy F(G(x)) ≈ x and G(F(y)) ≈ y, computed using the L1 distance:

$$Losscycle(G,F) = Ex \sim pdata(x)[\|F(G(x)) - x\|_1] + Ey \sim pdata(y)[\|G(F(y)) - y\|_1] \tag{4}$$

The full CycleGAN objective combines both terms:

$$L(G, F, DX, DY) = LGAN(G, DY, X, Y) + LGAN(F, DX, Y, X) + \lambda Losscycle(G, F) \tag{5}$$

Compared with Pix2Pix, CycleGAN utilises two CNN-based ResNet generators and two CNN-based PatchGAN discriminators, enabling more realistic translation while preserving structural fidelity without paired data from the same scene (Eason, 2023). In essence, CycleGAN's unpaired learning approach is especially valuable in the oil and gas sector, where paired daytime-nighttime pipeline leak datasets are scarce or unavailable. This further inspired its use in this research to address dataset and domain-adaptation challenges for robust pipeline leak-detection early-warning systems that mitigate economic crises in oil-producing states.

1.2. Evaluation Metrics for I2I Translation

I2I evaluation metric is used as a criterion in evaluating image production and image-to-image translation (Mehran *et al.*, 2025). Similarly, the I2I evaluation metric is a major determinant in attesting whether an I2I translation technique achieves realism or a consistent structure after image translation. The Inception Score (IS) computes the KL divergence between the conditional and marginal label distributions using a pretrained Inception network, with higher values indicating greater image quality and diversity (Salimans *et al.,* 2016):

$$IS = \exp(\ Ex[KL(p(y|x)\ ||\ p(y))]) \tag{6}$$

The Fréchet Inception Distance (FID) measures the distance between the feature distributions of real and generated images extracted through the Inception network (Szegedy *et al.,* 2015; Mehran *et al.,* 2025), with lower scores indicating greater realism and a score of 0 indicating identical distributions:

$$FID = \|\mu_r - \mu_g\|_2^2 + Tr(\Sigma_r + \Sigma_g - 2(\Sigma_r\Sigma_g)^{1/2}) \tag{7}$$

The Structural Similarity Index Measure (SSIM) assesses luminance, contrast, and structural similarity, providing a more comprehensive measure of perceived visual quality than pixel-wise metrics such as PSNR (Wang *et al.,* 2004; Mehran *et al.,* 2025).

This research utilises FID, SSIM and Visual Evaluation (VE) metrics. The chosen I2I evaluation metrics are most relevant for assessing anomaly preservation in the pipeline infrastructure inspection domain.

2. Methodology

2.1. Pipeline Image Dataset

The pipeline images dataset used in this study combines 204 pipeline images from Ukaegbu *et al.* (2021) with 1,000 images from Roboflow, yielding 1,204 images in total. The dataset is organised into two domains and four classes: a Daytime domain comprising Leak and Non-Leak classes, and a Nighttime domain comprising Night Non-Leak images. However, no nighttime pipeline-leak images exist in the acquired dataset or in publicly available repositories. Therefore, this research addresses that gap by generating the missing nighttime-leak class using CycleGAN, thereby providing a foundational contribution to oil and gas pipeline infrastructure inspection early warning systems.

2.2. Preprocessing of the Pipeline Images Dataset

The collected images, originally organised only into leak and non-leak classes, were reorganised into the two domains and four classes required for this study to address the missing night-leak class. All images were resized to a uniform 256×256 pixels to reduce mode collapse and overfitting, and then divided into training and test sets.

2.3. Daytime-to-Nighttime Pipeline Image Translation Using CycleGAN

This research review shows that no existing GAN-based I2I technique is tailored to pipeline infrastructure inspection. Similarly, Pix2Pix requires paired data, while traditional techniques (Retinex, HE) risk washing out pipeline anomaly features (leaks). Therefore, this research utilises the original CycleGAN by Zhu *et al.* (2017).

The architecture is implemented without modifications using Python and the PyTorch framework. Furthermore, Figure 1 below shows the CycleGAN architecture for daytime-to-nighttime cross-domain pipeline I2I translation.

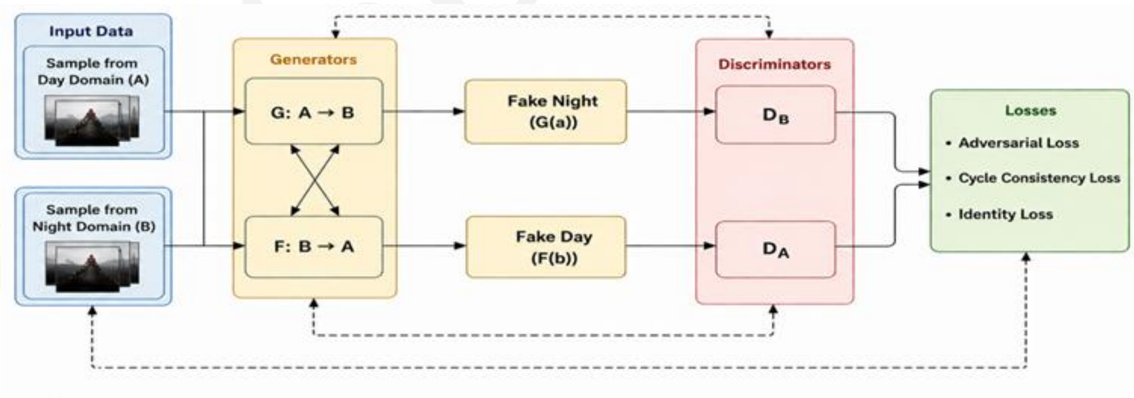

Figure 1. The CycleGAN Architecture for Daytime to Nighttime Cross-Domain Pipeline I2I Translation

The model follows the original CycleGAN architecture consisting of two generators and two discriminators. Similarly, Generator G**AB**: Translates images from Domain A (Daytime) to Domain B (Nighttime). While Generator G**BA**: Translates images from Domain B (Nighttime) to Domain A (Daytime), Discriminator D**A**: Distinguishes real vs generated Daytime images, and Discriminator D**B**: Distinguishes real vs generated Nighttime images.

Furthermore, each generator follows a ResNet architecture: an initial 7×7 convolution layer, two downsampling layers, nine residual blocks, two upsampling layers, and a Tanh-activated output layer, with instance normalisation used throughout to stabilise style transfer. Similarly, both discriminators utilise a 70×70 PatchGAN architecture, classifying image patches rather than the full image to emphasise high-frequency structures such as textures and edges. While Loss Functions training, combines three losses: (i) Least Squares GAN (LSGAN) adversarial loss, which stabilizes training and encourages realistic outputs; (ii) cycle-consistency loss (L1-based, Equation 4), which ensures Daytime-to-Nighttime and Nighttime-to-Daytime reconstruction fidelity; and (iii) identity loss, which discourages unnecessary changes when an image already belongs to the target domain, useful in preserving colour and structure of the translated image.

2.4. Model Training Configuration

The model was trained using the Adam optimiser (learning rate 0.0002, $\beta 1 = 0.5$), a batch size of 1, and 200 epochs with linear learning-rate decay after epoch 100. An image buffer of previously generated samples was used to improve discriminator stability, with checkpoints saved periodically for monitoring.

2.5. Model Performance Evaluation

The proposed model's generated nighttime pipeline images dataset was evaluated using FID and SSIM as quantitative metrics and VE as a qualitative metric, assessing both overall translation realism and the extent to which critical pipeline anomaly features were preserved.

3. Results and Discussion

3.1. Quantitative Evaluation

Table 1. Quantitative Results of Daytime-to-Nighttime Cross-Domain Pipeline I2I Translation Using CycleGAN

Metric	Value
FID	435.52
SSIM	0.13

The FID score of 435.52 indicates a considerable statistical distance between the generated nighttime images and the real nighttime distribution. Several factors likely contribute to this: the pipeline dataset contains highly repetitive industrial structures with limited scene diversity compared with benchmark I2I datasets such as horse-to-zebra; the Inception-V3 feature extractor used in FID computation is not well suited to pipeline inspection datasets; and the available nighttime images exhibit substantial illumination differences from daytime images, increasing domain discrepancy independent of translation quality. Because this research prioritises anomaly preservation over pure perceptual realism, optimising solely for FID was not the study aim.

The SSIM value of 0.13, while low relative to typical image-reconstruction benchmarks, is expected in an unpaired framework where no ground-truth nighttime pair exists for each daytime input. Since CycleGAN intentionally alters illumination, brightness, shadows, and colour to achieve domain adaptation, pixel-wise similarity is naturally reduced.

Despite high-level semantic content is preserved. SSIM should therefore be interpreted cautiously in unpaired pipeline translation tasks, as it does not directly reflect pipeline anomaly semantic preservation.

3.2. Qualitative Evaluation

Figure 2. The Visual Results of the Daytime-to-Nighttime Cross-Domain Pipeline I2I Translation Using CycleGAN

VE compared the generated nighttime pipeline images against real daytime pipeline images in Figure 2. Furthermore, despite the modest FID and SSIM scores, visual inspection shows that CycleGAN successfully transformed illumination characteristics to nighttime conditions while preserving pipeline structure, leak location, and surrounding environmental context.

No significant structural distortion or disappearance of leak features was observed, indicating that the model learned a domain mapping that alters illumination while retaining anomaly features critical to pipeline infrastructure inspection early warning systems.

3.3. Discussion

This study's aim extended beyond generating visually realistic nighttime images; it sought to determine whether CycleGAN could translate pipeline images from daytime to nighttime while preserving critical features of pipeline anomalies. The qualitative results suggest this aim was largely achieved, since FID and SSIM primarily capture global statistical and pixel-level correspondence rather than anomaly preservation. Similarly, in pipeline inspection early warning systems, preserving pipeline anomaly features is more important than photorealism. Furthermore, a visually realistic image that removes leak evidence is unsuitable for downstream detection.

However, an imperfectly realistic translated image that faithfully retains leak features remains valuable for optimal deep learning pipeline leak detection. The research findings indicate that CycleGAN can generate synthetic nighttime pipeline datasets that retain anomaly features necessary for automated leak-detection early warning systems. Despite the high FID and low SSIM values indicating considerable room to improve realism and structural consistency.

3. Conclusion

This research presents the first comprehensive analysis and taxonomy of GAN-based daytime-to-nighttime cross-domain I2I translation for pipeline infrastructure inspection early-warning systems, addressing the long-standing challenge of inadequate, inaccessible, and missing nighttime pipeline leak datasets. Similarly, the original CycleGAN framework, implemented without modification, was applied to a dataset that includes leak and non-leak daytime images and non-leak nighttime images to translate daytime pipeline images into nighttime images, tackling both dataset scarcity and domain-adaptation challenges relevant to deep learning pipeline leak-detection task.

Furthermore, the research findings show CycleGAN's potential to generate a synthetic nighttime pipeline image dataset that retains essential anomaly features needed for downstream pipeline leak detection. Although the FID score of 435.52 and SSIM of 0.13 indicate considerable room remains to improve realism and structural consistency. This research further analysed how translation quality affects downstream leak-detection performance, establishing a foundational benchmark and reviewing evaluation protocols relevant to safeguarding pipeline infrastructure whose failure enormously affects the GDP of oil-producing nations. The open challenges and future work include: difficulty in preserving the input image's semantic content of important pipeline infrastructure anomaly features after translation; no benchmark I2I evaluation metric tailored to the pipeline infrastructure domain to ensure pipeline leak features are preserved in the translated dataset; and also the translated nighttime pipeline leak images dataset needs to be applied to a deep learning leak-detection model.

Hence; to finally attest whether pipeline leak features are preserved based on the high-performance evaluation results obtained from the leak-detection model.

References

Agrawal, T., & Choudhary, P. (2023). Segmentation and classification on chest radiography: A systematic survey. *The Visual Computer*, *39*, 875–913.

Ali, M. A. H., Baggash, M., Rustamov, J., Abdulghafor, R., Abdo, N. A.-D. N., Abdo, M. H. G., Mohammed, T. S., Hasan, A. A., Abdo, A. N., Turaev, S., & Nukman, Y. (2023). An automatic visual inspection of oil tanks exterior surface using unmanned aerial vehicle with image processing and cascading fuzzy logic algorithms. *Drones*, *7*(2), 133. https://doi.org/10.3390/drones7020133

Binkowski, M., Sutherland, D., Arbel, M., & Gretton, A. (2018). Demystifying MMD GANs. *arXiv*. https://arxiv.org/abs/1801.01401

Binod, S., & Ashish, S. (2023). Petroleum dependencies in developing countries with an emphasis on Nepal and potential keys. *Energy Strategy Reviews*. https://doi.org/10.1016/j.esr.2023.101053

Eason, L. (2023). Comparative analysis of Pix2Pix and CycleGAN for image-to-image translation. *Highlights in Science, Engineering and Technology*, *39*.

Jabbar, M., Al-Battbootti, H., Goga, N., & Marin, I. (2022). Detection and analysis of oil spill using image processing. *International Journal of Advanced Computer Science and Applications*, *13*(4), 381–387. https://doi.org/10.14569/IJACSA.2022.0130445

Jiao, Z., Jia, G., & Cai, Y. (2019). A new approach to oil spill detection that combines deep learning with unmanned aerial vehicles. *Computers & Industrial Engineering, 135*, 1300–1311. https://doi.org/10.1016/j.cie.2018.11.008

Kerf, T. D., Gladines, J., S., S., & V., S. (2020). Oil spill detection using machine learning and infrared images. *Remote Sensing, 12*(24), 4090. https://doi.org/10.3390/rs12244090

Lee, J., Shiotsuka, D., Bang, G., Endo, Y., Nishimori, T., Nakao, K., & Kamijo, S. (2023). Day-to-night image translation via transfer learning to keep semantic information for driving simulator. *[Journal title not provided], 47*, 251–262.

Lee, Y., Go, Y., Lee, S., Son, D., & Lee, S. (2025). *Night-to-day image translation with road light attention training for traffic information detection*. [Publication details not provided].

Mehran, S., Behnaz, M., Hanieh, A., Nasrin, S., & Hamid, R. (2025). *Unpaired image-to-image translation with content preserving perspective: A review*. Preprint submitted to *Expert Systems with Applications*.

Muhammad, H. O., & Ali, H. H. (2023). Deep learning approach for oil pipeline leakage detection using image-based edge detection techniques. *Journal of Engineering and Applied Sciences, 56*(4), [page range not provided]. https://doi.org/10.18280/jesa.560416

Qian, W., Dong, S., Chen, L., & Ren, Q. (2024). Image enhancement method for low-light pipeline weld X-ray radiographs based on weakly supervised deep learning. *NDT & E International, 143*, 103049.

Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016). Improved techniques for training GANs. In *Advances in Neural Information Processing Systems* (Vol. 29).

Shatnawi, M., & B. M. Y. (2024). *An enhanced model for detecting and classifying emergency vehicles using a generative adversarial network (GAN)*. [Publication details not provided].

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–9).

Ukaegbu, U. F., Tartibu, L. K., Okwu, M. O., & Olayode, I. O. (2021). Integrating unmanned aerial vehicle and deep learning algorithm for pipeline monitoring and inspection in the oil and gas sector. *[Conference/journal title not provided]*, 1–6. https://doi.org/10.1145/3487923.3487924

Wang, Q. Y., Song, Y. H., Zhang, X. S., Dong, L. J., Xi, Y. C., Zeng, D. Z., Liu, Q. L., Zhang, H. L., Zhang, Z., Yan, R., & Luo, H. (2023). Evolution of corrosion prediction models for oil and gas pipelines: From empirical-driven to data-driven. *Engineering Failure Analysis, 146*, 107097.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing, 13*(4), 600–612.

Zhao, R., Han, Y., & Zhao, J. (2022). End-to-end Retinex-based illumination attention low-light enhancement network for autonomous driving at night. *[Journal title not provided]*. https://doi.org/10.1155/2022/4942420

Zhu, J., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 2223–2232).