Web-Based Cardiac Delayed Enhanced MRI Segmentation Using Deep Learning and Synthetic Data Augmentation

Omar N.[#1], Omar E.[#1], Khaled M.[#1], Hana H.[#1]

, Dr. Khaled A.[*2], Dr. Ahmed A..[#3], Dr. Tamer B.[#1]

[#]*Systems and Biomedical Engineering Department, Faculty of Engineering, Cairo University*
[1]omar.nabil.5386@gmail.com [2]omar.abdelbar02@gmail.com

[*]*Biomedical and Metabolic Imaging Branch, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), NIH,*
[5] khaled.z.abd.elmoniem@gmail.com

Abstract— **Accurate segmentation of myocardial infarction (MI) in cardiac magnetic resonance (CMR) images is crucial for diagnosis and treatment planning. We present a web-based platform that integrates a deep learning segmentation model directly into an interactive DICOM viewer, enabling clinicians to seamlessly upload studies, review, and refine infarction segmentations within a familiar clinical workflow. To address data scarcity, we leverage Generative Adversarial Networks (GANs) to generate realistic synthetic CMR images and employ a computational augmentation model to simulate diverse patterns of myocardial infarction, thus enhancing dataset diversity and model generalization. Our deep learning model, trained on both real and augmented data, delivers efficient, consistent, and high-quality detection of infarcted myocardium. By deploying this system within an accessible web-based DICOM viewer, we bridge the gap between AI research and clinical application, delivering rapid, reproducible, and interactive infarction analysis to support clinical decision-making and optimize patient care.**

I. INTRODUCTION

Cardiovascular disease remains the leading cause of global morbidity and mortality, driving the advancement of cardiac imaging for diagnosis, risk stratification, and intervention planning [1]. Cardiac Magnetic Resonance (CMR) imaging, particularly late gadolinium enhancement (LGE), is the clinical gold standard for non-invasive assessment of myocardial infarction, offering high spatial resolution and strong concordance with histopathology[3].

Precise segmentation of infarcted myocardium is critical for diagnosis and therapeutic planning, especially in procedures like ventricular tachycardia ablation, where individualized scar maps are essential [4],[5]. However, current practice is hindered by the challenges inherent in manual segmentation, including labor-intensive, reliant on expert input, and prone to inter- and intra-observer variability [2]. As cardiac imaging volumes increase and clinical complexity grows, there is an urgent need for automated, accurate, and accessible segmentation solutions to standardize and accelerate clinical workflows.

To address these limitations, we propose an end-to-end, data-driven framework as displayed in Fig.1 for automated myocardial infarction segmentation in LGE CMR images, integrating simulation-based data augmentation, generative adversarial networks (GANs), and a clinically accessible web-based platform. By synthesizing realistic and diverse training data and continuously optimizing segmentation performance, our system aims to deliver robust, reproducible, and widely adoptable solutions for myocardial infarction

delineation, supporting standardized clinical decision-making and expanding access to advanced cardiac imaging analytics.
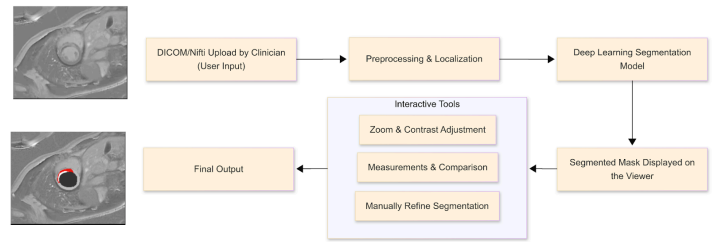

Fig.1 - Framework Diagram

II. BACKGROUND

As the clinical and technical importance of LGE CMR imaging continues to grow, so do the challenges in achieving robust, automated segmentation. This section outlines the technical principles of LGE, limitations in current segmentation efforts, and recent algorithmic advancements.

2.1 Late Gadolinium Enhancement in CMR Imaging
LGE is a CMR technique for detecting myocardial infarction and fibrosis by highlighting gadolinium accumulation in damaged tissue, which appears hyperintense on T1-weighted images, while viable myocardium is suppressed via inversion recovery [6]. Typically acquired 10–20 minutes post-contrast, and it requires accurate inversion time (TI) selection to avoid image artifacts [3], [6]

2.2 Limitations of LGE Data for Machine Learning
Despite its clinical advantages, LGE data presents several significant challenges for algorithmic analysis:
- **Annotation Scarcity**: Manual infarct segmentation is labor-intensive, requires expert input, and yields small, institution-specific datasets [2].
- **Protocol Variability:** Differences in acquisition parameters such as inversion time (TI), spatial resolution, and pulse sequences across institutions introduce heterogeneity that hinders model generalization [7].
- **Operator Dependence:** LGE quality and contrast are strongly affected by operator-controlled factors such as scan timing and reconstruction techniques [6].
- **Anatomical and Pathological Diversity:** Infarctions vary widely in location, transmurality, and shape, complicating both manual labeling and automated modeling.
- **Lack of Standardized Annotation Protocols:** The absence of unified annotation standards leads to inconsistency across datasets and clinical workflows [8].

- **Class Imbalance and Atypical Appearance:** Infarctions often constitute minority regions and may appear faint, irregular, or artifactually bright. Especially with suboptimal nulling-reducing algorithmic sensitivity [6].

These factors collectively hinder the development of generalizable segmentation models, underscoring the need for advanced data augmentation and standardized acquisition protocols.

III. METHODS

We developed an integrated pipeline displayed in Fig.2, combining data simulation, GAN-based augmentation, deep learning segmentation, and a web-based clinical interface for robust infarction segmentation from LGE CMR images.

Firstly, we used the EMIDEC dataset [9] to ensure clinical relevance. We generated synthetic infarction masks using the mathematical simulator described in Section **3.2**. A conditional Generative Adversarial Network (cGAN) [10] was then trained on real CMR images and corresponding masks to learn mask-to-image translation. This enabled controlled image generation. This model generated additional realistic CMR images from our simulated masks, effectively performing pixel-wise augmentation and expanding the labeled dataset. These augmented data were used to train a modified attention-residual U-Net (ResAtt-U-Net) [11], tailored for infarction detection, as detailed in Section 3.4.

Finally, the trained segmentation model was integrated into a web-based, DICOM-compatible viewer, as detailed in Section **3.5**. This platform allows clinicians to load cardiac MR studies, visualize segmentation overlays, interactively refine model outputs, and export results, all within their workflow..
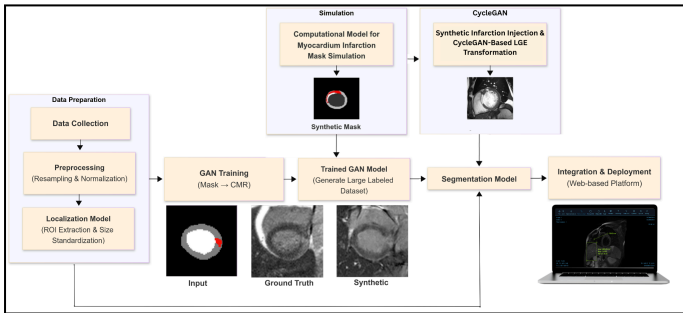

Fig.2 - Integrated Pipeline Diagram

3.1 Dataset
3.1.A Overview:

In this work, we employed the publicly available EMIDEC dataset, introduced by Lalande et al, which includes 100 training volumes in NIfTI format: 33 healthy controls and 67 patients with confirmed myocardial infarction (with some cases exhibiting microvascular obstruction/no-reflow) [9]. Each volume comprises a stack of short-axis slices covering the left ventricle, with expert-annotated voxel-wise masks for five semantic classes: background (0), cavity (1), healthy myocardium (2), infarct (3), and no-reflow (4)

3.1.B Spatial Normalization:
All images were resampled to a uniform voxel spacing of

$1 \times 1 \times 1$ mm³ (cubic interpolation for images; nearest-neighbor for masks).

3.1.C Slice Extraction & Naming:
3D volumes were converted to 2D short-axis slices and saved as **.npy** files using a consistent naming convention: **EMIDEC_Case_<N/P><ID>_slice_<#>_<Inf/NoInf>_<NoReflowY/N>.**
This encoded metadata supports easy filtering and stratification.

3.1.D Localization and ROI Extraction
To enhance segmentation efficiency, we implemented a modified nnU-Net[12] localization pipeline that extracts a fixed 128×128 region of interest (ROI) centered on the myocardium displayed in Fig.3. As CMR images cover large anatomical areas, limiting the model input to a consistent LV-centered ROI reduces complexity and improves focus.
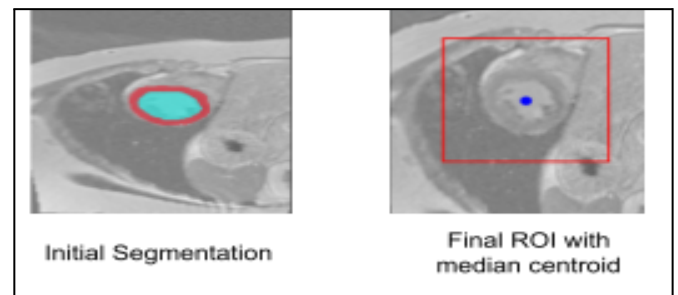

Fig.3 - Localization Sample

3.2 Mathematical Simulator
We designed a mathematical simulator to generate diverse, realistic myocardial infarction masks for data augmentation. The simulator mimics myocardial structures and infarct morphologies, enhancing segmentation robustness by expanding training variability. It employs three main strategies:

3.2.A Fully Synthetic Masks:
Myocardial masks are synthetically generated based on physiological diameter distributions and myocardium-to-blood-pool ratios. Infarct and no-reflow regions are simulated by placing seed points along the myocardium and expanding them via region-growing algorithms, allowing control over size, location, and heterogeneity.

3.2.B Hybrid Injection:
To avoid unrealistic synthetic shapes, infarct regions generated as above are injected into real myocardium masks from annotated datasets, preserving anatomical fidelity while maintaining infarct diversity.

3.2.C Cross-Subject Fusion:
Infarct regions from donor subjects are affinely transformed and merged into recipient myocardium masks selected via k-nearest neighbor matching based on diameter and shape descriptors. A cost function ensures anatomical realism and avoids infeasible overlaps, resulting in novel yet realistic infarction patterns.

3.2.D Deployment:
The simulator is packaged as a Docker image for reproducible,

batch, or expert-driven generation, supporting seamless integration into machine learning pipelines.

3.3 GAN Models

3.3.A Baseline Implementation and Results:

We began with a standard Pix2Pix setup, which consists of a U-Net generator and a PatchGAN discriminator as introduced by Isola et al. [12]. However, initial outputs were visually unrealistic and scored poorly across multiple quality metrics.: FID [13] > 500, low IS [14], PSNR [15], and LPIPS [16]. These results prompted architectural tuning and loss

3.3.B Addressing Class Imbalance

Our segmentation masks showed severe imbalance (e.g., background: 77%, no-reflow: <1%) shown in Figure.4 .
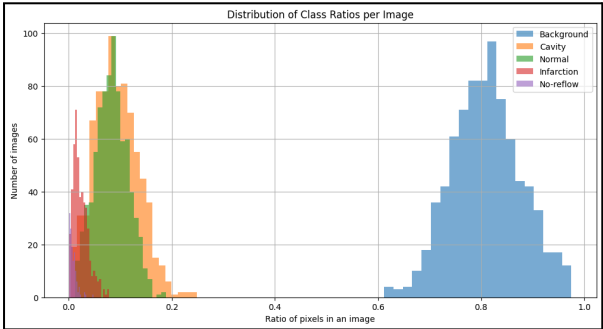

Fig.4 - Class Imbalance

To improve focus on clinically significant regions, we introduced a weighted L1 loss[12]. This shifted model attention to underrepresented regions, improving anatomical accuracy in generated images. Which resulted in the following **generator loss function**:

$$\mathcal{L}_{\text{gen}} = \mathcal{L}_{\text{GAN}} + \lambda \cdot \mathcal{L}_{\text{weighted L1}}$$

3.3.C Architectural & Hyperparameter Tuning

We iteratively modified key components as depicted in **Table I**:

- Generator: varied depth (6/5 layers), dropout, activation (tanh/sigmoid)
- Discriminator: shallow/deep PatchGANs, input variations
- Training: tuned λ (GAN loss weight), learning rate, norm layers, added structured noise to background

All models were evaluated with FID, IS, LPIPS, and PSNR.

TABLE I
SUMMARY OF PIX2PIX MODEL TRIALS AND KEY CONFIGURATIONS

Trial	Key Changes	FID ↓	IS ↑	LPIPS ↓	PSNR ↑
1	Baseline Pix2Pix	512.4	1.32	0.682	10.3
2	Weighted L1 (λ=60), 7/6 layers	423.9	1.75	0.598	11.4
3	λ=90, instance norm, no mask in D	348.4	1.68	0.581	15.3
4	λ=100, mask-only D	224.2	1.83	0.566	15.5
5	**Final: λ=100, 6/5 layers, weighted L1, PatchGAN**	**198.15**	**2.192**	**0.5382**	**16.42**

Final model used a U-Net (6/5 layers), weighted L1 loss (λ=100), tanh activation, and a moderately deep PatchGAN. It achieved the best performance across all metrics.

3.3.D Comparison with Other Models

TABLE II
PERFORMANCE COMPARISON BETWEEN OUR MODEL AND EXISTING GAN-BASED METHODS ON IMAGE SYNTHESIS TASKS

Model	FID ↓	IS ↑	PSNR ↑
CycleGAN	262.6	1.112	**32.4**
Pix2Pix	244.4	1.714	30.5
U-Net++ Pix2Pix	201.1	1.747	
Ours	**198.15**	**2.192**	16.42

Adapted from Wang et al. (2023); Lee et al. (2023).

Our model outperformed others in realism and diversity, though PSNR remained lower due to DE-MRI complexity. In this context, perceptual and structural accuracy (not pixel-level similarity) is more critical.
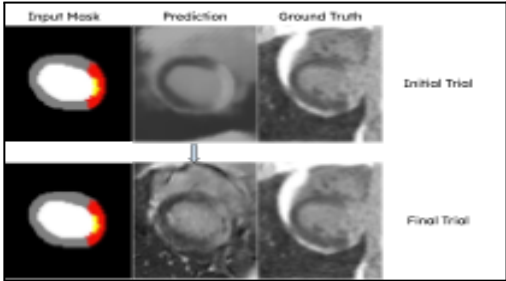

Fig.5 - Comparison of Baseline vs. Final Model Output

3.3.G Modified CycleGAN

To additionally address the scarcity of LGE CMR images highlighting infarction and no-reflow, we designed a modified CycleGAN [17] to translate cine CMR images (Module A) into synthetic LGE-like images (Module B). Cine images are plentiful but lack pathology; LGE images show infarcts but are limited.

Synthetic infarction and no-reflow masks were embedded into cine images via the Mathematical Simulator previously described in 3.2. Our modified CycleGAN then performed domain translation (A→B→A and B→A→B) displayed in Fig.6, guided by anatomical masks. A **mask-weighted cycle-consistency loss** was applied, focusing translation on relevant myocardial regions:

$$\mathcal{L}_{\text{cycle-mask}} = \|M \odot (G_{B \to A}(G_{A \to B}(x)) - x)\|_1$$
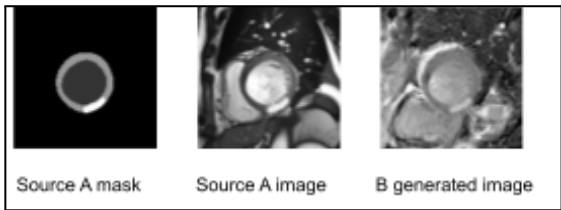

Fig.6 - CycleGAN workflow

The resulting synthetic LGE dataset provides high-quality labeled images for training infarction segmentation models.

3.4 Segmentation Model
3.4.A Attention Residual U-Net
We utilized an Attention Residual U-Net [18] for the semantic segmentation of cardiac DE-MRI. The model combines residual blocks to improve gradient flow and attention gates to enhance skip connections, focusing on class-relevant regions like infarction and no-reflow zones. It features five downsampling and upsampling stages, each with dual convolutional layers, batch normalization, and ReLU[19] activation.

3.4.B Loss Function: Class-Weighted Sparse Categorical Cross-Entropy
To mitigate the severe class imbalance observed in the dataset Figure.4, we adapted the loss function to include class-specific weights consistent with the weights in the GAN module.
This weighting scheme emphasizes clinically important but underrepresented classes, guiding the model to better distinguish infarcted and no-reflow regions.

3.4.C Evaluating GAN-Augmented Training

To quantify the effect of **synthetic data** from the GAN module, we designed a comparative experiment using two training datasets:

1. **Augmented Real Dataset (ARD)**: 10,000 slices generated by classical data augmentation from real cases.
2. **Mixed Dataset (MD)**: 5,000 real augmented slices + 5,000 GAN-generated slices.

Both datasets were class-balanced and trained using identical hyperparameters for the same amount of time. The model received 2D DE-MRI slices as input and produced five-channel segmentation maps. This setup isolates the contribution of synthetic data to segmentation performance, with results presented in **3.4** under the Results section.

3.5 Web Application and DICOM Viewer
To enable clinician interaction with AI-generated segmentations, we developed a web-based platform that combines AI-powered myocardial segmentation with an interactive DICOM viewer. It supports real-time visualization, manual refinement, and quantification of infarcted regions.

3.5.A DICOM Visualization and Interaction
The viewer supports full DICOM study uploads with interactive thumbnail navigation and multi-series display.
3.5.B Interactive Visualization and Measurement Tools
A comprehensive set of tools supports image manipulation, including real-time panning, zooming, and window-level adjustments. Measurement utilities allow for linear distance, angular assessment, and region-of-interest (ROI) analysis using rectangular and elliptical tools.

3.5.C Brush-Based Editing Interface
A brush-based tool is integrated into the platform to facilitate manual annotation and correction of segmentation outputs with pixel-level precision. Clinicians can save their edits in JSON format to re-upload them anytime.

3.5.D Deep Learning Model Integration
The platform incorporates the two U-Net-based models

described in sections 3.4 and 3.1.D to support infarction assessment. The resulting masks are rendered as overlays on the original DICOM slices.

3.5.E Manual Segmentation Refinement
Segmentation masks are internally converted into editable brush-stroke representations, allowing direct refinement using the brush tool while preserving spatial alignment.

IV. RESULTS

4.1 Segmentation Model
4.1.A Quantitative Results
We evaluated segmentation performance using the **Dice Similarity Coefficient (DSC)[20]**, measuring overlap between predicted and ground truth masks.

TABLE IV
DICE SCORE COMPARISON

Class	Description	Real Only	Real + GAN
1	LV Cavity	0.9350	0.9365
2	Normal Myocardium	0.8222	0.8319
3	Infarcted Myocardium	0.6126	0.6651
4	No-Reflow Area	0.5376	0.6676
—	Mean Dice Score	0.7268	0.7752

4.1.B Qualitative Comparison and Interpretation
TABLE IV illustrates a visual comparison of segmentation outputs from both datasets on one representative case. The **Mixed Dataset** model shows more consistent segmentation of no-reflow and infarcted zones, with fewer false positives and better boundary adherence.
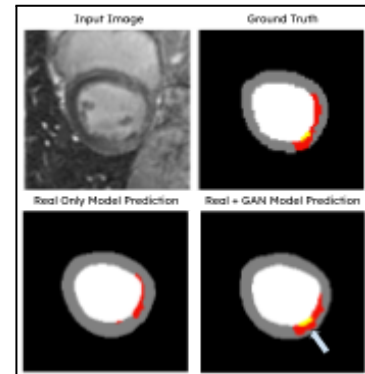

Fig.7 - Qualitative comparison of segmentation output from (A) Real-Only Dataset and (B) Mixed Dataset (Real + GAN). Arrow indicates improved delineation of small pathological zones.

These results confirm that incorporating synthetic data introduces beneficial variability, improving generalization on rare classes. Although performance on common classes remains similar, **the segmentation of underrepresented regions, like no-reflow, is significantly enhanced**, supporting the value of GAN-based augmentation in highly imbalanced medical image datasets.

References

[1] R. J. Kim *et al.*, "The use of contrast-enhanced magnetic resonance imaging to identify reversible myocardial dysfunction," *The New England Journal of Medicine*, vol. 343, no. 20, pp. 1445–1453, Nov. 2000, doi: 10.1056/NEJM200011163432003.

[2] R. Karim *et al.*, "Evaluation of state-of-the-art segmentation algorithms for left ventricle infarct from late Gadolinium enhancement MR images," *Medical Image Analysis*, vol. 30, pp. 95–107, May 2016, doi: 10.1016/j.media.2016.01.004.

[3] A. Wagner *et al.*, "Contrast-enhanced MRI and histopathology in acute and chronic myocardial infarction," *Circulation*, vol. 108, no. 11, pp. 1253–1258, Sep. 2003, doi: 10.1161/01.CIR.0000089082.60725.8E.

[4] H. L. Estner *et al.*, "Electroanatomic mapping and ablation in ventricular tachycardia: the role of MRI," *European Heart Journal*, vol. 32, no. 5, pp. 529–535, Feb. 2011, doi: 10.1093/eurheartj/ehq383.

[5] D. Andreu *et al.*, "Integration of late gadolinium enhancement–MRI scar maps into electroanatomic mapping systems: feasibility and impact on VT ablation," *Heart Rhythm*, vol. 8, no. 4, pp. 576–583, Apr. 2011, doi: 10.1016/j.hrthm.2010.11.027.

[6] V. M. Campello *et al.*, "Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge," *IEEE Transactions on Medical Imaging*, vol. 40, no. 12, pp. 3543–3554, Dec. 2021, doi: 10.1109/TMI.2021.3090082.

[7] P. Kellman and A. E. Arai, "Cardiac imaging techniques for physicians: Late enhancement," *Journal of Magnetic Resonance Imaging*, vol. 36, no. 3, pp. 529–542, Sep. 2012, doi: 10.1002/jmri.23699.

[8] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, pp. 203–211, 2021, doi: 10.1038/s41592-020-01008-z.

[9] P. Lalande, M.-P. Cervenansky, et al., "Standardized dataset for evaluation of myocardial infarction segmentation in late-enhancement cardiac MRI of the left ventricle (EMIDEC)," *Medical Image Analysis*, vol. 58, p. 101574, 2019, doi: 10.1016/j.media.2019.101574.

[10] M. Mirza and S. Osindero, "Conditional Generative Adversarial Nets," *arXiv preprint* arXiv:1411.1784, 2014. [Online]. Available: https://arxiv.org/abs/1411.1784

[11] Z. Li, H. Zhang, Z. Li, and Z. Ren, "Residual-Attention U-Net++: A Nested Residual-Attention U-Net for Medical Image Segmentation," *Applied Sciences*, vol. 12, no. 14, 2022, doi: 10.3390/app12147241.

[12] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 5967–5976, doi: 10.1109/CVPR.2017.632.

[13] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2017. [Online]. Available: https://arxiv.org/abs/1706.08500

[14] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen, "Improved Techniques for Training GANs," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2016. [Online]. Available: https://arxiv.org/abs/1606.03498

[15] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[16] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 586–595. [Online]. Available: https://arxiv.org/abs/1801.03924

[17] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 2223–2232. [Online]. Available: https://arxiv.org/abs/1703.10593

[18] O. Oktay *et al.*, "Attention U-Net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018. [Online]. Available: https://arxiv.org/abs/1804.03999

[19] V. Nair and G. E. Hinton, "Rectified Linear Units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, 2010, pp. 807–814.

[20] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, Jul. 1945.