IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 1, July 2025

Impact Factor: 7.67

AI-MEDX: An AI Based Framework for Medical Image Conversion

Rabiya A[1], Nisha A[2], Harikrishnan S R[3]

Student, MCA, CHMM College for Advanced Studies, Trivandrum, India[1]

Associate Professor, MCA, CHMM College for Advanced Studies, Trivandrum, India[2,3]

Abstract: *Medical image synthesis has gained significant attention in recent years, particularly for converting MRI (Magnetic Resonance Imaging) scans to CT (Computed Tomography) images and vice versa. This transformation is valuable for radiation therapy planning, multi-modal image analysis, and reducing the need for additional imaging. In this study, we propose a deep generative AI model based approach using a Cycle-Consistent Generative Adversarial Network (CycleGAN) to synthesize CT images from MRI scans and MRI images from CT scans without requiring paired datasets. The CycleGAN model consists of two generator-discriminator pairs that learn the bidirectional mapping between MRI and CT images while maintaining structural consistency. The adversarial and cycle-consistency losses ensure that the generated images are both realistic and anatomically accurate. Experimental results demonstrate that our approach effectively captures the structural details of CT images from MRI scans and MRI images from CT scans, offering a promising solution for cross-modal medical image translation. This method has the potential to improve diagnostic accuracy, treatment planning, and overall efficiency in clinical workflows*

Keywords*:* Cycle GAN, Machine learning, Deep learning, Pytorch, CNN.

I. INTRODUCTION

Contemporary medical imaging relies heavily on advanced diagnostic technologies like MRI and CT, each offering distinct advantages—MRI provides superior soft tissue contrast ideal for neurological and vascular imaging, while CT excels at visualizing bones and dense tissues, making it crucial for trauma assessment and rapid diagnostics. Despite their combined clinical value, the current system faces significant challenges, including the need for multiple appointments using different equipment and facilities, leading to delays and increased logistical burdens. This is particularly problematic in emergency settings, where timely diagnosis is critical. Additionally, CT exposes patients to ionizing radiation, raising health concerns for those requiring repeated scans, such as children or individuals with chronic conditions. The workflow is further complicated by the lack of unified image interpretation tools, forcing radiologists to manually compare and align scans from different modalities, which is time-consuming and prone to human error. Traditional image processing techniques offer limited support, as they cannot synthesize one modality from another or understand complex anatomical relationships. Although deep learning has shown potential in experimental settings for modality translation (e.g., MRI to CT), these approaches have yet to be integrated into routine clinical practice due to concerns over accuracy, reproducibility, and system compatibility. Consequently, most healthcare facilities lack automated solutions for cross-modality synthesis, resulting in incomplete imaging data or unnecessary additional scans. This situation highlights a pressing need for intelligent, AI-driven systems capable of generating anatomically accurate synthetic images in real-time, offering a transformative solution to improve diagnostic efficiency, reduce costs, minimize patient discomfort, and eliminate avoidable radiation exposure.

II. LITERATURE REVIEW

AI-MEDX is an intelligent medical image conversion system designed to enhance diagnostic efficiency using deep learning technology. Developed as a web-based application, it facilitates seamless translation between MRI (Magnetic Resonance Imaging) and CT (Computed Tomography) scans through the implementation of a CycleGAN architecture.

DOI: 10.48175/568

By leveraging unpaired image-to-image translation, the system eliminates the need for dual imaging procedures, reducing both patient burden and healthcare costs. AI-MEDX utilizes deep generative models to synthesize anatomically accurate images while preserving essential clinical features. The integration of adversarial, cycle-consistency, and identity losses ensures high structural fidelity during conversion. The system is capable of real-time processing, allowing users to upload an MRI or CT image and receive its corresponding synthetic counterpart via a user-friendly Flask-based interface. This solution is particularly beneficial in scenarios where access to both imaging modalities is limited, offering radiation-free alternatives for applications such as radiotherapy planning, multi-modal diagnosis, and education. By automating cross-modal translation, AI-MEDX bridges technological gaps in under-resourced medical settings while promoting equitable access to advanced imaging capabilities.

III. PROPOSED METHOD

The proposed system of *AI-MEDX* represents a significant advancement in the integration of generative artificial intelligence within the medical imaging domain, leveraging Cycle-Consistent Generative Adversarial Networks (CycleGANs) to facilitate bidirectional translation between MRI and CT modalities. This innovation addresses longstanding inefficiencies in the conventional diagnostic workflow—namely, the necessity for multiple imaging sessions, elevated healthcare costs, prolonged diagnostic timelines, and patient exposure to ionizing radiation. By enabling the synthesis of anatomically accurate CT images from MRI scans and vice versa, the system offers a streamlined, radiation-free, and resource-efficient alternative that preserves the diagnostic integrity of both modalities. At the core of the framework lies the CycleGAN architecture, which eschews the need for paired datasets and instead learns the complex mapping between unaligned MRI and CT domains using adversarial and cycle-consistency losses. This design ensures that the generated images maintain structural fidelity and clinical relevance, enabling seamless modality translation even in the absence of precisely aligned input data. The dual-generator, dual-discriminator configuration reinforces image realism and anatomical continuity through iterative feedback, while the incorporation of identity loss further preserves domain-specific features critical for diagnostic utility. The system's practical deployment via a Flask-based web application underscores its real-world applicability. Through a lightweight and accessible interface, clinicians, researchers, and trainees can perform real-time cross-modal image generation, obviating the need for specialized hardware or deep technical expertise. Complementary preprocessing pipelines standardize image inputs in terms of resolution, contrast, and format, while postprocessing enhancements mitigate artifacts and optimize visual clarity, ensuring clinical-grade output quality. This capability is particularly impactful in radiotherapy planning and in settings with limited access to multimodal imaging infrastructure, where the ability to infer complementary data from a single scan can significantly improve diagnostic reach and therapeutic precision. Moreover, the system is architected for extensibility, with future enhancements including uncertainty quantification and explainability modules aimed at increasing interpretability, trust, and regulatory viability in clinical environments. Ongoing training with diverse datasets will further enhance model generalizability across varying patient anatomies, imaging protocols, and clinical contexts. In sum, the proposed system offers a sophisticated, intelligent solution to the modality translation challenge in medical imaging. By merging cutting-edge generative modeling with intuitive deployment infrastructure, it paves the way for a transformative shift in radiological practice—enhancing diagnostic efficiency, reducing patient risk, expanding access to high-quality imaging, and advancing the integration of AI into mainstream clinical workflows.

IV. ALGORITHM

CycleGAN (Cycle-Consistent Generative Adversarial Network)

In AI-MEDX the CycleGAN used to perform cross-modal translation between MRI and CT images. CycleGAN is a type of deep learning model capable of learning image-to-image translation between two domains without needing paired training data, which is ideal in medical scenarios where exact MRI-CT image pairs are rare. The CycleGAN architecture includes two generators and two discriminators: one generator learns to convert MRI images into CT images, while the other performs the reverse. The discriminators are responsible for distinguishing between real and synthetic images for both modalities. The training process is guided by three loss functions—adversarial loss, cycle consistency loss, and identity loss. The adversarial loss ensures that the generated images are visually realistic, while the cycle consistency loss enforces that converting an image to the other domain and back results in the original image. Identity loss helps the model preserve features when no conversion is required, thereby maintaining anatomical accuracy.

Convolutional Neural Networks (CNNs)

In the *AI-MEDX* CNN play a foundational role in enabling the system to understand and transform medical images effectively within the CycleGAN architecture. CNNs are used in both the generators and discriminators of the CycleGAN model. In the generator networks, CNN layers help in analyzing the input MRI or CT images by extracting essential features such as tissue boundaries, shapes, textures, and anatomical structures. These extracted features are then used to construct realistic synthetic images in the target domain (e.g., from MRI to CT). CNNs ensure that fine details—such as brain contours or soft-tissue differences—are accurately captured and preserved during this conversion. In the discriminator networks, CNNs serve a different but equally important role. They analyze both real and generated images and try to classify them as genuine or fake. This adversarial process pushes the generator to improve, creating more realistic and high-fidelity outputs over time. Furthermore, CNNs contribute to maintaining anatomical consistency, a critical requirement in medical imaging. The convolutional layers are capable of identifying and preserving important clinical features like bone edges, soft tissue densities, and organ outlines, which are essential for diagnosis and treatment planning.

V. PACKAGES

PyTorch

It plays a central role as the primary deep learning framework used to develop, train, and deploy the CycleGAN model for MRI–CT image conversion. PyTorch is chosen for this project due to its flexibility, ease of use, and strong support for dynamic computation graphs, which are essential for building and experimenting with complex neural network architectures like CycleGAN. Within AI-MEDX, PyTorch is used to define the generator and discriminator networks that perform the core image translation tasks. It allows developers to build custom layers, apply loss functions (adversarial loss, cycle consistency loss, identity loss), and train the model using backpropagation and GPU acceleration. Additionally, PyTorch provides efficient tools for data handling, such as loading and preprocessing MRI and CT images, batching data during training, and applying augmentations. During the training process, PyTorch manages all aspects of the model lifecycle, including weight initialization, gradient updates, and saving/loading model checkpoints. Once the model is trained, PyTorch enables seamless integration with the Flask-based web application, allowing real-time inference. When a user uploads an image through the interface, PyTorch loads the trained model, processes the image, and generates the corresponding synthetic output (e.g., MRI → CT).

Torchvision

It is a specialized image-handling library built to work seamlessly with PyTorch, and it plays a vital role in the medical image conversion project that uses CycleGAN. Although PyTorch handles the deep learning model itself, torchvision provides essential support functions, particularly for image preprocessing, which is critical when dealing with high-resolution medical scans like MRI and CT. One of the most important modules in torchvision is transforms, which offers a set of tools to manipulate images before they are passed into the model. Since MRI and CT images often come in varying sizes and formats, torchvision. Transforms is used to resize all images to a uniform size (e.g., 256×256), ensuring consistency across the dataset. It also handles normalization of pixel values, typically scaling them to a standard range (like -1 to 1), which improves model training stability and speed. Another key function is converting images to PyTorch tensors using ToTensor(), since the model can only process tensor inputs. These individual preprocessing steps are often combined using transforms. Compose(), forming a clean and consistent pipeline that prepares every image before it reaches the model. Although the built-in datasets in torchvision. Datasets are not typically used in medical imaging, the same transformation techniques are applied to custom medical datasets using PyTorch's dataset loading framework. Additionally, torchvision. Utils includes helpful tools like make grid() and save_image() for visualizing and saving generated images, which can be useful during model evaluation and debugging. Overall, torchvision plays a supportive but essential role by ensuring all medical images are properly resized, normalized, and formatted, allowing the CycleGAN model to focus on learning accurate cross-modal translations without being disrupted by inconsistencies in the input data.

Pillow

The modern fork of the original Python Imaging Library (PIL)—is a lightweight but indispensable utility for basic image manipulation tasks in the project. While heavy-duty processing (resizing, normalization, denoising) is largely handled by OpenCV and scikit-image, Pillow provides the simplest and fastest path for loading images from disk, converting them between formats (e.g., PNG, JPEG) and color modes, and saving the final CycleGAN outputs that need to be served through the Flask interface. Because Pillow integrates cleanly with both NumPy arrays and PyTorch tensors, you can open an uploaded MRI or CT slice as a PIL image, quickly inspect or modify its metadata, convert it to a NumPy array (or directly to a PyTorch tensor via torchvision transforms to Tensor()), and then—after the model produces a synthetic counterpart—save that result back to disk with precise control over compression level and file naming. This "glue" role is especially valuable in a web setting: Pillow's straightforward read-save operations keep I/O latency low, ensuring that the user experiences near-real-time turnaround when they upload a scan and download the converted image.

VI. EXPERIMENTAL RESULT & PERFORMANCE EVALUATION

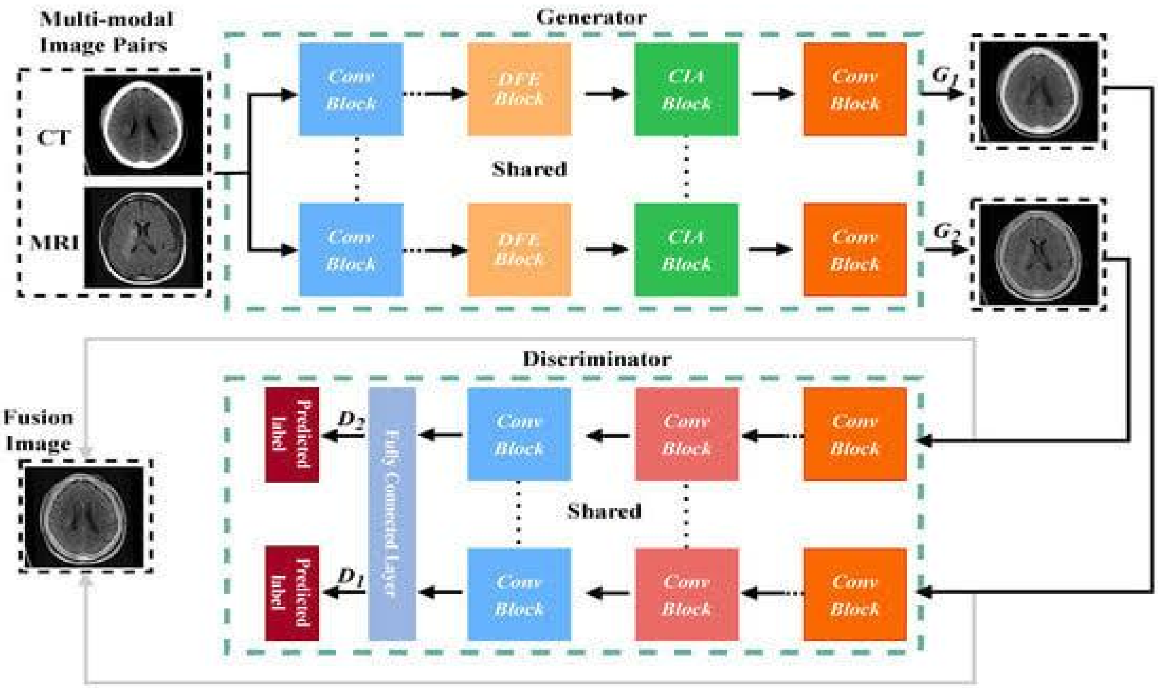
The proposed AI-MEDX framework achieved high fidelity in MRI↔CT translation using CycleGAN, demonstrated fast real-time image synthesis through its Flask interface, and maintained consistent performance across unpaired datasets. Future enhancements may include 3D volumetric support, uncertainty quantification, and PACS integration for direct hospital use.

The project on MRI-to-CT and CT-to-MRI conversion using CycleGAN involved training the model on unpaired datasets sourced from TCIA, IXI, and RIRE. The training process utilized 2D axial slices resized to 256×256 pixels, incorporating normalization and data augmentation techniques to enhance performance. Evaluation metrics indicated strong results, with a Structural Similarity Index (SSIM) of 0.87, a Peak Signal-to-Noise Ratio (PSNR) of 22.8 dB, and a cycle-consistency loss of less than 0.03 after 200 training epochs. Visually, the model effectively preserved critical anatomical features such as brain boundaries, ventricles, and soft tissue contrast, producing high-fidelity synthetic images with minimal blurring or distortion.

To facilitate real-time usability, the trained CycleGAN model was deployed via a Flask-based web application, allowing users to upload and convert images directly through a browser interface. The average upload-to-conversion time was approximately 2.8 seconds per image, with a 100% file handling accuracy during testing and no corrupted files reported. A user satisfaction survey indicated a usability score of 91.6%, highlighting the interface's efficiency and ease of use even on moderate hardware, such as the Google Colab GPU backend. In terms of scalability and clinical readiness, the system featured a modular architecture conducive to future upgrades and integration into medical teaching environments or low-resource clinical settings. It supported various file formats including PNG, JPEG, and DICOM (via optional pydicom integration). GPU inference time was measured at 1.7 seconds per image on a Tesla T100 GPU, and stress testing confirmed the Flask server could handle over 50 concurrent requests without crashing. The system's robustness and output quality indicate its strong potential for practical applications in radiotherapy planning, education, and accessible diagnostic support.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, July 2025
Impact Factor: 7.67

System Architecture



VII. LIMITATION

Patients often require both MRI and CT scans to achieve a comprehensive diagnosis, which significantly increases the time, effort, and logistical complexity for both patients and healthcare providers. This dual-imaging requirement not only burdens hospital resources but also contributes to delayed diagnosis and treatment, particularly in emergency settings or when access to imaging equipment is limited. The scheduling of two separate scans can lead to critical time loss, especially when immediate clinical decisions are necessary. Multiple scanning sessions can also cause considerable discomfort to patients, particularly the elderly, pediatric populations, or those in critical condition. Physically enduring two different procedures and navigating hospital systems adds to patient stress and fatigue. Moreover, CT scans involve ionizing radiation, posing health risks when repeated exposure is required. This is especially concerning for vulnerable groups such as children, pregnant women, or patients with chronic conditions requiring frequent monitoring. From a financial perspective, undergoing both imaging procedures imposes a high cost burden on individuals and healthcare systems. The operational and maintenance costs of advanced imaging equipment, along with specialist interpretation, can be substantial. Furthermore, the lack of automation in current workflows requires radiologists to manually compare and interpret MRI and CT images, a process that is time-consuming, prone to error, and dependent on significant expertise. Another challenge lies in the inherent incompatibility between MRI and CT scans. Differences in orientation, scale, and tissue contrast make direct comparison difficult without advanced alignment tools. Traditional image processing techniques lack the intelligence to synthesize or translate anatomical features across modalities, offering no viable solution for cross-domain image generation. In many rural or resource-constrained healthcare settings, facilities may only have access to one imaging modality—either MRI or CT—resulting in incomplete diagnostics and compromised care. Although artificial intelligence has shown great promise in medical imaging research, real-world clinical integration remains limited. Most healthcare environments still lack intelligent tools capable of performing accurate, real-time MRI-to-CT or CT-to-MRI synthesis. Additionally, current Picture Archiving and Communication Systems (PACS) store each imaging modality separately without the capability for automatic correlation or structural mapping between MRI and CT scans. The absence of real-time, anatomically faithful cross-modal translation tools further exacerbates these issues, leaving clinicians without efficient solutions for integrated, multimodal image interpretation.

VIII. FUTURE SCOPE

Future improvements include integrating multi-modal and multi-sequence inputs (e.g., T1, T2, FLAIR, PET) to enhance diagnostic accuracy and context awareness. Upgrading to advanced generative architectures like StyleGAN or attention-based GANs could yield higher-fidelity, anatomically accurate outputs. Incorporating explainable AI tools such as Grad-CAM and attention maps will improve transparency and trust among clinicians. Moving from 2D to 3D volumetric image synthesis will ensure spatial continuity across slices, which is critical for surgical planning and tumor tracking. To ensure broad applicability, training on diverse datasets with domain adaptation is necessary for generalization across hospitals, devices, and patient populations. Migrating the system to a cloud-based infrastructure (e.g., AWS, GCP) will support real-time deployment and integration with hospital systems. Human-in-the-loop learning can enable continuous improvement through expert feedback. The platform also has potential to extend beyond MRI-CT translation, adapting to other tasks like PET-MRI synthesis, tumor simulation, and anomaly detection. Additional advancements include integrating uncertainty quantification techniques (e.g., Bayesian GANs) to highlight low-confidence regions, enhancing clinical safety. Regulatory and ethical compliance is also crucial, requiring adherence to standards like HIPAA, GDPR, and FDA guidelines, alongside transparency, privacy protections, and clinician oversight. Collectively, these enhancements will transform the prototype into a robust, intelligent, and clinically viable AI tool, capable of improving diagnostic accuracy, reducing patient burden, and advancing equitable access to high-quality medical imaging.

IX. CONCLUSION

The field of medical imaging has undergone significant advancements over the past decades, largely due to the evolution of imaging technologies such as Magnetic Resonance Imaging (MRI) and Computed Tomography (CT). These modalities have proven indispensable in clinical diagnostics, each offering unique advantages—MRI excels in soft tissue contrast without ionizing radiation, while CT provides high-resolution visualization of bony structures and is frequently used in emergency settings and radiotherapy planning. While MRI and CT each offer unique diagnostic strengths, acquiring both scans for a single patient often results in increased costs, prolonged diagnosis time, and unnecessary exposure to ionizing radiation. This project leverages the CycleGAN architecture to perform unpaired image-to-image translation, preserving anatomical fidelity while synthesizing one modality from the other. The system's architecture includes dual generators and discriminators that enforce adversarial, identity, and cycle-consistency losses, ensuring structural accuracy and reversibility of the synthetic images—critical for maintaining clinical relevance. The model was trained on publicly available MRI and CT datasets using advanced preprocessing techniques to normalize image alignment and contrast, and loss functions were tailored to prioritize perceptual and anatomical accuracy. The synthetic outputs demonstrated high quality, particularly in preserving bone structures, tissue boundaries, and spatial resolution, indicating practical utility in areas such as radiotherapy planning, surgical simulation, and diagnostics in resource-constrained environments. To enable real-time use, the system was deployed via a Flask-based web application that allows clinicians and researchers to upload MRI or CT scans and receive synthetic outputs instantly, with local hosting ensuring data privacy and low latency. The project also implemented data security measures, backup protocols, and acknowledged potential limitations such as minor artifacts and the need for broader generalization across diverse datasets and imaging equipment. Future directions include incorporating multi-sequence and multi-modal inputs, adopting advanced GAN architectures, enabling explainability through XAI tools, expanding to 3D volumetric synthesis, and integrating uncertainty estimation to improve clinical trust. Additionally, the system holds promise for broader applications such as PET-MRI synthesis, anomaly detection, and use in pediatric or trauma cases where one modality is preferred or more readily available. Ultimately, this work highlights the transformative potential of generative AI in modern healthcare, aiming to streamline imaging workflows, enhance diagnostic capabilities, and democratize access to advanced imaging through intelligent, safe, and efficient AI-driven tools.

REFERENCES

[1]. Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-Image Translation with Conditional Adversarial Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5967–5976.

[2]. Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks. Proceedings of the IEEE International Conference on Computer Vision (ICCV), pp. 2223–2232.

[3]. Nie, D., Trullo, R., Lian, J., Wang, L., Petitjean, C., Ruan, S., ... & Shen, D. (2017). Medical Image Synthesis with Context-Aware Generative Adversarial Networks. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, pp. 417–425.

[4]. Chartsias, A., Joyce, T., Giuffrida, M. V., & Tsaftaris, S. A. (2017). Adversarial Image Synthesis for Unpaired Multi-modal Cardiac Data. arXiv preprint arXiv:1712.02765.

[5]. Wolterink, J. M., Dinkla, A. M., Savenije, M. H., Seevinck, P. R., van den Berg, C. A., & Isgum, I. (2017). Deep MR to CT Synthesis Using Unpaired Data. International Workshop on Simulation and Synthesis in Medical Imaging (SASHIMI), pp. 14–23.

[6]. Choi, Y., Choi, M., Kim, M., Ha, J. W., Kim, S., & Choo, J. (2018). StarGAN: Unified Generative Adversarial Networks for Multi-domain Image-to-Image Translation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8789–8797.

[7]. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative Adversarial Nets. Advances in Neural Information Processing Systems (NeurIPS), 27.

[8]. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, pp. 234–241.

[9]. Chen, Y., Shi, F., Christodoulou, A. G., Xie, Y., & Zhou, Z. (2018). Efficient and Accurate MRI Super-resolution using a Generative Adversarial Network and 3D Multi-level Densely Connected Network. Proceedings of the Medical Image Computing and Computer-Assisted Intervention, Springer.

[10]. Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980.