Advancing Healthcare: Innovative ML Approaches for Improved Medical Imaging in Data-Constrained Environments

Al Amin[1], Kamrul Hasan[1], Saleh Zein-Sabatto[1], Liang Hong[1], Sachin Shetty[2], Imtiaz Ahmed[3], Tariqul Islam[4]

[1] Tennessee State University, Nashville, TN, USA
[2] Old Dominion University, Norfolk, VA, USA
[3] Howard University, Washington, DC, USA
[4] Syracuse University, Syracuse, NY, USA

Email: {aamin2, mhasan1, mzein, lhong}@tnstate.edu, {sshetty}@odu.edu, {imtiaz.ahmed}@howard.edu, {mtislam}@syr.edu

Abstract—Healthcare industries face challenges when experiencing rare diseases due to limited samples. Artificial Intelligence (AI) communities overcome this situation to create synthetic data which is an ethical and privacy issue in the medical domain. This research introduces the CAT-U-Net framework as a new approach to overcome these limitations, which enhances feature extraction from medical images without the need for large datasets. The proposed framework adds an extra concatenation layer with downsampling parts, thereby improving its ability to learn from limited data while maintaining patient privacy. To validate, the proposed framework's robustness, different medical conditioning datasets were utilized including COVID-19, brain tumors, and wrist fractures. The framework achieved nearly 98% reconstruction accuracy, with a Dice coefficient close to 0.946. The proposed CAT-U-Net has the potential to make a big difference in medical image diagnostics in settings with limited data.

Index Terms—CAT-U-Net, Limited Dataset, Medical Image Analysis, Rare Diseases.

I. INTRODUCTION

The ability to extract features from medical images strongly suits Deep Learning (DL) architectures [1]–[4]. However, DL required large samples to reduce false positive and false negative rates [5]–[9]. Facing the challenge of data scarcity, especially in rare disease diagnostics, researchers and industry experts are increasingly turning to synthetic data generation and data augmentation techniques to bridge the gap [10]–[12]. In real-world clinical settings, AI models can be fooled by low-quality synthetic data when it comes to identifying the location and kind of diseases [13], [14]. Moreover, synthetic data generation in healthcare must navigate complex ethical and legal landscapes [15], [16], ensuring that privacy is preserved and inherent biases are not perpetuated [17]. To overcome this limitation, the proposed framework reduces false positive and negative rates in limited dataset scenarios also improving the dice coefficient by over 94%. Figure 1 represents the proposed CAT-Unet framework pipeline as an exemplary embodiment of innovation in medical imaging analysis, tailored to operate with high precision even when constrained to a minimal dataset of 100 positive samples. At the heart of its architecture lies a refined encoder-decoder framework: the encoder employs convolutional layers coupled with max pooling to distill critical features from images of brain tumors, COVID-19, and wrist fractures. After that, the decoder uses a new concatenation technique to improve and broaden these features; this merges the data's detailed and abstract properties, making the feature space much more robust. This method categorizes the decision according to a predetermined threshold value; scores of 50 or lower are regarded as positive, while scores exceeding 50 indicate negative outcomes.

In summary, the main contributions of this work are:

- The work presents the CAT-U-Net, a new variant of the U-Net architecture with supplementary concatenation layers. This advancement enhances the network's ability to discern intricate features in medical imaging, significantly increasing the accuracy of detecting rare pathological manifestations.
- This research achieved a validation accuracy of 95% and dice coefficient values of 0.946 by using several kinds of real-world medical imaging, such as MRI brain tumors, wrist fracture x-rays, and COVID-19 chest CT scans.

The remainder of this paper is structured as follows: Section II describes related works, and how different researchers solve data-constraints issues. Section III the proposed framework in detail. Section IV discusses dataset description, experimental setup, and result analysis Section V highlights the summaries of this research and future research directions.

II. RELATED WORK

Boosting the efficiency of DL methods in the face of limited data remains a crucial area of investigation. Conventional data augmentation practices are applicable in medical imaging [18]–[20]. However, the generation of poor-quality synthetic data poses significant risks to medical research, and synthetic medical data can also raise privacy concerns [21], [22]. To address these challenges, Chen et al. proposed the Anatomy-Regularized Representation Learning-based Generative Adver-
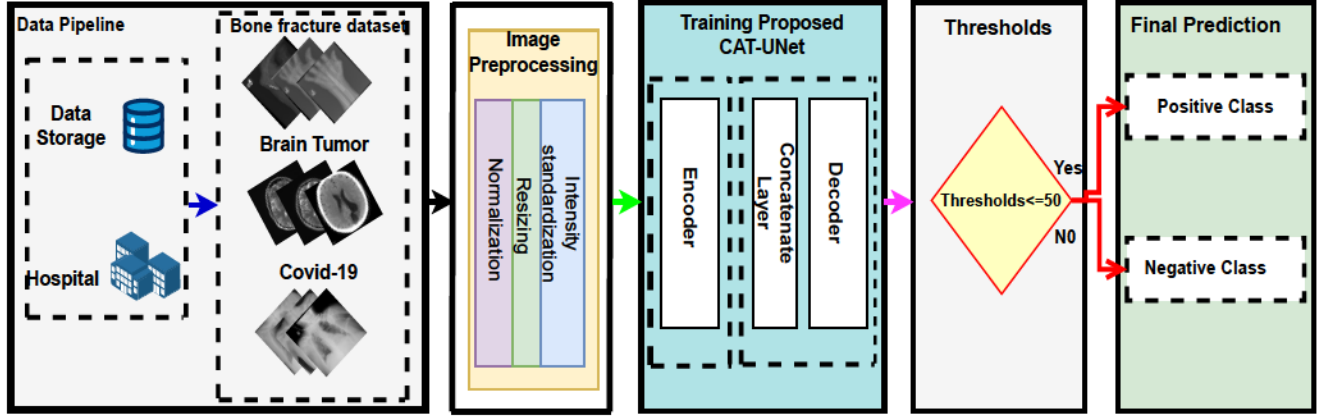
Fig. 1: The pipeline of proposed CAT-U-Net frameworks

sarial Network (ARL-GAN) to address the challenges of limited labeled data and diverse imaging modalities in medical image segmentation. While effective in tasks like skull and cardiac substructure segmentation, ARL-GAN faces overfitting and difficulty generalizing across varied modalities [23]. Similarly, Shetty et al. introduced RAD-DCGAN to improve radiology image classification with synthetic data, achieving a 4-5% accuracy boost over standard augmentation methods. However, RAD-DCGAN also struggles with ensuring data diversity and proper annotation, with risks to diagnostic reliability in real-world applications [24]. Similarly, Mahmood et al. tackled the issue of data scarcity by proposing an innovative approach using unsupervised reverse domain adaptation and adversarial training. This method enables models trained on synthetic data to interpret real-world medical images accurately, significantly improving depth estimation accuracy in monocular endoscopy images. Despite its promise, this approach struggles to simulate the complex anatomical diversity of human tissues, leading to difficulties in generalizing models trained on synthetic data to real-world medical imaging, which may adversely impact accuracy [25]. Rodriguez-Almeida et al. further explored synthetic data generation to address challenges associated with small and imbalanced medical datasets in disease prediction. Their framework showed that synthetic data can enhance classification performance, particularly in small datasets, but the results were mixed, with Gaussian Copula methods outperforming CTGANs in many cases. These findings highlight the limitations of relying solely on synthetic data, underscoring the need for approaches to improve diagnostic accuracy and maintain model reliability without over-reliance on synthetic data [26]. To tackle previous limitations, The suggested CAT-U-Net framework solves these problems by improving feature extraction and performing well across various datasets.

III. METHODOLOGY

A. Proposed CAT-U-Net framework approach

The proposed framework, depicted in Figure 2, employs a symmetrical encoder-decoder structure enhanced with concatenation layers for improved feature extraction and diagnostic precision. The model architecture can be mathematically conceptualized in the following steps:

1) Encoder: The encoder is responsible for capturing the hierarchical features of the input medical images. For an input image $I \in \mathbb{R}^{512 \times 512 \times C}$, the encoder applies a series of convolutional layers followed by max-pooling:

$$E_k(I) = P(\sigma(C(E_{k-1}(I)))) \qquad (1)$$

where $E_k(I)$ is the output of the k-th encoder block, C denotes a convolutional operation followed by a ReLU activation σ, and P denotes a max-pooling operation.

2) Concatenation Layers for Feature Extraction: At each level of the decoder, features from the encoder are concatenated with upsampled features to preserve high-resolution details:

$$F_k = \text{Concat}(U(D_{k-1}), E_{n-k}) \qquad (2)$$

Where F_k represents the concatenated features, U denotes the upsampling operation, D_{k-1} is the output of the previous decoder block, and E_{n-k} is the feature map from the encoder at the corresponding level.

3) Decoder: The decoder reconstructs the segmentation map from the compressed feature representation:

$$D_k(F) = \sigma(U(C(F_k))) \qquad (3)$$

Where $D_k(F)$ is the output of the k-th decoder block, and the operations are defined as before.

B. Robust Feature Learning with Concatenation Layers

The introduction of concatenation layers in the CAT-U-Net framework fosters a more expressive feature representation, enhancing the learning capacity of the network even with limited data. The concatenation operation can be mathematically formalized as a union of feature sets from different levels of the network, enabling the preservation and utilization of multi-scale information:

$$\mathcal{F}_{concat}^{(l)} = \mathcal{F}_{encoder}^{(l)} \oplus \mathcal{U}(\mathcal{F}_{decoder}^{(l)}) \qquad (4)$$

where $\mathcal{F}_{concat}^{(l)}$ represents the concatenated feature map at layer l, $\mathcal{F}_{encoder}^{(l)}$ is the feature map from the encoder, $\mathcal{F}_{decoder}^{(l)}$ is the feature map from the decoder, $\mathcal{U}$ denotes the upsampling operation, and $\oplus$ represents the concatenation operation.

C. Segmentation Map Generation

The final convolutional layer maps the concatenated features to the segmentation map:

$$S = \text{Conv}_{1\times 1}(D_n(F)) \tag{5}$$

Where S represents the final segmentation map and $\text{Conv}_{1\times 1}$ denotes a 1x1 convolution operation that consolidates the feature maps into the final output.

D. Binary Segmentation via Thresholding

The model employs a thresholding technique to binarize the segmentation map, distinguishing between positive and negative diagnoses:

$$B(x,y) = \begin{cases} 1 & \text{if } S(x,y) \leq T \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

Where B is the binarized segmentation map, (x,y) are the pixel coordinates, and T is the threshold value, empirically set to 50 in this study.

E. Leveraging CAT-U-Net for Precise Medical Diagnosis: A Threshold-Based Approach

This Algorithm 1 encapsulates the methodology employed by the CAT-U-Net model for diagnosing medical conditions from imaging data. The model exclusively trains on positive samples for a given condition, such as bone fractures. The initialization of the CAT-U-Net involves setting up an encoder-decoder architecture that includes concatenation layers to enhance feature extraction capabilities. Once initialized, the model undergoes training with preprocessed positive samples, employing mean squared error (MSE) as the loss function to optimize its parameters. This training process leverages the model's ability to reconstruct the training images accurately, honing its capacity to identify intricate features indicative of the medical condition. The evaluation phase introduces a mixed test dataset comprising positive and negative samples. The CAT-U-Net predicts these samples, and the reconstruction performance is quantified using MSE losses. A thresholding method is then applied to these MSE losses: samples resulting in a loss equal to or less than 50 are classified as positive (indicating the presence of the condition), while those with losses greater than 50 are classified as negative.

F. Optimization Framework for Limited Data

The CAT-U-Net model is optimized to extract the maximum amount of diagnostic information from a limited dataset, employing a robust loss function combined with a regularization strategy to ensure a comprehensive learning of features. The optimization objective balances the need for accurate

---

Algorithm 1: Proposed CAT-U-Net Based Medical Image Diagnosis Procedure

---

Data: positive_samples, mixed_test_samples
Result: final_disease_diagnosis
1 **Function** Main (*positive_samples, mixed_test_samples*):
2   **Step 1: Initialize CAT-U-Net Model**
3   cat_u_net ← initialize_CAT_U_Net()
4   **Step 2: Preprocess Positive Training Samples**
5   preprocessed_samples ← preprocess_images(positive_samples)
6   **Step 3: Training the CAT-U-Net Model**
7   X_train, X_test, y_train, y_test ← train_test_split(preprocessed_samples)
8   history ← cat_u_net.fit(X_train, y_train, validation_data=(X_test, y_test))
9   **Step 4: Model Evaluation with Mixed Test Samples**
10   preprocessed_test_samples ← preprocess_images(mixed_test_samples)
11   predicted_samples ← cat_u_net.predict(preprocessed_test_samples)
12   **Step 5: Calculate Mean Squared Error (MSE) Loss**
13   mse_losses ← calculate_mse(preprocessed_test_samples, predicted_samples)
14   **Step 6: Thresholding for Final Diagnosis**
15   threshold ← 50
16   final_predictions ← []
17   **foreach** *mse in mse_losses* **do**
18   **if** *mse ≤ threshold* **then**
19   final_predictions.append('Positive')
20   **else**
21   final_predictions.append('Negative')
22   **return** final_predictions
23   print("Final Disease Diagnosis:", final_predictions)

---

reconstruction against the complexity of the model, facilitating generalization from only 100 samples:

$$\theta^* = \arg\min_\theta \left(L_{\text{MSE}}(\theta) + \lambda\mathcal{R}(\mathcal{F}_{concat}; \theta) \right) \tag{7}$$

Where $L_{\text{MSE}}(\theta)$ is the mean squared error loss that measures the fidelity of image reconstruction, $\mathcal{R}$ represents the regularization term that promotes the learning of a diverse feature set $\mathcal{F}_{concat}$, and λ serves as a hyperparameter that ensures a balance between loss minimization and feature richness.

To achieve this objective, Stochastic Gradient Descent (SGD) is used to iteratively update the model parameters θ, reducing the loss on the training data:

$$\theta_{t+1} = \theta_t - \eta_t \cdot \nabla_\theta L_{\text{MSE}}(\theta_t) \tag{8}$$

Here, η_t is the learning rate at epoch t, dynamically adjusted according to a decay schedule contingent on the validation loss plateau, ensuring steady progress toward the optimal parameter set. This optimization framework is tailored for environments where data scarcity is a significant constraint, enabling the CAT-U-Net to perform with high diagnostic precision.

G. Model Robustness and Healthcare Implications

A theoretical framework supports the CAT-U-Net model's robustness, ensuring its performance with limited healthcare datasets governed by a specific boundedness condition on the feature space.

$$\|\mathcal{F}_{concat}\| \leq K \tag{9}$$

Where K is a predefined constant that bounds the complexity of the concatenated feature space, ensuring the model captures a comprehensive set of features from medical images without succumbing to overfitting. This is crucial in healthcare settings
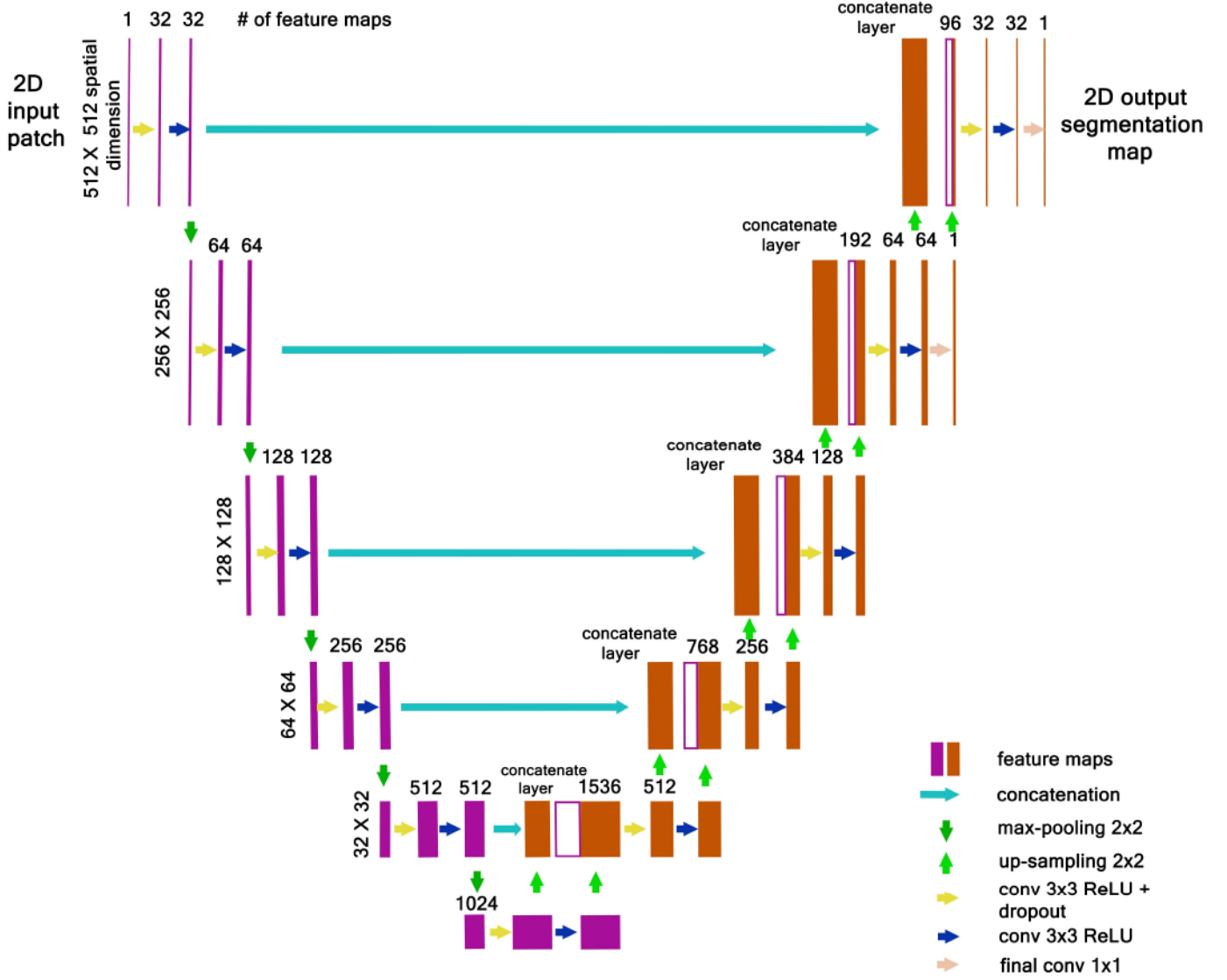

Fig. 2: The proposed CAT-U-Net framework architecture with concatenation layers

where the variance in medical imagery is high, but the volume of data is often low. Coupled with a compact training set, this constraint ensures that CAT-U-Net maintains a high generalization capability, as demonstrated by its predictive accuracy on unseen data.

The optimization goal of the model encapsulates the balance between minimizing reconstruction error and maintaining feature diversity, thus effectively utilizing limited data:

$$\theta^* = \arg\min_{\theta} \left(L_{\mathrm{MSE}}(\theta) + \lambda \mathcal{R}(\mathcal{F}_{concat}; \theta) \right), \quad (10)$$

$$\text{subject to } |D_{\mathrm{train}}| \leq 100 \quad (11)$$

Where λ is a hyperparameter that harmonizes the loss minimization and feature representation complexity. This equation articulates how CAT-U-Net is specifically designed to perform in environments where exhaustive data collection is challenging, a common scenario in medical diagnostics.

To finalize the diagnostic process, a threshold T on the MSE loss is used to discern positive from negative cases, encapsulating the diagnostic capability of CAT-U-Net within a decision boundary:

$$\text{Diagnosis}(x, y) = \begin{cases} \text{`Positive'} & \text{if } L_{\mathrm{MSE}}(\hat{I}(x, y), I(x, y)) \leq T \\ \text{`Negative'} & \text{otherwise.} \end{cases}$$

$$(12)$$

This threshold-based diagnostic criterion aligns with the clinical need for accurate and reliable decision-making based on quantifiable image analysis.

The mathematical rigor presented herein demonstrates the robustness of the CAT-U-Net framework and its practical applicability in future healthcare industries, where data-driven decision-making is paramount. The ability to provide accurate diagnoses with fewer samples translates to quicker, more

efficient medical care, potential reductions in healthcare costs, and broader accessibility of high-quality medical diagnostics.

IV. Proposed Framework Training and Optimization

The CAT-U-Net framework was implemented using the TensorFlow framework due to its comprehensive support for DL operations and GPU acceleration. The following subsection explains this study's implementation, training procedure, and evaluation protocol.

A. Training Procedure

Training the CAT-U-Net involved steps, starting with pre-processing the input images. Each image was resized to a uniform dimension of $256 \times 256 \times 4$ and normalized to have pixel values in the range [0, 1]. The network was then trained using the preprocessed images, with the Mean Squared Error (MSE) as the loss function. Stochastic Gradient Descent (SGD) with a learning rate of 0.01 was employed as the optimizer. To prevent overfitting, a dropout rate of 0.5 was applied after each convolutional layer in the decoder part of the network. The model was trained for 50 epochs with a batch size of 8, and the best-performing model on the validation set was saved for subsequent evaluation.

B. Hyperparameter Optimization

The model's hyperparameters were meticulously optimized to enhance performance, with a focus on the learning rate (η) adaptation strategy. To calibrate η, an exponential decay function was applied, governed by the following relation:

$$\eta_{t+1} = \eta_t \times \gamma^{\lfloor \frac{t}{\tau} \rfloor} \tag{13}$$

Where η_{t+1} is the learning rate for the subsequent epoch, η_t is the learning rate for the current epoch, γ is the decay rate, t is the current epoch number, and τ is the patience parameter, denoting the number of epochs after which the plateau condition is evaluated. The decay rate γ was set to 0.1, and the learning rate was reduced if no improvement in validation loss was observed throughout ten epochs ($\tau = 10$).

V. Results and Discussion

A. Experimental Setup

Experiments were conducted on a High-Performance Computing (HPC) system featuring 31 GB of DDR4 RAM and an NVIDIA GeForce GTX 3070 GPU with 8 GB of GDDR6 memory, operating on a Linux Ubuntu environment. This setup was chosen to effectively manage the computational requirements of the CAT-U-Net framework, particularly for processing medical imaging data and training the model under data-constrained conditions. The framework's robustness was validated across MRI brain tumor images, COVID-19 chest CT scans, and wrist bone fracture X-ray images.

B. Dataset Description

The CAT-U-Net framework underwent validation using three distinct medical imaging datasets from Kaggle, each addressing different diagnostic challenges: the Br35H dataset for MRI brain tumor detection [27], a chest CT scan dataset for COVID-19 identification [28], and an X-ray dataset for Wrist Bone Fracture [29]. Training utilized 100 images from each dataset, while testing employed 50 images, demonstrating the framework's robustness in detecting various pathological features, particularly in scenarios with limited data, a common challenge in medical diagnostics.

C. Performance Metrics for Evaluation

The proposed CAT-U-Net framework's effectiveness was assessed using reconstruction accuracy and the Dice coefficient to measure image precision and segmentation under limited data constraints.

1) Reconstruction Accuracy: Reconstruction accuracy was computed as the inverse of the mean squared error (MSE) between the original and reconstructed images across the test set. This metric reflects the model's ability to faithfully reconstruct key features in medical images, which is crucial for accurate diagnosis:

$$\text{Reconstruction Accuracy} = 1 - \frac{1}{N}\sum_{i=1}^{N}\|I_i - \hat{I}_i\|^2 \tag{14}$$

where I_i represents the original image, $\hat{I}_i$ is the reconstructed image by the model, and N is the number of images in the test dataset.

2) Dice Coefficient: The Dice coefficient, also known as the Dice similarity coefficient (DSC), was used to measure the overlap between the predicted segmentation and the ground truth, providing a spatial accuracy assessment of the segmentation:

$$\text{Dice Coefficient} = \frac{2 \times |X \cap Y|}{|X| + |Y|} \tag{15}$$

Where X represents the set of pixels in the ground truth segmentation, and Y denotes the set of pixels in the predicted segmentation by the model. A higher Dice score indicates more significant overlap and, thus, higher segmentation accuracy.

D. Quantitative Evaluation of CAT-U-Net Across Diverse Medical Imaging Datasets

The CAT-U-Net framework demonstrates high efficacy in medical image analysis, as evidenced by comprehensive statistical and visual evaluations across multiple datasets. In the Bone Fracture dataset, a confusion matrix (Figure 5) reveals the model's high predictive accuracy, correctly diagnosing 97 out of 99 positive cases, showcasing its sensitivity and specificity in detecting medical conditions. This accuracy is further supported by a high Dice coefficient of 0.946 (Figure 4), indicating an exceptional overlap between the model's predicted segmentations and the ground truth. Additionally, the training and validation loss and accuracy curves (Figure 6)

exhibit a steady convergence, highlighting the model's robust learning capability from a limited number of training samples. As depicted in Figure 3, the CAT-U-Net achieves consistently high performance across different datasets, with reconstruction accuracies of 94% for Brain Tumor (MRI), 93% for COVID-19 (CT Scan), and 98% for Wrist Fractures (X-ray). The corresponding Dice coefficients also remain high, at 93%, 92%, and 95%, respectively. These results confirm the model's adaptability and precision across diverse clinical scenarios, underscoring its potential as a reliable tool for medical image diagnostics.
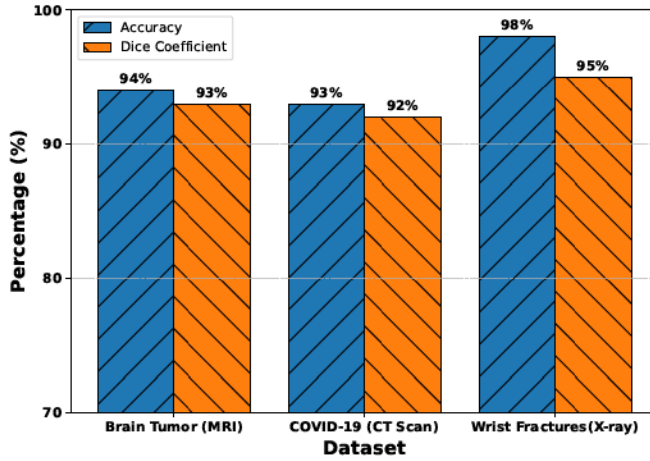
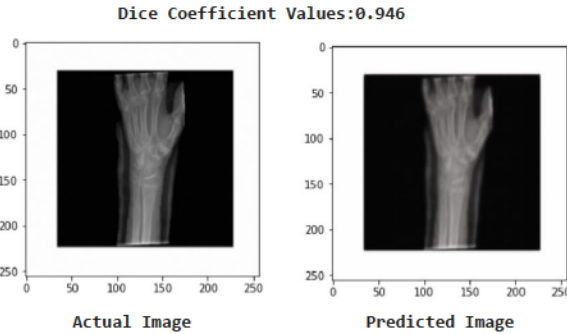Fig. 3: Performance Analysis of the Proposed Framework Across Different Datasets

Fig. 4: Dice-Cofficient value for Bone fracture dataset

E. Quantitative Comparison with Existing State-of-the-Art Methods

Figure 7 provides a comparative analysis of the proposed CAT-U-Net framework against several state-of-the-art models, including SIFA [23], ARL-GAN semi-supervised [23], U-Net [30], and PrEGAN [18], in terms of accuracy o Dice Similarity Coefficient (DSC). The proposed framework achieves superior performance with a DSC of 94.0%, outperforming other methods such as PrEGAN (88.32%), ARL-GAN semi-supervised (84.09%), SIFA (81.3%), and U-Net (76.99%). These results
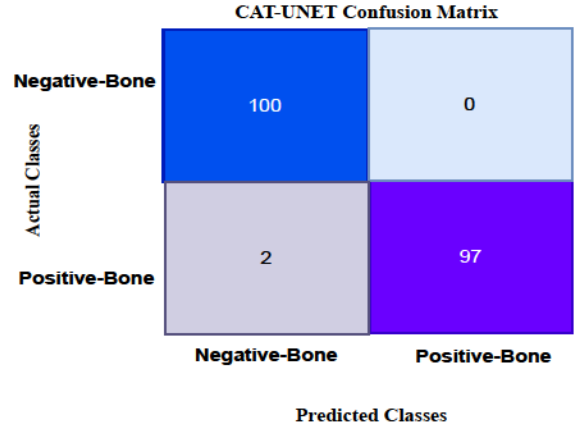

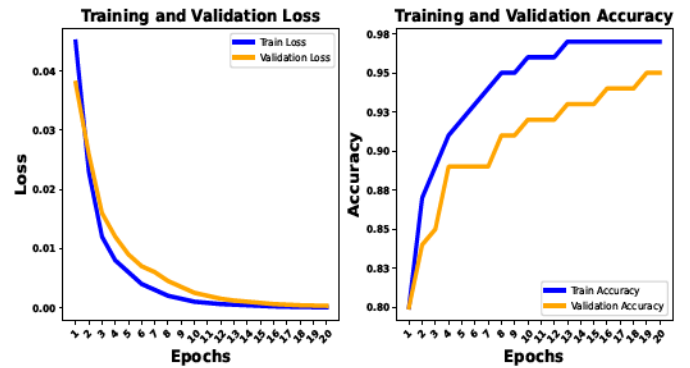Fig. 5: Proposed framework Confusion matrix

Fig. 6: Training and Validation loss of propose framework

demonstrate the efficacy of the CAT-U-Net framework in achieving higher segmentation accuracy and robustness, particularly in medical imaging tasks with limited data availability.

VI. CONCLUSION AND FUTURE WORK

In conclusion, while the CAT-U-Net framework presents a significant advancement in medical image analysis, its broader adoption in the healthcare industry will depend on addressing integration challenges and ensuring robustness across diverse clinical environments. Future research could focus on optimizing the framework for even smaller datasets and more complex imaging modalities and integrating advanced adversarial training methods to enhance robustness against adversarial attacks in clinical applications. Continued research in these areas will be essential to unlocking the full potential of this technology for improving patient outcomes.
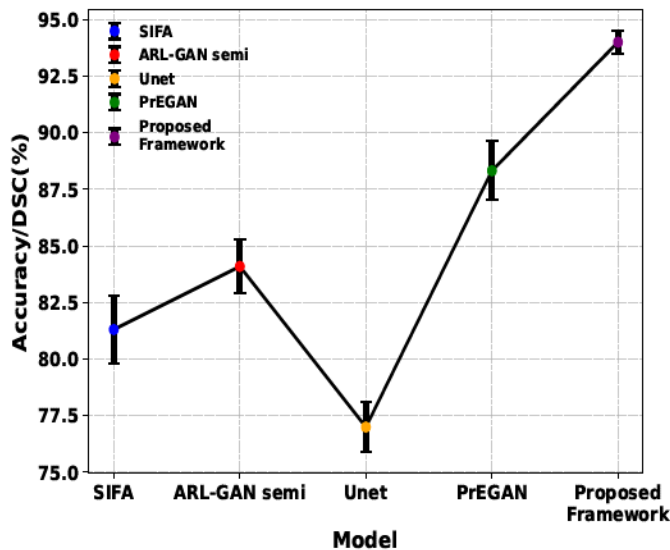
VII. ACKNOWLEDGMENT

Fig. 7: Comparative Performance Analysis of Proposed and Existing Models

REFERENCES

[1] V. Nath, D. Yang, B. A. Landman, D. Xu, and H. R. Roth, "Diminishing uncertainty within the training pool: Active learning for medical image segmentation," *IEEE Transactions on Medical Imaging*, vol. 40, no. 10, pp. 2534–2547, 2021.

[2] W. Li, J. Li, Z. Wang, J. Polson, A. E. Sisk, D. P. Sajed, W. Speier, and C. W. Arnold, "Pathal: An active learning framework for histopathology image analysis," *IEEE transactions on medical imaging*, vol. 41, no. 5, pp. 1176–1187, 2021.

[3] M. S. Ahmed and S. Giordano, "Pre-trained lightweight deep learning models for surgical instrument detection: Performance evaluation for edge inference," in *GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, pp. 3873–3878, 2023.

[4] S. Pattanaik, C. Chakraborty, S. Behera, S. K. Majhi, and S. K. Pani, "An miot framework of consumer technology for medical diseases prediction," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 3754–3761, 2024.

[5] A. S. Boehringer, A. Sanaat, H. Arabi, and H. Zaidi, "An active learning approach to deep learning glioma segmentation from brain mr images," in *2022 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pp. 1–2, IEEE, 2022.

[6] S. Liu, S. Yin, L. Qu, M. Wang, and Z. Song, "A structure-aware framework of unsupervised cross-modality domain adaptation via frequency and spatial knowledge distillation," *IEEE Transactions on Medical Imaging*, 2023.

[7] A. Farrag, G. Gad, Z. M. Fadlullah, and M. M. Fouda, "Mammogram tumor segmentation with preserved local resolution: An explainable ai system," in *GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, pp. 314–319, 2023.

[8] E. Haque, K. Hasan, I. Ahmed, M. S. Alam, and T. Islam, "Towards an interpretable ai framework for advanced classification of unmanned aerial vehicles (uavs)," in *2024 IEEE 21st Consumer Communications Networking Conference (CCNC)*, pp. 644–645, 2024.

[9] R. Wiencek and S. Ghosh, "Deep reinforcement learning for adaptive optimization of pi control for microgrid under fault and variable loading," in *2024 6th Global Power, Energy and Communication Conference (GPECOM)*, pp. 826–831, 2024.

[10] P. Chlap, H. Min, N. Vandenberg, J. Dowling, L. Holloway, and A. Haworth, "A review of medical image data augmentation techniques for deep learning applications," *Journal of Medical Imaging and Radiation Oncology*, vol. 65, no. 5, pp. 545–563, 2021.

[11] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 1–48, 2019.

[12] A. Amin, K. Hasan, S. Zein-Sabatto, D. Chimba, I. Ahmed, and T. Islam, "An explainable ai framework for artificial intelligence of medical things," in *2023 IEEE Globecom Workshops (GC Wkshps)*, pp. 2097–2102, 2023.

[13] Z. Wang, *Predictive Learning from Real-World Medical Data: Overcoming Quality Challenges*. PhD thesis, 2023.

[14] T. J. Jaspers, T. G. Boers, C. H. Kusters, M. R. Jong, J. B. Jukema, A. J. de Groof, J. J. Bergman, P. H. de With, and F. van der Sommen, "Robustness evaluation of deep neural networks for endoscopic image analysis: Insights and strategies," *Medical Image Analysis*, p. 103157, 2024.

[15] G. D'Acquisto, "Synthetic data and data protection laws," *Journal of Data Protection & Privacy*, vol. 6, no. 3, pp. 227–239, 2024.

[16] M. Giuffrè and D. L. Shung, "Harnessing the power of synthetic data in healthcare: innovation, application, and privacy," *NPJ Digital Medicine*, vol. 6, no. 1, p. 186, 2023.

[17] A. Amin, K. Hasan, S. Zein-Sabatto, D. Chimba, L. Hong, I. Ahmed, and T. Islam, "Empowering healthcare through privacy-preserving mri analysis," in *SoutheastCon 2024*, pp. 1534–1539, 2024.

[18] S. T. Ahmed, R. Sivakami, V. K. V, M. T. R, S. B. Khan, A. Mashat, and A. Almusharraf, "Pregan: Privacy enhanced clinical emr generation: Leveraging gan model for customer de-identification," *IEEE Transactions on Consumer Electronics*, pp. 1–1, 2024.

[19] S. Islam, M. T. Aziz, H. R. Nabil, J. R. Jim, M. F. Mridha, M. M. Kabir, N. Asai, and J. Shin, "Generative adversarial networks (gans) in medical imaging: Advancements, applications, and challenges," *IEEE Access*, vol. 12, pp. 35728–35753, 2024.

[20] A. Amin, K. Hasan, and M. S. Hossain, "Xai-empowered mri analysis for consumer electronic health," *IEEE Transactions on Consumer Electronics*, pp. 1–1, 2024.

[21] S. Abdel-Khalek, A. D. Algarni, G. Amoudi, S. Alkhalaf, F. M. Alhomayani, and S. Kathiresan, "Leveraging ai-generated content for synthetic electronic health record generation with deep learning-based diagnosis model," *IEEE Transactions on Consumer Electronics*, pp. 1–1, 2024.

[22] S. A. Naeini, L. Simmatis, D. Jafari, Y. Yunusova, and B. Taati, "Improving dysarthric speech segmentation with emulated and synthetic augmentation," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 12, pp. 382–389, 2024.

[23] X. Chen, C. Lian, L. Wang, H. Deng, T. Kuang, S. Fung, J. Gateno, P.-T. Yap, J. J. Xia, and D. Shen, "Anatomy-regularized representation learning for cross-modality medical image segmentation," *IEEE transactions on medical imaging*, vol. 40, no. 1, pp. 274–285, 2020.

[24] S. Shetty, V. Ananthanarayana, and A. Mahale, "Data augmentation vs. synthetic data generation: An empirical evaluation for enhancing radiology image classification," in *2023 IEEE 17th International Conference on Industrial and Information Systems (ICIIS)*, pp. 1–6, IEEE, 2023.

[25] F. Mahmood, R. Chen, and N. J. Durr, "Unsupervised reverse domain adaptation for synthetic medical images via adversarial training," *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2572–2581, 2018.

[26] A. J. Rodriguez-Almeida, H. Fabelo, S. Ortega, A. Deniz, F. J. Balea-Fernandez, E. Quevedo, C. Soguero-Ruiz, A. M. Wägner, and G. M. Callico, "Synthetic patient data generation and evaluation in disease prediction using small and imbalanced datasets," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 6, pp. 2670–2680, 2023.

[27] A. Hamada, "Brain tumor detection," 2021. Accessed: 2023-04-08.

[28] D. M. C. TAWSIFUR RAHMAN and A. Khandakar, "Covid-19 radiography database," 2022. Accessed: 2024-04-12.

[29] T. V. A. RAJ, "Wrist bone fracture segmentation," 2021. Accessed: 2024-04-12.

[30] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pp. 234–241, Springer, 2015.