Enhancing U-Net for Wrist Fracture Segmentation in X-ray Images using Adaptive Callbacks and Weighted Loss Functions

Teuku Radillah[1],[*] (ID), Sarjon Defit[2], (ID), Gunadi Widi Nurcahyo[3], (ID)

[1]*Institut Teknologi Mitra Gama, Jl. Khayangan No 99, Duri-Riau,28783, Indonesia*

[2,3]*Universitas Putra Indonesia YPTK Padang, Jl. Raya Lubuk Begalung, Sumatera Barat 25171, Indonesia*

Abstract

Wrist fracture detection using medical imaging remains a challenging task due to the subtle and varied nature of fracture appearances. This study aims to improve the segmentation performance for wrist fracture detection by enhancing the U-Net model through adaptive training and customized loss functions. The main contribution of this research lies in the integration of an adaptive callback mechanism and a dual-weighted loss strategy that combines Dice Loss and Binary Cross-Entropy Loss. The adaptive training mechanism dynamically adjusts training parameters based on model performance, improving generalization and preventing overfitting. Meanwhile, the linear and non-linear exponential weighting in the loss function enables balanced learning and focuses attention on complex fracture regions during training.The methodology involved training the enhanced U-Net on a dataset of 1,344 wrist X-ray images, comprising both normal and fractured cases. Seventy percent of the images were allocated for training, while the remaining thirty percent were used for testing. The images were resized to 256×256 pixels, and the model was trained for up to 100 epochs with a batch size of 8. Experimental results demonstrated that the proposed model achieved strong performance, with an accuracy of 91.78%, precision of 87.78%, recall of 86.70%, and F1-score of 87.17%. These results indicate that the combination of adaptive training and loss weighting significantly improves the sensitivity and robustness of the segmentation process. The system was also validated through training and validation metrics, where lower loss and improved accuracy confirmed its ability to generalize well on unseen data. The approach offers promising implications for clinical applications, where rapid and accurate identification of wrist fractures can enhance diagnostic workflows and patient outcomes. Future research may explore the extension of this method to other fracture types and the incorporation of real-time diagnostic tools in clinical settings.

Keywords: Wrist Fracture Segmentation, Enhanced U-Net, Adaptive Callback Training, Weighted Loss Function, X-Ray Image Analysis

1. Introduction

Wrist fractures represent a common type of musculoskeletal injury, particularly among the elderly population and individuals who experience direct trauma to the hand region [1]. Early and accurate detection of such fractures is essential to prevent long-term complications, including functional impairment, malunion, or the need for more invasive surgical interventions [2]. In clinical practice, interpreting radiographic (X-ray) images can be challenging for radiologists due to the visual variability of fractures, overlapping bone structures, and low contrast in the affected area [3]. With the advancement of artificial intelligence, deep learning methods, particularly the U-Net architecture, have shown great promise in medical image segmentation tasks, including X-ray analysis. U-Net is designed to perform well with limited data and is capable of capturing both spatial and contextual information through its symmetric encoder-decoder structure [4]. However, the application of the standard U-Net for wrist fracture detection still presents several challenges, including susceptibility to noise, risk of overfitting on limited datasets, and reduced effectiveness in class-imbalanced scenarios where fracture regions occupy significantly fewer pixels compared to non-fractured areas.

Adopting a more adaptive training method and a loss function that represents a more complex segmentation problem is necessary to overcome these challenges. Adaptive callbacks, including dynamic learning rate adjustment, validation-based early stopping, and performance monitoring during training, enable the system to automatically adjust the training process to the learning dynamics of the model. Consequently, the incorporation of loss functions such as Dice

Loss and Binary Cross-Entropy (BCE) Loss [5], with adaptive linear as well as exponential weighting, has been shown to improve segmentation performance by achieving a balance between pixel accuracy and target object structure [6]. In view of the previously mentioned background, the objective of this research is to improve the efficacy of the U-Net model in the identification of wrist fractures. This process is carried out by incorporating an adaptive callback training method and a weighted joint loss function design [7]. The method aims to enhance the precision and efficiency of the model while ensuring its generalization to variations in the morphology and intensity of fractures evident in medical images. This research contributes strategically to the development of image-based clinical decision support systems, particularly in the automated radiological diagnosis of limb bone fractures.

2. Related Works

Recent studies have explored various deep learning approaches for fracture detection and medical image segmentation. Classification-based methods using Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRUs) have shown notable performance in predicting wrist fractures, achieving accuracy levels between 83% and 90% [8]. While effective in classification, these methods lack spatial localization, making them less suitable for precise segmentation tasks required in fracture boundary identification. To improve localization, convolutional neural networks (CNNs), particularly ensemble models such as VGG19, ResNet152, MobileNet, and DenseNet, have been employed for fracture classification, yielding high precision and recall scores [9]. However, these models typically rely on coarse-level predictions and do not fully exploit pixel-level segmentation essential for clinical decision support.

Segmentation-based approaches using U-Net and its variants have gained popularity due to their ability to learn fine-grained spatial features. For instance, studies on osteoporotic fractures [10], pulmonary embolism [11], and brain or liver tumors [12], [13] utilized encoder-decoder networks combined with Dice Loss or Jaccard Loss to improve performance. These works emphasize the importance of using loss functions tailored for medical imaging. Nonetheless, they often overlook the challenges of class imbalance and sensitivity to noise, especially in small lesion areas such as wrist fractures. More advanced models like Residual U-Net and HAD-Net have been proposed to overcome limitations in basic U-Net architecture, yielding improved Dice coefficients across various organs [14], [15]. Similarly, research on hybrid loss functions combining Dice, BCE, and Jaccard losses has demonstrated superior results in applications like ovarian and abdominal organ segmentation [16], [17]. However, most of these models do not incorporate dynamic training strategies, such as adaptive callbacks or epoch-based loss weighting, which can further enhance generalization and convergence.

Several works have also explored segmentation using generalized or domain-specific loss functions in tasks involving CT or MRI images [18], [19], [20]. While these studies offer valuable insights into loss function design, they primarily focus on large or clearly defined anatomical structures, leaving a gap in the effective segmentation of subtle fractures in radiographic wrist images. In summary, although previous research has demonstrated the effectiveness of CNN-based and loss-optimized segmentation models across various medical domains, limited work has addressed the unique challenges of wrist fracture segmentation using adaptive training mechanisms and temporally weighted hybrid loss functions. This study aims to bridge that gap by enhancing the U-Net architecture with adaptive callbacks and a combined linear–exponential loss strategy, tailored to the class imbalance and morphological variability present in wrist X-ray images.

3. Materials and Methodology

The primary phase of this research included the acquisition of wrist fracture image through the implementation of preprocessing data, the subsequent labelling of data about the fracture area, and the segmentation of data using U-Net model. In the segmentation process, a comparison was made using a combination of the Dice and BCE loss methods, as implemented in U-Net model. Concerning the training process, the callback for the training feature was used in combining linear weighting and non-linear exponential weighting to achieve improved segmentation results. The outcomes of the weighting combination were represented by a new method, namely the adaptive weighted combined loss function.

3.1. Wrist Fracture

A wrist fracture is an injury to the wrist caused by a variety of traumatic events, including driving accidents, direct impact to the wrist with a blunt object, pinching, or active events such as sporting accidents and elderly activities related to osteoporosis. The following symptoms are indicative of a wrist fracture, namely inflammation, pain, swelling, and difficulty moving the fractured wrist area [22].

3.2. Wrist Fracture Dataset

The dataset used in this study consisted of X-ray images of fractured and non-fractured wrists, which were collected from various sources, including both public and private hospitals, such as RSUD Kota Dumai and Permata Hati Hospital in Duri City, as well as publicly available datasets obtained from the website https://www.kaggle.com/search?q=wrist+fracture using the keyword "wrist fracture". The dataset comprised a total of 1,344 X-ray images, including 924 images of fractured wrists and 424 images of normal wrists, each with varying pixel resolutions.

To facilitate the training and evaluation process, the dataset was divided into two subsets: 644 images were allocated for training, and 276 images were used for testing. All training images consisted of wrist X-rays that had not previously been used in any learning or model development phase, ensuring the validity of the evaluation process. However, the dataset lacked uniform metadata such as patient age, sex, or acquisition conditions. Additionally, ethical clearance was not explicitly required since publicly available datasets were used, and any patient data sourced from private hospitals were anonymized prior to use in accordance with applicable research ethics standards, as illustrated in figure 1 showing an example of a wrist fracture X-ray image.

Figure 1. Wrist Fracture

3.3. Data Preprocessing

Data preprocessing started with the collection of wrist fracture image, followed by cropping, then contrast stretching. Resizing to adjust the size of image or annotations that showed the fracture area was also performed, followed by augmentation and normalization. Augmentation is a method in image processing or machine learning to artificially expand the amount of training data by modifying existing image [23]. This method was used to improve the heterogeneity of the dataset, facilitate learning of diverse variations of the model, and mitigate the risk of overfitting during the training process. Subsequently, the image showing wrist fracture were labelled or annotated during the process. These annotations were necessary to be created by medical experts, including radiological and orthopedic specialists, as illustrated in figure 2 showing the preprocessing stages of wrist fracture.
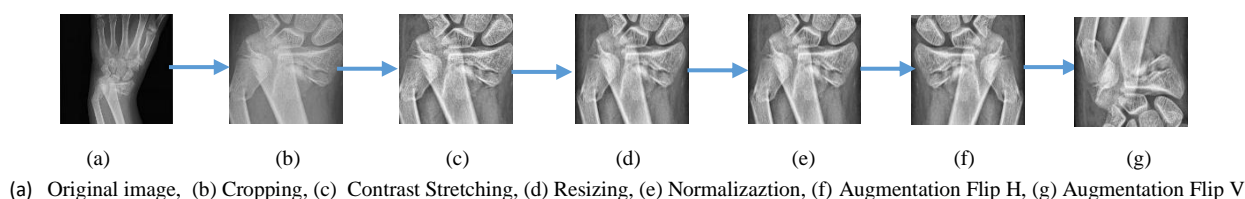

(a)   (b)   (c)   (d)   (e)   (f)   (g)

(a) Original image, (b) Cropping, (c) Contrast Stretching, (d) Resizing, (e) Normalizaztion, (f) Augmentation Flip H, (g) Augmentation Flip V

Figure 2. Preprocessing Stages of Wrist Fracture

3.4. Dice Loss

Dice Loss is a metric commonly used in segmentation tasks to evaluate the similarity between the predicted segmentation and the actual ground truth in an image [24]. It is particularly prevalent in the medical image segmentation domain [25], being derived from the Dice Coefficient, which measures the overlap between two datasets. This loss function evaluates the agreement between model predictions and ground truth labels [26], and is defined as follows [27].

$$Dice\ Loss\ =\ 1\ -\ \frac{2\ x\ TP}{2\ x\ TP+FP+FN} \tag{1}$$

TP (True Positives) = Number of pixels correctly predicted as positive; FP (False Positives) = Number of pixels incorrectly predicted as positive [28]; FN (False Negatives) = Number of positive pixels missed by the prediction. The application of Dice Loss was major to improve the performance of the model for segmentation, particularly in the domains of medical image processing and other tasks characterized by substantial class imbalance.

3.5. BCE Loss

BCE is commonly applied in different binary classification tasks, including object detection and image segmentation in deep learning [29]. The equation for BCE Loss during the analysis was presented as follows [30].

$$BCE\ Loss1\ =\ -\frac{1}{N}\sum_{i-1}^{N} yi\ .\log(pi)\ +\ (1-yi)\ .\log(1-pi) \tag{2}$$

yi = Ground truth label for the i-th pixel (0 or 1); pi = Predicted probability for the i-th pixel (ranging from 0 to 1); N = Total number of pixels. BCE Loss provides a quantitative measure of the discrepancy between the model's predicted probabilities and the actual binary labels [31]. It is particularly useful for differentiating between two distinct classes in binary classification scenarios

3.6. Adaptive Weighted Combined Loss Function

Adaptive Weighted Combined (AWC) Loss is a method that integrates Dice and BCE losses using dynamically adjusted weighting functions throughout the training process. While previous studies applied either fixed weights or a single adaptive strategy—such as linear or exponential weighting—this study explicitly combines both linear and nonlinear exponential weighting within a single loss formulation. The linear component provides a steady transition over training epochs, while the exponential component enables more responsive adaptation as training progresses. This dual-weighting mechanism is designed to leverage the strengths of both gradual and sharp adjustments in loss contribution, thereby enhancing the model's ability to cope with complex variations in data distribution. The proposed method aims to improve segmentation performance, particularly in medical images with challenging intensity and texture characteristics [32]. The formulation for linear weighting during the training process is defined as follows:

$$\alpha(t) = \alpha start + (\alpha end - \alpha start) \cdot (1 - e - \lambda \cdot t)$$
$$\beta(t) = \beta start + (\beta end - \beta start) \cdot e - \lambda \cdot (T - t) \tag{3}$$

The progression of the training allowed the weight to be adjusted by the functions α(t) and β(t). Moreover, the notations used during the process of the analysis were stated as follows. t = epoch or current iteration; T = The total number of epochs or training iterations; αstart, αend, βstart and βend = parameters that determined the initial and final values of the adaptive weights. The equation for exponential non-linear weighting used in this analysis was presented as follows.

$$\alpha(t) = \alpha start + (\alpha end - \alpha start) \cdot (1 - e - \lambda \cdot t)$$
$$\beta(t) = \beta start + (\beta end - \beta start) \cdot e - \lambda \cdot (T - t \tag{4}$$

αstart = The initial value of the parameter α; αend = The final value of the parameter α; t = The current time or iteration typically measured in units ranging from 0 to T; T = Total Number of Epochs; λ = The exponential scaling factor which was a critical element to control the rate of change and applied to the degree of non-linearity in this case; e = The base of the natural logarithm. The combination of linear and non-linear exponential weighting led to the following equation:

$$\alpha(t) = W1 \cdot (\alpha start + (\alpha end - \alpha start) \cdot \frac{t}{T}) + W2 \cdot (\alpha start + (\alpha end - \alpha start) \cdot (1 - e^{\wedge}(-\lambda t))) \tag{5}$$

$$\beta(t) = W1 \cdot (\beta start + (\beta end - \beta start) \cdot \frac{T - t}{T}) + W2 \cdot (\beta start + (\beta end - \beta start) \cdot e^{\wedge}(-\lambda(T - t)))$$

When W1 > W2, linear weighting was more dominant; When W2 > W1, the exponential non-linear weighting was more dominant; When W1 = W2, it was evident that both variables showed equivalent levels of contribution. Linear weighting: $\alpha(t) \propto \frac{t}{T}$ = Consistently increasing over time; $\beta(t) \propto \frac{T-t}{T}$ = Reduced consistently over time Exponential non-linear weighting: $\alpha(t) \propto (1-e^{\wedge}(-\lambda t)))$ = The initial rise was gradual but subsequently accelerated; $\beta(t) \propto e^{\wedge}(-\lambda(T-t)))$ = There was an instantaneous fall and later slowed down.

The parameters αstart, αend, βstart, βend , and λ were selected based on prior studies and empirical testing. The values aim to ensure smooth transitions in the loss weighting during training. A brief sensitivity analysis was conducted by varying λ and observing the effect on model convergence and segmentation accuracy. Results showed that moderate values (e.g., λ=5) yielded stable and improved performance. The combination of linear and exponential weighting allows the model to adapt dynamically during training, improving segmentation results over static approaches.

3.7. U-Net Model

Figure 3 shows the U-Net architecture, which is a deep neural network architecture that has been specifically designed for image segmentation tasks, especially in medical image processing [33]. U-Net architecture consisted of an encoder-decoder structure that included skip connections [34]. Additionally, the implementation of this architecture in a deep learning framework for image segmentation included the use of convolutional layers, batch normalization, and activation functions (e.g. ReLU) [35]. U-Net network had been developed from CNN for medical image segmentation [36]. The model was composed of a contracting (encoder) on the left and an expansion path (decoder) on the right side, as these two were interconnected via a bridge. The encoder functioned as a feature extractor, learning an abstract representation of the input image through a series of encoder blocks. Moreover, the decoder generated a semantic segmentation mask, using the abstract representation as the input.
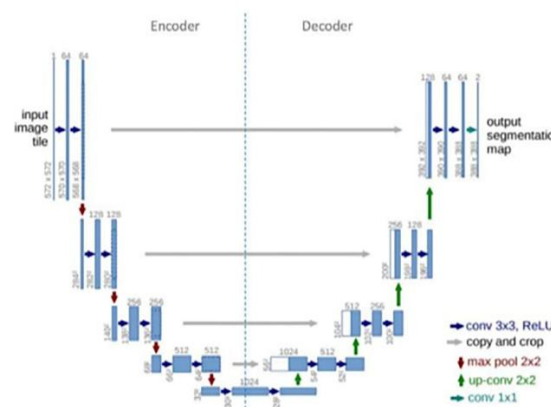

Figure 3. U-Net Architecture

U-Net model architecture comprised input layer, encoding path, bottleneck, decoding phase, output layer and a model.

3.7.1 Performance Development of U-Net

The development of the U-Net was an integral component of the training process, specifically the callback training feature. This feature was designed to improve the performance of U-Net segmentation by using an adaptive weighted combined loss function, which was applied to the Dice and BCE loss. Additionally, the use of adaptive weighting served to influence the degree of segmentation accuracy achieved by U-Net.

3.7.2 U-Net Training Performance

The training of the U-Net model typically employed the Cross-Entropy Loss function. The callback training feature was an integral component of U-Net segmentation, serving to regulate and optimize the training process. Callbacks were utilized in frameworks such as TensorFlow/Keras to enhance the efficiency of the training procedure. In this study, the adaptive weighted combined loss function method was implemented to improve the performance of the U-Net model during training. A total of 70% of the dataset was allocated for training, while the remaining 30% was

reserved for testing. Following preprocessing and analysis, the image size of the dataset was standardized to 256×256 pixels. The U-Net model was further developed by incorporating linear and non-linear exponential weighting methods, applied through callbacks during training. These callbacks were not part of the core U-Net architecture; rather, they operated externally and were only active throughout the training phase. The layer structure of the U-Net model remained unaffected by the callbacks. However, their implementation was observed to have a significant impact on the training outcomes. The development of the U-Net training performance is presented in table 1.

Table 1. Development of the U-Net Callback for Training model

Architecture Stages	Image Size (H x W x C)	Component	Description
Input Layer	$256 \times 256 \times 1$	Input Grayscale	X-ray image of wrist fracture
Encoder Block 1	$256 \times 256 \times 64$	$2\times$ Conv2D (3×3) + ReLU + MaxPool	Initial feature extraction
Encoder Block 2	$128 \times 128 \times 128$	$2\times$ Conv2D (3×3) + ReLU + MaxPool	Early intermediate level features
Encoder Block 3	$64 \times 64 \times 256$	$2\times$ Conv2D (3×3) + ReLU + MaxPool	Advanced intermediate level features
Encoder Block 4	$32 \times 32 \times 512$	$2\times$ Conv2D (3×3) + ReLU + MaxPool	Advanced feature extraction
Bottleneck	$16 \times 16 \times 1024$	$2\times$ Conv2D (3×3) + ReLU	Deepest layer, highly abstract features
Decoder Block 1	$32 \times 32 \times 512$	UpSampling + Concat + $2\times$ Conv2D	Concatenation with Encoder Block 4
Decoder Block 2	$64 \times 64 \times 256$	UpSampling + Concat + $2\times$ Conv2D	Concatenation with Encoder Block 3
Decoder Block 3	$128 \times 128 \times 128$	UpSampling + Concat + $2\times$ Conv2D	Concatenation with Encoder Block 2
Decoder Block 4	$256 \times 256 \times 64$	UpSampling + Concat + $2\times$ Conv2D	Concatenation with Encoder Block 1
Output Layer	$256 \times 256 \times 1$	Conv2D (1×1) + Sigmoid	Binary segmentation mask output (0 : no fracture, 1 : fracture)
Loss Function	Cross-Entropy Loss	Dice Loss + BCE Loss	Adaptive loss combinations are used
Loss weight (Epoch)	-	$\alpha(t), \beta(t)$	Weights adaptive with linear & exponential, changing with epochs
Callback Training	-	Custom Callback Function	Set loss weighting & monitoring during training (e.g. EarlyStopping)

4. Results and Discussion

The section showed the experimental analysis concerning the development of an adaptive loss function, namely the adaptive weighted combined loss function on X-ray image of different wrist fracture shapes. The implementation of the adaptive weighted combined loss function was used in the training process of U-Net model, specifically in the callback for training, which automatically updated the α and β values at the conclusion of each epoch. Furthermore, the callback for training function at the commencement of the process focused on BCE loss to accelerate per-pixel convergence and gradually increased the focus on the Dice loss to improve global segmentation quality.

4.1. Adaptive Loss Function Testing

The testing process included, comparing several U-net model variations, consisting of standard U-net, combination of U-net with dice and BCE loss, as well as U-net with a combined Dice and BCE loss using training callbacks. Additionally, U-net was compared with the same loss function improved by linear and non-linear exponential weighting and callbacks.

4.2. Testing Results Adaptive Weighted Combined Loss Function

The results demonstrated the development of the adaptive loss function method using the following parameters: input size of 256×256 pixels, maximum epochs of 100, batch size of 8, alpha start of 0.5, alpha end of 1.0, beta start of 0.5, beta end of 1.0, and a lambda parameter of 0.1. Figure 5 illustrates the training and testing phases by comparing various loss function methods with the proposed Adaptive Weighted Combined Loss, which integrates Dice Loss and BCE Loss in the U-Net model, as illustrated in figure 5 showing the implementation run using Python 3.11.1 for Windows.
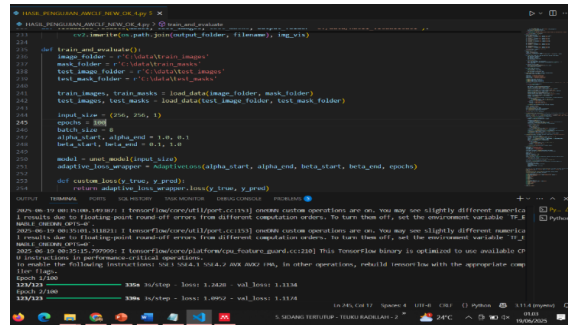
**Figure 5.** Run the App Using Python 3.11.1 For Windows

The execution process in the application, which used Python 3.11.1 for Windows, produced output in the form of training and validation loss graphs, confusion matrix, and metric table evaluation results, which comprised accuracy, recall, as well as F1_score. The following discussion focused on the results of the training and validation loss graphs, which were shown from several combinations of adaptive loss functions, as illustrated in figure 6 showing the Loss Per Epoch Graph (a, b, c, d).
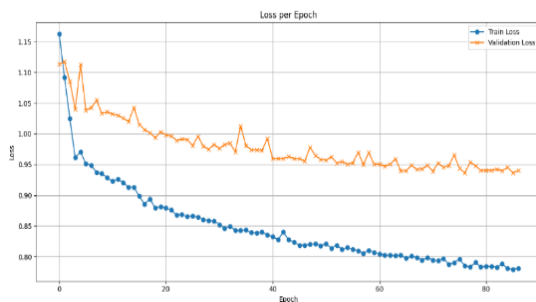
(a) U-Net Standart graph
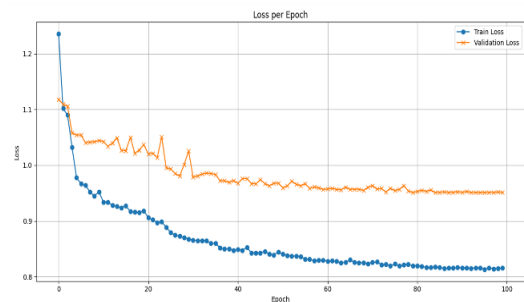
(b) U-Net with Dice and BCE loss combination graph
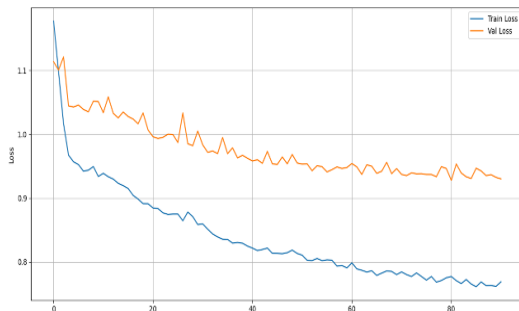
(c) U-Net with a Dice and BCE loss with Adaptive - Callbacks combination graph
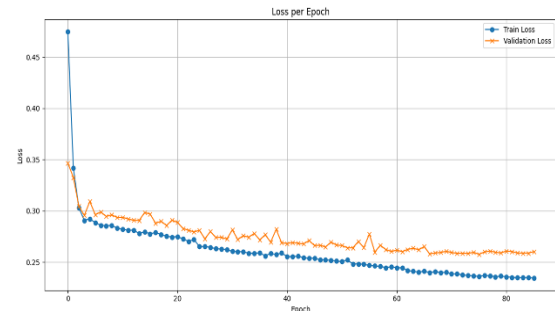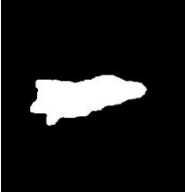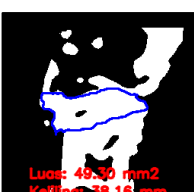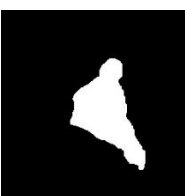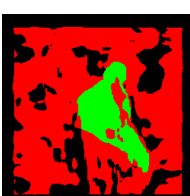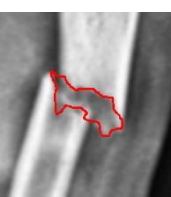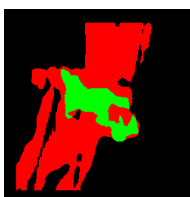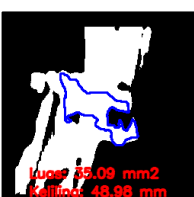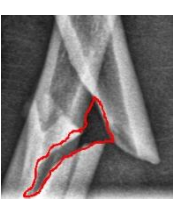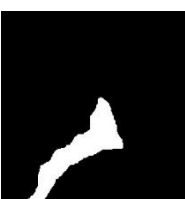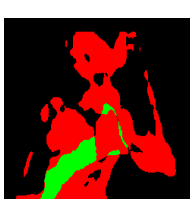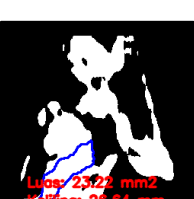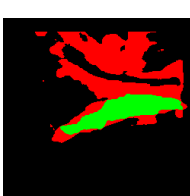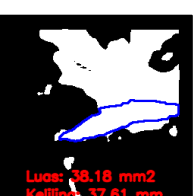
(d) U-Net with Dice dan BCE loss, adaptive linear weighted and adaptive exponential non-linear weighting algorithm graph.

Figure 6. Loss Per Epoch Graph (a,b,c,d)

4.2.1. The Following Image Showed the Results of the Visualization of Wrist Fracture Segmentation

Wrist fracture segmentation using a U-Net-based model, enhanced with adaptive training callbacks, and the Adaptive Weighted Combined Loss function, produced a segmentation visualization accompanied by bounding contours showing the fractured areas. This visualization method was shown to facilitate more accurate detection of fracture location, contributing to a more efficient diagnostic process for medical personnel. The results of the visualization sample data were shown in table 2.

Table 2. Wrist Fracture Segmentation Visualization Results

Original image of a wrist facture	Labeling	Ground truth	Binary Predictions	Overlay Prediction	ROI in the fracture area

The following table presents a comprehensive sequence of image segmentation stages for detecting wrist fractures using a medical image-based predictive modelling approach. Each column in the table signifies a pivotal element in the system workflow, ranging from the acquisition of inputs to the visual evaluation of the model's output. The initial column presents the Original Image of a Wrist Fracture, which functions as the fundamental input data employed in the fracture detection system. This image provides a visual representation of the condition of the patient's wrist bone prior to any processing.

The Labeling column illustrates the manually annotated regions created by medical experts or radiologists. At this stage, specific fracture areas are marked as references for model training and validation. The Ground Truth column contains binary mask images derived from the labeling process. These masks serve as the reference standard for evaluating the accuracy of the model's predictions, as they represent the actual location of the fracture. In the Binary Predictions column, the output of the model is shown as a binary segmentation image generated after training. The white regions in this image indicate the areas predicted by the model as potential fracture zones.

The Overlay Prediction column provides a visualization of the predicted segmentation results, which are superimposed onto the original image. The overlay's objective is to facilitate a more intuitive and informative comparison by illustrating the alignment between the predicted fracture areas and the anatomical structures. By presenting the entire

workflow in a structured and visual manner, this table offers a comprehensive overview of the segmentation system's performance and supports a detailed evaluation of the model's accuracy and effectiveness in detecting wrist fractures.

The Region of Interest (ROI) in the Fracture Area column highlights the overlap between the predicted binary segmentation and the ground truth mask. This representation enables precise localization of the fracture, offering a clearer assessment of the model's segmentation accuracy and its clinical applicability in wrist fracture detection.

4.3. Evaluation of Wrist Fracture Segmentation

The testing stage was the result of the detection of a wrist fracture by means of X-ray imaging. The 402 wrist X-ray image constituting the test set represented 30% of the total 1,340 X-ray, and the fracture image object data was labelled with a size of 256 X 256 pixels. Relating to the process, the test data had not been subject to training, and the testing process used the adaptive weighted combined loss function method in U-Net model. The evaluation of the test was based on a range of metrics, including accuracy, precision, recall, and F1-score. The results of applying the metric evaluation to the formula were shown in the confusion matrix as follows, as illustrated in figure 7 showing the confusion matrix (a, b, c, and d).

a). U-Net Standard

b). U-Net + Dice + BCE Loss

c). U-Net + Dice + BCE Loss + Callbacks for training

d). U-Net + Dice + BCE Loss +AWCLF + Callbacks for training
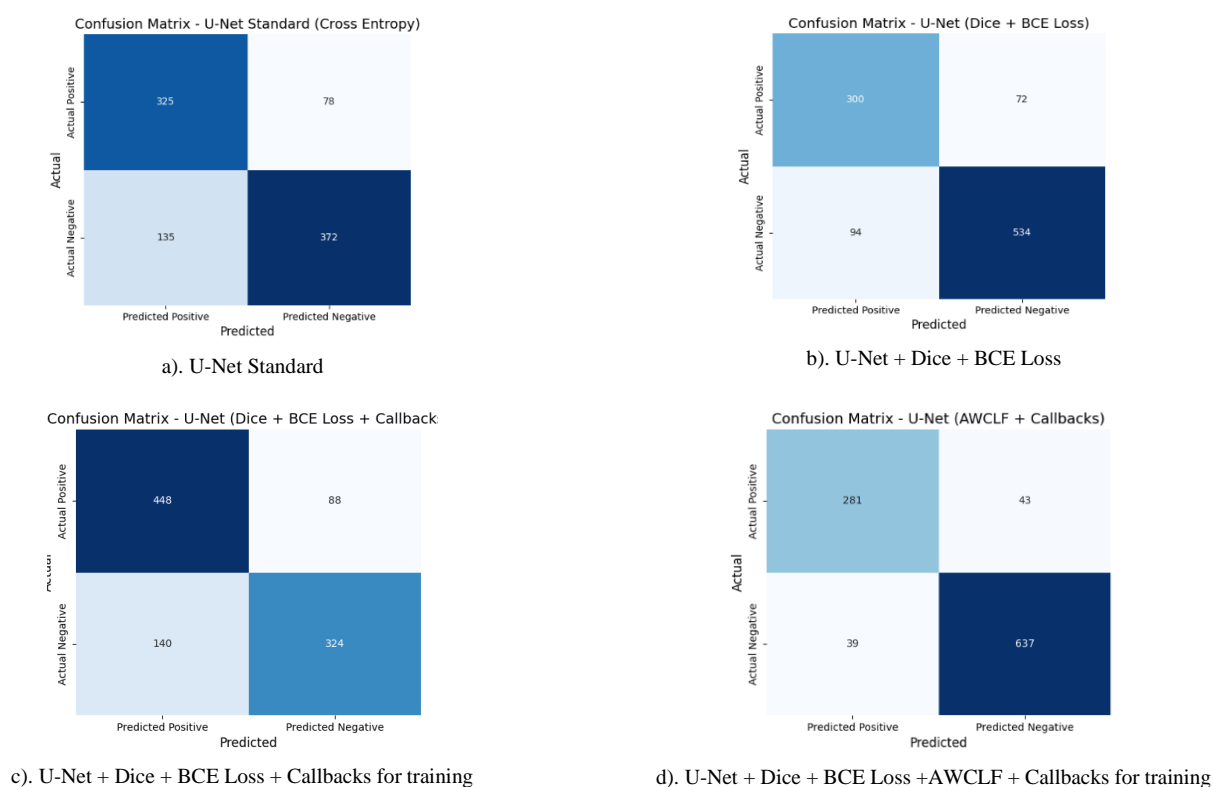
Figure 7. Confusion matrix (a,b,c and d)

The evaluation results using the confusion matrix revealed variations in performance across four U-Net model configurations for wrist fracture detection in medical images. The Standard U-Net model with Cross-Entropy Loss achieved 325 true positives (TP), 78 false negatives (FN), 135 false positives (FP), and 372 true negatives (TN). This performance indicates limited fracture detection capability, with a relatively high number of misclassifications in both FN and FP, suggesting the need for improvements in both sensitivity and specificity.

Furthermore, the U-Net with a combination of Dice Loss and BCE Loss produced 300 TP, 72 FN, 94 FP, and 534 TN. The integration of these two loss functions yielded enhanced performance, as evidenced by a notable reduction in FN and FP, reflecting improved sensitivity and overall predictive capability compared to the baseline model. The U-Net configuration combining Dice Loss, BCE Loss, and training callbacks achieved 448 TP, 88 FN, 140 FP, and 324 TN. The addition of callbacks, such as early stopping and learning rate scheduling, contributed to more stable training. This configuration demonstrated balanced performance, with a slight increase in TP, while FN values remained comparable to the previous model, indicating improved adaptability during the training process.

The final configuration, namely the U-Net with Dice Loss, BCE Loss, weighted combination (linear and non-linear exponential), and training callbacks, delivered the best performance with 281 TP, 43 FN, 39 FP, and 637 TN. This approach achieved the highest TP and TN values, along with a significant reduction in FN to only 60 cases, indicating highly accurate fracture detection. Moreover, the FP rate was the lowest among all configurations, demonstrating optimal specificity. The comparison results of U-Net improvement using adaptive callbacks and weighted loss combination for detecting wrist fracture was shown in table 3.

Table 3. U-Net Segmentation Improvement Comparison Results

Method	Dice Coefficient	Intersection over Union (IoU)	Accuracy	Precision	Recall	F1_Score
U-Net Standard - Cross Entropy	0.753	0.604	0.6966	0.7076	0.8063	0.7536
U-Net with Dice + BCE Loss combination	0.790	0.653	0.7619	0.7619	0.8063	0.7904
U-Net with a combination of Dice + BCE Loss + Callbacks for training	0.797	0.663	0.7724	0.7622	0.8363	0.7972
U-Net with Dice combination + BCE Loss + Weighted Combination (Linear and Non-linear exponential) + Callbacks for training	**0.871**	**0.772**	**0.9178**	**0.8778**	**0.8670**	**0.8717**

The research showed the efficacy of the linear and non-linear exponential weighting methods in Dice and BCE Loss in addressing the imbalance between the number of pixels representing the background and the object of interest in wrist fracture. This was achieved by adjusting the weights for each class, ensuring the model focused more intently on the minority class with high accuracy.

4.4. Confidence Interval Testing

Interval testing is a statistical technique used to estimate the range within which a true value of a parameter lies, based on sample observations. In the context of wrist fracture detection, it allows for the assessment of variability in segmentation results, such as the area and perimeter of the fracture, derived from automated image analysis. This approach helps quantify uncertainty, as illustrated in figure 8 showing the boxplot comparison of Standard U-Net and Weighted U-Net across performance metrics, enhances the reliability of the findings, and supports consistent comparison across different segmentation models or clinical cases. Employing confidence intervals ensures that conclusions drawn from the model are statistically valid and clinically interpretable.
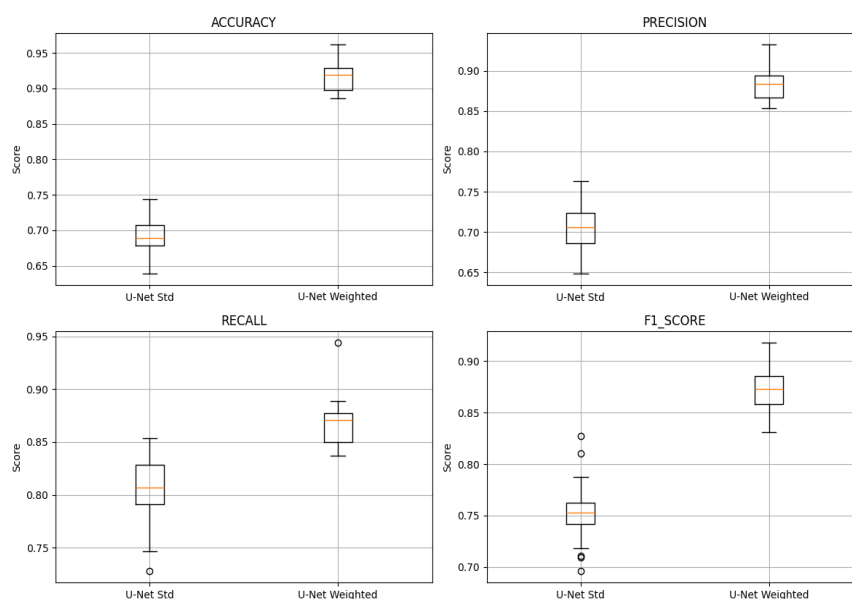

Figure 8. Boxplot Comparison of Standard U-Net and Weighted U-Net Across Performance Metrics

Figure 8 illustrates the boxplots for four key evaluation metrics: accuracy, precision, recall, and F1-score, comparing the standard U-Net model and the proposed weighted U-Net model. Each boxplot represents the distribution of 30 simulated values per metric, based on their respective means and standard deviations.

The weighted U-Net consistently outperforms the standard model across all metrics, as indicated by higher medians and tighter interquartile ranges. This suggests not only improved performance but also lower variability, highlighting the model's robustness. The statistical significance of these differences is supported by independent t-tests, with p-values below 0.05, confirming that the improvements are not due to random chance. This visual analysis reinforces the efficacy of the adaptive weighting strategy in enhancing segmentation performance for wrist fracture detection.

5. Conclusion

In conclusion, the enhancement of the U-Net model's capability for segmenting fractures in wrist X-ray images was achieved through the implementation of an adaptive training callback method combined with the development of adaptive loss functions. These functions incorporated both Dice and BCE loss with linear and non-linear exponential weighting strategies. The dataset consisted of 1,344 wrist X-ray images, both normal and fractured, with 70 percent (644 images) allocated for training and the remaining 30 percent (276 images) used for testing. The parameters applied during experimentation included an input size of 256 by 256 pixels, a maximum of 100 epochs, a batch size of 8, an initial alpha value of 0.5 (alpha_start), a final alpha value of 1.0 (alpha_end), a beta_start value of 0.5, a beta_end value of 1.0, and a lambda parameter set to 0.1. The model achieved an accuracy of 0.9178, a precision of 0.8778, a recall of 0.8670, and an F1-score of 0.8717. Throughout the training process, the loss value served as an essential metric indicating how well the model predicted outcomes, with lower loss values being more favorable. The validation accuracy ('val_accuracy') represented the model's performance on unseen data, while the validation loss ('val_loss') indicated its prediction error during validation. A lower value in validation loss signified better generalization. To enhance the precision of the U-Net model in segmenting wrist bone fractures from X-ray images, several technical strategies were employed, including data quality enhancement, callback adjustments, adaptive loss function design, training optimization, and comprehensive validation procedures.

6. Declarations

6.1. Author Contributions

Conceptualization: T.R., S.D., and G.W.N.; Methodology: S.D.; Software: T.R.; Validation: T.R., S.D., and G.W.N.; Formal Analysis: T.R., S.D., and T.R.; Investigation: T.R.; Resources: S.D.; Data Curation: S.D.; Writing Original Draft Preparation: T.R., S.D., and G.W.N.; Writing Review and Editing: S.D., T.R., and G.W.N.; Visualization: T.R.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] J. Li, H. J. Shan, and X. W. Yu, "Fracture detection of distal radius using deep- learning-based dual-channel feature fusion algorithm," *Chinese Journal of Traumatology - English Edition*, vol. 2025, no. Mar., pp. 1-13, 2025, doi: 10.1016/j.cjtee.2024.10.006.

[2] J. R. Zech, "Detecting pediatric wrist fractures using deep-learning-based object detection," *Pediatr Radiol*, vol. 53, no. 6, pp. 1125–1134, May 2023, doi: 10.1007/s00247-023-05588-8.

[3] S. Turk, O. Bingol, A. Coskuncay, and T. Aydin, "The impact of implementing backbone architectures on fracture segmentation in X-ray images," *Engineering Science and Technology, an International Journal*, vol. 59, no. Nov., pp. 1-20, 2024, doi: 10.1016/j.jestch.2024.101883.

[4] M. Widiasri, N. Suciati, C. Fatichah, E. R. Astuti, R. H. Putra, and A. Z. Arifin, "Alveolar Bone and Mandibular Canal Segmentation on Cone Beam Computed Tomography Images Using U-Net," *Proceedings of the 2023 International Conference on Instrumentation, Control, and Automation, ICA 2023,* vol. 2023, no. 008, pp. 36–41, 2023, doi: 10.1109/ICA58538.2023.10273108.

[5] J. Du, K. Guan, P. Liu, Y. Li, and T. Wang, "Boundary-Sensitive Loss Function With Location Constraint for Hard Region Segmentation," *IEEE J Biomed Health Inform*, vol. 27, no. 2, pp. 992–1003, 2023, doi: 10.1109/JBHI.2022.3222390.

[6] S. E. Roshan, J. Tanha, M. Zarrin, A. F. Babaei, H. Nikkhah, and Z. Jafari, "A deep ensemble medical image segmentation with novel sampling method and loss function," *Comput Biol Med*, vol. 172, no. Apr., pp. 1-12, 2024, doi: 10.1016/j.compbiomed.2024.108305.

[7] S. H. Choi, "Optimizing Deep Learning Models with Hybrid Nonlinear Loss Functions: Integrating Heterogeneous Nonlinear Properties for Enhanced Performance," *International Journal of Fuzzy Logic and Intelligent Systems*, vol. 24, no. 4, pp. 317–332, Dec. 2024, doi: 10.5391/IJFIS.2024.24.4.317.

[8] A. R. Hareendranathan, A. Tripathi, M. R. Panicker, J. Zhang, N. Boora, and J. Jaremko, "Deep Learning Approach for Automatic Wrist Fracture Detection Using Ultrasound Bone Probability Maps," *SN Compr Clin Med*, vol. 5, no. 1, pp. 1-13, Nov. 2023, doi: 10.1007/s42399-023-01608-8.

[9] A. Khanal, R. Rizk, and K. C. Santosh, "Ensemble Deep Convolutional Neural Network to Identify Fractured Limbs using CT Scans," in Proceedings - 2023 *IEEE Conference on Artificial Intelligence, CAI 2023, Institute of Electrical and Electronics Engineers Inc.,* vol. 2023, no. 1, pp. 156–157, 2023. doi: 10.1109/CAI54212.2023.00075.

[10] R. Dhanalakshmi, M. Thenmozhi, S. Saxena, and H. Mahalingam, "Convolutional Neural Network Model based Deep Learning Approach for Osteoporosis Fracture Detection," *in Proceedings of the 1st IEEE International Conference on Networking and Communications 2023, ICNWC 2023*, vol. 2023, no. 1, pp. 1-8, 2023. doi: 10.1109/ICNWC57852.2023.10127367.

[11] M. R, Vadhera, Sharma, "A novel hybrid loss-based Encoder–Decoder model for accurate Pulmonary Embolism segmentation," *Int. j. inf. tecnol*, vol. 2025, no. 1, pp. 1-12, 2025.

[12] M. Ali et al., "Segmentation of MRI tumors and pelvic anatomy via cGAN-synthesized data and attention-enhanced U-Net," *Pattern Recognit Lett*, vol. 187, no. November 2024, pp. 100–106, 2025, doi: 10.1016/j.patrec.2024.11.003.

[13] J. Sun, Y. Li, X. Wu, C. Tang, S. Wang, and Y. Zhang, "HAD-Net: An attention U-based network with hyper-scale shifted aggregating and max-diagonal sampling for medical image segmentation," *Computer Vision and Image Understanding*, vol. 249, no. September, pp. 1-21, 2024, doi: 10.1016/j.cviu.2024.104151.

[14] Q. Ming and X. Xiao, "Towards Accurate Medical Image Segmentation with Gradient-Optimized Dice Loss," *IEEE Signal Process Lett*, vol. 31, no. 1, pp. 191–195, 2024, doi: 10.1109/LSP.2023.3329437.

[15] E. Bosco, G. Magenes, and G. Matrone, "Echocardiographic Image Segmentation with Vision Transformers: A Comparative Analysis of Different Loss Functions," *2024 IEEE International Symposium on Medical Measurements and Applications*, MeMeA 2024 - Proceedings, vol. 2024, no. 1, pp. 1-8, 2024, doi: 10.1109/MeMeA60663.2024.10596874.

[16] R. Maurya, P. Thirwarni, T. Gopalakrishnan, and M. Karnati, "Combining Focal loss with Cross-entropy loss for Pneumonia Classification with a Weighted Sampling Approach," *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI),* vol. 2024, no. 1, pp. 1-5, 2024. doi: 10.1109/IATMSI60426.2024.10502684.

[17] F. Piri, N. Karimi, and S. Samavi, "Enhanced Segmentation in Abdominal CT Images: Leveraging Hybrid CNN-Transformer Architectures and Compound Loss Function," *2024 IEEE 5th World AI IoT Congress, AIIoT 2024*, vol. 2024, no. 1, pp. 363–369, 2024, doi: 10.1109/AIIoT61789.2024.10579036.

[18] L. Zeng and S. Guo, "Automatic Liver Tumor Segmentation from CT Images with Dual Attention Mechanism," *2024 9th International Conference on Electronic Technology and Information Science, ICETIS 2024,* vol. 2024, no. 1, pp. 477–480, 2024, doi: 10.1109/ICETIS61828.2024.10593773.

[19] Z. Song and L. Wang, "Robust Image Segmentation Algorithm Based on Local Correntropy-Based Generalized Loss Function," *2024 7th International Conference on Advanced Algorithms and Control Engineering, ICAACE 2024*, vol. 7, no. 1, pp. 12–16, 2024, doi: 10.1109/ICAACE61206.2024.10548610.

[20] Y. Song, J. Y. C. Teoh, K. S. Choi, and J. Qin, "Dynamic Loss Weighting for Multiorgan Segmentation in Medical Images," *IEEE Trans Neural Netw Learn Syst,* vol. 2023, no. 1, pp. 1-12, 2023, doi: 10.1109/TNNLS.2023.3243241.

[21] M. Geijer, E. Gunnlaugsson, L. Arvidsson, E. Österhed, and M. Tägil, "Outcome of follow - up computed tomography of suspected occult scaphoid fracture after normal radiography," *Emerg Radiol*, vol. 2024, no. 1, pp. 1-12, 2024, doi: 10.1007/s10140-024-02307-0.

[22] J. Sawant and A. D. Vibhute, "U-Net-based blood vessel segmentation using an improved data augmentation with green channel images," *Procedia Comput Sci*, vol. 260, no. 1, pp. 1071–1079, 2025, doi: 10.1016/j.procs.2025.03.292.

[23] Y. Zheng, "Adaptive boundary-enhanced Dice loss for image segmentation," *Biomed Signal Process Control*, vol. 106, no. Aug., pp. 1-12, Aug. 2025, doi: 10.1016/j.bspc.2025.107741.

[24] Y. Xu, "Automatic segmentation of prostate cancer metastases in PSMA PET/CT images using deep neural networks with weighted batch-wise dice loss," *Comput Biol Med*, vol. 158, no. May, pp. 1-9, May 2023, doi: 10.1016/j.compbiomed.2023.106882.

[25] Z. Zhou, L. Cai, P. Yin, X. Qian, Y. Dai, and Z. Zhou, "Narrow-band loss - A novel loss function focused on target boundary," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, vol. 2023, no. 1, pp. 1–4, 2023, doi: 10.1109/EMBC40787.2023.10340038.

[26] X. X. Ming Qi, "Toward Accurate Medical Image Segmentation With Gradien-Optimized Dice Loss," *IEEE Signal Process Lett*, vol. 31, no. 1, pp. 191–195, 2024, doi: DOI: 10.1109/LSP.2023.3329437.

[27] S. Kato and K. Hotta, "Adaptive t-vMF dice loss: An effective expansion of dice loss for medical image segmentation," *Comput Biol Med*, vol. 168, no. Jan., pp. 1-12, Jan. 2024, doi: 10.1016/j.compbiomed.2023.107695.

[28] S. Koles, S. Karakas, A. P. Ndigande, and S. Ozer, "Using Different Loss Functions with YOLACT++ for Real-Time Instance Segmentation," *2023 46th International Conference on Telecommunications and Signal Processing, TSP 2023*, vol. 46, no. 1, pp. 264–267, 2023, doi: 10.1109/TSP59544.2023.10197832.

[29] R. R, S. S, N. V. Kumar, R. S, and P. B V, "Integrated pixel-level crack detection and quantification using an ensemble of advanced U-Net architectures," *Results in Engineering*, vol. 25, no. Mar., pp. 1-12, 2025, doi: 10.1016/j.rineng.2024.103726.

[30] M. Attari, N. P. Nguyen, K. Palaniappan, and F. Bunyak, "Multi-Loss Topology-Aware Deep Learning Network for Segmentation of Vessels in Microscopy Images," *Proceedings - Applied Imagery Pattern Recognition Workshop*, vol. 2023, no. 1, pp. 1-8, 2023, doi: 10.1109/AIPR60534.2023.10440665.

[31] B. Gao, R. Yao, and Y. Li, "Physics-informed neural networks with adaptive loss weighting algorithm for solving partial differential equations," *Computers and Mathematics with Applications*, vol. 181, no. Mar., pp. 216–227, Mar. 2025, doi: 10.1016/j.camwa.2025.01.007.

[32] A. Akagic, M. Kapo, E. Kandić, M. Bećirović, and N. Kadrić, "Brain Tumor Segmentation of MRI Images with U-Net and DeepLabV3+," *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI),* vol. 2024, no. 1, pp. 1–6, 2024, doi: 10.1109/icmi60790.2024.10585749.

[33] M. Obayya, "A novel U-net model for brain tumor segmentation from MRI images," *Alexandria Engineering Journal*, vol. 126, no. 1, pp. 220–230, Jul. 2025, doi: 10.1016/j.aej.2025.04.051.

[34] R. Joshi, "Segmentation of Teeth in Panoramic X-ray Image Using U-net Algorithm," *2024 3rd International Conference on Artificial Intelligence for Internet of Things, AIIoT 2024,* vol. 3, no. *AIIoT*, pp. 1–6, 2024, doi: 10.1109/AIIoT58432.2024.10574786.

[35] R. Karthick Manoj, S. Aasha Nandhini, and M. Batumalay, "Automated Brain Tumor Analysis with Multimodal Fusion and Augmented Intelligence," *Journal of Applied Data Sciences*, vol. 6, no. 2, pp. 1277–1290, May 2025, doi: 10.47738/jads.v6i2.719.

[36] B. Garg, D. C. Lohani, and B. Rana, "A Comparative Study of U-Net-based Segmentation Architectures for Crack Detection," in *Procedia Computer Science, Elsevier B.V.*, vol. 2025, no. 1, pp. 3305–3314. doi: 10.1016/j.procs.2025.04.588.