

RESEARCH ARTICLE

A Unified Approach Addressing Class Imbalance in Magnetic Resonance Image for Deep Learning Models

LIJUAN CUI^{1,2}, DENGGAO LI^{1,2}, XIAOFENG YANG³, CHAO LIU⁴, AND XIAOTING YAN³¹College of Computer Science and Technology (College of Data Science), Taiyuan University of Technology, Taiyuan 030024, China²Key Laboratory of Big Data Fusion Analysis and Application of Shanxi Province, Taiyuan 030024, China³Department of Urology, First Hospital of Shanxi Medical University, Taiyuan, Shanxi 030001, China⁴Department of Bioengineering Research Center, First Hospital of Shanxi Medical University, Shanxi Medical University, Taiyuan, Shanxi 030001, China

Corresponding authors: Lijuan Cui (cuilijuan0171@link.tyut.edu.cn) and Denggao Li (lidenggao@tyut.edu.cn)

This work was supported in part by the Central Government Guided Local Science and Technology Development Fund Project under Grant YDZJSX2022A011, in part by the National Natural Science Foundation of China under Grant 62076177, in part by the National Major Scientific Research Instrument Development Project of China under Grant 62027819, in part by the Key Research and Development Projects of Shanxi Province under Grant 202102020101004 and Grant 202102020101006, in part by the Central Guidance on Local Science and Technology Development Fund of Shanxi Province under Grant YDZJSX2021C010, and in part by the Patents Transformation Project of Shanxi Province under Grant 202304008.

ABSTRACT Medical image datasets, particularly those comprising Magnetic Resonance (MR) images, are essential for accurate diagnosis and treatment planning. However, these datasets often suffer from class imbalance, where certain classes of abnormalities have unequal representation. Models trained on imbalanced datasets can be biased towards the prominent class, leading to misclassification. Addressing class imbalance problems is crucial to developing robust deep-learning MR image analysis models. This research focuses on the class imbalance problem in MR image datasets and proposes a novel approach to enhance deep learning models. We have introduced a unified approach equipped with a selective attention mechanism, unified loss function, and progressive resizing. The selective attention strategy identifies prominent regions within the underlying image to find the feature maps, retaining only the relevant activations of the minority class. Fine-tuning of the multiple hyperparameters was achieved using a novel unified loss function that plays a vital role in enhancing the overwhelming error performance for minority classes and accuracy for common classes. To address the class imbalances phenomenon, we incorporate progressive resizing that can dynamically adjust the input image size as the model trains. This dynamic nature helps handle class imbalances and improve overall performance. The research evaluates the effectiveness of the proposed approach by embedding it into five state-of-the-art CNN models: UNet, FCN, RCNN, SegNet, and DeeplabV3. For experimental purposes, we have selected five diverse MR image datasets, BUS2017, MICCAI 2015 head and neck, ATLAS, BRATS 2015, and Digital Database Thyroid Image (DDTI), to evaluate the performance of the proposed approach against state-of-the-art techniques. The assessment of the proposed approach reveals improved performance across all metrics for five different MR imaging datasets. DeepLabV3 demonstrated the best performance, achieving IoU, DSC, Precision, and Recall scores of 0.893, 0.953, 0.943, and 0.944, respectively, on the BUS dataset. These scores indicate an improvement of 5% in DSC, 6% in IoU, 4% in precision, and approximately 4% in recall compared to the baseline. The most significant increases were observed in the ATLAS and LiTS MICCAI 2017 datasets, with a 5% and 7% increase in IoU and DSC over the baseline (DSC = 0.628, DSC = 0.695) for the ATLAS dataset and a 5% and 9% increase in IoU and DSC for the LiTS MICCAI 2017 dataset.

The associate editor coordinating the review of this manuscript and approving it for publication was Essam A. Rashed¹.

INDEX TERMS Class imbalance, medical image analysis, diagnostic accuracy, MRI images, deep learning model.

I. INTRODUCTION

Medical image analysis is a rapidly growing field with the potential for diagnostic process automation for various diseases [1]. In recent years, deep learning models have shown remarkable achievement in analyzing medical images, including Magnetic Resonance (MR) images [2], [3], [4], [5], [6]. Deep learning-based CNN models have emerged as a powerful tool for analyzing medical images at the pixel-level. These models can potentially assist healthcare professionals in accurately diagnosing, planning treatment, and monitoring patients. However, one major challenge in developing effective deep-learning models for medical image analysis is the presence of a class imbalance in the datasets [1], [7], [8], [9], [10], [11]. Class imbalance in a medical image dataset refers to a situation where the number of images belonging to different classes or categories is significantly unequal [8].

Class imbalance occurs when specific categories of abnormalities in medical image datasets are prominent compared to the other classes. The class imbalance ratio for an image is determined by comparing the number of pixels in the background class (typically the most prevalent class) to the number of pixels in various object classes. The class imbalance ratio is presented as the average ratio across all its images for an entire dataset.

In the current era of big data, the abrupt progression in medical imaging data has led to the emergence of imbalanced datasets. The raw and imbalanced data is a significant challenge in developing effective models, especially deep learning models, for accurate classification and segmentation tasks. Ideally, these models should improve accuracy when dealing with positive and negative examples [12], [13]. However, previous research has demonstrated that class imbalance adversely affects the performance of commonly used models, including decision trees, support vector machines, artificial neural networks, and others [11], [14], [15], [16], [17]. As a result, deep learning models trained on imbalanced datasets tend to be biased towards the majority class and ignore the minority class, which leads to poor classification, detection and segmentation accuracy. Class imbalance issues in medical image analysis carry several challenges and problems. These challenges include:

Bias towards majority classes: Deep learning models trained on imbalanced datasets tend to prioritize accurately predicting the majority class while neglecting the minority classes. This bias may decrease sensitivity and accuracy during the diagnosis of medical abnormalities or conditions, which are crucial for accurate diagnosis.

Limited representation of minority classes: Imbalanced datasets have fewer examples of minority classes, making it difficult for models to learn their distinguishing features effectively. As a result, the models struggle to generalize and correctly classify instances belonging to the underrepresented classes.

Decreased predictive performance: Imbalanced datasets negatively impact deep learning models' overall prediction and segmentation performance. The lack of sufficient samples from minority classes hampers the models' ability to learn their characteristics accurately, leading to lower precision, recall, and F1 scores, which are essential for reliable diagnostic results.

Difficulty setting decision thresholds: Class imbalance affects the optimal decision thresholds for model predictions. Determining a threshold that balances sensitivity and specificity becomes challenging with a bias towards the majority class. Depending on the chosen threshold, this can increase false positives or negatives.

Limited generalizability: Deep learning models trained on imbalanced datasets tries to generalize well to unseen data, particularly for underrepresented classes. This limitation hampers the deployment of such models in real-world clinical settings where accurate detection and segmentation of various abnormalities are crucial.

Addressing the challenges related to class imbalance is essential for developing reliable and effective deep-learning models in medical image analysis. By mitigating class imbalance issues, one can improve diagnostic accuracy, enhance the identification of rare abnormalities, and achieve better performance in clinical applications.

Advancements in deep convolutional neural networks have led to various architectures for image segmentation, including GoogLeNet [18], ResNet [19], and SegNet [20]. While these models excel with natural images across diverse applications, they often underperform in medical image analysis due to their design for larger datasets.

Recently, researchers have proposed several techniques to address the class imbalance in medical image datasets [21], [22], [23], [24], [25], [26], [27], [28], [29]. These techniques try to enhance the representation of the minority class and reduce the representation of the majority class. Dice and cross entropy-based losses have been generalized by [17] to tackle the class imbalance in the segmentation of medical images. Adaptive Blended Consistency Loss (ABCL) has been introduced by [30] using a Semi-supervised learning approach to address the imbalanced data during the segmentation of retinal fundus glaucoma.

The consequences of class imbalance for deep learning models in medical image datasets are particularly critical, as misclassifying a rare but severe condition can have severe implications for patient outcomes. In scenarios where the focus is primarily on the majority class, the model may struggle to recognize and accurately predict instances of less common diseases, leading to delayed or missed diagnoses [31]. Imbalanced data can negatively affect the performance of models significantly. Many models that perform well on balanced datasets cannot achieve good performances when it comes to their imbalanced counterparts [32]. Addressing

class imbalance in medical image datasets is crucial for building reliable and clinically applicable deep learning models.

Previous research has shown satisfactory performance on certain datasets but has struggled to generalize effectively when applied to other medical image datasets. For instance, [1] designed DC-CNN, a two-stage deep learning framework that addresses class imbalance. It efficiently detects small lesions but struggles with larger lesions like retinal hemorrhage and mammography. Several research works are available for tackling class imbalance in medical images, but limited work is dedicated to MR images [7].

In this research, we have focused on the class imbalance in MR image datasets to enhance the performance and reliability of deep learning models. We propose a novel unified approach that combines selective attention strategy, unified loss function, and progressive resizing. The selective attention strategy uses the coefficient to identify the substantial regions in the underlying image to prune the feature responses, keeping only the activations relevant to the minority class. The unified loss function is used for fine-tuning of multiple hyperparameters and helps to enhance the suppressing error performance for minority classes and accuracy for common classes. Additionally, progressive resizing is used during the model's training to handle class imbalances. This unified approach is embedded into five state-of-the-art CNN models, i.e., UNet, FCN, RCNN, SegNet, and Deeplab-V3, to enhance their performance and analyze the impact of the proposed technique. We have conducted extensive experiments on diverse MR image datasets like BUS2017 [33], MICCAI 2015 head and neck [34], ATLAS [35], BRATS 2015 [36], and Digital Database Thyroid Image (DDTI) [37] to evaluate the performance of the proposed approach against state-of-the-art techniques.

The key contributions of this research work are as follows:

- We have designed a unified loss function that combines cross entropy loss and dice loss functions to address class imbalance in MRI datasets. This innovation contributes to more robust and accurate deep-learning models for MRI segmentation tasks, which can positively impact clinical diagnoses and treatment planning.
- We have introduced a novel selective attention mechanism capable of learning salient features while effectively suppressing the background. This innovation is especially advantageous in medical imaging, where precise identification of relevant structures is critical. By integrating this mechanism, we aim to improve the interpretability of deep learning models, making them more reliable in highlighting diagnostically relevant features in medical images.
- We have employed progressive resizing and transfer learning strategy to enhance feature diversity, generalizability and training efficiency. The impact of this contribution extends to improved model performance across different datasets, promoting the development

of more versatile and widely applicable deep learning models for medical image analysis.

- We have evaluated the impact of each contribution on the performance of deep learning models and compared it with existing methods.

II. RELATED WORK

Recently, several studies have focused on the class and data imbalance problem for segmentation and classification using various deep-learning models. This section summarizes the most relevant works in the field, highlighting their contributions, strengths, and limitations.

In [25], the authors target the class imbalance issue by rebalancing medical data using three methods. They combine the resampling methods, i.e., 1) synthetic minority oversampling technique and undersampling, 2) particle swarm optimization (PSO), and 3) MetaCost, and perform two experiments on nine different medical datasets. The outcome of these experiments reveals that the dataset with an Imbalance ratio of >9 must follow the undersampling for better decisions. While in the case of a ratio <9 , the model must consider the synthetic minority oversampling and undersampling techniques simultaneously for better classification. In [38] and [39], deep learning models have been applied to compare the evaluation metrics of osteoarthritis images and imbalances in medical image classification, respectively. Balanced Active Learning (BAL) was proposed by [39] to find the probability of majority and minority class samples in the dataset. They performed experiments on imbalanced CIFAR-10, ISIC2020, and Caltech256 datasets. Another study was done by [40], [41], and [42] to investigate the imbalance problem in nuclei data of histopathological images and COVID-19 images, respectively. In [40], the authors proposed an imbalance-aware nuclei segmentation model using enhanced lightweight U-Net architecture. The proposed model was evaluated using the Aggregated Jaccard Index (AJI) and Intersection of Union (IoU) metrics. While [41] work on the loss- and class imbalance-aware aggregation using federated learning. The proposed context Aggregator federated learning model was tested on the COVID-19 imaging dataset and gave better results than the standard federating average learning algorithms. To handle the imbalanced data in Raman spectroscopy, [22] proposed a hybrid sampling method of Raman-Gaussian distributed oversampling attached with random undersampling. The proposed method was applied to the dataset of malignant tumors, class B infectious diseases, and autoimmune diseases.

Analysis of Oversampling and under-sampling data distribution was performed by [43] using a semi-supervised hierarchical clustering algorithm (SSHC). SSHC model trained on labeled data that guide the clustering procedure on the whole dataset. The linear-exponential loss combined with the deep learning models to develop an asymmetric geometry interpretation model known as DLINEX [44]. This model was designed to pay more attention to the minority and hard-to-classify classes by uniquely adjusting

one parameter. DLINEX model was tested on CIFAR-10, STARE, CHASEDB1 and HAM10000 datasets for imbalance data issues. Anxiety detection is a major problem for classification using EEG in the presence of imbalanced data. Safe-level Synthetic Minority Oversampling Technique and CNN with Long Short-Term Memory Network (CNN-LSTM) were developed by [45], [46], and [47]. They embedded the KNN and SVM with their model and achieved 89.5% accuracy and the highest precision of 89.7% with enhanced class modalities. Transfer learning and active sampling techniques were used by [48] for handling imbalanced data problems in classification. The proposed transfer learning model includes three modules: 1) active sampling module, 2) real-time data augmentation module, and 3) DenseNet module. Imbalanced chest X-ray data was handled by [49] and [50]. In [49], Initially, the authors generated a heatmap of those areas of the images that are more relevant for classification. After that, they used EfficientNetB0 [51], DenseNet-201 [52], InceptionV3 [53], InceptionResNetV2 and Xception to classify enhanced data. Meanwhile, [50] used the CheXpert dataset to diagnose heart failure using multi-level classification and targeting data. Imbalancing situations with 84.44% accuracy. Previous studies reveal that classifier performance on imbalanced data mainly relies on the object's borderline within an image. In [54], the authors proposed a 0-order Takagi-Sugeno-Kang Fuzzy System (0-TSK-FS) system to accurately detect borderline. The 0-TSK-FS system outperforms classification performance and reasonable interpretability on imbalanced datasets. Another valuable work was carried out by [55] for the multi-classification of imbalanced data using a hierarchical belief rule-based model. The authors used the model's multiple belief rule base (BRB) systems, categorized as main-BRB and sub-BRB. During classification, the out of the main-BRB represents the approximated classification between confusable classes, while the XGBoost technique was used for feature selection. Class imbalance in graph data was handled by [56] and [57] using graph neural network node classification. They used the GNN-based Imbalanced Node Classification (GNN-INCM) Model to overcome class imbalance distribution. The GNN-INCM is equipped with two supportive modules, i.e., 1) Embedding Clustering-based Optimization (ECO) and 2) Graph Reconstruction-based Optimization (GRO). In [58], the authors try to overcome the imbalance of data distribution for glaucoma diagnosis using an adaptive rebalancing strategy in the feature space and Self-Ensemble Dual-Curriculum learning (SEDC). The proposed model pays attention to the minority samples and generates its augmentation to generate extra features that help to increase the minority class as equivalent to the majority class. The unified focal loss was introduced by [17] to handle the class imbalance in the segmentation of medical images by generalizing dice and cross-entropy losses. The authors in [59] and [60] produce generative adversarial networks (GAN) with classification enhancement and multibranch discriminator

respectively, to handle the imbalance data during classification. Imbalanced ultrasound image modalities was used by [61] to diagnose the breast cancer using Doubly supervised parameter transfer classifier.

Addressing class imbalance primarily involves adjusting either the training or input data sampling processes, with infrequent consideration given to adapting the loss function. However, commonly used methods like upsampling the underrepresented class inherently lead to an increase in false positive predictions. Moreover, intricate, often multi-stage training processes demand greater computational resources. Presently, two widely employed attention mechanisms are the weighted and self-attention mechanisms [62], [63], [64], [65], [66], [67]. The weighted attention technique entails globally squeezing the channel or spatial dimension to precisely enhance effective features while suppressing unnecessary ones. However, these methods lack selective induction, causing the network to prioritize the most globally salient effective features, neglecting secondary features with slightly lower weight values but equal importance. Similarly, the self-attention mechanism comes with the drawback of excessive computational overhead. In the context of medical images, achieving a fine-grained segmentation of regions of interest is crucial to prevent missed diagnoses. Consequently, both secondary salient features and the globally most salient features hold equal significance in the medical image segmentation process. Beyond architectural solutions, improvements in objective functions are crucial due to their direct impact on the model's learning process. The goal is to enhance model performance by enabling the loss function to penalize training parameters more for false classification compared to true classification, thereby promoting efficient learning of desired features. To address these challenges, we are actively developing a comprehensive strategy encompassing selective attention, an enhanced objective function named the Unified Loss Function, progressive resizing, and transfer learning. This approach aims to overcome the limitations of existing techniques, ensuring a more nuanced and effective approach to feature enhancement and suppression in the context of medical image analysis.

III. RESEARCH METHODOLOGY

Overview

Deep learning models have consistently delivered remarkable outcomes; however, they often grapple with the challenge of class imbalance. This issue is particularly pronounced when these models depend on large datasets, which are frequently limited in specialized fields such as medical imaging. In addition, the data that is available in this area is usually skewed, leading to suboptimal performance of deep learning algorithms.

To overcome this, we have devised a hybrid approach that integrates smoothly with state-of-the-art deep learning architectures. This solution effectively reduces the impact of class imbalance on the performance of these models, with a special

focus on MRI datasets. Our innovative strategy introduces a selective attention mechanism that hones in on essential features, while a focal parameter reduces the emphasis on less significant background regions. We have crafted a novel loss function, which we refer to as the unified loss function. This function combines the principles of cross-entropy loss and dice loss, and, by capitalizing on asymmetry, it allocates varying weights to different classes.

The proposed approach handles class imbalance in three main steps: 1) Progressive resizing as a preprocessing step, 2) a selective attention mechanism to prioritize regions of interest (ROIs), and 3) Application of a robust Unified loss function to handle the disparity between dominant and minority classes. To Present the effectiveness of this scheme, we have integrated it into various state-of-the-art deep-learning models for image segmentation, such as UNet, SegNet, RCNN, FCN, and DeepLab-V3. The detail description of each step is in the following.

A. PROGRESSIVE RESIZING

We have employed progressive resizing in the training of deep learning networks to tackle the problem of class imbalancing, specifically in the context of medical image segmentation as shown in Figure 1. During the training process, it involves sequentially resizing input images from smaller to larger sizes. We utilized this resizing technique in our approach by initially training the model on images of size $225 \times 225 \times 3$. Subsequently, in the second iteration, we trained the model on images of size $256 \times 256 \times 3$, followed by training on images of size $512 \times 512 \times 3$ in the third iteration. In each iteration, we incorporated the layers and weights from the previous small-scale model into the architecture of the current iteration. This strategy enables the model to address class imbalance, a common challenge in deep neural networks applied to medical image segmentation tasks.

B. SELECTIVE ATTENTION STRATEGY

The Selective Attention Suppression (SAS) progressively diminishes feature responses in irrelevant background regions without the need to crop an ROI between networks. This offers a promising avenue for improving CNNs' robustness, reducing model bias towards majority classes, and enhancing generalizability in medical image analysis. We have incorporated a selective attention strategy into networks. Let's consider A as the activation map of a selected layer l , denoted as $A = \{F_i^l\}_{i=1}^J$. Each F_i^l represents a feature vector at the pixel level, having a length of v_L (which corresponds to the number of channels). An Attention Gate (AG) is employed to calculate coefficients $a^l = \{a_i^l\}_1^n$, for each X_i^l , where a_i^l ranges from 0 to 1. These coefficients aim to identify significant regions within the image and prune the feature responses, keeping only the activations relevant to the specific task. The resulting output of the Attention Gate is $\hat{F} = \{a_i^l f_i^l\}_1^n$ for $i = 1$ to n , where each feature vector a_i^l is scaled by its corresponding attention coefficient.

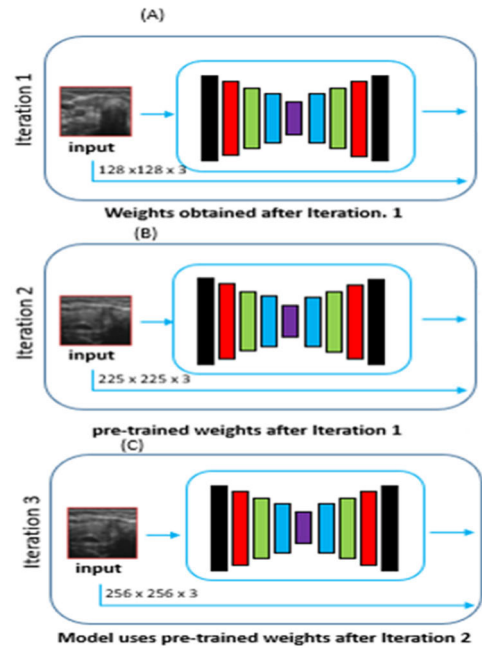


FIGURE 1. Presents overview of transfer learning and progressive resizing.

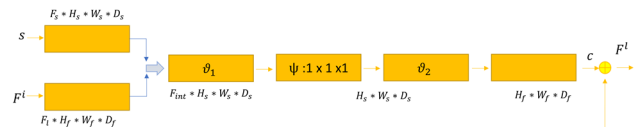


FIGURE 2. Shows pictorial overview of the SAS.

The attention coefficient, denoted as a_i^l , is calculated using the equation 1.

$$\alpha^l = \partial_1(\theta_1^l(z^l, p; \theta))$$

$$\theta_1^l = \psi^T \left(\partial_2 \left(L_z^T z_i^l + L_p^T p + b_{zg} \right) \right) + b_\psi \quad (1)$$

Here, activation non-linearity is denoted by $\partial_1(z)$ and normalization function is represented by $\partial_2(z)$. The attention Gate is symbolized by parameters θ , L_z and L_p indicates weights matrices for linear transformation on input p and z_i^l . b_{zg} and b_ψ shows biased terms related to a linear transformation.

Figure 2 presents a synoptic overview of the SAS. The inputs, denoted as F_i , undergo scaling with attention coefficient c within the SAS. Spatial areas are chosen through the analysis of both activations and contextual information by the selector S , which is derived from a coarser scale. Trilinear interpolation is then employed for grid resampling of attention coefficients.

C. UNIFIED LOSS FUNCTION

The Focal loss is a modification of the Cross Entropy loss (CE) designed to tackle the problem of data imbalancing by focusing more on the complex example during training, thereby facilitating the learning of more challenging examples. The simplified mathematical formulation for the Focal

loss function is shown in equation 2 below.

$$F_{loss(C_t)} = \alpha(1 - C_t)^\gamma \cdot BCE_{loss(c,t)},$$

$$\text{where } C_t = \begin{cases} c & \text{if } t = 1 \\ 1 - c & \text{if } t = 0 \end{cases} \quad (2)$$

where C_t is the probability assigned to the ground truth class, α is the hyperparameter that controls the weight assigned to each class. γ determines the degree of down-weighting, $BCE_{loss(c,t)}$, represents the binary cross-entropy loss between the predicted probabilities c and the true labels t . This loss function is characterized by parameters i.e. α and γ , which govern the weights of the class and the extent of reducing the influence of easily classifiable pixels. If parameter γ is set to 0, the Focal loss is reduced to the binary CE loss.

The Focal Tversky loss is mathematically describe in equation 3 below.

$$FT_{loss} = \sum_{k=1}^K (1 - T_i)^{\frac{1}{\gamma}} \quad (3)$$

where K represents class, T_i indicates Tversky index used to assess the similarity between two sets. Mathematically T_i is formulized as given in equation 4 below.

$$T_i = \frac{\sum_{i=1}^p F_{0i}B_{0i}}{\sum_{i=1}^p F_{0i}B_{0i} + \delta \sum_{i=1}^p F_{0i}B_{1i} + \beta \sum_{i=1}^p F_{1i}B_{0i}} \quad (4)$$

F_{0i} represents the likelihood of pixel being part of the foreground category, while F_{1i} denotes the likelihood of pixel belonging to the background category. The variable B_{1i} is assigned 1 if the pixel belongs to the foreground and 0 otherwise. Conversely, G_{1i} takes on value of 1 for the background and 0 value for the foreground.

The Dice Focal loss [68] and Combo loss [69] represent two composite loss functions that take advantages from CE loss and Dice loss functions. Nevertheless, unable to maximizes the complete advantages when dealing with skewed distribution of data or classes. Both the Dice Focal, and combo loss with adjustable parameters (δ and β) in the CE component losses, exhibit partial resilience to imbalanced output. However, neither approach adequately incorporates the balancing aspect of the Dice component, which equally weights positive and negative examples. Both losses share the weakness of their Dice component not handling input imbalance, but the Dice Focal loss provides a degree of counterbalance through the focal parameters. To address these issues, [70] developed the Hybrid loss, equipped with adjustable parameters for handling output imbalance within CE loss and Dice loss components. The Hybrid loss is mathematically articulated as shown in equation 5.

$$H_{FL} = \lambda F_{loss} + (1 - \lambda) FT_{loss} \quad (5)$$

λ belongs to the range [0 1] and estimates the proportional weight assigned to each component of loss functions.

The hybrid loss adjusts both the CE based loss and Dice functions to address class imbalance. However, using the hybrid loss in practical applications presents two challenges. First, there are 06 parameters to optimize: three (γ , α and β)

from focal Tversky loss function, two (α and β) from focal loss function and λ to govern the relative weight of two component losses. This provides greater flexibility and leads to a significantly expanded search space for hyperparameters. The second problem with focal loss functions is that the enhancement or suppression controlled by the focal parameter is employed universally to all classes equally, which is potentially be problematic at the later stages of training.

The hybrid loss function combines both cross-entropy loss and Dice loss to address the class imbalance. However, there are two key problems related to this loss function. Firstly, it has multiple hyperparameters (six parameters) that significantly increase the search space, making it challenging to fine-tune them effectively. Secondly, the suppressing and enhancing mechanism within this loss function makes achieving convergence during training quite difficult.

The Unified loss effectively tackles both issues. Firstly, it groups functionally equal hyperparameters, which makes it easier to tune the loss function. Secondly, it leverages asymmetry to effectively focus the effects of the focal parameters. As a result, the loss function improves performance in suppressing errors for rare classes and enhancing accuracy for majority classes.

We have replaced ϑ and μ in the Focal loss, Tversky loss, with a shared parameter called δ . This parameter helps control the class imbalance. Additionally, we have reformulated τ to allow both suppression in the Focal loss and enhancement in the Tversky loss simultaneously. Mathematically, these adjusted losses are shown in equations 6 and 7.

We use the terms ‘‘revised Focal loss’’ F_r and ‘‘revised Focal Tversky loss’’ T_r to refer to these revised versions respectively.

$$F_r = \delta (1 - G_t)^{1-\tau} \cdot L_{BCE}(X, Y) \quad (6)$$

$$T_r = \sum_{c=1}^c (1 - aT_i)^\tau \quad (7)$$

where G_t is the ground truth, X represents the predicted value, Y denotes ground-truth and τ is parameter to control focus strength and aT_i symbolizes adjusted Tversky index is mathematically shown in equation 8,

$$aT_i = \frac{\sum_{i=1}^p F_{0i}B_{0i}}{\sum_{i=1}^p P_{0i}B_{0i} + \delta \sum_{i=1}^p F_{0i}B_{1i} + (1 - \delta) \sum_{i=1}^p F_{0i}B_{0i}} \quad (8)$$

where F_{0i} represents foreground pixels of input image and F_{1i} represents background pixels. B_{0i} is an indicator function that takes 1 for pixels in the foreground and 0 for pixels in the background. On the other hand, B_{1i} assign 1 for the background pixels, while it is set to 0 for foreground pixels.

Hence, a symmetric version of Unified loss can be written as following equation 9.

$$C_{sl} = \delta F_a + (1 - \delta T_a) \quad (9)$$

The parameter δ , ranging from 0 to 1, plays a crucial role in determining the relative significance assigned to the two

types of losses. δ governs the weighting assigned to negative and positive instances, while τ controls both the amplification of the minority class and the reduction of the majority class. Lastly, δ determines the weights allocated to the two individual losses comprising the overall loss function.

The Focal loss effectively suppresses the background class but unintentionally suppresses the rare class because the focal parameters are utilized for all classes. Asymmetry provides a solution by allowing selective suppression or enhancement of specific classes using the focal parameter. By allocating variable losses to each class, the revised asymmetric Focal loss (F_{ra}) overcomes the issue of harmful suppression of the rare class while still maintaining suppression of the background. In this revised version, the focal parameter is discarded for the loss component associated with the rare class C_r while still preserving its effect on the background class. This phenomenon is shown in Equation 10.

$$F_{ra} = \frac{\delta}{P} Y_{i:C_r} \log(G_{t,C_r}) - \frac{1-\delta}{G} \sum_{c \neq C_r} (1-p_{t,c})^\tau \log(G_{t,C_r}) \quad (10)$$

Conversely, the revised Tversky loss takes a different approach. In this case, the focal parameter is discarded for the part of the loss that pertains to the background. However, the enhancement of the rare class (C_r) is still maintained. This leads to the definition of the revised asymmetric Focal Tversky loss (F_{raT}) is formulated in Equation 11:

$$F_{raT} = \sum_{c \neq C_r} (1-aT_i) + \sum_{c=C_r} (1-aT_i)^{1-\tau} \quad (11)$$

Therefore, the asymmetric version of the proposed Unified loss (L_c) is described in Equation 12.

$$L_c = \tau F_{ra} + (1-\tau) F_{raT} \quad (12)$$

The problem of suppression of loss that occurs with the Focal loss is addressed by combining it with the Focal Tversky loss in a complementary manner. The asymmetry in this combination allows for simultaneous attenuation of the background loss and amplification of the foreground loss.

By integrating concepts from preceding loss functions, the suggested loss generalizes both CE loss and Dice loss functions within a unified framework. It can be demonstrated that all previously described Dice and CE loss functions are specific instances of the unified loss. For instance, when γ is set to 0 and δ is set to 0.5, the DSC loss and the CE loss are retrieved by setting λ to 0 and 1 respectively. By elucidating the interrelation among these loss functions, optimizing the Comprehensive Focal loss is significantly more straightforward than independently experimenting with distinct loss functions. Moreover, it is more robust as it can handle both input and output imbalances. Notably, considering the efficiency of both the DSC loss and CE loss operations, and the minimal increase in time complexity introduced by the focal parameter, the Unified loss is not expected to extend training time beyond its constituent loss functions.

In practical applications, streamlining the optimization process for the Unified loss can be achieved by simplifying it to a single hyperparameter. Because the focal parameter has varied effects on each component loss, the role of λ is somewhat unnecessary. As a result, we suggest setting $\lambda = 0.5$, allocating equal weight to each component loss, a recommendation supported by empirical evidence [69]. Additionally, we propose setting $\delta = 0.6$ to address the tendency of the Dice loss to generate segmentations with high precision and low recall, particularly in the presence of class imbalance. This value is less than $\delta = 0.7$ in the Tversky loss, considering the influence of the CE loss component. By heuristically reducing the hyperparameter search space to the single γ parameter, the Unified loss becomes both effective and easy to optimize.

IV. DATASET DESCRIPTION

A. BUS2017 DATASET

Digital mammography is widely utilized as a primary screening modality for breast cancer detection. The BUS2017 dataset, specifically curated for breast cancer detection, comprises 163 ultrasound images. These images have dimensions of 760×570 pixels. Within the dataset, 110 images represent benign lesions, including 39 fibroadenomas, six from other benign categories, and 65 unspecified cysts. The remaining 53 ultrasound images illustrate cancerous masses, primarily consisting of invasive ductal carcinomas.

B. THE DIGITAL DATABASE THYROID IMAGE (DDTI)

The DDTI dataset serves as an extensive repository of ultrasound images specifically designed for the screening of thyroid cancer. The DDTI primarily emphasizes the B-mode UI of thyroid cancer, offering detailed diagnostic descriptions and annotations for each image. The dataset comprises 298 UI, with 270 images featuring women and 29 images featuring men.

C. LITS MICCAI 2017

This dataset comprises a collection of 260 MR images, each sized $512 \times 512 \times 3$, obtained from different patients diagnosed with liver cancer. Each case includes MRI scan and a corresponding segmentation mask that labels the liver and liver tumors. The tumor area constitutes the minority class, comprising less than 30% in each scan and exhibiting variability in size, shape, and texture, thereby presenting a challenging task for tumor segmentation.

D. ATLAS DATASET

The ATLAS dataset comprises 90 liver-focused MR images obtained from patients diagnosed with unresectable Hepatocellular Carcinoma (HCC). Each image has dimensions of $512 \times 512 \times 3$. The tumor region in each image spans 15-30%, with the extent varying depending on the stage of liver cancer.

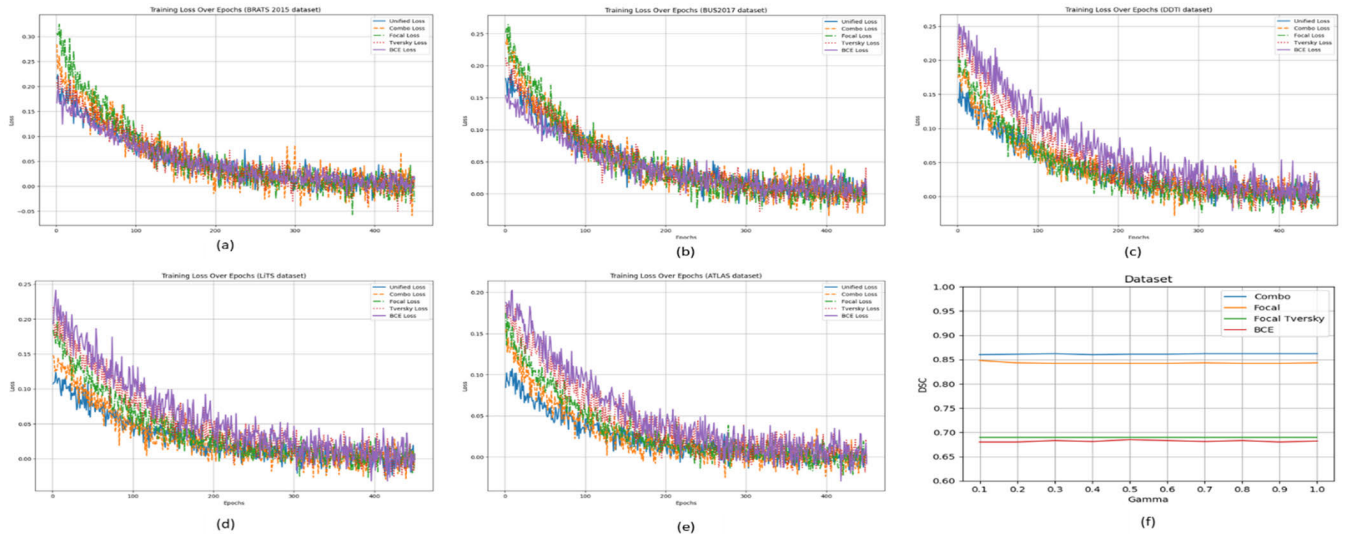


FIGURE 3. Shows the epochs vs different loss functions on MRI datasets during training. 3(f) illustrates DSC performance for each value of λ ranging from 0.1 to 0.9 for unified loss across all datasets.

E. BRATS 2015

The BRATS 2015 dataset consists of brain tumour MR images employed for the detection of brain tumour. The dataset includes a total of 220 Magnetic Resonance Imaging (MRI) scans of high-grade gliomas (HGG) and 54 MRI scans of low-grade gliomas (LGG).

V. EVALUATION METRICS

To assess the segmentation performance of the proposed strategy, we have utilized widely recognized metrics such as Intersection over Union (IoU), Dice Similarity Coefficient (DSC), Precision, and Recall. These metrics are defined by equations 13, 14, 15, and 16, as provided below.

$$IoU = \frac{2T^P}{2T^P + F^P + F^N} \quad (13)$$

$$DSC = \frac{T^P}{T^P + F^P + F^N} \quad (14)$$

$$Precision = \frac{T^P}{T^P + F^P} \quad (15)$$

$$Recall = \frac{T^P}{T^P + F^N} \quad (16)$$

VI. IMPLEMENTATION DETAIL

For our experiments, we utilized Python 3.6 along with libraries: Numpy, TensorFlow, and Keras, which are commonly employed for deep learning tasks. Additionally, we employed Matplotlib to create visualizations and plots for analyzing and presenting the results of our experiments. For each dataset, we resized the images to dimensions of $128 \times 128 \times 3$, $256 \times 256 \times 3$, and $512 \times 512 \times 3$ pixels. To normalize the pixel values, we applied Z-score normalization, which brings the values within the range of 0 to 1. Random assignment was used to perform five-fold cross-validation. We evaluated the baseline performance of several models,

including UNet, FCN, RCNN, SegNet, and Deeplab-V3. Subsequently, we examine the effects of progressive resizing, selective attention, and multiple loss functions on the performance of models as shown in Figure 3. We have integrated attention layers in the DeepLab-v3 model within the atrous spatial pyramid pooling module, in the case of FCN within the decoder part of the network, in CGAN within the generator, and similarly for R-CNN, we have incorporated them into the feature extraction part of the network. This enables the models to effectively capture multi-scale contextual features. The same hyperparameters were used during all experiments, while the parameters of models were initialized on the basis of their default settings. Various experiments were conducted to analyze the effectiveness of loss functions across five state-of-the-art deep learning models. For this purpose, the selected evaluation metrics include DSC, IoU, precision, and recall. The assessment involves extensively used loss functions for medical image segmentation, such as binary cross entropy loss, focal loss, Tversky loss, combo loss and unified. All models used the same batch size i.e. 32 and hyperparameters. Networks are trained using Stochastic Gradient Descent (SGD) with an initial learning rate of 0.001, momentum of 0.9, and weight decay of 0.0005, respectively. The model underwent training through a 5-fold cross-validation technique, with an incorporated early stopping mechanism to halt training in case there was no improvement in the loss functions.

VII. RESULTS ANALYSIS

To evaluate the performance of the proposed Unified approach, we conducted a series of experiments on multiple datasets i.e., BUS2017, DDTI, LiTS MICCAI 2017, ATLAS, and BRATS 2015. During experiments, the individual components of the Unified approach on the state-of-the-art

segmentation models such as UNet, FCN, RCNN, SegNet, and DeepLab-V3 were also evaluated. The objective was to systematically analyze the impact of each component of the proposed scheme on the segmentation results on selected datasets. Each component of the proposed technique was gradually incorporated with UNet, FCN, RCNN, SegNet, and DeepLab-V3 to monitor the performance. When all components were present, a significant improvement was observed, resulting in an Intersection over Union (IoU) score of 0.860 in comparison to the IoU scores for U-Net (0.820) without a class-imbalance strategy. The other models also demonstrated significant improvements in segmentation tasks across all datasets, which confirms the generalizability of the proposed approach for segmentation tasks in the medical image domain.

The results of all the models are presented in Table 1. DeepLab-V3 achieved the highest performance across all evaluation metrics and datasets. DeepLab-V3 exhibited outstanding results with a Dice Similarity Coefficient (DSC) of 0.953, Intersection over Union (IoU) of 0.893, precision of 0.943, and recall of 0.944. These scores reflect an improvement of 5% in DSC, 6% in IoU, 4% in precision, and around 4% in recall compared to baseline DeepLab-v3. The highest increase was noticed in ATLAS and LiTS MICCAI 2017 datasets. A 5%, 7% increase in IoU and DSC is observed over the previous extant (DeepLab-V3 DSC=0.628, DSC=0.695) for ATLAS dataset and 5%, 9% increase in IoU and DSC for LiTS MICCAI 2017 dataset for DeepLab-v3 model employing proposed strategy. Segmentation results for DeepLab-V3 are illustrated in Figure 4 using baseline and proposed strategy. The proposed scheme generalizes well with consistently accurate segmentation across different datasets. The images related to the poor delineation quality are either objectively challenging ROI to identify or, in many cases, poor-quality images.

Various experiments were conducted to investigate these components' contribution to medical image segmentation further; the experimental results are presented in Table 1, Table 2, and Table 3. The comprehensive description of each component is summarized in the section below.

A. IMPACT OF SELECTIVE ATTENTION STRATEGY

Based on the results of the ablation study, it is evident that the average segmentation scores of deep learning models performance compared to models with SAS, as shown in Table 2. Specifically, for the DeepLab-V3 model on the BUS2017 dataset, the scores for Intersection over Union (IoU), Dice Similarity Coefficient (DSC), Precision, and Recall are observed to be 0.871, 0.920, 0.933, and 0.921, respectively. Similarly, for the UNet model on the BUS2017 dataset, the IoU score reaches 0.739, and the DSC score reaches 0.921.

The scores achieved without SAS are considerably lower than those achieved with SAS, indicating a significant impact of SAS on the performance of these models for the given datasets. Additionally,

Table 2 presents the performance of other attention mechanisms, wherein it can be observed that the proposed mechanism outperforms others in most datasets. As a result, we have selected the proposed mechanism as the preferred choice in our proposed strategy.

B. IMPACT OF UNIFIED LOSS FUNCTION

Table 3 presents the impact of the proposed unified loss function evaluated across multiple metrics on all five imbalanced datasets. Consistently, the proposed Unified loss function demonstrated a significant impact across these datasets, i.e., Dice Similarity Coefficients (DSC) of 0.910, 0.865, 0.741, 0.673, and 0.791 on the BUS2017 DDTI, LiTS MICCAI 2017, DDTI, ATLAS dataset, and BRATS 2015 datasets, respectively. Table 3 also demonstrates the performance of other losses for DDTI, ATLAS, and BRATS 2015 datasets. The proposed loss is consistent across all datasets and demonstrated improvement in all segmentation metrics compared to others. The visual results of various loss functions, including the proposed ones, are depicted in Figure 5. Conversely, the cross-entropy-based losses performed comparatively worse, with the focal loss exhibiting even lower performance than the cross-entropy loss on the BRATS 2015 and MICCAI 2017 datasets. Notably, no significant differences were observed between the dice-based losses.

C. IMPACT OF PROGRESSIVE RESIZING

Table 4 provides a comprehensive overview of the effect of progressive resizing on the performance of deep learning models. The results of the proposed model highlight a substantial improvement in the segmentation outcomes across all the datasets. Table 4 demonstrates that progressive resizing simplifies the diversity in the training data, enhancing the model's capabilities to learn from various scales of the same input image. As a result, the models show better performance for both the majority and minority classes of the datasets.

VIII. DISCUSSION

The research community has conducted several studies to address the challenges associated with a class imbalance in computer vision and other image-processing domains. Previous medical research has focused on specific problems and targets the specific dataset for experimental purposes. Consequently, these approaches mostly fail to generalize effectively. In this research, we intend to develop a comprehensive and generalized model that can effectively handle class imbalance across various medical image datasets for segmentation tasks. The proposed scheme performs segmentation tasks in three key steps: 1) Progressive resizing as a preprocessing step, 2) A selective attention mechanism to prioritize regions of interest (ROIs), and 3) A robust Unified loss function to address the discrepancy between dominant and minority classes. To demonstrate the efficiency of our method, we merged it into several state-of-the-art deep-learning models for image segmentation, including UNet, SegNet, RCNN, FCN, and DeepLab-V3. After that, we

TABLE 1. Performance of different models using proposed class imbalance strategy on different MRI image datasets.

Model	IoU	DSC	Precision	Recall	Dataset
Models' score before application of class misbalancing strategy					
UNet	0.820	0.893	0.893	0.892	BUS2017
FCN	0.732	0.823	0.794	0.902	
RCNN	0.822	0.903	0.894	0.902	
SegNet	0.812	0.884	0.865	0.875	
DeepLab-V3	0.833	0.903	0.903	0.904	
Models' score after application of class misbalancing strategy					
Model	IoU	DSC	Precision	Recall	BUS2017
UNet	0.860	0.933	0.923	0.932	
FCN	0.752	0.843	0.814	0.912	
RCNN	0.862	0.923	0.914	0.912	
DeepLab-V3	0.893	0.953	0.943	0.944	
Models' score before application of class misbalancing strategy					
UNet	0.755	0.843	0.706	0.815	DDTI
FCN	0.712	0.703	0.672	0.672	
RCNN	0.732	0.833	0.701	0.801	
SegNet	0.7423	0.812	0.693	0.679	
DeepLab-V3	0.825	0.875	0.826	0.828	
Models' score after application of class misbalancing strategy					
Model	IoU	DSC	Precision	Recall	DDTI
UNet	0.785	0.873	0.746	0.845	
FCN	0.742	0.743	0.682	0.692	
RCNN	0.762	0.853	0.751	0.846	
DeepLab-V3	0.845	0.895	0.856	0.858	
Models' score before application of class misbalancing strategy					
UNet	0.733	0.653	0.622	0.812	LiTS MICCAI 2017
FCN	0.684	0.625	0.643	0.723	
RCNN	0.686	0.597	0.648	0.725	
SegNet	0.613	0.521	0.543	0.652	
DeepLab-V3	0.75	0.676	0.71	0.80	
Models' score after application of class misbalancing strategy					
Model	IoU	DSC	Precision	Recall	LiTS MICCAI 2017
UNet	0.775	0.674	0.684	0.845	
FCN	0.723	0.657	0.683	0.743	
RCNN	0.726	0.647	0.668	0.755	
DeepLab-V3	0.804	0.765	0.745	0.823	
Models' score before application of class misbalancing strategy					
UNet	0.742	0.624	0.703	0.795	ATLAS dataset
FCN	0.715	0.605	0.675	0.756	
RCNN	0.682	0.584	0.645	0.726	
SegNet	0.737	0.618	0.688	0.786	
DeepLab-V3	0.746	0.628	0.717	0.794	
Models' score after application of class misbalancing strategy					
Model	IoU	DSC	Precision	Recall	ATLAS dataset
UNet	0.776	0.664	0.745	0.832	
FCN	0.748	0.643	0.643	0.775	
RCNN	0.779	0.674	0.734	0.823	
DeepLab-V3	0.795	0.695	0.753	0.854	
Models' score before application of class misbalancing strategy					
UNet	0.893	0.772	0.874	0.833	BRATS 2015
FCN	0.875	0.756	0.857	0.806	
RCNN	0.865	0.746	0.8478	0.797	
SegNet	0.885	0.766	0.868	0.829	
DeepLab-V3	0.895	0.776	0.877	0.838	
Models' score after application of class misbalancing strategy					

TABLE 1. (Continued.) Performance of different models using proposed class imbalance strategy on different MRI image datasets.

Model	IoU	DSC	Precision	Recall	
UNet	0.932	0.803	0.894	0.862	BRATS 2015
FCN	0.903	0.783	0.863	0.843	
RCNN	0.942	0.822	0.893	0.843	
SegNet	0.921	0.783	0.893	0.853	
DeepLab-V3	0.921	0.821	0.890	0.883	

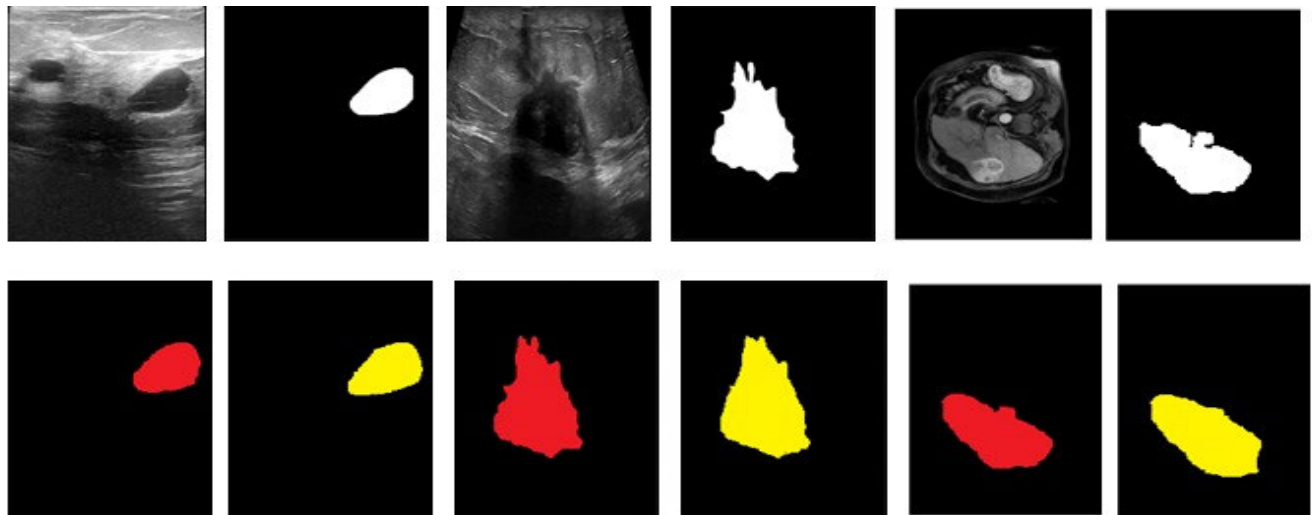


FIGURE 4. Segmentation results of DeepLab-v3 on different datasets. (The yellow color represents the deepLab-v3 generated mask without imbalancing scheme, and red represents the output of deepLab-v3 with the proposed scheme).

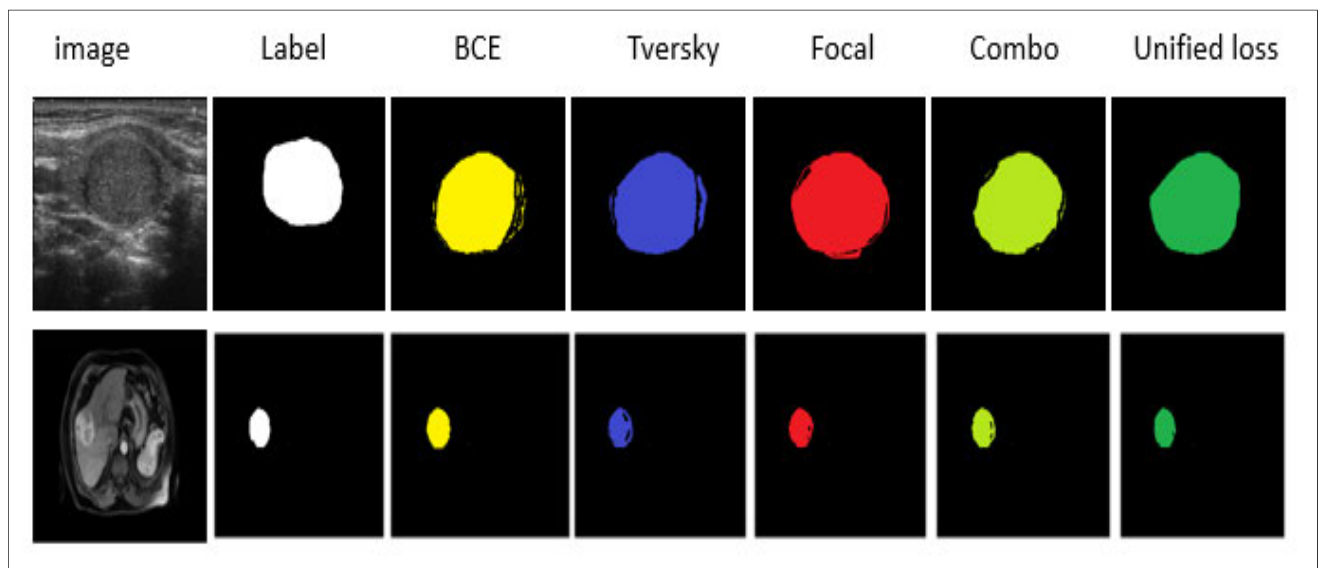


FIGURE 5. Instances of segmentation using a different loss function.

evaluated these models’ performance using five highly unbalanced image segmentation datasets.

After the processing step, the progressive resizing mechanism of input images from the entire dataset become active, such as for the first iteration, and we resize input images into $128 \times 128 \times 3$ pixels. For the second run of the experiment,

we used images of $256 \times 256 \times 3$; for the final run, we used $512 \times 512 \times 3$ pixels. We have employed the concept of transfer learning in which weights from the first run were used in the second run, and the second-run pretrained model was employed for the third round of experiments. This progressive resizing yielded considerable improvement. From Table 4,

TABLE 2. Impact of Selective attention strategy on the segmentation performance of different extant models.

Model	IoU	DSC	Precision	Recall	
UNet	0.849	0.921	0.901	0.913	
FCN	0.739	0.821	0.791	0.901	
RCNN	0.854	0.903	0.901	0.901	BUS2017
SegNet	0.801	0.892	0.871	0.881	
DeepLab-V3	0.871	0.920	0.933	0.921	
Model	IoU	DSC	Precision	Recall	
UNet	0.775	0.863	0.727	0.832	
FCN	0.723	0.721	0.661	0.671	DDTI
RCNN	0.753	0.849	0.740	0.838	
SegNet	0.747	0.831	0.701	0.692	
DeepLab-V3	0.823	0.873	0.834	0.842	
Model	IoU	DSC	Precision	Recall	
UNet	0.760	0.658	0.667	0.827	
FCN	0.711	0.645	0.666	0.731	LiTS MICCAI 2017
RCNN	0.704	0.624	0.645	0.734	
SegNet	0.642	0.543	0.579	0.672	
DeepLab-V3	0.796	0.751	0.728	0.811	
Model	IoU	DSC	Precision	Recall	
UNet	0.753	0.643	0.721	0.813	
FCN	0.734	0.632	0.632	0.753	ATLAS dataset
RCNN	0.754	0.663	0.721	0.803	
SegNet	0.726	0.629	0.674	0.793	
DeepLab-V3	0.775	0.685	0.739	0.837	
Model	IoU	DSC	Precision	Recall	
UNet	0.916	0.792	0.881	0.851	
FCN	0.893	0.762	0.851	0.831	BRATS 2015
RCNN	0.921	0.801	0.881	0.822	
SegNet	0.901	0.772	0.871	0.832	
DeepLab-V3	0.903	0.801	0.878	0.865	

you can notice a consistent improvement in evaluation metrics across all datasets for all models for the medical image segmentation task. Due to progressive resizing, the variation in scale can aid in capturing different levels of details and improving the model’s ability to distinguish between different classes. Secondly, by gradually increasing the size of the input images, the model becomes exposed to more complex and detailed information, which can be beneficial for accurately segmenting smaller or more challenging objects or classes. This exposure to larger images helps the model learn finer details and enhances its ability to capture the nuances of the minority class.

We have introduced a novel selective attention strategy, enabling deep learning models to learn salience regions integrated with focal parameters to control the suppression of irrelevant regions. Importantly, the suggested SAS performs consistently well across all models on all datasets, demonstrating the method’s capability to generalize to unseen data.

For a fair comparison, we have implemented several attention strategies—namely, spatial attention, channel attention,

and hybrid attention—into existing deep learning models, and the respective impacts of each strategy are presented in Table 5. The incorporation of these attention mechanisms has resulted in noticeable enhancements in the segmentation performance of the deep learning models. However, an issue arises with excessive upsampling, which leads to a reduction in the inter-pixel information. Additionally, these techniques encounter challenges when dealing with images that exhibit similar Regions of Interest (ROI) due to the presence of area similarity and positional changes. The suggested selective attention mechanism utilizes a dual attention approach, employing parallel processing to prevent the loss of important information and incorporating focal parameters to regulate the extent of background suppression.

To assess the effectiveness of the suggested Unified loss function for MR image segmentation, we conducted several experiments employing various segmentation models on five class-imbalanced datasets. Additionally, we compared the proposed loss function with five other losses. The results presented in Table 3 and Figure 4 confirm that the proposed

TABLE 3. Illustrate the effect of unified loss function using state-of-the-art deep learning models on highly imbalanced datasets.

Model	IoU	DSC	Precision	Recall	
UNet	0.837	0.919	0.897	0.908	BUS2017
FCN	0.730	0.815	0.782	0.893	
RCNN	0.847	0.893	0.896	0.893	
SegNet	0.793	0.889	0.865	0.878	
DeepLab-V3	0.865	0.910	0.925	0.901	
Model	IoU	DSC	Precision	Recall	
UNet	0.771	0.854	0.717	0.824	DDTI
FCN	0.718	0.716	0.653	0.663	
RCNN	0.743	0.834	0.735	0.827	
SegNet	0.735	0.826	0.696	0.686	
DeepLab-V3	0.818	0.865	0.801	0.832	
Model	IoU	DSC	Precision	Recall	
UNet	0.757	0.649	0.657	0.802	LITS MICCAI 2017
FCN	0.692	0.602	0.645	0.704	
RCNN	0.694	0.613	0.623	0.712	
SegNet	0.621	0.502	0.565	0.643	
DeepLab-V3	0.787	0.741	0.718	0.805	
Model	IoU	DSC	Precision	Recall	
UNet	0.742	0.632	0.702	0.801	ATLAS dataset
FCN	0.721	0.611	0.614	0.743	
RCNN	0.744	0.653	0.707	0.799	
SegNet	0.706	0.619	0.664	0.781	
DeepLab-V3	0.763	0.673	0.722	0.822	
Model	IoU	DSC	Precision	Recall	
UNet	0.901	0.782	0.873	0.847	BRATS 2015
FCN	0.882	0.753	0.841	0.823	
RCNN	0.882	0.799	0.823	0.801	
SegNet	0.886	0.763	0.861	0.805	
DeepLab-V3	0.893	0.791	0.872	0.854	

TABLE 4. Provides the effect of progressive resizing on the results of segmentation models.

Model	IoU	DSC	Precision	Recall	
UNet	0.735	0.629	0.697	0.798	ATLAS dataset
FCN	0.719	0.609	0.608	0.733	
RCNN	0.738	0.642	0.698	0.789	
SegNet	0.693	0.607	0.653	0.773	
DeepLab-V3	0.755	0.662	0.714	0.814	
Model	IoU	DSC	Precision	Recall	
UNet	0.893	0.779	0.867	0.836	BRATS 2015
FCN	0.876	0.743	0.836	0.819	
RCNN	0.871	0.781	0.817	0.791	
SegNet	0.875	0.751	0.853	0.794	
DeepLab-V3	0.882	0.772	0.861	0.841	

method consistently performs well for the segmentation task across all datasets and deep learning models. We observed

that this hyperparameter remains stable, making the optimization process relatively straightforward. The Unified loss

TABLE 5. Comparative analysis of different attention strategies.

Model	Attention mechanism	IOU	DSC	Precision	Recall	Dataset
DeepLab-V3	spatial	0.901	0.814	0.889	0.878	BRATS 2015
	channel	0.893	0.801	0.877	0.864	
	hybrid	0.911	0.813	0.887	0.876	
	SAS	0.921	0.821	0.890	0.883	
DeepLab-V3	spatial	0.789	0.758	0.729	0.808	LiTS MICCAI 2017
	channel	0.785	0.742	0.722	0.803	
	hybrid	0.793	0.752	0.731	0.812	
	SAS	0.804	0.765	0.745	0.823	

TABLE 6. Shows the results of different balancing techniques.

Model	Data imbalancing mechanism	IOU	DSC	Precision	Recall	Dataset
DeepLab-V3	Random Oversampling	0.879	0.803	0.859	0.858	BRATS 2015
	Synthetic Minority Oversampling (SMO)	0.873	0.805	0.869	0.854	
	SMO-Iterative Partitioning Filter	0.919	0.819	0.894	0.882	
	SAS	0.931	0.831	0.900	0.893	
DeepLab-V3	Random Oversampling	0.786	0.745	0.724	0.892	LiTS MICCAI 2017
	Synthetic Minority Oversampling (SMO)	0.787	0.747	0.734	0.817	
	SMO-Iterative Partitioning Filter	0.801	0.759	0.743	0.820	
	DeepLab-V3	0.824	0.775	0.765	0.833	

function offers practical advantages by simplifying hyperparameter tuning, enhancing training efficiency, improving convergence properties, and providing a robust solution for handling imbalances. These features contribute to the overall effectiveness of the model in various practical applications, particularly in medical image segmentation where imbalances in data distributions are common.

A. COMPARISON WITH OTHER BALANCING TECHNIQUES

To ensure an equitable comparison, we have conducted an assessment of the performance of various models by employing different techniques designed to address the data imbalance problem. The outcomes of the analysis are presented in Table 6, showcasing the segmentation results across three datasets.

IX. CONCLUSION

This research has devised a comprehensive approach to address the issue of medical image segmentation in MR images using a dataset with a significant class imbalance. The study has introduced a selective attention strategy focusing more on the Region of Interest (ROI) and a novel unified loss function. This loss function can suppress features from the majority classes (background) while highlighting features from the minority class (foreground). This helps mitigate any bias towards the background, making the model more balanced. The integration of progressive resizing with

transfer learning aids the models in capturing varying levels of detail, thereby enhancing their ability to differentiate between different classes. The effectiveness of this unified approach was evaluated on five distinct MR image datasets in multiple deep-learning models, revealing that the proposed strategy outperforms alternative techniques for addressing class imbalance. In the future, we are interested in applying it to other datasets having class imbalance issues.

DATA AVAILABILITY

The code is available at <https://github.com/LijuanCui0171/A-Unified-Approach-Addressing-Class-Imbalance-in-Magnetic-Resonance-Image-for-Deep-Learning-Models.git>. All the data associated with this work is publicly available.

COMPETING INTERESTS

The authors do not show any competing interests.

REFERENCES

- [1] A. Bria, C. Marrocco, and F. Tortorella, "Addressing class imbalance in deep learning for small lesion detection on medical images," *Comput. Biol. Med.*, vol. 120, May 2020, Art. no. 103735, doi: 10.1016/j.compbiomed.2020.103735.
- [2] A. Balasundaram, M. S. Kavitha, Y. Prathepan, D. Akshat, and M. V. Kaushik, "A foreground prototype-based one-shot segmentation of brain tumors," *Diagnostics*, vol. 13, no. 7, p. 1282, Mar. 2023, doi: 10.3390/diagnostics13071282.

- [3] J. Cheng, W. Huang, S. Cao, R. Yang, W. Yang, Z. Yun, Z. Wang, and Q. Feng, "Enhanced performance of brain tumor classification via tumor region augmentation and partition," *PLoS One*, vol. 10, no. 10, 2022, Art. no. e0140381, doi: [10.1371/journal.pone.0140381](https://doi.org/10.1371/journal.pone.0140381).
- [4] A. M. Gab Allah, A. M. Sarhan, and N. M. Elshennawy, "Classification of brain MRI tumor images based on deep learning PGGAN augmentation," *Diagnostics*, vol. 11, no. 12, p. 2343, Dec. 2021, doi: [10.3390/diagnostics11122343](https://doi.org/10.3390/diagnostics11122343).
- [5] H. Chen, Z. Qin, Y. Ding, L. Tian, and Z. Qin, "Brain tumor segmentation with deep convolutional symmetric neural network," *Neurocomputing*, vol. 392, pp. 305–313, Jun. 2020, doi: [10.1016/j.neucom.2019.01.111](https://doi.org/10.1016/j.neucom.2019.01.111).
- [6] M. A. Gómez-Guzmán, L. Jiménez-Beristáin, E. E. García-Guerrero, O. R. López-Bonilla, U. J. Tamayo-Perez, J. J. Esqueda-Elizondo, K. Palomino-Vizcaino, and E. Inzunza-González, "Classifying brain tumors on magnetic resonance imaging by using convolutional neural networks," *Electronics*, vol. 12, no. 4, pp. 1–22, 2023, doi: [10.3390/electronics12040955](https://doi.org/10.3390/electronics12040955).
- [7] L. Gao, L. Zhang, C. Liu, and S. Wu, "Handling imbalanced medical image data: A deep-learning-based one-class classification approach," *Artif. Intell. Med.*, vol. 108, Aug. 2020, Art. no. 101935, doi: [10.1016/j.artmed.2020.101935](https://doi.org/10.1016/j.artmed.2020.101935).
- [8] X. Liu, L. Guo, H. Wang, J. Guo, S. Yang, and L. Duan, "Research on imbalance machine learning methods for MRT₁WI soft tissue sarcoma data," *BMC Med. Imag.*, vol. 22, no. 1, pp. 1–13, Aug. 2022, doi: [10.1186/s12880-022-00876-5](https://doi.org/10.1186/s12880-022-00876-5).
- [9] F. Rustam, M. A. Siddique, H. U. R. Siddiqui, S. Ullah, A. Mehmood, I. Ashraf, and G. S. Choi, "Wireless capsule endoscopy bleeding images classification using CNN based model," *IEEE Access*, vol. 9, pp. 33675–33688, 2021, doi: [10.1109/ACCESS.2021.3061592](https://doi.org/10.1109/ACCESS.2021.3061592).
- [10] A. A. Reshi, F. Rustam, A. Mehmood, A. Alhossan, Z. Alrabiah, A. Ahmad, H. Alsuwailem, and G. S. Choi, "An efficient CNN model for COVID-19 disease detection based on X-ray image classification," *Complexity*, vol. 2021, pp. 1–12, May 2021, doi: [10.1155/2021/6621607](https://doi.org/10.1155/2021/6621607).
- [11] R. Rodríguez Outeiral, P. Bos, H. J. van der Hulst, A. Al-Mamgani, B. Jasperse, R. Simões, and U. A. van der Heide, "Strategies for tackling the class imbalance problem of oropharyngeal primary tumor segmentation on magnetic resonance imaging," *Phys. Imag. Radiat. Oncol.*, vol. 23, pp. 144–149, Jul. 2022, doi: [10.1016/j.phro.2022.08.005](https://doi.org/10.1016/j.phro.2022.08.005).
- [12] A. Bansal, A. Verma, S. Singh, and Y. Jain, "Combination of oversampling and undersampling techniques on imbalanced datasets," in *Proc. Int. Conf. Innov. Comput. Commun.*, in Lecture Notes in Networks and Systems, vol. 492, 2023, pp. 647–656, doi: [10.1007/978-981-19-3679-1_55](https://doi.org/10.1007/978-981-19-3679-1_55).
- [13] S. Sharma, A. Gosain, and S. Jain, "A review of the oversampling techniques in class imbalance problem," in *Proc. Int. Conf. Innov. Comput. Commun. (ICICC)*, vol. 1, 2022, pp. 459–472, doi: [10.1007/978-981-16-2594-7_38](https://doi.org/10.1007/978-981-16-2594-7_38).
- [14] F. Feng, K.-C. Li, E. Yang, Q. Zhou, L. Han, A. Hussain, and M. Cai, "A novel oversampling and feature selection hybrid algorithm for imbalanced data classification," *Multimedia Tools Appl.*, vol. 82, no. 3, pp. 3231–3267, Jan. 2023, doi: [10.1007/s11042-022-13240-0](https://doi.org/10.1007/s11042-022-13240-0).
- [15] Z. Huang, C. Yang, X. Chen, K. Huang, and Y. Xie, "Adaptive oversampling method for classification with application to imbalanced datasets in aluminum electrolysis," *Neural Comput. Appl.*, vol. 32, no. 11, pp. 7183–7199, Jun. 2020, doi: [10.1007/s00521-019-04208-7](https://doi.org/10.1007/s00521-019-04208-7).
- [16] J. Jani, J. Doshi, I. Kheria, K. Mehta, C. Bhadane, and R. Karani, "LayNet—A multi-layer architecture to handle imbalance in medical imaging data," *Comput. Biol. Med.*, vol. 163, Sep. 2023, Art. no. 107179, doi: [10.1016/j.compbiomed.2023.107179](https://doi.org/10.1016/j.compbiomed.2023.107179).
- [17] M. Yeung, E. Sala, C.-B. Schönlieb, and L. Rundo, "Unified focal loss: Generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation," *Computerized Med. Imag. Graph.*, vol. 95, Jan. 2022, Art. no. 102026, doi: [10.1016/j.compmedimag.2021.102026](https://doi.org/10.1016/j.compmedimag.2021.102026).
- [18] E. Irmak, "Multi-classification of brain tumor MRI images using deep convolutional neural network with fully optimized framework," *Iranian J. Sci. Technol., Trans. Electr. Eng.*, vol. 45, no. 3, pp. 1015–1036, Sep. 2021, doi: [10.1007/s40998-021-00426-9](https://doi.org/10.1007/s40998-021-00426-9).
- [19] M. Drozdal, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal, "The importance of skip connections in biomedical image segmentation," in *Proc. Int. Workshop Deep Learn. Med. Image Anal. Int. Workshop Large-Scale Annotation Biomed. Data Expert Label Synth.*, in Lecture Notes in Computer Science, Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, vol. 10008, 2016, pp. 179–187, doi: [10.1007/978-3-319-46976-8_19](https://doi.org/10.1007/978-3-319-46976-8_19).
- [20] M. Shahzad, A. I. Umar, M. A. Khan, S. H. Shirazi, Z. Khan, and W. Yousaf, "Robust method for semantic segmentation of whole-slide blood cell microscopic images," *Comput. Math. Methods Med.*, vol. 2020, pp. 1–13, Jan. 2020, doi: [10.1155/2020/4015323](https://doi.org/10.1155/2020/4015323).
- [21] F. Grina, Z. Elouedi, and E. Lefevre, "Re-sampling of multi-class imbalanced data using belief function theory and ensemble learning," *Int. J. Approx. Reasoning*, vol. 156, pp. 1–15, May 2023, doi: [10.1016/j.ijar.2023.02.006](https://doi.org/10.1016/j.ijar.2023.02.006).
- [22] C. Chen, X. Wu, E. Zuo, C. Chen, X. Lv, and L. Wu, "R-GDORUS technology: Effectively solving the Raman spectral data imbalance in medical diagnosis," *Chemometric Intell. Lab. Syst.*, vol. 235, Apr. 2023, Art. no. 104762, doi: [10.1016/j.chemolab.2023.104762](https://doi.org/10.1016/j.chemolab.2023.104762).
- [23] S. Iqbal, A. N. Qureshi, J. Li, I. A. Choudhry, and T. Mahmood, "Dynamic learning for imbalanced data in learning chest X-ray and CT images," *Heliyon*, vol. 9, no. 6, Jun. 2023, Art. no. e16807, doi: [10.1016/j.heliyon.2023.e16807](https://doi.org/10.1016/j.heliyon.2023.e16807).
- [24] E. Chamseddine, N. Mansouri, M. Soui, and M. Abed, "Handling class imbalance in COVID-19 chest X-ray images classification: Using SMOTE and weighted loss," *Appl. Soft Comput.*, vol. 129, Nov. 2022, Art. no. 109588, doi: [10.1016/j.asoc.2022.109588](https://doi.org/10.1016/j.asoc.2022.109588).
- [25] Y.-C. Wang and C.-H. Cheng, "A multiple combined method for rebalancing medical data with class imbalances," *Comput. Biol. Med.*, vol. 134, Jul. 2021, Art. no. 104527, doi: [10.1016/j.compbiomed.2021.104527](https://doi.org/10.1016/j.compbiomed.2021.104527).
- [26] Z. Li, K. Kamnitsas, and B. Glocker, "Analyzing overfitting under class imbalance in neural networks for image segmentation," *IEEE Trans. Med. Imag.*, vol. 40, no. 3, pp. 1065–1077, Mar. 2021, doi: [10.1109/TMI.2020.3046692](https://doi.org/10.1109/TMI.2020.3046692).
- [27] C. Zhang, H. Zheng, and Y. Gu, "Dive into the details of self-supervised learning for medical image analysis," *Med. Image Anal.*, vol. 89, Oct. 2023, Art. no. 102879, doi: [10.1016/j.media.2023.102879](https://doi.org/10.1016/j.media.2023.102879).
- [28] T. Zhou, S. Ruan, and S. Canu, "A review: Deep learning for medical image segmentation using multi-modality fusion," *Array*, vols. 3–4, Sep. 2019, Art. no. 100004, doi: [10.1016/j.array.2019.100004](https://doi.org/10.1016/j.array.2019.100004).
- [29] R. Wang, T. Lei, R. Cui, B. Zhang, H. Meng, and A. K. Nandi, "Medical image segmentation using deep learning: A survey," *IET Image Process.*, vol. 16, no. 5, pp. 1243–1267, 2022, doi: [10.1049/ipr2.12419](https://doi.org/10.1049/ipr2.12419).
- [30] T. Huynh, A. Nibali, and Z. He, "Semi-supervised learning for medical image classification using imbalanced training data," *Comput. Methods Programs Biomed.*, vol. 216, Apr. 2022, Art. no. 106628, doi: [10.1016/j.cmpb.2022.106628](https://doi.org/10.1016/j.cmpb.2022.106628).
- [31] J. L. Leevy, T. M. Khoshgoftar, R. A. Bauder, and N. Seliya, "A survey on addressing high-class imbalance in big data," *J. Big Data*, vol. 5, no. 1, Dec. 2018, doi: [10.1186/s40537-018-0151-6](https://doi.org/10.1186/s40537-018-0151-6).
- [32] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuan Yue, and G. Bing, "Learning from class-imbalanced data: Review of methods and applications," *Exp. Syst. Appl.*, vol. 73, pp. 220–239, May 2017, doi: [10.1016/j.eswa.2016.12.035](https://doi.org/10.1016/j.eswa.2016.12.035).
- [33] M. H. Yap, G. Pons, J. Martí, S. Ganau, M. Sentís, R. Zwigellaar, A. K. Davison, and R. Martí, "Automated breast ultrasound lesions detection using convolutional neural networks," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 4, pp. 1218–1226, Jul. 2018, doi: [10.1109/JBHI.2017.2731873](https://doi.org/10.1109/JBHI.2017.2731873).
- [34] P. F. Raudaschl et al., "Evaluation of segmentation methods on head and neck CT: Auto-segmentation challenge 2015," *Med. Phys.*, vol. 44, no. 5, pp. 2020–2036, May 2017, doi: [10.1002/MP.12197](https://doi.org/10.1002/MP.12197).
- [35] S.-L. Liew et al., "A large, open source dataset of stroke anatomical brain images and manual lesion segmentations," *Sci. Data*, vol. 5, no. 1, pp. 1–11, Feb. 2018, doi: [10.1038/sdata.2018.11](https://doi.org/10.1038/sdata.2018.11).
- [36] B. H. Menze, "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Trans. Med. Imag.*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015, doi: [10.1109/TMI.2014.2377694](https://doi.org/10.1109/TMI.2014.2377694).
- [37] L. Pedraza, C. Vargas, F. Narváez, O. Durán, E. Muñoz, and E. Romero, "An open access thyroid ultrasound image database," *Proc. SPIE*, vol. 9287, Jan. 2015, Art. no. 92870W, doi: [10.1117/12.2073532](https://doi.org/10.1117/12.2073532).

- [38] S. Liu, F. Roemer, Y. Ge, E. J. Bedrick, Z.-M. Li, A. Guermazi, L. Sharma, C. Eaton, M. C. Hochberg, D. J. Hunter, M. C. Nevitt, W. Wirth, C. K. Kwok, and X. Sun, "Comparison of evaluation metrics of deep learning for imbalanced imaging data in osteoarthritis studies," *medRxiv*, Sep. 2022, doi: [10.1016/j.joca.2023.05.006](https://doi.org/10.1016/j.joca.2023.05.006).
- [39] Q. Jin, M. Yuan, H. Wang, M. Wang, and Z. Song, "Deep active learning models for imbalanced image classification," *Knowl.-Based Syst.*, vol. 257, Dec. 2022, Art. no. 109817, doi: [10.1016/j.knsys.2022.109817](https://doi.org/10.1016/j.knsys.2022.109817).
- [40] E. Hancer, M. Traoré, R. Samet, Z. Yildirim, and N. Nemati, "An imbalance-aware nuclei segmentation methodology for H&E stained histopathology images," *Biomed. Signal Process. Control*, vol. 83, May 2023, Art. no. 104720, doi: [10.1016/j.bspc.2023.104720](https://doi.org/10.1016/j.bspc.2023.104720).
- [41] Q. Abbas, K. M. Malik, A. K. J. Saudagar, and M. B. Khan, "Context-aggregator: An approach of loss- and class imbalance-aware aggregation in federated learning," *Comput. Biol. Med.*, vol. 163, Sep. 2023, Art. no. 107167, doi: [10.1016/j.combiomed.2023.107167](https://doi.org/10.1016/j.combiomed.2023.107167).
- [42] J. H. Chan and C. Li, "Learning from imbalanced COVID-19 chest X-ray (CXR) medical imaging data," *Methods*, vol. 202, pp. 31–39, Jun. 2022, doi: [10.1016/j.ymeth.2021.06.002](https://doi.org/10.1016/j.ymeth.2021.06.002).
- [43] Z. Jiang, L. Zhao, Y. Lu, Y. Zhan, and Q. Mao, "A semi-supervised resampling method for class-imbalanced learning," *Expert Syst. Appl.*, vol. 221, Jul. 2023, Art. no. 119733, doi: [10.1016/j.eswa.2023.119733](https://doi.org/10.1016/j.eswa.2023.119733).
- [44] S. Fu, D. Su, S. Li, S. Sun, and Y. Tian, "Linear-exponential loss incorporated deep learning for imbalanced classification," *ISA Trans.*, vol. 140, pp. 279–292, Sep. 2023, doi: [10.1016/j.isatra.2023.06.016](https://doi.org/10.1016/j.isatra.2023.06.016).
- [45] S. N. S. S. Daud, R. Sudirman, and T. W. Shing, "Safe-level SMOTE method for handling the class imbalanced problem in electroencephalography dataset of adult anxious state," *Biomed. Signal Process. Control*, vol. 83, May 2023, Art. no. 104649, doi: [10.1016/j.bspc.2023.104649](https://doi.org/10.1016/j.bspc.2023.104649).
- [46] H. M. Rai, K. Chatterjee, and S. Dashkevych, "The prediction of cardiac abnormality and enhancement in minority class accuracy from imbalanced ECG signals using modified deep neural network models," *Comput. Biol. Med.*, vol. 150, Nov. 2022, Art. no. 106142, doi: [10.1016/j.combiomed.2022.106142](https://doi.org/10.1016/j.combiomed.2022.106142).
- [47] M. A. Ganaie and M. Tanveer, "KNN weighted reduced universum twin SVM for class imbalance learning," *Knowl.-Based Syst.*, vol. 245, Jun. 2022, Art. no. 108578, doi: [10.1016/j.knsys.2022.108578](https://doi.org/10.1016/j.knsys.2022.108578).
- [48] Y. Liu, G. Yang, S. Qiao, M. Liu, L. Qu, N. Han, T. Wu, G. Yuan, T. Wu, and Y. Peng, "Imbalanced data classification: Using transfer learning and active sampling," *Eng. Appl. Artif. Intell.*, vol. 117, Jan. 2023, Art. no. 105621, doi: [10.1016/j.engappai.2022.105621](https://doi.org/10.1016/j.engappai.2022.105621).
- [49] H. Liz, J. Huertas-Tato, M. Sánchez-Montañés, J. del Ser, and D. Camacho, "Deep learning for understanding multilabel imbalanced chest X-ray datasets," *Future Gener. Comput. Syst.*, vol. 144, pp. 291–306, Jul. 2023, doi: [10.1016/j.future.2023.03.005](https://doi.org/10.1016/j.future.2023.03.005).
- [50] D. Li, C. Zheng, J. Zhao, and Y. Liu, "Diagnosis of heart failure from imbalance datasets using multi-level classification," *Biomed. Signal Process. Control*, vol. 81, Mar. 2023, Art. no. 104538, doi: [10.1016/j.bspc.2022.104538](https://doi.org/10.1016/j.bspc.2022.104538).
- [51] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks mingxing," *Can. J. Emergency Med.*, vol. 15, no. 3, p. 190, 2013.
- [52] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 2261–2269, doi: [10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243).
- [53] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2016, Jun. 2016, pp. 2818–2826, doi: [10.1109/CVPR.2016.308](https://doi.org/10.1109/CVPR.2016.308).
- [54] Y. Zhang, G. Wang, X. Huang, and W. Ding, "TSK fuzzy system fusion at sensitivity-ensemble-level for imbalanced data classification," *Inf. Fusion*, vol. 92, pp. 350–362, Apr. 2023, doi: [10.1016/j.inffus.2022.12.014](https://doi.org/10.1016/j.inffus.2022.12.014).
- [55] G. Hu, W. He, C. Sun, H. Zhu, K. Li, and L. Jiang, "Hierarchical belief rule-based model for imbalanced multi-classification," *Expert Syst. Appl.*, vol. 216, Apr. 2023, Art. no. 119451, doi: [10.1016/j.eswa.2022.119451](https://doi.org/10.1016/j.eswa.2022.119451).
- [56] Z. Huang, Y. Tang, and Y. Chen, "A graph neural network-based node classification model on class-imbalanced graph data," *Knowl.-Based Syst.*, vol. 244, May 2022, Art. no. 108538, doi: [10.1016/j.knsys.2022.108538](https://doi.org/10.1016/j.knsys.2022.108538).
- [57] B. Chen, Y. Fan, W. Lan, J. Liu, C. Cao, and Y. Gao, "Fuzzy support vector machine with graph for classifying imbalanced datasets," *Neurocomputing*, vol. 514, pp. 296–312, Dec. 2022, doi: [10.1016/j.neucom.2022.09.139](https://doi.org/10.1016/j.neucom.2022.09.139).
- [58] R. Zhao, X. Chen, Z. Chen, and S. Li, "Diagnosing glaucoma on imbalanced data with self-ensemble dual-curriculum learning," *Med. Image Anal.*, vol. 75, Jan. 2022, Art. no. 102295, doi: [10.1016/j.media.2021.102295](https://doi.org/10.1016/j.media.2021.102295).
- [59] S. Suh, H. Lee, P. Lukowicz, and Y. O. Lee, "CEGAN: Classification enhancement generative adversarial networks for unraveling data imbalance problems," *Neural Netw.*, vol. 133, pp. 69–86, Jan. 2021, doi: [10.1016/j.neunet.2020.10.004](https://doi.org/10.1016/j.neunet.2020.10.004).
- [60] Z. Zheng, Z. Yu, Y. Wu, H. Zheng, B. Zheng, and M. Lee, "Generative adversarial network with multi-branch discriminator for imbalanced cross-species image-to-image translation," *Neural Netw.*, vol. 141, pp. 355–371, Sep. 2021, doi: [10.1016/J.NEUNET.2021.04.013](https://doi.org/10.1016/J.NEUNET.2021.04.013).
- [61] X. Fei, S. Zhou, X. Han, J. Wang, S. Ying, C. Chang, W. Zhou, and J. Shi, "Doubly supervised parameter transfer classifier for diagnosis of breast cancer with imbalanced ultrasound imaging modalities," *Pattern Recognit.*, vol. 120, Dec. 2021, Art. no. 108139, doi: [10.1016/j.patcog.2021.108139](https://doi.org/10.1016/j.patcog.2021.108139).
- [62] R. Gu, G. Wang, T. Song, R. Huang, M. Aertsen, J. Deprest, S. Ourselin, T. Vercauteren, and S. Zhang, "CA-Net: Comprehensive attention convolutional neural networks for explainable medical image segmentation," *IEEE Trans. Med. Imag.*, vol. 40, no. 2, pp. 699–711, Feb. 2021, doi: [10.1109/TMI.2020.3035253](https://doi.org/10.1109/TMI.2020.3035253).
- [63] Y. Li, Z. Wang, L. Yin, Z. Zhu, G. Qi, and Y. Liu, "X-Net: A dual encoding-decoding method in medical image segmentation," *Vis. Comput.*, vol. 39, no. 6, pp. 2223–2233, Jun. 2023, doi: [10.1007/s00371-021-02328-7](https://doi.org/10.1007/s00371-021-02328-7).
- [64] A. Vyas, A. Katharopoulos, and F. Fleuret, "Fast transformers with clustered attention," in *Proc. Adv. Neural Inf. Process. Syst.*, 2020, pp. 1–10.
- [65] Z. Huang, X. Wang, L. Huang, C. Huang, Y. Wei, and W. Liu, "CCNet: Criss-cross attention for semantic segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 603–612, doi: [10.1109/ICCV.2019.00069](https://doi.org/10.1109/ICCV.2019.00069).
- [66] Y. Xu, X. He, G. Xu, G. Qi, K. Yu, L. Yin, P. Yang, Y. Yin, and H. Chen, "A medical image segmentation method based on multi-dimensional statistical features," *Frontiers Neurosci.*, vol. 16, Sep. 2022, doi: [10.3389/fnins.2022.1009581](https://doi.org/10.3389/fnins.2022.1009581).
- [67] H. Zhu, B. Wang, X. Zhang, and J. Liu, "Semantic image segmentation with shared decomposition convolution and boundary reinforcement structure," *Appl. Intell.*, vol. 50, no. 9, pp. 2676–2689, Sep. 2020, doi: [10.1007/s10489-020-01671-x](https://doi.org/10.1007/s10489-020-01671-x).
- [68] Q. Zhu, B. Du, and P. Yan, "Boundary-weighted domain adaptive neural network for prostate MR image segmentation," *IEEE Trans. Med. Imag.*, vol. 39, no. 3, pp. 753–763, Mar. 2020, doi: [10.1109/TMI.2019.2935018](https://doi.org/10.1109/TMI.2019.2935018).
- [69] S. A. Taghanaki, Y. Zheng, S. K. Zhou, B. Georgescu, P. Sharma, D. Xu, D. Comaniciu, and G. Hamarneh, "Combo loss: Handling input and output imbalance in multi-organ segmentation," *Comput. Med. Imag. Graph.*, vol. 75, pp. 24–33, Jul. 2019, doi: [10.1016/J.COMPMEDIMAG.2019.04.005](https://doi.org/10.1016/J.COMPMEDIMAG.2019.04.005).
- [70] M. Yeung, E. Sala, C.-B. Schönlieb, and L. Rundo, "Focus U-Net: A novel dual attention-gated CNN for polyp segmentation during colonoscopy," *Comput. Biol. Med.*, vol. 137, Oct. 2021, Art. no. 104815, doi: [10.1016/j.combiomed.2021.104815](https://doi.org/10.1016/j.combiomed.2021.104815).



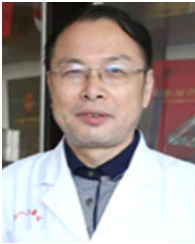
LIJUAN CUI was born in Shanxi, in 1994. She received the degree from the Taiyuan University of Technology. She has been a computer science and technology professional for many years and recently focused on big medical data research. Her research interests include medical image analysis and intelligent diagnosis.



DENGAO LI received the degree from the Taiyuan University of Technology. He is currently a Doctoral Supervisor and a Professor with the School of Big Data, Taiyuan University of Technology. He established the Key Laboratory of Big Data Fusion Analysis and Shanxi Province and the Intelligent Perception Engineering Technology Center of Shanxi. His research interests include big data fusion analysis and intelligent perception engineering.



CHAO LIU was born in Shandong, in 1989. He received the degree from the Shanxi University of Optical Engineering, in 2018. He is currently pursuing the Ph.D. degree with Shanxi Medical University. His research interests include biomedical photonics and biomedical imaging.



XIAOFENG YANG received the degree from the Hunan Medical University of Surgery, in 1998. He is currently the Chief Physician with the Department of Urology, First Hospital of Shanxi First Medical University. He is also the Doctoral Supervisor with Shanxi Medical University. He established Shanxi Siensi Company, focusing on the development of precision surgical medical equipment. He is also the Head of the Biomedical Engineering Research Center. His research interests include precision surgery and optical molecular images.



XIAOTING YAN was born in Shanxi, in 1982. He received the degree from the Shanxi Medical University of Clinical Medicine, in 2012.

He was with the First Hospital of Shanxi Medical University. His research interests include optical molecular imaging surgery, treatment techniques using nonthermal atmospheric pressure plasmas, fundamental study of plasma sources, and fabrication of micro- or nanostructured surfaces.

Mr. Yan's awards and honors include the Frew Fellowship (Australian Academy of Science), the I. I. Rabi Prize (APS), the European Frequency and Time Forum Award, the Carl Zeiss Research Award, the William F. Meggers Award, and the Adolph Lomb Medal (OSA).

...