

A Two-Phase Methodology leveraging X-DenseNet for Multi-Organ Segmentation in Abdominal CT Scans

Harinder Kaur¹, Navjot Kaur², Nirvair Neeru³

¹Research Scholar, ^{2,3}Assistant Professor

Department of Computer Science and Engineering, Punjabi University Patiala, 147002 Punjab, India
narain1293@gmail.com¹, navjot@pbi.ac.in², nirvair.ce@pbi.ac.in³

Article History:

Received: 06-08-2024

Revised: 16-09-2024

Accepted: 24-09-2024

Abstract:

Introduction: Multi organ segmentation (MOS) is crucial for various medical applications such as disease diagnosis, treatment planning, and surgery. Manual segmentation of CT organs is inefficient, laborious and time-consuming, necessitating the development of automated techniques. To address these challenges, automatic techniques based on deep learning models are explored in this field. These techniques require improvements to deal with patient-to-patient variability in organ size, location, and shape. Despite the transformative potential of deep learning techniques, research studies often inadequately address low contrast and overlapping structures within CT scans. This paper proposes a novel two phase methodology to integrate deep learning models and image enhancement strategies.

Methods: Our framework contributes in two stages named (a) Optimized Contrast Limited Adaptive Histogram Equalization-Weighted Grey Wolf Optimization (optiCLAHE-Weighted GWO), where the organs contrast is enhanced by employing weighted GWO for CLAHE clip limit selection (b) X-DenseNet architecture, where segmented regions of interest are produced with X-DenseNet model from the enhanced CT.

Results: For validation of proposed multi organ segmentation (PMOS) technique, experimentation and comparative analysis with existing models has been conducted on FLARE 22 challenge dataset. The Dice Score (DSC) of liver, aorta and spleen for proposed Multi organ segmentation (PMOS) approach is 90.4%, 99.38% and 98.33% respectively. The mean of Precision (Pre), Accuracy (Acc), F-score and DSC of the proposed approach are 95.05%, 95.55%, 95.29% and 96.06% respectively.

Keywords: Multiorgan Segmentation (MOS), Optimized Contrast Limited Adaptive Histogram Equalization (optiCLAHE), Grey Wolf Optimization (GWO), X-DenseNet

1. Introduction

Accurate organ segmentation in medical imaging plays a pivotal role in computer aided diagnostic (CAD) systems, clinical practices and treatment planning [1]. This enables radiologists to extract information such as area, size, shape, location and spatial relationships of organs. The detailed anatomical structure data produced from segmented images facilitates personalized treatments based on patient's anatomy. However, the labor-intensive and impractical manual segmentation necessitates the exploration of automated techniques. This not only significantly increases efficiency but also minimizes the susceptibility to errors, paving the way for more accurate and reliable organ segmentation.

The automated methods in this field has witnessed significant evolution, transitioning from traditional techniques like atlas-based methods, statistical shape models, and graph cut [2] to the dominance of deep learning models. Fully convolutional neural networks (FCNNs) like U-Net and V-Net have played a crucial role, with their encoder-decoder architecture effectively capturing global

context and high-level features [3], [4]. Furthermore, transformer architectures highlight the ongoing efforts to enhance model generalization and context understanding. Despite significant advancement in deep learning models, it has been analyzed that two major challenges persist in this field. Firstly, low contrast, poor image quality, overlapping objects of CT scans makes it problematic to segment organs and surrounding tissues, impacting their ability to capture intricate details. Secondly, it is difficult to detect small organs which lead to degradation in accuracy of deep learning models.

This study addresses these limitations by proposing a novel two-phased methodology. Each phase tackles a distinct aspect of the problem statement. The first phase is inspired from the review study [5], where a variation of CLAHE [6] gave optimum results. This stage incorporates an optimized Contrast Limited Adaptive Histogram Equalization (optiCLAHE) technique implemented with a novel weighted Grey Wolf Optimization (GWO) algorithm, collectively referred to as optiCLAHE-weighted GWO [7]. The proposed optiCLAHE-weighted GWO algorithm dynamically selects an optimal clip limit for CLAHE by utilizing weighted GWO, with the objective of maximizing entropy and enhancing CT scans. The second phase utilized dense blocks in X-net architecture hence, named X-DenseNet to segment liver, aorta and spleen from abdomen CT scan. By focusing on efficient feature extraction, feature reuse, facilitated information propagation and mitigating training challenges, dense blocks enhance the overall capability of the architecture to accurately segment organs. The distinctive structure of dense blocks improves segmentation performance not only for large organs like liver and spleen but also for smaller one e.g. aorta in this study.

The major objective of our research is to design an approach based on encoder-decoder architecture for better dice score and automatic segmentation of multiple organs. The contributions of the study are as given below:

- i. A novel approach to optimize and automate the clip selection limit of CLAHE algorithm by utilizing proposed weighted GWO algorithm. This will enhance the contrast by preserving organ shape and boundaries.
- ii. A dynamic weighting factors module is added to GWO in order to balance the dominant behavior of alpha, beta and delta wolves on other search agents. This will help the search agents to explore the more diverse search space.
- iii. The second phase of the proposed approach will incorporate three layer dense blocks to capture the more intrinsic features of the organs in order to improve the dice score. The optiCLAHE-weighted GWO is integrated with X-DenseNet architecture.
- iv. Our proposed approach shows superior performance for spleen and aorta when compared to the existing state-of-the-art methods. The mean dice score achieved for liver, aorta and spleen with the proposed approach is approx. 96%.
- v. This study also sheds light on limitation of trained model along with future directions in this field.

The remainder of the study is organized as follows: Section 2 provides the related work in the paper. Section 3 introduces the proposed methodology where in Section 3.1 details the first phase named optiCLAHE – Weighted GWO, Section 3.2 introduces the X-DenseNet architecture. Section 4 presents dataset, experimental setup, implementation details and comparison with state-of-the-art

methods. Section 5 demonstrated the ablation studies. Our study ends with detailed discussion along with findings, limitations and future directions in Section 6.

2. Related Work

The recent advancements in deep learning methods have overshadowed traditional medical image segmentation techniques such as atlas based methods [8], [9], [10], statistical shape models [11], [12], [13], [14], [15], [16], [17], graph cut [2], CRFs [18] etc. This shift shows the effectiveness and adaptability of deep learning methods in this field. While conventional techniques directly segment CT images with hand crafted features, whereas end-to-end data-driven learning models prove more suitable by facilitating automatic feature extraction for segmentation. Deep learning models such as CNNs, FCNs [19] excel in capturing detailed spatial features with their prior knowledge, crucial for delineating complex organs. However, the considerable diversity in anatomical structures across different patients and low intensity contrast of CT scans, poses a significant challenge in multi-organ segmentation [20].

FCNNs, including U-Net [21] and V-Net [22], have dominated medical segmentation, with encoders capturing global context through downsampling, and decoders upscaling for precise predictions via skip connections. The FCN majority voting method [23], a variant of FCN, leverages redundant segmentations to glean both high and low-level features like shape, size, and contour of organs. But the redundant segmentation require significant computational resources lead to new technique based on time implicit levels [24]. Another variation, FCN-DecNet [23] integrates FCN with multi scale weighted Probabilistic atlas (PA) and utilized convolution and deconvolution layer to achieve optimized results. But this study is semi-automated. A fully automated approach is developed [25] utilizing multi-atlas techniques for organ localization and employing a two-stage CNN for segmentation. There are other 3D variations such as 3D FCN [26], cascaded 3D FCN [27], 3D- U-JAPA-net [28] and Federated 3D [29] showcasing the versatility in this domain.

Dense V Net [13] performed extremely well on both small and large abdomen organs particularly emphasizing organs linked to the gastrointestinal tract, a critical aspect for navigation during endoscopic procedures. The contribution involved upsampling and downsampling with skip connections and batch wise spatial drop out. In a related study, deep multi-planar co-training stands out as it harnesses the full potential of unlabeled data to segment overlapped organs and delineate boundaries of organs [30]. Shape-Guided Ultralight Network (SGU-Net), a novel approach designed for extremely efficient computational performance. SGU-Net is an ultra light convolution method that aims to strike a balance between computational efficiency and segmentation accuracy, making it particularly well-suited for deployment in resource-constrained hardware environments [31]. Self-supervised learning, eliminating manual annotations, led to the self-supervised attention UNet for OAR segmentation [32], [33]. There are another U-Net based networks proposed named Residual U-net[34], nnUnet[35] etc. Further, deep learning based region growing technique is presented by [36]. FCNNs excel in medical image segmentation but struggle with long-range spatial context. Inspired by transformer success in NLP, numerous studies has been conducted in research such as Axial Fusion Transformer UNet [37], UNet TTransformers (UNETR), retaining the "U-shaped" architecture [38]. A study proposed the Cross-Convolutional Transformer Network (C2Former) as a solution for image segmentation challenges. A central innovation is the redesign of a cross-convolutional self-

attention mechanism that effectively integrates both local and global context information while modeling dependencies, thus enhancing the model's ability to understand the semantic features within images [39].

Inspired from the success of the U-Net encoder-decoder architecture, X-Net is utilized as baseline in the proposed work. X-net was never explored for this task and this study is the first to apply X-Net to abdominal organ segmentation. The accurate segmentation of small organs remains a significant challenge, which leads to performance degradation of the model. In PMOS model, additional feature extraction stage is provided and it avoids over fitting using L2 regularization. X-net excels in capturing spatial relationships within images, which is of paramount importance in multi-organ segmentation for accurately outlining organ boundaries. Secondly, its deep architecture enables the learning of hierarchical features from medical images [40]. Our study has incorporated three layer dense blocks to the X-Net hence named as X-DenseNet. The more detailed architecture is discussed in section 3.2. The comparison of different network variants of the proposed study is presented in ablation study section 5 where in dense blocks came out as critical part of the architecture.

3. Methods

This section proposes a two phased approach for multiorgan segmentation for abdomen CT images. Fig 1 shows the schematic diagram of the two phases of the proposed methodology. The two phases are (a) optiCLAHE – Weighted GWO (b) X-DenseNet. The more detailed methodology of two phases is explained in section 3.1 and section 3.2.

3.1 Optimized Contrast Limited Adaptive Histogram Equalization-Weighted Grey Wolf Optimization (optiCLAHE-Weighted GWO)

The low contrast in medical images poses a challenge for accurate multi-organ segmentation [4]. While various methods exist to address this issue, histogram-based techniques, particularly Contrast Limited Adaptive Histogram Equalization (CLAHE), have proven to be notably effective in enhancing contrast across diverse medical image anatomies [41]. CLAHE distinguishes itself from its predecessor, Adaptive Histogram Equalization (AHE), by thwarting excessive contrast amplification and noise in the image. This is accomplished by locally partitioning the image into tiles or sections (tn) and imposing a clip limit (CL) on each of them. The efficacy of CLAHE is influenced by several parameters, including the number of tiles, clip limit, and the redistribution of intensity range. A low clip value results in gradual slope of the equalized histogram, resulting in a modest enhancement of contrast. Conversely, a high clip limit sidesteps the redistribution of histogram bins but becomes more prone to noise [6]. Thus, it is crucial to make a careful selection of the clip limit.

This study focuses on automatic selection of the optimized clip limit. The search space, which refers to the range of possible solutions for clip limit, is vast and intricate, with numerous potential optimal points. This is where meta-heuristics, versatile problem-solving strategies, come into play. These methods begin with a diverse initial set of potential solutions. As the process unfolds through iterations, these solutions are refined and improved. Population-based algorithms, such as Swarm Intelligence (SI) algorithms, are particularly adept at this. They retain and utilize crucial information about the search space throughout the iterative process [7]. In this research study, SI algorithm named Grey Wolf is improved by focusing more on better search space, information preservation and convergence speed. Fig. 2g 2 depicts the schematic diagram of first phase of the methodology. The steps in the optiCLAHE-Weighted GWO algorithm are given below:

Step 1-Initialization. The fundamental parameter configurations in this experiment for weighted-GWO are as given below: Maximum number of iterations (T_{max}) is set to 100, No of wolves or search agent is designated as 50 and the dimensionality is established as 1. The lower bound (lb) is defined as 0; while the upper bound (ub) is stipulated as 100. In each dimension, stochastic values within the range of 0 to 1 are generated for all search agents. This yields search agents with random initial positions situated within the specified bounds, thereby setting the stage for the optimization process. Hence initialize the search agents or grey wolves say $X_i = 1 \dots n$ and coefficients a , A and C .

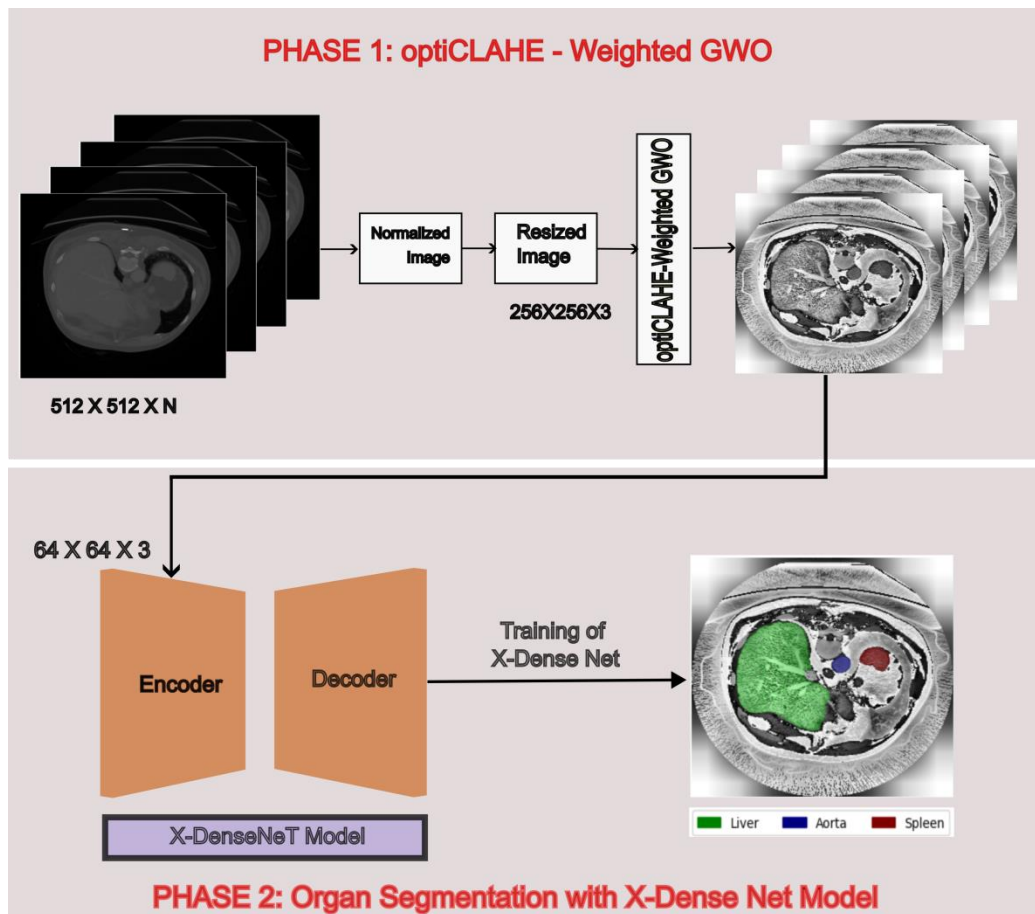


Fig. 1 Schematic diagram of proposed approach for multiorgan segmentation of Liver, Aorta and Spleen

Step 2- Objective Function. The objective function of this experiment is to maximize the fitness value, aiming to identify a set of parameters that yield the highest amalgamation of contrast and information content (entropy) in the equalized image. In essence, it seeks to achieve equilibrium between contrast enhancement and information preservation, as quantified by entropy. This algorithm leads to the maximization of contrast improvement, all while mitigating the potential hazard of **excessive** amplification of noise in the equalized image.

$$\vec{F} = \frac{1}{(\text{contrast} + \text{entropy})} \quad (1)$$

For each wolf or search agent, the fitness of the current position is evaluated using the objective function. Subsequently, the position of alpha ($\vec{X}_\alpha$), beta ($\vec{X}_\beta$) and delta wolf $\vec{X}_\delta$, is decided based on

the calculated fitness among all the wolfs. $\vec{F}_\alpha, \vec{F}_\beta, \vec{F}_\delta$ will store the fitness values of alpha, beta and delta wolves.

Step 3- Update Positions. The positions $\vec{X}_1, \vec{X}_2, \vec{X}_3$ are calculated based on the coefficients and positions of the alpha, beta, and delta wolves which are the leaders in the wolf pack hierarchy according to their fitness values.

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_\alpha \cdot (\vec{D}_\alpha) \quad (2)$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_\beta \cdot (\vec{D}_\beta) \quad (3)$$

$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_\delta \cdot (\vec{D}_\delta) \quad (4)$$

Where $\vec{X}_\alpha, \vec{X}_\beta, \vec{X}_\delta$ variables will hold the positions of the alpha, beta, and delta wolves, respectively. $\vec{D}_\alpha, \vec{D}_\beta, \vec{D}_\delta$ are the vectors representing the absolute differences between the current position (parameter set) and the positions of the alpha, beta, and delta wolves respectively. These vectors indicate the direction and magnitude of the update and calculated as follows:

$$\vec{D}_\alpha = |\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}| \quad (5)$$

$$\vec{D}_\beta = |\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}| \quad (6)$$

$$\vec{D}_\delta = |\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}| \quad (7)$$

$$\vec{C} = 2 \cdot \vec{r}_2 \quad (8)$$

where $\vec{C}_1, \vec{C}_2$ and $\vec{C}_3$ are coefficients calculated based on random variable $\vec{r}_2$.

Proposed modified equations to update wolf position: The coefficient 'A' influences the hunting behavior of the pack. The wolf positions $\vec{X}_1, \vec{X}_2$ and $\vec{X}_3$ are computed based on $\vec{A}_\alpha, \vec{A}_\beta, \vec{A}_\delta$ instead of only $\vec{A}$ unlike GWO. The coefficient $\vec{A}$ will be calculated as follows:

$$\vec{A}_\alpha = 2\vec{a}_\alpha \cdot \vec{r}_1 - \vec{a}_\alpha \quad (9)$$

$$\vec{A}_\beta = 2\vec{a}_\beta \cdot \vec{r}_1 - \vec{a}_\beta \quad (10)$$

$$\vec{A}_\delta = 2\vec{a}_\delta \cdot \vec{r}_1 - \vec{a}_\delta \quad (11)$$

The coefficients a_α, a_β , and a_δ control the rate at which the alpha, beta, and delta wolves respectively influence the search agents. By modifying these parameters based on the current iteration T and the maximum number of iterations T_{max} , the algorithm dynamically adjusts the exploration and exploitation phases. This helps in achieving a better balance between exploration (utilizing information from other wolves) and exploitation (utilizing information from the best wolves) of the search space. Early in the optimization process, there is more emphasis on exploration, while in later stages; there is more focus on exploitation. This adaptability can lead to faster convergence towards an optimal solution. Also, exponential region (maximum curvature) is the potential candidate for clip limit value [39]. The sinusoidal function $\sin(1/T_{max})$ in the equation for a_δ introduces a non-linear component that can enhance the robustness of the algorithm.

$$a_\alpha = 2 \left(1 - \frac{1}{T_{max}} \right) \quad (12)$$

$$a_\beta = 2 \left(1 - \frac{1}{(T_{max})^2} \right) \quad (13)$$

$$a_\delta = 2 \left(1 - \sin \left(\frac{1}{T_{max}} \right) \right) \quad (14)$$

Dynamic weights are calculated based on fitness value. Each calculated position is then combined with weights $\vec{W}_1, \vec{W}_2, \vec{W}_3$ and the positions. The current position of the search agent is updated using the weighted combination. The equations calculate weighting factors based on the fitness values of

the alpha, beta, and delta wolves relative to the best fitness value in the population. This allows the algorithm to favor the positions of wolves that have performed well, potentially guiding the search towards promising areas of the solution space. The weights $\vec{W}_1, \vec{W}_2, \vec{W}_3$ ensure that the influence of the alpha, beta, and delta wolves is balanced. This helps to avoid situations where one wolf dominates the influence on the search agents, promoting a more diverse exploration of the solution space. The utilization of dynamic weighting factors reduces the sensitivity of the algorithm to hyper-parameter tuning.

$$\vec{W}_1 = \frac{\vec{F}_\alpha}{F_{Best}} \quad (15)$$

$$\vec{W}_2 = \frac{\vec{F}_\beta}{F_{Best}} \quad (16)$$

$$\vec{W}_3 = \frac{\vec{F}_\delta}{F_{Best}} \quad (17)$$

$$\vec{X}(t+1) = \vec{X}_1\vec{W}_1 + \vec{X}_2\vec{W}_2 + \vec{X}_3\vec{W}_3 \quad (18)$$

Step 4 – Repeat: The algorithm rigorously assesses whether the current search agent surpasses the fitness levels of the alpha, beta, and delta wolves. In the event of superiority, the respective wolf's position and fitness records are updated. This iterative loop persists for the designated number of iterations, during which it systematically refines the positions of the search agents predicated on their individual fitness values and the coordinates of the alpha, beta, and delta wolves. The fitness metric of the alpha wolf ($\vec{F}_\alpha$) at the present iteration is meticulously logged in the Convergence Curve. Ultimately, the conclusive alpha position ($\vec{X}_\alpha$), its corresponding fitness reading ($\vec{F}_\alpha$), and the progression pathway ('Convergence_curve') collectively constitute the algorithm's output. As a result, the derived ($\vec{X}_\alpha$) value stands poised to serve as the optimal clip limit for the subsequent image enhancement phase.

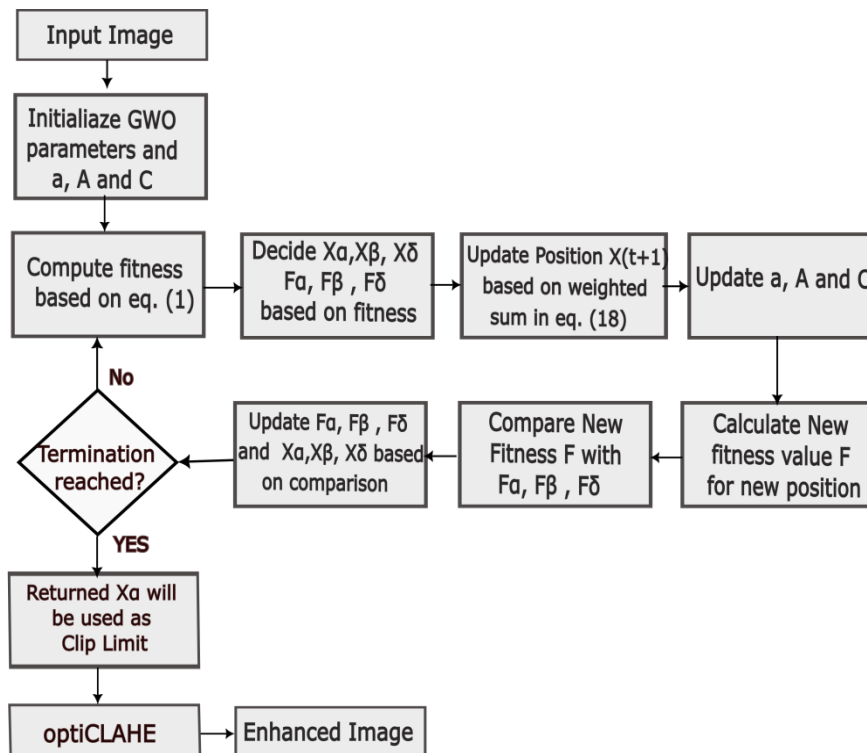


Fig. 2 Flow Diagram of first phase of i.e. optiCLAHE - Weighted GWO

3.2 X-Dense Net Architecture

This phase aims to segment multiple organs from abdomen CT scans. In this phase, the 2D CT enhanced axial slices generated in the previous phase are taken as input and 2D segmented axial slices are produced as outputs. This includes two steps: (i) Encoder Stage (ii) Decoder Stage. Initially, CT slices are enhanced with optiCLAHE - weighted GWO and the generated input is passed to the X-DenseNet architecture. The enhanced images have better entropy information. The detailed encoder and decoder stage of the model is given below:

(i) Encoder Stage

The primary goal of encoder stage is to extract fundamental features such as edges and gradients. It takes input images with dimensions of 64x64x3 and processes them via a series of convolutional layer. There are 15 convolutional layer in both encoder and decoder network with kernel size of 3 x 3. We have designated the combination of downsampling layers as block D, where Batch Normalization is employed after each convolutional operation to stabilize and accelerate the training process. This ensures that the model's parameters remain well-conditioned throughout training. Then, activation function, specifically ReLU (Rectified Linear Unit), is utilized to introduce non-linearity, aiding the model in learning complex relationships within the data. Following, Max Pooling layer with pool size of 2 x 2 is strategically placed to down-sample the feature maps, reducing spatial dimensions while retaining essential information. This allows the model to capture higher-level features in subsequent layers efficiently [42].

Inspired from [13] and [43], Dense Convolutional Network, or DenseNet layer is added to the architecture after max-pool layer. In dense block structure, one convolution layer is intricately connected to every subsequent convolution layer passing feature maps in a feed-forward manner, creating a highly interlinked network structure in dense block. The three-layer dense block is utilized and the schematic diagram for the same is shown in Fig 4. They effectively address the vanishing-gradient problem that has plagued deep neural networks, facilitating the unhindered flow of gradients during training. Additionally, they promote robust feature propagation, encouraging the seamless transfer of information across layers. This design also fosters the efficient reuse of features, thereby reducing redundancy and substantially curtailing the total number of model parameters. The filter depth for convolutions layers and dense block is (64,128,256,512). Subsequently, the model employs multi-resolution dense blocks, which are crucial in capturing features at different scales and able to capture the complex relationship and features of the organs. These blocks consist of densely connected convolutional layers, promoting effective feature reuse and propagation. It is designed to optimize the concatenation of features, ensuring that multiple copies of feature maps are not needlessly stored, ultimately achieving memory efficiency. In scenarios where memory constraints become a factor and retention of limited activation maps in memory, earlier layer information is stored only once but remains accessible for subsequent layers. These dense blocks help to mitigate memory overhead. Down-sampling is achieved through max-pooling layers after each dense block, allowing the model to focus on higher-level features. The n down-sampling blocks i.e. D blocks with feature map height h and width w can be denoted as follows:

$$D_{h,w} = \{D_{h,w}^0, D_{h,w}^2, \dots, D_{h,w}^n\} \quad (19)$$

where height and width of feature map is dependent on input feature map to the block (h_{input} , w_{input}) defined as given below:

$$\mathbf{h} = \frac{\mathbf{h}_{\text{input}}}{2}, \mathbf{w} = \frac{\mathbf{w}_{\text{input}}}{2} \quad (20)$$

It can be seen in architecture proposed in Fig 3 that down-sampling can be achieved through {D-dense blocks} which uses same equation to calculate the feature map size.

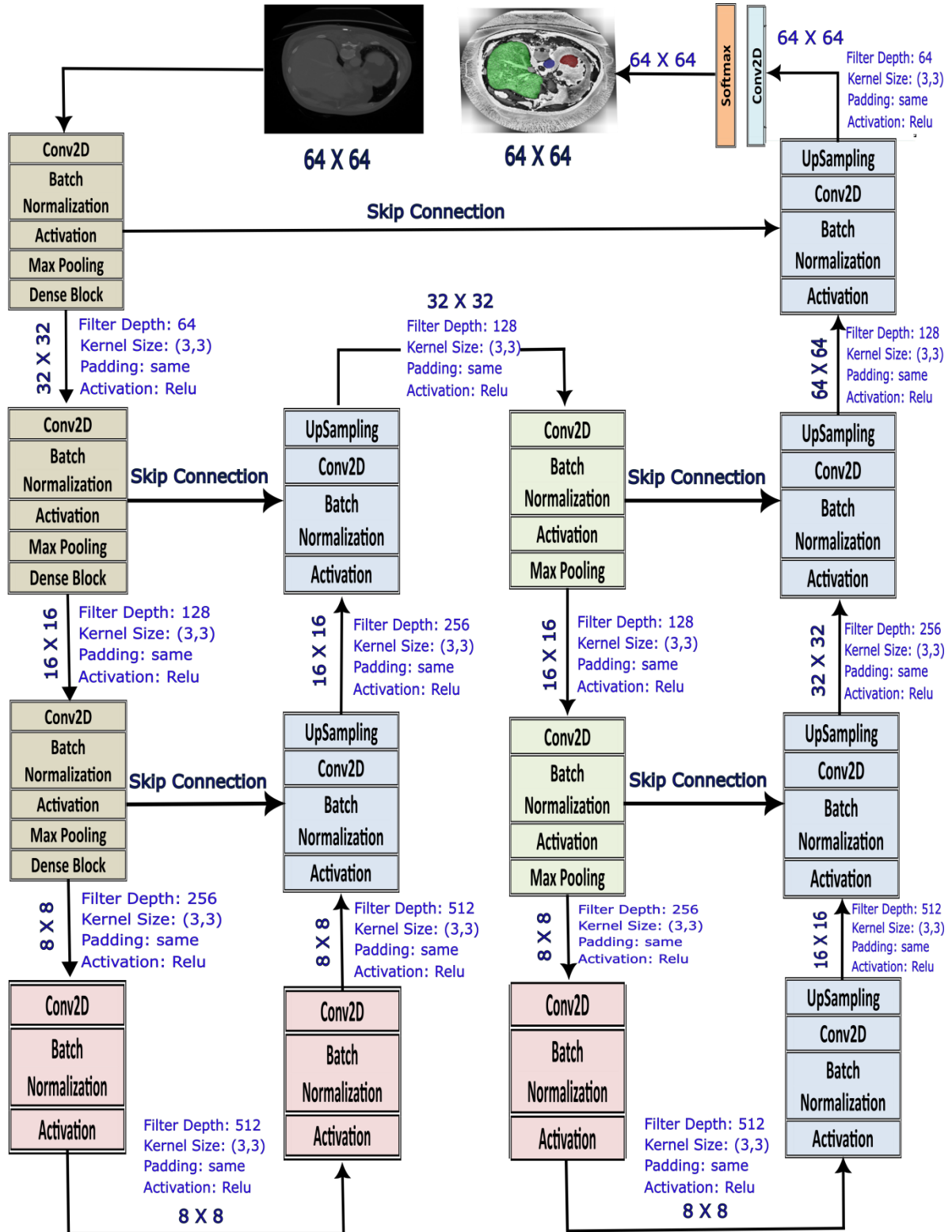


Fig. 3 Network Architecture of Proposed X-DenseNet Model

(ii) Decoder Stage

In the decoder stage, the model reconstructs the segmented regions based on the features extracted by the encoder. The N blocks dedicated for upsampling are $U_{H,W} = \{U_{H,W}^1, U_{H,W}^2, U_{H,W}^3 \dots U_{H,W}^N\}$. It starts with up-sampling layers to restore spatial information lost during the downsampling in the encoder. The concatenation of features from different resolutions is employed to preserve fine-grained details. This block comprises upsampling layer, then convolutional layer with same filter and kernel size as in encoder stage. These layers are followed by batch normalization and relu activation.

Then skip connections provides the feature concatenation and fine graded details. The decoder concludes with a 1×1 convolutional layer followed by a softmax activation, generating the final segmentation mask. The detailed training hyper-parameters are presented in Table 1. The provided skip connections will allow the low level features maps generated in encoder stage to flow through the network in later stages. The detailed features maps size and skip connections are presented in Fig 3. The height and width of feature maps of U blocks can be calculated as follows:

$$H = h_{input} * 2, \quad W = w_{input} * 2 \quad (21)$$

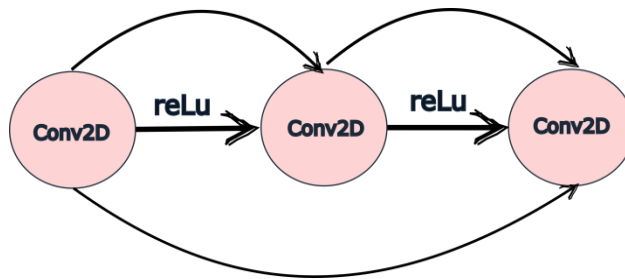


Fig. 4 Schematic Representation of three layer dense block

It is worth noting here that, X-DenseNet exhibits similar structural framework to Xnet [36], albeit with notable distinctions. Firstly, the dense blocks are the introduction in our proposed architecture unlike X-net model. Further feature map calculation is as given in encoder and decoder stage. Further, while the X-net model has demonstrated effectiveness when applied to X-ray images, we have adapted and modified the model for application on abdominal CT scans.

4. Experiments and Results

4.1 Dataset

Flare 22 dataset is the challenge dataset which is available in public domain. It consists of 50 abdomen CT scans [44], [45]. This dataset encompasses 13 distinct organs, namely: liver, spleen, pancreas, right kidney, left kidney, stomach, gallbladder, esophagus, aorta, inferior vena cava, right adrenal gland, left adrenal gland, and duodenum. Data augmentation is employed in order to extend the dataset. The Image Data Generator function from Keras facilitated rotations within a range of 10 degrees, as well as width and height shifts within a range of 0.1, along with zooming within a range of 0.2, and random horizontal and vertical flips. This augmented dataset is expanded with 500 images, subsequently partitioned into training and testing sets using a 7:3 ratio through a randomized split implemented with the 'train_test_split' function. For the purpose of this research, experiments were conducted specifically focusing on the liver, aorta, and spleen.

4.2 Experimental Setup and Model Training

Our PMOS technique is implemented with 11th Gen Intel core i5-1135G7, 8GB RAM, python 3.6 and training is implemented with T4 GPU with 16 GB RAM. During the training phase, categorical cross-entropy is employed as the loss function, suitable for multi-class segmentation tasks. The Adam optimizer, combined with a dynamic learning rate schedule, ensures efficient convergence during the training process. Training loss and accuracy per epoch is as shown in Fig 5.

Data Augmentation Methods	Rotation, width shift, height shift, zoom, horizontal flip and vertical flip
Initialization of the network	Glorot Uniform Intializer /Xavier Initialization
Batch size	16
Patch size	64 x 64 x 3
Total epochs	200
Kernel size	3 x 3
Optimizer	Adam
Initial learning rate	0.001
Stopping criteria, and optimal model selection criteria	When reached defined no of epochs (200)
Training loss	categorical cross-entropy
Training time	1-2 min/epoch

Table 1 Training Protocols

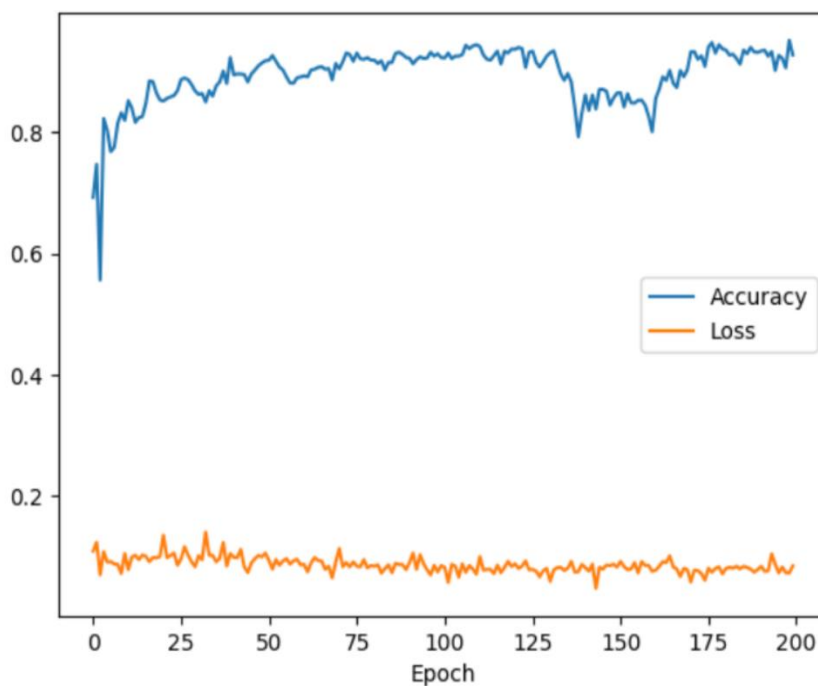


Fig. 5 Model Training History representing change in accuracy and loss with respect to epochs

4.3 Evaluation Metrics

In this research study, we have used various evaluation metrics named Dice Coefficient (DSC), Accuracy (Acc), Precision (Pr), Recall (Rec) and F-Score (FS) in order to confirm the segmentation results achieved by our proposed method. The equations of the aforementioned parameters are as given below:

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (22)$$

$$Pr = \frac{TP}{TP+FP} \text{ Range: } [0, 1] \quad (23)$$

$$Rc = \frac{TP}{TP+FN} \text{ Range: } [0, 1] \quad (24)$$

$$FS = 2 \times \frac{Precision \times Recall}{Precision + Recall} \text{ Range: } [0, 1] \quad (25)$$

$$DSC = \frac{2 \times Intersection}{Union + Intersection} \text{ Range: } [0, 1] \quad (26)$$

4.4 Comparison with state-of-the-art methods

In the literature, a variation of nnUnet known as Residual U-net has exhibited superior performance in literature. The incorporation of residual connections in nnUnet synergizes the architectural strengths of nnU-Net, particularly its context-preserving skip connections, with the benefits of residual connections, leading to enhanced accuracy and effectiveness in segmentation tasks [34]. A study [46] is proposed to deal with unlabeled data and enhance the model's generalization capabilities. The methodology involves generating pseudo labels through the utilization of the Swin Transformer. Subsequently, these pseudo labels are amalgamated with labeled data and input into the PHTrans framework for further processing. A study introduced a novel methodology termed 3D Cross-Pseudo Supervision (3D-CPS), which constitutes a semi-supervised network built upon nnU-Net, incorporating the Cross-Pseudo Supervision technique. Additionally, a dynamic adjustment of semi-supervised loss weights throughout epochs, aiming to promote linearity expansion [47]. A coarse-to-fine framework is employed to refine segmentation results in two sequential stages within a variation of nnUnet and USE-Net, termed Residual-USE-Net [43]. This innovative approach involves the integration of two teacher models and a student model to enhance the segmentation process [48]. The residual squeeze-and-excitation (SE) blocks are integrated to enhance the overall performance and efficiency of the architecture. All these studies have used FLARE 22 challenge dataset.

Reference	Network Architecture	Liver	Aorta	Spleen
[34]	Residual U-Net	0.9753	0.9566	0.9701
[46]	Self-Training and Hybrid Architecture	0.9761	0.9863	0.9572
[47]	3D-CPS	0.9733	0.9630	0.9202
[48]	Cross Teaching Teachers	0.9763±0.0154	0.9375±0.0469	0.9471±0.1386
	Proposed Approach	0.904670	0.993811	0.983357

Table 2 Dice Score comparison of competing methods with our proposed approach. Dice score of spleen and aorta is higher for our approach than competing methods.

The aforementioned four studies reported in literature are compared with the proposed approach for abdomen multi organ segmentation. The quantitative results of the proposed X-DenseNet architecture with the competing methods are presented in

Table ². The Dice Score of the proposed X-DenseNet for liver, aorta and spleen is 0.904, 0.9938, and 0.9833 respectively. The effectiveness of the proposed approach in segmenting the aorta and spleen surpasses that of state-of-the-art studies. Notably, the Dice Similarity Coefficient (DSC) for the aorta and spleen outperforms competing methods, although it is relatively lower for the liver. The proposed approach yields mean values of Precision (Pre), Accuracy (Acc), F-Score, and DSC as 95.05%, 95.55%, 95.29%, and 96.06%, respectively, as presented in Table 3.

Proposed Approach	Liver	Aorta	Spleen	Mean
Pr (%)	89.683130	98.769899	96.725836	95.059621
Acc (%)	88.944611	99.383046	98.349294	95.5589
F-Score (%)	89.280966	99.075524	97.530810	95.2957
DSC (%)	90.4670	99.3811	98.3357	96.06126

Table 3 Mean of evaluation metrics for liver, aorta and spleen for our proposed approach

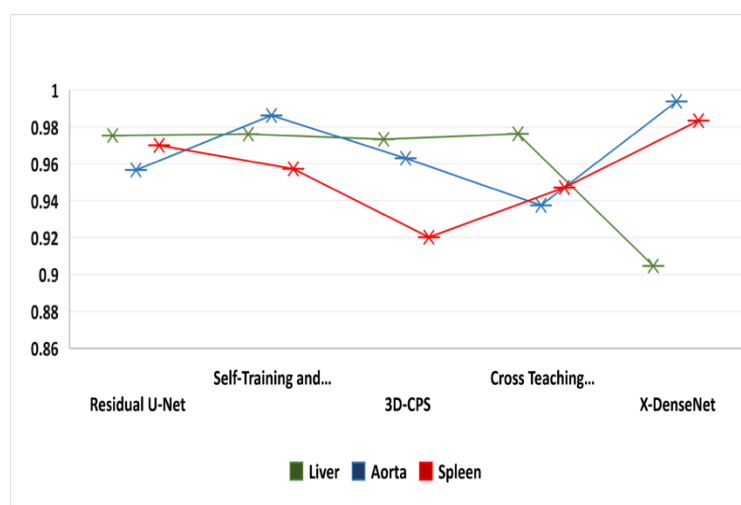


Fig. 6 Comparison of Dice Coefficient of different approaches with proposed X-DenseNet

5. Ablation Study

In our experiments, we have conducted ablation studies to assess the impact of various modules in our technique. Table 5 presents the dice score results of the different organs. For the sake of comparison, we adopted X-Net as our baseline, as we aimed to enhance its performance on our specific dataset. Specifically, we implemented the existing X-Net architecture without optiCLAHE-WeightedGWO on our dataset, resulting in dice scores of 0.2799, 0.97633, and 0.9759 for the liver,

aorta, and spleen, respectively. This baseline performance serves as a reference point for evaluating the enhancements we introduce in subsequent analyses.

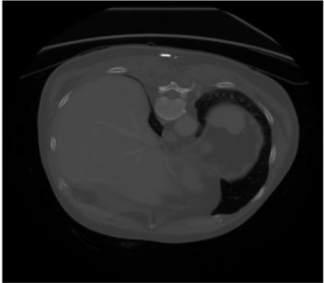
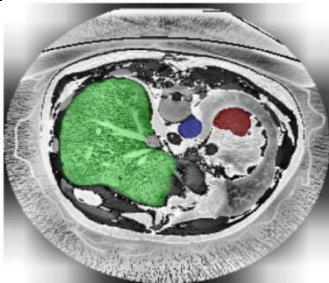
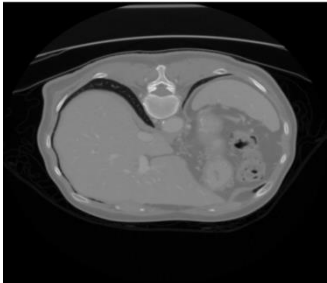
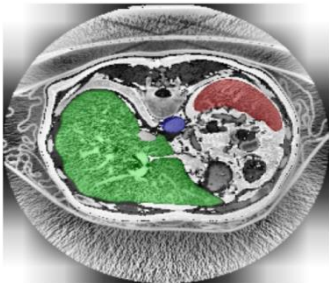
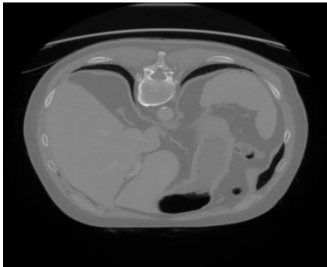
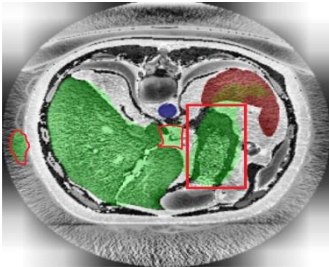
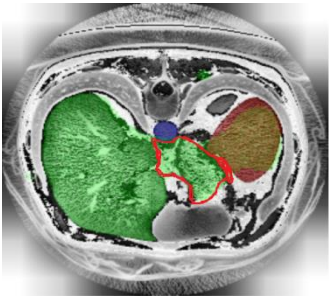
	Original Image	Segmented Image
True Positive (TP) and False Negative (FN)		
		
False Positive (FP)		
		

Table 4 Qualitative Results of Multiorgan Segmentation with Proposed Technique on 2D Axial Slices

Effect of optiCLAHE-WeightedGWO. The objective is to discern the impact of incorporating first phase into the existing X-Net framework. The results presented in the Table 5 clearly demonstrate that training the X-Net model on a dataset enhanced with optiCLAHE-Weighted GWO yields

superior performance compared to training the model without this enhancement algorithm. Specifically, the dice scores for the liver, aorta, and spleen exhibit significant improvement, reaching 0.481719, 0.985, and 0.981, respectively. These findings underscore the essential nature of the initial phase's inclusion in our methodology.

Experiments	Liver	Aorta	Spleen
X-Net	0.279913	0.976334	0.975982
X-Net with optiClahe- Weighted GWO	0.481719	0.985	0.981
First Phase + X-Net + 3 Dense Block on 100 iterations	0.633846	0.991	0.982
PROPOSED: First Phase +X-Net+3 Dense Block on 200 iterations	0.904670	0.993811	0.983357

Table 5 Quantitative Results of ablation study

Effect of Adding Dense Blocks. To gauge the significance of dense blocks in enhancing the X-Net architecture on a preprocessed dataset with optiCLAHE-Weighted GWO, a systematic exploration was conducted. Initially, one dense block was added incrementally until optimal results were obtained. Strikingly, the most favorable outcomes were achieved with the inclusion of three dense blocks. Beyond this point, the performance exhibited a noticeable decline upon introducing a fourth dense block. Consequently, after a series of experiments, the proposed X-denseNet with three dense blocks emerged as the architecture delivering superior results. This experimentation establishes the indispensability of employing three dense blocks for X-Net on our dataset. Further, experiments are performed with spatial dropouts to improve the complexity of architecture; however favorable outcomes are not achieved.

Effect of number of iterations. The appropriate number of epochs for segmentation model training depends on the dataset and model complexity, with a balance needed to avoid underfitting or overfitting. Regular monitoring of validation metrics, such as the Dice Similarity Coefficient (DSC) score, helps identify the optimal point for model convergence. We have performed experiments to get the best suitable count of iterations for this architecture and environmental variables. It has been observed that X-denseNet with 200 epochs yielded the most favorable results. While the number of epochs between 100 and 200 had a relatively minimal impact on the performance of the model concerning the aorta and spleen, a discernible effect was observed in the dice score for the liver. This underscores the critical role of epoch count in achieving optimal performance, particularly with respect to certain organs within our dataset.

6. Discussion

In this paper, a novel approach is proposed to accurately segment multiple organs from abdomen CT images. The experiments are conducted to segment liver, aorta and spleen.

6.1 Findings

The proposed approach consists of two phases (a) optiCLAHE-Weighted GWO (b) X-DenseNet. The proposed method, named optiCLAHE, employs a Swarm Intelligence (SI) algorithm called Grey Wolf Optimization (GWO) to automatically select an optimized clip limit for Contrast Limited Adaptive Histogram Equalization (CLAHE).

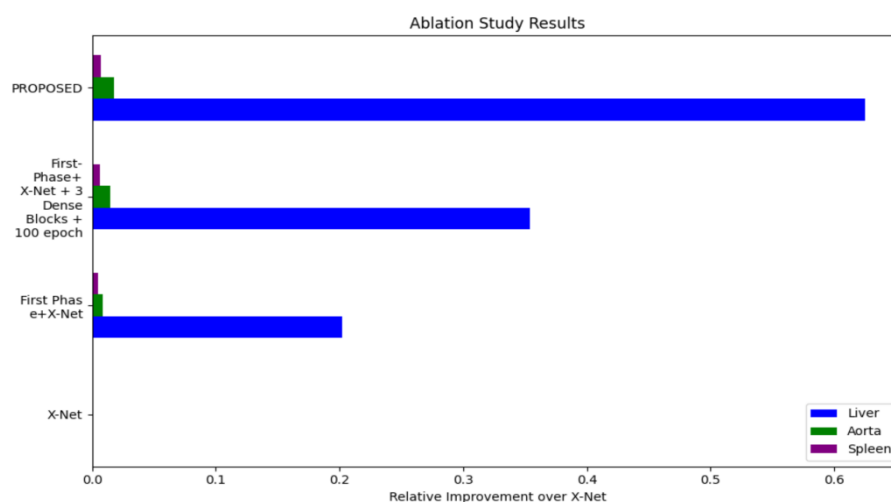


Fig. 7 Effect of different phases on the baseline X-Net

The paper then introduces segmentation architecture, X-DenseNet, which integrates enhanced images for multi-organ segmentation. We have utilized FLARE 22 Challenge dataset for training of our model. Further, we have used Dice Score (DSC), Precision (Pr), Accuracy (Acc) and F-Score (FS) to evaluate the proposed model. The DSC is used to perform the comparative study of existing studies in literature. The key steps of the first phase include initialization, objective function definition, updating positions based on the fitness values of alpha, beta, and delta wolves, and dynamic adjustment of weights for the search agents. The algorithm aims to strike a balance between exploration and exploitation in the search space, adapting its behavior based on the fitness values encountered during optimization. The experimental findings on architecture reveal three featured crucial modifications that significantly contribute to the improvement in Dice Score. Firstly, the changes proposed in coefficients a_α , a_β , and a_δ in first phase plays a pivotal role in achieving a more balance between exploration and exploitation of the search space. This adaptability facilitates to faster convergence towards an optimal solution. The sinusoidal function ' $\sin(1/T_max)$ ' introduces a non-linear component that can enhance the robustness of the algorithm. Secondly, incorporation of dynamic weights $\vec{W}_1, \vec{W}_2, \vec{W}_3$ allows the algorithm to prioritize the positions of wolves that have performed well, guiding the search towards promising areas of the solution space. The weights $\vec{W}_1, \vec{W}_2, \vec{W}_3$ ensure a balanced influence of the alpha, beta, and delta wolves. In the ablation study during the second phase, it becomes evident that there is room for improvement in the existing X-Net architecture when applied to our dataset.

Inspired form DenseNet a thought of introducing third crucial change i.e. dense block to existing X-Net, resulting in a notable improvement in the Dice Score. The model employs multi-resolution dense blocks, which are instrumental in capturing features at different scales and able to capture the complex relationship and features of the organs. This helps to detect small organs such as aorta in our study. Down-sampling is achieved through max-pooling layers after each dense block, allowing the model to focus on higher-level features. Combining the advancements from the first and second phases, the proposed approach yields mean values of Precision, Accuracy, F-Score, and Dice Similarity Coefficient (DSC) as 95.05%, 95.55%, 95.29%, and 96.06%, respectively, as presented in the table.

The relative improvement in dice score of liver, aorta and spleen with respect to the baseline X-Net is shown in Fig 7. It can be observed from the Figure 7 and Table 5 that the inclusion of both the phases impacts the DSC hence makes it mandatory for the required efficiency. The DSC score achieved for liver, aorta and spleen with the proposed model is 0.904670, 0.993811, and 0.983357 respectively. The results show that there is huge improvement from the baseline X-Net in DSC of liver however the percentage increase in performance for aorta and spleen is 1.79% and 0.76% respectively. Further, when the results are compared with the competing methods in Table 4 and Fig 6, it is noteworthy that competing methods such as [34], [46], [47] and [48] exhibit a superior Dice Similarity Coefficient (DSC) score for the liver compared to the proposed approach. However, the DSC scores for the aorta and spleen are higher in our proposed model than in the competing methods.

6.2 Limitations

The qualitative results of segmentation of liver, aorta and spleen are illustrated in Table 4. In Case 1 and 2, the True Positive (TP) and False Negative (FN) scenario is depicted, representing the best results where all three organs are segmented accurately. However, it is crucial to note that the proposed model, while promising, is not flawless. In Case 3 and 4, the False Positive (FP) situation is presented, with red squares and outline highlighting the FP pixels in the image. This indicates instances where the liver is not segmented accurately, and some pixels that should be negative are erroneously classified as positive. Consequently, these FP cases for the liver contribute to degradation in the mean dice of the proposed model. This observed limitation may arise from the architecture's inability to capture intricate features of the liver comprehensively. Considering a more advanced architecture, such as U-Net++, 3D U-Net, or similar, could potentially lead to an improvement in the Dice score by better capturing the detailed features of the liver. The experiments are conducted focusing on abdomen CT scans and three specific organs. The generalizability of the proposed approach to different datasets and organ systems needs to be explored. The experimental setup, including the use of T4 GPU, may limit the accessibility of the proposed methodology to researchers with less computational power.

6.3 Future directions

Firstly, our current study excludes pathological scans, such as those involving tumors or lesions. Further, more small organs such as pancreas, duodenum and esophagus can be included in the future researches because it is difficult to capture the complex shapes, structures and spatial relationships. In future work, we are actively engaged in developing methodologies to address the segmentation of these smaller organs. The existing literature highlights that techniques based on 2D models may face limitations in capturing the full three-dimensional information present in CT scans. Therefore, adopting more advanced models such as 3D UNet, Vision Transformer, ReXNet, among others, may lead to improved Dice scores by better leveraging the three-dimensional context. Moreover, exploring ensemble approaches involving different segmentation models or variations of the proposed model could further enhance accuracy and robustness. Ensemble methods often leverage the diversity of multiple models to achieve more reliable and generalized segmentation results.

References

- [1] E. Gibson *et al.*, "Automatic Multi-Organ Segmentation on Abdominal CT with Dense V-Networks," *IEEE Trans. Med. Imaging*, vol. 37, no. 8, pp. 1822–1834, 2018, doi: 10.1109/TMI.2018.2806309.
- [2] A. M. Ali, A. A. Farag, and A. S. El-Baz, "Graph Cuts Framework for Kidney Segmentation with Prior Shape

- Constraints,” *Med. Image Comput. Comput. Interv. – MICCAI 2007*, pp. 384–392, 2007, doi: 10.1007/978-3-540-75757-3_47.
- [3] U. K. Acharya and S. Kumar, “Genetic algorithm based adaptive histogram equalization (GAAHE) technique for medical image enhancement,” *Optik (Stuttg.)*, vol. 230, Mar. 2021, doi: 10.1016/j.ijleo.2021.166273.
- [4] K. T. Khudhair, F. H. Najjar, S. R. Waheed, H. M. Al-Jawahry, H. H. Alwan, and A. Al-Khaykan, “A novel medical image enhancement technique based on hybrid method,” in *Journal of Physics: Conference Series*, Institute of Physics, 2023, doi: 10.1088/1742-6596/2432/1/012021.
- [5] M. W. Mirza, A. Siddiq, and I. R. Khan, “A comparative study of medical image enhancement algorithms and quality assessment metrics on COVID-19 CT images,” *Signal, Image Video Process.*, vol. 17, no. 4, pp. 915–924, Jun. 2023, doi: 10.1007/s11760-022-02214-2.
- [6] “K. Zuiderveld, ” VIII.5. - Contrast Limited Adaptive Histogram Equalization”, Editor(s): Paul S. Heckbert, Graphics Gems, Academic Press, pp. 474-485, 1994, ISBN 9780123361561, <https://doi.org/10.1016/B978-0-12-336156-1.50061-6>.
- [7] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey Wolf Optimizer,” *Advances in Engineering Software*, Vol. 69, pp. 46-61, 2014, ISSN 0965-9978, <https://doi.org/10.1016/j.advengsoft.2013.12.007>.
- [8] T. Tong *et al.*, “Discriminative dictionary learning for abdominal multi-organ segmentation,” *Med. Image Anal.*, vol. 23, no. 1, pp. 92–104, 2015, doi: 10.1016/j.media.2015.04.015.
- [9] Z. Xu *et al.*, “Efficient multi-atlas abdominal segmentation on clinically acquired CT with SIMPLE context learning,” *Med. Image Anal.*, vol. 24, no. 1, pp. 18–27, 2015, doi: 10.1016/j.media.2015.05.009.
- [10] S. Saxena, N. Sharma, S. Sharma, S. Singh, and A. Verma, “An Automated System for Atlas Based Multiple Organ Segmentation of Abdominal CT Images,” *Br. J. Math. Comput. Sci.*, vol. 12, no. 1, pp. 1–14, 2016, doi: 10.9734/bjmcs/2016/20812.
- [11] Y. Wang, Y. Zhou, W. Shen, S. Park, E. K. Fishman, and A. L. Yuille, “Abdominal multi-organ segmentation with organ-attention networks and statistical fusion,” *Med. Image Anal.*, vol. 55, pp. 88–102, 2019, doi: 10.1016/j.media.2019.04.005.
- [12] S. Bai, X. Bai, Z. Zhou, Z. Zhang, Q. Tian, and L. J. Latecki, “GIFT: Towards Scalable 3D Shape Retrieval,” *IEEE Trans. Multimed.*, vol. 19, no. 6, pp. 1257–1271, 2017, doi: 10.1109/TMM.2017.2652071.
- [13] T. Kada, M. G. Linguraru, M. Hori, R. M. Summers, N. Tomiyama, and Y. Sato, “Abdominal multi-organ segmentation from CT images using conditional shape-location and unsupervised intensity priors,” *Med. Image Anal.*, vol. 26, no. 1, pp. 1–18, 2015, doi: 10.1016/j.media.2015.06.009.
- [14] J. J. Cerrolaza, M. Reyes, R. M. Summers, M. Á. González-Ballester, and M. G. Linguraru, “Automatic multi-resolution shape modeling of multi-organ structures,” *Med. Image Anal.*, vol. 25, no. 1, pp. 11–21, 2015, doi: 10.1016/j.media.2015.04.003.
- [15] R. Gauriau, R. Cuignet, D. Lesage, and I. Bloch, “Multi-organ localization with cascaded global-to-local regression and shape prior,” *Med. Image Anal.*, vol. 23, no. 1, pp. 70–83, 2015, doi: 10.1016/j.media.2015.04.007.
- [16] S. Tomoshige, E. Oost, A. Shimizu, H. Watanabe, and S. Nawano, “A conditional statistical shape model with integrated error estimation of the conditions; Application to liver segmentation in non-contrast CT images,” *Med. Image Anal.*, vol. 18, no. 1, pp. 130–143, 2014, doi: 10.1016/j.media.2013.10.003.
- [17] A. J. Asman and B. A. Landman, “Non-local statistical label fusion for multi-atlas segmentation,” *Med. Image Anal.*, vol. 17, no. 2, pp. 194–208, 2013, doi: 10.1016/j.media.2012.10.002.
- [18] C. Dong *et al.*, “Simultaneous Segmentation of Multiple Organs Using Random Walks,” *J. Inf. Process.*, vol. 24, no. 2, pp. 320–329, 2016, doi: 10.2197/ipsjip.24.320.
- [19] E. Shelhamer, J. Long, and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 4, pp. 640–651, 2017, doi: 10.1109/TPAMI.2016.2572683.
- [20] R. Gauriau, R. Ardori, D. Lesage, and I. Bloch, “Multiple template deformation application to abdominal organ segmentation,” *Proc. - Int. Symp. Biomed. Imaging*, vol. 2015–July, pp. 359–362, 2015, doi: 10.1109/ISBI.2015.7163887.
- [21] O. Ronneberger, P. Fischer, and T. Brox, “U-Net : Convolutional Networks for Biomedical,” pp. 234–241, 2015, doi: 10.1007/978-3-319-24574-4.
- [22] F. Milletari, “V-Net : Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation,” pp. 567–573, 2016, doi: 10.1109/3DV.2016.79.
- [23] X. Zhou, R. Takayama, and S. Wang, “Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting method,” vol. 44, no. 10, pp. 5221–5233, 2017, doi: 10.1002/mp.12480.
- [24] P. Hu, F. Wu, J. Peng, Y. Bao, F. Chen, and D. Kong, “Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 3, pp. 399–411, 2017, doi: 10.1007/s11548-016-1501-5.

- [25] M. Larsson, Y. Zhang, and F. Kahl, "Robust abdominal organ segmentation using regional convolutional neural networks," *Appl. Soft Comput. J.*, vol. 70, pp. 465–471, 2018, doi: 10.1016/j.asoc.2018.05.038.
- [26] S. Chen *et al.*, "Towards Automatic Abdominal Multi-Organ Segmentation in Dual Energy CT using Cascaded 3D Fully Convolutional Network," pp. 1–4, 2017, [Online]. Available: <http://arxiv.org/abs/1710.05379>
- [27] H. R. Roth *et al.*, "An application of cascaded 3D fully convolutional networks for medical image segmentation," *Comput. Med. Imaging Graph.*, vol. 66, no. October 2017, pp. 90–99, 2018, doi: 10.1016/j.compmedimag.2018.03.001.
- [28] H. Kakeya, T. Okada, and Y. Oshiro, "3D U-JAPA-Net: Mixture of Convolutional Networks for Abdominal Multi-organ CT Segmentation," vol. 1, pp. 426–433, 2018.
- [29] Z. Zheng, Y. Hayashi, M. Oda, T. Kitasaka, K. Misawa, and K. Mori, "Federated 3D multi-organ segmentation with partially labeled and unlabeled data," *Int. J. Comput. Assist. Radiol. Surg.*, 2024, doi: 10.1007/s11548-024-03139-6.
- [30] S. Irshad, D. P. S. Gomes, and S. T. Kim, "Improved Abdominal Multi-Organ Segmentation via 3D Boundary-Constrained Deep Neural Networks," *IEEE Access*, vol. 11, pp. 35097–35110, 2023, doi: 10.1109/ACCESS.2023.3264582.
- [31] T. Lei, R. Sun, X. Du, H. Fu, C. Zhang, and A. K. Nandi, "SGU-Net: Shape-Guided Ultralight Network for Abdominal Image Segmentation," *IEEE J. Biomed. Heal. Informatics*, vol. 27, no. 3, pp. 1431–1442, 2023, doi: 10.1109/JBHI.2023.3238183.
- [32] S. Pan *et al.*, "Deep learning-based multi-organ CT segmentation with adversarial data augmentation," p. 93, 2023, doi: 10.1117/12.2653970.
- [33] S. Francis, C. A. Minino, P. N. Pournami, P. Niyas, and P. B. Jayaraj, "Self-supervised approach for organs at risk segmentation of abdominal CT images," *ITM Web Conf.*, vol. 01003, pp. 1–9, 2023, [Online]. Available: https://www.itmconferences.org/articles/itmconf/abs/2023/04/itmconf_I3cs2023_01003/itmconf_I3cs2023_01003.html?utm_source=researcher_app&utm_medium=referral&utm_campaign=RESR_MRKT_Researcher_inbound
- [34] G. Krishnan Murugesan *et al.*, "Automatic Abdominal Multi Organ Segmentation using Residual UNet," no. DI, pp. 1–6, 2023, [Online]. Available: <https://doi.org/10.1101/2023.02.1> C. Li, "Abdominal multi-organ segmentation in CT using nnU-Net," 2022, <https://openreview.net/forum?id=7XxSp0DmoJq>. C. Li, "Abdominal multi-organ segmentation in CT using," pp. 1–7.
- [36] Z. Chen *et al.*, "Deep learning-aided 3D proxy-bridged region-growing framework for multi-organ segmentation," *Sci. Rep.*, vol. 14, no. 1, pp. 1–16, 2024, doi: 10.1038/s41598-024-60668-5.
- [37] X. Yan, H. Tang, S. Sun, H. Ma, D. Kong, and X. Xie, "AFTER-UNet: Axial Fusion Transformer UNet for Medical Image Segmentation," vol. c, pp. 3270–3280, 2022, doi: 10.1109/wacv51458.2022.00333.
- [38] A. Hatamizadeh *et al.*, "UNETR: Transformers for 3D Medical Image Segmentation," pp. 1748–1758, 2022, doi: 10.1109/wacv51458.2022.00181.
- [39] J. Wang, H. Zhao, W. Liang, S. Wang, and Y. Zhang, "Cross-convolutional transformer for automated multi-organs segmentation in a variety of medical images," *Phys. Med. Biol.*, vol. 68, no. 3, 2023, doi: 10.1088/1361-6560/acb19a.
- [40] J. Bullock, C. Cuesta-Lazaro, and A. Quera-Bofarull, "XNet: A convolutional neural network (CNN) implementation for medical X-Ray image segmentation suitable for small datasets," Dec. 2018, doi: 10.1117/12.2512451.
- [41] J. Joseph, J. Sivaraman, R. Periyasamy, and V. R. Simi, "An objective method to identify optimum clip-limit and histogram specification of contrast limited adaptive histogram equalization for MR images," *Biocybern. Biomed. Eng.*, vol. 37, no. 3, pp. 489–497, 2017, doi: 10.1016/j.bbe.2016.11.006.
- [42] C. Singla, R. Kaur, J. Singh, N. Nisha, A. K. Singh, and T. Singh, "Fine Tuned Pre-Trained Deep Neural Network for Automatic Detection of Diabetic Retinopathy Using Fundus Images," *Int. J. Intell. Syst. Appl. Eng.*, vol. 11, no. 9s, pp. 735–742, 2023.
- [43] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017–Janua, pp. 2261–2269, 2017, doi: 10.1109/CVPR.2017.243.
- [44] A. L. Simpson *et al.*, "A large annotated medical image dataset for the development and evaluation of segmentation algorithms," 2019, [Online]. Available: <http://arxiv.org/abs/1902.09063>
- [45] J. Ma *et al.*, "AbdomenCT-1K: Is Abdominal Organ Segmentation A Solved Problem," *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 1–19, 2021, doi: 10.1109/TPAMI.2021.3100536.
- [46] W. Liu, W. Xu, S. Yan, L. Wang, H. Li, and H. Yang, "Combining Self-training and Hybrid Architecture for Semi-supervised Abdominal Organ Segmentation," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13816 LNCS, pp. 281–292, 2022, doi: 10.1007/978-3-031-23911-3_25.

- [47] Y. Huang, H. Zhang, Y. Yan, and H. Hassan, “3D Cross-Pseudo Supervision (3D-CPS): A Semi-supervised nnU-Net Architecture for Abdominal Organ Segmentation,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13816 LNCS, pp. 87–100, 2022, doi: 10.1007/978-3-031-23911-3_9.
- [48] J. W. Choi, “Knowledge Distillation from Cross Teaching Teachers for Efficient Semi-supervised Abdominal Organ Segmentation in CT,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 13816 LNCS, pp. 101–115, 2022, doi: 10.1007/978-3-031-23911-3_10.