

# 基于粒子群优化的注意力自适应U-Net架构搜索

向 健

广东工业大学数学与统计学院, 广东 广州

收稿日期: 2026年4月23日; 录用日期: 2026年5月22日; 发布日期: 2026年5月29日

## 摘 要

U-Net通过跳跃连接对多尺度特征的有效融合, 已成为医学图像分割的主流方法。然而, 其架构超参数(卷积核大小、通道数)和注意力机制选择高度依赖人工设计, 缺乏自动化优化手段。若将注意力选择纳入搜索空间, 会导致维度爆炸和计算开销激增。针对此问题, 文章提出PSO-AttUNet, 通过粒子群优化(PSO)搜索U-Net架构超参数, 并结合基于计分的注意力机制为每层独立选择最合适的注意力类型(SE、CBAM、AG), 实现架构搜索与注意力选择的并行优化。在BUSI乳腺超声数据集上的实验表明, PSO-AttUNet在Dice、IoU、HD95等指标上优于U-Net、Attention U-Net等基线模型, 同时参数量降低至U-Net的1/47, 验证了方法的有效性。

## 关键词

医学图像分割, 神经网络架构搜索, 粒子群优化, U-Net, 注意力机制

# Attention-Adaptive U-Net Architecture Search via Particle Swarm Optimization

Jian Xiang

School of Mathematics and Statistics, Guangdong University of Technology, Guangzhou Guangdong

Received: April 23, 2026; accepted: May 22, 2026; published: May 29, 2026

## Abstract

U-Net has become the mainstream method for medical image segmentation due to its effective fusion of multi-scale features through skip connections. However, its architectural hyperparameters (kernel sizes, channel numbers) and attention mechanism selection heavily rely on manual design,

lacking automated optimization methods. Incorporating attention selection into the search space would lead to dimensional explosion and significant computational overhead. To address this issue, this paper proposes PSO-AttUNet, which searches U-Net architectural hyperparameters through Particle Swarm Optimization (PSO) and independently selects the optimal attention type (SE, CBAM, AG) for each layer using a score-based attention evaluation mechanism, achieving parallel optimization of architecture search and attention selection. Experiments on the BUSI breast ultrasound dataset demonstrate that PSO-AttUNet outperforms baseline models such as U-Net and Attention U-Net in metrics including Dice, IoU, and HD95, while reducing the parameter count to 1/47 of U-Net, validating the effectiveness of the proposed method.

## Keywords

Medical Image Segmentation, Neural Network Architecture Search, Particle Swarm Optimization, U-Net, Attention Mechanism

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

医学图像分割旨在自动识别并勾画出医学影像中特定的解剖结构或者病变区域，是计算机辅助诊断的核心环节。当前，常见的医学图像分割任务存在多种类型，包括 X 光图、超声图、CT 等医学图像。

近年来，研究者将深度学习方法应用于处理医学图像分割并取得了显著进展。卷积神经网络 (Convolutional Neural Network, CNN) 被逐渐应用在这类任务中，其中 U-Net [1] 凭借独特的结构和跳跃连接的应用，具备较强的空间表征能力和上下文建模能力，显著提升了分割性能，成为医学图像分割领域最具影响力的基础架构。

尽管 U-Net 具有较强的分割性能，但是其架构的超参数(如编码器的卷积核大小、输出通道数等)仍高度依赖人工设计，在面对不同任务时需要重新设计和反复调整。此外，有研究者尝试在跳跃连接处添加注意力机制以增强特征选择性表达[2]。神经网络架构搜索(NAS)为上述问题提供了自动化解决思路，其中粒子群优化(Particle Swarm Optimization, PSO) [3]作为一种群智能优化算法，存在实现简单、无需梯度信息、全局搜索能力强等优点，非常适合处理离散搜索空间中的优化问题[4]。但若将注意力类型的选择也纳入到搜索空间中，会导致搜索维度增大，带来计算开销的激增。

针对上述问题，本文提出 PSO-AttUNet，将基于 PSO 的架构搜索与基于计分的注意力评估机制相结合，通过这一设计实现了两者的并行搜索从而降低了计算开销。本文的核心贡献如下：

1) 自动化的超参数优化：将 U-Net 各层中的编码器、瓶颈层和解码器的卷积核大小与输出通道数编码并构成搜索空间，采用 PSO 在离散搜索空间中进行架构搜索，从而实现了超参数的自动优化。

2) 基于计分的注意力评估机制：在 PSO 迭代过程中，逐层评估跳跃连接上各类型注意力机制的性能，通过评估的得分确定每层最优注意力配置，实现与架构超参数的并行搜索，避免了将注意力的选择纳入 PSO 搜索空间带来的维度爆炸问题。

3) 系统性实验验证：在 BUSI 数据集上，与 U-Net、Attention U-Net 等多个基线模型进行了多项指标的全面对比，在 Dice 系数、IoU、HD95 及模型参数量等多个指标上均获得了提升。

## 2. 相关工作

本章主要介绍与本文相关的研究工作。U-Net [1]采用对称的编码器-解码器结构,与跳跃连接结合,已成为医学图像分割的标准基线模型。在此基础上,U-Net++ [5]通过嵌套的密集跳跃连接逐步缩小不同层级间的语义差距;ResUNet [6]将残差连接引入卷积块以缓解梯度消失问题。上述变体在特定场景下取得了一定的性能提升,但其架构设计仍依赖人工经验,各层的卷积核大小、通道数等关键超参数需要逐一调试,缺乏针对具体任务的自动化优化手段。

神经网络架构搜索(NAS)旨在通过自动化设计网络架构以节省计算成本,被广泛应用于 U-Net 的优化。基于进化算法的方法非常适合离散变量的优化:GAUNet [7]通过遗传算法自动优化 U-Net 各层的超参数;CTU-NAS [8]提出多目标进化 NAS 框架,同时优化分割精度与模型复杂度,EU-Net [9]通过引入不同的模块化结构和使用差分进化算法扩大了搜索空间从而提升性能。

粒子群优化(PSO) [3]是一种受鸟群觅食行为启发的群体智能算法,种群中的每个粒子通过个体最优和全局最优的引导在搜索空间中迭代更新位置与速度。与遗传算法相比,PSO 无需交叉和变异操作,实现更为简洁,收敛速度更快,尤其适合处理离散与混合搜索空间中的优化问题。在 CNN 架构搜索方面,Junior 等人[4]将 PSO 应用于深度 CNN 架构搜索,通过编码卷积层数量、滤波器数量等超参数取得了优异性能。Saifullah 等人提出了 PSO-UNet [10],通过 PSO 对 U-Net 的超参数进行优化,在 BraTS 2021 等数据集上取得了优异的分割性能,同时参数量也有明显减少,证明了 PSO 在 U-Net 架构搜索中的有效性。

注意力机制在医学图像分割中展现出显著优势,因此研究者们开始在 U-Net 中引入注意力机制。SE 注意力[11]通过全局平均池化与全连接网络学习通道间依赖关系,实现通道维度的自适应加权;CBAM [12]在通道注意力基础上进一步引入空间注意力,依次对通道和空间维度进行特征重标定。Zhang 等人将 CBAM 嵌入 U-Net 的跳跃连接中[13],优化编码器中的特征传输,减少语义间隙,进而提升分割性能。CA-Net [14]在通道、空间和尺度三个维度同时引入注意力机制,实验表明多种注意力的协同使用能够显著提升分割精度。Attention U-Net [2]将注意力门控(Attention Gate, AG)嵌入跳跃连接并取得了显著的性能提升。此外,随着 Transformer 的兴起,研究者将自注意力机制引入 U-Net 以建模长程依赖:TransUNet [15]将 Vision Transformer 嵌入编码器,弥补了纯卷积网络在全局上下文建模上的不足;Swin-UNet [16]以纯 Transformer 构建编解码器,利用移位窗口注意力实现层次化特征提取。

然而,上述方法通常在网络的所有层级统一使用同一种注意力机制,而不同层级的特征具有不同的语义粒度和空间分辨率,单一类型的注意力配置难以满足各层级的差异化需求。Hassanzadeh 等人[17]采用进化算法自动设计包含注意力机制的 CNN 架构,将注意力机制作为可搜索的网络组件用于医学图像分割。

CA-Net [14]的注意力机制无法实现自动化选择,基于进化的方法虽然实现了自动化,但如果需要把注意力机制的选择也纳入到搜索空间中,计算开销较大。面对这一难题,本文提出的基于计分的注意力评估机制能够实现神经网络架构与注意力机制的并行搜索,在不增加搜索空间的前提下实现注意力机制的自动化选择。通过将 PSO 与基于计分的注意力评估机制相结合,提出了一种高效的 U-Net 架构搜索方法,能够实现网络架构与注意力机制的并行优化,在降低计算成本的前提下显著提升了医学图像分割的性能。

## 3. 方法论

本章将详细介绍所提出的 PSO-AttUNet 的实现方法。图 1 展示了优化算法的整体框架,下文依次介绍搜索空间定义、PSO 算法、注意力评估机制和适应度评估方法。

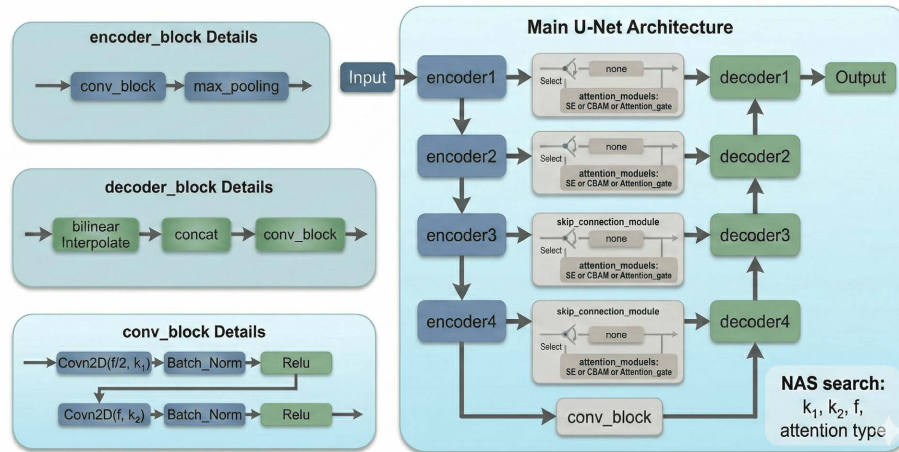


Figure 1. The framework of PSO-AttUNet  
图 1. PSO-AttUNet 整体框架

### 3.1. 搜索空间定义

采用深度固定为 4 层的 U-Net 作为基础框架，网络由 4 个编码器、1 个瓶颈层和 4 个解码器组成，共 9 个卷积块。每个卷积块均为 DoubleConv 结构(Conv-BN-ReLU-Conv-BN-ReLU)，其可搜索变量为两个卷积核大小( $k_1, k_2$ )和输出通道指数  $e$ ，对应实际通道数  $c = 2^e$ 。下采样使用  $2 \times 2$  最大池化，上采样使用双线性插值。

在 PSO-AttUNet 中，基于粒子群优化的方法仅作用于 U-Net 架构中的超参数，不直接更新注意力类型。种群中的每个粒子表示一个候选结构  $x_i = \{\text{encoders, bottleneck, decoders}\}$ ，其中每个卷积模块包含三个可搜索变量( $k_1, k_2, e$ )，分别对应两层卷积核大小与通道指数，实际输出通道数为  $2^e$ 。粒子速度  $v_i$  与结构同构，记录上述三个变量在各层的增量。

**粒子编码。**每个粒子表示一个候选 U-Net 结构，其位置向量  $x$  为结构化参数集合：

$$x = \{x_1^{\text{enc}}, \dots, x_4^{\text{enc}}, x^{\text{bn}}, x_1^{\text{dec}}, \dots, x_4^{\text{dec}}\} \quad (1)$$

其中每个卷积模块参数为三元组  $x_i^{(\cdot)} = \{k_1, k_2, e\}$ ，超参数取值范围如表 1 所示。

Table 1. Hyperparameter value ranges  
表 1. 超参数取值范围

超参数	符号	候选值
卷积核大小	$k_1, k_2$	{3, 4, 5, 6, 7}
输出通道数(指数形式)	$e$	{3, 4, 5, 6, 7}

单个架构对应搜索空间规模为  $125^9 = 5^{27} \approx 7.45 \times 10^{18}$ 。若在此基础上再将 4 个跳跃连接层的注意力类型的选择直接并入 PSO 搜索，搜索空间的扩大会导致计算成本的激增。

### 3.2. PSO 优化

**初始化策略。**种群规模设为  $N = 15$ ，最大迭代次数  $T = 40$ ，每个粒子的结构参数在离散候选集  $S$  上均匀随机采样初始化，速度向量初始化为零；个体最优  $p_{\text{best},i}$  初始化为粒子当前结构，个体最优适应度初始化为 0，全局最优适应度初始化为  $-\infty$ 。

**标准 PSO 更新。**在第  $t$  次迭代中，粒子  $i$  在维度  $d$  上的速度与位置更新为：

$$v_{i,d}^{t+1} = w^t \cdot v_{i,d}^t + c_1 r_1 (p_{\text{best},i,d} - x_{i,d}^t) + c_2 r_2 (g_{\text{best},d} - x_{i,d}^t) \quad (2)$$

$$x_{i,d}^{t+1} = x_{i,d}^t + v_{i,d}^{t+1} \quad (3)$$

其中,  $w$  为惯性权重,  $c_1, c_2$  为认知系数,  $r_1, r_2 \sim U(0,1)$ 。采用分段惯性权重: 前半段  $w=0.9$  增强全局探索, 后半段  $w=0.4$  加速局部收敛。设置  $c_1=0.5$ ,  $c_2=0.3$  以避免过早收敛。

**离散化映射。**因为卷积核大小以及通道数被设置为指数化表示, 所以搜索空间为离散集合, 连续更新后的变量需通过最近邻映射回合法值:

$$\hat{x}_{i,d} = \arg \min_{s \in \mathcal{S}_d} |x_{i,d}^{t+1} - s| \quad (4)$$

其中,  $\mathcal{S}_d$  为维度  $d$  的候选集合(卷积核大小与通道指数均为  $\{3, 4, 5, 6, 7\}$ )。

**早停机机制。**若全局最优连续 10 代未改进且刚好完成 4 层注意力评估, 则提前终止搜索。

### 3.3. 基于计分的注意力评估机制

当前的架构选择添加的注意力类型包括 SE (Squeeze-and-Excitation) [11]、CBAM (Convolutional Block Attention Module) [12]和 AG (Attention Gate) [2]。因为这三种注意力间以及与普通跳跃连接拼接间的差异难以直接用速度差异值进行表示, 进而导致难以通过 PSO 的速度更新机制实现优化。

为此, 本文提出一种基于计分的注意力评估机制(Count-Based Evaluation), 嵌入到 PSO 的优化流程中。其核心思想是: 在每次 PSO 迭代中, 以当前全局最优架构为基础, 轮询评估一个跳跃连接层的注意力类型, 通过累积计分来确定每层的最优注意力配置。在引入计分机制之前, 需要解决一个前置问题: 由于 PSO 搜索空间中各层通道数  $c=2^e$  ( $e \in \{3, 4, 5, 6, 7\}$ ) 是可变的, 标准 SE 注意力机制[11]中的某些固定参数在小通道数条件下会失效。因此, 本文对两种注意力机制进行了适配性修改, 使其能够灵活应用于不同参数配置的 U-Net 架构中。

**SE 的自适应降维。**因为架构搜索的过程中输出通道数是可变的, 所以使用固定的降维系数  $r=16$ , 可能会导致输出通道数变为 0。针对这一问题, 有研究者尝试用边界保护函数[18]确保压缩后的通道数不会低于一个最低值; 在 SASE [19]中研究者尝试将降维系数  $r$  也纳入到搜索空间中。

因为使用确定值可能会影响通道注意力的权重赋予, 且若将降维系数纳入到搜索空间中与本文的设计思路相悖, 所以本文设计了自适应降维函数, 根据实际通道数  $c$  动态计算降维系数:

$$r(c) = \begin{cases} 1 & c < 16 \\ 2 & 16 \leq c < 32 \\ 4 & 32 \leq c < 64 \\ 8 & 64 \leq c < 128 \\ 16 & c \geq 128 \end{cases} \quad (5)$$

这一设计确保了即使输入通道数可变, 也可以通过动态的降维系数在任意通道数下均能生成有意义的通道权重。

**AG 的动态参数配置。**在可变架构中, Attention Gate (AG) [2]的编码器通道数  $c_{\text{skip}}$  与解码器通道数  $c_{\text{dec}}$  由 PSO 搜索决定。但是两者的关系可能会影响到 AG 的特征提取能力, 对此 NAS-UNet [20]中尝试将中间通道数也纳入到搜索空间。

但这同样与本文设计思路相悖, 因此本文设计了 AG 的中间通道数需根据两者的关系动态确定:

$$c_{\text{inter}} = \begin{cases} c_{\text{dec}} & c_{\text{skip}} > c_{\text{dec}} \\ c_{\text{dec}}/2 & c_{\text{skip}} \leq c_{\text{dec}} \end{cases} \quad (6)$$

当编码器通道数比解码器通道数大时，保持中间通道数与解码器通道数一致有助于充分利用解码器的深层特征；当编码器通道数较小时，适当降低中间通道数可以减少冗余参数。

注意力评估机制通过轮询的形式完成对架构中每一层的注意力机制的评估，最后根据得分选择最优的注意力机制，具体的步骤如下：

1) 种群中的每个粒子都评估结束，更新个体最优以及全局最优后，以当前全局最优架构  $g_{best}$  为基础，构造 3 个变体粒子：仅对一层的跳跃连接添加注意力机制，包括 SE、CBAM 和 AG，其余层保持无注意力(none)。

2) 对 3 个变体粒子分别进行完整的训练 - 验证评估，计算每一个适应度分数。

3) 在三种注意力类型中选出适应度最高者  $a^* = \arg \max_{a \in \{SE, CBAM, AG\}} f_a$ ，在计分板中进行记录。

4) 在确保最大迭代次数为架构深度的整数倍的前提下，每一层的评估次数应保持一致，最后通过记录每一层的每一种注意力机制的得分，根据公式(8)可以选择每一层对应的最优注意力机制。

$$C_{l,a^*} \leftarrow C_{l,a^*} + \delta(t) \quad (7)$$

$$a_l^* = \arg \max_{a \in \mathcal{A}} C_{l,a}, l = 0, 1, \dots, L-1 \quad (8)$$

其中， $\delta(t)$  为计分权重，在早期迭代(前 20%)中设为 0.5，后期设为 1.0。之所以这样设置权重折扣是因为：搜索初期粒子尚未充分收敛，此时的注意力评估结果可靠性较低，应给予较小的权重避免早期不成熟的架构对最后注意力机制的选择造成较大的影响。

算法 1 展示了在 PSO 的一次迭代过程中，种群更新结束后嵌入的基于计分的注意力评估机制的算法流程：

---

#### 算法 1 单次迭代的注意力评估

---

**输入：** 全局最优架构  $g_{best}$ ，当前迭代  $t$ ，最大迭代数  $T$ ，网络深度  $L = 4$ ，注意力候选集  $A$ ，计分板  $C$

**输出：** 更新后的计分板  $C$

```

1:  $l \leftarrow (t - 1) \bmod L$  // 当前评估的注意力层
2: // 注意力层评估
3: for  $a \in \{SE, CBAM, AG\}$  do
4:   构造变体：将  $g_{best}$  第  $l$  层注意力替换为  $a$ 
5:    $f_a \leftarrow \text{Evaluate}(\text{变体})$ 
6: end for
7:  $a^* \leftarrow \arg \max_a f_a$  // 选择最优注意力
8: // 更新计分板
9:  $\delta(t) \leftarrow 0.5$  if  $t \leq 0.2T$  else 1.0
10:  $C_{l,a^*} \leftarrow C_{l,a^*} + \delta(t)$ 
11: return  $C$ 

```

---

这一设计使得注意力的评估始终基于当前最优架构进行，并不会参与到 PSO 的优化中，与架构的搜索是并行的。由于注意力机制不参与 PSO 的速度更新，搜索空间不会因此产生变化，计算成本仅增加每次迭代过程中的 3 次额外的适应度评估。

### 3.4. 适应度评估

适应度函数结合了 Dice 系数和 IoU:

$$\text{Dice} = \frac{2|Y \cap \hat{Y}|}{|Y| + |\hat{Y}|} \quad (9)$$

$$\text{IoU} = \frac{|Y \cap \hat{Y}|}{|Y \cup \hat{Y}|} \quad (10)$$

$$\text{fitness} = 0.5 \cdot \text{DSC} + 0.5 \cdot \text{IoU} \quad (11)$$

其中,  $Y$  为真值标注,  $\hat{Y}$  为预测结果。每个粒子训练 20 epoch 后计算适应度, 并更新个体最优和全局最优。

## 4. 实验结果

本章节主要介绍针对上一章节所述方法开展的实验设计与结果分析。包括 PSO 架构搜索的实验配置与搜索过程; 与多个基线模型的对比实验和验证基于计分的注意力评估机制对性能提升的消融实验。

### 4.1. PSO 架构搜索实验

**搜索配置。**粒子群算法的参数: 种群规模  $N = 15$ , 最大迭代  $T = 40$ , 惯性权重  $w = 0.9/0.4$  (前/后半段), 认知系数  $c_1 = 0.5$ ,  $c_2 = 0.3$ 。每粒子训练 20 epoch, 优化器 Adam (lr = 0.001), 批次大小 32。

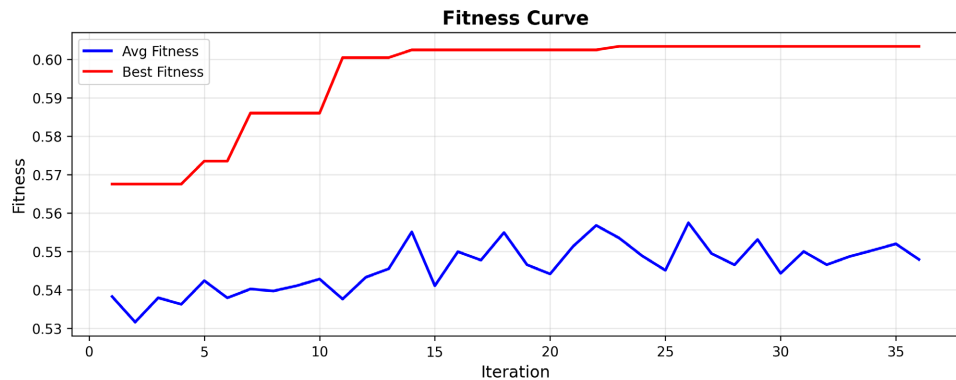


Figure 2. PSO search convergence curve

图 2. PSO 搜索收敛曲线

从图 2 中可以观察到: 种群在中期逐渐收敛, 适应度的提升速度放缓, 基本稳定在一定水平。全局最优在第 36 代持续了超过 10 代没有进行更新, 同时也完成了对四层的注意力机制评估而触发早停机制。

### 4.2. 对比实验

**实验配置。**搜索完成后, 使用最优架构与其他基线模型进行对比实验。实验参数: 训练轮数为 300 个 epoch、优化器为 SGD, 初始学习率 0.01, 采用 Poly 学习率衰减策略, 实验基于 PyTorch 框架实现, 上文的搜索部分的实验、基线模型以及搜索得到的最优架构的训练均在 NVIDIA 的 5090 GPU 上进行训练, 使用随机种子 42 进行随机性控制。搜索过程以及训练均采用 BCEDice 的组合损失函数:

$$\mathcal{L} = 0.5 \cdot \mathcal{L}_{\text{BCE}} + \mathcal{L}_{\text{Dice}} \quad (12)$$

其中,  $\mathcal{L}_{\text{BCE}}$  为交叉熵损失;  $\mathcal{L}_{\text{Dice}}$  为 Dice 损失。

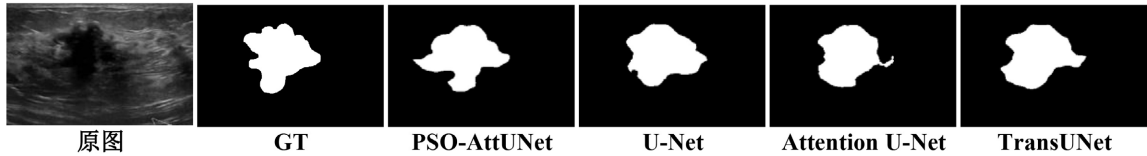
本文采用以下 7 个指标：Dice 系数、IoU、灵敏度 SE、精确率 PC、F1 分数、准确率 AC 和 HD95 (95% Hausdorff 距离)进行评估，并利用 7 个有代表性基线模型、一个基于差分进化算法的 U-Net 架构搜索方法 EU-Net [9]与本文提出的方法进行对比。

**Table 2.** Comparison results with baseline models

**表 2.** 基线模型对比实验结果

模型	Dice $\uparrow$	IoU $\uparrow$	HD95 $\downarrow$	SE $\uparrow$	PC $\uparrow$	F1 $\uparrow$	ACC $\uparrow$	Params (M) $\downarrow$	FLOPs (G) $\downarrow$
U-Net [1]	0.7574	0.6133	37.10	0.7049	<u>0.8286</u>	0.7574	0.9600	34.527 M	65.522 G
Attention U-Net [2]	<u>0.7662</u>	<u>0.6230</u>	37.51	0.7083	<b>0.8439</b>	<u>0.7662</u>	<b>0.9616</b>	34.879 M	66.632 G
ResUNet [6]	0.5867	0.4166	85.53	0.5309	0.6498	0.5867	0.9339	<u>13.043 M</u>	80.984 G
UNet++ [5]	0.7482	0.5995	33.15	0.7259	0.7820	0.7482	0.9569	26.904 M	37.626 G
ResUNet++ [21]	0.7212	0.5655	41.01	0.6902	0.7606	0.7212	0.9530	14.483 M	70.994 G
TransUNet [15]	0.7534	0.6067	<u>32.49</u>	<u>0.7391</u>	0.7768	0.7534	0.9572	93.231 M	32.229 G
EU-Net [9]	0.7296	0.5807	41.46	0.6824	0.8074	0.7296	0.9562	1.341 M	<b>9.175 G</b>
<b>PSO-AttUNet (Ours)</b>	<b>0.7752</b>	<b>0.6375</b>	<b>32.12</b>	<b>0.7550</b>	0.8049	<b>0.7752</b>	<u>0.9604</u>	<b>729.145 K</b>	<u>15.467 G</u>

从表 2 中可以观察到：PSO-AttUNet 在 Dice 和 IoU 上取得最优，HD95 最低(边界精度最高)，同时参数量仅为 U-Net 的 1/47。这表明 PSO 搜索的架构配置在提升分割精度的同时显著降低了模型复杂度。



**Figure 3.** Segmentation results visualization

**图 3.** 分割结果可视化

如图 3 所示，PSO-AttUNet 对病灶边界的捕捉更精确，相比 U-Net 和 Attention U-Net 减少了过分割，相比 TransUNet 减少了欠分割。

### 4.3. 注意力评估消融实验

在 PSO 架构搜索实验中，最优架构中的跳跃连接未配置注意力机制。本节将通过消融实验验证在 PSO 搜索过程中引入的基于计分的注意力评估机制对分割性能提升的贡献。实验中参与评估的三种候选注意力类型分别是：SE、CBAM、AG。

**Table 3.** Attention evaluation score results

**表 3.** 注意力评估得分结果

跳跃连接层	SE	CBAM	AG
第 1 层	1.0	5.0	2.0
第 2 层	1.5	2.0	4.5
第 3 层	2.5	4.5	1.0
第 4 层	2.5	3.0	2.5

从表 3 中可以得到以下观察：SE 的整体表现比较差，CBAM 在大部分深度的表现都较好，同时 AG 在第二层的得分最高。

**Table 4.** Ablation study of different attention configurations  
**表 4.** 不同注意力配置的消融实验

注意力配置	Dice $\uparrow$	IoU $\uparrow$	HD95 $\downarrow$
N,N,N,N	0.7752	0.6375	32.12
C,N,N,N	0.7626	0.6189	31.16
N,G,N,N	0.7680	0.6257	31.66
C,G,N,N	0.7842	0.6462	31.06
C,G,C,N	0.7533	0.6071	32.71
C,G,C,C	0.7554	0.6164	32.06
对比配置(验证机制的有效性)			
S,S,S,S	0.7619	0.6178	33.21
C,C,C,C	0.7596	0.6162	30.09
G,G,G,G	0.7792	0.6415	29.51

从表 4 的第一组实验中可以观察到：1) 某一层引入独立的注意力机制不足以带来性能增益，同时注意力机制也并非越多越好，错误的选择会导致性能下降。2) C,G,N,N 取得了最好的表现，这一结果表明两种注意力类型之间可能存在协同效应：CBAM 增强浅层特征图的空间定位能力，AG 在中间层筛选跳跃连接中与分割目标相关的信息，二者功能互补从而提升了整体分割性能。

#### 4.4. 局限性分析

实验结果证明了 PSO-AttUNet 的有效性，这一方法搜索得到的架构获得了很好的表现，并且实现了注意力机制评估与架构搜索的并行搜索。但是这一方法仍然存在部分局限性：PSO 作为一种群智能算法，其效果会受到种群的影响，初始化策略的选择可能会让算法陷入局部最优而无法获得良好的效果；基于计分的注意力评估机制无法兼容太多的注意力机制，候选的注意力机制增多可能会导致计算开销激增，且注意力机制的选择也非常关键，例如实验中的 SE 表现不佳，对实验结果可能会产生影响。

## 5. 结论

针对医学图像分割任务中神经网络架构手工设计难度大、注意力机制选择引入自动化搜索后计算开销激增等问题，本文提出了一种基于粒子群优化的 U-Net 架构搜索方法，并设计了一种基于计分的注意力机制评估方法(PSO-AttUNet)，并行搜索最优的神经网络架构及对应的注意力机制。该方法的核心贡献在于：

- 1) 设计了面向 U-Net 的结构化粒子编码方案，将各层卷积核大小和通道数指数编码为粒子位置，并通过标准 PSO 的速度 - 位置更新在离散搜索空间中进行全局优化。动态惯性权重策略和最近邻离散化映射确保了搜索过程的稳定性和有效性。
- 2) 提出了基于计分的注意力评估机制，将注意力类型选择从 PSO 的连续更新中解耦，通过轮询评估和累积计分为每个跳跃连接层独立确定最优注意力机制。并通过降低早期权重的策略提高了评估结果的可靠性。

通过在 BUSI 数据集上的实验表明, PSO-AttUNet 搜索得到的架构在多项分割指标上优于 U-Net、Attention U-Net、U-Net++等基线模型。消融实验证明了基于计分的注意力评估方法将注意力机制的选择脱离出搜索空间实现并行搜索的有效性。未来工作将从以下方向展开: 1) 实验过程中仍有部分的注意力机制表现不佳, 可能需要添加其他类型的注意力机制参与搜索, 比如自注意力机制等; 2) 尝试结合其他优化策略来解决搜索过程中期几乎收敛的问题; 3) 在更多医学图像分割数据集上验证方法的泛化能力。

## 参考文献

- [1] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W. and Frangi, A., Eds., *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, Springer, 234-241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [2] Oktay, O., Schlemper, J., Le Folgoc, L., Lee, M., Heinrich, M., Misawa, K., et al. (2018) Attention U-Net: Learning Where to Look for the Pancreas. arXiv: 1804.03999.
- [3] Kennedy, J. and Eberhart, R. (1995) Particle Swarm Optimization. *Proceedings of ICNN'95—International Conference on Neural Networks*, Perth, 27 November-1 December 1995, 1942-1948. <https://doi.org/10.1109/icnn.1995.488968>
- [4] Fernandes Junior, F.E. and Yen, G.G. (2019) Particle Swarm Optimization of Deep Neural Networks Architectures for Image Classification. *Swarm and Evolutionary Computation*, **49**, 62-74. <https://doi.org/10.1016/j.swevo.2019.05.010>
- [5] Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N. and Liang, J. (2018) Unet++: A Nested U-Net Architecture for Medical Image Segmentation. In: Stoyanov, D., et al., Eds., *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 3-11. [https://doi.org/10.1007/978-3-030-00889-5\\_1](https://doi.org/10.1007/978-3-030-00889-5_1)
- [6] Diakogiannis, F.I., Waldner, F., Caccetta, P. and Wu, C. (2020) ResUNet-a: A Deep Learning Framework for Semantic Segmentation of Remotely Sensed Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, **162**, 94-114. <https://doi.org/10.1016/j.isprsjprs.2020.01.013>
- [7] Khouy, M., Jabrane, Y., Ameer, M. and Hajjam El Hassani, A. (2023) Medical Image Segmentation Using Automatic Optimized U-Net Architecture Based on Genetic Algorithm. *Journal of Personalized Medicine*, **13**, Article 1298. <https://doi.org/10.3390/jpm13091298>
- [8] Ying, W., Zheng, Q., Wu, Y., Yang, K., Zhou, Z., Chen, J., et al. (2023) Efficient Multi-Objective Evolutionary Neural Architecture Search for U-Nets with Diamond Atrous Convolution and Transformer for Medical Image Segmentation. *Applied Soft Computing*, **148**, Article ID: 110869. <https://doi.org/10.1016/j.asoc.2023.110869>
- [9] Yu, C., Wang, Y., Tang, C., Feng, W. and Lv, J. (2023) EU-Net: Automatic U-Net Neural Architecture Search with Differential Evolutionary Algorithm for Medical Image Segmentation. *Computers in Biology and Medicine*, **167**, Article ID: 107579. <https://doi.org/10.1016/j.combiomed.2023.107579>
- [10] Saifullah, S. and Drezewski, R. (2025) Particle Swarm-Optimized U-Net Framework for Precise Multimodal Brain Tumor Segmentation. *Proceedings of the Genetic and Evolutionary Computation Conference Companion*, Malaga, 14-18 July 2025, 323-326. <https://doi.org/10.1145/3712255.3726561>
- [11] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/cvpr.2018.00745>
- [12] Woo, S., Park, J., Lee, J. and Kweon, I.S. (2018) CBAM: Convolutional Block Attention Module. In: Ferrari, V., Hebert, M., Sminchisescu, C. and Weiss, Y., Eds., *Computer Vision—ECCV 2018*, Springer, 3-19. [https://doi.org/10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1)
- [13] Zhang, B., Qiu, S. and Liang, T. (2024) Dual Attention-Based 3D U-Net Liver Segmentation Algorithm on CT Images. *Bioengineering*, **11**, Article 737. <https://doi.org/10.3390/bioengineering11070737>
- [14] Gu, R., Wang, G., Song, T., Huang, R., Aertsen, M., Deprest, J., et al. (2021) CA-Net: Comprehensive Attention Convolutional Neural Networks for Explainable Medical Image Segmentation. *IEEE Transactions on Medical Imaging*, **40**, 699-711. <https://doi.org/10.1109/tmi.2020.3035253>
- [15] Chen, J.N., Lu, Y.Y., Yu, Q.H., Luo, X.D., Adeli, E., Wang, Y., Lu, L., Yuille, A.L. and Zhou, Y.Y. (2021) TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation. arXiv: 2102.04306.
- [16] Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., et al. (2023) Swin-Unet: Unet-Like Pure Transformer for Medical Image Segmentation. In: Karlinsky, L., Michaeli, T. and Nishino, K., Eds., *Computer Vision—ECCV 2022 Workshops*, Springer, 205-218. [https://doi.org/10.1007/978-3-031-25066-8\\_9](https://doi.org/10.1007/978-3-031-25066-8_9)
- [17] Hassanzadeh, T., Essam, D. and Sarker, R. (2021) Evolutionary Deep Attention Convolutional Neural Networks for 2D and 3D Medical Image Segmentation. *Journal of Digital Imaging*, **34**, 1387-1404. <https://doi.org/10.1007/s10278-021-00526-2>

- 
- [18] Alansari, M., Hay, O.A., Javed, S., Shoufan, A., Zweiri, Y. and Werghi, N. (2023) GhostFaceNets: Lightweight Face Recognition Model from Cheap Operations. *IEEE Access*, **11**, 35429-35446. <https://doi.org/10.1109/access.2023.3266068>
  - [19] Wang, H., Li, Y., Wu, Z., Wang, H. and Zhang, Y. (2024) SASE: A Searching Architecture for Squeeze and Excitation Operations. arXiv: 2411.08333.
  - [20] Weng, Y., Zhou, T., Li, Y. and Qiu, X. (2019) Nas-Unet: Neural Architecture Search for Medical Image Segmentation. *IEEE Access*, **7**, 44247-44257. <https://doi.org/10.1109/access.2019.2908991>
  - [21] Jha, D., Smedsrud, P.H., Riegler, M.A., Johansen, D., Lange, T.D., Halvorsen, P., *et al.* (2019) ResUNet++: An Advanced Architecture for Medical Image Segmentation. 2019 *IEEE International Symposium on Multimedia (ISM)*, San Diego, 9-11 December 2019, 225-230. <https://doi.org/10.1109/ism46123.2019.00049>