

# DRWKV: 聚焦于对象边缘的低光照图像增强

Xuecheng Bai<sup>1\*</sup>, Yuxiang Wang<sup>2\*</sup>, Boyu Hu<sup>3</sup>, Qinyuan Jie<sup>1</sup>,  
Chuanzhi Xu<sup>2†</sup>, Hongru Xiao<sup>4</sup>, Kechen Li<sup>5</sup>, Vera Chung<sup>2</sup>

<sup>1</sup>Shenyang Ligong University, Shenyang, China

<sup>2</sup>The University of Sydney, NSW, Australia

<sup>3</sup>University of International Business and Economics, Beijing, China

<sup>4</sup>Tongji University, Shanghai, China

<sup>5</sup>Nanjing University of Aeronautics and Astronautics, Nanjing, China

**Abstract**—低光图像增强仍然是一项具有挑战性的任务，特别是在极端光照退化下保持物体边缘连续性和细小结构细节方面。在本文中，我们提出了一种新颖的模型，DRWKV（详细接纳加权键值），它整合了我们提出的全局边缘视网膜（GER）理论，从而实现了光照和边缘结构的有效解耦，以增强边缘保真度。其次，我们引入了演化的WKV注意力，这是一种螺旋扫描机制，可以更有效地捕捉空间边缘连续性和建模不规则结构。第三，我们设计了双向谱对齐器（Bi-SAB）和定制的MS<sup>2</sup>-Loss，以共同对齐亮度和色度特征，改善视觉自然度并减轻伪影。在五个LLIE基准上的广泛实验表明，DRWKV在PSNR、SSIM和NIQE方面实现了领先的性能，同时保持低计算复杂性。此外，DRWKV在低光多目标跟踪任务中提升了下游性能，验证了其泛化能力。

**Index Terms**—Low-light Image Enhancement, Receptance Weighted Key Value, Global Edge Retinex theory

低光照图像增强（LLIE）技术被认为是一种以低成本亮化低光照图像的主动方法。最近，学术界主要探索了诸如全局注意力建模、局部特征增强和选择性状态更新等突出方法。基于稳健的Retinex理论，低光照图像的组成可以解释为：

其中表示场景反射率，代表光照组件，而表示内在噪声。

在构建低光场景时，对于图像，除了噪声 $N$ 的增加外，照明成分 $L$ 被分解为全局和局部照明。在光线反射下，全局光谱和轮廓光场将动态变化。在低光图像质量提升的背景下，仅关注全局照明会放大背景噪声[?]，并导致边缘失真[?]，而仅关注轮廓则会导致突兀的过渡和整体曝光不足的结果。因此，在全局曝光可控的情况下，低光图像增强的本质在于物体边缘的精细化。如图??(c)所示，ZeroDCE[?]存在边缘失真，而RetinexMamba[?]表现出细节模糊，表明现有方法在处理极低光环境中存在局限性。

为了解决这些问题，我们提出了Detailed Receptance Weighted Key Value (DRWKV)模型，该模型优化了VRWKV (ICLR 2025中)的双向WKV注意力，这是视觉RWKV领域的经典模型。我们引入了演变WKV注意力，以增强在物体边缘提取精细结构的能力。与双向WKV注意力相比，演变WKV注意力的空间拓扑感知能力能够最大化对物体边缘连续空间的捕捉，从而加强对不规则空间结构的建模。在此基础上，我们设计了一种辅助技术——双边光谱对齐器（Bi-SAB）。通过对亮度和颜色

光谱特征的对齐，SpAligner解决了由独立处理引起的过曝和颜色失真的问题。结合MS<sup>2</sup>-Loss，SpAligner使得增强图像的光照和颜色更自然。我们在五个基准数据集上进行了广泛的图像增强和低光物体跟踪实验。如图??(a)和??(b)所示，我们的方法在PSNR、SSIM和NIQE分数上取得了领先，显示出在边缘连续性和细节保真度方面的显著优越性。

我们的贡献可以总结如下：

- 我们提出了详细接收加权键值（DRWKV）模型，首次将我们提出的全局边缘Retinex（GER）理论与VRWKV模型结合。它有效地增强了低光图像的边缘。
- 我们设计了一种新颖的应用于WKV注意力机制的动态扫描（ES）机制，也称为动态WKV注意力。它从内到外挖掘边缘细节，增强DRWKV在建模不规则空间结构边缘方面的能力。
- 我们引入双向光谱对齐器（Bi-SAB），通过协调光谱特征并结合我们提出的MS<sup>2</sup>-Loss，丰富增强低光图像的色域。

## I. 相关工作

### A. 对象边缘感知与提取

边缘感知和提取技术是基于传统方法[?], [?], [?], [?]和深度学习方法[?], [?]的。它们通过融合解耦的边缘特征与上下文信息，在低光场景中实现对象边缘的细粒度恢复。

与依赖于局部卷积来提取边缘特征的传统基于CNN的方法（如Retinex-Net[?]和LLNet[?]）不同，新兴的边缘感知方法引入了拓扑建模[?], [?], [?]。这些方法将边缘视为“空间序列”，并通过状态累积来捕获跨区域特征。这一设计通过调整边缘梯度一致性约束和动态阈值，实现平衡的过渡和局部优化，从而在增强弱边缘信号方面提供了核心优势。

### B. 低光物体边缘恢复与增强

先前在复杂的低光照环境中提高物体边缘准确性的努力主要集中在两个方向：探索多样的模型结构和优化边缘增强算法。

诸如EDTER[?]和Retinex-RAWMamba[?]之类的技术，通过利用自注意力来捕获跨区域依赖关系，并通过线性扫描来优化时序积累，从而探索多样的架构。

\*Equal contribution.

†Corresponding author: Chuanzhi Xu  
(chuanzhi.xu@sydney.edu.au)

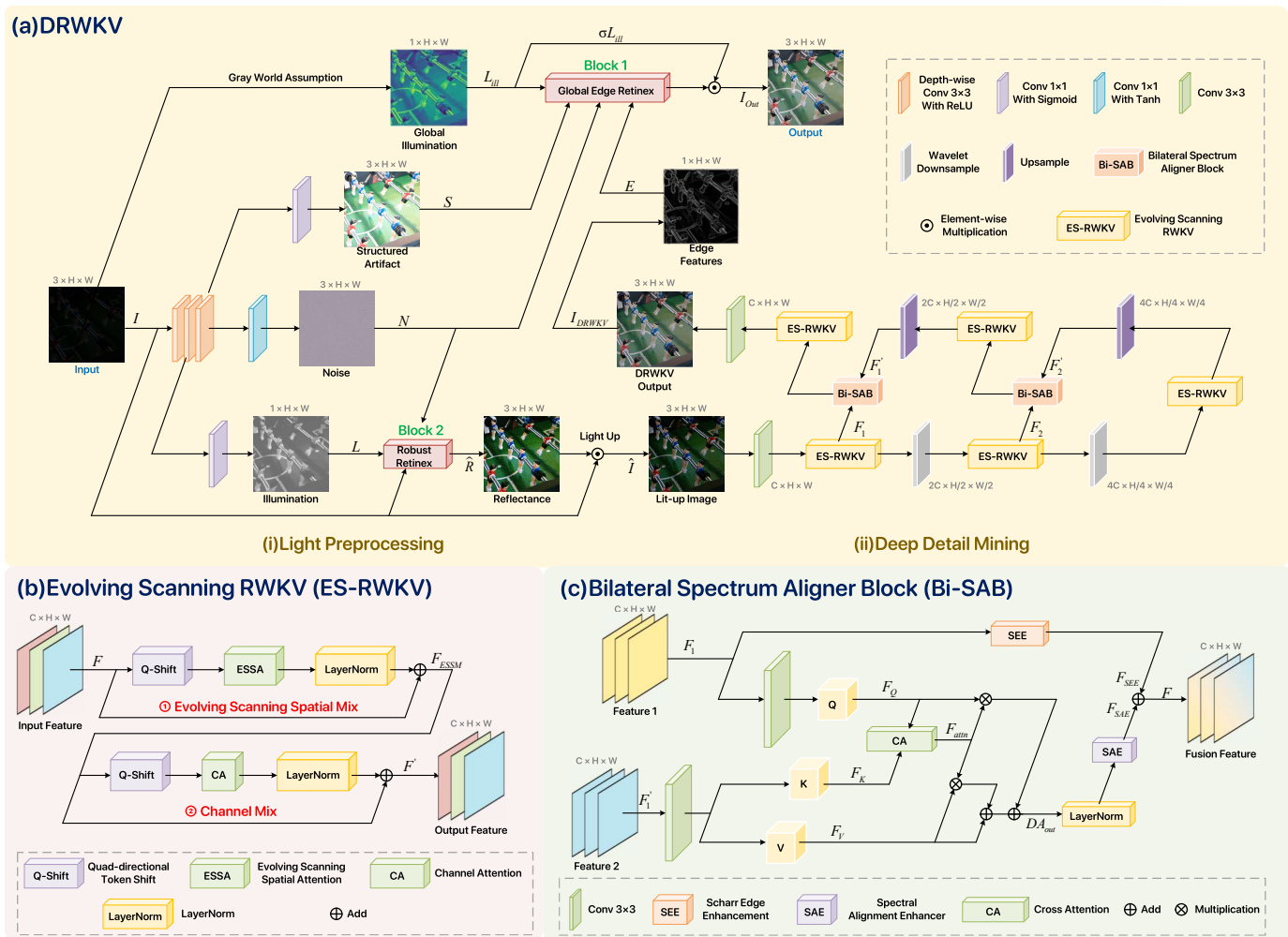


Fig. 1. DRWKV 模型概述。(a) DRWKV 通过两个步骤处理低光图像：(i) 光预处理和 (ii) 深度细节挖掘。Block1 和 Block2 的位置对应于两种 Retinex 理论的布局设置。(b) 演进扫描 RWKV (ES-RWKV) 的模块结构，嵌入了两个核心混合模块：演进扫描空间混合和通道混合。(c) 双谱对齐块 (Bi-SAB) 的模块结构。

通过频域特征过滤 [?]、多模态边缘互补 [?], [?] 和基于对抗学习的边缘约束 [?] 实现稳健的边缘提取，进一步精确增强边缘特征。

尽管取得了这些进展，但在视觉 RWKV 框架下的物体边缘一致性建模仍然受到有限的关注，尤其是在复杂的低光场景中。这表明在基于 RWKV 的架构中集成边缘感知保真度建模是一个有前景的方向。

## II. 方法论

重新思考 VRWKV 中的边缘保真度建模：边缘保真度的核心在于结构和特征之间的强一致性。分析了 VRWKV 在低光场景中的局限性，从两个方面进行探讨：

(1) **Neural Architecture.** VRWKV 采用了单一的递归结构，这与边缘的层次化特征不匹配。这导致浅层和深层特征混淆（表现为边缘模糊），无法将光照变化与边缘表示分离，并导致在低光下明亮/黑暗区域中噪声与边缘之间的频繁误判。

(2) **Computational Geometry.** 边缘是高维像素空间中的低维流形，由于光照剧烈变化，在低光条件下变得更加弯

曲。然而，VRWKV 假设特征在欧几里德空间中是均匀分布的，而边缘流形实际上位于弯曲的黎曼空间中。这导致欧几里德距离度量不准确（导致边缘特征的错误分类），以及在低光条件下卷积的平移不变性与边缘尺度变化之间的不匹配。

根据我们的推理结果，我们提出了两个假设：(1) 分层细节增强依赖于具有照明不变性的梯度引导机制。(2) 低光边缘连续性的建模需要一种与边缘空间拓扑结构兼容的时间编码方案。

**DRWKV 模型：**不同于现有的主流方法，我们采用分层优化策略，优先结合灰色世界假设与由噪声  $N$ 、伪影  $S$  和光照  $L$  组成的三支通道，来实现 (II-A) 光照预处理，从而得到亮度恢复的图像  $\hat{I}$ 。这遵循了与 Retinexformer [?] 类似的策略。RWKV 结构由四组 ES-RWKV 和采样块组成的集群构成。对于编码器，(II-B) 演化扫描首次被应用，结合小波变换的下采样以完成边缘梯度特征提取。解码器通过 (II-C) 双边谱校准块 (Bi-SAB) 集成采样信息  $F$  和  $F'$ ，使微观纹理到宏观轮廓的逐步增强成为可能。

新的设计考虑到了非线性边缘拓扑和信噪比的限制，平衡了弱光细节挖掘的深度和计算成本。我们的 DRWKV 模型的总体结构见图 1。

#### A. 轻预处理

轻量预处理结构采用了双路径并行引导方法。从技术上讲，我们使用灰世界假设  $L_{ill}$  重新计算 RGB 分量：

$$L_{ill} = [R' \ G' \ B'] = \text{diag} \left( \frac{\frac{1}{3} \sum_{i=1}^N |c_i|}{\sum_{i=1}^N c_i} \right) \cdot c \quad (1)$$

。其中， $c_i = [R_i, G_i, B_i]^T$  重构了像素的三通道向量。详细过程包括通过  $\text{diag}(\cdot)$  将向量转换为对角矩阵，然后计算与原始通道向量  $c$  的内积。结合  $L_{ill}$  与伪影分量  $S$ ，我们明确定义了需要避免的干扰边界，使全球边缘视网膜增强 (GER) 在不放大伪影的情况下提高亮度，从而实现了干净基线增强的第一路径。

我们认识到将图像信息分离成一个光照成分  $L$  和一个噪声成分  $N$ ，能够实现更精细的表示：

$$\hat{I} = I \times \hat{R} = I \times (I - N) / L \quad (2)$$

在第二条路径中，我们通过计算重建的反射分量  $\hat{R}$  与输入的低光照图像  $I$  的乘积来处理不同的信息分量，从而得到一个保留原始细节信息的低光增强图像  $\hat{I}$ 。

总之，轻度预处理在平衡亮度、保留细节的同时，还确保了低光增强结果的基线质量，以便为随后的 RWKV 阶段做好准备。

#### B. 深度细节挖掘

##### Global Edge Retinex Theory

稳健的 Retinex 理论存在一些局限性：缺乏层次化的边缘表示，以及照明和边缘之间的强耦合。假设 (1) 认为，在低光照条件下，细节优化的关键在于在一个照明不变的特征基础上解耦组件，并结合梯度信息对浅层和深层细节进行差异化增强。我们提出了全球边缘 Retinex (GER) 理论来证实这一点，其核心公式为：

$$I = (R + \alpha \cdot E) \odot L + \beta \cdot N + \gamma \cdot S \quad (3)$$

相较于传统的 Retinex 模型，GER 的改进体现在三个方面：引入一个新的边缘特征项  $\alpha E$  来建模层次化细节；加入一个伪影项  $\gamma S$  以抑制重建噪声；并且将  $N$  重新定义为空间异质噪声。

具体细分如下：深度边缘结构表现出光照不变性，因此 GER 使用反射分量  $R$  作为稳定基础：

$$R = \frac{I - N}{L + \epsilon} \in [0, 1], \quad (4)$$

。这里， $R$  通过  $I - N$  去除空间异质性噪声，并通过  $L + \epsilon$  中和光照干扰，提供一个用于细节增强的“干净”基准。

为了解决浅细节模糊的问题，GER 引入了边缘特征  $E$  来构建增强反射图：

$$R_{\text{enh}} = R + \alpha \cdot E, \quad (5)$$

。以  $R$  为锚点， $E$  指导对浅层和深层细节进行差异增强，缓解了 VRWKV 中的单次递归引起的特征混叠，同时结合梯度信息以加强边缘连续性。

在光预处理步骤中，通过灰度世界假设估计照明分量  $L$ ： $L = \text{Illumination}(I) \in (0, 2)$ 。为了消除低光下明暗区域的边缘误判，GER 设置了伪影残差抑制项  $S_{\text{supp}} = \gamma \cdot S$ 。最终，通过对所有成分进行协同估计，获得增强图像：

$$\hat{I} = R_{\text{enh}} \times L + \beta \cdot N + S_{\text{supp}}. \quad (6)$$

#### Evolving Scanning Mechanism

进化扫描 (ES) 技术通过将边缘流形的几何连续性转换为 VRWKV 传播机制可捕获的时间连续性，来解决假设 (2)。其数学基础在于阿基米德螺旋方程：

$$P(\theta) = (r(\theta) \cos \theta, r(\theta) \sin \theta), \quad (7)$$

。在此， $r(\theta) = a + b\theta$  控制螺旋的初始半径 ( $a$ ) 和扩展速率 ( $b$ )。路径从起始点统一扩展，自然适应从局部到全局区域传播的边缘特性。值得注意的是，我们将单个螺旋扩展为四向螺旋系统，该系统从特征图的四个角启动，并支持顺时针和逆时针方向，以实现低光边缘问题的针对性适应。

为了确保螺旋路径保留边缘点的几何邻接性，引入了严格的拓扑保留约束。约束规则如下：对于边缘点集  $E$  中任意两点  $e_i, e_j$ ，如果它们的黎曼几何距离满足  $\|e_i - e_j\|_{\text{geo}} < \delta$  (其中  $\delta$  为几何邻近阈值)，则它们在螺旋路径生成的一维序列中的时间位置  $t_i, t_j$  必须满足  $|t_i - t_j| < \tau$  (其中  $\tau$  为时间邻近阈值)。

在网络层面，ES-RWKV 是 ES 技术和 RWKV 结构的融合。它弥合了低光边缘的黎曼流形特性与 VRWKV 的时间建模能力之间的差距，有效地增强了边缘连续性的建模性能。

#### C. 双边光谱对准器

边缘细节优化可以在低光照图像中确保特征聚焦，但其优化程度和特征匹配度尤为关键。具体而言，过度强调细节会导致低光区域噪声突显和伪影生成。与亮度和色彩特征的错配会降低视觉感知。为了解决这个问题，我们提出的光谱对齐技术被嵌入到 DRWKV 中，以双边光谱对齐块 (Bi-SAB) 的形式。

在结构上，Bi-SAB 通过交叉注意力 (CrossAttention) 实现亮度和颜色分支之间的特征交互。它利用光谱对齐增强器 (SAE) 来增强图像亮度和减少噪声，并使用 Scharr 算子提取边缘细节。完整的结构块为亮度、颜色、细节和噪声相关问题提供了可行的解决方案。具体过程如算法 1 所示。

#### D. 损失函数

新定义的  $MS^2$ -Loss 关注模型的五个关键属性：结构、边缘、照明、伪影和权重。

受压缩感知理论的启发 [?]，它将 L1 范数应用于图像重建：

$$L_{\text{recon}} = \|I - \hat{I}\|_1, \quad (8)$$

其中  $\|\cdot\|_p$  表示  $p$ -范数。 $L_{\text{recon}}$  专注于平衡像素级错误和结构信息。

Edge sparsity loss: 使用 L1 损失来控制边缘轮廓的稀疏性，并抑制由噪声引起的错误边缘：

$$L_{\text{sparse}} = \|E\|_1, \quad (9)$$

TABLE I

在 LSRW-HUAWEI, LOLv2 和 SDDS 测试基准上的实验结果。背景颜色从深到浅渐变, 其中 红色 表示最优结果, 橙色 表示次优结果, 黄色 依次排列。

Method	GFLOPs	LSRW-Huawei			LOLv2_Real			LOLv2_Syn			SDDS_indoor			SDDS_outdoor		
		PSNR↑	SSIM↑	NIQE↓	PSNR↑	SSIM↑	NIQE↓	PSNR↑	SSIM↑	NIQE↓	PSNR↑	SSIM↑	NIQE↓	PSNR↑	SSIM↑	NIQE↓
RetinexNet [?]	587.47	15.61	0.414	7.235	16.10	0.407	9.425	17.14	0.756	5.405	20.84	0.617	7.924	20.96	0.629	9.947
KinD [?]	34.99	15.77	0.548	5.517	18.67	0.772	9.221	15.22	0.542	6.124	20.98	0.597	7.221	21.65	0.621	8.714
RRDNet [?]	2.10	14.66	0.541	7.725	15.21	0.514	9.912	16.67	0.667	4.894	21.07	0.601	6.921	20.04	0.647	8.997
Zero-DCE [?]	5.21	14.86	0.559	4.215	14.12	0.512	8.652	14.93	0.531	5.507	21.23	0.752	6.941	20.39	0.691	8.014
SCI [?]	0.06	14.78	0.526	3.667	17.30	0.541	8.077	14.96	0.721	4.899	20.43	0.714	6.574	21.33	0.701	7.778
LLFormer [?]	22.52	16.23	0.642	3.947	20.56	0.801	9.162	24.42	0.914	4.614	25.66	0.832	6.627	28.45	0.821	8.011
UHDFormer [?]	48.37	21.07	0.604	4.621	21.59	0.804	7.230	22.60	0.903	5.476	28.13	0.875	7.021	22.75	0.732	6.657
Retinexformer [?]	15.51	22.24	0.701	2.976	21.65	0.835	4.735	25.10	0.925	3.971	28.96	0.879	5.441	28.96	0.896	7.201
IAT [?]	1.44	20.12	0.694	4.217	20.30	0.752	5.232	22.96	0.856	5.512	19.97	0.713	4.177	19.97	0.711	6.417
EnlightenGAN [?]	61.01	17.42	0.651	3.176	18.35	0.617	6.591	17.21	0.725	4.631	20.10	0.602	4.596	19.67	0.623	5.147
MambaIR [?]	60.66	21.14	0.704	3.004	20.11	0.802	4.928	24.75	0.922	4.522	25.11	0.873	3.661	26.35	0.510	4.477
RetinexMamba [?]	34.75	20.88	0.629	3.104	22.34	0.826	4.771	24.71	0.932	4.503	28.21	0.893	3.976	26.22	0.866	4.011
MambaLLIE [?]	20.85	20.64	0.627	3.047	22.14	0.821	4.021	24.81	0.940	4.071	29.74	0.902	4.417	28.92	0.869	4.417
DRWKV (Ours)	1.67	22.34	0.703	2.944	24.12	0.832	3.926	25.02	0.947	3.941	30.26	0.922	3.441	28.96	0.891	3.954

### Algorithm 1 Bi-SAB: 双边频谱对齐块

Input: Low-level Feature  $F_1 : (B, C_i, H_i, W_i)$   
High-level Feature  $F_1' : (B, C_i, H_i, W_i)$   
Output: Bi-SAB Fusion Feature  $F : (B, C_i, H_i, W_i)$   
/\* 注意: 技术细节在补充材料中提供。\*/  
1: /\* 计算查询 (Q)、键 (K) 和值 (V) 的表示。\*/  
2:  $F_Q : (B, C_-, H_-, W_-) \leftarrow \text{Conv}_{3 \times 3}(F_1)$   
3:  $F_K : (B, C_-, H_-, W_-) \leftarrow \text{Conv}_{3 \times 3}(F_1')$   
4:  $F_V : (B, C_-, H_-, W_-) \leftarrow \text{Conv}_{3 \times 3}(F_1)$   
5: 计算机交叉注意力 (CA) 特征表示。  
6:  $F_{\text{attn}} : (B, C_-, H_-, W_-) \leftarrow \text{CA}(F_Q, F_K)$   
7: /\* 计算机特性差异调整 (FDA) 特性表示。\*/  
8:  $Att_{\text{out}} : (B, C_-, H_-, W_-) \leftarrow F_{\text{attn}} \times F_K$   
9:  $F_{DA} : (B, C_-, H_-, W_-) \leftarrow \lambda \cdot Att_{\text{out}} + F_V, \lambda = 0.2$   
10:  $DA_{\text{out}} : (B, C_-, H_-, W_-) \leftarrow F_{\text{attn}} \times F_Q + DA_{\text{out}}$   
11: /\* 计算 Scharr 边缘增强 (SEE) 特征表示。\*/  
12:  $F_{SEE} : (B, C_-, H_-, W_-) \leftarrow \text{SEE}(F_1)$   
13: 计算机光谱对齐增强器 (SAE) 特征表示。  
14:  $F_{SAE} : (B, C_-, H_-, W_-) \leftarrow \text{SAE}(\text{LN}(DA_{\text{out}}))$   
15: 计算机双 SAB 特征表示。  
16:  $F : (B, C_-, H_-, W_-) \leftarrow F_{SAE} + F_{SEE}$   
17: return F

在此, 边缘图  $E$  仅在真实轮廓处非零, 而在其他位置应接近 0, 以满足稀疏性先验条件。

在自然场景中, 光照  $L$  展示了空间平滑性和突然边缘变化的特征, 需要针对性地处理:

$$L_{\text{smooth}} = |\nabla L| \cdot \exp(-\lambda |\nabla I|), \quad (10)$$

, 其中  $\nabla L$  是光照梯度,  $|\nabla I|$  表示图像区域。高  $|\nabla I|$  表

示边缘区域, 此时对  $L$  的约束应放宽, 以强调光照的自然性; 低  $|\nabla I|$  表示平坦区域, 此时应惩罚  $L$  的剧烈变化, 以避免光照中无效的波动。

提出了一种选择正则化的伪影抑制损失, 以最小化伪影能量并确保空间平滑性: 其中  $\delta$  是一个权重参数。当  $\delta > 1$  时,  $\mathcal{L}_{\text{noise}}$  被 TV 正则化项主导, 抑制诸如光晕和条纹的高频振荡; 当  $\delta < 1$  时, 损失由 L1 正则化项主导, 控制残差的幅度。

参数正则化损失: 针对可学习参数  $\alpha$ 、 $\beta$  和  $\gamma$  的优化, L2 正则化等同于权重衰减, 迫使参数值接近 0, 其公式如下:

$$L_{\text{reg}} = \alpha^2 + \beta^2 + \gamma^2 \quad (11)$$

总损失结构由给出, 总损失结构: , 其中

$$\mathcal{L}_{\text{MS2Loss}} = \lambda_1 L_{\text{recon}} + \lambda_2 L_{\text{sparse}} + \lambda_3 L_{\text{smooth}} + \lambda_4 L_{\text{artifact}} + \lambda_5 L_{\text{reg}}, \quad (12)$$

,  $\lambda_i$  是预定义的权重。

我们在 NVIDIA A800-80GB 平台上实施模型训练和推理。具体参数设置如下: 训练周期数  $E$  设为 500; 输入图像大小  $S$  为  $128 \times 128$ ; DRWKV 结构的块数量分别设为  $N_1 = 4$  和  $N_2 = 8$ , 输入通道数  $C$  为 16。模型优化器使用 Adam [?], 参数设为  $\beta_1 = 0.9$  和  $\beta_2 = 0.999$ ; 初始学习率为  $2 \times 10^{-4}$ , 通过余弦退火策略 [?] 逐渐降低到  $1 \times 10^{-6}$ 。

Datasets:

(1) LSRW-Huawei 是 LSRW [?] 数据集的一个子集, 由华为智能手机拍摄的图像组成, 包括 2450 对训练图像和 30 对测试图像。(2) LOLv2 [?] 分为两个子集: LOLv2 Real 和 LOLv2 Synthetic。前者是一个实际低光场景的数据集, 包含 689 对训练图像和 100 对测试图像, 涵盖了各种环境如校园和俱乐部。后者是一个合成数据集, 由 900

对训练图像和 100 对测试图像组成。(3) SDSD [?] 数据集有两个子集：室内和室外。室内子集包括 62 对低光/正常光视频用于训练，以及 6 对低光/正常光视频用于测试。室外子集包含 116 对低光/正常光视频用于训练，以及 10 对低光/正常光视频用于测试。

指标：我们使用传统质量(失真)指标 PSNR 和 SSIM [?] 来评估低光图像增强的结果，旨在测量增强图像与参考图像之间的像素级相似度和失真。此外，使用 NIQE [?] 来评估增强图像的无参考感知质量。

### E. 比较实验

模型比较结果：基于五个测试基准的定量结果，DRWKV 模型在极低 GFLOPS 的情况下，相对于多个基线方法实现了性能改进，如表 I 所示。具体来说，在 LSRW-Huawei 上，DRWKV 在自然性方面表现出色，其特点是较低的 NIQE 基线值和较高的优化率（比第二名方法高 1.1 %）。此外，在 LOLv2 上，DRWKV 在实际和合成测试基准间表现出不同的优化强度：在实际场景中，其 PSNR 平均提高 5.00 dB，在合成场景中 SSIM 提高 0.152，平衡了图像质量的增强和结构信息的保留。在低光场景视频中，通过在 SDSD 中独立增强低光帧，模型在室内和室外场景分别表现出高达 9.42 dB/8 dB 的平均 PSNR 差异，确保图像细节的高“恢复精度”。有关可视化结果，请参见图 2；我们 DRWKV 模型生成的结果展示了连续的目标边缘和高保真细节，与对象集中的边缘高度一致。

TABLE II  
在 LOLv2 实际基准数据集上不同 BLOCK 组合的比较。

Config	Block1	Block2	PSNR↑	SSIM↑	NIQE↓
(I)	Robust Retinex	Robust Retinex	22.57	0.741	4.124
Ours	Robust Retinex	GER	24.12	0.832	3.926

Retinex 配置结果：本实验关注低光照下的边缘失真矛盾。在网络结构方面，预设了可替换结构 Block1 和 Block2 以辅助配置 Retinex，使得 DRWKV 能够达到全局最优。由于 GER 理论的输入端需要接收边缘特征  $E$ （来自 DRWKV 输出端的参数），Block1 受到结构限制。为了最小化对网络架构的干扰，最终采用了单方面固定。将 GER 理论部署到 Block2 中使 PSNR、SSIM 和 NIQE 分别显著提高了 6.9 % / 12.3 % / 4.8 %，如表 II 所示。这强烈验证了 GER 理论中梯度导向的全局边缘耦合核心思想，展示了其在解决低光照图像边缘问题上的强大实用性，并进一步确认了假设 (1)。

### F. 消融研究

损失性能结果：对提出的损失函数组合进行了一项有针对性的消融研究。在图 3 (a) 中，缺少分解一致性损失  $L_{recon}$  导致重建图像整体较暗，边缘信息不完整，间接验证了其在结构和亮度重建中的作用。在图 3 (b) 中，去除边缘稀疏性损失  $L_{sparse}$  导致高斯噪声污染平坦区域，从而生成虚假边缘，突显其在保持和增强低光边缘中的贡献。图 3 (c) 和 3 (d) 描述了去除光照平滑损失  $L_{smooth}$  和去伪影损失  $L_{artifact}$  分别导致的不均匀光照和过度去噪效果，确认了其平衡角色。图 3 (五) 展示了当去除参数正则化损失  $L_{reg}$  时导致的图像伪影现象，表明有效

的正则化可以提高模型对未知数据的自适应性。此外，模型的损失参数如下所示： $\lambda_1 = 1$ 、 $\lambda_2 = 0.01$ 、 $\lambda_3 = 0.1$ 、 $\lambda_4 = 0.05$  和  $\lambda_5 = 1e-4$ 。

组件消融结果：本实验旨在验证假设 (2) 中提出的用于低光图像边缘的进化扫描技术的有效性。在结构上，采用逐级组件集成策略，以 LOLv2 实作为测试基准，基本结构作为测试框架，构建五级渐进验证系统。每一层的组件配置、消融变量和验证结果如表 III 所示。基础结构方面，我们保留了 VRWKV 的 ViT 风格块叠加结构，并引入 SSIM+L1 损失，使其初步具备处理低光问题的能力。然而，该模型在光照降解场景中特征提取几乎无效，表现为极低的 SSIM (0.415) 和 PSNR (12.57 dB) 值。

验证 (1)：在嵌入了演化扫描模块 (ES-RWKV) 后，SSIM 和 PSNR 的性能分别提高了 42.4 % 和 29.4 %。这明确验证了所提出的螺旋扫描与边缘的空间拓扑相兼容，能够确保低光对象边缘的连续性。

验证 (2)：通过将网络从 ViT 风格的块堆叠方案修改为 UNet 风格的方案，模型性能优化关注于细节精度。在参数数量和计算负载有一定增加的代价下，PSNR 指标提升了 3.16 dB。

验证 (3)：增加一个浅层照明恢复阶段导致图像质量评估指标略有提升，表明此阶段有助于改善像素误差和结构。

验证 (4)：嵌入双 SAB 模块并采用高性能的多模块级联显著优化了模型结构。参数数量和计算负荷分别减少了 49.3 % 和 43.8 %，而 SSIM 和 PSNR 分别提高了 0.028 和 0.8 dB，使模型具有基本的轻量化特性。

验证 (5)：将 SSIM+L1 损失替换为提出的  $MS^2$ -Loss，特别优化了训练过程，其中 SSIM 和 PSNR 分别优化到 0.832 和 24.12 dB。

此时，方法论中提出的假设已得到了充分验证。本文从四个维度构建了一个完整的研究闭环：理论、技术、实验和应用。

TABLE III

在 LOLv2 REAL 基准数据集上进行消融研究。如果不选择 RWKV 块和训练损失，将默认使用 Bi-RWKV 和 SSIM+L1 损失。✓✗ 表示使用 ViT 风格的块堆叠结构，而 ✓ 则表示采用 UNet 风格的架构。

ES-RWKV	SIR	Bi-SAB	MS2Loss	SSIM↑	PSNR↑	Params.(M)	GFLOPs
				0.415	12.57	10.51	0.02
✓				0.591	16.27	14.57	2.65
✓				0.607	19.43	16.36	2.95
✓	✓			0.748	22.67	16.36	2.97
✓	✓	✓		0.776	23.47	8.29	1.67
✓	✓	✓	✓	0.832	24.12	8.28	1.67

### G. 扩展验证

为了评估泛化能力，我们将 DRWKV 应用于 UAVDark135 基准上的低光多目标跟踪，该基准包含 135 个在 30 FPS 和 1920×1080 分辨率下的不同场景序列。

主流的跟踪算法在有和没有 DRWKV 的情况下进行了测试，使用 MOTA、IDF1 和 HOTA 作为指标。DRWKV 始终提高了性能，尤其是在身份保留方面。ByteTrack 的性能提高了 5.189 % / 3.076 % / 2.705 %，而 BoT-SORT 和其他算法的性能提高了 1.5 % - 7 %。

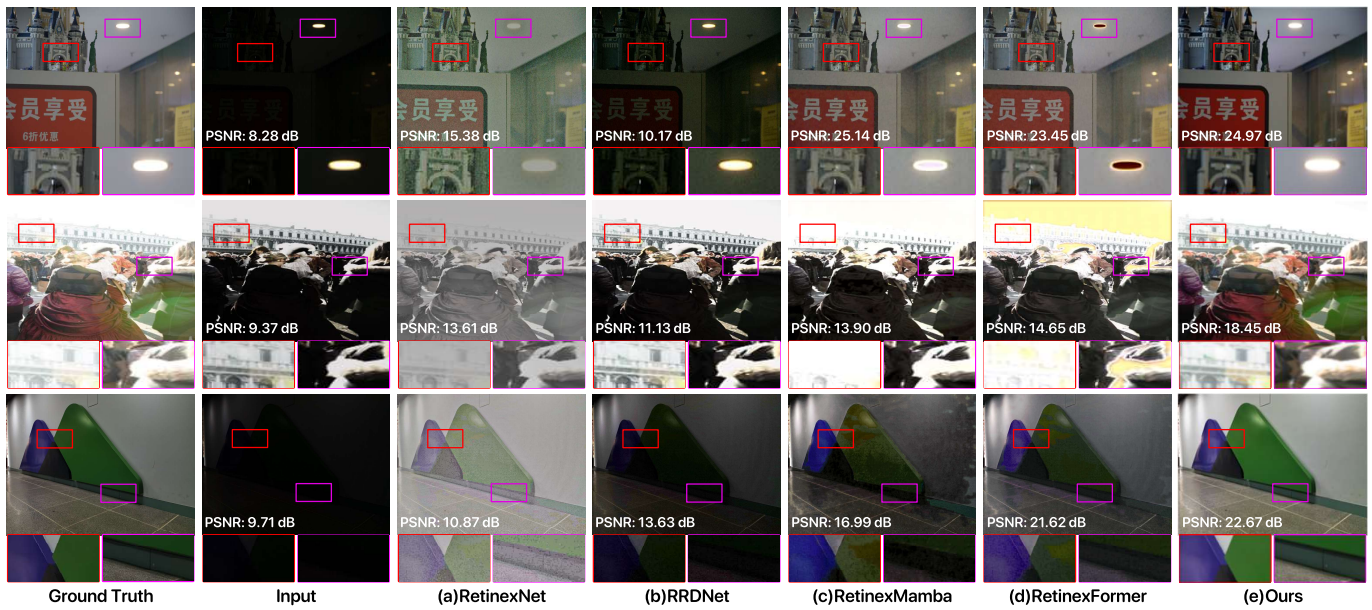


Fig. 2. RetinexNet, RRDNet, RetinexMamba, RetinexFormer 和 DRWKV 在 LOLv2-Real、LOLv2-Synthetic 和 LSRW-Huawei 数据集上的可视化结果（放大显示每个模型的详细信息）。



Fig. 3. 对所提出的分解一致性损失  $L_{recon}$ 、边缘稀疏性损失  $L_{sparse}$ 、光照平滑性损失  $L_{smooth}$ 、伪影抑制损失  $L_{artifact}$  和参数正则化损失  $L_{reg}$  进行消融研究。

相反，Deep OC-SORT 显示出混合结果，MOTA 增加但 IDF1 减少，这表明对增强特征有结构敏感性。实施细节见附录。

TABLE IV  
DRWKV 在低光物体跟踪性能上的扩展验证结果。MOTA [?]、IDF1 [?] 和 HOTA [?] 是物体跟踪常用的性能评估指标。

Methods	w/o LLIE Module			DRWKV (Ours)		
	MOTA(%)	IDF1(%)	HOTA(%)	MOTA(%)	IDF1(%)	HOTA(%)
ByteTrack [?]	23.434	62.776	50.117	28.623	65.852	51.432
BoT-SORT [?]	18.242	62.916	49.556	25.496	64.049	51.435
Hybrid-SORT [?]	20.684	62.054	49.196	29.355	66.365	52.905
OC-SORT [?]	21.062	62.967	49.469	28.102	66.000	52.501
Deep OC-SORT [?]	19.514	63.205	49.857	24.929	58.016	49.124
StrongSORT [?]	20.923	53.583	45.553	25.148	64.677	51.842

在这项研究中，我们引入了 DRWKV 模型来解决低光图像增强中保持边缘细节的挑战。通过整合全局边缘 Retinex

(GER) 理论，该模型解耦了光照和边缘结构，确保即使在极端低光条件下也能保持高保真的边缘连续性。进化 WKV 注意机制增强了空间边缘连续性，而双边谱对齐器 (Bi-SAB) 和  $MS^2$ -Loss 优化了亮度和色度对齐，减少了伪影并提高了视觉质量。在基准数据集上的大量实验表明，DRWKV 在 PSNR、SSIM 和 NIQE 方面优于现有方法，而且计算成本更低。此外，它在低光多目标跟踪等任务中也证明了其有效性，尤其是在身份识别方面提高了性能。未来的工作将侧重于将 DRWKV 适配于移动平台，并在各种低光场景中进一步提高其鲁棒性，缩小学术研究与实际应用之间的差距。

# DRWKV: Focusing on Object Edges for Low-Light Image Enhancement

## Appendices & Supplementary Material

在本附录中，我们首先对主要文本中介绍的 DRWKV 模型的核心计算流程进行了全面的阐述，并详细说明了每个关键模块的实现细节。随后，为了严格验证所提出的 DRWKV 模型声称的增强能力，我们设计了一系列扩展实验；实验方案和结果的相应分析在接下来的部分中详细说明。

在本节中，我们详细展示了在第三节中介绍的 Light Preprocessing 流程、ES-RWKV 计算流程及其组成的细粒度模块。

Light Preprocessing 以低光图像  $I \in \mathbb{R}^{3 \times H \times W}$  作为输入。首先，在灰世界假设下，它计算出全局光照成分  $L_{ill} \in \mathbb{R}^{C \times H \times W}$ 。该假设认为，在颜色平衡的场景中，所有像素的平均反射率趋向于中性灰色。在这一假设的指导下，表征整体光照强度空间分布的全局光照图  $L_{ill}$  从输入图像  $I$  中推导出来，并作为最终增强的基础。

输入图像  $I$  最初经过一个具有 ReLU 激活的深度可分  $3 \times 3$  卷积层处理。然后，结果特征图被并行输入到三个独立的  $1 \times 1$  卷积模块中，每个模块具有不同的激活函数以生成特定的组件：一个具有 Sigmoid 激活的模块输出结构化伪影组件  $S \in \mathbb{R}^{3 \times H \times W}$ ，旨在捕捉潜在的结构化干扰模式；一个具有 Sigmoid 激活的模块输出噪声估计组件  $N \in \mathbb{R}^{3 \times H \times W}$ ，用于量化噪声污染水平；以及一个具有 Tanh 激活的模块输出局部光照估计组件  $L \in \mathbb{R}^{1 \times H \times W}$ ，提供了更精细、空间自适应的光照强度表示。其过程如下：

$$B(I) = \text{ReLU}(\text{DepthConv}_{3 \times 3}(I)), \quad (13)$$

$$S = \text{Sigmoid}(\text{Conv}_{1 \times 1}(B^3(I))), \quad (14)$$

$$L = \text{Sigmoid}(\text{Conv}_{1 \times 1}(B^3(I))), \quad (15)$$

$$N = \text{Tanh}(\text{Conv}_{1 \times 1}(B^3(I))), \quad (16)$$

，其中  $\text{DepthConv}_{3 \times 3}(\cdot)$  表示  $3 \times 3$  深度卷积， $\text{Conv}_{1 \times 1}(\cdot)$  表示  $1 \times 1$  卷积， $\text{ReLU}(\cdot)$ 、 $\text{Sigmoid}(\cdot)$  和  $\text{Tanh}(\cdot)$  分别是激活函数，而  $B^3(\cdot)$  表示连续三次应用  $B(\cdot)$ 。

噪声估计  $N$ 、光照估计  $L$  和原始输入  $I$  通过

$$\hat{R} = \frac{I - N}{L}, \quad (17)$$

结合，得到重建的反射率  $\hat{R} \in \mathbb{R}^{3 \times H \times W}$ 。然后， $I$  和  $\hat{R}$  结合得到增强图像  $\hat{I} \in \mathbb{R}^{3 \times H \times W}$ ：

$$\hat{I} = I \odot \hat{R}, \quad (18)$$

，其中  $\odot$  表示元素级乘法。

### H. 演变扫描 RWKV

与传统的空间建模方法不同，所提出的 ES-RWKV 突破了空间大小限制，保持了结构完整性，增强了建模能力，并实现了跨空间信息交互，同时显著减少了处理复杂空间特征的计算开销，从而为视觉任务中的空间特征提取和融合提供了一种有效的解决方案。ES-RWKV 模块分解为两个连续的阶段：演变空间混合和通道混合。演变扫描 RWKV 结构如图所示。4

1) 演变的的空间混合：他们解决了空间尺寸限制和结构完整性之间的困境。通道级的 Q-Shift 使模型能够灵活地适应不同的任务而无需额外计算。结合 EV-WKV 机制，引入了一种嵌入的八螺旋扫描模式，以更全面和有效地捕捉空间依赖性，抑制与空间提取无关的冗余。随后，通过 LayerNorm 巩固空间信息，使模型能够自动强调显著特征和减少噪声，从而促进下游层学习有效的空间模式。

对于输入特征  $X \in \mathbb{R}^{T \times C}$ ，我们通过一个通道分区的 Q-Shift 机制来增强标准的顺序处理框架。通过在通道轴上切片和连接，我们构建了一个位移特征  $X^\dagger$ 。结合可学习向量  $\mu_{(*)}$ ，我们动态地插值并融合原始特征  $X$  和位移特征  $X^\dagger$ ，进一步突破传统方法的空间尺寸限制，保持空间结构的完整性，并增强模型捕捉跨空间位置信息的能力。这个过程生成了用于空间混合计算的成分：

$$Q\text{-Shift}_{(*)}(X) = X + (1 - \mu_{(*)})X^\dagger \quad (19)$$

$$\begin{aligned} X^\dagger[h, w] = & \text{Concat}(X[h - 1, w, 0 : C/4], \\ & X[h + 1, w, C/4 : C/2], \\ & X[h, w - 1, C/2 : 3C/4], \\ & X[h, w + 1, 3C/4 : C]) \end{aligned} \quad (20)$$

在这个上下文中， $(*) \in \{R, K, V\}$  表示对  $X$  和  $X^\dagger$  执行的三种插值操作。这些操作由  $\mu$  控制，并用于随后 R、K 和 V 的计算。在这里， $h$ 、 $w$  和  $C$  分别表示高度、宽度和通道数。

随后，在 Q-Shift 操作之后， $X$  被分解为三个独立的组件  $R_s, K_s, V_s \in \mathbb{R}^{T \times C}$

$$(*)_s = Q\text{-Shift}_{(*)}(X)W_{(*)} = (X + (1 - \mu_{(*)})X^\dagger)W_{(*)} \quad (21)$$

$W_{(*)}$  表示权重。为了确保数值稳定性，防止训练发散并加速训练收敛，我们引入了 LN 层。Evolving Spatial Mix 块的总输出如下：

$$O_s = X + \text{LN}((\sigma(R_s) \odot wkv)W_{O_s}), \quad (22)$$

$$\text{wherewkv} = \text{EV} - \text{WKV}(K_s, V_s), \quad (23)$$

这里， $\sigma$  表示 S 形激活函数， $\odot$  代表逐元素乘法运算，LN 表示层归一化， $W_{O_s}$  表示权重。EV-WKV 是一种基于双向 WKV 的线性扫描公式  $wkvt$  提出的演变扫描机制。

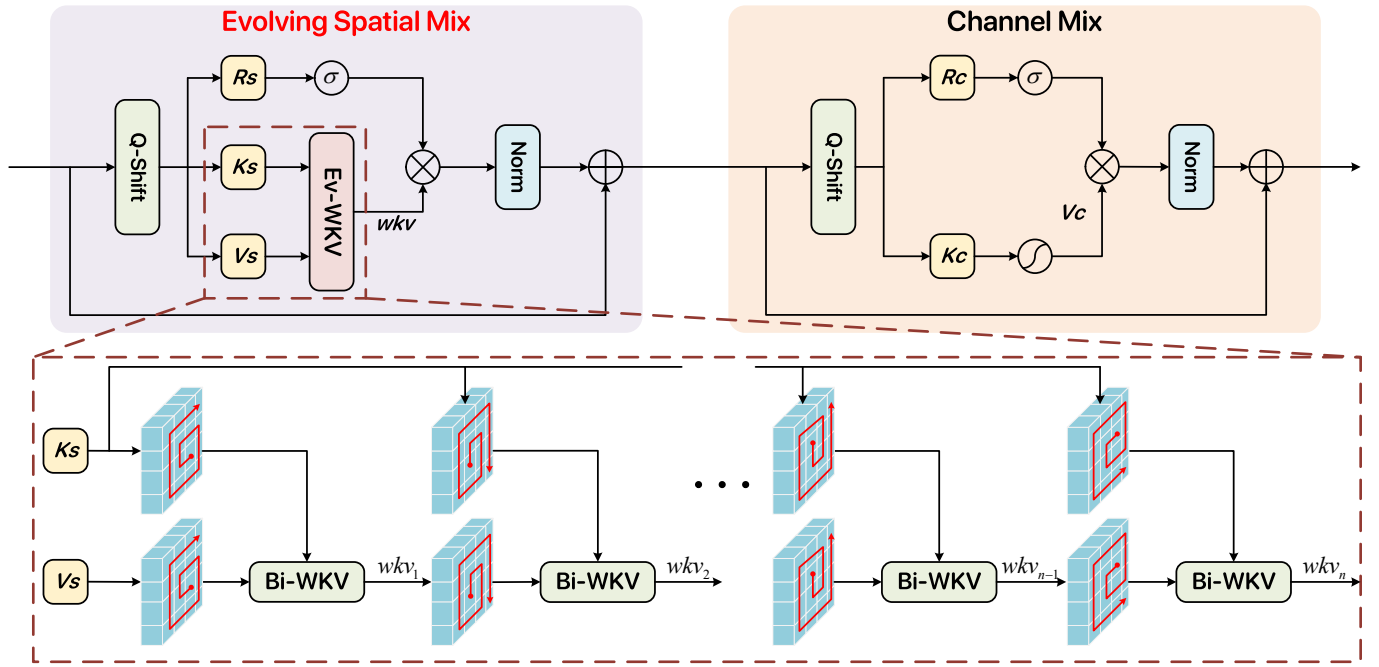


Fig. 4. 演化的扫描 RWKV 结构由两部分组成：演化的空间混合和通道混合。

我们重新引入了 Q-Shift 机制，并将其与 SquaredReLU 激活函数和线性投影相结合。这旨在解决第一阶段空间混合后可能存在的不足之处，例如对通道特征之间相关性的不足探索以及过于简单的特征表达。它确保通道维度上的特征融合更加充分，增强特征的非线性表达能力和辨识能力，并通过层归一化 (LN) 进一步保持数值分布的稳定性，从而为模型深层的特征学习提供更丰富和更稳健的输入。

首先，我们将第一阶段的输出特征  $O_s$  作为输入，并通过 Q-Shift 机制分别构造用于通道交互的查询特征  $R_c$  和键特征  $K_c$ 。计算过程如下：

$$\begin{aligned} R_c &= Q - Shift_R(O_s)W_R, \\ K_c &= Q - Shift_K(O_s)W_K \end{aligned} \quad (24)$$

这里， $Q - Shift_R$  和  $Q - Shift_K$  在原始特征  $O_s$  及其偏移版本  $O_s^\dagger$  之间动态插值，使用可学习的参数  $\mu_{(R)}$  和  $\mu_{(K)}$ 。这一设计不仅保留了原始特征中的空间结构信息，还引入了跨位置的通道依赖，使得模型能够自适应地关注不同通道之间的长距离相关性。

随后，为了增强特征的非线性表达能力并缓解梯度消失问题，应用 *SquaredReLU* 激活函数进行  $K_c$  变换。

此外，转换后的特征通过线性投影投影到值空间，并通过残差连接和层规范化操作生成通道混合模块的输出。详细过程如下：

$$O_c = O_s + LN((\sigma(R_c) \odot V_c)W_{O_c}), \quad (25)$$

$$\text{where } V_c = SR((O_s + (1 - \mu_{(k)})O_s^\dagger)W_K)W_V \quad (26)$$

### I. 双边光谱对齐

本节重点介绍双向光谱对齐器的三个核心组件：光谱对齐增强器 (SAE)、交叉注意力 (CA) 和 Scharr 边缘增强 (SEE)。

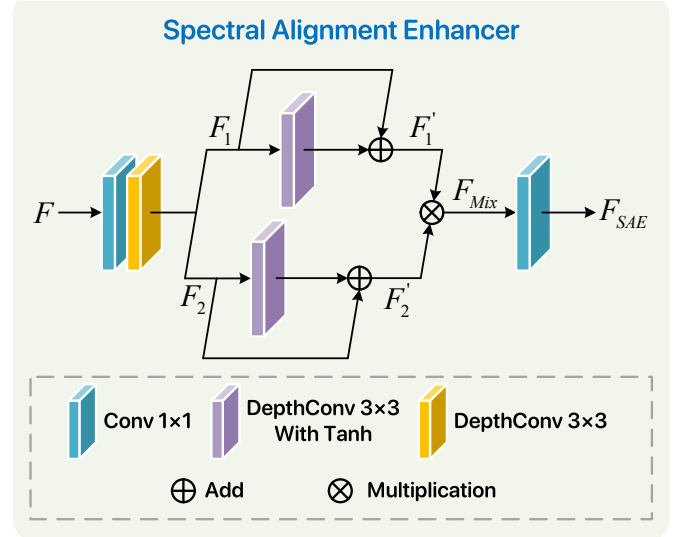


Fig. 5. 谱对齐增强器的结构。

1) 谱对齐增强器 (SAE)：为了增强低光照图像增强模型在不同光照条件下对光谱特征的感知和处理，我们提出了一种自适应特征增强模块，称为光谱对齐增强器 (SAE)，如图 5 所示。具体而言，该模块通过三个连续阶段实现对低光照特征的精确优化：特征扩展、分离增强和交互融合。

最初，输入特征  $F \in \mathbb{R}^{C \times H \times W}$  通过  $1 \times 1$  卷积进行通道扩展，然后分成两个分支，从而保留了足够的表现能力，



以捕捉低光环境中的微弱光谱细节。

$$F_1, F_2 = \text{Split}(\text{DepthConv}_{3 \times 3}(\text{Conv}_{1 \times 1}(F))) \quad (27)$$

这里,  $\text{Split}(\cdot)$  表示通道均匀的减半。

随后, 将两个特征流  $F_1, F_2 \in \mathbb{R}^{C_1/2 \times H \times W}$  分别通过深度可分离卷积处理, 以聚焦于局部空间模式, 接着是一个非线性  $\text{Tanh}$  激活和残差连接, 从而得到精炼特征  $F'_1, F'_2 \in \mathbb{R}^{C_1/2 \times H \times W}$ 。这个设计使得 SAE 能够增强关键弱光线索 (如边缘和纹理) 的提取, 同时防止由于过多变换导致的细节丢失。该过程可以表达为:

$$F'_1 = F_1 + \text{Tanh}(\text{DepthConv}_{3 \times 3}(F_1)) \quad (28)$$

$$F'_2 = F_2 + \text{Tanh}(\text{DepthConv}_{3 \times 3}(F_2)) \quad (29)$$

最后, 这两条精细化特征流逐元素相乘以实现动态交互, 自适应地增强信息线索。随后, 一个  $1 \times 1$  卷积压缩结果回到原始通道维度, 生成优化后的特征  $F_{SAE} \in \mathbb{R}^{C \times H \times W}$ 。这个过程被公式化为:

$$F_{SAE} = \text{Conv}_{1 \times 1}(F'_1 \times F'_2) \quad (30)$$

在低光增强中, 亮度流  $x$  和色度流  $y$  在空间统计和语义内容上通常是不对齐的。我们提出了一个轻量级的交叉注意力 (CA) 模块, 让  $x$  通过多头、L2 归一化的交叉协方差注意力从  $y$  中查询相关信息, 该注意力仅通过深度卷积实现。整个过程如下总结, 并严格对应于发布的代码, 如图 ?? 所示。

首先, 给定两个特征图:

$$x, y \in \mathbb{R}^{C \times H \times W} \quad (31)$$

, 其中  $x$  来自亮度子网,  $y$  来自颜色子网。

查询 / 键 / 值生成

所有线性投影被实现为  $1 \times 1$  卷积, 随后是  $3 \times 3$  深度卷积, 以在参数效率高的同时保留局部上下文。

$$Q = \text{DWConv}_{3 \times 3}(\text{Conv}_{1 \times 1}(x)) \in \mathbb{R}^{C \times H \times W} \quad (32)$$

$$KV = \text{DWConv}_{3 \times 3}(\text{Conv}_{1 \times 1}(y)) \in \mathbb{R}^{2C \times H \times W} \quad (33)$$

然后, 将  $KV$  张量按通道分割为键和值:

$$K, V = \text{split}(KV, \text{dim} = 1), K, V \in \mathbb{R}^{C \times H \times W} \quad (34)$$

多头重塑和 L2-归一化

为了利用互补的注意力模式, 我们将每个张量重新塑造为  $\text{XMATHX}_h$  个头:

$$Q_{\text{head}} = \text{rearrange}(Q) \in \mathbb{R}^{n_h \times c_h \times (H \cdot W)} \quad (35)$$

$$K_{\text{head}} = \text{rearrange}(K) \in \mathbb{R}^{n_h \times c_h \times (H \cdot W)} \quad (36)$$

$$V_{\text{head}} = \text{rearrange}(V) \in \mathbb{R}^{n_h \times c_h \times (H \cdot W)} \quad (37)$$

为了确保训练的稳定性, 在最后一个维度上应用 L2 归一化:

$$\hat{Q} = \frac{Q_{\text{head}}}{\|Q_{\text{head}}\|_2}, \hat{K} = \frac{K_{\text{head}}}{\|K_{\text{head}}\|_2} \quad (38)$$

交叉协方差注意力

我们计算  $x$ -query 和  $y$ -key 之间的交叉协方差注意力矩阵:

$$\text{attn} = \text{softmax}\left(\frac{\hat{Q}\hat{K}^T}{\tau}\right) \in \mathbb{R}^{n_h \times c_h \times c_h} \quad (39)$$

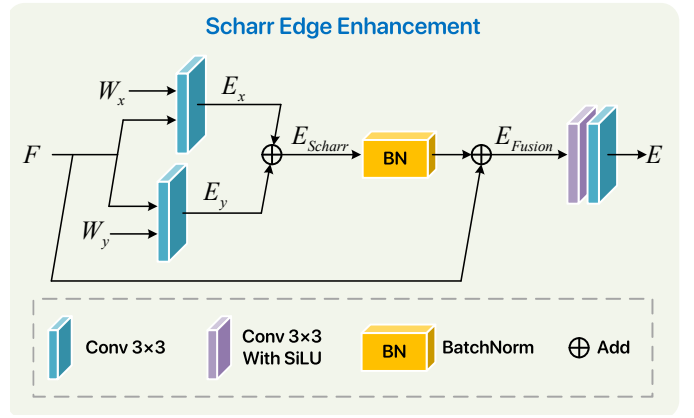


Fig. 6. Scharr 边缘增强的结构。

其中温度  $\tau$  是一个可学习的标量, 初始值为  $d$  然后通过与  $y$ -value 的矩阵乘法来聚合注意到的特征:

$$\text{out} = \text{attn} \cdot V_{\text{head}} \in \mathbb{R}^{n_h \times c_h \times (H \cdot W)} \quad (40)$$

输出投影

最后, 多头输出被重新排列, 并通过一个  $1 \times 1$  卷积投影回原始通道维度:

$$\text{Output} = \text{Conv}_{1 \times 1}(\text{rearrange}(\text{out})) \in \mathbb{R}^{C \times H \times W} \quad (41)$$

得到的张量按元素逐项添加到亮度分支上, 从而生成一个既感知颜色又抑制噪声的增强特征。

2) Scharr 边缘增强 (SEE): 在低光图像增强领域, 精确保留边缘轮廓和完整恢复纹理细节一直是限制算法性能的核心挑战。在低光环境中, 图像信噪比 (PSNR) 显著降低, 导致边缘区域的梯度信息被噪声淹没, 从而变弱。同时, 传统的增强方法在提高亮度的过程中, 常因过度的平滑操作或噪声放大效应进一步加剧边缘模糊和纹理丢失, 最终导致增强结果在视觉上表现为“亮度提高但细节缺失”的失真, 如图 6 所示。为了应对上述问题, 我们提出了一个 Scharr 边缘增强 (SEE) 模块。通过整合传统图像处理的先验知识与深度学习的特征学习能力, 它为低光场景中的结构感知增强提供了有效的解决方案。具体来说, 对于给定的输入特征  $F \in \mathbb{R}^{C \times H \times W}$ , 我们首先使用预设的 Scharr 卷积核和卷积权重  $W_x$ 、 $W_y$  进行按通道过滤操作, 以模拟传统图像处理中梯度提取的过程, 分别获得  $x$  方向梯度特征  $E_x \in \mathbb{R}^{C \times H \times W}$  和  $y$  方向梯度特征  $E_y \in \mathbb{R}^{C \times H \times W}$ 。

随后, 为了抑制噪声干扰并聚合多方向边缘信息, 我们使用 L1 范数融合  $E_x$  和  $E_y$  来计算 Scharr 边缘强度特征  $E_{\text{Scharr}} \in \mathbb{R}^{C \times H \times W}$ 。

然后, 边缘强度特征通过残差机制与原始输入特征进行融合, 以获得融合特征  $E_{\text{Fusion}} \in \mathbb{R}^{C \times H \times W}$ 。一方面, 该操作保留了原始特征中的结构先验; 另一方面, 它避免了在边缘信息指导下由于特征转换过度而引起的“边缘过度曝光”, 从而确保了边缘增强的自然性。

最后, 为了实现特征维度的自适应调整和语义信息的深度融合, 使用  $1 \times 1$  卷积和  $3 \times 3$  卷积进行维度缩减和恢复, 最终得到输出  $E \in \mathbb{R}^{C \times H \times W}$ 。

具体的计算过程可以表达如下：

$$E_x = Conv_{3 \times 3}(F, W_x), E_y = Conv_{3 \times 3}(F, W_y) \quad (42)$$

$$E_{Scharr} = \|E_x\|_1 + \|E_y\|_1 \quad (43)$$

$$E_{Fusion} = F + BN(E_{Scharr}) \quad (44)$$

$$E = Conv_{3 \times 3}(SiLU(Conv_{1 \times 1}(E_{Fusion}))) \quad (45)$$

这里,  $Conv_{3 \times 3}(\cdot, \cdot)$  表示具有固定权重的  $3 \times 3$  卷积,  $BN(\cdot)$  代表 BatchNorm 操作, 而  $SiLU(\cdot)$  表示 SiLU 激活函数。

为了在复杂的真实世界条件下严格评估增强算法的感知质量, 我们设计并进行了一项用户主观评价实验。从公开可用的 LSRW 华为基准中随机选择了十张低光图像, 并通过七种增强模型进行处理: RetinexNet、Zero-DCE、RRDNet、RetinexFormer、RetinexMamba、MambaIR 以及我们提出的 DRWKV 来生成相应的增强结果。然后招募了一百名年龄和专业背景各异的参与者。在一种独立的双盲协议下, 参与者在整体视觉质量、局部细节恢复和全局色彩保真度这三大关键感知维度上, 使用五分制 Likert 量表从一分 (最差) 到五分 (最佳) 进行评分。平均意见评分报告在表 ?? 中。在所有三个维度上, 所提出的 DRWKV 模型都获得了最高评分, 这表明其在人类视觉角度下具有出色的感知增强性能。

#### J. 低光照物体跟踪实验

上述实验已经表明, DRWKV 在基本实验场景中已经表现出卓越的性能。为了研究其泛化能力, 我们设计了一项低光多目标跟踪任务。在数据集方面, 我们采用了无人机低光跟踪基准 UAVDark135。该数据集总共包含 135 个视频序列, 涵盖了各种跟踪场景, 如交叉路口和高速公路, 以及跟踪目标, 包括行人、船只和车辆。视频以 30 帧每秒的帧率拍摄, 分辨率为  $1920 \times 1080$ 。

在方法学方面, 我们使用 MOTA (多目标跟踪精度)、IDF1 (身份识别 F1 得分) 和 HOTA (高阶跟踪精度) 作为评估指标, 比较了主流目标跟踪算法在集成和未集成 DRWKV 弱光增强模块时的性能。

实验结论表明, DRWKV 普遍提高了多种主流跟踪算法的准确性, 尤其在多目标跟踪和身份保持方面表现突出。同时, 该模块在优化 Deep OC-SORT 方面的局限性提醒我们, 面对网络结构的低适应性和弱调整性, 可能无法同时提高检测准确性和关联能力。