1. Junjie, T. A. O., et al. "Image defocus deblurring method based on gradient difference of boundary neighborhood." Virtual Reality & Intelligent Hardware 5.6 (2023): 538-549. 基于边界邻域梯度差的图像去焦去模糊方法

刊名：[虚拟现实和智能硬件](https://www.sciencedirect.com/journal/virtual-reality-and-intelligent-hardware)

本文是一篇关于基于边界邻域梯度差异的图像散焦去模糊方法的研究论文。

研究背景为静态多深度层场景下，现有散焦图像去模糊方法存在因模糊量估计不准确导致边缘环伪影或去模糊不足，以及非盲反卷积先验知识不强致使图像细节恢复困难等问题。

为此，研究提出了基于边界邻域梯度差异的散焦图像模糊图估计方法，该方法利用边界邻域梯度差异准确获取模糊量，可防止边界环伪影；再通过所得模糊图进行模糊检测，判断图像是否需要去模糊，能提高去模糊效率且无需人工干预；最后设计基于模糊量选择策略和稀疏先验的非盲反卷积算法实现图像去模糊。

实验结果显示，该方法与现有方法相比，PSNR和SSIM平均分别提高4.6%和7.3%，能更好地解决散焦图像去模糊中的边界环伪影和细节信息保留问题。

此外，论文还介绍了研究的相关工作、方法概述、详细阐述、实验结果与分析等内容，最后总结了研究成果并提及未来计划。

1. UnderwaterImage2IR: Underwater impulse responsegeneration via dual-path pre-trained networks andconditional generative adversarial networks通过双路径预训练网络和条件生成对抗网络生成水下脉冲响应

刊名：计算机动画和虚拟世界

本文提出**UnderwaterImage2IR**方法，通过**双路径预训练网络**和**条件生成对抗网络（CGAN）** 从水下图像生成声学脉冲响应（IR），实现低成本、高精度的水下视觉图像到声学信息的跨模态转换。该方法利用一个分支提取图像空间特征，另一个分支识别水下特征，输入 CGAN 生成匹配场景的声学 IR。实验表明，其在生成水下声学 IR 方面优势显著，在非水下场景也保持高精度，消融实验验证了各组件有效性，专家评估也证实其性能。

1. Highlight mask-guided adaptive residual network for singleimage highlight detection and removal高光掩模引导自适应残差网络，用于单图像高光检测和去除

刊名：计算机动画和虚拟世界

本文提出**高亮掩码引导的自适应残差网络（HMGARN）**，用于单图像高光检测与去除。该网络包含**detection-net**、**自适应去除网络（AR-Net）** 和**reconstruct-net**三个主要组件。detection-net 结合**CBAM 注意力机制**精准预测高光掩码；AR-Net 在掩码引导下自适应去除高光；reconstruct-net 通过**多尺度特征融合残差模块（MSFM）** 优化结果，去除残留高光并恢复细节。在**SHIQ 数据集**上的实验表明，其高光检测准确率达**0.98**，高光去除的 SSIM 达**0.95**，优于现有方法，消融实验验证了各组件有效性。

**研究背景与目标**

* **背景**：高光会扭曲图像颜色、模糊物体轮廓和纹理，影响目标检测、图像分割等计算机视觉任务。传统方法依赖手动参数和先验假设，在复杂场景中表现有限；现有深度学习方法泛化能力不足，难以应对多样的高光情况。
* **目标**：提出高光掩码引导的自适应残差网络（HMGARN），实现单图像高光的精准检测与去除，同时保留图像细节和质量。

1. SAT-Net:Structure-Aware Transformer-Based Attention Fusion Network for Low-Quality Retinal Fundus Images Enhancement基于结构感知变换的注意力融合网络在低质量眼底图像增强中的应用

刊名：ISTEX类文件杂志（IEEE里面的）

本文提出了一种名为**SAT-Net**的新型结构感知 Transformer 注意力融合网络，用于低质量视网膜眼底图像增强。该网络包含**Transformer-based 注意力融合模块**，能同时捕捉全局空间依赖关系并强调重要特征通道；引入**跨质量知识蒸馏技术**，通过高性能教师网络指导轻量级学生网络，捕捉低质量图像细节；还设计了**结构感知多尺度损失函数**，利用可训练子网络获取不同尺度边缘结构，约束关键血管和毛细血管细节。实验表明，SAT-Net 在合成和真实眼底图像数据集上优于其他先进方法，且在血管分割、视盘 / 杯检测任务中也表现出色

5. EAPT: Efficient Attention Pyramid Transformer for Image Processing用于图像处理的高效注意力金字塔转换器

刊名：ISTEX类文件杂志（IEEE里面的）

本文提出了Efficient Attention Pyramid Transformer（EAPT），旨在解决基于 patch 的 Transformer 模型在视觉任务中存在的固定尺寸 patch 破坏语义信息、patch 间信息交流受限以及位置编码不适应视觉数据等问题。EAPT 包含可变形注意力（Deformable Attention）、编解码通信模块（En-DeC module）和多维连续混合描述符（MCMD） 三个关键组件，在图像分类、目标检测和语义分割任务上的实验表明，其性能优于现有先进方法，消融实验也验证了各组件的有效性。

**一、研究背景**

现有基于 patch 的 Transformer 模型在视觉任务中存在以下问题：

* **固定尺寸 patch 的局限**：视觉元素大小和形状各异，固定尺寸的 patch 难以完整覆盖，可能破坏语义信息。
* **patch 间信息交流受限**：传统方法仅在单个 patch 或局部窗口内计算注意力，限制了全局信息的传递。
* **位置编码不适应视觉数据**：现有方法多采用 NLP 中的低维编码技术，无法有效捕捉高维视觉数据的位置信息，或增加学习成本。

为解决这些问题，本文提出了 Efficient Attention Pyramid Transformer（EAPT）。

**二、EAPT 核心组件**

1. **可变形注意力（Deformable Attention）**
   * 为 patch 中的每个位置学习一个偏移量，使固定尺寸的 patch 能更好地覆盖不同大小和形状的视觉元素。
   * 通过边界约束核限制偏移后的位置在特征边界内，采用双线性插值计算虚拟位置的值，不增加额外计算成本。
2. **编解码通信模块（En-DeC module）**
   * 基于编解码架构实现所有 patch 间的信息交流，首先通过编码压缩所有 patch 的信息，再通过解码恢复并将结果添加到注意力计算中。
   * 与 Swin Transformer 的局部窗口移位机制不同，实现全局信息交流。
3. **多维连续混合描述符（MCMD）**
   * 专为视觉 Transformer 设计的位置编码，独立编码不同维度后进行混合，适应任意维度和长度的 patch。
   * 具有归纳性、数据驱动性、对称性、并行性和参数高效性等特性，通过高斯函数描述位置信息，且编码与特征值相关联。