MS-SCANet：一种用于无参考图像质量评价的基于多尺度 Transformer 的双重注意力架构

Mayesha Maliha R. Mithila 和 Myl`ene C.Q.
Farias德克萨斯州立大学计算机科学系美国圣马科斯
elx12@txstate.edu, mylene@txstate.edu

arXiv:2602.04032v1 [eess.IV] 3 Feb 2026

摘要—**我们提出了多尺度空间通道注意力网络（MS-SCANet），这是一种专为无参考图像质量评价（IQA）设计的基于 Transformer 的架构。MS-SCANet 采用双分支结构，在多个尺度上处理图像，从而有效捕捉精细和粗略细节，改进了传统的单尺度方法。通过整合量身定制的空间和通道注意力机制，我们的模型能够突出重要特征，同时降低计算复杂度。MS-SCANet 的一个关键组成部分是其跨分支注意力机制，该机制增强了不同尺度间的特征融合，克服了以往方法的局限性。我们还提出了两种新的 一致性损失函数，即跨分支一致性损失和自适应池化一致性损失，它们在特征缩放过程中保持空间完整性，性能优于传统的线性和双线性技术。在 KonIQ-10k、LIVE、LIVE Challenge 和 CSIQ 等数据集上进行的大量评估表明，MS-SCANet 的表现始终优于最先进的方法，为与人类主观评分具有更强相关性的稳健框架提供了支持。**

索引词—**盲图像质量评价、空间注意力、通道注意力、跨分支注意力、自适应平均池化。**

I. 引言

图像质量评价（IQA）在众多应用中对于分析和提升图像的感知质量发挥着至关重要的作用，这些应用包括图像压缩、复原和增强。SSIM 和 PSNR 等传统技术难以预测并泛化到各种失真 [1], [2]。在过去十年中，卷积神经网络（CNN）凭借其提取视觉感知 [3]关键特征的卓越能力，已广泛应用于图像质量评价领域。尽管卷积神经网络具有独特优势，但由于主要关注局部特征，它们在识别全局上下文方面受到限制 [4]。

Transformer 近年来在图像质量评价中展现出尤为出色的效果，这得益于其通过自注意力技术捕捉长距离依赖关系的能力。 [5]视觉 Transformer（ViT） [6] 通过将图像块之间的关系建模为序列数据，并利用自注意力取得了显著成果，展示了这一方法。然而，ViT 的自注意力机制会带来巨大的计算开销，且该开销会随着输入尺寸的增大而呈二次方增长。为解决这些问题，人们引入了 Swin Transformer [7]等混合模型，以优化计算效率并捕捉全局上下文。尽管具备这些优势，这些模型在复杂场景下仍表现欠佳，因为其层次结构可能无法捕捉不同尺度上的所有特征。

CrossViT [8] 通过采用独立分支处理不同尺寸的图像块，并使用交叉注意力机制融合这些图像块，从而提升特征表示能力。MAMIQA [9] 和双分支 ViT [10],等多尺度方法展示了在图像质量评价中整合多尺度特征的优势。然而，TOPIQ [11], 等一些技术倾向于独立处理多尺度特征，或采用自底向上的方式，这可能会限制层级特征的有效整合。另一方面，SCA-Net [12], 将通道注意力机制与空间注意力机制相结合，用于医学图像分割。同样，双重视觉 Transformer（DaViT） [13]也运用这些技术来改进图像分类。然而，这些模型日益增加的复杂度导致计算需求更高，尤其是在处理高分辨率图像时。

为了解决以往方法的不足，我们提出了用于盲图像质量评价的多尺度空间通道注意力网络（MS-SCANet）。我们的模型将空间和通道注意力机制融入多尺度变换器，从而增强其捕捉和突出不同尺度关键特征的能力，包括局部和全局结构。与现有用于分类令牌和图像块标记的跨注意力方法不同，我们创新设计的跨分支注意力机制旨在整合多个尺度下不同分支的特征。我们还为图像质量评价提出了两种新颖的一致性损失函数：跨分支一致性损失，旨在确保不同尺度间的特征整合保持一致；以及自适应池化一致性损失，用于在下采样过程中保留特征完整性。通过整合这些改进，我们的方法提升了图像质量评价的可靠性，并通过在图像质量评估基准上开展的全面评估验证了这一点。MS-SCANet 的实现可在以下地址免费获取

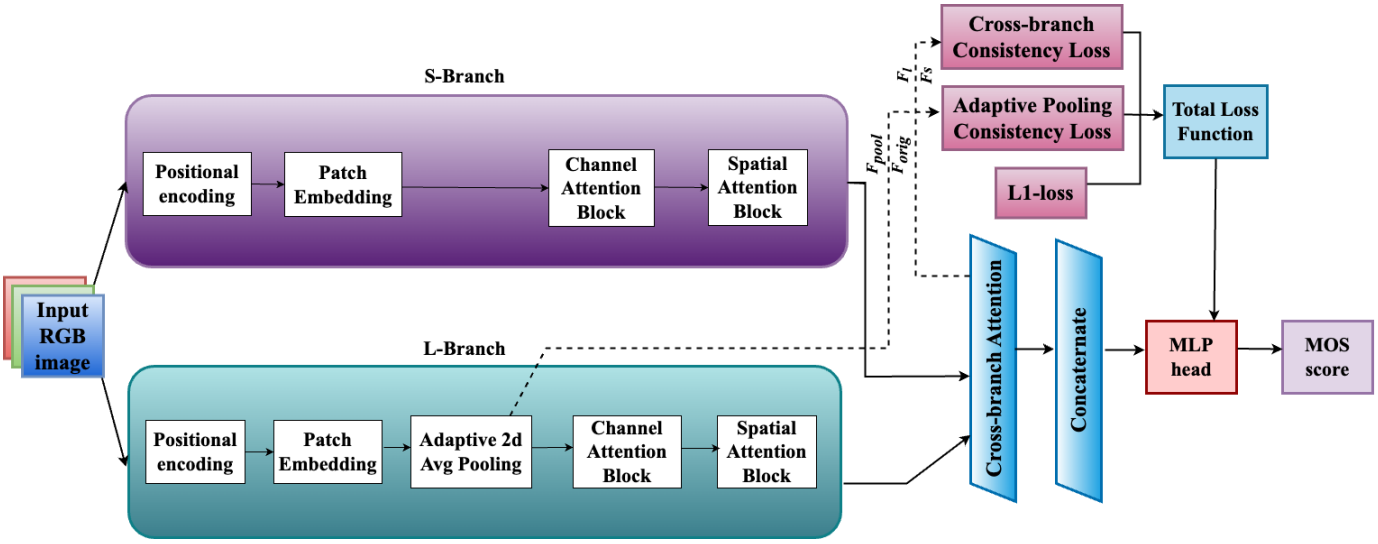


图 1：所提出的多尺度 SCANet（MS-SCANet）框架架构

https://github.com/mithila442/MS-SCANet.

II. 所提出的模型

图 1 展示了 MS-SCANet 的结构，其最终层用于预测平均意见分（MOS）。MS-SCANet 采用双分支架构处理多个尺度的数据，这对于确定感知图像质量至关重要。每个分支都将输入划分为不同尺度的图像块，并将其映射到三维空间。位置编码用于编码空间数据。

输入被划分为非重叠窗口，并对每个片段应用自注意力 [13]。该方法通过关注局部区域而非在全局范围内应用注意力，降低了计算复杂度。在每个窗口 W内，注意力的计算方式如下：

$$\text{Att}_W(\mathbf{Q}_W, \mathbf{K}_W, \mathbf{V}_W) = \text{softmax}\left(\frac{\mathbf{Q}_W \mathbf{K}_W^T}{\sqrt{d_k}}\right)\mathbf{V}_W, \quad (1)$$

窗口 W 的查询、键和值矩阵分别为 $\mathbf{Q}_W$、$\mathbf{K}_W$和 $\mathbf{V}_W$，键向量的维度为 d_k。注意力权重通过对 $\mathbf{Q}_W$ 和 $\mathbf{K}_W^T$的缩放点积执行 softmax 操作来计算。

利用压缩-激励网络 [14], 进行通道注意力，可突出关键特征通道。该过程通过全局平均池化和卷积层生成激励图，对各通道的特征重新加权，随后利用该激励图调制输入特征图。首先，压缩后的特征图$\mathbf{F}_{\text{sq}}$ 按如下方式计算：

$$\mathbf{F}_{\text{sq}} = \text{Conv2D}_1(\text{GAP}(\mathbf{F})), \quad (2)$$

其中$\mathbf{F}$ 是输入特征图，GAP 表示全局平均池化操作，而 Conv2D_1 是一个 1×1 卷积层。随后计算激励图$\mathbf{F}_{\text{ex}}$ ：

$$\mathbf{F}_{\text{ex}} = \sigma(\text{Conv2D}_2(\text{ReLU}(\mathbf{F}_{\text{sq}}))), \quad (3)$$

其中，Conv2D_2 是另一个 1×1 卷积层。ReLU 是修正线性单元激活函数。 σ 是 sigmoid 激活函数。最后，通道注意力机制$\mathbf{F}_{\text{ch}}$ 的输出为：

$$\mathbf{F}_{\text{ch}} = \mathbf{F} \cdot \mathbf{F}_{\text{ex}}. \quad (4)$$

为高效融合来自不同图像块尺寸分支的特征，我们提出了一种受 CrossViT [8] 启发的跨分支注意力机制。与专注于整合 CLS 标记和图像块标记以进行图像分类的 CrossViT 不同，我们的方法支持不同尺度图像块标记之间的直接交互。这一概念专为无参考图像质量评估（NR-IQA）设计，因为在多个尺度上捕捉细粒度和粗粒度失真至关重要。该机制定义如下：

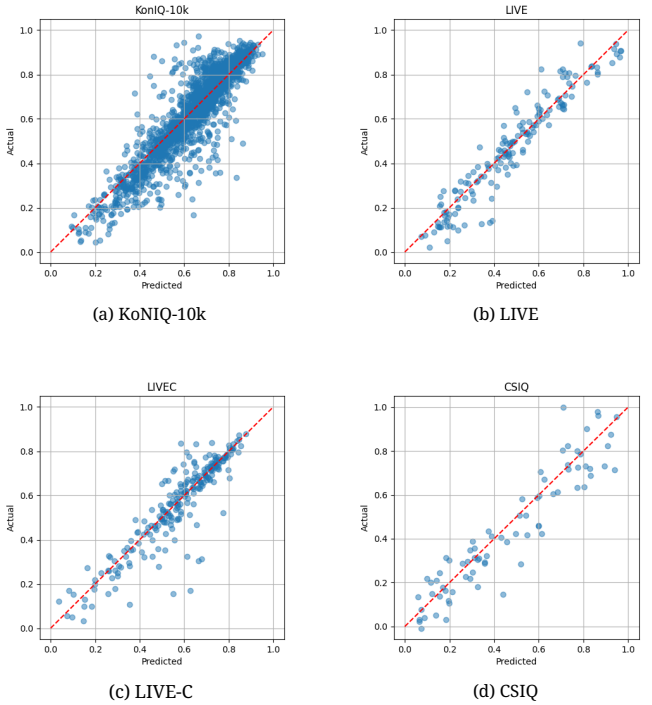
$$\mathbf{F}_{\text{cross}} = \text{softmax}\left(\frac{\mathbf{Q}_s \mathbf{K}_l^T}{\sqrt{d_k}}\right)\mathbf{V}_l + \text{softmax}\left(\frac{\mathbf{Q}_l \mathbf{K}_s^T}{\sqrt{d_k}}\right)\mathbf{V}_s, \quad (5)$$

其中，$\mathbf{Q}_s$、$\mathbf{K}_s$ 和 $\mathbf{V}_s$ 是由短分支特征得到的查询、键和值矩阵，而 $\mathbf{Q}_l$、$\mathbf{K}_l$ 和 $\mathbf{V}_l$ 是由长分支特征得到的对应矩阵。softmax 操作生成交叉注意力权重，使一个分支能够关注另一个分支的特征，从而成功实现跨分支特征整合。

MS-SCANet 将其计算复杂度降低至 $O(N_w^2 \cdot d)$，采用了基于窗口的注意力、更少且尺寸更大的图像块（16x16 和 32x32），以及更小的嵌入维度（256，相较于 768）。相比之下，视觉 Transformer 的全局注意力以 $O(N^2 \cdot d)$ 的速率呈二次增长，其中 N 是图像块总数。因此，MS-SCANet 的效率更高，每个令牌约需 14.7M 浮点运算次数，相比之下，SwinT 为 71.8M 浮点运算次数/令牌，TRIQ 为 92.9M 浮点运算次数/令牌，视觉 Transformer 为 185.7M 浮点运算次数/令牌。基于窗口的注意力不仅限制了交互，还提高了效率，超越了 SwinT 和 TRIQ 等模型；尽管这些模型针对特定任务进行了优化，但其计算成本仍然更高。

表 I：我们的方法与最先进的 NR-IQA 算法在基准数据集上的皮尔逊线性相关系数和斯皮尔曼秩序相关系数比较，最高、次高和第三高的结果分别以红色、橙色和蓝色标出。

方法	KonIQ-10k		LIVE		LIVE-C		CSIQ		平均值	
	皮尔逊线性相关系数	斯皮尔曼秩序相关系数	皮尔逊线性相关系数	斯皮尔曼秩序相关系数	皮尔逊线性相关系数	斯皮尔曼秩序相关系数	皮尔逊线性相关系数	斯皮尔曼秩序相关系数	皮尔逊线性相关系数	斯皮尔曼秩序相关系数
DIIVINE [15]	0.558	0.546	0.908	0.892	0.591	0.588	0.776	0.804	0.708	0.708
BRISQUE [16]	0.685	0.681	0.940	0.929	0.629	0.629	0.748	0.812	0.751	0.763
MEON [17]	0.628	0.611	0.955	0.951	0.710	0.697	0.942	0.922	0.809	0.795
DBCNN [18]	0.884	0.875	0.971	0.969	0.869	0.869	0.959	0.946	0.921	0.914
TRIQ [19]	0.903	0.892	0.965	0.949	0.861	0.845	0.838	0.828	0.892	0.879
HyperIQA [20]	0.885	0.872	0.966	0.962	0.842	0.844	0.942	0.923	0.909	0.900
TreS [21]	0.918	0.915	0.968	0.969	0.877	0.846	0.942	0.922	0.926	0.913
MAMIQA [9]	0.937	0.926	0.981	0.981	0.895	0.874	0.972	0.962	0.946	0.936
Ours	0.903	0.909	0.968	0.964	0.903	0.895	0.945	0.925	0.928	0.923



(a) KoNIQ-10k (b) LIVE

(c) LIVE-C (d) CSIQ

图2：所提出的 MS-SCANet 方法在基准数据集上的散点图和线性回归拟合线。

III. 所提出的损失函数

为促进不同分支之间的协同特征学习，我们提出了两种一致性损失函数：跨分支一致性损失和自适应池化一致性损失。跨分支一致性损失（CB损失）的主要目标是通过最小化分支特征之间的均方误差（MSE）来保持分支间的特征一致性，其定义如下：

$$\mathcal{L}_{CB} = \alpha \cdot \text{MSE}(\mathbf{F}_s, \mathbf{F}_l), \tag{6}$$

其中，$\mathbf{F}_s$ 和$\mathbf{F}_l$ 分别表示来自短分支和长分支的特征，α 则作为加权系数。

自适应池化一致性损失（AP损失）旨在统一不同尺度下的特征图，确保下采样阶段的特征整合保持一致。与线性变换和双线性插值方法不同，[22],这些方法在重缩放时可能引入伪影并扭曲特征空间；自适应池化则通过根据特征图内容调整池化过程，在保持特征质量的同时维护其结构完整性。该损失函数定义如下：

$$\mathcal{L}_{AP} = \beta \cdot \text{MSE}(\mathbf{F}_{orig}, \mathbf{F}_{pool}), \tag{7}$$

其中，$\mathbf{F}_{orig}$ 和$\mathbf{F}_{pool}$ 分别表示池化前后的特征，β 为相应的加权系数。AP损失最初用于质量评价，在目标检测中保留了空间关系；而CB损失则增强了语义分割的尺度均匀性。

L1损失函数也称为平均绝对误差（MAE），用于计算预测值与真值之间的绝对差异。它强调稀疏性，且与L2损失等其他损失函数相比，不易受离群值影响。通过最小化 L1损失，模型旨在减小预期图像质量分数与实际图像质量分数之间的绝对误差，从而获得更稳定、可预测的性能。总训练损失由核心 L1损失和一致性损失函数共同构成。

$$\mathcal{L}_{total} = \mathcal{L}_{L1} + \mathcal{L}_{CB} + \mathcal{L}_{AP}. \tag{8}$$

我们使用 PyTorch 实现该模型，并在 NVIDIA RTX 4000 Ada GPU 上进行训练，初始学习率为 1×10^{-4}。训练采用余弦退火调度，批大小为 16，共训练 600 轮次，并据此调整学习率；一致性损失权重 α 和 β 均设为 0.5。

IV. 实验结果

我们在四个广受欢迎的基准数据集上评估了 MS-SCANet 模型的性能：KonIQ-10k [24], LIVE [25], LIVE Challenge [26], 和 CSIQ [27]。KonIQ-10k 包含 10,073 张图像，存在噪声、模糊、压缩伪影和块效应等失真。LIVE 数据集包含 779 张图像，具有 JPEG2000、JPEG、高斯噪声和模糊等视觉伪影。LIVE Challenge 数据集包含 1,162 张图像，这些图像来自用户生成内容，带有真实世界失真。最后，CSIQ 数据集包含 866 张具有各种合成失真的图像，包括加性噪声、JPEG压缩等。评估遵循

表 II：注意力机制消融实验结果。最佳结果以粗体显示。

配置	LIVE		KonIQ-10k		LIVE-C		CSIQ	
	皮尔逊线性相关系数	斯皮尔曼秩序相关系数	皮尔逊线性相关系数	斯皮尔曼秩序相关系数	皮尔逊线性相关系数	斯皮尔曼秩序相关系数	皮尔逊线性相关系数	斯皮尔曼秩序相关系数
单分支 双重注意力	0.939	0.935	0.887	0.875	0.865	0.854	0.922	0.910
单分支空间注意力	0.955	0.950	0.846	0.856	0.873	0.862	0.929	0.918
单分支通道注意力	0.951	0.948	0.844	0.844	0.869	0.860	0.927	0.916
多分支双重注意力（所提出的方法）	**0.968**	**0.964**	**0.903**	**0.909**	**0.903**	**0.895**	**0.937**	**0.925**

表 III：不同损失函数的消融实验结果。最佳结果以粗体显示。

配置	LIVE		KonIQ-10k		LIVE-C		CSIQ	
	皮尔逊线性相关系数	斯皮尔曼秩序相关系数	皮尔逊线性相关系数	斯皮尔曼秩序相关系数	皮尔逊线性相关系数	斯皮尔曼秩序相关系数	皮尔逊线性相关系数	斯皮尔曼秩序相关系数
仅使用 L1损失（基线）	0.968	0.964	0.903	0.909	0.903	0.895	0.937	0.925
L1 + CB损失	0.971	0.969	0.910	0.916	0.907	0.902	0.942	0.930
L1 + AP损失	0.969	0.969	0.915	0.917	0.908	0.899	0.940	0.928
L1 + CB损失 + AP损失（完整模型）	**0.972**	**0.973**	**0.921**	**0.923**	**0.909**	**0.907**	**0.945**	**0.933**

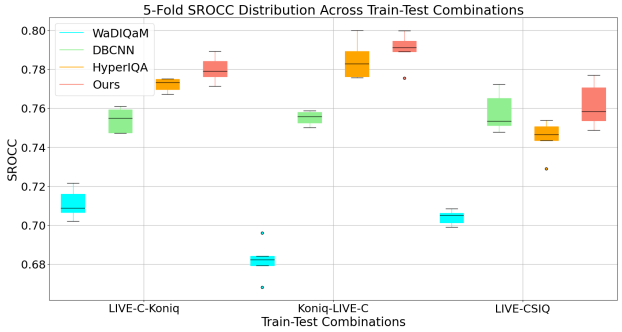


图 3：不同模型在跨数据集组合上的斯皮尔曼秩序相关系数箱线图：WaDIQaM [23],DBCNN [18], HyperIQA [20], 和 MS-SCANet 在三种训练-测试组合上的表现。x轴上的标签“数据集1-数据集2”表示在数据集1上训练，在数据集2上测试。

标准图像质量评价协议，使用皮尔逊线性相关系数（PLCC）和斯皮尔曼秩序相关系数（SROCC）来衡量模型准确性。

我们将 MS-SCANet 与多种先进的无参考图像质量评价模型进行了比较。所比较的方法分为以下几类：自然场景统计方法中的 DIVINE 和 BRISQUE；基于卷积神经网络的方法中的 MEON、DBCNN、HyperIQA 和 MAMIQA；以及基于 Transformer 的方法中的 TRIQ 和 TreS。如表 I 所示，MS-SCANet 在皮尔逊线性相关系数和斯皮尔曼秩序相关系数方面始终位列前三（最佳结果分别以红色、橙色和蓝色表示）。实证分析表明， 6×6 窗口大小与 256 的嵌入维度能够在皮尔逊线性相关系数和斯皮尔曼秩序相关系数之间实现均衡。研究结果显示，该模型在极具挑战性的 LIVE挑战数据集上表现稳健，并取得了最高相关性。图 2 展示了 MS-SCANet 在所有数据集上的散点图。这些结果表明，MS-SCANet 能够处理噪声、模糊、压缩伪影和块效应等合成和自然视觉失真。该模型的双分支设计和跨分支注意力机制使其能够处理细粒度和粗粒度的图像特征，从而检测局部和全局伪影。为评估 MS-SCANet 在多个数据集之间的泛化能力，我们采用了五折跨数据集验证流程，即在一个数据集上训练，并在另一个数据集上测试。图 3 展示了不同训练-测试场景下斯皮尔曼秩序相关系数值的分布箱线图，其中x轴上的标签表示训练-测试组合。MS-SCANet 展现出优异的性能和较低的变异性，尤其是在 KonIQ-10k 上训练并在 LIVE-C 上测试时，这表明其对各种失真具有较强的鲁棒性。

我们进行了消融分析，以确定 MS-SCANet 的各个组件对整体性能的贡献。表 II 比较了采用空间注意力、通道注意力和双重注意力的模型与所提出的双重注意力 MS-SCANet。多尺度双重注意力配置始终优于其他配置，凸显了跨尺度整合注意力机制以捕获局部和全局特征的重要性。此外，我们分别分析了各一致性损失函数单独使用及组合使用时的影响。表 III 显示，跨分支一致性损失（CB-Loss）与自适应池化一致性损失（AP-Loss）的组合，通过在下采样过程中改善特征整合，显著提升了性能。

V. 结论

我们提出了 MS-SCANet，这是一种用于盲图像质量评估的多尺度双重注意力 Transformer 模型，采用双分支设计来捕获细粒度和粗粒度特征，并通过空间与通道注意力机制加以增强。跨分支注意力机制能够有效整合不同尺度的特征表示。此外，新增的两种一致性损失策略——跨分支一致性和自适应池化一致性——提升了特征学习与多尺度融合效果。在四个基准数据集上的大量实验表明，MS-SCANet 的性能优于当前方法；跨数据集评估进一步证实了其较强的泛化能力。未来，我们计划进一步提升计算效率，并探索所提出模型在视频质量评估任务中的潜在适用性。

参考文献

[1] Z. Wang、A. C. Bovik、H. R. Sheikh 和 E. P. Simoncelli，《图像质量评估：从误差可见性到结构相似性》，IEEE图像处理汇刊，第 13 卷，第 4 期，第 600–612 页，2004 年。[2] A. Mittal、A. K. Moorthy 和 A. C. Bovik，《构建“完全盲”的图像质量分析器》，IEEE信号处理快报，第 20 卷，第 3 期，第 209–212 页，2013 年。[3] L. Kang、P. Ye、Y. Li 和 D. Doermann，《用于无参考图像质量评价的卷积神经网络》，载于 CVPR 论文集，2014 年，第 1733–1740 页。[4] R. Zhang、P. Isola、A. A. Efros、E. Shechtman 和 O. Wang，《深层特征作为感知度量的非凡有效性》，载于 CVPR 论文集，2018 年，第 586–595 页。[5] A. Vaswani、N. Shazeer、N. Parmar、J. Uszkoreit、L. Jones、A. N. Gomez、L. Kaiser 和 I. Polosukhin，《你所需要的只是注意力》，载于神经信息处理系统进展，2017 年，第 5998–6008 页。[6] A. Dosovitskiy、L. Beyer、A. Kolesnikov、D. Weissenborn、X. Zhai、T. Unterthiner、M. Dehghani、M. Minderer、G. Heigold、S. Gelly、J. Uszkoreit 和 N. Houlsby，《一幅图像胜过 16×16 个词：用于大规模图像识别的 Transformer》，载于 ICLR 论文集，2021 年。[7] Z. Liu、Y. Lin、Y. Cao、H. Hu、Y. Wei、Z. Zhang、S. Lin 和 B. Guo，《Swin Transformer：使用移位窗口的层次化视觉 Transformer》，载于 ICCV 论文集，2021 年，第 10 012–10 022 页。[8] C.-F. Chen、Q. Fan、R. Panda、S. Sadiq、S. Saxena 和 S. N. Lim，《CrossViT：用于图像分类的交叉注意力多尺度视觉 Transformer 》，载于 ICCV 论文集，2021 年，第 357–366 页。[9] L. Yu 、Y. Yang、X. Guan、K. Gu 和 Z. Pan，《MAMIQA：基于结合自然场景统计的多尺度注意力机制的无参考图像质量评价》，IEEE信号处理快报，第 30 卷，第 588–592 页，2023 年。[10] S.-H. Lee 和 S.-W. Kim，《用于盲图像质量评估的双分支视觉 Transformer》，《视觉通信与图像表示杂志》，第 94 卷，文章编号 103850，2023 年。[11] C. Chen、Y. Liu 和 H. Huang，《TOPIQ：从语义到失真的自顶向下图像质量评价方法》，IEEE图像处理汇刊，2024 年。[12] A. G. Roy、N. Navab 和 C. Wachinger，《SCA-Net：用于医学图像分割的空间与通道注意力网络》，IEEE医学成像汇刊，2018 年。[13] H. Ding、J. Xia、A. Yuille 和 L. Shao，《DaViT：双重视觉 Transformer》，载于 NeurIPS 论文集，2022 年。[14] J. Hu、L. Shen 和 G. Sun，《挤压激励网络》，载于 CVPR 论文集，2018 年，第 7132–7141 页。[15] M. A. Saad、A. C. Bovik 和 C. Charrier，《盲图像质量评估：DCT域中的自然场景统计方法》，IEEE图像处理汇刊，第 21 卷，第 8 期，第 3339–3352 页，2012 年。[16] A. Mittal、R. Soundararajan 和 A. C. Bovik，《空间域中的无参考图像质量评价》，IEEE 图像处理汇刊，第 21 卷，第 12 期，第 4695–4708 页，2012 年。[17] K. Ma 、 W. Liu 、 K. Wang、D. Zeng、Z. Wang、H. Wang 和 W. Lin，《使用深度神经网络进行端到端盲图像质量评估》，IEEE图像处理汇刊，第 27 卷，第 3 期，第 1202–1213 页，2017 年。[18] W. Zhang、G. Zhai、X. Min、K. Gu 和 X. Yang，《使用深度双线性卷积神经网络进行盲图像质量评估》，IEEE电路与系统视频技术汇刊，第 30 卷，第 1 期。[19] J. You 和 J. Korhonen，《用于图像质量评估的 Transformer》，IEEE图像处理汇刊，第 31 卷，第 477–490 页，2021 年。[20] S. Su、G. Liu、X. Zeng、X. Zhu、S. Wang、G. Zhai 和 X. Yang，《由自适应超网络引导的野外图像盲质量评估》，载于 CVPR 论文集，2020 年，第 3667–3676 页。[21] S. A. Golestaneh 和 K. M. Kitani，《通过 Transformer、相对排序和自一致性进行无参考图像质量评价》，载于 WACV 论文集，2022 年，第 1220–1230 页。[22] K. He、X. Zhang、S. Ren 和 J. Sun，《用于视觉识别的深度卷积网络中的空间金字塔池化》，IEEE模式分析与机器智能汇刊，第 37 卷，第 9 期，第 1904–1916 页，2015 年。[23] S. Bosse、D. Maniry、K.-R. M¨uller、T. Wiegand 和 W. Samek，《用于无参考和全参考图像质量评价的深度神经网络》，IEEE图像处理汇刊，第 27 卷，第 1 期，第 206–219 页，2018 年。[24] H. Lin、J. Fu、X. Wang、Y. Jiang 和 J. Liu，《KonIQ-10k：用于盲图像质量评估深度学习的生态有效数据库》，IEEE图像处理汇刊，第 28 卷，第 1 期，第 1048–1061 页，2019 年。[25] H. R. Sheikh、M. F. Sabir 和 A. C. Bovik，《近期全参考图像质量评价算法的统计评估》，*IEEE*图像处理汇刊，第 15 卷，第 11 期，第 3440–3451 页，2006 年。[26] D. Ghadiyaram 和 A. C. Bovik，《真实环境图像质量挑战数据库》，*IEEE*信号处理快报，第 21 卷，第 3 期，第 145–149 页，2016 年。[27] D. M. Chandler，《最明显失真：全参考图像质量评价及策略的作用》，《电子成像杂志》，第 19 卷，第 1 期，文章编号 011006，2010 年。