arXiv:2603.00938v1 [cs.CV] 1 Mar 2026

超越8位：HDR-UGC视频的主观与客观质量评估

Shreshth Saini[1]、**Bowen Chen**[1]、**NeilBirkbeck**[2]、**YilinWang**[2]、**Balu Adsumilli**[2]、**AlanC.Bovik**[1†]
图像与视频工程实验室（LIVE），德克萨斯大学奥斯汀分校。 [2]Google/YouTube

高动态范围（HDR）用户生成内容（UGC）视频正在社交平台上迅速普及，但大多数感知视频质量评估（VQA）系统仍针对标准动态范围（SDR）进行设计。HDR更高的位深、宽广的色域和更大的亮度范围，会暴露出近黑压暗、高光裁剪、色带和曝光闪烁等失真，从而放大UGC伪影，并对SDR模型提出挑战。为推动该领域发展，我们整理了**Beyond8Bits**，这是一个大规模主观数据集，包含来自6.5K个来源的 ∼44K个视频以及 >1.5M条众包评分，涵盖多样化的场景、拍摄条件和压缩设置。我们进一步提出**HDR-Q**，这是首个用于HDR-UGC视频质量评估的多模态大语言模型（MLLM）。我们提出：（i）一种新型HDR 感知视觉编码器，用于生成HDR 敏感嵌入；（ii）HDR感知策略优化（HAPO），一种将推理锚定于HDR线索的强化学习微调框架。HAPO通过HDR–SDR对比KL增强GRPO，促使模型依赖HDR输入，并通过高斯加权回归奖励实现细粒度的MOS校准。在 **Beyond8Bits** 和公开HDR- VQA基准测试中，**HDR-Q** 均取得了当前最先进的性能。

日期： 2026年3月3日
页面： https://shreshthsaini.github.io/Beyond8Bits通
信作者： Shreshth Saini，邮箱：saini.2@utexas.edu

1 引言

YouTube、TikTok 和 Instagram 等平台上的用户生成内容（UGC）爆炸式增长，改变了数字媒体生态系统（99Firms，2024；Omnicore，2024；Mohsin，2020）。与此同时，高动态范围（HDR）视频已成为主流，相较于标准动态范围（SDR），它具有更高的位深、更宽的色域和更大的亮度范围（Chen等人，2025a；Shang 等人，2023；Saini 等，2024，2025）。这些特性提升了感知真实感，但也凸显了在 SDR 中不太明显的失真，例如近黑压暗、高光裁剪、色带和曝光闪烁；而这些问题往往还会与 UGC 中常见的压缩或拍摄伪影相互叠加（图 2）。因此，评估 HDR-UGC 内容的感知质量仍是一项尚未解决且研究不足的挑战。现有的视频质量评估（VQA）模型在这一领域表现不佳。基于专业制作的HDR数据集（Shang 等人，2023；Chen等人，2025a）或 SDR-UGC视频（Lu 等人，2024；Zhang等人，2023）训练的方法，无法泛化到真实 HDR-UGC 中多样的拍摄条件、设备差异和非受控失真。此外，当前的 HDR主观数据集规模较小、范围有限，主要关注合成失真或经过筛选的专业内容（Wang 等人，2024；Ebenezer 等人，2024a）。缺乏大规模、真实世界的 HDR-UGC标注，是开发能够与人类感知判断相契合的模型的一大障碍。

与此同时，多模态大语言模型（MLLMs）已成为融合感知与语言的强大推理系统，在可解释图像和视频质量评估方面展现出良好前景（Wu 等人，2023d、2024a、b；You 等人，2025、2024b；Wu 等人，2025；Li 等人，2025）。然而，将其直接应用于 HDR-UGC 视频质量评估面临着若干关键障碍：(i) 标准视觉编码器是在标准动态范围数据上预训练的，无法捕捉 HDR 特定线索；(ii) 在下一词元预测范式下，准确预测连续平均意见分仍颇具挑战，因为离散等级或回归头方法（Wu 等人，2025；You 等人，2025；Wu 等人，2023d；Duan 等人，2025；Zhu 等人，2024）缺乏细粒度校准；(iii) 若缺少明确的

†LIVE-YouTube Beyond8Bits 数据集。

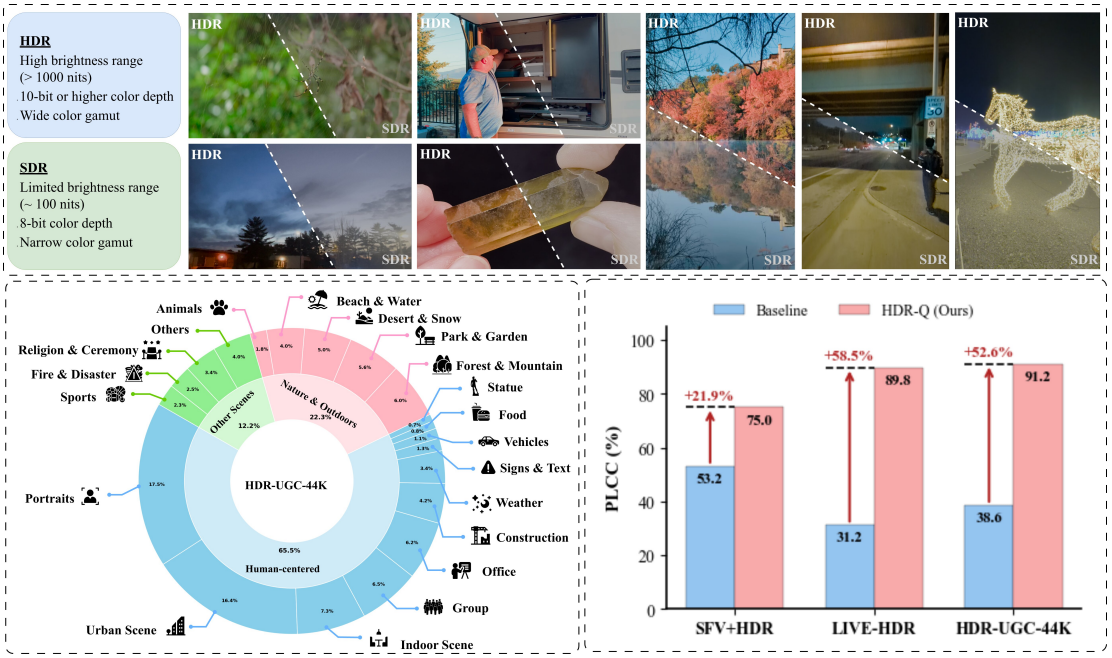


图1 我们的数据集与性能评估概览。 上方：高动态范围与标准动态范围帧的示例对比，展示了不同场景中亮度范围、色彩深度和视觉细节的差异。左下：**Beyond8Bits** 数据集中的视频类别分布，涵盖以人为中心的内容、自然与户外场景以及其他各类真实场景。右下：我们提出的HDR-Q模型与基线方法在三个数据集上的性能对比，其中HDR-Q在皮尔逊线性相关系数上取得了显著提升。

激励机制方面，相关策略往往忽视 HDR输入并依赖文本先验，这是一种模态忽视（Zheng等人，2025；Chen等人，2025c；Wang等人，2023年）。

为应对这些挑战，我们推出了**Beyond8Bits**，首个大规模众包主观HDR-UGC质量数据集，包含 ~44K个视频，来自6,861个不同来源，并拥有超过150万条人工评分。该数据集为训练和评估能够反映真实世界HDR感知现象的模型提供了必要基础。在此基础上，我们提出了**HDR-Q**，首个专为HDR-UGC质量评估设计的多模态大语言模型。**HDR-Q**集成了两个新颖组件：（i）HDR 感知视觉编码器，在保持语义对齐的同时学习HDR敏感表征；（ii）HDR感知策略优化（HAPO），一种强化学习框架，通过HDR–SDR对比KL项强制HDR grounding，稳定熵，并借助熵加权优势细化词元级信用分配。高斯回归奖励进一步实现细粒度MOS校准，而组级自奖励则提升推理一致性。我们的贡献如下：

- 我们推出了**Beyond8Bits**，这是规模最大的主观高动态范围用户生成内容质量数据集。
- 我们提出了**HDR-Q**，这是一种新颖的基于MLLM的视觉问答模型，将全新的HDR感知视觉编码器与我们提出的HAPO策略相结合；HAPO是一种专为高动态范围条件下感知推理设计的强化学习微调范式。
- 在**Beyond8Bits**和公开HDR基准上的大量实验表明，**HDR-Q** 在MOS预测方面达到了当前最优水平，并生成简洁且基于HDR的理由，从而为HDR感知视觉质量评估研究开辟了新方向。

2 相关工作

2.1 高动态范围视频质量评估：数据集与模型

早期的VQA数据集，例如CVD2014（Nuutinen等人，2016年）、LIVE-VQA（Seshadrinathan等人，2010年）、LIVE-VQC（Sinno和Bovik，2019年）、LSVQ（Ying等人，2021年）、MDVQA（Zhang等人，2023年）以及Maxwell（Wu等人，

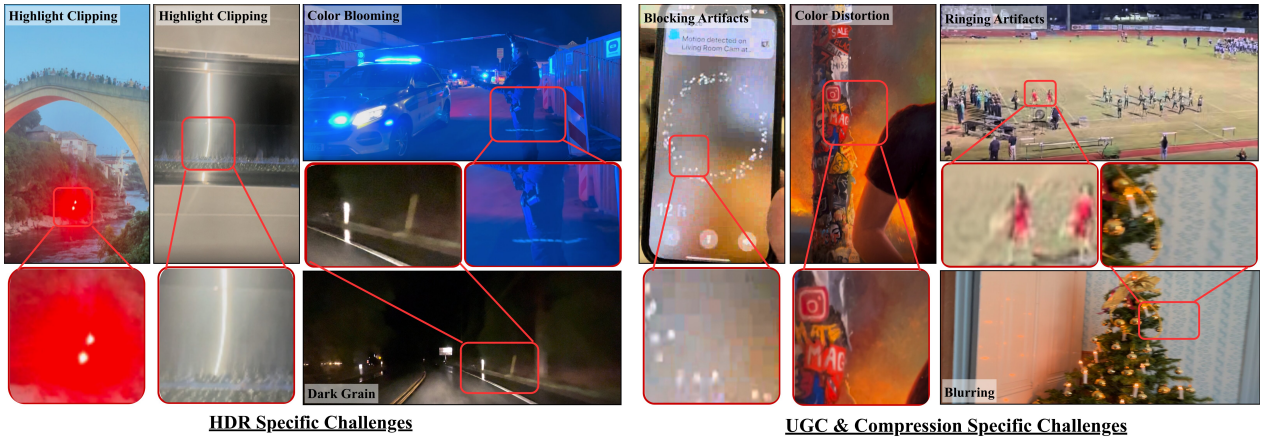


图2 典型挑战 在HDR-UGC视频中，包括HDR特定问题（例如 .g.，高光裁剪、色彩光晕、暗部颗粒）以及压缩相关伪影（例如，色彩失真、块效应、振铃效应、模糊）。

2023b）既支持手工设计的VQA模型（Mittal等人，2013，2012；Saad等人，2014；Korhonen，2019；Mittal等人，2016；Ebenezer等人，2021），也支持深度VQA模型（Li 等人，2019；Wu 等人，2022a、b；Madhusudana等人，2022；He等人，2024；Wu 等，2022c），但这些模型仍面向SDR，由于亮度和色调映射存在根本差异，因此不适用于HDR。HDR特定主观数据集（Azimi等人，2021；Pan等人，2018；Baroncini等人，2016；Rerabek等人，2015；Athar等人，2019）弥补了这一空白，但其中许多数据集已经过时或受到限制。近期发布的LIVE-HDR（Shang 等人，2023）（包含310个已标注视频）和SFV+HDR（Wang 等人，2024）（包含2,000个片段，其中300个经过评分）为现代HDR算法提供了更可靠的基准测试。相应地，HDR-VQA模型也开始出现：全参考指标HDR-VQM（Narwaria等人，2015）、HDR-BVQM（Aamir等人，2021）和PU21（Mantiuk和Azimi，2021）采用感知亮度或感知均匀变换，但依赖参考视频，并且难以应对多样化的HDR失真。HDR-ChipQA（Ebenezer 等，2024b）和HIDRO-VQA（Saini 等，2024）等无参考方法，则通过非线性亮度映射或大规模无标注HDR数据扩展了ChipQA和CONTRIQUE（Madhusudana等人，2022）。尽管如此，现有数据集和模型仍无法捕捉HDR UGC的异质退化，因此有必要探索新的数据和建模策略。

2.2 基于多模态大语言模型的感知质量评估

近期，研究者开始探索将多模态大语言模型用于图像质量评估/视频质量评估。Q-Bench (Wu 等人, 2023c) 等基准测试揭示了多模态大语言模型与人类判断之间的巨大差距，推动了指令调优（Q-Instruct (Wu 等人, 2024a)）和描述性失真推理（DepictQA、DepictQA-Wild (You 等人, 2024b,a)）的发展。比较方法和基于排序的方法（Compare2Score (Zhu 等人, 2024)、VisualQuality-R1 (Wu 等人, 2025)）进一步提升了与人类判断的一致性；与此同时，Q-Align (Wu 等人, 2023d)、DeQA-Score (You 等人, 2025) 和 Q-Insight (Li 等人, 2025) 致力于提升可解释性、回归保真度和联合退化推理能力。视频领域的扩展包括 Q-Bench-Video (Zhang等人, 2025) 和 MVQA-68K (Pu 等人, 2025)，它们提供了大规模、多维标注和文本理由，用于训练视频感知的MLLM质量评估器。

3 数据集：Beyond8Bits

现有的高动态范围视觉问答数据集（Shang 等人，2023年；Wang等人，2024；Saini等人，2025；Ebenezer等人，2024a；Chen等人，2025b）在规模、多样性或动态范围方面存在局限，且主要关注专业制作的内容。相比之下，真实世界中的高动态范围用户生成视频（HDR-UGC）呈现出更加广泛的亮度、运动和压缩特性，且通常是在非受控条件下拍摄的。为弥合这一差距，我们推出了**Beyond8Bits**，这是迄今规模最大、最多样化的高动态范围视觉问答数据集，专为真实世界 HDR-UGC 质量评估而设计。

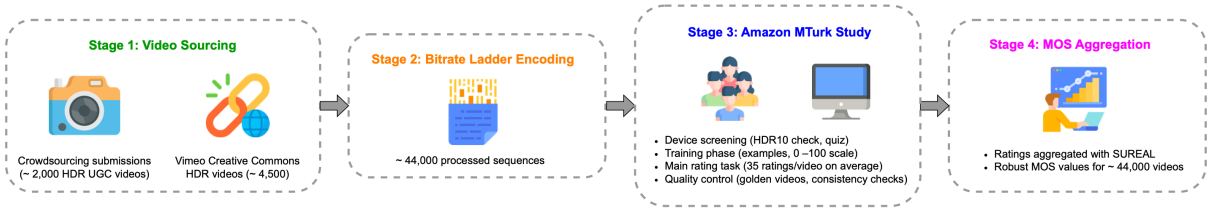


图3 Beyond8Bits 的构建流程。

3.1 数据收集与处理

我们从两个互补来源收集了 6,861 个独特的高动态范围源视频：（1）一项专门的众包活动，用户在研究知情同意的前提下，贡献了使用消费级设备（iPhone、Pixel、Galaxy等）拍摄的高动态范围片段，共 2,253 个视频；（2）来自 Vimeo、采用知识共享许可的公开高动态范围视频，共 4,608 个视频。这种组合确保了数据集涵盖以人物为中心的场景、自然场景和弱光场景，同时具有丰富的设备内和设备间差异。

我们对每个源视频的高动态范围元数据（PQ传递函数、10位HEVC、BT.2020色域）进行了验证，并筛除重复视频、静态帧和不适宜内容。视频片段被裁剪至最长10秒，并按照模拟真实流媒体条件的码率梯度进行转码（Google 支持，2024；苹果公司，2024），涵盖多种分辨率（1080p–360p）和码率（0.2–5 Mbps），详见附录D。所有版本均保留完整的高动态范围信号标记，总计处理了 ∼44,276个视频片段；条件允许时，横屏和竖屏视频保持数量相等。图1 展示了数据集的构成及示例内容的多样性。

3.2 主观质量研究

我们在亚马逊Mechanical Turk（AMT）上开展了一项大规模主观研究，这是首次达到这一规模的高动态范围大规模众包视觉问答研究。为确保显示保真度，我们仅允许使用经过验证、具备高动态范围显示能力的设备和浏览器的工作者参与。人工智能任务（HIT）的设计纳入了多项质量控制措施。每项HIT均以说明和资格测验开始，以检查参与者的高动态范围显示能力及其对任务的理解。参与者首先通过包含代表性示例的训练和校准阶段，随后按照ITU-R BT.500-14（国际电信联盟，2019年）指南，使用连续的0–100李克特量表对成批视频片段进行评分。为确保可靠性，我们嵌入了隐藏的质量控制视频（重复视频和在试点研究中确定了已知质量范围的金标准视频）。我们根据参与者在控制视频上的一致性检查结果、显示位深、网速、任务完成时间及所报告的观看条件，严格执行参与者筛选和淘汰标准。我们嵌入金标准视频和重复视频，以评估受试者内和受试者间的一致性。经过严格的质量控制后，共收集了超过150万条有效评分。每个视频平均获得 ∼35 条独立评分。这种大规模设计能够捕捉真实高动态范围观看条件下的感知差异。

3.3 平均意见分聚合

为将主观评分聚合为可靠的平均意见分，我们采用了主观可靠性（SUREAL）方法（Li 等人，2020年）。SUREAL 通过对个体受试者评分（ψ_j）进行建模，并考虑受试者偏差（S_{ij}）和不一致性（Δ_i），从而得出真实视频质量（ν_i）的最大似然估计（MLE）。该模型如下：

$$S_{ij} = \psi_j + \Delta_i + \nu_i X, \quad X \sim \mathcal{N}(0,1) \tag{1}$$

通过最大化对数似然来估计参数。所得平均意见分表现出较强的受试者间相关性（斯皮尔曼等级相关系数中位数为 0.90），证实了研究的可靠性。

Beyond8Bits 涵盖多种内容类别（人物、室内、室外、夜景、运动密集型场景）、不同的亮度分布以及广泛的平均意见分（10–95）。附录D中提供了关键统计数据，包括空间/时间复杂度以及与现有数据集的比较。**Beyond8Bits** 为现代高动态范围感知训练模型奠定了重要基础，例如 **HDR-Q**。

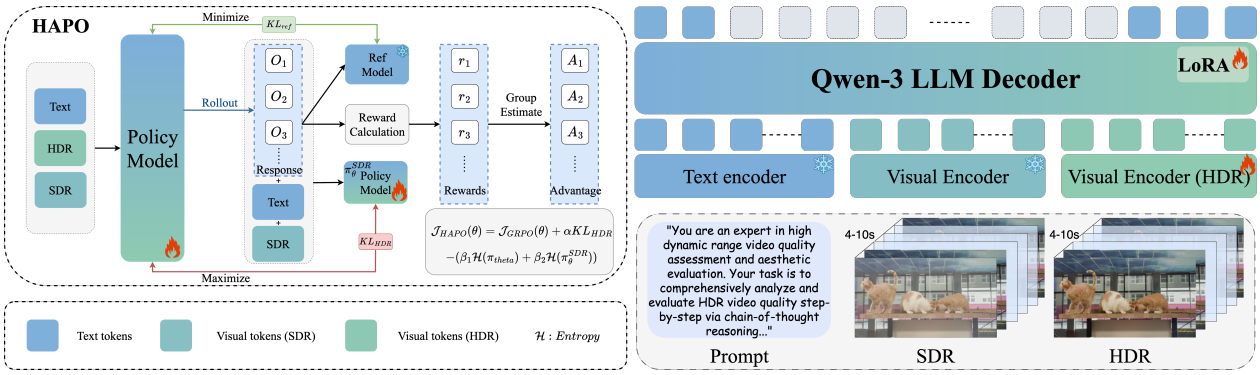


图 4 HDR-Q 结合 HAPO的概览。左：HAPO 对比高动态范围输入（文本 ＋标准动态范围 ＋高动态范围词元）与缺少高动态范围的路径（仅文本 ＋ 标准动态范围）的 rollout，并最大化二者的 KL 散度，以确保 HDR grounding；同时采用双熵正则化，防止奖励投机。组级奖励包括平均意见分/属性准确率、推理质量和自奖励。右：经过 LoRA 微调的大语言模型解码高动态范围感知推理；视觉输入来自标准编码器和我们的高动态范围感知适配器。

4 预备知识

大规模多模态强化学习需要稳定的优化方法，同时避免基于评论家的方法（如近端策略优化（Schulman等人，2017）所带来的高方差。组相对策略优化（GRPO）（Shao 等人，2024）通过在采样得到的响应组内对奖励进行归一化，实现了这一点；它无需训练价值网络，同时保留了样本效率。GRPO（Shao 等人，2024）已成为现代大语言模型和多模态大语言模型后训练流程中的关键组成部分（Yu等人，2025；Chu等人，2025），尤其是在无法直接进行奖励建模时。

形式化。给定一个多模态数据集 $\mathcal{D} = \{(q,\ I,\ a)\}$，其中包含输入提示 q、多模态输入 I 和目标答案 a，我们从先前策略 $\pi_{\theta_{\text{old}}}$中采样 K 个候选补全 $\{o_i\}_{i=1}^{K}$。每个补全都会获得一个标量奖励 R_i，其归一化的组相对优势计算如下：

$$\hat{A}_i = \frac{R_i - \mu_R}{\sigma_R + \epsilon}, \quad \mu_R = \frac{1}{K}\sum_{j=1}^{K} R_j, \\ \sigma_R = \sqrt{\frac{1}{K}\sum_{j=1}^{K}(R_j - \mu_R)^2}. \tag{2}$$

裁剪代理目标变为：

$$\mathcal{J}_{\text{GRPO}}(\theta) = \mathbb{E}_{(q,I)\sim\mathcal{D},\, o_i\sim\pi_{\theta_{\text{old}}}} \frac{1}{K}\sum_{i=1}^{K}\frac{1}{|o_i|}\sum_{t}\Big[\min\big(\rho_{i,t}\hat{A}_i,\ \text{clip}(\rho_{i,t}, 1-\epsilon, 1+\epsilon)\hat{A}_i\big) - \beta\, D_{\text{KL}}(\pi_\theta\|\pi_{\text{ref}})\Big]. \tag{3}$$

其中 $\rho_{i,\ t} = \pi_\theta(o_{i,\ t}|q,\ I,\ o_{i,\ <t})/\pi_{\theta_{\text{old}}}(o_{i,\ t}|q,\ I,\ o_{i,\ <t})$ 表示词元级重要性比率，而 π_{ref}是一个用于锚定稳定性的冻结参考策略。

局限性。尽管原始 GRPO 具有稳定性（Shao 等人，2024），但它缺少明确机制来确保学习到的策略将其行为锚定于输入模态的感知线索（Wang等人，2025）。在 HDR-UGC 视频质量评估这类以感知为主的任务中，这会导致模态忽视（Zheng等人，2025；Wang等人，2025），即策略虽然能够生成连贯性很高的文本，却忽略了 HDR视觉信息。此外，它对所有输出词元一视同仁，忽视了多模态推理任务中至关重要的词元级不确定性和推理结构问题。这些局限促使我们提出HDR感知策略优化（HAPO）（第 5.2节），在 GRPO（Shao 等人，2024）的基础上引入 HDR–SDR对比 grounding、双熵正则化和熵加权优势塑形。

5 方法：HDR-Q

我们提出 **HDR-Q**，这是一种多模态大语言模型（MLLM），专为 HDR 用户生成视频的感知质量评估而设计。该框架将 HDR 感知视觉编码器与强化学习（RL）目标 HDR感知策略优化（HAPO）相结合，后者能够显式地强化 HDR grounding，稳定学习过程以避免奖励投机，并提升推理保真度。如图 4所示，HDR-Q 融合了感知和推理两条路径：(i) HDR 感知编码器生成 HDR 敏感嵌入，以捕捉亮度极值和色彩体积保真度；(ii) HAPO 通过对比式、熵正则化的强化学习微调策略，使其依赖这些线索。

5.1 HDR 感知视觉编码器

令 $v = \{x_t\}_{t=1}^{T}$ 表示采用 PQ（BT.2020）的 10 位 HDR 视频。我们以全精度保留 HDR 信号，不进行色调压缩，以保留近黑结构、高光动态和广色域色彩关系。为进行对比式监督，通过确定性色调映射算子 $TM(\cdot)$（PQ→γ映射、量化和 BT.709 收缩）生成 SDR 对应版本 $v^{SDR} = \{TM(x_t)\}_{t=1}^{T}$。

我们对预训练的 SigLIP-2 编码器（Tschan- nen 等人，2025） $\mathcal{E}_\psi$ 使用 HDR 帧–描述对（x_t，c_t）进行适配，其中描述由 Qwen2.5-VL-72B（Bai 等人，2025）生成。目标是获得语义保持对齐、同时对 HDR 变化具有内在敏感性的嵌入。

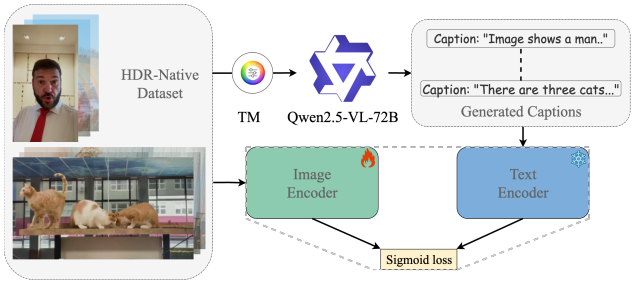


图 5 HDR 感知视觉编码器微调。我们使用由 Qwen2.5-VL-72B 生成描述的 HDR–SDR 帧–描述对，对 SigLIP-2（Tschannen 等人，2025）进行适配，以促进在感知上对齐的 HDR 嵌入。

双域监督。一个关键挑战是，通用描述 c_t 对于 x_t 和x_t^{SDR}同样成立，这可能导致表征坍塌，使高动态范围与标准动态范围嵌入重叠。为避免这一问题，我们引入双域监督。对于每个高动态范围帧 x_t，我们生成 x_t^{SDR}并施加对比式分离约束：高动态范围嵌入与其描述的距离必须小于标准动态范围嵌入与其描述的距离：

$$\mathcal{L}_{\text{contrast}} = \max\Big(0,\ \delta - D\big(\mathcal{E}_\psi(x_t),\ \mathcal{E}_\psi(c_t)\big) + D\big(\mathcal{E}_\psi(x_t^{SDR}),\ \mathcal{E}_\psi(c_t)\big)\Big). \tag{4}$$

其中，$D(\cdot,\ \cdot)$ 表示余弦距离，δ 是间隔。完整的编码器损失结合了对齐和高动态范围判别：

$$\mathcal{L}_{\text{enc}} = \mathcal{L}_{\text{Sigmoid}}(x_t, c_t) + \lambda_{\text{ctr}}\mathcal{L}_{\text{contrast}} \tag{5}$$

确保学习到的嵌入在保持语义忠实性的同时，也能感知高动态范围对比度和亮度线索（见图 5）。

5.2 高动态范围感知策略优化（HAPO）

虽然 GRPO（Shao 等人，2024）能够稳定多模态强化学习，但无法保证策略利用的是视觉线索，而非文本先验。HAPO 在 GRPO（Shao 等人，2024）的基础上扩展出三个高动态范围专属组件，以明确强制实现模态接地（见附录 E）：

(i) 高动态范围–标准动态范围对比*KL*。为防止模态忽视（Zheng等人，2025），我们对比包含和不包含HDR词元时的推演结果：

$$\mathcal{K}_{\text{HDR}}(\theta) = D_{\text{KL}}\big(\pi_\theta^{HDR} \,\|\, \pi_\theta^{SDR}\big) \tag{6}$$

其中， π_θ^{HDR}和 π_θ^{SDR}分别表示使用和不使用HDR输入的策略。最大化 $\mathcal{K}_{\text{HDR}}$ 可确保移除HDR词元会显著扰动解码分布，从而激励模型利用HDR特定信息，而不是退化为仅依赖标准动态范围的推理。

（*ii*）双熵正则化。对比KL最大化中一个众所周知的陷阱是熵膨胀：策略可以通过生成过度不确定的输出来轻易满足目标（Rafailov 等人，2023年；Zeng 等人，2024）。为防止这种情况，我们在高动态范围和标准动态范围路径上都引入策略熵正则化：

$$\mathcal{H}_{\text{dual}}(\theta) = \mathbb{E}_{o\sim\pi_{\theta_{old}}} \frac{1}{K}\sum_{i,t}\Big[\eta_1\,\mathcal{H}\big(\pi_\theta^{HDR}(o_{i,t})\big) + \eta_2\,\mathcal{H}\big(\pi_\theta^{SDR}(o_{i,t})\big)\Big]. \tag{7}$$

其中， $\mathcal{H}$ 表示词元级熵，即 $\mathcal{H}(\pi_\theta) = \log \pi_\theta$， η_1 和 η_2 为超参数。这可以防止模式崩溃，同时保留清晰且以HDR为锚定的分布。

（*iii*）高熵加权（*HEW*）。GRPO 为同一补全中的所有词元分配相同的归一化优势 $\hat{} A_i$ ， o_i不论其信息量如何。然而，近期研究（Cui 等人，2025）表明，强化学习能够从聚焦于促进探索的词元策略梯度中获益，从而改善推理；而遵循固定推理路径的词元提供的信号很少。在 HDR-UGC 视频质量评估中，当模型必须识别或校准 HDR 特定失真（例如渐变色带、高光裁剪、近黑压暗）时，通常会出现高熵词元。通过增强这些词元上的学习信号，HEW 将策略优化引导至信息量最大的推理步骤，从而实现更强的 HDR $\hat{}$锚定和更精确的平均意见分预测。随后，我们将组归一化优势 A_i 重新缩放为词元特定优势：

$$w_{i,t} = \text{clip}\left(1 + \lambda_{\text{HEW}} \frac{H_{i,t}}{\frac{1}{|o_i|}\sum_{t'=1}^{|o_i|} H_{i,t'}},\, w_{\min},\, w_{\max}\right), \quad \tilde{A}_{i,t} = w_{i,t}\cdot\hat{A}_i. \tag{8}$$

其中$H_{i,\ t}$ 为逐词元熵。

HAPO 完整目标。 结合这些项可得：

$$\begin{aligned}\mathcal{J}_{\text{HAPO}}(\theta) = \mathbb{E}_{o\sim\pi_{\theta_{\text{old}}}}\Bigg[\frac{1}{K}\sum_{i,t}\min\Big(\rho_{i,t}\tilde{A}_{i,t},\, \text{clip}(\rho_{i,t}, 1-\epsilon, 1+\epsilon)\tilde{A}_{i,t}\Big)\Bigg] \\ -\beta\, D_{\text{KL}}\big(\pi_\theta^{HDR}\,\|\,\pi_{\text{ref}}\big) + \gamma\,\mathcal{K}_{\text{HDR}}(\theta) - \mathcal{H}_{\text{dual}}(\theta).\end{aligned} \tag{9}$$

这会在保持优化稳定的同时，确保进行高动态范围感知推理。

互信息视角。 我们的 HDR–SDR对比KL 可解释为在高动态范围输入与模型输出之间建立一种信息论依赖关系。令 v 表示高动态范围视频， v^{SDR} 表示其标准动态范围色调映射对应版本， o 表示输出序列。通过应用变分互信息界（Ish- mael Belghazi 等人，2018；Ma 等人，2023年），我们得到$\mathbb{E}_{v,\ v^{SDR}}[\mathcal{K}_{\text{HDR}}(\theta)]$ 如下：

$$\mathbb{E}_{v,v^{SDR}}\,\mathbb{E}_{o\sim\pi_\theta(\cdot\,|\,v,v^{SDR})}\left[\log\frac{\pi_\theta(o\,|\,v,v^{SDR})}{\pi_\theta(o\,|\,v^{SDR})}\right] \;\geq\; I_\theta\big(o;\,v,v^{SDR}\,\big|\,v^{SDR}\big) - \kappa_\theta. \tag{10}$$

其中 $I_\theta(o; v, v^{SDR}|v^{SDR})$ 是在 π_θ 下的条件互信息，而 κ_θ 刻画了因以 v^{SDR} 为条件而产生的不匹配。该结果表明，最大化公式 6 可被证明会提高HDR信息量，从而确保策略依赖HDR 特定线索，而不是退化为仅SDR推理。

5.3 奖励与训练流程

HAPO联合优化三种奖励信号：格式（R_{fmt}）、回归精度（R_{sc}）（Li 等人，2025；You 等人，2024b；Wu 等人，2025），以及自一致性（R_{self}）（Zhou 等人，2025），组合为

$$\mathcal{R}_i = w_{\text{fmt}} R_{\text{fmt}} + w_{\text{sc}} R_{\text{sc}} + w_{\text{self}} R_{\text{self}}. \tag{11}$$

高斯加权分数奖励可稳定细粒度MOS预测，而自奖励则巩固组内一致性。

两阶段强化学习训练。 我们的训练遵循基于强化学习的两阶段范式（Chen等人，2025c；Dai等人，2025），两者均以相同目标进行优化，但服务于不同目的：

- **阶段1（模态对齐）：** 通过短程HAPO运行对齐HDR词元与投影层。
- **阶段2（完整RFT）：** 在HDR-UGC语料上应用完整的HAPO优化，平衡失真多样性与推理质量。

总体而言，**HDR-Q** 统一了感知灵敏度与推理稳定性，HDR编码器注入物理亮度感知，对比KL强化锚定，熵正则化抑制不确定性，而HEW则细化词元级学习，从而产生准确、可解释的HDR感知质量判断。

表 1 在 **Beyond8Bits**上的表现。 **蓝色粗体**表示最佳结果，次佳结果带下划线

模型	SRCC(↑)	PLCC(↑)	RMSE(↓)	KRCC(↑)
深度学习模型				
BRISQUE（Mittal 等人，2012）	0.4096	0.4689	11.7019	0.2797
CONTRIQUE（Madhusudana 等人，2022）	0.6245	0.6054	15.0224	0.4464
RE-IQA（Saha 等人，2023年）	0.5698	0.5441	17.9049	0.4038
VBLIINDS（Saad 等人，2014）	0.4440	0.4397	11.7234	0.3044
CONVIQT（Madhusudana 等人，2023年）	0.7987	0.8099	8.4807	0.6095
FastVQA（Wu 等人，2022a）	0.4909	0.4193	26.1325	0.3398
FasterVQA（Wu 等人，2022b）	0.4808	0.3224	29.6357	0.3367
DOVER（Wu 等人，2022c）	0.5094	0.5037	16.7176	0.3548
COVER（He 等人，2024）	0.6645	0.6645	16.8597	0.4870
HDRMAX（Shang 等人，2023年）	0.6054	0.6070	10.1400	0.4277
HDRChipQA（Ebenezer 等人，2024b）	0.7180	0.7290	8.2987	0.5282
HIDROVQA (Saini 等人，2024)	0.8508	0.8784	<u>6.0875</u>	0.6694
多模态大语言模型基座模型				
Qwen2.5-VL(7B) (Bai 等人，2025)	0.3089	0.3228	27.8899	0.2432
GLM-4.1V-Thinking(9B) (Hong 等人，2025)	0.2641	0.3944	23.8883	0.2924
Ovis2.5(9B)(Lu 等人，2025)	0.3423	0.3860	26.7570	0.2823
OmniLong-Qwen2.5-VL(7B) (Song 和 Wu，2025)	0.3472	0.3595	25.7616	0.2677
多模态大语言模型视觉问答模型				
Q-Align(Wu 等人，2023d)	0.4615	0.3673	20.3411	0.3257
Q-Insight(Li 等人，2025)	0.5170	0.5621	20.7832	0.4138
Q-Instruct (Li 等人，2025)	0.5035	0.4712	19.6567	0.3496
DeQA (You 等人，2025)	0.5064	0.4642	19.5772	0.3586
Visual-Quality-Q1 (Wu 等人，2025)	0.3909	0.3617	23.6462	0.2809
HDR-Q（标准动态范围）	<u>0.8914</u>	<u>0.8895</u>	7.4240	<u>0.7052</u>
HDR-Q（完整）	**0.9206**	**0.9118**	**5.1594**	**0.7218**

6 实验

6.1 实验设置

数据集。 我们在精心整理的 **HDR-Q** 上进行评估，并在两个公开的高动态范围视频质量评估数据集上测试泛化能力：LIVE-HDR (Shang 等人,2023年) 和 SFV+HDR (Wang等人, 2024)。**Beyond8Bits** 按源身份划分为 70%/20%/10% 的训练集/验证集/测试集，以避免重叠。

指标。 遵循视觉问答的惯例 (Madhusudana 等人, 2023年; Lu 等人, 2024; Saini 等人, 2025, 2024), 我们报告斯皮尔曼等级相关系数 (SRCC)、皮尔逊线性相关系数 (PLCC) 和肯德尔等级相关系数 (KRCC) (↑越高越好), 以及均方根误差 (↓ 越低越好), 并与平均意见分进行比较。

基线方法。我们比较了四类方法。(i) 无参考视频质量评估：BRISQUE（Mittal 等人, 2012）、VBLI- INDS（Saadet al., 2014）、FastVQA（Wu 等人, 2023a）、FasterVQA（Wu 等人, 2022b）、DOVER（Wu 等人, 2022c）、CONVIQT（Madhusudana etal., 2023年）、COVER（He et al., 2024）；(ii) 高动态范围视频质量评估：HDR- MAX（Ebenezer etal., 2023年）、HDR-ChipQA（Ebenezer et al., 2024b）、HIDRO-VQA（Saini et al., 2024）；(iii) 多模态大语言模型／视觉语言模型视频质量评估：Q-Align（Wu 等人, 2023d）、Q-Instruct（Wu 等人, 2024a）、Q-Insight（Li 等人, 2025）、DeQA（You 等人, 2025）、Visual-Quality-R1（Wu 等人, 2025）；(iv) 基础多模态大语言模型：Qwen2.5-VL（Bai 等人, 2025）、GLM-4.1V-Thinking（Hong 等人, 2025）、Ovis2.5（Lu 等人, 2025）、OmniLong-Qwen2.5-VL（Song 和 Wu, 2025）。在适用情况下，我们按照作者提供的流程在**Beyond8Bits**上重新训练方法；其余方法则以发布版本进行评估。

实现细节。HDR-Q 基于 Ovis2.5（Lu 等人, 2025）构建，并使用秩为 4 的 LoRA 适配器（Hu 等人, 2022）。帧以原生 10 位感知量化格式输入（不进行线性缩减）。每个片段均匀采样为 $T = 8$帧；来自 $\mathcal{E}_\psi(x_t)$ 的视觉令牌和来自 $\mathcal{E}_\psi(x_t^{\mathrm{SDR}})$ 的标准动态范围令牌通过可学习投影输入语言解码器。在 HAPO 中，组大小为 $K = 8$；片段范围为 $\epsilon = 0.1$（片段较高）；参考 KL 权重为 $\beta = 0.02$；HDR–SDR对比KL 权重为 $\gamma = 0.5$；策略熵为 η_1、$\eta_2 = 0.01$、0.05；HEW 调节为 $\lambda_{\mathrm{HEW}} = 0.3$，其中 $w_{\min} = 0.5$、$w_{\max} =$ 为 2.0。在高斯分数奖励 R_{sc} 中，我们使用 $\sigma = 3$和 $\alpha = 1$；权重（w_{fmt}、w_{sc}、w_{self}）通过验证集调优。我们使用 AdamW（Loshchilov 和 Hutter, 2017），学习率为 1×10^{-5}，批量大小为 4。训练使用四块 NVIDIA H200 GPU。



图 6 对于同一段高动态范围视频，OVIS 2.5 会做出多个错误判断。相比之下，我们的 HAPO 增强型 **HDR-Q** 能够基于高动态范围内容进行推理。（放大查看效果最佳）

表 2 在 LIVE-HDR（Shang 等人, 2023年）和 SFV+HDR（Wang等人, 2024）数据集上的跨数据集性能比较。

模型	LIVE-HDR				SFV+HDR			
	斯皮尔曼秩相关系数(↑)	皮尔逊线性相关系数(↑)	均方根误差(↓)	肯德尔秩相关系数(↑)	斯皮尔曼秩相关系数(↑)	皮尔逊线性相关系数(↑)	均方根误差(↓)	肯德尔秩相关系数(↑)
深度学习模型								
BRISQUE (Mittal 等人, 2012)	0.7251	0.7139	12.6404	0.3424	0.4664	0.4186	0.3811	0.3165
CONTRIQUE (Madhusudana 等人, 2022)	0.8170	0.7875	11.2514	0.5876	0.5901	0.5959	0.3368	0.4204
RE-IQA (Saha 等人, 2023年)	0.7196	0.6883	15.1653	0.5197	0.5822	0.5998	0.3072	0.4145
VBLIINDS(Saad 等人, 2014)	0.7483	0.7193	12.7794	0.2541	0.3335	0.2713	0.3988	0.2300
CONVIQT (Madhusudana 等人, 2023年)	0.7922	0.8001	11.9681	0.6041	0.5736	0.6017	0.3412	0.4170
FastVQA (Wu 等人, 2022a)	0.5182	0.5727	18.8379	0.3822	0.7130	0.7295	0.7467	0.5193
FasterVQA (Wu 等人, 2023a)	0.3385	0.4114	22.1425	0.2282	0.6948	0.6889	0.3081	0.5089
DOVER (Wu 等人, 2022c)	0.6303	0.6832	17.0005	0.4692	0.6001	0.6154	0.5750	0.4270
COVER (He 等人, 2024)	0.5022	0.5013	21.3297	0.3731	0.6613	0.7048	0.6831	0.4705
HDRMAX (Shang 等人, 2023年)	0.6308	0.5088	15.4146	0.4509	0.5371	0.5463	0.3495	0.3821
HDRChipQA (Ebenezer 等人, 2024b)	0.8250	0.8344	9.8038	0.4501	0.6296	0.6508	0.3271	0.4440
HIDROVQA (Saini 等人, 2024)	0.8793	0.8678	8.8743	0.6919	0.7003	0.7320	0.2735	0.5156
多模态大语言模型基座模型								
Qwen2.5-VL(7B) (Bai 等人, 2025)	0.3099	0.3630	30.2082	0.2411	0.2925	0.2696	0.7480	0.2270
GLM-4.1V-Thinking(9B) (Hong 等人, 2025)	0.4513	0.5517	26.3800	0.3464	0.5971	0.6066	0.4591	0.4484
Ovis2.5(9B) (Lu 等人, 2025)	0.2948	0.3124	29.7789	0.2154	0.5909	0.5317	0.7016	0.4528
OmniLong-Qwen2.5-VL(7B) (Song 和 Wu, 2025)	0.2403	0.2223	29.7394	0.1853	0.2363	0.2212	0.7403	0.1823
多模态大语言模型视觉问答模型								
Q-Align (Wu 等人, 2023d)	0.3346	0.3604	19.8287	0.2313	0.6968	0.6709	0.5097	0.4991
Q-Insight(Li 等人, 2025)	0.3675	0.3825	25.0578	0.2820	0.6266	0.4685	0.6636	0.4747
Q-Instruct (Wu 等人, 2024a)	0.4083	0.4340	23.1015	0.2839	0.5830	0.5501	1.0250	0.3975
DeQA (You 等人, 2025)	0.3321	0.3809	19.3193	0.2298	0.6850	0.6721	0.4452	0.4845
Visual-Quality-Q1(Wu 等人, 2025)	0.4824	0.5394	20.8971	0.3564	0.5955	0.5577	0.5878	0.4416
HDR-Q（标准动态范围）	0.8542	0.8445	12.4121	0.6681	0.6971	0.7019	0.3075	0.4885
HDR-Q（完整）	**0.9081**	**0.8978**	**7.6031**	**0.7363**	**0.7251**	**0.7502**	**0.2514**	**0.5261**

表3 Beyond8Bits 的组件消融。 √=启用，=禁用。CoT 长度：思维链长度；Tok. H：平均 Token 熵。

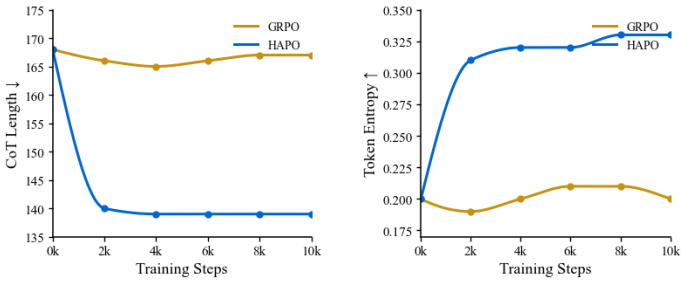
变体	HDR 编码器	HAPO	HDR–SDR KL 散度	双重熵	HEW	自评	PLCC	SRCC	RMSE	KRCC	思维链长度	词元 H
GRPO 基线	✗	✗	✗	✗	✗	✗	0.79	0.81	10.73	0.56	168	0.20
GRPO + HDR 编码器	✓	✗	✗	✗	✗	✗	0.81	0.83	8.96	0.61	161	0.24
HAPO（无 HDR–SDR KL 散度）	✓	✓	✗	✓	✓	✓	0.84	0.86	7.10	0.64	142	0.29
HAPO（无双重熵）	✓	✓	✓	✗	✓	✓	0.89	0.91	5.82	0.71	148	0.26
HAPO（无 HEW）	✓	✓	✓	✓	✗	✓	0.87	0.88	6.11	0.68	155	0.27
HAPO（无自奖励）	✓	✓	✓	✓	✓	✗	0.90	0.92	5.22	0.71	140	0.31
HDR-Q（完整）	✓	✓	✓	✓	✓	✓	**0.91**	**0.92**	**5.15**	**0.72**	**137**	**0.33**

6.2 主要结果

表 1 报告了在 **Beyond8Bits** 上的定量结果。**HDR-Q** 在各项相关性指标上始终优于所有基于标准动态范围、高动态范围和多模态大语言模型的基线，在均方根误差方面也取得了显著提升。与 HDR-ChipQA（Ebenezer 等人，2024b）和 HIDRO-VQA（Saini 等人，2024）相比，**HDR-Q** 的斯皮尔曼等级相关系数和皮尔逊线性相关系数更高，均方根误差更低，表明高动态范围感知嵌入与 HAPO 对齐具有优势。与 FastVQA（Wu 等人，2023a）和 DOVER（Wu 等人，2022c）相比，**HDR-Q**在多样的用户生成内容采集流程下仍保持稳健。相较于所有多模态大语言模型和视觉语言模型，**HDR-Q**的优势源于在 10 位感知量化格式上进行高动态范围感知编码器微调（不采用线性标准动态范围缩放），以及采用 HDR–SDR对比KL（以避免模态忽视）。

为检验鲁棒性，我们在 LIVE-HDR（Shang 等人，2023年）和 SFV+HDR（Wang 等人，2024）（表 2）。**HDR-Q**无需重新训练，便能保持较高的相关性和较低的误差，这一证据表明，其高动态范围感知编码器和 HAPO 对齐所生成的表征能够泛化至用户生成内容和专业生成内容的高动态范围领域。

HDR-Q 生成简洁且具备高动态范围感知能力的推理（图 6），能够识别“自然室内场景”“色度偏移导致的可能色调”或“抖动”，并将其与感知判断联系起来。图7 显示，HAPO 稳定了思维链长度，而 HEW 将梯度集中于信息丰富的词元，从而实现高效且可解释的推理。



(a) 思维链长度　　**(b)** 词元熵

图 7 训练迭代过程中思维链（CoT）长度与词元熵的分析。(a) 显示思维链长度的减少，(b) 显示词元熵相应的增加。

6.3 消融研究

表 3量化了各组件的贡献。移除高动态范围微调会显著降低斯皮尔曼等级相关系数，证实 10 位线索至关重要。省略高动态范围–标准动态范围 KL 散度会导致模态忽视，而禁用熵正则化则会使推理变得不稳定且冗长。HEW 能改善词元级信用分配，自奖励则能增强模型在噪声样本上的稳定性。图 7 显示，HAPO 会随着时间推移减少不必要的思维链长度，同时保持或提升准确率，这表明它更善于利用视觉证据，而非增加模板化推理。

6.4 复杂度和吞吐量

HAPO 仅在训练期间额外执行一次 SDR 路径前向传播。推理成本相当于单次 HDR 路径解码，同时在 NVIDIA H200 GPU 上保持了有竞争力的吞吐量。

7 结论

我们攻克了快速增长的 HDR 用户生成视频领域中感知质量评估这一关键挑战。我们推出了 **Beyond8Bits**，这是规模最大的真实世界 HDR 内容众包主观数据集，涵盖多样的场景、设备和压缩设置。我们还提出了 **HDR-Q**，首个用于 HDR 质量评估的多模态大语言模型。该模型结合了我们新颖的 HDR 感知视觉编码器与 HDR感知策略优化（HAPO），后者是一种强化学习框架，通过实施 HDR–SDR感知对齐，并借助双熵正则化和熵加权信用分配来稳定推理。HAPO 能够实现准确、可解释且对 HDR 敏感的质量推理。**HDR-Q** 在 **Beyond8Bits**、LIVE-HDR 和 SFV+HDR 上与人类意见评分的对齐程度均达到最先进水平。通过发布该数据集，我们希望推动 HDR 感知、评估和生成模型对齐领域的未来研究。

8 致谢

本研究得到美国国家科学基金会机器学习基础人工智能研究所（IFML）第2019844号资助。作者感谢德克萨斯大学奥斯汀分校的德克萨斯高级计算中心（TACC）提供 VISTA 计算基础设施，为本文的部分研究成果作出了贡献。

参考文献

99Firms。Facebook 视频统计数据，2024年。https://99firms.com/blog/facebook-video-statistics/。 [在线]。

Naima Aamir、Junaid Mir、Imran Fareed Nizami、Furqan Shaukat 和 Muhammad Majid。Hdr-bvqm：高动态范围盲视频质量模型。《多媒体工具与应用》，80：27701–27715，2021年。https: //api.semanticscholar.org/CorpusID:236339560。

苹果公司 . Apple 设备的 HLS 创作规范，2024 . https://developer.apple.com/documentation/http-live-streaming/hls-authoring-specification-for-apple-devices。访问日期：2024年2月。

Shahrukh Athar、Thilan Costa、Kai Zeng 和 Zhou Wang。UHD-HDR-WCG 视频的感知质量评估。载于*2019* 年 *IEEE* 国际图像处理会议（*ICIP*），第 1740–1744 页。IEEE，2019年。Maryam Azimi 等。PU21：一种新颖的感知均匀编码方法，用于调整现有质量指标以适用于 HDR。载于*2021* 年图像编码研讨会（*PCS*），第 1–5 页。IEEE，2021年。Shuai Bai、Keqin Chen、Xuejing Liu、Jialin Wang、Wenbin Ge、Sibo Song、Kai Dang、Peng Wang、Shijie Wang、Jun Tang 等。Qwen2. 5-vl 技术报告。*arXiv* 预印本 *arXiv:2502.13923*，2025。V Baroncini、K Andersson、AK Ramasubramonian 和 G Sullivan。使用 HEVC Main 10 配置进行 HDR/WCG 视频编码的验证测试报告。载于第 *24* 届 *JCT-VC* 会议论文集 *JCTVC-X1018*，第 293–303 页，2016年。Bowen Chen、Cheng-han Lee、Yixu Chen、Zaixi Shang、Hai Wei 和 Alan C Bovik。Hdrsdr-vqa：用于 HDR 和 SDR 比较评估的主观视频质量数据集。*arXiv* 预印本 *arXiv:2505.21831*，2025a。Bowen Chen、Cheng-han Lee、Yixu Chen、Zaixi Shang、Hai Wei 和 Alan C Bovik。Hdrsdr-vqa：用于 HDR 和 SDR 比较评估的主观视频质量数据集。*arXiv* 预印本 *arXiv:2505.21831*，2025b。Hardy Chen、Haoqin Tu、Fali Wang、Hui Liu、Xianfeng Tang、Xinya Du、Yuyin Zhou 和 Cihang Xie。SFT 还是强化学习？对训练类似 R1*11468*、2025*2504*推理型大型视觉语言模型的初步研究。*arXiv* 预印本 *arXiv:* .c。Tianzhe Chu、Yuexiang Zhai、Jihan Yang、Shengbang Tong、Saining Xie、Dale Schuurmans、Quoc V Le、Sergey Levine 和 Yi Ma。SFT 会记忆，强化学习能泛化：基础模型后训练的比较研究。*arXiv* 预印本 *arXiv:2501.17161*，2025。

Ganqu Cui、Yuchen Zhang、Jiacheng Chen、Lifan Yuan、Zhi Wang、Yuxin Zuo、Haozhan Li、Yuchen Fan、Huayu Chen、Weize Chen 等。《推理语言模型强化学习的熵机制》。网址 *https://arxiv. org/abs/2505.22617*，2025。

Wei Dai、Peilin Chen、Chanakya Ekbote 和 Paul Pu Liang。《Qoq-med：通过领域感知 GRPO 训练构建多模态临床基础模型》。arXiv 预印本 arXiv:2506.00711，2025。Huiyu Duan、Qiang Hu、Jiarui Wang、Liu Yang、Zitong Xu、Lu Liu、Xiongkuo Min、Chunlei Cai、Tianxiao Ye、Xiaoyun Zhang 等。《Finevq：细粒度用户生成内容视频质量评估》。载于《计算机视觉与模式识别会议论文集》，第 3206–3217 页，2025。Joshua P Ebenezer、Zaixi Shang、Yongjun Wu、Hai Wei、Sriram Sethuraman 和 Alan C Bovik。《使视频质量评估模型对位深具有鲁棒性》。IEEE 信号处理快报，30：488–492，2023。Joshua P Ebenezer、Zaixi Shang、Yixu Chen、Yongjun Wu、Hai Wei、Sriram Sethuraman 和 Alan C Bovik。《高动态范围还是标准动态范围？缩放与压缩视频的主观和客观研究》。IEEE 图像处理汇刊，2024a。Joshua P Ebenezer、Zaixi Shang、Yongjun Wu、Hai Wei、Sriram Sethuraman 和 Alan C Bovik。《HDR-ChipQA：高动态范围视频的无参考质量评估》。信号处理：图像通信，129：117191，2024b。Joshua Peter Ebenezer、Zaixi Shang、Yongjun Wu、Hai Wei、Sriram Sethuraman 和 Alan C Bovik。《ChipQA：通过时空芯片进行无参考视频质量预测》。IEEE 图像处理汇刊，30：8059–8074，2021。Google 支持。《推荐的上传编码设置》，2024。https://support.google.com/youtube/answer/ 1722171?hl=en。访问日期：2024 年 2 月。Chenlong He、Qi Zheng、Ruoxi Zhu、Xiaoyang Zeng、Yibo Fan 和 Zhengzhong Tu。《COVER：全面的视频质量评估器》。载于《2024 IEEE/CVF 计算机视觉与模式识别会议研讨会（CVPRW）》，第 5799–5809 页，2024。doi：10.1109/CVPRW63382.2024.00589。Wenyi Hong、Wenmeng Yu、Xiaotao Gu、Guo Wang、Guobing Gan、Haomiao Tang、Jiale Cheng、Ji Qi、Junhui Ji、Lihang Pan 等。《GLM-4.1 V-Thinking：借助可扩展强化学习实现多样化多模态推理》。arXiv 电子印本，第 arXiv–2507 页，2025。Edward J Hu、Yelong Shen、Phillip Wallis、Zeyuan Allen-Zhu、Yuanzhi Li、Shean Wang、Lu Wang、Weizhu Chen 等。《LoRA：大语言模型的低秩适配》。ICLR，1(2)：3，2022。国际电信联盟。《电视图像质量主观评估方法》。技术报告 BT.500-14，国际电信联盟，2019。https://www.itu.int/rec/ R-REC-BT.500-14-201910-I/en。Mohamed Ishmael Belghazi、Aristide Baratin、Sai Rajeswar、Sherjil Ozair、Yoshua Bengio、Aaron Courville 和 R Devon Hjelm。《MINE：互信息神经估计》。arXiv 电子印本，第 arXiv–1801 页，2018。Jari Korhonen。《无参考消费级视频质量评估的两级方法》。IEEE 图像处理汇刊，28(12)：5923–5938，2019。Dingquan Li、Tingting Jiang 和 Ming Jiang。《野外视频的质量评估》。载于《ACM 多媒体》，第 2351–2359 页。ACM，2019。Weiqi Li、Xuanyu Zhang、Shijie Zhao、Yabin Zhang、Junlin Li、Li Zhang 和 Jian Zhang。《Q-Insight：通过视觉强化学习理解图像质量》。arXiv 预印本 arXiv:2503.22679，2025。Zhi Li、Christos G. Bampis、Lucjan Janowski 和 Ioannis Katsavounidis。《主观实验中受试者行为的简单模型》。载于《电子成像》，第 2020 卷，第 131–1–131–14 页，2020。doi：10.2352/ISSN.2470-1173.2020.11.HVEI-131。https://doi.org/10.2352/ISSN.2470-1173.2020.11.HVEI-131。Ilya Loshchilov 和 Frank Hutter。《解耦权重衰减正则化》。arXiv 预印本 arXiv:1711.05101，2017。Shiyin Lu、Yang Li、Yu Xia、Yuwei Hu、Shanshan Zhao、Yanqing Ma、Zhichao Wei、Yinglun Li、Lunhao Duan、Jianshan Zhao 等。《Ovis2.5 技术报告》。arXiv 预印本 arXiv:2508.11737，2025。Yiting Lu、Xin Li、Yajing Pei、Kun Yuan、Qizhi Xie、Yunpeng Qu、Ming Sun、Chao Zhou 和 Zhibo Chen。《KVQ：快手短视频质量评估》。载于《IEEE/CVF 计算机视觉与模式识别会议论文集》，第 25963–25973 页，2024。Xiao Ma、Bingyi Kang、Zhongwen Xu、Min Lin 和 Shuicheng Yan。《互信息正则化离线强化学习》。载于《神经信息处理系统进展》，36：19058–19072，2023。

Pavan C. Madhusudana、Neil Birkbeck、Yilin Wang、Balu Adsumilli 和 Alan C. Bovik。使用对比学习进行图像质量评估。IEEE 图像处理汇刊，31：4149–4161，2022。Pavan C Madhusudana、Neil Birkbeck、Yilin Wang、Balu Adsumilli 和 Alan C Bovik。CONVIQT：对比视频质量估计器。IEEE 图像处理汇刊，32：5138–5152，2023。RK Mantiuk 和 M Azimi。PU21：一种新颖的感知均匀编码方法，可用于调整现有质量指标以适配高动态范围。2021。doi：10.17863/CAM.74563。https://www.repository.cam.ac.uk/handle/1810/327114。Anish Mittal、Anush Krishna Moorthy 和 Alan Conrad Bovik。空间域中的无参考图像质量评估。IEEE 图像处理汇刊，21(12)：4695–4708，2012。doi：10.1109/TIP.2012.2214050。Anish Mittal、Rajiv Soundararajan 和 Alan C. Bovik。打造“完全盲”的图像质量分析器。IEEE 信号处理快报，20(3)：209–212，2013。doi：10.1109/LSP.2012.2227726。Anish Mittal、Michele A. Saad 和 Alan C. Bovik。完全盲的视频完整性预言机。IEEE 图像处理汇刊，25(1)：289–300，2016。Maryam Mohsin。2020年每位营销人员都应了解的10项YouTube统计数据，2020。https://www.oberlo.com/blog/youtube-statistics。 [在线]。Manish Narwaria、Matthieu Perreira Da Silva 和 Patrick Le Callet。HDR-VQM：一种高动态范围视频的客观质量度量。信号处理：图像通信，35：46–60，2015。ISSN 0923-5965。doi：https://doi.org/10.1016/j.image.2015.04.009。https://www.sciencedirect.com/science/article/pii/S0923596515000703。Mikko Nuutinen、Toni Virtanen、Mikko Vaahteranoksa、Tero Vuori、Pirkko Oittinen 和 Jukka Häkkinen。CVD2014——用于评估无参考视频质量评估算法的数据集。IEEE 图像处理汇刊，25(7)：3073–3086，2016。Omnicore。TikTok数据一览，2024。https://www.omnicoreagency.com/tiktok-statistics/。 [在线]。Xiaofei Pan、Jiaqi Zhang、Shanshe Wang、Shiqi Wang、Yun Zhou、Wenhua Ding 和 Yahui Yang。高动态范围视频质量评估：压缩高动态范围视频的感知评估。视觉通信与图像表示期刊，57：76–83，2018。Yanyun Pu、Kehan Li、Zeyi Huang、Zhijie Zhong 和 Kaixiang Yang。MVQA-68K：一个具有质量可解释性、带有多维标注和因果标注的视频评估数据集。arXiv预印本 arXiv:2509.11589，2025。Rafael Rafailov、Archit Sharma、Eric Mitchell、Christopher D Manning、Stefano Ermon 和 Chelsea Finn。直接偏好优化：你的语言模型其实是一个奖励模型。神经信息处理系统进展，36：53728–53741，2023。Martin Rerabek、Philippe Hanhart、Pavel Korshunov 和 Touradj Ebrahimi。高动态范围视频压缩的主观与客观评估。载于第9届国际消费电子视频处理与质量指标研讨会（VPQM），2015。Michele A. Saad、Alan C. Bovik 和 Christophe Charrier。自然视频质量的盲预测。IEEE 图像处理汇刊，23(3)：1352–1365，2014。Avinab Saha、Sandeep Mishra 和 Alan C. Bovik。RE-IQA：用于野外图像质量评估的无监督学习。载于 IEEE/CVF 计算机视觉与模式识别会议，CVPR 2023，加拿大不列颠哥伦比亚省温哥华，2023年6月17–24日，第5846–5855页。IEEE，2023。doi：10.1109/CVPR52729.2023.00566。https: // doi.org/10.1109/CVPR52729.2023.00566。Shreshth Saini、Avinab Saha 和 Alan C Bovik。HIDRO-VQA：用于视频质量评估的高动态范围预言机。载于 IEEE/CVF冬季计算机视觉应用会议论文集，第469–479页，2024。Shreshth Saini、Alan C Bovik、Neil Birkbeck、Yilin Wang 和 Balu Adsumilli。Chug：众包用户生成高动态范围视频质量数据集。载于2025年 IEEE 国际图像处理会议（ICIP），第2504–2509页。IEEE，2025。John Schulman、Filip Wolski、Prafulla Dhariwal、Alec Radford 和 Oleg Klimov。近端策略优化算法。arXiv预印本 arXiv:1707.06347，2017。Kalpana Seshadrinathan、Rajiv Soundararajan、Alan Conrad Bovik 和 Lawrence K. Cormack。视频主观与客观质量评估研究。IEEE 图像处理汇刊，19(6)：1427–1441，2010。doi：10.1109/TIP.2010.2042111。

Zaixi Shang、Joshua P Ebenezer、Abhinau K Venkataramanan、Yongjun Wu、Hai Wei、Sriram Sethuraman 和 Alan C Bovik。高动态范围视频主观与客观质量评估研究。IEEE 图像处理汇刊，33：42–57，2023。Zhihong Shao、Peiyi Wang、Qihao Zhu、Runxin Xu、Junxiao Song、Xiao Bi、Haowei Zhang、Mingchuan Zhang、YK Li、Yang Wu 等。DeepSeekMath：突破开放语言模型数学推理能力的极限。arXiv预印本 arXiv:2402.03300，2024。Zeina Sinno 和 Alan Conrad Bovik。大规模感知视频质量研究。IEEE 图像处理汇刊，28 (2)：612–627，2019。Yin Song 和 Chen Wu。aws-prototyping/omnilong-qwen2.5-vl-7b，2025。https://huggingface.co/aws-prototyping/OmniLong-Qwen2.5-VL-7B。Michael Tschannen、Alexey Gritsenko、Xiao Wang、Muhammad Ferjad Naeem、Ibrahim Alabdulmohsin、Nikhil Parthasarathy、Talfan Evans、Lucas Beyer、Ye Xia、Basil Mustafa 等。SigLIP 2：具备更强语义理解、定位和密集特征能力的多语言视觉语言编码器。arXiv预印本 arXiv:2502 14786，2025。。Yilin Wang、Joong Gon Yim、Neil Birkbeck 和 Balu Adsumilli。YouTube sfv+ 高动态范围质量数据集。载于2024年 IEEE 国际图像处理会议（ICIP），第96–102页。IEEE，2024。Zhenhailong Wang、Xuehang Guo、Sofia Stoica、Haiyang Xu、Hongru Wang、Hyeonjeong Ha、Xiusi Chen、Yangyi Chen、Ming Yan、Fei Huang 等。面向多模态推理的感知感知策略优化。arXiv预印本 arXiv:2507.06448，2025。Haoning Wu、Chaofeng Chen、Jingwen Hou、Liang Liao、Annan Wang、Wenxiu Sun、Qiong Yan 和 Weisi Lin。Fast-VQA：通过片段采样实现高效端到端视频质量评估。载于《计算机视觉——ECCV 2022：第17届欧洲计算机视觉会议》，以色列特拉维夫，2022年10月23–27日，论文集第VI部分，第538–554页，柏林、海德堡，2022a。施普林格出版社。ISBN 978-3-031-20067-0。doi：10.1007/978-3-031-20068-7_31。https: //doi.org/10.1007/978-3-031-20068-7_31。Haoning Wu、Chaofeng Chen、Liang Liao、Jingwen Hou、Wenxiu Sun、Qiong Yan、Jinwei Gu 和 Weisi Lin。用于高效端到端视频质量评估的邻域代表性采样，2022b。Haoning Wu、Liang Liao、Chaofeng Chen、Jingwen Hou、Annan Wang、Wenxiu Sun、Qiong Yan 和 Weisi Lin。分离用户生成内容视频质量评估中的美学与技术效果。CoRR，abs/2211.04894，2022c。Haoning Wu、Chaofeng Chen、Liang Liao、Jingwen Hou、Wenxiu Sun、Qiong Yan、Jinwei Gu 和 Weisi Lin。用于高效端到端视频质量评估的邻域代表性采样。IEEE 模式分析与机器智能汇刊，45(12)：15185–15202，2023a。doi：10.1109/TPAMI.2023.3319332。https://doi.org/10.1109/TPAMI.2023.3319332。Haoning Wu、Erli Zhang、Liang Liao、Chaofeng Chen、Jingwen Hou、Annan Wang、Wenxiu Sun、Qiong Yan 和 Weisi Lin。迈向可解释的真实场景视频质量评估：数据集与语言提示方法。载于 Abdulmotaleb El-Saddik、Tao Mei、Rita Cucchiara、Marco Bertini、Diana Patricia Tobon Vallejo、Pradeep K. Atrey 和 M. Shamim Hossain 编，《第31届 ACM 国际多媒体会议论文集》，MM 2023，加拿大安大略省渥太华，2023年10月29日–11月3日，第1045–1054页。ACM，2023b。doi：10.1145/3581783.3611737。https://doi.org/10.1145/3581783.3611737。Haoning Wu、Zicheng Zhang、Erli Zhang、Chaofeng Chen、Liang Liao、Annan Wang、Chunyi Li、Wenxiu Sun、Qiong Yan、Guangtao Zhai 等。Q-Bench：面向低层视觉通用基础模型的基准测试。arXiv预印本 arXiv:2309.14181，2023c。Haoning Wu、Zicheng Zhang、Weixia Zhang、Chaofeng Chen、Liang Liao、Chunyi Li、Yixuan Gao、Annan Wang、Erli Zhang、Wenxiu Sun 等。Q-Align：通过离散文本定义等级指导大型多模态模型进行视觉评分。arXiv预印本 arXiv:2312.17090，2023d。Haoning Wu、Zicheng Zhang、Erli Zhang、Chaofeng Chen、Liang Liao、Annan Wang、Kaixin Xu、Chunyi Li、Jingwen Hou、Guangtao Zhai 等。Q-Instruct：提升多模态基础模型的低层视觉能力。载于 IEEE/CVF 计算机视觉与模式识别会议论文集，第25490–25500页，2024a。Haoning Wu、Hanwei Zhu、Zicheng Zhang、Erli Zhang、Chaofeng Chen、Liang Liao、Chunyi Li、Annan Wang、Wenxiu Sun、Qiong Yan 等。迈向开放式视觉质量比较。载于欧洲计算机视觉会议，第360–377页。施普林格，2024b。

Tianhe Wu、Jian Zou、Jie Liang、Lei Zhang 和 Kede Ma。Visualquality-r1：通过强化学习进行排序，以推理驱动图像质量评估。*arXiv* 预印本 *arXiv:2505.14460*，2025年。Zhenqiang Ying、Maniratnam Mandal、Deepti Ghadiyaram 和 Alan Bovik。Patch-vq：“修补”视频质量问题。载于 *IEEE/CVF* 计算机视觉与模式识别会议论文集，第 14019–14029 页，2021年。Zhiyuan You、Jinjin Gu、Zheyuan Li、Xin Cai、Kaiwen Zhu、Chao Dong 和 Tianfan Xue。野外描述性图像质量评估。*arXiv* 预印本 *arXiv:2405.18842*，2024a。Zhiyuan You、Zheyuan Li、Jinjin Gu、Zhenfei Yin、Tianfan Xue 和 Chao Dong。超越分数的刻画：通过多模态语言模型推进图像质量评估。载于 欧洲计算机视觉会议，第 259–276 页。施普林格，2024b。Zhiyuan You、Xin Cai、Jinjin Gu、Tianfan Xue 和 Chao Dong。利用分数分布教大语言模型回归精确的图像质量分数。载于 计算机视觉与模式识别会议论文集，第 14483–14494 页，2025年。Qiying Yu、Zheng Zhang、Ruofei Zhu、Yufeng Yuan、Xiaochen Zuo、Yu Yue、Weinan Dai、Tiantian Fan、Gaohong Liu、Lingjun Liu 等。Dapo：一个可扩展的开源大语言模型强化学习系统。*arXiv* 预印本 *arXiv:2503 14476*，2025年。。Qingcheng Zeng、Mingyu Jin、Qinkai Yu、Zhenting Wang、Wenyue Hua、Zihao Zhou、Guangyan Sun、Yanda Meng、Shiqing Ma、Qifan Wang 等。不确定性很脆弱：操纵大语言模型中的不确定性。*arXiv* 预印本 *arXiv:2407.11282*，2024。Zicheng Zhang、Wei Wu、Wei Sun、Dangyang Tu、Wei Lu、Xiongkuo Min、Ying Chen 和 Guangtao Zhai。Md-vqa：用户生成内容直播视频的多维质量评估，2023年。https://arxiv.org/abs/2303.14933。Zicheng Zhang、Ziheng Jia、Haoning Wu、Chunyi Li、Zijian Chen、Yingjie Zhou、Wei Sun、Xiaohong Liu、Xiongkuo Min、Weisi Lin 等。Q-bench-video：对大型多模态模型的视频质量理解进行基准测试。载于 计算机视觉与模式识别会议论文集，第 3229–3239 页，2025年。Xu Zheng、Chenfei Liao、Yuqian Fu、Kaiyu Lei、Yuanhuiyi Lyu、Lutao Jiang、Bin Ren、Jialei Chen、Jiawen Wang、Chengxin Li 等。多模态大语言模型深受模态偏差影响。*arXiv* 预印本 *arXiv:2505.18657*，2025年。Xin Zhou、Yiwen Guo、Ruotian Ma、Tao Gui、Qi Zhang 和 Xuanjing Huang。内部奖励模型的自一致性可提升自奖励语言模型。*arXiv* 预印本 *arXiv:2502.08922*，2025年。Hanwei Zhu、Haoning Wu、Yixuan Li、Zicheng Zhang、Baoliang Chen、Lingyu Zhu、Yuming Fang、Guangtao Zhai、Weisi Lin 和 Shiqi Wang。通过教大型多模态模型进行比较，实现自适应图像质量评估。神经信息处理系统进展，37:32611–32629，2024。