



图像增强中的图像质量评估

汇报人：张文剑



Augmenting Perceptual Super-Resolution via Image Quality Predictors

Fengjia Zhang*

Samrudhdi B. Rangrej*

Tristan Aumentado-Armstrong*

Afsaneh Fazly

Alex Levinshtein

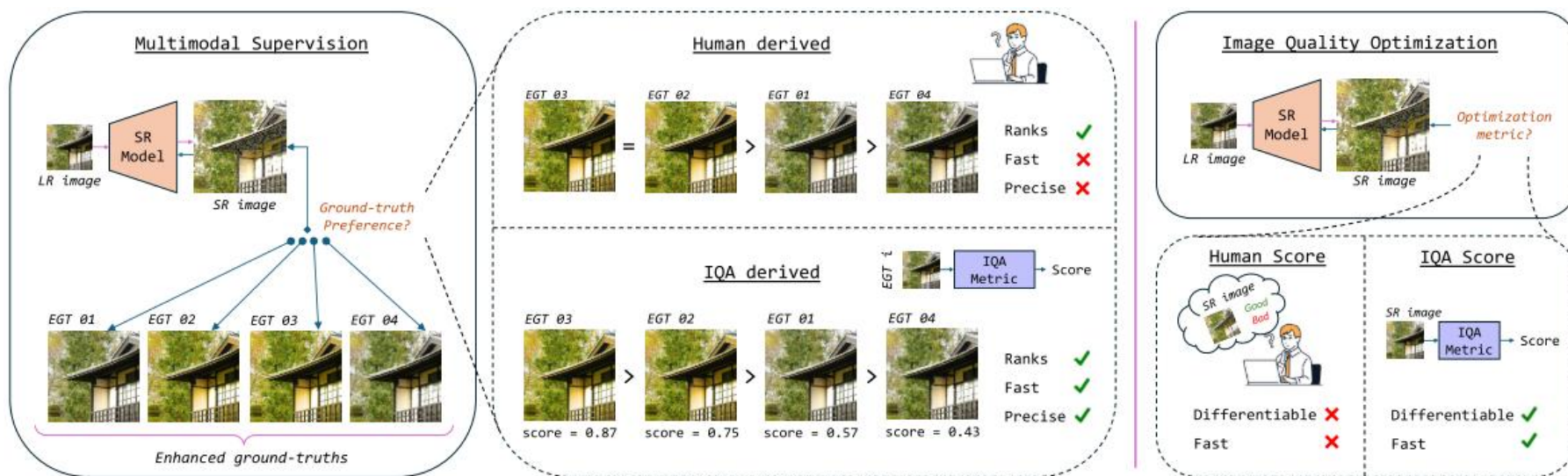
AI Center – Toronto, Samsung Electronics

`{f.zhang2, s.rangrej, tristan.a, a.fazly, alex.lev}@samsung.com`

论文于4月挂在Arxiv上，针对由最小化像素误差获得的分布并不是超分辨率渴望的分布这一问题，为了获得代表更高图像质量的分布，作者探索了NR-IQA指标应用于SR任务的两种方法。

Motivation

- 在单图像超分辨率SISR任务中，给定低分辨率图像不仅仅是从找到的一个高分辨率分布中采样到“平均”样本，而是采样到在感知质量上的可信样本。
- 基于像素损失训练的早期模型，有效地对像素空间中可能的解决方案进行了平均，具有高的PSNR和SSIM值，但是却只能输出模糊的图像结果；
 - 后续引入感知度量和对抗性损失的模型，大大提高了输出图像的质量，人类对后者的偏好表明了“理想”的图像质量是要高于基于像素损失训练产找到的一种在所有可能解决方案的“平均”。
 - 受到在T2I任务中采用人类反馈指导的启发，文章尝试采用NR-IQA指标用来改善SISR任务。



Contributions

- 文章展示了在两个humanderived SR数据集上NR-IQA指标详细的分析，表明确定的指标对于提高SR图像质量通常是有用的；
- 探索了NR-IQA指标在SISR任务中的两种应用方式：通过NR-IQA评级的方式从多个GTs采样，以及直接由NR-IQA进行微调优化模型；
- 与使用手动人类反馈相比，使用NR-IQA自动评估的方式在感知质量上达到了相当甚至更好的效果；

选择NR-IQA：

使用图像质量评估指标(IQA)可以用来评价超分辨率图像的感知质量，相比于全参考指标，无参考指标更加适合在真实参考图像不可用的现实场景

NR-IQA指标分析

分析了各种NR-IQA指标和人类判断在超分辨率图像感知质量之间的一致性，选用了公开数据集SBS 180K和HGGT

SBS180k数据集

SBS180K是一个大型的关于图像超分辨率的人类偏好数据集，数据集包括167019、9421个训练和测试图像。每一对图像通过[0~1]的值用来描述在人类注释中偏爱图像对中第二张图像的比例。由于数据集的规模比较大，作者分两个阶段进行训练：

- 阶段一：分析了20个NR-IQA指标及其变体（共计42个指标）在一个训练集子集上的表现；
- 阶段二：选择在阶段一中表现最好的前7个指标在完整的训练集和测试集上作进一步分析；

阶段一:

训练集中包括404对不同超分模型的对比, 随机选择3张LQ图像, 得到1212对SR图像, 选择了: Q-Align ,LIQE , ARNIQA , TOPIQ , TReS , CLIP-IQA(+) , MANIQA , MUSIQ , DBCNN ,PaQ-2-PiQ , HyperIQA , NIMA , WaDIQaM, CNNIQA , NRQM , PI (Perceptual Index), BRISQUE , ILNIQE and NIQE , and PIQE及变体, 测试这些指标与人类偏好的一致性来评估每个指标的准确性

Method	PaQ-2-PiQ [97]	NIMA [†] [74]	MUSIQ [∇] [41]	LIQE [♡] [101]	ARNIQA-TID* [1]	Q-Align [◇] [88]	TOPIQ-NR [11]
Accuracy (%)	76.41	74.91	74.47	74.03	74.03	73.77	73.06

阶段二:

Method	PaQ-2-PiQ [97]	NIMA [†] [74]	MUSIQ [∇] [41]	LIQE [♡] [101]	ARNIQA-TID* [1]	Q-Align [◇] [88]	TOPIQ-NR [11]
Train Acc. (%)	80.41	79.32	79.96	77.70	77.74	80.00	78.30
Test Acc. (%)	80.57	81.37	82.73	77.45	77.07	80.68	81.28

HGGT 数据集

HGGT数据集中包含20193个五元组（GT+4个SR结果），五元组中的每一个增强结果都由人类注释与GT的对比结果positive、similar、negative。在数据集中选择了1270张图像组成的子集，这些子集中至少有一个是negative的评级。

评价方式：使用SRCC系数和正（负）未对齐率

Method	PaQ-2-PiQ [97]	NIMA [†] [74]	MUSIQ [¶] [41]	LIQE [♡] [101]	ARNIQA-TID* [1]	Q-Align [◇] [88]	TOPIQ-NR [11]
SRC ↑	0.10	0.17	0.17	0.03	0.28	0.20	0.09
PM ↓	0.26	0.28	0.16	0.58	0.51	0.39	0.21
NM ↓	0.85	0.79	0.95	0.64	0.45	0.63	0.97

正未对齐率：至少一个positive的评级低于similar;

负未对齐率：至少一个negative的评级高于similar;

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

基于NR-IQA评估的GT采样:


- Softmax-All (SMA): 使用所有 GT 候选, 对这些 GTs 的 IQA 分数进行 softmax 加权后采样;
 - Softmax-Positives (SMP): 只在人类标注为“positive”的 GTs 中进行 softmax 加权采样;
 - Argmax-online (AMO): 调换从多个 GT 中选一个, 再从该 GT 中随机选一个 patch 的顺序;
-

基于NR-IQA评估进行微调:

先前工作已经表明, 直接使用NR-IQA指标作为损失微调模型会出现“hack”现象, 结果容易出现“伪影”或者“油画感”, 因此通过仅在LoRA权重上增加NR-IQA损失来进行正则化优化消除“hack”现象。

$$\mathcal{L}(\theta|\hat{I}, I) = \lambda_{\ell_1} \|I - \hat{I}\|_1 + \lambda_P d_P(\hat{I}, I) + \lambda_A D(\hat{I}),$$

$$\tilde{\mathcal{L}}(\phi|\hat{I}, I) = \mathcal{L}(\phi|\hat{I}, I) - \lambda_Q Q(\hat{I}),$$

Model		FR Low-Lev. Dist.		FR Mid-Lev. Dist.			NR High-Lev. Perceptual Quality			
		PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	LPIPS-ST \downarrow	DISTS \downarrow	MUSIQ \uparrow	NIMA \uparrow	Q-Align \uparrow	TOPIQ \uparrow
Gold Standard	✗	—	—	—	—	—	69.64	5.28	3.78	0.69
SwinIR-OrigOnly	✓	22.72	0.652	0.227	0.174	0.162	59.47	4.87	3.17	0.48
SwinIR-Rand	✓	22.45	0.650	0.180	0.139	0.131	65.27	5.11	3.52	0.59
SwinIR-UPos*	✗	22.30	0.647	0.169	0.129	0.123	66.39	5.16	3.56	0.62
SwinIR-SMA	✓	22.27	0.646	0.171	0.129	0.124	66.73	5.16	3.60	0.63
SwinIR-SMP	✗	22.29	0.647	0.171	0.130	0.124	66.83	5.17	3.62	0.62
SwinIR-AMO	✓	22.08	0.641	0.167	0.124	0.123	68.08	5.21	3.67	0.66
SwinIR-UPos + FT _{HP}	✗	22.17	0.642	0.166	0.123	0.122	68.38	5.23	3.64	0.65
SwinIR-UPos + FT _{IG}	✗	22.03	0.635	0.168	0.122	0.123	69.37	5.24	3.69	0.66
SwinIR-UPos + FT	✗	22.01	0.633	0.169	0.123	0.124	69.70	5.26	3.70	0.67
SwinIR-AMO + FT	✓	21.77	0.624	0.174	0.121	0.128	70.81	5.29	3.75	0.70
RESRGAN-OrigOnly	✓	22.10	0.618	0.283	0.229	0.185	57.91	4.84	2.99	0.46
RESRGAN-Rand	✓	21.66	0.611	0.234	0.190	0.160	64.82	5.18	3.40	0.60
RESRGAN-UPos*	✗	21.54	0.608	0.233	0.192	0.158	65.93	5.25	3.47	0.63
RESRGAN-SMA	✓	21.46	0.606	0.227	0.182	0.157	65.87	5.23	3.46	0.63
RESRGAN-SMP	✗	21.44	0.607	0.226	0.182	0.156	66.66	5.24	3.51	0.64
RESRGAN-AMO	✓	21.28	0.602	0.224	0.178	0.156	67.86	5.29	3.56	0.66
RESRGAN-UPos + FT _{HP}	✗	21.30	0.595	0.226	0.175	0.158	70.28	5.32	3.65	0.69
RESRGAN-UPos + FT _{IG}	✗	21.14	0.586	0.236	0.182	0.160	72.01	5.35	3.70	0.70
RESRGAN-UPos + FT	✗	21.09	0.580	0.235	0.179	0.163	72.69	5.37	3.69	0.71
RESRGAN-AMO + FT	✓	21.02	0.581	0.228	0.169	0.161	71.67	5.35	3.68	0.71

“origonly”是指没有多模式监督（没有SR增强的GT），“rand”表示随机GT选择（包括SR增强的和原始的），而“upos”表示从“positive”选择。“ft”是指微调（直接优化）；“FT_{HP}”表示使用较高的感知体重（ λ_p ）和“FT_{IG}”包含GAN损失

AdaQual-Diff: Diffusion-Based Image Restoration via Adaptive Quality Prompting

Xin Su

Fuzhou University
China
suxin4726@gmail.com

Chen Wu

University of Science and Technology
of China
China

Yu Zhang

University of the Chinese Academy of
Sciences
China

Chen Lyu

Shandong Normal University
China

Zhuoran Zheng[✉]

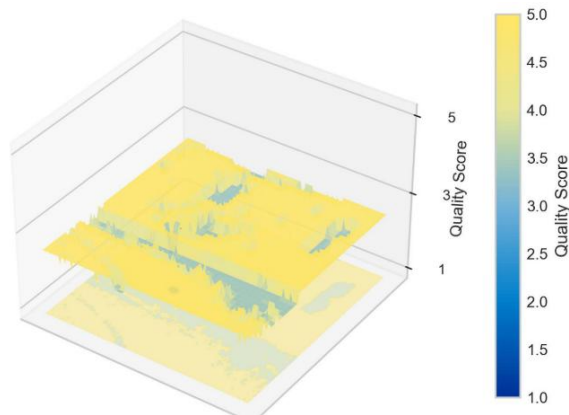
Sun Yat-sen University
China
zhengzr@njust.edu.cn

论文4月份挂在Arxiv上，针对All-In-One图像复原任务，在基于扩散的框架中直接集成感知质量评估用于生成式的恢复过程。

Motivation



Rainy



给定一张类似左图的恶劣天气图像进行图像复原任务，传统的All-in-one方法通常在整个图像上应用统一的处理，给定的退化提示往往是针对全局的情况，比如说：

“Remove raindrops from the image”、“bad, Rainy day, Remove”等，不考虑空间分布和不同区域内退化的不同严重程度。

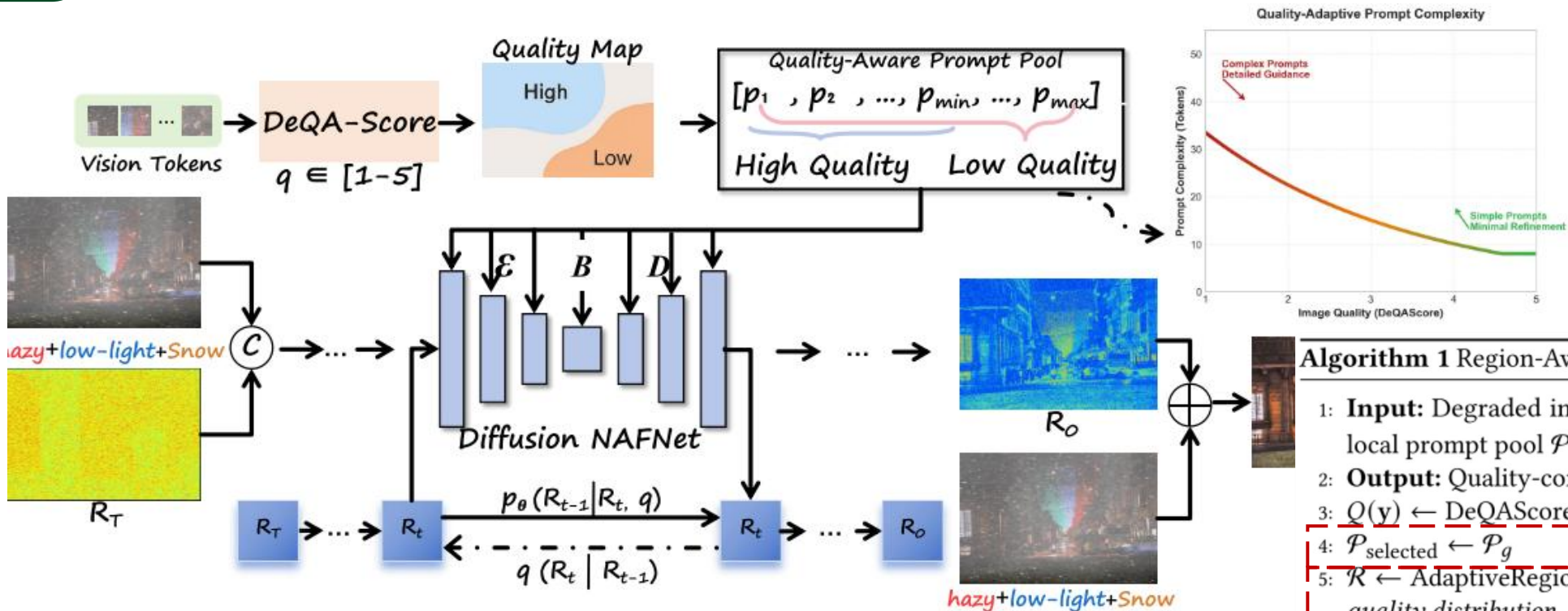
传统的基于扩散模型的方采用的全局调节信号无法充分捕获异质退化的局部统计数据。



造成扩散过程的空间均匀性和真实世界退化的空间变化性之间的**不匹配**；
一般性的先验对局部结构具有不同退化模式的图像仅提供**次优的优化结果**；

Contributions

- 1、建立了一个理论框架，正式连接感知质量评估与扩散模型中的最佳指导复杂性，证明提示复杂性应与局部图像质量成反比，以最大限度地提高恢复效果。
- 2、引入了自适应质量优化，一种新的算法机制，通过动态调制引导扩散过程的文本提示的语义结构和词汇复杂性来实现这种理论关系，从而精确控制特定区域的恢复强度。
- 3、开发了一种双调节策略，在数学上结合了质量驱动的提示与内容特定性信息，创建一个空间上感知的指导场，精确地定位和量化整个图像不同区域的退化程度。



$$\mathbf{x}_{t-1} = \mathcal{F}_{\theta}(\mathbf{x}_t, t, \mathbf{y}, \mathcal{P}(Q(\mathbf{y}))),$$

$$C_p(q) = C_{min} + (C_{max} - C_{min}) \cdot \left(1 - \frac{q - q_{min}}{q_{max} - q_{min}}\right),$$

去噪

计算复杂度

Algorithm 1 Region-Aware Quality-Adaptive Prompt Selection

- 1: **Input:** Degraded image \mathbf{y} , global prompt pool $\mathcal{P}_g \in \mathbb{R}^{L_g \times D}$, local prompt pool $\mathcal{P}_l \in \mathbb{R}^{P \times L_{max} \times D}$, quality threshold τ
- 2: **Output:** Quality-conditional prompt set $\mathcal{P}_{selected}$
- 3: $Q(\mathbf{y}) \leftarrow \text{DeQAScore}(\mathbf{y})$ ▷ Generate quality map
- 4: $\mathcal{P}_{selected} \leftarrow \mathcal{P}_g$ ▷ Initialize with global prompts
- 5: $\mathcal{R} \leftarrow \text{AdaptiveRegionPartition}(Q(\mathbf{y}))$ ▷ Partition based on quality distribution
- 6: **for** each region $r \in \mathcal{R}$ **do**
- 7: $q_r \leftarrow \text{Mean}(Q(r))$ ▷ Calculate regional quality statistic
- 8: $C_p \leftarrow C_{min} + (C_{max} - C_{min}) \cdot (1 - \frac{q_r - q_{min}}{q_{max} - q_{min}})$ ▷ Determine prompt complexity
- 9: $\mathcal{P}_{pool} \leftarrow \begin{cases} \mathcal{P}_{high} & \text{if } q_r > \tau \\ \mathcal{P}_{low} & \text{otherwise} \end{cases}$ ▷ Select appropriate prompt pool
- 10: $\mathcal{P}_r \leftarrow \text{TopK}(\text{Similarity}(\mathcal{P}_{pool}, \mathcal{F}_e(r)), C_p)$ ▷ Select prompts via feature matching
- 11: $\mathcal{P}_{selected} \leftarrow \mathcal{P}_{selected} \cup \{\mathcal{P}_r\}$
- 12: **end for**
- 13: **return** $\mathcal{P}_{selected}$

Loss

为了应对局部质量具有空间多样性的特点，设计了质量权重损失(quality-weighted loss):

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{noise}} + \lambda_1 \mathcal{L}_{\text{quality}} + \lambda_2 \mathcal{L}_{\text{percep}},$$

$$\mathcal{L}_{\text{noise}} = \mathbb{E}_{\mathbf{x}_0, \mathbf{y}, \epsilon, t} [\|\epsilon - \epsilon_{\theta}(\mathbf{x}_t, t, \mathbf{y}, \mathcal{P}(Q(\mathbf{y})))\|_2^2],$$

扩散去噪损失

$$\mathcal{L}_{\text{quality}} = \mathbb{E}_{\mathbf{x}_0, \mathbf{y}, \epsilon, t} \left[\sum_{i,j} w(Q(\mathbf{y})_{i,j}) \cdot \|\epsilon_{i,j} - \epsilon_{\theta}(\mathbf{x}_t, t, \mathbf{y}, \mathcal{P}(Q(\mathbf{y})))_{i,j}\|_2^2 \right] \quad w(q) = (q_{\max} - q) / (q_{\max} - q_{\min})$$

显式空间权重
损失

$$\mathcal{L}_{\text{percep}} = \sum_{r \in \mathcal{R}_{\text{low}}} \sum_l \|\phi_l(\mathbf{x}_0^r) - \phi_l(\mathbf{x}_{\text{pred}}^r)\|_1,$$

选择性区域感
知损失

VGG的第l层特征


































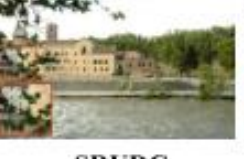





Types	Methods	Venue & Year	PSNR \uparrow	SSIM \uparrow
	Input		16.00	0.6008
One-to-One	MIRNet [58]	ECCV2020	25.97	0.8474
	MPRNet [59]	CVPR2021	25.47	0.8555
	Restormer [57]	CVPR2022	26.99	0.8646
	DGUNet [31]	CVPR2022	26.92	0.8559
	NAFNet [3]	ECCV2022	24.13	0.7964
	SRUDC [39]	ICCV2023	27.64	0.8600
	Fourmer [62]	ICML2023	23.44	0.7885
	OKNet [9]	AAAI2024	26.33	0.8605
One-to-Many	AirNet [23]	CVPR2022	23.75	0.8140
	TransWeather [42]	CVPR2022	23.13	0.7810
	WeatherDiff [33]	TPAMI2023	22.49	0.7985
	PromptIR [34]	NIPS2023	25.90	0.8499
	WGWSNet [66]	CVPR2023	26.96	0.8626
One-to-Composite	OneRestore	ECCV2024	28.47	0.8784
	OneRestore [†]	ECCV2024	28.72	0.8821
	AdaQualDiff (Ours)		30.11	0.9001

CDD-11dataset, 包括单一退化和复合退化

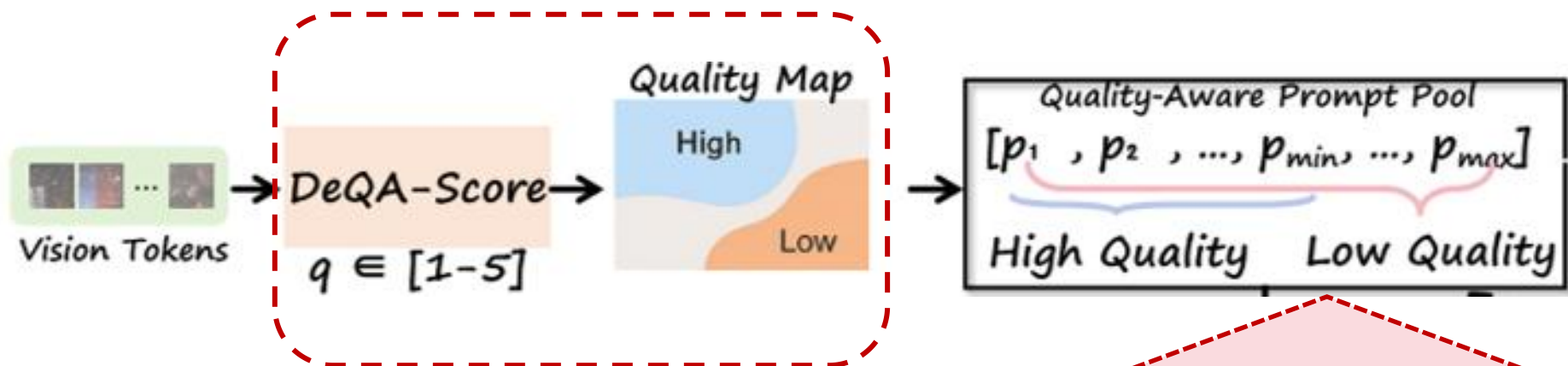
Outdoor-Rain [25]			
CycleGAN	ICCV2017 [65]	17.62	0.6560
pix2pix [18]	ICCV2017 [18]	19.09	0.7100
HRGAN	CVPR2019 [25]	21.56	0.8550
PCNet	TIP2021 [20]	26.19	0.9015
MPRNet	CVPR2021 [59]	28.03	0.9192
NAFNet	ECCV2022 [3]	29.59	0.9027
Restormer	CVPR2022 [57]	29.97 Δ	0.9215 Δ
All-in-One	CVPR2020 [26]	24.71	0.8980
TransWeather	CVPR2022 [42]	28.83	0.9000
TKL&MR	CVPR2022 [8]	29.92	0.9167
WeatherDiff ₆₄	PAMI2023 [32]	29.64	0.9312
WeatherDiff ₁₂₈	PAMI2023 [32]	29.72	0.9216
AWRCP	ICCV2023 [55]	31.39	0.9329
T^3 -DiffWeather	ECCV2024 [6]	31.99	0.9365
AdaQual-Diff (Ours)	-	31.81	0.9370

Snow100K-S [29]			
SPANet	CVPR2019 [44]	29.92	0.8260
JSTASR	ECCV2020 [7]	31.40	0.9012
RESCAN	ECCV2018 [27]	31.51	0.9032
DesnowNet	TIP2018 [29]	32.33	0.9500
DDMSNet	TIP2021 [60]	34.34	0.9445
MPRNet	CVPR2021 [59]	34.97	0.9457
NAFNet	ECCV2022 [3]	34.79	0.9497
Restormer	CVPR2022 [57]	35.03 Δ	0.9487 Δ
All-in-One	CVPR2020 [26]	-	-
TransWeather	CVPR2022 [42]	32.51	0.9341
TKL&MR	CVPR2022 [8]	34.80	0.9483
WeatherDiff ₆₄	PAMI2023 [32]	35.83	0.9566
WeatherDiff ₁₂₈	PAMI2023 [32]	35.02	0.9516
AWRCP	ICCV2023 [55]	36.92	0.9652
T^3 -DiffWeather	ECCV2024 [6]	37.51	0.9664
AdaQualDiff (Ours)	-	37.55	0.9687

RainDrop [35]			
pix2pix	ICCV2017 [18]	28.02	0.8547
DuRN	CVPR2019 [28]	31.24	0.9259
RaindropAttn	ICCV2019 [37]	31.44	0.9263
AttentiveGAN	CVPR2018 [35]	31.59	0.9170
CCN	CVPR2021 [36]	31.34	0.9286
IDT	PAMI2022 [51]	31.87	0.9313
UDR-S ² Former	ICCV2023 [5]	32.64 Δ	0.9427 Δ
All-in-One	CVPR2020 [26]	31.12	0.9268
TransWeather	CVPR2022 [42]	30.17	0.9157
TKL&MR	CVPR2022 [8]	30.99	0.9274
WeatherDiff ₆₄	PAMI2023 [32]	30.71	0.9312
WeatherDiff ₁₂₈	PAMI2023 [32]	29.66	0.9225
AWRCP	ICCV2023 [55]	31.93	0.9314
T^3 -DiffWeather	ECCV2024 [6]	32.66	0.9411
AdaQual-Diff (Ours)	-	32.74	0.9330

17.670/0.700	26.485/0.894	26.675/0.909	27.102/0.899	26.635/0.903	27.078/0.911	21.205/0.837	23.205/0.847	27.221/0.915	PSNR/SSIM
									
16.828/0.487	23.757/0.718	24.021/0.736	26.152/0.736	25.552/0.717	25.652/0.747	23.339/0.591	23.539/0.711	25.898/0.748	PSNR/SSIM
									
14.260/0.569	23.254/0.755	24.122/0.759	23.777/0.755	23.690/0.753	24.264/0.778	20.273/0.670	20.598/0.693	26.378/0.820	PSNR/SSIM
									
14.924/0.550	23.254/0.755	24.486/0.784	25.029/0.777	24.470/0.772	26.092/0.810	20.997/0.662	20.935/0.735	25.108/0.845	PSNR/SSIM
									
Input	MIRNet	Restormer	SRUDC	PromptIR	OneRestore	T ³ -DiffWeather	WeatherDiff	AdaQualDiff (Ours)	GT

low+haze+rain (the upper two examples) and low+haze+snow (the lower two examples)



退化的程度



质量评估



Prompt的选择

```

6: for each region  $r \in \mathcal{R}$  do
7:    $q_r \leftarrow \text{Mean}(Q(r))$   $\triangleright$  Calculate regional quality statistic
8:    $C_p \leftarrow C_{\min} + (C_{\max} - C_{\min}) \cdot (1 - \frac{q_r - q_{\min}}{q_{\max} - q_{\min}})$   $\triangleright$  Determine
      prompt complexity
9:    $\mathcal{P}_{\text{pool}} \leftarrow \begin{cases} \mathcal{P}_{\text{high}} & \text{if } q_r > \tau \\ \mathcal{P}_{\text{low}} & \text{otherwise} \end{cases}$   $\triangleright$  Select appropriate prompt
      pool
10:   $\mathcal{P}_r \leftarrow \text{TopK}(\text{Similarity}(\mathcal{P}_{\text{pool}}, \mathcal{F}_e(r)), C_p)$   $\triangleright$  Select prompts
      via feature matching
11:   $\mathcal{P}_{\text{selected}} \leftarrow \mathcal{P}_{\text{selected}} \cup \{\mathcal{P}_r\}$ 
12: end for
13: return  $\mathcal{P}_{\text{selected}}$ 

```