

# VisualQuality-R1: Reasoning-Induced Image Quality Assessment via Reinforcement Learning to Rank

Tianhe Wu<sup>1,2</sup>, Jian Zou<sup>1</sup>, Jie Liang<sup>2</sup>, Lei Zhang<sup>2,3\*</sup>, and Kede Ma<sup>1\*</sup>

<sup>1</sup>Department of Computer Science, City University of Hong Kong

<sup>2</sup>OPPO Research Institute

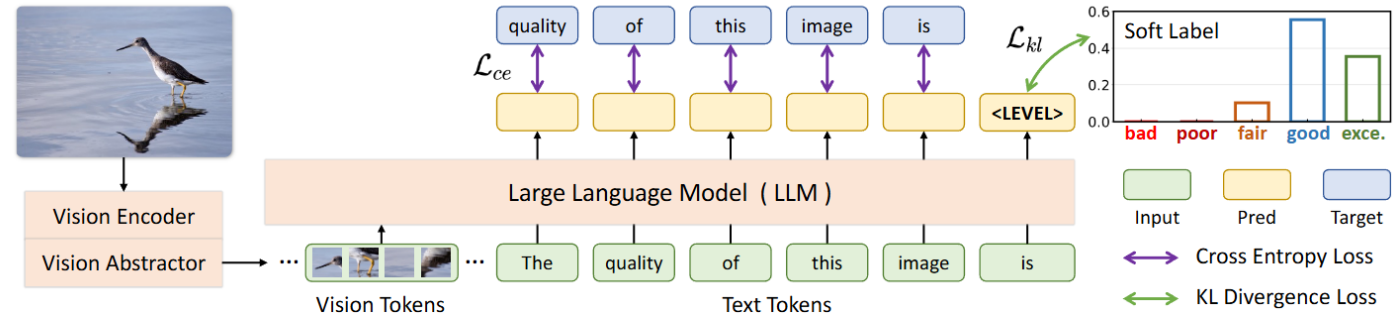
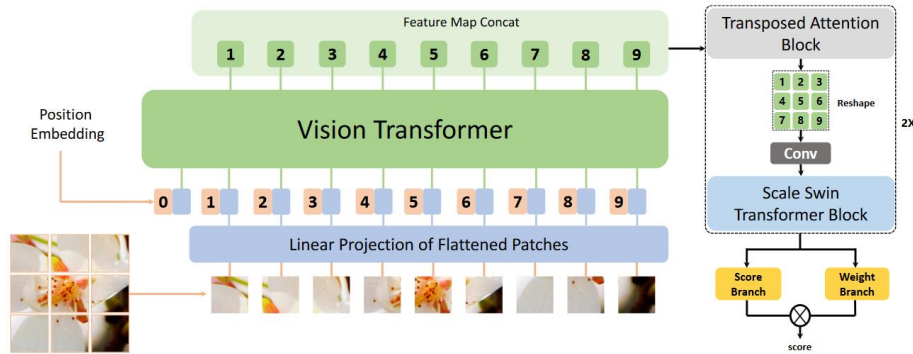
<sup>3</sup>Department of Computing, The Hong Kong Polytechnic University

<https://github.com/TianheWu/VisualQuality-R1>

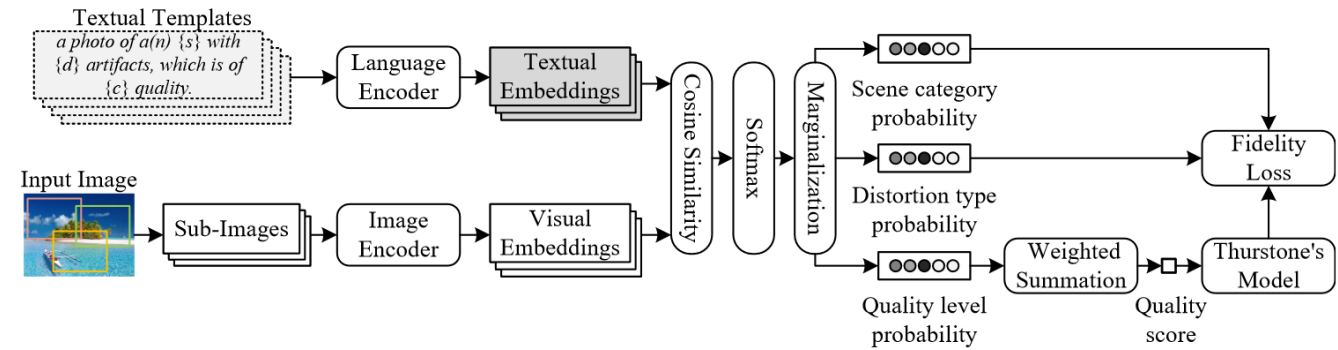
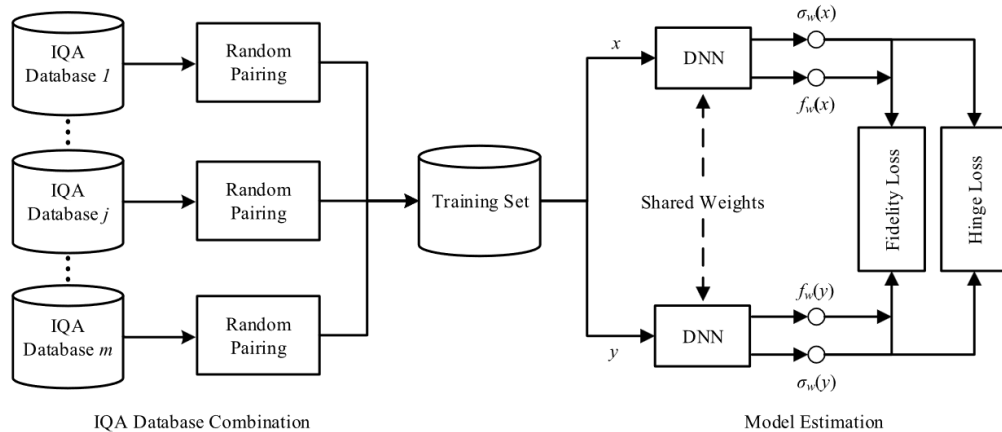


# Evolution of Image Quality Assessment

- Quality Regression with L1 / L2 loss



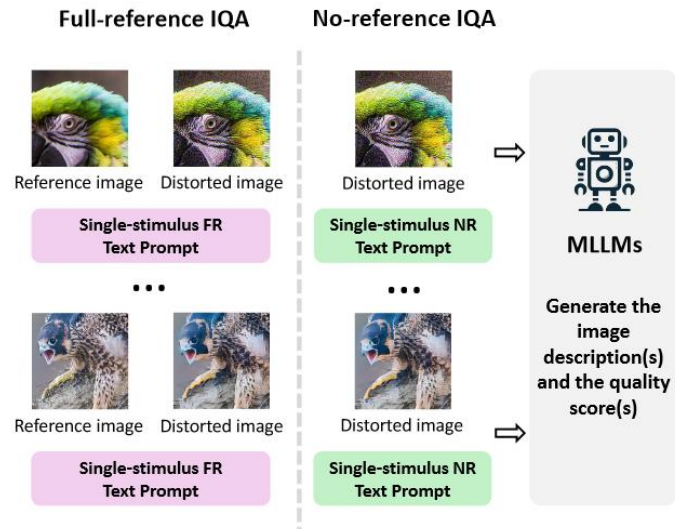
- Quality Ranking with fidelity loss



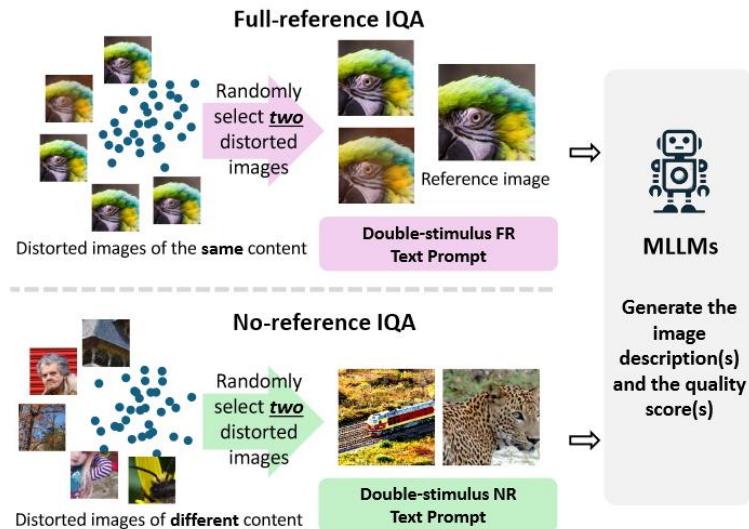
# Quality Perception with Large Vision-Language Models

- Generalization (model backbone and pretraining data)
- **SFT-based** IQA models can not deliver accurate descriptions (catastrophic forgetting)

## (a) Single-stimulus Method



## (b) Double-stimulus Method



## Input Prompt



Please describe this image.

## Answers

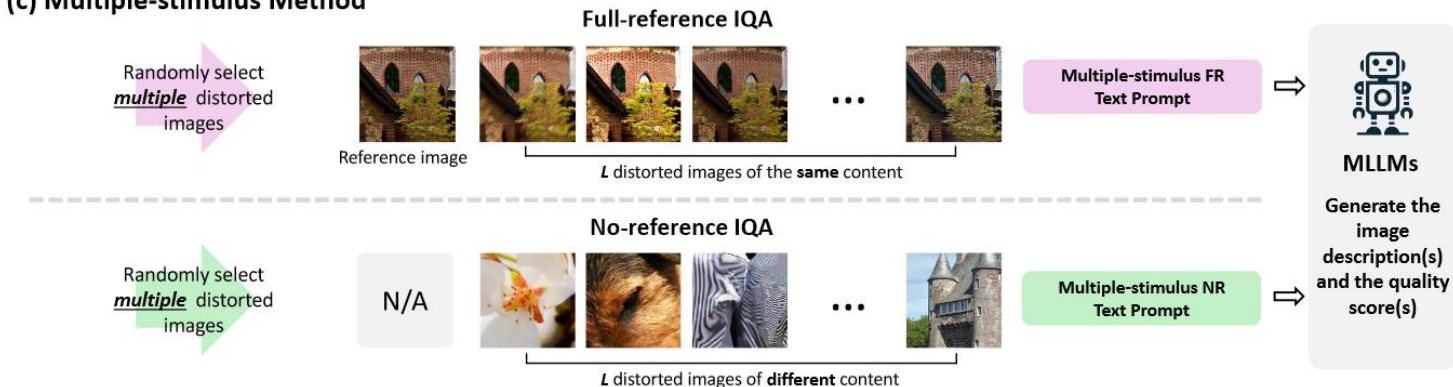
### LLaVA-v1.5 (Vicuna-v1.5-7B)

The image shows a tiger in a natural setting, likely a zoo or wildlife park. The tiger is standing on its hind legs, with its front paws raised as if it's stretching or perhaps in a playful or curious posture. The tiger's fur is a mix of orange and black stripes, typical of the species, and it appears to be looking directly at the camera with a relaxed expression. The background is blurred, but it suggests a grassy area with trees, indicating a habitat designed to mimic the tiger's natural environment. The lighting in the image is bright, suggesting it might be a sunny day.

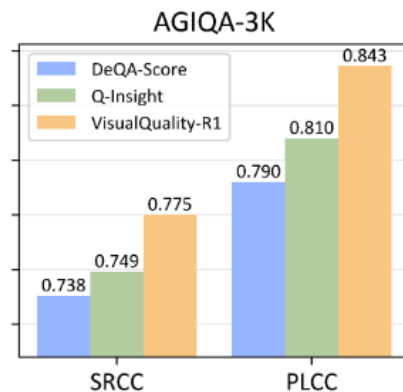
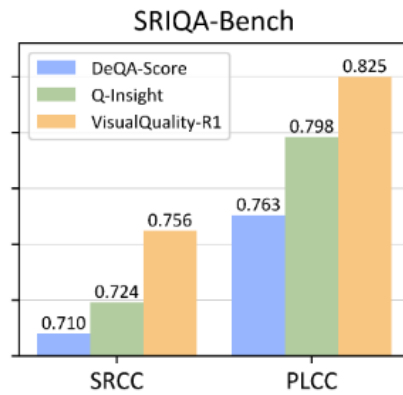
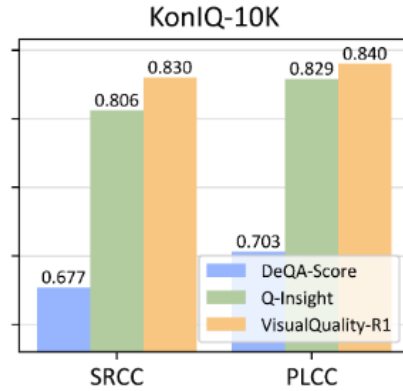
### Q-Instruct (based on LLaVA-v1.5 (Vicuna-v1.5-7B))

y area, and there is a tree in the background. The scene is set in a zoo, where the tiger is likely to be cared for and monitored by zookeepers.

## (c) Multiple-stimulus Method



# SFT vs RL for Quality Model Learning



Please describe the visual quality of this image.



## Q-Instruct



Clear, and the composition is centered. The main subject, a bird in the middle, is clearly defined. **Therefore, the quality is good.**

## VisualQuality-R1



The image is of high quality, featuring a clear and sharp depiction of a bird standing in water with its reflection visible below. The details of the bird's feathers, beak, and legs are **well-defined**, indicating **good focus and lighting**. The water's surface is smooth, creating a clean and symmetrical reflection that enhances the overall composition. The background appears to be a natural setting, possibly a wetland or marsh, **which adds context to the scene without distracting from the main subject**. The colors are vibrant, with the bird's brown and white plumage contrasting nicely against the blue tones of the water. Overall, the image is well-composed and captures the essence of the bird in its natural habitat.



Please describe the visual quality of this image.



## Q-Instruct



**Are rich**, and the **lighting is sufficient**. The texture details are not clear. The composition is centered, and the main subject is clearly the man in the middle. **Therefore, the quality is poor.**

## VisualQuality-R1

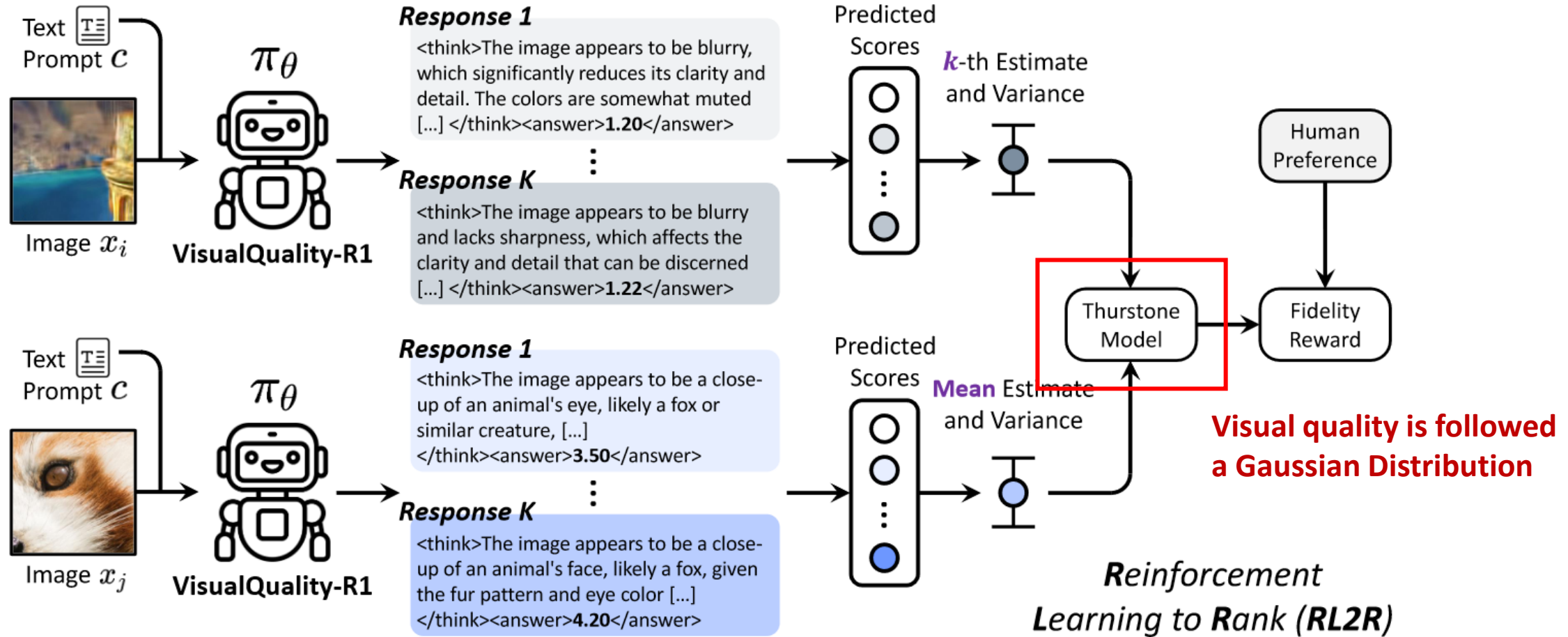


The image appears to be of low quality, with several noticeable issues:

- Blurriness:** The subject and background are not in sharp focus, making it difficult to discern fine details.
- Lighting:** The lighting seems uneven, **with some areas overexposed and others underexposed**, particularly around the subject's face and the microphone.
- Color Balance:** The colors appear somewhat muted or distorted, possibly **due to the lighting conditions or the camera settings**.
- Noise:** There is visible noise or graininess in the image, which can be seen as small specks or dots, **especially in darker areas**.

Overall, the image lacks clarity and detail, which affects its quality. **Improvements could include better focus, more consistent lighting, and potentially adjusting the color balance and reducing noise.**

# VisualQuality-R1 via RL2R



$$p_k(x_i, x_j) = \Phi \left( \frac{q_k(x_i) - \mu(q(x_j))}{\sqrt{\sigma^2(q(x_i)) + \sigma^2(q(x_j)) + \gamma}} \right), \quad \text{for } i \neq j, \quad r_k(x_i) = \frac{1}{B-1} \sum_{j \neq i} \left( \sqrt{p(x_i, x_j)p_k(x_i, x_j)} + \sqrt{(1-p(x_i, x_j))(1-p_k(x_i, x_j))} \right)$$

# Main Results

Table 2: SRCC and PLCC results of NR-IQA models trained on KADID-10K. Exceptions include Q-Insight<sup>†</sup> and VisualQuality-R1<sup>†</sup>, which use a combined training set (KADID-10K and SPAQ). Top two results are highlighted in **bold** and underline, respectively.

Method	Imaging-Related Distortion				Processing-Related Distortion				Avg
	BID	CLIVE	KonIQ	SPAQ	De-blurring	Super-Res.	De-hazing	Image Gen.	
<b>SRCC</b>									
<i>Handcrafted</i>									
NIQE [33]	0.515	0.450	0.421	0.676	0.360	0.557	0.343	0.533	0.482
BRISQUE [32]	0.522	0.314	0.385	0.614	0.389	0.482	0.242	0.497	0.431
<i>Discriminative Deep-Learning-based</i>									
UNIQUE [56]	0.412	0.470	0.649	0.751	0.669	0.649	0.577	0.608	0.598
MUSIQ [19]	0.327	0.284	0.473	0.720	0.656	0.404	0.458	0.494	0.477
MANIQA [52]	0.420	0.487	0.213	0.745	0.726	0.263	0.608	0.422	0.486
<i>VLM-based</i>									
LIQE [57]	0.677	0.719	0.684	0.815	0.797	0.743	<b>0.646</b>	0.653	0.717
Q-Align [49]	0.576	0.554	0.573	0.767	0.761	0.684	0.455	0.682	0.632
DeQA-Score [53]	0.702	0.743	0.677	0.852	0.785	0.710	<u>0.643</u>	0.738	0.731
Qwen2.5-VL-7B [1]	0.711	0.733	0.754	0.848	0.820	0.603	0.458	0.735	0.708
Q-Insight [21]	0.784	0.761	0.806	0.872	0.831	0.724	0.601	0.749	0.766
Q-Insight <sup>†</sup>	<u>0.806</u>	<u>0.804</u>	0.812	<u>0.907</u>	<b>0.846</b>	0.700	0.539	0.657	0.759
VisualQuality-R1	0.790	0.750	<u>0.830</u>	0.875	0.838	<b>0.756</b>	0.598	<b>0.775</b>	<u>0.777</u>
VisualQuality-R1 <sup>†</sup>	<b>0.811</b>	<b>0.811</b>	<b>0.855</b>	<b>0.913</b>	<u>0.845</u>	<u>0.752</u>	0.588	<u>0.754</u>	<b>0.791</b>
<b>PLCC</b>									
<i>Handcrafted</i>									
NIQE [33]	0.527	0.494	0.439	0.683	0.376	0.587	0.482	0.560	0.519
BRISQUE [32]	0.528	0.362	0.400	0.624	0.380	0.556	0.468	0.541	0.482
<i>Discriminative Deep-Learning-based</i>									
UNIQUE [56]	0.385	0.472	0.590	0.708	0.654	0.668	0.578	0.581	0.580
MUSIQ [19]	0.280	0.325	0.435	0.666	0.563	0.441	0.455	0.434	0.450
MANIQA [52]	0.512	0.571	0.257	0.753	0.728	0.243	0.663	0.448	0.522
<i>VLM-based</i>									
LIQE [57]	0.680	0.726	0.652	0.814	0.712	0.775	0.661	0.653	0.709
Q-Align [49]	0.651	0.643	0.612	0.779	0.802	0.713	0.525	0.705	0.679
DeQA-Score [53]	0.743	0.795	0.703	0.858	0.838	0.763	<b>0.688</b>	0.790	0.772
Qwen2.5-VL-7B [1]	0.725	0.760	0.810	0.854	0.852	0.653	0.553	0.810	0.752
Q-Insight [21]	0.796	0.795	0.829	0.872	0.857	0.798	0.669	0.810	0.803
Q-Insight <sup>†</sup>	<u>0.818</u>	<u>0.837</u>	0.809	<u>0.912</u>	0.861	0.779	0.626	0.705	0.793
VisualQuality-R1	0.806	0.794	<u>0.840</u>	0.878	<u>0.872</u>	<b>0.825</b>	0.651	<b>0.843</b>	<u>0.814</u>
VisualQuality-R1 <sup>†</sup>	<b>0.820</b>	<b>0.844</b>	<b>0.870</b>	<b>0.917</b>	<b>0.879</b>	<u>0.824</u>	<u>0.674</u>	<u>0.820</u>	<b>0.831</b>

- Qwen2.5-VL-7B has powerful generalization across multiple scenarios
- Better Generalization under **synthetic data training**
- Do not require perceptual scale realignment

# Ablation Studies and Further Analysis

Table 3: PLCC results of VisualQuality-R1 with varying  $K$  in GRPO. The default setting is highlighted in **bold**.

# of Generated Responses	Imaging-Related Distortion				Processing-Related Distortion				Avg
	BID	CLIVE	KonIQ	SPAQ	De-blurring	Super-Res.	De-hazing	Image Gen.	
$K = 4$	0.805	0.795	0.839	0.875	0.875	0.815	0.643	<b>0.844</b>	0.811
$K = 5$	<b>0.806</b>	<b>0.804</b>	<b>0.840</b>	<b>0.879</b>	0.867	<b>0.826</b>	0.639	0.840	0.813
$K = 6$	<b>0.806</b>	0.794	<b>0.840</b>	0.878	<b>0.872</b>	0.825	<b>0.651</b>	0.843	<b>0.814</b>

Table 4: Comparison of different Thurstone model variants [40] in the GRPO of VisualQuality-R1.

Method	Imaging-Related Distortion				Processing-Related Distortion				Avg
	BID	CLIVE	KonIQ	SPAQ	De-blurring	Super-Res.	De-hazing	Image Gen.	
<b>SRCC</b>									
Q-Insight [21]	0.784	<b>0.761</b>	0.806	0.872	0.831	0.724	0.601	0.749	0.766
Binary Reward	0.780	0.756	0.821	<b>0.877</b>	0.834	0.748	0.587	0.771	0.772
Probability Average (Eq. (7))	0.785	<b>0.761</b>	<b>0.836</b>	0.875	0.835	0.747	0.574	<b>0.775</b>	0.774
Fixed Variance of One	0.778	0.750	0.818	0.871	0.830	0.744	<b>0.606</b>	0.760	0.770
VisualQuality-R1	<b>0.790</b>	0.750	0.830	0.875	<b>0.838</b>	<b>0.756</b>	0.598	<b>0.775</b>	<b>0.777</b>
<b>PLCC</b>									
Q-Insight [21]	0.796	0.795	0.829	0.872	0.857	0.798	<b>0.669</b>	0.810	0.803
Binary Reward	0.790	0.792	0.833	0.876	0.867	<b>0.825</b>	0.646	0.840	0.809
Probability Average (Eq. (7))	0.796	<b>0.797</b>	<b>0.844</b>	0.875	0.861	0.817	0.621	0.831	0.805
Fixed Variance of One	0.791	0.785	0.817	0.873	0.852	0.802	0.655	0.818	0.799
VisualQuality-R1	<b>0.806</b>	0.794	0.840	<b>0.878</b>	<b>0.872</b>	<b>0.825</b>	0.651	<b>0.843</b>	<b>0.814</b>

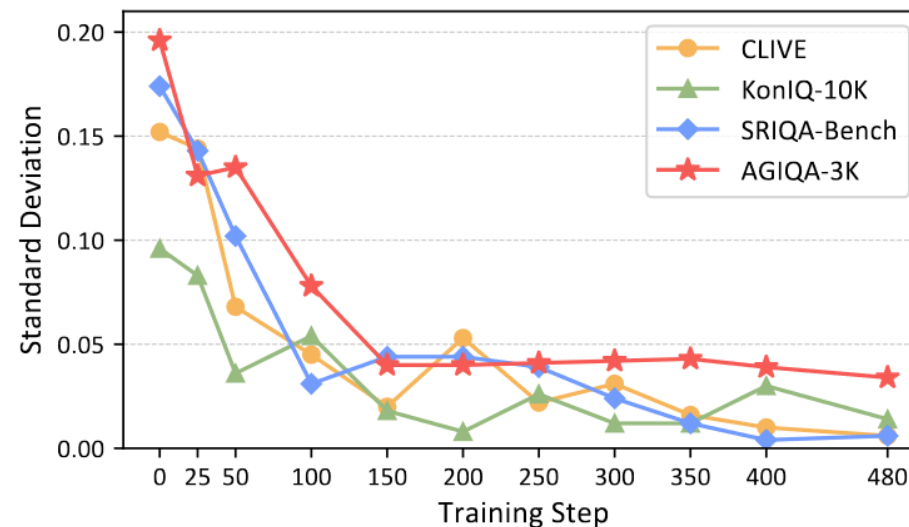


Figure 3: Prediction variability decreases during GRPO. We randomly select 20 images from each of CLIVE [11], KonIQ-10k [15], SRIQA-Bench [6], and AGIQA-3K [20]. At successive training steps, we generate multiple responses per image, compute the std of the predicted quality scores, and plot the average std across images. The uniformly downward trend confirms that VisualQuality-R1 becomes steadily more stable in assessing image quality as training progresses.

# Ablation Studies and Further Analysis



## *Q-Insight*

<think>**The image appears to be somewhat blurry**, which affects the clarity and detail that can be discerned. The colors are vibrant, but the lack of sharpness reduces the overall quality. There's also a slight **overexposure** in some areas, which might be due to lighting conditions or camera settings.</think>  
<answer>2.00</answer>

## *Qwen2.5-VL-7B (0-th Step)*

<think>The image appears to be of high quality with clear details and vibrant colors. The texture and patterns on the objects are well-defined, suggesting good resolution and lighting. **There are no visible distortions or blurriness** that would typically indicate lower quality. The contrast and saturation levels seem appropriate, enhancing the visual appeal without overwhelming the viewer.</think><answer>4.80</answer>

## *VisualQuality-R1 (50-th Step)*

<think>The image appears to be somewhat **abstract or stylized**, which might not be ideal for a realistic quality assessment. However, it is clear and the **details are well-defined**, suggesting good resolution. The colors are vibrant and the contrast is high, which can be visually appealing but may not reflect naturalistic quality. The image does not contain any discernible text or identifiable objects that could provide further context.</think><answer>3.50</answer>

## *VisualQuality-R1 (200-th Step)*

<think>The image appears to have a high level of detail and clarity, with distinct patterns and colors that are well-defined. However, there is a **slight lack of sharpness around the edges**, which might be due to the texture or the way the image was captured. The lighting seems even, and the colors are vibrant, contributing positively to the overall quality.</think><answer>3.80</answer>

## *VisualQuality-R1 (Last Step)*

<think>The image appears to be a close-up of **some kind of fruit or vegetable with a patterned surface**. The colors are vibrant, but the image seems to have a filter applied that gives it a somewhat **surreal appearance**. **The details are not very sharp, which might be due to the nature of the subject or the processing of the image**. However, the saturation and contrast are high, making the image visually striking.</think><answer>3.00</answer>



# Thanks

Tianhe Wu

2025.11.4

[tianhewu-c@my.cityu.edu.hk](mailto:tianhewu-c@my.cityu.edu.hk)