AU-IQA Supplementary Material

1 Dataset Construction Details

- **Data Annotation Interface.** Annotators utilize a dedicated visualization interface to label AI-UGC, shown in Fig. 1, with the annotation outcomes systematically recorded in CSV file.
- **Validation of Annotation Accuracy.** Our quality scores are annotated by five annotators simultaneously, and we calculate the cross-correlation matrix (in Table. 1) to validate the annotations' reliability.

Table 1: The cross-correlation of the scores annotated by the five annotators, represented by PLCC.

	annotator1	annotator2	annotator3	annotator4	annotator5
annotator1	1.000	0.524	0.509	0.568	0.524
annotator2	-	1.000	0.629	0.558	0.722
annotator3	-	-	1.000	0.539	0.663
annotator4	-	-	-	1.000	0.607
annotator5	-	-	-	-	1.000

2 Prompts in SFT Dataset

To fully leverage the performance of LMMs, we need to modify prompts based on the features of the image, ensuring that the features in the prompt correspond as closely as possible to the image.

- For the SFT dataset consisting of UGC images, the prompt is as follows:
  prompt: Please evaluate the overall quality of this user-generated image, considering factors such as noise, blur, compression artifacts, and color accuracy. Provide a score between 1 and 5, with 1 being poor quality and 5 being excellent quality.
- For the SFT dataset consisting of AIGC images, the prompt is as follows:
  prompt: Please assess the visual quality of this AI-generated image, considering factors like clarity, visual coherence, and artistic quality. Provide a score between 1 and 5, where 1 represents low quality and 5 represents high quality.
- For the SFT dataset consisting of both UGC and AIGC images, the prompt is as follows:
  prompt: Please evaluate the visual quality of this image, considering aspects such as clarity, realism, coherence, and overall aesthetic appeal. Provide a score between 1 and 5, where 1 represents poor quality and 5 represents excellent quality.

These prompts are used in fine-tuning the LMMs for directly outputting quality scores. For another fine-tuning method, we replace the response to output scores with to output ratings (excellent, good, fair, poor, bad). For example,

- *Please evaluate the overall quality of this user-generated image, considering factors such as noise, blur, compression artifacts, and color accuracy. Then give your rating on the visual quality of this image. Your rating should be chosen from the following five categories: Excellent, Good, Fair, Poor and Bad.*

Figure 1: Data Annotation Interface.

3 Parameters in Finetuning LMMs

In this paper, we utilize the *ms-swift* to fine-tune LMMs. This experiment was conducted on 8 NVIDIA RTX 4090 (24G) GPUs. The parameter settings are shown in Table. 2.

Table 2: Parameter settings during fine-tuning.

Parameters	***Value***
num_train_epochs	*3*
sft_type	*lora*
batch_size	*2*
learning_rate	*1e-4*
lora_rank	*8*
lora_alpha	*32*
lora_dropout	*0.05*
max_length	*2048*