CoT1: Natural Human Language

Q: <image>Rate the aesthetics and quality of this image step by step.
A: The image presents a captivating scene with an arched entrance that serves as a natural frame, ... The composition is well-thought-out, earning a score of 7, and the overall eye-catching quality is strong, with a score of 8 ... While the subject is slightly cut off, this does not detract significantly from the overall impact. The combination of the well-composed elements, the clear and properly exposed image, and the evocative atmosphere results in a visually pleasing and engaging photograph. **The final score is <6.59>.**

CoT2: Regular Expression Extraction

Q: <image> Rate the aesthetics and quality of this image step by step with the regular form.
A: 1.1 Reason for Composition Score: ..., 1.2 Composition Score: ...
1.3 Reason for Eye-Catching Score: ..., 1.4 Eye-Catching Score: ...
(ignore other attributes for the best view) **The final score is <6.59>.**

(c) AVA Datasets by Incorporating two types of CoT

High Level

Eye-Catching Reason/Score
Composition Reason/Score

Low → *High*

Middle Level

Subject and Background
Subject Integrity
Subject Clutter
Background Clutter
Level Shot

Bad → *Good*

Low Level

Image Clarity
Exposure
Saturation

Poor → *Excellent*

(a) Fine-grained Attributes

(b) Realistic image Quality and Aesthetic dataset

Img A
Eye-Catching Score: 3
Reason for Eye-Catching Score
The photo shows a **messy bed**, which might appeal to those who prefer a lived-in look, but for most people, **a neat bed is more comforting**.
Subject: bed **Subject Clutter**: **cluttered**
Subject Integrity: partially complete (slightly cut off or slightly obscured)

Img B
Eye-Catching Score: 5
Reason for Eye-Catching Score
The photo shows a neat and comfortable hotel room, but lacks standout features or uniqueness.
Subject: bed **Subject Clutter**: **uncluttered**
Subject Integrity: partially complete (slightly cut off or slightly obscured)

	Img A	Img B	
Q-Align	**5.76**	>	5.12 😨
Ours	5.67	<	**6.80** 😄

(d) Application in the Real World