

Peer Review

Review of: "Grounding-IQA: Multimodal Language Grounding Model for Image Quality Assessment"

Shiguang Liu¹

1. Tianjin University, China

This paper proposes a novel image quality evaluation paradigm called Grounding IQA, which combines multimodal indication and localization capabilities with IQA tasks to achieve more refined quality perception and evaluation. The authors designed two subtasks: GIQA-DES, which focuses on providing detailed descriptions of image quality by accurately locating key objects; and GIQA-VQA, a visual Q&A task, which solves Q&A questions about specific area quality attributes.

In this paper, the idea of introducing multimodal indication and localization technology into IQA tasks has been proposed, which solves the limitations of existing methods in fine quality assessment and has high innovative value. They constructed the GIQA-160K dataset and reduced labor costs through automated annotation processes, while ensuring the quality and diversity of annotations.

The writing of this paper is easy to follow.

The scale of the GIQA Bench benchmark is small (only 100 images).

Are there any cases of failure? For example, in complex or blurry images?

Recently, blind image quality assessment has received much attention. Some works in this field are missing. For example,

MTIQA360: An easily trainable multitasking network for blind omnidirectional image quality assessment. In ICPR, 2024

Blind Image Quality Assessment Based on Geometric Order Learning. In CVPR, 2024

Declarations

Potential competing interests: No potential competing interests to declare.