

# Panoptic Pairwise Distortion Graph [ICLR 2026]

Muhammad Kamran Janjua, Abdul Wahab, Bahador Rashidi



**ICLR**  
International Conference On  
Learning Representations



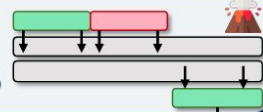
# Motivation



Pairwise Images

Region Query

Multimodal  
large language  
model  
(e.g., Co-Instruct)



```
<sys> Compare two images A:
<img1>, B: <img2> in terms of its
regions <reg1>...<regN> and for same
regions in each image answer (i)
which image is better, (ii) what
distortion is present, (iii) what
is the severity of the distortion,
(iv) and assign a quality score...
```

## Answer (Region Query):

- Person:** The image has a moderate level of distortion, with pixelation and blurring. The colors appear to be washed out, and there is a loss of detail in the person's face and clothing. The quality is average. ❌
- Sheep:** The image has significant distortion, with pixelation and blurring that make it difficult to discern fine details. The colors are also somewhat muted. The quality of this image is poor.
- Sky:** ? Missed
- Mountain:** Image has minimal distortion, with a clear mountain landscape in the background. The quality is good with no noticeable distortion.

No Comparative Assessment/Score & Wrong Answer!

## SFT Query

```
The first image <img>\n The
second image <img>Which image
has better quality, and why?
```

The first image has accurate exposure, rich colors, and sufficient lighting, which contribute to a clearer and more detailed image. Although the focus is not accurate, resulting in some distortion and blurriness, the main subject is still recognizable. On the other hand, the second image has poor lighting, dark colors, and a lot of noise, which make the image less clear and detailed. The main subject is also blurry, and the texture details are lost, making it difficult to identify the subject. **Answer (SFT Query)** ❌

No Grounding & No Region Information!

> This trend is visible in frontier LLMs (GPT, Gemini), open-source general-purpose and distortion-specific LLMs

> With SFT on distortion data, the open-source LLMs much worse  
> parrot template responses

## What To Do?

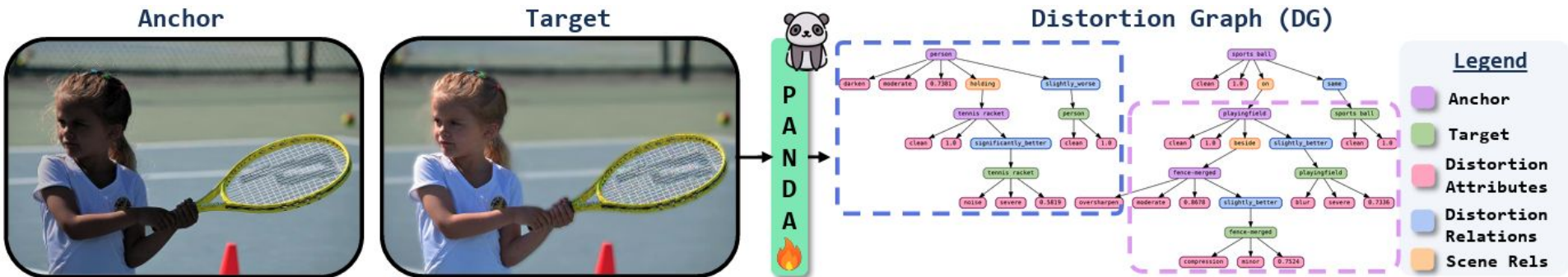
> Can we represent two images in terms of their regions in a structured manner?

> MLLMs struggle with region-level understanding for distortion analysis

> Even when given explicit regions cues (Set-of-Mark, BBoxes, Masks), they still cannot understand regions

> This worsens when task is to compare two images in terms of their regions

# Distortion Graph: A New Task

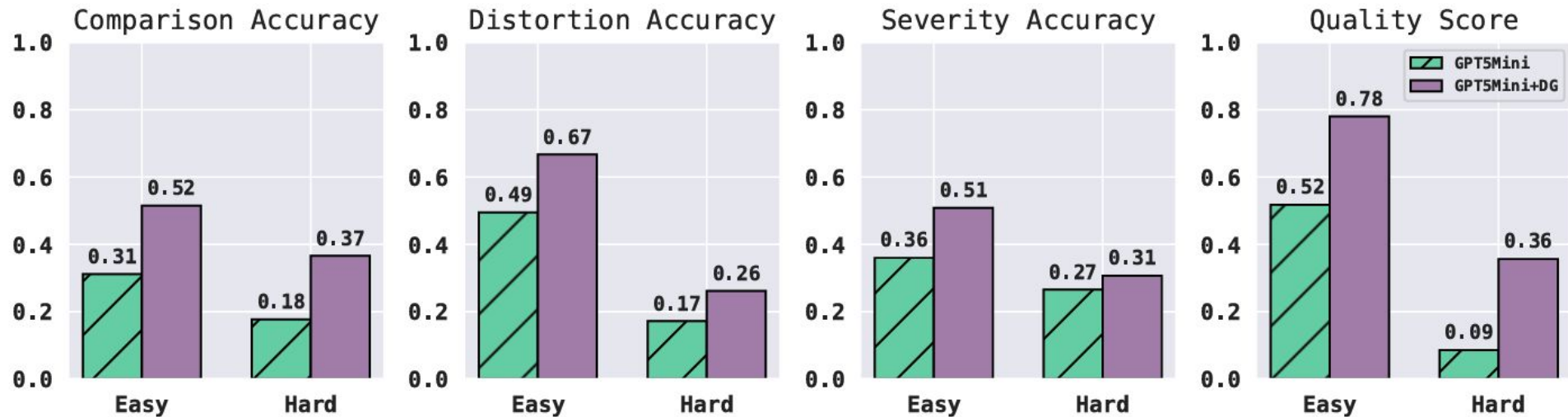


- > A topological structure grounded in regions -- between two images under comparison
- > "nodes" denote regions, "edges" between nodes denote comparative relations, and distortion signals are node "attributes"
- > distortion signals: score, distortion type, severity of distortion, etc.





# DG → Emergent Distortion Understanding



> GPT-5 Mini On Its Own Struggles [Green Bars]

> GPT-5 Mini + Predicted Distortion Graph (DG) → performance improves [Purple Bars]

> GPT-5 Mini + Predicted DG also Encourages Reasoning about Distortion [Distortion Understanding]

# DG → Emergent Distortion Understanding

Ground Truth

Anchor

Target

Region on Anchor



Predicted DG (Hint to GPT-5 Mini)

```
{“anchor”: {“distortion”: “clean”,  
“severity”: 0, “score”: 0.9678}, “target”:  
{“distortion”: “brightness”, “severity”: 1,  
“score”: 0.9241}, “comparison”: “same”}
```

GPT-5 Mini Prediction

```
{“anchor”: {“distortion”: “darken”,  
“severity”: 2, “score”: 0.8915}, “target”:  
{“distortion”: “brightness”, “severity”: 1,  
“score”: 0.9241}, “comparison”: “same”}
```

Predicted DG (Hint to GPT-5 Mini)

```
{“anchor”: {“distortion”:  
“clean”,...}, “target”:  
{“distortion”: “clean”,...},  
“comparison”: “same”}
```

GPT-5 Mini Prediction

```
{“anchor”: {“distortion”:  
“oversharpener”,...}, “target”:  
{“distortion”: “clean”,...},  
“comparison”: “significantly_worse”}
```



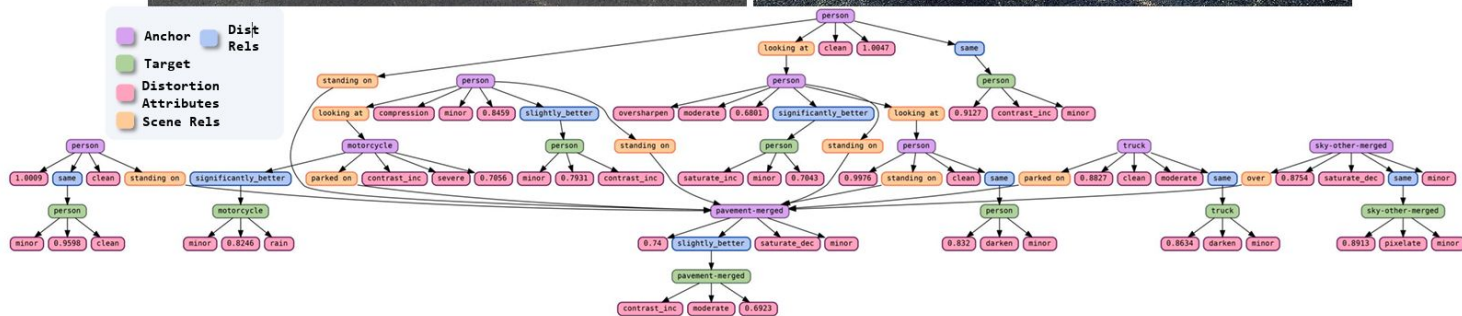
- > GPT-5 Mini uses Predicted DG as a hint to reason about region-wise comparisons
- > If pixel evidence disagrees, GPT-5 Mini discards DG, otherwise uses it and reasons with it

# Distortion Graph → General-Purpose Formalism

> Lots of problems are comparative in nature

> Human decisions are comparative in nature

> DG offers a general-purpose formalism to impose grounded structure on those comparative tasks



> A Complicated Example

> Many Regions Each with Varying Distortions & Severity

# Panoptic Pairwise Distortion Graph [ICLR 2026]

Muhammad Kamran Janjua, Abdul Wahab, Bahador Rashidi

**Paper:** <https://openreview.net/pdf?id=VDfF7NqJJ1>

**Code:** [github.com/AISmartPerception/distortion-graphs](https://github.com/AISmartPerception/distortion-graphs)

**Dataset:** [huggingface.co/datasets/kjanjua26/pandabench](https://huggingface.co/datasets/kjanjua26/pandabench)



**ICLR**  
International Conference On  
Learning Representations

