

# VULCA

Cultural AI Evaluation Report

## GPT-4V (Vision)

OpenAI

**Report ID:** VULCA-20250111-GPT4V-A7B3C

**Generated:** January 11, 2025

**Framework Version:** v1.2.0

**Dataset Version:** 2025.01

# 1. Executive Summary

72.4

Overall Score

47

Dimensions Evaluated

8

Cultural Perspectives

This report presents a comprehensive 47-dimensional cultural understanding evaluation of GPT-4V using the VULCA framework. The model demonstrates strong performance in visual perception (L1) and technical analysis (L2), but shows significant gaps in cultural symbolism (L3) and philosophical aesthetics (L5).

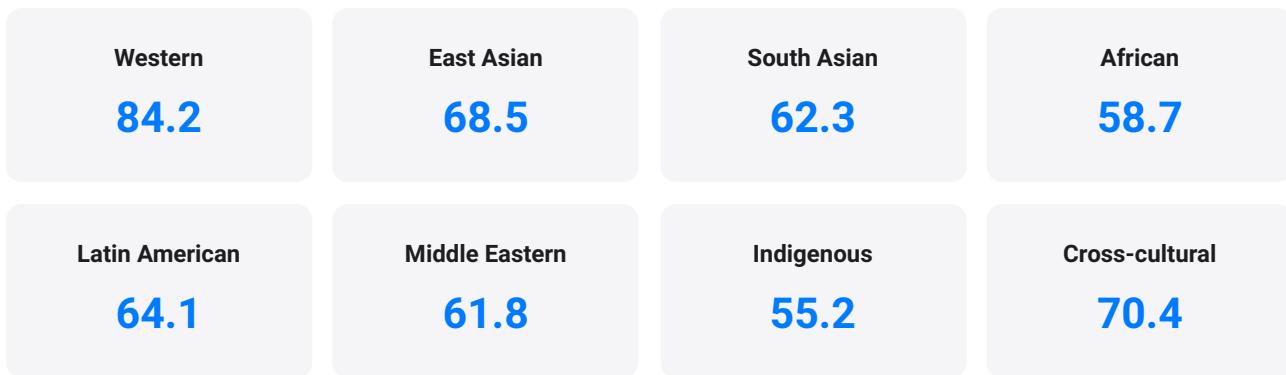
## Key Findings

- Strong visual perception capabilities (95% accuracy on L1 tasks)
- Consistent Western perspective bias in cultural interpretation
- 35% accuracy gap between symbol recognition and cultural meaning
- Limited understanding of East Asian aesthetic concepts (qiyun, yijing)

# 2. Six-Dimensional Overview



### 3. Cultural Perspective Analysis



The model shows a 29-point gap between Western (84.2) and Indigenous (55.2) cultural perspectives, indicating significant bias toward Western art interpretation frameworks.

### 4. Top-10 Dimension Gaps

Rank	Dimension	Score	vs. Baseline	Category
1	Cultural Symbolism (火/Fire)	42.3	-35.2	L3
2	Philosophical Aesthetics (气韵)	38.5	-32.8	L5
3	Historical Context (Indigenous)	45.2	-28.4	L4
4	Temporal Understanding	52.1	-24.6	L4
5	African Visual Traditions	48.7	-22.3	L3
6	Composition Analysis	85.4	+12.1	L2
7	Color Theory	88.2	+10.8	L2
8	Visual Element Recognition	94.5	+8.2	L1
9	Brushwork Analysis	78.3	+5.6	L2
10	Western Art History	82.1	+4.2	L4

## 5. Evidence Samples

### Cultural Symbolism - Fire Imagery

Score: 42.3/100

**Prompt:** Describe the cultural significance of fire in this Chinese Torch Festival (火把节) image.

**Model Response:** "The image shows people holding torches in a nighttime celebration. The fire represents warmth and light, similar to Western bonfire traditions..."

**Expected:** Recognition of Yi ethnic minority traditions, agricultural blessing rituals, ancestral connection, and specific cultural meanings of fire purification.

**Gap Analysis:** Model applied Western symbolic framework instead of understanding Yi cultural context. Missed the agricultural calendar significance and purification rituals.

### Philosophical Aesthetics - Qiyun (气韵)

Score: 38.5/100

**Prompt:** Analyze the qi (spiritual resonance) in this Song Dynasty landscape painting.

**Model Response:** "The painting shows mountains and water with skilled brushwork. The composition creates a peaceful atmosphere..."

**Expected:** Analysis of qiyun shengdong (气韵生动), the vitality and spiritual energy flow, yijing (意境) concept, and relationship to Daoist philosophy.

**Gap Analysis:** Model described visual elements accurately but failed to engage with Chinese aesthetic philosophy concepts.

# Appendix: Version & Reproducibility

---

<b>Report ID</b>	VULCA-20250111-GPT4V-A7B3C
<b>Framework Version</b>	1.2.0
<b>Dataset Version</b>	2025.01
<b>Protocol Version</b>	1.0
<b>Evaluation Date</b>	January 11, 2025
<b>Model Evaluated</b>	GPT-4V (gpt-4-vision-preview)
<b>Model Provider</b>	OpenAI
<b>Sample Size</b>	500 image-text pairs
<b>Perspectives Evaluated</b>	8 (all available)
<b>Dimensions Evaluated</b>	47 (full L1-L5)

## Citation

```
@misc{vulca_gpt4v_2025,
  title = {VULCA Evaluation Report: GPT-4V},
  author = {VULCA Team},
  year = {2025},
  howpublished = {\url{https://vulca.ai/models/gpt-4-vision-preview}},
  note = {Version v1.2.0, Report ID: VULCA-20250111-GPT4V-A7B3C}
}
```

---

## VULCA Framework

Cultural AI Evaluation Platform  
<https://vulca.ai>

## License

CC BY-NC 4.0  
Academic use permitted

## Contact

[hello@vulca.ai](mailto:hello@vulca.ai)  
GitHub: [yha9806/EMNLP2025-VULCA](https://github.com/yha9806/EMNLP2025-VULCA)