

Peer Review

Review of: "Vision-Language In-Context Learning Driven Few-Shot Visual Inspection Model"

Leanne Nortje¹

1. Independent researcher

Summary

The paper presents a novel approach to improving in-context learning (ICL) capabilities in vision-language models (VLMs). While the method appears innovative, the distinction between this approach and prior VLM-based ICL enhancements is not clearly articulated. The authors demonstrate improved performance, but further justification is needed to explain why their method outperforms existing alternatives.

Strengths

Novelty: The approach appears to introduce a unique enhancement for ICL in VLMs.

Performance Gains: The method shows improvements over baselines, indicating its effectiveness.

Relevance: Addresses a crucial challenge in improving ICL for VLMs, making the work valuable to the field.

Weaknesses

Lack of Comparative Justification: The paper does not explicitly differentiate its approach from other VLM methods designed to improve ICL. It is unclear why this method works better.

Insufficient Model and Selection Details: More details should be provided on the model architecture and the selection algorithm to better understand the design choices and their impact.

Poor Grammar and Clarity: Many sentences are not written in full, making the paper difficult to read and reducing its overall clarity.

Declarations

Potential competing interests: No potential competing interests to declare.