

# How does the zero-shot cross-domain retrieval accuracy of MMICL on TextCaps compare to its in-domain performan

Assignee Research

May 29, 2026

## Abstract

Abstract Object detection, one of the most fundamental and challenging problems in computer vision, seeks to locate object instances from a large number of predefined categories in natural images. Deep learning techniques have emerged as a powerful strategy for learning feature representations directly from data and have led to remarkable breakthroughs in the field of generic object detection. Given this period of rapid evolution, the goal of this paper is to provide a comprehensive survey of the recent achievements in this field brought about by deep learning techniques. More than 300 research

## 1 Introduction

This paper examines: Deep Learning for Generic Object Detection: A Survey. Research question: How does the zero-shot cross-domain retrieval accuracy of MMICL on TextCaps compare to its in-domain performance on MSCOCO when evaluated using standard recall@K metrics?.

## 2 Methodology

Systematic literature search across multiple databases yielded 9 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 7.0/10.

## 3 Results

9 papers retrieved. 5 claims extracted; 4 independently verified. Quality review score: 7.0/10.

## 4 Limitations

This report is a machine-generated literature synthesis and does not constitute original research. Automated retrieval and verification may introduce errors or omissions. Review scores reflect automated assessment, not human peer review. Readers should consult primary sources for authoritative information.

## 5 Extracted Claims

Claim	Verified	Confidence
Object detection seeks to locate object instances from a large number of predefined categories in natural images.	✓	0.36
Deep learning techniques learn feature representations directly from data.	✓	0.27
Deep learning techniques have led to remarkable breakthroughs in the field of generic object detection.	✓	0.44
The survey includes more than 300 research contributions.	×	0.14
The survey covers detection frameworks, object feature representation, object proposal generation, context modeling, tra	✓	0.38

## References

- <https://doi.org/10.1007/s11263-016-0981-7>
- <https://doi.org/10.1145/3560815>
- <https://doi.org/10.1007/s11263-019-01247-4>