

# Continuous Contrastive Pretraining with Distributionally Robust Optimization for Cross-Modal Retrieval

Assignee Research

June 9, 2026

## Abstract

This report synthesises findings from 14 peer-reviewed papers addressing the following research question: Can contrastive learning combined with distributionally robust optimization improve the cross-modal retrieval performance of multimodal models evaluated on the COCO text-image retrieval benchmark. 10 claims were extracted from source literature; 2 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 5.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: COCO-DR: Combating Distribution Shifts in Zero-Shot Dense Retrieval with Contrastive and Distributionally Robust Learning. Research question: Can contrastive learning combined with distributionally robust optimization improve the cross-modal retrieval performance of multimodal models evaluated on the COCO text-image retrieval benchmark?.

## 2 Methodology

Systematic literature search across multiple databases yielded 14 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 5.2/10.

## 3 Results

14 papers retrieved. 10 claims extracted; 2 independently verified. Quality review score: 5.2/10.

## 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
Dense retrieval has shown strong advantages over sparse retrieval methods, although the advantages are more observed in	×	0.12
ANCE initialized from coCondenser still underperforms BM25 on BEIR tasks where distribution shifts are severe.	×	0.07
COCO-DR introduces two training techniques: COntinuous COntrastive pretraining (COCO) and implicit Distributionally Robu	✓	0.16
COCO addresses the challenge of distribution shifts via continuously pretraining the language model on the target corpora	✓	0.19
COCO-DR Base model achieves an nDCG@10 score of 0.789 on the TREC-COVID dataset.	×	0.10
COCO-DR Large model achieves an nDCG@10 score of 0.771 on the TREC-COVID dataset.	×	0.09
COCO-DR Base model without iDRO achieves an nDCG@10 score of 0.763 on the TREC-COVID dataset.	×	0.10
COCO-DR Base model achieves an nDCG@10 score of 0.429 on the BioASQ dataset.	×	0.10
COCO-DR Large model achieves an nDCG@10 score of 0.424 on the BioASQ dataset.	×	0.10
COCO-DR Base model without iDRO achieves an nDCG@10 score of 0.353 on the BioASQ dataset.	×	0.11

## References

- <http://arxiv.org/abs/2601.18190v1>
- <http://arxiv.org/abs/2210.15212v2>

- <http://arxiv.org/abs/2406.15306v1>