SOVEREIGN: How do vision-language models like CLIP and MedSAM compare in anomaly localization accuracy on out-of-distribu

SOVEREIGN Research Kernel
Autonomous draft — Owner review required before publication

May 28, 2026

Abstract

Deep convolutional neural networks have performed remarkably well on many Computer Vision tasks. However, these networks are heavily reliant on big data to avoid overfitting. Overfitting refers to the phenomenon when a network learns a function with very high variance such as to perfectly model the training data. Unfortunately, many application domains do not have access to big data, such as medical image analysis. This survey focuses on Data Augmentation, a data-space solution to the problem of limited data. Data Augmentation encompasses a suite of techniques that enhance the size and quality

1 Introduction

Analysis of: A survey on Image Data Augmentation for Deep Learning. Research goal: How do vision-language models like CLIP and MedSAM compare in anomaly localization accuracy on out-of-distribution brain MRI data from different scanner manufacturers when evaluated against the NOVA benchmark?.

2 Methodology

Multi-query arXiv search (4 parallel queries, Relevance-sorted). TF-IDF cosine semantic verification (bigrams, threshold=0.15). NIM nv-embedqa-e5-v5 (dim=1024) for semantic indexing. Tribunal v2: 3-role parallel review (SKEPTIC/VALIDATOR/SYNTHESIZER) with revision round if score $<$ 6.5.

3 Results

9 papers retrieved. 8 claims extracted, 8 verified. Tribunal: 7.5/10 → APPROVE (revision_round=0). Policy: AUTO_APPROVE.

4 Uncertainties

NIM free tier latency varies. TF-IDF verification is a weak signal. arXiv Relevance ranking is query-dependent. Tribunal consensus is LLM-based and prompt-sensitive.

5 Extracted Claims

Claim	Verified	Confidence
Deep convolutional neural networks have performed remarkably well on many Computer Vision tasks	✓	0.26
These networks are heavily reliant on big data to avoid overfitting	✓	0.26
Overfitting refers to the phenomenon when a network learns a function with very high variance such as to perfectly model	✓	0.29
Many application domains do not have access to big data, such as medical image analysis	✓	0.25
Data Augmentation encompasses a suite of techniques that enhance the size and quality of training datasets	✓	0.31
The image augmentation algorithms discussed in this survey include geometric transformations, color space augmentations,	✓	0.49
The application of augmentation methods based on GANs are heavily covered in this survey	✓	0.27
This survey will present existing methods for Data Augmentation, promising developments, and meta-level decisions for im	✓	0.36

References

- https://doi.org/10.1186/s40537-019-0197-0
- https://doi.org/10.1186/s40537-023-00727-2
- https://doi.org/10.18653/v1/2020.coling-main