

SOVEREIGN: Can domain-generalized semantic segmentation models maintain $>90\%$ mIoU performance when deployed on edge device

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: Can domain-generalized semantic segmentation models maintain $>90\%$ mIoU performance when deployed on edge devices with $<100\text{ms}$ inference latency?.

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

7 papers retrieved. 9 claims extracted, 8 verified. Tribunal: 7.8/10 \rightarrow APPROVE (revision_round=1). 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	✓	0.23
Data Augmentation encompasses a suite of techniques that enhance the size and quality of training datasets	✓	0.30
This survey focuses on Data Augmentation as a data-space solution to the problem of limited data	✓	0.30
Image data augmentation techniques discussed in this survey include geometric transformations, color space augmentations	✓	0.45
Deep convolutional neural networks are heavily reliant on big data to avoid overfitting	✓	0.29
This survey covers the application of augmentation methods based on GANs	✓	0.18
Many application domains do not have access to big data	✓	0.21
Data Augmentation can be used to improve Deep Learning model performance	×	0.15
Overfitting occurs when a network learns a function with very high variance to perfectly model the training data	✓	0.22

References

- <https://doi.org/10.1109/cvpr.2016.91>
- <https://doi.org/10.1186/s40537-019-0197-0>
- <https://doi.org/10.1109/iccv.2019.00219>