

SOVEREIGN: Can AlphaX framework be adapted to improve sample efficiency in neural architecture search by incorporating un

SOVEREIGN Research Kernel

Autonomous draft — Owner review required before publication

May 29, 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 AlphaX framework be adapted to improve sample efficiency in neural architecture search by incorporating uncertainty quantification through Bayesian neural networks with fewer Monte Carlo samples?.

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

8 papers retrieved. 20 claims extracted, 10 verified. Tribunal: 6.7/10
\$\\rightarrow\$ REWISE (revision_round=1). Policy: ESCALATE_TO_OWNER.

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 are heavily reliant on big data to avoid overfitting.	✓	0.31
Overfitting is defined as a phenomenon where a network learns a function with very high variance to perfectly model the	✓	0.25
Medical image analysis is an application domain that often lacks access to big data.	✓	0.17
Data Augmentation is a data-space solution to the problem of limited data.	✓	0.25
Data Augmentation encompasses techniques that enhance the size and quality of training datasets.	✓	0.27
The survey discusses geometric transformations as an image augmentation algorithm.	×	0.11
The survey discusses color space augmentations as an image augmentation algorithm.	×	0.14
The survey discusses kernel filters as an image augmentation algorithm.	×	0.11
The survey discusses mixing images as an image augmentation algorithm.	×	0.11
The survey discusses random erasing as an image augmentation algorithm.	×	0.11
The survey discusses feature space augmentation as an image augmentation algorithm.	×	0.14
The survey discusses adversarial training as an image augmentation algorithm.	×	0.13
The survey discusses generative adversarial networks (GANs) as an image augmentation algorithm.	✓	0.16
The survey discusses neural style transfer as an image augmentation algorithm.	×	0.14
The survey discusses meta-learning as an image augmentation algorithm.	✓	0.18
The survey heavily covers the application of augmentation methods based on GANs.	✓	0.17
The survey discusses test-time augmentation as a characteristic of Data Augmentation.	✓	0.15
The survey discusses resolution impact as a characteristic of Data Augmentation.	×	0.14
The survey discusses final dataset size as a characteristic of Data Augmentation.	✓	0.17
The survey discusses curriculum learning as a characteristic of Data Augmentation.	×	0.14

References

- <https://doi.org/10.1186/s40537-019-0197-0>
- <https://doi.org/10.1609/aaai.v35i12.17233>
- <https://openalex.org/W3163842339>