KANs vs. Transformers on ImageNet-1K Under Matched Computational Budgets

Assignee Research

June 4, 2026

Abstract

This report synthesises findings from 12 peer-reviewed papers addressing the following research question: What is the accuracy gap between KANs and transformers on the ImageNet-1K benchmark when trained with identical computational budgets. 9 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.5/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: KKANs: Kurkova-Kolmogorov-Arnold Networks and Their Learning Dynamics. Research question: What is the accuracy gap between KANs and transformers on the ImageNet-1K benchmark when trained with identical computational budgets?.

2 Methodology

Systematic literature search across multiple databases yielded 12 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 3.5/10.

3 Results

12 papers retrieved. 9 claims extracted; 0 independently verified. Quality review score: 3.5/10.

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
cKANs could not handle a significantly higher number of parameters without degrading performance, thereby limiting their	×	0.04
The model's performance was evaluated on a 256 $\times$ 256 uniform grid, and training was performed using 10,000 points sample	×	0.02
The KKAN model achieves a relative L2 error of 5.86 $\times$ 10-3 for the discontinuous function approximation.	×	0.06
The function exhibits a wide range of magnitudes, with outputs spanning from -5 to 25.	×	0.02
KKANs demonstrated significantly better performance on the testing dataset, achieving a relative L2 error of 5.86 $\times$ 10-3	×	0.05
MLPs were slightly faster, averaging 2.36 ms per iteration compared to 2.64 ms for cKANs and 2.77 ms for KKANs.	×	0.03
KKANs converge significantly faster than MLP and cKAN.	×	0.03
The training process minimizes the data residuals: rD(x, θ) = u(x, θ) - u(x), x $\in$ D.	×	0.04
The corresponding loss function is: L = LD = N X i=1 λD,ir2 D(xi, θ), xi $\in$ D.	×	0.03

References

- http://arxiv.org/abs/2601.03290v1
- http://arxiv.org/abs/2405.12832v2
- http://arxiv.org/abs/2412.16738v1