

# Resolution Diversity in Fine-T2I Datasets and Stable Diffusion Perceptual Quality

Assignee Research

June 9, 2026

## Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: How does the resolution diversity of Fine-T2I fine-tuning datasets impact the perceptual quality of Stable Diffusion-generated images as measured by Inception Score (IS) compared to other. 5 claims were extracted from source literature; 1 was independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 4.7/10. This report is a machine-generated literature synthesis and does not constitute original research.

## 1 Introduction

This paper examines: Fine-T2I: An Open, Large-Scale, and Diverse Dataset for High-Quality T2I Fine-Tuning. Research question: How does the resolution diversity of Fine-T2I fine-tuning datasets impact the perceptual quality of Stable Diffusion-generated images as measured by Inception Score (IS) compared to other text-to-image models?.

## 2 Methodology

Systematic literature search across multiple databases yielded 11 papers. Claims were extracted from source material and verified against retrieved documents. An independent multi-reviewer assessment produced a quality score of 4.7/10.

## 3 Results

11 papers retrieved. 5 claims extracted; 1 independently verified. Quality review score: 4.7/10.

## 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
Fine-tuning LlamaGen on Fine-T2I achieves an 80.7% win rate for visual quality compared to a counterpart without further	×	0.13
Fine-tuning LlamaGen on Fine-T2I achieves a 65.3% win rate for text-image alignment.	✓	0.16
Fine-T2I boosts the visual quality the most, likely due to the extremely high aesthetic quality of the dataset.	×	0.12
Models fine-tuned on Fine-T2I generate much better results than those fine-tuned on T2I-2M and BLIP3o-60k datasets on bo	×	0.12
Fine-T2I enhances controllability on standard automatic protocols as shown by improvements in GenEval for both LlamaGen	×	0.04

## References

- <http://arxiv.org/abs/1406.2018v1>
- <http://arxiv.org/abs/2506.10825v1>
- <http://arxiv.org/abs/2602.09439v1>