Computational Efficiency of FRD and FID in High-Dimensional Medical Imaging Datasets

Assignee Research

June 9, 2026

Abstract

This report synthesises findings from 11 peer-reviewed papers addressing the following research question: What is the impact of dataset size and dimensionality on the computational efficiency of FRD compared to perceptual metrics like FID when applied to medical imaging tasks. 16 claims were extracted from source literature; 0 were independently verified against retrieved documents. An automated multi-reviewer quality assessment produced a score of 3.2/10. This report is a machine-generated literature synthesis and does not constitute original research.

1 Introduction

This paper examines: Frchet Radiomic Distance (FRD): A Versatile Metric for Comparing Medical Imaging Datasets. Research question: What is the impact of dataset size and dimensionality on the computational efficiency of FRD compared to perceptual metrics like FID when applied to medical imaging tasks?.

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 3.2/10.

3 Results

11 papers retrieved. 16 claims extracted; 0 independently verified. Quality review score: 3.2/10.

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
FRDv0 features over ImageNet or RadImageNet results in noticeably improved average accuracy and sensitivity, and on-par	×	0.04
FRD improves on FRDv0 noticeably in AUC and sensitivity, and is roughly on-par for accuracy and specificity.	×	0.02
For datasets which have images from multiple domains of the same patient, e.g., BraTS, random sampling is used to ensure	×	0.04
The Breast MRI dataset (Duke Breast Cancer MRI) consists of 12K/2.4K/2.6K train/val/test images from Siemens$\rightarrow$GE (T1 MRI)	×	0.01
The Brain MRI dataset (BraTS) consists of 28K/6K/6K train/val/test images from T1$\rightarrow$T2 modality domains.	×	0.02
The Lumbar spine dataset (TotalSegmentator and in-house MRIs) consists of 2K/0.6K/0.6K train/val/test images from T1 MRI	×	0.02
The CHAOS dataset (Abdom. MRI & CT) consists of 1.8K/1.1K/0.6K train/val/test images from CT$\rightarrow$T1 MRI (in-phase) domains.	×	0.03
FRD achieves an average AUC of 0.94 across Breast MRI, Brain MRI, Lumbar, and CHAOS datasets.	×	0.02
FRD achieves an average accuracy of 0.85 across Breast MRI, Brain MRI, Lumbar, and CHAOS datasets.	×	0.03
FRD achieves an average sensitivity of 0.92 across Breast MRI, Brain MRI, Lumbar, and CHAOS datasets.	×	0.03
FRD achieves an average specificity of 0.93 across Breast MRI, Brain MRI, Lumbar, and CHAOS datasets.	×	0.02
In almost all cases, there is a drop in average performance on test data that was detected as OOD using the binary thres	×	0.02
FRD outperforms other metrics in ranking which of different OOD datasets will result in worse downstream task performanc	×	0.12
The common approach of comparing medical image distributions in terms of the performance of some downstream task such as	×	0.13
Perceptual metrics like FID are commonly used to evaluate image quality relative to real images in computer vision, yet	×	0.13
Many applications of medical image translation and generation rely on FID (or the related Kernel Inception Distance/KID)	×	0.14

References

- http://arxiv.org/abs/1808.05205v1
- http://arxiv.org/abs/2106.08575v1
- http://arxiv.org/abs/2412.01496v2