The Intrinsic Manifolds of Radiological Images and their Role in Deep Learning

Nicholas Konz[1] Hanxue Gu[1] Haoyu Dong[2] Maciej Mazurowski[1,2,3,4]

[1]Department of Electrical and Computer Engineering, [2]Department of Radiology, [3]Department of Computer Science, [4]Department of Biostatistics & Bioinformatics, Duke University, Durham, North Carolina, USA

Introduction

- **The Manifold Hypothesis (MH)**: *High dimensional data can be well described by a much smaller number of degrees of freedom, or* **intrinsic dimensions**.
  - The intrinsic dimension of a dataset estimates its information content.
  - Neural networks can learn to convert raw data to abstract, informative features that are *intrinsic* to the dataset [1].

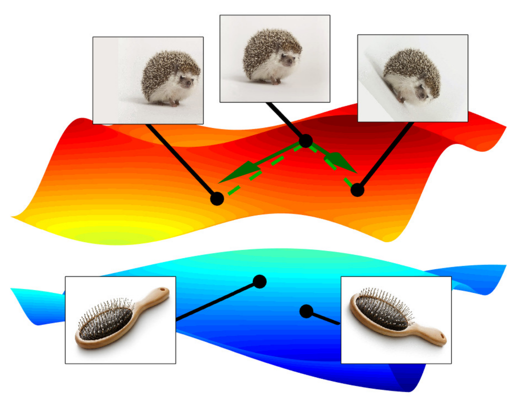

Figure 1: Visualization of intrinsic low-dimensional image manifolds from [2].

- **Why study the intrinsic dimension of medical images?**
  - Medical vs. natural image datasets: different relevant semantics, yet we lack understanding of how networks learn differently between the two domains.
  - Due to the MH, understanding the intrinsic structure of medical image datasets is key to analyzing how networks learn from them.
  - We can answer questions such as "how does the intrinsic dimension of a dataset affect how challenging it is for a network to learn from it?"

Objectives

1. Estimate the intrinsic dimensions of common radiology datasets, and compare to natural image datasets.
2. Evaluate the relationship of dataset intrinsic dimension with network generalization ability; comparing within and between the domains of radiological and natural images.

Estimating the Intrinsic Dimension of Image Manifolds

- The MH says that our d-dimensional data lies on a manifold $\mathcal{M} \subseteq \mathbb{R}^d$ such that $\dim \mathcal{M} = m \ll d$.
- We can estimate m via **maximum likelihood**:
  - assume that manifold volume scales exponentially with m as we move away from a point.
  - model volume with k-nearest neighbor distance T_k.
  - Model dataset as being sampled from a Poisson Process, and find m via MLE:

$$\hat{m} = \left[\frac{1}{N(k-1)}\sum_{i=1}^{N}\sum_{j=1}^{k-1}\log\frac{T_k(x_i)}{T_j(x_i)}\right]^{-1}$$

Datasets and Experimental Settings

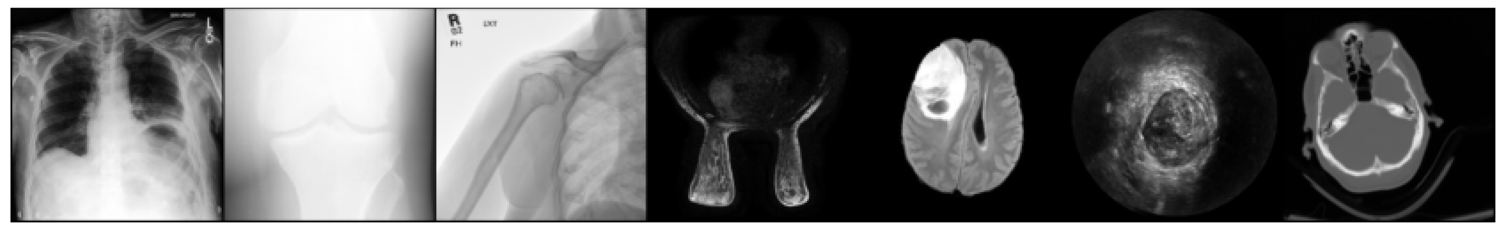

Figure 2: Samples from our seven evaluated datasets.

- We analyze seven common radiology datasets from different modalities.
- **ID estimation:**
  - For each dataset, we use a sample of 7500 images, evenly class-balanced according to our chosen binary classification task for each.
- **ID and Generalization Ability**:
  - We train a classification network for each dataset, and test on 750 unseen data points.
  - We evaluate many training set sizes, neural network models, and task choices. All networks are maximally fit to the training set.

Result 1: Radiological vs. Natural Image Intrinsic Dimension

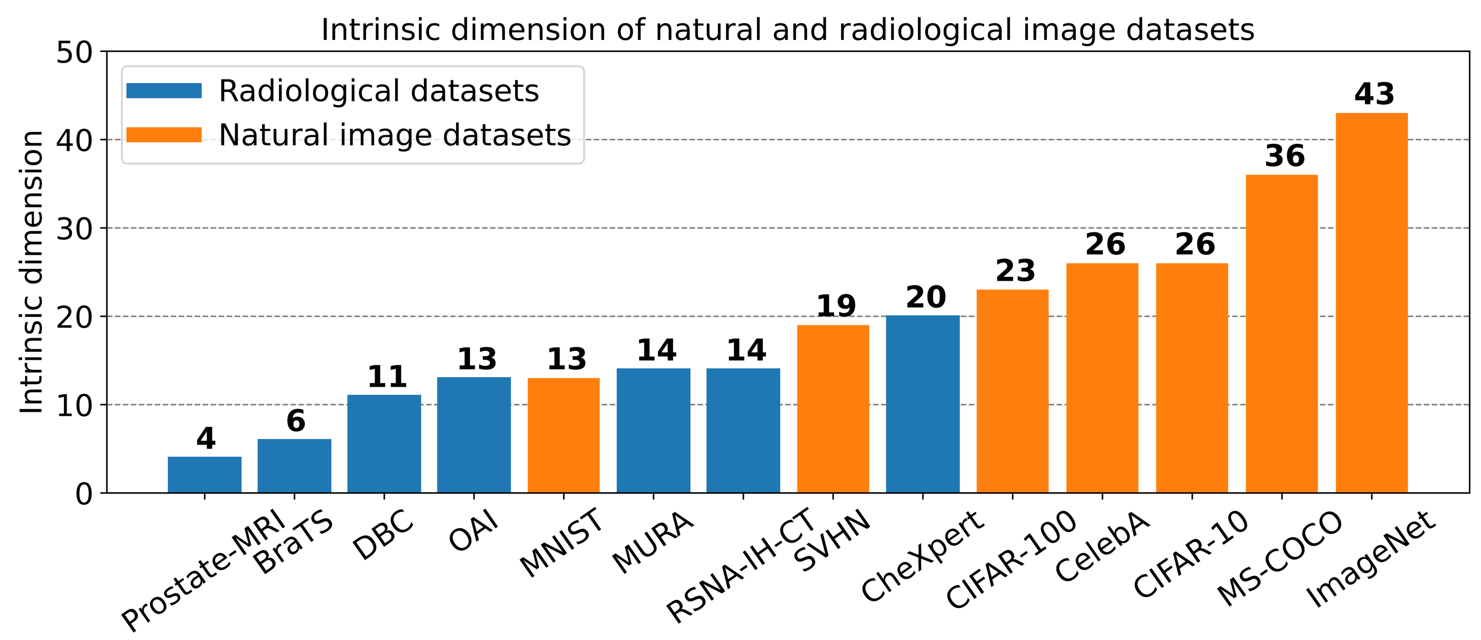

Figure 3: Intrinsic dimension of radiological and natural [3] image datasets.

- **Findings:**
  - **Radiological image datasets tend to have lower intrinsic dimension (ID) than natural image datasets!**
  - ID $\ll$ extrinsic dimension (ED/num. of pixels).
    - Intuitively, modifying ED (resizing images) didn't affect ID.

Main Result: Intrinsic Dimension and Generalization Ability

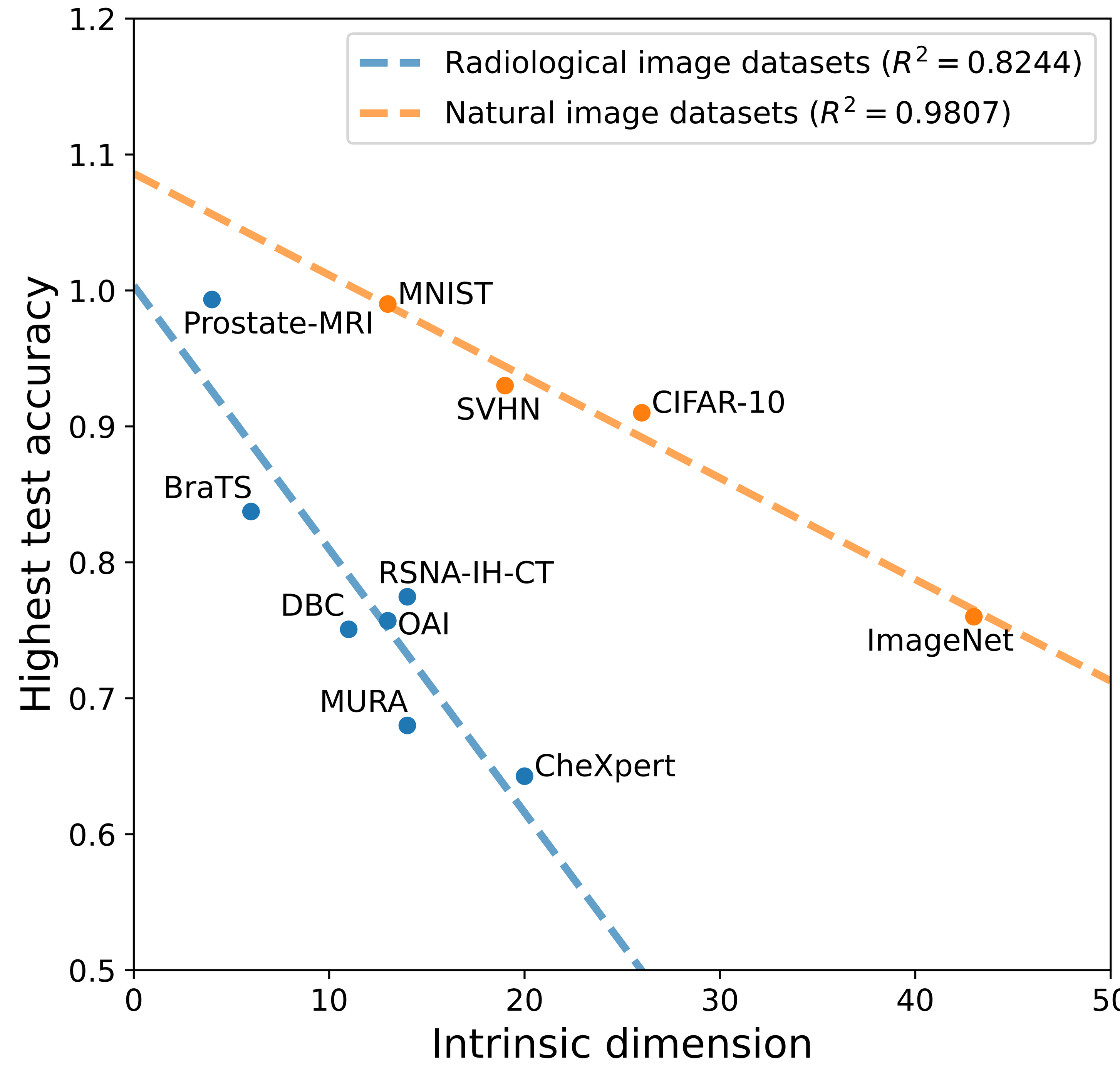

Figure 4: Linearity of model generalization ability with respect to dataset intrinsic dimension, for radiological and natural image datasets ($N_{\text{train}} = 2000$ on ResNet-18).

- Vestibulum nisl, quis euismod velit eros in ligula.
  - Cras rhoncus quam et augue convallis in elementum urna tincidunt.
- Proin ut vestibulum augue.
  - Donec dapibus sagittis neque eu ultrices.

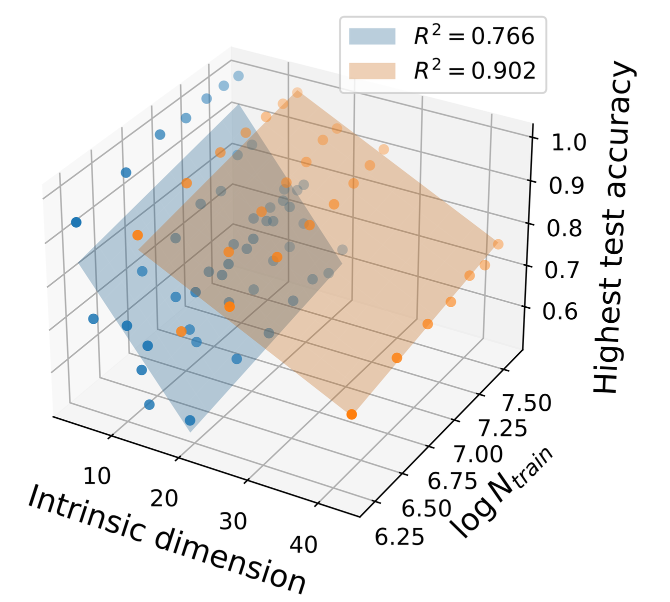

Figure 5: Fig. 1, including training set size.

{numerical evidence showing how these relationships didn't change with model, dataset size, task, etc (see paper). Also additional results from "Contributions" in paper.}

Open Questions/Future Work

- Investigate theoretical reasons for linearity and difference in sharpness of correlation between the two domains
- Determine further uses of ID estimate for predictions, experiments