MaNo: Exploiting Matrix Norm for Unsupervised Accuracy Estimation

Renchunzi Xie*[1] Ambroise Odonnat*[2,3] Vasilii Feofanov*[2] Weijian Deng[4] Jianfeng Zhang[2] Bo An[1]

*Equal contribution [1]NTU [2]Huawei Noah's Ark Lab [3] Inria [4] ANU

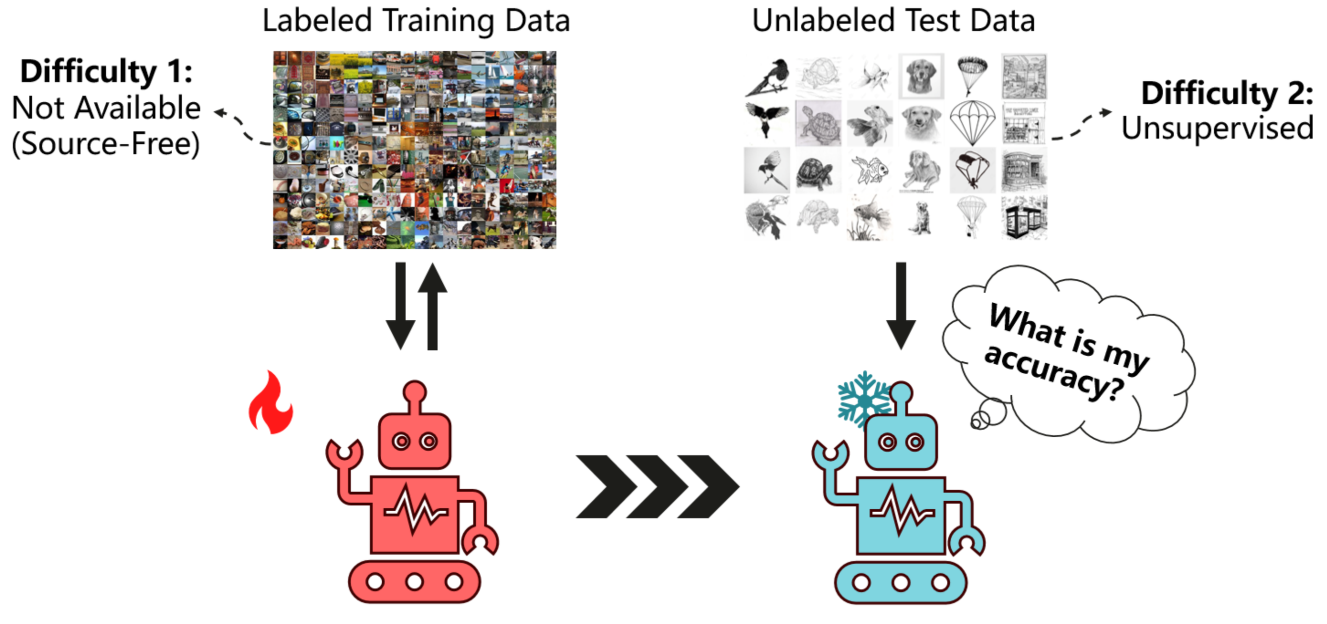
TL;DR

- Predicting generalization performance under distribution shifts is challenging
- Most methods use logits without dealing with miscalibration cases
- We propose MaNo, a **theoretically grounded** estimation approach
- It automatically takes into account miscalibration scenarios
- It can be applied to ResNets, ConvNext, and ViT architectures
- Benefits: **SOTA**, **efficient**, **architecture agnostic**, **robust**

Problem Setup

Goal: given a pre-trained model f, predicts its performance on a test set $\mathcal{D}_{\text{test}}$.

- Takes as input a pre-trained model f and test data $\mathcal{D}_{\text{test}}$.
- Distribution shifts, i.e., $p_S \neq p_T$ where training data $\sim p_S$ and test data $\sim p_T$.
- Outputs an estimation score $\mathcal{S}(f, \mathcal{D}_{\text{test}})$ linearly correlated with the true accuracy.



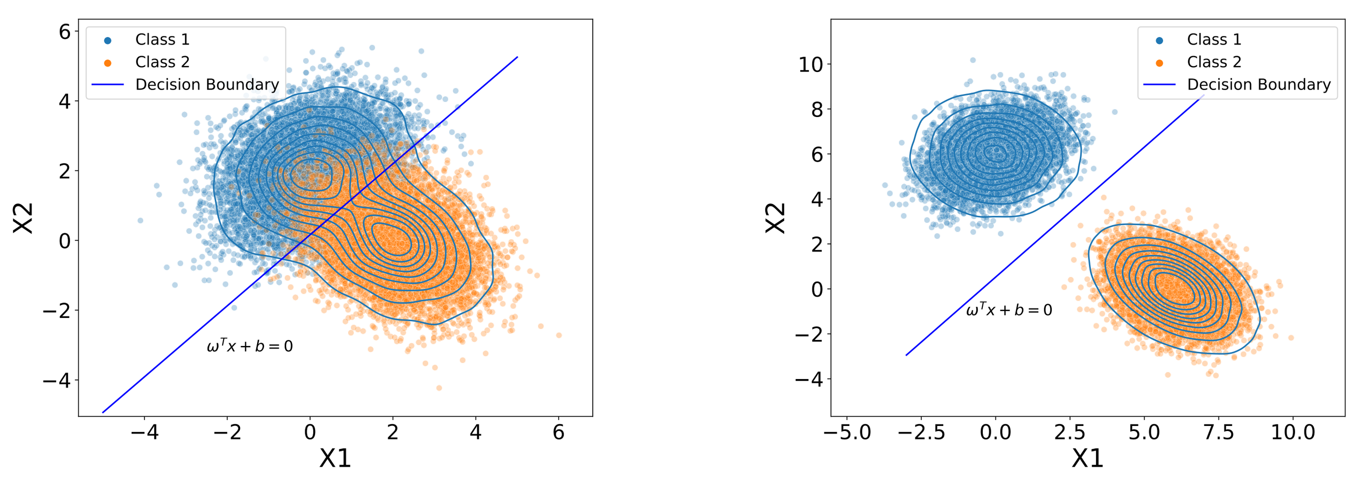
This is a challenging task often occurring in real-world scenarios.

Motivation

Question 1: *Why are logits informative of generalization performance?*
Question 2: *How to alleviate the overconfidence issues of logits-based methods?*

Logits Reflect Distances to Decision Boundaries

- Decision boundary of class k is the hyperplane $\{\mathbf{z}' \in \mathbb{R}^q | \boldsymbol{\omega}_k^\top \boldsymbol{z}' = 0\}$,
- Distance from a point $\mathbf{z}$ this hyperplane is $\mathrm{d}(\boldsymbol{\omega}_k, \mathbf{z}) = |\boldsymbol{\omega}_k^\top \boldsymbol{z}| / \|\boldsymbol{\omega}_k\|$,
- Logits reflects decision to decision boundary as $|\mathbf{q}_k| = |\boldsymbol{\omega}_k^\top \mathbf{z}| \propto d(\boldsymbol{\omega}_k, \mathbf{z})$,
- **Low-density assumption**: misclassified samples are closer to decision boundaries.



High-density region Low-density region

Logits (in absolute values) positively correlated to generalization performance.

MaNo: A Simple Three-Step Recipe

- **Input:** Pre-trained model f, test dataset $\mathcal{D}_{\text{test}} = \{\mathbf{x}_i\}_{i=1}^N$.
- **Inference:** Recover logits $\mathbf{q}_i = f(\mathbf{x}_i)$.
- **Criterion:** $\Phi(\mathcal{D}_{\text{test}}) = \mathrm{KL}(\text{uniform}||\text{softmax proba})$

$$1)\quad v(\mathbf{q}_i) = \begin{cases} 1 + \mathbf{q}_i + \frac{\mathbf{q}_i^2}{2}, & \text{if } \Phi(\mathcal{D}_{\text{test}}) \leq \eta \\ \exp(\mathbf{q}_i), & \text{if } \Phi(\mathcal{D}_{\text{test}}) > \eta \end{cases}$$

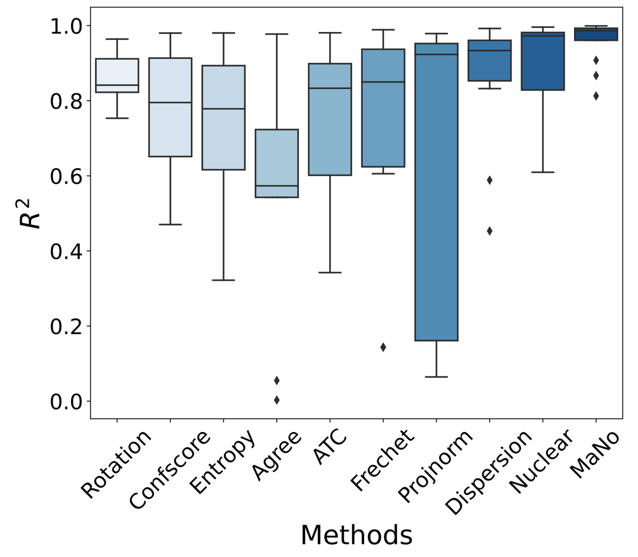
$$2)\quad \sigma(\mathbf{q}_i) = \frac{v(\mathbf{q}_i)}{\Sigma_{k=1}^K v(\mathbf{q}_i)_k} \in \Delta_K$$

$$3)\quad \mathcal{S}(f, \mathcal{D}_{\text{test}}) = \frac{1}{\sqrt[p]{NK}} \|\mathbf{Q}\|_p = \left(\frac{1}{NK} \sum_{i=1}^N \sum_{k=1}^K |\sigma(\mathbf{q}_i)_k|^p\right)^{\frac{1}{p}}$$

MaNo is simple yet efficient and theoretically grounded as we prove that it captures the model's uncertainty.

Robustness Comparison

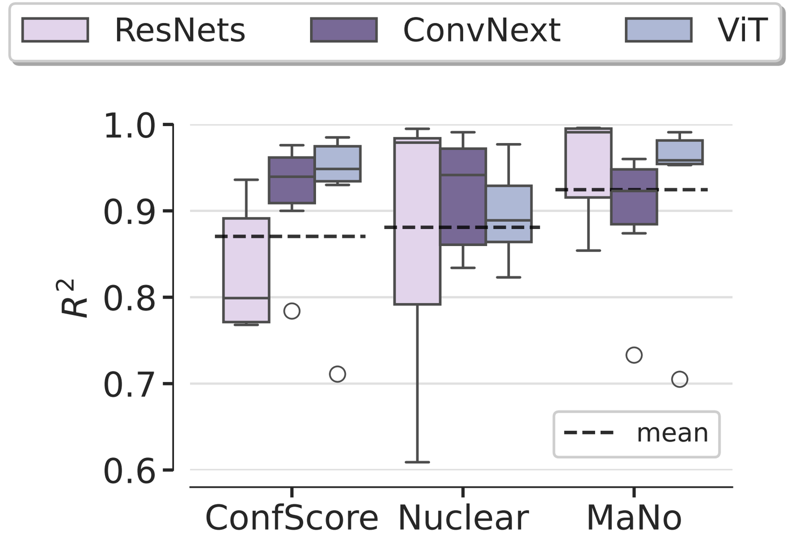
We compare our method with its competitors on all the distribution shifts.



Overall, MaNo leads to the best and most robust estimations.!

Beyond ResNets: ConvNext and Vision Transformers

To ensure the efficiency and versatility of MaNo, we apply it to 2 other SOTA architectures.



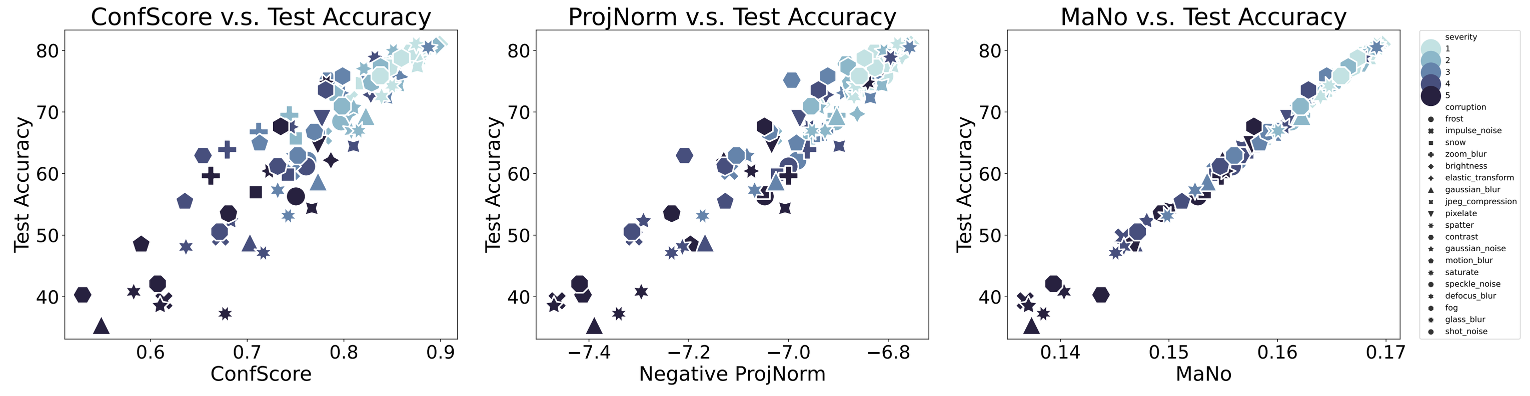
MaNo is the best approach with SOTA architectures!

Experimental Results: Better, Faster, Stronger

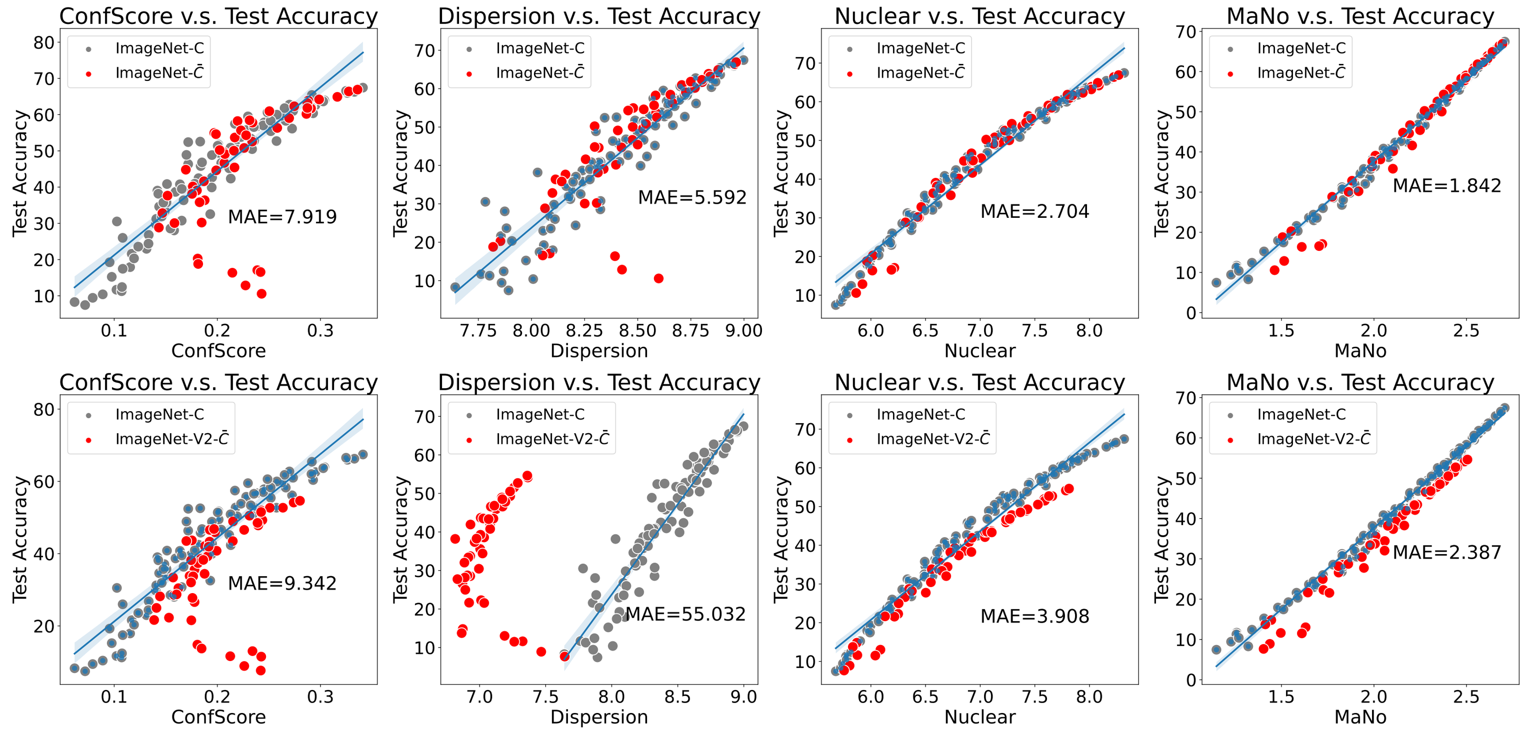
- Comparison between MaNo and its competitors with metrics ρ and R^2,
- Comparison across several architectures: ResNets, ConvNext, ViT,
- Extensive evaluation with common benchmarks on various distribution shifts.

Shift	MaNo -	COT 2024	MDE 2024	Nuclear 2023	Dispersion 2023	ProjNorm 2022
Synthetic	**0.991**	0.988	0.947	0.982	0.960	971
Subpopulation	**0.983**	0.962	0.920	0.973	0.909	897
Natural	**0.905**	0.871	0.436	0.455	0.410	382
Overall improvement		2%	25%	6%	26%	28%

MaNo outperforms all the baselines while being training-free.



MaNo linearly correlates with ground-truth test accuracy on Entity-18.



MaNo significantly surpasses its competitors on variants of ImageNet.

Main References

- **Odonnat et al.** - AISTATS 2023 *$\mathcal{T}$-similarity*
- **Deng et al.** - ICML 2023 *Nuclear*
- **Xie et al.** -**NeurIPS 2024 (this work)** MaNo

Renchunzi Xie Ambroise Odonnat

Take Home Message

Predicting generalization performance under distribution shifts is challenging.
→ Start using MaNo for an **efficient** and **accurate** estimation!

Want to Know More?

paper code

Correspondence to renchunzi.xie@ntu.edu.sg and ambroiseodonnattechnologie@gmail.com