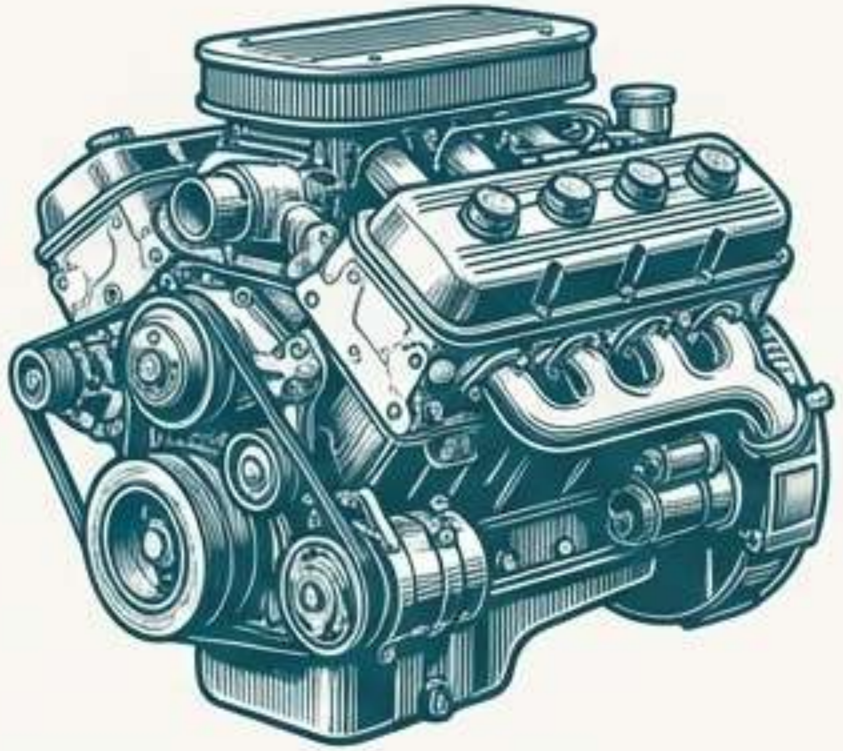
Towards Efficient Foundation Models: A Novel Time Series Embedding

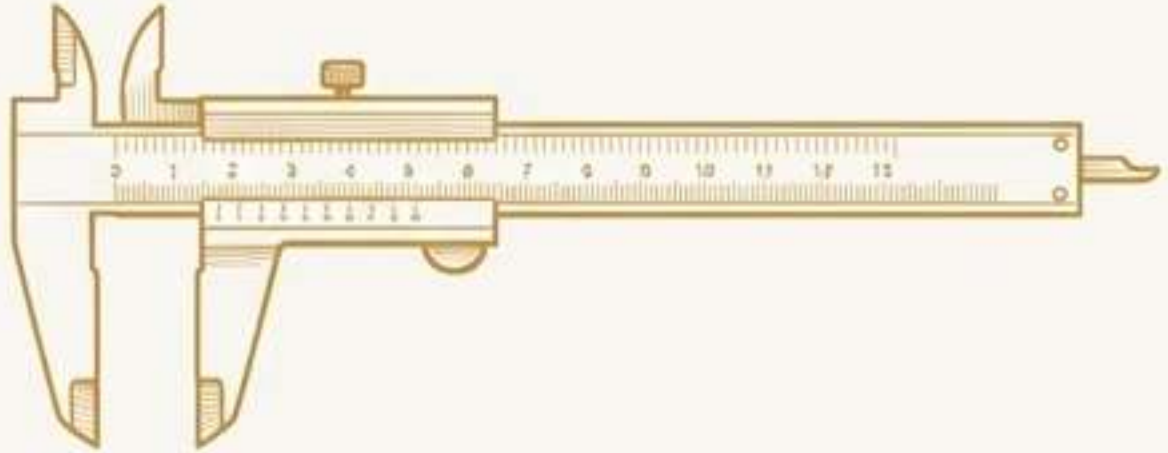
Jessy Xinyi Han, Arth Dharaskar, Nathaniel Lanier, Abdullah Omar Alomar, Aditya Agrawal, Angela Yuan, Jocelyn Hsieh, Ishan Shah, Muhammad Jehangir Amjad, Devavrat Shah

Massachusetts Institute of Technology & Ikigai Labs

The Current Landscape: A Clash of Two Paradigms



VS.



The Foundation Model Paradigm

Massive, universal models pre-trained on vast datasets.

Examples: TimesFM, Moirai, Chronos

The Traditional Model Paradigm

Resource-efficient models fit to individual time series.

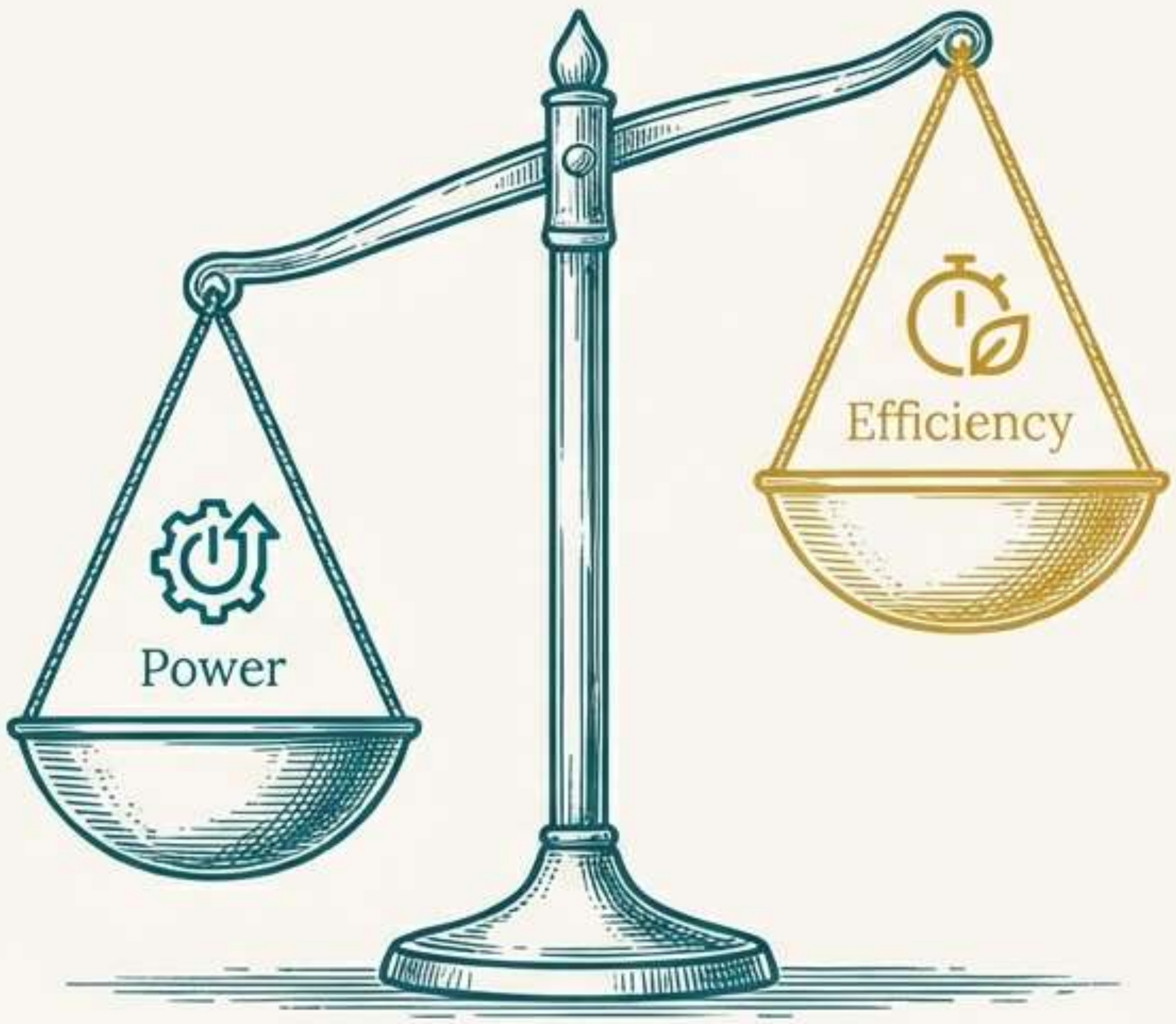
Example: Prophet

Every Approach Comes with a Fundamental Trade-Off

The Foundation Model Challenge

- Resource-intensive training and inference.
- Requires a sufficiently rich pre-training dataset.

High Power, High Cost.



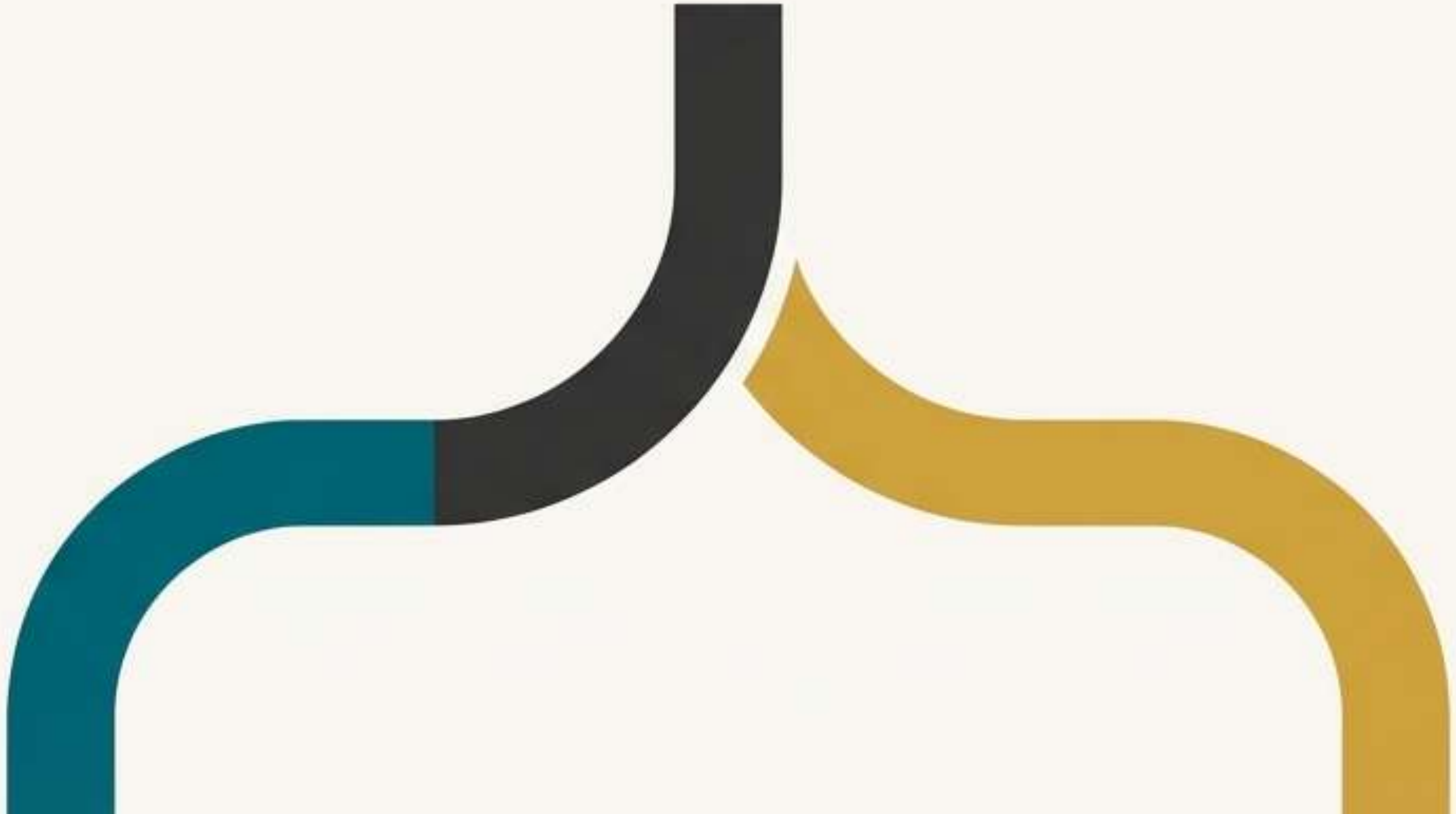
The Traditional Model Challenge

- Experimentation is ineffective, especially for short time series.
- Unable to incorporate information from pre-training data.

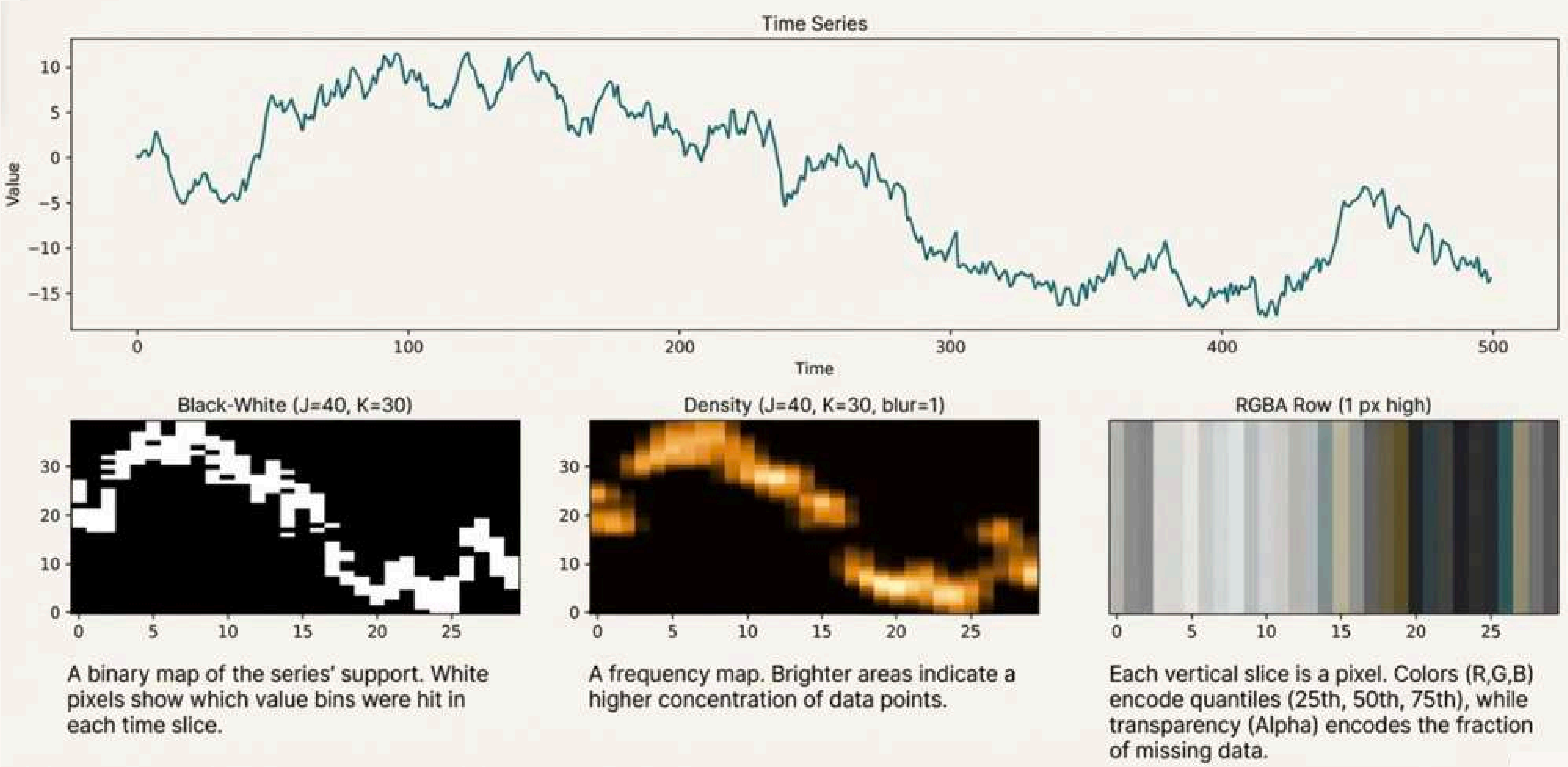
Low Cost, Limited Insight.

Can we achieve the best of both worlds?

Is it possible to have the resource efficiency of traditional approaches while retaining the ability to incorporate information from pre-training data?



The Unifying Insight: Treat a Time Series as an Image



We introduce a novel, universal embedding that maps any time series—of any length, scale, or with missing values—to a fixed-size 2D image.

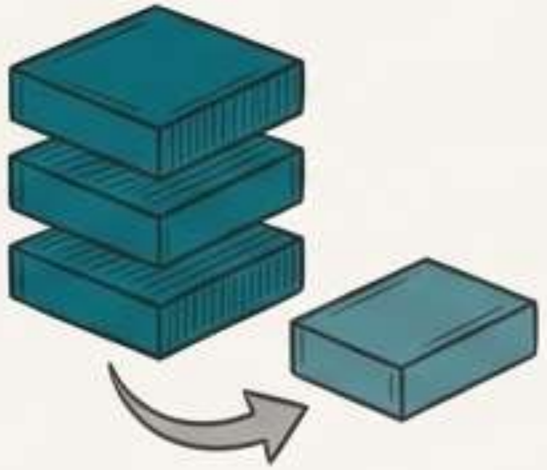
NotebookLM

The Benchmark: A Model Identification Challenge

Objective

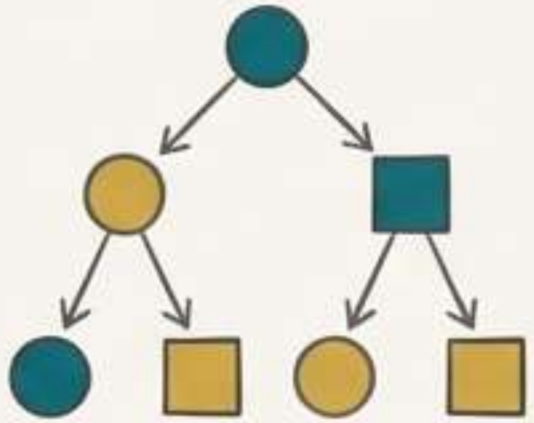
To evaluate the efficacy of the embeddings, we test their ability to identify the true generating model class for a given time series.

Dataset



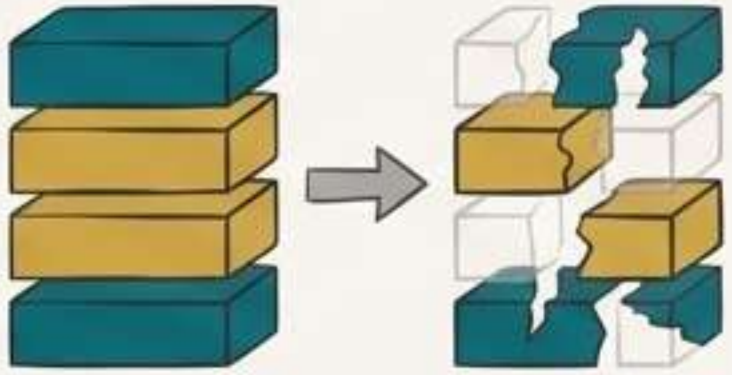
- Synthetic time series from classical models (SARIMA, Linear, Bernoulli, Harmonic) labeled as 'Pure' classes.
- Compositional time series created by mixing two pure classes, labeled as 'Mixed' classes.

The Test



A random forest classifier is trained on the embeddings to predict the model class.

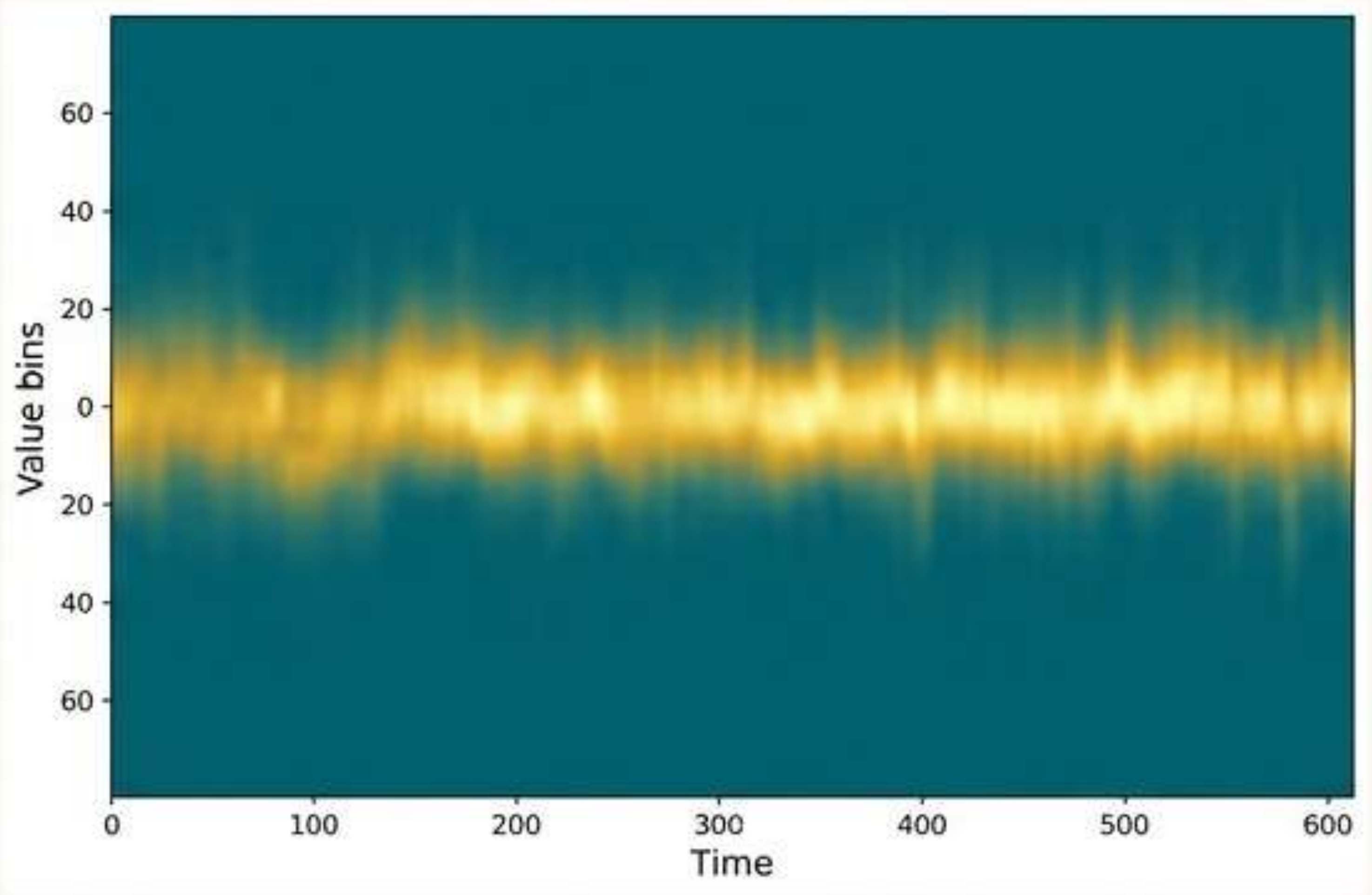
Robustness Check



We inject random missing data (10% and 20%) to test performance under pressure.

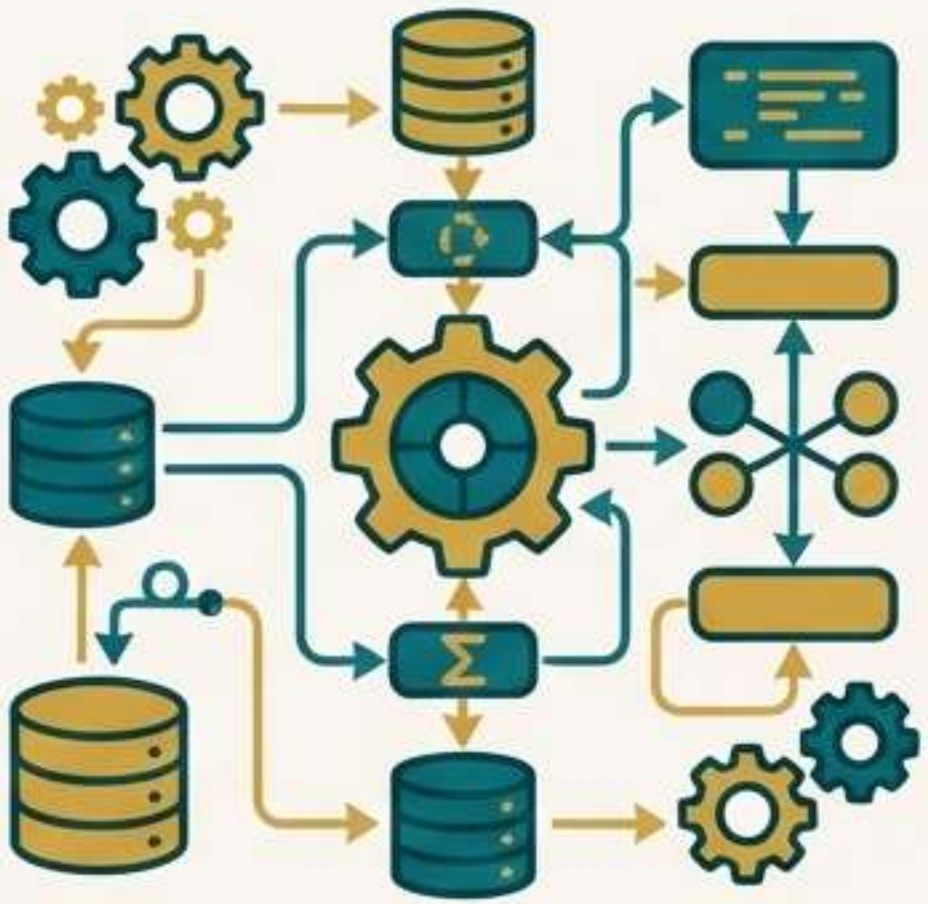
The Head-to-Head Challenge

Contender 1: The Image-Based Challenger



Density Heatmap (with Gaussian Blur). Our best-performing image embedding, capturing value frequency over time. Cross-validated for optimal parameters (J=40, K=32, blur=1).

Contender 2: The Foundation Model Champion



TimesFM. A 50-layer Transformer TSFM (google/timesfm-2.0-500m-pytorch) that produces a learned embedding vector.

Round 1: Performance on Complete Data (0% Missing)

The lightweight Heatmap embedding achieves performance nearly identical to the heavyweight TimesFM.

Comparable predictive power without the need for massive pre-training.

Round 2: Performance Under Pressure (20% Missing)

As data quality degrades, the Heatmap embedding demonstrates superior robustness in strict classification.

The image-based approach excels in strict classification under moderate data corruption.

A Complete Performance Breakdown

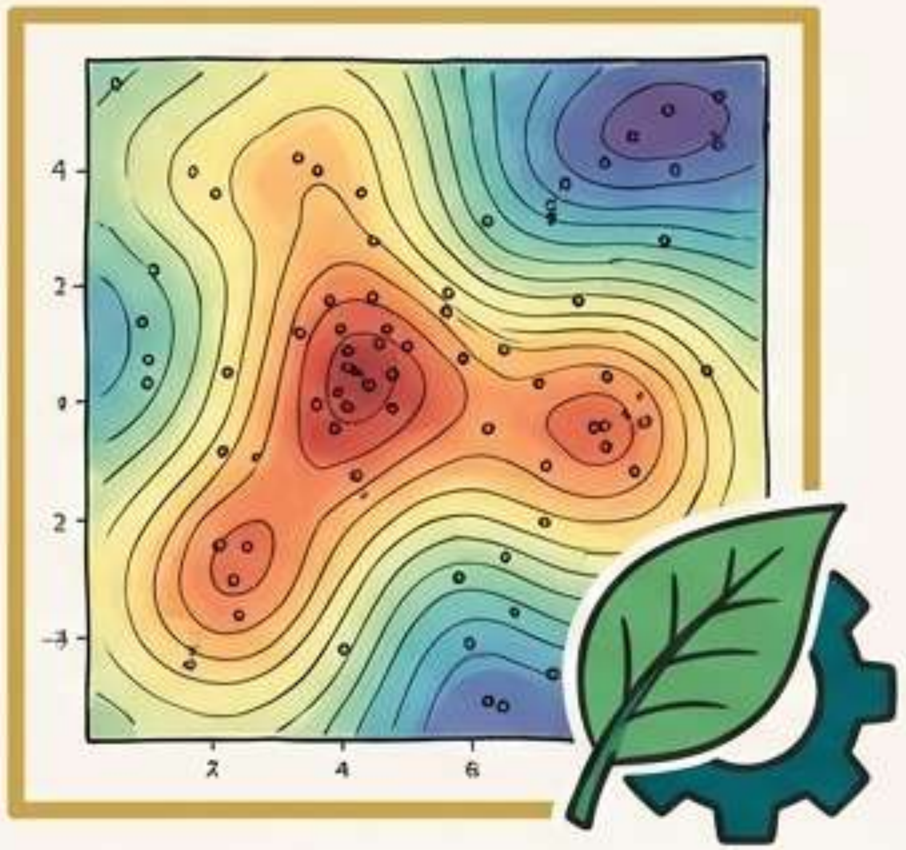
Embedding	Missingness (p)	Strict Accuracy (Overall)	Overlap Accuracy (Overall)
Heatmap	0.0	0.775	0.988
TimesFM	0.0	**0.781**	**0.997**
Heatmap	0.1	**0.759**	0.983
TimesFM	0.1	0.748	**0.996**
Heatmap	0.2	**0.738**	0.983
TimesFM	0.2	0.721	**0.992**

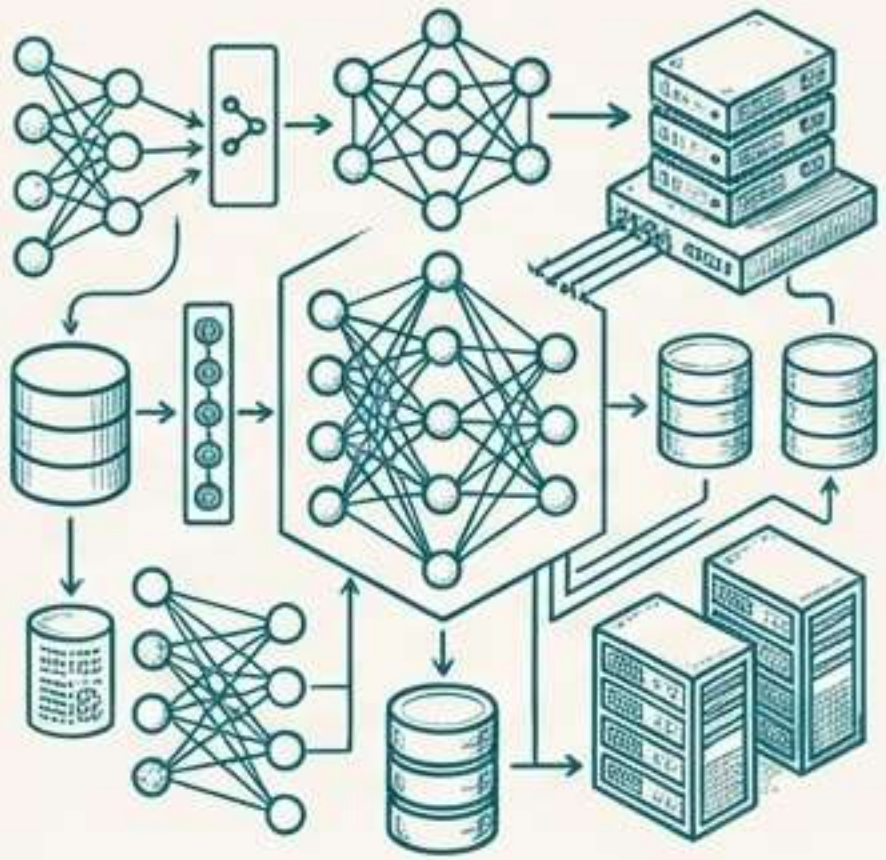
Heatmap excels in strict accuracy as missingness increases.

TimesFM maintains a consistent edge in relaxed "overlap" accuracy, especially for mixed classes.

The analysis reveals our embedding is able to achieve similar performance as the TSFM embedding but with a fraction of computational cost.

Comparable Power. A Fraction of the Cost.





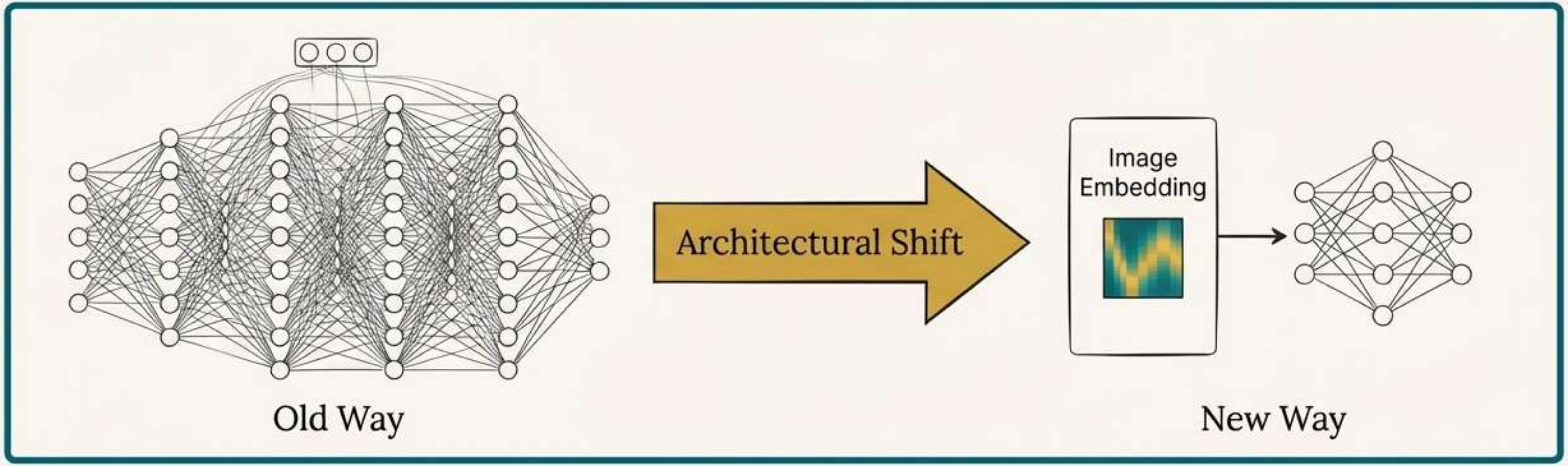
High-Fidelity Representation

Image-based embeddings can achieve discriminative power that is competitive with—and in some cases superior to—a state-of-the-art Time Series Foundation Model.

Radical Resource Efficiency

This performance is achieved via a simple, universal transformation, eliminating the need for resource-intensive pre-training and inference required by large TSFMs.

A New Blueprint for Efficient Time Series Models



Core Idea: This work suggests that simple, resource-efficient encoders can rival complex foundation models on structured tasks.

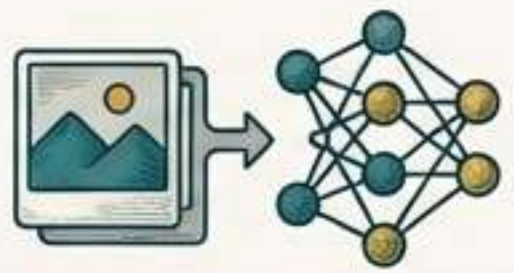
Implication: Image-based embeddings can serve as powerful, lightweight, and effective building blocks for the next generation of compute-efficient Time Series Foundation Models.

The Path Forward

Extending the Evaluation

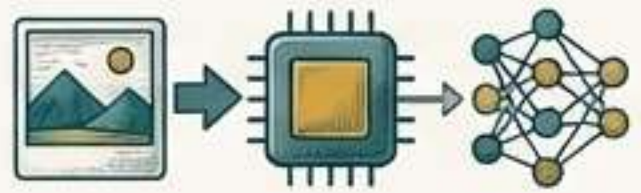
Apply and **validate** these embeddings on diverse, real-world datasets beyond the synthetic benchmark.

Exploring Hybrid Encoders



Combine the strengths of image-based features with learned representations in novel hybrid architectures.

Integrating with Vision Transformers



Leverage the power of pre-trained vision models (e.g., ViT) by feeding them these novel time series images.