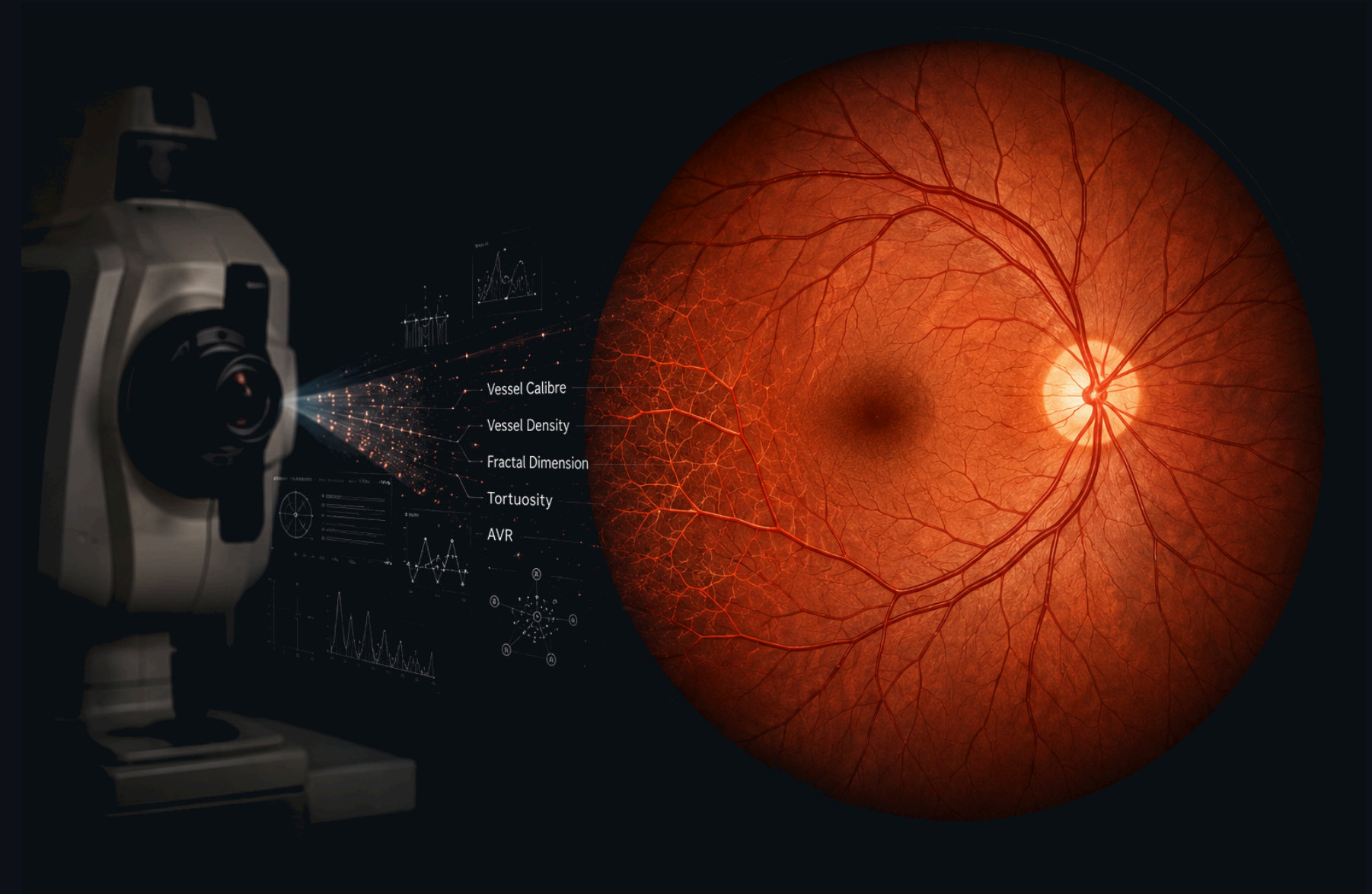


Feature-Space Translation for Cross-Device Retinal Biomarkers

Aligning heterogeneous fundus cameras and decoding foundation-model embeddings into clinically meaningful vascular biomarkers.

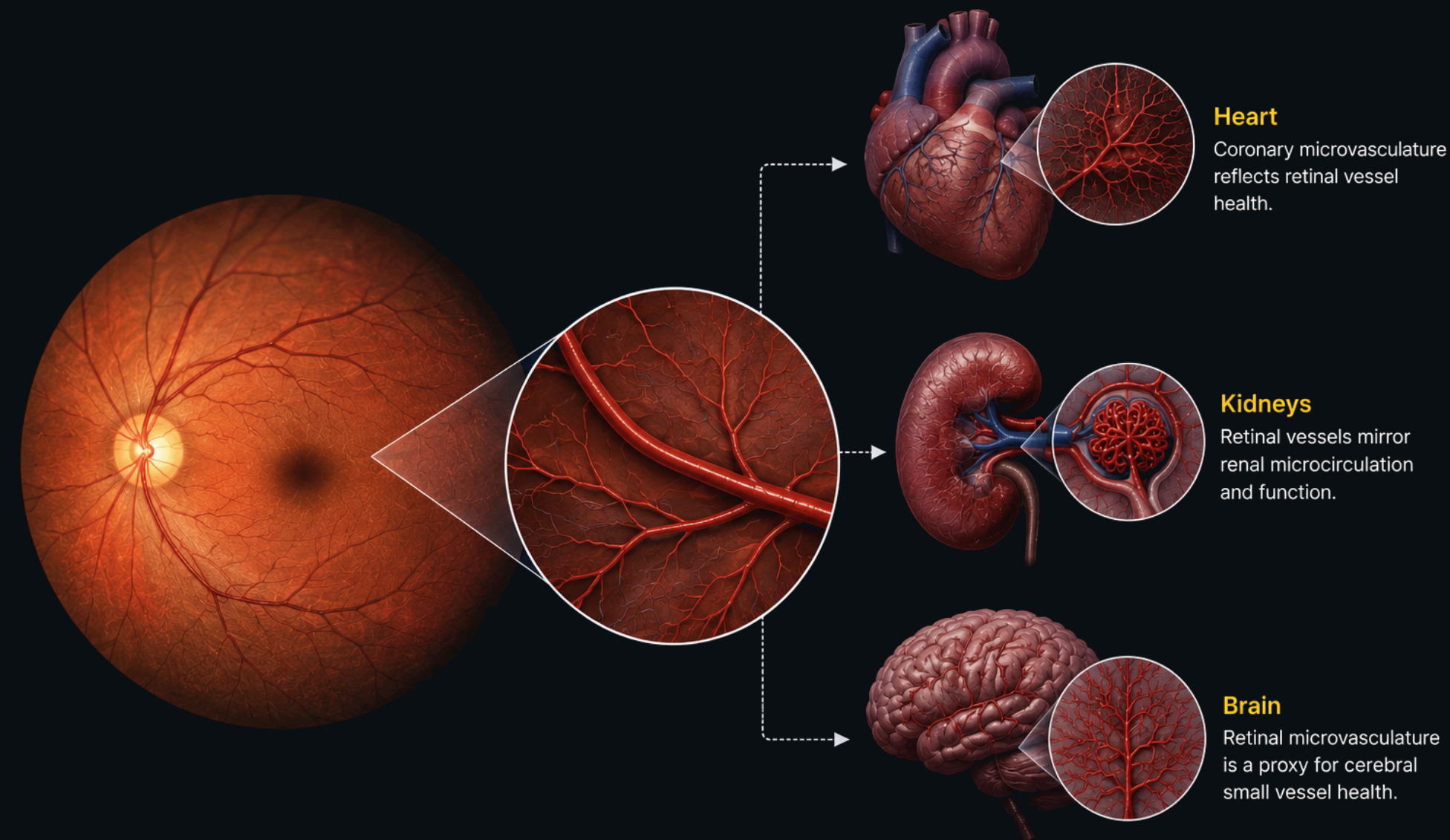


— Sparsh Rastogi et al.

MOTIVATION

The retina is a window into systemic health

The only place to directly, non-invasively image microvasculature — and its vessels mirror the beds feeding the heart, kidneys, and brain.



THE GAP

Two barriers to clinical deployment

BARRIER 01

01



≠



CAMERA 1

CAMERA 2

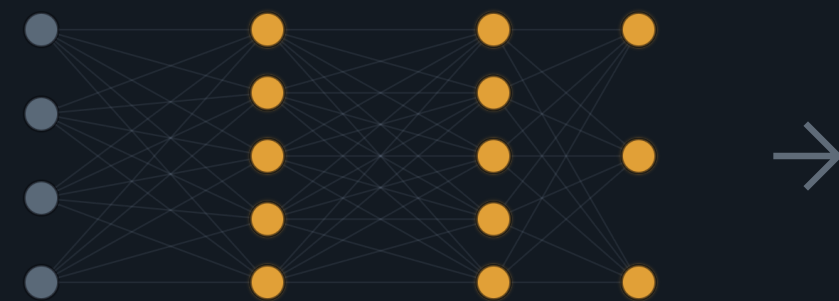
Device-induced variability

Camera heterogeneity introduces systematic OOD shift in extracted features.

↘ Model drift across devices & populations

BARRIER 02

02



FOUNDATION-MODEL
EMBEDDING

MEANING

Uninterpretable embeddings

Foundation-model representations have no direct link to clinical biomarkers.

Ⓢ Limited transparency & validation

OUR APPROACH

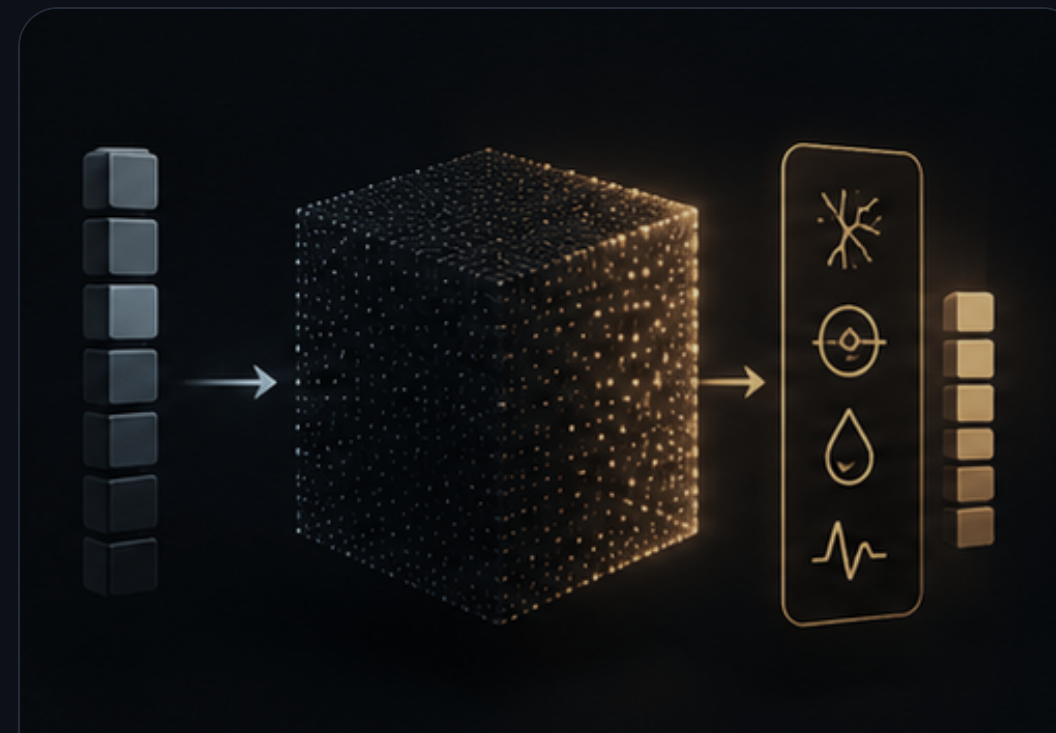
A unified feature-space translation framework



STAGE 1

Cross-device translation

Align feature distributions across imaging devices.



STAGE 2

Biomarker extraction

Decode embeddings into clinically meaningful biomarkers.



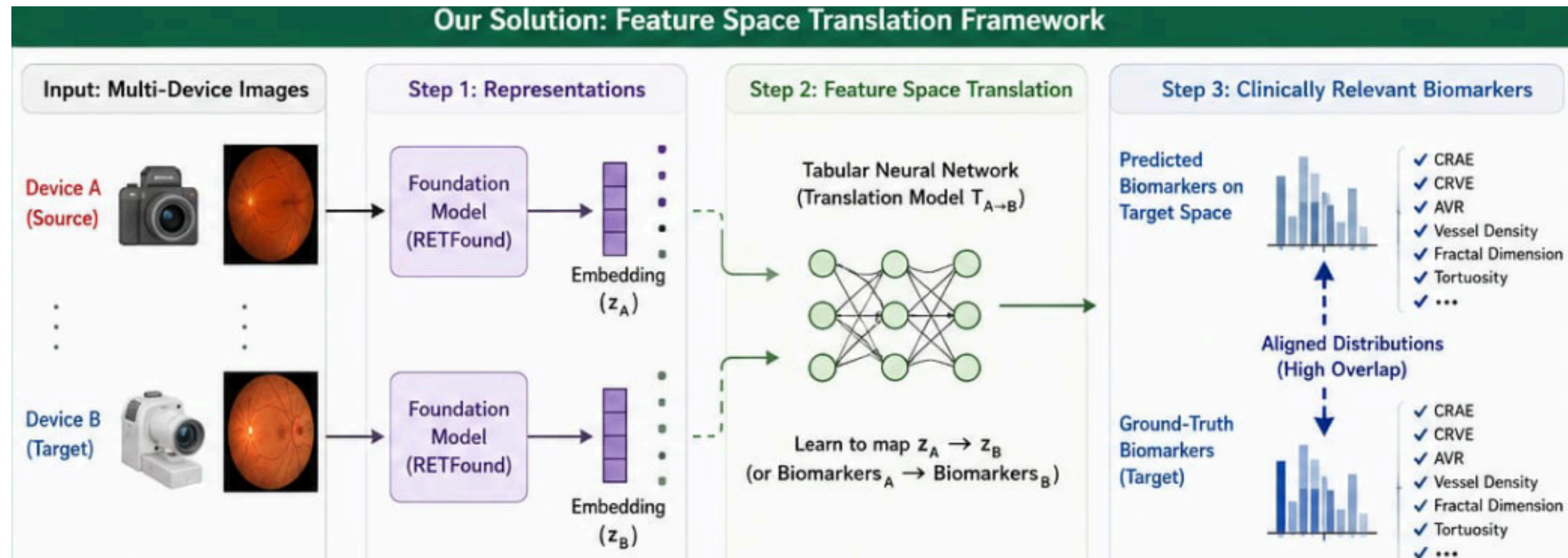
STAGE 3

Cross-device estimation

Predict target-device biomarkers directly from embeddings.

FRAMEWORK OVERVIEW

The end-to-end translation pipeline



TRAINING OBJECTIVE

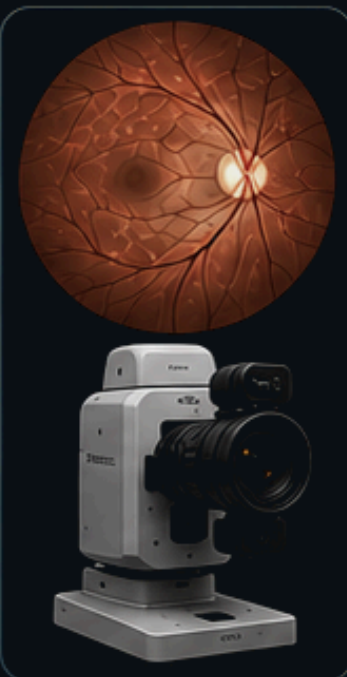
$$L = ||\mathbf{f}(X_S) - X_T||^2 + \lambda \cdot W_1(P_{\text{pred}}, P_{\text{target}})$$

reconstruction loss + distributional alignment

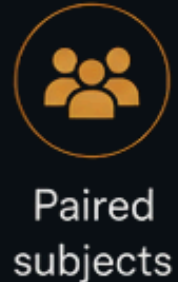
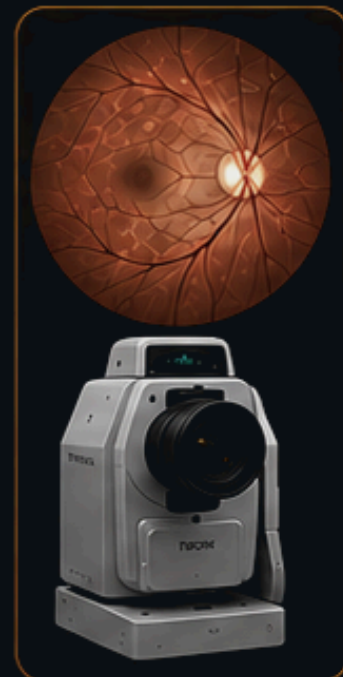
Experimental setup

1 DATASET • AI-READI

Triton
(source device)



Maestro2
(target device)



Paired
subjects



1,067

Participants



12

Biomarkers



1024-D

RETFound
embeddings



70/15/15

5-fold CV
(split)

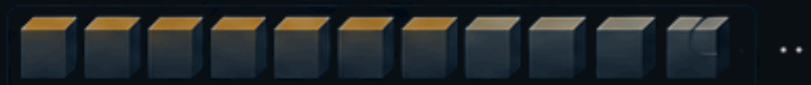
2 OUR PIPELINE

Retinal images



RETFound
Foundation model

1024-D embedding



Tabular model
(e.g., TabPFN)

12 clinical biomarkers



4 EVALUATION METRICS



Prediction accuracy

R^2

MAE



Distribution alignment

Wasserstein
distance
(lower is better)

KL
divergence
(lower is better)

JS
divergence
(lower is better)

3 BENCHMARK • 8 TABULAR MODELS

Classical ML



Linear
Regression



Ridge
Regression



Gradient
Boosting



CatBoost

Tabular Neural Nets



TabM



RealMLP



Lightweight
MLP



TabPFN
(best)

Structural biomarkers align across devices

78.5%

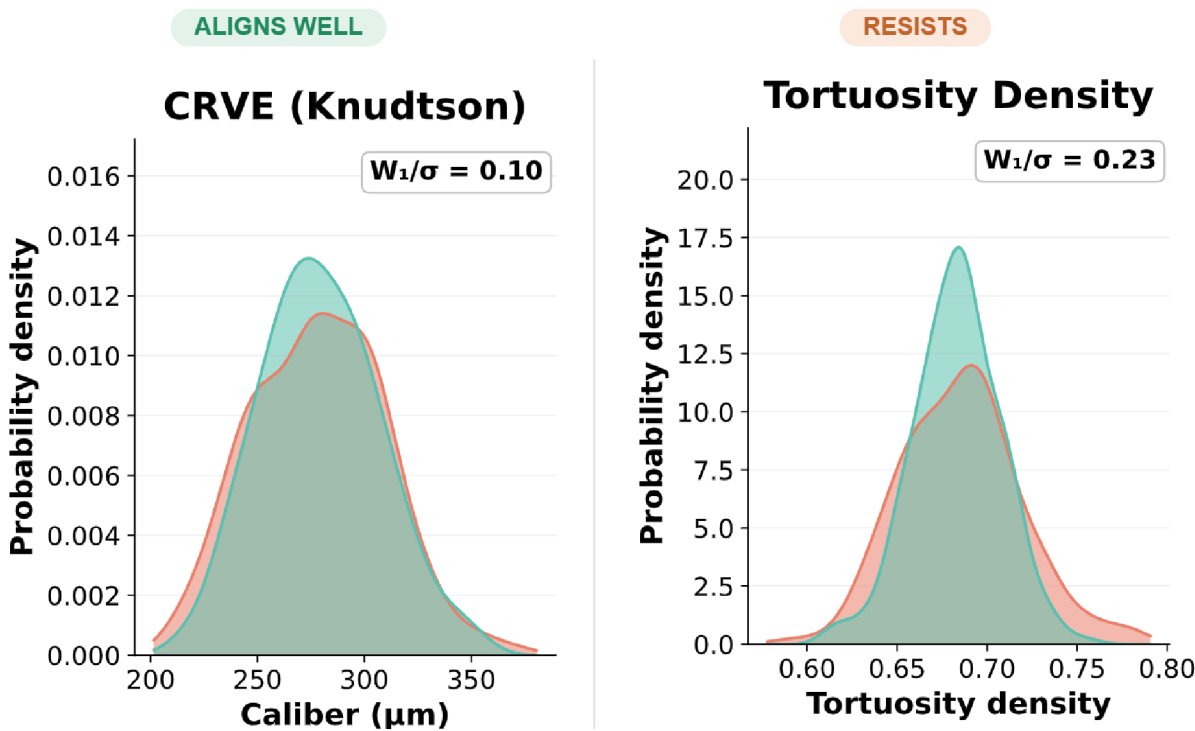
Best cross-device R^2 • CRVE (vessel caliber)
Structural biomarkers translate between devices with high fidelity.

• RESULTS • R^2 (%) • TOP 6 MODELS

Biomarker	TabPFN	L-MLP	CatB	TabM	RMLP	Ridge
AVR (Knudtson)	59.7	<u>59.2</u>	53.0	<u>59.2</u>	56.8	57.8
Artery dist. tort.	29.6	25.6	26.6	<u>27.8</u>	24.6	25.6
Artery fractal dim.	68.8	66.9	63.3	<u>68.3</u>	68.2	67.5
Average width	67.4	66.4	61.0	<u>67.3</u>	64.8	66.8
CRAE (Knudtson)	72.9	<u>73.5</u>	72.5	72.3	73.7	73.0
CRVE (Knudtson)	<u>78.2</u>	78.0	76.8	77.7	78.1	78.5
Distance tort.	42.9	41.2	39.5	<u>42.0</u>	40.5	41.5
Fractal dimension	70.1	69.4	67.5	<u>69.9</u>	69.2	69.7
Tortuosity density	51.2	49.8	47.0	<u>50.5</u>	48.8	50.0
Vein dist. tort.	31.8	30.2	28.1	<u>31.0</u>	29.5	30.5
Vein fractal dim.	64.5	63.2	60.1	<u>64.0</u>	62.8	63.6
Vessel density	58.9	57.5	54.2	<u>58.4</u>	56.7	57.9

Table 1 • Maestro2 → Triton • **bold** best, underline 2nd • full 8-model table in paper

• DISTRIBUTION ALIGNMENT



Prediction vs. Triton target (n = 241). Lower W_1/σ (Wasserstein distance) indicates tighter alignment. Full 12-biomarker figure in the paper.

Embeddings encode clinical morphology

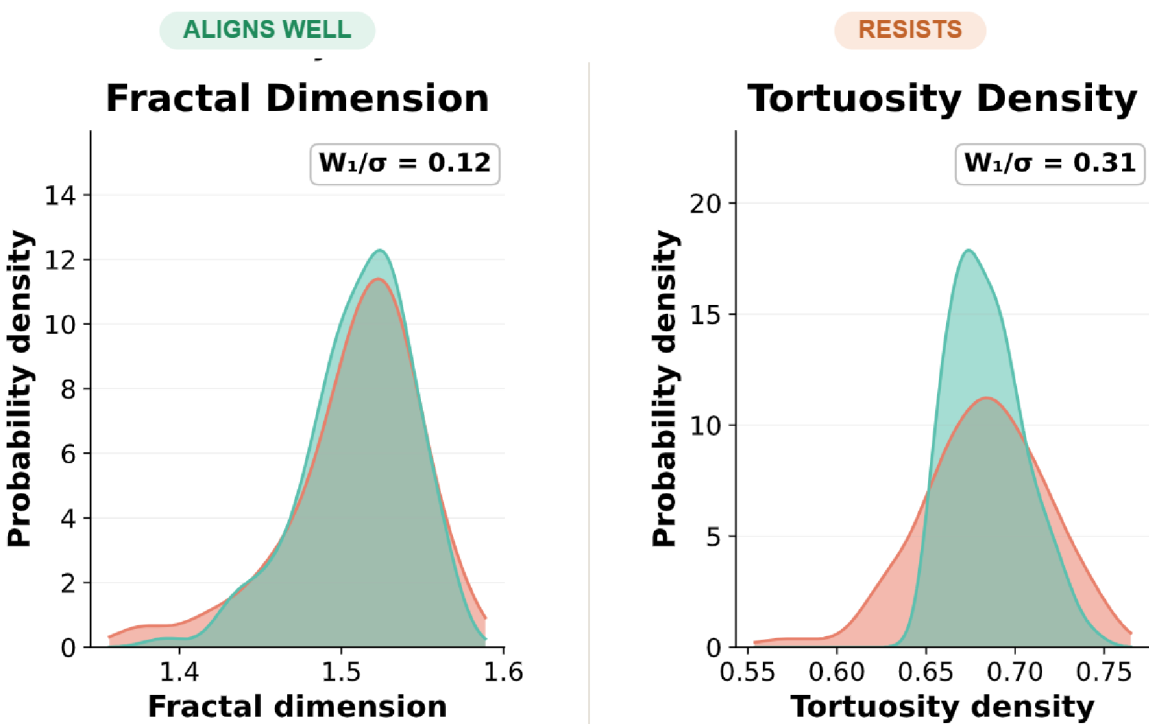
87.5% Best embedding → biomarker R² • Fractal dimension
RETFound embeddings directly encode vascular morphology.

• RESULTS • R² (%) • 4 MODELS

Biomarker	TabPFN	Ridge	ElasticNet	PLS
AVR (Knudtson)	<u>5.8</u>	2.7	3.3	8.3
Artery dist. tort.	-3.6	-34.2	<u>0.0</u>	1.9
Artery fractal dim.	74.9	75.1	78.9	<u>75.4</u>
Average width	47.5	57.6	<u>53.9</u>	42.6
CRAE (Knudtson)	48.7	53.7	50.4	<u>51.1</u>
CRVE (Knudtson)	49.0	54.0	50.1	<u>53.9</u>
Distance tort.	23.3	12.6	<u>18.1</u>	12.9
Fractal dimension	<u>85.7</u>	87.5	83.5	83.0
Tortuosity density	<u>24.1</u>	-3.5	14.9	24.3
Vein dist. tort.	-4.6	-15.9	3.9	<u>-3.1</u>

Table 2 • embedding → biomarker (Maestro2) • **bold** best • underline 2nd

• DISTRIBUTION ALIGNMENT



Prediction vs. Maestro2 target (n = 118). Lower W₁/σ (Wasserstein distance) indicates tighter alignment. Full 12-biomarker figure in the paper.

Signal survives the compositional pipeline

76.0%

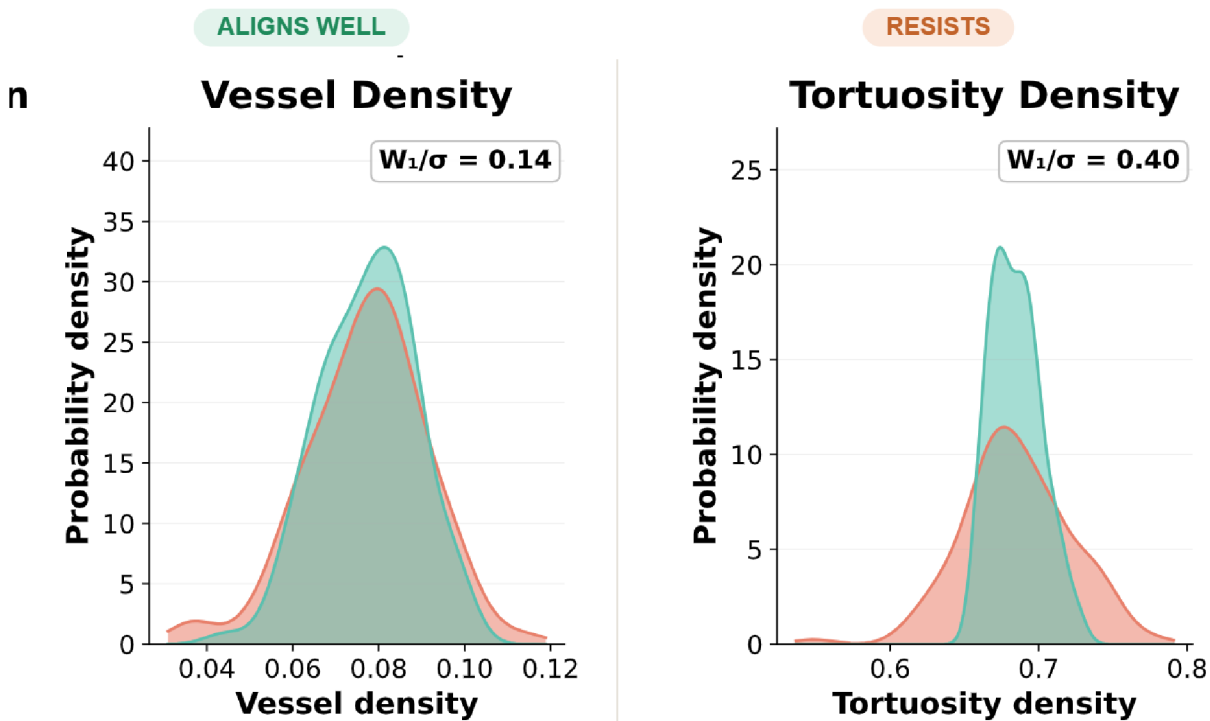
Best end-to-end R^2 • Vessel density
Clinically relevant signal survives the full compositional pipeline.

RESULTS • R^2 (%) • 4 MODELS

Biomarker	TabPFN	CatBoost	L-MLP	Ridge
AVR (Knudtson)	<u>4.2</u>	1.7	5.2	4.0
Artery dist. tort.	<u>6.5</u>	2.3	-0.6	8.5
Artery fractal dim.	<u>57.3</u>	57.0	56.7	57.7
Average width	31.8	33.8	<u>33.1</u>	32.2
CRAE (Knudtson)	44.3	44.2	46.1	<u>45.2</u>
CRVE (Knudtson)	<u>44.5</u>	44.2	43.7	54.0
Distance tort.	21.6	19.1	11.6	<u>19.6</u>
Fractal dimension	<u>71.5</u>	70.3	71.6	70.7
Tortuosity density	<u>19.4</u>	17.6	14.7	19.5
Vein dist. tort.	-8.1	-11.0	<u>-7.1</u>	-6.0
Vein fractal dim.	<u>71.9</u>	69.9	73.1	<u>71.9</u>
Vessel density	76.0	74.7	<u>75.2</u>	75.0

Table 3 • compositional cross-device • **bold** best • underline 2nd

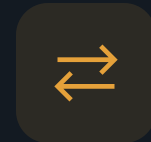
DISTRIBUTION ALIGNMENT



Prediction vs. Triton target ($n = 118$). Lower W_1/σ (Wasserstein distance) indicates tighter alignment. Full 12-biomarker figure in the paper.

INSIGHTS

What it means in practice



Portable

Task 1

Structural biomarkers translate between imaging devices.



So a single model trained on any type of fundus images can serve every downstream camera in the clinic.



Interpretable

Task 2

Opaque embeddings decode into named vascular biomarkers.



So black-box foundation-model features gain clinical, auditable meaning.



Deployable

Task 3

Target-device biomarkers estimated straight from source embeddings.



So a model trained on one device can serve another with no target-device labels.

💡 The signal lives in structural morphology — caliber, fractal dimension, density; tortuosity stays hardest to recover throughout.

CONCLUSION

Robust and interpretable oculomics, no new architecture



Device-agnostic

Feature-space translation removes device-induced variability while preserving biomarker distributions.



Interpretable

RETFound embeddings decode into clinically meaningful vascular biomarkers.



Plug-and-play

Model-agnostic and drop-in — minimal computational overhead, no new architecture.

➤ FUTURE WORK

01 Bidirectional & multi-device

Translate across many devices, in both directions.

02 Larger, more diverse cohorts

Expanded AI-READI and external datasets.

03 Real-world robustness

Hold up on lower-quality, routine-clinic imaging.

TEAM

The people behind the work

Seven researchers bridging retinal imaging and machine learning.



Sparsh Rastogi
TIET · Meesho



Akshat Bakshi
TIET



Jidugu Sriharisesh
TIET



Imran Razzak
MBZUAI



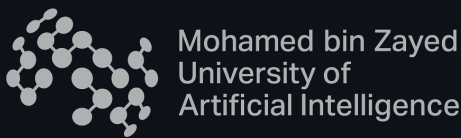
Chinmaya Panigrahy
TIET



Jatin Bedi
TIET



Sahil Thakur
Mediwhale · SERI



Thank you!!

CORRESPONDENCE

✉ srastogi_be22@thapar.edu

✉ sahil.thakur@mediwhale.com

📍 PRESENTED AT ICML 2026 WORKSHOPS

- › Structured Data for Health
- › Foundational Models for Structured Data



 Scan for the paper

