

# MEASURING TRUST

## Why XAI cannot replace clinical diagnosis (yet)

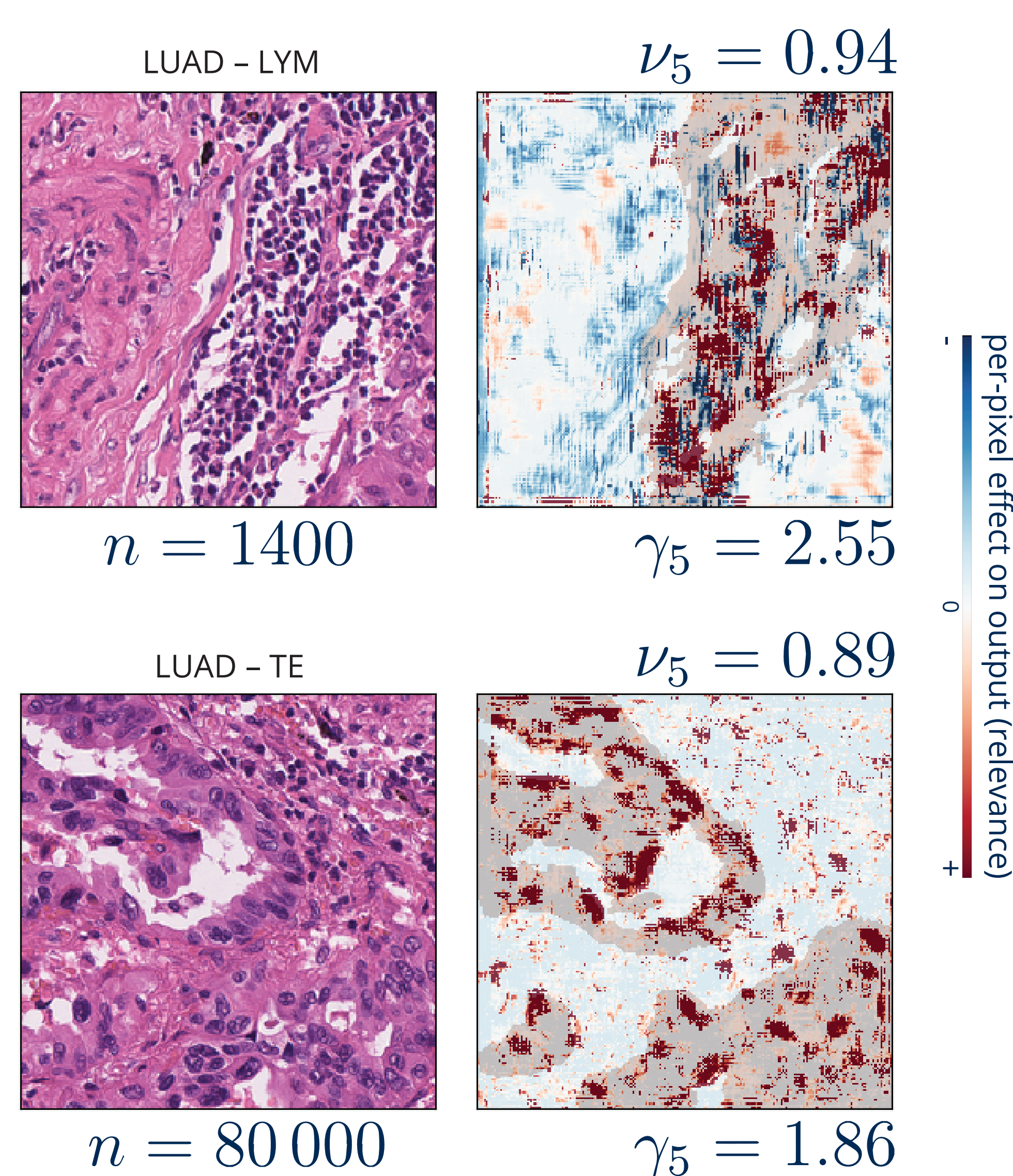


### Explainability is just a buzzword

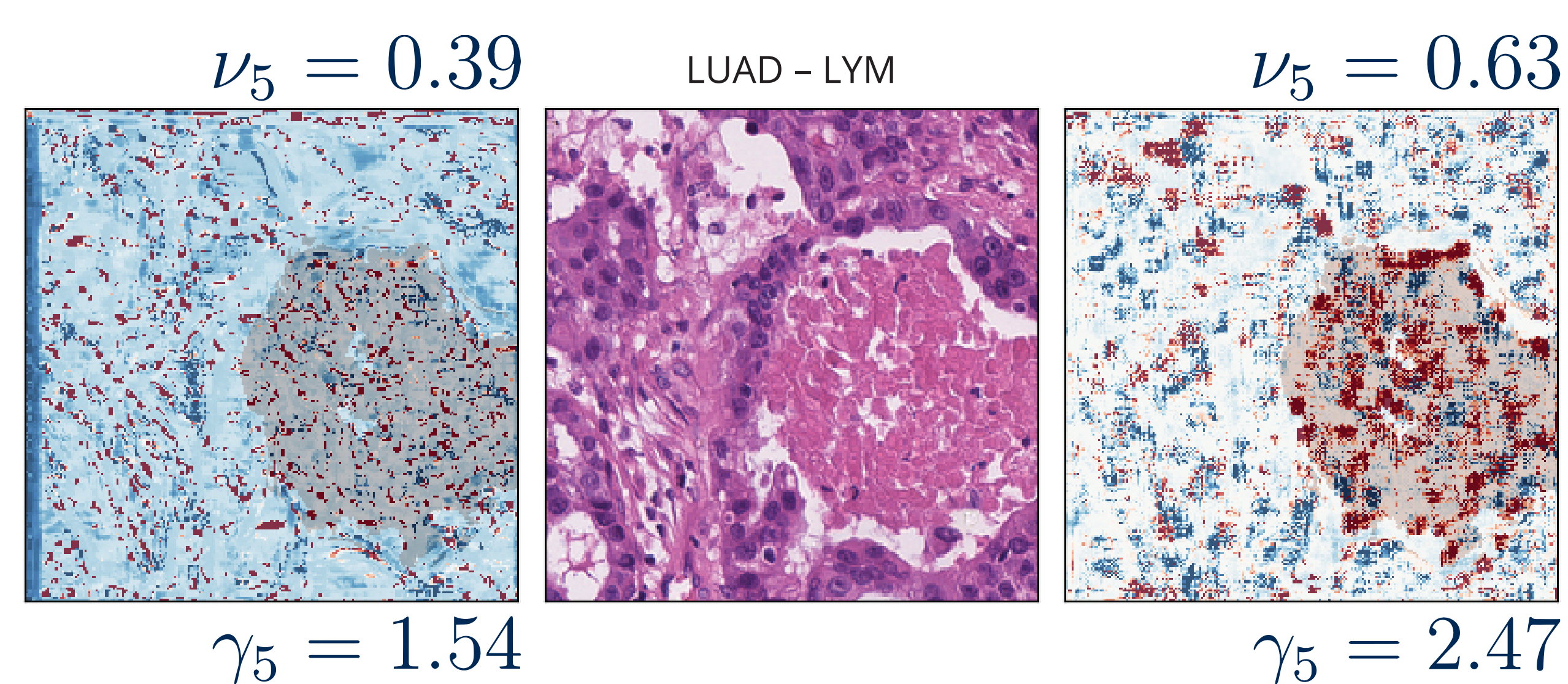
- For real-world application, explainable methods should aim at earning **clinicians' trust**.
- But they are mostly made by engineers for engineers:
  - accuracy is rigorously optimized,
  - clinical validity of decisions is often **overlooked**.

Q: How to **measure** XAI model trustworthiness?

### Optimizing for valid decision-making

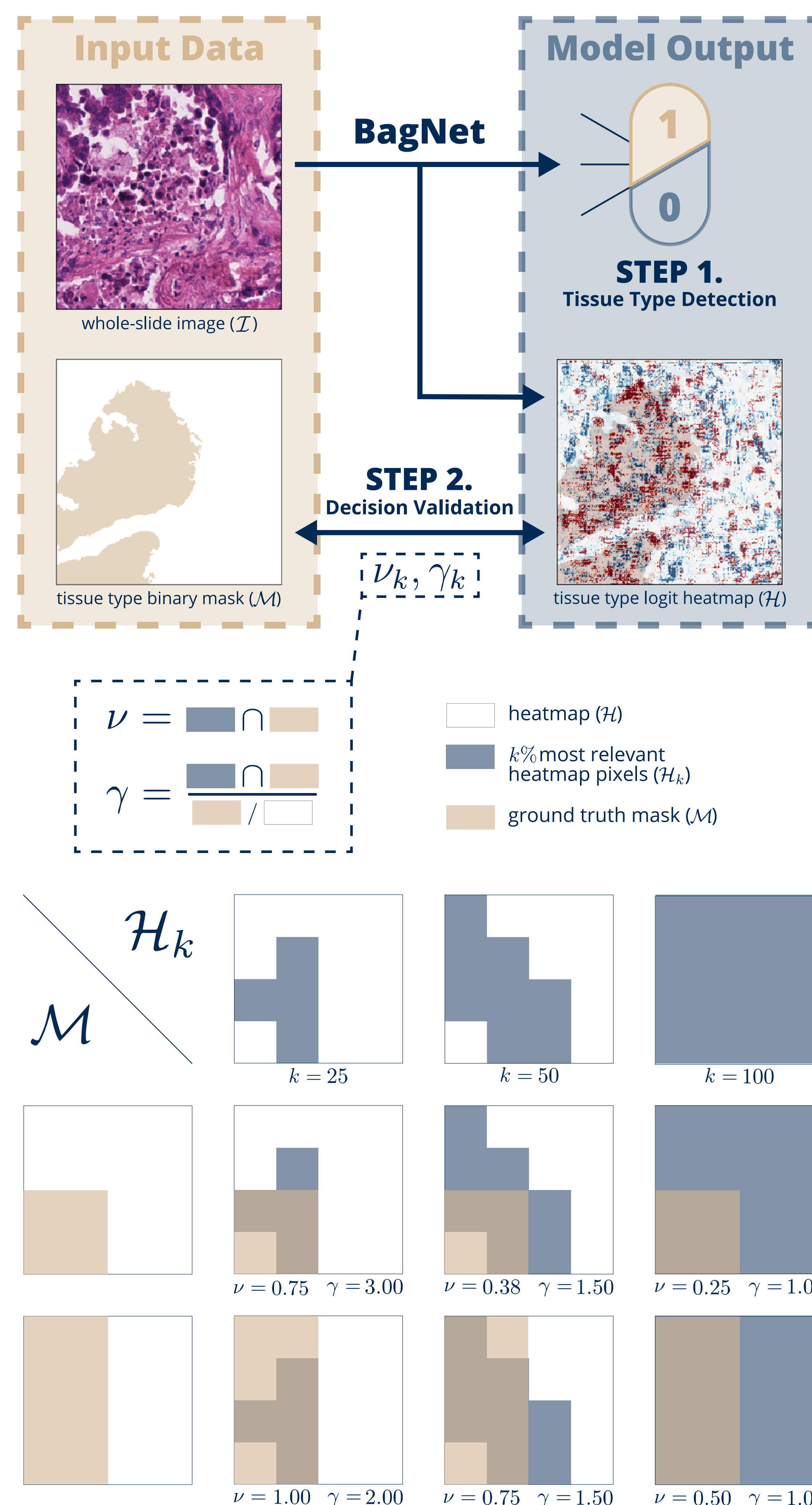


Very large **training data size** ( $n$ ) does not have a significant impact on enforcing clinical decision-making between different tissue types.



**Data augmentation** (right) leads to better correspondence of decisions with clinical guidelines.

### Measuring clinical validity of decisions



**Case study:** tissue type detection in whole-slide scans.

### Quantifying explanations creates trust

- The proposed quantitative scores **measure the clinical validity** of model decisions (i.e. trustworthiness).
- This enables model fine-tuning from a **decision validity perspective**.
- Further usage of validity scores include:
  - comparing trustworthiness of different approaches,
  - fine-tuning models during training.

Gergő Galiger

galiger.gergo@inf.elte.hu

Supported by the ÚNKP-23-3 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund.