DISTILLO

ERGO SUM

HETEROGENEOUS
KNOWLEDGE DISTILLATION
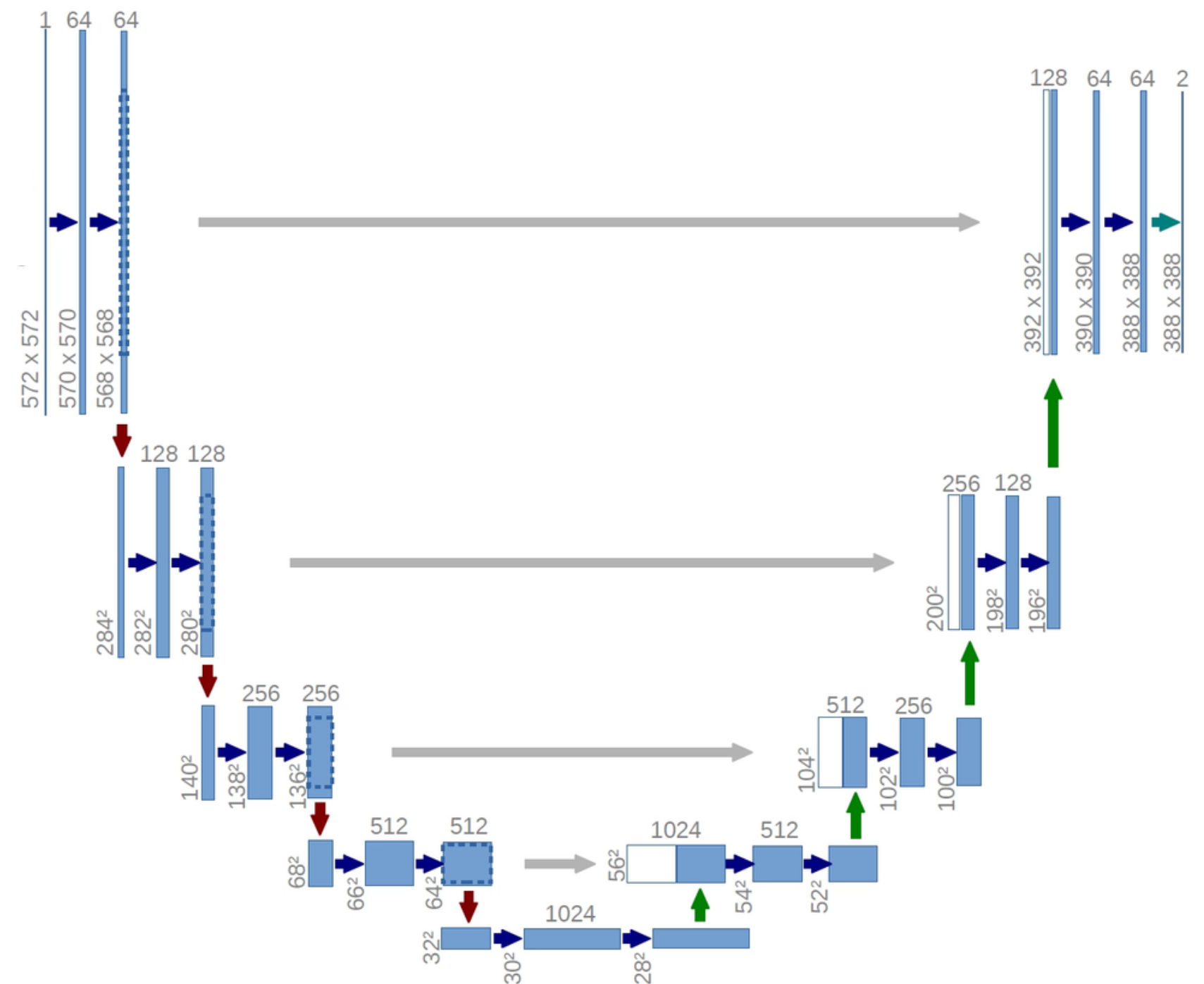### IN MEDICAL COMPUTER VISION

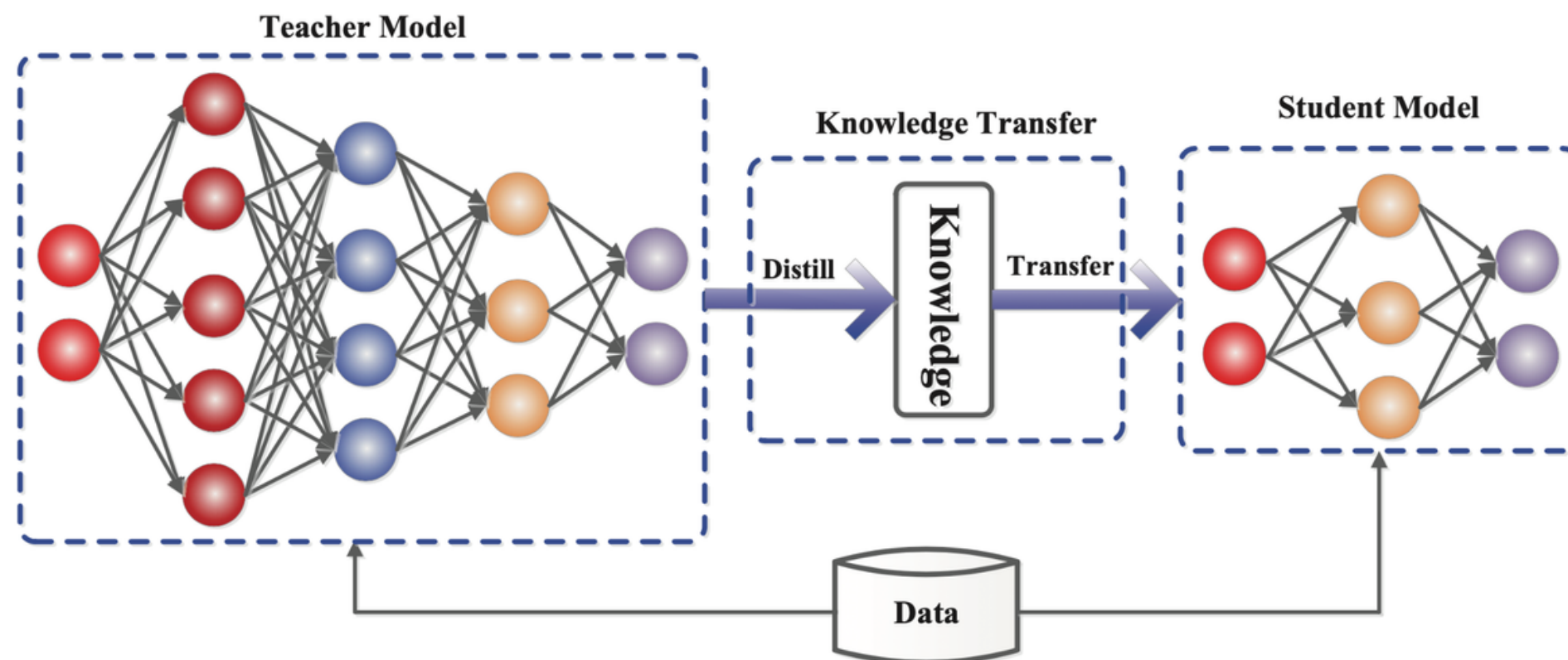Giacomo Bellini, Ludovica Mazza, Matteo Rampolla, Daniele Solombrino

Facts

- **TransUNet**

- **CNNs** can't effectively capture **global context**.

- **Transformer** proficient in capturing global context.

- **Knowledge Distillation:**
  - High performances, few parameters.
  - "Distilling the Knowledge in NNs"
  - "Knowledge Distillation: A Survey"

U-Net architecture

Goals

- **Heterogeneous Knowledge Distillation (HKD):** from Transformer to CNN

- Hypothesis testing:
  - HKD makes manifold **learning invariant** w.r.t. model **inductive bias**.
  - **Impact** of HKD on **contexts** attended by CNN

Our mission

Baseline starting results

Pre-train SqueezeU-Net on ImageNet
Fine-tune SqueezeU-Net on Synapse

→

Heterogeneous knowledge distillation

Transformer → SqueezeU-Net
Logits matching

Quantitative and qualitative **analysis**
I-O-U, Dice score and accuracy **metrics**
Attention visualization
Saliency maps

←

Literature research

Not working?

Global context captured?