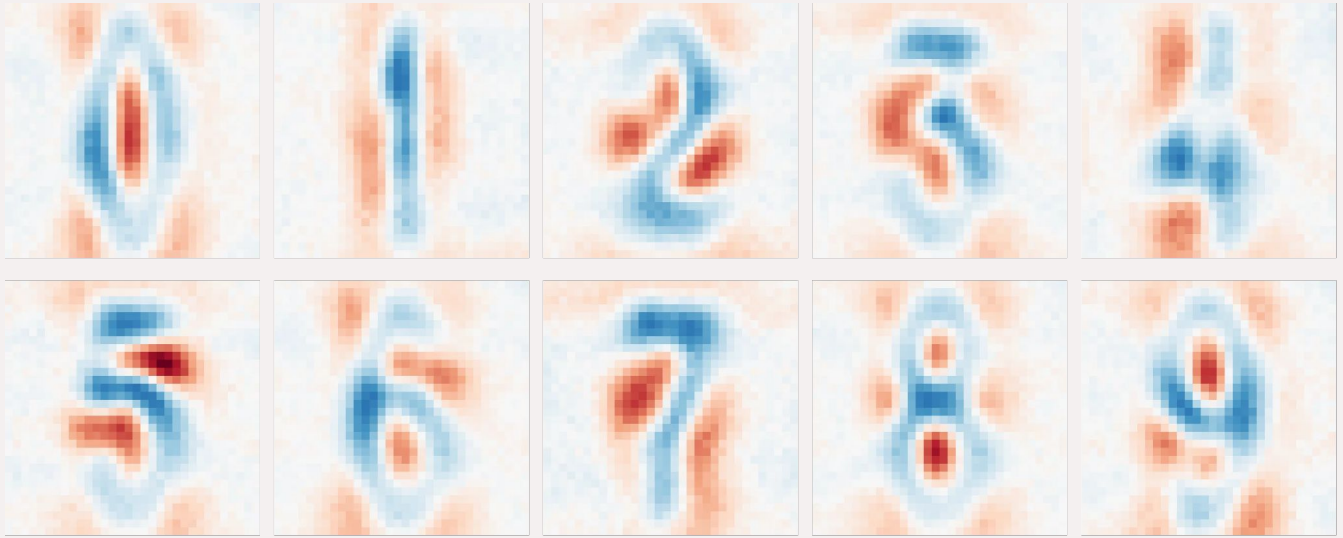
Introducing X-nets

Deep interpretability without compromise

Thomas Dooms & Ward Gauderis

Geraint Wiggins & Jose Oramas

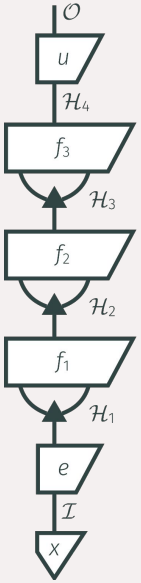
Deep X-net eigenvectors are interpretable. They correspond to prototypical digits.



X-nets combine the strengths of neural and tensor networks.

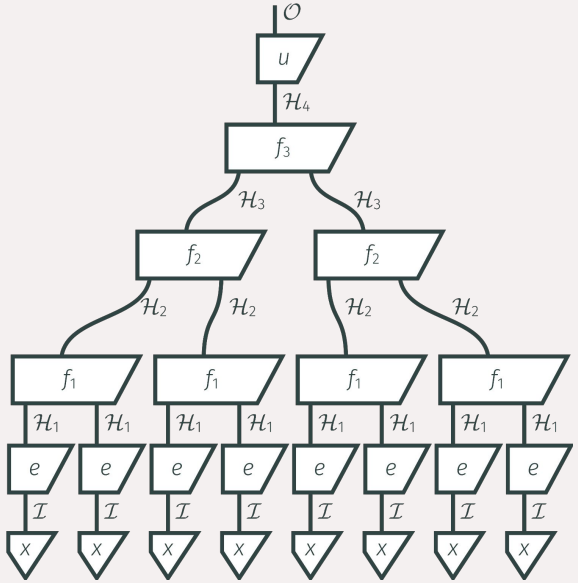
Neural network

Efficient training & evaluation



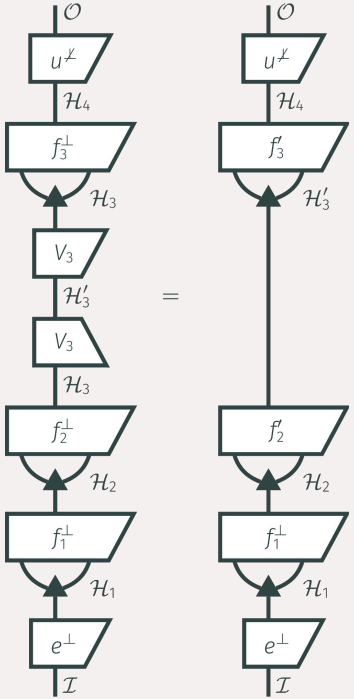
Tensor network

Easy decomposition & interpretation



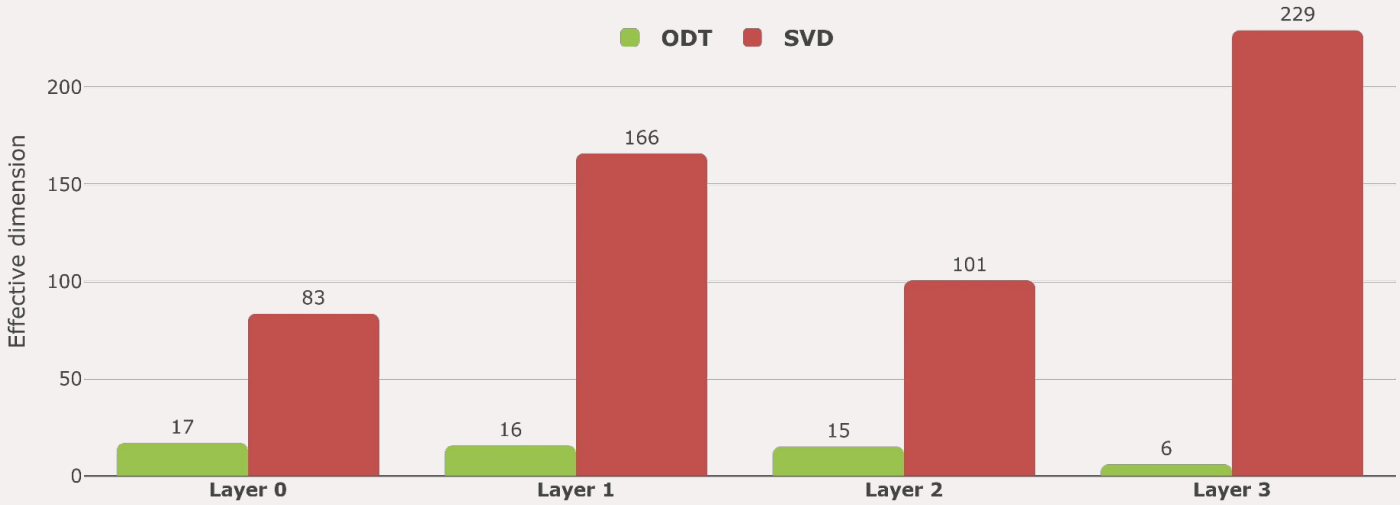
We propose the global ODT algorithm to find the most important dimensions per layer.

Compute eigenvectors
This aggregates global info

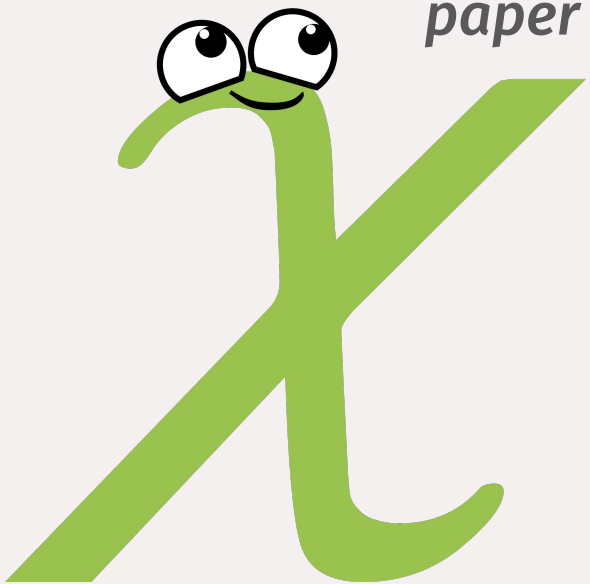


Fold into the network
This diagonalises the layer

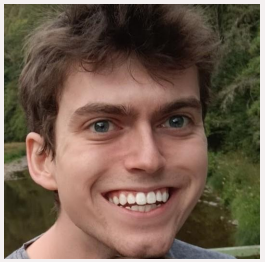
Our algorithm exposes global low-rank structure much better than local SVD.



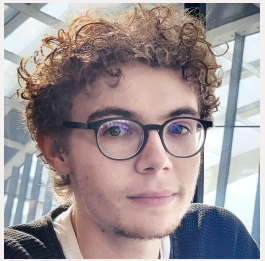
Compositionality unlocks deep interpretable models



Read our paper here!



Thomas Dooms



Ward Gauderis

Universiteit Antwerpen

VUB ARTIFICIAL INTELLIGENCE RESEARCH GROUP