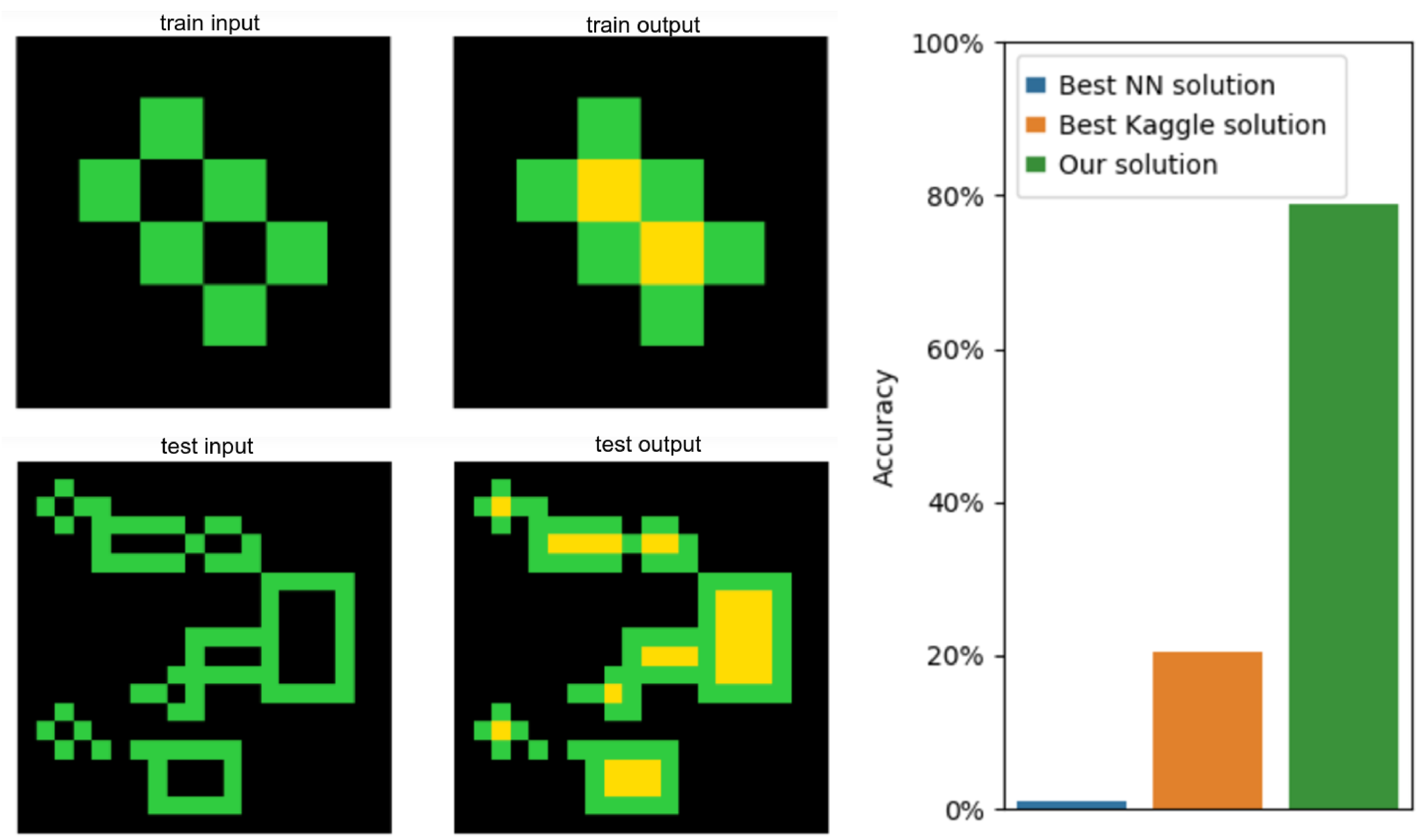
Neural Abstract Reasoner

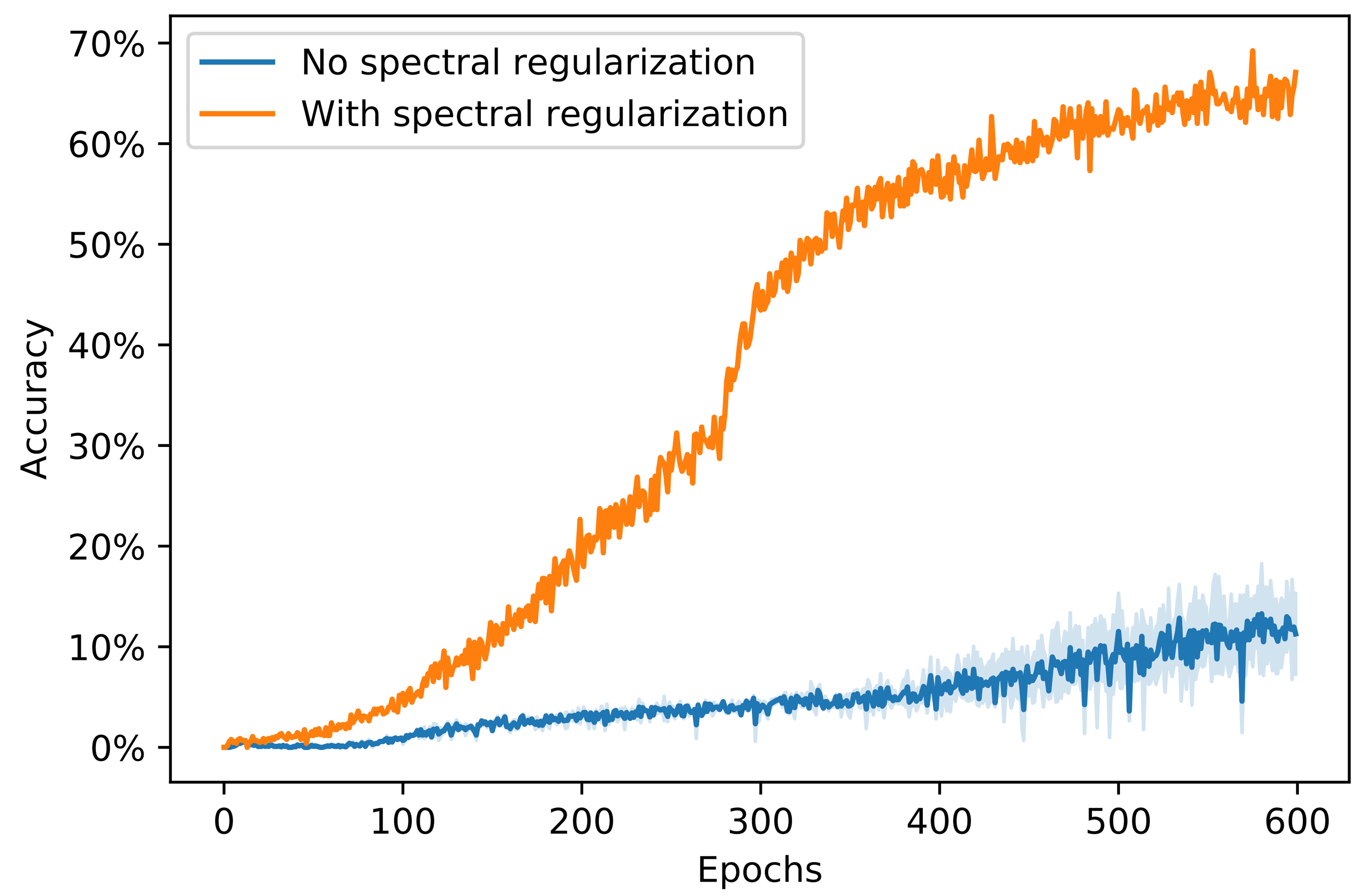
Victor Kolev[1] **Bogdan Georgiev**[2] **Svetlin Penkov**[3]

Sofia High School of Mathematics [1] Fraunhofer IAIS [2] Sciro Reseach [3]

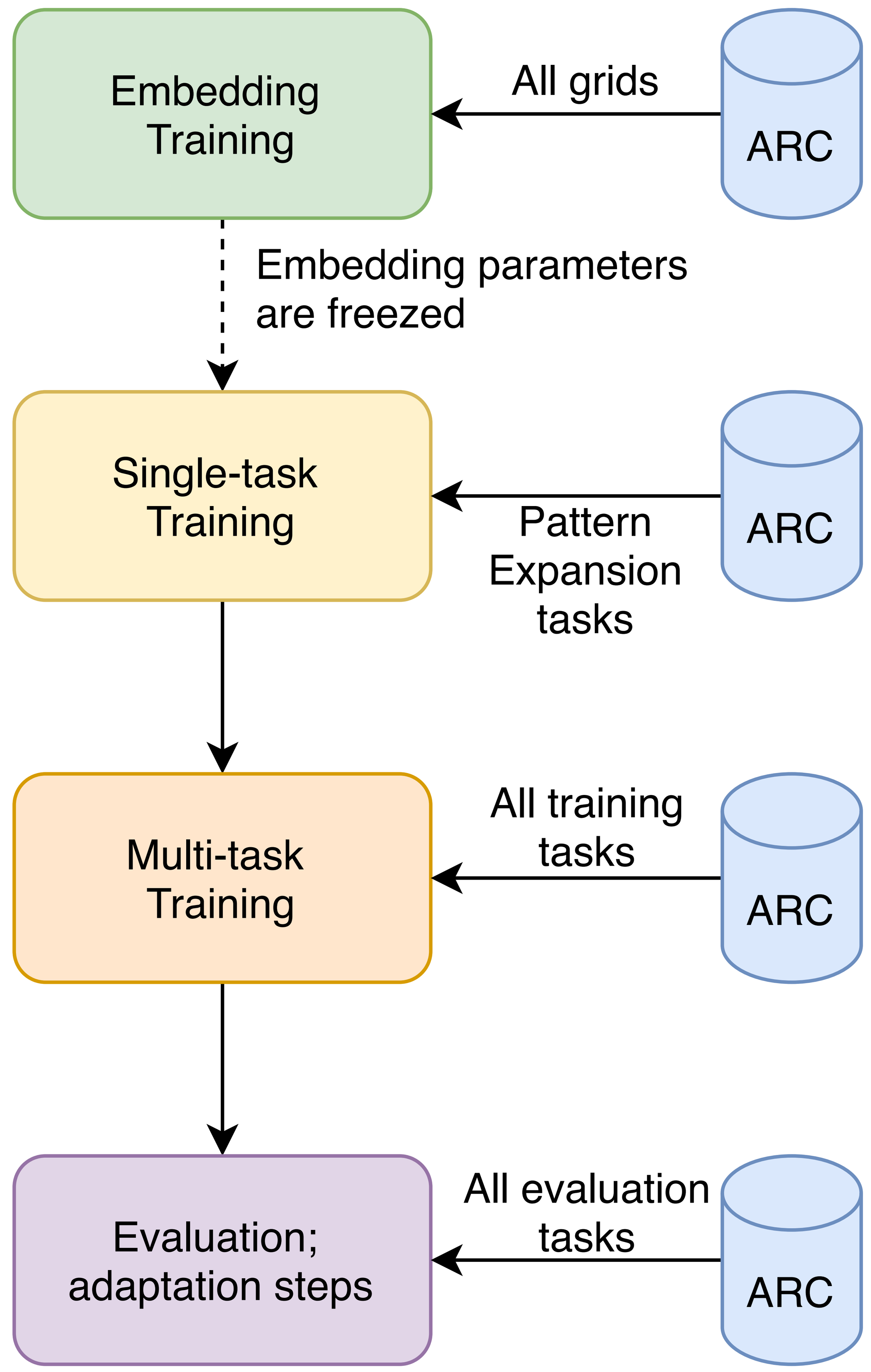
Abstraction and Reasoning Corpus



Spectral Regularization



Training Procedure

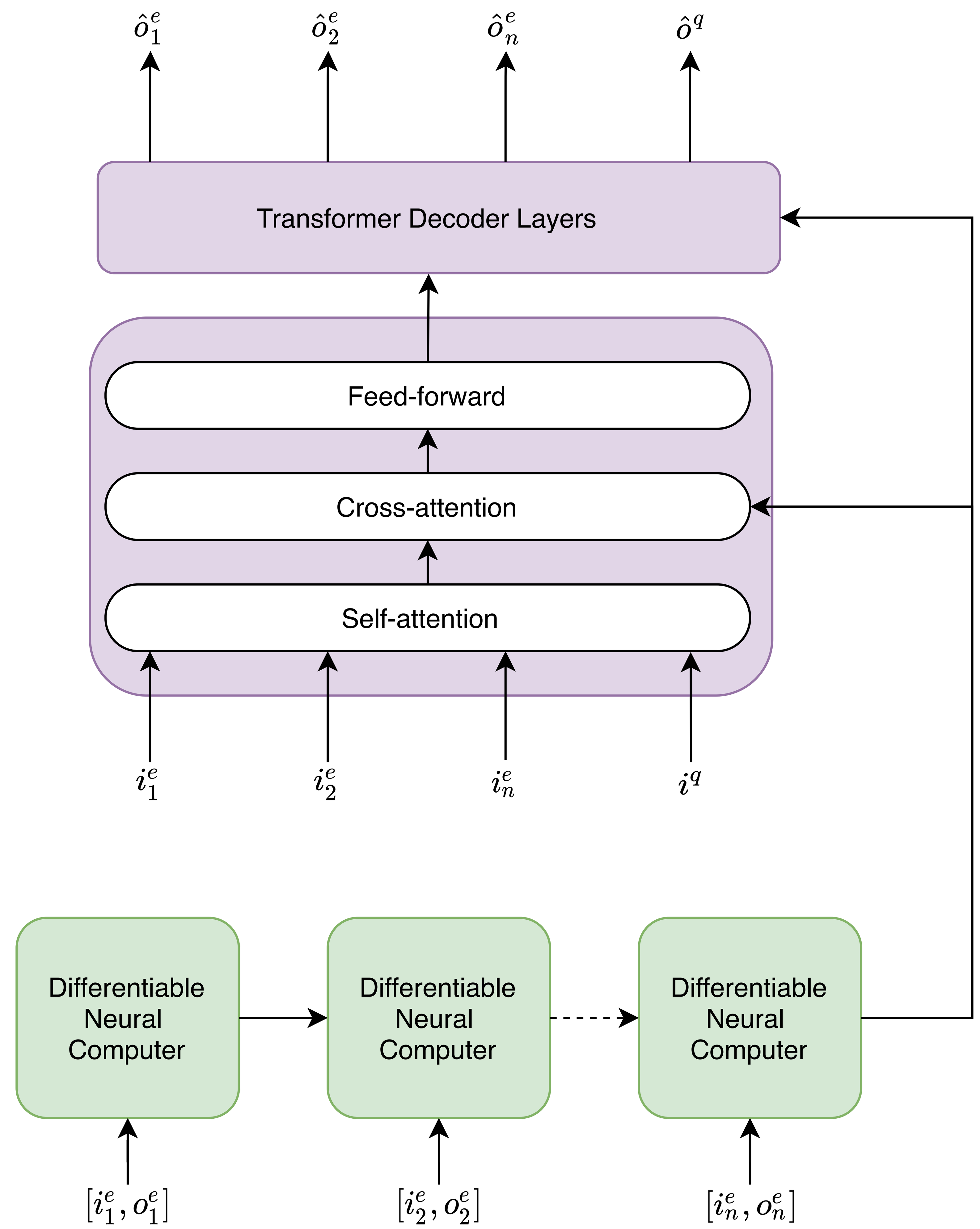


The embedding network is an autoencoder with different convolution sizes to allow for a variety of pattern scales, similar to InceptionNet. We linearly interpolate the different convolution with a separate "relevance" network that places a weight on each kernel size, depending on how relevant it deems it to the final embedding.

We found that, at the last step of the procedure, the additional adaptation steps during evaluation (similar to the MAML algorithm) resulted in zero gradient for the DNC module. We can therefore postulate that spectral regularization enabled the DNC to learn a general algorithm for task acquisition, acting as a memory-based meta-learner.

DNC-Transformer Hybrid Architecture

Example input-output pairs are first processed by the DNC, thus extracting task context (task acquisition step). Subsequently, all inputs, including the query input, are processed by the Transformer Decoder Stack, which cross-attends to the DNC-derived task context. The self-attention mechanisms of the Transformer allow it to extract as much information from the data as possible, since it relates the query input to the example inputs.



LaTeX TikZposter