

Summary

Learning Arbitrary-Graph Tensor Networks for Image Reconstruction

Candidate: Abolfazl Javidian (s314441)

Tensor networks (TN) provide a structured framework for representing high-dimensional tensors using a network of smaller interconnected tensors. In many applications such as image processing and scientific computing, data naturally appears as high-order tensors whose size grows exponentially with the number of dimensions. Directly storing or processing these tensors quickly becomes computationally infeasible. Tensor networks address this challenge by factorizing a large tensor into a set of lower-order tensors connected through shared indices. This factorized representation dramatically reduces the number of parameters while preserving important structural relationships within the data.

A number of tensor network architectures have been proposed to exploit this idea. One of the most widely used models is the Tensor Train (TT) decomposition, also known as the Matrix Product State representation. In TT decomposition, a high-order tensor is represented as a sequence of third-order tensors connected in a linear chain. Each connection has an associated rank parameter that determines the expressive capacity of the model. TT decomposition is computationally attractive because it allows efficient tensor contractions and scalable optimization algorithms. However, the linear chain structure imposes strong structural constraints on how tensor dimensions interact, which can limit the model's ability to capture complex dependencies present in natural images.

Another important tensor network architecture is the Tensor Ring (TR) decomposition. Tensor Ring extends the Tensor Train model by connecting the first and last cores to form a cyclic structure. This circular topology removes the boundary effects that appear in TT and allows information to propagate more freely across the network. While Tensor Ring models can improve representation flexibility compared to TT, they still rely on a fixed topology that must be specified in advance. As a result, the model structure may not always match the intrinsic structure of the data.

More recent approaches aim to move beyond fixed network structures. Methods such as Tensor Network Graph Algorithms (TNGA) and Tensor Network Learning with Adaptive Edges (TnLAE) attempt to discover tensor network structures automatically rather than relying on predefined architectures. These approaches treat the topology of the tensor network as a variable that can be optimized alongside the tensor parameters. However, discovering an efficient arbitrary-graph tensor network remains a challenging problem because the number of possible graph structures grows combinatorially. In practice, heuristic strategies such as greedy topology search are often used. Greedy methods iteratively modify the network by adding or removing edges based on local improvements in an objective function.

Despite these advances, existing approaches still suffer from several limitations. First, most methods assume predefined structures or restrict the search space to a limited set of candidate topologies. Second, topology search often relies on simple heuristics that do not explicitly balance multiple objectives such

as reconstruction quality and compression efficiency. Third, training procedures for tensor networks can be unstable when the model structure changes during optimization. These challenges motivate the development of more systematic approaches for topology learning in tensor networks.

The goal of this thesis is to develop a framework for learning arbitrary-graph tensor network structures for image reconstruction. The problem begins with the tensorization of images. Each grayscale image is resized to a resolution of 256×256 , resulting in 65,536 pixels. Instead of treating the image as a simple matrix, the image is reshaped into a higher-order tensor representation. Tensorization allows the image to be decomposed using tensor network techniques while preserving spatial relationships between pixels.

The reconstruction objective is formulated as an optimization problem in which the tensor network parameters are chosen to approximate the original image tensor. Given a tensor network representation $\hat{X}$ of the image tensor X , the reconstruction error is measured using the mean squared error between the original and reconstructed tensors. The goal is therefore to find a tensor network structure and parameters that minimize this reconstruction error while also maintaining strong compression.

In the proposed framework, the tensor network is represented as an arbitrary graph whose nodes correspond to tensor cores and whose edges represent contracted indices. Each edge is associated with a rank that determines the dimensionality of the corresponding connection. Unlike classical tensor network models such as TT or TR, this representation does not impose a fixed topology. Instead, the connectivity between tensor cores can vary, allowing the network to capture complex interactions between tensor dimensions.

To evaluate candidate tensor network structures, three categories of evaluation metrics are considered. The first category is reconstruction accuracy, measured using the mean squared error between the original and reconstructed images. The second category is compression efficiency, which measures the number of parameters required by the tensor network representation relative to the original image tensor. The third category consists of perceptual quality measures that assess how well the reconstructed image preserves visually meaningful structures.

Because these metrics often conflict with each other, the thesis introduces a robust scalarization method based on a Min–Max objective formulation. Instead of combining metrics through a weighted sum, the Min–Max approach minimizes the worst normalized objective among the considered metrics. Formally, if f_1 , f_2 , and f_3 denote normalized measures of reconstruction error, parameter count, and perceptual distortion, the optimization objective can be written as minimizing

$$\max\{f_1(\theta), f_2(\theta), f_3(\theta)\}$$

where θ denotes the tensor network parameters and topology. This formulation encourages balanced solutions that perform reasonably well across all evaluation criteria. The Min–Max strategy prevents degenerate solutions that achieve extremely strong compression at the cost of very poor reconstruction quality, or vice versa. By focusing on the worst performing metric, the optimization process produces

robust solutions that maintain acceptable performance across multiple objectives.

To explore the space of possible tensor network structures, a topology search algorithm is proposed. The algorithm begins with an initial network structure and iteratively proposes modifications such as adding edges, removing edges, or adjusting ranks. Each candidate topology is evaluated using the Min–Max objective. If the modification improves the objective value, it is accepted and the search continues from the new structure.

A key component of the search process is feasibility-first local exploration. Before evaluating a candidate topology, the algorithm checks whether it satisfies structural constraints such as rank bounds, parameter budgets, and graph connectivity conditions. Only feasible candidate structures are evaluated using the objective function. This strategy reduces computational cost by eliminating invalid configurations early in the search process and focusing the evaluation on promising candidates.

The optimization of tensor parameters for a fixed topology is performed using a hybrid training routine. Initially, the Adam optimizer is used because of its strong performance during early training stages. Adam provides adaptive learning rates that help stabilize optimization when gradients vary significantly. After this exploratory phase, the optimization switches to the L-BFGS algorithm, a quasi-Newton method that often achieves faster convergence near optimal solutions. To ensure numerical stability, gradient clipping is applied to prevent excessively large parameter updates.

Another important feature of the proposed framework is topology adaptation during training. Instead of performing topology search only once, the algorithm periodically revisits the network structure and evaluates whether alternative topologies could improve the objective function. If a better structure is discovered, the network is updated accordingly and parameter optimization continues. This dynamic adaptation allows the model to progressively refine its structure while learning the tensor parameters.

Experimental results demonstrate that the proposed framework can discover tensor network structures that achieve strong trade-offs between compression and reconstruction quality. Compared to classical architectures such as Tensor Train and Tensor Ring, the learned arbitrary-graph tensor networks often achieve similar reconstruction accuracy while using significantly fewer parameters. The Min–Max objective plays a crucial role in maintaining balanced performance across multiple evaluation metrics.

In summary, this thesis introduces a systematic framework for topology learning in tensor networks applied to image reconstruction. By combining arbitrary-graph parameterization, multi-objective optimization through Min–Max scalarization, and efficient topology search strategies, the proposed method provides a flexible approach for discovering efficient tensor network representations.

Future work may explore more advanced topology search strategies, including stochastic optimization or reinforcement learning methods for structure discovery. Another promising direction is integrating tensor network representations with deep learning architectures to combine the interpretability and compression benefits of tensor networks with the representational power of neural networks.