Quantum-assisted learning for convolutional deep belief networks

Gabriel M. Bianconi[1,2] Maxwell P. Henderson[2] Mykel J. Kochenderfer[1]

[1]Stanford University [2]QxBranch

Abstract

Problem: Practical applications of quantum Boltzmann machines have been limited by the architecture of existing quantum annealers.

Goal: We present a novel quantum-assisted learning algorithm for convolutional deep belief networks with the goal of enabling larger problems to be modeled.

Results: While our approach benefits from faster convergence during the pre-training step, classification accuracy at the discriminative step does not outperform the classical implementation.

Motivation

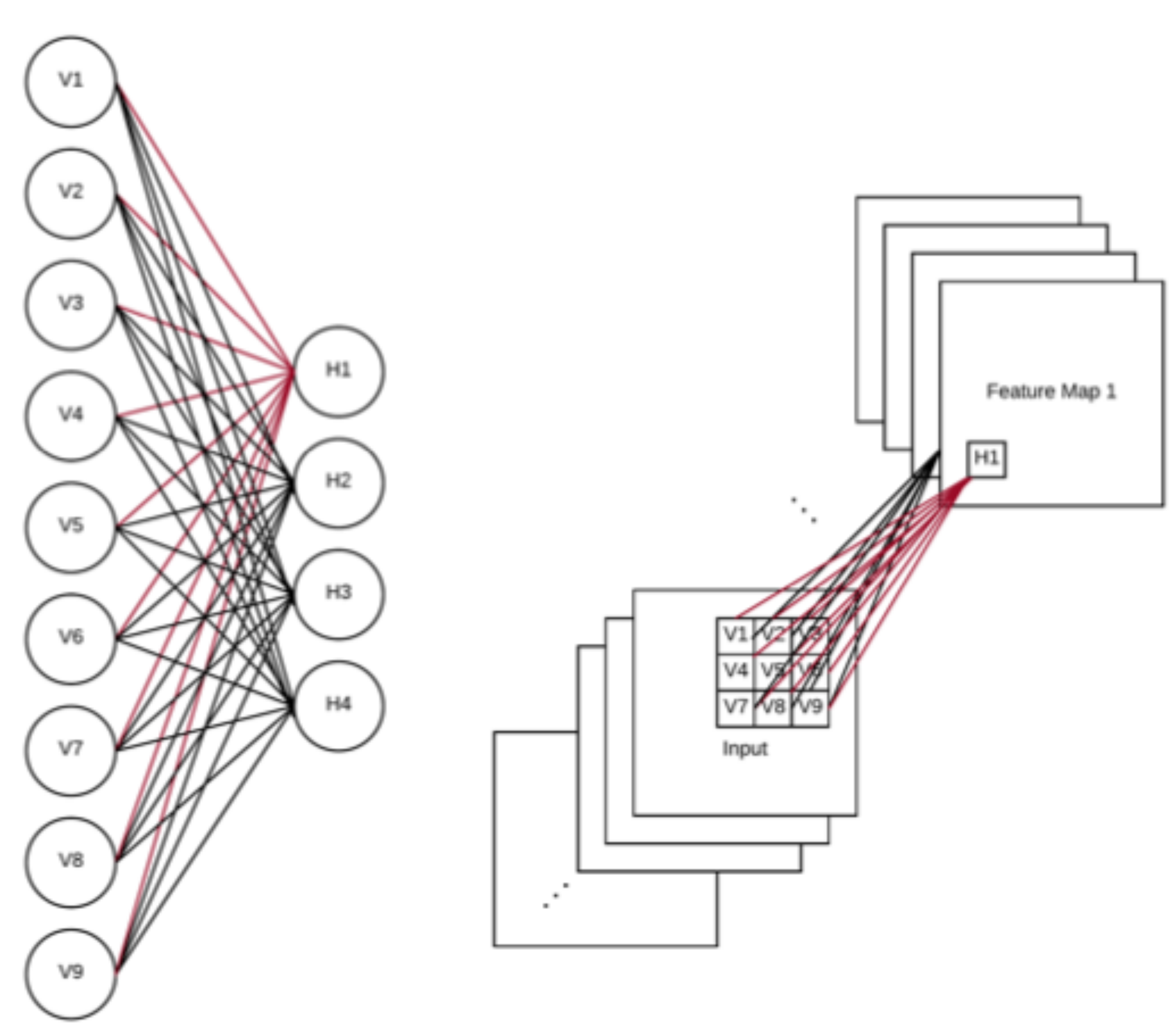
- Previous research investigated whether quantum sampling can help train Boltzmann machines. Adachi & Henderson (2015) proposed an embedding scheme for Restricted Boltzmann Machines. They showed that a quantum-assisted (hybrid) approach outperformed its classical implementation.
- Limitations in current quantum annealers hinder practical applicability. Embedding Boltzmann machines onto the hardware requires a compromise between connectivity and model size. Until larger hardware becomes available, such approaches are unlikely to be applicable to large-scale problems.
- With the limitation posed by problem size in mind, we present a novel network architecture and hybrid learning algorithm inspired by convolutional deep belief networks (Lee et al., 2009) that incorporate relatively small Boltzmann machines to learn models for large problems, particularly in computer vision.

References:

- Adachi, Steven H., and Maxwell P. Henderson. "Application of quantum annealing to training of deep neural networks." *arXiv preprint arXiv:1510.06356* (2015).
- Lee, Honglak, et al. "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations." *International Conference on Machine Learning* (2009).

Model

- Our goal is to create an architecture that can learn arbitrarily large problems by splitting the problem into small sub-problems that can benefit from quantum annealing.
- We construct a small RBM that we convolve through the input image, recording the hidden activations at each step. Each of the hidden units is an extracted featured for the processed patch; we then compose the extracted features as feature maps. As a result, the RBM acts as a feature extractor for arbitrarily large inputs, while the RBM size stays constant.

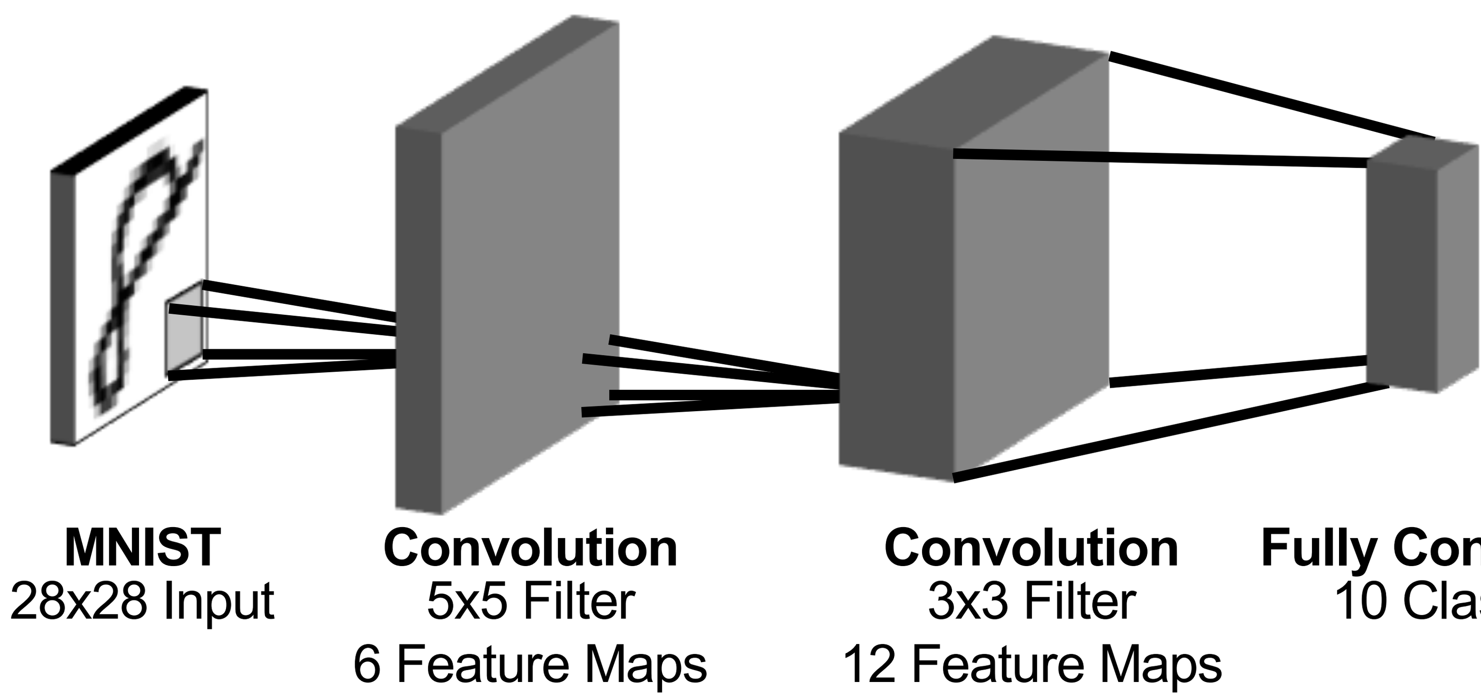


- A deep network can be composed by stacking multiple RBMs. The deeper layers learn progressively more abstract features, and increase the representation power of the model.
- We use quantum sampling during pre-training, and subsequently fine-tune the model with backpropagation or fit a classifier on the hidden activations of the last layer.

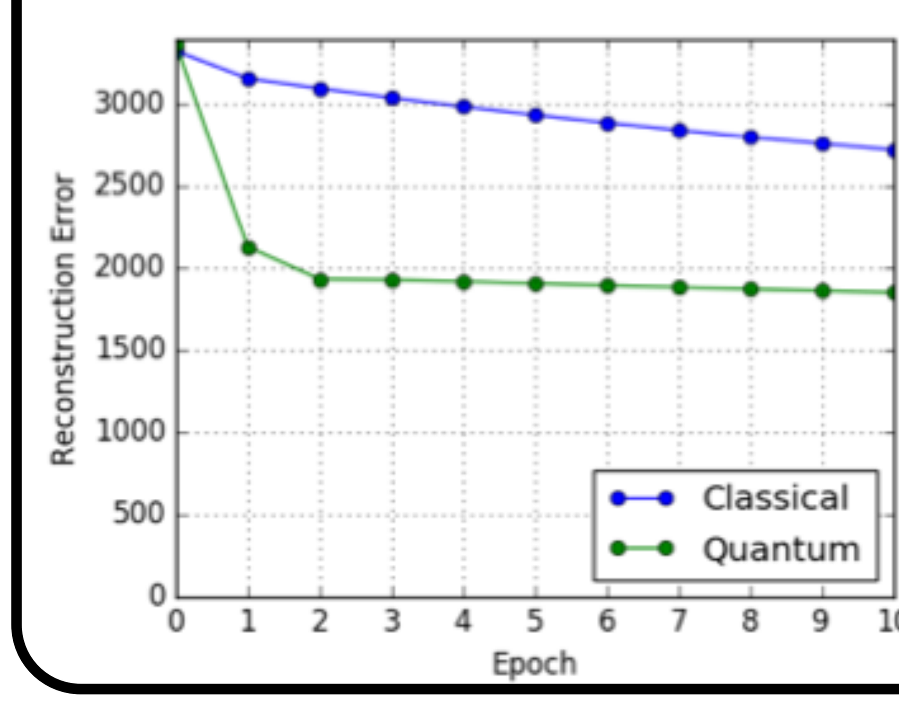
Pre-Training		
Pre-Training Extract Features *Unsupervised*	⇒	**Fine-Tuning** Fully Connected Layer & Backprop *Supervised*
	⇒	**Fitting** Classifier on Hidden Activations *Supervised*

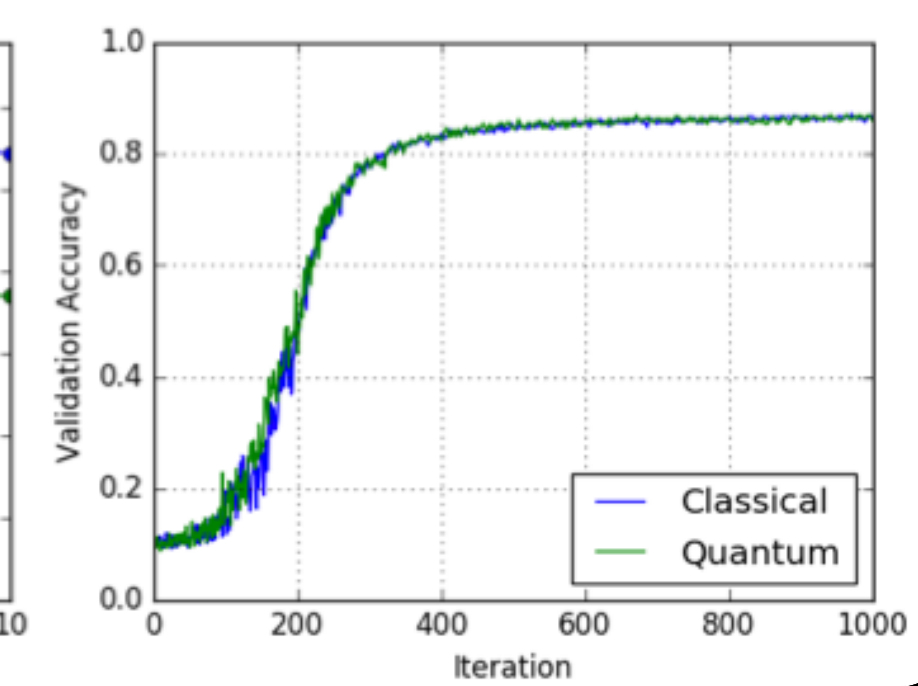
Experimental Results & Discussion

- We compared models pre-trained with contrastive divergence and quantum sampling in the D-Wave 2X.



- Reconstruction error of the filters fell considerably faster with quantum sampling during pre-training.
- Validation accuracy after fine-tuning the models (or applying logistic regression to hidden activations) was similar.
- In a classical setting, pre-training beyond one epoch does not lead to a stronger classifier, suggesting that quantum sampling would not have led to a substantial benefit.





Conclusion

- We presented a convolutional architecture for deep belief networks that aimed to leverage existing quantum annealers to train models on large vision problems.
- While our hybrid model reduced its reconstruction error faster during pre-training, it did not outperform a classical implementation during the discriminative step.
- Further work is required to explore in which conditions, if any, faster pre-training convergence leads to stronger classifiers.