

Shot-Based Quantum Encoding: A Data-Loading Paradigm for Quantum Neural Networks

Basil Kyriacou, Viktoria Patapovich, Maniraman Periyasamy, Alexey Melnikov
Terra Quantum AG, St. Gallen, Switzerland

Efficient data loading remains a bottleneck for near-term quantum machine learning: angle, amplitude, and basis encodings either underuse the exponential Hilbert-space dimension or require circuit depths beyond noisy intermediate-scale quantum (NISQ) coherence budgets. We introduce Shot-Based Quantum Encoding (SBQE), which embeds a datum $\mathbf{x}$ in a data-dependent classical distribution $\mathbf{p}(\mathbf{x})$ over a fixed basis-state pool, realised by allocating shots rather than coherent gates. The encoded state is the mixed state $\rho(\mathbf{x}) = \text{diag}(\mathbf{p}(\mathbf{x}))$, and measured expectation values are linear in $\mathbf{p}(\mathbf{x})$; they therefore compose with classical nonlinearities, so a single SBQE layer is structurally equivalent to a perceptron layer whose weights are realised by a variational circuit. No data-dependent encoding gates are used, so the coherent depth is set solely by the variational ansatz.

The map from datum to distribution is learned rather than prescribed. For image inputs, each image is reduced to its $d = 8$ leading principal components, rescaled to the unit interval, passed through a trainable affine layer $\mathbf{z} = W\mathbf{x} + \mathbf{b} \in \mathbb{R}^{2^q}$, and projected onto the probability simplex,

$$p_j(\mathbf{x}) = \frac{|z_j|}{\sum_k |z_k|}, \quad p_j \geq 0, \quad \sum_j p_j = 1.$$

Taking the initial-state pool to be the 2^q computational-basis states ($q = 8$, so $2^q = 256$), the entries of $\mathbf{p}(\mathbf{x})$ are the populations of the basis states, and $\rho(\mathbf{x}) = \text{diag}(\mathbf{p}(\mathbf{x}))$ is realised physically by starting a fraction p_j of the N_{tot} shots in $|j\rangle$ (prepared with single-qubit X gates) before the variational circuit is applied. Applying $U(\boldsymbol{\theta})$ and measuring observables $\{O_i\}$ gives

$$f_i(\mathbf{x}, \boldsymbol{\theta}) = \sum_j W_{ij}(\boldsymbol{\theta}) p_j(\mathbf{x}), \quad W_{ij}(\boldsymbol{\theta}) = \langle j | U^\dagger(\boldsymbol{\theta}) O_i U(\boldsymbol{\theta}) | j \rangle,$$

a map that is linear in $\mathbf{p}(\mathbf{x})$ and mirrors a classical perceptron, except that the weight matrix $W(\boldsymbol{\theta})$ is read out from the circuit rather than stored directly. The affine weights $W, \mathbf{b}$ and the circuit angles $\boldsymbol{\theta}$ are trained jointly by back-propagation, and stacking layers through a probability-preserving activation gives a multilayer network with all weights in $\boldsymbol{\theta}$.

We benchmark SBQE against amplitude encoding and a width-matched classical network on three image-classification tasks (Semeion, Fashion-MNIST, and MNIST) under an identical protocol: eight qubits, four variational layers of single-qubit $\text{Rot}(\alpha, \beta, \gamma)$ rotations followed by a linear CNOT chain, the classical optimiser over 20 000 epochs on stratified 70/15/15 splits, and ten random initialisations with best-validation weights, all as exact statevector simulations; the three models are matched for trainable-parameter count.

SBQE reaches $89.1\% \pm 0.9\%$ on Semeion, $80.95\% \pm 0.10\%$ on Fashion-MNIST, and $90.25\% \pm 0.18\%$ on MNIST (mean $\pm$ standard deviation over ten seeds). It exceeds amplitude encoding on all three datasets (+2.0 and +2.1 percentage points on Fashion-MNIST and MNIST, a 5.3% relative error reduction on Semeion; two-tailed paired t -test $p < 0.01$, and $p \approx 2 \times 10^{-10}$ on MNIST) and matches or slightly exceeds the classical network. To isolate the circuit’s role, we replace it with a same-shape classical low-rank map of $\mathbf{p}(\mathbf{x})$ and sweep its rank: the quantum mixing layer reaches its accuracy with only 120 trainable angles, matching classical maps with an order of magnitude more parameters, whereas the smallest reach only 57–67%. The class separation is therefore carried by the circuit, not by the dimensionality reduction.

Shot allocation thus offers a practical, gate-free channel for data encoding on near-term hardware. The poster develops the encoding scheme and perceptron correspondence, a worked example on real data, the full benchmark suite, and the finite-shot and scaling considerations for hardware.

Corresponding author: map@terraquantum.swiss.