Feature Selection and Feature Extraction for Satellite Hyper-spectral Imagery Data

Hyperspectral imaging, utilized extensively in Earth observation, produces copious amounts of data containing intricate spectral information across wide range of wavelengths. When coupled with machine learning, this data aids in inferring critical details about environment. However, the sheer volume of information generated by hyperspectral imaging poses challenges due to its high dimensionality, making it cumbersome for machine learning algorithms. To alleviate this issue, preprocessing becomes crucial. Feature selection and feature extraction methods play a central role in this process, allowing to condense the data while preserving its essential characteristics. By identifying and extracting relevant features or reducing redundant ones, these techniques effectively streamline the data, enhancing computational efficiency and maintaining the data's relevance and utility for subsequent machine learning tasks. It is expected that quantum computing will improve the feature selection and feature extraction procedures. The gain from utilizing quantum devices may come from two sources: speedup in running classical hard problems on quantum computers; and more accurate predictions about relevant features deriving from non-trivial correlations found with the use of quantum device.

In the study of the use case, six possible candidates for utilizing quantum computing for feature selection and feature extraction of hyperspectral data were identified. The two **long-term** candidates which promise exponential or strong polynomial speedups are: Recursive Feature Elimination (RFE) based on Quantum Support Vector Machines (QSVM); and Quantum Principal Component Analysis (QPCA). The above methods utilize best quantum algorithms in terms of computational complexity, yet require the technological advancement of quantum devices which probably will not be achieved in the next 15 years. The four **mid-term** candidates are expected to be possibly realized in the coming 3-5 years. The candidates rely mostly on the hybrid quantum-classical routines, which distribute tasks accordingly to the classical and quantum devices. They include variational version of the above long-term algorithms: variational RFE and variational QPCA. Additionally, two other **mid-term** candidates were identified: quantum optimization which can be also implemented on the quantum annealing device; and the quantum variational version of the neural autoencoder architecture.

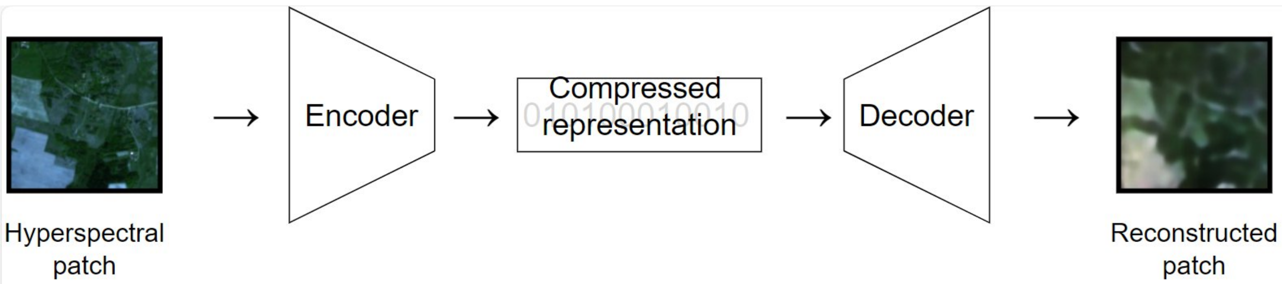

Figure I: Quantum feature extraction and selection pipeline for processing satellite images shown in the compressed representation pipeline.

Table I: We presented the main quantum parameters of a superconducting quantum machine for a quantum feature selection algorithm. [theoretical speculation, subject to change]. Here, theoretical speculation refers to gate error rate and gate fidelities, whereas the number of physical qubits are promised the quantum roadmap provided by industry.

	Quantum Optimization	RFE for QSVM	RFE for VQAs
Time horizon	present day	3-5 years	5-15 years
Quantum platform	anealing/gate-based	gate-based	gate-based
Resource requirement	low/moderate	high	moderate
Number of physical qubits	$> 10^5$	$\sim 10^2$	$\sim 10^2$
Error rate/threshold error rate (p/p_th)	0.0149	0.0051	0.1823
1Q Fidelity	99.99%	99.99%	>99.99%
2Q Fidelity	99.97%	99.98%	>99.99%

Table II: We presented the main quantum parameters of a superconducting quantum machine for a quantum feature extraction algorithm. [theoretical speculation, subject to change]. Here, theoretical speculation refers to gate error rate and gate fidelities, whereas the number of physical qubits are promised the quantum roadmap provided by industry.

	Variational QPCA	QAutoencoders	QPCA
Time horizon	Present day	3-5 years	5-15 years
Quantum platform	gate-based	gate-based	gate-based
Resource requirement	moderate	moderate	high
Number of physical qubits	$\sim 10^2$	$\sim 10^2$	$\sim 10^2$
1Q Fidelity	99.99%	99.99%	>99.99%
2Q Fidelity	99.97%	99.98%	>99.99%