

CompressedMediQ: Hybrid Quantum Machine Learning Pipeline for High-Dimensional Neuroimaging Data

Kuan-Cheng Chen^{†‡*}, Yi-Tien Li^{§¶}, Tai-Yu Li^{||}, Chen-Yu Liu^{**}, Po-Heng (Henry) Lee[‡] Cheng-Yu Chen^{§‡‡^{xxi}}

[†]Department of Electrical and Electronic Engineering, Imperial College London, London, UK

[‡]Centre for Quantum Engineering, Science and Technology (QuEST), Imperial College London, London, UK

[§]Translational Imaging Research Center, Taipei Medical University Hospital, Taipei, Taiwan

[¶]Research Center for Neuroscience, Taipei Medical University, Taipei, Taiwan

^{**}Ph.D. Program in Medical Neuroscience, College of Medical Science and Technology,

Taipei Medical University, Taipei, Taiwan

^{||} National Center for High-Performance Computing, NARLabs, Hsinchu, Taiwan

^{**}Graduate Institute of Applied Physics, National Taiwan University, Taipei, Taiwan

^{††}Department of Radiology, School of Medicine, College of Medicine, Taipei, Taiwan

^{‡‡}Department of Medical Imaging, Taipei Medical University Hospital, Taipei, Taiwan

^{xxi}Department of Radiology, National Defense Medical Center, Taipei, Taiwan

Abstract—This paper introduces CompressedMediQ, a novel hybrid quantum-classical machine learning pipeline specifically developed to address the computational challenges associated with high-dimensional multi-class neuroimaging data analysis. Standard neuroimaging datasets, such as large-scale MRI data from the Alzheimer’s Disease Neuroimaging Initiative and Neuroimaging in Frontotemporal Dementia, present significant hurdles due to their vast size and complexity. CompressedMediQ integrates classical HPC nodes for advanced MRI pre-processing and Convolutional Neural Network (CNN)-PCA-based feature extraction and reduction, addressing the limited-qubit availability for quantum data encoding in the NISQ era. This is followed by Quantum Support Vector Machine (QSVM) classification. By utilizing quantum kernel methods, the pipeline optimizes feature mapping and classification, enhancing data separability and outperforming traditional neuroimaging analysis techniques. Experimental results highlight the pipeline’s superior accuracy in dementia staging, validating the practical use of quantum machine learning in clinical diagnostics. Despite the limitations of NISQ devices, this proof-of-concept demonstrates the transformative potential of quantum-enhanced learning, paving the way for scalable and precise diagnostic tools in healthcare and signal processing.

Index Terms—Quantum Machine Learning, Quantum Neural Networks, MRI, Neuroimaging Data

I. INTRODUCTION

Magnetic Resonance Imaging (MRI) is a cornerstone of neuroscience, providing critical insights into brain structure and function across a variety of neurological conditions, including Alzheimer’s disease (AD) [1], [2]. Traditional approaches to MRI analysis have predominantly relied on hand-crafted feature extraction methods combined with classifiers such as Support Vector Machines (SVM) [3]. Although these

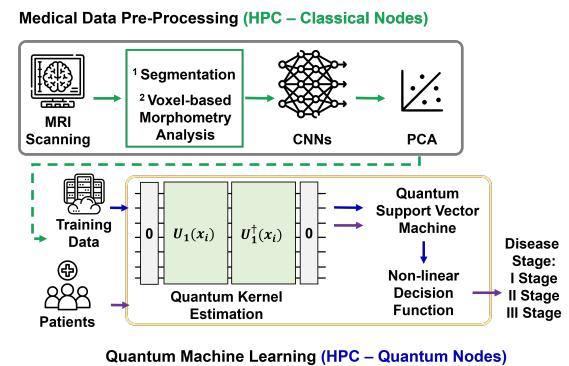


Fig. 1. Overview of the CompressedMediQ Pipeline for High-Dimensional Neuroimaging Data Analysis. The pipeline integrates classical and quantum computing nodes to process MRI data for disease classification. The workflow begins with data pre-processing on HPC classical nodes, including segmentation and voxel-based morphometry analysis, followed by feature extraction using CNNs and dimensionality reduction via PCA. The extracted features are then input into the quantum machine learning stage, where quantum kernel estimation and a QSVM are employed to perform multi-class classification of disease stages.

methods offer some success, they often struggle with the high-dimensional and limited sample size nature of MRI data, leading to challenges such as overfitting, prolonged training times, and diminishing accuracy as data dimensionality increases [4]–[6]. Recently, deep learning models, particularly Convolutional Neural Networks (CNNs), have demonstrated improved performance by automatically extracting high-level features from complex neuroimaging data [7]. However, these models are not without limitations, often suffering from excessive computational demands, reduced scalability, and sensitivity to noise, making them suboptimal for large-scale, multi-class

*Corresponding Author: kuan-cheng.chen17@imperial.ac.uk

classification problems typical in MRI applications [6], [8].

This paper introduces CompressedMediQ, a versatile hybrid quantum machine learning (QML) pipeline specifically designed to address the computational challenges associated with high-dimensional MRI data analysis. Standard MRI datasets, often featuring spatial resolutions of about 1 mm and typical dimensions of 256 x 256 x 150 voxels, result in approximately 10 million individual voxels per scan. These large data sizes, ranging from several hundred megabytes to several gigabytes per scan session, pose significant challenges for traditional machine learning approaches in terms of computational efficiency and accuracy [9]. While quantum machine learning has shown superior performance in various applications [10]–[14], the limited qubit availability for data encoding in the NISQ (Noisy Intermediate-Scale Quantum) era remains a major constraint [15], [16]. CompressedMediQ addresses these issues by integrating quantum computing with classical ML techniques, enhancing processing speed and accuracy [17], [18]. The pipeline leverages quantum properties such as superposition and entanglement to improve the efficiency of handling massive datasets. As depicted in Fig. 1, CompressedMediQ consists of two primary stages: (1) MRI pre-processing using classical high-performance computing (HPC) nodes, and (2) a hybrid quantum-classical workflow that incorporates Quantum Support Vector Machines (QSVM) with quantum feature mapping to encode classical data into quantum states. This approach enables the exploration of high-dimensional feature spaces with enhanced separability between classes, effectively overcoming the limitations of traditional methods [19], [20].

In this paper, we detail the architecture of the CompressedMediQ pipeline, outline the specific quantum algorithms used, and present comparative results highlighting its advantages. Our findings underscore the potential of hybrid QML frameworks in overcoming the scalability and accuracy limitations of classical methods, offering a new paradigm for large-scale neuroimaging data analysis. By demonstrating a proof-of-concept application in dementia classification—a particularly challenging problem in neuroimaging research [21]—this work validates the practical benefits of integrating quantum-enhanced learning techniques. CompressedMediQ not only advances the field of MRI-based AD classification but also sets the foundation for future research in applying quantum-enhanced algorithms to complex medical data challenges, paving the way for scalable and accurate diagnostic solutions [8].

II. METHODOLOGY

A. MRI Data Collection

The data used in this study were sourced from two primary databases: the Alzheimer’s Disease Neuroimaging Initiative (ADNI) [22] and Neuroimaging in Frontotemporal Dementia (NIFD) [23]. Subjects were categorized into six distinct groups based on their Clinical Dementia Rating Sum of Boxes (CDR-SB) scores [24], which served as the classification criteria. These groups included: normal controls (CDR-SB score of

0); individuals with questionable cognitive impairment (CDR-SB scores ranging from 0.5 to 2.5); subjects with very mild dementia (CDR-SB scores between 3 and 4); those with mild dementia (CDR-SB scores from 4.5 to 9); individuals with moderate dementia (CDR-SB scores between 9.5 and 15.5); and subjects with severe dementia (CDR-SB scores ranging from 16 to 18). This stratification provided a structured framework for analyzing various stages of cognitive decline, supporting the study’s objective of exploring neuroimaging data across the full spectrum of dementia severity. For this research, 1,000 data points were used for training, while 200 were reserved for testing.

B. MRI Data Pre-Processing

Prior to inputting the data into our specialized auto-segmentation neural network models, the MRI scans underwent a rigorous preprocessing pipeline designed to optimize data quality and reliability. This preprocessing approach was essential for standardizing the structural MRI data, correcting potential artifacts, and accurately isolating relevant brain tissues for further analysis. The following subsections provide a detailed account of each step in this comprehensive preprocessing protocol.

1) Brain tissue segmentation: The 3D T1-weighted structural MRI scans were acquired from all participants and underwent a comprehensive preprocessing pipeline using the unified segmentation algorithm available in the SPM12 software package [25]. SPM12 is a MATLAB-based toolset designed for advanced statistical analysis of neuroimaging data. This preprocessing phase is crucial for standardizing the MRI scans, ensuring that subsequent analyses are both accurate and reliable.

The segmentation process began with bias field correction, which compensates for intensity inhomogeneities caused by magnetic field distortions during MRI acquisition. Following this, the algorithm performed tissue classification to segment the brain into distinct tissue classes: gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). Additionally,

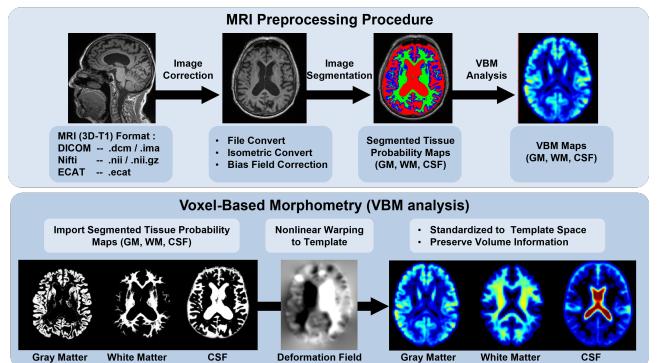


Fig. 2. Overview of the MRI Preprocessing Procedure and VBM Analysis, illustrating the steps from image correction and segmentation to VBM map generation, and nonlinear warping of segmented tissue probability maps to standardized template space for preserving volumetric information of GM, WM, and CSF.

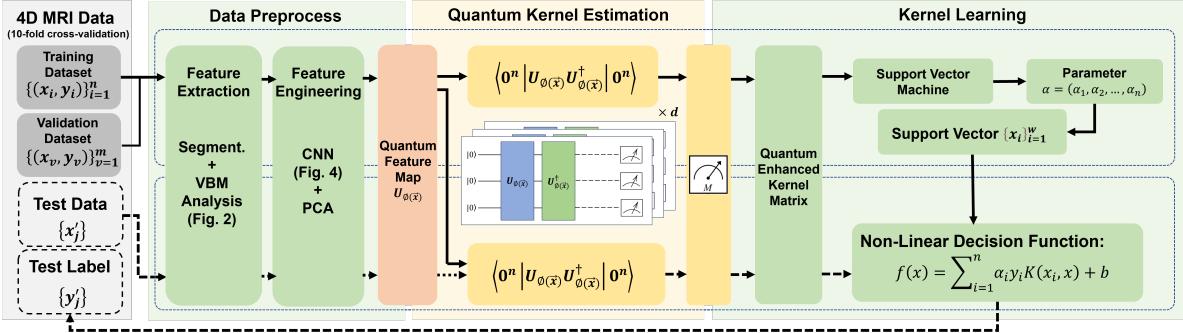


Fig. 3. Overview of the CompressedMediQ pipeline illustrating the hybrid quantum-classical workflow, including data preprocessing with CNN and PCA, quantum kernel estimation, and kernel learning using QSVM for enhanced multi-class classification of high-dimensional MRI data.

probability maps were generated for non-brain tissues, such as the skull and scalp. The segmentation utilized medium bias regularization and incorporated a reference brain probability map provided by SPM12, which enhances the precision and consistency of the tissue classification. This rigorous approach ensures high-quality data that is essential for subsequent statistical analyses.

2) *Voxel-based morphometry (VBM) analysis:* Following the initial segmentation, the MRI data underwent further refinement using the Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra (DARTEL)-based spatial normalization technique [26], implemented via SPM12 software. This advanced normalization method is particularly advantageous for studies involving multi-subject comparisons, as it aligns the structural brain images into a common anatomical framework. The precise alignment facilitated by DARTEL helps in reducing inter-subject variability, ensuring that each image conforms closely to a standard template, thereby enhancing the robustness of subsequent analyses.

During the normalization process, certain brain regions may expand or contract to match the template's anatomical structure. To retain the original volumetric characteristics of these regions, a modulation step is employed, which scales each voxel in the segmented images by the Jacobian determinant derived from the spatial transformation. This modulation step preserves the integrity of tissue volume information (e.g., GM, WM, CSF) by ensuring that the overall signal remains consistent before and after spatial normalization. This meticulous preprocessing approach is designed to minimize variability from individual anatomical differences, thus enabling more precise and reliable statistical analyses of the structural MRI data.

C. Hybrid Classic-Quantum Network Architecture

The CNN-PCA-QSVM architecture proposed in this study combines classical CNNs with QSVMs to enhance the classification of neuroimaging data, as illustrated in Figure 1. The CNN component of the architecture accepts modulated GM and CSF maps as inputs, derived from MRI data to extract essential features relevant to dementia classification. The CNN comprises two primary sections: a down-convolutional section

for reducing spatial dimensions and extracting key features, and a fully connected section designed to model the non-linear relationships among these features. Each basic convolutional module within the down-convolutional section starts with a $3 \times 3 \times 3$ convolutional layer followed by batch normalization, with shortcut connections merging the input layers and outputs of the stacked convolutional layers to facilitate identity mapping and mitigate the vanishing gradient problem. The Scaled Exponential Linear Unit (SELU) is employed as the activation function, followed by a $2 \times 2 \times 2$ max pooling operation with a stride of 2 for down-sampling. These convolutional modules are replicated four times throughout the architecture, with the 1st, 3rd, and 5th modules augmented by a Convolutional Block Attention Module (CBAM) to refine features adaptively before max pooling.

Following the convolutional layers, a dedicated convolution module consolidates the feature maps, which are then flattened into a one-dimensional array and connected to four fully connected layers, each containing 32 nodes. Demographic data, including subjects' age and sex, are integrated with the feature nodes after the first fully connected layer to enrich the feature space. To reduce the high-dimensional output from the CNN, Principal Component Analysis (PCA) is utilized to condense the data into a reduced feature space with a dimension of 8, crucial for optimizing the subsequent quantum encoding phase. This PCA step not only compresses the data into a manageable

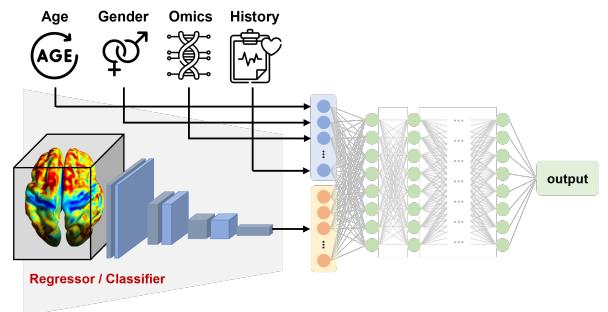


Fig. 4. The hybrid framework integrates MRI-derived features with demographic, genomic, and clinical history inputs into a combined regressor/classifier model for enhanced neuroimaging analysis and prediction.

size but also enhances computational efficiency, particularly in the NISQ era.

After dimensionality reduction, the data are encoded into both classical and quantum feature spaces. Quantum kernel methods complement the CNN-based feature extraction by leveraging advanced quantum computing techniques for classification. SVMs are particularly effective in high-dimensional, non-linear classification tasks, extending their capabilities through kernel functions that project data into higher-dimensional feature spaces. QSVMs extend this concept further by mapping classical data into quantum states within the Hilbert space using quantum feature maps [27]. This process involves rotational entanglement operations:

$$U_\Phi(\mathbf{x}) = \exp \left(i \sum_{S \subseteq [n]} \phi_S(\mathbf{x}) \prod_{k \in S} Z_k \right), \quad (1)$$

which encode the classical features into quantum states, enabling the formation of quantum kernels that capture complex, high-dimensional relationships with enhanced computational efficiency. The quantum kernel between two data points is calculated as: $K(\mathbf{x}_i, \mathbf{x}_j) = |\langle \psi(\mathbf{x}_i) | \psi(\mathbf{x}_j) \rangle|^2 = |\langle 0^{\otimes N} | U^\dagger(\mathbf{x}_i) U(\mathbf{x}_j) | 0^{\otimes N} \rangle|^2$, utilizing quantum parallelism to achieve significant computational speedups, often scaling as $O(\log(n))$ compared to classical methods for kernel calculations involving n^2 data point pairs [28]. This hybrid CNN-QSVM approach effectively combines deep learning's feature extraction capabilities with quantum-enhanced kernel methods, pushing the boundaries of data dimensionality handling and computational efficiency in neuroimaging classification tasks.

III. RESULT

The evaluation of the hybrid CNN-QSVM pipeline, as illustrated in Fig. 5, highlights the superior performance of the quantum model compared to the classical approach in dementia classification. The quantum model achieved high classification accuracy, particularly in distinguishing early stages of cognitive decline, such as normal controls, questionable cognitive impairment, and very mild dementia, with precision rates of 98%, 99%, and 91%, respectively. It maintained strong performance across all categories, with minimal misclassification between mild, moderate, and severe dementia stages. These results demonstrate the model's ability to effectively manage complex neuroimaging data, providing precise classification across a broad spectrum of dementia severities.

In contrast, the classical SVM model, also depicted in Fig. 5, struggled significantly, particularly in distinguishing mild and moderate dementia, as evident from the confusion matrix. The model frequently misclassified cases within these stages, with up to 36% of instances confused between mild and moderate dementia, highlighted within the marked region of the matrix. This pattern reflects the limitations of the classical approach in capturing subtle differences in intermediate stages of cognitive impairment. The comparison underscores the enhanced accuracy and reliability of the quantum model, highlighting its

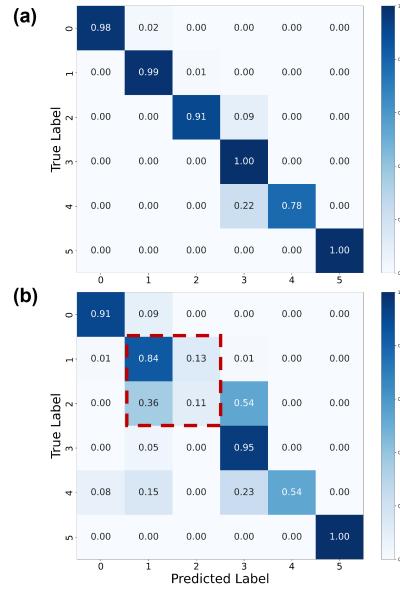


Fig. 5. Confusion matrices of the testing dataset compare the Quantum Support Vector Machine (SVM) model (a) and the Classical SVM model (b), highlighting the superior accuracy and reduced misclassification achieved by the Quantum model (96.1%) compared to the Classical SVM model (78.8%). Notably, the Quantum model demonstrates improved performance in the marked region with fewer misclassifications.

potential as a valuable tool for neuroimaging-based diagnostics in dementia. These findings validate the integration of QML within clinical pipelines, offering a significant advancement in diagnostic precision for neurodegenerative diseases.

IV. CONCLUSION

This study demonstrates the feasibility of combining classical and quantum computing to handle the computational complexity of high-dimensional neuroimaging data, particularly in dementia staging. By effectively reducing data dimensionality through CNN-based feature compression and leveraging a QSVM for quantum-enhanced classification, our hybrid framework delivers both improved diagnostic accuracy and reduced computational overhead. These findings underscore the potential of quantum machine learning in healthcare and signal processing, even within the constraints of NISQ-era hardware [29], [30]. Future research will focus on extending this quantum-classical pipeline to larger and more heterogeneous datasets, employing techniques such as quantum error mitigation [31], [32], distributed quantum computing [33], and refined resource management [34] to further enhance reliability and scalability. In doing so, we aim to establish a robust methodological foundation for quantum HPC systems [35] that can address a broader range of neuroimaging tasks and clinical challenges.

ACKNOWLEDGMENT

This work was supported by a grant from Imperial Seed grant (CIEW F14020) and the Engineering and Physical Sciences Research Council (EPSRC) grant (EP/W032643/1).

REFERENCES

- [1] N. K. LOGOTHETIS, *What we can do and what we cannot do with fmri*, Nature **453** (2008), pp. 869–878.
- [2] E. L. DENNIS AND P. M. THOMPSON, *Functional brain connectivity using fmri in aging and alzheimer's disease*, Neuropsychology review **24** (2014), pp. 49–62.
- [3] Y.-D. ZHANG, S. WANG AND Z. DONG, *Classification of alzheimer disease based on structural magnetic resonance imaging by kernel support vector machine decision tree*, Progress In Electromagnetics Research **144** (2014), pp. 171–184.
- [4] K. SPECHT, *Current challenges in translational and clinical fmri and future directions*, Frontiers in psychiatry **10** (2020), p. 924.
- [5] P. A. BANDETTINI, L. HUBER AND E. S. FINN, *Challenges and opportunities of mesoscopic brain mapping with fmri*, Current Opinion in Behavioral Sciences **40** (2021), pp. 189–200.
- [6] N. BEHESHTI AND L. JOHNSON, *Classification of 4d fmri images using ml, focusing on computational and memory utilization efficiency*, in MICCAI Workshop on Resource-Efficient Medical Image Analysis, Springer, 2022, pp. 55–64.
- [7] J. WEN ET AL., *Convolutional neural networks for classification of alzheimer's disease: Overview and reproducible evaluation*, Medical image analysis **63** (2020), p. 101694.
- [8] M. SUDHARAN AND G. THAILAMBAL, *Alzheimer's disease prediction using machine learning techniques and principal component analysis (pca)*, Materials Today: Proceedings **81** (2023), pp. 182–190.
- [9] M. MAKKIE ET AL., *A distributed computing platform for fmri big data analytics*, IEEE transactions on big data **5** (2018), pp. 109–119.
- [10] K.-C. CHEN, X. LI, X. XU, Y.-Y. WANG AND C.-Y. LIU, *Quantum-classical-quantum workflow in quantum-hpc middleware with gpu acceleration*, in 2024 International Conference on Quantum Communications, Networking, and Computing (QCNC), IEEE, 2024, pp. 304–311.
- [11] C.-H. A. LIN, C.-Y. LIU AND K.-C. CHEN, *Quantum-train long short-term memory: Application on flood prediction problem*, arXiv preprint arXiv:2407.08617 (2024).
- [12] K. T. M. HO, K.-C. CHEN, L. LEE, F. BURT, S. YU ET AL., *Quantum computing for climate resilience and sustainability challenges*, arXiv preprint arXiv:2407.16296 (2024).
- [13] C.-Y. LIU, C.-H. A. LIN, C.-H. H. YANG, K.-C. CHEN AND M.-H. HSIEH, *Qtrl: Toward practical quantum reinforcement learning via quantum-train*, arXiv preprint arXiv:2407.06103 (2024).
- [14] T. LI, V. R. MEKALA, K. NG AND C. SU, *Classification of tumor metastasis data by using quantum kernel-based algorithms*, in 2022 IEEE 22nd International Conference on Bioinformatics and Bioengineering (BIBE), IEEE, 2022, pp. 351–354.
- [15] G. LI, R. YE, X. ZHAO AND X. WANG, *Concentration of data encoding in parameterized quantum circuits*, Advances in Neural Information Processing Systems **35** (2022), pp. 19456–19469.
- [16] K. BHARTI ET AL., *Noisy intermediate-scale quantum algorithms*, Reviews of Modern Physics **94** (2022), p. 015004.
- [17] H.-Y. HUANG ET AL., *Power of data in quantum machine learning*, Nature communications **12** (2021), p. 2631.
- [18] M. C. CARO ET AL., *Generalization in quantum machine learning from few training data*, Nature communications **13** (2022), p. 4919.
- [19] K.-C. CHEN, X. XU, H. MAKHANOV, H.-H. CHUNG AND C.-Y. LIU, *Quantum-enhanced support vector machine for large-scale multi-class stellar classification*, in International Conference on Intelligent Computing, Springer, 2024, pp. 155–168.
- [20] K.-C. CHEN ET AL., *cutn-qsvm: cutensornet-accelerated quantum support vector machine with cuquantum sdk*, 2024. URL: <https://arxiv.org/abs/2405.02630>, arXiv:2405.02630.
- [21] M. C. TARTAGLIA, H. J. ROSEN AND B. L. MILLER, *Neuroimaging in dementia*, Neurotherapeutics **8** (2011), pp. 82–92.
- [22] R. C. PETERSEN ET AL., *Alzheimer's disease neuroimaging initiative (adni) clinical characterization*, Neurology **74** (2010), pp. 201–209.
- [23] J. D. ROHRER AND H. J. ROSEN, *Neuroimaging in frontotemporal dementia*, International Review of Psychiatry **25** (2013), pp. 221–229.
- [24] S. E. O'BRYANT ET AL., *Staging dementia using clinical dementia rating scale sum of boxes scores: a texas alzheimer's research consortium study*, Archives of neurology **65** (2008), pp. 1091–1095.
- [25] J. ASHBURNER ET AL., *Spm12 manual*, Wellcome Trust Centre for Neuroimaging, London, UK **2464** (2014).
- [26] Z. CHEN, L. LI, J. SUN AND L. MA, *Mapping the brain in type ii diabetes: voxel-based morphometry using dartel*, European journal of radiology **81** (2012), pp. 1870–1876.
- [27] M. SCHULD AND N. KILLORAN, *Quantum machine learning in feature hilbert spaces*, Physical review letters **122** (2019), p. 040504.
- [28] G. GENTINETTA, A. THOMSEN, D. SUTTER AND S. WOERNER, *The complexity of quantum support vector machines*, Quantum **8** (2024), p. 1225.
- [29] K. KOUROU, T. P. EXARCHOS, K. P. EXARCHOS, M. V. KARAMOUZIS AND D. I. FOTIADIS, *Machine learning applications in cancer prognosis and prediction*, Computational and structural biotechnology journal **13** (2015), pp. 8–17.
- [30] S. UDDIN, A. KHAN, M. E. HOSSAIN AND M. A. MONI, *Comparing different supervised machine learning algorithms for disease prediction*, BMC medical informatics and decision making **19** (2019), pp. 1–16.
- [31] A. STRIKIS, D. QIN, Y. CHEN, S. C. BENJAMIN AND Y. LI, *Learning-based quantum error mitigation*, PRX Quantum **2** (2021), p. 040330.
- [32] K.-C. CHEN, *Short-depth circuits and error mitigation for large-scale ghz-state preparation, and benchmarking on ibm's 127-qubit system*, in 2023 IEEE International Conference on Quantum Computing and Engineering (QCE), vol. 2, IEEE, 2023, pp. 207–210.
- [33] F. BURT, K.-C. CHEN AND K. LEUNG, *Generalised circuit partitioning for distributed quantum computing*, arXiv preprint arXiv:2408.01424 (2024).
- [34] K.-C. CHEN, X. XU, F. BURT AND C.-Y. LIU, *Noise-aware distributed quantum approximate optimization algorithm on near-term quantum hardware*, arXiv preprint arXiv:2407.17325 (2024).
- [35] T. S. HUMBLE, A. MCCASKEY, D. I. LYAKH, M. GOWRISHANKAR, A. FRISCH AND T. MONZ, *Quantum computers for high-performance computing*, IEEE Micro **41** (2021), pp. 15–23.