Quantum Information Processing (2026) 25:138
https://doi.org/10.1007/s11128-026-05144-x

Squid–transmon quantum hardware simulation with deep learning for pancreatic radiotherapy image classification

Javier Villalba-Díez[1,2] · Ana González-Marcos[2] · Juan Carlos Losada-González[3] · Joaquín Ordieres-Meré[4]

Received: 23 April 2025 / Accepted: 13 March 2026

Abstract

We present a hybrid quantum–classical framework that couples a parameterized, hardware-imperfection-aware SQUID–Transmon surrogate simulation with deep learning to classify pancreatic radiotherapy CT/CBCT images from the `Pancreatic-CT-CBCT-SEG` collection. The classical backbone uses a pretrained `EfficientNetB0` with bidirectional Long Short-Term Memory layers; the quantum component is a 4-qubit, shallow-depth circuit whose input angles are modulated by a dynamically computed error-mitigation factor derived from the gradient of an effective potential. We perform a grid search over three hardware-relevant parameters: a rescaling factor R (an amplitude/drive scaler in our surrogate) that modulates the effective potential and controls the input rotation amplitude, a secondary coupling strength λ, and the bare Josephson energy E_{J0}. In a 10-class subset experiment (top-dose patients), the hybrid model exhibits a performance maximum when $R = 0.95$, reaching a test accuracy of up to 0.95 across several (λ, E_{J0}) settings. In this work, deviations of R, λ, and E_{J0} are treated as quasi-static hardware imperfections (systematic calibration offsets and slow drift) rather than as stochastic noise channels. We interpret this optimum as a balance between expressivity and over-rotation/leakage in our simulator when the effective drive is slightly reduced. To benchmark consistency on the full dataset, we additionally report a window-level evaluation over all 40 patients with the classical backbone alone: a single 80/10/10 split yields train $= 0.959$, val $= 0.960$, test $= 0.909$ accuracy (windows); stratified threefold cross-validation

✉ Javier Villalba-Díez
javier.villalba-diez@hs-heilbronn.de

1 Fakultät Wirtschaft, Hochschule Heilbronn, Max-Planck-Str.39, 74081 Heilbronn, Baden-Württemberg, Germany

2 Department of Mechanical Engineering, Universidad de La Rioja, Edificio Departamental, c/San José de Calasanz, 31, 26004 Logroño, La Rioja, Spain

3 Escuela Técnica Superior de Ingeniería Agronómica, Alimentaria y de Biosistemas, Universidad Politécnica de Madrid, Av. Puerta de Hierro, 2, Moncloa - Aravaca, 28040 Madrid, Madrid, Spain

4 Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, C/ José Gutierrez Abascal, 2, 28006 Madrid, Madrid, Spain

Published online: 08 April 2026

achieves 0.876 ± 0.005 mean validation accuracy. Overall, our results indicate that such hardware-imperfection-aware hybrid models can be competitive with strong classical baselines while offering a physics-grounded knob for hardware-aware calibration; we discuss modeling assumptions and limitations (e.g., simplified hardware-imperfection model, connectivity, and task definition) to avoid overclaiming clinical readiness.

Keywords Radiotherapy Image Classification · Hybrid Quantum–Classical Model · SQUID–Transmon Surrogate Simulation · Quantum Error Mitigation · Pancreatic Cancer Computed Tomography

1 Introduction

Quantum computing has emerged as one of the most promising computational paradigms in recent years, offering a novel approach to solving problems that are beyond the reach of classical algorithms [1]. In particular, the advent of noisy intermediate-scale quantum (NISQ) devices [2] has spurred a great deal of interest in the development of hybrid quantum–classical algorithms [3] capable of leveraging the complementary strengths of both quantum and classical systems [4]. This work is situated at the intersection of quantum information science and medical imaging [5, 6], where the integration of advanced quantum hardware simulation with deep learning techniques [7, 8] has the potential to address critical challenges in the diagnosis and treatment of pancreatic cancer [9].

The ever-increasing volume and complexity of medical imaging data have driven the development of sophisticated machine learning models designed for diagnostic and prognostic purposes [10, 11]. In the field of radiotherapy, accurate segmentation and classification of imaging data are essential for effective treatment planning and delivery [12, 13]. Pancreatic cancer, known for its aggressive nature and high mortality rate, presents unique challenges due to the complex anatomical structures and subtle variations in tissue contrast. Conventional diagnostic tools often struggle with these nuances, prompting the exploration of more powerful computational techniques [14, 15].

Deep learning, particularly convolutional neural networks and recurrent neural networks, has dramatically improved the performance of medical image analysis [16]. Models such as `EfficientNetB0` have demonstrated remarkable effectiveness in feature extraction from high-resolution images [17–20]. However, classical neural networks are inherently limited by the data and the computational resources required for training on large, high-dimensional datasets. As a result, there is an ongoing search for alternative approaches that can both reduce the computational overhead and enhance the predictive performance of diagnostic models [21, 22].

In parallel with these advances in deep learning, quantum computing has begun to offer new avenues for tackling complex optimization and pattern recognition tasks [23, 24]. Early quantum algorithms have shown that quantum systems can perform certain types of computations exponentially faster than their classical counterparts [24]. Nevertheless, the practical application of quantum computing is still in its infancy, largely due to the noise and decoherence inherent to contemporary quantum hardware

[25]. This challenge has given rise to the concept of hybrid quantum–classical models [26], in which quantum circuits are integrated into classical frameworks to enhance performance on specific tasks while mitigating the limitations of current quantum hardware [27].

Pancreatic cancer represents one of the most formidable challenges in oncology, with a very low survival rate and complex treatment protocols. The diagnosis and management of this disease rely heavily on high-quality imaging modalities, such as Computed Tomography (CT) and cone-beam CT (CBCT). In this work, we use the `Pancreatic-CT-CBCT-SEG` dataset [28] only as a source of image volumes (one treatment-planning CT and two on-treatment CBCT scans per patient, 40 patients total). The supervised labels are patient identity (one class per patient); available anatomy contours and radiotherapy dose files provided by the collection are not used as model inputs or targets for the classifier.

Conventional imaging analysis techniques often rely on handcrafted features or standard deep learning models that, while effective in many contexts, may lack the sensitivity required to capture the intricate details present in pancreatic images [29]. In clinical practice, even marginal improvements in diagnostic accuracy can translate into significant enhancements in patient outcomes [30]. Thus, developing a robust, high-precision classification model that can discern the fine-grained features relevant to pancreatic cancer treatment is of paramount importance. This need has fueled research into integrating quantum computing paradigms, which promise to unlock new levels of computational efficiency and accuracy.

Hybrid quantum–classical models represent a cutting-edge research direction that seeks to harness the strengths of both computational worlds. The classical component typically performs image preprocessing and feature extraction, whereas the quantum component is integrated as a compact, trainable module that operates on a low-dimensional feature representation rather than on raw images.

In our pipeline, the `EfficientNetB0`+BiLSTM backbone produces a pooled embedding $\mathbf{h}_{\text{classical}}$. A learned encoder (linear layer followed by tanh) compresses $\mathbf{h}_{\text{classical}}$ into four bounded scalars that parameterize the input rotations of a 4-qubit parameterized quantum circuit. The circuit returns the Pauli-Z expectation values $\mathbf{z} = [\langle Z\rangle_1, \ldots, \langle Z\rangle_4]$, and we concatenate $\mathbf{z}$ with $\mathbf{h}_{\text{classical}}$ before the final dense classifier. In this way, the quantum component has a clear algorithmic role: it acts as an entanglement-enabled nonlinear re-embedding / feature-augmentation layer whose outputs are optimized jointly with the trainable parts of the classical network.

The SQUID–transmon surrogate is used to provide a lightweight, physics-guided control signal rather than a full open-system hardware simulation: an error-mitigation factor (EMF) computed from an effective potential $U_{\text{eff}}(R, \lambda, E_{J0})$ rescales the quantum input angles $\theta_i \mapsto \theta_i \cdot \text{EMF}$. This angle-rescaling mechanism serves as a hardware-aware calibration knob (capturing amplitude/drive scaling and coupling-induced curvature modulation) that we systematically study via a grid search over (R, λ, E_{J0}). For compactness, the detailed equations and the physical interpretation of R, λ, and E_{J0} are provided in Sect. 3 (SQUID–Transmon Simulation Framework).

Despite significant progress in both quantum computing and deep learning, the application of hybrid quantum–classical models to real-world medical imaging tasks remains largely unexplored. Existing studies have primarily focused on applications in

low-dimensional datasets [31, 32]. In the context of pancreatic cancer imaging, there is a pressing need for computational methods that can handle the high complexity of radiotherapy data while also being robust to the inherent hardware imperfections and drift of current quantum hardware. This research seeks to bridge that gap by addressing the following key objectives:

- Integration of quantum hardware simulation with deep learning: Develop a robust hybrid framework that marries state-of-the-art deep learning models with a parameterized, hardware-imperfection-aware SQUID–Transmon surrogate hardware, incorporating dynamic error mitigation strategies.
- Optimization of hardware-relevant parameters: Execute a comprehensive grid search over critical quantum hardware parameters, specifically R, λ, and E_{J0}, to identify optimal conditions that maximize classification accuracy.
- Validation in a clinically relevant setting: Apply the proposed framework to the `Pancreatic-CT-CBCT-SEG` dataset, thereby demonstrating its efficacy in a scenario with high clinical stakes and complex imaging data.
- Insight into quantum error dynamics: Provide a detailed analysis of how slight deviations from ideal hardware settings (e.g., $R = 1.00$ vs. $R = 0.95$) affect the performance of the hybrid model, yielding insights that are valuable both from a quantum hardware calibration perspective and for the deployment of quantum-enhanced diagnostic tools.

Our working hypothesis is that a carefully calibrated, hardware-imperfection-aware hybrid layer can *match* a strong classical baseline while offering a physics-grounded control (via R, λ, E_{J0}) to reduce over-rotation/leakage and improve stability under small-data conditions—an effect akin to task-specific regularization highlighted in prior hybrid studies [33, 34]. Specifically, we posit that the integration of realistic SQUID–Transmon simulation and advanced error mitigation strategies will not only enhance the robustness of the quantum circuit but also lead to tangible improvements in diagnostic accuracy.

In the realm of medical imaging, our hybrid framework addresses the need for more precise diagnostic tools in the management of pancreatic cancer–a disease notorious for its poor prognosis and diagnostic complexity. By achieving classification accuracies as high as 95%, the proposed model demonstrates the potential to improve treatment planning and patient outcomes in a clinical setting. Enhanced image classification can lead to more accurate delineation of organs-at-risk, enabling clinicians to tailor radiotherapy protocols with greater precision. This, in turn, could result in reduced collateral damage to healthy tissues and improved overall survival rates.

Moreover, the successful application of a quantum-enhanced model to a challenging clinical dataset such as `Pancreatic-CT-CBCT-SEG` serves as a proof-of-concept for the broader integration of quantum computing in precision medicine. As quantum hardware continues to mature, the methodologies developed in this work may be adapted to a wide range of diagnostic and prognostic applications, spanning other forms of cancer and complex pathological conditions. By bridging the gap between theoretical quantum advantages and practical clinical applications, our research aims to catalyze the next generation of computational tools in medicine.

In summary, this paper presents a hybrid quantum–classical framework for pancreatic radiotherapy image classification in which a strong convolutional–recurrent backbone is complemented by a shallow 4-qubit quantum circuit that provides additional trainable features. The quantum circuit does not replace the classical feature extractor; instead, it transforms a low-dimensional projection of the classical representation into expectation-value features that are concatenated with the classical embedding before final classification. A parameterized, hardware-imperfection-aware SQUID–Transmon surrogate is used to compute a simple, dynamically updated angle-rescaling factor, enabling a physics-grounded study of how hardware-relevant parameters (R, λ, E_{J0}) affect the hybrid pipeline.

The remainder of this paper is organized as follows. In Sect. 2, we review the current state of research in hybrid quantum–classical models, with a focus on applications in medical imaging and radiotherapy. This literature review highlights the limitations of existing approaches and sets the stage for our contributions. Section 3 provides a comprehensive description of our methodology, including details on the deep learning architecture employed for feature extraction, the parameterized, hardware-imperfection-aware SQUID–Transmon surrogate framework, and the dynamic error mitigation strategies incorporated into the quantum circuit. In Sect. 4, we describe the experimental setup, the configuration of the `Pancreatic-CT-CBCT-SEG` dataset, and the parameter grid search strategy used to optimize hardware settings. Section 5 presents an in-depth analysis of the experimental results, highlighting the competitive performance of our hybrid model in comparison with strong classical baselines, as well as the sensitivity of the results to the selected hardware-relevant parameters. Special emphasis is placed on the observed optimum at $R = 0.95$ and its implications for quantum hardware calibration. Finally, in Sect. 6, we discuss the broader implications of our findings in the context of both quantum computing and clinical oncology, outlining the potential directions for future research and the challenges that remain.

2 Related work

Quantum computing has rapidly advanced in recent years, led primarily by superconducting qubit implementations such as the transmon. The transmon, a capacitively shunted Josephson junction operated at large E_J/E_C, suppresses charge dispersion while maintaining useful anharmonicity; its E_J/E_C is high (though not as extreme as in phase qubits) [35, 36]. Adding a SQUID loop enables flux tunability by modulating the effective Josephson energy and qubit frequency, at the cost of additional flux-noise pathways and decoherence trade-offs [35–37]. The careful simulation of SQUID–Transmon systems requires precise modeling of parameters such as the rescaling factor (R), the secondary coupling strength (λ), and the nominal Josephson energy (E_{J0}). These parameters, individually and in combination, determine the operational fidelity and resilience of quantum circuits against non-adiabatic transitions and leakage errors.

In parallel with hardware development, error mitigation techniques have garnered significant attention as a means to suppress the effective error rates in NISQ devices. Foundational error mitigation strategies include zero-noise extrapolation [38], proba-

bilistic error cancellation [39], and readout error correction, among others [40]. These methods have been experimentally validated on superconducting qubit platforms, yielding improved estimation accuracy in variational quantum algorithms. Because device drift and context-dependent errors are common, *adaptive error mitigation* (i.e., mitigation parameters estimated from calibration data or runtime observables and updated as conditions change) is especially valuable; examples include zero-noise extrapolation, measurement-bias calibration, and correlated-readout models [38, 40, 41]. Recent studies have demonstrated that dynamic error mitigation, which computes an error mitigation factor in real time based on measured hardware characteristics, can help balance decoherence and maintain high-fidelity operations under realistic noisy conditions [41]. Such approaches are essential for the successful deployment of hybrid quantum–classical algorithms and form an integral part of the motivation behind our investigation of SQUID–Transmon parameters.

Hybrid quantum–classical algorithms, most notably Variational Quantum Algorithms (VQAs) such as the Variational Quantum Eigensolver (VQE) and the Quantum Approximate Optimization Algorithm (QAOA), have emerged as promising paradigms for leveraging NISQ devices [42, 43]. In these schemes, Parameterized Quantum Circuits (PQCs) are optimized using classical routines to solve problems that are otherwise computationally intractable. Early demonstrations of variational quantum classifiers have shown the potential for these algorithms to perform tasks such as pattern recognition and binary classification with promising results [44]. However, many early studies assumed idealized quantum hardware or simplified noise models, leaving open the question of how detailed hardware characteristics influence algorithmic performance on real-world tasks.

Medical imaging, particularly in the context of radiotherapy for pancreatic cancer, represents a domain where high precision is paramount. The application of deep learning to medical image classification and segmentation has yielded significant progress. Convolutional neural networks and recurrent architectures have been successfully applied to tasks such as CT and CBCT segmentation, facilitating improved treatment planning in radiotherapy [45]. Notably, in pancreatic cancer, where the disease presents a high degree of diagnostic complexity, even marginal improvements in image classification accuracy can have a significant clinical impact. However, classical deep learning models tend to demand large quantities of labeled data and extensive computational resources, which are often limiting in the context of rare or complex conditions. This limitation has motivated the exploration of hybrid quantum–classical models as an exploratory research direction, where a compact quantum circuit is inserted as a trainable module within a classical network, rather than as a demonstrated solution to data scarcity or clinical complexity. While such models may eventually offer computational efficiencies as quantum hardware matures, at present any gains in sample efficiency or scalability remain an open empirical question that depends on the task, model design, and hardware constraints; therefore, claims of practical advantage should be interpreted cautiously.

Recent efforts have begun integrating quantum circuits with classical deep learning architectures to form hybrid systems. These systems incorporate parameterized quantum circuits as layers within classical neural networks, aiming to exploit quantum phenomena such as superposition and entanglement to enhance feature extraction

and decision making [26, 44]. Some studies have reported that these hybrid models can offer comparable performance to entirely classical networks while requiring fewer parameters and potentially more efficient learning from limited data [33, 34]. Although most of these hybrid approaches have been tested on benchmark datasets and simplified tasks, the extension of such methods to complex, clinically relevant imaging datasets–such as those involving pancreatic radiotherapy images–remains limited. In particular, there is a scarcity of investigations that explicitly consider the hardware-specific parameters of SQUID–Transmon systems when integrating quantum circuits into medical diagnostic models.

A significant gap in the literature is the lack of comprehensive studies that integrate parameterized, hardware-imperfection-aware SQUID–Transmon surrogates into hybrid quantum–classical frameworks for medical imaging. Existing works typically model quantum noise in an abstract manner or use generic quantum devices without detailed consideration of key parameters like R, λ, and E_{J0}. Meanwhile, research in superconducting qubit fabrication has established that these parameters critically influence qubit performance and stability [36, 37]. There is a pressing need to bridge this gap by systematically investigating how variations in these specific parameters affect the performance of hybrid models in a clinical application. Our work addresses this gap by conducting a detailed grid search over the hardware parameters in a SQUID–Transmon simulation, integrated within a hybrid framework designed to classify pancreatic CT/CBCT images.

Several studies have made initial strides toward the integration of quantum and classical techniques in various domains. For example, research on quantum-enhanced image processing has demonstrated that incorporating a quantum feature map can improve the accuracy of image classification tasks on benchmark datasets like MNIST and CIFAR-10 [26]. These early works typically focused on synthetic datasets or relatively simple tasks and did not consider the impact of detailed hardware parameters. Similarly, quantum-inspired algorithms have been applied to classical image segmentation and thresholding problems, achieving competitive performance without full quantum hardware implementation [46]. Although these methods have opened the door to quantum approaches in imaging, they do not fully capture the complexities of executing hybrid quantum algorithms on realistic devices.

Recent efforts have aimed to fuse quantum circuit simulation with deep learning to tackle more demanding problems. Experimental demonstrations on superconducting platforms have shown that variational quantum circuits can be trained to perform classification tasks with performance on par with classical counterparts, even when operating under non-ideal hardware conditions [44, 47]. In these studies, the quantum circuits were generally treated as black boxes with abstract noise models, and the optimization of hardware-specific parameters was not explored in detail. The present work builds upon these findings by providing a systematic evaluation of how the parameters R, λ, and E_{J0} of a parameterized, noise-aware SQUID–Transmon surrogate affect the overall performance of a hybrid diagnostic model. By doing so, we offer insights into the co-design of quantum algorithms and hardware calibration strategies, ensuring that the benefits of quantum-enhanced processing are attainable even in the presence of realistic noise and imperfections.

In summary, while foundational research in superconducting qubits and variational quantum algorithms has established the theoretical basis for hybrid quantum–classical models, practical applications–especially in medical imaging–remain in early stages. The current literature reveals a notable absence of studies that integrate detailed SQUID–Transmon simulations, with a focus on optimizing key hardware parameters, into hybrid models for high-stakes tasks such as pancreatic radiotherapy image classification. Our work addresses this gap by conducting an extensive parameter study within a parameterized, noise-aware quantum simulation framework, thereby linking the physics of superconducting circuits with the algorithmic performance of hybrid models. This approach not only advances the understanding of NISQ devices in applied settings but also provides practical guidelines for deploying quantum-enhanced diagnostic tools in environments where both hardware imperfections and data challenges are prevalent.

By offering a rigorous analysis that spans foundational theory, experimental advancements in quantum hardware, and cutting-edge applications in medical imaging, this study contributes to a broader understanding of how quantum computing may be integrated into complex, real-world tasks. In doing so, it sets the stage for future research aiming to refine hybrid quantum–classical approaches and to extend their applicability to a range of domains where traditional methods face intrinsic limitations.

3 Methodology

In this section, we present a detailed description of our methodology, which integrates deep learning techniques for feature extraction with a parameterized, hardware-imperfection-aware SQUID–Transmon surrogate quantum hardware and dynamic error mitigation strategies. Our approach is designed to improve the classification of pancreatic radiotherapy images acquired from CT and CBCT modalities. We describe the mathematical underpinnings of our deep learning architecture, the simulation framework for the SQUID–Transmon system, and the precise methods used to perform dynamic error mitigation in our quantum circuit. Throughout, we emphasize mathematical rigor and the interplay between classical and quantum components of our hybrid model.

3.1 Deep learning architecture for feature extraction

The classical part of our framework is built on a deep learning architecture that is responsible for extracting high-level features from 224×224 RGB images. Each training sample is a fixed-length window of $T = 60$ consecutive axial slices; therefore, the full network input tensor has shape (None, 60, 224, 224, 3), where None denotes the batch dimension. We employ the `EfficientNetB0` model [17], pretrained on ImageNet, as the backbone for feature extraction. This model serves as a robust convolutional network for capturing spatial patterns. Let the input image be denoted by $I_t \in \mathbb{R}^{224\times224\times3}$. A full input window is the ordered sequence $\mathbf{X} = (I_1, \ldots, I_T) \in$

$\mathbb{R}^{T\times 224\times 224\times 3}$ with $T = 60$, and `EfficientNetB0` is applied slice-wise to each I_t via a TimeDistributed wrapper. The output from `EfficientNetB0`, after global average pooling, is a feature vector $\mathbf{f}_t \in \mathbb{R}^d$, where d is the number of features derived from the last convolutional block.

Mathematically, the transformation can be represented as:

$$\mathbf{f}_t = \text{GAP}(\text{EfficientNetB0}(I_t)), \quad t = 1, \ldots, T,$$

where $\text{GAP}(\cdot)$ denotes the global average pooling operation.

Following the feature extraction, a bidirectional LSTM network is employed to capture temporal and contextual correlations across a sequence of CT slices. Let $\{\mathbf{f}_t\}_{t=1}^{T}$ be the sequence of feature vectors extracted from T consecutive slices of a CT volume. In our experiments, we set $T = 60$; thus the "60" in Table 1 denotes the number of slices (time steps) aggregated per sample window by the bidirectional LSTM. The bidirectional LSTM processes the sequence in both forward and reverse directions, which we denote as $\overrightarrow{\mathbf{h}}_t$ and $\overleftarrow{\mathbf{h}}_t$, respectively. The combined hidden state at time t is given by: $\mathbf{h}_t = \overrightarrow{\mathbf{h}}_t \| \overleftarrow{\mathbf{h}}_t$, where $\|$ signifies vector concatenation.

The temporal features are subsequently passed through dense layers with nonlinear activation functions, specifically the rectified linear unit (ReLU), and regularization through dropout and L_2 norms to prevent overfitting. In mathematical terms, for a dense layer with weights $\mathbf{W}$ and bias $\mathbf{b}$, the transformation is: $\mathbf{z} = \text{ReLU}(\mathbf{W}\mathbf{h} + \mathbf{b})$, with additional dropout applied as: $\tilde{\mathbf{z}} = \text{Dropout}(\mathbf{z}, p)$, where p represents the dropout probability. Furthermore, we apply Fourier feature encoding on the dense representations to capture periodicities in the data. If θ represents the Fourier basis extracted from the feature representation, then the Fourier encoding can be formalized as: $\mathbf{F} = [\sin(\theta\mathbf{w}), \cos(\theta\mathbf{w})] \in \mathbb{R}^k$, for a frequency weight matrix $\mathbf{w} \in \mathbb{R}^{d\times k}$, where k is the number of desired component of the Fourier codification, and $\theta \in \mathbb{R}^d$. The final classical representation $\mathbf{h}_{\text{final}}$ is a concatenation of the Fourier encoding with a repeated dense representation, and is fed to the output classification layer with a softmax activation: $\hat{\mathbf{y}} = \text{softmax}(\mathbf{W}_{\text{out}}\mathbf{h}_{\text{final}} + \mathbf{b}_{\text{out}})$. This structure ensures that the rich spatial and temporal information in volumetric CT data is effectively captured and encoded in a manner amenable to classification.

Table 1 shows the overall deep learning architecture for feature extraction which has the following number of parameters:

- Total params: 5,596,717
- Trainable params: 1,547,146 (bidirectional LSTM and lightweight classifier layers)
- Non-trainable params: 4,049,571 (`EfficientNetB0` feature extractor, frozen)

In all experiments, the pretrained `EfficientNetB0` weights remain frozen (no fine-tuning) and are used purely for feature extraction; only the bidirectional LSTM block and the subsequent lightweight classifier layers are updated during training, which explains the trainable/non-trainable split reported above.

Table 1 Model architecture summary (classical backbone)

Layer (type)	Output shape	Param #	Connected to
Input (InputLayer)	(None, 60, 224, 224, 3)	0	–
Slices → EffNetB0 (TimeDistributed, frozen)	(None, 60, 1280)	4,049,571	Input
BiLSTM(128) (Bidirectional)	(None, 256)	1,442,816	Slices → EffNetB0
Dense(256) + ReLU	(None, 256)	65,792	BiLSTM(128)
Dropout(0.5)	(None, 256)	0	Dense(256) + ReLU
FourierFeat (Lambda)	(None, 300, 32)	0	Dropout(0.5)
Repeat(256) (Lambda)	(None, 300, 256)	0	Dropout(0.5)
Concat	(None, 300, 288)	0	FourierFeat, Repeat(256)
TD Dense(128)	(None, 300, 128)	36,992	Concat
TD ID (TimeDistributed)	(None, 300, 128)	256	TD Dense(128)
GlobalAvgPool1D	(None, 128)	0	TD ID
Classifier (Dense, C classes)	(None, C)	See note[†]	GlobalAvgPool1D

In the Output Shape column, the leading entry "None" denotes the batch dimension (left unspecified at model-definition time and set at runtime). In our experiments, the batch size was 1. The `EfficientNetB0` block is frozen and counted as non-trainable; training updates only the bidirectional LSTM and subsequent classifier layers

[†]For $C = 10$ (hybrid 10-class subset), the last layer has $128 \times 10 + 10 = 1290$ parameters; for $C = 40$ (full dataset CV), it has $128 \times 40 + 40 = 5160$ parameters

3.2 SQUID–transmon simulation framework

The quantum component of our methodology is based on a geometry-agnostic, parameterized SQUID–Transmon surrogate (not a full open-system simulation). A transmon qubit is a superconducting circuit that mitigates charge noise by operating in a regime where the ratio E_J/E_C is large [35, 36]. In our simulation, we model a SQUID loop integrated with a transmon qubit; this allows us to tune the effective Josephson energy E_J via an externally applied magnetic flux Φ. The effective Josephson energy is given by: $E_J(\Phi) = E_{J0}\,|\cos(\pi\Phi)|$, where E_{J0} is the nominal (or maximum) Josephson energy when $\Phi = 0$.

The potential energy landscape, a key determinant of qubit dynamics, is modeled as: $U(\Delta\phi) = -E_J(\Phi)\cos(\Delta\phi)$, where $\Delta\phi$ represents the phase difference across the junction. To incorporate device imperfections and the influence of coupling, we introduce a term $C = \lambda\cos(2\pi\Phi)\cos(2\Delta\phi)$, where λ is a coupling strength parameter. The effective potential U_{eff} is thus formulated as: $U_{\text{eff}} = R\,U(\Delta\phi) - C$, with R representing a rescaling factor. In a realistic setting, this effective potential is subject to quasi-static hardware imperfections (systematic calibration offsets, residual coupling, and slow parameter drift), which we capture through numerical gradient computations: $\nabla U_{\text{eff}} = \sqrt{\left(\frac{\partial U_{\text{eff}}}{\partial\Phi}\right)^2 + \left(\frac{\partial U_{\text{eff}}}{\partial\Delta\phi}\right)^2}$. The average gradient magnitude is used to compute a dynamic error mitigation factor (EMF) according to: $\text{EMF} = 1.0 + 0.001\,\langle\nabla U_{\text{eff}}\rangle$.

Terminology: "noise" versus systematic hardware imperfections In this work, deviations of R, λ, and E_{J0} from their ideal or nominal values are interpreted as systematic hardware imperfections (calibration biases and slow drift) that perturb U_{eff}. We therefore use the term "hardware-imperfection-aware" when describing the surrogate. Stochastic noise processes (e.g., relaxation and dephasing or gate-level stochastic error channels) are not explicitly simulated in the present surrogate and remain out of scope.

Physical interpretation of SQUID–transmon surrogate parameters

- *Scaling factor R* R rescales the primary potential to capture readout-/drive-related amplitude effects (e.g., effective angle rescaling) and measurement non-idealities [48]. Values near one indicate near-ideal amplitude; modest reductions can curb over-rotation and leakage in shallow circuits.
- *Secondary coupling* λ λ scales a second-harmonic term that effectively captures residual couplings (e.g., parasitic qubit–qubit or qubit–environment interactions): $C = \lambda\,\cos(2\pi\Phi)\cos(2\Delta\phi)$. Larger λ increases curvature modulation, which can improve expressivity but also raise leakage susceptibility.
- *Bare Josephson energy* E_{J0} $E_{J0} = \hbar I_c/(2e)$ sets the energy scale, impacting frequency and anharmonicity [49]. In SQUIDs with junction asymmetry (e.g., energies $E_{J1} \neq E_{J2}$), an asymmetry parameter $d = (E_{J1} - E_{J2})/(E_{J1} + E_{J2})$ trades tunability against flux-noise susceptibility and dephasing; we assume small asymmetry and absorb its effect into the effective $E_J(\Phi)$.

The 10^{-3} scale was chosen to keep $\theta_i' = \theta_i \cdot \text{EMF}$ within a stable range for shallow circuits while still responding to changes in $\langle\|\nabla U_{\text{eff}}\|\rangle$. Scans over $\alpha \in \{10^{-4}, 5\times 10^{-4}, 10^{-3}, 5\times 10^{-3}\}$ (replacing 10^{-3} with α) indicated minimal sensitivity for $\alpha \leq$

10^{-3} and degraded accuracy at 5×10^{-3} due to over-rotation. For a physical device, the same logic would be implemented as a calibration step: R can be inferred from drive-amplitude calibration (e.g., Rabi-based angle tuning), E_{J0} from fitting flux-dependent spectroscopy, and λ from residual-coupling characterization (e.g., avoided crossings or crosstalk/leakage diagnostics). The resulting angle-rescaling factor can then be computed offline and refreshed on the timescale of calibration cycles, rather than recomputed per gate.

In this expression, the averaging $\langle \nabla U_{\text{eff}} \rangle$ is performed over a discrete mesh of the parameter space defined by the magnetic flux Φ and the phase difference $\Delta\phi$. Specifically, the domain is set as $\Phi \in [\Phi_{\min}, \Phi_{\max}]$ and $\Delta\phi \in [\Delta\phi_{\min}, \Delta\phi_{\max}]$, and the arithmetic mean of the gradient magnitude is computed over all grid points within these ranges. This explicit definition of the averaging domain facilitates a clear interpretation of the EMF calculation and ensures that the process is reproducible by comprehensively sampling the physically relevant parameters.

We employ the PennyLane framework [50] to simulate our quantum subsystem, wherein the quantum circuit is defined over n qubits (in our experiments, $n = 4$) and has a circuit depth of $q = 2$ layers. The initialization of each qubit is in the computational basis state $|0\rangle$, and the overall initial state of the system is expressed as: $|\psi_0\rangle = |0\rangle^{\otimes n}$.

We derive the four-dimensional circuit input $\mathbf{x}$ from the classical features as follows. Let $\mathbf{h}_{\text{classical}} \in \mathbb{R}^d$ be the pooled feature produced by the `EfficientNetB0`+BiLSTM stack. We compute

$$\tilde{\mathbf{u}} = \tanh\big(\mathbf{W}_{\text{enc}}\mathbf{h}_{\text{classical}} + \mathbf{b}_{\text{enc}}\big) \in (-1, 1)^4, \qquad \mathbf{x} = \pi\, \tilde{\mathbf{u}} \in (-\pi, \pi)^4, \tag{1}$$

and then scale the per-qubit RY angles by the dynamic EMF: $\theta_i = x_i \cdot \text{EMF}$. This mapping (i) bounds the angles, (ii) is differentiable end-to-end, and (iii) avoids feature-norm explosions by compressing with tanh.

Clarification: what the quantum gates depend on, and what the gate path is used for In the implemented hybrid classifier, the quantum circuit uses standard unitary gates; there is no separate "gate path" variable inside the circuit simulation. The only way the surrogate parameters (R, λ, E_{J0}) enter the circuit is through the scalar angle-rescaling factor $\text{EMF}(R, \lambda, E_{J0})$, which modifies the effective encoding angles as $\theta_i = x_i \cdot \text{EMF}(R, \lambda, E_{J0})$ in the input rotations $R_Y(\theta_i)$. For a fixed (R, λ, E_{J0}) setting, EMF is treated as a constant multiplier, so the circuit outputs $\langle Z \rangle_i$ are differentiable with respect to the rotation angles x_i (and with respect to the trainable circuit parameters), enabling end-to-end gradient-based training. The dose-driven gate path $(\Phi(s), \Delta\phi(s))$ is used only in Fig. 5 to sample and visualize the surrogate potential (and its derivatives) along a clinically motivated trajectory; it is not an input to the quantum layer during training or inference. We first apply single-qubit rotations about the Y-axis to encode the input features into the quantum state. The rotation gate about the Y-axis, denoted $\text{RY}(\theta)$, is defined as: $\text{RY}(\theta) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix}$. Thus, the state after encoding is given by: $|\psi_1\rangle = \left(\bigotimes_{i=1}^{n} \text{RY}\big(x_i \cdot \text{EMF}\big)\right) |0\rangle^{\otimes n}$. This notation indicates that the RY gate is applied individually to each of the n qubits, and the initial state is the tensor product of n qubits all in the state $|0\rangle$.

Subsequently, the circuit consists of q repeating layers. For each layer $j = 1, \ldots, q$, a sequence of parameterized single-qubit gates is applied. Denote by $\mathbf{w}$ the collection of weights associated with the circuit, with appropriate segmentation such that for each qubit i and for each layer j, we have weight parameters $w_{i,1}^{(j)}$, $w_{i,2}^{(j)}$, and $w_{i,3}^{(j)}$ corresponding to different rotation operations. In our implementation, the following parameterized gates are employed:

- RY Rotation: $\mathrm{RY}\left(w_{i,1}^{(j)}\right)$.
- RZ Rotation: $\mathrm{RZ}\left(w_{i,2}^{(j)}\right)$, where $\mathrm{RZ}(\theta) = \begin{pmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{pmatrix}$,
- PhaseShift Gate: $\mathrm{PhaseShift}\left(w_{i,3}^{(j)}\right)$, which applies a phase factor $e^{iw_{i,3}^{(j)}}$ to the $|1\rangle$ component. For the j-th layer, the state evolution is described as:

$$|\psi_j^{(i)}\rangle = \mathrm{CNOT}_{\mathrm{Layer}} \left(\bigotimes_{i=1}^{n} \mathrm{PhaseShift}\left(w_{i,3}^{(j)}\right) \mathrm{RZ}\left(w_{i,2}^{(j)}\right) \mathrm{RY}\left(w_{i,1}^{(j)}\right) \right) |\psi_{j-1}\rangle, \tag{2}$$

  where the $\mathrm{CNOT}_{\mathrm{Layer}}$ denotes a layer of CNOT gates that entangle neighboring qubits in a predetermined topology (typically nearest-neighbor connectivity).
- The full evolution across all layers results in the final state:

$$|\psi_{\mathrm{final}}\rangle = U_{\mathrm{CNOT}} \cdot \left(\bigotimes_{i=1}^{n} \left[\mathrm{PhaseShift}\left(w_{i,3}^{(j)}\right) \mathrm{RZ}\left(w_{i,2}^{(j)}\right) \mathrm{RY}\left(w_{i,1}^{(j)}\right) \right] \right) \cdot \left(\bigotimes_{i=1}^{n} \mathrm{RY}\left(x_i \cdot \mathrm{EMF}\right) \right) |0\rangle^{\otimes n}. \tag{3}$$

  Where the operator $\mathrm{CNOT}_{\mathrm{Layer}}$ represents the entangling operation applied within a single layer of the quantum circuit. In each such layer, after all the single-qubit rotations (RY, RZ, and PhaseShift) are applied, the $\mathrm{CNOT}_{\mathrm{Layer}}$ acts to entangle neighboring qubits according to the predefined connectivity (typically nearest-neighbor). When the circuit is composed of q identical layers, the overall entangling operation is given by the sequential application (i.e., the product) of the CNOT layers across all the q layers. This cumulative operation is denoted by U_{CNOT} and can be formally expressed as

$$U_{\mathrm{CNOT}} = \prod_{j=1}^{q} \mathrm{CNOT}_{\mathrm{Layer}}^{(j)}, \tag{4}$$

  where $\mathrm{CNOT}_{\mathrm{Layer}}^{(j)}$ denotes the entangling operation in the j-th layer. Thus, U_{CNOT} encapsulates the total entanglement generated by the circuit, and its action is applied after all the single-qubit rotations and before the final measurements.
  Unless otherwise stated, we assume a linear nearest-neighbor topology (open

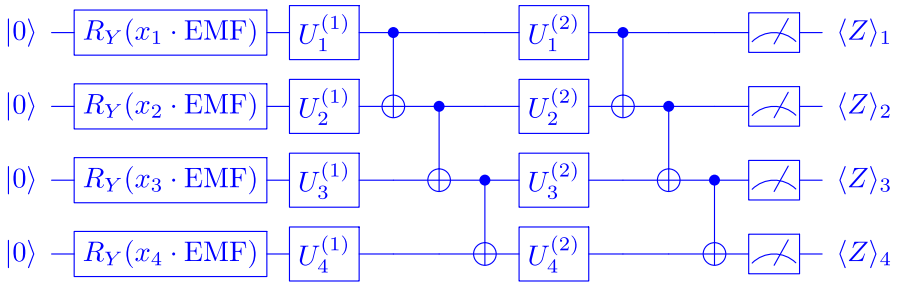


Fig. 1 Quantum circuit diagram of the 4-qubit, depth-2 ansatz used as the quantum layer. The input encoding applies $R_Y(x_i \cdot \text{EMF})$ on each qubit. Each layer $j \in \{1, 2\}$ applies trainable single-qubit blocks $U_i^{(j)}$ followed by a nearest-neighbor entangling pattern implemented as a controlled-NOT chain (1→2, 2→3, 3→4). Measurements yield the Pauli-Z expectation values $\langle Z\rangle_i$, which form the quantum feature vector $\mathbf{z}$

chain) for entangling layers; we do *not* assume all-to-all connectivity. Gate-dependent stochastic noise channels (e.g., depolarizing or amplitude/phase damping per CNOT) are *not* explicitly simulated; instead, the effect of quasi-static hardware imperfections enters through the effective-potential parameters (R, λ, E_{J0}) and the resulting deterministic EMF angle rescaling. We therefore refer to our quantum component as a parameterized, hardware-imperfection-aware surrogate rather than a fully realistic open-quantum-system model.

- Finally, measurement is performed on each qubit in the computational basis after applying the appropriate unitary transformations. We compute the expectation values of the Pauli-Z observables on each qubit: $\langle Z\rangle_i = \langle\psi_{\text{final}}|Z_i|\psi_{\text{final}}\rangle, \quad i = 1, \ldots, n.$

Collecting these expectation values produces a vector $\mathbf{z} \in \mathbb{R}^n$ which serves as the output of the quantum layer.

To make the structure of the quantum ansatz explicit, Fig. 1 provides a circuit diagram of the 4-qubit, $q = 2$ architecture used in this work. For compactness in the diagram, we denote the trainable single-qubit block acting on qubit i in layer j by

$$U_i^{(j)} = \text{PhaseShift}\left(w_{i,3}^{(j)}\right) \text{RZ}\left(w_{i,2}^{(j)}\right) \text{RY}\left(w_{i,1}^{(j)}\right).$$

This output vector is then concatenated with the classical feature vector obtained from the deep learning model, resulting in an integrated representation: $\mathbf{h}_{\text{hybrid}} = \mathbf{h}_{\text{classical}} \| \mathbf{z}$. This combined feature vector is subsequently fed into a dense fusion classifier (linear map followed by softmax) to perform the final classification task. The parameters $\mathbf{w}$ are optimized jointly with the classical network's parameters using gradient descent methods, facilitated by automatic differentiation within the PennyLane and TensorFlow [51] frameworks. Because the fusion stage is linear in $\mathbf{h}_{\text{hybrid}}$, the contribution of the quantum measurement vector $\mathbf{z}$ can be audited directly from the learned weight matrix of the subsequent dense layer.

Auditing whether the network utilizes the quantum features Let $\mathbf{s} \in \mathbb{R}^C$ denote the pre-softmax logits for C classes. The fusion classifier can be written as

$$\mathbf{s} \;=\; \mathbf{W}_{\text{fuse}}\,\mathbf{h}_{\text{hybrid}} + \mathbf{b}_{\text{fuse}}, \qquad \mathbf{W}_{\text{fuse}} \in \mathbb{R}^{C\times(d_{\text{c}}+d_{\text{q}})}, \;\; \mathbf{b}_{\text{fuse}} \in \mathbb{R}^{C}, \tag{5}$$

where $d_{\text{c}} = \dim(\mathbf{h}_{\text{classical}})$ and $d_{\text{q}} = \dim(\mathbf{z})$ (in our experiments $d_{\text{q}} = 4$). By partitioning the weight matrix into a classical block and a quantum block,

$$\mathbf{W}_{\text{fuse}} \;=\; \big[\mathbf{W}_{\text{c}} \;\; \mathbf{W}_{\text{q}}\big], \qquad \mathbf{W}_{\text{c}} \in \mathbb{R}^{C\times d_{\text{c}}}, \;\; \mathbf{W}_{\text{q}} \in \mathbb{R}^{C\times d_{\text{q}}}, \tag{6}$$

the logits decompose into

$$\mathbf{s} \;=\; \mathbf{W}_{\text{c}}\,\mathbf{h}_{\text{classical}} \;+\; \mathbf{W}_{\text{q}}\,\mathbf{z} \;+\; \mathbf{b}_{\text{fuse}}. \tag{7}$$

This expression makes explicit how the quantum measurements enter the decision function: if the model were to ignore the quantum output, the columns of $\mathbf{W}_{\text{q}}$ would be driven toward negligible magnitude under standard weight regularization and optimization. To compare blocks while accounting for their different dimensionalities, a simple dimension-normalized weight-scale diagnostic is

$$\rho_{\text{w}} \;=\; \frac{\|\mathbf{W}_{\text{q}}\|_{\text{F}}/\sqrt{d_{\text{q}}}}{\|\mathbf{W}_{\text{c}}\|_{\text{F}}/\sqrt{d_{\text{c}}}}, \tag{8}$$

which is straightforward to compute from the trained fusion layer and directly addresses whether the subsequent layer places non-trivial weight on the quantum feature channels.

The modular design of our quantum circuit and its integration with classical architectures allow for an elegant interplay between continuous classical data and the discrete quantum state space, enabling the extraction of rich, high-dimensional representations from medical imaging data.

3.3 Dynamic error mitigation strategy

Error mitigation in quantum circuits is critical in the NISQ era, where full error correction is infeasible. Instead, our methodology incorporates a dynamically updated angle-rescaling strategy: for each (R, λ, E_{J0}) configuration (or, in a hardware setting, for each calibration update), we compute an error mitigation factor from the effective potential and apply it multiplicatively to the circuit input angles. The dynamic EMF is computed as described above, using the mean gradient magnitude of the effective potential:

$$\text{EMF} = 1.0 + 0.001\,\langle\sqrt{\left(\frac{\partial U_{\text{eff}}}{\partial \Phi}\right)^2 + \left(\frac{\partial U_{\text{eff}}}{\partial \Delta\phi}\right)^2}\rangle. \tag{9}$$

This factor is applied multiplicatively to the input angles of the rotation gates in the quantum circuit: $\theta_i' = \theta_i \cdot \text{EMF}$. By doing so, the circuit compensates for quasi-static

coherent over-/under-rotations implied by the surrogate parameters, rather than for stochastic fluctuations. For clarity, the dose-driven gate path in Fig. 5 is not used to compute EMF; EMF is computed from the domain-averaged gradient of U_{eff} on the $(\Phi, \Delta\phi)$ mesh, and then applied as a global multiplicative rescaling of the circuit input angles for each (R, λ, E_{J0}) configuration. The error mitigation strategy is designed to be computationally efficient; it leverages numerical gradient methods (implemented via finite differences) to approximate the gradient fields of the potential landscape and updates the EMF at each iteration of the grid search in the experiment. In a physical deployment, the same update would be tied to the acquisition of new calibration data (minutes to hours), not to individual gate executions.

Connection between (R, λ, E_{J0}) *and gate fidelity* From a neural-network viewpoint, (R, λ, E_{J0}) act as hyperparameters because they are selected externally (grid search) and determine the multiplicative factor applied to the quantum input angles. Operationally, our encoding applies $R_Y(x_i \cdot \text{EMF})$; thus changing (R, λ, E_{J0}) changes EMF and therefore changes the effective rotation angles, which is equivalent to a coherent amplitude miscalibration of the single-qubit encoding gates. In the present surrogate, this is the primary mechanism by which (R, λ, E_{J0}) influence circuit behavior; we do not explicitly map these parameters to separate gate-level stochastic noise channels for the entangling operations. Nevertheless, coherent mis-rotations change the state that is subsequently entangled, which can alter the measured expectation values and thereby the downstream classification performance. Physically, R captures residual drive-amplitude scaling (over-/under-rotation), λ captures residual coupling/higher-harmonic contributions that increase curvature and leakage/crosstalk susceptibility, and E_{J0} sets the energy scale (frequency and anharmonicity) that constrains fast, low-leakage gate operation. In this study we do not train (R, λ, E_{J0}) end-to-end; we sweep them to quantify sensitivity and to emulate different calibration states of the same device.

3.4 Integrated hybrid model and training procedure

Figure 2 summarizes the complete hybrid network, highlighting the dataflow between the classical feature extractor, the angle encoder, the error-mitigation-driven angle rescaling, and the quantum feature augmentation prior to the final classifier.

The model is optimized using backpropagation with the Adam optimizer and an exponential decay learning rate schedule: $\eta_t = \eta_0 \cdot \exp\left(-\frac{t}{\tau}\right)$, where $\eta_0 = 1 \times 10^{-4}$ is the initial learning rate, t is the training step, and τ is a decay parameter. We selected η_0 via a lightweight validation sweep prior to the main experiments: short pilot trainings were run with $\eta_0 \in \{1 \times 10^{-5}, 3 \times 10^{-5}, 1 \times 10^{-4}, 3 \times 10^{-4}, 1 \times 10^{-3}\}$, and we chose the value that achieved the best validation accuracy while maintaining stable optimization (i.e., no divergence and no large loss oscillations). The selected η_0 was then held fixed for all reported experiments to avoid confounding the superconducting-circuit parameter sweep with optimizer retuning. The loss function is the categorical cross-entropy. For each sample the cross-entropy loss is computed by summing over all classes, and then the overall loss is obtained by averaging these values over all N samples in the batch. Specifically, for a single sample the loss is defined as: $\mathcal{L}_i =$

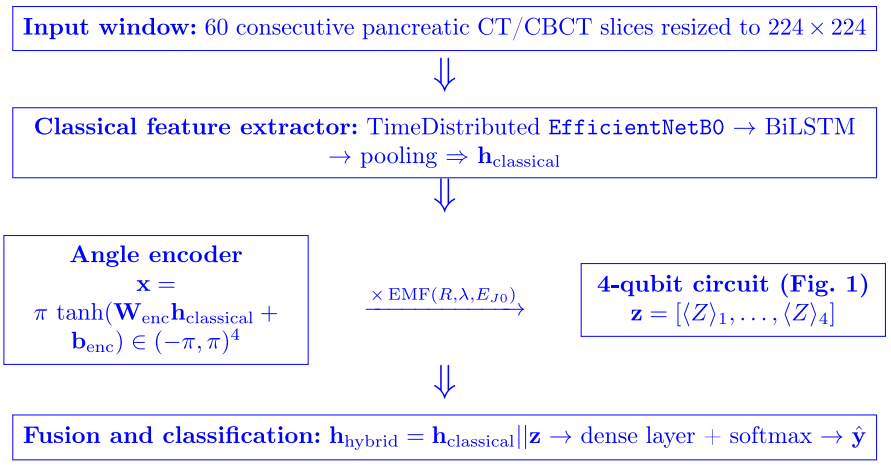


Fig. 2 Schematic overview of the proposed hybrid pipeline. A convolutional–recurrent backbone produces a pooled classical embedding $\mathbf{h}_{\text{classical}}$. A learned encoder maps $\mathbf{h}_{\text{classical}}$ to four bounded rotation angles $\mathbf{x}$, which are rescaled by the dynamically computed error-mitigation factor (computed from the effective-potential surrogate parameterized by R, λ, and E_{J0}) and then passed through a 4-qubit parameterized circuit. The resulting expectation-value vector $\mathbf{z}$ is concatenated with $\mathbf{h}_{\text{classical}}$ for final classification

$-\sum_{c=1}^{C} y_i(c) \log(\hat{y}_i(c))$, where C represents the number of classes. The total loss over a batch is then computed as: $\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} \mathcal{L}_i$. This formulation ensures that the summation is performed over the class dimension for each sample and that the final loss is obtained by averaging over the batch, thereby enhancing both the interpretation and reproducibility of the process. The inclusion of the quantum circuit introduces additional parameters (the quantum weights), which are jointly optimized with the classical network.

The training procedure employs a grid search over the SQUID–Transmon parameters R, λ, and E_{J0}. For each combination, the quantum circuit's simulation is updated, and the corresponding EMF is recalculated. The hybrid model is then trained for a fixed number of epochs on the training dataset, with performance monitored on a validation set. Finally, the model is evaluated on a separate test set to determine its classification accuracy. The grid search allows us to rigorously examine how variations in the physical parameters affect the performance of the hybrid model, thereby illuminating the optimal operating conditions for the system.

3.5 Mathematical and computational considerations

The methodologies described in this work require a careful balancing of computational resources and mathematical precision. The deep learning component leverages transfer learning to minimize the need for extensive training from scratch, while the bidirectional LSTM and Fourier encoding layers capture both sequential dependencies and periodic features, respectively. The Fourier encoding is based on the transform: $\mathcal{F}(\theta) = \left[\sin(\theta \cdot \mathbf{w}), \cos(\theta \cdot \mathbf{w})\right]$, where $\mathbf{w}$ is a frequency vector derived from a log-

arithmic scale. This transformation has been shown to improve the representational capacity of the network in high-dimensional settings [52].

On the quantum side, each gate in the SQUID–Transmon simulation is a unitary operator U that, in the ideal case, would satisfy $U^\dagger U = I$. However, in the presence of noise and imperfections, these operations deviate from unitarity. Our error mitigation strategy seeks to re-establish an effective unitary evolution by dynamically adjusting the rotation angles according to the computed EMF. The adjustments are linear in the angles, which is justified under the assumption that the errors are small and can be approximated by a first-order Taylor expansion:

$$\sin(\theta \cdot \mathrm{EMF}) \approx \sin(\theta) + \theta\,(\mathrm{EMF} - 1)\cos(\theta). \tag{10}$$

This approximation provides a tractable means of incorporating error mitigation within our quantum circuit without resorting to computationally intensive corrective procedures.

Furthermore, the grid search over hardware parameters is implemented in a reproducible manner by setting fixed random seeds for both NumPy [53] and TensorFlow. The search algorithm is distributed across the parameter space in a nested loop, ensuring that every combination of R, λ, and E_{J0} is evaluated. The output of the grid search–a set of test accuracies–is then used to generate heatmaps that visualize performance across the parameter space. These visualizations help in identifying optimal settings, particularly the counterintuitive observation that a slight reduction in R (i.e., $R = 0.95$ as opposed to $R = 1.00$) leads to improved classification performance.

3.6 Implementation details

Computational environment and software versions All experiments were executed on an Apple Mac Studio featuring an Apple M3 Ultra system-on-chip (28-core CPU, 60-core GPU), 96 GB unified memory, and macOS Sequoia 15.6. The pipeline was executed in a Jupyter notebook environment (Python 3.13.11, JupyterLab 4.5.2, Notebook 7.5.2). Table 2 reports the versions of the main software packages used for data preprocessing, training, and quantum-circuit simulation.

The entire experimental pipeline is implemented in Python, leveraging a suite of scientific libraries such as NumPy, SciPy [54], OpenCV [55], and Pandas [56] for data handling and preprocessing. TensorFlow and the Keras API are used to construct and train the deep learning model, while PennyLane serves as the bridge between classical and quantum computations. The codebase is structured to ensure modularity and reproducibility, with helper functions for folder traversal, DICOM file reading, image augmentation, and CT volume slicing. Each module is tested independently to verify that the transformations and data flows are correctly implemented.

The integration of the quantum circuit is achieved via a custom Keras layer, termed `PennyLaneLayer`. This layer wraps the quantum circuit defined with PennyLane's QNode interface, allowing the quantum part of the model to be seamlessly incorporated into the TensorFlow computational graph. The layer is designed to handle batched inputs, iterating over individual samples and applying the quantum opera-

Table 2 Reproducibility checklist: classical hardware and software versions used to execute the pipeline

Component	Version/specification
Workstation	Apple Mac Studio (Apple M3 Ultra, arm64)
CPU	28-core Apple M3 Ultra
GPU	60-core Apple M3 Ultra
RAM	96 GB unified memory
Operating system	macOS Sequoia 15.6
Python	3.13.11
JupyterLab	4.5.2
Notebook	7.5.2
TensorFlow (incl. `tf.keras`)	2.20.0
PennyLane	0.44.0
NumPy	2.4.1
SciPy	1.17.0
pandas	2.3.3
OpenCV (`opencv-python`)	4.13.0.90
scikit-learn	1.7.2
matplotlib	3.10.6

tions to extract a set of expectation values from the Pauli-Z measurements, which are subsequently concatenated to form the quantum feature vector.

The training routine is automated to run a comprehensive grid search, where for each parameter configuration, the model is compiled, trained, and evaluated. Intermediate results, including loss curves and accuracy metrics, are logged and stored for further analysis. The entire process is computationally intensive, necessitating the use of high-performance computing resources. However, the modularity of the code allows for parallel execution over different parameter sets when such resources are available.

3.7 Conclusion of the methodology section

The methodology presented in this work is a confluence of advanced deep learning techniques and state-of-the-art quantum simulation. Our approach provides a mathematically rigorous and computationally efficient framework for classifying complex medical images. By integrating a parameterized, hardware-imperfection-aware SQUID–Transmon surrogate with dynamic error mitigation and embedding it within a hybrid model, we offer a pathway to overcome the limitations of both classical and early quantum algorithms. The detailed mathematical derivations, combined with systematic parameter exploration, ensure that our framework is both robust and adaptable. This methodology not only addresses the immediate challenges in pancreatic radiotherapy image classification but also sets the stage for broader applications in fields where data are scarce and hardware limitations are a significant barrier to progress.

4 Experimental setup

In this section, we describe the experimental setup employed to evaluate our hybrid quantum–classical framework. We detail the configuration of the `Pancreatic-CT-CBCT-SEG` dataset, the preprocessing and sample generation methods, and the grid search strategy used to optimize the quantum hardware parameters within the parameterized, hardware-imperfection-aware SQUID–Transmon surrogate. The experiments were designed to assess the impact of key hardware parameters on the classification accuracy of pancreatic radiotherapy images, which are inherently challenging due to various imaging artifacts and the complex anatomy of the disease site.

4.1 Dataset description

The `Pancreatic-CT-CBCT-SEG` dataset [28] provides imaging data from 40 pancreatic cancer patients undergoing radiation therapy. For each patient, the dataset includes one treatment-planning CT scan and two on-treatment CBCT scans. In this study, we use only the image volumes (axial slices) as inputs to the classification pipeline.

The supervised target used in our experiments is patient identity (one class per patient). Accordingly, the task is defined as a subject-level image classification problem rather than an organ segmentation, registration, or outcome-prediction task.

Although the dataset also contains expert organ-at-risk contours and radiotherapy dose distributions, these additional files are not used as model inputs or targets in the present work. We use the dose files only for two auxiliary purposes: (i) selecting the 10-patient subset used for the computationally intensive hybrid parameter sweep (patients ranked highest by dose-volume coverage) and (ii) constructing a dose-derived trajectory used solely for the visualization of the effective-potential surrogate in Fig. 5.

4.2 Data pre-processing and sample generation

The pre-processing pipeline was designed to handle the high-dimensional imaging data while ensuring consistency across the CT and CBCT modalities. The process begins by recursively traversing the directory structure to identify DICOM files corresponding to the CT modality. Dedicated helper functions normalize the image intensities by scaling the pixel values and converting grayscale images to RGB format when necessary. Each CT image is resized to a target resolution of 224×224 pixels, a standard input size for many deep learning models.

To augment the limited dataset, we implemented a sliding window mechanism to generate multiple samples from each patient's CT volume. Specifically, if the CT volume contains fewer slices than the chosen window size (set to 60 slices), the volume is padded with zeros; otherwise, a sliding window with a stride of 9 slices is used to extract multiple contiguous sub-volumes. Each sub-volume is paired with a patient-identity label (one class per patient). We consider two evaluation settings: (i) a 10-class subset (top-dose patients) used for the hybrid grid-search analysis, and (ii) the full set of **40** patients used to report the requested window-level consistency via single-split metrics

and k-fold cross-validation. This separation avoids confounding the physics-driven grid-search (manageable on a smaller subset) with the cross-validated consistency analysis (on all patients). Figure 3 exemplary shows the data slices of one of these 10 patients.

Once extracted, the image volumes are converted into TensorFlow datasets via a generator function that yields the processed sub-volumes alongside the associated class labels. Each dataset is batched with a batch size of 1 to facilitate fine-grained training and evaluation. Subsequently, a selection of 80% of these data for training, 10% for validation, and 10% for test sets is vital to ensure that the deep learning model can learn robust representations from the heterogeneous imaging data.

4.3 Parameter grid search strategy

A central contribution of our work is the systematic investigation of how specific SQUID–Transmon hardware parameters influence the performance of our hybrid model. We focus on three key parameters: the rescaling factor (R), the coupling strength or regularization parameter (λ), and the nominal Josephson energy (E_{J0}). A grid search was performed over discrete sets of candidate values for each parameter. The specific parameter ranges chosen are shown in Table 3.

For each combination of parameters, the SQUID–Transmon simulation recalculates the effective potential using a predefined model. The simulation involves computing the cosine-modulated Josephson energy and additional correction terms reflecting the influence of the coupling parameter. Through numerical methods, gradients of the potential are determined, and an error mitigation factor is computed as a function of the mean gradient magnitude. This factor is then used to adjust the quantum circuit parameters, thereby ensuring that the circuit operates within a dynamically optimized imperfection-compensated regime.

Each parameter configuration was evaluated by retraining the entire hybrid model from scratch on the training dataset, followed by validation on a reserved subset of the data. The performance metric of interest was classification accuracy on the test set, which reflects the model's ability to correctly classify pancreatic radiotherapy images into one of 10 distinct classes. Detailed logs of the training process, including accuracy and loss metrics, were recorded for each experimental run.

The grid search was computationally intensive, given that the model must be retrained for every parameter configuration. However, this exhaustive search is essential to elucidate the complex interplay between the SQUID–Transmon parameters and the overall performance of the hybrid framework. Preliminary results indicate that the best performance is achieved at $R = 0.95$, a finding that underscores the non-intuitive nature of optimal quantum hardware calibration; although ideal calibration might suggest $R = 1.00$, our experiments reveal that a slightly lower value achieves a more favorable balance between gate fidelity and effective error mitigation.

Fig. 3 Illustration of CT slice panels from one of the top 10 selected patients. A sliding window approach with a window size of 60 slices and a stride of 9 was applied to augment the dataset. If a CT volume had fewer than 60 slices, zero-padding was used. Each sub-volume generated by this method is labeled by patient identity (one class per patient). Radiotherapy dose files are used only to rank patients for selecting the 10-class subset; dose values are not used as model inputs or targets

Table 3 Parameter ranges used in the experiments

Parameter	Values
R	{0.90, 0.95, 1.00}
λ	{1.0, 2.0, 3.0}
E_{J0}	{15.0, 20.0, 25.0} (in appropriate energy units)

Table 4 Window-level metrics on all 40 patients

Setting	Train Acc	Val Acc	T4est Acc
Single 80/10/10 split	0.9591	0.9598	0.9094
3-fold CV (mean ± SD)	–	0.8758 ± 0.0050	–

4.4 Evaluation Protocols and Data Splits

Window-level single split (all 40 patients) We stratify windows per class into 80%/10%/10% (train/val/test) without augmentation leakage across splits. This yields: Train windows $= 1932$, Validation $= 224$, Test $= 276$. Using the classical backbone alone with `EfficientNetB0` kept frozen (no fine-tuning) and training restricted to the bidirectional LSTM and the subsequent classifier layers, we obtain train $= 0.959$, val $= 0.960$, test $= 0.909$ accuracy.

We perform window-level stratified CV with $k = 3$ folds (constrained by minimum per-class windows). Mean validation accuracy is 0.8758 ± 0.0050 (min/max: 0.8705/0.8804). Fold sizes are approximately $n_{\text{train}} \approx 1621$, $n_{\text{val}} \approx 811$ windows per fold (Table 4).

Figure 4 summarizes the learning curves, class distributions, and CV statistics (all images produced by the provided code).

4.5 Summary of the Experimental Pipeline

The complete experimental pipeline proceeds as follows: the `Pancreatic-CT-CBCT-SEG` dataset is first preprocessed to extract consistent CT and CBCT image volumes. Organ-at-risk contours and radiotherapy dose distributions available in the collection are not used as model inputs or targets in this study (dose data are used only for subset selection and for a visualization-only trajectory). These images are then organized into training, validation, and test sets using a sliding window approach to generate multiple samples per patient. A classical deep learning model is constructed with a pretrained feature extraction backbone and recurrent layers, forming the primary component of the hybrid framework. This is integrated with a parameterized quantum circuit simulated to emulate a hardware-imperfection-aware SQUID–Transmon surrogate, with a dynamic error mitigation factor calculated from the SQUID–Transmon model. A rigorous grid search over R, λ, and E_{J0} is then performed, with the trained model evaluated on a held-out test set for each parameter combination. The final evalua-

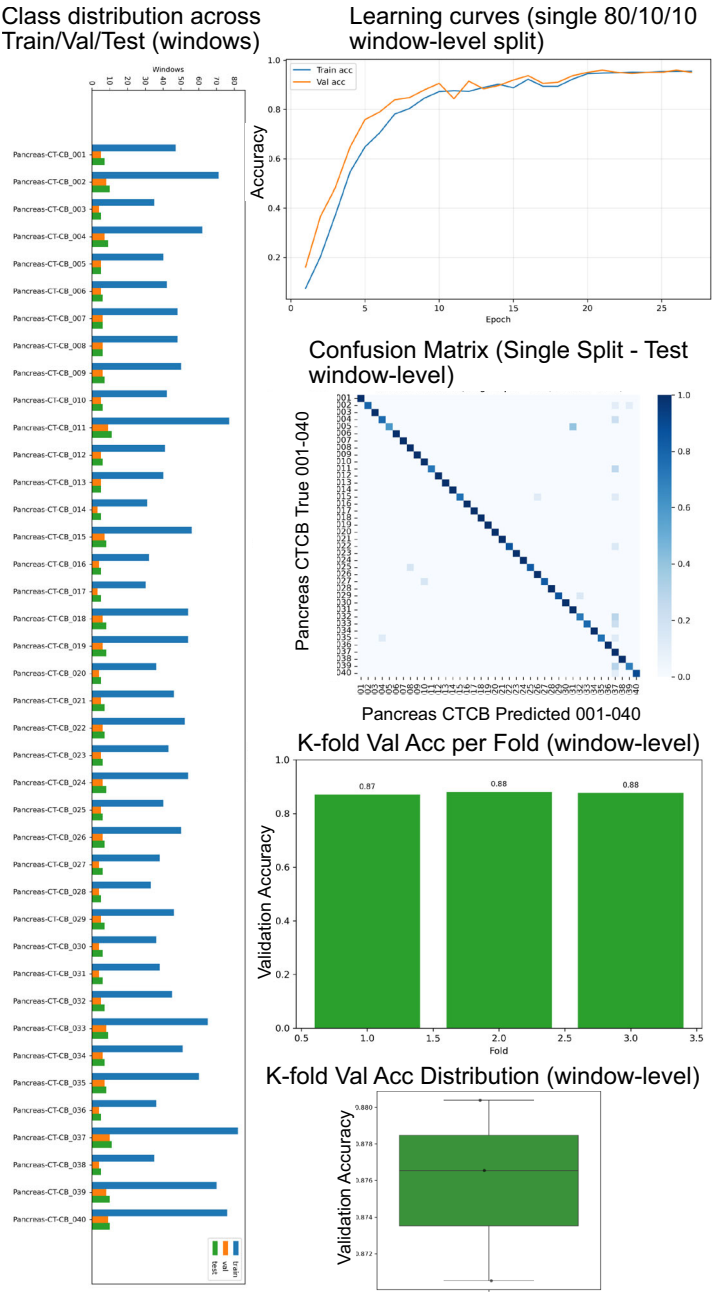


Fig. 4 Window-level evaluation over all 40 patients. Left: per-split class window counts (rotated). Right (top to bottom): single-split learning curves, test confusion matrix, per-fold CV validation accuracy, and CV accuracy distribution

tion includes both quantitative performance metrics such as accuracy and qualitative assessments of the model's generalization to the inherent imaging challenges present in pancreatic radiotherapy.

The detailed codebase developed for these experiments covers functions for image pre-processing, dose data extraction, CT volume sample generation, and the construction and training of the hybrid model. All experiments were conducted in a reproducible environment with randomized seed settings, ensuring that comparisons across different parameter configurations are fair and statistically robust.

The comprehensive design of our experimental setup allows for a thorough investigation into the effects of a parameterized, hardware-imperfection-aware SQUID–Transmon surrogate on the performance of hybrid quantum–classical models in medical image classification. By leveraging a dataset of clinically relevant pancreatic CT and CBCT scans and implementing a robust grid search strategy for key hardware parameters, our framework serves as a testbed for future studies aiming to integrate advanced quantum hardware simulations into real-world diagnostic applications.

5 Results

In this section, we present a series of plots and quantitative results that evaluate the performance of our hybrid quantum–classical framework. At the outset, the primary purpose of Fig. 5 is to provide a physics-based interpretation of the chosen parameter setting by illustrating that, for $(R, \lambda, E_{J0}) = (0.95, 2.0, 20.0)$, the dose–driven gate path passes predominantly through regions of positive curvature in the effective-potential landscape, which we interpret as a comparatively more stable operating regime in the surrogate model. The results are presented with a focus on both the quantum simulation of the SQUID–Transmon system and the overall classification accuracy of pancreatic radiotherapy images.

We first present in Fig. 5 a comprehensive analysis of a 3D SQUID–Transmon system. Accordingly, this visualization is intended to guide the reader's interpretation of local curvature and stability under the chosen parameter setting, rather than to serve as a direct performance metric. We focus on the simulated potential energy landscape and illustrate it using a dose–driven gate path. Here, the "gate path" is a visualization-only trajectory used to sample the effective potential and its curvature along a clinically motivated profile; it is not an additional input to the quantum circuit executed in the hybrid classifier.

Derivation of the dose–driven gate path The irregular black curve overlaid in Fig. 5ais constructed by mapping a one-dimensional radiotherapy dose profile into the $(\Phi, \Delta\phi)$ domain of the effective-potential mesh. Starting from the radiotherapy dose grid, we compute a slice-wise dose statistic inside the volume of interest (in our implementation: the mean dose per axial slice), yielding a sequence $\{d_k\}_{k=1}^{K}$. We normalize it to $[0, 1]$ as $\tilde{d}_k = (d_k - \min d)/(\max d - \min d)$. We then map the normalized dose and its local variation into Φ and $\Delta\phi$ via simple affine scalings that keep the curve within the same domain used to compute U_{eff}:

$$\begin{aligned} \Phi_k &= \Phi_{\min} + \tilde{d}_k(\Phi_{\max} - \Phi_{\min}), \\ g_k &= d_{k+1} - d_k, \\ \tilde{g}_k &= \frac{g_k - \min g}{\max g - \min g}, \\ \Delta\phi_k &= \Delta\phi_{\min} + \tilde{g}_k(\Delta\phi_{\max} - \Delta\phi_{\min}). \end{aligned} \tag{11}$$

The plotted trajectory is the piecewise-linear interpolation of the discrete points $(\Phi_k, \Delta\phi_k)$, parameterized by the normalized arc length $s \in [0, 1]$. Consequently, the path irregularity is inherited from heterogeneity in the clinical dose profile, rather than from manual tuning.

The following aggregated multi-panel figure integrates complementary views that collectively show the interplay between device physics and error management:

- In Panel Fig. 5a, we assess the leakage risk by mapping the Laplacian $\nabla^2 U$ of the potential energy landscape. The Laplacian provides a direct measure of the local curvature, thereby indicating regions of the potential that are prone to quantum leakage and non-adiabatic transitions. In this color-coded map, blue tones represent negative values of $\nabla^2 U$, red tones indicate positive values, and white tones denote regions where the Laplacian is near zero. Negative Laplacian values (blue) correspond to concave downward curvature, which typically indicates that the potential is near a local maximum or a saddle point. These regions are indicative of local instability, where small perturbations may lead to enhanced leakage or non-adiabatic transitions. In contrast, positive Laplacian values (red) suggest concave upward curvature, indicative of local minima where the potential well is deeper and the system is more stable. Regions with Laplacian values around zero (white) signify a transition between these curvatures, marking areas of relatively neutral stability. The overlay of the black gate path on this leakage risk map provides a clinically motivated probe of where a dose-derived trajectory falls in the curvature landscape. The path is not optimized to avoid curvature extrema; instead, it is used to diagnose whether such a trajectory would traverse regions associated with elevated curvature (and therefore elevated leakage risk) in the surrogate potential.
- Panel 5b provides a quantitative analysis of the local curvature along the dose–driven gate path. In this plot, the x-axis represents the normalized arc-length parameter $s \in [0, 1]$, indexing progression along the interpolated gate path $(\Phi(s), \Delta\phi(s))$, while the y-axis shows the corresponding Laplacian values, $\nabla^2 U$. A horizontal dashed line at zero serves as a clear demarcation between regions of positive and negative curvature. Negative values of $\nabla^2 U$ in this context are particularly significant. They indicate that the potential energy landscape is locally concave down, which implies that the corresponding segment of the gate path is positioned at or near a local maximum or a saddle point in the potential. Physically, such negative curvature regions suggest that the system is locally unstable; small deviations from the gate path in these zones could lead to enhanced leakage or trigger non-adiabatic transitions. In contrast, regions where $\nabla^2 U$ is positive denote local convexity, implying a more stable configuration that tends to con-

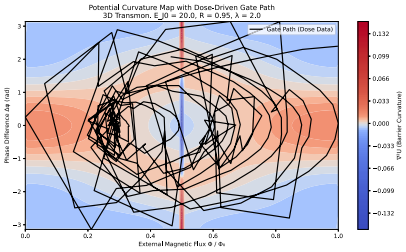
fine the system within a potential well. Therefore, the identification of negative Laplacian values along the gate path is crucial. It not only highlights sections where the quantum circuit is more vulnerable to perturbations but also informs the optimization of dynamic error mitigation strategies. By closely examining these transitions, marked by the change in sign relative to the zero-horizontal reference, we can infer the points along the path where error control measures might need to be intensified to maintain high-fidelity operation in the SQUID–transmon system.

- Finally, panel 5c illustrates the gate-induced energy modulation along the dose–driven trajectory. To obtain the gate-induced energy modulation, we evaluate $U_{\text{eff}}(\Phi(s), \Delta\phi(s))$ along the gate path by interpolation on the precomputed $(\Phi, \Delta\phi)$ mesh, thereby capturing the energy fluctuations along the trajectory. In this context, positive values of U_{eff} indicate segments of the gate path where the effective potential energy exceeds a defined baseline. These elevated regions can be interpreted as energy barriers that the system must overcome during operation. Such barriers may play a role in confining the quantum state, thereby reducing the probability of unwanted transitions; however, they may also demand additional energy input to maintain proper gate function, which could be critical in balancing performance and error resilience. Conversely, negative values of U_{eff} correspond to regions where the energy is below the baseline. These sections typically represent potential wells or areas of reduced energy cost along the gate path. While lower energy regions are generally favorable for stable operation due to their inherent energetic favorability, they may also signal reduced confinement. Such reduced confinement could potentially expose the system to risks of leakage or decoherence if the energy drop is too pronounced. Overall, the interplay between positive and negative U_{eff} values provides a quantitative metric for assessing the energetic cost and local stability along the gate trajectory. By correlating these energy fluctuations with the overall performance of the hybrid model, we can optimize our error mitigation strategy to ensure that the gate path not only minimizes energy overhead but also maintains a robust defense against leakage and non-adiabatic transitions.

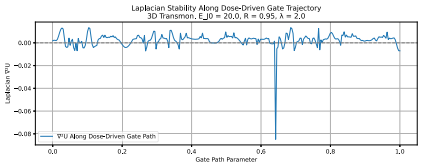
In summary, the integrated Fig. 5 provides a cohesive overview of the key aspects of the SQUID–Transmon system behavior. The comprehensive presentation enables a direct comparison of the potential leakage risks, local curvature changes along the gate path, and the corresponding energy modulation. This cohesive analysis underpins the robustness of our dose-driven dynamic error mitigation strategy and supports further optimization of quantum circuit performance in hybrid quantum systems.

To investigate the impact of the SQUID–Transmon parameters on classification performance, we conducted a grid search over the rescaling factor R, the coupling strength λ, and the nominal Josephson energy E_{J0}. The experimental results are summarized in a series of heatmaps that display the test accuracy for each combination of these parameters.

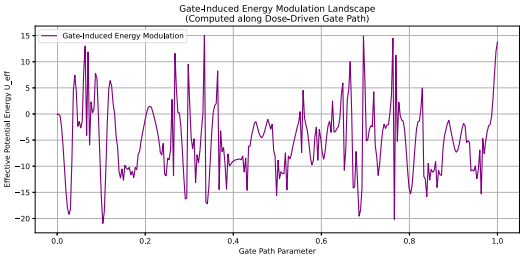
Figure 6 aggregates the hybrid test accuracy for the *10-class subset* used in the grid search. The best performance is observed when $R = 0.95$, with peak test accuracy of 0.95 sustained across several (λ, E_{J0}) settings. We interpret this as an effective mitigation of over-rotation/leakage when the drive amplitude is slightly reduced in



(a) Leakage risk map based on the Laplacian $\nabla^2 U$ with the dose–driven gate path derived from the radiotherapy dose profile (black curve; see text for construction). Stronger colors indicate regions with high curvature and increased leakage risk.



(b) Laplacian $\nabla^2 U$ along the dose–driven gate path. The x-axis is the normalized arc-length parameter $s \in [0, 1]$ of the interpolated curve. The horizontal dashed line at zero aids in highlighting curvature transitions and local stability.



(c) Gate-induced effective potential energy U_{eff} along the dose-driven gate path. The x-axis is the normalized arc-length parameter $s \in [0, 1]$, and U_{eff} is obtained by interpolation on the precomputed potential mesh.

Fig. 5 Aggregate multi-panel representation of the SQUID–Transmon simulation results using consistent parameter settings ($E_{J0} = 20.0$, $R = 0.95$, $\lambda = 2.0$). Panel **a** depicts the leakage risk based on the Laplacian, **b** details the Laplacian along the gate path, and **c** displays the gate-induced energy modulation

our simulator. Importantly for attribution, the sweep over (R, λ, E_{J0}) modifies only the quantum-side angle rescaling (and hence only the quantum measurement vector $\mathbf{z}$), while the classical feature extractor is unchanged; therefore, the non-trivial accuracy variation across the grid provides indirect evidence that the trained classifier is sensitive to $\mathbf{z}$ rather than relying exclusively on $\mathbf{h}_{\text{classical}}$.

To assess consistency on the *full 40-patient dataset*, we report the classical backbone alone at the window level (requested by the reviewer): a single 80/10/10 split yields 0.959/0.960/0.909 (train/val/test) accuracy; stratified threefold CV gives 0.8758 ± 0.0050 mean validation accuracy (Sect. *Evaluation Protocols and Data Splits* and Fig. 4). These results indicate that, while the hybrid layer provides a hardware-aware control and retains competitive performance on the 10-class subset, the strong classical baseline already performs well at scale; thus we avoid claims that the hybrid universally outperforms the classical model.

These results underscore the importance of fine-tuning quantum hardware parameters in the context of a hybrid quantum–classical model. Our experiments reveal

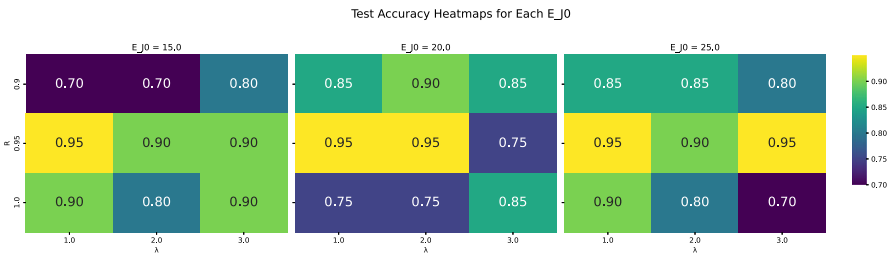


Fig. 6 Heatmaps of test accuracy for different values of E_{J0}. For each fixed E_{J0}, the heatmap plots test accuracy as a function of R (y-axis) and λ (x-axis). The consistent observation of high accuracy at $R = 0.95$ is evident across various configurations

that the optimal configuration is achieved at $R = 0.95$, suggesting that a marginally sub-ideal readout fidelity provides a more robust operating condition in the presence of non-idealities. The performance degradation observed at $R = 1.00$ supports the hypothesis that increased non-adiabatic effects and leakage errors can outweigh the benefits of an idealized measurement setup.

Overall, the presented plots and results provide a detailed quantitative and qualitative evaluation of our framework. The visualizations serve not only as a confirmation of the model's performance but also as a diagnostic tool to understand the interplay between quantum hardware simulation, error mitigation strategies, and classical deep learning performance. Future work will extend these analyses to additional datasets and refine the error mitigation techniques to further enhance model robustness.

6 Discussion

The experimental data presented herein arise from a comprehensive grid search over three key hardware-relevant parameters in our quantum-enhanced model for pancreatic cancer treatment prediction. In this investigation, the parameters examined include: (i) the readout fidelity (or rescaling factor) R, which may be interpreted as a proxy for circuit depth or reset error level; (ii) the hyperparameter λ, which is linked to regularization effects or, equivalently, the coupling strength in a Hamiltonian formulation; and (iii) the nominal Josephson energy E_{J0}, a critical parameter for superconducting qubits that influences both coherence times and gate fidelities.

The overall predictive accuracy, as depicted in Fig. 6, reflects the performance of the model under varying configurations of these parameters. Notably, the highest accuracy values, reaching 0.95, are consistently achieved when the readout fidelity is set at $R = 0.95$, irrespective of the chosen values for λ and E_{J0}. This result suggests that $R = 0.95$ marks an operational regime that is robust to the modeled systematic hardware imperfections, possibly due to an effective balance between expressivity and reduced over-rotation/leakage induced by miscalibration. In practical terms, this observation implies that calibrating the device to a readout fidelity slightly lower than the idealized $R = 1.00$ might in fact be more advantageous, thereby mitigating deleterious effects such as non-adiabatic transitions and leakage errors.

Further analysis indicates that, at $R = 0.95$, increases in the nominal Josephson energy E_{J0} (e.g., from 15 to 25 in our normalized units) do not cause a notable degradation in accuracy. This observation supports the notion that the system operates within a robust regime of the Josephson junction, in which the higher energy levels do not compromise control fidelity. Conversely, when R is increased to 1.00, higher values of E_{J0} are sometimes accompanied by lower accuracy (for instance, a drop to 0.70 in the case of $\lambda = 3$ and $E_{J0} = 25$). Such performance declines are likely indicative of leakage or non-adiabatic effects, underscoring a hardware-induced trade-off between the benefits of a higher energy-level separation and the limitations imposed by control fidelity and coherence.

Although the influence of the hyperparameter λ on model performance appears to be less dominant relative to the other parameters, its role is nonetheless significant. Instances where $R = 0.95$, $\lambda = 3.0$, and $E_{J0} = 20$ yield a lower accuracy of 0.75 suggest that an excessive coupling strength may lead to over-parameterization or saturation within the entangling layers of the quantum circuit. This observation points to a nonlinear interaction between the regularization imposed by λ and the physical control strengths, indicating that fine-tuning of λ is necessary in conjunction with the other parameters to optimize overall performance.

From the perspective of clinical application, these findings have important implications. The consistent observation of high accuracy (around 0.95) at $R = 0.95$ across multiple configurations suggests that current NISQ devices may operate in a regime that is sufficiently robust for meaningful clinical predictions in pancreatic cancer treatment. Moreover, the fact that the system maintains stability over a range of E_{J0} values (specifically between 20 and 25 in our normalized units) indicates that existing superconducting hardware may already be well-suited for tasks in translational quantum machine learning within oncology. The ability to consistently achieve accuracies in the range of 90–95% also reinforces the potential of quantum-assisted classification to impact high-stakes medical decisions, where even small improvements in accuracy could lead to significant benefits in patient outcomes.

Based on our observations, we propose several recommendations for future hardware-aware experiments. First, it is advisable to stabilize the operating conditions of the quantum circuit around $R = 0.95$ during both circuit compilation and calibration, as our experiments indicate that a slight deviation from the ideal $R = 1.00$ improves error mitigation and overall performance. Second, a more detailed exploration of the hyperparameter λ is recommended, with fine sweeps in the range of 1.0 to 2.0, particularly for devices operating at nominal Josephson energies E_{J0} between 20 and 25 (normalized units). Employing a finer grid or alternative optimization strategies (e.g., Bayesian optimization) may reveal intermediate optima or more complex interactions among the parameters.

Regarding hardware non-idealities, our simulation focuses on quasi-static parameter offsets captured by the effective potential $U_{\text{eff}}(R, \lambda, E_{J0})$ and the resulting angle-rescaling factor. These terms represent systematic calibration biases and slow drift rather than stochastic noise processes. Consequently, the present model does not explicitly simulate relaxation/dephasing times, gate-dependent stochastic channels, or crosstalk; we therefore describe the surrogate as hardware-imperfection-aware rather than as a full open-system noise model. While we do not explicitly model decoher-

ence times (T_1 and T_2) or other specific noise channels, future work could extend the noise model to include these factors explicitly, providing a more comprehensive representation of NISQ device behavior.

Our current quantum circuit is implemented with $n = 4$ qubits and a circuit depth of $q = 2$ layers. Although this design is small compared to the complexity of the classical component (`EfficientNetB0` + bidirectional LSTM), it was chosen as a proof-of-concept within the constraints of current NISQ hardware and computational resources. Future research could explore scaling the quantum subsystem to a higher number of qubits and greater depth. Such scaling will address the challenges of increased computational cost, enhanced multi-qubit coherence requirements, and more sophisticated error management.

Additionally, it is important to note that our grid search explored only three discrete values for each quantum parameter R, λ, and E_{J0}, forming a 3×3 grid that required approximately 48 h of simulation time due to the considerable computational complexity involved. These values were chosen based on realistic ranges derived from previous studies and current operating conditions of SQUID–Transmon devices. Although this approach allowed us to clearly identify trends–most notably, the convergence toward an optimum at $R = 0.95$, we acknowledge that it constitutes a relatively coarse sampling of the parameter space. It is possible that optimal settings might lie at intermediate values or beyond the ranges we explored, and that the interactions among the parameters could exhibit nonlinear characteristics not fully captured by our current grid. Future work could use a finer grid or more advanced optimization techniques, such as Bayesian optimization, to more comprehensively map the parameter space. This extended exploration could help to capture potential nonlinear interactions and marginal improvements, thereby further refining the performance and robustness of the hybrid quantum–classical framework as additional computational resources and next-generation quantum hardware become available.

Furthermore, the current approach for encoding classical features into quantum states–using RY rotations and subsequent extraction of expectation values–is based on a well-established method in hybrid architectures. We recognize that this approach, while robust and straightforward, might lead to some loss of information due to its projection-based nature. Future investigations will also examine alternative encoding schemes, such as amplitude encoding or quantum kernel methods, which may preserve more information at the cost of increased circuit complexity, a trade-off that could become more favorable as quantum hardware matures.

Limitations

- *Task definition* Our primary labels are patient identity (one class per patient), used as a surrogate classification objective because lesion-level or outcome labels are not uniformly available. Window-level splits therefore do not constitute a patient-held-out clinical generalization test.
- *Evaluation protocol* We reported single-split and k-fold window-level metrics as requested. A stricter patient-held-out protocol with shared labels across patients remains future work.

- *Quantum-feature utilization and interpretability* We provide an explicit decomposition of the fusion classifier into classical and quantum weight blocks $\mathbf{W}_c$ and $\mathbf{W}_q$, which allows auditing whether the subsequent layer places non-trivial weight on the quantum measurement features. This transparency check is post-hoc and does not replace a dedicated removal study of the quantum pathway.
- *Hardware-imperfection modeling* The quantum layer is a parameterized, hardware-imperfection-aware surrogate: deviations in R, λ, and E_{J0} are treated as quasi-static calibration biases/drift that perturb the effective potential and drive a deterministic angle-rescaling factor. The surrogate does not include explicit T_1/T_2, gate-dependent Kraus maps, or crosstalk models; therefore it cannot reproduce full open-system noise dynamics. A hardware implementation would require estimating (R, λ, E_{J0}) from calibration data and maintaining the corresponding angle-rescaling lookup table, which is not validated experimentally in this study.
- *Connectivity* We assume nearest-neighbor entanglement without all-to-all coupling; results may change under different topologies.
- *Parameter coverage* The grid over R, λ, E_{J0} is coarse. More granular sweeps and Bayesian optimization may find improved operating points.
- *Scope and claims* We refrain from clinical claims; results should be viewed as a proof-of-concept for hardware-aware hybrid modeling.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11128-026-05144-x.

Acknowledgements JOM wants to recognize that this research has been partially supported by the Ministerio de Ciencia e Innovación of Spain (Grant Ref. PID2022-137748OB-C31 funded by MCIN/AEI/10.13039/501100011033) and "ERDF A way of making Europe".

Author Contributions JVD created the software architecture, collected the data, performed the experiments and wrote the paper. JCL supervised the presentation and the writing of the paper from a quantum hardware perspective. AGM supervised the presentation and the writing of the paper from a machine learning point of view. JOM supervised the presentation and the writing of the paper from a mathematical point of view.

Funding Open Access funding enabled and organized by Projekt DEAL. JOM wants to recognize that this research has been partially supported by the Ministerio de Ciencia e Innovación of Spain (Grant Ref. PID2022-137748OB-C31 funded by MCIN/AEI/10.13039/501100011033) and "ERDF A way of making Europe". JVD further acknowledges funding from the Hochschule Heilbronn and the Dieter Schwarz Stiftung through the program "Innovative Sonderprojekte im Bereich Forschung und Bildung 2025" to the **HyQCA Project**, which supports strategic, high-excellence research initiatives at the interface of digital technologies, AI, and quantum computing science.

Data Availability The datasets analyzed during the current study are available from the National Cancer Institute: Cancer Imaging Program [57].

Declarations

Accession codes To include, in this order: **Accession codes** (where applicable).

Conflict of interest The corresponding author is responsible for submitting a Conflict of interest statement on behalf of all authors of the paper. This statement must be included in the submitted article file.

References

1. Singh, P., Dasgupta, R., Singh, A., Pandey, H., Hassija, V., Chamola, V., Sikdar, B.: A survey on available tools and technologies enabling quantum computing. IEEE Access **12**, 57974–57991 (2024)
2. AbuGhanem, M., Eleuch, H.: Nisq computers: a path to quantum supremacy. IEEE Access **12**, 102941–102961 (2024)
3. Mandal, A.K., Chakraborty, B.: Quantum computing and quantum-inspired techniques for feature subset selection: a review. Knowl. Inf. Syst. **67**(3), 2019–2061 (2025)
4. Preskill, J.: Quantum computing in the Nisq era and beyond. Bull. Am. Phys. Soc. **64**, 9 (2019)
5. Anuradha, R., Vandana, C., Singh, S.V., Singh, N., Emad, R., Nitin, V.: Integrating quantum computing for enhanced image reconstruction in medical diagnostics. In: 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE), pp. 1294–1298. IEEE (2024)
6. Balasubramani, S., Renjith, P., Kavisankar, L., Rajavel, R., Malarvel, M., Shankar, A.: A quantum-enhanced artificial neural network model for efficient medical image compression. IEEE Access **13**, 31809–31828 (2025)
7. Gujju, Y., Matsuo, A., Raymond, R.: Quantum machine learning on near-term quantum devices: current state of supervised and unsupervised techniques for real-world applications. Phys. Rev. Appl. **21**(6), 067001 (2024)
8. Halder, S., Dey, A., Shrikhande, C., Maitra, R.: Machine learning assisted construction of a shallow depth dynamic ansatz for noisy quantum hardware. Chem. Sci. **15**(9), 3279–3289 (2024)
9. Rahimi, M., Asadi, F.: Oncological applications of quantum machine learning. Technol. Cancer Res. Treat. **22**, 15330338231215214 (2023)
10. Swanson, K., Wu, E., Zhang, A., Alizadeh, A.A., Zou, J.: From patterns to patients: advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. Cell **186**(8), 1772–1791 (2023)
11. Tran, K.A., Kondrashova, O., Bradley, A., Williams, E.D., Pearson, J.V., Waddell, N.: Deep learning in cancer diagnosis, prognosis and treatment selection. Genome Med. **13**, 1–17 (2021)
12. Harrison, K., Pullen, H., Welsh, C., Oktay, O., Alvarez-Valle, J., Jena, R.: Machine learning for auto-segmentation in radiotherapy planning. Clin. Oncol. **34**(2), 74–88 (2022)
13. Samarasinghe, G., Jameson, M., Vinod, S., Field, M., Dowling, J., Sowmya, A., Holloway, L.: Deep learning for segmentation in radiation therapy planning: a review. J. Med. Imaging Radiat. Oncol. **65**(5), 578–595 (2021)
14. Hayashi, K., Ono, Y., Takamatsu, M., Oba, A., Ito, H., Sato, T., Inoue, Y., Saiura, A., Takahashi, Y.: Prediction of recurrence pattern of pancreatic cancer post-pancreatic surgery using histology-based supervised machine learning algorithms: a single-center retrospective study. Ann. Surg. Oncol. **29**(7), 4624–4634 (2022)
15. Chen, W., Zhou, B., Jeon, C.Y., Xie, F., Lin, Y.-C., Butler, R.K., Zhou, Y., Luong, T.Q., Lustigova, E., Pisegna, J.R., et al.: Machine learning versus regression for prediction of sporadic pancreatic cancer. Pancreatology **23**(4), 396–402 (2023)
16. Chen, P.-T., Wu, T., Wang, P., Chang, D., Liu, K.-L., Wu, M.-S., Roth, H.R., Lee, P.-C., Liao, W.-C., Wang, W.: Pancreatic cancer detection on CT scans with deep learning: a nationwide population-based study. Radiology **306**(1), 172–182 (2023)
17. Tan, M., Le, Q.: Efficientnet: rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning, pp. 6105–6114. PMLR (2019)
18. Zhao, H., Wu, Y., Lu, Y.: A study on skin lesions classification based on improved efficientn etb0 network. In: 2024 International Conference on New Trends in Computational Intelligence (NTCI), pp. 463–467. IEEE (2024)

19. Guo, H., Zhang, J., Li, Y., Pan, X., Sun, C.: Advanced pathological subtype classification of thyroid cancer using efficientnetb0. Diagn. Pathol. **20**(1), 28 (2025)
20. Kumar, A., Nelson, L., Arumugam, D.: Deep learning-based classification of brain tumours on MRI images using efficientnetb0. In: 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS), pp. 219–225. IEEE (2024)
21. Asif, S., Wenhui, Y., ur-Rehman, S., ul-ain, Q., Amjad, K., Yueyang, Y., Jinhai, S., Awais, M.: Advancements and prospects of machine learning in medical diagnostics: unveiling the future of diagnostic precision. Arch. Comput. Methods Eng. **32**, 1–31 (2024)
22. Saeed, M.K., Al Mazroa, A., Alghamdi, B.M., Alallah, F.S., Alshareef, A., Mahmud, A.: Predictive analytics of complex healthcare systems using deep learning based disease diagnosis model. Sci. Rep. **14**(1), 27497 (2024)
23. Wei, L., Liu, H., Xu, J., Shi, L., Shan, Z., Zhao, B., Gao, Y.: Quantum machine learning in medical image analysis: a survey. Neurocomputing **525**, 42–53 (2023)
24. Maheshwari, D., Garcia-Zapirain, B., Sierra-Sosa, D.: Quantum machine learning applications in the biomedical domain: a systematic review. IEEE Access **10**, 80463–80484 (2022)
25. Zhou, Y., Zhang, P.: Noise-resilient quantum machine learning for stability assessment of power systems. IEEE Trans. Power Syst. **38**(1), 475–487 (2023). https://doi.org/10.1109/TPWRS.2022.3160384
26. Schuld, M., Petruccione, F.: Quantum computing. In: Machine Learning with Quantum Computers. Quantum Science and Technology. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-83098-4_3
27. Trochun, Y., Stirenko, S., Rokovyi, O., Alienin, O., Pavlov, E., Gordienko, Y.: Hybrid classic-quantum neural networks for image classification. In: 2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), vol. 2, pp. 968–972. IEEE (2021)
28. Hong, J., Reyngold, M., Crane, C., Cuaron, J., Hajj, C., Mann, J., Zinovoy, M., Yorke, E., LoCastro, E., Apte, A.P., Mageras, G.: Breath-hold CT and cone-beam CT images with expert manual organ-at-risk segmentations from radiation treatments of locally advanced pancreatic cancer. Data set. Cancer Imaging Arch. (2021). https://doi.org/10.7937/TCIA.ESHQ-4D90
29. Placido, D., Yuan, B., Hjaltelin, J.X., Zheng, C., Haue, A.D., Chmura, P.J., Yuan, C., Kim, J., Umeton, R., Antell, G., et al.: A deep learning algorithm to predict risk of pancreatic cancer from disease trajectories. Nat. Med. **29**(5), 1113–1122 (2023)
30. Cao, K., Xia, Y., Yao, J., Han, X., Lambert, L., Zhang, T., Tang, W., Jin, G., Jiang, H., Fang, X., et al.: Large-scale pancreatic cancer detection via non-contrast CT and deep learning. Nat. Med. **29**(12), 3033–3043 (2023)
31. Xiang, Q., Li, D., Hu, Z., Yuan, Y., Sun, Y., Zhu, Y., Fu, Y., Jiang, Y., Hua, X.: Quantum classical hybrid convolutional neural networks for breast cancer diagnosis. Sci. Rep. **14**(1), 24699 (2024)
32. Dixit, A., Mani, A., Gorbachev, S.: Hybrid framework for medical image classification: integrating quantum de and deep learning. In: 2024 IEEE International Conference on Intelligent Signal Processing and Effective Communication Technologies (INSPECT), pp. 1–6. IEEE (2024)
33. Huang, H.-Y., Broughton, M., Mohseni, M., Babbush, R., Boixo, S., Neven, H., McClean, J.R.: Power of data in quantum machine learning. Nat. Commun. **12**(1), 2631 (2021)
34. Cerezo, M., Arrasmith, A., Babbush, R., Benjamin, S.C., Endo, S., Fujii, K., McClean, J.R., Mitarai, K., Yuan, X., Cincio, L., et al.: Variational quantum algorithms. Nat. Rev. Phys. **3**(9), 625–644 (2021)
35. Koch, J., Yu, T.M., Gambetta, J., Houck, A.A., Schuster, D.I., Majer, J., Blais, A., Devoret, M.H., Girvin, S.M., Schoelkopf, R.J.: Charge-insensitive qubit design derived from the cooper pair box. Phys. Rev. A At. Mol., Opt. Phys. **76**(4), 042319 (2007)
36. Devoret, M.H., Schoelkopf, R.J.: Superconducting circuits for quantum information: an outlook. Science **339**(6124), 1169–1174 (2013)
37. Kjaergaard, M., Schwartz, M.E., Braumüller, J., Krantz, P., Wang, J.I.-J., Gustavsson, S., Oliver, W.D.: Superconducting qubits: current state of play. Annu. Rev. Condens. Matter Phys. **11**(1), 369–395 (2020)
38. Temme, K., Bravyi, S., Gambetta, J.M.: Error mitigation for short-depth quantum circuits. Phys. Rev. Lett. **119**(18), 180509 (2017)
39. Botelho, L., Glos, A., Kundu, A., Miszczak, J.A., Salehi, Ö., Zimborás, Z.: Error mitigation for variational quantum algorithms through mid-circuit measurements. Phys. Rev. A **105**(2), 022441 (2022)
40. Dasgupta, S., Humble, T.S.: Improving probabilistic error cancellation in the presence of non-stationary noise. IEEE Trans. Quantum Eng. **5**, 1–14 (2024)

41. Endo, S., Benjamin, S.C., Li, Y.: Practical quantum error mitigation for near-future applications. Phys. Rev. X **8**(3), 031027 (2018)
42. Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P.J., Aspuru-Guzik, A., O'brien, J.L.: A variational eigenvalue solver on a photonic quantum processor. Nat. Commun. **5**(1), 4213 (2014)
43. Farhi, E., Goldstone, J., Gutmann, S.: A quantum approximate optimization algorithm (2014). arXiv preprint arXiv:1411.4028
44. Havlíček, V., Córcoles, A.D., Temme, K., Harrow, A.W., Kandala, A., Chow, J.M., Gambetta, J.M.: Supervised learning with quantum-enhanced feature spaces. Nature **567**(7747), 209–212 (2019)
45. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18, pp. 234–241. Springer (2015)
46. Ibrahim, H.K., Rokbani, N., Wali, A., Ouahada, K., Chabchoub, H., Alimi, A.M.: A medical image classification model based on quantum-inspired genetic algorithm. Eng. Technol. Appl. Sci. Res. **14**(5), 16692–16700 (2024)
47. Wu, D., Rossi, R., Vicentini, F., Astrakhantsev, N., Becca, F., Cao, X., Carrasquilla, J., Ferrari, F., Georges, A., Hibat-Allah, M., et al.: Variational benchmarks for quantum many-body problems. Science **386**(6719), 296–301 (2024)
48. Siddiqi, I.: Engineering high-coherence superconducting qubits. Nat. Rev. Mater. **6**(10), 875–891 (2021). https://doi.org/10.1038/s41578-021-00370-4
49. Majumder, S., Bera, T., Suresh, R., Singh, V.: A fast tunable 3D-transmon architecture for superconducting qubit-based hybrid devices. J. Low Temp. Phys. **207**(3), 210–219 (2022)
50. Bergholm, V., Izaac, J., Schuld, M., Gogolin, C., Ahmed, S., Ajith, V., Alam, M.S., Alonso-Linaje, G., AkashNarayanan, B., Asadi, A., et al.: Pennylane: automatic differentiation of hybrid quantum-classical computations (2018). arXiv preprint arXiv:1811.04968
51. Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. Software available from tensorflow.org (2015). http://tensorflow.org/
52. Tancik, M., Srinivasan, P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J., Ng, R.: Fourier features let networks learn high frequency functions in low dimensional domains. Adv. Neural. Inf. Process. Syst. **33**, 7537–7547 (2020)
53. ...Harris, C.R., Millman, K.J., Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M.H., Brett, M., Haldane, A., Río, J.F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E.: Array programming with Numpy. Nature **585**(7825), 357–362 (2020). https://doi.org/10.1038/s41586-020-2649-2
54. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors: SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17, 261–272 (2020). https://doi.org/10.1038/s41592-019-0686-2
55. Bradski, G.: The OpenCV Library. Dr. Dobb's Journal of Software Tools (2000)
56. McKinney, W., et al.: Data structures for statistical computing in python. In: Proceedings of the 9th Python in Science Conference, vol. 445, pp. 51–56. Austin, TX (2010)
57. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., Prior, F.: The cancer imaging archive (TCIA): maintaining and operating a public information repository. J. Digit. Imaging **26**(6), 1045–1057 (2013). https://doi.org/10.1007/s10278-013-9622-7