# Quantum Fusion Learning for 6G Networks with Kolmogorov-Arnold Networks

Usama Inam Paracha, Uman Khalid, Syed Muhammad Abuzar Rizvi, Trung Q. Duong, *Fellow, IEEE*, Moe Z. Win, *Fellow, IEEE*, and Hyundong Shin, *Fellow, IEEE*

***Abstract*—Driven by the vision of intelligent and adaptive connectivity in the sixth-generation (6G) networks, future wireless systems are expected to support advanced applications such as cognitive resource allocation, semantic communication (SC), and low-latency integrated sensing and communication (ISAC) or digital twins. However, realizing these capabilities requires learning frameworks capable of efficiently modeling complex, high-dimensional, and dynamic communication environments. To address this challenge, we propose a Kolmogorov-Arnold network (KAN)-enhanced quantum fusion learning (QFL) model designed to advance intelligent wireless communication systems. Specifically, we introduce a hybrid quantum-classical (HQC) architecture that integrates parametrized quantum circuits with deep learning (DL) models based on residual networks (ResNets) and shifted window transformers (SwinTs), further augmented with KAN layers. This HQC framework potentially addresses the limitations of conventional DL (ResNets/SwinTs) by accurately modeling nonlinear relationships, capturing high-order dependencies, and flexibly adapting to dynamic wireless environments. We evaluate the proposed QFL framework on three representative 6G tasks: 1) satellite remote-sensing scene classification (RSSCN7 dataset), 2) ISAC beamforming prediction (DeepSense 6G dataset)—which supports cognition-level resource allocation and low-latency ISAC—and 3) generative SC classification (GenSC-6G dataset)—relevant to semantic environment modeling and digital twin updates. Numerical results demonstrate that the designed QFL architecture delivers competitive performance, underscoring its potential for enabling intelligent and secure 6G systems.**

***Index Terms*—Beamforming, ISAC, KAN learning, QFL models, remote-sensing classification, SC classification.**

## I. Introduction

The fundamental research described in this paper was supported, in part, by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) under RS-2025-00556064 and by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2026-RS-2021-II212046) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation), in part by the Canada Excellence Research Chair (CERC) Program CERC-2022-00109, by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant Program RGPIN-2025-04941, and by the NSERC CREATE program (Grant number 596205-2025), and in part by the National Science Foundation under Grant CCF-1956211. *(Corresponding author: Hyundong Shin.)*

U. I. Paracha, U. Khalid, S. M. A. Rizvi, and H. Shin are with the Department of Electronics and Information Convergence Engineering, Kyung Hee University, 1732 Deogyeong-daero, Giheung-gu, Yongin-si, Gyeonggi-do 17104, Republic of Korea (e-mail: hshin@khu.ac.kr).

T. Q. Duong is with the Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL A1C 5S7, Canada, and is also with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, U.K. (e-mail: tduong@mun.ca).

M. Z. Win is with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139 USA (e-mail: moewin@mit.edu).

QUANTUM fusion learning (QFL) is emerging as a core enabler of the sixth-generation (6G) wireless networks, which are envisioned as foundational infrastructures for intelligent environments, including autonomous vehicle ecosystems, digital twin networks, and real-time, privacy-sensitive environmental monitoring systems [1], [2], [3]. The 6G systems facilitate transformative applications such as high-resolution remote sensing, beamforming prediction, vehicular localization, medical robotics, non-terrestrial networks (NTNs), and semantic communication (SC) [4], [5], [6], [7]. Notably, satellite remote-sensing scene classification, beamforming prediction for vehicle-to-infrastructure (V2I) communications, and SC for 6G seamless immersive experience have emerged as crucial tasks for intelligent 6G wireless systems, demanding exceptional model accuracy, perpetual-data processing, adaptability to diverse geographic and channel conditions, and robustness against hardware and signal imperfections [8]. However, achieving such an advanced intelligence within resource-constrained computing platforms and highly dynamic wireless environments remains significantly challenging, thereby driving exploration of more efficient and adaptive learning paradigms [9].

In recent years, deep learning (DL) techniques, including convolutional neural networks (CNNs) such as residual networks (ResNets) and attention-based models like transformers, have shown strong potential in addressing core 6G tasks [10]. ResNets are particularly effective at extracting local features from visual data robustly, while transformer architectures leverage global self-attention mechanisms to capture spatial dependencies crucial for integrated sensing and communication (ISAC) applications such as beamforming and remote sensing [11]. Despite their advantages, these architectures typically rely on static and manually designed activation functions, e.g., rectified linear unit (ReLU) and Gaussian error linear unit (GELU) functions, which can limit their adaptability to learn complex and task-specific nonlinear patterns. Additionally, conventional machine learning (ML) models face persistent challenges related to generalization, interpretability, privacy, and scalability, especially when deployed in real-time, resource-constrained 6G environments [12]. This limitation underscores the growing need to augment classical DL architectures with more adaptive and expressive learning mechanisms capable of meeting the high-dimensional, dynamic, and heterogeneous demands of 6G systems.

To address these limitations, an emerging line of research explores a new avenue using Kolmogorov-Arnold networks (KANs), which substitute conventional fixed activation functions with learnable, data-driven functional mappings [13], [14]. This architectural innovation enables each network layer

TABLE I
EXPANSIONS OF ACRONYMS

| Acronym | Expansion | Acronym | Expansion |
|---|---|---|---|
| 6G | Sixth-Generation | AI | Artificial Intelligence |
| BASNet | Boundary-Aware Salient Object Detection Network | BSRBF | B-Spline Radial-Basis Function |
| CNN | Convolutional Neural Network | CNOT | Controlled-NOT |
| DL | Deep Learning | DNN | Deep Neural Network |
| DoA | Direction-of-Arrival | DRL | Deep Reinforcement Learning |
| FC | Fully Connected | GAN | Generative Adversarial Network |
| GELU | Gaussian Error Linear Unit | GPU | Graphics Processing Unit |
| GR-KAN | Group-Rational Kolmogorov-Arnold Network | HHL | Harrow–Hassidim-Lloyd |
| HQC | Hybrid Quantum-Classical | IoT | Internet of Things |
| ISAC | Integrated Sensing and Communication | KAN | Kolmogorov-Arnold Network |
| LLM | Large Language Model | LSTM | Long Short-Term Memory |
| ML | Machine Learning | MLP | Multi-Layer Perceptron |
| MSTUnet | Masked Swin Transformer U-Net | NISQ | Noisy Intermediate-Scale Quantum |
| NTN | Non-Terrestrial Network | ODE | Ordinary Differential Equations |
| PCA | Principal Component Analysis | PDE | Partial Differential Equation |
| PQC | Parametrized Quantum Circuit | QAOA | Quantum Approximate Optimization Algorithm |
| QCNN | Quantum Convolutional Neural Network | QFL | Quantum Fusion Learning |
| QML | Quantum Machine Learning | QNN | Quantum Neural Network |
| QPU | Quantum Processing Unit | QRL | Quantum Reinforcement Learning |
| ReLU | Rectified Linear Unit | ResNet | Residual Network |
| RL | Reinforcement Learning | RSSCN | Remote-Sensing Scene Classification |
| SC | Semantic Communication | SGRS | Satellite-Ground Remote Sensing |
| SineKAN | Sinusoidal Kolmogorov-Arnold Network | SVM | Support Vector Machine |
| SwinT | Shifted Window Transformer | UAV | Unmanned Aerial Vehicle |
| URLLC | Ultra-Reliable Low-Latency Communication | V2I | Vehicle-to-Infrastructure |
| V2X | Vehicle-to-Everything | ViT | Vision Transformer |
| VQA | Variational Quantum Algorithm | VQE | Variational Quantum Eigensolver |

to dynamically adapt task-specific nonlinear transformations in learning, thereby significantly enhancing representational flexibility, generalization ability, and expressive power of both residual and transformer-based architectures [15], [16]. However, despite such benefits, KAN-enhanced models tend to be often computationally intensive, and their inherently non-convex optimization landscapes pose challenges related to parallel computation, scalability, and privacy-preserving learning, especially in real-time, resource-constrained 6G environments [17], [18]. Therefore, addressing these constraints demands further architectural innovation, particularly quantum-enhanced learning for emerging multi-functional low-altitude wireless networks with ISAC, where highly dynamic aerial-ground environments require future wireless infrastructures to jointly support sensing, communication, control, computation, energy transfer, and intelligent decision-making [19].

The transformative potential of quantum computing in this context is rooted in two pivotal breakthroughs that demonstrated its capacity to solve classically intractable problems. Shor's algorithm revealed that quantum computers can factor large integers exponentially faster than classical algorithms, highlighting the computational power of quantum parallelism. In parallel, Grover's algorithm provided a quadratic speedup for unstructured search tasks, establishing early evidence of quantum acceleration in generic computational settings [20], [21]. Building upon these foundations, the Harrow–Hassidim–Lloyd (HHL) algorithm laid the groundwork for quantum ML (QML) by enabling exponentially faster matrix inversion under conditions of data sparsity and efficient state preparation, thereby inspiring the rapid growth of QML research aimed at overcoming the scalability and privacy limitations of classical DL [22].

Recent advancements in QML offer promising solutions to overcome these scalability and privacy limitations inherent to classical DL approaches [23], [12], [24], [25], [26]. Quantum computing provides notable advantages, including inherent parallelism and information-theoretic security, making it well suited for computation-intensive and privacy-sensitive applications envisioned in 6G networks [27], [28]. In this context, QFL frameworks, specifically hybrid quantum-classical (HQC) architectures that integrate parametrized quantum circuits (PQCs) with DL, large language models (LLMs), or large artificial intelligence (AI) models, have emerged as a compelling paradigm for harnessing quantum advantages within existing AI workflows [29]. In this work, we propose a novel KAN-enhanced QFL framework that integrates KAN-augmented residual and transformer models with PQC modules. This integrated framework effectively captures complex high-order feature correlations and nonlocal dependencies, thereby addressing critical challenges in privacy preservation, computational scalability, and learning efficiency for intelligent and secure 6G systems.

In the proposed framework, the ResNet or shifted window transformer (SwinT) first extracts discriminative classical representations using group-rational Kolmogorov-Arnold network (GR-KAN) activation functions, while the PQC further processes the learned feature embedding before the final classifier. Therefore, the role of the quantum module is not only to provide immediate accuracy improvement, but also to explore whether quantum-inspired feature interactions can complement KAN-based classical learning for future 6G intelligence [30], [31]. Although the performance gain may vary across datasets and backbone architectures, the quantum fusion stage provides an additional representation space for modeling complex feature dependencies, which is particularly relevant for emerging HQC learning systems. The main contributions of this paper are outlined as follows.

- *Survey of QML and QFL*: We provide a comprehensive literature review tracing the QML evolution, outlining its foundational principles, inherent limitations, and the emerging motivation for developing QFL frameworks. We highlight the application of QFL models across diverse 6G communication and networking scenarios, emphasizing their emerging relevance in distributed and intelligent wireless systems. In parallel, we review recent advances in KAN and their potential for efficient and interpretable learning paradigms.
- *KAN-Enhanced QFL Models*: We propose a novel KAN-enhanced QFL framework, seamlessly integrating KANs into classical AI models—namely, ResNets and SwinTs—further augmented by quantum processing unit (QPU) modules (i.e., PQCs) to form a robust HQC architecture tailored for 6G environments. KAN layers enable adaptive, data-driven nonlinear transformations that improve generalization capabilities, while PQC modules capture high-order semantic correlations crucial for remote sensing and ISAC beamforming.
- *Experimental Validation on 6G Benchmarks*: We evaluate the designed QFL architecture on three representative 6G benchmark datasets: 1) RSSCN7 [3] for satellite remote-sensing scene classification, 2) DeepSense 6G [32] for ISAC beamforming prediction, and 3) GenSC-6G [33] for generative SC classification. The experimental results show the QFL framework's promising performance, suggesting its potential to address challenges in learning efficiency, privacy preservation, and real-time adaptability in 6G networks.

The rest of the paper is organized as follows. Section II provides a detailed overview of QML and reviews the existing QFL models for 6G systems, with a particular focus on both classical and quantum AI approaches, as well as recent advances in KANs. Section III presents the proposed KAN-enhanced QFL architecture, detailing its classical, quantum, and KAN components. Section IV provides experimental results on three representative 6G tasks. Section V concludes the paper. A list of acronyms used in this paper is summarized in Table I.

## II. Related Work

This section presents an overview of QML and its evolution toward QFL. We then explore recent QFL developments within the context of 6G applications. Furthermore, we provide an in-depth review of KANs.

### A. QML to QFL Overview

*1) Classical QML Methods:* QML can be broadly categorized into two complementary approaches: non-neural-network-based and neural-network-based methods. The former focuses on developing quantum algorithms for fundamental linear algebra and statistical operations, such as HHL for solving linear systems, performing principal component analysis (PCA), and implementing quantum Boltzmann machines, often revealing potential quantum speedups when data access is structured and efficiently organized [22], [34], [35]. In parallel, large-margin and distance-based classification methods have also been proposed, including quantum support vector machines (SVMs), quantum nearest-neighbor classifiers, and early clustering formulations [36], [37], [38]. For instance, large-margin methods such as classical SVMs aim to find decision boundaries that maximize the separation between classes and are often implemented using kernel functions that map data into high-dimensional feature spaces. In this context, quantum SVMs extend this classical kernel framework by employing quantum states to define and evaluate kernels efficiently within quantum feature spaces. These algorithmic directions have been further discussed in recent surveys that also clarify assumptions on data loading and query access [39], [40], [41].

*2) Variational and Neural Quantum Learning Methods:* The latter category introduces neural network perspectives for quantum systems, giving rise to PQC learning within variational frameworks [42]. As quantum hardware matures, such variational quantum algorithms (VQAs) have become a primary near-term strategy. Notable examples include the variational quantum eigensolver (VQE) for determining the lowest eigenvalue or ground-state energy of a Hamiltonian in quantum chemistry, and the quantum approximate optimization algorithm (QAOA) for solving discrete combinatorial optimization problems through alternating layers of cost and mixing unitaries. These algorithms employ PQCs to learn optimal parameters on noisy intermediate-scale quantum (NISQ) devices, thereby establishing the methodological backbone for optimization on current hardware [43], [44]. These variational approaches have been extensively analyzed in terms of circuit design, optimization techniques, and hardware integration [42], [45], [46]. Within this landscape, quantum neural networks (QNNs) have been explored as quantum analogues of classical neural networks for discriminative tasks such as classification and regression [47], [48]. These models leverage PQCs to learn decision boundaries or latent representations directly from quantum-embedded classical data, often trained through HQC optimization loops. Early studies have demonstrated the feasibility of encoding classical data into quantum states and training QNNs to perform supervised classification, extending architectures such as perceptrons and convolutional networks into the quantum domain [49].

However, both the variational and kernel-based approaches in QML face challenges, including optimization instability, noise sensitivity, and data encoding bottlenecks [39], [50]. Specifically, converting classical data into quantum states can be computationally expensive, often incurring exponential overheads [51], [52]. Furthermore, the expressivity-trainability trade-off arises as highly expressive quantum ansätze approach random-state behavior that leads to flatter cost landscapes and diminished gradient magnitudes, while more structured circuits preserve stronger gradient information and better trainability [53], [54]. Among these limitations, a central issue within variational approaches arises from gradient concentration, leading to barren plateaus where trainability deteriorates with increasing circuit depth, system size, or unsuitable cost functions [55]. Additionally, hardware noise, including gate imperfections, decoherence, and measurement errors, can further suppress gradients and hinder convergence in optimization on NISQ devices [56]. To mitigate these issues, structured quantum CNNs (QCNNs) have shown promise in preserving gradient signal and avoiding barren plateaus [57]. Subsequent work has focused on refining ansatz design [54], [58], [59], [60], [61], developing hardware-efficient ansätze [62], and implementing geometry-aware optimization and analytical gradient methods [63], [64]. Furthermore, error mitigation techniques, such as extrapolation and probabilistic error cancellation, have shown partial success at the cost of additional resource overhead [65], [66]. From the kernel perspective, concentration effects and data access limitations constrain potential quantum advantages, while recent studies emphasize the importance of informative datasets and controlled experimental demonstrations [38], [52], [67]. These insights motivate hybrid architectures in which classical networks handle most of representation learning, while quantum modules act as compact trainable components contributing nonclassical transformations under limited resources, paving the way toward QFL.

*3) QFL Methods:* The QFL framework defines hybrid pipelines that couple classical deep neural networks (DNNs)—such as ResNets and SwinTs—with PQC or QNN components, aiming to leverage the strengths of both modalities. Typically, the classical modules extract robust local and global features, while quantum layers manipulate data in high-dimensional Hilbert spaces using learned quantum feature maps or variational unitaries, potentially yielding improved sample efficiency and expressivity for specific tasks. Concrete QFL strategies include 1) pre-quantum feature extraction, where classical models such as ResNets or vision transformers (ViTs) generate embeddings for PQC-based classifiers or regressors, 2) interleaved hybrid designs, where quantum blocks are embedded within classical layers, and 3) post-QFL designs, where quantum layers act as representation learners whose outputs are processed by classical decision heads [31], [68], [69]. These patterns are particularly attractive when classical encoders can reduce dimensionality and enforce inductive biases, allowing quantum modules to introduce non-classical decision boundaries with minimal parameter overhead.

A central conceptual foundation is the notion of a quantum feature space, where PQCs serve as learnable feature maps that encode classical data into high-dimensional Hilbert spaces. Within this framework, quantum-enhanced feature spaces can be combined with classical layers to form hybrid models that support end-to-end trainability. This contrasts with quantum-enhanced kernels, which evaluate fixed inner products between non-trainable embeddings [70], [30], [71]. Recent theoretical work has clarified the conditions for quantum advantage and provided design guidelines for integrating quantum layers effectively [72], [73]. Several recurring architectural patterns have been observed across the existing QFL models. Data re-uploading injects classical features repeatedly across circuit layers for stronger expressivity [74]. Bosonic implementations broaden hardware platforms available for applying the same principle [75]. QCNN models introduce hierarchical locality-aware processing that aligns with vision and physics tasks [49], [76]. Resource-efficient QCNN variants further reduce quantum circuit depth and parameter count, which is crucial when the quantum layer is only one component in larger HQC stacks [77]. Quantum transformers extend attention mechanisms into the quantum domain and can be fused with classical self-attention in hybrid pipelines [78].

### *B. QML and QFL for 6G*

Classical AI models, although effective in many scenarios, face scalability challenges when handling high-dimensional inputs, real-time decision-making demands, and federated deployment constraints inherent to 6G networks [12]. To address these challenges, QML has emerged as a promising solution, leveraging quantum phenomena such as superposition and entanglement to potentially enhance computational efficiency [28]. These quantum properties facilitate accelerated model training, reduced inference latency, and enhanced learning performance, particularly in dynamic resource-constrained wireless environments [30], [79]. In [12], the quantum-enhanced learning framework was introduced utilizing VQAs and QAOA to address ultra-reliable low-latency communication (URLLC) bottlenecks. This approach demonstrated faster convergence and more reliable decision-making under stringent end-to-end latency constraints, underscoring the QFL efficacy in supporting real-time communication tasks. Similarly, the QNN architecture was proposed in [80] for non-orthogonal multiple access user-grouping. Further extending this work, the federated teleportation-based QML framework was developed in [9] to preserve quantum-state fidelity during federated learning processes by leveraging quantum teleportation, thereby ensuring data privacy.

The versatility of QML has also been demonstrated in beam management tasks. Recently, hybrid quantum-transformer networks were developed for beam prediction in V2I communications [1]. By fusing multimodal sensor data—such as GPS, LiDAR, and camera feeds—with quantum-enhanced vision-based transformers, these QML models achieved state-of-the-art performance on real-world datasets (DeepSense 6G). The reported results highlight the potential of quantum models to enhance robustness and generalization of 6G ISAC beamforming solutions in dynamic vehicular environments. Additionally, the quantum-enhanced deep reinforcement learning (DRL) model was proposed in [81] for joint direction-of-arrival estimation and task offloading. By integrating quantum policy optimization into the DRL framework, this QML model achieved considerable improvements in convergence rates and estimation accuracy, demonstrating the promise of QML for multi-objective 6G ISAC applications where sensing accuracy and computational efficiency are both critical. Finally, the QFL framework was introduced in [3], fusing quantum computing with classical DL to enhance both data processing efficiency and communication security in integrated satellite-ground remote sensing (SGRS) systems. Although prior QML and QFL studies have demonstrated the potential of quantum models for optimization, classification, feature mapping, and 6G communication tasks [9], [17], [28], [39], [40], [41], most existing works focus on standalone quantum learning models, variational quantum circuits (VQCs), or hybrid frameworks where quantum modules are used mainly as auxiliary feature processors [48], [51], [52]. In contrast, the proposed framework explicitly combines KAN-enhanced deep feature extraction with quantum feature learning for 6G-oriented ISAC tasks. This design differs from prior works in that learnable KAN-based nonlinear mappings strengthen the classical backbone representation before quantum processing, while the quantum module further transforms the learned features into a quantum-enhanced representation space [16]. Therefore, the proposed approach bridges adaptive classical feature learning and quantum feature transformation in a unified pipeline, rather than treating them as separate or independently optimized components. An evolutionary timeline of QML toward QFL is illustrated in Fig 1.

### C. KAN Developments

Recent ML developments have stimulated growing interest in KAN-enhanced architectures, which replace conventional linear weights with learnable spline-based activation functions on network edges to improve expressivity and interpretability. The foundational KAN framework demonstrated competitive accuracy in regression and partial differential equation (PDE) solving tasks. Rather than altering the spline placement or core architecture, KAN 2.0 emphasizes KANs's role in scientific discovery by introducing key tools: MultKAN (KANs with multiplicative nodes), the kanpiler (a compiler converting symbolic expressions into KAN architectures), and a tree converter for modular and interpretable representation learning [82]. These tools enable KANs to synergize symbolic reasoning with neural modeling, thereby facilitating applications in equation discovery, modular learning, and interpretable AI. Building on these foundations, subsequent research has focused on improving efficiency and stability. For instance, simplified variants like ReLU-KAN [83] employ ReLU-based activations and matrix-friendly operations for significant speedups. By leveraging element-wise nonlinearities and streamlined linear transformations, ReLU-KAN significantly reduces runtime overhead, especially on graphics processing unit (GPU)-accelerated environments, enabling practical scalability for large-scale graph and grid data applications. P1-KAN [84] and B-spline radial-basis function (BSRBF)-KAN [85] introduce piecewise-linear and radial-basis expansions, respectively, to ensure robust performance in high-dimensional function approximation. The piecewise-linear expansions in P1-KAN provide fine-grained adaptability to local features, boosting approximation fidelity, while BSRBF-KAN offers smooth isotropic feature modeling and enhanced generalization in complex domains. Similarly, smooth activation formulations enhance structural reliability in demanding scientific and biomedical applications [86]. By incorporating continuously differentiable activations, this approach reduces gradient discontinuities and improves convergence stability in highly sensitive simulation and diagnostic pipelines, thus safeguarding against spurious artifacts and ensuring dependable outcomes.

Beyond spline-based functions, the KAN paradigm has been extended through specialized activation designs. Sinusoidal KAN (SineKAN) [87], for example, utilizes sinusoidal activations for more lightweight deployments. Other works have integrated Chebyshev polynomials to enhance nonlinear modeling capacity [88]. These architectural innovations are complemented by the early theoretical work [13], which provides a basis for understanding the approximation power of such networks. This line of research has also led to hybrid integrations, including KAN-ordinary differential equations (ODEs), which embed KAN operators into neural ODE solvers to uncover hidden physical dynamics [89]. Comprehensive benchmarking and large-scale evaluations have been crucial in positioning KANs relative to traditional multi-layer perceptrons (MLPs). Their performances have been compared in operator learning tasks, demonstrating that KANs often match or even surpass MLPs in approximating partial differential operator mappings with fewer parameters [90]. Notably, in learning non-local operators or integral transforms, KANs demonstrate improved sample complexity and robustness to noisy training data, especially when the underlying operator kernel exhibits smoothness. In low-data regimes, KANs with trainable activation expansions (e.g., B-splines or piecewise-linear bases) have shown superior adaptability compared to standard MLPs, capturing non-linear feature interactions with fewer training samples [91]. However, these experiments also revealed that when data is extremely scarce, their relatively high parameter count in KANs can lead to overfitting unless regularization (such as sparsity or smoothness constraints) is applied carefully. In federated environments, KANs have been evaluated in settings where data is distributed across clients with heterogeneous distributions, and results show that federated KAN (Fed-KANs) variants often outperform Fed-MLPs in convergence speed and generalization across clients, thanks to their ability to model client-specific nonlinearities [92]. However, these gains are sensitive to communication overhead and synchronization strategy. Optimal performance often requires balancing activation complexity with the frequency of model aggregation and local update steps.

In computer vision, early KAN adaptations achieved state-of-the-art results in medical imaging [93], [94]. Building on these successes (med-KANs), research has progressed beyond simple backbones to deeply integrate KAN principles with core DL structures. Examples include hybrid transformer-KAN models and architectural primitives such as ConvKAN, which replaces standard convolutions with KAN-based operators [16], [95]. To enable deeper networks, KAN-ResNet has been introduced for straightforward incorporation of KAN layers into classical DNNs such as ResNets [96]. The paradigm has been extended to non-Euclidean data with graph KANs (G-KANs) for graph-structured data, demonstrating a clear trend towards embedding

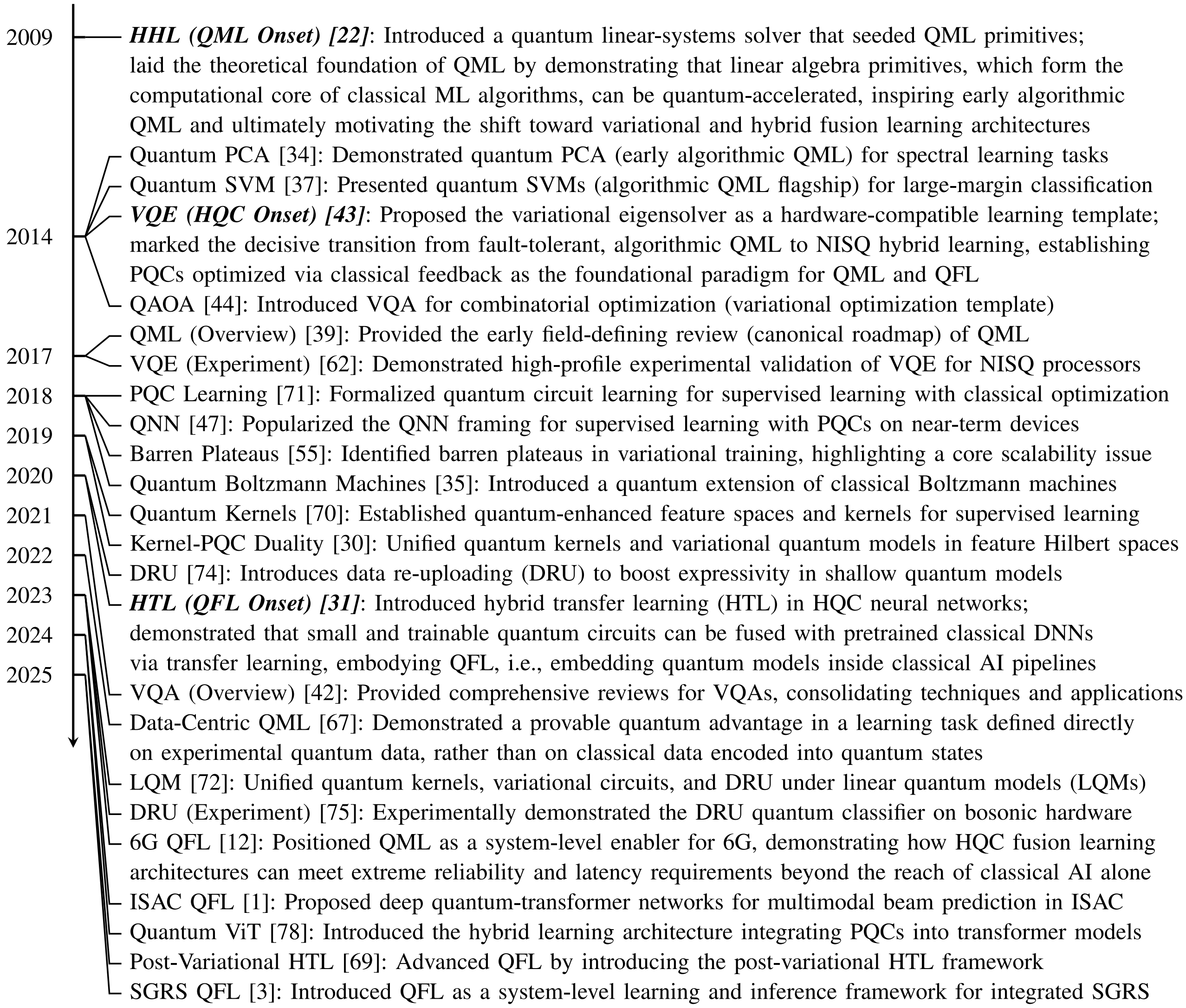


Fig. 1. A timeline illustrating the evolution of QML towards QFL.

learnable activation functions as a fundamental component of DL architectures [97]. A consolidated comparison of recent KAN variants is summarized in Table II.

## III. KAN-Enhanced QFL

This section introduces the KAN-enhanced QFL framework and outlines its key processing components, including the KAN activation mechanism, the classical processing stage, and the quantum processing stage.

### A. Classical Processing Layer

The classical processing layer comprises KAN-driven classical AI modules, where KAN activations are embedded within both residual and transformer network architectures to extract and transform rich semantic features from the input data prior to quantum processing, thereby enabling effective HQC fusion. To ensure a compact and architecture-compatible KAN integration, this work adopts GR-KAN activations rather than grid-based B-spline KAN functions [16]. Accordingly, grid size and spline degree are not introduced as tunable hyperparameters in the proposed implementation. The grouped rational formulation provides learnable nonlinear mappings while controlling parameter growth through group-wise parameterization. This design reduces the number of KAN-specific hyperparameters and facilitates integration into different backbone models using a consistent activation-replacement strategy.

*1) KAN Activation:* The group-rational KAN (GR-KAN) activation function is an advanced alternative to conventional static activation functions such as ReLU and GELU. This activation introduces a learnable group-wise rational function applied over structured groups of input channels, thereby enhancing representational capacity and computational efficiency. Consider an input vector $\boldsymbol{z} = (z_1, z_2, \dots, z_N)$ with $N$ input channels. These channels are partitioned into $N_0$ groups, where each group contains $N/N_0$ channels. For an input channel indexed by $k$, the corresponding group index is given by $g_k = \lfloor (k-1)\, N_0/N \rfloor + 1$, ensuring channel-to-group consistency, where $\lfloor x \rfloor$ denotes the greatest integer less than or equal to $x$.

TABLE II
A STRUCTURED OVERVIEW OF KAN DEVELOPMENTS

| Class | Model | Feature | Advantage | Limitation |
|---|---|---|---|---|
| Foundational | KAN [14] | Learnable spline-based activation functions replacing linear weights | Expressivity and interpretability; strong in regression and PDE solving | Computationally expensive compared to standard MLPs |
| | KAN 2.0 [82] | MultKAN (multiplicative nodes), kanpiler, tree converter for symbolic representation | Symbolic–neural synergy, equation discovery, and modular interpretability | Functionality rather than speed optimization |
| Efficiency oriented | ReLU-KAN [83] | ReLU functions and matrix-friendly operations replacing splines | Fast GPU execution; scalable on grid and graph data | Spline-level smooth loss; fine-tuning requirement |
| | P1-KAN [84] | Piecewise-linear (P1) basis expansions | Fine-grained adaptation; stable high-dimensional approximation | More segments required for global smoothness |
| | BSRBF-KAN [85] | Radial-basis spline expansions for isotropic modeling | Smooth generalization; effective function approximation | Computationally heavy for dense radial-basis functions |
| Smooth and functional | Smooth-KAN [86] | Continuously differentiable activations for smooth gradients | Improved convergence, reliability in biomedical or scientific tasks | Slightly higher training time |
| | SineKAN [87] | Sinusoidal activations for lightweight deployments | Efficient periodic signal modeling; compact structure | Less suitable for non-periodic data |
| | Chebyshev-KAN [88] | Chebyshev polynomial activations | Improved nonlinear modeling capacity | Possible numerical instability at high polynomial degrees |
| Applied and theoretical | KAN-ODE [89] | KAN operators embedded in neural ODE solvers | Interpretable modeling of hidden physical dynamics | Computationally demanding due to ODE integration |
| | Benchmark [90] | Large-scale comparison between KAN and MLP in operator learning | Accuracy with fewer parameters; robust to noise | Slow training due to spline operations |
| | Low-Data [91] | KAN with limited samples | High adaptability and sample efficiency | Overfitting risk without regularization |
| | Fed-KAN [92] | Federated KAN for heterogeneous clients | Fast convergence; improved cross-client generalization | Communication and synchronization overhead |
| Architectural integrations and applications | Med-KAN [93], [94] | KAN integrated into biomedical imaging models | State-of-the-art interpretability and diagnostic accuracy | High computational cost for high-dimensional imaging |
| | KAN-Transformer [16] | KAN attention modules combined with transformers | Enhanced feature learning in vision and sequential tasks | More complex optimization; high training cost |
| | ConvKAN [95] | KAN operators replacing convolution kernels | More expressive convolutional layers | Inference-latency increase |
| | KAN-ResNet [96] | KAN layers embedded into ResNet architectures | Deeper and more stable KAN-based networks | Careful initialization required for convergence |
| | G-KAN [97] | KAN extended to graph and non-Euclidean data | Strong graph learning; effective edge modeling | Scalability challenges on large graphs |

The GR-KAN activation for the input vector $\boldsymbol{z}$ is given by [16]

$$\text{GR-KAN}\left(\boldsymbol{z}\right) = \boldsymbol{W}\boldsymbol{\phi}\left(\boldsymbol{z}\right) \tag{1}$$

where $\boldsymbol{W} \in \mathbb{R}^{M\times N}$ is a learnable weight matrix with $M$ output channels, consisting of scaling coefficients associated with each input-output edge. The vector $\boldsymbol{\phi}\left(\boldsymbol{z}\right)$ comprises rational activation values computed individually per channel but parameterized per group:

$$\boldsymbol{\phi}\left(\boldsymbol{z}\right) = \begin{bmatrix}\phi_{g_1}\left(z_1\right) & \phi_{g_2}\left(z_2\right) & \cdots & \phi_{g_N}\left(z_N\right)\end{bmatrix}^T \tag{2}$$

where the superscript $T$ denotes the transpose. For a given group $g_k \in \{1, 2, \ldots, N_0\}$, the corresponding rational nonlinearity is defined as

$$\phi_{g_k}\left(z_k\right) = \frac{\sum_{i=0}^{K_1} a_{g_k,i} z_k^i}{1 + \left|\sum_{j=1}^{K_2} b_{g_k,j} z_k^j\right|} \tag{3}$$

where $K_1$ and $K_2$ denote the orders of the rational activation; and the rational function coefficients $a_{g_k,i}$ and $b_{g_k,j}$ are shared across all channels within the same group $g_k$. These parameters are initialized to closely approximate the standard activation function (e.g., GELU) at the onset. All learnable parameters—including scaling weights $W_{i,j}$ in $\boldsymbol{W}$ and the rational function coefficients $a_i$ and $b_j$—are trained through back-propagation to optimize the activation functions for the target tasks.[1] To ensure architectural consistency, a recursive replacement routine traverses the module tree and systematically substitutes every instance of conventional activation functions (such as ReLU or GELU) with the `GroupKANActivationWrapper`, thereby incorporating the GR-KAN activation within the KAN-driven QFL framework.

[1]We denote the $(i, j)$th entry of a matrix $\boldsymbol{A}$ by $A_{i,j}$. The same indexing convention is adopted for higher-dimensional tensors.

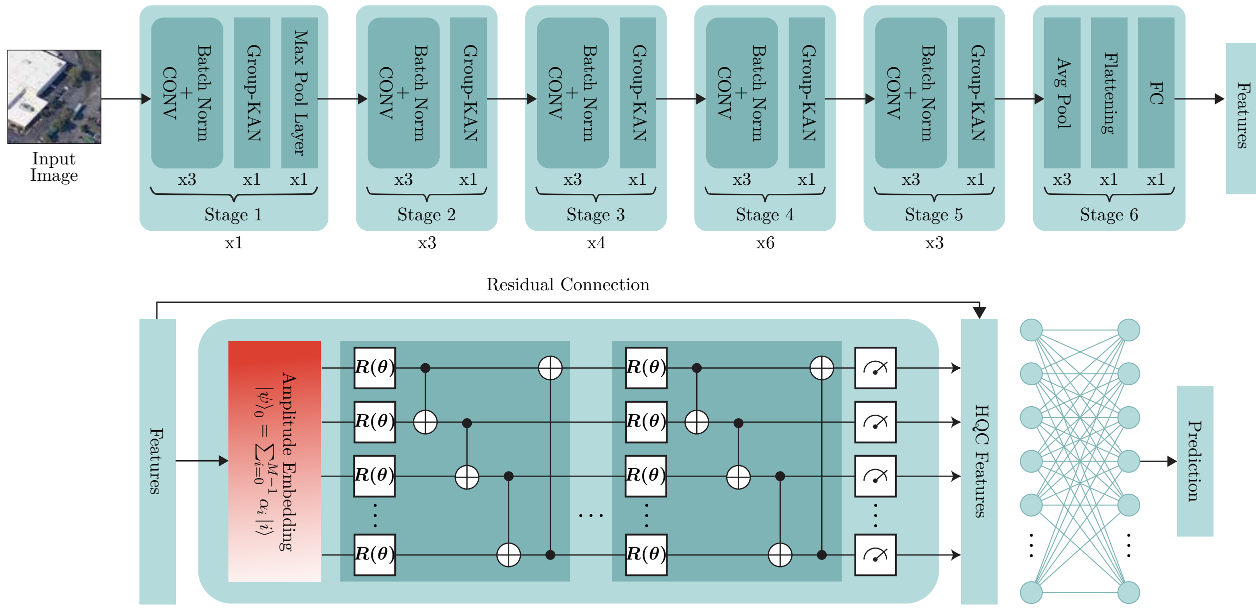


Fig. 2. A quantum-KAN-ResNet variant for QFL. The input image is processed through multi-stage convolutional blocks with residual connections and grouped KAN-based learnable activations for enhanced feature learning. The extracted features are amplitude-embedded into the PQC module, forming HQC features that are combined with a residual connection from pre-quantum features and passed to a fully connected layer for final prediction.

*2) KAN-Enhanced Residual Networks:* Integrating KAN activation into ResNet architectures significantly enhances their ability to model complex nonlinear relationships, particularly crucial for 6G ISAC tasks such as image classification and semantic feature extraction (see Fig. 2). The core components of this architecture comprise convolutional layers, batch normalization, KAN nonlinear activation, residual learning connections, flattening operations, and fully connected dense layers. Let $\boldsymbol{X}$ be the classical input data typically representing grayscale image data. For multichannel inputs such as red-green-blue (RGB) images or video streams, the corresponding image tensor $\boldsymbol{X} \in \mathbb{R}^{N_1 \times N_2 \times N}$ is fed into a convolution layer employing a filter tensor $\boldsymbol{Q} \in \mathbb{R}^{M \times N \times K \times K}$, where $N_1$ and $N_2$ denote the spatial height and width of the input feature map, respectively, and $K$ denotes the spatial kernel size of the convolution filter. The convolution operation produces the output feature map $\boldsymbol{Y} \in \mathbb{R}^{N_1' \times N_2' \times M}$ according to:

$$Y_{i,j,k} = \sum_{\ell=1}^{N} \sum_{m=1}^{K} \sum_{n=1}^{K} Q_{k,\ell,m,n} X_{i+m,j+n,\ell} \tag{4}$$

where $N_1'$ and $N_2'$ denote the spatial height and width of the output feature map, respectively. Following convolution, batch normalization stabilizes and accelerates the training process by normalizing the feature map:

$$Z_{i,j,k} = \gamma_k \frac{Y_{i,j,k} - \mu_k}{\sqrt{\sigma_k^2 + \epsilon}} + \beta_k \tag{5}$$

where $\mu_k$ and $\sigma_k^2$ are the batch-wise mean and variance for output channel $k$ computed from the convolution output feature map $\boldsymbol{Y}$, respectively, while $\gamma_k$ and $\beta_k$ are learnable scaling and shifting affine parameters. Here, $\epsilon$ is a small positive constant added for numerical stability. After reshaping the tensor $\boldsymbol{Z}$ into vector form $\boldsymbol{z}$, each normalized activation scalar $Z_{i,j}$ within the tensor $\boldsymbol{Z}$ is further processed by a learnable GR-KAN function, as defined in (1). In ResNet architectures, the residual unit incorporates an identity mapping that combines the original input features $\boldsymbol{X}$ with the KAN-activated transformation, which can be expressed in matrix form as follows:

$$\boldsymbol{V} = \boldsymbol{X} + \text{GR-KAN}\left(\boldsymbol{Z}\right). \tag{6}$$

This residual connection enhances gradient flow during backpropagation, thereby enabling efficient training of deeper networks by countering vanishing gradients. The aforementioned KAN-enhanced residual block can be repeated multiple times for enhanced learning capacity. Finally, the resulting feature tensor $\boldsymbol{V}_\star$ is flattened and fed into a fully connected (dense) layer for downstream classification tasks as follows:

$$\boldsymbol{\zeta} = \text{softmax}\left(\boldsymbol{W}_{\text{fc}}\, \text{flatten}\left(\boldsymbol{V}_\star\right) + \boldsymbol{b}_{\text{fc}}\right) \tag{7}$$

where $\boldsymbol{W}_{\text{fc}}$ and $\boldsymbol{b}_{\text{fc}}$ denote the trainable weights and bias of the fully connected layer. The resulting feature vector $\boldsymbol{\zeta}$ can be directly utilized by subsequent quantum processing layers to perform sophisticated tasks, including scene classification or predictive beamforming within 6G ISAC frameworks.

*3) KAN-Enhanced Transformer Networks:* The SwinT is a hierarchical vision-based transformer that employs self-attention within shifted local windows, effectively capturing spatial dependencies with computational efficiency. To further enhance its representational power for sophisticated 6G ISAC tasks, we introduce a KAN-enhanced variant (see Fig. 3). Specifically, we replace conventional activation functions (e.g., GELU) in both the SwinT blocks and MLP submodules with learnable KAN activations, enabling data-adaptive and highly expressive nonlinear transformations tailored to task-specific features. The designed module is characterized by three key enhancements: i) the frame-level representation

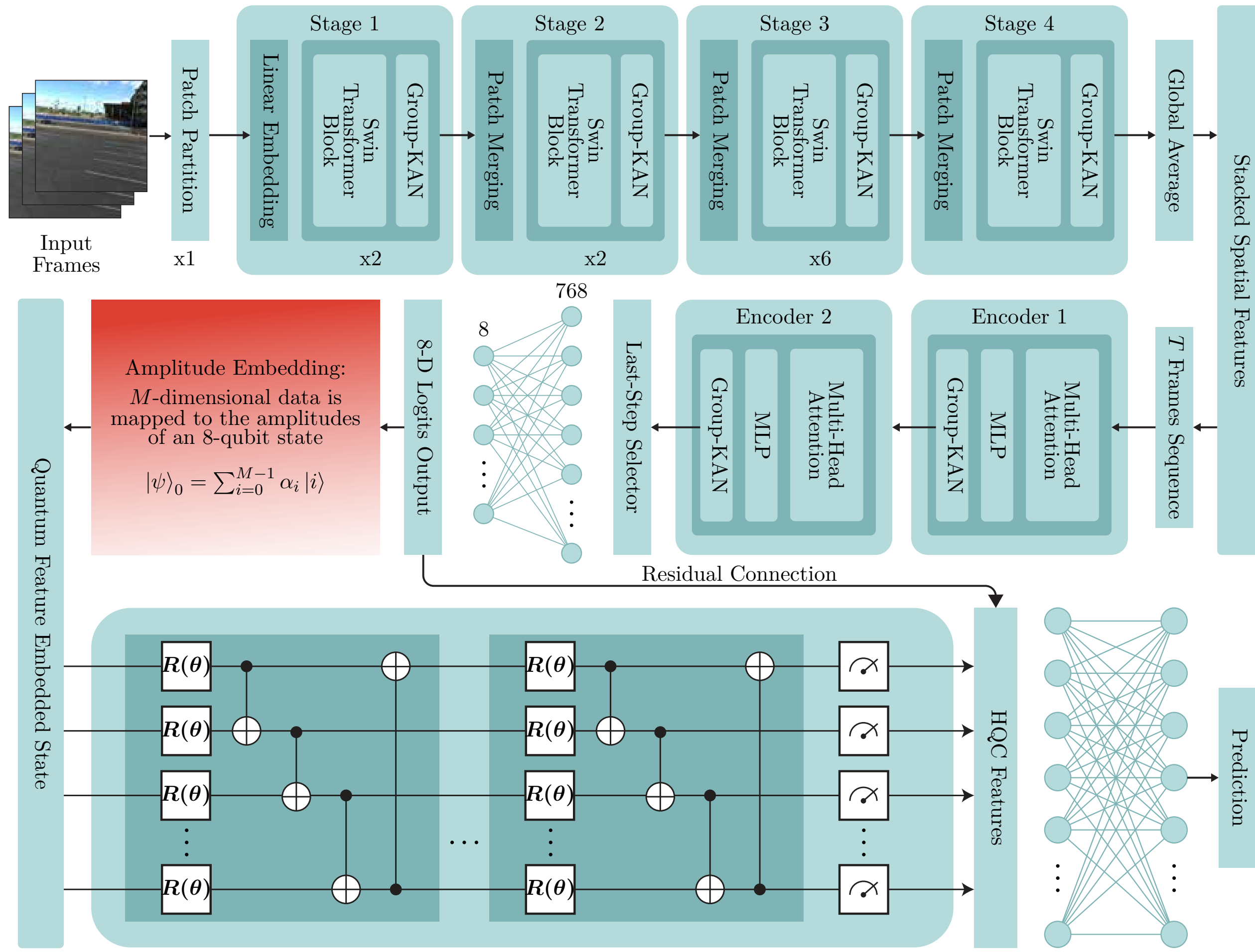


Fig. 3. A quantum-KAN-SwinT model variant for QFL. Input frames are partitioned and processed through hierarchical transformer stages with patch merging, spatial attention, and grouped KAN-based transformer blocks to extract stacked spatial and temporal features. These features are amplitude-embedded into the PQC module for enhanced feature learning in quantum space, forming HQC features that are combined with a residual connection and passed to a fully connected layer for final prediction.

mechanism, ii) the integration of KAN-based activation layers, and iii) the incorporation of temporal self-attention to capture cross-frame dependencies. Consider a batch of RGB video frames represented by a tensor $\boldsymbol{X} \in \mathbb{R}^{M_1 \times M_2 \times N_1 \times N_2 \times 3}$, where $M_1$ and $M_2$ denote the batch size and the temporal sequence length, respectively. For frame-level processing, the data is reshaped by flattening the temporal dimension to $\mathrm{reshape}\left(\boldsymbol{X}\right) \in \mathbb{R}^{M_1 M_2 \times N_1 \times N_2 \times 3}$, merging batch and temporal dimensions. Each frame $\boldsymbol{X}_{b,t}$ of this reshaped tensor is then individually processed by the SwinT module with KAN-enhanced feedforward sublayers, producing feature embeddings of a predefined dimensionality as follows:

$$\boldsymbol{f}_{b,t} = \mathrm{SwinT}_{\mathrm{KAN}}\left(\boldsymbol{X}_{b,t}\right) \tag{8}$$

for $b = 1, 2, \ldots, M_1$ and $t = 1, 2, \ldots, M_2$. These embeddings are subsequently reorganized into the stacked feature sequence tensor $\boldsymbol{F} = \mathrm{stack}\left(\boldsymbol{f}_{b,t}\right) \in \mathbb{R}^{M_1 \times M_2 \times M_3}$, where $M_3$ is the output embedding dimension. Within each MLP or feed-forward block of the SwinT architecture, standard activation functions are replaced with KAN activations, as defined in (1). These learnable rational activations dynamically adapt to complex, data-specific nonlinearities, significantly enhancing the network's ability to model intricate patterns crucial to ISAC tasks. To effectively capture temporal dynamics across frames, the sequence of frame embeddings $\boldsymbol{F}$ is further processed by a transformer encoding module comprising multiple attention heads (e.g., $8$ heads) and a large hidden dimension (e.g., $2,048$). This temporal encoder produces enriched contextual embeddings, formally defined as:

$$\boldsymbol{F}_{\star} = \mathrm{TemporalEncoder}\left(\boldsymbol{F}\right). \tag{9}$$

The embedding corresponding to the final time-step token $\boldsymbol{F}_{\star}^{(\mathrm{t})}$ serves as a global temporal context representation. These extracted high-dimensional embeddings are further condensed into a compact, actionable feature vector by a linear projection operation as follows:

$$\boldsymbol{\zeta} = \boldsymbol{W}_{\mathrm{p}}\, \mathrm{flatten}\big(\boldsymbol{F}_{\star}^{(\mathrm{t})}\big) + \boldsymbol{b}_{\mathrm{p}} \tag{10}$$

where $\boldsymbol{W}_{\mathrm{p}}$ and $\boldsymbol{b}_{\mathrm{p}}$ are learnable projection parameters. This compact representation $\boldsymbol{\zeta}$ is employed directly in downstream classification or regression tasks and can also be leveraged as input features for quantum-enhanced processing modules, supporting applications such as beam prediction, scene classification, or anomaly detection within ISAC frameworks.

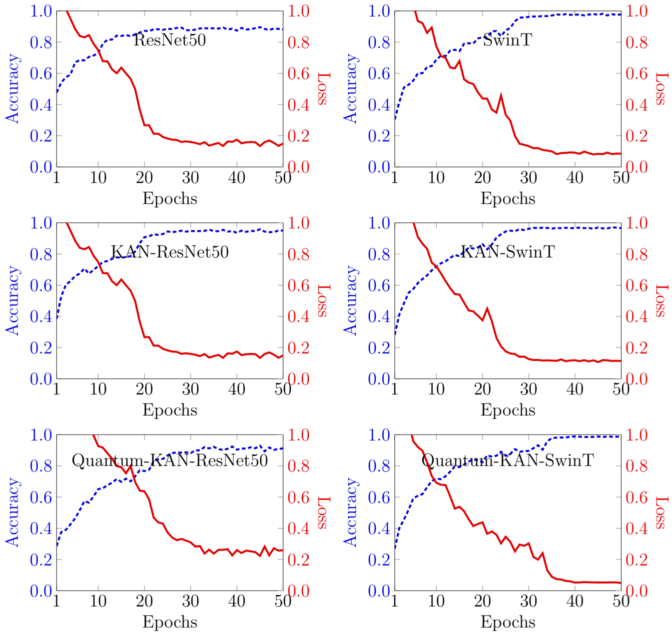


Fig. 4. Training accuracy and loss curves of ResNet50, KAN-ResNet50, and Quantum-KAN-ResNet50 models (left) as well as SwinT, KAN-SwinT, and Quantum-KAN-SwinT models (right) as a function of epochs for the RSSCN7 satellite-scene classification.

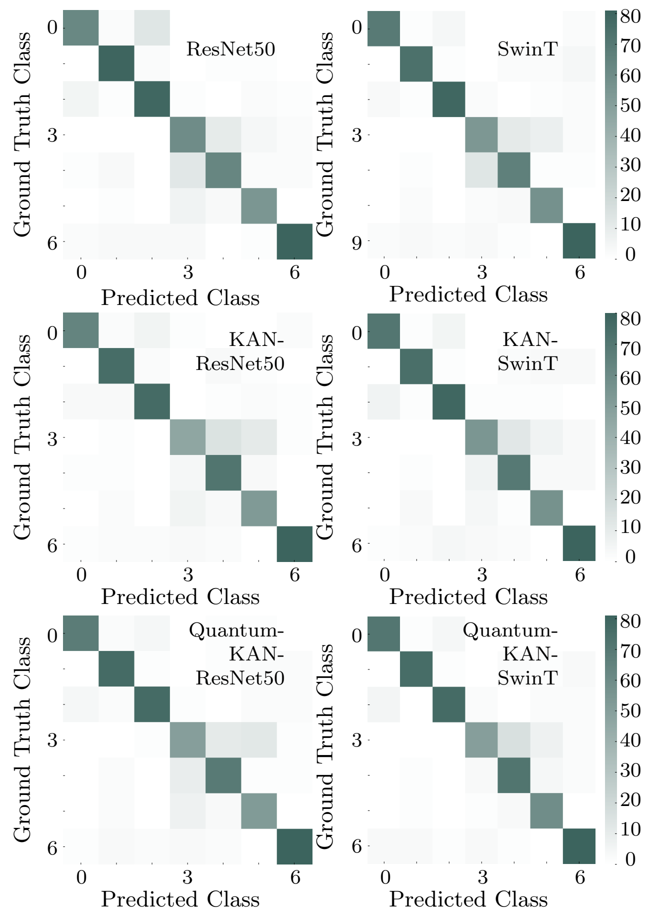


Fig. 5. Confusion matrices of ResNet50, KAN-ResNet50, and Quantum-KAN-ResNet50 models (left) as well as SwinT, KAN-SwinT, and Quantum-KAN-SwinT models (right) for the RSSCN7 satellite-scene classification.

### *B. Quantum Processing Layer*

The quantum processing layer operates on the extracted classical features by performing quantum feature embedding, entangling circuit operations, and quantum measurement, followed by post-processing to support downstream tasks such as satellite scene classification and ISAC beam prediction.

*1) Feature Embedding:* The output features of the classical processing layer are encoded into the PQC using amplitude embedding [1]. Specifically, the normalized feature data $\boldsymbol{\alpha} = \boldsymbol{\zeta}/\|\boldsymbol{\zeta}\|$ are encoded into the probability amplitudes of a quantum state composed of $\lceil \log_2 M \rceil$ qubits as follows:

$$|\psi\rangle_0 = \sum_{i=0}^{M-1} \alpha_i \, |i\rangle \tag{11}$$

where $\lceil x \rceil$ denotes the smallest integer greater than or equal to $x$ and $|i\rangle$ is the $i$th computational basis state. As discussed in [1, Section IV-A1], amplitude embedding enables compact representation of high-dimensional features but incurs a state-preparation cost scaling with $O(M)$, which motivates careful circuit-depth control in near-term implementations.

*2) Entangling Layer:* In this layer, the single-qubit rotation gates $\boldsymbol{R}(\theta)$ along with controlled-NOT (CNOT) gates play a pivotal role in inducing entanglement within the PQC where $\theta$ is a trainable rotation parameter [1]. The single-qubit rotations are implemented using Pauli-x, -y and/or -z rotation gates:

$$\boldsymbol{R}(\theta) = e^{-\imath \boldsymbol{\sigma}_{\mathrm{p}} \theta/2} \tag{12}$$

where $\imath = \sqrt{-1}$, $\mathrm{p} \in \{\mathrm{x}, \mathrm{y}, \mathrm{z}\}$, and $\boldsymbol{\sigma}_{\mathrm{x}} = |0\rangle\langle 1| + |1\rangle\langle 0|$, $\boldsymbol{\sigma}_{\mathrm{z}} = |0\rangle\langle 0| - |1\rangle\langle 1|$, and $\boldsymbol{\sigma}_{\mathrm{y}} = \imath \boldsymbol{\sigma}_{\mathrm{x}} \boldsymbol{\sigma}_{\mathrm{z}}$ are the Pauli operators. Following the HQC construction in [1, Section IV-A2], the embedded state is evolved through a stack of $L$ entangling layers, each consisting of trainable single-qubit rotations $\boldsymbol{R}(\theta)$ interleaved with the CNOT entangling operations [1, eq. (14)]. This design induces nonclassical correlations among qubits while maintaining a favorable balance between circuit expressiveness and hardware efficiency. A cyclic CNOT topology is adopted to promote entanglement across all qubits and to limit circuit depth.

*3) Quantum Measurement and Readout:* After the variational circuit prepares the PQC state, single-qubit measurements in the computational (e.g., Pauli-z) basis are performed. The expectation values of the corresponding observables are evaluated and collected into the classical feature vector, which serves as the quantum-enhanced representation. This output is then post-processed to support downstream tasks such as scene classification and ISAC beam prediction, following the HQC measurement-and-readout strategy in [1, Section IV-A3].

## IV. Numerical Results

This section presents numerical examples for the designed KAN-enhanced QFL framework across three representative 6G ISAC tasks: SGRS scene classification, beamforming prediction, and generative SC classification. The experimental results are presented using an ablation-style comparison among three model levels: the original classical backbone, the KAN-enhanced backbone, and the complete Quantum-KAN model. The classical

TABLE III
CLASSIFICATION PERFORMANCE OF CLASSICAL AND KAN-ENHANCED QFL MODELS ON THE RSSCN7 DATASET

| Model | Parameter | Accuracy | Precision | Recall | F1-Score | Loss |
|---|---|---|---|---|---|---|
| ResNet50 | 23,518,277 | 0.8446 | 0.8473 | 0.8446 | 0.8447 | 0.6185 |
| KAN-ResNet50 | 23,518,923 | **0.8482** | **0.8516** | **0.8482** | **0.8474** | 0.5630 |
| Quantum-KAN-ResNet50 | 23,519,003 | 0.8464 | 0.8480 | 0.8464 | 0.8463 | **0.5547** |
| SwinT | 27,581,999 | 0.8571 | 0.8550 | 0.8562 | 0.8553 | 0.5535 |
| KAN-SwinT | 27,582,911 | 0.8589 | 0.8565 | 0.8579 | 0.8566 | **0.5071** |
| Quantum-KAN-SwinT | 27,582,991 | **0.8696** | **0.8698** | **0.8693** | **0.8671** | 0.5446 |

backbone serves as the baseline without KAN and without PQC. The KAN-enhanced model measures the effect of replacing conventional activation functions with GR-KAN activation functions. The complete Quantum-KAN model quantifies the additional contribution of the PQC-based fusion module. This structure separates the effects of classical feature extraction, KAN-based nonlinear representation, and quantum fusion learning.

### A. RSSCN7 [3], [98]: SGRS Scene Classification

To assess the semantic scene understanding capability, we evaluate the framework on the RSSCN7 dataset, which consists of SGRS images spanning diverse land-cover categories.

*1) Dataset Composition:* We employ the RSSCN7 dataset, a benchmark for remote-sensing scene classification, which comprises seven diverse scene categories, including farmland, forest, and industrial areas. Each category encompasses images captured under varying weather conditions and from distinct geographic regions, making the classification task both challenging and representative of real-world complexity. To ensure reproducibility and consistency in performance evaluation, the dataset is partitioned using a fixed random seed. Specifically, $80\,\%$ of the available images (excluding the official test set) are used for model training, while the remaining $20\,\%$ are reserved for validation purposes. In addition, the separate official test set is used exclusively for the final evaluation of model generalization.

*2) Model Considerations and Training Parameters:* Model architectures and training configurations follow unified experimental setups, including ResNet50 and SwinT backbones in their classical, KAN-enhanced, and quantum-integrated variants. For the RSSCN7 remote-sensing scene classification task, single RGB images are used as inputs without temporal stacking. All models are trained for 50 epochs using the adaptive moment estimation (Adam) optimizer with an initial learning rate of $10^{-3}$ under identical optimization settings to ensure fair comparison.

To ensure stable and efficient convergence, we utilize the `ReduceLROnPlateau` learning-rate scheduler, which reduces the learning rate by a factor of $0.1$ whenever the validation loss fails to improve for five consecutive epochs. Training is conducted using mini-batch gradient descent with a batch size of 32. In addition, to stabilize the training of the HQC architecture, we adopt a conservative optimization strategy based on a relatively low initial learning rate and adaptive learning-rate scheduling. This is particularly important for the Quantum-KAN-ResNet50 variant, where the deep residual backbone, GR-KAN activation functions, and the PQC introduce multiple levels of trainable nonlinear transformations. An excessively high learning rate may lead to unstable parameter updates, poor convergence, or degraded validation performance. Therefore, the adaptive learning-rate strategy gradually adjusts the update scale during training, allowing the model to refine both classical and quantum trainable parameters more smoothly. Although no additional explicit regularization technique is introduced in this study, this optimization strategy helps improve training stability and reduces the risk of unstable convergence in the proposed HQC framework.

To complement the performance evaluation, we also analyze the model complexity in terms of trainable parameters. The proposed KAN-based modules introduce additional learnable nonlinear mappings, but their grouped formulation helps limit parameter growth compared with fully independent channel-wise mappings. Thus, the parameter count comparison provides a hardware-independent indication of the additional classical model complexity introduced by the KAN integration. We note that classical floating-point operations per second (FLOPS)-based profiling is not directly transferable to the simulated quantum processing stage, since quantum execution is more appropriately characterized by circuit-level and hardware-dependent metrics such as circuit depth, number of shots, quantum volume, and circuit-layer operations per second (CLOPS) [99]. Therefore, detailed FLOPS/runtime and quantum-hardware-aware profiling, which depend on the target QPU and execution backend, are left for future deployment-oriented evaluation.

*Quantum-KAN-ResNets:* We evaluate three model variants derived from the ResNet50 architecture. The first variant, designated as baseline ResNet50, employs the standard convolutional structure with conventional ReLU activation functions. The second variant, termed KAN-ResNet50, integrates KAN-based activation functions in place of ReLU, facilitating the learning of more flexible and data-driven nonlinear representations. The third variant, Quantum-KAN-ResNet50, further extends KAN-ResNet50 by incorporating PQC modules into its feature-extraction pipeline (see Fig. 2). This quantum enhancement is designed to capture complex higher-order interactions within extracted features using 8-qubit and 10-layer PQCs, potentially augmenting classification performance.

*Quantum-KAN-SwinTs:* The SwinT architecture utilizes hierarchical self-attention mechanisms with shifted window partitions, providing strong inductive biases particularly suited for capturing spatial features in remote sensing imagery. To ensure a fair and consistent evaluation, we adopt the same fixed training–validation split used in the ResNet experiments. We examine three SwinT-based model variants. The baseline SwinT model

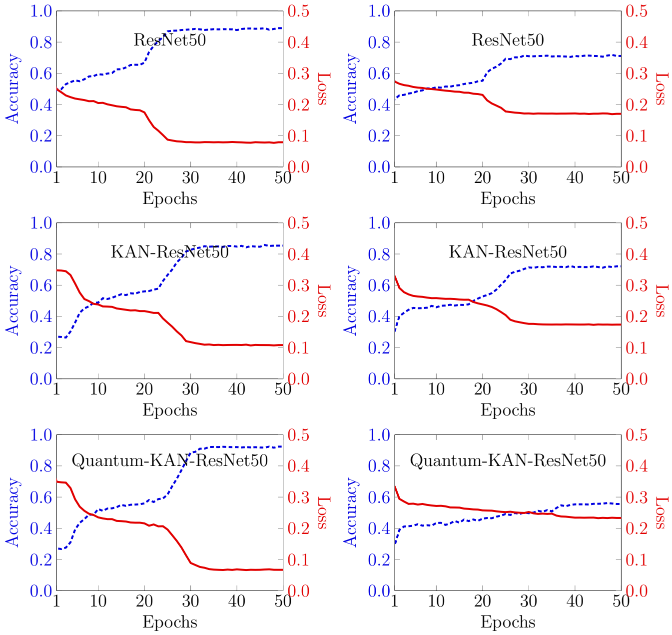


Fig. 6. Training accuracy and loss curves of ResNet50, KAN-ResNet50, and Quantum-KAN-ResNet50 as a function of epochs for the DeepSense 6G dataset Scenarios: SC-8 (left) and SC-9 (right).

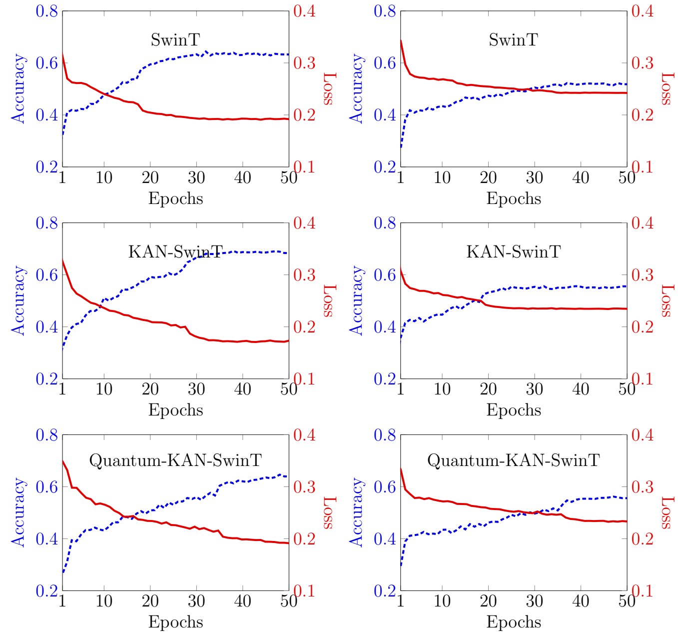


Fig. 7. Training accuracy and loss curves of SwinT, KAN-SwinT, and Quantum-KAN-SwinT as a function of epochs for the DeepSense 6G dataset Scenarios: SC-8 (left) and SC-9 (right).

employs the standard architecture with GELU activation functions. The second variant, termed KAN-SwinT, replaces all GELU activations with KAN-based activation functions, which provide more flexible and data-adaptive nonlinear transformations. Finally, Quantum-KAN-SwinT extends this KAN-SwinT by integrating PQC modules that transform the learned feature representations using amplitude encoding and quantum entanglement to effectively capture complex feature correlations (see Fig. 3).

*3) Classification Performance:* Fig. 4 and Fig. 5 illustrate the training performance curves and confusion matrices of KAN-ResNet50, Quantum-KAN-ResNet50, KAN-SwinT, and Quantum-KAN-SwinT in comparison with their respective baseline models, respectively. The baseline ResNet50 model achieves a test accuracy of $84.46\,\%$. By introducing KAN-based activations, the KAN-ResNet50 variant improves performance, reaching a test accuracy of $84.82\,\%$. Further extending KAN-ResNet50 with the PQC module, the Quantum-KAN-ResNet50 variant attains a test accuracy of $84.64\,\%$. As summarized in Table III, these results demonstrate that both the KAN-based activation functions and the PQC-enhanced module yield competitive improvements over the conventional ResNet50 baseline, with KAN-ResNet50 providing the most notable accuracy gain. In contrast, the transformer-based models exhibit more pronounced improvements. The baseline SwinT achieves a test accuracy of $85.71\,\%$, while KAN-SwinT improves this performance to $85.89\,\%$, reflecting better generalization enabled by KAN-based nonlinearities. Notably, Quantum-KAN-SwinT achieves the highest test accuracy of $86.96\,\%$, demonstrating the advantage of quantum-enhanced feature modeling. Overall, these results suggest that KAN-based activations enhance generalization performance by providing adaptive and expressive nonlinear transformations, while the incorporation of PQC modules further boosts accuracy by capturing higher-order feature dependencies that are difficult to model using purely classical architectures.

### *B. DeepSense 6G [32]: ISAC Beam Prediction*

To evaluate ISAC beamforming prediction performance, we use the DeepSense 6G dataset, focusing specifically on image-based input data to predict optimal beam indices under realistic vehicular communication scenarios.

*1) Dataset Composition:* We leverage the DeepSense 6G dataset, specifically utilizing Scenarios 6, 8, and 9, to assess our vision-aided beam prediction framework. The dataset offers synchronized RGB imagery and corresponding millimeter-wave (mmWave) power measurements across various urban and vehicular environments, enabling in-depth investigation of visually-guided wireless communication systems. The selected scenarios differ in scale and complexity. Scenario 6 (SC-6) comprises $915$ instances, Scenario 8 (SC-8) includes $4,043$ instances, and Scenario 9 (SC-9) contains $5,964$ instances. We again partition each scenario into $80\,\%$ training and $20\,\%$ validation subsets, maintaining a fixed random seed to guarantee reproducibility.

The original task involves selecting the optimal beam index from $64$ mmWave power values. To enable the integration of PQC modules within our model, we reduce the output dimensionality by selecting the $8$ most frequently activated beam indices from the original set of mmWave beam measurements [32]. Accordingly, the considered task remains a beam prediction problem, where the model predicts the optimal beam among the selected representative candidates. This reduced beam-index formulation alleviates output sparsity and enables a controlled comparison of different learning architectures under the same classification setting. Since the same beam-index mapping, data partitioning, and evaluation protocol are applied to all baseline and proposed models, the relative performance

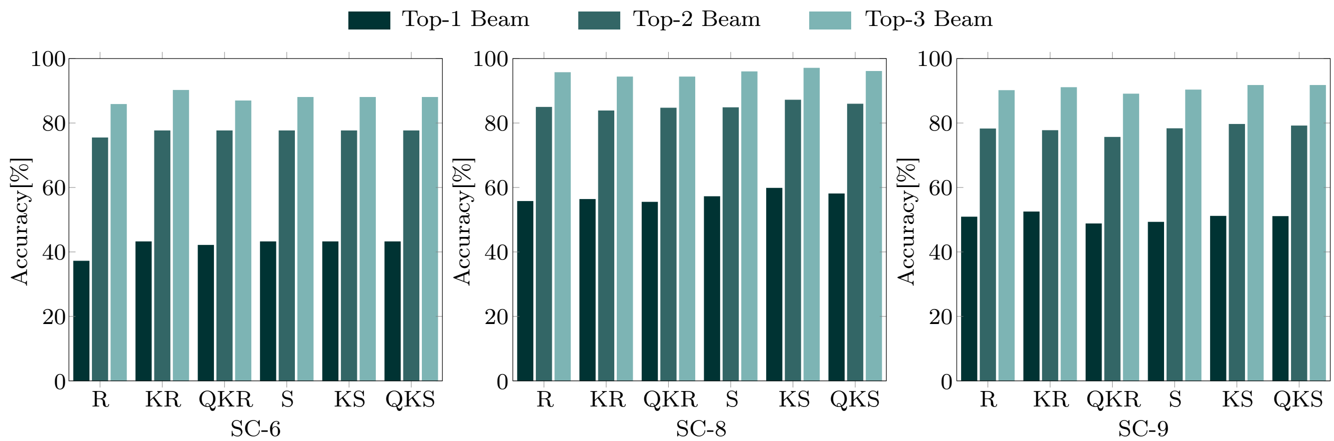


Fig. 8. Beamforming prediction accuracy (top-1, top-2, and top-3) of classical baselines and KAN-enhanced QFL models across the DeepSense 6G dataset Scenarios (SC-6, SC-8, and SC-9). R, KR, QKR, S, KS, and QKS stand for ResNet50, KAN-ResNet50, quantum-KAN-ResNet50, SwinT, KAN-SwinT, and quantum-KAN-SwinT models, respectively.

comparison remains consistent and fair. The reported results should be interpreted as beam prediction over a selected set of representative beam indices, which is relevant to hierarchical and low-complexity beam management in 6G systems.

*2) Model Considerations and Training Parameters:* To investigate vision-aided beam prediction for vehicular communication scenarios, we adapt and extend both residual network and transformer-based architectures to effectively process temporal input sequences derived from the DeepSense 6G dataset. Unless otherwise stated, the model architectures and training procedures for this task are identical to those described in Section IV-A2. Unlike RSSCN7, this task requires explicit temporal modeling. Accordingly, all models process inputs structured as sequences of five consecutive RGB frames, stacked along the channel dimension to form tensors of shape $15 \times N_1 \times N_2$, facilitating effective utilization of short-term temporal dynamics for beam prediction tasks. Owing to the additional memory and computational overhead introduced by temporal processing and quantum module integration, each mini-batch for training and validation consists of eight temporal samples, each containing five sequential frames. This sequential data structure allows the models to effectively harness spatiotemporal information, thereby enhancing beam prediction accuracy. The `ReduceLROnPlateau` scheduler employs three consecutive epochs. For ResNet-based models, the initial learning rate is set to $10^{-3}$, while for SwinT-based models, the rate of $10^{-4}$ is used to accommodate the increased optimization sensitivity of transformer architectures.

*3) Prediction Performance:* Fig. 6 and Fig. 7 depict training accuracy and loss curves for benchmark models in DeepSense 6G Scenarios 8 and 9. Top-1, -2, and -3 beam prediction accuracies (in percentages) are plotted in Fig. 8, while Table IV reports the corresponding accuracies (in ratios), validation losses, and parameter counts for all model variants across Scenarios 6, 8, and 9 in the DeepSense 6G dataset. For Scenario 6, the KAN-ResNet50 model achieves the highest top-1 accuracy of $43.17\,\%$, surpassing both the baseline ResNet50 ($37.16\,\%$) and its quantum-integrated counterpart, Quantum-KAN-ResNet50, which reaches $42.08\,\%$. Notably, Quantum-KAN-ResNet50 demonstrates a slightly lower validation loss of $0.2630$, indicating improved convergence despite minor reductions in top-3 accuracy. For Scenario 8, KAN-ResNet50 again shows superior performance with a top-1 accuracy of $56.31\,\%$ and the lowest validation loss of $0.2563$. While Quantum-KAN-ResNet50 achieves a competitive top-2 accuracy of $84.65\,\%$, its overall performance slightly decreases, likely due to quantum module sensitivity to noise in large-scale datasets. In Scenario 9, KAN-ResNet50 maintains its leading performance with a top-1 accuracy of $52.43\,\%$ and a top-3 accuracy of $91.02\,\%$. The Quantum-KAN-ResNet50 variant shows slightly lower accuracy across all metrics, potentially due to reduced regularization effects from the quantum layer, leading to mild overfitting. The transformer-based models exhibit consistently robust performance, most notably in Scenario 8, where the KAN-SwinT achieves the highest overall accuracy with a top-1 accuracy of $59.78\,\%$ and a top-3 accuracy of $97.03\,\%$. Quantum-KAN-SwinT closely follows, attaining $58.04\,\%$ top-1 and $96.04\,\%$ top-3 accuracy, while also achieving the lowest validation loss of $0.2152$, highlighting the effectiveness of quantum integration in refining learned feature representations. Across all evaluated scenarios, KAN-enhanced models consistently outperform baselines, underscoring the benefits of data-adaptive nonlinear activations. Moreover, quantum-integrated models further exhibit competitive performance and often improved performance, particularly in Scenarios 6 and 8, by effectively capturing intricate spatiotemporal dependencies. These findings highlight the potential of HQC architectures in enhancing ISAC beam prediction accuracy within temporally structured and resource-constrained contexts.

### *C. GenSC-6G [33]: Generative SC Classification*

To evaluate the effectiveness of KAN-enhanced QFL models for generative SC tasks, we conduct supervised classification experiments on the GenSC-6G dataset using plain (noise-free) images. This example focuses on the semantic classification of realistic vehicular imagery with provided ground-truth labels.

*1) Dataset Composition:* The GenSC-6G dataset is designed to support a wide range of SC and ML tasks, including classification, segmentation, object detection, and edge LLM applications.

TABLE IV
PREDICTION PERFORMANCE OF CLASSICAL AND KAN-ENHANCED QFL MODELS ON THE DEEPSENSE 6G DATASET

| Scenario | Model | Parameter | Accuracy | | | Loss |
|---|---|---|---|---|---|---|
| | | | Top-1 | Top-2 | Top-3 | |
| SC-6 | ResNet50 | 23,562,056 | 0.3716 | 0.7541 | 0.8579 | 0.2727 |
| | KAN-ResNet50 | 23,562,702 | **0.4317** | 0.7760 | **0.9016** | 0.2635 |
| | Quantum-KAN-ResNet50 | 23,562,854 | 0.4208 | **0.7760** | 0.8689 | **0.2630** |
| | SwinT | 38,613,810 | 0.4317 | 0.7760 | 0.8798 | 0.2626 |
| | KAN-SwinT | 38,614,722 | 0.4317 | 0.7760 | 0.8798 | 0.2626 |
| | Quantum-KAN-SwinT | 38,614,874 | 0.4317 | 0.7760 | 0.8798 | 0.2659 |
| SC-8 | ResNet50 | 23,562,056 | 0.5569 | **0.8490** | **0.9567** | 0.2987 |
| | KAN-ResNet50 | 23,562,702 | **0.5631** | 0.8379 | 0.9431 | **0.2563** |
| | Quantum-KAN-ResNet50 | 23,562,854 | 0.5545 | 0.8465 | 0.9431 | 0.3184 |
| | SwinT | 38,613,810 | 0.5718 | 0.8478 | 0.9592 | 0.2175 |
| | KAN-SwinT | 38,614,722 | **0.5978** | **0.8713** | **0.9703** | 0.2197 |
| | Quantum-KAN-SwinT | 38,614,874 | 0.5804 | 0.8589 | 0.9604 | **0.2152** |
| SC-9 | ResNet50 | 23,562,056 | 0.5084 | **0.7819** | 0.9010 | 0.2711 |
| | KAN-ResNet50 | 23,562,702 | **0.5243** | 0.7768 | **0.9102** | **0.2561** |
| | Quantum-KAN-ResNet50 | 23,562,854 | 0.4874 | 0.7559 | 0.8901 | 0.2739 |
| | SwinT | 38,613,810 | 0.4924 | 0.7827 | 0.9027 | 0.2519 |
| | KAN-SwinT | 38,614,722 | **0.5109** | **0.7961** | **0.9169** | **0.2471** |
| | Quantum-KAN-SwinT | 38,614,874 | 0.5101 | 0.7911 | 0.9169 | 0.2489 |

TABLE V
CLASSIFICATION PERFORMANCE OF CLASSICAL AND KAN-ENHANCED QFL MODELS ON THE GENSC-6G DATASET

| Model | Parameter | Accuracy | | | Training Loss |
|---|---|---|---|---|---|
| | | Training | Validation | Test | |
| ResNet50 | 23,518,277 | **0.99677** | 0.78516 | 0.64635 | **0.03213** |
| KAN-ResNet50 | 23,518,923 | 0.98086 | **0.81250** | 0.61146 | 0.08061 |
| Quantum-KAN-ResNet50 | 23,519,003 | 0.99138 | 0.79688 | **0.68385** | 0.05015 |
| SwinT | 27,581,999 | 0.99452 | 0.70614 | 0.61290 | 0.04967 |
| KAN-SwinT | 27,582,911 | **0.99671** | 0.74123 | **0.63441** | **0.02767** |
| Quantum-KAN-SwinT | 27,582,991 | 0.99123 | **0.74561** | 0.61290 | 0.06569 |

The full dataset contains ground-truth annotated data for $15$ semantic classes, with $4,829$ instances in the training split and $1,320$ instances in the testing split. Each instance consists of an image paired with its semantic class label, primarily representing common vehicle types in both military and civilian domains. In this example, we use the dataset for image classification, selecting the first five classes from the original label space due to current NISQ limitations. This reduced five-class setup ensures computational feasibility while preserving the essential characteristics required to demonstrate the SC framework. Only plain image data and ground-truth class labels are used without introducing additive noise or feature-level perturbations. From the selected subset, a total of $1,142$ instances are retained. The original training portion, comprising approximately $1,133$ instances, is further divided into the training ($80\,\%$) and validation ($20\,\%$) subsets. As a result, the dataset consists of 913 training images, 229 validation images, and 373 test images.

*2) Model Considerations and Training Parameters:* The experimental configuration for this example follows the same model architectures and training procedures described in Section IV-A2, except that the initial learning rate for the ResNet-based variants is set to $10^{-4}$.

*3) Classification Performance:* The classification results for the GenSC-6G task are summarized in Table V. The baseline ResNet50 model achieves strong training performance with $99.68\,\%$ training accuracy and $78.52\,\%$ validation accuracy, resulting in a test accuracy of $64.64\,\%$. Incorporating the KAN activations into the ResNet architecture yields comparable results, with KAN-ResNet50 reaching $98.09\,\%$ training accuracy, $81.25\,\%$ validation accuracy, and $61.15\,\%$ test accuracy. When further enhanced with quantum-inspired PQC modules, the Quantum-KAN-ResNet50 model maintains stable training behavior and attains the highest test accuracy among the ResNet variants at $68.38\,\%$, highlighting the potential benefits of quantum-enhanced feature learning. For the transformer-based models, the baseline SwinT achieves $99.45\,\%$ training accuracy, $70.61\,\%$ validation accuracy, and $61.29\,\%$ test accuracy. Introducing KAN activations improves training accuracy to $99.67\,\%$ and validation accuracy to $74.12\,\%$, achieving $63.44\,\%$ test accuracy with the lowest training loss among all models. The Quantum-KAN-SwinT variant attains the highest validation accuracy of $74.56\,\%$. Overall, these results demonstrate that both KAN-based activations and quantum-inspired extensions deliver competitive performance in generative SC classification, even under reduced-class and NISQ-constrained settings.

## V. Conclusion

This paper has put forth a KAN-enhanced QFL framework tailored for 6G networks. By combining PQCs with KAN-enhanced ResNet and SwinT architectures, the designed HQC models effectively capture intricate nonlinearities, high-order feature interactions, and adaptive representations required in dynamic wireless environments. Extensive evaluations on the remote-sensing satellite scene classification (RSSCN7 dataset), ISAC beamforming prediction (DeepSense 6G dataset), and generative SC classification (GenSC-6G dataset) demonstrate comparable task-dependent improvement in predictive accuracy and generalization capability over purely classical AI-based models. These results highlight the QFL potential for enabling semantic-aware communications and digital-twin intelligence in 6G networks. Future work will focus on enhancing the scalability and hardware efficiency of quantum-integrated, KAN-enhanced large AI architectures as well as realizing practical co-deployment of PQCs and generative diffusion models in multimodal 6G environments.

## References

[1] S. Tariq, B. E. Arfeto, U. Khalid, S. Kim, T. Q. Duong, and H. Shin, "Deep quantum-transformer networks for multimodal beam prediction in ISAC systems," *IEEE Internet Things J.*, vol. 11, no. 18, pp. 29 387–29 401, Sep. 2024.

[2] J. Yao, M. Jin, T. Wu, M. Elkashlan, C. Yuen, K.-K. Wong, G. K. Karagiannidis, and H. Shin, "FAS-driven spectrum sensing for cognitive radio networks," *IEEE Internet Things J.*, vol. 12, no. 5, pp. 6046–6049, Mar. 2025.

[3] U. Khalid, U. I. Paracha, Z. Naveed, T. Q. Duong, M. Z. Win, and H. Shin, "Quantum fusion intelligence for integrated satellite-ground remote sensing," *IEEE Wireless Commun.*, vol. 32, no. 3, pp. 46–55, Jun. 2025.

[4] A. Faisal, I. Al-Nahhal, K. Lee, O. A. Dobre, and H. Shin, "Conditional generative adversarial networks for channel estimation in RIS-assisted ISAC systems," *IEEE Trans. Commun.*, vol. 73, no. 9, pp. 7828–7841, Sep. 2025.

[5] S. Kolb, A. Madden, N. Kröger, F. Mehmeti, F. Jurosch, L. Bernhard, W. Kellerer, and D. Wilhelm, "6G in medical robotics: Development of network allocation strategies for a telerobotic examination system," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 20, no. 1, pp. 167–178, Sep. 2025.

[6] F. Naeem, M. Ali, G. Kaddoum, C. Huang, and C. Yuen, "Security and privacy for reconfigurable intelligent surface in 6G: A review of prospective applications and challenges," *IEEE Open J. Commun. Soc.*, vol. 4, pp. 1196–1217, May 2023.

[7] F. Wang, S. Zhang, H. Yang, and T. Q. Quek, "Non-terrestrial networking for 6G: Evolution, opportunities, and future directions," *Engineering*, vol. 54, pp. 56–68, Nov. 2025.

[8] S. Singh, K. Singh, S. K. Singh, H. Shin, and T. Q. Duong, "Beamforming design towards sum-rate maximization for holographic active RIS-aided uplink near-field communications," *IEEE Internet Things J.*, vol. 12, no. 12, pp. 18 582–18 595, Jun. 2025.

[9] B. Narottama and S. Y. Shin, "Federated quantum neural network with quantum teleportation for resource optimization in future wireless communication," *IEEE Trans. Veh. Technol.*, vol. 72, no. 11, pp. 14 717–14 731, Nov. 2023.

[10] A. Faisal, I. Al-Nahhal, O. A. Dobre, T. M. N. Ngatched, and H. Shin, "Deep reinforcement learning for RIS-aided full-duplex systems: Advances and challenges," *IEEE Commun. Mag.*, vol. 63, no. 6, pp. 82–89, Jun. 2025.

[11] Q. Li, P. Sisk, A. Kannan, T. Yoo, T. Luo, G. Shah, B. Manjunath, C. Samarathungage, M. T. Boroujeni, H. Pezeshki, and H. Joshi, "Machine learning based time domain millimeter-wave beam prediction for 5G-advanced and beyond: Design, analysis, and over-the-air experiments," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1787–1802, Jun. 2023.

[12] F. Zaman, A. Farooq, M. A. Ullah, H. Jung, H. Shin, and M. Z. Win, "Quantum machine intelligence for 6G URLLC," *IEEE Wireless Commun.*, vol. 30, no. 2, pp. 22–30, Apr. 2023.

[13] A. Polar and M. Poluektov, "Probabilistic Kolmogorov-Arnold network," *arXiv:2104.01714*, Apr. 2021.

[14] Z. Liu, Y. Wang, S. Vaidya, F. Ruehle, J. Halverson, M. Soljačić, T. Y. Hou, and M. Tegmark, "KAN: Kolmogorov-Arnold networks," *arXiv:2404.19756*, Apr. 2024.

[15] M. Shokrnezhad, H. Mazandarani, and T. Taleb, "Fairness-utilization trade-off in wireless networks with explainable Kolmogorov-Arnold networks," *arXiv:2411.01924*, Nov. 2024.

[16] X. Yang and X. Wang, "Kolmogorov-Arnold transformer," *arXiv:2409.10594*, Sep. 2024.

[17] F. Zaman, U. Khalid, T. Q. Duong, H. Shin, and M. Z. Win, "Quantum full-duplex communication," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 9, pp. 2966–2980, Sep. 2023.

[18] S. N. Paing, J. W. Setiawan, T. Q. Duong, D. Niyato, M. Z. Win, and H. Shin, "Quantum anonymous networking: A quantum leap in privacy," *IEEE Netw.*, vol. 38, no. 5, pp. 131–145, Sep. 2024.

[19] J. Wu, W. Yuan, X. Zhang, Y. Yu, Y. Cui, F. Liu, G. Sun, J. Wang, D. Niyato, and D. I. Kim, "Toward multi-functional LAWNs with ISAC: Opportunities, challenges, and the road ahead," *IEEE Wireless Commun.*, pp. 1–8, Mar. 2026.

[20] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *Proc. 35th annual symposium on foundations of computer science*, Santa Fe, NM, USA, Nov. 1994, pp. 124–134.

[21] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proc. of the 28th annual ACM symposium on Theory of computing*, Philadelphia, PA, USA, Jul. 1996, pp. 212–219.

[22] A. W. Harrow, A. Hassidim, and S. Lloyd, "Quantum algorithm for linear systems of equations," *Phys. Rev. Lett.*, vol. 103, no. 15, p. 150502, Oct. 2009.

[23] U. Khalid, M. S. Ulum, A. Farooq, T. Q. Duong, O. A. Dobre, and H. Shin, "Quantum semantic communications for Metaverse: Principles and challenges," *IEEE Wireless Commun.*, vol. 30, no. 4, pp. 26–36, Aug. 2023.

[24] M. S. Ulum, U. Khalid, J. W. Setiawan, T. Q. Duong, M. Z. Win, and H. Shin, "Variational anonymous quantum sensing," *IEEE J. Sel. Areas Commun.*, vol. 42, no. 9, pp. 2275–2291, Sep. 2024.

[25] A. Khan, U. Khalid, J. ur Rehman, and H. Shin, "Quantum anonymous private information retrieval for distributed networks," *IEEE Trans. Commun.*, vol. 70, no. 6, pp. 4026–4037, Jun. 2022.

[26] S. M. A. Rizvi, U. Khalid, S. Chatzinotas, T. Q. Duong, and H. Shin, "Controlled quantum semantic communication for industrial CPS networks," *IEEE Trans. Netw. Sci. Eng.*, vol. 13, pp. 996–1009, 2026.

[27] U. Khalid, J. ur Rehman, S. N. Paing, H. Jung, T. Q. Duong, and H. Shin, "Quantum network engineering in the NISQ age: Principles, missions, and challenges," *IEEE Netw.*, vol. 38, no. 1, pp. 112–123, Jan. 2024.

[28] T. Q. Duong, J. A. Ansere, B. Narottama, V. Sharma, O. A. Dobre, and H. Shin, "Quantum-inspired machine learning for 6G: Fundamentals, security, resource allocations, challenges, and future research directions," *IEEE Open J. Veh. Technol.*, vol. 3, pp. 375–387, Aug. 2022.

[29] S. Tariq, U. Khalid, B. E. Arfeto, T. Q. Duong, and H. Shin, "Integrating sustainable big AI: Quantum anonymous semantic broadcast," *IEEE Wireless Commun.*, vol. 31, no. 3, pp. 86–99, Jun. 2024.

[30] M. Schuld and N. Killoran, "Quantum machine learning in feature Hilbert spaces," *Phys. Rev. Lett.*, vol. 122, no. 4, p. 040504, Feb. 2019.

[31] A. Mari, T. R. Bromley, J. Izaac, M. Schuld, and N. Killoran, "Transfer learning in hybrid classical-quantum neural networks," *Quantum*, vol. 4, p. 340, Oct. 2020.

[32] A. Alkhateeb, G. Charan, T. Osman, A. Hredzak, J. Morais, U. Demirhan, and N. Srinivas, "Deepsense 6G: A large-scale real-world multi-modal sensing and communication dataset," *IEEE Commun. Mag.*, vol. 61, no. 9, pp. 122–128, Jun. 2023.

[33] B. E. Arfeto, S. Tariq, U. Khalid, T. Q. Duong, and H. Shin, "GenSC-6G: A prototype testbed for integrated generative AI, quantum, and semantic communication," *IEEE Commun. Mag.*, vol. 63, no. 10, pp. 28–35, Oct. 2025.

[34] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum principal component analysis," *Nat. Phys.*, vol. 10, no. 9, pp. 631–633, Jul. 2014.

[35] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy, and R. Melko, "Quantum Boltzmann machine," *Phys. Rev. X*, vol. 8, no. 2, p. 021050, May 2018.

[36] N. Wiebe, A. Kapoor, and K. M. Svore, "Quantum algorithms for nearest-neighbor methods for supervised and unsupervised learning," *Quantum Inform. Comput.*, vol. 15, no. 3–4, pp. 316–356, Mar. 2015.

[37] P. Rebentrost, M. Mohseni, and S. Lloyd, "Quantum support vector machine for big data classification," *Phys. Rev. Lett.*, vol. 113, no. 13, p. 130503, Sep. 2014.

[38] S. Lloyd, M. Mohseni, and P. Rebentrost, "Quantum algorithms for supervised and unsupervised machine learning," *arXiv:1307.0411*, Jul. 2013.

[39] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, no. 7671, pp. 195–202, Sep. 2017.

[40] V. Dunjko and H. J. Briegel, "Machine learning & artificial intelligence in the quantum domain: A review of recent progress," *Rep. Prog. Phys.*, vol. 81, no. 7, p. 074001, Jun. 2018.

[41] M. Schuld, I. Sinayskiy, and F. Petruccione, "An introduction to quantum machine learning," *Contemp. Phys.*, vol. 56, pp. 172–185, Sep. 2015.

[42] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, "Variational quantum algorithms," *Nat. Rev. Phys.*, vol. 3, no. 9, pp. 625–644, Aug. 2021.

[43] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor," *Nat. Commun.*, vol. 5, no. 1, p. 4213, Jul. 2014.

[44] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," *arXiv:1411.4028*, Nov 2014.

[45] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, "The variational quantum eigensolver: A review of methods and best practices," *Phys. Rep.-Rev. Sec. Phys. Lett.*, vol. 986, pp. 1–128, Nov. 2022.

[46] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L. C. Kwek, and A. Aspuru-Guzik, "Noisy intermediate-scale quantum (NISQ) algorithms," *Rev. Mod. Phys.*, vol. 94, no. 1, p. 015004, Feb. 2022.

[47] E. Farhi and H. Neven, "Classification with quantum neural networks on near-term processors," *arXiv:1802.06002*, Feb. 2018.

[48] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, "Parameterized quantum circuits as machine learning models," *Quantum Sci. Technol.*, vol. 4, no. 4, p. 043001, Nov. 2019.

[49] I. Cong, S. Choi, and M. D. Lukin, "Quantum convolutional neural networks," *Nat. Phys.*, vol. 15, no. 12, pp. 1273–1278, Aug. 2019.

[50] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, "Cost-function dependent barren plateaus in shallow parameterized quantum circuits," *Nat. Commun.*, vol. 12, p. 1791, Mar. 2021.

[51] M. Schuld, R. Sweke, and J. J. Meyer, "Effect of data encoding on the expressive power of variational quantum-machine-learning models," *Phys. Rev. A*, vol. 103, no. 3, p. 032430, Mar. 2021.

[52] H.-Y. Huang, M. Broughton, M. Mohseni, R. Babbush, S. Boixo, H. Neven, and J. R. McClean, "Power of data in quantum machine learning," *Nat. Commun.*, vol. 12, no. 1, p. 2631, May 2021.

[53] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, "Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms," *Adv. Quantum Technol.*, vol. 2, no. 12, p. 1900070, Dec. 2019.

[54] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, "Connecting ansatz expressibility to gradient magnitudes and barren plateaus," *PRX Quantum*, vol. 3, no. 1, p. 010313, Jan. 2022.

[55] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," *Nat. Commun.*, vol. 9, no. 1, p. 4812, Nov. 2018.

[56] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, "Noise-induced barren plateaus in variational quantum algorithms," *Nat. Commun.*, vol. 12, no. 1, p. 696, Mar. 2021.

[57] A. Pesah, M. Cerezo, S. Wang, T. Volkoff, A. T. Sornborger, and P. J. Coles, "Absence of barren plateaus in quantum convolutional neural networks," *PRX Quantum*, vol. 11, no. 4, p. 041011, Oct. 2021.

[58] Y. Du, M.-H. Hsieh, T. Liu, and D. Tao, "Expressive power of parameterized quantum circuits," *Phys. Rev. Research*, vol. 2, no. 3, p. 033125, Jul. 2020.

[59] E. Grant, L. Wossnig, M. Ostaszewski, and M. Benedetti, "An initialization strategy for addressing barren plateaus in parameterized quantum circuits," *Quantum*, vol. 3, p. 214, Dec. 2019.

[60] R. Puig, M. Drudis, S. Thanasilp, and Z. Holmes, "Variational quantum simulation: a case study for understanding warm starts," *PRX Quantum*, vol. 6, no. 1, p. 010317, Jan. 2025.

[61] V. Heyraud, Z. Li, K. Donatella, A. L. Boité, and C. Ciuti, "Efficient estimation of trainability for variational quantum circuits," *PRX Quantum*, vol. 4, no. 4, p. 040335, Dec. 2023.

[62] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," *Nature*, vol. 549, no. 7671, pp. 242–246, Sep. 2017.

[63] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware," *Phys. Rev. A*, vol. 99, no. 3, p. 032331, Mar. 2019.

[64] J. Stokes, J. Izaac, N. Killoran, and G. Carleo, "Quantum natural gradient," *Quantum*, vol. 4, p. 269, May 2020.

[65] K. Temme, S. Bravyi, and J. M. Gambetta, "Error mitigation for short-depth quantum circuits," *Phys. Rev. Lett.*, vol. 119, no. 18, p. 180509, Nov. 2017.

[66] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, "Extending the computational reach of a noisy superconducting quantum processor," *arXiv:1805.04492*, May 2018.

[67] H.-Y. Huang, M. Broughton, J. Cotler, S. Chen, J. Li, M. Mohseni, H. Neven, R. Babbush, R. Kueng, J. Preskill, and J. R. McClean, "Quantum advantage in learning from experiments," *Science*, vol. 376, no. 6598, pp. 1182–1186, Jun. 2022.

[68] G. Cho and D. Kim, "Machine learning on quantum experimental data toward solving quantum many-body problems," *Nat. Commun.*, vol. 15, no. 1, p. 7552, Aug. 2024.

[69] K. Yogaraj, B. Quanz, T. Vikas, A. Mondal, and S. Mondal, "Post-variational classical quantum transfer learning for binary classification," *Sci. Rep.*, vol. 15, no. 1, p. 23682, Jul. 2025.

[70] V. Havlíček, A. D. Córcoles, K. Temme, A. W. Harrow, A. Kandala, J. M. Chow, and J. M. Gambetta, "Supervised learning with quantum-enhanced feature spaces," *Nature*, vol. 567, no. 7747, pp. 209–212, Mar. 2019.

[71] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," *Phys. Rev. A*, vol. 98, no. 3, p. 032309, Sep. 2018.

[72] S. Jerbi, L. J. Fiderer, H. P. Nautrup, J. M. Kübler, H. J. Briegel, and V. Dunjko, "Quantum machine learning beyond kernel methods," *Nat. Commun.*, vol. 14, no. 1, p. 517, Jan. 2023.

[73] S. Thanasilp, S. Wang, M. Cerezo, and Z. Holmes, "Exponential concentration in quantum kernel methods," *Nat. Commun.*, vol. 15, no. 1, p. 5200, Jun. 2024.

[74] A. Pérez-Salinas, A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre, "Data re-uploading for a universal quantum classifier," *Quantum*, vol. 4, p. 226, Feb. 2020.

[75] T. Ono, W. Roga, K. Wakui, M. Fujiwara, S. Miki, H. Terai, and M. Takeoka, "Demonstration of a bosonic quantum classifier with data reuploading," *Phys. Rev. Lett.*, vol. 131, no. 1, p. 013601, Jul. 2023.

[76] S. Y.-C. Chen, T.-C. Wei, C. Zhang, H. Yu, and S. Yoo, "Quantum convolutional neural networks for high energy physics data analysis," *Phys. Rev. Research*, vol. 4, no. 1, p. 013231, Mar. 2022.

[77] Y. Song, J. Li, Y. Wu, S. Qin, Q. Wen, and F. Gao, "A resource-efficient quantum convolutional neural network," *Front. Phys.*, vol. 12, p. 1362690, Apr. 2024.

[78] I. Kerenidis, N. Mathur, J. Landman, M. Strahm, and Y. Y. Li, "Quantum vision transformers," *Quantum*, vol. 8, p. 1265, Feb. 2024.

[79] M. Schuld and F. Petruccione, *Machine learning with quantum computers*, 2nd ed. Berlin, Germany.: Springer, 2021.

[80] B. Narottama and S. Y. Shin, "Quantum neural networks for resource allocation in wireless communications," *IEEE Trans. Wireless Commun.*, vol. 21, no. 2, pp. 1103–1116, Feb. 2022.

[81] A. Paul, K. Singh, A. Kaushik, C.-P. Li, O. A. Dobre, T. Q. Duong, and M. D. Renzo, "Quantum-enhanced DRL optimization for DoA estimation and task offloading in ISAC systems," *IEEE J. Sel. Areas Commun.*, vol. 43, no. 1, pp. 364–381, Jan. 2025.

[82] Z. Liu, P. Ma, Y. Wang, W. Matusik, and M. Tegmark, "KAN 2.0: Kolmogorov-Arnold networks meet science," *arXiv:2408.10205*, Aug. 2024.

[83] Q. Qiu, T. Zhu, H. Gong, L. L. Chen, and H. Ning, "ReLU-KAN: New Kolmogorov-Arnold networks that only need matrix addition, dot multiplication, and ReLU," *arXiv:2406.02075*, Jun. 2024.

[84] X. Warin, "P1-KAN: An Effective Kolmogorov-Arnold network with application to hydraulic valley optimization," *arXiv:2410.03801*, Oct. 2024.

[85] H. T. Ta, "BSRBF-KAN: A combination of B-splines and radial basic functions in Kolmogorov-Arnold networks," *arXiv:2406.11173*, Jun. 2024.

[86] M. E. Samadi, Y. Müller, and A. Schuppert, "Smooth Kolmogorov Arnold networks enabling structural knowledge representation," *arXiv:2405.11318*, May 2024.

[87] E. Reinhardt, D. Ramakrishnan, and S. Gleyzer, "SineKAN: Kolmogorov-Arnold networks using sinusoidal activation functions," *Front. Artif. Intell.*, vol. 7, p. 1462952, Jan. 2025.

[88] S. SS, K. AR, and A. KP, “Chebyshev Polynomial-Based Kolmogorov-Arnold networks: An efficient architecture for nonlinear function approximation,” *arXiv:2405.07200*, May 2024.
[89] B. C. Koenig, S. Kim, and S. Deng, “KAN-ODEs: Kolmogorov–Arnold network ordinary differential equations for learning dynamical systems and hidden physics,” *Comput. Meth. Appl. Mech. Eng.*, vol. 432, p. 117397, Dec. 2024.
[90] K. Shukla, J. D. Toscano, Z. Wang, Z. Zou, and G. E. Karniadakis, “A comprehensive and FAIR comparison between MLP and KAN representations for differential equations and operator networks,” *Comput. Meth. Appl. Mech. Eng.*, vol. 431, p. 117290, Nov. 2024.
[91] F. Pourkamali-Anaraki, “Kolmogorov-Arnold networks in low-data regimes: A comparative study with multilayer perceptrons,” *arXiv:2409.10463*, Sep. 2024.
[92] A. M. Sasse and C. M. de Farias, “Evaluating federated Kolmogorov-Arnold networks on non-IID data,” *arXiv:2410.08961*, Oct. 2024.
[93] I. Drokin, “Kolmogorov-Arnold convolutions: Design principles and empirical studies,” *arXiv:2407.01092*, Jul. 2024.
[94] C. Li, X. Liu, W. Li, C. Wang, H. Liu, Y. Liu, Z. Chen, and Y. Yuan, “U-KAN makes strong backbone for medical image segmentation and generation,” in *Proc. of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 5, Philadelphia, PA, USA, Nov. 2025, pp. 4652–4660.
[95] A. D. Bodner, A. S. Tepsich, J. N. Spolski, and S. Pourteau, “Convolutional Kolmogorov-Arnold networks,” *arXiv:2406.13155*, Jun. 2024.
[96] R. C. Yu, S. Wu, and J. Gui, “Residual Kolmogorov-Arnold network for enhanced deep learning,” *arXiv:2410.05500*, Oct. 2024.
[97] M. Kiamari, M. Kiamari, and B. Krishnamachari, “GKAN: Graph Kolmogorov-Arnold networks,” *arXiv:2406.06470*, Jun. 2024.
[98] Q. Zou, L. Ni, T. Zhang, and Q. Wang, “Deep learning based feature selection for remote sensing scene classification,” *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2321–2325, Sep. 2015.
[99] A. P. Wack, H. Paik, A. Javadi-Abhari, P. Jurcevic, I. Faro, J. M. Gambetta, and B. R. Johnson, “Quality, speed, and scale: three key attributes to measure the performance of near-term quantum computers,” *arXiv:2110.14108*, Oct. 2021.

**Usama Inam Paracha** received his B.S. degree in Computer Science from National University of Sciences and Technology (NUST), Islamabad, Pakistan, in 2021. Currently, he is pursuing his Ph.D. with the Department of Electronics and Information Convergence Engineering, Kyung Hee University, South Korea. His research interests include quantum computing, quantum information science and quantum machine learning.

**Uman Khalid** received his B.S. degree in Electronics Engineering from the Ghulam Ishaq Khan (GIK) Institute, Topi, Pakistan, in 2015, and his Ph.D. in Electronics Engineering from Kyung Hee University, South Korea, in Feb. 2023. Currently, he is a Research Professor with the Department of Electronics and Information Convergence Engineering, Kyung Hee University, South Korea. Before joining as a Research Professor, he served as a Postdoctoral Fellow with the same department from Mar. 2023 to Feb. 2026. His research interests include quantum information science, quantum semantic communications, and quantum networks.

**Syed Muhammad Abuzar Rizvi** received his B.S. degree in Electrical Engineering from the National University of Sciences and Technology (NUST), Islamabad, Pakistan, in 2018. He is currently pursuing the Ph.D. degree with the Department of Electronics and Information Convergence Engineering, Kyung Hee University, South Korea. His research interests include quantum computing, quantum machine learning, quantum communication, and quantum information science.

**Trung Q. Duong** (Fellow, IEEE) is currently a Canada Excellence Research Chair (CERC) and a Full Professor with Memorial University, St. John’s, NL, Canada. He is also an Adjunct Professor with Queen’s University Belfast, Belfast, U.K., and Kyung Hee University, Yongin-si, South Korea. From 2017 to 2019, he was a Distinguished Advisory Professor with Inje University, Gimhae, South Korea. His research interests include wireless communications, quantum machine learning, and quantum optimization.

Dr. Duong was an Editor/Guest Editor of the IEEE Transactions on Wireless Communications, IEEE Transactions on Communications, IEEE Transactions on Vehicular Technology, IEEE Transactions on Network Science and Engineering, IEEE Communications Letters, IEEE Wireless Communications Letters, IEEE Wireless Communications, IEEE Communications Magazines, IEEE Network, and IEEE Journal on Selected Areas in Communications. He is the Editor-in-Chief of IEEE Communications Surveys & Tutorials and an IEEE ComSoc Distinguished Lecturer. He was the recipient of the Best Paper Award at the IEEE VTC-Spring 2013, IEEE ICC 2014, IEEE GLOBECOM 2016, 2019, 2022, IEEE DSP 2017, IWCMC 2019, 2023, and 2024, and IEEE CAMAD 2023, 2024. He received two prestigious awards from the Royal Academy of Engineering (RAEng): RAEng Research Chair and the RAEng Research Fellow, and the prestigious Newton Prize 2017. He is also a fellow of the Engineering Institute of Canada (EIC), Canadian Academy of Engineering (CAE), Institution of Engineering and Technology (IET), and Asia-Pacific Artificial Intelligence Association (AAIA).

**Moe Z. Win** (Fellow, IEEE) is a Professor at the Massachusetts Institute of Technology (MIT) and the founding director of the Wireless Information and Network Sciences Laboratory. Prior to joining MIT, he was with AT&T Research Laboratories and with NASA Jet Propulsion Laboratory.

His research encompasses fundamental theories, algorithm design, and network experimentation for a broad range of real-world problems. His current research topics include ultra-wideband systems, network localization and navigation, network interference exploitation, and quantum information science. He has served the IEEE Communications Society as an elected Member-at-Large on the Board of Governors, as elected Chair of the Radio Communications Committee, and as an IEEE Distinguished Lecturer. Over the last two decades, he held various editorial positions for IEEE journals and organized numerous international conferences. Recently, he has served on the SIAM Diversity Advisory Committee.

Dr. Win is an elected Fellow of the AAAS, the EURASIP, the IEEE, and the IET. He was honored with two IEEE Technical Field Awards: the IEEE Kiyo Tomiyasu Award (2011) and the IEEE Eric E. Sumner Award (2006, jointly with R. A. Scholtz). His publications, co-authored with students and colleagues, have received several awards. Other recognitions include the MIT Frank E. Perkins Award (2024), the MIT Everett Moore Baker Award (2022), the IEEE Vehicular Technology Society James Evans Avant Garde Award (2022), the IEEE Communications Society Edwin H. Armstrong Achievement Award (2016), the Cristoforo Colombo International Prize for Communications (2013), the Copernicus Fellowship (2011) and the *Laurea Honoris Causa* from the Università degli Studi di Ferrara (2008), and the U.S. Presidential Early Career Award for Scientists and Engineers (2004). He is an ISI Highly Cited Researcher.

**Hyundong Shin** (Fellow, IEEE) received the B.S. degree in Electronics Engineering from Kyung Hee University (KHU), Yongin-si, Korea, in 1999, and the M.S. and Ph.D. degrees in Electrical Engineering from Seoul National University, Seoul, Korea, in 2001 and 2004, respectively. During his postdoctoral research at the Massachusetts Institute of Technology (MIT) from 2004 to 2006, he was with the Laboratory for Information Decision Systems (LIDS). In 2006, he joined the KHU, where he is now the Dean of the College of Electronics & Information as well as the College of Software. His research interests include quantum information science, wireless communication, and machine intelligence. Dr. Shin received the IEEE Communications Society’s Guglielmo Marconi Prize Paper Award and William R. Bennett Prize Paper Award. He served as the Publicity Co-Chair for the IEEE PIMRC and the Technical Program Co-Chair for the IEEE WCNC and the IEEE GLOBECOM. He was an Editor of IEEE Transactions on Wireless Communications and IEEE Communications Letters.