Review

Not peer-reviewed version

A Systematic Review of Quantum Deep Learning Architectures for Image Classification

Saleha Daraan * and Wael Alghamdi

Posted Date: 12 June 2026

doi: 10.20944/preprints202606.1015.v1

Keywords: QDL; image classification; QML; hybrid quantum-classical models; QNNs; QViT; quantum attention mechanism

Review

A Systematic Review of Quantum Deep Learning Architectures for Image Classification

Saleha Daraan * and Wael Alghamdi

Taif University
* Correspondence: s44680610@students.tu.edu.sa

Abstract

Quantum deep learning (QDL) for image classification has emerged as a rapidly growing research area at the intersection of quantum computing and computer vision. However, the literature remains fragmented across diverse model architectures, datasets, and evaluation settings, making it difficult to obtain a focused understanding of the field. This study presents a systematic review of QDL models for image classification published between 2020 and 2025. The review was conducted using a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)-based selection process supported by a mapping perspective. Relevant studies were identified from major academic databases, screened using predefined inclusion and exclusion criteria, and analyzed through a taxonomy-based framework. A total of 47 studies were included and organized into four main architectural categories: Quantum Neural Networks (QNNs), Quantum Convolutional Neural Networks (QCNNs), Quantum Vision Transformers (QViT), and Quantum Attention Mechanisms. The findings show that most existing approaches rely on hybrid quantum-classical designs, where quantum components are integrated selectively into specific stages of the learning pipeline rather than replacing the full classical architecture. Several studies report encouraging results, such as improved parameter efficiency, competitive ACC (accuracy), and robustness in specific settings. Current progress remains constrained by hardware noise, limited qubit availability, encoding overhead, scalability challenges, and inconsistent benchmarking practices. It remains a promising but still emerging direction, with future progress expected to depend on more scalable architectures, fairer evaluation protocols, and broader validation on realistic application.

Keywords: QDL; image classification; QML; hybrid quantum-classical models; QNNs; QViT; quantum attention mechanism

---

1. Introduction

QDL has recently emerged as an active research area. Among its various application tasks, image classification provides a particularly suitable focus for review because it is a well-established and widely studied problem with strong relevance across multiple domains, while also offering more consistent benchmarks for comparing different model architectures[1,2]. In this context, this review presents a structured analysis of QDL models for image classification.

The motivation for this review is driven by the rapid growth of research at the intersection of quantum computing and visual deep learning, accompanied by increasing diversity in model architectures, evaluation settings, and application domains. Although this expansion reflects strong scientific interest, it also makes the literature increasingly difficult to interpret in a coherent way, especially when studies differ substantially in their quantum components, datasets, and experimental assumptions.

The problem addressed in this review lies in the rapidly expanding scope of QDL, where research is increasingly dispersed across multiple architectures, tasks, and application domains. As a result, the field has become broad and fast-evolving, making it difficult to obtain a focused understanding of specific directions within it.

 doi:10.20944/preprints202606.1015.v1

To guide the review process, a set of research questions (RQs) and mapping questions (MQs) was formulated. While the MQs provide a structured overview of the selected studies, the RQs focus on deeper analytical issues related to model design, practical value, and current limitations in QDL for image classification.

- **RQ1:** How is the quantum component integrated into image classification models in the reviewed studies?
- **RQ2:** Do current QDL models demonstrate a real practical quantum advantage in image classification?
- **RQ3:** What are the main limitations and open challenges reported in the selected studies?

- **MQ1:** What are the selected studies, and what selection strategy was followed in this review?
- **MQ2:** What are the main model architectures investigated in the selected studies?
- **MQ3:** How are the selected studies distributed across the proposed taxonomy categories?
- **MQ4:** What are the most frequently used datasets in the selected studies, and what are their main characteristics?

This study aim to provide a structured and focused analysis of QDL research for image classification by systematically identifying, classifying, and examining the selected studies. In particular, the review aims to organize the literature through a taxonomy-based perspective in order to clarify the main model architectures, research trends, and current directions within this evolving area.

Table 4 presents a comparative overview of the existing review studies in relation to the present review. As shown, prior surveys have largely addressed quantum machine learning (QML) from broader or domain-specific perspectives, while limited attention has been given to a focused review of QDL for image classification. In contrast, the present study is designed as a systematic review with a mapping perspective and concentrates specifically on QDL models for image classification over the 2020–2025 period. Moreover, this study combines a PRISMA-based selection strategy with a taxonomy-based analysis and a dedicated dataset analysis, thereby providing a more structured and focused synthesis of the field. To address the current need for a focused and systematic understanding of this rapidly evolving area, this review makes four main contributions to the field of QDL for image classification:

- Provides a focused systematic review of recent studies on QDL for image classification, covering the period from 2020 to 2025.
- Organizes the reviewed literature through a taxonomy-based perspective, offering a clearer and more structured view of the main architectural categories explored in this area.
- Presents a systematic mapping of the selected studies in terms of model architectures, study distribution, and the most frequently used datasets and their characteristics.
- Critically examines the integration of quantum components, the reported practical advantages, and the main limitations and open challenges that currently shape this research area.

The remainder of this manuscript is organized as follows. Section 2 presents the background of the study, Section 3 outlines the methodology, and Section 4 provides the taxonomy-based analysis. Section 5 discusses the main findings, limitations, and future directions, while Section 6 concludes the manuscript. Figure 1 illustrates the hierarchical structure of the entire manuscript.

Table 1. Summary of QML Review Papers.

Ref./ Year	Review Type	Review Scope	Time Span	No. of included studies	PRISMA Based	Taxonomy Based Analysis	Datasets Reported Analysis
[2], 2025	Systematic	General QML and its applications	2023–2025	30	✓	✓	×

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2026 doi:10.20944/preprints202606.1015.v1

Table 1. *Cont.*

Ref./ Year	Review Type	Review Scope	Time Span	No. of included studies	PRISMA Based	Taxonomy Based Analysis	Datasets Reported Analysis
[3], 2025	Systematic	QML applications in real-world classification	2013–2023	23	✓	✓	×
[4], 2025	Review / tutorial-style review	Fundamentals, algorithms, techniques, and real-world applications of QML/QDL	N/A	N/A	×	×	×
[5], 2024	Systematic	QML/QDL in Healthcare	2018–2023	49	✓	×	✓
[6], 2022	Systematic	QML in Biomedical	2013–2021	30	✓	×	✓
[7], 2025	Systematic	QML for Digital Health	2015–2024	46	✓	×	✓
[8], 2025	Scoping Survey	QML and QDL for real-world applications	N/A	N/A	×	×	✓
[9], 2025	Comprehensive tutorial and survey	QML for integrated networks	N/A	N/A	×	✓	✓
[10], 2024	Survey	QML for Computer Vision	N/A	N/A	×	×	✓
[11], 2023	Survey	QML/ QDL for Image Classification	2015–2022	27	×	✓	✓
This Review, 2026	**Systematic**	**QDL for Image Classification**	**2020–2025**	**47**	✓	✓	✓

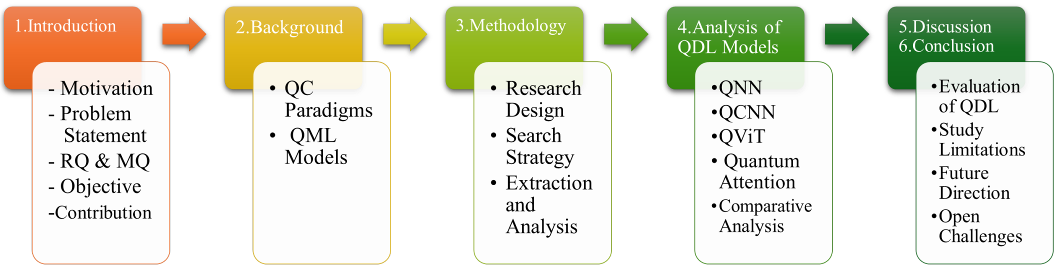


Figure 1. Hierarchical overview of the entire manuscript.

2. Background of QML

This section provides the theoretical background necessary for this study. It covers the paradigms of quantum computing, and the integration of quantum computing with machine learning. It also introduces parameterized and variational quantum circuits (VQCs), in addition to traditional QML models, which together form the basis for the advanced models discussed later in this research. Readers with a strong background in these areas can easily skip this section.

2.1. Gate-Based Quantum Computing

The most common and widespread form of quantum computing is the quantum equivalent of classical computing based on logic circuits. In this Paradigm, information is represented using qubits and is processed through an ordered sequence of quantum gates, which are reversible unitary operations that control the evolution of the quantum state [12]. This paradigms also relies on quantum circuits designed to perform specific algorithms such as:

- Shor's factorization algorithm,
- Grover's search algorithm,

This approach is programmable using languages such as Qiskit, Cirq, and PennyLane, and it forms the basis of most current Noisy Intermediate-Scale Quantum (NISQ) platforms.

2.2. *Adiabatic Quantum Computing (AQC)*

It represents a fundamentally different paradigm of quantum computation compared to the gate-based paradigm. While gate-model quantum computing relies on a discrete sequence of unitary operations applied to qubits, AQC is based on the continuous and gradual evolution of a quantum system under a time-dependent Hamiltonian[1]. The core principle behind AQC is the adiabatic theorem, which states that a quantum system initialized in the ground state of a Hamiltonian H_0 will remain in the instantaneous ground state of a smoothly varying Hamiltonian $H(t)$, provided the evolution occurs sufficiently slowly [1].

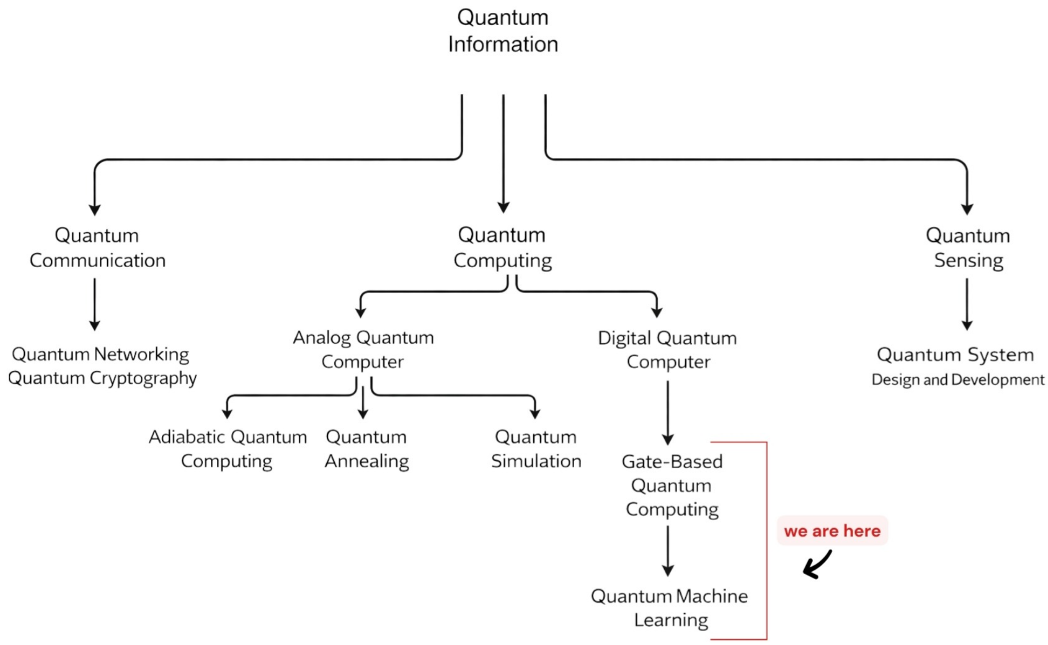


Figure 2. Overview of the quantum information landscape, illustrating its main domains : quantum communication, quantum computing, and quantum sensing. With a focus on gate-based quantum computing as the foundation for QML. Adapted from [13].

Quantum hardware development has witnessed significant progress in recent years, driven by advances in different physical implementations such as superconducting qubits, trapped ions, and quantum annealing systems[14]. Major technology companies and research institutions, including IBM, Google, Rigetti, and Quantinuum, have developed increasingly sophisticated quantum processors with growing qubit counts and improved performance. These hardware platforms form the foundation for executing quantum algorithms and enabling practical quantum applications. Table 2 presents representative examples of current quantum hardware systems, highlighting their corresponding technologies and qubit capacities.

Table 2. Examples of leading quantum hardware platforms, highlighting quantum processors, qubit counts, and underlying technologies across different paradigms.

Quantum Processor	Organization	Qubit Count	Technology
IBM Condor[15]	IBM Quantum	1,121 qubits	Superconducting
IBM Heron[16]	IBM Quantum	156 qubits	Superconducting
IBM Eagle[17]	IBM Quantum	127 qubits	Superconducting
Google Willow[18]	Google Quantum AI	105 qubits	Superconducting
Quantinuum H2[19]	Quantinuum	56 qubits	Trapped-Ion
Rigetti Aspen-M[20]	Rigetti Computing	~80 qubits	Superconducting
D-Wave Advantage2[21]	D-Wave	>5000 qubits	Quantum Annealing

doi:10.20944/preprints202606.1015.v1

As quantum hardware matured, a parallel transformation occurred in quantum software ecosystems. Modern quantum programming frameworks aim to abstract low-level quantum operations and provide high-level interfaces for algorithm development, simulation, and benchmarking. Table 3 provides a concise overview of the major quantum programming frameworks, including their developers, primary programming languages, and supported hardware backends.

Table 3. Major quantum programming frameworks, their developers, primary programming languages, and hardware support.

Framework	Developer	Primary Language	Hardware Support
Qiskit[22]	IBM	Python	IBM Quantum processors
Cirq[23]	Google	Python	Google Quantum AI hardware
PyQuil[24]	Rigetti	Python	Rigetti QPUs via cloud
Q#[25]	Microsoft	Q#	Azure Quantum ecosystem

2.3. Integration of Quantum Computing and Machine Learning

Quantum Artificial Intelligence (QAI) is the intersection of quantum computing and artificial intelligence. It is a field that examines how the rules of quantum mechanics can enhance intelligent systems and how AI can accelerate quantum technologies [2,3]. Quantum computing offers fundamentally new ways to perform calculations through superposition, entanglement, and quantum parallelism. These new ways could, "at least in theory," speed up several types of AI activities, especially optimization, machine learning, and reasoning. The main subfields emerging at the crossroads of quantum computing and AI as shown in Figure 3. This helps to put QAI in context. The graphic shows that QAI includes several study topics, each of which combines quantum concepts with known AI fields. The integration of quantum computing and machine learning has become one of the most exciting areas of research in QAI. The goal of this convergence is to bring together the expressive learning capabilities of classical models with the computing benefits of quantum systems, especially in high-dimensional representation, sampling, and optimization. Recent studies have shown that the two domains are connected in both directions: quantum processors can accelerate learning, and machine learning methods can support quantum circuit planning, noise modeling, and device calibration[4,26].

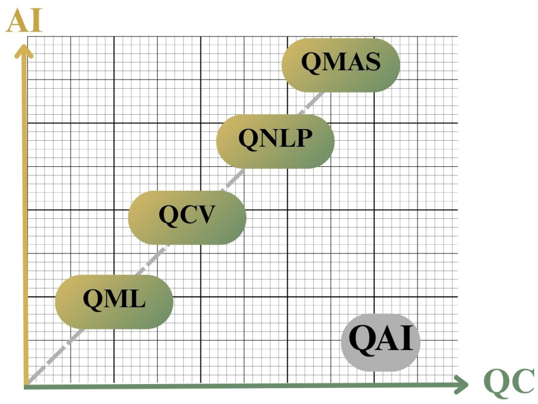


Figure 3. Interaction between Artificial Intelligence and Quantum Computing, highlighting major areas such as QML, Quantum Computer Vision, Quantum Natural Language Processing and Quantum Multi-Agent Systems.

Quantum data encoding is one of the most critical and sensitive stages in quantum learning models, as it represents the bridge between raw data, whether classical or quantum, and the Hilbert space in which quantum algorithms operate. Although it may seem like a simple step, choosing the right encoding scheme is a fundamental and well-known problem in QML. We can see in Figure 4 four scenarios between Data Type and Algorithm Type. It clearly outlines the space in which a researcher can work. This description illustrates that coding becomes necessary only when dealing with classical data using a quantitative model, the most prevalent scenario in QML today. This is where

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2026 doi:10.20944/preprints202606.1015.v1

the challenges of encoding become most apparent[8,27]. The rapid development of QML has produced a diverse set of algorithmic models. These models can be grouped into several core categories, the following subsections provide a concise overview of these key frameworks.

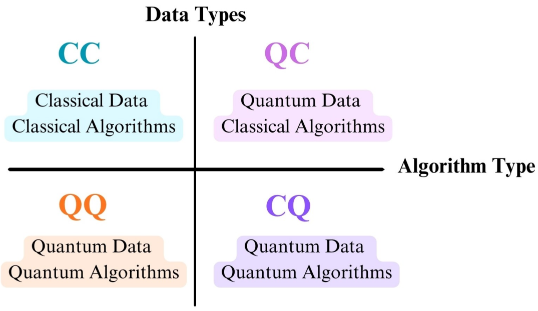


Figure 4. QML paradigms based on Data Type and Algorithm Type, illustrating hybrid approaches that combine classical and quantum data and algorithms.

2.3.1. Parameterized and VQCs for Data Encoding and Learning Process

Parameterized Quantum Circuits (PQCs) are among the most commonly used models in today's QML landscape, especially within NISQ-era hybrid systems. They provide a structured way to combine classical data with quantum operations by introducing trainable parameters inside a quantum circuit [11]. One of the main advantages of PQCs is their ability to model rich and expressive representations that may be difficult to capture using classical neural architectures. However, they also come with practical challenges, such as the risk of barren plateaus (where gradients vanish) and the need to carefully control the depth of the circuit to avoid noise accumulation on real quantum hardware [3]. Regarding to VQCs it contains trainable parameters, and these parameters are optimized using a classical algorithm. VQCs are flexible and can approximate complex functions, making them suitable for tasks such as classification, regression, and variational eigensolvers. However, they also face challenges like noise sensitivity, barren plateaus (vanishing gradients), and the need to carefully balance circuit depth with hardware limitations[3]. As illustrated in Figure 5, PQCs and VQCs share a common foundation, where both are constructed using parameterized quantum gates. However, VQCs extend PQCs by incorporating a classical optimization loop that updates the circuit parameters iteratively to minimize a cost function.

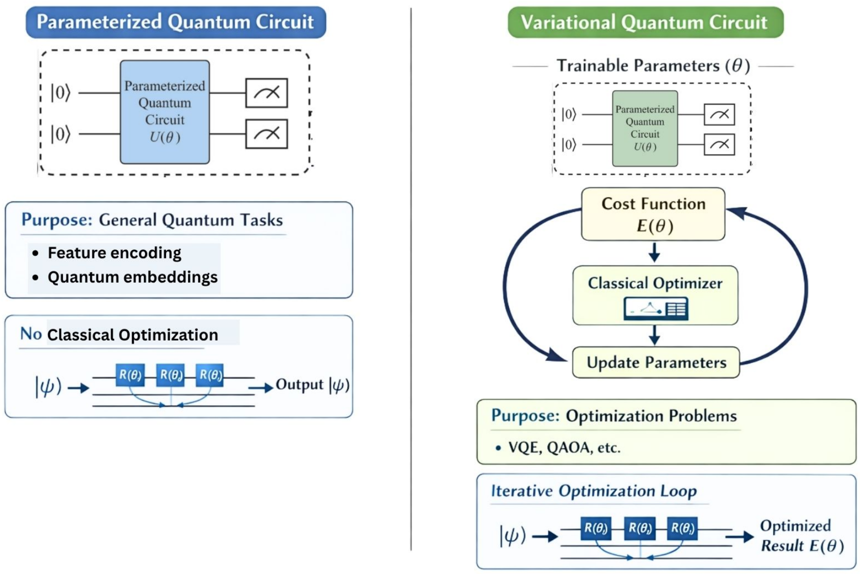


Figure 5. The relationship between PQCs and VQCs, where VQCs build upon PQCs by introducing trainable parameters optimized through a classical feedback loop.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2026 doi:10.20944/preprints202606.1015.v1

2.3.2. Main Algorithms in QML

This section provides a brief overview of selected traditional QML methods. As shown in Table 4, it summarizes representative studies on QSVM, QKNN, QRF, Q-means, and QPCA with respect to their main ideas, key results, and limitations. To reflect recent developments in these methods, 11 papers published since 2019 were selected, illustrating how continuous refinements and methodological improvements have led to noticeable progress in their reported results and practical effectiveness.

Table 4. Summary of representative QML methods, highlighting core descriptions, key results, and limitations.

Method	Ref	Year	Description	Key Result	Limitations
Quantum Vector Machine	[28]	2019	Proposes QSVM tailored to NISQ hardware with optimized preprocessing, tomography-free kernel oracle, and HHL-based classifier.	High ACC (≈98–99.5% OCR, ≈97–98% Iris), reduced circuit depth.	Limited to linearly separable data; small training sets.
	[29]	2021	Quantum kernel estimation mapping data into high-dimensional Hilbert space.	∼99% ACC, no QRAM, robust to noise.	Requires fault-tolerant QC; synthetic datasets only; high sampling cost.
	[30]	2024	Anti-noise QSVM using weighted hinge loss and iterative quantum optimization.	100% ACC on noisy digit datasets (vs 87.5% classical).	Deep circuits; complex optimization; limited scalability.
Quantum k-Nearest Neighbor (QKNN)	[31]	2024	Quantum k-NN using Euclidean distance estimation with Bell-H circuit and amplitude encoding.	Complexity $O(\log d + N)$ with QRAM; classical-equivalent performance in simulation.	Performance drops with limited shots; depends on QRAM.
	[32]	2024	Benchmark of quantum k-NN using amplitude-encoded state overlap metric.	≥50% qubit reduction; comparable or improved NISQ performance.	Simulation-based; noise-sensitive; dataset-dependent.
Quantum Random Forest	[33]	2024	Ensemble of quantum decision trees using kernel-based SVM splits via Nyström approximation.	Reduced sampling complexity $O(N^3L)$; better than QSVM.	Curse of dimensionality; depends on inductive bias.
	[34]	2022	Hybrid QRF + QSVM for COVID-19 classification.	QRF 75%, QSVM 78%, outperforming classical baselines.	Limited qubits and coherence time.
Quantum K-means	[35]	2019	Quantum clustering using distance estimation and quantum linear algebra.	Polylogarithmic scaling; ACC ≈0.891.	Sensitive embeddings; kernel vanishing with scaling.
	[36]	2024	Hybrid quantum k-means with parallel distance computation.	High similarity to classical k-means; SIL = 0.81.	High shot cost; limited advantage on NISQ devices.
Quantum PCA	[37]	2024	Variational hardware-efficient QPCA for medical image recognition.	ACC 92.31% / 87.17%; fidelity > 0.9998.	Limited by circuit depth; state preparation issues.
	[38]	2022	Low-complexity QPCA reducing circuit depth via phase estimation.	3× reduced depth vs prior methods; similar PCA performance.	Sensitive to state preparation; small matrices only.

In general, these methods demonstrate promising theoretical and experimental results. Nevertheless, recent research has increasingly moved toward QDL and transformer-based architectures, which constitute the main focus of our systematic review.

 doi:10.20944/preprints202606.1015.v1

3. Methodology

This section presents the methodology adopted in this review to systematically identify, select, and analyze the relevant studies on QDL for image classification. The review process was designed to ensure a structured and transparent investigation of the literature, covering the research design, search strategy, study selection process, and data extraction and analysis procedures. In addition, this section outlines the development of the proposed taxonomy, along with the descriptive and bibliometric analyses used to examine publication trends and research patterns in the selected studies.

3.1. Research Design

This study was designed as a systematic review supported by a mapping perspective to investigate the current research landscape of QDL for image classification. The adopted design enables a structured and transparent process for identifying, selecting, classifying, and analyzing the relevant studies in this emerging field.

This review was conducted to provide a comprehensive understanding of the existing studies in this domain, with particular attention to the employed model architectures, publication trends, datasets, and methodological characteristics. The choice of this review design was also motivated by the broad diversity of tasks, methods, model architectures, and application domains reported in the literature. Since QDL has been explored across multiple directions, this review narrows its focus to image classification as a well-established and widely studied task with strong relevance across many fields. This focused scope helps provide a more consistent and meaningful analysis of the selected studies.

The design of this review was guided by the MQs and RQs presented in the Introduction, which structured both the analytical process and the interpretation of the findings.

3.2. Research Design

To identify the relevant studies, the search process was conducted across several major academic sources that broadly cover the fields of computer science, artificial intelligence, and quantum computing. The selected sources included Scopus, Web of Science, IEEE Xplore, and SpringerLink as the primary databases, in addition to Google Scholar as a supplementary search source. Citation searching was also employed to identify potentially relevant studies that may not have been captured through the direct database search. These sources were selected to ensure broad coverage of peer-reviewed journals, conference proceedings, and preprint records relevant to the scope of this review.

The search strings were constructed using combinations of keywords related to the main concepts of this review, namely quantum computing, deep learning, and image classification. To improve the comprehensiveness of the search, multiple related terms and model-specific keywords were considered, such as QML, QNN, QCNN,QViT, and quantum attention. These terms were combined using Boolean operators such as AND and OR to connect the main concepts and their alternatives. This strategy helped capture a broad range of relevant studies while keeping the search focused on the scope of the review.

The screening and study selection process was primarily conducted by the first author through title–abstract screening followed by full-text eligibility assessment based on the predefined inclusion and exclusion criteria. To improve consistency and reduce selection bias, uncertain cases were rechecked through supervisory discussion and validated against the review objectives and taxonomy scope. The search process was carried out by applying the developed search strings across the selected databases and search sources. To maintain consistency with the scope of the review, the search was limited to studies relevant to image classification within the broader area of QDL. Additional filtering was applied where needed based on publication relevance and the alignment of the retrieved studies with the objectives of the review. In the case of Google Scholar, search filtering was used to refine the retrieved results and improve their relevance. Overall, the search process was designed to balance breadth and specificity in order to identify the most relevant studies for subsequent screening and analysis.

After the initial search stage, a structured study selection procedure was performed to refine the retrieved records and identify the final studies included in the review. This procedure consisted of duplicate removal, preliminary screening based on titles and abstracts, full-text eligibility assessment, and final study inclusion. To ensure transparency in the selection workflow, the review followed the PRISMA framework [39], and the complete selection process is presented in Figure 6.

In total, 391 records were initially identified from the selected sources. After removing 68 duplicates and 29 records for other reasons, 294 records remained for screening. Following title and abstract screening, 165 records were excluded, and 129 reports were sought for full-text retrieval. Of these, 18 reports were not retrieved, resulting in 111 studies assessed for eligibility. Finally, 47 studies were included in the review.

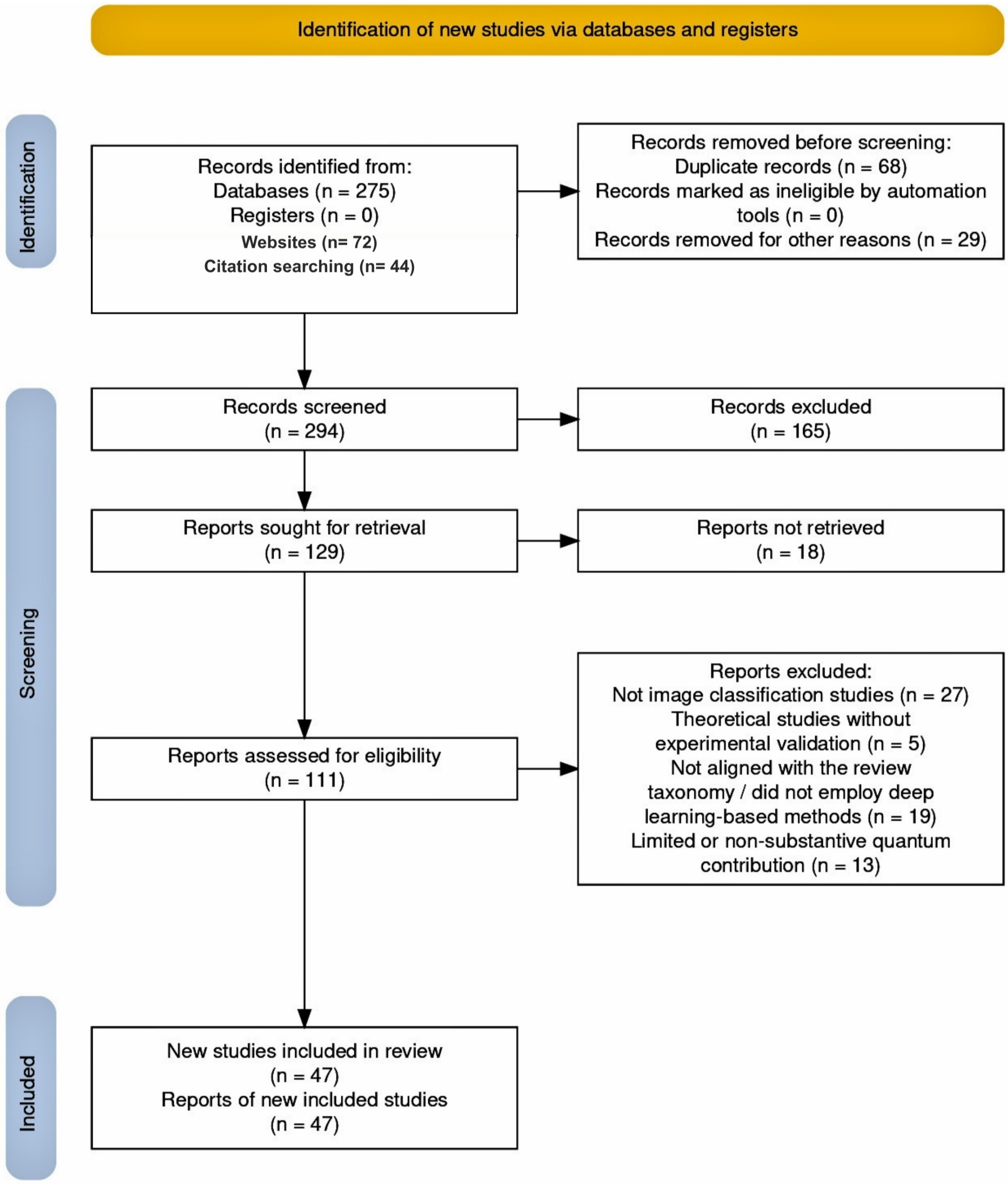


Figure 6. Flowchart of the article selection strategy based on the PRISMA style.

To ensure a rigorous and unbiased selection of studies, clearly defined inclusion and exclusion criteria were formulated and applied throughout the screening process. These criteria served to evaluate whether each retrieved study was sufficiently relevant to the focus of this review and methodologically appropriate for inclusion. More specifically, the inclusion criteria required that studies address QDL, or closely related QML models, within the context of image classification, while employing relevant architectural paradigms such as QNN, QCNN,QViTs, or quantum attention-based

 doi:10.20944/preprints202606.1015.v1

approaches. In addition, only studies providing experimental validation and reported results were retained. Conversely, studies published before 2020, lacking a substantive quantum component, written in languages other than English, or not based on deep learning methods were excluded. As summarized in Table 5. These criteria provided a structured basis for narrowing the retrieved literature to the most relevant studies. This step also contributes to addressing MQ1 by clarifying the study selection procedure adopted in this review. Overall, applying these criteria strengthened the consistency of the selection process and ensured that the final analysis was based on studies closely aligned with the review scope and quality expectations.

Table 5. Inclusion and exclusion criteria used in the study selection process.

Inclusion criteria (IC)	**Exclusion criteria (EC)**
IC-1 Studies related to QDL or QML, employing relevant model architectures such as QNN, QCNN,QViTs, or quantum attention-based models.	EC-1 Studies published before 2020.
IC-2 Studies related to image classification.	EC-2 Studies with limited or non-substantive quantum contribution.
IC-3 Studies that include experimental validation and report results.	EC-3 Studies not written in English.
	EC-4 Studies that do not employ deep learning-based methods.

The application of these criteria helped refine the selected studies and ensured that only relevant and high-quality research aligned with the scope of this review was included in the final analysis.

3.3. Data Extraction and Analysis

After the final selection of studies, data were extracted from each paper based on the aspects most relevant to the objectives of this review. Particular attention was given to the model architecture adopted in each study, the way the quantum component was integrated into the proposed framework, the specific role played by the quantum part, and whether its inclusion contributed to improving image classification performance.

The outcomes sought from each study included the reported image classification performance, the role and position of the quantum component within the model, and the claimed advantages or limitations associated with its use. When multiple results were reported within a study, the most relevant results aligned with the main image classification task and the core proposed architecture were extracted for comparison. In addition to the main outcomes, other variables were extracted from each study to support the comparative and descriptive analysis. These variables included publication year, publication type, application domain, dataset used, model category, encoding strategy, hardware or simulation setting, and the main limitations reported by the authors. When some information was unclear or not explicitly stated, it was interpreted cautiously based on the study context when possible; otherwise, it was treated as not reported.

Before synthesis and presentation, the extracted information was standardized into a consistent comparative format across studies. When studies reported multiple experiments, metrics, or model variants, the most representative results related to the main image classification task and the core proposed architecture were retained to ensure comparability.

Data extraction was performed manually by authors using a structured extraction framework covering model architecture, dataset, quantum integration strategy, application domain, and reported performance. The extracted information was cross-checked for consistency during the comparative analysis stage to ensure reliable synthesis across the selected studies.

The publication years of the selected studies were analyzed to provide a temporal overview of the literature included in this review. As shown in Figure 7, most of the selected studies were published in recent years, with the highest concentration observed in 2024 and 2025. In contrast, only a small number of studies were published between 2020 and 2022. This distribution indicates that the studies included in this review are largely concentrated in the most recent period. The distribution of the selected studies by publication type was also examined to provide further insight into the characteristics of the included literature. As illustrated in Figure 8, journal articles represent the largest proportion of the selected studies, followed by conference papers and preprint publications (e.g., arXiv). This distribution reflects the variety of publication venues associated with the studies included in this review.

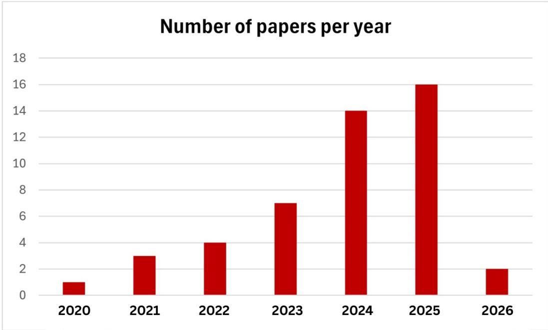


Figure 7. Distribution of the selected studies by publication year.

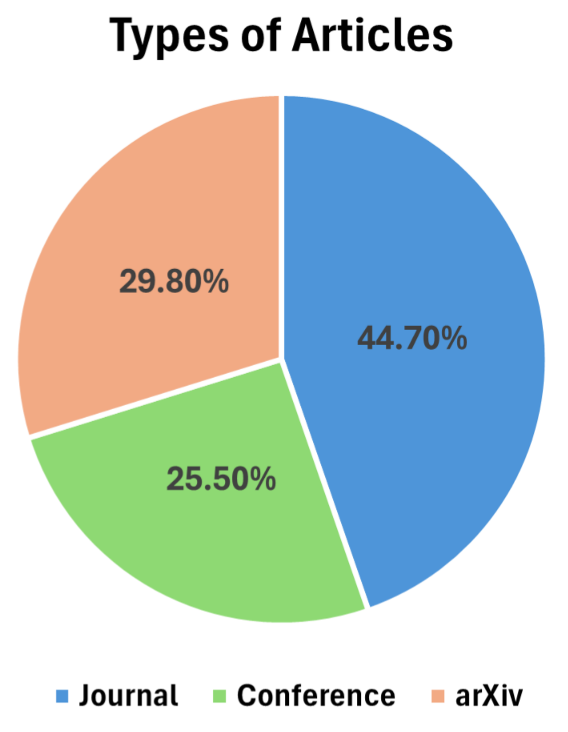


Figure 8. Types of selected published articles.

This part of the analysis examines the datasets used in the selected studies to addresses MQ4 for better understand the commonly adopted benchmarks in QDL for image classification and to addresses MQ4. As illustrated in Figure 9. MNIST is the most frequently used dataset among the selected studies, followed by MedMNIST and CIFAR-10, while other datasets such as Fashion-MNIST, Brain Magnetic Resonance Imaging (MRI), and Dogs vs Cats are used less frequently. Figure 10 presents sample images from the most commonly used datasets, providing a visual overview of their characteristics and variability. The datasets differ significantly in terms of complexity, ranging from simple grayscale images, such as MNIST, to more complex color images, such as CIFAR-10. In

 doi:10.20944/preprints202606.1015.v1

addition, they vary in their classification settings, where some are designed for binary classification tasks, while others involve multi-class classification, reflecting different levels of difficulty. The main characteristics of these datasets are summarized in Table 6 3.3.1, including their domain, size, and image resolution. The analysis shows that a large portion of the selected studies rely on relatively simple and standardized datasets, particularly MNIST, due to its low complexity and suitability for early-stage quantum experiments. In contrast, more complex and domain-specific datasets, such as medical imaging datasets, are used less frequently, although they indicate emerging efforts toward real-world applications.

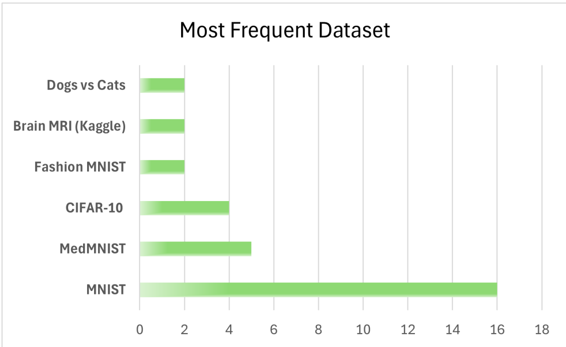


Figure 9. Most frequent datasets used in the selected studies.

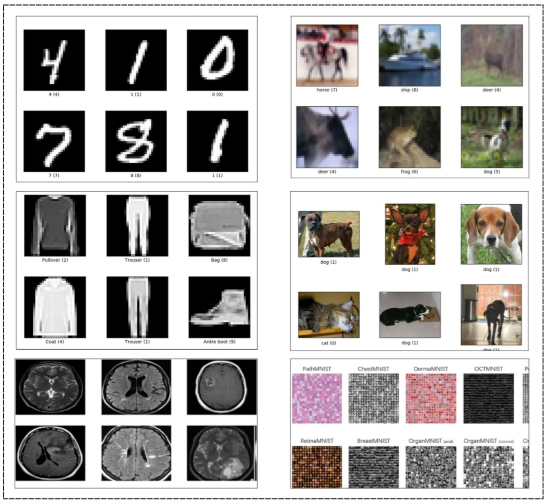


Figure 10. Sample images from commonly used datasets, illustrating variations in image complexity, data types and classification tasks.

Table 6. Common datasets used in image classification tasks.

Dataset	Type	Domain	No. of Images	Image Size
MNIST [40]	Grayscale images	Handwritten digits	70,000	28×28
MedMNIST [41]	Medical images	Healthcare	∼70,000+	28×28
CIFAR-10 [42]	RGB images	Object classification	60,000	32×32
Fashion-MNIST [43]	Grayscale images	Fashion items	70,000	28×28
Brain MRI [44]	Medical images	Healthcare	Varies	Varies
Dogs vs Cats [45]	RGB images	Binary classification	25,000	Varies

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2026 doi:10.20944/preprints202606.1015.v1

3.4. Architecture-Based Taxonomy of QDL Models for Image Classification

To further organize and analyze the selected studies, a taxonomy-based classification was developed based on the underlying model architectures. As illustrated in Figure 11, the selected studies are categorized into four main groups: QNNs, QCNNs,QViTs, and Quantum Attention-based models. The first level of the taxonomy represents the main classification of the selected studies based on their underlying model architectures, while the subsequent levels illustrate specific models and representative examples derived from the included studies.

This taxonomy provides a structured view of the different approaches adopted in QDL for image classification and highlights the diversity of model designs within the selected studies. By grouping the studies according to their architectural characteristics, the taxonomy facilitates a clearer understanding of the research landscape and enables a more systematic comparison between different methods. The distribution of studies across these categories is further illustrated in Figure 12, showing the relative prevalence of each approach. This classification also serves as the foundation for the detailed analysis presented in the following section, where each category is discussed individually.

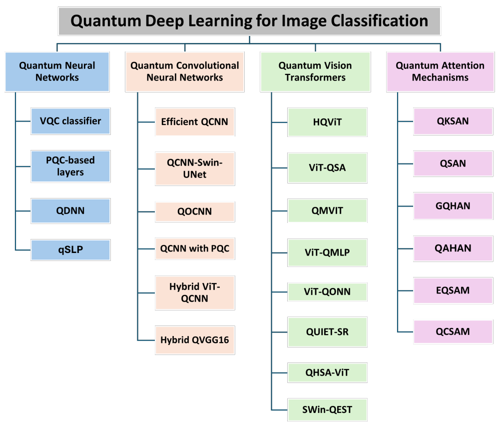


Figure 11. Proposed taxonomy of QDL methods for image classification, highlighting the main architectural categories at the top level and representative models from the selected studies at lower levels.

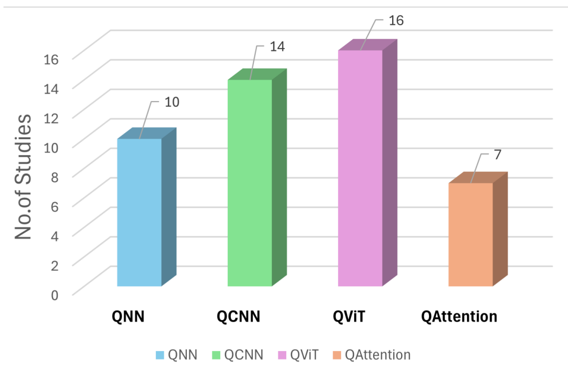


Figure 12. Distribution of the selected studies across the proposed taxonomy categories.

4. Analysis of QDL Models Based on Taxonomy

This section presents a taxonomy-based analysis of the selected studies in QDL for image classification. The proposed taxonomy serves as a structured framework to systematically organize and examine the existing literature according to their underlying model architectures. Specifically, the studies are categorized into four main groups: QNNs, QCNNs,QViTs, and quantum attention-based models. For each category, the analysis focuses on the general characteristics of the models, the nature of the proposed architectures, and the datasets and evaluation strategies adopted in the selected studies. In addition, representative works are summarized and compared to highlight similarities, differences, and emerging trends within each category. This structured analysis facilitates a clearer

 doi:10.20944/preprints202606.1015.v1

understanding of the research landscape and provides a foundation for the comparative and critical discussions presented in the subsequent sections.

4.1. Quantum Neural Network (QNN)

QNNs represent one of the earliest and most fundamental approaches in QML, aiming to extend classical neural network concepts into the quantum domain. By leveraging quantum properties such as superposition and entanglement, QNN-based models seek to enhance the representational capacity and computational efficiency of learning systems. In the context of image classification, QNNs have been widely explored through hybrid quantum-classical architectures, where PQCs are integrated within classical learning pipelines.

Early work on QNNs can be traced back to the study by Altaisky (2001)[46], where the concept of a QNN was formally introduced by combining principles of quantum information processing with classical neural network structures. In this work, the author proposed a quantum perceptron model in which input and output neurons are represented by qubits, while the network weights are implemented using optical components such as beam splitters and phase shifters. One of the key challenges highlighted in this study is the incompatibility between classical neural networks and quantum systems, particularly due to the reliance of classical models on nonlinear activation functions. However, the model was not designed for practical large-scale applications and did not consider modern deep learning architectures or real-world datasets. Recent studies on QNN-based models for image classification can be broadly categorized into hybrid architectures, theoretical frameworks, and application-specific implementations. The majority of the literature focuses on hybrid quantum-classical models, where classical neural networks are combined with PQCs to leverage the strengths of both paradigms.

Several studies aim to recover classical neural network expressivity in quantum settings, but differ in their approaches. QDNN [47] provides a more comprehensive theoretical framework, including universal approximation, backpropagation (BP), and practical NISQ considerations. In contrast, qSLP [48] focuses on architecture, showing how a single-layer perceptron can achieve exponential scaling via quantum superposition, while highlighting challenges like non-linear activation.

From the VQC perspective, learning is achieved by optimizing trainable gate parameters using a classical optimizer and a loss function. Study [49] demonstrates this through a parameterized circuit enabling effective classification with few qubits, while[48] presents a hybrid quantum-classical model where a variational circuit $U(x;\Theta)$ produces an output $f(x;\Theta)$, and an updating rule for Θ is iteratively refined via measurement and parameter tuning, similar to classical neural network training.

Hybrid classical–quantum architectures offer a practical path for quantum-enhanced machine learning under NISQ constraints. Study [47] proves they can approximate any continuous function and surpass classical Deep Neural Networks (DNNs) in representational power. Building on this, [50] applies hybrid transfer learning for brain tumor MRI classification using ResNet-18 with a VQC classifier, achieving 96.7% ACC. Similarly, [51] proposes a CNN–PQC model for iris tumor classification, reaching 98.4% ACC, showing that quantum properties can help mitigate data scarcity. In terms of application domains, the reviewed studies can be classified into three principal categories:

- **Medical and Healthcare:** The convergence of QML and medical image analysis has emerged as a promising research frontier. Several studies show that hybrid classical–quantum architectures can achieve competitive diagnostic performance across different medical imaging tasks. This approach has been applied to brain tumor [50], iris tumor [51], breast cancer [52], pneumonia, and retinopathy classification [53], with reported accuracies ranging from about 83% to 98.4%, depending on the task, dataset size, and circuit depth. A key finding is that quantum models may be especially useful in data-scarce settings, where superposition and entanglement can help improve generalization [53].
- **General Image and Noisy Classification:** QML has shown strong potential for image classification. PQCs are integrated into classical deep learning pipelines, improving feature discrimination

through superposition and entanglement while reducing trainable parameters. In some cases, hybrid models achieve comparable or better ACC with up to eight times fewer parameters and remain robust against various noise types [54]. These methods have been applied to handwritten digits, fashion and Kuzushiji images, medical imaging, and natural scenes [54,55]. However, a major limitation remains the drop in performance on real NISQ hardware compared with simulators, due to noise, limited coherence, and restricted quantum volume [55].
- **Remote Sensing:** In the classification of multispectral satellite images from the EuroSAT dataset, study [49] shows that quantum models can be extended beyond conventional benchmark tasks to real-world geospatial data analysis. A key domain-specific challenge is the high dimensionality of remote sensing imagery, which requires classical feature extraction and dimensionality reduction before quantum processing. This makes the study a relevant example of how hybrid quantum–classical models can be adapted for practical remote sensing applications [49].

4.2. Quantum Convolutional Neural Network

QCNNs extend classical neural networks by using quantum principles, particularly VQCs, for feature extraction and image classification. To address NISQ hardware limitations such as restricted qubit counts, recent studies increasingly adopt hybrid architectures in which classical networks perform initial feature extraction, while quantum circuits handle deeper processing and classification, showing strong performance in radiological image analysis [56]. This integration has also been expanded to advanced deep learning models, including Vision Transformers (ViT) [57], VGG16 [58], and Swin-UNet [59], reflecting ongoing efforts to combine quantum circuits with modern, efficient AI architectures for accurate image processing.

Although hybrid models are essential for coping with current hardware limitations, developing robust native quantum frameworks remains a major research goal. Several foundational studies have focused on designing and validating QCNN structures that can directly encode and process classical data, demonstrating their feasibility for standard image classification without extensive classical pre-processing as you can see in Figure 13 [60,61]. An important step in this direction is the emergence of quanvolutional neural networks, where classical convolutional filters are replaced by VQCs that act on local image patches, enabling image recognition through quantum transformations [62]. Building on these ideas, later models further connected quantum computing with deep learning to achieve accurate and computationally efficient baseline classification [63]. More recently, researchers have also explored new hardware directions, such as the Quantum Optical Convolutional Neural Network (QOCNN) [64], which extends quantum image recognition beyond traditional gate-based qubit systems and broadens the architectural possibilities of the field.

As theoretical QCNN frameworks mature, the research landscape has increasingly prioritized rigorous benchmarking against state-of-the-art classical Convolutional Neural Networks (CNNs). A significant portion of this comparative analysis is dedicated to evaluating performance discrepancies in complex, real-world tasks. Direct comparative studies have been conducted to rigorously assess the proficiency of standard CNNs versus QCNNs in binary classification tasks, such as analyzing breast cancer histopathological images [65]. Similarly, researchers have executed comprehensive comparative studies on various implementation methods to identify the optimal approach for enhanced skin lesion classification [66].

Beyond mere classification ACC, researchers are heavily focused on the operational viability and computational efficiency of these quantum operations. Studies have demonstrated that deeply integrated QML frameworks can achieve highly accurate and efficient image classification [63]. More critically, substantial research is currently addressing the practical deployment challenges by engineering efficient QCNN models specifically designed to overcome the hardware constraints inherent in contemporary NISQ devices [67]. Furthermore, the evaluation of QCNNs has expanded beyond standard ACC metrics to include deeper analytical characteristics. Comprehensive investigations into the models' robustness against noise and adversarial conditions, coupled with their overall explainability, are actively being prioritized. These deeper benchmarks aim to clearly delineate the specific

 doi:10.20944/preprints202606.1015.v1

advantages and analytical limitations of quantum models when compared to traditional deep learning methodologies [68].

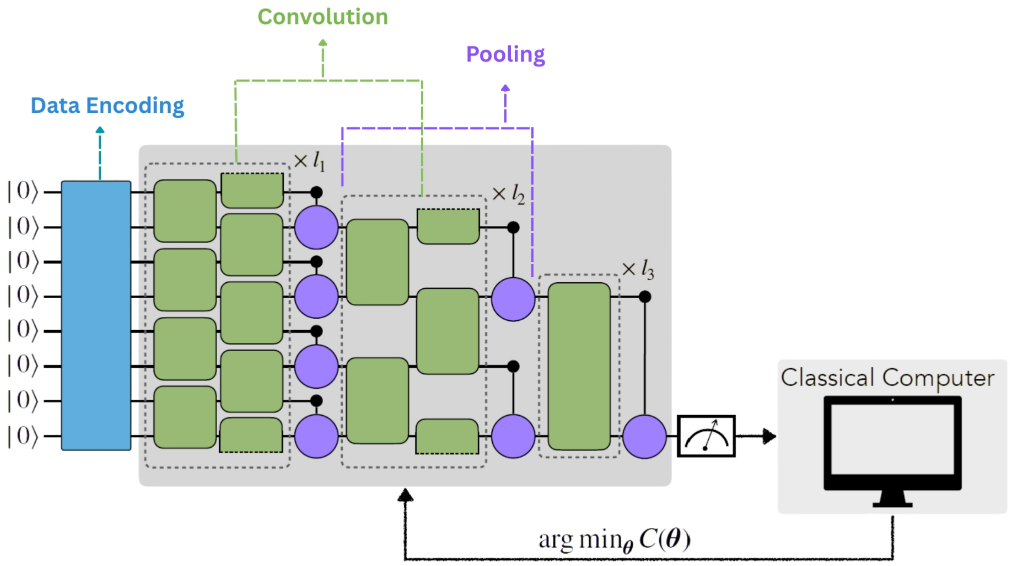


Figure 13. General framework of QCNN for classical data classification, showing the integration of quantum data encoding, hierarchical convolution and pooling operations, measurement, and feedback-based classical optimization. Adapted from [60].

QCNN use across a wide range of application domains, highlighting their flexibility in both specialized medical tasks and broader computer vision problems.

- **Medical Imaging Domain**
  Medical imaging is one of the most important application areas of QCNNs, where accurate and robust pattern recognition is essential. Their quantum-based feature extraction capabilities make them well suited for analyzing complex medical images. Recent studies have applied these models to liver tumor segmentation and classification [59], general radiological image classification [56], breast cancer histopathology classification [65], and skin lesion recognition [66]. These results demonstrate the strong potential of QML for improving medical diagnostic systems.
- **Object and Emotion Detection** Beyond medical imaging, QCNNs have also been applied to more general computer vision tasks. Hybrid quantum models have been used to enhance emotion detection for improving human-computer interaction and personalized digital services [58]. In addition, recent frameworks have been developed for object detection and classification in dynamic environments [69]. This expansion shows the adaptability of quantum models and their growing relevance in real-world visual recognition applications.

Table 7 below summarizes a subset of representative studies reviewed in this section to provide a concise and non-redundant comparison. The selected papers were chosen because they offer clearer methodological diversity and better reflect the range of application domains, datasets, main research focus, key contributions, and limitations discussed within this category.

 doi:10.20944/preprints202606.1015.v1

Table 7. Summary of representative QCNN-based studies, including their application domains, datasets, main focus, key contributions, and reported limitations.

Ref./ Year	Application Domain	Dataset	Focus	Key Contribution	Limitations
[67], 2025	General Image Classification	MNIST, Fashion-MNIST	Efficient QCNN with reduced input dimensionality encoding for NISQ devices; automated PQC selection framework	49-qubit QCNN with no classical dimensionality reduction; achieved 96.08% ACC on IBM Heron r2, outperforming classical CNN (71.74%)	Tested only on binary MNIST classification; limited to NISQ hardware with current qubit counts
[59], 2025	Medical Imaging (Liver Tumor Segmentation & Classification)	3D-IRCADb, LiTS17, MSD Task03	Lightweight hybrid framework combining optimized Swin-UNet (segmentation) with QCNN (classification) for edge deployment	Integration of SAR metaheuristic optimization and Focal AUC loss; 80.25% model size reduction (64.16 MB) enabling Jetson Nano deployment	Still heavy for stricter edge-device memory constraints
[64], 2021	General Image Classification	MNIST	Quantum Optical CNN combining quantum computing and optical neural networks	Proposed QOCNN framework leveraging quantum optical circuits; benchmarked against CNN on MNIST	Only MNIST tested; impractical optical quantum hardware; lacks real-world validation
[65], 2024	Medical Imaging (Breast Cancer)	BreakHis dataset	QCNN vs CNN comparison for histopathology classification	QCNN achieved AUC 0.96 vs CNN 0.91; better performance under class imbalance	Small dataset; binary task only; no clinical-scale validation
[57], 2025	General Image Classification	CIFAR-10, MNIST	ViT + QCNN hybrid (ViT-QCNN-FT); entanglement study	Achieved 99.77% on CIFAR-10; quantum noise improved ACC (+2.71%) in some cases	Computationally expensive ViT preprocessing; quantum part only simulated
[58], 2025	Emotion Detection	RAF-DB	QCNN vs VGG16 for facial emotion recognition	4.5% ACC gain over VGG16 and 15–25% faster processing using hybrid pipeline	Only RAF-DB used; low-resolution quantum inputs; simulation-only evaluation
[63], 2025	General Classification	MNIST, Wine dataset	Compact 8-qubit QCNN for NISQ devices	97.26% MNIST (binary), 97.22% Wine classification with low-depth circuits	Heavy dimensionality reduction; simulation-based only

4.3. Quantum Vision Transformer (QViT)

Quantum Transformer architectures aim to combine the strengths of transformer-based learning with the computational properties of quantum systems. In this section, the reviewed studies are discussed from three complementary perspectives: architectural design, computational complexity and efficiency, and application-oriented analysis.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2026 doi:10.20944/preprints202606.1015.v1

The classical ViT forms the basis of all QViT models discussed here. Its self-attention mechanism captures global dependencies between image patches, but its computational cost scales as $O(nd^2 + n^2d)$, which motivates the search for quantum alternatives [70], which replaces standard attention with quantum linear-algebra primitives and introduces several variants. It also reduces parameters significantly, using 80 parameters per attention layer compared with 512 in the classical version [70].

A more practical hybrid design is presented by Hybrid Quantum Vision Transformer (HQViT) [71], which quantizes only the attention coefficient computation, the main classical bottleneck with cost $O(T^2d)$. Using whole-image amplitude encoding and swap-test-based similarity evaluation, it requires only $O(\log_2 N)$ qubits and $O(\log_2 d)$ parameterized gates, while achieving better results than classical ViT on several datasets, including up to 10.9% improvement on MNIST [71]. In a different hybrid direction, QMViT [72] keeps classical multi-head attention but introduces quantum processing in the feed-forward stage, using the distance-based attention formula $A_{ij} = -(Q_i - K_j)^2$. On mushroom classification, it achieves 92.33% species ACC and 99.24% edibility ACC [72]. Finally, the naive QViT [73] offers a simpler proof of concept by replacing dot-product attention with a VQC, showing performance comparable to classical ViT on MNIST [73].

A central challenge in quantum transformer design is reducing the cost of classical attention while staying compatible with quantum hardware. Since standard self-attention scales as $O(n^2)$, it becomes a natural target for quantum reformulation. In [74] address this by replacing the linear projections in ViT attention with parameterized QNNs, reducing parameter scaling from $O(n^2)$ to $O(n)$ while preserving expressive power, and achieving near-SOTA performance on RetinaMNIST with 99.99% fewer parameters than comparable classical models. In [75] they focus on deeper multi-layer quantum architectures, proposing a modular hybrid framework in which high-dimensional operations are assigned to Quantum Linear Algebra Modules (QLAMs) and lower-dimensional structured operations to Quantum Arithmetic Modules (QAMs). In contrast [76] researchers target the normalization stage, replacing softmax with a VQC that directly produces doubly stochastic attention matrices, avoiding the instability of iterative classical methods such as Sinkhorn and scaling as $O(\log_2(T))$ in qubit count with sequence length T. These studies highlight complementary strategies for managing transformer complexity: reducing attention parameters, enabling scalable multi-layer quantum depth, and reformulating the normalization operator.

From an application perspective, the reviewed quantum transformer studies can be grouped into several major domains based on the primary tasks and real-world problems they address.

- **High Energy Physics Applications**
  The projected data scale of the HL-LHC has motivated the use of hybrid QViT for high-energy physics classification. In quark–gluon jet classification using CMS Open Data, [77] integrated VQCs into both the attention and MLP blocks of a QViT, achieving a test AUC of 0.775 versus 0.793 for the classical ViT, with fewer parameters. On the same task, [78] compared several pooling variants and found that the column-max model came closest to the classical baseline, while the class-token version failed to converge, revealing strong architectural sensitivity. In [79] replaced classical projections with quantum attention mechanisms, obtaining strong competitive performance while maintaining parameter efficiency.
- **Medical and Biomedical Imaging**
  Recent studies show that QViT-based models are promising for biomedical image classification with strong parameter efficiency. Study [70] demonstrates the relevance of QViT-based models to biomedical image classification through evaluation on multiple biomedical datasets, highlighting their competitive performance and strong parameter efficiency for resource-constrained healthcare settings. While in [80] introduces the Quantum-Enhanced Swin Transformer (QEST), which replaces the final classifier with a VQC, reducing parameters by 62.5% and improving resistance to overfitting. It also validates the model on a 72-qubit superconducting quantum computer, supporting feasibility in the NISQ era.

 doi:10.20944/preprints202606.1015.v1

- **Advanced Recognition and Image Enhancement**
  From an application domain perspective, these studies illustrate the breadth of quantum transformer research beyond standard image classification to extend the paradigm to image enhancement, like single-image super-resolution [81], where the proposed model is used to reconstruct high-resolution images from degraded low-resolution inputs across general and medical imaging datasets. Also [82] demonstrates the applicability of quantum transformers to intelligent transportation systems, where the model is designed for traffic sign recognition in intelligent vehicular networks, emphasizing its role in traffic scene understanding and autonomous driving perception.

Table 8 below highlights a selection of recent studies discussed in this section, they represent the latest contributions published in 2025 and capture the current landscape of this category through their covered application areas, datasets, research objectives, main findings, and identified limitations.

Table 8. Summary of selected recent quantum transformer studies (2024–2025).

Ref./Year	Application Domain	Dataset	Focus	Key Contribution	Limitations
[[70] (2024)	Medical Image Classification	MedMNIST	Design of QViTs by translating attention layers into quantum circuits; orthogonal patch-wise and transformer variants	Introduces 3 QViT variants; Compound Transformer achieves $O(d \log d)$ parameters vs $O(d^2)$ classical; demonstrated on 6-qubit hardware	Deep circuits limit NISQ scalability; requires all-to-all connectivity; no large-scale dataset validation
[70] (2025)	General Image Classification	MNIST, MedMNIST, CIFAR-10, CIFAR-100, Mini-ImageNet	HQViT with whole-image processing and quantum self-attention	Amplitude encoding reduces qubits to $O(\log_2 Td)$ and PQGs to $O(\log_2 d)$; up to 11% MNIST improvement	Encoding and measurement overhead increases with stacked quantum blocks
[74] (2025)	Biomedical Image Classification	RetinaMNIST + 7 MedMNIST datasets	Quantum self-attention ViTs with knowledge distillation (KD)	Reduces self-attention parameters from $O(n^2)$ to $O(n)$; first KD study in QViTs; near-SOTA on RetinaMNIST	Limited to simulation and 4-/8-qubit setups; KD weak for low-capacity models
[76] (2025)	Image Classification	MNIST, FashionMNIST, MedMNIST datasets	Quantum doubly stochastic attention (QDSFormer) replacing softmax using QontOT	Improved training stability; outperforms ViT and Sinkformer; scalable quantum-based normalization	Simulation-only; hardware noise not modeled; limited large-scale evaluation
[79] (2025)	High Energy Physics (Jet Classification)	CMS Open Data (933K quark-gluon jets)	Hybrid QViT with quantum orthogonal neural networks in self-attention	Comparable performance to classical ViT; introduces quantum-enhanced attention for jet classification	Marginal gain over classical ViT; limited projection dimension; simplified experimental setup
[80] (2025)	Breast Cancer Classification	Cohort A (2,601 FFDM cases), IN-breast	VQC integrated into Swin Transformer classifier (QEST)	Reduced parameters; comparable or slightly improved performance; improved generalization	Limited scalability; shallow 8-/16-qubit circuits; only classifier quantumized
[82] (2025)	Traffic Sign Recognition	GTSRB, MNIST, KMNIST, FashionMNIST, CIFAR-10	QViT with quantum DFT-based feature extraction and hierarchical attention fusion	Average +9.01% ACC and +8.48% F1 over QC-CNN and QViT baselines	Simulation-only evaluation; high circuit complexity; hardware scalability not demonstrated

4.4. *Quantum Attention Mechanism Based Models*

Attention mechanisms are designed to enhance representation learning by enabling models to focus on the most informative relationships among input features rather than processing all elements uniformly. Within the reviewed quantum literature, the selected studies can be broadly grouped into soft/self-attention models and hard-attention models. The soft-attention category includes the foundational self-attention formulations QKSAN [83] and QSAN[84], as well as more advanced extensions such as SASQuaTCh[85], EQSAM[86], and QCSAM[87]. More specifically, QKSAN [83] and SASQuaTCh[85] represent kernel-based attention designs, QSAN[84] introduces a logic-based near-term self-attention framework. The hard-attention category is represented by GQHAN[88] and QAHAN[89], which employ discrete selection strategies through Grover-inspired and quantum-annealing-based optimization, respectively. Table 9 provides a structured comparison of these studies in terms of their architectural category, framework design, dataset, quantum encoding strategy, computational complexity, and key contribution.

Table 9. Overview of selected quantum attention-based models in terms of architectural category, framework design, quantum encoding strategy, complexity-related advantages, and limitations.

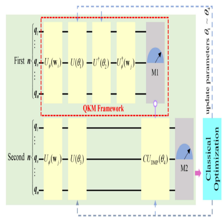
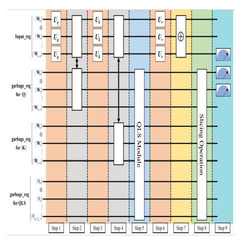
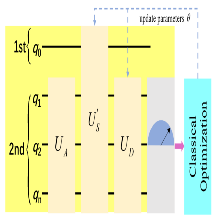
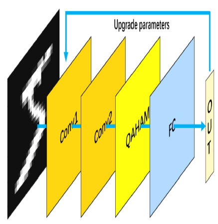
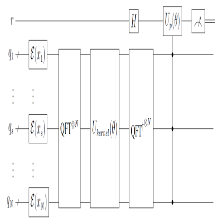
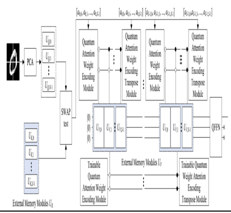
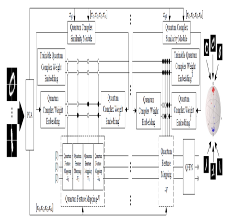
Category	Ref./ Year	Framework Diagram	Dataset	Quantum Encoding	Computational Complexity	Key Contribution
Basic Quantum Self-Attention Models (Soft Attention)	QKSAN [83], 2024		MNIST, Fashion-MNIST (Binary)	Amplitude & Angle Encoding	O(N log N) via Quantum Kernel Hilbert Space; reduced resource via Deferred Measurement Principle (DMP)	First integration of QKMwith Self-Attention; achieves >99% ACC with significantly fewer parameters than classical models
	QSAN [84], 2024		MNIST, CIFAR-10 (Binary)	Basis Encoding via Quantum Logic Similarity (QLS)	One-step quantum execution; measurement compression; near-term achievable; converges 1.7× faster than HE-ansatz	Proposes QLS and QBSASM to realize near-term achievable quantum self-attention in a single holistic framework
Quantum Hard Attention Models	GQHAN [88], 2024		Fashion-MNIST (Binary)	Amplitude encoding	Exponential FO scaling (2^n discrete primitives); ADO adds linear overhea	First application of Grover's Search for Hard Attention; introduces (FO) + (ADO); outperforms soft attention with ≥ 98% ACC
	QAHAN [89], 2024		MNIST, CIFAR-10	Classical Feature Maps (CNN)	Binary quadratic optimization solved on D-Wave hardware	Combines QAHAM with a CNN backbone to improve convergence speed, learning stability, and robustness

Table 9. *Cont.*

Category	Ref./ Year	Framework Diagram	Dataset	Quantum Encoding	Computational Complexity	Key Contribution
Enhanced Quantum Attention Models (Soft Attention)	SASQuaTCh [85], 2024		MNIST (1 vs 3 and 3 vs 8) Synthetic lines	Angle or amplitude embedding after classical patch and position embedding	QFT contributes $O(q \log q)$; overall complexity depends on the embedding scheme; demonstrated with only 9 qubits on one MNIST setup.	Introduces a variational quantum transformer with QFT-based kernel self-attention entirely inside the quantum circuit
	EQSAM [86], 2025		MNIST, Fashion-MNIST	Rx data encoding + trainable ZZ/Ry layers	Reduced from O(N²) to O(N) via fixed-size external memory modul	First quantum self-attention framework using complex-valued similarities, quantum multi-head attention
	QCSAM [87], 2025		MNIST, Fashion-MNIST	Rx data encoding with trainable ZZ and Ry variational layers	Pairwise query-key attention over N patches (effectively O(N^2)); uses CLCUs + improved Hadamard test	First complex-valued attention weights capturing amplitude AND phase; extends LCUs to CLCUs with complex coefficients; quantum multi-head attention supported

As shown in Table 10 that no single model is universally best, because the studies use different task settings, datasets, and platforms. Even so, some patterns are clear. On MNIST, the strongest overall performers are QSAN [84] and QCSAM [87], since both reach perfect classification in their reported binary settings; however, QSAN [84] appears stronger from an optimization perspective because it also converges faster than the hardware-efficient and QAOA baselines, making its advantage more than just a final-ACC result. QKSAN [83] remains highly competitive, but it is generally slightly below QSAN on MNIST and below the best Fashion-MNIST results, while its experiments rely on heavy dimensional compression, which limits how directly its results can be interpreted against richer input settings. On Fashion-MNIST, GQHAN [88] is particularly notable, as it slightly surpasses the main soft-attention baselines in the reported binary task and shows stronger convergence behavior; still, its evidence is confined to Fashion-MNIST, so its superiority cannot be generalized too broadly. EQSAM [86] is arguably the most balanced recent model, because it combines very strong ACC with a clear scalability advantage by reducing attention complexity from quadratic to linear, although this benefit depends on selecting a suitable number of external memory modules. QCSAM [87] is also very strong, especially because it achieves top-tier performance with only a small qubit count and introduces complex-valued, multi-head attention; however, its own analysis indicates that performance declines as the task becomes harder, so its advantage is most convincing in simpler settings. By contrast,

 doi:10.20944/preprints202606.1015.v1

QAHAN [89] shows very favorable convergence and stable learning, but it is less directly comparable because it is a 10-class CNN–quantum annealing hybrid, not a standard gate-based self-attention model. Finally, SASQuaTCh [85] is methodologically important as a kernel-based quantum transformer, yet its evaluation is narrower, being mainly demonstrated on selected digit-pair classification tasks rather than a broader shared benchmark. Overall, QSAN [84] and QCSAM [87] appear strongest in raw MNIST performance, GQHAN [88] is the most competitive hard-attention model on Fashion-MNIST, and EQSAM [86] offers the clearest trade-off between performance and scalability, while the main limitations across the group remain restricted evaluation settings, compressed inputs, and incomplete comparability between experimental protocols.

Table 10. Experimental comparison of selected quantum attention-based models evaluated on MNIST and Fashion-MNIST in terms of task setting, qubit count, reported ACC, and convergence behavior. Reported results are taken from different experimental setups.

Model	**MNIST Task Type / Platform**	**# Qubits**	**Train / Test ACC**	**Convergence Behavior**	**Fashion-MNIST Task Type / Platform**	**# Qubits**	**Train / Test ACC**	**Convergence Behavior**
GQHAN [88]	N/A	N/A	N/A	N/A	Binary (PennyLane)	4	98.65 / 98.59	Converges step 19; loss = 0.219
SAS QuaTCh [85]	Binary class / PennyLane	9	NR / 96.8	200 epochs; sensitive to init.	N/A	N/A	N/A	N/A
QKSAN [83]	Binary class / (PennyLane + IBM Qiskit)	4	$\sim$ 99% / 99.00%	Best submodel = AmHE; smallest final loss; trails comparator QSAN by 42 steps	Binary (PennyLane / IBM Qiskit)	4	98.52 / 98.05	Best submodel = AmHE; smallest final loss; trails comparator QSAN by 50 steps
QSAN [84]	Binary class / PennyLane	18	100 / 100	Converges at step 130; 1.7× faster than HE-ansatz and 2.3× faster than QAOA	N/A	N/A	N/A	N/A
QAHAN [89]	10-class / D-Wave + PyTorch	NR (annealer - based)	Train: 0.99–0.999 / Test: 0.99– 1.0	Loss drops from epoch 2 and falls below 0.0049 after epoch 5	N/A	N/A	N/A	N/A
QCSAM [87]	Binary, Ternary class / PennyLane	4	100 / 100	no single convergence value reported; dual-head consistently outperforms single-head	Binary class / TensorCircuit + TensorFlow	4	98.4 ±0.55 / 99.2 ±0.74	no single convergence value reported; dual-head consistently outperforms single-head
EQSAM [86]	Binary, Ternary class / Tensor Circuit + TensorFlow	6	99.68 ±0.17 / 99.92 ±0.16	trained for 200 epochs; performance saturates near the optimal number of external modules	Binary, Ternary class / TensorCircuit + TensorFlow	6	98.89± 0.26 / 98.36± 0.63	trained for 200 epochs; performance saturates near the optimal number of external modules

 doi:10.20944/preprints202606.1015.v1

4.5. Comparative Analysis Across Model Categories

This section provides a cross-category comparative analysis of the reviewed model families, focusing on their architectural design, representation capability, computational complexity, application suitability, and practical limitations. QNNs may be viewed as general quantum learning models in which classical data are encoded into quantum states and processed through PQCs for prediction. By contrast, QCNNs are more structurally specialized for image data, as they adopt hierarchical quantum convolution and pooling operations to capture local spatial patterns and progressively compress features . Thus, the main distinction is that QNNs emphasize general quantum representation learning, whereas QCNNs emphasize locality and hierarchical feature extraction. In the comparative Fashion-MNIST study [[90], the proposed QNN achieved faster training, lower loss, and higher final ACC than the corresponding QCNN (90% vs. 86%), although this result should be interpreted cautiously because the images were reduced to 4×4 during evaluation. From an architectural perspective, QCNN-based models use the VQC mainly as a feedforward feature-processing module, whereas QViT-based models inherit the global self-attention mechanism of ViT, allowing them to model long-range dependencies more effectively. This helps explain why transformer-based quantum hybrids perform better on complex tasks. In particular, [91] reports that HQC-ViT retained about 86 million parameters while improving ACC by roughly 4% over the classical ViT baseline. The difference becomes clearer when data sensitivity and initialization are considered. According to [92], under hybrid learning without pretraining, Hybrid CNN achieved 95.50%, while Hybrid ViT reached only 68.79%, a gap of nearly 27 percentage points. However, with transfer learning, the pattern changes: Q-TL CaiT achieved 98.98%, slightly surpassing Q-TL ResNet at 98.62%. This suggests that transformer-based quantum models benefit much more from pretrained representations. In terms of efficiency, [91] also shows that a 4×4 quantum kernel reached 95.23% ACC, exceeding the 94.07% of the classical CNN baseline. Overall, the evidence suggests that QNN/QCNN-style models are better suited to limited-data or simpler settings, while QViT models combined with transfer learning are more suitable for complex classification tasks that require stronger global contextual modeling.

Table 11 summarizes the selectedQViT-based studies discussed in this section. It provides a concise comparison in terms of task setting, qubit usage, encoding method, main quantum component, reported performance, classical baseline, and the specific quantum contribution of each work.

The comparative results presented in Table 11 indicate that no single QDL category can be considered universally superior; rather, each model family demonstrates distinct strengths depending on the task setting, architectural objective, and hardware constraints. QNN-based models generally exhibit strong representational capacity and competitive ACC with relatively few qubits, making them particularly attractive for compact hybrid learning scenarios. QCNN-based models, in contrast, appear more structurally aligned with image data due to their hierarchical convolution and pooling operations, which allow them to capture local spatial features more effectively and, in some cases, outperform classical CNN baselines even under real-hardware conditions.

Meanwhile, QViT-based models introduce a fundamentally different advantage by incorporating quantum mechanisms into transformer-style attention, thereby offering stronger potential for modeling global dependencies and improving performance in more complex visual learning settings. However, this advantage often comes at the cost of greater architectural complexity, higher simulation dependence, and additional encoding or optimization overhead. Overall, Table 11 suggests that QNN and QCNN models currently provide a more practical and hardware-aware path for near-term quantum image classification, whereas QViT models represent a promising but still emerging direction for achieving more expressive and globally informed quantum vision architectures.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 June 2026 doi:10.20944/preprints202606.1015.v1

Table 11. Comprehensive comparison of MNIST-based studies across QNN, QCNN, and QViT.

Categories	Ref./ Year	Task / Platform	# Qubits	Encoding Method	Main Quantum Component	Performance ACC (%)	Classical Baseline (%)	Quantum Contribution
QNN	[47], 2021	Binary (0 vs 1) / Julia Yao.jl simulator	8 / 6 / 4	PQC-based encoder after resizing 28×28 to 8×8	QDNN: Multi-layer QNNLs built on PQCs	99.57% (test, ∞ shots)	N/A	Multi-layer QNN with BP-compatible training; strong representation power and high MNIST ACC with few qubits.
	[48], 2022	Binary (0 vs 9, 3 vs 8) / Qiskit + IBM Real Hardware	Varies by encoding and control qubits	Amplitude encoding (single-qubit or padded)	Quantum Single Layer Perceptron with control-register superposition	Simulator: up to 91% / Real device: up to 84%	N/A	Exponential hidden neurons in superposition; amplitude encoding for high-dim data.
	[55], 2023	4-class noisy / PennyLane + IBM Real Hardware	5	Amplitude encoding via Hadamard + Rx/Rz rotation gates	VQC	98.2% / ∼59.7%	DSNN:97% / RQNN:96% / ResNet-18:82%	Improves noisy-image classification over baselines with fewer quantum parameters.
	[54], 2024	Multi-class / PennyLane	5 Parallel, 4 Quanv	Angle Embedding (Rx rotation, Bloch sphere X-axis)	Parallel PQCs / Quanvolutional kernel	HQNN-Parallel: 99.21% / HQNN-Quanv: 67%	CNN ≈ 98.71%	HQNN-Parallel beats CNN with ∼8× fewer parameters.
QCNN	[62], 2020	Multi-class / QxBranch Simulator	9	Threshold-to-basis-state encoding	Quanvolutional layer	≈97.5–98.0	CNN ≈ 99%	Introduced quanvolutional layer as quantum analogue of convolution.
	[64], 2021	Multi-class / PyTorch simulation	392 (theoretical; Fock-state encoding)	Quantum-optical encoding into complex amplitudes	QOCNN	97.4%	CNN 97.6% / ONN 97.9%	Quantum optical CNN with strong robustness.
	[60], 2022	Binary / PennyLane	8	Hybrid encodings (amplitude, angle, PCA)	QCNN with quantum pooling	98.4%	97.0%	Entanglement-based feature extraction.
	[69], 2023	Binary / Qiskit IBM	N/A	ZZFeatureMap encoding	QCNN layers (conv + pooling)	52–61%	CNN 99.9%	Shows hardware limitations and noise impact.
	[67], 2025	Real hardware QCNN	49 (hardware); 2–16 (hybrid)	Fragment encoding (WUE best)	QCNN + pooling + PQC blocks	98.7% / 96.08%	CNN 93.4% / 71.74%	First QCNN outperforming CNN on real hardware.
QViT	[73], 2024	Multi-class / PennyLane	10	Ry encoding after projection + α scaling	Variational quantum self-attention	∼95–96%	∼95–96%	Matches classical ViT performance.
	[71], 2025	TensorCircuit simulation	8–10	Amplitude encoding	Quantum transformer blocks	93.1%	N/A	Swap-test reduces $O(T^2 d)$ complexity.
	[76], 2025	Qiskit + IBM hardware	16 total (4 data + $\geq$ 4 ancilla)	Matrix injection $f(\theta, M) = \theta \odot \bar{M}$	QontOT softmax replacement	98.8%	N/A	Quantum attention replaces softmax for stability.

 doi:10.20944/preprints202606.1015.v1

5. Discussion and Future Direction

This section discusses the main findings of the review in relation to the RQs and provides a broader interpretation of the current state of QDL for image classification. It evaluates the reported quantum advantages and practical limitations of the reviewed models, reflects on the limitations of the present study, and outlines future research directions and open challenges that may shape the development of this field.

5.1. Evaluation of Quantum Advantage and Limitations in Image Classification

The reviewed studies reveal that QDL for image classification is still evolving through diverse architectural directions, with most efforts focusing on how quantum components can be integrated into classical learning pipelines under current hardware constraints.

Regarding RQ1, the most quantum component is integrated into image classification models predominantly through hybrid quantum–classical architectures[57–59,62,64,66,69,71,74] rather than fully quantum end-to-end designs[47,48,52]. In most cases, classical modules are retained for data preprocessing, feature extraction, or dimensionality reduction[59,71,74], while the quantum part is introduced in a more selective manner through parameterized or VQCs. This pattern is consistently observed across the main taxonomy categories. In QNN-based models, the quantum circuit is commonly employed as a trainable learning layer or classifier[51,53,54]; in QCNN-based models, it is integrated into convolutional and pooling operations or placed after classical feature extractors[58,63,66]; and in QViT and quantum attention-based models, the quantum component is mainly used to reformulate attention-related operations such as projection, similarity estimation, or normalization [78,79,86,87]. The findings suggest that the dominant trend in the reviewed literature is not to replace classical image classification pipelines entirely, but rather to embed the quantum component into the most computationally or representationally critical parts of the model in a way that remains compatible with current NISQ-era hardware constraints.

To address RQ2, this review critically examines whether the reported performance of current QDL models reflects a real practical quantum advantage in image classification. Although several works report encouraging gains in specific settings, such as higher ACC on simplified benchmarks[65,67,80], improved parameter efficiency[58,67,70,71,74,83], better robustness to noise [55,81], or reduced overfitting[80], these gains remain largely context-dependent rather than generalizable. However, many of these results are obtained on binary or heavily reduced tasks, after substantial classical preprocessing or dimensionality reduction, or through simulation rather than scalable quantum execution. In addition, other studies show the opposite pattern, where classical models still clearly outperform quantum counterparts under comparable conditions[90–92]. This interpretation is also supported by broader methodological reflections in "Is Quantum Advantage the Right Goal for QML?"[93], which argues that demonstrating practical advantage in machine learning is particularly difficult because classical baselines are already extremely strong and realistic benchmarking remains limited, and by "Better than classical?[94] which shows that, across systematic small-scale benchmarks, classical models often outperform quantum ones and that benchmarking outcomes are highly sensitive to experimental design. Therefore, the most defensible conclusion is that current QDL models show promising signs of task-specific or architectural benefit, but they still fall short of demonstrating a robust, scalable, and practically verified quantum advantage for image classification at the present stage.

To evaluate RQ3, the reviewed studies reveal several recurring limitations and open challenges:

- Persistent hardware constraints, including noise, decoherence, limited qubit counts, and restricted quantum volume, which still hinder reliable real-device performance [67,81].
- Scalability and trainability issues, such as barren plateaus, optimization difficulty, deep-circuit cost, and encoding/measurement overhead[62,79].
- Methodological limitations, since many studies are still evaluated on simplified or binary tasks, reduced inputs, or simulators rather than fully realistic large-scale settings[47,48,60,63,67,69,83,85,86].

- Benchmarking and generalization challenges, as reported gains often depend strongly on the chosen dataset, preprocessing pipeline, and baseline, making cross-study comparison and broad claims of practical quantum superiority still difficult[90–92].

Overall, the literature suggests that the field remains promising, yet its progress is still constrained by both current NISQ-era hardware and the lack of standardized large-scale evaluation protocols.

5.2. Study Limitations

This study has several limitations. It was intentionally restricted to QDL studies on image classification, which improved focus but excluded other related QML tasks and model families. A formal risk-of-bias assessment tool was not applied in this review, as the included studies were highly heterogeneous in terms of model architectures, datasets, evaluation settings, and experimental designs. Instead, methodological limitations and potential sources of bias were considered qualitatively during the comparative analysis and discussion. The selection process was also limited to English-language studies published between 2020 and 2025 and filtered by specific inclusion and exclusion criteria, so some relevant work may have been omitted. In addition, although multiple major databases and a PRISMA-based workflow were used, some studies may not have been fully captured. A few potentially relevant papers were also inaccessible because of paywalls or unavailable full texts. Finally, the strong heterogeneity of the selected studies in datasets, preprocessing, evaluation settings, and use of simulation versus real hardware limits direct comparison and broad generalization.

5.3. Future Direction and Open Challenges

Future research should move beyond proof-of-concept demonstrations toward more scalable, application-oriented, and fairly benchmarked QDL models. In particular, the reviewed literature suggests several promising directions that may help strengthen both the practical value and the methodological maturity of this field.

- **Exploring hybrid quantum architectures based on hierarchical Swin Transformer:** It represents a particularly promising direction, since most current quantum transformer studies focus on standard ViT-style formulations, integrating quantum components into shifted-window attention and multi-scale hierarchical representations may provide a more practical and efficient path for image classification, especially on more complex visual datasets.
- **Investigating different attention mechanisms:** It remains an important open direction, future studies may compare multiple attention designs, including self-attention, window-based attention, sparse attention, and quantum-enhanced attention variants, across different application domains to better understand where each mechanism provides the most meaningful improvement.
- **Improving scalability and hardware compatibility:** Particularly by reducing encoding overhead, limiting circuit depth, and developing quantum modules that are more robust to current NISQ constraints.
- **More standardized benchmarking:** It is still needed to enable fairer comparison with classical baselines under consistent evaluation settings.

6. Conclusion

This systematic review presented a focused analysis of QDL for image classification by examining 47 selected studies published between 2020 and 2025 through a PRISMA-based selection process and a taxonomy-based framework. The reviewed studies were organized into four main categories: QNNs, QCNNs,QViTs, and quantum attention-based models. This classification helped provide a clearer understanding of the current research landscape, the main architectural trends, and the datasets most frequently used in this field. The review showed that most existing studies rely on hybrid quantum–classical architectures, where the quantum component is integrated selectively into specific parts of the learning pipeline rather than replacing the full classical model. Although several studies reported promising results, such as improved parameter efficiency, competitive ACC, or robustness

 doi:10.20944/preprints202606.1015.v1

in specific settings, the findings do not yet support a robust practical quantum advantage in image classification. This is mainly due to persistent limitations related to hardware noise, limited qubit availability, scalability issues, encoding overhead, and inconsistent benchmarking settings. QDL for image classification remains a promising but still emerging research direction, with strong potential for future development as both quantum hardware and model design continue to mature.

Author Contributions: Saleha Daraan: Conceptualization, Methodology, Investigation, Data collection, Formal analysis, Writing—original draft. Wael Alghamdi: Conceptualization, Supervision, Validation, Writing—review and editing. All authors reviewed the results and approved the final version of the manuscript.

Abbreviations

ACC	Accuracy
AQC	Adiabatic Quantum Computing
BP	Backpropagation
DNN	Deep Neural Network
DSNN	Deep Spiking Neural Network
DSQ-Net	Deep Spiking Quantum Neural Network
HQNN	Hybrid Quantum Neural Network
HQViT	Hybrid Quantum Vision Transformer
MQ	Mapping Question
MRI	Magnetic Resonance Imaging
NISQ	Noisy Intermediate-Scale Quantum
PQC	Parameterized Quantum Circuit
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QAI	Quantum Artificial Intelligence
QAM	Quantum Arithmetic Module
QCNN	Quantum Convolutional Neural Network
QDL	Quantum Deep Learning
QEST	Quantum-Enhanced Swin Transformer
QKNN	Quantum k-Nearest Neighbor
QLAM	Quantum Linear Algebra Module
QML	Quantum Machine Learning
QNN	Quantum Neural Network
QRF	Quantum Random Forest
QSVM	Quantum Support Vector Machine
QViT	Quantum Vision Transformer
RQ	Research Question
ViT	Vision Transformer
VQC	Variational Quantum Circuit
qSLP	Quantum Single-Layer Perceptron
QSAN	Quantum Self-Attention Network
QKSAN	Quantum Kernel Self-Attention Network
QHAN	Quantum Hard Attention Network
QAHAN	Quantum Annealing Hard Attention Network
KD	Knowledge Distillation
QOCNN	Quantum Optical Convolutional Neural Network

References

1. Albash, T.; Lidar, D.A. Adiabatic quantum computation. *Reviews of Modern Physics* **2018**, *90*.
2. Sekmen, Y.; PARLAK, B. A Systematic Literature Review of Quantum Machine Learning and Its Applications. *Computers and Computational Science* **2025**, *1*, 34–58.
3. Mohammadisavadkoohi, E.; Shafiabady, N.; Vakilian, J. A systematic review on quantum machine learning applications in classification. *IEEE transactions on artificial intelligence* **2025**.

4. Revythi, M.; Koukiou, G. Quantum Machine Learning and Deep Learning: Fundamentals, Algorithms, Techniques, and Real-World Applications. *Machine Learning and Knowledge Extraction* **2025**, *7*, 75.
5. Ullah, U.; Garcia-Zapirain, B. Quantum machine learning revolution in healthcare: a systematic review of emerging perspectives and applications. *IEEE Access* **2024**, *12*, 11423–11450.
6. Maheshwari, D.; Garcia-Zapirain, B.; Sierra-Sosa, D. Quantum machine learning applications in the biomedical domain: A systematic review. *Ieee Access* **2022**, *10*, 80463–80484.
7. Gupta, R.S.; Wood, C.E.; Engstrom, T.; Pole, J.D.; Shrapnel, S. A systematic review of quantum machine learning for digital health. *npj Digital Medicine* **2025**, *8*, 237.
8. Aishwarya, C.; Venkatesan, M.; Prabhavathy, P. A Scoping Survey of Quantum Machine Learning and Deep Learning for Real-World Applications. *Procedia Computer Science* **2025**, *258*, 633–646.
9. Hassan, S.S.; Khan, L.U.; Park, Y.M.; Guizani, M.; Han, Z.; Ratnarajah, T.; Hong, C.S. Quantum machine learning for 6g space-air-ground integrated networks: A comprehensive tutorial and survey. *IEEE Communications Surveys & Tutorials* **2025**.
10. Islam, M.M.; He, J.S. Quantum machine learning for computer vision: a survey. In Proceedings of the 2024 International Conference on Machine Learning and Applications (ICMLA). IEEE, 2024, pp. 1827–1832.
11. Kharsa, R.; Bouridane, A.; Amira, A. Advances in quantum machine learning and deep learning for image classification: A survey. *Neurocomputing* **2023**, *560*, 126843.
12. Klusch, M.; et al. Quantum artificial intelligence: a brief survey. *KI-K"unstliche Intelligenz* **2024**, *38*.
13. Ayoade, O.; Rivas, P.; Orduz, J. Artificial intelligence computing at the quantum level. *Data* **2022**, *7*, 28.
14. Huang, H.L.; Wu, D.; Fan, D.; Zhu, X. Superconducting quantum computing: a review. *Science China Information Sciences* **2020**, *63*, 180501.
15. IBM. IBM Quantum Roadmap 2033: Condor. https://www.ibm.com/quantum/blog/quantum-roadmap-2033, 2024.
16. IBM. IBM Quantum: Heron Processor. https://quantum.cloud.ibm.com/computers?processorType=Heron, 2024.
17. IBM Quantum. IBM Eagle Processor, 2023. Accessed: 2026-03-31.
18. Google. Google Willow Quantum Chip, 2024. Accessed: 2026-03-31.
19. Quantinuum. Quantinuum System Model H2, 2024. Accessed: 2026-03-31.
20. Rigetti Computing. Rigetti Documentation (PyQuil / QPU Access), 2024. Accessed: 2026-03-31.
21. D-Wave. D-Wave Quantum Systems, 2024. Accessed: 2026-03-31.
22. IBM. Qiskit: An Open-Source Quantum Computing Framework, 2024. Accessed: 2026-03-31.
23. Google Quantum AI. Cirq: A Python Framework for Quantum Circuits, 2024. Accessed: 2026-03-31.
24. Rigetti Computing. PyQuil Documentation, 2024. Accessed: 2026-03-31.
25. Microsoft. Q#: Quantum Programming Language Overview, 2024. Accessed: 2026-03-31.
26. Zhang, H.; Zhao, Q.; Zhou, M.; Feng, L.; Niyato, D.; Zheng, S.; Chen, L. A Survey of Quantum Transformers: Architectures, Challenges and Outlooks. *arXiv preprint arXiv:2504.03192* **2025**.
27. Wang, Y.; Liu, J. A comprehensive review of quantum machine learning: from NISQ to fault tolerance. *Reports on Progress in Physics* **2024**, *87*, 116402.
28. Yang, J.; Awan, A.J.; Vall-Llosera, G. Support vector machines on noisy intermediate scale quantum computers. *arXiv preprint arXiv:1909.11988* **2019**.
29. Liu, Y.; Arunachalam, S.; Temme, K. A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics* **2021**, *17*, 1013–1017.
30. Li, J.; Li, Y.; Song, J.; Zhang, J.; Zhang, S. Quantum support vector machine for classifying noisy data. *IEEE Transactions on Computers* **2024**, *73*, 2233–2247.
31. Zardini, E.; Blanzieri, E.; Pastorello, D. A quantum k-nearest neighbors algorithm based on the Euclidean distance estimation. *Quantum Machine Intelligence* **2024**, *6*, 23.
32. Guerrero-Estrada, A.Y.; Quezada, L.; Sun, G.H. Benchmarking quantum versions of the kNN algorithm with a metric based on amplitude-encoded features. *Scientific Reports* **2024**, *14*, 16697.
33. Srikumar, M.; Hill, C.D.; Hollenberg, L.C. A kernel-based quantum random forest for improved classification. *Quantum Machine Intelligence* **2024**, *6*, 10.
34. Ullah, U.; Maheshwari, D.; Gloyna, H.H.; Garcia-Zapirain, B. Severity Classification of COVID-19 Patients Data using Quantum Machine Learning Approaches. In Proceedings of the 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), 2022, pp. 1–6. https://doi.org/10.1109/ICECCME55909.2022.9987991.

 doi:10.20944/preprints202606.1015.v1

35. Kerenidis, I.; Landman, J.; Luongo, A.; Prakash, A. q-means: A quantum algorithm for unsupervised machine learning. *Advances in neural information processing systems* **2019**, *32*.
36. Poggiali, A.; Berti, A.; Bernasconi, A.; Del Corso, G.M.; Guidotti, R. Quantum clustering with k-means: A hybrid approach. *Theoretical Computer Science* **2024**, *992*, 114466.
37. Lin, Z.; Liu, H.; Tang, K.; Liu, Y.; Che, L.; Long, X.; Wang, X.; Fan, Y.a.; Huang, K.; Yang, X.; et al. Hardware-efficient quantum principal component analysis for medical image recognition. *Frontiers of Physics* **2024**, *19*, 51202.
38. He, C.; Li, J.; Liu, W.; Peng, J.; Wang, Z.J. A low-complexity quantum principal component analysis algorithm. *IEEE transactions on quantum engineering* **2022**, *3*, 1–13.
39. PRISMA Group. PRISMA Statement, 2026. Accessed: 2026-03-31.
40. Deng, L. The MNIST database of handwritten digit images for machine learning research. *IEEE Signal Processing Magazine* **2012**, *29*, 141–142.
41. Yang, J.; Shi, R.; Wei, D.; Liu, Z.; Zhao, L.; Ke, B.; Pfister, H.; Ni, B. MedMNIST v2: A large-scale lightweight benchmark for 2D and 3D biomedical image classification. *Scientific Data* **2023**, *10*, 41.
42. Krizhevsky, A.; Hinton, G. CIFAR-10 dataset. https://www.cs.toronto.edu/~kriz/cifar.html, 2009. Accessed: 2026-03-31.
43. Xiao, H.; Rasul, K.; Vollgraf, R. Fashion-MNIST: a novel image dataset for benchmarking machine learning algorithms. *arXiv preprint arXiv:1708.07747* **2017**.
44. Dataset, K. Brain Tumor MRI Dataset. https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset, 2020. Accessed: 2026-03-31.
45. Parkhi, O.M.; Vedaldi, A.; Zisserman, A.; Jawahar, C. Cats and Dogs. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012, pp. 3498–3505.
46. Altaisky, M. Quantum neural network. *arXiv preprint quant-ph/0107012* **2001**.
47. Zhao, C.; Gao, X.S. QDNN: deep neural networks with quantum layers. *Quantum Machine Intelligence* **2021**, *3*, 15.
48. Macaluso, A.; Orazi, F.; Klusch, M.; Lodi, S.; Sartori, C. A variational algorithm for quantum single layer perceptron. In Proceedings of the International Conference on Machine Learning, Optimization, and Data Science. Springer, 2022, pp. 341–356.
49. Otgonbaatar, S.; Datcu, M. Classification of remote sensing images with parameterized quantum gates. *IEEE Geoscience and Remote Sensing Letters* **2021**, *19*, 1–5.
50. Kanimozhi, T.; Sridevi, S.; et al. Brain tumor recognition based on classical to quantum transfer learning. In Proceedings of the 2022 International Conference on Innovative Trends in Information Technology (ICITIIT). IEEE, 2022, pp. 1–5.
51. Kanimozhi, T.; Sridevi, S.; et al. Iris tumor recognition based on hybrid classical and quantum neural network. In Proceedings of the 2023 International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS). IEEE, 2023, pp. 608–611.
52. Díaz-Santos, S.; Escanez-Exposito, D. Classical vs. Quantum machine learning for breast cancer detection. In Proceedings of the 2023 19th International Conference on the Design of Reliable Communication Networks (DRCN). IEEE, 2023, pp. 1–5.
53. Mathur, N.; Landman, J.; Li, Y.; Strahm, M.; Kazdaghli, S.; Prakash, A.; Kerenidis, I. Medical image classification via quantum neural networks. arXiv 2021. *arXiv preprint arXiv:2109.01831*.
54. Senokosov, A.; Sedykh, A.; Sagingalieva, A.; Kyriacou, B.; Melnikov, A. Quantum machine learning for image classification. *Machine Learning: Science and Technology* **2024**, *5*, 015040.
55. Konar, D.; Aggarwal, V.; Sarma, A.D.; Bhandary, S.; Bhattacharyya, S.; Cangi, A. Deep spiking quantum neural network for noisy image classification. In Proceedings of the 2023 International joint conference on neural networks (IJCNN). IEEE, 2023, pp. 1–10.
56. Matic, A.; Monnet, M.; Lorenz, J.M.; Schachtner, B.; Messerer, T. Quantum-classical convolutional neural networks in radiological image classification. In Proceedings of the 2022 IEEE International Conference on Quantum Computing and Engineering (QCE). IEEE, 2022, pp. 56–66.
57. Wang, M.; Shang, Y. Hybrid Vision Transformer and Quantum Convolutional Neural Network for Image Classification. *arXiv preprint arXiv:2510.12291* **2025**.
58. Florestiyanto, M.Y.; Kaswidjanti, W.; Fariszy, R. Evaluation of Emotion Detection Using CNN VGG16 and Hybrid QCNN for Enhancing Digital Content Personalization. In Proceedings of the RSF Conference Series: Business, Management and Social Sciences. Research Synergy Foundation, 2025, Vol. 5, p. 519.

59. Idress, W.M.; Zhao, Y.; Abouda, K.A.; Elhag, H.M. QCNN-Swin-UNet: Quantum Convolutional Neural Network Integrated with Optimized Swin-UNet for Efficient Liver Tumor Segmentation and Classification on Edge Devices. *Journal of Imaging Informatics in Medicine* **2025**, pp. 1–21.
60. Hur, T.; Kim, L.; Park, D.K. Quantum convolutional neural network for classical data classification. *Quantum Machine Intelligence* **2022**, *4*, 3.
61. Chen, G.; Chen, Q.; Long, S.; Zhu, W.; Yuan, Z.; Wu, Y. Quantum convolutional neural network for image classification. *Pattern Analysis and Applications* **2023**, *26*, 655–667.
62. Henderson, M.; Shakya, S.; Pradhan, S.; Cook, T. Quanvolutional neural networks: powering image recognition with quantum circuits. *Quantum Machine Intelligence* **2020**, *2*, 2.
63. Prajapat, S.; Tomar, M.; Kumar, P.; Kumar, R.; Vasilakos, A.V. Quantum Computing Meets Deep Learning: A QCNN Model for Accurate and Efficient Image Classification. *Mathematics* **2025**, *13*, 3148.
64. Parthasarathy, R.; Bhowmik, R.T. Quantum optical convolutional neural network: a novel image recognition framework for quantum computing. *IEEE access* **2021**, *9*, 103337–103346.
65. Tasnim, T.; Rahman, M.; Wu, F. Comparison of CNN and QCNN performance in binary classification of breast cancer histopathological images. In Proceedings of the 2024 IEEE International Conference on Big Data (BigData). IEEE, 2024, pp. 3780–3787.
66. Reka, S.S.; Karthikeyan, H.L.; Shakil, A.J.; Venugopal, P.; Muniraj, M. Exploring quantum machine learning for enhanced skin lesion classification: A comparative study of implementation methods. *IEEE Access* **2024**.
67. R¨oseler, P.; Schaudt, O.; Berg, H.; Bauckhage, C.; Koch, M. Efficient quantum convolutional neural networks for image classification: Overcoming hardware constraints. *arXiv preprint arXiv:2505.05957* **2025**.
68. Chen, G.; Long, S.; Yuan, Z.; Li, W.; Peng, J. Robustness and explainability of image classification based on QCNN. *Quantum Engineering* **2023**, *2023*, 2842217.
69. Meedinti, G.N.; Srirekha, K.S.; Delhibabu, R. A quantum convolutional neural network approach for object detection and classification. *arXiv preprint arXiv:2307.08204* **2023**.
70. Kerenidis, I.; Mathur, N.; Landman, J.; Strahm, M.; Li, Y.Y.; et al. Quantum vision transformers. *Quantum* **2024**, *8*, 1265.
71. Zhang, H.; Zhao, Q.; Zhou, M.; Feng, L. Hqvit: Hybrid quantum vision transformer for image classification. *arXiv preprint arXiv:2504.02730* **2025**.
72. Dutta, S.; Singh, H.; Shankhdhar, K.; Iyer, S. QMViT: A Mushroom is worth 16x16 Words. *arXiv preprint arXiv:2407.04708* **2024**.
73. Proceedings of the 26th International Conference on Computing in High Energy and Nuclear Physics (CHEP 2023). *EPJ Web of Conferences* **2024**, *295*, 8. Published online 06 May 2024, https://doi.org/10.1051/epjconf/202429512003.
74. Boucher, T.; Whittle, J.; Mazomenos, E.B. From O (n 2) to O (n) parameters: Quantum self-attention in vision transformers for biomedical image classification. In Proceedings of the International Workshop on Efficient Medical Artificial Intelligence. Springer, 2025, pp. 112–122.
75. Xu, X.F.; Xue, C.; Zhuang, X.N.; Wang, Y.J.; Sun, T.P.; Fang, Y.; Wang, J.C.; Liu, H.Y.; Wu, Y.C.; Chen, Z.Y.; et al. Towards Fault-Tolerant Quantum Deep Learning: Designing and Analyzing Quantum ResNet and Transformer with Quantum Arithmetic and Linear Algebra Primitives. *arXiv preprint arXiv:2402.18940* **2024**.
76. Born, J.; Skogh, F.; Rhrissorrakrai, K.; Utro, F.; Wagner, N.; Sobczyk, A. Quantum doubly stochastic transformers. *arXiv preprint arXiv:2504.16275* **2025**.
77. Comajoan Cara, M.; Dahale, G.R.; Dong, Z.; Forestano, R.T.; Gleyzer, S.; Justice, D.; Kong, K.; Magorsch, T.; Matchev, K.T.; Matcheva, K.; et al. Quantum vision transformers for quark–gluon classification. *Axioms* **2024**, *13*, 323.
78. Unlu, E.B.; Comajoan Cara, M.; Dahale, G.R.; Dong, Z.; Forestano, R.T.; Gleyzer, S.; Justice, D.; Kong, K.; Magorsch, T.; Matchev, K.T.; et al. Hybrid quantum vision transformers for event classification in high energy physics. *Axioms* **2024**, *13*, 187.
79. Tesi, A.; Gleyzer, S.; Matchev, K.T.; Matcheva, K.; Kong, K.; Dahale, G.R.; Magorsch, T. Quantum attention for vision transformers in high energy physics. In Proceedings of the International Workshop on Quantum Computing and Artificial Intelligence. Springer, 2025, pp. 16–31.
80. Xie, Z.; Yang, X.; Zhang, S.; Yang, J.; Zhu, Y.; Zhang, A.; Sun, H.; Dai, Q.; Li, L.; Liu, H.; et al. Quantum integration in swin transformer mitigates overfitting in breast cancer screening. *Scientific Reports* **2025**, *15*, 31589.
81. Dutta, S.; Innan, N.; Najafi, K.; Yahia, S.B.; Shafique, M. QUIET-SR: Quantum image enhancement transformer for single image super-resolution. *arXiv preprint arXiv:2503.08759* **2025**.

doi:10.20944/preprints202606.1015.v1

82. Qu, Z.; Zhou, M.; Sun, L.; Yu, Y.; Muhammad, G. QHSA-ViT: A Quantum Discrete Fourier Transform-Based Hierarchical Self-Attention Fusion Vision Transformer for Traffic Sign Recognition in Intelligent Vehicular Networks. *IEEE Internet of Things Journal* **2025**.
83. Zhao, R.X.; Shi, J.; Li, X. Qksan: A quantum kernel self-attention network. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2024**, *46*, 10184–10195.
84. Shi, J.; Zhao, R.X.; Wang, W.; Zhang, S.; Li, X. QSAN: A near-term achievable quantum self-attention network. *IEEE Transactions on Neural Networks and Learning Systems* **2024**, *36*, 13995–14008.
85. Evans, E.N.; Cook, M.; Bradshaw, Z.P.; LaBorde, M.L. Learning with sasquatch: a novel variational quantum transformer architecture with kernel-based self-attention. *arXiv preprint arXiv:2403.14753* **2024**.
86. Chen, F.; Feng, L.; Hu, Z.; Ren, Y. External Quantum Self-Attention Model. *IEEE Access* **2025**.
87. Chen, F.; Zhao, Q.; Feng, L.; Tang, L.; Lin, Y.; Huang, H. Quantum complex-valued self-attention model. *arXiv preprint arXiv:2503.19002* **2025**.
88. Zhao, R.X.; Shi, J.; Li, X. GQHAN: A Grover-inspired quantum hard attention network. *arXiv preprint arXiv:2401.14089* **2024**.
89. Zhao, R.X. QAHAN: A quantum annealing hard attention network. *arXiv preprint arXiv:2412.20930* **2024**.
90. Reejisha, A.; Mohan, A. Improving image classification with quantum neural network. In Proceedings of the 2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3). IEEE, 2023, pp. 1–12.
91. Rizvi, S.M.A.; Paracha, U.I.; Khalid, U.; Lee, K.; Shin, H. Quantum Machine Learning: Towards Hybrid Quantum-Classical Vision Models. *Mathematics* **2025**, *13*, 2645.
92. Kati, B.E.; K"uç"uksille, E.U.; Sarıman, G. Comparison of quantum deep learning methods for image classification. *M"uhendislik Bilimleri ve Tasarım Dergisi* **2025**, *13*, 90–106.
93. Schuld, M.; Killoran, N. Is quantum advantage the right goal for quantum machine learning? *Prx Quantum* **2022**, *3*, 030101.
94. Bowles, J.; Ahmed, S.; Schuld, M. Better than classical? the subtle art of benchmarking quantum machine learning models. *arXiv preprint arXiv:2403.07059* **2024**.