International Journal of Research in Engineering and Innovation (IJREI)
Journal home page: https://ijrei.com/

REVIEW PAPER

Variational quantum circuits for visual inspection in intelligent manufacturing: A critical review through a parameter-accounting framework

Tarina Jayant[1], Rakesh Kumar[2], Shankar Singh[2]

[1]*Department of Computer Science and Engineering, Vivekananda Institute of Professional Studies, Delhi, India*
[2]*Department of Mechanical Engineering, Sant Longowal Institute of Engineering and Technology, Longowal, Sangrur, Punjab, India*

Article Information

Received: 14 July 2026
Revised: 23 August 2026
Accepted: 01 September 2026
Available online: 03 September 2026

Keywords:

Variational Quantum Circuits
Quantum Machine Learning
Automated Visual Inspection
Additive Manufacturing
Parameter Efficiency
NISQ Noise Resilience

Abstract

Automated visual inspection underpins quality assurance in intelligent manufacturing. However, the convolutional and transformer models used for this purpose can be costly to deploy at the network edge, where latency and energy budgets are tight. Variational quantum circuits (VQCs) are often advanced as a low-parameter alternative, the reasoning being that n qubits carry a state space of size 2^n, so that comparatively few trainable angles might still describe an intricate decision surface. This selective narrative review asks whether that reasoning survives once resources are tallied from sensor input through to decision. A framework proposed that separates task-specific trainable parameters, stored non-trainable parameters, and fitted quantities, preprocessing compression, and quantum encoding reach, and apply it to representative hybrid pipelines. Three findings emerge. First, claimed quantum gains tend to arise in studies where the classical baseline differs from the quantum pipeline in parameter count, preprocessing applied, or the amount of tuning each received. A prominent controlled benchmark for small binary classification tasks generally favors classical models, although direct extrapolation of that result to industrial vision remains unproven. Second, the selected examples show few trainable angles within the circuit, but the representational burden is largely shifted onto classical preprocessing or frozen backbones. Third, scalability is constrained by both encoding reach and state preparation: a standard video clip can contain nearly 5×10^6 values, and generic amplitude loading may require millions of operations. These constraints are particularly relevant to video and sequential-image process monitoring. Six reporting requirements are proposed to make future comparisons reproducible.

1. Introduction

Automated visual inspection sits at the core of intelligent manufacturing. Cameras mounted inside laser powder-bed fusion and directed energy deposition machines capture melt-pool and layer-wise imagery, from which vision models identify porosity, balling, and keyholing [1]. Semi-supervised networks also monitor selective laser melting directly from video [2], and recent surveys document the rapid adoption of such methods for adaptive quality control [3]. Deployment is constrained, however, because inference occurs at the edge of the machine under strict latency and energy budgets. This constraint motivates the practical question addressed in this review: whether parameter-frugal alternatives can provide an effective solution. The models have grown steadily more expensive, from roughly 60 million weights in AlexNet [4] to 86 million in ViT-B/16 [5]. Video architectures add temporal computation and activation-memory costs, although their parameter counts do not necessarily scale linearly with clip length [6]. Classical work has responded with parameter-

Corresponding author: Tarina Jayant
Email Address: tarinajayant@gmail.com
https://doi.org/10.36037/IJREI.2026.10502

efficient designs [7]; attention has also turned to substrates whose capacity is not linear in the number of stored weights.

Variational quantum circuits (VQCs) are the leading quantum candidate [8]. In a VQC, a unitary $U(\theta)$ whose gates carry free parameters acts on a state that represents the input; measurement outcomes then drive a classical optimizer that revises θ. The attraction is essentially arithmetic: because n qubits span 2^n dimensions, a circuit holding only O(poly(n)) angles still operates on an exponentially large space, and might therefore rival a convolutional network's accuracy while storing far fewer weights.

The argument remains unsettled. Across 160 binary classification datasets, a systematic comparison of 12 quantum models found that classical baselines with no extensive task-specific tuning were usually the stronger performers [9]. More consequential for the present argument, stripping entanglement from the circuits typically costs little or nothing in terms of accuracy. These results apply to the small binary classification tasks studied and should not be generalized directly to industrial vision. Quantum convolutional networks have been shown separately to be efficiently simulable on the quantum data benchmarks used to validate them [10]. Both findings expose weaknesses in unmatched comparisons. A central confound is parameter accounting: a "60-parameter quantum classifier" fed by a frozen ResNet-18 has only 60 trainable quantum parameters but still retains an approximately 11.7-million-parameter backbone, while preprocessing and tuning asymmetries introduce further bias.

This review, therefore, (i) formalizes a parameter-accounting and matching framework and applies it to representative pipelines; (ii) evaluates the image and video evidence using this framework; (iii) characterizes how data encoding governs scalability; and (iv) assesses whether any efficiency advantage would survive noisy intermediate-scale Quantum (NISQ) hardware and examines the implications for inspection at the manufacturing edge.

2. Literature Review

2.1 Classical baselines

For a parameter-efficiency comparison, the relevant classical reference is not the largest available model but the efficiency frontier: EfficientNet-B0 [7] reaches 77.1% ImageNet top-1 accuracy with 5.3 million parameters. Industrial inspection models are drawn from this frontier rather than from research-scale architectures [3], so quantum claims must be read against it. Video adds a temporal axis: 3D convolutional networks [6] operate on clips of tens to hundreds of frames, decisive in Section 4.3.

2.2 Variational circuits and data encoding

A representative VQC maps the input onto a quantum state, acts on that state with tunable rotations and entangling gates, and then estimates a chosen set of observables [8]. Its expectation-value derivatives can be evaluated on hardware by the parameter-shift rule [11]. Under amplitude encoding, a D-dimensional input x maps to the normalized state

$$|\psi(x)\rangle = (1 / \|x\|_2) \sum_{i=0}^{D-1} x_i |i\rangle, \quad n = \lceil \log_2 D \rceil \qquad (1)$$

So, ten qubits suffice to represent a 1024-dimensional vector, but not necessarily to load it efficiently. The saving in register width is genuine, yet loading an arbitrary amplitude vector exactly costs $O(2^n)$ gates in the general case [12]: the exponential migrates from width into depth. Inputs with exploitable structure, or those required only approximately, may be prepared more cheaply. Angle encoding assigns one feature per rotation at constant encoding-layer depth. The encoding fixes the frequency spectrum of the learnable function, and the variational layers' coefficients [13], so expressive power depends on the encoding as much as on the parameters. The raw parameter count is a poor proxy for capacity.

2.3 Quantum architectures for visual recognition

The quantum convolutional neural network (QCNN) [14] interleaves convolution and pooling unitaries while the register progressively contracts. Its original target was quantum data, the motivating task being the identification of phases of matter. Its adaptation to classical images [15] applies the architecture in a setting where its original physical motivation does not directly apply, a point that matters in Section 4.4. Quantum convolutional layers [16] replace individual filters with small circuits over image patches; equivariant designs exploit label symmetry [17]; and quantum vision transformers adapt attention mechanisms to quantum circuits [18]. Transfer-learning hybrids [19], [20] dominate applied vision research and represent the setting in which parameter accounting is most often incomplete.

2.4 Prior reviews and the present gap

Several surveys map this territory. Kharsa et al. [21] synthesize quantum machine learning and deep learning for image classification, and Peral-García et al. [22] review quantum machine learning more broadly; both organize the field by algorithm family and report published accuracies. Neither imposes a common end-to-end parameter-accounting criterion, so reported differences cannot readily be attributed to the quantum component rather than to the classical scaffolding around it. Two further gaps motivate this review: scalability is often framed primarily in terms of qubit availability, even though data loading can dominate vision pipelines, and video evidence remains comparatively sparse in the selected literature.

3. Materials and Methods

Because this is a critical review rather than an experimental study, the "materials" are the published studies surveyed, and the "methods" are the accounting framework applied to them.

3.1 Scope and selection of studies

This review is narrative and critical rather than systematic: no exhaustive screening protocol or meta-analysis was applied.

Literature was retrieved through Jul. 19, 2026, from arXiv, IEEE Xplore, ScienceDirect, SpringerLink, and the APS journal collection. Keyword combinations included ("variational quantum circuit" OR "quantum neural network") AND ("image classification", "computer vision", "video classification", "visual inspection", "manufacturing", "parameter efficiency", "barren plateau", "noise resilience" or "classical simulability"). Backward and forward citation chaining were used to identify foundational and benchmarking papers. We prioritized work published since 2018 that trained a parameterized circuit on classical visual data and reported a trainable parameter count or included a classical comparator. Earlier foundational studies on classical methods, error mitigation, and state preparation were retained for context. Tensor-network, annealing, and quantum-native-input methods were excluded from the applied comparison. No formal deduplication or record-count log was retained; consequently, the review does not claim corpus completeness or report study frequencies. Section 4, therefore, describes patterns in the selected evidence.

3.2 The parameter-accounting and matching framework

The framework combines trainable-parameter accounting, stored-parameter accounting, preprocessing compression, and quantum encoding reach:

$$P_{train} = P_{pre} + P_q + P_{post} \quad (2)$$

where P_q is the number of task-specific trainable circuit angles, P_{pre} the trainable parameters of any classical front end, and P_{post} those of the readout head. For deployment accounting, frozen pre-trained parameters and other stored non-trainable fitted quantities (P_{froz})—including centering vectors and fixed circuit angles—must be included because they remain resident during inference. Parameter count is not, by itself, a measure of memory footprint; deployment comparisons should also report the numerical precision of each parameter group. If group j contains Pj parameters stored at bj bits each, its total parameter-storage requirement is ΣjPjbj/8 bytes:

$$P_{stored} = P_{train} + P_{froz} \quad (3)$$

$$|P_{train}{}^{Q} - P_{train}{}^{C}| / P_{train}{}^{C} \leq \delta,\ \delta = 0.10,\ P_{train}{}^{C} > 0 \quad (4)$$

Here δ = 0.10 is an author-proposed operational criterion, not a community standard. Tuning budgets are considered matched when Rtune = max(BQ, BC)/min(BQ, BC) ≤ 2, where BQ and BC denote the total numbers of evaluated hyperparameter configurations for the Quantum and classical pipelines, respectively. This ratio is also an author-proposed criterion. Results close to the boundary should also be reported at δ = 0.05 and δ = 0.20. Equation (4) assumes $P_{trainC} > 0$. The criterion assesses matching in terms of task-specific trainable parameters; P_{stored} is reported separately to evaluate storage and deployment efficiency. Comparisons failing these requirements are termed unmatched; those with insufficient information are indeterminate. Preprocessing, compression, and quantum encoding are measured separately:

$$\rho_{pre} = D_{native} / D_{circuit} \quad (5)$$

$$\rho_{enc} = D_{circuit} / D_{reach} \quad (6)$$

where D_{native} is the dimensionality before task-specific reduction, $D_{circuit}$ is the number of features presented to one circuit instance, and Dreach = 2^n for amplitude encoding or n for one angle-encoding layer. The latter is a per-layer operational definition: data re-uploading [23] can encode additional features across multiple layers, but it increases circuit depth and must be reported separately. Thus, ρ_{pre} measures classical compression, while ρ_{enc} measures whether the circuit input fits the chosen encoding.

4. Results and Discussion

4.1 Characteristics of the surveyed literature

Four features characterize the selected literature. MNIST and Fashion-MNIST dominate, usually as binary or four-class subsets; register sizes range from four to twelve qubits; most evaluations are simulation-based; and comparatively few studies represent video. QCNN [15] and quanvolutional [16] designs are evaluated predominantly on MNIST-scale inputs. At the same time, transfer-learning hybrids place small circuits behind frozen classical front ends, ranging from compact task-trained CNNs [19] to ImageNet-scale backbones containing 10^6–10^7 parameters [20].

4.2 The parameter-efficiency claim under end-to-end accounting

Table 1 applies Eqs. (2) and (3), while Table 2 applies Eqs. (5) and (6) and records whether Eq. (4) can be evaluated. These are analytical examples computed from disclosed assumptions rather than empirical results copied from individual papers. The PCA-8 example uses $P_{pre} = 0$ trainable parameters, an 8 × 784 projection matrix plus a 784-element centering vector in P_{froz}, $P_q = 48$, and $P_{post} = 90$. The fixed quanvolutional circuit contributes 24 stored angles, and the downstream CNN is represented as $P_{post} \approx 10^5$. The ResNet-18 hybrid uses P_{pre} = 4,104 for the 512-to-8 projection; it applies to a 3,072-dimensional CIFAR-10-scale input, $P_{post} = 90$, $P_q = 48$, and an approximately 11.69-million-parameter frozen backbone. Circuit-level trainable-parameter frugality is visible, but Eq. (4) is not applicable because no task-matched comparator is specified.

In the controlled benchmark examined here, the apparent advantage diminished when comparisons were more closely matched. The benchmark in [9] found classical models generally superior on its small binary-classification tasks and reported that ablating entanglement often left performance unchanged. Entanglement is one potential source of non-classical representational structure in a VQC; unchanged performance after its ablation therefore weakens, but does not eliminate, evidence that a result depends specifically on quantum resources. Because the benchmark does not evaluate industrial image or video inspection, extrapolation to vision remains indirect.

Table 1: Illustrative trainable- and stored-parameter accounting under Eqs. (2) and (3). Values are calculated from the assumptions stated in Section 4.2; ≈ denotes an order-of-magnitude value.

Configuration	P_{pre}	P_q	P_{post}	P_{froz}	P_{train}	P_{stored}
PCA-8 + VQC	0	48	90	7,056	138	7,194
Fixed quanvolution + CNN	0	0	$\approx 10^5$	24	$\approx 10^5$	≈100,024
Frozen ResNet-18 + VQC	4,104	48	90	11.69×10^6	4,242	$\approx 11.694 \times 10^6$
EfficientNet-B0 scale ref.	-	-	-	0	5.3×10^6	5.3×10^6

Table 2: Preprocessing compression, encoding, reach, and applicability of the parameter-matching criterion. The VQC examples use single-layer angle encoding.

Configuration	ρ_{pre}	ρ_{enc}	Eq. (4) status
PCA-8 + 8-qubit VQC	98	1	N/A—no task-matched comparator
Four-feature quanvolutional patch	1/patch	1	N/A—no task-matched comparator
ResNet projection + 8-qubit VQC	384	1	N/A—no task-matched comparator
EfficientNet-B0 scale reference	—	—	Scale reference only

4.3 Scalability and the encoding bottleneck

Qubit availability is the limitation most often named; in vision, however, the tighter constraints are usually how the data are loaded and how deep the resulting circuit becomes. A 28×28 grayscale image is 784-dimensional, a CIFAR-10 image 3,072-dimensional, a 224×224 RGB crop 150,528-dimensional, a 16-frame 112×112 RGB clip 602,112-dimensional, and a 32-frame 224×224 RGB clip 4,816,896-dimensional. The largest clip requires $\lceil \log_2 4{,}816{,}896 \rceil = 23$ qubits merely to represent its amplitudes. Generic exact preparation of an arbitrary 23-qubit state can require on the order of $2^{23} \approx 8.39 \times 10^6$ elementary operations [12]. Structured or approximate state-preparation methods may reduce this cost, but their assumptions and approximation error must be reported. Under a fixed-width 20-qubit angle-encoding layer, the native-input-to-circuit ratio increases from 784/20 = 39.2 for MNIST to 4,816,896/20 ≈ 240,845 for the largest clip—approximately 3.8 orders of magnitude.

For the small registers available today, classical reduction is therefore difficult to avoid. The circuit actually classifies the reduced representation, and the quality of that reduction can govern the accuracy that gets reported. A stronger reduction may improve tractability while weakening attribution of performance to the quantum component. Reporting both ρ_{pre} and ρ_{enc} makes this distinction explicit.

4.4 Trainability and noise resilience

Where the ansatz is sufficiently expressive, the variance of the cost gradient decays exponentially in the qubit count [24]; holding relative precision fixed then demands exponentially more samples, at least under the assumptions on which the barren-plateau analysis rests. Local costs in shallow circuits can retain O(1/poly(n)) variance [25], and certain QCNN-style ansätze have been shown to avoid exponentially vanishing gradients under specified locality and initialization assumptions [26]. However, under particular assumptions, provable trainability has also been associated with efficient classical simulability [10], [27]. Trainability and quantum advantage may therefore be in tension for some model families, but no universal equivalence has been established.

Hardware noise compounds the limit. The IBM superconducting processors considered here contain approximately 120–156 physical qubits [28], [29], and IBM's roadmap describes Nighthawk workloads that can execute up to 5,000 two-qubit gates before large-scale fault tolerance [28]. As a deliberately simplified illustration, an independent two-qubit error probability of 10^{-3} applied independently across 5,000 two-qubit gates gives $(1 - 10^{-3})^{5000} \approx 6.7 \times 10^{-3}$. This is not a circuit-fidelity prediction: coherent and correlated errors, crosstalk, idling, measurement noise, and device topology are omitted. Layer error per gate is likewise an aggregate benchmarking metric and should not be used directly to substitute for an independent gate-failure probability [29]. Error mitigation [30] incurs sampling overhead that grows with noise, while noise-induced plateaus [31] can arise independently of cost locality. IBM targets its first fault-tolerant system for 2029 [28]; this is a vendor roadmap rather than a demonstrated capability at present.

4.5 Video-specific evidence and open research questions

Video-focused variational-circuit studies remain comparatively limited in the selective literature examined here [32]. Because this review did not use an exhaustive systematic search, it does not claim that no full-resolution benchmark exists. Three approaches emerge: frame-wise circuits with classical temporal aggregation, which leave the temporal model classical; quantum recurrent architectures [33], which have not yet been demonstrated at an inspection-relevant visual scale; and quantum 3D convolution [32], which must address the T-fold increase in native data volume discussed in Section 4.3, although streaming or recurrent implementations need not load all T frames into one circuit simultaneously. Several considerations nevertheless justify continued investigation. Current benchmarks may be too small for any separation to appear: registers that fit current hardware correspond to feature spaces that classical models readily handle. Hence, a null result at eight qubits provides weak evidence about performance at eighty qubits. Equivariant circuits [17] are motivated by inductive bias rather than parameter counting, an argument unaffected by Section 4.2. Furthermore, rigorous separations exist for constructed problems [34], making the question empirical rather than settled.

4.6 Implications for inspection at the manufacturing edge

Table 3 compares different manufacturing-inspection modalities in terms of their input structure, temporal characteristics, and implications for visual quality control resource accounting. It distinguishes melt-pool, layer-wise,

surface-defect, and volumetric inspection approaches, emphasizing data dimensionality, processing requirements, compression needs, contextual information, and computational considerations for effective monitoring. For visual inspection, the consequences are direct. Many in situ monitoring applications use video or sequential image streams [2], whose native data volume increases by a factor of T relative to a single frame when spatial resolution and channel count are fixed; streaming or frame-wise architectures may nevertheless process the frames sequentially. Melt-pool imagery may be acquired at high frame rates [1], [2], whereas circuit inference and expectation estimation require repeated shots. Putting a variational head after a frozen backbone leaves deployed storage untouched for as long as that backbone must stay resident. Parameter count is therefore only one component of deployment efficiency: state-preparation depth, one- and two-qubit gate counts, total circuit depth, shots per inference, circuit evaluations per gradient, classical–quantum communication latency, error-mitigation overhead, memory, and infrastructure-level energy consumption must also be reported. On present evidence, compact classical networks on the relevant task-specific efficiency frontier [7] remain the better-supported route to parameter-efficient inspection. Therefore propose six reporting requirements: (1) disclose P_{pre}, P_q, P_{post}, P_{froz}, P_{train} and P_{stored}; (2) report Eq. (4) at $\delta = 0.05$, 0.10 and 0.20 or state that it is not applicable; (3) disclose ρ_{pre} and ρ_{enc} and benchmark against a task-appropriate efficiency-frontier model [7]; (4) report state-preparation depth, gate counts and shot budgets per inference and gradient step; (5) include an entanglement-ablation control [9]; and (6) report a stimulability check where the architecture admits one [10].

Table 3: Manufacturing-inspection modalities and VQC resource implications [1-3].

Inspection modality	Native input	Temporal structure	Accounting implication
Melt-pool monitoring	H × W × C per frame	High-rate stream	Report per-frame loading, shots, and end-to-end latency
Layer-wise powder-bed inspection	H × W × C per layer	Sequential layers	Report compression and retained inter-layer context
Surface-defect inspection	H × W × C crop	Still image or short sequence	Compare against task-matched compact CNNs
Volumetric/ CT inspection	H × W × Z × C	3D volume	Report dimensional reduction and state-preparation complexity

5. Conclusions

This selective narrative review examined whether variational quantum circuits offer a parameter-efficient path to visual recognition and inspection at the manufacturing edge once the full resource ledger is in place. Application of Eqs. (2), (3), (5) and (6), together with assessment of the applicability of Eq. (4), show that circuit-level parameter frugality does not by itself establish training or deployment efficiency: classical preprocessing, readout layers, frozen backbones, data loading, and repeated measurement must be reported separately. A 32-frame 224×224 RGB clip contains 4,816,896 values, requires at least 23 amplitude-encoding qubits, and may require approximately 8.39×10^6 elementary operations for generic exact state preparation [12]. Evidence of practical noise resilience at an inspection-relevant scale remains insufficient. The contribution of this review is therefore a transparent resource-accounting and reporting framework rather than a claim of demonstrated quantum advantage.

Three limitations qualify these conclusions: coverage is selective, and substantial positive-result publication bias may exist; the literature is dominated by simulation, so hardware behavior is inferred; and Tables 1 and 2 are illustrative and computed from stated assumptions rather than surveyed parameter counts, because such counts are seldom fully disclosed.

Four research directions follow. A parameter-matched benchmark suite, with mandatory disclosure of frozen parameters and shot budgets, would enable the field's central claim to be tested rather than asserted. Encoding schemes that permit sub-exponential state preparation for structured or approximately represented inputs constitute a high-priority research target; generic arbitrary inputs do not receive this scaling without additional assumptions. Video-specific designs that exploit temporal redundancy deserve attention for process monitoring, where successive frames are highly correlated. Finally, quantum-native sensing data may avoid the classical-to-quantum state-preparation bottleneck, returning VQCs to the setting for which Quantum convolutional architectures were devised [14], although interface fidelity, noise, measurement, and readout costs remain.

Acknowledgements and Declarations

The authors thank the Department of Computer Science and Engineering, SLIET Longowal, for institutional support.

Funding: None

Conflicts of interest: none declared.

References

[1] L. Scime and J. Beuth, "Anomaly detection and classification in a laser powder bed additive manufacturing process using a trained computer vision algorithm," Additive Manufacturing, vol. 19, pp. 114–126, 2018, doi: 10.1016/j.addma.2017.11.009.

[2] B. Yuan, B. Giera, G. Guss, I. Matthews, and S. McMains, "Semi-supervised convolutional neural networks for in-situ video monitoring of selective laser melting," in Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV), Waikoloa, HI, USA, 2019, pp. 744–753, doi: 10.1109/WACV.2019.00084.

[3] L. Chen et al., "In-situ process monitoring and adaptive quality enhancement in laser additive manufacturing: A critical review," Journal of Manufacturing Systems, vol. 74, pp. 527–574, 2024, doi: 10.1016/j.jmsy.2024.04.013.

[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. Adv. Neural Inf. Process. Syst. (NeurIPS), Lake Tahoe, NV, USA, 2012, pp. 1097–1105.

[5] A. Dosovitskiy et al., "An image is worth 16×16 words: Transformers for image recognition at scale," in Proc. Int. Conf. Learn. Represent. (ICLR), Virtual Conference, 2021.

[6] J. Carreira and A. Zisserman, "Quo vadis, action recognition? A new

model and the Kinetics dataset," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, 2017, pp. 4724–4733, doi: 10.1109/CVPR.2017.502.

[7] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proc. Int. Conf. Mach. Learn. (ICML), Long Beach, CA, USA, 2019, pp. 6105–6114.

[8] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, "Parameterized quantum circuits as machine learning models," Quantum Science and Technology, vol. 4, no. 4, Art. no. 043001, 2019, doi: 10.1088/2058-9565/ab4eb5.

[9] J. Bowles, S. Ahmed, and M. Schuld, "Better than classical? The subtle art of benchmarking quantum machine learning models," arXiv:2403.07059, 2024, doi: 10.48550/arXiv.2403.07059.

[10] P. Bermejo, P. Braccia, M. S. Rudolph, Z. Holmes, L. Cincio, and M. Cerezo, "Quantum convolutional neural networks are effectively classically simulable," PRX Quantum, vol. 7, no. 2, Art. no. 020304, 2026, doi: 10.1103/8qt9-72ts.

[11] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," Physical Review A, vol. 98, no. 3, Art. no. 032309, 2018, doi: 10.1103/PhysRevA.98.032309.

[12] M. Möttönen, J. J. Vartiainen, V. Bergholm, and M. M. Salomaa, "Transformation of quantum states using uniformly controlled rotations," Quantum Information and Computation, vol. 5, no. 6, pp. 467–473, 2005, arXiv:quant-ph/0407010.

[13] M. Schuld, R. Sweke, and J. J. Meyer, "Effect of data encoding on the expressive power of variational quantum-machine-learning models," Physical Review A, vol. 103, no. 3, Art. no. 032430, 2021, doi: 10.1103/PhysRevA.103.032430.

[14] I. Cong, S. Choi, and M. D. Lukin, "Quantum convolutional neural networks," Nature Physics, vol. 15, no. 12, pp. 1273–1278, 2019, doi: 10.1038/s41567-019-0648-8.

[15] T. Hur, L. Kim, and D. K. Park, "Quantum convolutional neural network for classical data classification," Quantum Machine Intelligence, vol. 4, no. 1, Art. no. 3, 2022, doi: 10.1007/s42484-021-00061-x.

[16] M. Henderson, S. Shakya, S. Pradhan, and T. Cook, "Quanvolutional neural networks: Powering image recognition with quantum circuits," Quantum Machine Intelligence, vol. 2, no. 1, Art. no. 2, 2020, doi: 10.1007/s42484-020-00012-y.

[17] P. San Sebastian Sein, M. Cañizo, and R. Orús, "Image classification with rotation-invariant variational quantum circuits," Physical Review Research, vol. 7, no. 1, Art. no. 013082, 2025, doi: 10.1103/PhysRevResearch.7.013082.

[18] E. A. Cherrat, I. Kerenidis, N. Mathur, J. Landman, M. Strahm, and Y. Y. Li, "Quantum vision transformers," Quantum, vol. 8, Art. no. 1265, 2024, doi: 10.22331/q-2024-02-22-1265.

[19] J. Kim, J. Huh, and D. K. Park, "Classical-to-quantum convolutional neural network transfer learning," Neurocomputing, vol. 555, Art. no. 126643, 2023, doi: 10.1016/j.neucom.2023.126643.

[20] A. Mari, T. R. Bromley, J. Izaac, M. Schuld, and N. Killoran, "Transfer learning in hybrid classical-quantum neural networks," Quantum, vol. 4, Art. no. 340, 2020, doi: 10.22331/q-2020-10-09-340.

[21] R. Kharsa, A. Bouridane, and A. Amira, "Advances in quantum machine learning and deep learning for image classification: A survey," Neurocomputing, vol. 560, Art. no. 126843, 2023, doi: 10.1016/j.neucom.2023.126843.

[22] D. Peral-García, J. Cruz-Benito, and F. J. García-Peñalvo, "Systematic literature review: Quantum machine learning and its applications," Computer Science Review, vol. 51, Art. no. 100619, 2024, doi: 10.1016/j.cosrev.2023.100619.

[23] A. Pérez-Salinas, A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre, "Data re-uploading for a universal quantum classifier," Quantum, vol. 4, Art. no. 226, 2020, doi: 10.22331/q-2020-02-06-226.

[24] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," Nature Communications, vol. 9, Art. no. 4812, 2018, doi: 10.1038/s41467-018-07090-4.

[25] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, "Cost function dependent barren plateaus in shallow parametrized quantum circuits," Nature Communications, vol. 12, Art. no. 1791, 2021, doi: 10.1038/s41467-021-21728-w.

[26] A. Pesah, M. Cerezo, S. Wang, T. Volkoff, A. T. Sornborger, and P. J. Coles, "Absence of barren plateaus in quantum convolutional neural networks," Physical Review X, vol. 11, no. 4, Art. no. 041011, 2021, doi: 10.1103/PhysRevX.11.041011.

[27] M. Cerezo et al., "Does provable absence of barren plateaus imply classical simulability?," Nature Communications, vol. 16, Art. no. 7907, 2025, doi: 10.1038/s41467-025-62360-2.

[28] IBM, "IBM delivers new quantum processors, software, and algorithm breakthroughs on path to advantage and fault tolerance," IBM Newsroom, Nov. 12, 2025. [Online]. Available: https://newsroom.ibm.com/2025-11-12-ibm-delivers-new-quantum-processors,-software,-and-algorithm-breakthroughs-on-path-to-advantage-and-fault-tolerance. Accessed: Jul. 19, 2026.

[29] IBM Quantum, "Our first Nighthawk QPU and latest Heron QPU are now available," IBM Quantum Documentation, Jun. 5, 2026. [Online]. Available: https://quantum.cloud.ibm.com/docs/guides/latest-updates. Accessed: Jul. 19, 2026.

[30] K. Temme, S. Bravyi, and J. M. Gambetta, "Error mitigation for short-depth quantum circuits," Physical Review Letters, vol. 119, no. 18, Art. no. 180509, 2017, doi: 10.1103/PhysRevLett.119.180509.

[31] S. Wang et al., "Noise-induced barren plateaus in variational quantum algorithms," Nature Communications, vol. 12, Art. no. 6961, 2021, doi: 10.1038/s41467-021-27045-6.

[32] K. Blekos and D. Kosmopoulos, "A quantum 3D convolutional neural network with application in video classification," in Advances in Visual Computing (ISVC 2021), Lecture Notes in Computer Science, vol. 13017. Cham, Switzerland: Springer, 2021, pp. 601–612, doi: 10.1007/978-3-030-90439-5_47.

[33] S. Y.-C. Chen, S. Yoo, and Y.-L. L. Fang, "Quantum long short-term memory," in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Singapore, 2022, pp. 8622–8626, doi: 10.1109/ICASSP43922.2022.9747369.

[34] Y. Liu, S. Arunachalam, and K. Temme, "A rigorous and robust quantum speed-up in supervised machine learning," Nature Physics, vol. 17, no. 9, pp. 1013–1017, 2021, doi: 10.1038/s41567-021-01287-z.

Cite this article as: Tarina Jayant, Rakesh Kumar, Shankar Singh, Variational quantum circuits for visual inspection in intelligent manufacturing: A critical review through a parameter-accounting framework, International Journal of Research in Engineering and Innovation, 10 (5), (2026), 196-201. https://doi.org/10.36037/IJREI.2026.10502