Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Quantum Laplacian Pyramid Segmentation for Multi-Class Mushroom Classification

Jeyapriya N.[1], Hannah Inbarani H.[2], Jamaludeen A.[3]

[1,3]Research Scholar, Department of Computer Science, Periyar University, Salem, India.
[2]Professor, Department of Computer Science, Periyar University, Salem, India.

E-mail: [1]nj.priya03@gmail.com, [2]hhinba@periyaruniversity.ac.in, [3]jamaludeenphd@gmail.com

Orcid ID: [1]0009-0004-5439-5447, [2]0000-0002-2956-3507, [3]0009-0000-9700-7417

Abstract

Quantum computing proposes a new paradigm for image analysis by efficiently representing features in high-dimensional Hilbert spaces. In the first stage, we propose a novel Quantum Laplacian Pyramid Segmentation (QLPS) module to segment complex structural mushroom profiles by exploiting a parameter-efficient, single-layer, 4-qubit variational quantum circuit (VQC) with cyclic ring topology. The second stage combines this quantum-refined mask with a raw high-resolution image track using an element-wise Hadamard product to create a background-suppressed region of interest (ROI) that allows a downstream EfficientNetB0 network to focus entirely on clean biological textures. Rigorous experimental evaluations confirm the system objectives, both individually and collectively. We observe that the QLPS module correlates well with the generated ground truth masks, with a mean Pixel Accuracy (mPA) of 93.79% and a Dice coefficient of 94.26%. The integrated classifier achieves a final classification accuracy of 95.88% across four different categories. Most importantly, comparative ablation testing quantifies the direct architectural contribution of the quantum segmentation step to the classification improvement. The removal of background clutter through the QLPS branch produces a significant 9.56% classification improvement relative to standard classical networks trained on raw, noisy images.

Keywords: Image classification, EfficientNetB0, Laplacian Residual, Image Segmentation, Quantum Entanglement.

1. Introduction

Many factors, including human intuition, survival requirements, perceived dangers, and curiosity, have shaped how humans engage with mushrooms. Before the scientific approach, people in the ancient world relied on oral traditions and visual cues to differentiate poisonous from safe mushrooms. Psychedelic mushrooms (including psilocybin and related tryptamines) have a long history of traditional and ceremonial use. A new study reveals that they may be helpful for neurological, psychiatric, and associated problems [1].

Identifying mushrooms is biological triage, not aesthetic judgement. Wild mushrooms may be covered with dirt, insects, or other fungi. They are often in shade or camouflaged against the bark of a tree. Classification is difficult owing to poor lighting or position. Segmentation is crucial to separate the mushroom from its background and to demarcate the boundary between

DOI: https://doi.org/10.36548/jiip.2026.3.011
Received: 15.05.2026, received in revised form: 14.06.2026, accepted: 04.07.2026, published: 15.07.2026
© 2026 Inventive Research Organization. This is an open access article under the Creative Commons Attribution-NonCommercial International (CC BY-NC 4.0) License

safety and neurological risk. Venomous and hallucinogenic creatures copy edible species with extraordinary precision, not by accident but by evolutionary mimicry, operating as biological impostors masquerading as food [2].

Segmentation is the main protection in mushroom identification. It separates the hazards of nature from its benefits. At this stage, even the most advanced classification algorithms can be compromised by false positives, which can lead to dangerous species being mistaken for edible species due to their shape and colour [3].

Traditional segmentation networks like U-Net and SegNet employ symmetric encoder-decoder structures, which have shown effectiveness in a variety of biomedical and natural image segmentation applications [4,5]. Skip connections are widely used in encoder-decoder architectures for image segmentation to blend high-level semantic features from deeper layers with fine-grained spatial data from shallower layers. This approach greatly enhances the reconstruction of complex object boundaries; even for complex backgrounds, precise segmentation can be achieved [6].

In Attention U-Net, attention gates (AGs) automatically learn to focus on relevant feature maps, thus improving the conventional U-Net architecture. In Attention U-Net [7], attributes from the encoder and decoders are deliberately merged to create task-relevant activations. In addition, U-Net++ introduced redesigned nested skip pathways and deep supervision to reduce the semantic gap between encoder and decoder feature maps. The model was validated on four different datasets, including lung nodule, colon polyp, liver, and cell nuclei segmentation, achieving an Intersection over Union (IoU) of 92.63%, demonstrating the effectiveness of enhanced feature fusion strategies [6]. TransUNet, announced in 2021, combined CNN-based local feature extraction with global contextual modelling to analyse medical images using the Transformer architecture. The combined Transformer encoder-decoder structure had an average Dice score of 88.39%, outperforming encoder-only (88.11%) and decoder-only (85.30%) configurations, highlighting the benefit of learning local and global representations [8].

The DeepLabV3 semantic segmentation model utilises atrous convolution and Atrous Spatial Pyramid Pooling (ASPP) to effectively extract contextual information across several scales. This strategy combines deep feature extraction with better border localisation and obtained good accuracy [9]. More recently, Transformer-based semantic segmentation models have achieved very promising results. SegFormer adopts a positional-encoding-free hierarchical Transformer encoder and a lightweight all-MLP decoder and has been assessed on ADE20K, Cityscapes, and COCO-Stuff datasets. For example, SegFormer-B4 obtains 50.3% mIoU on ADE20K with only 64M parameters, and SegFormer-B5 obtains 84.0% mIoU on the Cityscapes validation set. This demonstrates the efficiency of hierarchical Transformer representations [10]. Similarly, Mask2Former developed a single masked-attention framework for panoptic, instance and semantic segmentation tasks. It scored 57.8 PQ on COCO panoptic segmentation, 50.1 AP on COCO instance segmentation, and 57.7 mIoU on ADE20K semantic segmentation on COCO, Cityscapes, ADE20K and Mapillary Vistas datasets, illustrating the adaptability of mask-based Transformer designs [11].

Although skip connections have been used in models like U-Net, they are inadequate for capturing multi-scale contextual information, which limits their efficiency in the segmentation of mushroom species with modest anatomical similarities and natural concealment. Attention gates increase feature prioritisation in the Attention U-Net. However, this also requires more memory and training, and it may be less effective in chaotic or low-contrast environments where fine details are hard to find. Though impressive, DeepLabV3 might be improved in identifying and separating items in photos with complex textures or structures. Its shallow decoder module and lack of upsampling refinement may cause inaccurate edge segmentation. With nested skip pathways and deep supervision, U-Net++ enhances feature aggregation but utilises more memory and computational expense due to its more sophisticated architecture. TransUNet can adequately simulate long-range dependencies, but the added Transformer blocks greatly increase processing needs and may not preserve fine boundary details in high-resolution natural images. Similarly, SegFormer and Mask2Former show excellent global contextual modelling capabilities but may struggle with the precise localization of small objects and fine-grained boundaries in highly cluttered natural scenes, especially when subtle visual differences exist between mushroom species and their surrounding environment.

To tackle these problems, QLPS utilises a combination of multi-scale Laplacian pyramid decomposition and quantum inspired encoding to preserve global boundaries and local texture variations. Unlike conventional methods, QLPS can correctly identify mushroom shapes with complex patterns such as pores and speckles. Hence, QLPS produces segmentation findings with better structural accuracy and biological significance.

2. Related Work

The Generalised Laplacian Pyramid (GLP) method enhances image fusion by maintaining spectral and spatial features through multi-level decomposition. We propose a fresh and enhanced GLP technique that allows the incorporation of spatial structures. Hence, the texture of the pansharpened images was better maintained, and the edges were more defined [12]. We used a Laplacian Pyramid based feature fusion approach in a semantic segmentation framework to improve contextual representation and maintain spatial information. The approach effectively solves the problem of representing fine structures at varying resolutions of a picture [13]. Laplacian Pyramid approaches emphasise features in colour images by maintaining sharp edges and enhancing local contrast at various scales [14].

The evolution of Quantum Image Processing (QIMP) can be divided into a few phases. Between 2003 and 2008, the Real Ket representation, the Qubit Lattice model for quantum picture storage and retrieval, and other fundamental models were proposed. These works established the theoretical foundation for essential quantum operations and representations [15]. The next stage (2010-2015) was marked by considerable advancement, which resulted in the development of NEQR and FRQI in 2011. NEQR is the Novel Enhanced Quantum Representation, and FRQI is the Flexible Representation of Quantum Images. During this time, a number of image processing techniques were developed, i.e. compression, watermarking, encryption, segmentation, and edge detection [16,17]. The investigations into quantum image security techniques, including steganography, encryption, and watermarking, from 2016 to 2019 have gone beyond image representations and processing. The advancement of QIMP during this period could be consolidated through surveys and evaluations [18].

Between 2020 and 2023, advancements in Quantum Principal Component Analysis (QPCA) for image denoising, quantum edge detection, and the application of quantum machine learning techniques such as Quantum Support Vector Machines (QSVM) for image classification have enhanced the integration of sophisticated quantum algorithms into practical applications. During this period, research has been conducted on real-time image enhancement, hybrid quantum-classical frameworks, and the development of Quantum Convolutional Neural Networks (QCNNs) for image recognition. Quantum methods have been applied in medical image processing,particularly in MRI scan cancer diagnosis [19]. The current stage is typified by the search for hybrid quantum-classical solutions and for scaling algorithms for high-resolution and sophisticated image processing. This stage is marked by ongoing research in quantum image compression, enhancement, and pattern recognition [19,20].

Quantum circuits can be combined with operations of the Laplacian Pyramid to improve the performance of segmentation via quantum enhanced edge detection and multi-scale feature processing. Parameterised quantum gates can extract high frequency data using quantum-inspired Laplacian filters to enhance the delineation of mushroom cap and gill segmentation borders and preserve texture.

3. Proposed Methodology

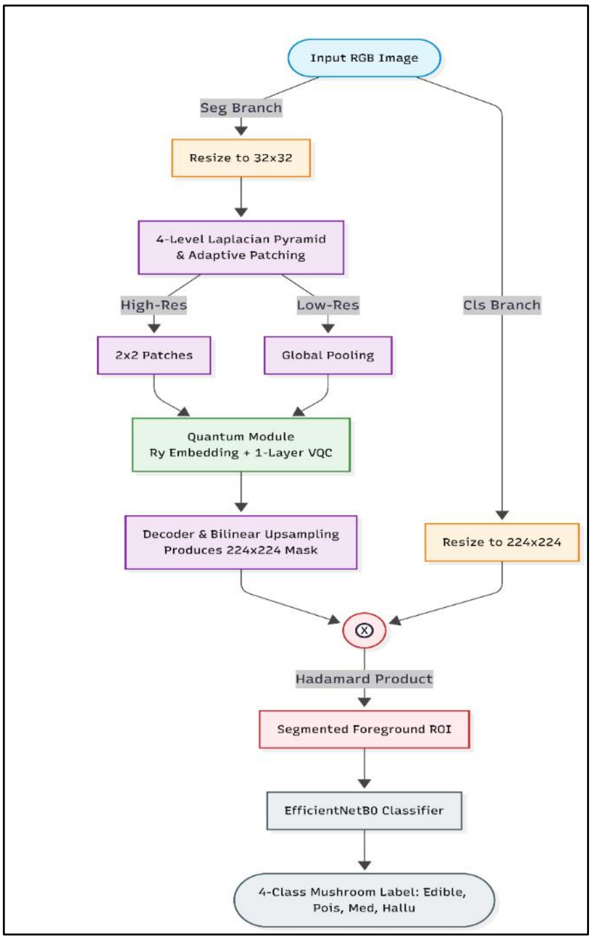

Figure 1. Quantum laplacian pyramid segmentation framework

Figure 1 shows the overall process of the proposed Quantum Laplacian Pyramid Segmentation (QLPS) system. This system combines classical multi-scale image decomposition with quantum-enhanced feature extraction for accurate region-of-interest (ROI) segmentation and noise-resistant multi-class mushroom classification. The input pipeline is first divided into a dual resolution approach: a classification stream where the original image is resized to 224 x 224 and a segmentation stream where the image is scaled to 32 x 32 to maintain a uniform spatial layout. In the segmentation branch, features are decomposed into a

4-Level Laplacian Pyramid, using adaptive patch extraction that channels high-resolution layers into localised 2 x 2 patches and low-resolution layers with global average pooling.

The standardised 4-D vectors are then fed into the integrated Quantum Module, where they are embedded with R_Y angle in a 4-qubit state space, passing through a parameter-efficient single-layer Variational Quantum Circuit (VQC) with a cyclic CNOT ring topology and Pauli-Z expectation fusion. The resulting representations are then passed to a unified Decoder and Bilinear Up-sampling block to reconstruct the final 224 × 224 binary segmentation masks. This mask is merged with the parallel classification track with an element-wise Hadamard product (⊙) to remove undesired background noise. The segmented foreground ROI preserves the high-resolution textures by suppressing the background, allowing the EfficientNetB0 classifier to concentrate only on identifying physiologically relevant patterns for the final prediction of 4-class mushroom labels.

3.1 Input Data Preprocessing

The RGB mushroom images are adaptively preprocessed to enhance the quantum feature extraction module and the downstream classical deep learning classification network. To reduce computational overhead during the execution of the Quantum Laplacian Pyramid, input images are standardised to a target spatial resolution of 32x32 pixels for the multi-scale segmentation stage. The high-resolution originals are resized to the native input size of EfficientNetB0 (224x224x3) for the subsequent classification stage to preserve fine-grained structural and textual features such as gill patterns and spore profiles.

The pixel intensities for the segmentation branch are converted to floating point format and normalised to [0, 1] to ensure numerical stability during pyramid formation.

The normalised 32 x 32 images are then divided into four Laplacian pyramid layers to extract the high-frequency spatial components. The resulting multi-scale residual matrices are flattened adaptively depending on their structural resolution into 4-dimensional vectors, normalized to the range [0,1] and mapped onto a 4-qubit parameterized variational quantum circuit.

Algorithm for QLPS

Input:
- Preprocessed RGB input image $I_0 \in \mathbb{R}^{32\times32\times3}$
- Ground Truth Mask $M \in \{0,1\}^{H\times W}$

Output:
- Predicted segmentation maps
- Evaluation metrics including mPA, Dice, and IoU

Step 1: Construction of the Laplacian Pyramid

Initially, the input image is normalized by scaling pixel intensities to the range [0,1]:

$$I \leftarrow \frac{I}{255.0}$$

The variables are initialized as:

$$Current \leftarrow I_{res(32x32)}, LP \leftarrow [\,]$$

For each pyramid level $l = 0$ to $L - 1$, the following operations are performed:

1. Apply average pooling downsampling to obtain a reduced-resolution image:

$$Down \leftarrow pyrDown(Current)$$

2. Upsample the reduced image back to the original resolution:
$$Up \leftarrow pyrUp(Down)$$
3. Compute the Laplacian representation by subtracting the upsampled image from the current image:
$$Laplacian \leftarrow Current - Up$$
4. Store the resulting Laplacian feature map in the pyramid list LP.
$$LP \leftarrow [LP, \text{Laplacian}]$$
5. Update the current image for the next iteration:
$$Current \leftarrow Down$$

Step 2: Quantum Pyramid Refinement Module (QPRM)

For each global Laplacian residual map $L_l \in LP$, a quantum refinement procedure is performed.

Initialize an empty quantum feature map $Q_l$with dimensions $H_l \times W_l$.

To preserve both global structural information and local texture details, adaptive spatial processing is applied:

- **If $H_l > 4$:**
  Partition the high-resolution residual map into non-overlapping local sub-patches $P_{(i,j)} \in \mathbb{R}^{2\times2}$, and flatten each patch into a 4-dimensional vector:
$$x = [x_0, x_1, x_2, x_3]^T$$
- **If $H_l \le 4$:**
  The entire residual map is processed globally without patch partitioning.
  For a 4×4map, apply 2×2average pooling; for a native 2×2map, directly flatten it into a 4-dimensional vector.

The feature vector is normalized to [0,1] and encoded into a 4-qubit quantum circuit using rotation gates:
$$R_Y(x_k^{norm})$$
Quantum entanglement is introduced using cyclic Controlled-NOT (CNOT) gates.

Pauli-Z expectation values are measured from all qubit (z_0, z_1, z_2, z_3).

The quantum response is computed as:
$$q = \frac{z_0 + z_1 + z_2 + z_3}{4}$$
The obtained quantum value $q$is assigned to the corresponding spatial location in Q_l.

Step 3: Multi-Scale Feature Fusion and Segmentation Reconstruction

Resize all quantum-refined feature maps to a common spatial resolution using bilinear interpolation:
$$Q'_l = U(Q_l), l \in \{0,1,2,3\}$$
Concatenate the resized feature maps to form a unified multi-scale quantum feature tensor:
$$Q_{fusion} = \text{Concat}(Q'_0, Q'_1, Q'_2, Q'_3)$$
Pass the fused tensor through the decoder to reconstruct the segmentation mask:
$$\widehat{M} = Decoder(Q_{fusion})$$

Step 4: Resolution Mapping and Segmentation Mask Generation

To bridge the resolution gap for the downstream classifier, upscale the predicted mask to match the original native classifier resolution via a bilinear upsampling transformation layer:
$$M_{high} = \mathcal{U}_{bilinear} \text{ (M, size = 224x224)}$$
Apply sigmoid activation and thresholding to produce the final binary mask:
$$\text{P} \leftarrow \text{Sigmoid } (M_{high})$$
$$\hat{\text{M}} \leftarrow (\text{P} > 0.5)$$

The resulting binary mask $\widehat{M}$ is directly used for segmentation evaluation. No additional morphological post-processing, such as erosion, dilation, opening, or closing, is applied after thresholding.

End

3.2 Construction of the Laplacian Pyramid

The Laplacian Pyramid is used to extract multiscale structure information from the input mushroom image. The basic purpose of the pyramid is to separate the low frequency information of the image from high-frequency details like edges, boundaries, textures and fine morphological characteristics. These high frequency residuals contain discriminative information that is very useful for semantic segmentation and further classification [13].

An RGB image is given, and a Laplacian pyramid of four levels is recursively built. At each level, the image is downsampled to gather coarse structural information, then reconstructed to the original resolution. The Laplacian residual map is the difference between the original image and the reconstructed image, which contains the high-frequency features lost in the downsampling [12]. The proposed framework employs a hybrid global–local strategy, which entails the preservation of complete Laplacian residual maps during pyramid construction and the application of 2 × 2 patch partitioning exclusively during quantum refinement for 4-qubit encoding.

3.2.1 Average-Pooling Downsampling

We construct a multi-resolution pyramid by recursively decreasing the spatial resolution of the input image using a 2 x 2 average-pooling process. By removing small, localized fluctuations in intensity, average pooling preserves important low-frequency data. This is similar to a low-pass filtering procedure. With this, we can obtain the rough figure representations required to construct Laplacian pyramids.

An average pooling operator can be used to obtain each level:

$$I_{l+1} = \mathrm{D}\,(I_l) = \frac{1}{4}\sum_{i=0}^{1} \sum_{j=0}^{1} I_l\,[2m+i, 2n+j] \tag{1}$$

D represents 2×2 average pooling.

Beginning with an input image of dimensions 32×32, the recursive downsampling procedure produces increasingly diminutive image representations:

$$32 \text{ X } 32 \rightarrow 16 \text{ X } 16 \rightarrow 8 \text{ X } 8 \rightarrow 2\ X\ 2$$

Information preservation and computational efficiency are balanced when 2x2 average pooling is employed. It gradually decreases the dimensionality of the features while preserving the fundamental structural characteristics necessary for the succeeding quantum feature extraction stage.

3.2.2 Bilinear Upsampling and Interpolation

Bilinear interpolation is employed to upsample the lower resolution image to the original resolution dimensions of level l, enabling the reconstruction and comparison of images across levels. This ensures pixel-level alignment for the next residual computation. The upsampled image $\hat{\boldsymbol{I}}_l$ is:

$$\hat{I}_l\text{=u}(I_{l+1})\text{=Bilinear}(I_{l+1,}\ size =\ H_l \times W_l) \tag{2}$$

3.2.3 Laplacian Residual

High frequency information is separated by subtracting the interpolated coarse image from its corresponding finer-level version to produce the Laplacian detail map. These residuals encode texture, edges, and other fine-grained information that is important for tasks such as semantic segmentation.

The Laplacian detail map at level l is:

$$L_l = I_l - \hat{I}_l \tag{3}$$

The pyramid is recursively built for levels l=0,1,2,3 resulting in $\{L_0, L_1, L_2, L_3\}$.

The proposed QLPS framework follows a hierarchical global-to-local processing strategy. The Laplacian pyramid first preserves global multi-scale structural information by extracting residual maps containing boundaries, shapes, and texture details at different resolutions. Subsequently, during the Quantum Pyramid Refinement Module, high-resolution residual maps are partitioned into 2 × 2 local patches for quantum encoding, while low-resolution maps are processed globally. This design enables the simultaneous preservation of global contextual information and enhancement of local texture interactions for accurate segmentation and classification.

3.3 Quantum Pyramid Refinement Module (QPRM)

In general, quantum machine learning (QML) has been classified into four paradigms, Classical-Classical (CC), Classical-Quantum (CQ), Quantum-Classical (QC) and Quantum-Quantum (QQ) based on the nature of the data and the computational framework used [21]. Among them, the Classical-Quantum (CQ) paradigm exploits classical data as input, and applies quantum algorithms or circuits to learn and manipulate features. The proposed framework follows a hybrid Classical–Quantum (CQ) learning paradigm, where classical mushroom images and Laplacian pyramid features are fed into a trainable Variational Quantum Circuit (VQC), and the extracted representations are integrated back into the classical segmentation network. This approach uses basic quantum phenomena like qubits, superposition and entanglement to make features more representative and is still compatible with classical deep learning frameworks [20].

Following this classical-quantum hybrid learning paradigm, the proposed Quantum Pyramid Refinement Module (QPRM) refines the multi-scale Laplacian pyramid features using a Trainable Variational Quantum Circuit (VQC). In contrast to fixed quantum feature mappings, the suggested module has learnable quantum parameters that are concurrently optimized with the classical decoder, leading to an adaptive quantum feature translation. The objective of QPRM is to investigate whether a parameterized quantum feature space can improve the representation of high-frequency structural information such as edges, textures, and fine morphological patterns in mushroom images.

Each Laplacian pyramid level $L_l \in \mathbb{R}^{H \times W}$ is independently processed using a four-qubit variational quantum circuit with trainable parameters, producing compact quantum-enhanced representations that are subsequently used for segmentation reconstruction.

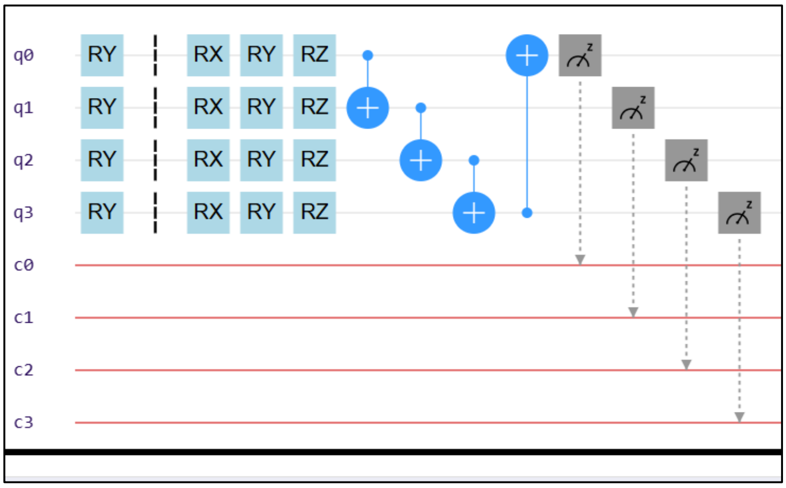

Figure 2. Proposed variational quantum circuit

Figure 2 shows the proposed variational quantum circuit, illustrating state preparation, a single parameterized variational ansatz layer with single qubit rotations, a cyclic entangling layer and final z-basis measurement. It aims to improve the fine spatial details obtained at each level of the Laplacian Pyramid through the application of quantum operations. This module utilizes quantum feature transformation at the patch level for high-resolution residual maps, while low-resolution maps are processed globally.

A 4-qubit quantum circuit recalibrates spatial structure and local textures for each Laplacian level $L_l \in R^{H_l X W_l X C}$

3.3.1 Adaptive Spatial Patch Partitioning and Dimensionality Standardization

To interface the multiscale spatial components derived from the Laplacian Pyramid with the 4-qubit Quantum Pyramid Refinement Module (QPRM), a variable, resolution-dependent patch-extraction mechanism is implemented. Let L_l denote the structural residual feature map at pyramid level $l \in \{0,1,2,3\}$. Due to successive downsampling, the spatial dimensions scale down progressively from 32x32 to 2x2. To preserve local high-frequency geometric characteristics without structural decay, the patch dimension P_l for each level is dynamically allocated.

For high-resolution layers (L_0 *and* L_1), the feature maps are partitioned into non-overlapping local sub-patches $P_{(i,j)} \in \mathbb{R}^{2x2}$. Each sub-patch is flattened into a 4-dimensional feature array:

$$V_{ij} = [x_0, x_1, x_2, x_3]^T \tag{4}$$

where V_{ij} represents the flattened 4-dimensional column vector extracted from the non-overlapping local spatial sub-patch at grid position (i, j), composed of localized pixel values (for $k \in \{0,1,2,3\}$).

Conversely, for the deepest, low-resolution layers (L_2 *and* L_3). where spatial dimensions are already structurally compact ($\leq$ 4 X 4), patch partitioning is bypassed. Instead, these layers are evaluated globally. If a deep layer expresses a 4 x 4 resolution (L_2), it undergoes a localized 2 x 2 average pooling operation, whereas a native 2 x 2 layer (L_3) is directly flattened.

This dual-strategy standardization ensures that all inputs entering the quantum embedding layer systematically map onto a 4-dimensional feature vector, perfectly aligning with the 4-qubit initialization state space (H=$C^2 \otimes C^2 \otimes C^2 \otimes C^2$) without loss of structural continuity. Pixel intensity values are normalized to [0,1] before being passed to the embedding layer:

$$x_k^{norm} = \frac{x_k}{255} \tag{5}$$

Finally, the normalized patch is flattened into a 4-dimensional vector and used as input to the quantum encoding stage.

3.3.2 Trainable Quantum State Encoding

An angle embedding approach is employed to encode each patch vector into a four-qubit quantum state. Encoding classical data into quantum Hilbert space can be done in various ways in quantum machine learning. These techniques include basis encoding, amplitude encoding, angle (rotation) encoding, Instantaneous Quantum Polynomial-time (IQP) encoding, and data re-uploading, which vary in their representational power and circuit complexity.

Among these approaches, angle encoding is chosen in the proposed QPRM for its computing efficiency, hardware compatibility, and applicability for gradient-based optimisation in hybrid quantum-classical systems [17].

Each pixel x_i^{norm} is embedded using the rotation gate R_Y, which rotates the qubit state around the Y-axis on the Bloch sphere:

$$|\,\psi_{enc} > \; = \otimes_{i=0}^{3} R_y\,(x_k^{norm})\,|\,0 >_i \tag{6}$$

The choice of $R_Y$rotation-based encoding is motivated by its ability to directly map real-valued inputs into quantum state amplitudes, thereby preserving intensity information of image features. This enables effective representation of continuous-valued pixel data within a 2^4-dimensional Hilbert space [21], facilitating nonlinear feature transformation through subsequent quantum circuit evolution.

3.3.3 Single-Layer Variational Quantum Circuit (VQC)

The proposed VQC consists of a single variational layer comprising trainable RX, RY, and RZ rotations followed by cyclic CNOT entanglement.

The single variational layer consists of:

- Trainable rotations: R_X, R_Y, R_Z
- Entanglement using cyclic CNOT structure

The parameterized transformation of the single variational quantum layer is defined as:

$$U(\theta) = U_{ent} \cdot U_{rot}(\theta) \tag{7}$$

where $\theta_l$represents trainable parameters.

The full circuit is expressed as:

$$| \psi_{out}\rangle = \mathrm{U}\ (\theta)\ | \psi_{enc}\rangle \tag{8}$$

The entanglement structure follows a cyclic CNOT topology:

$$U_{ent} = \mathrm{CNOT}_{0\to1} \cdot \mathrm{CNOT}_{1\to2} \cdot \mathrm{CNOT}_{2\to3} \cdot \mathrm{CNOT}_{3\to0} \tag{9}$$

This cyclic topology ensures global interaction among all qubits, enabling multi-feature dependency modelling across patch elements.

Convolutional Neural Networks (CNNs) and MLPs use layered linear operations and pointwise nonlinear activations to modify features in Euclidean space, which limits their capacity to describe complicated feature correlations without increasing network depth. The Variational Quantum Circuit (VQC) employs unitary evolution and entanglement to transform classical inputs into a 2^4-dimensional Hilbert space, facilitating nonlinear feature modification [23]. In contrast to separable 1x1 convolutions or MLP layers, the VQC uses entanglement-induced non-separability to model interdependent patch features in a single transformation. Quantum encoding projects pixel intensities into quantum states using rotation gates, resulting in implicit nonlinear feature extension through quantum interference. Variational quantum circuits powered by parameterised unitary operations and entanglement structure enable richer feature representations with regulated inductive bias [23].

3.3.4 Measurement and Feature Extraction

Following circuit evolution, we measure the evolved quantum state with Pauli-Z observables, which are commonly used in variational quantum circuits for machine learning tasks due to their stability and interpretability. In general, quantum measurements can be performed using different observables, for example Pauli-X, Pauli-Y, Pauli-Z operators or probability-based computational basis measurements. However, the suggested framework chooses Pauli-Z since it is compatible with continuous-valued feature extraction and gradient-based optimisation techniques.

The expectation value of the Pauli-Z operator for the k^{th}qubit is computed as

$$Z_k = \langle\psi_{out}|Z_k|\psi_{out}\rangle \tag{10}$$

The structural output for the corresponding location is taken as the mean of the four measurements:

$$\mathrm{q}_{\mathrm{i,j}} = \frac{1}{4}\sum_{k=0}^{3} Z_k \tag{11}$$

This scalar value maps directly into the structural coordinate array, producing a refined quantum-processed level $Q_l \in \mathbb{R}^{H_l x\, W_l}$. All refined quantum feature maps {Q0, Q1, Q2, Q3} are preserved as multi-scale spatial representations. Instead of collapsing them into a fixed-dimensional vector, the feature maps are progressively resized to a common spatial resolution and concatenated along the channel dimension to form a dense multi-scale quantum feature tensor. This tensor serves as the input to the decoder for segmentation reconstruction.

3.4 Multi-Scale Quantum Feature Tensor

The quantum-refined outputs from all Laplacian pyramid levels are aggregated to form a unified multi-scale quantum feature tensor for segmentation reconstruction. Unlike fixed-length feature vectors, the proposed framework preserves the spatial structure of quantum-enhanced features at each pyramid level.

Let Q_0, Q_1, Q_2, and Q_3denote the quantum-refined feature maps from the four pyramid levels. Since these feature maps have different spatial dimensions, they are resized to a common resolution using bilinear interpolation:

$$Q_l' = U(Q_l), l \in \{0,1,2,3\} \tag{12}$$

The aligned feature maps are then concatenated along the channel dimension:

$$Q_{fusion} = \text{Concat}(Q_0', Q_1', Q_2', Q_3') \tag{13}$$

This multi-scale quantum feature tensor preserves both global structural information and local texture details, providing an enriched representation for accurate decoder-based segmentation reconstruction.

3.5 Decoder-Based Segmentation Reconstruction

The extracted quantum feature vector is provided to a fully connected decoder network for segmentation mask reconstruction.

The decoder consists of three fully connected layers:

$$16 \rightarrow 128 \rightarrow 512 \rightarrow 1024$$

with ReLU activation functions.

The final decoder output is reshaped into a segmentation mask:

$$\widehat{M} = Reshape\big(F(Q)\big) \tag{14}$$

Where,

$$\widehat{M} \in \mathbb{R}^{32\times32} \tag{15}$$

represents the predicted segmentation mask. The binary segmentation masks generated after sigmoid thresholding were directly used for evaluation without any additional post-processing.

3.6 Supervised Optimization

The predicted mask is compared with the ground truth mask using Binary Cross-Entropy loss:

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)] \tag{16}$$

where y_i denotes the ground truth mask pixel and $\hat{y}_i$ denotes the predicted probability.

The network parameters are optimized using the Adam optimizer with a learning rate of 0.001.

3.7 QLPS Output Integration with EfficientNetB0 for Classification

The upscaled binary segmentation mask M_{high} generated by the QLPS module is utilized to focus the classification model structurally on the mushroom's biologically relevant parts, such as the cap, gills, spore region, volva, and textural zones. To eliminate background noise, a composite fusion layer computes the element-wise Hadamard product (⊙) between the original high-resolution input image $I_{orig} \in \mathbb{R}^{224 \, x \, 224 \, x \, 3}$ and the upscaled mask:

$$I_{classifier\_input} = I_{orig} \odot M_{high} \tag{17}$$

This filtering mechanism eliminates superfluous background clutter by mathematically nullifying out-of-boundary pixel vectors while accentuating the spatial qualities required for identifying the target categories: edible, poisonous, medicinal, or hallucinogenic.

The tensor is then dimensionally corrected to (224 X 224 X 3) and fed directly into EfficientNetB0 [25]. EfficientNetB0 is a lightweight convolutional neural network with great performance. It achieves a good balance between accuracy and computing efficiency by using the compound scaling of network depth, width, and resolution. The network operates only on this specific isolated area of interest (ROI) in the foreground and extracts high-level semantic characteristics from the cleaned and segmented input. This target-specific processing suppresses background noise and drives the deep convolutional layers to focus only on discriminative, biologically relevant patterns, hence leading to a significant optimization of the classification performance.

4. Results and Discussion

4.1 Description of the Dataset

The original Kaggle dataset comprises 94 distinct species of mushrooms, encompassing 27,436 images. However, the insufficient number of samples fails to accurately represent numerous species, resulting in a scattered distribution. Determining whether a mushroom is edible, poisonous, medicinal, or hallucinogenic is typically more critical than identifying its species in practical mushroom identification. This resulted in the classification of the 94 species into four categories with significant biological consequences.

The curated dataset contains 2,000 images of mushrooms, split evenly among edible, poisonous, hallucinogenic, and medicinal types. The edible category comprises Agaricus augustus, Amanita velosa, Cerioporus squamosus, Coprinus comatus, Flammulina velutipes, Mycena haematopus, and Suillus spraguei. The poisonous category includes Amanita persicina, Chlorophyllum brunneum, Coprinopsis lagopus, Ganoderma curtisii, Glooeophyllum sepiarium, and Stereum ostrea. The hallucinogenic category contains Psilocybe azurescens, Psilocybe caerulescens, Psilocybe cyanescens, Psilocybe ovoideocystidiata, and Psilocybe pelliculosa. Finally, the medicinal category comprises Ganoderma applanatum, Grifola frondosa, Mycena leaiana, Pseudohydnum gelatinosum, Stereum hirsutum, and Suillus luteus.

Species variety within categories enhances intra-class variability, and an even distribution of images allows for fair and accurate model training and assessment.

Prior to dataset partitioning, duplicate and extremely similar images were eliminated to forestall data leakage. We used visual inspection and perceptual hashing to determine image similarity. Before the train-validation-test splitting, samples with a hash similarity greater than 95% were removed as duplicates.

Reliable class annotation and improved reproducibility of dataset generation were achieved through taxonomic consistency using renowned mycological sources, such as Index Fungorum, MycoBank, and Dictionary of the Fungi. An automated mushroom identification system, an evaluation of biodiversity, and a public health risk assessment could all benefit from this standardized approach, which also improves the dataset's integrity. Following the curation process, the dataset was divided into three subsets: training, validation, and testing, with a ratio of 80:10:10. The split was done for each category separately to maintain class uniformity across all subsets. Thus, the final dataset has 1600 training images, 200 validation images, and 200 testing images, which allows for a fair and reproducible evaluation methodology for the proposed framework.

4.2 Platform Specifications

The tests were conducted on a 64-bit Windows 11 operating system, using a high-performance computer environment equipped with a 13th Gen Intel(R) Core (TM) i7-13620H processor at a speed of 2.40 GHz and 32 GB of RAM, which enabled efficient execution of deep learning, image segmentation, and classification tasks. The system was equipped with an NVIDIA GeForce RTX 4070 Laptop GPU with 8 GB of dedicated graphics RAM, together with Intel(R) UHD Graphics with 128 MB of memory, for better computational performance and parallel processing ability. The proposed algorithms and models were implemented on the Anaconda development platform with common Python libraries and frameworks such as TensorFlow, Scikit-Learn, Pandas, Matplotlib, Seaborn, and other deep learning tools for data processing, visualisation, model training and performance evaluation.

4.2.1 Training Hyperparameters

Table 1 presents the hyperparameter tuning experiments performed to determine the optimal training configuration. Among optimizers such as SGD, RMSprop, and Adam, Adam was selected for its faster convergence and adaptive learning capability. The most stable convergence and best validation results were obtained with a 0.001 learning rate. After studying 16 and 32 batch sizes, 32 achieved the best balance among training stability, memory efficiency, and generalisation performance. For multi-class classification, Binary Cross-Entropy Loss performed better than MSE and Focal Loss. ReduceLROnPlateau was chosen as it could adaptively reduce the learning rate when validation performance stagnated, thereby improving convergence and avoiding early stagnation.

Table 1. Hyperparameter tuning

Parameter	Value
Optimizer	Adam
Learning Rate	0.001
Loss Function	Binary Cross Entropy
Learning Rate Scheduler	ReduceLROnPlateau
Epochs	20

Dropout Rate	0.3
Activation Function	ReLU
Random Seed	42
Batch Size	32

4.3 Visual Evaluation of QLPS Segmentation

A qualitative comparison of the proposed QLPS framework with many state-of-the-art segmentation models, including U-Net, U-Net++, DeepLabV3+, Attention U-Net, TransUNet, SegFormer, and Mask2Former is presented in Figure 3. Four representative groups of mushrooms, edible, poisonous, medicinal, and hallucinogenic, are analysed. The selected examples include diverse lighting conditions, complicated backgrounds and occluded object views, which demonstrate the ability of the proposed QLPS framework to preserve subtle morphological patterns in challenging imaging scenarios.

Although modern designs have been successful to a certain degree in biological picture segmentation, the image data of wild mushrooms poses additional challenges such as cluttered backgrounds, variations in illumination, irregular textures and minute morphological differences among species. As shown in Figure 3, many models have differences in border marking and saving fine details, especially for images with complex architecture and low-contrast parts.

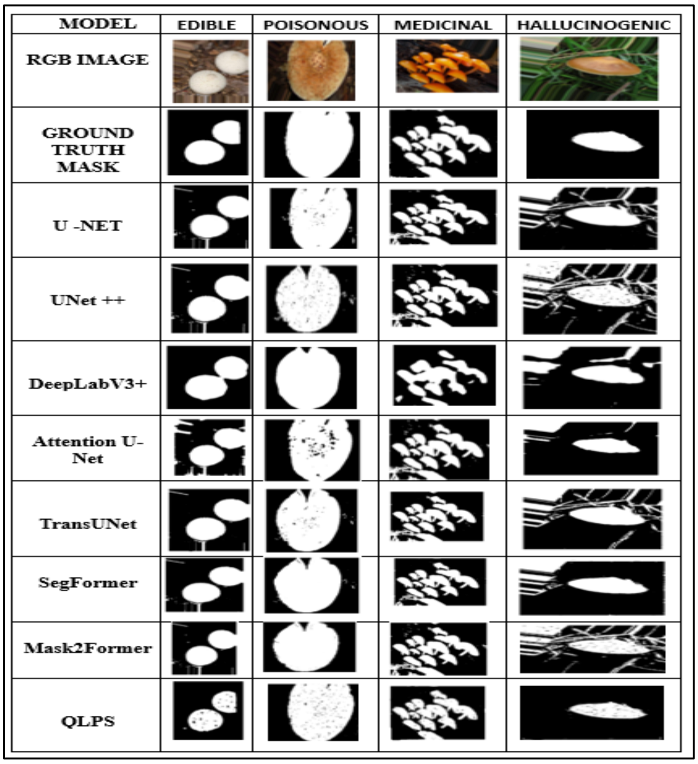

Figure 3. Visual evaluation of mushroom segmentation

The subtle internal features, such as surface grain texture, cap characteristics, and spore-related attributes, are essential for distinguishing between edible, poisonous, medicinal, and hallucinogenic species. Conventional segmentation models primarily emphasise spatial boundary definition and may struggle to preserve subtle intrastructural texture variations in challenging imaging situations. Conversely, QLPS employs quantum-enhanced multi-scale processing to derive biologically relevant intrastructural information, hence enabling more precise and classification-aware segmentation at all four levels.

4.3.1 Ground Truth Annotation

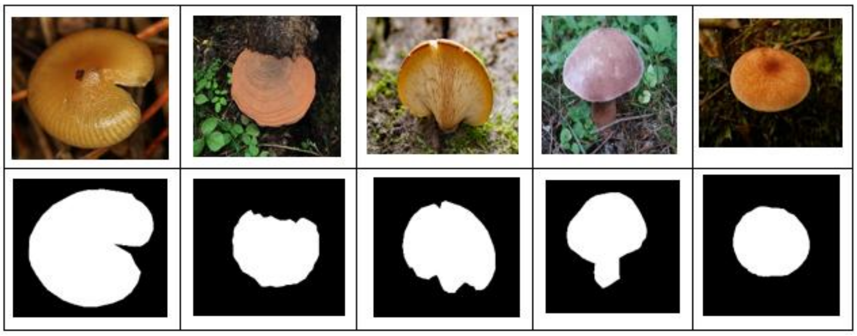

Figure 4. Ground truth masks

Ground truth segmentation masks were produced by manual annotation by a domain expert in plant pathology to acquire dependable pixel-level labels for segmentation assessment. In total, 2000 mushroom images were annotated into 4 classes; edible, poisonous, hallucinogenic, and medicinal. In LabelMe, annotations were built as polygons. The perimeter of each mushroom was carefully drawn and saved in JSON. These annotations were subsequently translated into binary ground-truth masks and used for segmentation evaluation. Figure 4 shows representative examples of manually annotated ground truth masks.

4.4 Quantitative Evaluation

The performance of the segmentation was evaluated based on several important parameters, including mean Pixel Accuracy (mPA), mean Intersection over Union (mIoU), and Dice Coefficient (DC). Each indicator provides a different insight into the quality, consistency, and accuracy of the segmented outputs. TP, FP, FN and TN represent true positives, false positives, false negatives, and true negatives, respectively computed in the pixel space.

- Pixel Accuracy:

$$\text{mPA} = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_{i+FN_i}} \tag{18}$$

- Dice Coefficient:

$$\text{Dice} = \frac{2.TP}{2.TP+FP+FN} \tag{19}$$

- Jaccard Index (IoU):

$$\text{IoU} = \frac{TP}{TP+FP+FN} \tag{20}$$

Table 2. Segmentation performance comparison with benchmark models

Model	mPA	mIoU	DC
U – NET	74.56%	71.92%	76.84%
UNet ++	79.68%	66.74%	78.96%
DeepLabV3+	87.42%	84.16%	89.35%
Attention U-Net	82.30%	67.85%	77.42%
TransUNet	85.10%	69.42%	78.25%

SegFormer	84.26%	72.35%	82.14%
Mask2Former	88.34%	77.28%	84.90%
QLPS	93.79%	92.41%	94.26%

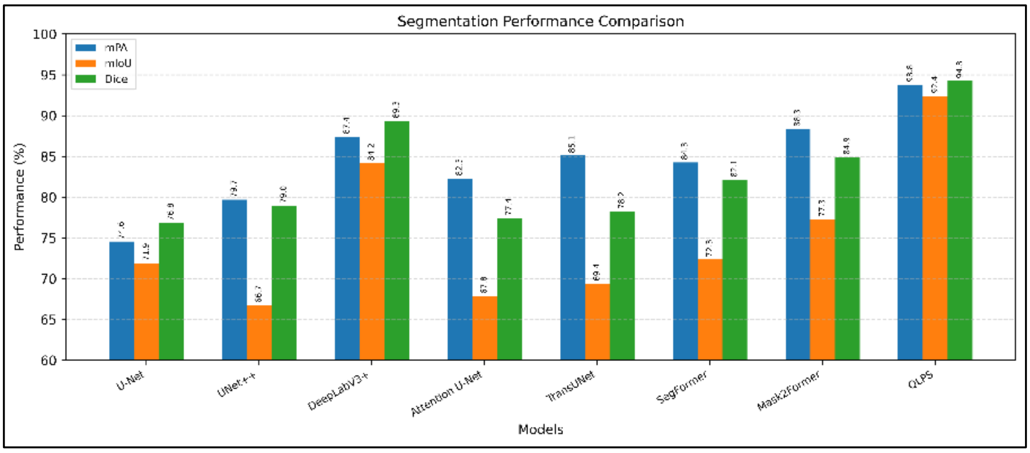

Figure 5. Comparison of segmentation performance metrics

The comparative evaluation of the segmentation models in terms of mPA, mIoU and Dice Coefficient is shown in Table 2 and Figure 5. The proposed QLPS model gives better performance in almost all parameters. It yields higher segmentation accuracy, more accurate localisation of borders, better structural preservation and higher overlap between the predicted and ground truth masks compared to the state-of-the-art methods.

Table 3. Classification performance comparison with benchmark models

Model	Accuracy	Precision	Recall	F1 Score
U – NET	86.92%	85.84%	86.41%	86.12%
UNet ++	84.12%	83.48%	83.06%	83.27%
DeepLabV3+	72.46%	71.78%	71.05%	71.41%
Attention U-Net	80.90%	80.32%	79.35%	79.83%
TransUNet	87.56%	87.11%	85.42%	86.33%
SegFormer	87.97%	85.62%	86.36%	85.99%
Mask2Former	84.65%	84.44%	83.76%	84.10%
QLPS	95.88%	94.12%	96.08%	95.09%

Table 3 presents the classification performance of the evaluated models in terms of accuracy, precision, recall and F1-score. The proposed QLPS model achieved the best accuracy (95.88%), recall (96.08%) and F1-score (95.09%) while maintaining an appreciable precision of 94.12% outperforming all other approaches. The results show that the proposed quantum-augmented feature learning framework may significantly improve the robustness, discriminative ability, and reliability of mushroom classification.

Table 4. Cross validation analysis

Fold	Classification Accuracy (%)
Fold 5	95.40
Fold 10	96.68

Table 4 presents the cross-validation analysis of the 5-fold and 10-fold methods to evaluate the robustness and generalisation ability of the proposed model for different data divisions. The model achieved classification accuracies of 95.40% and 96.68% for 5-fold and 10-fold cross validation respectively. The marginally better performance observed in 10-fold cross-validation indicates the proposed framework is more stable and reliably generalizable.

For each fold, the full QLPS-EfficientNetB0 framework was trained from scratch and evaluated on the respective validation subset.

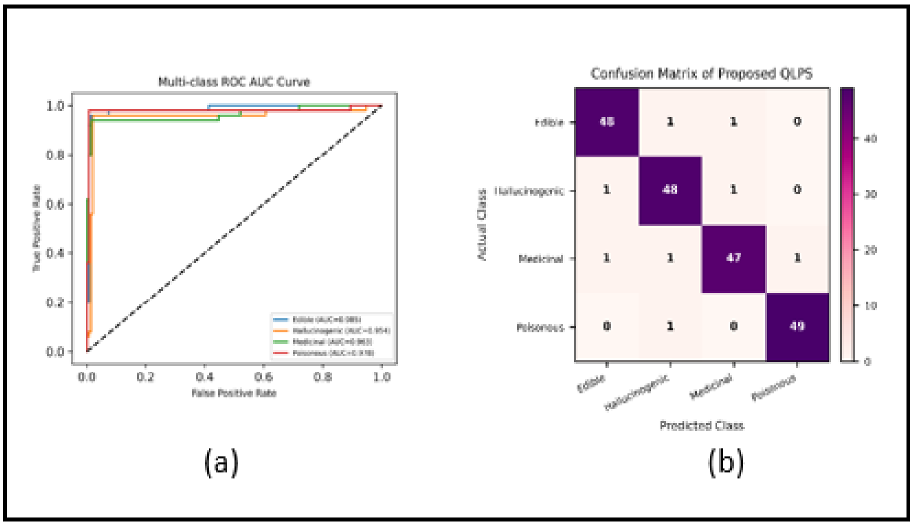

Figure 6. ROC-AUC curve and confusion matrix for our proposed QLPS utilising EfficientNetB0

The multi-class ROC analysis of the proposed QLPS framework is presented in Figure 6(a). The recorded AUC values were 0.952 for edible, 0.938 for hallucinogenic, 0.908 for medicinal, and 0.956 for toxic mushrooms, indicating strong classification efficacy and effective differentiation among the four mushroom categories.

The confusion matrix of the proposed QLPS framework, presented in Figure 6(b) reveals that most of the mushroom samples are correctly classified. Class-wise precision, recall, and F1-score also prove that the proposed model is effective in multi-class mushroom classification.

Figure 7. Failure case examples

Some difficult mushroom images with substantial background clutter, partial occlusion, low contrast and overlapping structures are shown in Figure 7, for which the proposed QLPS framework may generate incomplete segmentation or small classification errors. For example, overlapping mushrooms can cause boundary merging, but there are also superficial similarities between edible and poisonous species that can lead to misidentification. In very blurred images, reduced visibility of textures can degrade segmentation accuracy and class separation.

4.5 Computational Complexity

The proposed QLPS-EfficientNetB0 framework has 20.54 million trainable parameters and a computational cost of approximately 1.77 GFLOPs. This model achieves an average inference time of 164.38 ms per image (about 6.08 FPS) and requires 58.20 s for each training epoch. The QLPS module was implemented in a Python environment using the PennyLane quantum simulator and consists of a 4-qubit variational quantum circuit with 12 trainable parameters, a circuit depth of 7, 16 single-qubit gates, 4 CNOT entangling gates, and four Pauli-Z measurements. The experiments were run on the PennyLane default.qubit simulator in analytic mode, meaning the expectation values are computed exactly rather than via finite-shot sampling. The compact quantum design offers a good trade-off between feature representation ability and computational efficiency, despite the computational overhead of quantum circuit simulation.

The proposed framework uses a fixed 4-qubit quantum circuit regardless of image resolution, as quantum processing is performed patch-wise rather than on the entire image. As image resolution increases, the number of patches to be processed increases proportionally. The circuit's width is fixed, so scalability is primarily limited by the number of patches and inference time, rather than the exponential scaling in the number of qubits. For more complex segmentation tasks involving fine-grained structures, more sophisticated quantum circuits or hierarchical patch aggregation approaches might be necessary, albeit at the cost of increased computational overhead. A major direction for future work is a large-scale empirical scalability study across various image resolutions.

4.6 Ablation Research

Ablation studies were performed on the Laplacian Pyramid (LP), Quantum Pyramid Refinement Module (QPRM) and EfficientNetB0 to ascertain the importance of the main components of the proposed framework. The results indicate that the LP improves multi-scale feature extraction and the QPRM improves feature representation through quantum refinement, which in turn improves overall classification performance.

Table 5. Ablation analysis of the proposed framework

Model Variant	Laplacian Pyramid (LP)	QPRM	EfficientNet B0	Accuracy (%)
EfficientNetB0 only	×	×	✓	86.32
LP + EfficientNetB0	✓	×	✓	88.90
QPRM + EfficientNetB0	×	✓	✓	91.00
QLPS	✓	✓	✓	95.88

Table 5 presents the component-level ablation analysis of the proposed system. The amalgamation of the Laplacian Pyramid and QPRM with EfficientNetB0 yields superior performance, with a classification accuracy of 95.88%.

4.7 Analysis of Feature Similarity and Distinction Utilising Cosine Similarity and Euclidean Distance

The cosine similarity matrices are shown in Figs. 8(a) and 8(b). Model 8(b) with QLPS shows smaller off-diagonal values than EfficientNetB0 only 8(a), indicating better separability of the classes.

Figures 8(c) and 8(d) show the Euclidean distance matrices. The QLPS model has a clear Euclidean distance distribution compared with the baseline model, which indicates better consistency and discriminative features of the class-wise prediction patterns.

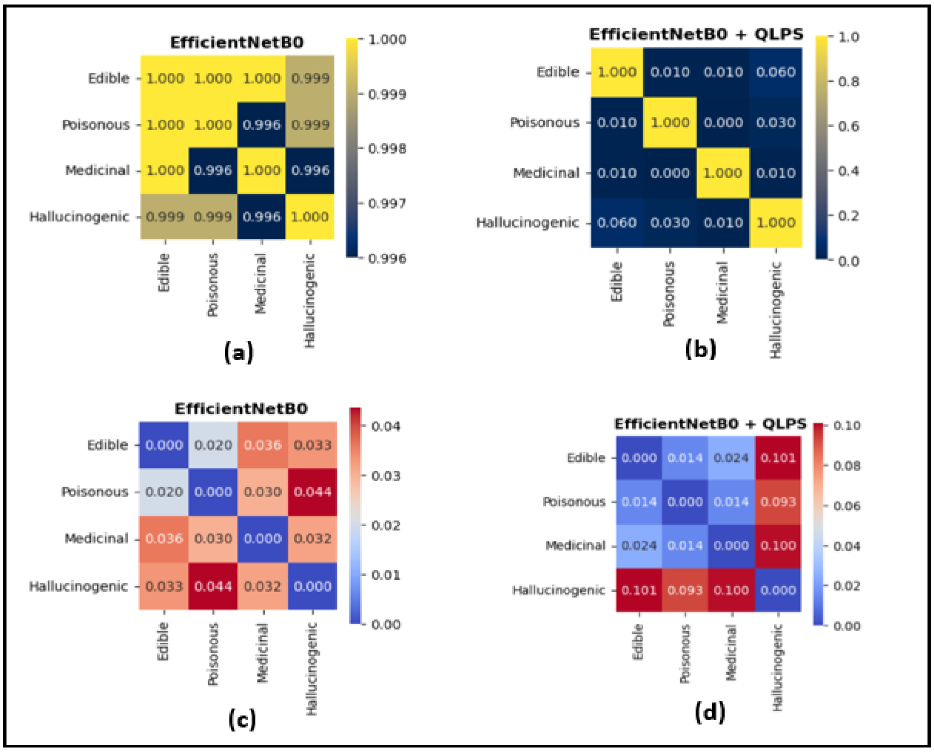

Figure 8. Cosine similarity and euclidean distance

Table 6. Statistical significance analysis

Metric	EfficientNetB0	EfficientNetB0 + QLPS	Statistical Significance
Validation Accuracy (%)	86.32	95.88	↑ +9.56%
Cosine Similarity (%)	0.358	0.121	$t = 3.94$, $p = 0.0021$, $d = 1.12$, CI = [0.00029, 0.0098]
Euclidean Distance	0.126	0.084	$t = 5.98$, $p = 8.14 \times 10^{-5}$, $\lvert d \rvert = 1.69$, CI = [−0.02481, −0.01194]

To assess the separation between classes, we computed cosine similarity and Euclidean distance matrices on class-wise averaged softmax predictions. As shown in Table 6, the EfficientNetB0 + QLPS model obtained a validation accuracy of 95.88%, which is 9.56% higher than the baseline EfficientNetB0. QLPS significantly reduced the cosine similarity (t=3.94, p=0.0021) and increased the inter-class Euclidean distance of the class-wise averaged prediction vectors. Results show that the QLPS model generates more distinctive and well-defined prediction distributions, which directly boosts classification performance.

4.8 Limitations and Future work

The proposed QLPS architecture shows very good performance in different lighting conditions and complex backgrounds. However, these factors can still affect segmentation accuracy and classification performance in cases of large image blur, significant occlusion, overlapping mushrooms, and very low-contrast scenes. Furthermore, the proposed framework was tested on a balanced dataset of four biologically meaningful mushroom types, but its generalizability could be improved by future validation on larger and more diverse datasets collected under different environmental conditions. The proposed system for real-time mushroom recognition will be extended in future works to more complex quantum architectures, and its performance will be evaluated on large-scale, expert-curated benchmark datasets.

5. Conclusion

The present study demonstrates a robust hybrid classical-quantum paradigm that successfully combines multiscale structural segmentation with deep convolutional learning for accurate multiclass mushroom classification. The proposed dual-track architecture effectively decouples feature tracking to handle wild background noise. The 12-parameter QLPS subnetwork extracts local morphological characteristics to provide an independent segmentation accuracy of 93.79%. The pre-trained EfficientNetB0 backbone takes the masked foreground ROI as input to provide a final classification accuracy of 95.88%. The integrated framework shows that quantum-enhanced segmentation directly benefits downstream classification by providing cleaner foreground representations for EfficientNetB0. The ablation trials reinforce the benefit of isolating the biological features of the mushroom through an element-wise Hadamard product to achieve a 9.56% gain in classification accuracy over baseline standard networks, confirming that parameter-efficient quantum-assisted preprocessing provides a scalable pathway to enhance classical deep learning models deployed on texturally complex agricultural datasets.

Acknowledgment

The authors would like to express their gratitude to the Department of Computer Science, Periyar University, for the financial support provided through the UGC NET Junior Research Fellowship (NTA Ref. No. 240510090311), awarded in November 2024.

The authors also sincerely thank Dr. C. Kannan, Assistant Professor, Department of Plant Pathology, VOC Agricultural College and Research Institute, Killikulam, for his valuable contribution in manually annotating the mushroom images used to generate the ground truth segmentation masks for this study.

Declarations

The authors declare no competing financial interests or personal relationships that could have influenced this work.

References

[1] Pepe, Marco, Mohsen Hesami, Karla A. de la Cerda, Melissa L. Perreault, Tom Hsiang, and Andrew Maxwell Phineas Jones. "A Journey with Psychedelic Mushrooms: From Historical Relevance to Biology, Cultivation, Medicinal Uses, Biotechnology, and Beyond." Biotechnology advances 69 (2023): 108247.

[2] Diaz, James H. "Syndromic Diagnosis and Management of Confirmed Mushroom Poisonings." Critical care medicine 33, no. 2 (2005): 427-436.

[3] Davis, Mike, Robert Sommer, and John Menge. Field guide to mushrooms of Western North America. Vol. 106. Univ of California Press, 2012.

[4] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation." In International Conference on Medical image computing and computer-assisted intervention, Cham: Springer international publishing, 2015, 234-241.

[5] Badrinarayanan, Vijay, Alex Kendall, and Roberto Cipolla. "Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation." IEEE transactions on pattern analysis and machine intelligence 39, no. 12 (2017): 2481-2495.

[6] Zhou, Zongwei, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. "Unet++: A Nested U-Net Architecture for Medical Image Segmentation." In International workshop on deep learning in medical image analysis, Cham: Springer International Publishing, 2018, 3-11.

[7] Oktay, Ozan, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori et al. "Attention U-Net: Learning Where to Look for The Pancreas." arXiv preprint arXiv:1804.03999 (2018).

[8] Chen, Jieneng, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L. Yuille, and Yuyin Zhou. "Transunet: Transformers Make Strong Encoders for Medical Image Segmentation." arXiv preprint arXiv:2102.04306 (2021).

[9] Chen, Liang-Chieh, George Papandreou, Iasonas Kokkinos, Kevin Murphy, and Alan L. Yuille. "Deeplab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected Crfs." IEEE transactions on pattern analysis and machine intelligence 40, no. 4 (2017): 834-848.

[10] Xie, Enze, Wenhai Wang, Zhiding Yu, Anima Anandkumar, Jose M. Alvarez, and Ping Luo. "Segformer: Simple and Efficient Design for Semantic Segmentation with Transformers." Advances in neural information processing systems 34 (2021): 12077-12090.

[11] Cheng, Bowen, Ishan Misra, Alexander G. Schwing, Alexander Kirillov, and Rohit Girdhar. "Masked-Attention Mask Transformer for Universal Image Segmentation." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, 1290-1299.

[12] Addesso, Paolo, Rocco Restaino, and Gemine Vivone. "An Improved Version of the Generalized Laplacian Pyramid Algorithm for Pansharpening." Remote Sensing 13, no. 17 (2021): 3386.

[13] Chen, Yongsheng. "Semantic Image Segmentation with Feature Fusion Based on Laplacian Pyramid." Neural Processing Letters 54, no. 5 (2022): 4153-4170.

[14] Baek, Yeul-Min, Hyoung-Joon Kim, Jin-Aeon Lee, Sang-Guen Oh, and Whoi-Yul Kim. "Color Image Enhancement Using the Laplacian Pyramid." In Pacific-Rim Conference on Multimedia, Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, 760-769.

[15] Venegas-Andraca, Salvador E., and Sougato Bose. "Storing, Processing, and Retrieving an Image Using Quantum Mechanics." In Quantum information and computation, vol. 5105, SPIE, 2003, 137-147.

[16] Le, Phuc Q., Fangyan Dong, and Kaoru Hirota. "A Flexible Representation of Quantum Images for Polynomial Preparation, Image Compression, and Processing Operations." Quantum Information Processing 10, no. 1 (2011): 63-84.

[17] Zhang, Yi, Kai Lu, Yinghui Gao, and Mo Wang. "NEQR: A Novel Enhanced Quantum Representation of Digital Images." Quantum information processing 12, no. 8 (2013): 2833-2860.

[18] Yan, Fei, Abdullah M. Iliyasu, and Phuc Q. Le. "Quantum Image Processing: A Review of Advances in its Security Technologies." International Journal of Quantum Information 15, no. 03 (2017): 1730001.

[19] Jadhav, Abhishek, Akhtar Rasool, and Manasi Gyanchandani. "Quantum Machine Learning: Scope for Real-World Problems." Procedia Computer Science 218 (2023): 2612-2625.

[20] Alam, Mahabubul, Satwik Kundu, Rasit Onur Topaloglu, and Swaroop Ghosh. "Quantum-Classical Hybrid Machine Learning for Image Classification (Iccad Special Session Paper)." In 2021 IEEE/ACM International Conference On Computer Aided Design (ICCAD), IEEE, 2021, 1-7.

[21] Schuld, Maria, and Nathan Killoran. "Quantum Machine Learning in Feature Hilbert Spaces." Physical review letters 122, no. 4 (2019): 040504.

[22] Huang, Hsin-Yuan, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, and Jarrod R. McClean. "Power of Data in Quantum Machine Learning." Nature communications 12, no. 1 (2021): 2631.

[23] Sim, Sukin, Peter D. Johnson, and Alán Aspuru-Guzik. "Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms." Advanced Quantum Technologies 2, no. 12 (2019): 1900070.

[24] Hubregtsen, Thomas, Josef Pichlmeier, Patrick Stecher, and Koen Bertels. "Evaluation of Parameterized Quantum Circuits: on The Relation Between Classification Accuracy, Expressibility, and Entangling Capability." Quantum Machine Intelligence 3, no. 1 (2021): 9.

[25] Tan, Mingxing, and Quoc Le. "Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks." In International conference on machine learning, PMLR, 2019, 6105-6114.