

Received 28 June 2024, accepted 11 July 2024, date of publication 16 July 2024, date of current version 25 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3429413

RESEARCH ARTICLE

Robust Inverse Quantum Fourier Transform Inspired Algorithm for Unsupervised Image Segmentation

TAOREED A. AKINOLA^{1,2}, (Member, IEEE), XIANGFANG LI^{1,2}, (Member, IEEE),
RICHARD WILKINS¹, (Member, IEEE), PAMELA H. OBIOMON^{1,2}, (Member, IEEE),
AND LIJUN QIAN^{1,2}, (Senior Member, IEEE)

¹Department of Electrical and Computer Engineering, Prairie View A&M University, Prairie View, TX 77446, USA

²CREDIT Center, Prairie View A&M University, Prairie View, TX 77446, USA

Corresponding author: Taoreed A. Akinola (takinola2@pvamu.edu)

This research work is supported by the IBM-HBCU Quantum Center and the NSF Award 2205891 and 2302469.

ABSTRACT Image segmentation is a crucial task in computer vision. In this study, we propose a method based on inverse quantum Fourier transform (IQFT) and develop a robust IQFT-inspired algorithm for unsupervised image segmentation. The proposed method leverages the underlying mathematical mechanism of the IQFT to cluster the input image pixels automatically and efficiently into different segments. Specifically, by considering the correlation between the within-cluster mean sum of squared error (MSSE) and the probability of quantum measurements, the proposed robust algorithm significantly improves the segmentation performance. It is an unsupervised method with characteristics similar to k -means, i.e., the proposed method does *not* require training. Extensive evaluation of the proposed method has been carried out, showing that it outperforms the classical k -means on the PASCAL VOC 2012 segmentation benchmark by as much as 4.21%, the Flowers dataset by as much as 4.4%, and the xVIEW2 challenge dataset by as much as 11.1% in terms of Intersection-Over-Union (IoU). It is also demonstrated that the proposed method has comparable or mixed performance compared to recent more complex approaches. However, compared to approaches such as GrabCut, which require a measure of the user interaction, and deep learning-based methods, which require generative models and deep feature extraction algorithms, the proposed method does *not* require training or user involvement. This makes it a promising choice for applications that do not have access to data before deployment or have very limited training data.

INDEX TERMS Computer vision, image segmentation, inverse quantum Fourier transform, unsupervised learning, k -means.

I. INTRODUCTION

Image segmentation is a key approach in the field of computer vision and image analysis that helps us understand an image's visual contents more deeply. Fundamentally, image segmentation can be formulated as the problem of simplifying the complexity of an image by separating it into unique sections or segments, each associated with a significant object, texture, or area of interest. At the pixel

The associate editor coordinating the review of this manuscript and approving it for publication was Tai Fei¹.

level, image segmentation allows for the exact delineation of object boundaries and the separation of distinct entities inside the image, in contrast to classification, which assigns a single label to an entire image. Image segmentation is important because it gives machines the ability to comprehend the visual environment in a way similar to humans. By segmenting an image into significant parts, computers can precisely detect and extract particular objects or regions, irrespective of their shape, size, or context. This technology has applications in various industrial processes, including industrial automation, autonomous vehicles, video surveillance, content-based

image retrieval image object recognition and object detection, and medical diagnostic imaging.

The development of deep learning and artificial intelligence has considerably accelerated the evolution of image segmentation algorithms [1], [2]. Convolutional neural networks (CNNs) [3] and other cutting-edge architectures have replaced the traditional techniques in many critical and challenging tasks, particularly in applications where correctly labeled training datasets are readily available. The segmentations produced by these state-of-the-art approaches are surprisingly accurate because they make use of enormous volumes of manually labeled data to discover complex patterns and relationships inside images. However, in applications involving large, diversified datasets where labeled data are not available, unsupervised approaches are exploited. These alternative techniques utilize the image's statistical, geometrical, or texture-based qualities to automatically reveal subtle regions of interest and solve image segmentation problems without the need for any previous knowledge. Unsupervised techniques find application in image decomposition into more general regions such as "Foreground" and "background".

Unsupervised image segmentation is pivotal in fields where labeled data is sparse or inaccessible, yet accurate image comprehension is essential. Unlike supervised learning, which necessitates costly and time-intensive annotation efforts, unsupervised methods capitalize on inherent data patterns to autonomously segment images. This adaptability is particularly advantageous in environments lacking annotated datasets or requiring continual adaptation to new conditions [4]. For example, in space missions [5], where obtaining labeled data for image analysis is impractical, unsupervised segmentation facilitates the categorization of celestial body features captured by rovers or satellites, aiding in geological exploration and potential life detection. Similarly, in medical imaging, such as pathology and radiology, unsupervised segmentation supports the identification of anomalies in scans where annotated data is scarce, enhancing diagnostic accuracy and treatment planning [6]. Moreover, in environmental sciences, unsupervised techniques enable automated feature extraction from remote sensing data, crucial for monitoring deforestation or wildlife habitats in regions where ground truth data collection is challenging or impractical. Driven by applications in space exploration, medical imaging, and environmental monitoring, this research focuses on advancing unsupervised segmentation techniques to eliminate the reliance on labeled training data.

It is pertinent to note that the literature encompasses different types of image segmentation methodologies, including binary segmentation, and semantic segmentation, among others. However, this study primarily concentrates on binary segmentation, a fundamental approach wherein an image is partitioned into two distinct regions: foreground and background. Each pixel within the image is rigorously categorized as either a part of the object of interest (foreground)

or not (background). Unlike semantic segmentation, which assigns specific class labels to pixels, our focus on binary segmentation underscores a generalized partitioning devoid of explicit reference to particular objects [7].

In this work, we focus on the problem of automatically dividing an image into prominent or significant regions without prior information. By using the pixel's RGB color features, we aim to cluster an image into regions of interest. Because the overall segmentation output from these color features depends on the colors defining the region of interest, there is a need to adequately describe this region in terms of color and texture [8]. This study extends our previous work in [9].

Our previous work provides insights into the applicability of the inverse quantum Fourier transform (IQFT) for unsupervised image segmentation. The segmentation output observed using the previously proposed algorithm is dependent on three heuristically determined angular parameters which do not offer enough flexibility required to tune the segmentation mask obtained to an optimal output. This current work represents a substantial advancement over our previous research, with significant enhancements in robustness, theoretical analysis, experimental validation, and experimentation efforts.

Robustness is a major concern of machine learning algorithms, including image segmentation. It is paramount for ensuring the effectiveness and reliability of algorithms across various real-world scenarios. Robustness is the ability of an algorithm to maintain high performance and stability even when faced with unforeseen or challenging conditions, such as noisy data, outliers, changes in distribution, or adversarial attacks. For a machine learning algorithm, robustness enhances its trustworthiness, instilling confidence in its reliability and performance, even in challenging conditions [10], [11]. This is especially important in safety-critical applications such as healthcare, autonomous vehicles, and finance. For this reason, *this work considers the robustness of the IQFT method for image segmentation which was not mentioned in our previous work.*

Additionally, in our previous algorithm, pixel classification relied solely on the state with the maximum probability. However, in this current work, we refine this approach by considering the probability distribution of the other states as well. This enhancement aims to ensure the region of interest is achieved with a good degree of certainty. Specifically, we introduce a probability threshold delineating a boundary of certainty. Pixels falling below this threshold undergo further analysis using similarity measures such as cosine similarity or Euclidean distance. This process contributes to more precise segmentation results. Unlike our previous work, where the probability value was determined by random guessing until satisfactory results were obtained, this study introduces a heuristic approach involving probability search using the average sum of squared error (SSE) to achieve optimal or near-optimal results. This eliminates

user involvement and ensures consistency in segmentation outcomes.

Furthermore, we expand the scope of evaluation by validating the proposed method on four benchmark datasets, including Flowers [12] and Caltech-UCSD Birds-200-2011 (CUB-200-2011) [13], which were not considered in our previous work. This broader evaluation provides a more comprehensive assessment of the algorithm's performance across diverse datasets and scenarios. Moreover, we have enriched the analysis of our method by comparing it with additional state-of-the-art techniques, such as ReDo [14], GrabCut [15], Inpainting Error Maximization (IEM) [16], and other complex algorithms utilizing deep learning networks. These comparisons offer valuable insights into the strengths and weaknesses of our approach relative to existing methods.

The contributions of this paper are:

- 1) A robust, IQFT-inspired algorithm is proposed for unsupervised image segmentation.
- 2) The proposed approach has much less computational cost when compared to many deep learning-based image segmentation methods and is thus suitable for real-time applications because it does *not* require training.
- 3) We demonstrate the potential of the robust, IQFT-inspired method in solving image segmentation problems, and compare its effectiveness with k -means, which is a de facto standard method matching our approach in terms of the structural simplicity of the algorithm involved.

The remainder of this paper is structured as follows: Related works are reviewed in Section II. Section III introduces the quantum Fourier transform and its inverse. Section IV provides the details of the proposed approach. The experimental results, including observations and insights, are given in Section V. Section VI concludes the paper.

II. RELATED WORKS

A. IMAGE SEGMENTATION IN CLASSICAL COMPUTING

As a key process in image processing and analysis, image segmentation has been well-studied in the classical computing domain. Many methods for solving image segmentation problems have emerged over the years. These methods vary widely depending on the specific application since a single method is not sufficient for different images with varying characteristics in terms of sharpness, texture, noise presence, and the degree of overlapping objects [18]. Traditional techniques used for image segmentation are categorized as thresholding-based technique [19], region-based technique [20], edge-based technique [21], clustering-based technique [22], and watershed technique [23]. These methods vary widely depending on the specific application since they all have their limitations based on the underlying principles. For instance, employing an unsupervised method like the k -means for image segmentation has a major

drawback which is the requirement for the optimal number of clusters to be specified before the algorithm is applied [24]. The segmentation error decreases as the number of clusters increases and there is no theoretical means of obtaining the optimal number of clusters to be used. Similarly, Otsu's thresholding technique does not consider the spatial information of the image, and this makes it sensitive to the unevenness and noise in a grayscale image [25].

Although the advent of deep learning has brought about new classes of image segmentation techniques that have become widely available [26], and usually have very high computational complexity due to the required training and retraining. Traditional deep CNN-based segmentation methods depend heavily on large datasets, which limits their effectiveness in various scenarios. Few-shot segmentation addresses this limitation by learning to generalize from only a few labeled examples [27]. This is often achieved through techniques like prototype learning and advanced network architectures to enhance accuracy and robustness [28]. Despite its potential, current methods can suffer from overfitting and imprecise boundaries. Ongoing research aims to address these issues by introducing innovative solutions, such as Holistic Prototype Activation (HPA) and other variants [29], [30], [31], to improve performance in low-data regimes. These techniques are supervised methods that require training, however small. On the contrary, the proposed method does not require training and has low computational cost compared to the deep learning-based methods, thus making it suitable for real-time applications. In our analysis, we will compare the performance of our proposed technique with k -means Clustering [32], which has similar advantages in computational efficiency.

An emerging unsupervised clustering algorithm gaining attention in recent years is mean-shift clustering. Its distinct advantages over traditional methods like k -means lie in its non-parametric approach, leveraging kernel density estimates within its algorithm. Unlike k -means, mean-shift can effectively model intricate clusters with non-convex shapes, requiring users to set only one parameter: the bandwidth. This bandwidth parameter dictates the scale of the clusters, autonomously determining the number of clusters based on data distribution. In contrast to k -means, where users explicitly specify the number of clusters, the outcome from using mean-shift's clustering is solely determined by the bandwidth, eliminating the need for multiple runs with different initialization values. However, in certain applications such as foreground-background separation or medical image segmentation, users may necessitate a predefined number of clusters. Regrettably, mean-shift lacks direct control over the cluster count. Achieving a specific number of clusters entails an exhaustive search over bandwidth values, incurring significant computational overhead [33]. Due to this computational complexity in achieving a predetermined number of clusters, mean-shift segmentation is not directly comparable to our proposed approach.

B. IMAGE SEGMENTATION IN QUANTUM COMPUTING

In recent years, some image segmentation algorithms have emerged in the quantum domain [34] to exploit the properties of quantum computing to improve the performance of classical techniques and, subsequently, their applications. Caraiman et al. [35] proposed an algorithm based on histogram to achieve a significant speedup by using the performance of the quantum Fourier transform and quantum amplitude amplification. However, this algorithm uses a Grover operator that requires a quantum oracle, which was not explicitly specified. This makes it difficult to simulate the algorithm. In the subsequent year, Caraiman et al. [36] proposed another single threshold-based algorithm for image segmentation. This time, the oracle circuit was specified and implementable with about fifty qubits, which made it difficult to simulate in the present quantum computer. In [37], Li et al. proposed an algorithm for image segmentation using a quantum search algorithm. The specific oracle operator employed for amplitude amplification was, however, missing. Again, this made simulation impossible. In general, most of these proposed methods involved the implementation of some oracle operators which are either fully theoretical or hard to simulate due to the number of qubits required or some missing information about their implementation. In [34], Yuan et al. proposed the dual-threshold quantum image segmentation algorithm and showed the effectiveness of their algorithm with an 8×8 grayscale image. The actual time consumption for the segmentation process was not reported concerning the theoretical time complexity of their algorithm.

In conclusion, it can be established from the literature that

- Many image segmentation methods have been proposed in the classical domain.
- A universally accepted method that satisfies all image requirements does not exist because all the proposed methods have their own challenges.
- Quantum application in image segmentation is still being developed and not many works have been done effectively in this quantum domain.

It is against this backdrop that this work seeks to propose a novel segmentation technique that, by benefiting from both classical and quantum computing, will be a promising method to solve an image segmentation problem.

III. QUANTUM FOURIER TRANSFORM (QFT) AND INVERSE QUANTUM FOURIER TRANSFORM (IQFT)

Quantum Fourier Transform (QFT) transforms a computational basis state $|x\rangle$ as shown in equation (1) [38], where $N = 2^n$, and n is the number of qubits.

$$QFT(|x\rangle_n) = \frac{1}{\sqrt{N}} \otimes_{k=1}^n \left(|0\rangle + e^{i\frac{2\pi x}{2^k}} |1\rangle \right) \quad (1)$$

For example, when three qubits are considered, $n = 3$ equation (1) becomes:

$$QFT(|x\rangle_3) = \frac{1}{\sqrt{8}} \left(|0\rangle + e^{i\frac{2\pi x}{2}} |1\rangle \right) \otimes \left(|0\rangle + e^{i\frac{2\pi x}{4}} |1\rangle \right)$$

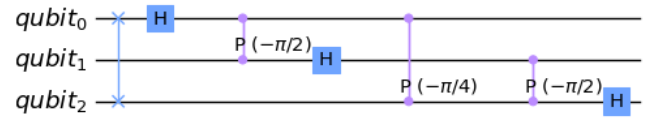


FIGURE 1. IQFT quantum circuit.

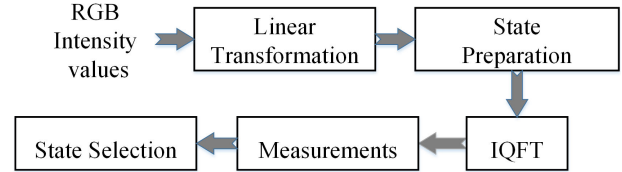


FIGURE 2. Framework of the IQFT method for RGB image segmentation.

$$\otimes \left(|0\rangle + e^{i\frac{2\pi x}{8}} |1\rangle \right) \quad (2)$$

Considering this mathematical formulation of the QFT for a system of three qubits, it becomes obvious that a superposition of states, with some phase information, can be transformed into a single state representation through the instrumentality of the inverse quantum Fourier transform (IQFT) as shown in equation (3).

$$IQFT \left[\frac{1}{\sqrt{8}} \left(|0\rangle + e^{i\alpha} |1\rangle \right) \otimes \left(|0\rangle + e^{i\beta} |1\rangle \right) \otimes \left(|0\rangle + e^{i\gamma} |1\rangle \right) \right] = |k\rangle_3 \quad (3)$$

where α , β and γ are the relative phases.

In this study, three qubits are considered to match the three channels of the RGB color space. Hence, a three-qubit IQFT circuit [39] employed is shown in Figure 1. It is worth stating here that the IQFT method is not limited by the number of qubits since the desired number of qubits can be chosen based on the available number of inputs. However, using a high number of qubits will increase the computational complexity, and hence, the runtime of the algorithm.

IV. PROPOSED ALGORITHMS FOR RGB IMAGE SEGMENTATION

A. IQFT METHOD FOR RGB IMAGE SEGMENTATION IN THE QUANTUM DOMAIN

According to equation (4), where n is the number of qubits, with three qubits, it is possible to partition an image into a maximum of eight segments which is sufficient for most practical applications.

$$\text{Number of segments} \leq 2^n \quad (4)$$

The IQFT framework in Figure 2 consists of five stages. The linear transformation stage involves mapping from RGB intensity values to a corresponding set of relative phases (α , β , γ) according to equation (5), where $I_{range} = I_{max} - I_{min}$ is the range of the intensity values in channel, $I \in \{R, G, B\}$, and $\{\theta_1, \theta_2, \theta_3\}$ is a set of angle parameters to be chosen based on the application.

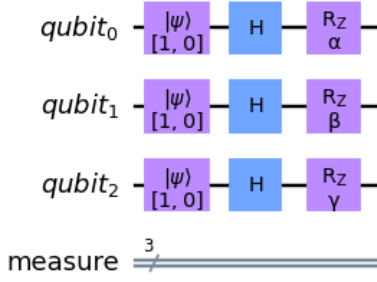


FIGURE 3. State preparation circuit. All qubits are initialized to state $|0\rangle$.

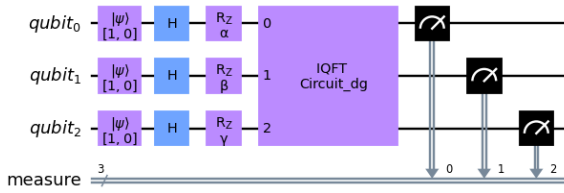


FIGURE 4. Complete quantum circuit for RGB image segmentation.

$$\begin{aligned} \left(\frac{R - R_{min}}{R_{range}} \right) \times \theta_1 &\longrightarrow \alpha; \left(\frac{G - G_{min}}{G_{range}} \right) \times \theta_2 \longrightarrow \beta; \\ \left(\frac{B - B_{min}}{B_{range}} \right) \times \theta_3 &\longrightarrow \gamma \end{aligned} \quad (5)$$

At the state preparation stage, the quantum circuit in Figure 3 converts the phase angles (α, β, γ) to a three-qubit state $\frac{1}{\sqrt{8}}(|0\rangle + e^{i\alpha}|1\rangle) \otimes (|0\rangle + e^{i\beta}|1\rangle) \otimes (|0\rangle + e^{i\gamma}|1\rangle)$, which is the input of the IQFT circuit in Figure 1. The Hadamard gates produce superposition of states $|0\rangle$ and $|1\rangle$ while the R_z gates perform single-qubit rotations around the z -axis [40] to establish the required relative phase values. Figure 4 presents a complete quantum circuit for the IQFT method for RGB image segmentation in the quantum domain.

The output segmentation label is realized as explained in the following example. Consider a set of random relative phases $\{\alpha, \beta, \gamma\} = \{0.2464, 2.464, 3.080\}$ from a linear transformation of a pixel's RGB values. When the circuit in Figure 4 was simulated in IBM Qiskit [42] using “qasm” simulator backend set to 2000 shots, the output probability distribution in Figure 5 was realized. Since the most probable state is “001”, the pixel is labeled 1. This process is repeated for all the pixels in a given image data for a complete label.

B. IQFT-INSPIRED ALGORITHM FOR RGB IMAGE SEGMENTATION IN CLASSICAL DOMAIN

The requirement of three qubits makes the quantum method suitable in this noisy intermediate-scale quantum (NISQ) era. However, the major drawback, like other practical quantum methods, is the high computation cost for transforming the classical image intensities into the quantum states. To eliminate the computation cost for classical-to-quantum data transformation, without loss of performance, an IQFT-Inspired algorithm for RGB Image segmentation is proposed.

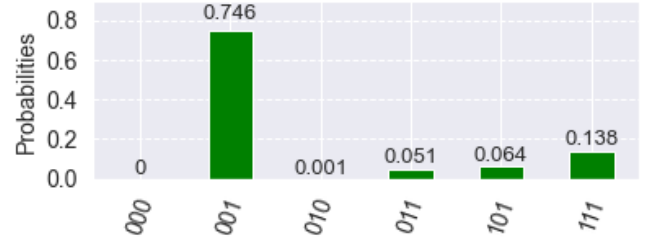


FIGURE 5. Probability distribution for a set of random relative phases $\{\alpha, \beta, \gamma\}$.

Expanding the left-hand side (LHS) of equation (3) results in equation (6). Further simplification of equation (6) yields equation (7).

$$\begin{aligned} RHS = \frac{1}{\sqrt{8}} &\times [IQFT|000\rangle + e^{i\gamma}IQFT|001\rangle \\ &+ e^{i\beta}IQFT|010\rangle + e^{i(\beta+\gamma)}IQFT|011\rangle \\ &+ e^{i\alpha}IQFT|100\rangle + e^{i(\alpha+\gamma)}IQFT|101\rangle \\ &+ e^{i(\alpha+\beta)}IQFT|110\rangle + e^{i(\alpha+\beta+\gamma)}IQFT|111\rangle] \end{aligned} \quad (6)$$

where $IQFT(|abc\rangle) = \frac{1}{\sqrt{8}} \sum_{x=0}^7 \omega^{-x(d)} |x_2\rangle$ is the IQFT of a three-qubit basis state, $|abc\rangle$. d and x_2 are base-10 equivalents of abc_2 and x , respectively.

$$\begin{aligned} RHS = P|000\rangle + Q|001\rangle + R|010\rangle + S|011\rangle + T|100\rangle \\ + U|101\rangle + V|110\rangle + W|111\rangle \end{aligned} \quad (7)$$

where probability amplitudes P through W are given by the matrix representation in equation (8).

$$[P \ Q \ R \ S \ T \ U \ V \ W]^T \equiv \frac{1}{8} \times M \times N \quad (8)$$

where

$$M = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \omega^{-1} & \omega^{-2} & \omega^{-3} & \omega^{-4} & \omega^{-5} & \omega^{-6} & \omega^{-7} \\ 1 & \omega^{-2} & \omega^{-4} & \omega^{-6} & 1 & \omega^{-2} & \omega^{-4} & \omega^{-6} \\ 1 & \omega^{-3} & \omega^{-6} & \omega^{-1} & \omega^{-4} & \omega^{-7} & \omega^{-2} & \omega^{-5} \\ 1 & \omega^{-4} & 1 & \omega^{-4} & 1 & \omega^{-4} & 1 & \omega^{-4} \\ 1 & \omega^{-5} & \omega^{-2} & \omega^{-7} & \omega^{-4} & \omega^{-1} & \omega^{-6} & \omega^{-3} \\ 1 & \omega^{-6} & \omega^{-4} & \omega^{-2} & 1 & \omega^{-6} & \omega^{-4} & \omega^{-2} \\ 1 & \omega^{-7} & \omega^{-6} & \omega^{-5} & \omega^{-4} & \omega^{-3} & \omega^{-2} & \omega^{-1} \end{bmatrix}$$

$$N = [1, e^{i\gamma}, e^{i\beta}, e^{i(\beta+\gamma)}, e^{i\alpha}, e^{i(\alpha+\gamma)}, e^{i(\alpha+\beta)}, e^{i(\alpha+\beta+\gamma)}]^T$$

The matrix formulation of equation (8) enables a straightforward implementation of the proposed IQFT method in a classical computation domain such as Python. This completely classical, IQFT-inspired algorithm is presented in Algorithm 1. The input P_m is a 3D vector of RGB intensities of the m th pixel, T is the total number of pixels, M is a complex 8 by 8 matrix in equation (8), θ_1, θ_2 , and θ_3 are angle parameters for mapping pixel intensities into phase values, and the output is the required pixel label $l_m \in \{0, 1, 2, \dots, 7\}$. The segmentation algorithm involves a linear transformation

in Line 1 using equation (5), dimensional transformation from 3D to 8D vector in Line 2, following equation (8), and probability measure in Line 3. Finally, a pixel is classified according to the basis vector with the highest probability in line 4. The overall computational complexity of this algorithm is $\mathcal{O}(N \cdot 2^n)$, where N is the number of pixels in the given image, and n , is the number of qubits.

Algorithm 1 : IQFT-Inspired Algorithm for RGB Image Segmentation

Input:

$$P_m \in \mathbb{R}^3, m = [1, T]$$

$$\theta_1, \theta_2, \theta_3 \in \mathbb{R}$$

$$M \in \mathbb{C}^{8 \times 8}$$

Output: $\mathcal{L} \mathcal{D} \{l_m \in \mathbb{Z}\}$
for $m=1$ **to** T **do**

$$1. \{\gamma_m, \beta_m, \alpha_m\} \leftarrow \left\{ \frac{I_m - I_{\min}}{I_{\max}} \times \theta \right\}, I \in \{R, G, B\}$$

$$2. \{N_m\} = \begin{bmatrix} 1 \\ e^{i\gamma} \\ e^{i\beta} \\ e^{i(\beta+\gamma)} \\ e^{i\alpha} \\ e^{i(\alpha+\gamma)} \\ e^{i(\alpha+\beta)} \\ e^{i(\alpha+\beta+\gamma)} \end{bmatrix} \leftarrow \begin{bmatrix} \gamma_m \\ \beta_m \\ \alpha_m \end{bmatrix}$$

$$3. \{S_m\} \leftarrow [\text{abs}(\text{Dot Product}(N_m, M) / 8)]^2$$

$$4. \{l_m\} \leftarrow \{\text{argmax}\{S_m\}\}$$

In addition, the rows of M matrix in equation (8) represent eight state basis vectors shown in Figure 6, where each vector is visualized as a set of points on a unit circle. With this insight, the labeling of a pixel using the proposed IQFT-inspired algorithm involves a transformation from a 3D RGB vector to a set of 8 points on a unit circle using the N column vector in equation (8) and choosing a label based on the most similar basis vector. For example, the representation of a random case of $\alpha = 0.246, \beta = 0.025, \gamma = 0.250$ using equation (8) is shown in Figure 7. By visual inspection, the corresponding pixel is labeled 0 since the plots on the unit circle tend to a single point like state $|000\rangle$.

C. ROBUST, IQFT-INSPIRED ALGORITHM FOR RGB IMAGE SEGMENTATION IN THE CLASSICAL DOMAIN

Figure 8 shows two sets of arbitrary probability distributions corresponding to two random pixels. Using the proposed IQFT algorithm, the distribution in Figure 8(A) will result in label 0 which stands out with a large probability margin. However, a label of 0 for the distribution in Figure 8(B) might be a wrong prediction because of the small probability margin between states 000, 001, and 100. This slight difference could be due to the random nature of the quantum measurements or noise in the image data set. To add a level of confidence by setting a probability margin, a new probability parameter (p) is introduced for consistent and reliable

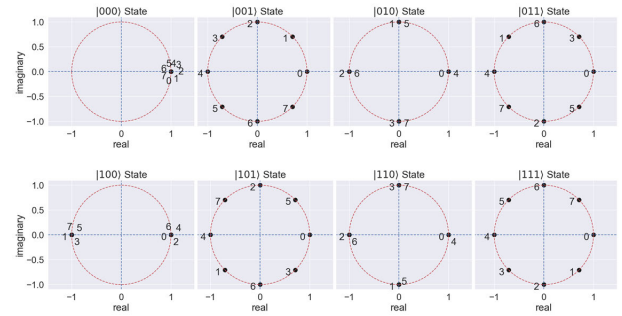


FIGURE 6. Visualization of the state basis vectors for three qubits: each vector is represented by a set of 8 points on a unit circle.

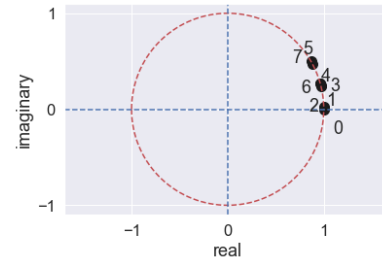


FIGURE 7. Representation of phase angles $\alpha = 0.246, \beta = 0.025, \gamma = 0.250$ on a unit circle using equation (8).

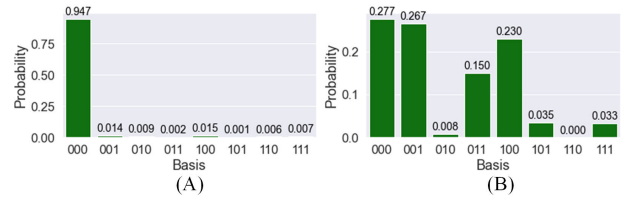


FIGURE 8. Probability distributions from two random pixels. Using the IQFT-inspired algorithm, pixel A is labeled 0 with a high degree of confidence. A label of 0 for pixel B is not reliable because of the small probability difference shown by states $|000\rangle$ and $|001\rangle$.

label predictions. With this adjustment, the IQFT-inspired algorithm will be used to automatically cluster the image data into significant regions with a specified probability threshold, and the remaining pixels for which the maximum probability falls below the specified value are then classified into the established segments by using a similarity measure like cosine similarity and Euclidean distance. Also, to ensure that our algorithm does not require the user's effort, we define the initial probability threshold $p \in \{0, 1\}$ as the closest value of p that gives the average sum of squared-error (Avg. SSE) given by equation (9). The robust, IQFT-inspired algorithm is presented in Algorithm 2.

$$\text{Avg. SSE} = \frac{1}{N} \sum_{k=1}^K \sum_{x_i \in C_k} \|x_i - \mu\|^2 \quad (9)$$

where C_k is the set of pixels in segment k , K is the total number of segments, μ_k is the mean vector of segment k calculated using $\mu = \frac{1}{|C_k|} \sum_{x_i \in C_k} x_i$.

Algorithm 2 Robust IQFT-Inspired Algorithm for RGB Image Segmentation

Input:

$$P_m \in \mathbb{R}^3, m = [1, T]$$

$$\theta_1, \theta_2, \theta_3 p \in \mathbb{R} \quad // \text{ input parameters}$$

$$M \in \mathbb{C}^{8 \times 8} \quad // \text{ equation (8)}$$

$$\text{Output: } \mathcal{L} \mathcal{D} \{l_m \in \mathbb{Z}\} \quad // \text{ Final label}$$

for m = 1 **to** T **do**:

$$1. \{\gamma_m, \beta_m, \alpha_m\} \leftarrow \left\{ \frac{l_m - l_{\min}}{l_{\max} - l_{\min}} \times \theta \right\}, I \in \{R, G, B\}$$

$$2. \{N_m\} = \begin{bmatrix} 1 \\ e^{i\gamma} \\ e^{i\beta} \\ e^{i(\beta+\gamma)} \\ e^{i\alpha} \\ e^{i(\alpha+\gamma)} \\ e^{i(\alpha+\beta)} \\ e^{i(\alpha+\beta+\gamma)} \end{bmatrix} \leftarrow \begin{bmatrix} \gamma_m \\ \beta_m \\ \alpha_m \end{bmatrix}$$

$$3. \{S_m\} \leftarrow [\text{abs}(\text{Dot Product}(N_m, M) / 8)]^2$$

for p = 0 **to** 1, **step** = 0.1 **do**:

$$4. \text{ if } \max\{S_m\} \geq p :$$

$$5. \{l_m\} \leftarrow \{\text{argmax}\{S_m\}\}$$

$$6. \text{ if } \max\{S_m\} < p :$$

$$7. \{l_m = 9\} // \text{ Temporary mask: arbitrary label, say, 9.}$$

$$8. \{0, 1, \dots, 7\} \xleftarrow{\text{similarity}} \{\mathcal{L} == 9\} // \text{ update the label}$$

$$9. \text{ Calculate SSE}$$

$$10. \text{ Calculate Avg. SSE}$$

$$11. \text{ Select the optimal } p \text{ value and the corresponding label}$$

V. RESULTS AND ANALYSIS**A. DATASET**

We used the following publicly available benchmark datasets to demonstrate the performance of our proposed technique, and, whenever possible, we compare our results with other methods in the domain of unsupervised techniques for image segmentation, especially k -means which matches our method in terms of algorithmic simplicity.

- 1) Flowers dataset [12] comprises 8,189 images of 102 different classes of flowers. For a fair comparison, we followed the data split adopted in ReDO [14] and IEM [16] and used only 1020 test images. Since our method does not require training, we did not use the validation and training images. Also, out of the 1020 test images, 32 images were not used because they have empty ground truth masks which can negatively affect the values of the performance metrics. Specifically, our method was tested on 988 images.
- 2) Caltech-UCSD Birds-200-2011 (CUB-200-2011) dataset [13] consists of 11,788 images of a collection

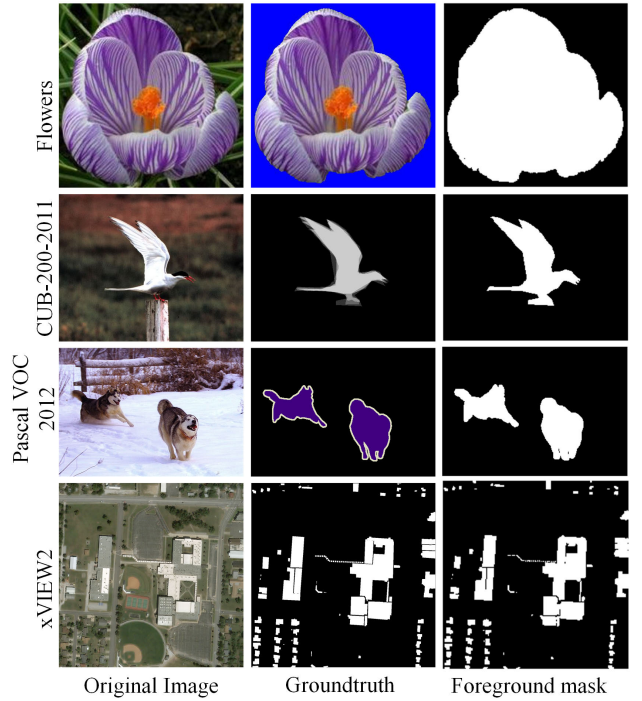


FIGURE 9. Samples from benchmark datasets: Flowers, CUB-200-2011, PASCAL VOC 2012 and xVIEW2 challenge datasets.

of 200 bird species. Again, 1000 test images were used in alignment with ReDO [14] and IEM [16].

- 3) PASCAL VOC 2012 segmentation benchmark [8] consists of 2913 labeled dataset, focusing only on the segmentation category which contains 2913 labeled set of some well-known objects like ‘birds’, ‘dog’, ‘car’, etc. The test data without annotation is ignored.
- 4) The xVIEW2 challenge dataset [41] is a large-scale collection of satellite imagery containing a mixture of pre- and post-disaster RGB satellite images for assessment of building damage. For this study, the pre-disaster images for the “Joplin-tornado” disaster are used due to the availability of ground-truth masks. It contains 148 RGB satellite images (with tiny objects) and the corresponding ground-truth masks, each of size 1024×1024 pixels.

Figure 9 shows samples of these datasets and their ground-truth masks.

B. EXPERIMENTAL SETUP

All images are RGB for all experiments. To evaluate the performance of our proposed method, we compared our results with those of the existing methods such as k -means clustering [32], ReDO [14], IEM [16], PerturbGAN [17], and BigBiGAN [44]. Flowers and CUB-200-2011 datasets are resized and center-cropped to 128×128 pixels while the dimensions of the PASCAL VOC 2012 and xVIEW2 datasets are kept. For repeatability, the k -means algorithm is developed using Scikit-learn library [45] with default

settings. All the algorithms used in this study were coded using Python on a MacBook Pro with 8-Core Intel Core i9 running at 2.3 GHz. All quantum circuits were evaluated on the IBM Qiskit. Here, Aer simulator using *qasm* simulator backend was used.

C. EVALUATION METRIC

An extensive literature review reveals that the Intersection over Union (IoU) metric, often synonymous with Jaccard's Index, holds significant importance in evaluating segmentation tasks. IoU quantifies the segmentation model's ability to delineate objects from their backgrounds within an image, making it a widely adopted evaluation metric for comparing image segmentation techniques [43].

In line with standard practices and to ensure a robust comparison of our proposed method with prior approaches, we utilized the standard average IoU metric. Given that the state-of-the-art techniques referenced in this study report results in IoU, this metric serves as a common benchmark for evaluation. Additionally, we furnish results in terms of the DICE score and per-pixel accuracy for comprehensive analysis. It's essential to note that for all these metrics, higher values indicate superior performance.

- For ease of comparing our work with previous work, the intersection over union (IoU) score was computed as the IoU of the foreground according to equation (10) or as the mean-IoU (*mIoU*) of both the foreground and the background, according to equation (11). To apply equation (10) to an image, two *IoU* values are calculated for each of the predicted segments using the given ground-truth foreground. In line with [8], the higher of the two values is taken as the *IoU* of the image. For N images, the average *IoU* of the foreground is $(\sum_1^N IoU(\text{foreground}))/N$.
- *DICE* score is defined in equation (12), all symbols remain as defined in equation (10). The mean-*DICE* score of the foreground was calculated following a similar procedure for mean-*IoU*.
- Accuracy is defined as the proportion of labels that have been correctly assigned to the foreground and background.

$$IoU(\text{foreground}) = \frac{|F_{GT} \cap F_P|}{|F_{GT} \cup F_P|} \quad (10)$$

where F_{GT} and F_P are the groundtruth foreground and predicted foreground labels, $|\epsilon|$ is equal to the number of elements in ϵ .

$$mIoU = \frac{IoU(\text{foreground}) + IoU(\text{background})}{2} \quad (11)$$

$$\begin{aligned} DICE(\text{foreground}) \\ = 2 \times \frac{|F_{GT} \cap F_P|}{|F_{GT}| + |F_P|} \end{aligned} \quad (12)$$

In all cases of the labels, the foreground and background regions are defined by binary 1 and binary 0, respectively. All ground-truth masks are binarized using OpenCV library. The

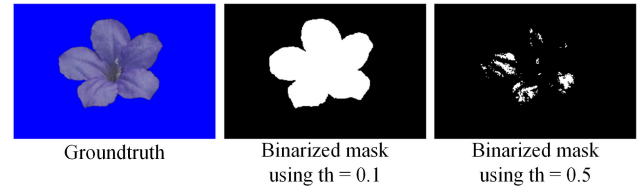


FIGURE 10. Effect of the threshold value on a binarized mask: wrong value can result in loss of the ground-truth foreground region, and, hence, performance evaluation error. For this ground-truth, a threshold value of 0.1 appears better than 0.5.

binarization threshold is set according to the dataset as 0.1 for Flowers, PASCAL VOC 2012 and xVIEW2 dataset, and 0.5 for CUB-200-2011. These threshold values are carefully chosen so that the resulting binarized mask is visually identical to the given ground-truth masks. Figure 10 shows that chosen a threshold of 0.1 is better than 0.5 which visibly results in loss of ground-truth foreground region and, hence, errors in performance evaluation. For ease of reproducibility, all the metrics used are coded using the Scikit-learn library. In the case of PASCAL VOC 2012 dataset [46], pixels around the border of an object that are marked 'void' in the ground-truth are not used in the assessment of our method.

D. EXPERIMENTAL RESULTS

To demonstrate the effectiveness of the proposed methods for image segmentation and evaluate the effects of design choices on performance, several experiments were conducted on benchmark datasets, and the results obtained are presented in this section.

1) IMAGE SEGMENTATION USING THE IQFT METHOD IN THE QUANTUM DOMAIN

Figure 11 presents a piece of visual evidence for the clustering capability of the IQFT method in the quantum domain. By visual inspection, even at 200 shots, the three dogs are automatically segmented out by the IQFT algorithm, without previously specifying the number of segments. For the upper bird image, $\theta_1 = \theta_2 = \theta_3 = \pi$ and the number of segments is 7. Setting $\theta_1 = \theta_2 = \theta_3 = 3\pi/2$ for the lower dog image resulted in 6 segments.

A visual comparison of the outputs with those from *k*-means, reveals that the IQFT quantum algorithm can spontaneously split an image into some significant segments. However, as revealed in Figure 11, the segments appear "noisy" below 500 shots for the dog and bird cases because a good number of shots is necessary for the simulator in Qiskit to build a probability distribution which is the backbone of the IQFT method. Figure 12 shows that the IQFT in the quantum domain and the IQFT-inspired algorithm have similar segmentation results. This conclusion is supported by the quantitative results in Table 1 for a 333×500 input image, where the values of the performance metrics are approximately equal for both techniques, except

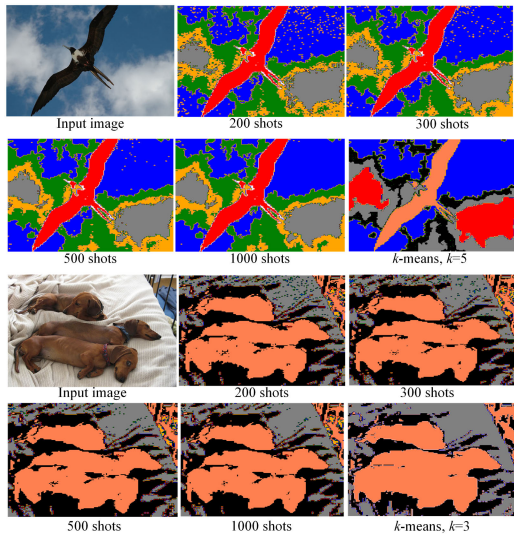


FIGURE 11. A visual evidence for the clustering capability of the IQFT method in the quantum domain. Different segments are displayed in different colors.

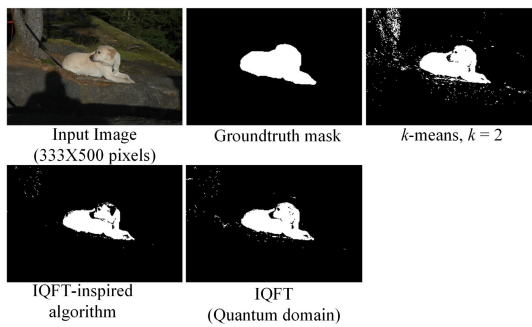


FIGURE 12. Comparing the performance of IQFT (quantum domain) with IQFT-inspired algorithm.

TABLE 1. Comparing the IQFT method in the quantum domain and IQFT-inspired algorithm in terms of IoU, DICE score and accuracy.

Performance metric	IQFT		
	Quantum Domain	IQFT-Inspired	<i>k</i> -means, <i>k</i> =2
IoU (%)	82.6	83.1	67.9
DICE score (%)	90.5	90.8	80.9
Accuracy (%)	98.6	98.7	96.6
Runtime (sec.)	3644.6	4.2	0.63

for the large runtime associated with the quantum domain implementation. The results from using *k*-means method are also provided for completeness. Because implementing of the IQFT method in Qiskit is computationally expensive, all subsequent experiments were carried out using the IQFT-inspired algorithm.

2) IMAGE SEGMENTATION USING ROBUST, IQFT-INSPIRED ALGORITHM IN THE CLASSICAL DOMAIN

Figure 13 shows an unsupervised image segmentation into foreground and background using the robust, IQFT-inspired method with $\theta_1 = \pi$, $\theta_2 = \pi/4$ and $\theta_3 = \pi/4$.

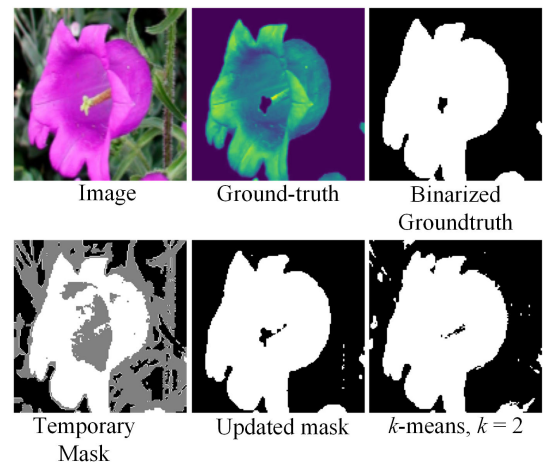


FIGURE 13. Behavior of our robust, IQFT-inspired algorithm on image dataset.

A temporary mask resulting from specifying a probability range of $p \geq 0.8$ is also shown. The white and black colors in the temporary mask identify the pixels that are labeled as foreground and background, respectively, with a probability of at least 80%. The gray color highlights approximately 39% of the total pixels which cannot be reliably determined by the IQFT method due to the specified probability threshold. The updated mask is adjusted by grouping the “gray” into either foreground or background based on their similarity score. In this work, cosine similarity is used because it has generally shown better performance in this application when compared to other similarity measures like Euclidean distance. The mask produced by the *k*-means algorithm is also provided for comparison.

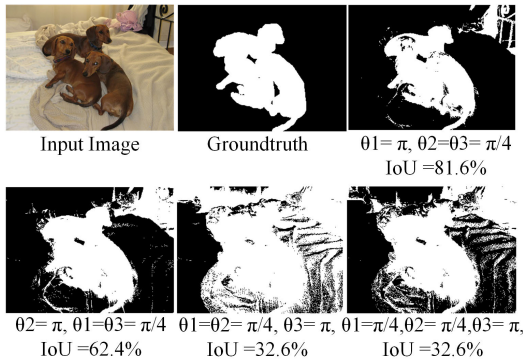
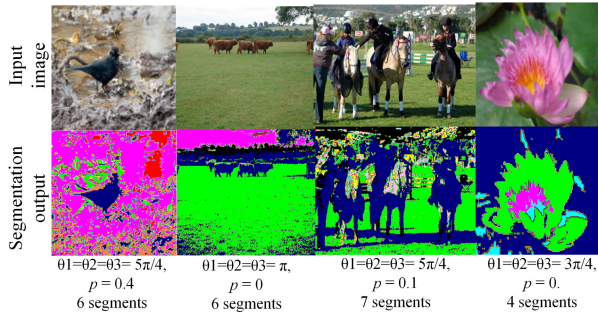
E. EFFECTS OF DESIGN CHOICES ON THE PERFORMANCE OF THE PROPOSED METHODS

1) EFFECTS OF ANGLE PARAMETERS ($\theta_1, \theta_2, \theta_3$) ON THE PERFORMANCE OF THE PROPOSED METHOD

To study the effects of the angle parameters on the proposed method, we generated 100,000 $\times$ 3 random numbers between 0 and 1 as normalized RGB values. We determined the maximum number of segments using different sets of $(\theta_1, \theta_2, \theta_3)$. The results in Table 2 reveal that number of segments varies with $(\theta_1, \theta_2, \theta_3)$. This variation occurs due to the modification of the input pattern, discussed in Section IV-B, as the angle parameters are varied. The last four rows in Table 2 show the combinations that usually output two segments. These are especially useful for cases like foreground-background segmentation, where two segments are required. Figure 14 illustrates the practical effect of the angle parameters on a real image. The performance of the IQFT method depends on the angle parameters and, in this example, $(\pi, \pi/4, \pi/4)$ achieves the highest IoU value of 81.6% for the foreground-background segmentation. Also, Figure 15 demonstrates that multiple segments are possible by setting the values of $(\theta_1, \theta_2, \theta_3)$ for such outputs.

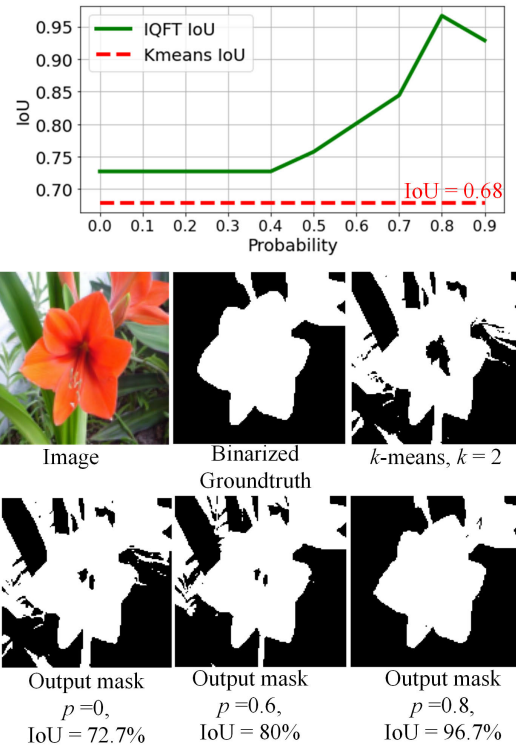
TABLE 2. Angle parameters ($\theta_1, \theta_2, \theta_3$) and the possible number of segments.

Parameter ($\theta_1, \theta_2, \theta_3$)	Maximum Number of Segments	Basis States
$(\pi/4, \pi/4, \pi/4)$	1	$ 000\rangle$
$(\pi/2, \pi/2, \pi/2)$	3	$ 000\rangle, 001\rangle, 010\rangle$
$(3\pi/4, 3\pi/4, 3\pi/4)$	5	$ 000\rangle, \dots, 100\rangle$
(π, π, π)	6	$ 000\rangle, \dots, 101\rangle$
$(5\pi/4, 5\pi/4, 5\pi/4)$	8	$ 000\rangle, \dots, 111\rangle$
$(3\pi/2, 3\pi/2, 3\pi/2)$	8	$ 000\rangle, \dots, 111\rangle$
$(2\pi, 2\pi, 2\pi)$	8	$ 000\rangle, \dots, 111\rangle$
$(\pi/4, \pi/2, \pi)$	2	$ 000\rangle, 001\rangle$
$(\pi, \pi/4, \pi/4)$	2	$ 000\rangle, 100\rangle$
$(\pi/4, \pi, \pi/4)$	2	$ 000\rangle, 010\rangle$
$(\pi/4, \pi/4, \pi)$	2	$ 000\rangle, 001\rangle$

**FIGURE 14.** Different two-segment outputs for different combinations of ($\theta_1, \theta_2, \theta_3$).**FIGURE 15.** Examples of segmentation outputs with multiple segments.

2) EFFECTS OF THE PROBABILITY PARAMETER (p) ON THE PERFORMANCE OF THE IQFT METHOD

To investigate how the choice of p value affects the quality of segmentation, we applied our algorithm to some random images and recorded the IoU scores as the probability value $p \in [0, 1]$ was varied while keeping the angle parameters constant. We observed that different segmentation qualities can be achieved by varying the probability value. A visual reference of this observation is given in Figure 16, showing the plot of IoU as a function of probability for the given image and the predicted segmentation masks at different probability values. For this image, the optimal result is given

**FIGURE 16.** Variation of the segmentation quality with probability: probability of $p = 0.8$ yields an optimum segmentation quality.

at $p = 0.8$ for $[\theta_1, \theta_2, \theta_3] = [\pi, \pi/4, \pi/4]$. Similar results for a different image are shown in Figure 17. In this case, $(\theta_1, \theta_2, \theta_3) = (\pi/4, \pi/2, \pi)$ yields an optimal result at $p = 0.9$, and reveals that the inclusion of the probability parameter in the robust, IQFT inspired algorithm can indeed significantly improve the segmentation quality of the IQFT algorithm. The performance of the k -means algorithm is also depicted for comparison.

3) EFFECTS OF THE PROBABILITY PARAMETER (p) ON THE SUM OF SQUARED ERROR (SSE) OF THE SEGMENTS

The SSE is a widely used criterion measure for intra-cluster homogeneity which determines the compactness of a cluster. Specifically, in minimum variance partitioning algorithms, clusters are determined by minimizing the within-cluster SSE criterion. To eliminate the user's effort in our algorithm, we employed SSE, as stated in Section IV-C, to automatically determine the value of p for an optimum or near optimum result. Figure 18 shows the graph of SSE for 10 different values of p . The avg. SSE in equation (9) was found to be 0.084. Since this value is closest to 0.082, the probability value for the required optimum value was set to $p = 0.7$. The optimum segmentation mask is also displayed in Figure 18.

F. OVERALL PERFORMANCE COMPARISON

The effectiveness of the robust, IQFT-inspired algorithm was validated by performing foreground-background

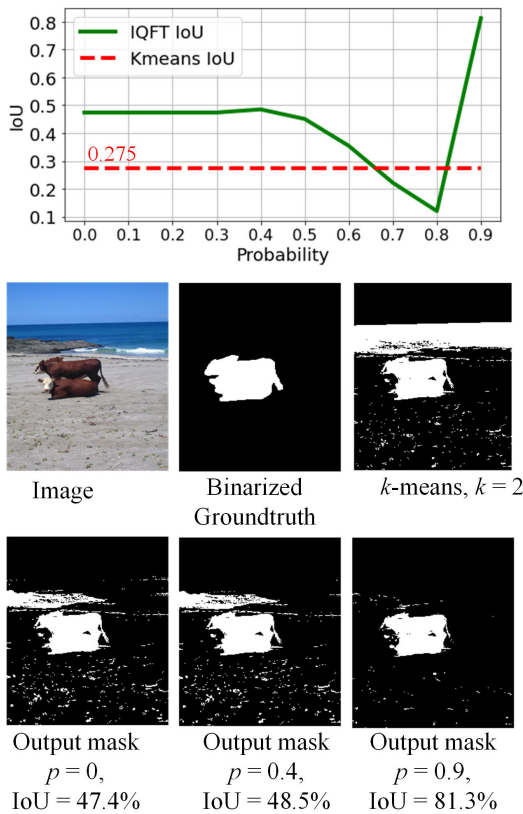


FIGURE 17. Variation of the segmentation quality with probability. In this case, optimum quality is achieved at probability of $p = 0.7$.

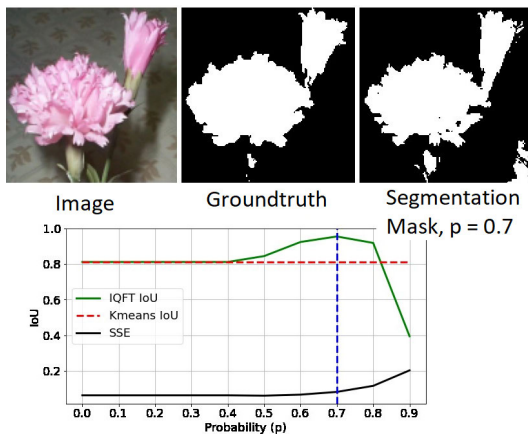


FIGURE 18. Effects of probability p on the SSE of the segments.

segmentation on the benchmark datasets in Section V-A and comparing the quantitative results to prior studies in terms of the IoU score, DICE score and accuracy.

Table 3 captures our foreground-background segmentation results on Flowers dataset and results from existing unsupervised segmentation methods: k -means, ReDO [14], GrabCut [15] and BigBiGAN [44] algorithms. Here, we set the parameters as $\theta_1 = \pi$, $\theta_2 = \pi/4$ and $\theta_3 = \pi/4$. Our improved, IQFT-inspired algorithm achieved the best average

TABLE 3. Comparing the unsupervised segmentation results on the Flowers dataset in terms of average IoU score, DICE score, and Accuracy.

Method	Avg. IoU (%)	Avg. DICE (%)	Accuracy (%)
k -means, $k=2$	71.2	81.2	85.4
GrabCut [15]	69.2	79.1	82.0
ReDo [14]	76.4	-	87.9
BigBiGAN [44]	54.0	-	76.5
IEM [16]	76.8	84.6	88.3
IQFT-inspired [9]	71.1	81.2	85.6
Robust, IQFT-inspired (ours)	74.2	83.2	86.7

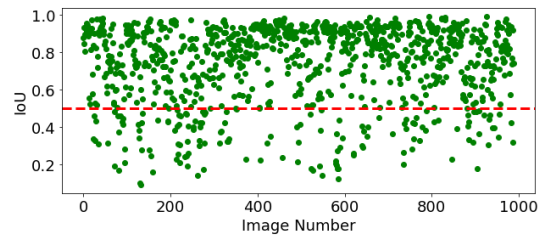


FIGURE 19. Performance of the robust, IQFT-inspired algorithm on Flowers dataset.

IoU score, DICE score and accuracy. Figure 19 shows that the algorithm is successful for many flower segmentation, especially when there are clear color differences between the background (leaves, ground, and others) and the foreground (flower). The robust method gives $\text{IoU} \geq 0.5$ for about 85% of the entire images used in this experiment. Considering the segmentation results on the Flowers dataset in Figure 20, the robust, IQFT-inspired algorithm outperforms both k -means and the IQFT-inspired algorithm in [9].

In Table 4 comparative results are presented for the unsupervised image segmentation on PASCAL VOC 2012 benchmark dataset. For these results, the parameters were set as $\theta_1 = \pi/4$, $\theta_2 = \pi/2$, $\theta_3 = \pi$ which are known in Section V-E1 to give two segments. The distribution of the IoU values for all the 2913 images in the segmentation category are shown in Figure 21, which reveals that about 21.5% of all the images have foreground $\text{IoU} \geq 0.5$. The DFC technique shows a better performance. While our method depends completely on the RGB values of the input image, the DFC technique utilizes a convolutional neural network (CNN) to extract deep features that contribute to its observed performance.

The graph of the average IoU score against probability in Figure 22 shows about 3.4% increase in the average IoU score for probability change from 0 to 0.9. This observation, again, underscores the significance of the probability parameter. Figure 21 shows some examples of foreground-background image segmentation results on PASCAL VOC 2012.

Table 5 shows comparative segmentation results on Caltech-UCSD Birds-200-2011 and Figure 23 displays the corresponding performance of the robust, IQFT algorithm on each image data. These figures reveal that our method, in general, shows good performance with $\text{IoU} \geq 0.5$ for

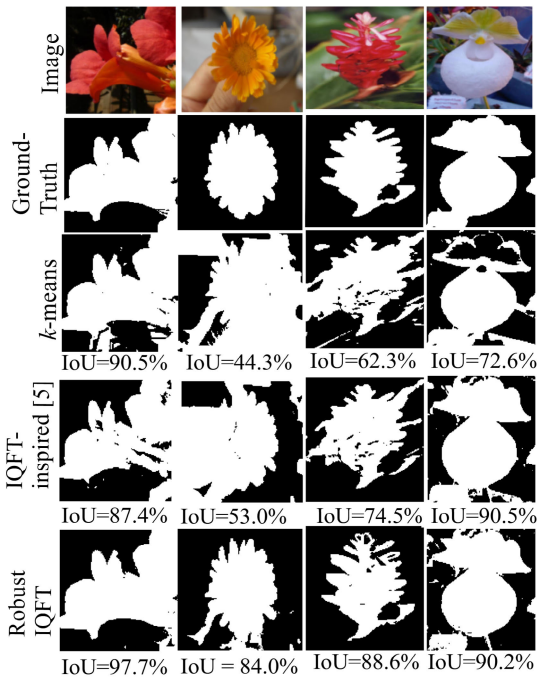


FIGURE 20. Some examples of segmentation results using robust, IQFT-inspired algorithm on Flowers dataset.

TABLE 4. Comparing the unsupervised segmentation results on PASCAL VOC 2012 dataset in terms of average IoU score, DICE score and accuracy.

Method	Avg. IoU (%)	Avg. DICE (%)	Accuracy (%)
<i>k</i> -means [8], <i>k</i> =2	31.7	46.4	63.0
GS [47], $\tau = 100$	26.8	-	-
GS [47], $\tau = 500$	36.5	-	-
IIC [48] <i>k</i> = 2	27.3	-	-
IIC [48] <i>k</i> = 20	20.1	-	-
DFC [8] w/ continuity loss, $\mu = 5$	35.2	-	-
IQFT-inspired [9]	32.6	45.3	59.7
Robust,IQFT-inspired (ours)	33.1	45.8	61.7

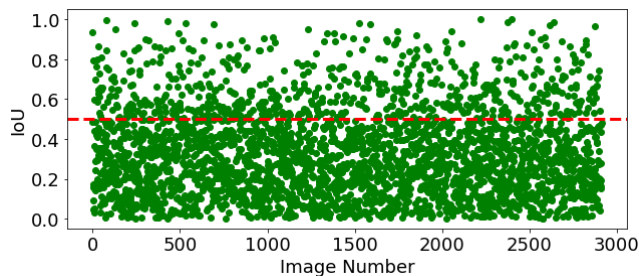


FIGURE 21. Performance of the robust, IQFT-inspired algorithm on PASCAL VOC 2012 dataset.

about 14% of 1000 test images from the CUB-200-2011 dataset. This low percentage is as a result of using only RGB color features which are not very effective for separating the foreground from the background for most images in the CUB-200-2011 dataset. PerturbGAN [17] and BigBiGAN [44] methods show better results because they benefited from

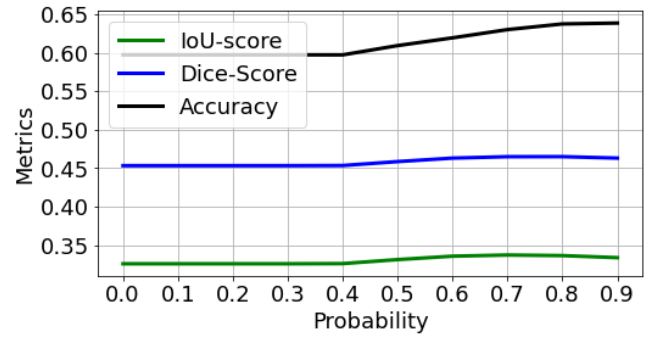


FIGURE 22. Performance of the robust, IQFT-inspired algorithm on PASCAL VOC 2012 dataset.

TABLE 5. Comparing the unsupervised segmentation results on CUB-200-2011 dataset in terms of average IOU score, DICE score, and accuracy.

Method	IoU (%)	DICE (%)	Accuracy (%)
<i>k</i> -means, <i>k</i> =2	28.1	40.2	68.6
GrabCut [15]	36.0	48.7	72.3
ReDo [14]	42.6	-	84.5
PerturbGAN [17]	38.0	-	-
BigBiGAN [44]	68.3	-	93.0
IEM [16]	52.2	66.0	88.6
IQFT-inspired [9]	27.3	40.0	71.4
IQFT-Robust,IQFT-inspired (ours)	24.9	37.0	64.1

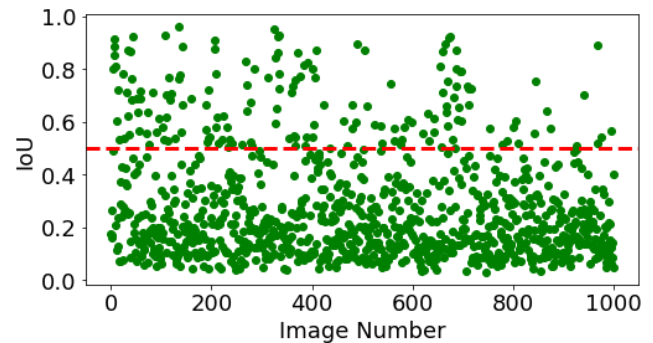


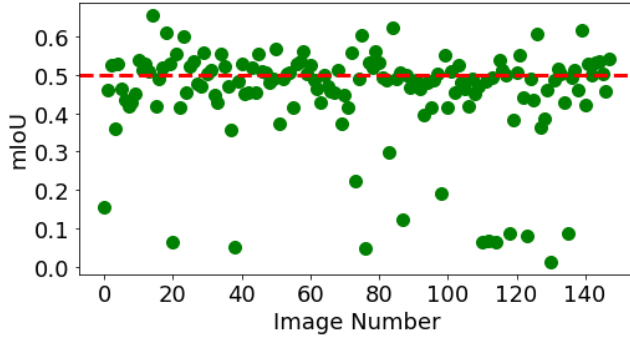
FIGURE 23. Performance of the robust, IQFT-inspired algorithm on CUB-200-2011 dataset.

the pre-trained generative model in their architecture. Also, IEM [16] outperforms our method because of the extra features gained from the CNN network.

The low resolution of the images in xVIEW2 dataset makes its segmentation a more challenging task. Nonetheless, we perform foreground-background segmentation on it to further evaluate the performance of our unsupervised methods, and Table 6 compares our method with *k*-means. Our method shows superior performance than *k*-means according to the mean-IoU of both the foreground and background, mean-DICE score, and mean-accuracy. From Figure 24, displaying the mean-IoU values of individual images, our method shows significant $mIoU \geq 0.5$ in about 42% of the entire images. Examples of foreground-background segmentation of the pre-disaster images for the “Joplin-tornado” disaster are

TABLE 6. Comparing the unsupervised segmentation results on the xVIEW2 challenge dataset in terms of mIoU score, mDICE score, and accuracy.

Method	mIoU (%)	mDICE (%)	Accuracy (%)
<i>k</i> -means, <i>k</i> =2	40.7	52.4	67.7
IQFT-inspired [9]	44.7	53.4	73.9
Robust, IQFT-inspired (ours)	45.2	55.1	74.8

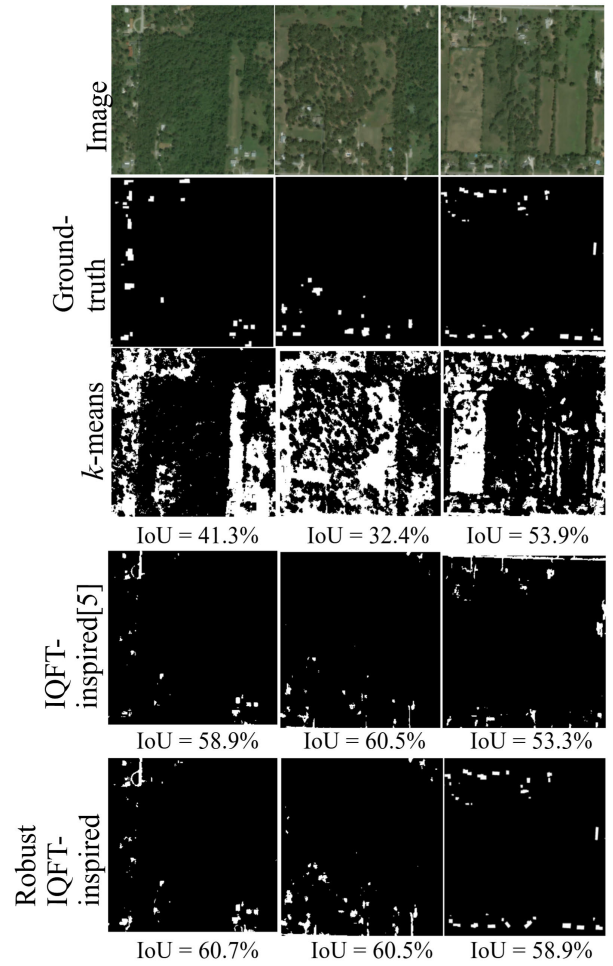
**FIGURE 24.** Performance of the improved, IQFT-inspired algorithm on xVIEW2 challenge dataset.

presented in Figure 25, where our method exhibits superior performance as shown by the mIoU values. For the robust IQFT-inspired algorithm, the measured average running time for the 1024×1020 is $284sec$ which is about 14 times greater than the running time for *k*-means.

G. COMPUTATIONAL PERFORMANCE EVALUATION

To comprehensively evaluate the efficiency of the proposed algorithm, we conducted a computational performance assessment focusing on key metrics: parameter count, memory usage, and inference time. The results of this evaluation are presented in Table 7, alongside a comparative analysis with the *k*-means clustering algorithm.

- 1) Parameter count: The fundamental parameter count for the *k*-means algorithm is kd [45], where *k* is the number of segments (clusters) and *d* is the number of dimensions in the data. For the proposed method, the parameter count is $d + 1$. Focusing on two segments and the RGB color space (with 3 dimensions) considered in this work, the parameter count for the *k*-means algorithm becomes 6 while for the proposed method it becomes 4: three angle parameters ($\theta_1, \theta_2, \theta_3$) and one probability parameter discussed in Section IV.
- 2) Memory usage: To evaluate the feasibility of deploying the proposed algorithm on different hardware, we utilized *tracemalloc*, a standard Python library, to track memory allocation and capture memory usage patterns during the execution of both the *k*-means algorithm and the proposed algorithm. The peak memory allocation during the measurement period is reported in Table 7.
- 3) Inference Time: In Table 7, we present the average inference time of the proposed algorithm for a single

**FIGURE 25.** Examples of foreground-background segmentation of the pre-disaster images for "Joplin-tornado" disaster.**TABLE 7.** Computational performance metrics of the proposed algorithm vs. *k*-means Clustering, where *k* and *d* represent the number of segments and dimensionality of the input data, respectively.

Method	Parameter Count	Memory Usage (MB)	Average Inference Time (sec)
<i>k</i> -means	kd	19.74	0.20
Robust IQFT	$d + 1$	21.58	0.98

128×128 image input, offering insights into its real-time application capability. Additionally, we compare this with the processing time of the *k*-means algorithm

H. ADVANTAGES AND LIMITATIONS OF THE PROPOSED METHOD

Overall, the proposed method outperforms the classical *k*-means and some state-of-the-art algorithms considered in this work by enhancing the number of features for each pixel from three to eight. These additional features provide a richer source of information, potentially enhancing the discriminative power and improving the clustering process.

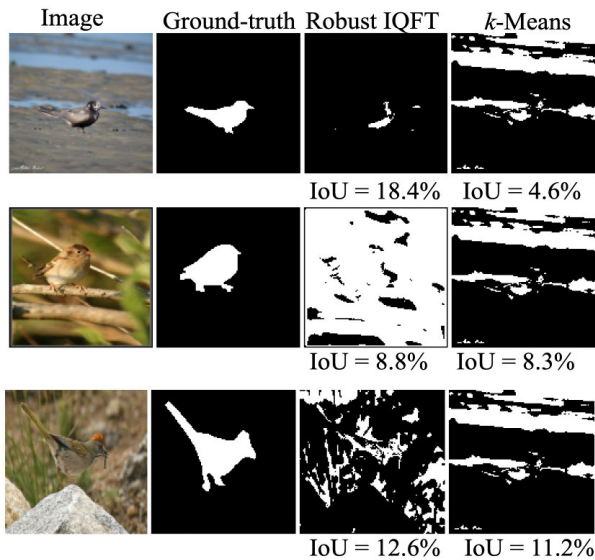


FIGURE 26. Instances highlighting the limitations of the proposed method. Since the proposed method relies only on the color features, it under-performs in cases where foreground colors significantly overlap with those of the background.

Unlike *k*-means, which relies solely on RGB features, the proposed method leverages these additional features to achieve superior performance, particularly in discerning complex patterns and structures within the image data. The conclusion gains further support from the significant improvement in segmentation performance achieved by applying the *k*-means method to the datasets after transforming them from 3D to 8D, as detailed in Section IV.

Section V shows the dependency of the proposed method's performance on the values of angle parameters θ_1 , θ_2 , θ_3 , and the probability parameter (p), impacting both the number of segments and segmentation quality. Table 2 presents diverse combinations of angle parameters yielding two segments. However, segmentation quality remains dependent on the dataset, defining a limitation of our approach. Similarly, selecting an appropriate probability parameter value can yield over 60% performance improvement. This study employs the average sum of squared error to automate the algorithm and eliminate user involvement. Nonetheless, this heuristic doesn't always ensure optimal results, as evidenced by IoU values for the Flower dataset differing between the optimum probability value of 0.7 (IoU = 78.1%) and the average SSE value (IoU = 74.2%), showcasing another limitation.

It is noteworthy that our method relies solely on the RGB color features of the input image to distinguish the foreground from the background. As evidenced by the overall performance results shown in Figures 19, 21 and 23 for images in the Flowers, PASCAL VOC 2012, and CUB-200-2011 datasets, respectively, our method may significantly fall short of optimal performance, particularly in scenarios where the foreground color distribution significantly overlaps with that of the background. An illustration of this challenge is presented in Figure 26. In such cases, incorporating additional

features may be required to enhance the performance of our method.

VI. CONCLUSION

In this work, we extended our previous work [9] and proposed a robust, IQFT-inspired algorithm for unsupervised image segmentation. Particularly, in addition to the angle parameters in [9], a probability parameter is introduced in robust, IQFT-inspired algorithm to fine-tune an output segmentation mask to realize a better result. The average sum of squared error (SSE) is also included to automatically determine the value of the probability parameter so that it is fully automatic and does not require the user's efforts.

Through experimental results from four benchmark datasets, it is observed that the inclusion of the probability parameter in the robust, IQFT-inspired algorithm achieved a better performance over the method in [9] and some recent computationally complex methods, especially when the RGB color features are sufficient to differentiate the foreground from the background like in the Flowers dataset. A thorough evaluation of the proposed method across four diverse datasets demonstrates its superiority over classical *k*-means. Specifically, our method exhibits significant performance improvements on various benchmarks: surpassing classical *k*-means by up to 4.21% on the PASCAL VOC 2012 segmentation benchmark, up to 4.4% on the Flowers dataset, and up to 11.1% on the xVIEW2 challenge dataset in terms of IoU.

It is worth mentioning that the proposed robust IQFT-inspired algorithm has a low computational cost when compared to the GAN-based methods and does *not* require training or pre-trained model. Furthermore, experimental results obtained for image samples from Flowers, xVIEW2, and PASCAL VOC 2012 datasets demonstrated that the proposed method outperforms *k*-means in all cases and by a big margin when the task becomes more challenging such as for the xVIEW2 dataset. This makes the proposed method a promising choice for mission-critical applications.

While our proposed method demonstrates promising results, there are avenues for further improvement and refinement. In particular, addressing the limitations identified in this study will be a focus of future research endeavors.

- 1) Exploring Alternative Heuristics: One avenue for future research entails exploring alternative heuristic approaches to achieve optimal segmentation results. Although the sum of squared error (SSE) has been employed in our methodology due to its computational efficiency, it may not consistently yield the best segmentation outcomes. Investigating alternative heuristic methods or optimization techniques could potentially enhance segmentation accuracy and robustness
- 2) Optimizing Angle Parameter Selection: Another promising avenue for future investigation involves identifying the optimal combination of angle parameters (θ_1 , θ_2 , θ_3) to achieve optimal segmentation results.

Our future work will focus on developing automated strategies to determine the best angle parameter combination based on dataset properties. By doing so, we aim to enhance the performance and applicability of our segmentation methodology, enabling more accurate and robust image segmentation techniques across a wide range of real-world applications.

REFERENCES

- [1] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988.
- [2] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 12, pp. 2481–2495, Dec. 2017.
- [3] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.
- [4] M. Noroozi and P. Favaro, "Unsupervised learning of visual representations by solving jigsaw puzzles," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 69–84.
- [5] K. H. Williford, K. A. Farley, K. M. Stack, A. C. Allwood, D. Beaty, L. W. Beegle, R. Bhartia, A. J. Brown, M. de la Torre Juarez, S.-E. Hamran, M. H. Hecht, J. A. Hurowitz, J. A. Rodriguez-Manfredi, S. Maurice, S. Milkovich, and R. C. Wiens, "The NASA Mars 2020 rover mission and the search for extraterrestrial life," in *From Habitability to Life on Mars*. Amsterdam, The Netherlands: Elsevier, 2020, pp. 275–308.
- [6] A. Zhao, G. Balakrishnan, F. Durand, J. V. Guttag, and A. V. Dalca, "Data augmentation using learned transformations for one-shot medical image segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 8535–8545.
- [7] S. Hao, Y. Zhou, and Y. Guo, "A brief survey on semantic segmentation with deep learning," *Neurocomputing*, vol. 406, pp. 302–321, Sep. 2020.
- [8] W. Kim, A. Kanezaki, and M. Tanaka, "Unsupervised learning of image segmentation based on differentiable feature clustering," *IEEE Trans. Image Process.*, vol. 29, pp. 8055–8068, 2020.
- [9] T. Akinola, X. Li, R. Wilkins, P. Obiomon, and L. Qian, "Inverse quantum Fourier transform inspired algorithm for unsupervised image segmentation," 2023, *arXiv:2301.04705*.
- [10] M. A. Hanif, F. Khalid, R. V. W. Putra, S. Rehman, and M. Shafique, "Robust machine learning systems: Reliability and security for deep neural networks," in *Proc. IEEE 24th Int. Symp. On-Line Test. Robust Syst. Design*, Jul. 2018, pp. 257–260.
- [11] T. Freiesleben and T. Grote, "Beyond generalization: A theory of robustness in machine learning," *Synthese*, vol. 202, no. 4, p. 109, Sep. 2023.
- [12] M.-E. Nilsback and A. Zisserman, "Delving into the whorl of flower segmentation," in *Proc. BMVC*, 2007, pp. 1–10.
- [13] C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie, "The Caltech-UCSD birds-200-2011 dataset," California Inst. Technol., Pasadena, CA, USA, Jul. 2011.
- [14] M. Chen, T. Artières, and L. Denoyer, "Unsupervised object segmentation by redrawing," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–20.
- [15] C. Rother, V. Kolmogorov, and A. Blake, "'GrabCut' interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, 2004.
- [16] P. Savarese, S. S. Y. Kim, M. Maire, G. Shakhnarovich, and D. McAllester, "Information-theoretic segmentation by inpainting error maximization," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 4028–4038.
- [17] A. Bielski and P. Favaro, "Emergence of object segmentation in perturbed generative models," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 1–28.
- [18] R. Pandey and R. Lalchhanhima, "Segmentation techniques for complex image: Review," in *Proc. Int. Conf. Comput. Perform. Eval.*, Jul. 2020, pp. 804–808.
- [19] S. Uk Lee, S. Yoon Chung, and R. Hong Park, "A comparative performance study of several global thresholding techniques for segmentation," *Comput. Vis., Graph., Image Process.*, vol. 51, no. 2, p. 218, Aug. 1990.
- [20] H. G. Kaganami and Z. Beiji, "Region-based segmentation versus edge detection," in *Proc. 5th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process.*, Sep. 2009, pp. 1217–1221.
- [21] M. A. Wani and B. G. Batchelor, "Edge-region-based segmentation of range images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 3, pp. 314–319, Mar. 1994.
- [22] Y. Zou and B. Liu, "Survey on clustering-based image segmentation techniques," in *Proc. IEEE 20th Int. Conf. Comput. Supported Cooperat. Work Design (CSCWD)*, May 2016, pp. 106–110.
- [23] N. Salman, "Image segmentation based on watershed and edge detection techniques," *Int. Arab J. Inf. Technol.*, vol. 3, no. 2, pp. 104–110, 2006.
- [24] D. T. Pham, S. S. Dimov, and C. D. Nguyen, "Selection of K in K-means clustering," *Proc. Inst. Mech. Engineers, Part C, J. Mech. Eng. Sci.*, vol. 219, no. 1, pp. 103–119, 2005.
- [25] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, Sep. 1993.
- [26] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, and J. Garcia-Rodriguez, "A review on deep learning techniques applied to semantic segmentation," 2017, *arXiv:1704.06857*.
- [27] C. Lang, G. Cheng, B. Tu, and J. Han, "Learning what not to segment: A new perspective on few-shot segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 8047–8057.
- [28] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few-shot learning," *ACM Comput. Surveys*, vol. 53, no. 3, pp. 1–34, May 2021.
- [29] G. Cheng, C. Lang, and J. Han, "Holistic prototype activation for few-shot segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 4, pp. 4650–4666, Apr. 2023.
- [30] C. Lang, G. Cheng, B. Tu, C. Li, and J. Han, "Base and meta: A new perspective on few-shot segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 1, no. 1, pp. 1–18, Jul. 2023.
- [31] C. Lang, G. Cheng, B. Tu, and J. Han, "Few-shot segmentation via divide-and-conquer proxies," *Int. J. Comput. Vis.*, vol. 132, no. 1, pp. 261–283, Jan. 2024.
- [32] J. MacQueen, "Classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, 1967, pp. 281–297.
- [33] M. Á. Carreira-Perpiñán, "A review of mean-shift algorithms for clustering," 2015, *arXiv:1503.00687*.
- [34] S. Yuan, C. Wen, B. Hang, and Y. Gong, "The dual-threshold quantum image segmentation algorithm and its simulation," *Quantum Inf. Process.*, vol. 19, no. 12, pp. 1–21, Dec. 2020.
- [35] S. Caraiman and V. I. Manta, "Histogram-based segmentation of quantum images," *Theor. Comput. Sci.*, vol. 529, pp. 46–60, Apr. 2014.
- [36] S. Caraiman and V. I. Manta, "Image segmentation on a quantum computer," *Quantum Inf. Process.*, vol. 14, no. 5, pp. 1693–1715, May 2015.
- [37] L. K. Grover, "Quantum mechanics helps in searching for a needle in a haystack," *Phys. Rev. Lett.*, vol. 79, no. 2, pp. 325–328, Jul. 1997.
- [38] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [39] A. Abbas et al., "Learn quantum computation using Qiskit," 2020. [Online]. Available: <https://qiskit.org/textbook>
- [40] R. S. Sutor, *Dancing With Qubits: How Quantum Computing Works and How It Can Change the World*. Birmingham, U.K.: Packt Publishing Ltd, 2019.
- [41] R. Gupta, R. Hosfelt, S. Sajeve, N. Patel, B. Goodman, J. Doshi, E. Heim, H. Choset, and M. Gaston, "XBD: A dataset for assessing building damage from satellite imagery," 2019, *arXiv:1911.09296*.
- [42] D. S. Steiger, T. Häner, and M. Troyer, "ProjectQ: An open source software framework for quantum computing," *Quantum*, vol. 2, p. 49, Jan. 2018.
- [43] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 658–666.
- [44] A. Voynov, S. Morozov, and A. Babenko, "Object segmentation without labels with large-scale generative models," in *Proc. Int. Conf. Mach. Learn. (ICML)*, vol. 139, 2021, pp. 10596–10606.
- [45] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Nov. 2011.

- [46] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The Pascal visual object classes challenge: A retrospective," *Int. J. Comput. Vis.*, vol. 111, no. 1, pp. 98–136, Jan. 2015.
- [47] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, Sep. 2004.
- [48] X. Ji, A. Vedaldi, and J. Henriques, "Invariant information clustering for unsupervised image classification and segmentation," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 9864–9873.



TAOREED A. AKINOLA (Member, IEEE) received the B.S. degree in electrical and electronic engineering from Obafemi Awolowo University, Ife, Nigeria, in 2008, and the M.S. degree in electrical engineering from Prairie View A&M University, Prairie View, TX, USA, in 2022, where he is currently pursuing the Ph.D. degree in electrical engineering.

Since 2021, he has been a Research Assistant with the Center of Excellence in Research and

Education for Big Military Data Intelligence (CREDIT Center). His research interests include machine learning and quantum machine learning.



XIANGFANG (LINDSEY) LI (Member, IEEE) received the M.S. and Ph.D. degrees in computer engineering from Rutgers University, in 2003 and 2007, respectively.

Before joining Prairie View A&M University (PVAMU) as an Assistant Professor, she was a TEES Associate Research Scientist with Texas A&M University, College Station, TX, USA. Currently, she is an Associate Professor with the Department of Electrical and Computer Engineering,

PVAMU, and a member of Texas A&M University System. She actively participated in the Bioinformatics Training Program sponsored by NIH and has co-authored over 100 peer-reviewed publications. Her research interests include computer networking and communication, computational and systems biology, systems pharmacology, data ethics, and artificial intelligence.

Dr. Li was a recipient of the Outstanding Researcher of the Year Award from the Roy G. Perry College of Engineering, PVAMU, in 2017.



RICHARD WILKINS (Member, IEEE) received the B.S. degree in physics from the University of Pittsburgh and the Ph.D. degree in physics from the University of Michigan, in 1991, where he performed some of the first experiments on single electron tunneling in nano-scale metal structures and tunneling spectroscopy on high-T_c superconductors.

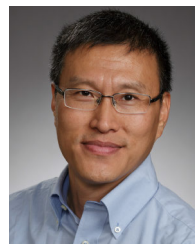
He is currently the Associate Dean of Research for the Roy G. Perry College of Engineering, the Director of the NASA Center for Radiation Engineering and Science for Space Exploration, and a Professor with the Department of Electrical and Computer Engineering, Prairie View A&M University. He has co-authored over 50 journal publications, including several articles on radiation effects on SiC power devices, novel electronic materials, and quantum and nano-scale devices. His research interests include radiation effects on electronic devices and materials and radiation dosimetry.



PAMELA H. OBIOMON (Member, IEEE) received the B.S. degree in electrical engineering from The University of Texas at Arlington, the M.S. degree in engineering from Prairie View A&M University, and the Ph.D. degree in electrical engineering from Texas A&M University.

She has been a Faculty Member with the Department of Electrical and Computer Engineering, since 2003. In 2013, she took on the role of the Department Head and began her role as the Dean of the Roy G. Perry College of Engineering, in September 2018. Before coming to PVAMU, she taught courses with Rochester Institute of Technology, Rochester, NY, USA. She was with the Shuttle Avionics Integration Laboratory, Johnson Space Center, Houston, TX, USA.

Dr. Obiomon holds memberships in IEEE, IEEE Women in Engineering, and Eta Kappa Nu.



LIJUN QIAN (Senior Member, IEEE) received the B.S. degree from Tsinghua University, Beijing, China, the M.S. degree from the Technion—Israel Institute of Technology, Haifa, Israel, and the Ph.D. degree from Rutgers University, USA.

He was a Visiting Professor with Aalto University, Finland. He is currently a Regents Professor and holds the AT&T Endowment with the Department of Electrical and Computer Engineering, Prairie View A&M University (PVAMU), Prairie View, TX, USA. He is also the Founder and the Director of the Center of Excellence in Research and Education for Big Military Data Intelligence (CREDIT Center). His research interests are in the areas of artificial intelligence, machine learning, big data analytics, wireless communications and mobile networks, network security, and computational systems biology.

Dr. Qian was a member of Technical Staff of the Bell-Laboratories Research, Murray Hill, NJ, USA.

...