

# 3D Quantum-inspired Self-supervised Tensor Network for Volumetric Segmentation of Brain MR Images

Debanjan Konar, *MIEEE*, Siddhartha Bhattacharyya, *SMIEEE*, Tapan K. Gandhi, *SMIEEE*,  
Bijaya K. Panigrahi, *SMIEEE*, and Richard Jiang, *SMIEEE*

**Abstract**—This paper introduces a novel shallow self-supervised tensor neural network for volumetric segmentation of brain MR images obviating training or supervision. The proposed network is a 3D version of the Quantum Inspired Self Supervised Neural Network (QIS-Net) architecture and is referred to as 3D Quantum-inspired Self-supervised Tensor Neural Network (3D-QNet). The underlying architecture of 3D-QNet is composed of a trinity of volumetric layers viz. input, intermediate and output layers inter-connected using a 26-connected third-order neighborhood-based topology for voxel-wise processing of 3D MR image data suitable for semantic segmentation. Each of the volumetric layers contains quantum neurons designated by *qubits* or quantum bits. The incorporation of tensor decomposition in quantum formalism leads to faster convergence of the network operations to preclude the inherent slow convergence problems faced by the self-supervised networks. The segmented volumes are obtained once the network converges. The suggested 3D-QNet is tailored and tested on the BRATS 2019 data set extensively in the experiments carried out. 3D-QNet has achieved promising dice similarity while compared with the intensively supervised convolutional network-based models 3D-UNet, Vox-ResNet, DRINet, and 3D-ESPNet, thus facilitating annotation free semantic segmentation using a self-supervised shallow network.

**Index Terms**—Quantum Computing, Volumetric Medical Image Segmentation, QIS-Net, 3D-UNet, Vox-ResNet

## I. INTRODUCTION

Automatic volumetric brain Magnetic Resonance (MR) image segmentation assisted by contextual information yields Volumes of Interest (VOIs), which are critical to glioma patients. Deeply supervised Convolutional Neural Networks (CNN) have achieved respectable accuracy in 2D medical image segmentation [1]. However, in automatic 3D MR image data segmentation, deeply supervised 3D-CNNs suffer from manually affected challenges viz. acquiring sufficient 3D annotated data for suitable training, high heterogeneity

and dimensionality of 3D MR images, complex anatomical environments and the need for optimizing the 3D neural networks [2], [3]. Hence, structural neuro-imaging research calls for self-supervised learning for accurate and fast segmentation of brain images with different modalities.

Given 3D MR image data, the primary aim of our proposed 3D-QNet architecture is to perform volumetric segmentation for brain tumor identification alleviating supervision or training. Our proposed 3D-QNet architecture is centered on the self-supervised bi-directional counter propagation of the quantum states obviating the time-intensive quantum back-propagation algorithm for faster convergence. The network hyper-parameters associated with the gray-level thresholding are adaptive in nature, and voxel-wise context-sensitive information is exhibited in quantum formalism as reported in this article. The current voxel-wise segmentation work has significant contributions over 2D brain image segmentation [4]–[6] as given below.

- 1) We propose a shallow voxel-wise quantum-inspired self-supervised neural network referred to as 3D-QNet which has significant relevance on volumetric MR image segmentation.
- 2) In this work, 26-connected quantum fuzzy context-sensitive voxel information is processed to integrate the appearance of low-level and high-level local image features with wide intensity variations and implicit shape of the VOIs, thereby enabling accurate volumetric segmentation of 3D MR images.
- 3) A novel generalized quantum-inspired self-supervised learning is proposed using a tensor representation of weight vector for high dimensional data and employed in our suggested 3D-QNet for brain MR image segmentation.
- 4) The convergence analysis of the proposed 3D-QNet is also demonstrated with super-linearity. The primary aim at incorporating quantum computing in our proposed 3D network architecture is to exploit the features of quantum correlation and to accelerate the speed of convergence of the network operation, simultaneously improving the discrimination ability to yield fast and accurate segmentation.

The organization of the remaining sections of the

D. Konar, T. K. Gandhi, and B. K. Panigrahi are with the Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, India, Email: konar.debanjan@ieee.org, tgandhi@iitd.ac.in and bkpanigrahi@ee.iitd.ernet.in

S. Bhattacharyya is with the Department of Department of Computer Science and Engineering, CHRIST University, Bangalore, India, Email: dr.siddhartha.bhattacharyya@gmail.com India

R. Jiang is with the School of Computing and Communications, Lancaster University, Lancaster, UK, Email: r.jiang2@lancaster.ac.uk

manuscript is as follows: a comprehensive literature review about various deep learning-based volumetric segmentation of medical images and the challenges are presented in Section II. Section III illustrates the fundamental concepts of quantum computing. The novel self-supervised 3D-QNet architecture with quantum-inspired tensor network model for voxel-wise segmentation of three dimensional MR images is introduced in Section IV. The experimental outcome and discussions are provided in Section V. Section VI states the concluding remarks of the proposed work and sheds light into the future directions of research.

## II. RELATED WORKS

Recent years have witnessed a surge in the application of deep learning networks in brain lesion segmentation [7]–[10]. However, in contrast to automated volumetric segmentation of brain MR images, 2D convolutional neural network architectures (CNNs) [7], [9], [10] process the MR images in 2D independent slice-wise fashion which leads to non-optimal use of 3D contextual feature information of volumetric MR image data. In turn, 3D CNN based architectures extract rich spatial and contextual features and perform voxel-wise segmentation of volumetric brain MR images [11]–[13]. Kamnitsas *et al.* [11] suggested a dual path 3D CNN incorporating local and larger contextual feature information to obviate the computationally complex 3D MR image processing and to exhibit dense inference on medical image segmentation. A flexible network, 3D-UNet architecture [13] achieved remarkable success on brain MR image semantic segmentation. Of late, to exploit the 3D contextual information, Brebisson *et al.* [14] employed 2D CNNs on three orthogonal 2D patches and formed 3D patches in combination to reduce the memory requirements. However, 3D CNN networks suffer from slow convergence problems owing to computationally exhaustive 3D convolution operations and extensive training procedures. However, despite popularity among medical and computer vision researchers, U-Net architectures [13] fall short in scalability and are unable to distinguish the distinctive features (shape, size, intensity, location etc.) learned at the convolutional layers. Moreover, it suffers from the vanishing gradient problem when the number of feature layers is increased for better representation of the features. Various deeper network architectures obviating the vanishing gradient problem have been proposed concurrently for voxel-wise medical image segmentation including VoxResNet [15], DRINet [16] and 3D-ESPNet [17]. However, these deeply supervised network architectures suffer in computational complexity and slow-convergence with an increase in the number of feature layers in the network architecture. Currently, self-supervised/semi-supervised/weakly supervised networks have gained significant attention among computer vision and medical research community due to lack of annotated images for deep supervision [18], [19]. Nevertheless, these self-supervised networks [18], [19] for volumetric brain segmentation rely on pre-trained 3D CNN models, and hence these are not fully self-supervised networks. It inspires us to develop 3D self-supervised neural network architectures for volumetric brain segmentation.

The main problem with the classical self-supervised neural network models lies in the fact that they do not converge fast and hence the segmented outcome is distorted due to slow convergence problems [20]–[22]. Numerous quantum neural networks have been evolved in the last few decades replicating classical neural networks and offering faster processing while compared with the classical counterparts [23]–[28]. The quantum version of the classical self-supervised neural network architectures [29]–[32] offer a potential solution for faster and efficient image segmentation and surpasses the classical counterparts. Konar *et al.* recently developed quantum-inspired neural network models referred to as QIS-Net [4] and QIBDS-Net [5] suitable for brain MR image segmentation. These networks have been found to attain promising outcome in complete brain tumor segmentation. The optimized version of the network (Opti-QIBDS Net) [6] is also proposed for optimal segmentation of brain tumors and serves the motivation behind assimilation of quantum-inspired computing in the current 3D-QNet architecture.

### A. Motivation

Despite the remarkable success achieved in volumetric brain MR image segmentation, 3D CNN architectures still face some inherent challenges owing to deeper and complex network architectures [11], [13]–[15]:

- 1) Owing to complex anatomical properties of 3D MR brain images, majority of 3D CNN architectures require a large number of parameters to capture 3D contextual representative feature information.
- 2) High computational (GPU) and memory resources required for large scale 3D CNNs pose a potential concern in widespread clinical applications.
- 3) Automated volumetric brain MR image segmentation using 3D CNN architectures often confronts overfitting, slow-convergence and vanishing gradient problems. Moreover, it is a paramount task to manipulate the hyper-parameters of the underlying 3D architecture.
- 4) In addition, the availability of 3D annotated data for training a 3D CNN is not sufficient and very expensive resulting in lack of image-specific adaptability.

## III. FUNDAMENTALS OF QUANTUM COMPUTING

The basic concept of quantum computing deals with the principles of quantum mechanics and offers to demonstrate the quantum computing algorithms which rely on quantum bits having quantum operations on *qubits* [33].

### A. Quantum Bits and Tensor Products

The basic element equivalent to classical bits in quantum computing is known as quantum bit or *qubit* and is represented using *Dirac* notations  $|0\rangle$  and  $|1\rangle$ . However, unlike classical computing, quantum bits are expressed as a linear combination of probability amplitudes often known as superposition as follows [28].

$$|\phi\rangle = \cos \frac{\alpha}{2} |0\rangle + e^{i\frac{\theta}{2}} \sin \frac{\alpha}{2} |1\rangle \quad (1)$$

where,  $0 \leq \alpha \leq \pi$  and  $0 \leq \theta \leq 2\pi$ .

Hence, *qubits* reside in the Hilbert space parametrized by the continuous variables  $\theta$  and  $\alpha$ . In quantum formalism, the *tensor products* of the subspace form the full Hilbert Hyperspace,  $H$  as

$$H = \otimes_{t=1}^n H_t \quad (2)$$

A set of  $n$  basis states (designated as  $|\phi_j\rangle$ ) comprising  $0 - 1$  can form a *qubit* system  $|\psi\rangle$ , of size  $\log n$  in the Hilbert space,  $H$  as follows.

$$|\psi\rangle = \sum_j^n p_j |\phi_j\rangle \quad (3)$$

where,  $p_j$  is the probability amplitude and  $|\phi_j\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \otimes \dots \otimes |\phi_n\rangle$ . For example, using two *qubits*, four distinct tensor sub-spaces can be created as basis  $|0\rangle \otimes |0\rangle$ ,  $|0\rangle \otimes |1\rangle$ ,  $|1\rangle \otimes |0\rangle$  and  $|1\rangle \otimes |1\rangle$  often represented as  $|00\rangle$ ,  $|01\rangle$ ,  $|10\rangle$ , and  $|11\rangle$ , respectively.

### B. Input Data Encoding and Tensor Decomposition

A tensor product basis relies on the local input feature map  $\{\Phi^{d_j}(\alpha_j)\}$  in the Hilbert space of functions over  $\alpha_j \in [0, 1]$  as [27].

$$|\Phi^{d_1, d_2, \dots, d_N}(\alpha)\rangle = |\phi^{d_1}(\alpha_1)\rangle \otimes |\phi^{d_2}(\alpha_2)\rangle \otimes \dots \otimes |\phi^{d_N}(\alpha_N)\rangle \quad (4)$$

where,  $d_j$  varies from  $1 \dots N$  ( $N$ -dimensional vector). A function,  $f^l(\alpha)$  can be realized using the tensor product of input local feature map  $\phi^{d_j}(\alpha_j)$  and the network weight decomposition  $\Psi$ , as follows.

$$f^l(\alpha) = \Psi^l \cdot \Phi(\alpha) = |\Psi_{d_1 d_2 \dots d_N}\rangle |\phi(\alpha_1^{d_1})\rangle \otimes |\phi(\alpha_2^{d_2})\rangle \dots \otimes |\phi(\alpha_N^{d_N})\rangle \quad (5)$$

Hence, the local feature map  $\phi^{d_j}(\alpha_j)$  forms a basis for a Hilbert space of functions defined over  $\alpha \in [0, 1]$  and the tensor product basis  $\Phi^{d_1, d_2, \dots, d_N}(\alpha)$  forms a Hilbert space of functions defined over  $\alpha \in [0, 1]^N$ . Considering the dimensions of the input feature vector restricted to  $N = 2$ ,  $\phi(\alpha)$  is defined as

$$\begin{aligned} \phi(0) &= [0, 1] \\ \phi(1) &= [1, 0] \end{aligned} \quad (6)$$

In order to enhance and extract the contextual information from high dimensional data, Trucker tensor decomposition is suitable for neural network layer decomposition [34].

## IV. 3D QUANTUM-INSPIRED SELF-SUPERVISED TENSOR NEURAL NETWORK (3D-QNET) ARCHITECTURE

In this article, a 3D version of the QIS-Net [4] referred to as 3D Quantum-inspired Self-supervised Tensor Neural Network (3D-QNet) with self-supervised tensor learning is proposed for automatic voxel-wise segmentation of MR images. The 3D-QNet comprises trinity of volumetric layers of quantum neurons arranged as input, intermediate and output layers. A schematic outline of the proposed 3D-QNet architecture is shown in Figure 1. The input volume ( $M \times N \times P$ ) is normalized and propagated from the 3D input layer to

the successive 3D hidden and output layers of the 3D-QNet architecture for processing through  $3 \times 3 \times 3$  voxels. Each of the three volumetric layers of 3D-QNet architecture is fully intra-linked with *qubits* using a 3D-matrix representation. Each 3D layer of the proposed architecture is intra-connected through quantum neurons with intra-connection strengths set to  $\frac{\pi}{2}$  (Quantum 1 logic). The basic processing unit of each volumetric layer of the 3D-QNet architecture is a 26-connected neighborhood-based voxel-wise orientation of each candidate neuron as illustrated in Figure 1.

The relevant details about the principle of operation of the proposed 3D-QNet are provided in the following subsections using a self-supervised Tensor learning model in quantum formalism.

### A. Quantum-Inspired Self-supervised Tensor Network Model

In the suggested 3D-QNet architecture, the high dimensional weight vector  $\Psi$  is represented using tensor to optimize the network operations and to facilitate the extraction of significant semantic feature information in the quantum-inspired supervised model. The internal kernels associated with the network operate in parallel, thereby accelerating convergence of the 3D-QNet. The input quantum neurons containing the pixel intensity are expressed as *qubits* and the inter-connection weights are modified using quantum rotation gates. The classical intensity of any  $i^{th}$  normalized gray-scale image pixel of MR volume (denoted as  $\alpha_i \in [0, 1]$ ) is transformed into quantum state using a mapping function  $\phi(\alpha_i)$  as follows.

$$\phi(\alpha_i) = \left[ \cos\left(\frac{\pi}{2}\alpha_i\right) \sin\left(\frac{\pi}{2}\alpha_i\right) \right] \forall i = 1, \dots, M, j = 1, \dots, N \quad (7)$$

The angle of rotation is measured using the relative difference of fuzzy intensity of the candidate pixel and the neighborhood pixels in quantum formalism. This relative measure helps to segment the foreground and background regions of an image.

$$\omega_{i,j} = 1 - (\alpha_i - \alpha_{i,j}); j \in \{1, 2, 3, \dots, 26\} \quad (8)$$

Hence,  $\omega_{i,j}$  is designated as the angle of rotation and measured as relative intensity difference between the candidate pixel ( $\alpha_i$ ) and one of its neighborhood pixels  $\alpha_{i,j}$ . The strength of inter-connection between neuron  $j$  (neighborhood of the candidate neuron  $i$ ) of a layer to the corresponding candidate neuron of the adjacent layer is mapped using  $\varphi$ . The classical interconnection weight  $[0, 1]$  is transformed into quantum formalism as

$$|\varphi(\omega_i)\rangle = \left[ \cos\left(\frac{\pi}{2}\omega_i\right) \sin\left(\frac{\pi}{2}\omega_i\right) \right] \quad (9)$$

In this proposed tensor network model, the 3D-QNet layer is decomposed as voxel (core tensor) using Tucker Tensor decomposition [34] to reduce the input dimensions and the interconnection weight as factor matrices. Let us consider tensor  $\mathcal{V}, \Psi \in \mathbb{R}^{m \times n \times p}$ , where  $\mathcal{V}$  is the voxel-wise input of 3D MR images and the corresponding inter-connection 3D weight matrix as evaluated in Eq. 9, respectively ( $m, n, p$  denote the row, column and slice number and  $\mathcal{V}, \Psi$  are third order tensors

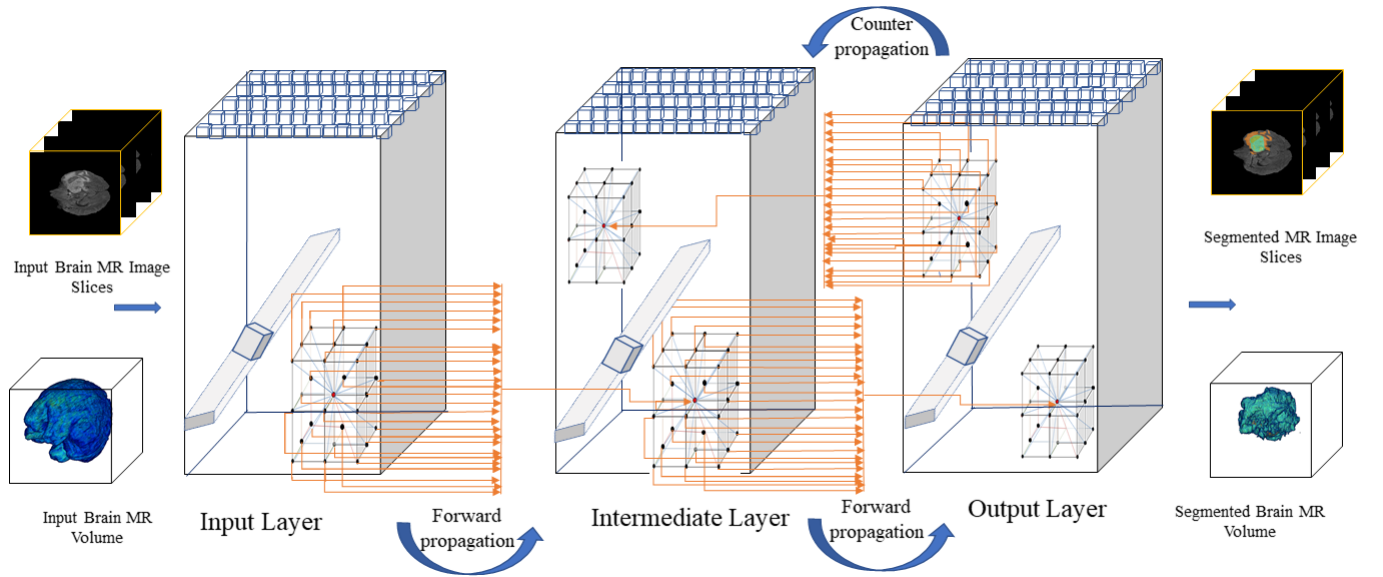


Fig. 1: 3D Quantum-inspired Self-supervised Tensor Neural Network (3D-QNet) architecture (Only three Inter-layer connection is illustrated for better visibility).

( $1 \leq m \leq M, 1 \leq n \leq N, 1 \leq p \leq P$ )). According to Tucker Tensor decomposition [34],

$$\mathcal{X} = \mathcal{V} \times_1 \Psi^1 \times_2 \Psi^2 \times_3 \Psi^3 \quad (10)$$

where,  $\mathcal{X} \in \mathbb{R}^{m \times n \times p}$  the tensor outcome,  $\Psi^n$  is the weight matrix in terms of  $n$  factor matrix and  $\times_n$  is the  $mod - n$  product of tensor with a matrix. Each layer of the proposed 3D-QNet architecture is transformed to lower dimensional tensors. Such types of  $M \times N \times P$  tensors in terms of voxel are formed for each layer in the underlying network architecture. Each volumetric layer of the 3D-QNet architecture forms  $M \times N \times P$  volumetric patches (voxels) of size  $3 \times 3 \times 3$  corresponding to the candidate pixels as

$$\mathbf{v} = vox(\mathcal{V}) \quad (11)$$

Here,  $\mathcal{V}$  comprises all 3D-patches (voxels),  $\mathbf{v} \in \mathbb{R}^{m \times n \times p}$  of a network layer in the proposed 3D-QNet architecture. The spatial features in terms of neighborhood pixels of every seed pixel at the network layer are extracted and propagated to next subsequent layers as input guided by a Quantum-inspired voxel-wise multi-level Sigmoidal (Vox-QSig) activation function,  $\sigma_{3D-QNet}$  as follows.

$$\mathbf{y}^l = \sigma_{3D-QNet}(\mathbf{v}^{l-1} * \boldsymbol{\varphi}^l(\omega)) \quad (12)$$

where,  $\mathbf{v}^{l-1} \in \mathbb{R}^{M \times N \times P}$ ,  $\boldsymbol{\varphi}^l(\omega) \in \mathbb{R}^{M \times N \times P \times K}$  at the network layer  $l = 2, 3$ ,  $\mathbf{y}^l \in \mathbb{R}^K$  and  $*$  is the inner product operator. The fuzzy context-sensitive activation (designated as  $\chi_i$ ) for semantic segmentation in quantum formalism is defined as follows.

$$|\chi_i\rangle = \begin{bmatrix} \cos \vartheta_i \\ \sin \vartheta_i \end{bmatrix} \quad (13)$$

where, the angle of rotation,  $\vartheta_i$  is evaluated using the summation of the intensities of third order 26-connected neighborhood pixels (denoted as  $\alpha_{i,j}, j = 1, 2, \dots, K$ ) of a candidate

pixel  $i$  (neuron) in quantum formalism using the following equation.

$$\vartheta_i = 2\pi \times \left( \sum_{j=1}^K \alpha_{i,j} \right) \quad (14)$$

Quantum fuzzy context-sensitive thresholding determines the bi-directional propagation of quantum information between the layers of the 3D-QNet architecture by means of self-organization of the inter-linked weight matrices. Reduction of feature dimensions using tensor decomposition followed by voxel-wise information processing of the proposed 3D-QNet architecture is inspired by the basic quantum neural network input-output model [6] as follows:

$$|\phi^l(\alpha_i^d)\rangle = \sigma_{3D-QNet} \left( \sum_j^{m \times n \times p} f^{l-1}(\alpha_i^d) \langle \varphi_j^l | \chi_i^{l,d} \rangle \right) \quad (15)$$

where,  $|\phi^l(\alpha_i^d)\rangle$  denotes the intermediate output of the  $i^{th}$  seed quantum neuron at the 3D network layer in the  $l^{th}$  sample with depth (slice#)  $d = 1, 2, \dots, P$ .  $\sigma_{3D-QNet}$  is the Quantum-inspired voxel-wise multi-level Sigmoidal activation (QSig) function with activation as  $|\chi_i^{l,d}\rangle$  described in the following subsection IV-B. The output  $|\phi^l(\alpha_i^d)\rangle$  can be written as

$$|\phi^l(\alpha_i^d)\rangle = f\left(\frac{\pi}{2} \delta_i^{l,d} - \arg\left\{ \sum_j^{m \times n \times p} f^l(\omega_{j,i}^{l,d}) f^{l-1}(\alpha_i^d) - f^l(\chi_i^d) \right\}\right) = \sigma_{3D-QNet} \left( \sum_j^{m \times n \times p} f^{l-1}(\alpha_i^d) (\cos(\omega_{j,i}^{l,d} - \vartheta_i^l) + \gamma \sin(\omega_{j,i}^{l,d} - \vartheta_i^l)) \right) \quad (16)$$

Here, the designated rotation angles associated with the inter-connection weights between input neuron  $j$  to output neuron  $i$  are represented by  $\omega_{j,i}^{l,d}$  and  $\delta_i^{l,d}$  is the phase transfer parameter. The true classical output state ( $|1\rangle$ ) from the  $i^{th}$  quantum



neuron is obtained considering the imaginary section ( $\sin$ ) of the above expression where,  $\gamma$  is an imaginary unit. Assume that the inter-connection weights between the input and hidden layer of the 3D-QNet architecture are denoted by  $|\Psi_{k,j}^{l,d}\rangle$  and for the hidden layer to output layer are denoted by  $|\Psi_{j,i}^{l,d}\rangle$  in the  $l^{th}$  sample sets. The activation at the hidden and output layers are designated using  $|\chi_j^{l,d}\rangle$  and  $|\chi_i^{l,d}\rangle$ , respectively. Considering any quantum seed neuron  $k$  from the sample of input neurons at the input layer, the corresponding seed neuron at the hidden layer be  $j$  and the output seed neuron be  $i$ , the response at the  $i^{th}$  neuron with depth  $d$  in the  $l^{th}$  sample sets is expressed as

$$\begin{aligned} |\phi^l(\alpha_i^d)\rangle &= \sigma_{3D-QNet}(\sum_j^{m \times n \times p} f(\frac{\pi}{2} y_j^{l,d}) \langle \varphi_{ji}^{l,d} | \vartheta_j^{l,d} \rangle) \\ &= \sigma_{3D-QNet}(\sum_j^{m \times n \times p} f(\frac{\pi}{2} \times \sigma_{3D-QNet}(\sum_k^{m \times n \times p} f(\frac{\pi}{2} y_j^{l,d}) \\ &\quad \langle \varphi_{kj}^{l,d} | \vartheta_k^{l,d} \rangle) \langle \varphi_{ji}^{l,d} | \vartheta_j^{l,d} \rangle)) \end{aligned} \quad (17)$$

i.e.,

$$\begin{aligned} |\phi^l(\alpha_i^d)\rangle &= \sigma_{3D-QNet}(\sum_j^{m \times n \times p} \\ &\quad f(\frac{\pi}{2} \times \sigma_{3D-QNet}(\sum_k^{m \times n \times p} f(\frac{\pi}{2} y_j^{l,d}) \cos(\omega_{k,j}^{l,d} - \vartheta_j^{l,d}) \\ &\quad \cos(\omega_{j,i}^{l,d} - \vartheta_i^{l,d}) + \gamma \sin(\omega_{k,j}^{l,d} - \vartheta_j^{l,d}) \sin(\omega_{j,i}^{l,d} - \vartheta_i^{l,d}))) \end{aligned} \quad (18)$$

### B. Quantum-inspired Voxel-wise multi-level Sigmoidal (Vox-QSig) activation function

In this 3D-QNet architecture, a modified version of the QMSig [4] activation function is suggested and referred to as Quantum-inspired voxel-wise multi-level Sigmoidal (Vox-QSig) activation function for voxel-wise processing of 26-connected spatially oriented neighborhood-based pixels. The Vox-QSig activation function,  $\sigma_{3D-QNet}$  with slope  $\lambda$  and activation  $v$ , is defined as

$$\sigma_{3D-QNet}(x) = \frac{1}{\beta_\tau + e^{-\lambda(x-v)}}, 0 \leq \beta_\tau \leq \frac{\pi}{2} \quad (19)$$

where,  $\beta_\tau$  describes the multi-level class responses exhibited by the 26-connected third order neighborhood pixels expressed as

$$\beta_\tau = \frac{\chi_N}{\rho_\tau - \rho_{\tau-1}} \quad (20)$$

where,  $\rho_\tau$  and  $\rho_{\tau-1}$  are the  $\tau^{th}$  and  $(\tau-1)^{th}$  class outcomes, respectively and the contribution of the 26-connected neighborhood gray-level pixels is  $\chi_N$ . The generalized form of Vox-QSig is obtained by leveraging the activation function hyper-parameters employed in Equation 19 as

$$\sigma_{Vox-QSig}(x; \beta_\tau, \rho_\tau) = \sum_{\tau=1}^L \frac{1}{\beta_\tau + e^{-\lambda(x-(\tau-1)\rho_{\tau-1}-v)}} \quad (21)$$

where,  $L$  corresponds to the number of class levels. The multi-class responses for various hyper-parameters employed in Vox-QSig activation functions are provided in Figure 2. Brain MR volumes exhibit heterogeneous responses over

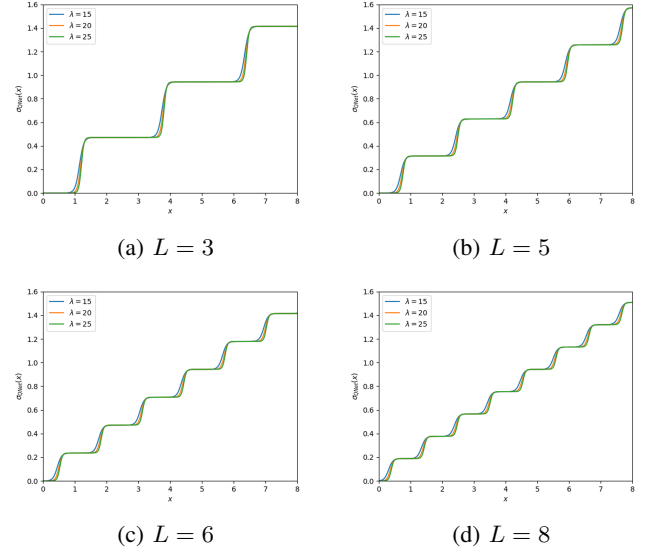


Fig. 2: Multi-level class response of Vox-QSig activation function for  $\lambda = 15, 20, 25$ .

the local intensities in the 26-connected neighborhood regions, owing to the wide variations of gray-levels. Inspired by the authors' previous works [4], [22], [35], [36], the proposed Vox-QSig activation function employs four different adaptive thresholding schemes suitable for efficient gray-scale segmentation in the 3D-QNet architecture.

- (1) Activation guided by  $\beta$ -distribution of the intensity of 26-connected neighborhood voxels ( $v_\beta$ ).
- (2) Activation guided by 26-connected voxels based on Skewness ( $v_\chi$ ).
- (3) Activation guided by 26-connected heterogeneous voxel intensities ( $v_\epsilon$ ).
- (4) Activation guided by 26-connected fuzzy voxel cardinality estimates ( $v_\kappa$ ).

In addition, to investigate a number of optimal thresholds  $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_{C_l-1}\}$  in multi-class image thresholding, Otsu's method [37] is explored. The optimal thresholds maximize the class variance as follows [37].

$$\mathcal{O} = h_n\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_{C_l-1}\} \theta_i(\mu_i - \omega) \quad (22)$$

where  $C_l$  represents the number of defined classes in  $\mathcal{C} = \{C_1, C_2, \dots, C_{C_l}\}$  and

$$\theta_i = \sum_{i \in C_l} p_i, \mu_i = \sum_{i \in C_l} ip_i / \theta_i \quad (23)$$

where, the  $i^{th}$  pixel is defined as  $p_i$ . The probability of class  $C_i$  is represented as  $\mu_i$  and its mean value is given by  $\mu_i$ .  $\omega$  is known to be the mean of class  $\mathcal{C}$ .

### C. Adjustment of Inter-connection Weights of 3D-QNet and Loss Function

Each inter-connection link of the 3D weight matrix of size  $3 \times 3 \times 3$  for each candidate pixel of brain MR volume and its corresponding activation is updated using quantum rotation gates thereby enabling faster convergence of the proposed 3D-QNet architecture. The inter-connection weight,  $\varphi^{l,d}$  and its activation,  $\chi^{l,d}$  are updated as follows.

$$|\varphi^{l+1,d}\rangle = \begin{pmatrix} \cos \Delta\omega^{l+1,d} & -\sin \Delta\omega^{l+1,d} \\ \sin \Delta\omega^{l+1,d} & \cos \Delta\omega^{l+1,d} \end{pmatrix} |\varphi^{l,d}\rangle \quad (24)$$

$$|\chi^{l+1,d}\rangle = \begin{pmatrix} \cos \Delta\vartheta^{l+1,d} & -\sin \Delta\vartheta^{l+1,d} \\ \sin \Delta\vartheta^{l+1,d} & \cos \Delta\vartheta^{l+1,d} \end{pmatrix} |\chi^{l,d}\rangle \quad (25)$$

where,

$$\omega^{l+1,d} = \omega^{l,d} + \Delta\omega^{l,d} \quad (26)$$

and

$$\vartheta^{l+1,d} = \vartheta^{l,d} + \Delta\vartheta^{l,d} \quad (27)$$

Equations 26 and 27 refer to updating the angles of rotation and activation, respectively. The error or loss function,  $\zeta(\omega^{l,d}, \vartheta^{l,d})$  in the suggested 3D-QNet is evaluated in terms of Root Mean Square Error (RMSE) of the 3D-weight matrices at depth  $d$  (or the slice  $\#d$ ) in the  $l^{th}$  epoch and is defined on the phase angles  $\omega^{l,d}, \vartheta^{l,d}$  as

$$\zeta(\omega^{l,d}, \vartheta^{l,d}) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{26} \left[ \varphi_{ij}(\omega_{ij}^{l+1,d}, \vartheta_i^{l+1,d}) - \varphi_{ij}(\omega_{ij}^{l,d}, \vartheta_i^{l,d}) \right]^2 \quad (28)$$

The convergence analysis of the proposed 3D-QNet is illustrated in Appendix Section A.

## V. RESULTS AND DISCUSSION

### A. Data Set

The proposed 3D-QNet is validated extensively using the BRATS 2019 data-sets [38]. The BRATS 2019 dataset is composed of 315 (239 HGG and 76 LGG) 3D MRI volumes. Each MRI volume comprises 155 slices of resolution  $240 \times 240$  with the ground truth segmented labels and includes four different modalities of 3D MR images viz.  $T1$ ,  $T1$  with Contrast-Enhanced ( $T1 - CE$ ),  $T2$  and  $FLAIR$ . The segmented labels are annotated with three distinct tumour sub-regions, viz. tumor core (TC), tumor enhancing (TE), necrosis and non-enhancing core region. These three annotations form a complete tumor (WT). The BRATS 2019 data set is divided into 8 : 2 ratio for training (252) and testing (63) due to GPU limitations.

### B. Experimental Setup

Intensive experiments have been carried out using 3D-QNet on 3D brain MR volumes collected from the BRATS 2019 data set of size  $240 \times 240$  using a Nvidia RTX2070 GPU with MATLAB 2020a and Python3.6 (Pytorch). The proposed 3D-QNet is implemented with the multi-level gray-scale images using distinct multi-class levels  $L = 4, 6$ , and 8 characterized by the Vox-QSig activation function. The steepness  $\lambda$  is

varied in the range 0.23 to 0.24 with step size 0.001. It has been observed that in majority of cases,  $\lambda = 0.232$  yields optimal performance. Moreover, the Vox-QSig is guided by four distinct activation schemes ( $v_\beta, v_\xi, v_\zeta, v_\kappa$ ) [4], [22], [35]. Highly representative volumetric segmentation is followed by  $k$ -means algorithm [39] for false-positive reduction in brain tumor detection and to refine the segmentation accuracy and dice score. The lesion or complete brain tumor detection mask is binarized using a threshold of 0.5. Experiments have also been performed using the 3D-UNet [13] architecture, Deep Voxel-wise Residual Network (VoxResNet) [15], Dense-Res-Inception Net (DRINet) [16], and 3D-ESPNet [17] on the BRATS 2019 data set [38]. We have trained 3D-UNet [13] and VoxResNet [15] rigorously using the Stochastic Gradient Descent (SGD) algorithm on *Caffe* library<sup>1</sup>. The 3D-ESPNet is implemented using *Pytorch* from the code available in Github<sup>2</sup> with 100 epochs using *adam* optimizer with a learning rate of 0.0001. The DRINet is implemented using *adam* optimizer with learning rate of 0.001 and kernel size of  $3 \times 3$ . The segmented output images resemble in size with the dimensions of the binary mask and the outcome 1 is considered as tumor region and 0 as background in detecting complete tumor. The pixel by pixel comparison with the manually segmented regions of interest or lesion mask allows evaluating the dice similarity (DS) [9], which is considered as a standard evaluation procedure in automatic medical image segmentation. The evaluation process involves the manually segmented lesion mask as ground truth, and each 2D pixel is predicted as either True Positive ( $T_{RP}$ ) or True Negative ( $T_{RN}$ ) or False Positive ( $F_{RP}$ ) or False Negative ( $F_{LN}$ ). The empirical goodness measures [Positive Predictive Value ( $PV$ ), Sensitivity ( $SS$ ), Accuracy ( $AC$ ) and Dice Similarity ( $DS$ ) [9]] are assessed to evaluate the results.

### C. Experimental Results

Extensive experiments have been performed in the current setup, and experimental outcomes are reported with the demonstration of numerical and statistical analyses using the proposed 3D-QNet, 3D-UNet [13], VoxResNet [15], DRINet [16], and 3D-ESPNet [17]. It is evident from the experimental data reported in Table I that the proposed 3D-QNet performs optimally for complete brain tumor segmentation of four different modalities of MR volumes (viz.  $T1$ ,  $T1 - CE$ ,  $FLAIR$ , and  $T2$ ) using the activation guided by 26-connected heterogeneous voxel intensities ( $v_\xi$ ) with  $L = 8$  in comparison with other thresholding schemes under the four evaluation parameters ( $AC, DS, PV, SS$ ) [9]. The 3D-QNet segmented brain MR slices collected from two different volumes *BRATS19-CBICA-AAG* and *BRATS19-CBICA-AAB* using class level  $L = 8$  with activation schemes,  $v_\xi$  are shown in Figure 3 and 4, respectively. The human expert annotated ground truth slices for all the four different modalities are illustrated in Figure 5. It has been observed from the segmented MR slices that our 3D-QNet is suitable in segmenting the correct position and size of the complete tumor while

<sup>1</sup><https://doi.org/10.1145/2647868.2654889>

<sup>2</sup><https://github.com/sacmehta/3D-ESPNet>

compared with the ground truth segmentation. However, it is not efficient in mapping the sharp contour of the core and enhanced tumor sub-regions outlined in the annotated slices. Table II presents the quantitative results reported using the proposed 3D-QNet, 3D-UNet [13], VoxResNet [15], DRINet [16], and 3D-ESPNet [17] on evaluating the average accuracy ( $AC$ ), dice similarity score ( $DS$ ), positive prediction value ( $PV$ ), and sensitivity ( $SS$ ) [9]. It has been observed from the 3D-QNet segmented brain MR slices and the results reported in Table II, that optimal segmentation is achieved for *FLAIR* and reported an average of 0.821 dice score ( $DS$ ) for *FLAIR*. The proposed 3D-QNet marginally outperforms the convolutional based architectures (3D-UNet [13], VoxResNet [15], DRINet [16], and 3D-ESPNet [17]) in predicting complete brain tumor detection. However, it may be noted that our 3D-QNet does not intend to predict the core, enhanced tumor and necrosis sub regions owing to lack of optimization of the parameters in the suggested 3D-QNet. The box plots are also demonstrated in Figure 6 citing the outcome reported in Tables II. Moreover, to show the effectiveness of our proposed 3D-QNet over 3D-UNet [13], VoxResNet [15], DRINet [16], and 3D-ESPNet [17], we have conducted one-sided two-sample Kolmogorov-Smirnov (KS) [40] test with significance level  $\alpha = 0.05$ . It is interesting to note that despite being characterized by a fully self-supervised quantum learning, the 3D-QNet has shown similar accuracy ( $AC$ ) and dice similarity ( $DS$ ) in comparison to 3D-UNet [13], VoxResNet [15], DRINet [16], and 3D-ESPNet [17]. Hence, the performance of the 3D-QNet model on the BRATS 2019 data set is statistically significant and offers a promising solution to self-supervised deep learning for 3D-medical image segmentation.

## VI. CONCLUSION

A 3D Quantum-inspired Self-supervised Tensor Neural Network (3D-QNet) architecture characterized by 26-connected voxel-wise processing for fully automated semantic segmentation of Brain MR volume is presented in this work. Intensive validation using the BRATS 2019 data set shows the efficacy of the proposed self-supervised 3D-QNet to promote automatic semantic segmentation of Brain MR volumes in real-time with minimum human intervention which is still considered as an uphill task in the field of medical image segmentation. The incorporation of quantum-inspired computing and tensor-based learning in the suggested network model aims to provide faster convergence of the 3D-QNet, thereby enabling accurate segmentation results. Despite being a 3D self-supervised network model, 3D-QNet achieved similar dice similarity score on complete tumor detection as deeply supervised 3D-UNet, VoxResNet, DRINet and ESPNet, thus promoting self-supervised network learning for volumetric segmentation of medical images. In principle, the proposed 3D-QNet is a general self-supervised network architecture and to be extended in many other 3D medical image segmentation avenues, where the segmented annotations are limited. However, the 3D-QNet fails to yield optimal outcome for multi-level segmentation on the BRATS 2019 data sets. Authors are currently engaged in extending the 3D-QNet by up-scaling the intermediate

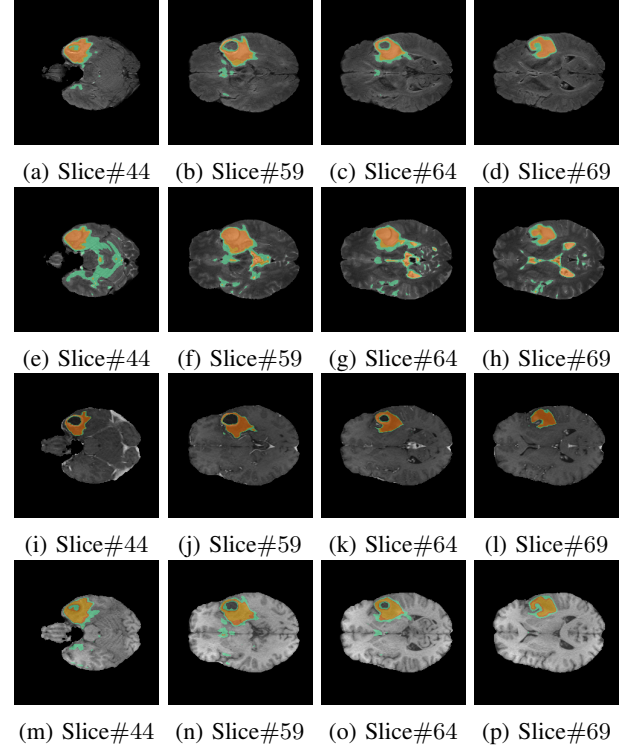


Fig. 3: 3D-QNet segmented Brain MR volume (a – d) BraTS19-CBICA-AAG-1-flair, (e–h) BraTS19-CBICA-AAG-1-t2, (i – l) BraTS19-CBICA-AAG-1-t1ce, (m – p) BraTS19-CBICA-AAG-1-t1 from the BRATS 2019 data set [38]

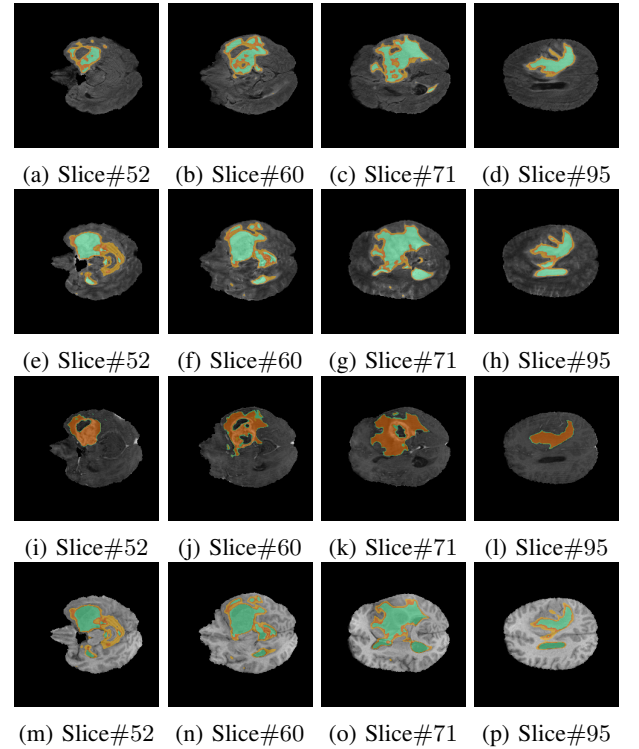


Fig. 4: 3D-QNet segmented Brain MR volume (a – d) BraTS19-CBICA-AAB-1-flair, (e–h) BraTS19-CBICA-AAB-1-t2, (i – l) BraTS19-CBICA-AAB-1-t1ce, (m – p) BraTS19-CBICA-AAB-1-t1 from the BRATS 2019 data set [38]

TABLE I: Results obtained using proposed 3D-QNet for complete tumor detection (CT) on BraTS19-CBICA-AAG-1-flair-slice#69

Level	Modality	AC $\overline{T}_{RP+T_{RN}}$				DS $\overline{2T}_{RP}$				PV $\overline{T}_{RP}$				SS $\overline{T}_{RP}$			
		$\overline{T}_{RP+T_{RN}+F_{LP}+F_{LN}}$				$\overline{2T}_{RP+T_{RN}+F_{LP}+F_{LN}}$				$\overline{T}_{RP+T_{RN}+F_{LP}+F_{LN}}$				$\overline{T}_{RP+T_{RN}+F_{LP}+F_{LN}}$			
		$v_\beta$	$v_\chi$	$v_\xi$	$v_\kappa$	$v_\beta$	$v_\chi$	$v_\xi$	$v_\kappa$	$v_\beta$	$v_\chi$	$v_\xi$	$v_\kappa$	$v_\beta$	$v_\chi$	$v_\xi$	$v_\kappa$
$L = 4$	$T1$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	0.62	<b>0.80</b>	<b>0.80</b>	0.79	0.44	<b>0.66</b>	<b>0.66</b>	0.65	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
	$T1 - CE$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	0.79	<b>0.80</b>	<b>0.79</b>	<b>0.80</b>	<b>0.66</b>	<b>0.66</b>	<b>0.66</b>	<b>0.66</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
	$FLAIR$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	0.62	<b>0.80</b>	<b>0.80</b>	0.79	0.44	<b>0.66</b>	<b>0.66</b>	0.65	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
	$T2$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	0.62	<b>0.80</b>	<b>0.80</b>	0.79	0.44	<b>0.66</b>	<b>0.66</b>	0.65	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
$L = 6$	$T1$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	0.62	<b>0.80</b>	<b>0.80</b>	0.79	0.44	<b>0.66</b>	<b>0.66</b>	0.65	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
	$T1 - CE$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	0.62	<b>0.80</b>	<b>0.80</b>	0.79	0.44	<b>0.66</b>	<b>0.66</b>	0.65	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
	$FLAIR$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	0.62	<b>0.80</b>	<b>0.80</b>	0.79	0.44	<b>0.66</b>	<b>0.66</b>	0.65	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
	$T2$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	0.62	<b>0.80</b>	<b>0.80</b>	0.79	0.44	<b>0.66</b>	<b>0.66</b>	0.65	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
$L = 8$	$T1$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.82</b>	<b>0.82</b>	<b>0.82</b>	<b>0.81</b>	<b>0.70</b>	0.69	0.69	0.68	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
	$T1 - CE$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.81</b>	<b>0.81</b>	<b>0.81</b>	<b>0.81</b>	<b>0.68</b>	<b>0.68</b>	<b>0.68</b>	<b>0.68</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>
	$FLAIR$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.84</b>	<b>0.84</b>	<b>0.84</b>	<b>0.84</b>	<b>0.73</b>	<b>0.73</b>	<b>0.73</b>	0.72	<b>0.99</b>	<b>0.99</b>	0.98	0.98
	$T2$	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.82</b>	<b>0.82</b>	<b>0.82</b>	<b>0.82</b>	0.69	0.69	<b>0.70</b>	<b>0.70</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>	<b>0.99</b>

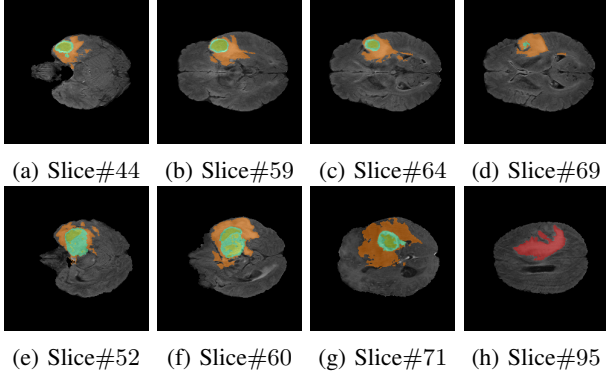


Fig. 5: Annotated Brain MR volume (a – d) BraTS19-CBICA-AAG-1-seg, (e – h) BraTS19-CBICA-AAB-1-seg from the BRATS 2019 data set [38]

volumetric features in the network and optimizing its hyper-parameters to yield optimal segmentation outcome.

## APPENDIX

### A. Convergence Analysis of 3D-QNet

Let us assume the optimal phase angles at depth  $d$  for the weighted matrix and the activation are denoted as  $\omega^d$  and  $\vartheta^d$ , respectively. Now, consider

$$\mathcal{W}^{l,d} = \omega^{l,d} - \overline{\omega^d} \quad (29)$$

$$\mathcal{V}^{l,d} = \vartheta^{l,d} - \overline{\vartheta^d} \quad (30)$$

and

$$\mathcal{D}^{l,d} = \omega^{l+1,d} - \omega^{l,d} = \mathcal{W}^{l+1,d} - \mathcal{W}^{l,d} \quad (31)$$

$$\mathcal{P}^{l,d} = \vartheta^{l+1,d} - \vartheta^{l,d} = \mathcal{V}^{l+1,d} - \mathcal{V}^{l,d} \quad (32)$$

The loss function  $\zeta(\omega^{l,d}, \vartheta^{l,d})$  is differentiable with respect to  $\omega^{l,d}$  and  $\vartheta^{l,d}$  as

$$\frac{\partial \zeta(\omega^{l,d}, \vartheta^{l,d})}{\partial \omega_{ij}^{l,d}} = \frac{2}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} \sum_{j=1}^{26} \Delta \varphi_{ij}^{l,d}(\omega_{ij}^{l,d}, \vartheta_{ij}^{l,d}) \left[ \frac{\partial \varphi_{ij}^{l+1,d}(\omega_{ij}^{l+1,d}, \vartheta_{ij}^{l+1,d})}{\partial \omega_{ij}^{l+1,d}} - \frac{\partial \varphi_{ij}^{l,d}(\omega_{ij}^{l,d}, \vartheta_{ij}^{l,d})}{\partial \omega_{ij}^{l,d}} \right] \quad (33)$$

$$\frac{\partial \zeta(\omega^{l,d}, \vartheta^{l,d})}{\partial \vartheta_{ij}^{l,d}} = \frac{2}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} \sum_{j=1}^{26} \Delta \varphi_{ij}^{l,d}(\omega_{ij}^{l,d}, \vartheta_{ij}^{l,d}) \left[ \frac{\partial \varphi_{ij}^{l+1,d}(\omega_{ij}^{l+1,d}, \vartheta_{ij}^{l+1,d})}{\partial \vartheta_{ij}^{l+1,d}} - \frac{\partial \varphi_{ij}^{l,d}(\omega_{ij}^{l,d}, \vartheta_{ij}^{l,d})}{\partial \vartheta_{ij}^{l,d}} \right] \quad (34)$$

TABLE II: Comparative analysis of proposed 3D-QNet with 3D-UNet [13], VoxResNet [15], DRINet [16], and 3D-ESPNet [17] [The bold values reflect evaluation metrics with significance level  $\alpha = 0.05$  conducted with one sided two sample KS test [40]]

Methods	Modality	AC	DS	PV	SS
3D-UNet [13]	$T1$	0.990	0.811	0.736	0.941
	$T1 - CE$	0.990	0.807	0.732	0.938
	$FLAIR$	<b>0.992</b>	<b>0.823</b>	0.737	0.943
	$T2$	0.989	0.812	0.735	0.944
VoxResNet [15]	$T1$	0.990	0.810	0.737	0.937
	$T1 - CE$	0.989	0.813	0.732	0.943
	$FLAIR$	<b>0.991</b>	<b>0.822</b>	<b>0.751</b>	0.942
	$T2$	0.990	0.807	0.729	0.944
DRINet [16]	$T1$	0.989	0.793	0.701	0.958
	$T1 - CE$	0.988	0.800	0.711	0.959
	$FLAIR$	0.989	0.805	0.708	<b>0.969</b>
	$T2$	0.987	0.789	0.700	0.958
3D-ESPNet [17]	$T1$	0.989	0.801	0.709	0.961
	$T1 - CE$	0.989	0.813	0.721	<b>0.966</b>
	$FLAIR$	0.989	0.800	0.715	0.959
	$T2$	0.988	0.802	0.714	0.957
3D-QNet	$T1$	0.989	0.801	0.736	<b>0.965</b>
	$T1 - CE$	0.989	0.811	0.740	0.957
	$FLAIR$	<b>0.991</b>	<b>0.821</b>	<b>0.751</b>	0.957
	$T2$	0.990	0.814	0.736	0.960

where,

$$\Delta \varphi_{ij}^{l,d}(\omega_{ij}^{l,d}, \vartheta_{ij}^{l,d}) = |\varphi_{ij}^{l+1,d}(\omega_{ij}^{l+1,d}, \vartheta_{ij}^{l+1,d}) - \varphi_{ij}^{l,d}(\omega_{ij}^{l,d}, \vartheta_{ij}^{l,d})| \quad (35)$$

The following equations evaluate the change in phase or angles ( $\Delta \omega$  and  $\Delta \alpha$ ) of the rotation gates as

$$\Delta \omega_{ij}^{l,d} = -\rho_{ij}^{l,d} \left\{ \frac{\partial \varphi(\omega^{l,d}, \vartheta^{l,d})}{\partial \omega_{ij}^{l,d}} \varphi(\omega^{l,d}, \vartheta^{l,d}) \right\}^{\frac{1}{\tau}} \quad (36)$$

$$\Delta \vartheta_{ij}^{l,d} = -\kappa_{ij}^{l,d} \left\{ \frac{\partial \varphi(\omega^{l,d}, \vartheta^{l,d})}{\partial \vartheta_{ij}^{l,d}} \varphi(\omega^{l,d}, \vartheta^{l,d}) \right\}^{\frac{1}{\tau}} \quad (37)$$

where,  $\rho_{ij}$  and  $\kappa_{ij}$  refer to the learning rates for the adjustments of weights and activation, respectively and are evaluated as

$$\begin{aligned} \rho_{ij}^{l,d} &= \mathcal{X}_{ij}^{l,d} - \mathcal{X}_{ij}^{l-1,d} \forall j = 1, 2 \dots 8 \\ \kappa_{ij}^{l,d} &= \left( \sum_j \mathcal{X}_{ij}^{l,d} \right) \forall j = 1, 2 \dots 8 \end{aligned} \quad (38)$$

The sequences of  $\{\omega^{l,d}\}$  and  $\{\vartheta^{l,d}\}$  converge super-linearly subject to the following conditions [41].

$$\lim_{l \rightarrow \infty} \frac{\|\omega^{l+1,d} - \overline{\omega^d}\|}{\|\omega^{l,d} - \overline{\omega^d}\|} \leq 1 \quad (39)$$



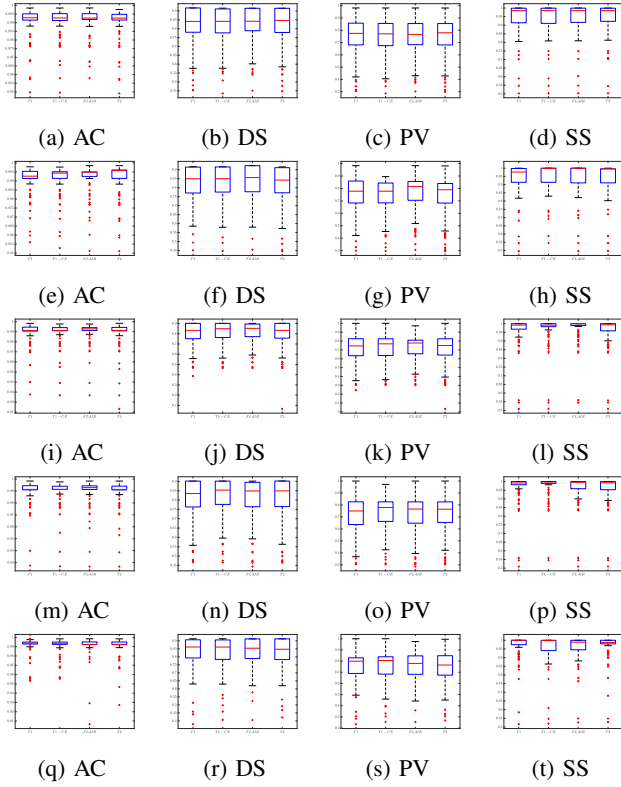


Fig. 6: Illustration of Box plots for the results reported using (a – d) 3D-UNet [13], (e – h) VoxResNet [15], (i – l) DRINet [16], (m – p) 3D-ESPNet [17] and (q – t) 3D-QNet

and

$$\|\mathcal{W}^{l+1,d}\| = O\|\mathcal{D}^{l,d}\| \quad (40)$$

Also,

$$\lim_{l \rightarrow \infty} \frac{\|\vartheta^{l+1,d} - \bar{\vartheta}^d\|}{\|\vartheta^l - \bar{\vartheta}^d\|} \leq 1 \quad (41)$$

and

$$\|\mathcal{V}^{l+1,d}\| = O\|\mathcal{P}^{l,d}\| \quad (42)$$

The convergence of the sequence  $\{\omega^{l,d}\}$  according to  $L$ -Lipschitz continuity is illustrated as [42]

$$\begin{aligned} \zeta(\omega^{l+1,d}) &\leq \zeta(\omega^{l,d}) + \langle \nabla_{\omega} \zeta(\omega^{l,d}), \omega^{l+1,d} - \omega^{l,d} \rangle + \\ &\quad \frac{L}{2} \|\omega^{l+1,d} - \omega^{l,d}\|^2 \\ &= \zeta(\omega^{l,d}) + \langle \nabla_{\omega} \zeta(\omega^{l,d}) - \rho \nabla_{\omega} \zeta(\omega^{l,d}), \omega^{l+1,d} - \omega^{l,d} \rangle + \frac{L}{2} \|\omega^{l+1,d} - \omega^{l,d}\|^2 \\ &= \zeta(\omega^{l,d}) - \rho \|\nabla_{\omega} \zeta(\omega^{l,d})\|^2 + \rho^2 \frac{L}{2} \|\nabla_{\omega} \zeta(\omega^{l,d})\|^2 \\ &= \zeta(\omega^{l,d}) - \rho(1 - \rho \frac{L}{2}) \|\nabla_{\omega} \zeta(\omega^{l,d})\|^2 \\ &\leq \zeta(\omega^{l,d}) - \frac{\rho}{2} \|\nabla_{\omega} \zeta(\omega^{l,d})\|^2 \quad (\text{Assuming, } \rho \in (0, \frac{1}{L})) \\ &\leq \zeta(\bar{\omega}^d) + \langle \nabla_{\omega} \zeta(\omega^{l,d}), \omega^{l,d} - \bar{\omega}^d \rangle - \frac{\rho}{2} \|\nabla_{\omega} \zeta(\omega^{l,d})\|^2, \\ &\quad (\zeta \text{ is convex}) \\ &= \zeta(\bar{\omega}^d) + \langle \nabla_{\omega} \zeta(\omega^{l,d}), \omega^{l,d} - \bar{\omega}^d \rangle - \frac{\rho}{2} \|\nabla_{\omega} \zeta(\omega^{l,d})\|^2 + \\ &\quad \frac{1}{2\rho} (\|\omega^{l,d} - \bar{\omega}^d\|^2 - \|\omega^{l+1,d} - \bar{\omega}^d\|^2) \\ &= \zeta(\bar{\omega}^d) + \frac{1}{2\rho} (\|\omega^{l,d} - \bar{\omega}^d\|^2 - (\|\omega^{l,d}\|^2 - 2\langle \omega^{l,d}, \bar{\omega}^d \rangle + \|\bar{\omega}^d\|^2 - 2\rho \langle \nabla_{\omega} \zeta(\omega^{l,d}), \omega^{l,d} - \bar{\omega}^d \rangle + \rho^2 \|\nabla_{\omega} \zeta(\omega^{l,d})\|^2)) \\ &= \zeta(\bar{\omega}^d) + \frac{1}{2\rho} (\|\omega^{l,d} - \bar{\omega}^d\|^2 - (\|\omega^{l,d} - \rho \nabla_{\omega} \zeta(\omega^{l,d})\|^2 - \\ &\quad 2\langle \omega^{l,d} - \nabla_{\omega} \zeta(\omega^{l,d}), \bar{\omega}^d \rangle + \|\bar{\omega}^d\|^2)) \\ &= \zeta(\bar{\omega}^d) + \frac{1}{2\rho} (\|\omega^{l,d} - \bar{\omega}^d\|^2 - \|\omega^{l+1,d} - \bar{\omega}^d\|^2) \\ \therefore, \zeta(\omega^{l+1,d}) - \zeta(\bar{\omega}^d) &\leq \frac{1}{2\rho} (\|\omega^{l,d} - \bar{\omega}^d\|^2 - \|\omega^{l+1,d} - \bar{\omega}^d\|^2) \end{aligned}$$

Similarly, it can also be shown that

$$\zeta(\vartheta^{l+1,d}) - \zeta(\bar{\vartheta}^d) \leq \frac{1}{2\rho} (\|\vartheta^{l+1,d} - \bar{\vartheta}^d\|^2 - \|\vartheta^{l+1,d} - \bar{\vartheta}^d\|^2) \quad (43)$$

Now, according to Thaler formula

$$\begin{aligned} \zeta(\omega^{l+1,d}, \vartheta^{l+1,d}) - \zeta(\omega^{l,d}, \vartheta^{l,d}) &= \quad (44) \\ \left[ \begin{array}{cc} \Delta \omega_{ij}^{l,d} & \Delta \vartheta_j^{l,d} \end{array} \right] \left[ \begin{array}{c} \frac{\partial \zeta(\omega^{l,d}, \vartheta^{l,d})}{\partial \omega_{ij}^{l,d}} \\ \frac{\partial \zeta(\omega^{l,d}, \vartheta^{l,d})}{\partial \vartheta_j^{l,d}} \end{array} \right] + O \left[ \|\Delta \omega_{ij}^{l,d} \quad \Delta \vartheta_j^{l,d}\| \right] \\ \approx \left[ \left\{ -\rho_{ij}^{l,d} \frac{\partial \zeta(\omega^{l,d}, \vartheta^{l,d})}{\partial \omega_{ij}^{l,d}} \right\}^2 + \left\{ -\kappa_j^{l,d} \frac{\partial \zeta(\omega^{l,d}, \vartheta^{l,d})}{\partial \vartheta_j^{l,d}} \right\}^2 \right] \left\{ \zeta(\omega^{l,d}, \vartheta^{l,d}) \right\}^{\frac{1}{2}} \end{aligned} \quad (45)$$

It is obvious that  $(\zeta(\omega^{l+1,d}, \vartheta^{l+1,d}) - \zeta(\omega^{l,d}, \vartheta^{l,d})) \leq 0$  and the sequences of  $\{\omega^{l,d}\}$  and  $\{\vartheta^{l,d}\}$  are monotonically decreasing as

$$\lim_{l \rightarrow \infty} \zeta(\omega^{l,d}, \vartheta^{l,d}) = (\bar{\omega}^d, \bar{\vartheta}^d) \quad (46)$$

and

$$\lim_{l \rightarrow \infty} \frac{\|\zeta(\omega^{l+1,d}, \vartheta^{l+1,d}) - (\bar{\omega}^d, \bar{\vartheta}^d)\|}{\|\zeta(\omega^{l,d}, \vartheta^{l,d}) - (\bar{\omega}^d, \bar{\vartheta}^d)\|} \leq 1 \quad (47)$$

## REFERENCES

- [1] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017. DOI: <https://doi.org/10.1016/j.media.2017.07.005>
- [2] F. Milletari, N. Navab, and S. A. Ahmadi, “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation” *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571, 2016 DOI: [10.1109/3DV.2016.79](https://doi.org/10.1109/3DV.2016.79).
- [3] O. Ronneberger, P. Fischer, T. Brox, “U-Net: convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*. LNCS, vol. 9351, pp. 234–241. 2016. DOI: [https://doi.org/10.1007/978-3-319-46723-8\\_49](https://doi.org/10.1007/978-3-319-46723-8_49)
- [4] D. Konar, S. Bhattacharyya, T. K. Gandhi and B. K. Panigrahi, “A quantum-inspired self-supervised Network model for automatic segmentation of brain MR images,” *Applied Soft Computing*, vol. 93, 2020, DOI: <https://doi.org/10.1016/j.asoc.2020.106348>.

- [5] D. Konar, S. Bhattacharyya and B. K. Panigrahi, "QIBDS Net: A Quantum-Inspired Bi-Directional Self-supervised Neural Network Architecture for Automatic Brain MR Image Segmentation," *Proc. 8th International Conference on Pattern Recognition and Machine Intelligence (PREMI 2019)*, vol. 11942, pp. 87–95, 2019, DOI: [https://doi.org/10.1007/978-3-030-34872-4\\_64](https://doi.org/10.1007/978-3-030-34872-4_64).
- [6] D. Konar, S. Bhattacharyya, S. Dey, and B. K. Panigrahi, "Opti-QIBDS Net: A Quantum-Inspired Optimized Bi-Directional Self-supervised Neural Network Architecture for Automatic Brain MR Image Segmentation," *Proc. 2019 IEEE Region 10 Conference (TENCON)*, pp. 761–766, 2019, DOI: [10.1109/TENCON.2019.8929585](https://doi.org/10.1109/TENCON.2019.8929585).
- [7] M. Havaei et al., "Brain tumor segmentation with deep neural networks," *Medical Image Analysis*, vol. 35, pp. 18–31, 2017, DOI: <https://doi.org/10.1016/j.media.2016>.
- [8] E. Shelhamer, J. Long, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640–651, 2017.
- [9] S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, 2016, DOI: [10.1109/TMI.2016.2538465](https://doi.org/10.1109/TMI.2016.2538465).
- [10] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. Vries, M. J. N. L. Benders and I. Išgum, "Automatic Segmentation of MR Brain Images With a Convolutional Neural Network," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, 2016, DOI: [10.1109/TMI.2016.2548501](https://doi.org/10.1109/TMI.2016.2548501).
- [11] K. Kamnitsas, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61–78, 2017, DOI: <https://doi.org/10.1016/j.media.2016.10.004>.
- [12] Q. Dou, H. Chen, L. Yu, L. Zhao, J. Qin, D. Wang, V. C. Mok, L. Shi, and P. A. Heng, "Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks," *IEEE transactions on medical imaging*, vol. 35, no. 5, pp. 1182–1195, 2016, DOI: [10.1109/TMI.2016.2528129](https://doi.org/10.1109/TMI.2016.2528129).
- [13] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2016)*, pp. 424–432, 2016, DOI: [https://doi.org/10.1007/978-3-319-46723-8\\_49](https://doi.org/10.1007/978-3-319-46723-8_49).
- [14] A. Brebisson, and G. Montana, "Deep Neural Networks for Anatomical Brain Segmentation," In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 20–28, 2015, DOI: [10.1109/CVPRW.2015.7301312](https://doi.org/10.1109/CVPRW.2015.7301312).
- [15] H. Chen, Q. Dou, L. Yu, J. Qin, and P. A. Heng, 2017. "VoxResNet: Deep voxelwise residual networks for brain segmentation from 3D MR images," *NeuroImage*, vol. 170, pp. 446–455, 2017, DOI: <https://doi.org/10.1016/j.neuroimage.2017.04.041>.
- [16] L. Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara, and D. Rueckert, "DRINet for Medical Image Segmentation," *IEEE Transaction on Medical Imaging*, vol. 37, no. 11, 2018, DOI: [10.1109/TMI.2018.2835303](https://doi.org/10.1109/TMI.2018.2835303).
- [17] N. Nuechterlein and S. Mehta, "3D-ESPNet with Pyramidal Refinement for Volumetric Brain Tumor Image Segmentation," *International MICCAI Brainlesion Workshop (BrainLes 2018)*, pp. 245–253, 2018, DOI: [https://doi.org/10.1007/978-3-030-11726-9\\_22](https://doi.org/10.1007/978-3-030-11726-9_22).
- [18] G. Wang et al., "Interactive Medical Image Segmentation Using Deep Learning With Image-Specific Fine Tuning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 7, 2018, DOI: [10.1109/TMI.2018.2791721](https://doi.org/10.1109/TMI.2018.2791721).
- [19] X. Zhuang, Y. Li, Y. Hu, K. Ma, Y. Yang, and Y. Zheng, "Self-supervised Feature Learning for 3D Medical Images by Playing a Rubik's Cube," *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI 2019)*, pp. 420–428, [https://doi.org/10.1007/978-3-030-32251-9\\_46](https://doi.org/10.1007/978-3-030-32251-9_46).
- [20] A. Ghosh, N. R. Pal, and S. K. Pal, "Self organization for object extraction using a multilayer neural network and fuzziness measures," *IEEE Transactions on Fuzzy Systems*, vol. 1, no. 1, pp. 54–68, 1993, DOI: [10.1109/TFUZZ.1993.390285](https://doi.org/10.1109/TFUZZ.1993.390285).
- [21] S. Bhattacharyya, P. Dutta and U. Maulik, "Binary object extraction using bi-directional self-organizing neural network (BDSOINN) architecture with fuzzy context -sensitive thresholding," *Pattern Anal. Applic.*, vol. 10, pp. 345–360, 2007, DOI: <https://doi.org/10.1007/s10044-007-0072-z>.
- [22] S. Bhattacharyya, P. Dutta and U. Maulik, "A parallel bi-directional self-organizing neural network (PBDSONN) architecture for color image extraction and segmentation," *Neurocomputing*, vol. 86, pp. 1–23, 2012, DOI: <https://doi.org/10.1016/j.neucom.2011.11.025>.
- [23] S. Kak, "On quantum neural computing," *Information Sciences*, vol. 83, pp. 143–160, 1995, DOI: [https://doi.org/10.1016/0020-0255\(94\)00095-s](https://doi.org/10.1016/0020-0255(94)00095-s).
- [24] G. Purushothaman, N. B. Karayiannis, "Quantum neural networks (QNNs): inherently fuzzy feedforward neural networks," *IEEE Transactions on Neural Networks*, vol. 8, no. 3, 1997, DOI: [10.1109/72.572106](https://doi.org/10.1109/72.572106).
- [25] T. C. Lu, G. R. Yu, and J. C. Juang, "Quantum-based algorithm for optimizing artificial neural networks," *IEEE Transaction on Neural Network and Learning Systems*, vol. 24, no. 8, pp. 1266–1278, 2013, DOI: [10.1109/TNNLS.2013.2249089](https://doi.org/10.1109/TNNLS.2013.2249089).
- [26] Nam-H. Nguyen ; E. C. Behrman, A. Moustafa, J. E. Steck, "Benchmarking Neural Networks For Quantum Computations," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–10, 2019, DOI: [10.1109/TNNLS.2019.2933394](https://doi.org/10.1109/TNNLS.2019.2933394).
- [27] E. M. Stoudenmire, and D. J. Schwab, "Supervised Learning with Quantum-Inspired Tensor Networks," <https://www.arxiv-vanity.com/papers/1605.05775/>, 2020.
- [28] P. Li, H. Xiao, F. Shang, X. Tong, X. Li, and M. Cao, "A hybrid quantum-inspired neural networks with sequence inputs," *Neurocomputing*, vol. 117, pp. 81–90, 2013, DOI: <https://doi.org/10.1016/j.neucom.2013.01.029>.
- [29] N. Masuyama, C. K. Loo, M. Seera, and N. Kubota, "Quantum-Inspired Multidirectional Associative Memory With a Self-Convergent Iterative Learning," *IEEE Transaction on Neural Network and Learning Systems*, vol. 29, no. 4, pp. 1058–1068, 2018, DOI: [10.1109/TNNLS.2017.2653114](https://doi.org/10.1109/TNNLS.2017.2653114).
- [30] S. Bhattacharyya, P. Pal and S. Bhowmick, "Binary Image Denoising Using a Quantum Multilayer Self Organizing Neural Network," *Applied Soft Computing*, vol. 24, pp. 717–729, 2014, DOI: <https://doi.org/10.1016/j.asoc.2014.08.027>.
- [31] D. Konar, S. Bhattacharya, B. K. Panigrahi, K. Nakamatsu "A quantum bi-directional self-organizing neural network (QBDSOINN) architecture for binary object extraction from a noisy perspective," *Applied Soft Computing*, vol. 46, pp. 731–752, 2016, DOI: <https://doi.org/10.1016/j.asoc.2015.12.040>.
- [32] D. Konar, S. Bhattacharya, U. Chakraborty, T. K. Gandhi, B. K. Panigrahi, "A quantum parallel bi-directional self-organizing neural network (QPBDSONN) architecture for extraction of pure color objects from noisy background," *Proc. IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2016, pp. 1912–1918, 2016, DOI: [10.1109/ICACCI.2016.7732330](https://doi.org/10.1109/ICACCI.2016.7732330).
- [33] M. A. Nielson and I. L. Chung, *Quantum computation and quantum information*, Cambridge University press, 2002.
- [34] T. G. Kolda, B. W. Bader, "Tensor Decompositions and Applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009, DOI: <https://doi.org/10.1137/0707011X>.
- [35] S. Bhattacharyya, P. Dutta and U. Maulik, "Multilevel image segmentation with adaptive image context based thresholding," *Applied Soft Computing*, vol. 11, no. 1, pp. 946–962, 2011, DOI: <https://doi.org/10.1016/j.asoc.2010.01.015>.
- [36] S. Bhattacharyya, P. Dutta and U. Maulik, "A parallel bi-directional self-organizing neural network (PBDSONN) architecture for color image extraction and segmentation," *Neurocomputing*, vol. 86, pp. 1–23, 2012, DOI: <https://doi.org/10.1016/j.neucom.2011.11.025>.
- [37] N. Otsu, "A threshold selection method from gray level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979, DOI: [10.1109/TSMC.1979.4310076](https://doi.org/10.1109/TSMC.1979.4310076).
- [38] B. H. Menze et al., "The multimodal brain tumor image segmentation benchmark (BRATS)," *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993–2024, 2015, DOI: [10.1109/TMI.2014.2377694](https://doi.org/10.1109/TMI.2014.2377694).
- [39] H. P. Ng, S. H. Ong, K. W. C. Foong, P. S. Goh, and W. L. Nowinski, "Medical Image Segmentation Using K-Means Clustering and Improved Watershed Algorithm," *2006 IEEE Southwest Symposium on Image Analysis and Interpretation*, 2006, DOI: [10.1109/SSIAI.2006.1633722](https://doi.org/10.1109/SSIAI.2006.1633722).
- [40] M. H. Gail and S. B. Green, "Critical values for the one-sided two-sample Kolmogorov–Smirnov statistic," *J. Am. Stat. Assoc.*, vol. 71, pp. 757–760, 1976.
- [41] D. Mu, Z. Guan, and H. Zhang, "Learning Algorithm and Application of Quantum Neural Networks with Quantum Weights," *International Journal of Computer Theory and Engineering*, vol. 5, no. 5, pp. 788–792, 2013, DOI: [10.7763/IJCTE.2013.V5.797](https://doi.org/10.7763/IJCTE.2013.V5.797).
- [42] H. Kim, J. Kang, W. Park, S. Ko, Y. Cho, D. Yu, Y. Song, and J. Choi, "Convergence Analysis of Optimization Algorithms," 2017, [arXiv:1707.01647](https://arxiv.org/abs/1707.01647).