A Novel 2-Phase U-net Algorithm Combined with Optimal Mass Transportation for 3D Brain Tumor Detection and Segmentation

Wen-Wei Lin

National Yang Ming Chiao Tung University

Jia-Wei Lin

National Yang Ming Chiao Tung University

Tsung-Ming Huang (✉ min@ntnu.edu.tw)

National Taiwan Normal University

Tiexiang Li

Southeast University

Mei-Heng Yueh

National Taiwan Normal University

Shing-Tung Yau

Harvard University

Research Article

Keywords: Novel 2-Phase U-net Algorithm , Optimal Mass Transportation, 3D Brain Tumor Detection, Segmentation

Posted Date: January 3rd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1215123/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A Novel 2-Phase U-net Algorithm Combined with Optimal Mass Transportation for 3D Brain Tumor Detection and Segmentation

Wen-Wei Lin^{1,4}, Jia-Wei Lin¹, Tsung-Ming Huang^{2,*}, Tiexiang Li^{3,4,*}, Mei-Heng Yueh², and Shing-Tung Yau⁵

¹Department of Applied Mathematics, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan

²Department of Mathematics, National Taiwan Normal University, Taipei 116, Taiwan

³School of Mathematics, Southeast University, Nanjing 211189, People's Republic of China

⁴Nanjing Center for Applied Mathematics, Nanjing 211135, People's Republic of China

⁵Department of Mathematics, Harvard University, Cambridge, USA

*Corresponding authors

ABSTRACT

Utilizing the optimal mass transportation (OMT) technique to convert an irregular 3D brain image into a cube, a required input format for the U-net algorithm, is a brand new idea for medical imaging research. We develop a cubic volume-measure-preserving OMT (V-OMT) model for the implementation of this conversion. The contrast-enhanced histogram equalization grayscale of fluid attenuated inversion recovery (FLAIR) in a brain image creates the corresponding density function. We then propose an effective two-phase U-net algorithm combined with the V-OMT algorithm for training and validation. First, we use the U-net and V-OMT algorithms to precisely predict the whole tumor (WT) region. Second, we expand this predicted WT region with dilation and create a smooth function by convoluting the step-like function associated with the WT region in the brain image with a $5 \times 5 \times 5$ blur tensor. Then, a new V-OMT algorithm with mesh refinement is constructed to allow the U-net algorithm to effectively train Net1–Net3 models. Finally, we propose ensemble voting postprocessing to validate the final labels of brain images. We randomly choose 1000 and 251 brain samples from the BraTS 2021 training dataset, which contains 1251 samples, for training and validation, respectively. The Dice scores of the WT, tumor core (TC) and enhanced tumor (ET) regions for validation computed by Net1–Net3 were 0.93705, 0.90617 and 0.87470, respectively. A significant improvement in brain tumor detection and segmentation with higher accuracy is achieved.

Due to the rapid development of the architectural designs of convolutional neural networks (CNNs), deep CNNs have become one of the most widely used AI technologies and have undoubtedly provided surprisingly powerful approaches for tasks, such as object detection, feature extraction, image classification and segmentation, language translation, chemical molecular structure prediction, and other significant tasks in medical and computational sciences. Deep CNNs are composed of three key components, namely, input data, computation steps and models. The last two items include the design of the optimization algorithms and model structure weight training. The former is largely fueled by inputting large amounts of data into the CNN for the development of a more capable prediction system. In the past decade, innovations in high-performance computing, such as GPU accelerators, have promoted powerful advancements in deep learning. However, with Moore's law coming to an end, the reduction in the gratis cost of computing and storage may cease in the near future. Although, thus far, expanding the size of trillion-parameter models and adding a large amounts of training data has been successful, and these models have effectively obtained excellent prediction performance. However, because of the limitation of Moore's law, the calculation of these super models will become extremely expensive and inefficient.

In recent years, two benchmark datasets, MSD2018^{1,2} and BraTS 2020^{3–5}, which contain 484 and 369 labeled 3D brain image samples for brain tumor segmentation, respectively, have provided a challenging platform and have attracted enormous attention and interest from researchers in this field. The brain samples were scanned with four modalities, namely, fluid-attenuated inversion recovery (FLAIR), T1-weighted (T1), T1-weighted contrast-enhanced (T1CE), and T2-weighted (T2), by multiparametric magnetic resonance imaging (mpMRI). The challenge is the evaluation of state-of-the-art methods for the task of brain tumor segmentation of whole tumor (WT, labeled as $\{2,1,4\}$), tumor core (TC, labeled as $\{1,4\}$), and enhanced tumor (ET, labeled as $\{4\}$) regions in the human brain. To address this issue, in the early years, random forest algorithms and machine learning techniques were used to perform image classification^{6–8} and segmentation^{6,9–11}. After that, CNN structures with two layers¹² and eight layers¹³ were proposed and made good progress in brain tumor segmentation. Then, a more sophisticated

multiple CNN architecture, called the U-net model, was first developed in¹⁴ and improved in¹⁵ by assembling two full CNNs and U-net. The merits of applying the U-net model to the challenge of MSD 2018^{1,2} were first proposed by¹⁶. Recently, BraTS 2021^{3,4,17} expands a large number of new brain samples in the database and provides 1251 labeled samples for training and 219 unlabeled samples for validation.

For this reason, preprocessing to effectively represent large amounts of input data for CNNs is crucial. For example, taking an irregular 3D physical brain image obtained from MRI, which is generally composed of 1.5 million vertices, by randomly selecting several cubes (e.g., 16 cube filters were used in¹⁸) with seamless coverage to overlay the irregular brain image is a natural way to fit the input format of tensors for a U-net system. An efficient two-stage optimal mass transportation (OMT) algorithm, newly proposed by¹⁹, was designed to first transform an irregular 3D brain image into a unit ball and then into a cube with minimal distortion and transport costs. This strategy can greatly reduce the capacity of input data, so there are more opportunities to expand various types of training data, and then, the existing computing resources can be effectively used to improve the expected prediction accuracy. Thus, we are motivated to consider transforming an irregular brain image into a cube. In practical applications, using a CNN to make segmentation predictions for human brain tumors, we only need to use a suitable cube to represent a brain image without losing many important features and conversion accuracy. In this way, the storage capacity and the computational cost of the computer environment can be greatly economized.

OMT is a very old optimization problem that was raised by Monge in 1781 (see²⁰ for details) to find an optimal solution that minimizes the transport cost and preserves the local mass ratios between two spaces. The existence and uniqueness of a solution to the OMT problem was proven by Kantorovich²⁰ by relaxing the probability measure with a joint probability distribution. The regularity condition for the solution of the OMT problem was first shown by Caffarelli¹¹, and an elegant theoretical survey paper “Optimal Transport: Old and New”, which summarized the achievements of predecessors, was published by Villani²¹. For numerical methods, Brenier¹⁰ proposed an alternative scheme for solving the OMT problem with a quadratic cost function for a special class of convex domains. Based on Brenier’s approach and the variational principle²², Su et al.²³ developed a volume-preserving parameterization from a 3-manifold $\mathcal{M}$ with a spherical boundary to a unit ball $\mathcal{B}^3$. Recently, Yueh et al.²⁴ proposed a novel algorithm to compute a volume-preserving parameterization from $\mathcal{M}$ to $\mathcal{B}^3$ by modifying the denominators of the coefficients of the corresponding Laplacian matrix by imposing the local volume stretch factor at each iteration step and adopted the projected gradient method (PGM) combined with the homotopy technique in²⁵ to find the OMT map between $\mathcal{M}$ and $\mathcal{B}^3$. In addition, a two-stage OMT (2SOMT) procedure from $\mathcal{M}$ to $\mathcal{B}^3$ and from $\mathcal{B}^3$ to a cube was efficiently developed by Lin et al.¹⁹ and applied prior to U-net training and inference in 3D brain tumor segmentation.

In this paper, we study the applicability of mapping an irregular 3D image (i.e., a human brain) to a canonical domain (i.e., a cube or a cuboid), which minimizes the transport cost and preserves the local mass ratios. First, based on the homotopy technique, a direct one-stage OMT approach from a 3-manifold $\mathcal{M}$ with a genus-zero boundary to a cube is developed for 3D U-net training and inference to improve the higher conversion loss of 2SOMT¹⁹ from $\mathcal{M}$ to $\mathcal{B}^3$ and $\mathcal{B}^3$ to a cube. Thus, we can construct a one-to-one correspondence between the input data of irregular images and the associated cubic tensors. Without any the conversion loss between OMT maps, the capacity of the training data of the 3D U-net model is greatly reduced, and it is our belief that 3D U-net training can easily find a local minimum and achieve better performance.

Next, we propose a two-phase U-net with OMT (2P-Unet-OMT) algorithm utilizing the density distribution of brain tumor features and train four related networks to detect tumor regions and segment tumor labels. Given an irregular 3D brain, in Phase I, we first construct the associated density map at each vertex according to the normalized contrast-enhanced histogram equalization (CEHE) grayscale values of the FLAIR modality of a brain image by MRI. Then, we compute OMT maps from brain images to cubes for the training set and train Net0 by the U-net algorithm for the detection of possible tumor regions. In fact, there are no clues at the beginning; the CEHE grayscales of FLAIR, which typically reflect the distribution of WT, should be an effective way to detect tumor regions. Next, we cover these possible tumor regions by 5 voxels with dilation. In Phase II, because $ET \subset TC \subset WT$, we construct a smooth density function by convoluting the step-like function with $\exp(\text{FLAIR})$ on the expanding WT region and 1.0 on the others, with a $5 \times 5 \times 5$ box blur tensor. We remesh the tetrahedron with finer meshes in the higher density region in the brain so that the target tumor region can be enlarged in the cube by OMT and better viewed and learned by U-net. We then train Net1 for WT, Net2 for TC and Net3 for ET by U-net. In practice, for the testing issue, Net0 can help by first detecting a WT region as much as possible. As above, by covering this region with 5 voxels with dilation and creating a similar smooth density function with finer meshes on the raw brain image, we compute the corresponding OMT map and call Net1–Net3 combined with ensemble voting postprocessing to make the final label prediction and image segmentation.

Contribution. The 2P-Unet-OMT procedure transforms an irregular 3D brain image into a cube with density estimates and mesh refinement to fit the input format of the U-net algorithm. Unlike the previous methods, 2P-Unet-OMT minimizes the transport cost and preserves the global features of input data to surpass the other methods. The main contributions of this paper are summarized below.

1. Our proposed 2P-Unet-OMT algorithm transforms an irregular 3D brain image into a cube to satisfy the input format of

U-net while preserving the local mass ratios between two domains and minimizing the transport cost and the distortion. These advantages for 2SOMT were highlighted in¹⁹. However, 2SOMT did not make full use of estimating the distribution of the density function, so U-net could not infer the target object accurately. 2P-Unet-OMT fully grasps the distribution of the associated density function for creating an effective OMT map from an irregular 3D domain to a cube and provides it to U-net for training a high-performance prediction network. 2P-Unet-OMT inherits the advantages of 2SOMT in that it only needs to use a cube to represent an irregular 3D brain image without losing many important features and conversion accuracy. In this way, the computational cost and the computer environment can be greatly economized during U-net training and used for data augmentation, which exactly considers the limitation of the memory capacity. Nevertheless, 2P-Unet-OMT greatly increases the prediction accuracy via a precise estimate of the density function for the tumor distribution.

2. One of the characteristics of the OMT map is to preserve the local mass. With this peculiar feature, in Phase II of 2P-Unet-OMT, we apply mesh refinement on the expanding WT region detected during Phase I. The mesh refinement technique can increase the number of tetrahedrons in a specific region in the brain and enlarge the portion of volume appearing in the target domain; that is, using the U-net algorithm is similar to using a magnifying glass to view and learn how to mark the segmentation labels well. The numerical experiment with the trained Net1–Net3 model combined with ensemble voting shows that the Dice scores of validation for WT, TC and ET can reach 0.93705, 0.90617 and 0.87470, respectively; hence, this approach significantly boosts the accuracy of brain tumor detection and segmentation.
3. Because the OMT approach must convert the labels predicted by U-net to a brain image and to evaluate the Dice score more precisely, we propose a new conversion technique with ensemble voting postprocessing to convert the predicted labels on the cube back to each voxel of the brain by using the multivalues on the cube validated by various models to make a precise evaluation of labels corresponding to voxels in the brain image. From the expressively high validation Dice scores on the BraTS 2021 validation data, it can be said that using a cube for representing an irregular 3D brain image by OMT is indeed an innovative idea and the most streamlined approach for CNN training and prediction.

This paper is organized as follows. In “Discrete OMT Problems and Cubic OMT Maps”, we introduce the discrete OMT problem and the spherical-cubic area-measure-preserving and cubic volume-measure-preserving OMT maps. In “Two-Phase U-net with OMT for Training and Validation”, we propose a two-phase U-net model with OMT maps for training and validation. For the evaluation of high Dice scores, we develop an effective conversion technique to convert the predicted labels on the cube back to the brain image using all related probability information corresponding to each voxel in the brain image. In “Results and Discussions”, we show the improvement in the Dice score obtained by the U-net models in Phase II with mesh refinements on the expanding WT region provided by Phase I and the ensemble voting postprocessing for the label evaluation. Finally, a concluding remark is given in “Conclusions”.

Discrete OMT Problems and Cubic OMT Maps

Let $\mathcal{M}$ be a simplicial 3-complex that describes an irregular 3D brain image with a genus-zero boundary. $\mathcal{M}$ is generally composed of sets of vertices $\mathbb{V}(\mathcal{M})$, edges $\mathbb{E}(\mathcal{M})$, faces $\mathbb{F}(\mathcal{M})$ and tetrahedrons $\mathbb{T}(\mathcal{M})$. A discrete OMT problem is to find a bijective function that maps $\mathcal{M}$ to a canonical simple domain with minimal distortion. The canonical shape could be a ball $\mathcal{B}^3$ or a unit cube $\mathcal{C}^3$. Since a tensor form is necessary for the input of the U-net algorithm, a cube or a cuboid is the target domain for $\mathcal{M}$. In this section, we propose a one-stage OMT approach to map $\mathcal{M}$ to $\mathcal{C}^3$.

Discrete OMT Problem. Let ρ be a density map on $\mathbb{V}(\mathcal{M})$. The piecewise linear density functions of ρ on $\mathbb{F}(\partial\mathcal{M})$ and $\mathbb{T}(\mathcal{M})$ are defined by

$$\rho(\alpha) = \frac{1}{3} \sum_{i=1}^3 \rho(\hat{v}_i), \quad \rho(\tau) = \frac{1}{4} \sum_{i=1}^4 \rho(v_i), \quad (1)$$

respectively, where $\hat{v}_i \in \mathbb{V}(\alpha)$, $\alpha \in \mathbb{F}(\partial\mathcal{M})$, $v_i \in \mathbb{V}(\tau)$, and $\tau \in \mathbb{T}(\mathcal{M})$. Furthermore, we define the local area-/volume-measures (i.e., local mass) by

$$a_\rho(\hat{v}) := \frac{1}{3} \rho(\hat{v}) \sum_{\hat{v} \subset \alpha} |\alpha|, \quad m_\rho(v) := \frac{1}{4} \rho(v) \sum_{v \subset \tau} |\tau|, \quad (2)$$

respectively, where $|\alpha|$ and $|\tau|$ are the area and volume of α and τ , respectively.

Denote

$$\mathbf{G}_\rho = \{g : \partial\mathcal{M} \rightarrow \partial\mathcal{C}^3 \mid \rho(\alpha)|\alpha| = |g(\alpha)|, \forall \alpha \in \mathbb{F}(\partial\mathcal{M})\} \quad (3a)$$

and

$$\mathbf{F}_\rho = \{f : \mathcal{M} \rightarrow \mathcal{C}^3 \mid \rho(\tau)|\tau| = |f(\tau)|, \forall \tau \in \mathbb{T}(\mathcal{M})\} \quad (3b)$$

as the sets of all area-/volume-measure-preserving (i.e., mass-preserving) piecewise linear maps from $\partial\mathcal{M}$ to $\partial\mathcal{C}^3$ and from $\mathcal{M}$ to $\mathcal{C}^3$, respectively, in which the bijective maps between α and $g(\alpha)$, as well as τ and $f(\tau)$, are determined by the barycentric coordinates on α and τ , respectively. For given $g \in \mathbf{G}_\rho$ and $f \in \mathbf{F}_\rho$, we define the transport costs of g and f , respectively, by

$$d_\rho(g) = \sum_{\hat{v} \in \mathbb{V}(\partial\mathcal{M})} \|\hat{v} - g(\hat{v})\|_2^2 a_\rho(\hat{v}), \quad c_\rho(f) = \sum_{v \in \mathbb{V}(\mathcal{M})} \|v - f(v)\|_2^2 m_\rho(v), \quad (4)$$

where $a_\rho(\hat{v})$ and $m_\rho(v)$ are the local area-/volume-measures at $\hat{v} \in \mathbb{V}(\partial\mathcal{M})$ and $v \in \mathbb{V}(\mathcal{M})$, respectively, as in (2). The discrete OMT problems on $\partial\mathcal{M}$ and $\mathcal{M}$ with respect to $\|\cdot\|_2$ are to find a $g_\rho^* \in \mathbf{G}_\rho$ and $f_\rho^* \in \mathbf{F}_\rho$ that solve optimal problems

$$g_\rho^* = \operatorname{argmin}_{g \in \mathbf{G}_\rho} d_\rho(g), \quad f_\rho^* = \operatorname{argmin}_{f \in \mathbf{F}_\rho} c_\rho(f), \quad (5)$$

where $d_\rho(g)$ and $c_\rho(f)$ are given in (4). Without loss of generality, hereafter, each simplicial 3-complex $\mathcal{M}$ is centralized and normalized so that the center of mass is located at the origin and the mass is one.

Cubic Area-Measure-Preserving OMT Maps. Let $\mathcal{M}$ be a simplicial 3-complex with a genus-zero boundary and of mass one with density functions $\rho : \mathbb{T}(\mathcal{M}) \rightarrow \mathbb{R}$ and $\rho : \mathbb{F}(\partial\mathcal{M}) \rightarrow \mathbb{R}$ defined on the tetrahedrons of $\mathcal{M}$ and triangles of $\partial\mathcal{M}$, respectively. We define the area-weighted stretch energy on $\partial\mathcal{M}$ with $m = \#\mathbb{V}(\partial\mathcal{M})$. For $\hat{v}_i \in \mathbb{V}(\partial\mathcal{M})$, $g(\hat{v}_i) = (g_i^1, g_i^2, g_i^3)$, $i = 1, \dots, m$, $\mathbf{g}^t = (g_i^t, \dots, g_m^t)^\top$, $t = 1, 2, 3$ and $\mathbf{g} = [\mathbf{g}^1, \mathbf{g}^2, \mathbf{g}^3] \in \mathbb{R}^{m \times 3}$. The piecewise linear function g on $\partial\mathcal{M}$ is given by the barycentric coordinates, g is called the induced function by $\mathbf{g}$ and $\mathbf{g}$ is the inducing matrix for g . The area-weighted stretch energy²⁶ on $\partial\mathcal{M}$ is defined as

$$E_S(g) = \frac{1}{2} \sum_{t=1}^3 (\mathbf{g}^t)^\top L_S(g) \mathbf{g}^t, \quad (6a)$$

where $L_S(g)$ is the area-weighted Laplacian matrix with

$$[L_S(g)]_{i,j} = \begin{cases} w_{i,j}(g), & [v_i, v_j] \in \mathbb{E}(\partial\mathcal{M}), \\ w_{i,i}(g) = -\sum_{\ell \neq i} w_{i,\ell}(g), & i = j, \\ 0, & \text{otherwise,} \end{cases} \quad (6b)$$

and

$$w_{i,j}(g) = -\frac{1}{2} \left(\frac{\cot \theta_{i,j}(g)}{\sigma_{g^{-1}}([v_i, v_j, v_\ell])} + \frac{\cot \theta_{j,i}(g)}{\sigma_{g^{-1}}([v_j, v_i, v_m])} \right), \quad i \neq j, \quad (6c)$$

where $\theta_{i,j}(g)$ and $\theta_{j,i}(g)$ are two angles opposite to edge $g([v_i, v_j])$ and $\sigma_{g^{-1}}(\alpha) = \rho(\alpha)|\alpha|/|g(\alpha)|$, for $\alpha \in \mathbb{F}(\partial\mathcal{M})$, is the local area-measure stretch factor.

To compute the cubic area-measure-preserving OMT (A-OMT) map from $\partial\mathcal{M}$ to $\partial\mathcal{C}^3$, we utilize the PGM proposed in²⁵, which can be used to efficiently compute the A-OMT maps $h_\rho^* : \partial\mathcal{M} \rightarrow \mathcal{S}^2$, where $\mathcal{S}^2$ denotes the unit sphere in $\mathbb{R}^3$ and $h_1^* : \partial\mathcal{C}^3 \rightarrow \mathcal{S}^2$ ($\rho = 1$), respectively. Then, the composition map $g_\rho^* = (h_1^*)^{-1} \circ h_\rho^* : \partial\mathcal{M} \rightarrow \partial\mathcal{C}^3$, as shown in Figure 1, is the desired A-OMT map. The computational procedure is summarized in Algorithm 1.

Cubic Volume-Measure-Preserving OMT Maps. In this section, we will develop the OMT algorithm for solving the cubic OMT map f_ρ^* , as in (5), from $\mathcal{M}$ to $\mathcal{C}^3$ directly. Let g_ρ^* be the cubic A-OMT map from $\partial\mathcal{M}$ to $\partial\mathcal{C}^3$ computed by Algorithm 1. We now construct a homotopy $g_\zeta : \partial\mathcal{M} \rightarrow \mathbb{R}^3$ for the boundary maps by

$$g_\zeta(\hat{v}) = (1 - \zeta)\hat{v} + \zeta g_\rho^*(\hat{v}), \quad \hat{v} \in \mathbb{V}(\partial\mathcal{M}), \quad \zeta \in [0, 1]. \quad (7)$$

Figure 1. A diagram illustrating the construction of the A-OMT map g_ρ^* between $(\partial\mathcal{M}, \rho)$ and $(\partial\mathcal{C}^3, \delta \equiv 1)$.

Algorithm 1 Cubic A-OMT from $\partial\mathcal{M}$ to $\partial\mathcal{C}^3$

Input: A genus-zero simplicial 2-complex $\partial\mathcal{M}$ of mass one and a piecewise linear density function ρ on $\partial\mathcal{M}$.

Output: A cubic A-OMT map $g_\rho^* : \partial\mathcal{M} \rightarrow \partial\mathcal{C}^3$.

- 1: Construct a Delaunay triangular mesh of $\partial\mathcal{C}^3$.
 - 2: Compute the A-OMT map $h_\rho^* : \partial\mathcal{M} \rightarrow \mathcal{S}^2$ by the PGM.
 - 3: Compute the A-OMT map $h_1^* : \partial\mathcal{C}^3 \rightarrow \mathcal{S}^2$ by the PGM.
 - 4: **return** $g_\rho^* = (h_1^*)^{-1} \circ h_\rho^*$.
-

Then, $[0, 1]$ is uniformly partitioned by $\{0 = \zeta_0 < \zeta_1 < \dots < \zeta_p = 1\}$. Let $L_V(f)$ be the mass-weighted Laplacian matrix with the (i, j) th coefficient

$$w_{i,j}(f) = -\frac{1}{9} \sum_{\substack{\tau \in \mathbb{T}(\mathcal{M}) \\ [v_i, v_j] \cup [v_\ell, v_m] \subset \tau \\ [v_i, v_j] \cap [v_\ell, v_m] = \emptyset}} \frac{|f([v_i, v_\ell, v_m])| |f([v_j, v_m, v_\ell])| \cos \theta_{i,j}^{\ell,m}(f)}{\rho(\tau) |\tau|}, \quad (8)$$

as in²⁴, where $\theta_{i,j}^{\ell,m}(f)$ is the dihedral angle between $f([v_i, v_\ell, v_m])$ and $f([v_j, v_m, v_\ell])$ in tetrahedron $f([v_i, v_j, v_\ell, v_m])$. For $k = 1, \dots, p$, we compute the interior map by solving the linear system

$$[L_V(f^{(k-1)})]_{\mathbf{I}, \mathbf{I}} \mathbf{f}_{\mathbf{I}}^{(k)} = -[L_V(f^{(k-1)})]_{\mathbf{I}, \mathbf{B}} [g_{\zeta_k}(\hat{v})]_{\hat{v} \in \mathbb{V}(\partial\mathcal{M})}, \quad (9)$$

where $f^{(0)} = \text{id.}$, $n = \#(\mathbb{V}(\mathcal{M}))$, $\mathbf{B} = \{i \mid \hat{v}_i \in \partial\mathcal{M}\}$, and $\mathbf{I} = \{1, \dots, n\} \setminus \mathbf{B}$. The map $f^{(p)} : \mathcal{M} \rightarrow \mathcal{C}^3$ is the desired cubic V-OMT map f_ρ^* . The corresponding computational procedure is stated in Algorithm 2.

Algorithm 2 Cubic V-OMT from $\mathcal{M}$ to $\mathcal{C}^3$

Input: A simplicial 3-complex $\mathcal{M}$ with a genus-zero boundary of mass one and a piecewise linear density function ρ on $\mathcal{M}$.
 $p = 11$.

Output: A cubic V-OMT map $f_\rho^* : \mathcal{M} \rightarrow \mathcal{C}^3$.

- 1: Let $g : \partial\mathcal{M} \rightarrow \partial\mathcal{C}^3$ be a cubic A-OMT map computed by Algorithm 1.
 - 2: Solve $[L_V(\text{id.})]_{\mathbf{I}, \mathbf{I}} \mathbf{f}_{\mathbf{I}}^t = -[L_V(\text{id.})]_{\mathbf{I}, \mathbf{B}} \mathbf{f}_{\mathbf{B}}^t$, $t = 1, 2, 3$.
 - 3: Let f be induced by $\mathbf{f}$.
 - 4: **for** $\zeta = 1/p, 2/p, \dots, 1$ **do**
 - 5: Let $\mathbf{f}_{\mathbf{B}} = (1 - \zeta)\hat{\mathbf{v}} + \zeta\mathbf{g}$, $\hat{\mathbf{v}} = [\hat{v}_1, \dots, \hat{v}_m]^\top$, and $\hat{v}_i \in \mathbb{V}(\partial\mathcal{M})$.
 - 6: Solve $[L_V(f)]_{\mathbf{I}, \mathbf{I}} \mathbf{f}_{\mathbf{I}}^t = -[L_V(f)]_{\mathbf{I}, \mathbf{B}} \mathbf{f}_{\mathbf{B}}^t$, $t = 1, 2, 3$.
 - 7: Let f be induced by $\mathbf{f}$.
 - 8: **end for**
 - 9: **return** $f_\rho^* \leftarrow f$. ($\mathbf{f}_\rho^* \leftarrow \mathbf{f} = [\mathbf{f}_{\mathbf{I}}^\top, \mathbf{f}_{\mathbf{B}}^\top]^\top$.)
-

To study the partition number p of homotopy in step (9), we define the total mass distortion and the local mass ratio as

$$d_{\mathcal{M}}(f) = \sum_{v \in \mathbb{V}(\mathcal{M})} \sum_{\tau \in \mathcal{N}(v)} \left| \frac{\rho(\tau) |\tau| - |f(\tau)|}{4} \right|, \quad r_f(v) = \sum_{\tau \in \mathcal{N}(v)} \frac{\rho(\tau) |\tau|}{|f(\tau)|}, \quad (10)$$

respectively, where $\mathcal{N}(v)$ is the set of 1-ring neighboring tetrahedrons of v .

Two-Phase U-net with OMT for Training and Validation

A brain image scanned by mpMRI typically provides four modalities, namely, FLAIR, T1, T1CE and T2, with various grayscale values ranging from 0 to 65535 on each voxel of four $240 \times 240 \times 155$ cuboids, denoted by $\{I_s\}_{s=1}^4$. For a training brain image, let $\mathcal{L}$ denote the $240 \times 240 \times 155$ labeled cuboid, where $WT = \{2, 1, 4\}$, $TC = \{1, 4\}$, $ET = \{4\}$ and $\{0\}$ for others. The grayscale values on I_s can be normalized in $[0, 1]$ by (grayscale value – mean)/variance. In practice, these grayscale values are adjusted by CEHE in $[0, 1]$, denoted by $\bar{I}_s$.

A physical brain image is contained in I_s and accounts for approximately 12%–20% of the voxels. Suppose $\mathcal{M} \subseteq I_s$ is a simplicial 3-complex with a genus-zero boundary composed of tetrahedral meshes representing a brain image. Since $\bar{I}_1$ records the adapted CEHE grayscale values of FLAIR and, in general, the FLAIR modality typically reflects the distribution of $WT = \{2, 1, 4\}$, the adapted CEHE grayscale values on the voxel $\bar{I}_1(i, j, k)$ can help with defining the density map on $\mathbb{V}(\mathcal{M})$ by

$$\rho_\gamma(v) = \exp(\gamma \bar{I}_1(i, j, k)), \quad v \in \bar{I}_1(i, j, k), \quad (11)$$

where γ is some value chosen from $[1, 2]$.

Two-Phase U-Net with OMT for Training. For the given samples in the training set of 3D brain images, we propose a 2P-Unet-OMT algorithm with density function estimates to construct an effective input tensor for the U-net algorithm. In general, a real brain image roughly contains 1.5 million vertices. It is reasonable to cover a brain image with 128^3 voxels.

Phase I We first construct training tensors by using the OMT algorithm with the density $\rho_\gamma(v)$, as in (11). By computing the OMT map $f_{\rho_\gamma}^*$ with Algorithm 2 from $\mathcal{M}$ to a $128 \times 128 \times 128$ cube $\mathcal{N}_0^\gamma$, we obtain four $128 \times 128 \times 128$ cubes $\{\mathcal{N}_{0,s}^\gamma\}_{s=1}^4$, one $128 \times 128 \times 128$ cube $\mathcal{L}_0^\gamma$ corresponding to the grayscale values of $\mathcal{M} \subseteq I_s$, $s = 1, \dots, 4$, and labels in $\mathcal{M} \subset \mathcal{L}$, as shown in Figure 2a. Via the OMT map $f_{\rho_\gamma}^*$, we define the grayscale and the label on each voxel $u \in \mathcal{N}_{0,s}^\gamma$ and $\mathcal{L}_0^\gamma$ by $f_{\rho_\gamma}^{*-1}(c_u)$ in some voxels v of I_s and $\mathcal{L}$, respectively, where c_u is the center of u , for $s = 1, \dots, 4$. Then, as shown in Figure 2b, we use U-net to train Net0 with $\{\mathcal{N}_{0,s}^\gamma\}_{s=1}^4$ and $\mathcal{L}_0^\gamma$ as the input tensors, where $\mathcal{L}_0^\gamma$ denotes the $(0, 1)$ labels, which are given as $\{0 = \{0\}, 1 = \{2, 1, 4\}\}$ from the $240^2 \times 155$ MRI image.

(a) OMT map $f_{\rho_\gamma}^*$, $\{\mathcal{N}_{0,s}^\gamma\}_{s=1}^4$ for grayscale values and $\mathcal{L}_0^\gamma$ for labels of $\mathcal{M}$.

(b) Net0 for the detection of the possible region of WT.

Figure 2. The procedure of Phase I. (a) Compute the OMT map $f_{\rho_\gamma}^*$ from $\mathcal{M}$ to $\mathcal{N}_0^\gamma$ and construct $\{\mathcal{N}_{0,s}^\gamma\}_{s=1}^4$ of $128 \times 128 \times 128 \times 4$ grayscale values for FLAIR, T1, T1CE, and T2 and $\mathcal{L}_0^\gamma$ with labels of $128 \times 128 \times 128$ for $\mathcal{M}$ by MRI. (b) Input format $\{\mathcal{N}_{0,s}^\gamma\}_{s=1}^4$ and $\mathcal{L}_0^\gamma$ for the U-net and the trained Net0.

Net0 is designed to detect the possible tumor region of WT and then used to construct a new density function for enlarging the tumor region for Phase II.

Phase II For a given training brain image, we cover the possible tumor region of WT that is labeled $\mathbf{1} = \{2, 1, 4\}$ by m voxels with morphology dilation; that is, $\mathbb{T} \subseteq \mathcal{M}$. Let $\rho_\gamma(v)$ be a step-like functions defined as

$$\rho_\gamma(v) = \begin{cases} \exp(\gamma \bar{I}_1(i, j, k)), & \text{if } v \in \bar{I}_1(i, j, k) \subseteq \mathbb{T}, \\ 1.0, & \text{if } v \notin \mathbb{T}. \end{cases} \quad (12)$$

Then, we construct new smooth density functions using the image filtering technique by convoluting $\rho_\gamma(v)$ with a $m \times m \times m$ blur box tensor, as follows:

$$\tilde{\rho}_\gamma(v) \leftarrow \rho_\gamma(v) \otimes \frac{\mathbb{1}_{m \times m \times m}}{m^3}, \quad v \in \mathcal{M}, \quad (13)$$

where $\rho_\gamma(v)$ is given in (12). As shown in Figure 3a, we compute the OMT map $f_{\tilde{\rho}_\gamma}^*$ from $\mathcal{M}$ to a $128 \times 128 \times 128$ cube $\mathcal{N}_1^\gamma$ and obtain four $128 \times 128 \times 128$ cubes $\{\mathcal{N}_{1,s}^\gamma\}_{s=1}^4$ corresponding to the grayscale values of $\mathcal{M} \subseteq I_s$ via OMT $f_{\tilde{\rho}_\gamma}^*$, as in Phase I. We set three $128 \times 128 \times 128$ cubes, $\mathcal{L}_1^\gamma$, $\mathcal{L}_2^\gamma$ and $\mathcal{L}_3^\gamma$, which are associated with the labels of $\{\mathbf{0} = \{0\}, \mathbf{1} = \{2, 1, 4\}\}$, $\{\mathbf{0} = \{0, 2\}, \mathbf{1} = \{1, 4\}\}$, and $\{\mathbf{0} = \{0, 2, 1\}, \mathbf{1} = \{4\}\}$, respectively, of $\mathcal{M}$ by MRI. Here, $\mathcal{L}_1^\gamma$, $\mathcal{L}_2^\gamma$ and $\mathcal{L}_3^\gamma$ with $\mathbf{0}$ and $\mathbf{1}$ elements are the new labels of WT, TC, and ET, respectively. Then, as shown in Figure 3b, we use U-net to train Net1, Net2 and Net3 for WT, TC and ET, respectively, where $\{\mathcal{N}_{1,s}^\gamma\}_{s=1}^4$ applies $\mathcal{L}_1^\gamma$, $\mathcal{L}_2^\gamma$ and $\mathcal{L}_3^\gamma$ as the initial input tensors.

Figure 3. The procedure of Phase II. (a) Compute OMT $f_{\tilde{\rho}_\gamma}^*$ from $\mathcal{M}$ to $\mathcal{N}_1^\gamma$ and $\{\mathcal{N}_{1,s}^\gamma\}_{s=1}^4$ of $128 \times 128 \times 128 \times 4$ grayscale values for $\{I_s\}_{s=1}^4$. Let $\mathcal{L}_1^\gamma$, $\mathcal{L}_2^\gamma$ and $\mathcal{L}_3^\gamma$ be labels of $128 \times 128 \times 128 \times 3$ for $\mathcal{M}$ corresponding to WT, TC and ET, respectively, where $\tilde{\rho}_\gamma$ is given in (13). (b) Input format $\{\mathcal{N}_{1,s}^\gamma\}_{s=1}^4$ with $\mathcal{L}_1^\gamma$, $\mathcal{L}_2^\gamma$ and $\mathcal{L}_3^\gamma$ for U-net and the trained Net1, Net2 and Net3 models for WT, TC and ET, respectively.

Net0 and Net1–Net3 for Validation. Once we have computed Net0 and Net1–Net3 in Phase I and Phase II, respectively, we use Net0 to detect the possible tumor region of $\text{WT} = \{2, 1, 4\}$ with the density function $\rho_\gamma(v)$ defined in (11) and cover WT by m voxels with dilation, that is, $\mathbb{T} \subseteq \mathcal{M}$, and construct a new density function $\tilde{\rho}_\gamma$, as in (13) and (12). We compute four $128 \times 128 \times 128$ cubes $\{\mathcal{N}_{1,s}^\gamma\}_{s=1}^4$ for grayscale values of FLAIR, T1, T1CE and T2 via OMT $f_{\tilde{\rho}_\gamma}^*$ and use Net1, Net2 and Net3 to validate three $128 \times 128 \times 128$ cubes, $\mathcal{L}_1^\gamma$, $\mathcal{L}_2^\gamma$ and $\mathcal{L}_3^\gamma$, for the predicted labels. Here, each voxel u_i^t in $\mathcal{L}_i^\gamma$, $t = 1, 2, 3$, is predicted by a probability $0 \leq \tilde{p}_i^t \leq 1$. The validation of a brain image in the testing set proceeds as follows. Let $\mathcal{L}_* \supset \mathcal{M}$ be a $240 \times 240 \times 155$ cuboid with each voxel in $\mathcal{L}_*$ set to $\{0\}$.

- (i) For each voxel $v_j \in \mathcal{M} \subset \mathcal{L}_*$, let $f_{\tilde{p}_\gamma}^{*-1}(c_{u_i^t}) \in v_j$ for $t = 1, 2, 3$, where $c_{u_i^t}$ is the center of u_i^t for $i = 1, \dots, n(j) \neq 0$ and $f_{\tilde{p}_\gamma}^{*-1}(c_{u_k^t})$ be the closest to the center of v_j if $n(j) = 0$. Then we define the probability p_j^t of v_j as

$$p_j^t = \begin{cases} \sum_{i=1}^{n(j)} \tilde{p}_i^t / n(j), & \text{if } n(j) \neq 0, \\ \tilde{p}_k^t, & \text{if } n(j) = 0. \end{cases} \quad (14)$$

- (ii) For $t = 1, 2, 3$, at each voxel $v_j \in \mathcal{M}$:
 Net1 evaluates $\mathbf{0} = \{0\}$, if ① $p_j^1 < 1/2$ and $\mathbf{1} = \{2, 1, 4\}$, if ② $p_j^1 \geq 1/2$;
 Net2 evaluates $\mathbf{0} = \{0, 2\}$, if ③ $p_j^2 < 1/2$ and $\mathbf{1} = \{1, 4\}$, if ④ $p_j^2 \geq 1/2$;
 Net3 evaluates $\mathbf{0} = \{0, 2, 1\}$, if ⑤ $p_j^3 < 1/2$ and $\mathbf{1} = \{4\}$, if ⑥ $p_j^3 \geq 1/2$.
 (iii) The voxel $v_j \in \mathcal{M} \subset \mathcal{L}_*$ is labeled by the procedure in Figure 4.

Figure 4. Flowchart of predicting the label of $v_j \in \mathcal{M} \subset \mathcal{L}_*$.

We denote GT as the ground truth of $WT \supset TC \supset ET$ and PD as the prediction of WT, TC and ET, respectively, by (i)–(iii) above. The associated relationship of between sets of GT and PD is plotted in Figure 5. For numerical experiments, we define the Dice score as

$$\text{Dice} = \frac{2|GT \cap PD|}{|GT| + |PD|}. \quad (15)$$

In (ii), we see that for each voxel v_j in $\mathcal{M}$, we utilize the multivalues $p_i^t, i = 1, \dots, n(j)$, on the cube to define the most likely probability, which can be used to make a more precise evaluation of the label prediction.

Figure 5. The relationships between WT, TC and ET of the ground truth and prediction.

Furthermore, if we define GT_p and PD_p as the probability density tensors of GT and PD, respectively, we can define

$$\text{Loss function} = \text{Dice loss} + \text{Cross entropy loss} \equiv \left(1 - \frac{2 \sum_i (GT_p \cdot PD_p)_i}{\sum_i (GT_p + PD_p)_i} \right) - \sum_i (GT_p)_i \log(PD_p)_i. \quad (16)$$

The Dice loss in (16) can help with checking the convergence of the training procedure for WT, TC and ET vs. epochs by the U-net algorithm.

Improvement in Dice Scores with Mesh Refinement and Ensemble Voting Postprocessing. In this subsection, we propose two methods to improve the Dice scores of WT, TC, and ET. One is the mesh refinement on the WT region for the OMT map, and the other is the ensemble voting postprocessing.

(a) Mesh refinement. With the merit of 2P-Unet-OMT, the density distribution of interesting regions in a brain image computed by Phase I can be enlarged with finer meshes and can be better viewed in Phase II for U-net training.

One of the most important features of the OMT map is that the density can be increased in the region of interest, and then the region can be remeshed by the mesh refinement technique. In this way, due to the mass-preserving property of OMT, the region of interest can be enlarged in the cube, which enables U-net to learn more efficiently and achieve high-performance prediction results.

(b) Ensemble voting postprocess. We propose an ensemble voting postprocessing approach to determine the final labels in the brain image for validation. The main purpose of this postprocessing step is to modify the probability p'_i , $t = 1, 2, 3$, in steps (i)–(iii) of paragraph “Net0 and Net1–Net3 for Validation”. We first select the three best models $\{\text{Net1}_v, \text{Net2}_v, \text{Net3}_v\}_{v=1}^3$ for WT, TC and ET from the training procedure. For each $128 \times 128 \times 128$ brain tensor ($\mathcal{R}_0$) for validation, we further build four $128 \times 128 \times 128$ tensors with 90 degree counterclockwise rotations ($\mathcal{R}_1$), mirroring from the left to the right ($\mathcal{R}_2$), mirroring from the top to the bottom ($\mathcal{R}_3$) and mirroring from the left to the right followed by a 90 degree counterclockwise rotation ($\mathcal{R}_4$).

For $t = 1, 2, 3$, at each voxel u_i^t in $\mathcal{R}_0, \dots, \mathcal{R}_4$, Net1_t , Net2_t and Net3_t can predict 15 probabilities $\{p'_{\mu,v}(i) \mid \mu = 0, \dots, 4, v = 1, 2, 3\}$. We compute

$$\text{Dice}([p'_{0,1}], [p'_{\mu,v}]) \equiv D_{\mu,v}, \alpha'_{\mu,v} = \begin{cases} 1, & D_{\mu,v} \geq 0.8, \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

and

$$p'_i = \sum_{v=1}^3 \sum_{\mu=0}^4 \alpha_{\mu,v} p'_{\mu,v}(i) / N, \quad (18)$$

where $[p'_{\mu,v}] =$ The set of $\{0, 1\}$ by taking Gaussian sybmols of $p'_{\mu,v}(i)$ and $N = \sum_{v=1}^3 \sum_{\mu=0}^4 \alpha_{\mu,v}$. Then, we perform the conversion in steps (i)–(iii) to convert the associated labels back to the brain image.

The various rotations $\mathcal{R}_1, \dots, \mathcal{R}_4$ of the brain tensor $\mathcal{R}_0$ constructed above indeed help with improving the Dice scores with the ensemble voting technique developed in (17) and (18).

Results and Discussions

Based on the CNN technique, the U-net algorithm is designed to learn an effective network from training data using an optimization process that requires decreasing the model error of the loss function on the training and validation sets. We adopt the U-net algorithm and set the hyperparameters as follows: encoder depth: 3, initial learning rate: $\alpha_0 = 1.0 \times 10^{-4}$, learning rate drop factor: $F = 0.95$, learning rate drop period: $P = 10$, L_2 -regularization: 1.0×10^{-4} , minimum batch size: 8.

For the 1251 brain image samples in the BraTS 2021 challenge database^{3,4,17}, we randomly fix 1000 samples for training and 251 for validation. The training is carried out on a server equipped with an NVIDIA Tesla V100S PCIe 32 GB×4 GPU. All calculations are implemented in MATLAB R2020a.

Partition number p in Algorithm 2. We select *BraTS0002* as an $\mathcal{M}$ from the BraTS 2021 dataset and compute the cubic V-OMT from $\mathcal{M}$ to $\mathcal{C}^3$ by Algorithm 2. In Figure 6, we plot the statistical summary of the local mass distortion $\sum_{\tau \in \mathcal{N}(v)} |\rho(\tau)|\tau| - |f(\tau)|/4$, as in (10), and $r_f(v)$ for all $v \in \mathbb{V}(\mathcal{M})$ versus the partition number p of homotopy. In each box, the red centerline indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The dotted lines extend to the most extreme data points that are not considered outliers, and the outliers are represented separately with “+” signs. Furthermore, in Figure 7, we also plot the statistical summary of the total mass distortion $d_{\mathcal{M}}(f)$ and the mean and standard deviation (SD) of $r_f(v)$ vs. the partition number p for the first 1000 brain samples from the BraTS 2021 dataset.

Figure 6 shows that when $p = 11$, the cubic V-OMT between *BraTS0002* and $\mathcal{C}^3$ has the smallest local mass distortion and the closest local mass ratio to one. Moreover, in Figure 7, when $p = 11$, the first 1000 brain samples of BraTS 2021 have the smallest total mass distortion and the best mean and SD of the local mass ratios. Therefore, we choose $p = 11$ in Algorithm 2.

Dimension m of blur box tensor. We now discuss the dimension m of blur box tensor in (13), which covers the WT region by m voxels. To choose a suitable number m for the covering voxels with dilation for WT, we apply the recall and precision metrics, which are defined as

$$\text{Recall} = \frac{|\text{PD} \cap \text{GT}|}{|\text{GT}|}, \text{ Precision} = \frac{|\text{PD} \cap \text{GT}|}{|\text{PD}|}, \quad (19)$$

where PD denotes the prediction of $\{\text{WT covered by } m \text{ voxels with dilation}\}$ by Net0.

In fact, the recall metric in (19) indicates how many voxels lie in the prediction, and the precision metric in (19) indicated how precise the prediction is. Thus, we want to make both recall and precision as large as possible. In Figure 8, we plot the mean, minimum, median and maximum values of the recall and precision metrics of the WT validation vs. the numbers of covering voxels with dilation. We find that $m = 5$ is a suitable number to balance the recall and precision values for the validation data.

Figure 6. Statistical summary of the (a) local mass distortion and (b) $r_f(v)$ for all vertices of *BraTS0002* vs. the partition number p of homotopy.

Figure 7. Statistical summary of the (a) $d_{\mathcal{M}}(f)$; (b) mean and (c) SD of $r_f(v)$ vs. the partition number p of homotopy for the first 1000 brain samples of BraTS 2021.

Figure 8. Mean, minimum, median and maximum values of the recall and precision metrics of the WT validation vs. the numbers of covering voxels.

BraTS 2021	Transport Cost		#Fold.		Enlarged Ratio	
	Mean	SD	Mean	SD	Mean	SD
Validation (1000)	0.028	0.004	17.13	9.49	1.480	0.078
Testing (251)	0.029	0.003	18.62	9.24	1.444	0.079

Table 1. Mean and SD values of the transport costs, folding numbers and enlarged ratios of $f_{\hat{\rho}_1}^*$ from $\mathcal{M}$ to $\mathcal{C}^3$.

For fixed $m = 5$, in Table 1, we list the mean, minimum, median, maximum and SD values of the transport costs, folding numbers and enlarged ratios for both the 1000 training samples and the 251 validation samples. The enlarged ratio is defined by (the ratio of WT in the raw data)/(the ratio of WT in the cube).

In Table 1, we observe that the numerical results of the transport costs, folding numbers and enlarged ratios for the 1000 training and 251 validation samples computed by $f_{\hat{\rho}_1}^*$ in (13) are in line with what we expected.

Dice Scores and Loss Functions. We first compare 2P-Unet-OMT developed in Section “Two-Phase U-net with OMT for Training and Validation” with one-phase Unet-OMT (1P-Unet-OMT); i.e., the density functions of (11) with $\gamma = 1.0$ and 1.5 are used for training Net1. We learn Net1 by using 2P- and 1P-Unet-OMT with 310 epochs. In Figures 9a and 9b, we plot the Dice scores of WT for training and validation by 2P- and 1P-Unet-OMT, respectively. We observe that for both the training and validation scores, 2P-Unet-OMT is obviously much better than 1P-Unet-OMT. Therefore, in the following numerical experiments, we prefer to adopt 2P-Unet-OMT.

Figure 9. Dice scores of WT for training and validation by 2P- and 1P-Unet-OMT.

To expand the training data in Phase II, we use three different density functions $\tilde{\rho}_1(v)$, $\tilde{\rho}_{1.5}(v)$ and $\tilde{\rho}_2(v)$ for $v \in \bar{I}_1(i, j, k) \subseteq \mathcal{T}$, as in (13), to create 3000 augmentation brain images for training. We now use 2P-Unet-OMT to train Net0 and Net1-Net3 on 3000 training samples. Then, we utilize them to obtain predictions on the 251 validation samples. In Figure 10, we plot the Dice scores with blue “o” and “x” symbols and the loss functions with red “o” and “x” symbols vs. the epoch numbers for the training and validation sets of WT, TC and ET, respectively. Note that the Dice scores for WT, TC and ET are defined by (15) and the loss function is defined by (16). The predicted labels of WT, TC and ET in a brain image are evaluated by steps (i)–(iii), which are precisely determined by the probability value $p_j^t = \sum_{i=1}^{n(j)} p_i^t / n(j)$, ($t = 1, 2, 3$) in each voxel $v_j \in \mathcal{M}$.

Figure 10. Dice scores (blue) and loss functions (red) for the training (o) and validation (x) sets of WT, TC and ET vs. the epoch number.

We see that the training and validation Dice scores for WT, TC and ET increase very fast during the first 50 epochs but

then do not increase significantly and reach (0.9720, 0.9673, 0.9330) and (0.9325, 0.8965, and 0.8614), respectively, after 310 epochs. On the other hand, the training and validation loss function for WT, TC and ET decrease very fast during the first 50 epochs and approach (7.008×10^{-2} , 7.067×10^{-2} , and 8.678×10^{-2}) and (8.006×10^{-2} , 9.487×10^{-2} , and 9.957×10^{-2}), respectively, after 310 epochs. The trends of both the Dice score and loss function value indicate the typical training and validation history. Thus, based on the clear tendency of the curves of the Dice scores and loss functions, in our experiment, we run U-net for 310 epochs.

Merit of Mesh Refinement on the WT Region. Based on the discussion of the merit of mesh refinement on the expanding WT region in the brain image, we compute OMT maps with the smooth density functions $\tilde{\rho}_\gamma(v)$, $\gamma = 1.0, 1.5, 1.75, 2.0$, in (13) for 1000 brain samples to obtain 4000 augmented brain cubes and use U-net to train Net1–Net3. Furthermore, for validation, we compute 2P-OMT for 251 brain samples with the density function $\tilde{\rho}_{1.75}(v)$ on the expanding WT region by Phase I with mesh refinement.

We train U-net for 310 epochs on 4000 augmented brain cubes. From epoch 10 to 310, for every 10 epochs, we validate the Dice scores on the 251 samples of validation data for WT, TC and ET. In Table 2, we show the top 3 validation Dice scores for WT, TC and ET by steps (i)–(iii) at epochs (150, 140, 100), (170, 120, 70) and (130, 170, 80), respectively. The corresponding training Dice scores for WT, TC and ET are listed in the first three columns of Table 2. We see that the validation Dice scores for WT, TC, and ET for the brain image reach 0.93469, 0.90251 and 0.86912, respectively, which is a satisfactory result.

Epochs (WT,TC,ET)	Dice Scores					
	Training			Validation		
	WT	TC	ET	WT	TC	ET
(150, 140, 100)	0.95388	0.95411	0.90912	0.93451	0.90251	0.86912
(170, 120, 70)	0.95363	0.95038	0.90496	0.93469	0.90167	0.86720
(130, 170, 80)	0.95241	0.95475	0.89994	0.93355	0.90187	0.86697

Table 2. Dice scores of WT, TC and ET for the brain image with mesh refinement on the training and validation sets.

Dice Scores with Ensemble Voting Postprocessing. In this subsection, we show the improvement in dice scores with the mesh refinement and the ensemble voting postprocessing approach to determine the final labels in the brain image for validation. We first select the three best models for WT, TC and ET at epochs (150, 170, 130), (140, 120, 170) and (100, 70, 80) from the training procedure, as shown in Table 2, and call them Net1_v, Net2_v, and Net3_v for $v = 1, 2, 3$.

In Figures 11a, 11b, and 11c, we plot the histograms of the Dice scores with and without ensemble voting postprocessing in blue and green lines for WT, TC, and ET, respectively, vs. the epoch. Furthermore, the associated increments of the Dice scores are plotted with red lines in Figures 11a, 11b, and 11c. We see that the Dice scores for WT, TC, and ET with the ensemble voting technique are much better than those without voting postprocessing. In addition, the Dice score curves for WT, TC and ET have a relatively stable upward trend.

Figure 11. Dice scores with (blue line) and without (green line) the ensemble technique and increments (red line) for (a) WT; (b) TC; and (c) ET vs. epoch.

Finally, in Table 3, we show the Dice, sensitivity, specificity, and 95 percentile of the Hausdorff distance (HD95) scores of 251 validation samples for WT, TC, and ET in brain images by Net1_v–Net3_v, $v = 1, 2, 3$, with the ensemble voting technique. We see that Net1_v–Net3_v with mesh refinement and ensemble voting postprocessing, as well as with the precise conversion of

steps (i)–(iii), significantly boosts the validation Dice scores (251) on BraTS 2021. This result is very promising for brain tumor detection and segmentation.

2P-Unet-OMT + Mesh refinement + Ensemble voting

Validation	Dice score	Sensitivity	Specificity	HD95
WT	0.93705	0.93641	0.99947	6.0129
TC	0.90617	0.90886	0.99974	7.9799
ET	0.87470	0.86945	0.99984	9.8300

Table 3. Dice, sensitivity, specificity and HD95 scores on 251 validation samples for WT, TC and ET with the mesh refinement and the ensemble voting techniques.

Conclusions

In this paper, we mainly introduce 2P-Unet-OMT with density estimates for 3D brain tumor detection and segmentation. We first propose a cubic volume-measure-preserving OMT algorithm in Section to compute an OMT map for transforming an irregular 3D brain image to a cube while preserving the local mass ratios and keeping the minimal deformation. Furthermore, OMT is bijective and minimizes the transport cost. The concept of expressing an irregular brain image with a cube with minimal distortion is proposed for the first time in this research field, and these cubes are typically adequate for the tensor input format of the U-net algorithm that creates validation networks. Representing 3D brain images as cubes significantly reduces the effective brain images from sizes of $240 \times 240 \times 155$ to cubes of sizes $128 \times 128 \times 128$ and preserves the global information of tumor features. This novel OTM preprocessing technique can save a large amount of input data and reduce the computational time for training. In addition, the ensemble voting technique proposed in (17)–(18) and the robust conversion steps (i)–(iii) of paragraph “Net0 and Net1–Net3 for Validation” from cubes ($128 \times 128 \times 128$) with predicted labels back to brain images ($240 \times 240 \times 155$) considerably increase the Dice scores for brain images compared to those for cubes on the 1251 brain image samples.

One of the characteristics of the OMT map is that it can control the densities of tumor regions in brain images, and then, via mass-preserving OMT, the high-density areas can be enlarged in the cube so that the U-net algorithm can strengthen the cognition and learning in the high-density regions. In fact, 2P-Unet-OMT in Section is designed for this purpose. Phase I can first catch the possible region of WT and then cover this region by 5 voxels with dilation. Next, Phase II reconstructs new smooth density functions, as in (13), and performs mesh refinement on the range estimated by Phase I. With the advantage of the mass preservation of OMT, the portion of the possible WT region can be enlarged in the cube. Then, the U-net algorithm is utilized to train more effective Net1–Net3 models for tumor prediction and validation.

The Dice scores of WT, TC and ET by Net0 and Net1_v–Net3_v for $v = 1, 2, 3$, with mesh refinement and ensemble voting postprocessing reach 0.93705, 0.90617 and 0.87470 for validation, respectively. 2P-Unet-OMT with mesh refinement sufficiently utilizes the mass-preserving property to significantly improve the tumor detection and segmentation accuracy.

In future work, because an irregular 3D brain image only needs to be represented by a cube in our approach, we have much room to expand the augmented data with various density settings, such as in (12); these settings include rotating, mirroring, shearing and cropping, and will allow for more opportunities to boost the prediction accuracy. In addition, we believe that for a 3D image provided by real 3D scanning instruments that may be developed in the future, the use of OMT to represent an irregular 3D object must retain the structure of the global information. This 3D OMT representation would take advantage of a precise conversion in the three directions in space and be beneficial to the input format of CNN algorithms. We believe this is a cross-trend research direction for medical images in the near future.

Acknowledgments. This work was partially supported by the Ministry of Science and Technology (MoST), the National Center for Theoretical Sciences, the Big Data Computing Center of Southeast University, the Nanjing Center for Applied Mathematics, the ST Yau Center in Taiwan, and the Shing-Tung Yau Center at Southeast University. W.-W. Lin, T.-M. Huang, and M.-H. Yueh were partially supported by MoST 110-2115-M-A49-004-, 110-2115-M-003-012-MY3, and 109-2115-M-003-010-MY2 and 110-2115-M-003-014-, respectively. T. Li was supported in part by the National Natural Science Foundation of China (NSFC) 11971105.

Author contributions. Yau is the leadership of the research project. Li, W.-W. Lin, Yau, and Yueh developed and analyzed the volume-preserving parametrizations of 3-manifolds. Huang, Li, W.-W. Lin, and J.-W. Lin developed and analyzed the 2P-Unet-OMT algorithm. Huang and Yueh implemented the programs. W.-W. Lin wrote the main manuscript text. Huang, J.-W. Lin, and Yueh prepared figures and tables. All authors reviewed the manuscript.

Competing interests. The authors declare no competing interests.

References

1. Antonelli, M. *et al.* The Medical Segmentation Decathlon. <http://medicaldecathlon.com/> (2021). [2106.05735](#).
2. Simpson, A. L. *et al.* A large annotated medical image dataset for the development and evaluation of segmentation algorithms. *arXiv* **1902.09063** [cs.CV] (2019). [1902.09063](#).
3. Menze, B. H. *et al.* The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imag.* **34**, 1993–2024, DOI: [10.1109/TMI.2014.2377694](#) (2015).
4. Bakas, S. *et al.* Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci. Data* **4**, DOI: [10.1038/sdata.2017.117](#) (2017).
5. Bakas, S. *et al.* Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge (2019). [1811.02629](#).
6. Bauer, S. *et al.* Segmentation of brain tumor images based on integrated hierarchical classification and regularization. In *Proc. MICCAI-BRATS*, 10–13 (2012).
7. Tustison, N., Wintermark, M., Durst, C. & Avants, B. ANTs and Årboles. In *Proc. NCI MICCAI-BRATS*, 47–50 (2013).
8. Reza, S. & Iftekharuddin, K. M. Multi-class abnormal brain tissue segmentation using texture features. In *Proc. NCI MICCAI-BRATS*, 38–42 (2013).
9. Bonnotte, N. From Knothe’s rearrangement to Brenier’s optimal transport map. *SIAM J. Math. Anal.* **45**, 64–87, DOI: [10.1137/120874850](#) (2013).
10. Brenier, Y. Polar factorization and monotone rearrangement of vector-valued functions. *Commun. Pure Appl. Math.* **44**, 375–417, DOI: [10.1002/cpa.3160440402](#) (1991).
11. Caffarelli, L. A. The regularity of mappings with a convex potential. *J. Am. Math. Soc.* **5**, 99–104 (1992).
12. Zikic, D., Ioannou, Y., Criminisi, A. & Brown, M. Segmentation of brain tumor tissues with convolutional neural networks. In *Proc. MICCAIBRATS, 2014*, 36–39 (2014).
13. Randhawa, R. S., Modi, A., Jain, P. & Warier, P. Improving boundary classification for brain tumor segmentation and longitudinal disease progression. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, 65–74 (Springer International Publishing, 2016).
14. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 234–241 (Springer International Publishing, 2015).
15. Kamnitsas, K. *et al.* Ensembles of multiple models and architectures for robust brain tumour segmentation. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, 450–462 (Springer International Publishing, 2018).
16. Isensee, F., Kickingereder, P., Wick, W., Bendszus, M. & Maier-Hein, K. H. No new-net. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, 234–244 (Springer International Publishing, 2019).
17. Baid, U. *et al.* The RSNA-ASNR-MICCAI BraTS 2021 benchmark on brain tumor segmentation and radiogenomic classification. *arXiv:2107.02314* (2021).
18. Isensee, F., Kickingereder, P., Wick, W., Bendszus, M. & Maier-Hein, K. H. Brain tumor segmentation and radiomics survival prediction: Contribution to the BRATS 2017 challenge. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, 287–297 (Springer International Publishing, 2018).
19. Lin, W.-W. *et al.* 3D brain tumor segmentation using a two-stage optimal mass transport algorithm. *Sci. Rep.* **11**, 14686 (2021).
20. Kantorovich, L. V. On a problem of Monge. *Uspekhi Mat. Nauk* **3**, 225–226 (1948).
21. Villani, C. *Optimal Transport: Old and New*, vol. 338 of *Grundlehren der mathematischen Wissenschaften* (Springer-Verlag Berlin Heidelberg, 2009).
22. Gu, X., Luo, F., Sun, J. & Yau, S.-T. Variational principles for Minkowski type problems, discrete optimal transport, and discrete Monge-Ampère equations. *Asian J. Math.* **20**, 383–398, DOI: [10.4310/AJM.2016.v20.n2.a7](#) (2016).
23. Su, K. *et al.* Volume preserving mesh parameterization based on optimal mass transportation. *Comput. Aided Des.* **82**, 42–56, DOI: [10.1016/j.cad.2016.05.020](#) (2017).

24. Yueh, M.-H., Li, T., Lin, W.-W. & Yau, S.-T. A novel algorithm for volume-preserving parameterizations of 3-manifolds. *SIAM J. Imag. Sci.* **12**, 1071–1098, DOI: [10.1137/18M1201184](https://doi.org/10.1137/18M1201184) (2019).
25. Yueh, M.-H., Huang, T.-M., Li, T., Lin, W.-W. & Yau, S.-T. Projected gradient method combined with homotopy techniques for volume-measure-preserving optimal mass transportation problems. *J. Sci. Comput.* **88**, DOI: [10.1007/s10915-021-01583-z](https://doi.org/10.1007/s10915-021-01583-z) (2021).
26. Yueh, M.-H., Lin, W.-W., Wu, C.-T. & Yau, S.-T. A novel stretch energy minimization algorithm for equiareal parameterizations. *J. Sci. Comput.* **78**, 1353–1386, DOI: [10.1007/s10915-018-0822-7](https://doi.org/10.1007/s10915-018-0822-7) (2019).