Research Article ISSN Online: 2783-4859 Computational Mathematics and Computer Modeling with Applications

Received 13 Mar 2023
Accepted 2 June 2023

DOI: 10.52547/CMCMA.1.2.86
AMS Subject Classification: 68U10; 65D18

On two convex variational models and their iterative solutions for selective segmentation of images with intensity inhomogeneity

Liam Burrows[a], **Ke Chen**[a] **and Francesco Torella**[b]

Treating images as functions and using variational calculus, mathematical imaging offers to design novel and continuous methods, outperforming traditional methods based on matrices, for modelling real life tasks in image processing. Image segmentation is one of such fundamental tasks as many application areas demand a reliable segmentation method. Developing reliable selective segmentation algorithms is particularly important in relation to training data preparation in modern machine learning as accurately isolating a specific object in an image with minimal user input is a valuable tool. When an image's intensity is consisted of mainly piecewise constants, convex models are available. Different from previous works, this paper proposes two convex models that are capable of segmenting local features defined by geometric constraints for images having intensity inhomogeneity. Our new, local, selective and convex variants are extended from the non-convex Mumford-Shah model intended for global segmentation. They have fundamentally improved on previous selective models that assume intensity of piecewise constants. Comparisons with related models are conducted to illustrate the advantages of our new models.

Keywords: Variational calculus; Inverse problems; Image segmentation; Mumford-Shah; Intensity inhomogeneity; Geometric constraints; Iterative methods.

1. Introduction

Image science is a relatively new research area, arising to tackle the emerging and increasingly challenging problems from the rapidly developing imaging technology. As the number of image modalities increases, the level of difficulties and requirements from sciences and engineering is constantly going up. To provide high quality of processed results, it is natural to build geometry into models by treating images as functions rather than matrices and tensors, and using variational calculus rather than matrix analysis (linear transforms and filtering). Lying in the heart of image sciences is mathematical imaging that offers to design novel and continuous methods, called variational methods, for modelling real life tasks in image processing. See [39, 19] and many references therein. Most problems in imaging are inverse problems due to uniqueness issues and hence use of regularization is common.

Segmentation is one important imaging task that requires partitioning an image into segments in order to identify and classify objects. Over the last few decades, variational segmentation methods have been well studied. Among them, the model by Mumford and Shah (MS) [31] is the most famous and fundamental, aiming to approximate an input image by a piecewise-smooth function. This model is difficult to solve exactly, and so many attempts to simplify or approximate it have been proposed, with the simplification by the Chan and Vese (CV) model [17] thw most well-known. The CV method and variants of it make the assumption of piecewise-constant intensities which limits the performance, particularly on complicated images. This paper considers how to approximate the MS model without the assumption of piecewise-constant intensities and develops a variational model that is convex and can segment one desirable feature only of a given image with intensity inhomogeneity.

An approximation of the MS model is by the two-stage method be Cai et al. [11], who proposed a convex variant of the MS model. Critically their model is able to tackle images with non-constant intensities, and employs a so-called "smoothing

[a] *Centre for Mathematical Imaging Techniques and Department of Mathematical Sciences, University of Liverpool, Liverpool L19 7ZL, United Kingdom. Email: k.chen@liverpool.ac.uk.*
[b] *Liverpool Vascular & Endovascular Service, Royal Liverpool and Broadgreen University Hospitals NHS Trust, Liverpool, L7 8XP, United Kingdom.*
[*] *Correspondence to: K. Chen: www.liv.ac.uk/∼cmchenke.*

and thresholding" (SaT) segmentation paradigm, in which a clean approximation is found in the first stage, and segmentation is achieved by thresholding in a second stage. More investigation of SaT was discussed in [10], where the authors linked the piecewise-constant MS model and the Rudin-Osher-Fatemi (ROF) denoising model [38], consolidating the connection between segmentation and denoising models. Alberti, Bouchitte and Dal Maso [1] proposed another convex relaxation of the MS functional. This is also intended for global segmentation.

Models explicitly addressing segmentation in the presence of intensity inhomogeneity have been studied, such as the model by Li et al. [25] who introduced such local intensity information in the form of a region scaling fitting energy using the CV framework. Another model based on local intensity clustering [27] assumes that the image contains a bias field causing the inhomogeneity. The above two models are non-convex and hence are initialisation dependent. The local CV model by Wang et al. [42] can deal with intensity inhomogeneity and is less sensitive to initialisation, but struggles in the presence of noise. The challenge in these models is heavy dependence on the initialisation of the contour due to use of level sets or non-convexity.

In many practical applications, segmentation of all features or all homogeneous regions is first challenging and second unnecessary. It can be more useful to segment a particular region or regions of interest only. This is known as selective segmentation [18], and is the task of concern in this paper.

An easy way of obtaining a selective segmentation method is by starting initialisation (around a region of interest) or early stopping of the optimisation in a non-convex model. For example, we may use the CV model (if intensity inhomogeneity is not strong), or the more robust distance regularised level set evolution (DRLSE) model by Li et al. [26] (if there is intensity inhomogeneity and weak noise). Equally one may use the edge based active contours method by Kass et al. [24] or Caselles et al. [12] by suitable initialisation of contours. However, this use of such approaches is not robust.

An improvement can be made to enforce selectiveness directly by incorporating geometrical constraints into models. An early such method was the edge-based approach by Gout et al. [23] and, to improve the model to deal with noisy images, Badshah and Chen [5] combined this edge based idea with the CV fidelity. Further refinements include imposing area constraints to prevent the evolving contour from growing too much and adding different and improved distance penalties [35, 41, 36, 9].

There are recent works attempting to extend the above selective models to images of non-piecewise constant intensity. The works by Ali et al. [2] and Mabood et al. [29] have tackled intensity inhomogeneity in selective models. Their idea is to transform the input image to new images that are better approximated by fidelity fitting of intensities with piece-wise constant [2] or linear distribution [29]. Owning to non-convexity, unless a good initialisation is provided, these models may get stuck in unwanted local minima, and hence are not robust. The model by Liu et al. [28] applied geometrical constraints to the SaT paradigm, but their model is not yet robust due to requiring too many markers.

The key observations motivating this work are: (i) the above mentioned selective methods, due to many using the CV fidelity (with the assumption of a piecewise-constant distribution), are not adequate for segmenting real images. (ii) a non-convex model is too strongly dependent of initialisation. In this paper, we overcoming these two problems by proposing two selective and convex segmentation models to tackle intensity inhomogeneity in the MS framework (instead of the CV type) and to segment selectively one feature of the image requiring few markers.

The rest of this paper is organised as follows: In §2 with review the relevant segmentation literature. In §3 we introduce the proposed selective models based on convex variants of the MS model, give some simple theories and detail their numerical solution methods. In §4 we show numerical experiments demonstrating the advantages of our proposed models. Finally in §5 we draw a conclusion.

2. Related works

In this section we review some related literature. We first look at the Mumford-Shah model and convex variants of it. Then we review some recent selective segmentation models, followed by models suited to segmenting images displaying inhomogeneity. Throughout this paper, we denote an input image as z, defined over a rectangular image domain $\Omega \subset \mathbb{R}^2$. Extension to 3D is immediate.

Features in an image present themselves by way of sudden changes in gradients ∇z that are called jumps. If ∇z is slowly and smoothly varying, there are no edges or features to detect when z must appear blurry. This is where imaging science or vision science prefers 'discontinuous' to continuous and differentiable functions for best visual effects or differs from other non-visual sciences.

The aim of **global** segmentation is to decompose z in the entire domain Ω into K subregions Ω_i, $i = 1, ..., K$, with a discontinuity set Γ (where large jumps in ∇z occur) separating these subregions, denoted as $\Omega = \Omega_1 \cup \Omega_2 \cup ... \cup \Omega_K \cup \Gamma$, or equivalently to find the edges Γ. Given a set $\mathcal{M} = \{\mathbf{x}_i \in \Omega, 1 \leq i \leq m\}$ of marker points $\mathbf{x}_i$ placed on Ω, the aim of **selective** segmentation is to identify some **local** image region $\Omega_{\mathcal{M}}$ that is defines an interested object nearest to the set $\mathcal{M}$.

2.1. Image Global Segmentation by Mumford-Shah

For global segmentation, the MS model [31] aims to decompose the image domain Ω into K subregions Ω_i, $i = 1, ..., K$, with a discontinuity set Γ separating these subregions, such that $\Omega = \Omega_1 \cup \Omega_2 \cup ... \cup \Omega_K \cup \Gamma$. The MS model minimises the functional:

$$F_{MS}(u, \Gamma) = \frac{\lambda}{2} \int_{\Omega} (z - u)^2 \, d\mathbf{x} + \int_{\Omega \setminus \Gamma} |\nabla u|^2 \, d\mathbf{x} + \nu \mathcal{H}^1(\Gamma), \tag{1}$$

where $u : \Omega \to \mathbb{R}^2$ is a piecewise-smooth approximation of the input (potentially noisy with Gaussian noise) image $z : \Omega \to \mathbb{R}^2$ and $\mathcal{H}^1(\Gamma)$ is the Hausdorff measure (length of Γ). Here the first term is for denoising (in case z has noise) and simply states that the new and clean image u should be near the given image z; the second term is the main term for locating the jumps of ∇z and the third term is to ensure solvability i.e. only boundaries having the least length are eligible. Clearly (1) is non-convex.

A key concept used widely is called a level set function that represents the purely geometric quantity of length (or area or volume) of a subregion by a function, in order to facilitate analysis and computation ([40]). A level set function is associated with the idea of 'lifting' where a 'difficulty' in low dimension is represented by an 'easy' quantity in a high dimension.

Since it is challenging to compute the last two terms of (1), many attempts have been taken to approximate (1). A well-known approach by Ambrosio and Tortorelli [3, 4] aimed to approximate (1) by a sequence of elliptic variational problems by replacing the edge set Γ with a $2D$ function and replacing the final term in (1) with a phase field energy. The modified model is only quasi-convex, not convex. Non-local approximations, in which a family of continuous, non-decreasing functions are used to avoid the explicit computation of the edge set Γ can be found in [22, 14]. The CV model by [17] simplifies (1) by restricting u to a piecewise-constant intensity distribution, has been intensively studied in the literature but will not be reviewed here. Instead, we focus on how to solve (1) by reformulating it to a convex model.

The more recent work by Cai et al. [11] is a two-stage approach for segmentation. In the first stage a convex variant of the MS model is minimised, and thresholding is applied in stage two to achieve segmentation. Using u to approximate the Heaviside of the above-mentioned level set function (in a relaxation approach), the key observation was that the term $\int_{\Gamma} |\nabla u|^2 d\mathbf{x}$ appearing in (1) is 0 if $u \in W^{1,2}(\Omega)$; hence the second term of MS becomes $\int_{\Omega \setminus \Gamma} |\nabla u|^2 d\mathbf{x} = \int_{\Omega} |\nabla u|^2 d\mathbf{x}$. Their proposed model is given as:

$$F_C(u) = \frac{\lambda}{2} \int_{\Omega} (z - u)^2 d\mathbf{x} + \frac{\mu}{2} \int_{\Omega} |\nabla u|^2 d\mathbf{x} + \int_{\Omega} |\nabla u| d\mathbf{x}, \tag{2}$$

where the smooth approximation u is constrained to be in the Sobolev space $W^{1,2}(\Omega)$, and the total variation is used to approximate $\mathcal{H}^1(\Gamma)$, inspired by [16]. The second-stage of their approach involves thresholding the minimiser u with appropriate values (by using, for example, k-means). The output is a global segmentation of the image, split up into K phases.

Another convex approach by Alberti, Bouchitte and Dal Maso [1] calibrates the optimizer of a new convex functional to coincide with an absolute minimizer of the original (non-convex) MS functional, by finding a lower bound of MS and optimizing it. To introduce the model, we define the characteristic function, $\mathbb{1}_u : \Omega \times \mathbb{R} \to \{0, 1\}$, of the subgraph of u: $\mathbb{1}_u(\mathbf{x}, t) = \begin{cases} 1, & t < u(\mathbf{x}), \\ 0, & \text{else.} \end{cases}$. Denote a vector field $\varphi = (\varphi^{\mathrm{x}}, \varphi^t) \in C_0(\Omega \times \mathbb{R})$. Let u^-, u^+ be limits of u from two sides of Γ with $u^+ > u^-$. Denote by ν_u the unit normal oriented from the u^- side to u^+ side along Γ. Then the convex relaxed version of the MS functional [1] can be written as:

$$F_{CMS}(u) = \sup_{\varphi \in K} \int_{\Omega \times \mathbb{R}} \varphi D\mathbb{1}_u. \tag{3}$$

where the set K is convex, defined by

$$K = \Bigg\{ \varphi \in C_0(\Omega \times \mathbb{R}) : \frac{|\varphi^{\mathrm{x}}(\mathbf{x}, t)|^2}{4} - \varphi^t(\mathbf{x}, t) \le \frac{\lambda}{2}(t - z(\mathbf{x}))^2, \\ \Big| \int_{t_1}^{t_2} \varphi^{\mathrm{x}}(\mathbf{x}, s) ds \Big| \le \nu, \forall \mathbf{x} \in \Omega, t, t_1, t_2 \in \mathbb{R} \Bigg\}. \tag{4}$$

Here the connection of (3) to MS, to be explained shortly, is not yet obvious. A key property of (3) for a general φ is the following [1]:

$$\int_{\Omega \times \mathbb{R}} \varphi D\mathbb{1}_u = \int_{\Omega \setminus \Gamma} (\varphi^{\mathrm{x}} \cdot \nabla u - \varphi^t) d\mathbf{x} + \int_{\Gamma} \Big(\int_{u^-}^{u^+} \varphi^{\mathrm{x}} dt \Big) \nu_u d\mathcal{H}^1. \tag{5}$$

Combining the cosine formula for vectors, $\varphi^{\mathrm{x}} \cdot \nabla u \le |\varphi^{\mathrm{x}}|^2/4 + |\nabla u|^2$, with the first constraint in (4), $|\varphi^{\mathrm{x}}(\mathbf{x}, u)|^2/4 - \varphi^t(\mathbf{x}, u) \le \frac{\lambda}{2}(u - z(\mathbf{x}))^2$, we get $\varphi^{\mathrm{x}} \cdot \nabla u - \varphi^t \le |\nabla u|^2 + \frac{\lambda}{2}(u - z(\mathbf{x}))^2$, which clearly connects to the fitting term and the regularisation term of the MS. Hence if $\varphi \in K$ from (4), the first term in (5) becomes:

$$\int_{\Omega \setminus \Gamma} (\varphi^{\mathrm{x}}(\mathbf{x}, u) \cdot \nabla u - \varphi^t(\mathbf{x}, u)) d\mathbf{x} \le \int_{\Omega \setminus \Gamma} |\nabla u|^2 d\mathbf{x} + \frac{\lambda}{2} \int_{\Omega} (u - z(\mathbf{x}))^2 d\mathbf{x}.$$

The second term in (5) satisfies $\int_\Gamma \left(\int_{u^-}^{u^+} \varphi^x dt \right) \nu_u d\mathcal{H}^1 \leq \nu \int_\Gamma \nu_u d\mathcal{H}^1 = \nu \mathcal{H}^1(\Gamma)$. So, overall, the convex functional (3) provides a lower bound for MS (1):

$$\int_{\Omega\times\mathbb{R}} \varphi D\mathbb{1}_u \leq \frac{\lambda}{2}\int_\Omega (z-u)^2 d\mathbf{x} + \int_{\Omega\setminus\Gamma} |\nabla u|^2 d\mathbf{x} + \nu\mathcal{H}^1(\Gamma), \tag{6}$$

since the right hand side is exactly the MS functional. As remarked before, this particular lower bound has the same minimiser as the MS model. This theoretical formulation was not tested in [1] but was later studied in [34] by introducing an efficient primal dual algorithm for solving it.

2.2. Local and Selective Image Segmentation

Unlike **global** segmentation methods as reviewed above, selective segmentation methods aim to segment a particular and **local** region(s) of interest only. One method of doing so is to use a classic edge based model, initialise inside the region of interest, evolve the contour and stop when suitable criteria has been met (either convergence or early stopping when the region of interest has been segmented, to prevent leaking).

The distance regularised level set evolution (DRLSE) model introduced by Li et al. [26] is one such edge based segmentation method, involving driving a contour to a stopping condition informed by an edge detector. The DRLSE model is given by the following:

$$F_{DRLSE}(\phi) = \mu\int_\Omega p(|\nabla\phi|)d\mathbf{x} + \lambda\int_\Omega g\delta(-\phi)|\nabla\phi|d\mathbf{x} + \alpha\int_\Omega gH(\phi)d\mathbf{x}, \tag{7}$$

where g is an edge detector, H is the Heaviside function and δ is the Dirac delta function. The authors suggest two different definitions for p given as: $p_1(s) = \frac{1}{2}(s-1)^2$, $p_2(s) = \frac{1}{(2\pi)^2}(1-\cos 2\pi s)$, if $s \leq 1$ and $\frac{1}{2}(s-1)^2$ if $s \geq 1$. Generally $p = p_2$ is preferred. Finally, the DRLSE model uses a level set function ϕ [32], where the contour Γ is the zero level set of ϕ i.e. $\Gamma = \{\mathbf{x}\in\Omega : \phi(\mathbf{x}) = 0\}$, inside$(\Gamma) = \{\mathbf{x}\in\Omega : \phi(\mathbf{x}) > 0\}$ and outside$(\Gamma) = \{\mathbf{x}\in\Omega : \phi(\mathbf{x}) < 0\}$. This model is also suitable for images of intensity inhomogeneity because it is not reliant on intensity information directly, i.e. provided that the inhomogeneity in the image is not too drastic, segmentation results can be good. A downside is that the model is non-convex, and moreover is not necessarily designed for selective segmentation. A good selective result is heavily dependent on initialisation and suitable stopping condition.

More robust selective segmentation methods, such as the models we review from now and in our proposed methods shortly, make use of a set of marker points as described earlier. These points are usually defined by a user clicking on the image, and the points can be used to construct geometric constraints in a model in order to segment the object of interest only. This approach is more reliable than using an edge based model (relying on stopping criteria and good initialisation), as selectiveness is enforced directly in the model.

Liu et al. [28] proposed a selective model using the SaT segmentation paradigm, based on Cai et al. [11], Gout et al. [23] and Badshah and Chen [5]. Their model incorporating $\mathcal{M}$ in term ω is as follows:

$$F_L(u) = \frac{\lambda}{2}\int_\Omega \omega^2(z-u)^2 d\mathbf{x} + \frac{\mu}{2}\int_\Omega |\nabla u|^2 d\mathbf{x} + \int_\Omega |\nabla u| d\mathbf{x}, \tag{8}$$

where $\omega^2 = 1 - gd$, where d and g are distance and edge detection functions respectively, used in [23, 5], given by: $d(\mathbf{x}) = \prod_{i=1}^m \left(1 - e^{-\frac{|\mathbf{x}-\mathbf{x}_i|^2}{2\sigma^2}}\right)$, and $g(|\nabla z(\mathbf{x})|) = \frac{1}{1+\beta_g|\nabla z(\mathbf{x})|^2}$, for $\mathbf{x}\in\Omega, \mathbf{x}_i \in \mathcal{M}$. In this approach, one first solves for u in (8) and then thresholding u gives the selective segmentation. The model is interesting but robustness is an issue due to strong requirements on markers (e.g. close to edges), as observed from [28] and also tests in §4.

The recent approach for selective segmentation by Roberts and Chen [36] makes use of an edge-weighted geodesic distance to achieve selectiveness. Their convex model is:

$$F_{RC}(u) = \int_\Omega g|\nabla u| d\mathbf{x} + \lambda\int_\Omega \left((z-c_1)^2 - (z-c_2)^2\right) u d\mathbf{x} + \theta\int_\Omega \mathcal{D}u d\mathbf{x}, \tag{9}$$

where c_1 and c_2 are the average intensities inside and outside Γ respectively. $\mathcal{D}$ is the geodesic distance defined as $\mathcal{D}(\mathbf{x}) = \mathcal{D}_0(\mathbf{x})/||\mathcal{D}_0(\mathbf{x})||_{L^\infty}$ where:

$$|\nabla\mathcal{D}_0(\mathbf{x})| = \varepsilon + \beta_G|\nabla S^L(z(\mathbf{x}))|^2 + \theta_G\mathcal{D}_E(\mathbf{x}),\ \mathbf{x}\in\Omega, \tag{10}$$

with $\mathcal{D}_0(\mathbf{x}) = 0$ for $\mathbf{x}\in\mathcal{M}$, $\mathcal{D}_E(\mathbf{x})$ is the Euclidean distance from $\mathcal{M}$ and in [36] the parameters are set to be $\varepsilon = 10^{-3}, \beta_G = 1000, \theta_G = 0.1$. The equation (10) offers us a way to prescribe a desirable distance as well as where the edges are expected to be. Here $S^L(z)$ denotes the solution after L cheap fixed point Gauss-Seidel iterations to the anisotropic TV denoising [13, 33] model: $S^L(z) = \arg\min_\varphi \frac{\iota}{2}\int_\Omega(\varphi - z)^2 d\Omega + \hat{\mu}\int_\Omega g(|\nabla z|)|\nabla\varphi| d\Omega$. To solve the Eikonal equation (10) efficient numerical schemes such as fast sweeping [43] or fast marching [40] can be implemented.

2.3. Intensity Inhomogeneity

There are innovative works that apply the CV model indirectly, extending its capability to deal with intensity inhomogeneity (for certain images). In 2016, Ali et al. [2] proposed a variational segmentation model capable of segmenting images with intensity inhomogeneity. They considered the minimising functional $F_{Ali}(\phi, c_1, c_2, d_1, d_2)$, defined by

$$\int_\Omega \Big(\lambda_1(w-c_1)^2+\lambda_2(w^*-w-d_1)^2\Big)H(\phi)\,d\mathbf{x}+\mu\int_\Omega \delta(\phi)|\nabla\phi|\,d\mathbf{x} + \int_\Omega \Big(\lambda_1(w-c_2)^2+\lambda_2(w^*-w-d_2)^2\Big)(1-H(\phi))\,d\mathbf{x}, \tag{11}$$

where $z^* = g_k * z$ with g_k an averaging convolution operator of window size k, and $w = zz^*$. Here, c_1 and c_2 are the average intensities of w inside and outside Γ, and similarly d_1, d_2 are the average intensities of $w^* - w$ inside and outside Γ. The idea is that the new images $w, w^* - w$ are more likely to have a distribution of piecewise-constant intensities, though z may not.

Additionally, Mabood et al. [29] proposed the following selective segmentation functional capable of segmenting objects of intensity inhomogeneity:

$$F_{Mabood}(\phi, a_0, a_1, a_2, b_0, b_1, b_2) = \lambda_1\int_\Omega (z^*-(a_0+a_1x+a_2y))^2H(\phi)\,d\mathbf{x} + \lambda_2\int_\Omega (z^*-(b_0+b_1x+b_2y))^2(1-H(\phi))\,d\mathbf{x} \ +\mu\int_\Omega dg\delta(\phi)|\nabla\phi|\,d\mathbf{x}, \tag{12}$$

where the new image $z^* = \frac{1}{9}\sum_{i=1}^{9} J^E_{\sigma,i}$ is assumed to have a distribution of piecewise-linear intensities, with $J^E_\sigma = K_\sigma * [(z_x, z_y, z)^T(z_x, z_y, z)]$; here the Gaussian kernel K_σ with standard deviation σ is applied point-wise to the 3×3 matrix. The values a_i and b_i are coefficients of linear functions approximating the intensities of z^* inside and outside the contour Γ respectively, where we use the notation $\mathbf{x} = (x, y)$. Requiring the new images (or their transforms) to have a specific form of intensity distribution amounts to demanding the input image z is in a certain class, and so implicitly imposing a condition on robustness.

3. Proposed Models and Implementation

We now present two convex and selective models that are capable of segmenting images with piecewise smooth intensities (not piecewise constants) and designed to overcome the shortcomings of related models. Though building on existing works, our models offer more robustness (i.e. less parameters' sensitivity) than (8), do not require as many marker points from a user, and do not require as strict positions of where these markers should be inside the object. They differ from the less-widely used calibration method (3) in that local selectivity is achieved, not aiming to segment all objects.

3.1. Model 1 by two-stage selective segmentation

Our first proposed selective model is given by the following functional:

$$\min_u F_1(u) = \frac{\lambda}{2}\int_\Omega (z-u)^2\,d\mathbf{x}+\frac{\mu}{2}\int_\Omega |\nabla u|^2\,d\mathbf{x}+\int_\Omega |\nabla u|\,d\mathbf{x}+\theta\int_\Omega \mathcal{D}u\,d\mathbf{x} \tag{13}$$

$$s.t. \quad |\nabla\mathcal{D}(\mathbf{x})| = \varepsilon+\beta_G|\nabla S^L(z(\mathbf{x}))|^2+\theta_G\mathcal{D}_E(\mathbf{x}), \quad u\in W^{1,2}(\Omega) \tag{14}$$

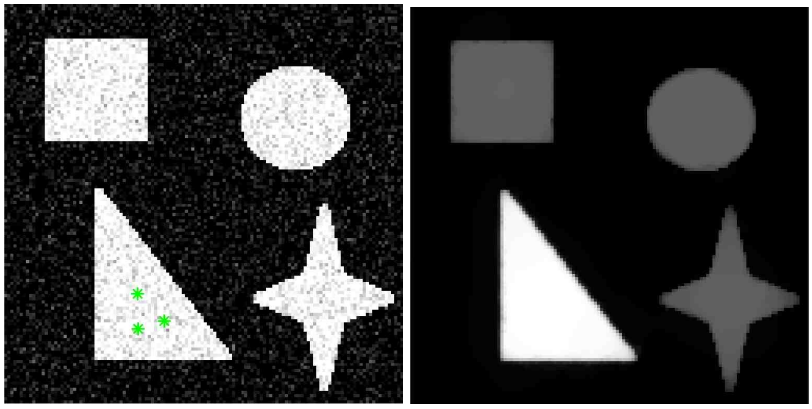
where details of incorporating the geometric prior by computing the the geodesic distance $\mathcal{D}$ can be found in (10). We see that the model approximates the MS equipped with a distance constraint, because if $\theta = 0$ it is reduced to the Cai et al. [11] model i.e. (2) that is based on the original MS model (1) for $u \in W^{1,2}(\Omega)$. Therefore, with the addition of the selective (distance) constraint, our proposed model can be viewed as a selective convex variant of the Mumford-Shah model.

The addition of the term $\int_\Omega \mathcal{D}u\,d\mathbf{x}$ encourages u to be small away from the user input. We see the effect of the model in Figure 1, where we show an example output of our model on a noisy synthetic image. We see that, in addition to denoising the image, the output encourages the values of pixels away from user input $\mathcal{M}$ to be small. This is exactly the effect that adding our selective constraint has. Where $\mathcal{D}$ is small (i.e. close to $\mathcal{M}$) the constraint doesn't apply, whereas where $\mathcal{D}$ is large (i.e. away from $\mathcal{M}$) the constraint is applied and this information is filtered out (i.e. the intensity of the output is encouraged to be smaller). The result is an output in which only the region of interest has notable intensity, and the background can be thresholded out in the second stage.

Once the minimiser u^* is found from solving (14), we can obtain an appropriate segmentation result by thresholding u^* suitably. The binary image representing our segmentation result, Σ, can be defined by $\Sigma = \{\mathbf{x}\in\Omega : u^*(\mathbf{x}) \geq \rho\}$, where $\rho \in [0, 1]$ is the threshold value. Typically, the best results occur when the user selects a value of ρ manually through tuning empirically, as suggested in [11].

In order to state a theory of the existence of a unique minimiser, we impose the constraint $u \geq 0$ on the energy (14) but in practice this is not required. Hence the model takes the form $F_1(u) =$

$$\frac{\lambda}{2}\int_\Omega (z-u)^2d\mathbf{x}+\frac{\mu}{2}\int_\Omega |\nabla u|^2d\mathbf{x}+\int_\Omega |\nabla u|d\mathbf{x}+\theta\int_\Omega \mathcal{D}ud\mathbf{x}+\chi_C(u), \tag{15}$$



(a) Image with user input $\mathcal{M}$. (b) Proposed model output.

Figure 1. Example output of our proposed model on a noisy synthetic image. Here we see the effect of the distance constraint on the output of the model. The background objects (pixels away from $\mathcal{M}$) have low intensity (relative to the region of interest) and can therefore be easily thresholded out to achieve segmentation.

where χ_C is the characteristic function of the set $C = \{u \in W^{1,2} : u \geq 0\}$.

Proposition 1 *Let $z \in L^2(\Omega)$. Then model (15) is strictly convex and there exists a unique minimiser $u(\mathbf{x}) \in W^{1,2}(\Omega)$.*

Proof. To begin, we first show that $0 \leq \inf_{u \in W^{1,2}(\Omega)} F_1(u) < \infty$, as done in [11]. To prove the upper bound, take $u_0 = 0$ and obtain the inequality

$$\inf_{u \in W^{1,2}(\Omega)} F_1(u) \leq F_1(u_0) = \frac{\lambda}{2} \int_\Omega z^2 d\mathbf{x} < \infty$$

since the lower bound is guaranteed by the projection onto the set C.

We know the first three terms of $F_1(u)$ are convex and lower semi-continuous from [11]. Additionally, we know from [36] that the distance term (linear) is convex, and the set C is a convex set, so $F_1(u)$ is fully convex. To show $F_1(u)$ is lower semi-continuous, we need to show that the final two terms are lower semi-continuous. Clearly the fourth term is lower semi-continuous as $\forall u, u_0 \in W^{1,2}$:

$$\lim_{u \to u_0} \inf \int_\Omega \mathcal{D}u \, d\mathbf{x} = \int_\Omega \mathcal{D}u_0 \, d\mathbf{x}.$$

Finally, the characteristic function of any set is lower semi-continuous if and only if the set is closed. Clearly the set C is closed, and so we can conclude that our model $F_1(u)$ is lower semi-continuous.

Next we prove coercivity by showing that $||u||_{W^{1,2}(\Omega)}$ is bounded by $\sqrt{F_1(u)}$. First we see that $||\nabla u||_{L^2(\Omega)} = \left(\int_\Omega |\nabla u|^2 \, d\mathbf{x} \right)^{\frac{1}{2}} \leq \sqrt{\frac{2}{\mu} F_1(u)}$, and also

$$||u||_{L^2(\Omega)} \leq ||u - z||_{L^2(\Omega)} + ||z||_{L^2(\Omega)} \leq \sqrt{\frac{2}{\lambda} F_1(u)} + ||z||_{L^2(\Omega)}.$$

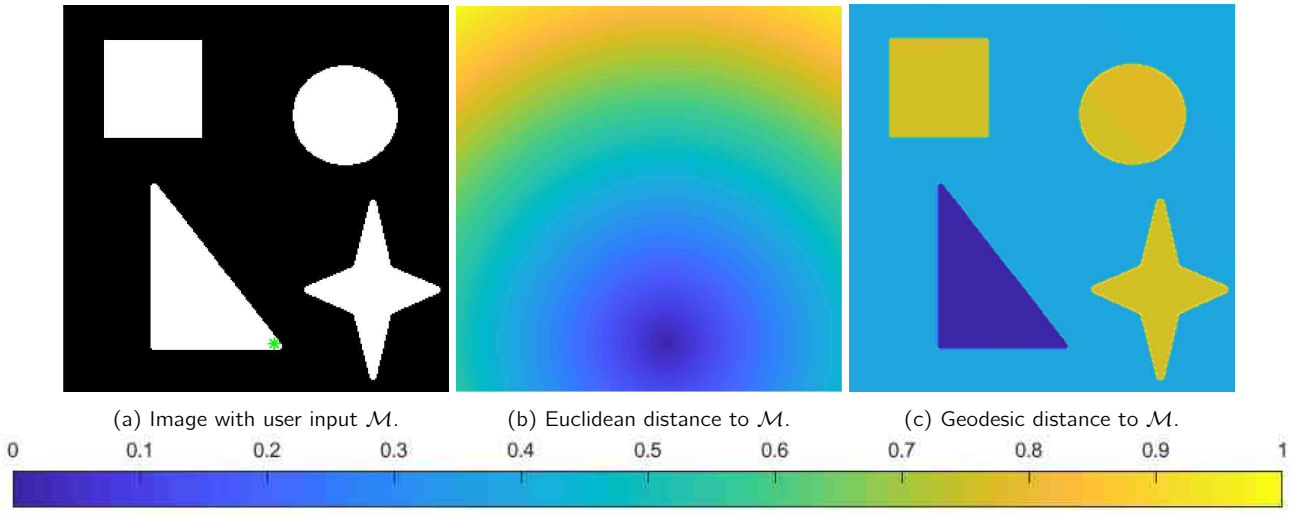
Therefore, combining the above yields

$$\begin{aligned} ||u||_{W^{1,2}(\Omega)} &= ||u||_{L^2(\Omega)} + ||\nabla u||_{L^2(\Omega)} \leq ||u - z||_{L^2(\Omega)} + ||z||_{L^2(\Omega)} + ||\nabla u||_{L^2(\Omega)} \\ &\leq \left(\sqrt{\frac{2}{\lambda}} + \sqrt{\frac{2}{\mu}} \right) \sqrt{F_1(u)} + ||z||_{L^2(\Omega)}, \end{aligned}$$

and so $F_1(u)$ is coercive. Using this fact, combined with the knowledge that $W^{1,2}(\Omega)$ is a reflective Banach space and $F_D(u)$ is convex and lower semi-continuous, then we can conclude that a minimiser exists in $W^{1,2}(\Omega)$. Uniqueness of the minimiser can be implied by the convexity of $F_1(u)$. □

Compared to previous selective segmentation models such as [35, 36], who assume a piecewise-constant distribution, our model can handle non-piecewise constant images, as a key feature of the SaT paradigm. Because of this, we are not limited to a specific class or classes of image. The proposed model is quite capable of segmenting natural, synthetic and medical images.

The model by Liu et al. [28] given in (8) is a selective model that also follows the SaT paradigm, but they use a Euclidean-like distance as a geometric constraint, which restricts the flexibility of the placement of marker points $\mathcal{M}$. Their method requires many marker points, usually spanning the length of the object, and better results come if markers are placed along the perimeter



(a) Image with user input $\mathcal{M}$. (b) Euclidean distance to $\mathcal{M}$. (c) Geodesic distance to $\mathcal{M}$.

Figure 2. Comparison of Geodesic and Euclidean distance constraints.

of the region of interest. As we will show in §4, the use of geodesic distance in our model allows us to be much less strict in terms of the placement of user input. Moreover, overall less user input is required for our model, often only $1-3$ clicks is needed. This is a clear advantage over the models that employ an Euclidean-like distance, such as [28, 41, 5].

See Fig. 2 for a visual comparison of Euclidean and geodesic distances. We see that, with the input placed in the bottom right of the triangle, the Euclidean distance is actually placing a larger constraint on pixels at the top of the triangle when compared with pixels in the left most part of the star. This is not desirable, as this would mean we could not segment the entire triangle without including the star. In contrast, the geodesic distance applies the same constraint (close to 0) for the entire triangle, and so segmentation of the region of interest is easily achievable. Of course with more input spanning the length of the triangle the Euclidean distance would be suitable, however this toy example highlights the downside of a Euclidean-like distance with minimal input, as will be further shown in §4.

In images with drastic intensity inhomogeneity the geodesic distance also requires more input from the user to cover the range of intensities, however this increased requirement is still less than what a Euclidean-like distance would require on the same difficult image.

3.2. Model 2 by inequality constraint

Building on the calibration method (3), we propose a new selective model:

$$F_{PD}(u) = \sup_{\varphi \in K_D} \int_{\Omega\times\mathbb{R}} \varphi D\mathbb{1}_u, \tag{16}$$

where K_D is a convex set such that:

$$K_D = \Big\{ \varphi \in C_0(\Omega\times\mathbb{R}) : \varphi^t(\mathbf{x},t) \geq \frac{\varphi^x(\mathbf{x},t)^2}{4} - \frac{\lambda}{2}(t-z)^2 - \theta\mathcal{D}t, \\ \Big| \int_{t_1}^{t_2} \varphi^x(x,s)ds \Big| \leq \nu, \forall \mathbf{x}\in\Omega, t, t_1, t_2 \in \mathbb{R} \Big\}, \tag{17}$$

where θ is a parameter and $\mathcal{D}$ is the geodesic distance constraint defined in (10). The set K_D is similar to the set K defined in (4), but includes the selective constraint. We can show that the above formulation satisfies, similar to (6),

$$\int_{\Omega\times\mathbb{R}} \varphi D\mathbb{1}_u \leq \frac{\lambda}{2}\int_\Omega (z-u)^2 d\mathbf{x} + \int_{\Omega\setminus\Gamma} |\nabla u|^2 d\mathbf{x} + \nu\mathcal{H}^1(\Gamma) + \theta\int_\Omega \mathcal{D}u d\mathbf{x}.$$

Similarly to model 1, segmentation is achieved by thresholding the minimiser from the model.

3.3. Numerical Solutions

The above two models are presented as nonlinear optimization. To solve them there are two general approaches: (i) discretize the functional as a nonlinear function for discrete variables; (ii) find the Euler-Lagrange equations and then discrete the partial differential equations. Below we take the second approach for both models.

To implement **Model 1** by (14), we adapt the Chambolle-Pock primal dual method in [15] designed for convex and non-smooth functionals. Our model (14) in the discrete setting is given as:

$$\min_u ||\nabla u||_1 + \frac{\lambda}{2}||z-u||_2^2 + \frac{\mu}{2}||\nabla u||_2^2 + \theta\langle u, \mathcal{D}\rangle,$$

where $||\nabla u||_1 = \sum_{(i,j)\in\Omega}\sqrt{(\nabla_x u)_{i,j}^2 + (\nabla_y u)_{i,j}^2}$ and we use backward difference to approximate the discrete gradient operator ∇ (so $u, \mathcal{D}$ are now in matrix form). Following [15], we pose the model into a primal-dual formulation as follows:

$$\min_u \max_p -\langle u, \mathrm{div} p\rangle + \frac{\lambda}{2}||z-u||_2^2 + \frac{\mu}{2}||\nabla u||_2^2 + \theta\langle u, \mathcal{D}\rangle - \chi_P(p),$$

where p is the dual variable and the convex set P is given by $P = \{p : ||p||_\infty \le 1\}$, where $||p||_\infty = \max_{i,j}|p_{i,j}|$, and $|p_{i,j}| = \sqrt{(p_x)_{i,j}^2 + (p_y)_{i,j}^2}$ (with $p = (p_x, p_y)$). The function χ_P is the characteristic function of P.

The primal-dual algorithm is given as follows:

$$\begin{cases} \hat{p}^k = p^k + \sigma\nabla\bar{u}^k, & p_{i,j}^{k+1} = \frac{\hat{p}_{i,j}^k}{\max(1,|\hat{p}_{i,j}|)}, \\ \hat{u}^k = u^k - \tau\nabla^* p^{k+1}, & u^{k+1} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}\left(\hat{u}^k + \tau(\lambda z - \theta\mathcal{D})\right)}{\tau\lambda + 1 + \tau\mu\mathcal{F}(\nabla^*\nabla)}\right), \\ \bar{u}^{k+1} = u^{k+1} + \eta(u^{k+1} - u^k), \end{cases}$$

where $\mathcal{F}$ is the fast Fourier transform operator, $\tau = 0.01$ and $\sigma = 1/(8\tau)$ are fixed time-step parameters, and $\eta = 0.9$ is a fixed parameter.

For **Model 2**, we utilise the primal-dual algorithm proposed by Pock et al. [34], where instead of projecting onto the convex set K defined in (4), we project onto the convex set K_D enforcing the selective constraint defined in (17). To introduce the details, we first replace $\mathbb{1}_u(\mathbf{x}, t)$ from (3) by a generic function $v(\mathbf{x}, t) : \Omega \times \mathbb{R} \to [0, 1]$ such that

$$\lim_{t\to-\infty} v(\mathbf{x}, t) = 1, \qquad \lim_{t\to+\infty} v(\mathbf{x}, t) = 0.$$

Thus, we solve the following problem:

$$\min_v \sup_{\varphi\in K_D} \int_{\Omega\times\mathbb{R}} \varphi Dv. \tag{18}$$

Following the work of [34], we discretise on the grid and still denote by v and φ respectively as their discretised versions. Then, we write (18) as the following primal-dual problem $\min_{v\in C}\max_{\varphi\in K_D}\langle Av, \varphi\rangle$, where A is the discrete gradient operator and $C = \{v : v \in [0, 1]\}$. The algorithm for solving this, as presented by Pock et al., is given as the following iterative system:

$$\begin{cases} \varphi^{k+1} &= \Pi_{K_D}(\varphi^k + \sigma A\bar{v}^k), \\ v^{k+1} &= \Pi_C(v^k - \tau A^*\varphi^{k+1}), \\ \bar{v}^{k+1} &= 2v^{k+1} - v^k, \end{cases}$$

where $\tau, \sigma > 0$ are two time-steps. Projection onto the set C is achieved by a simple clipping operation. Projection onto the set K_D is more difficult as it is the intersection of two convex sets involving local and non-local constraints. In order to do this, we use Dykstra's projection algorithm [7]. To handle the first constraint in K_D, we project onto a parabola imposed by the constraint using Newton's algorithm. To handle the second non-local constraint in K_D, we use a soft shrinkage scheme ([21] which refers to the analytical solution of $\min_x \|x-y\|_2^2/2 + \gamma\|x\|_1$.

4. Numerical Experiments

In this section we conduct numerical experiments in order to show the advantages of our proposed models (**Model 1** as **Model 2**) comparing with five other models discussed in §2, namely:

- Ali [2] – The model by Ali et al. [2] given in (11).
- Mabood [29] – The model by Mabood et al. [29] given in (12).
- DRLSE [26] – The DRLSE model by Li et al. [26] given in (7).
- Roberts [36] – The model by Roberts and Chen [36], given in (9).
- Liu [28] – The model by Liu et al. [28] given in (8).

The first three models on the list are non-convex, implementing level set methods, whereas the latter two and ours are convex. In all experiments we supply the methods with a set of marker points $\mathcal{M}$ input by the user. These marker points are used to design

geometric constraints for the various methods, and we also use them as initialisation when performing the optimisation; this is particularly important for the non-convex methods to achieve a desired result in the region of interest. We use the numerical algorithms discussed in the original works to implement each of the models.

The structure of the experiments is as follows: we initially show our models as well as Liu's model on relatively simple images. This is to demonstrate the intermediate steps of the SaT segmentation paradigm for selective segmentation. Here, we also highlight the advantages of our methods in terms of the robustness of user input. We remark here that while both of our proposed models show advantages, we prefer **Model 1** as **Model 2** is computationally expensive. Then we show some qualitative results on some images (both medical and natural) displaying some intensity inhomogeneity. Finally we show some quantitative comparisons for a dataset where ground truth is available.

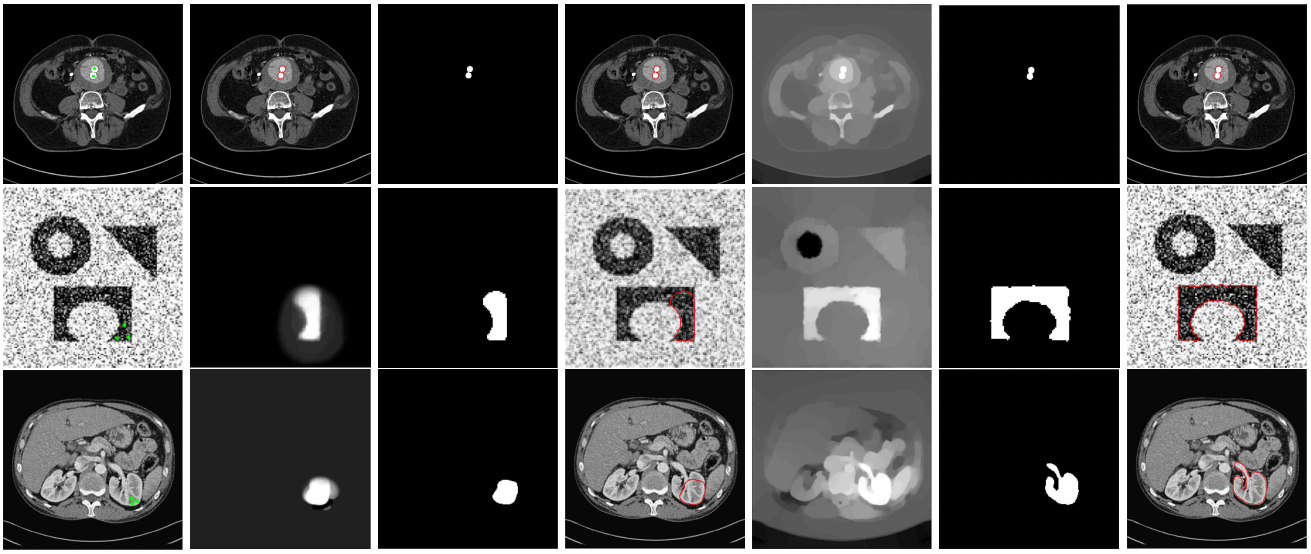


Figure 3. An intuitive display of the different steps of models employing the "smoothing and thresholding" paradigm for selective segmentation: Liu's [28] and our Model 1. In the first column we show the image with the user input $\mathcal{M}$. Column two shows the output u of Liu's model. Column three shows binary segmentation result of Liu's after thresholding. Column four shows the binary segmentation overlaid on the original image. Similarly, columns five, six and seven show: the output of our Model 1, the result after thresholding and the result overlaid on the original image respectively. The examples in the second and third row demonstrate how our approach is more robust to the placement of user input than Liu's.

All experiments were conducted in MATLAB R2021a on a PC with AMD Ryzen 5 3600 3.6 GHz; DDR4 16GB @ 3200 MHz. The user input (marker set $\mathcal{M}$) was taken manually in each case using the `roipoly` command[†].

4.1. Visual Comparisons of SaT models

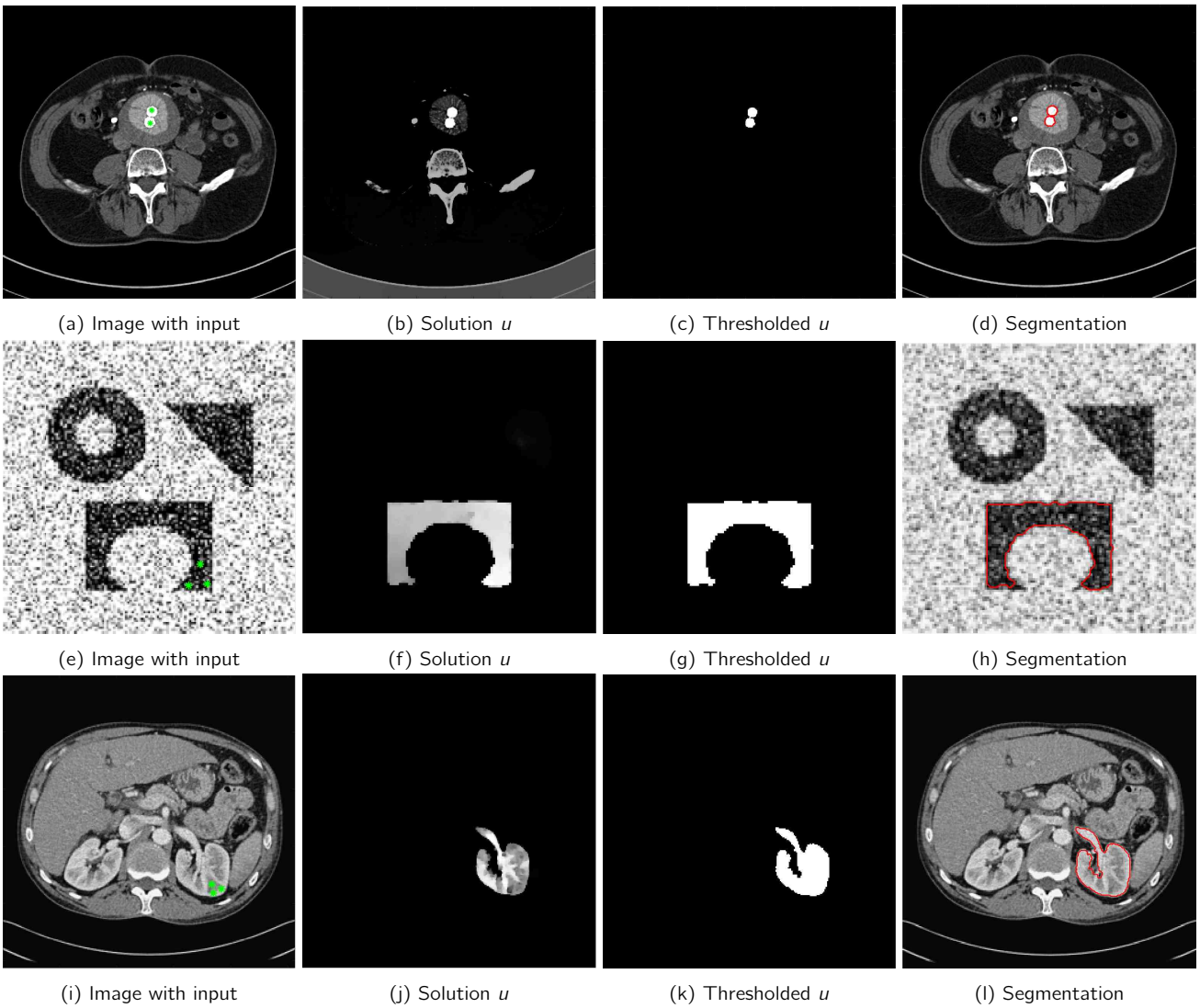
To begin the experiments, we show in Figure 3 a comparison between the model by Liu et al. [28] and our **Model 1**. Both use the SaT segmentation paradigm, and so we have displayed the intermediate steps of each model to given an understanding of how a segmentation result is acquired. To be clear, for each example we show the input to both of the models, then show three steps for each model: the output u of the model, the segmentation output after thresholding (a binary image), and the segmentation overlaid on the original image.

As a comparison, we show the examples from Figure 3 in the **Model 2** framework in Figure 4. We note that while providing similar results, in comparison to the **Model 2** method, the **Model 1** method is mutltitudes faster (we can expect a 256×256 image to take just seconds to process for **Model 1**, however **Model 2** may take a couple of hours for a single image), and so we focus our attention in the rest of the experiments to just **Model 1** method.

We also compare both of our **Model 1** with Liu's in Figure 5. Figures 3 and 5 highlight the advantages of our model compared with Liu's in terms of robustness of user input. Liu's model requires input to span the length of the object for good results, whereas our method requires minimal clicking at any point in the object. This is particularly highlighted in both of the examples in Figure 3 where we have input the marker set $\mathcal{M}$ in the corner of the objects. Clearly the model by Liu struggles to identify the entire object, whereas our **Model 1** is capable of segmenting the full object, because of the geodesic distance constraint.

Remark 1 *For the second example in Figure 3, we note that the object selected has intensity close to* 0. *For the selective variant of the SaT paradigm, it does not make sense to minimise the relevant function to obtain a u where the desired object has a value close to* 0. *When the object of interest has low intensity we replace z as the input image by $1 - z$, where $z \in [0, 1]$.*

[†]Examples may be found from https://www.liverpool.ac.uk/~cmchenke/softw.htm



(a) Image with input (b) Solution u (c) Thresholded u (d) Segmentation

(e) Image with input (f) Solution u (g) Thresholded u (h) Segmentation

(i) Image with input (j) Solution u (k) Thresholded u (l) Segmentation

Figure 4. An intuitive display of the different steps of our Model 2 on examples first presented in Figure 3. First column: the input image. Second column: the 0.5 isosurface of solution u of the model. Third column: The binary segmentation output after thresholding. Fourth column: Segmentation result overlaid onto the original image.

4.2. Qualitative Comparison of Images

Here, we now compare our **Model 1** with the five models detailed at the beginning of this section.

First we show two examples of Liver segmentation in Figure 6. The example in the top row is relatively easy, with edges well defined and only slight inhomogeneity. We intentionally placed input in a narrow region at the top to demonstrate robustness against placement of marker set compared with the model by Liu again. Despite the Liu model struggling, the model by Mabood and the DRLSE model perform well, largely due to well defined edges. However, in the example in the second row we see that none of the competing models successfully segment the liver with the drastic inhomogeneity. The models by Ali and Mabood fail to segment the inhomogeneous part and are susceptible to leaking to the nearby object in the upper right of the liver. The DRLSE model stops at the artificial edge created by the dark region on the organ, and the Roberts model who assume a piecewise constant intensity distribution fails to capture the intricacies of the region of interest. Moreover, the Liu model fails to capture the liver without leaking to a nearby object. We remark that the Liu model would be successful if given more precise input, however given the limited marker set it fails to effectively segment the object. Clearly our **Model 1** outperforms the competitors here.

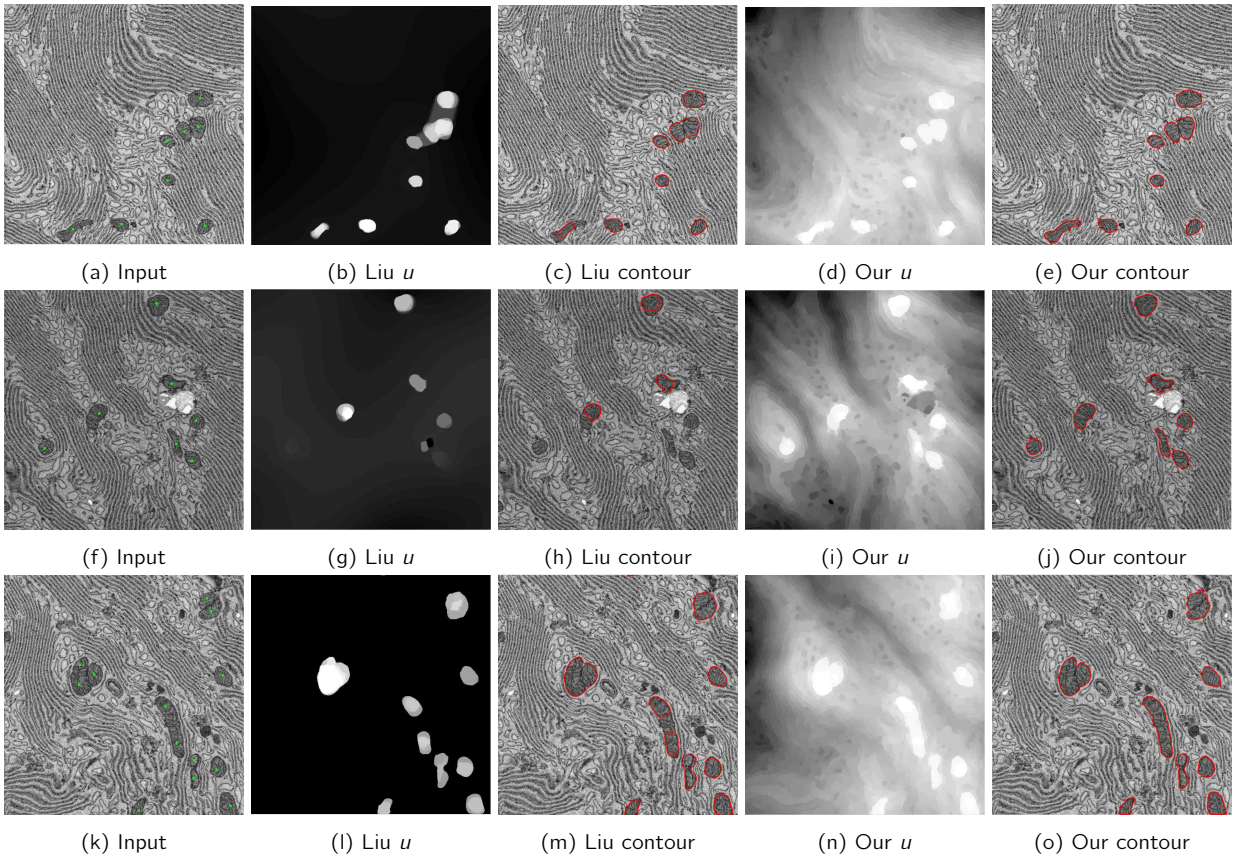


Figure 5. Comparison of our Model 1 with Liu et al. [28], showing advantages of minimal input. Liu's method requires the user input to span the entire length of the object, in all corners of edges whereas ours needs minimal clicking.

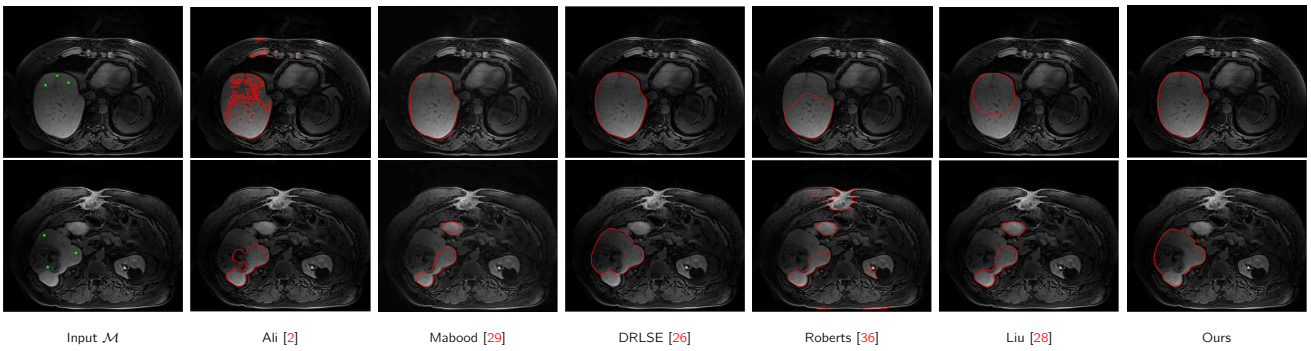


Figure 6. CT liver segmentation. The top example shows only slight inhomogeneity and most models perform decently. In the second example, the competing models fail to adequately segment the liver, whereas our Model 1 performs well. Each image is of size 270 × 320.

In Figure 7, we show some examples of lung segmentation of CT scans of patients with COVID-19 [‡]. The images display inhomogeneity when segmenting the whole lung, due to scarring tissue. We find that the models by Ali and Mabood struggle to effectively segment the lung well. The DRLSE model performs better, though we see that given the initialisation, the curve usually evolves past the boundary of the lung at an area of low contrast before completely segmenting the lung in all places. The Roberts model, interestingly, struggles to segment these images. When aiming to segment the lung the intensity of the region of interest is similar to the intensity of the background (i.e. $c_1 \approx c_2$), and so the CV fidelity has little to no effect, thus Roberts model on this class of images was temperamental (in fact, in [37], they reformulate the CV term to be better suited to cases like this). The Liu model almost segments the lungs well, however fails to capture all of the object due to limited input. Here is more evidence that our **Model 1** needs less input, as our result is clearly the best segmentation here.

[‡]Test data from https://coronacases.org/

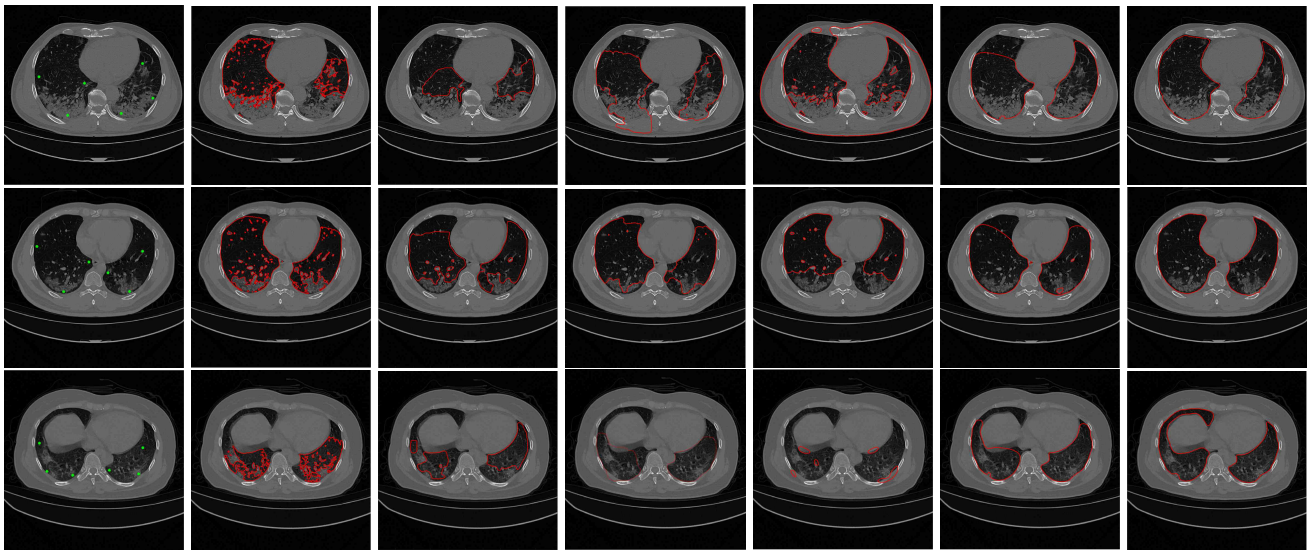


Figure 7. CT lung segmentation of lungs with scarring tissue, providing some inhomogeneity. Each image is of size 512×512.

We now show some results on natural images, displayed in Figure 8. The image on the top row of the mushroom presents some intensity inhomogeneity over the whole object. The results from Ali and Mabood show the segmentation struggling to segment the stalk. The DRLSE method stops at the edge between the cap and the stalk, and the input provided for the Liu method is not suitable to segment the entire object. The Roberts model does well, but due to the limited CV fidelity, it does not segment the darker part of the cap of the mushroom, whereas our model segments the object well.

Similarly for the vase and the dog, inhomogeneity is present throughout the region of interest. For both Ali and Mabood, we either tune the fidelity parameters to be too restrictive within the object and no leakage outside the object, or make the fidelity too dominant and thus segment regions outside the object of interest. The DRLSE is susceptible to areas of low contrast, for example at the bottom of the vase either side where it begins to curve. The Roberts model is unable to segment the objects due to the restriction of piecewise-constant intensities, and as before the Liu model struggles due to limited input.

The last two rows show segmentation of sand dunes. These were selected due to the piecewise-smooth region of the dunes, and the low relatively low contrast at the bottom right at each of the dunes. Only the Roberts model is comparable here, and yet that model fails to segment the darker region in the bottom left of the second image, due to the fidelity.

4.3. Quantitative Comparison

Finally, we evaluate our model on images from both the BraTS dataset [30] and the LiTS dataset [6], both providing ground truth segmentations hand drawn by experts[§].

We assess the segmentation accuracy using a number of metrics given in terms of the four quantities: True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). These quantities are defined using the segmentation output of the relevant model, Σ and the ground truth, GT.

- The DICE coefficient [20]: $DICE(\Sigma, GT) = \dfrac{2TP}{2TP + FP + FN}$.
- The Jaccard coefficient: $JAC(\Sigma, GT) = \dfrac{TP}{TP + FP + FN}$.
- Segmentation accuracy $ACC(\Sigma, GT) = \dfrac{TP + TN}{FN + FP + TP + TN}$.
- Matthews correlation coefficient:

$$MCC(\Sigma, GT) = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}.$$

For DICE, Jaccard and Accuracy, a value of 1 is a perfect match and the quality of the segmentation result as the value reduces towards 0. For MCC, a perfect result is 1, and the minimum value possible is -1.

We run a comparison on 51 images from the BraTS dataset, and show a few select results in Figure 9, where we see our model outperform the competing methods, particularly on the final three selected images. The DICE scores for these images

[§]Data from https://www.kaggle.com/datasets/awsaf49/brats2020-training-data and https://www.kaggle.com/datasets/andrewmvd/liver-tumor-segmentation

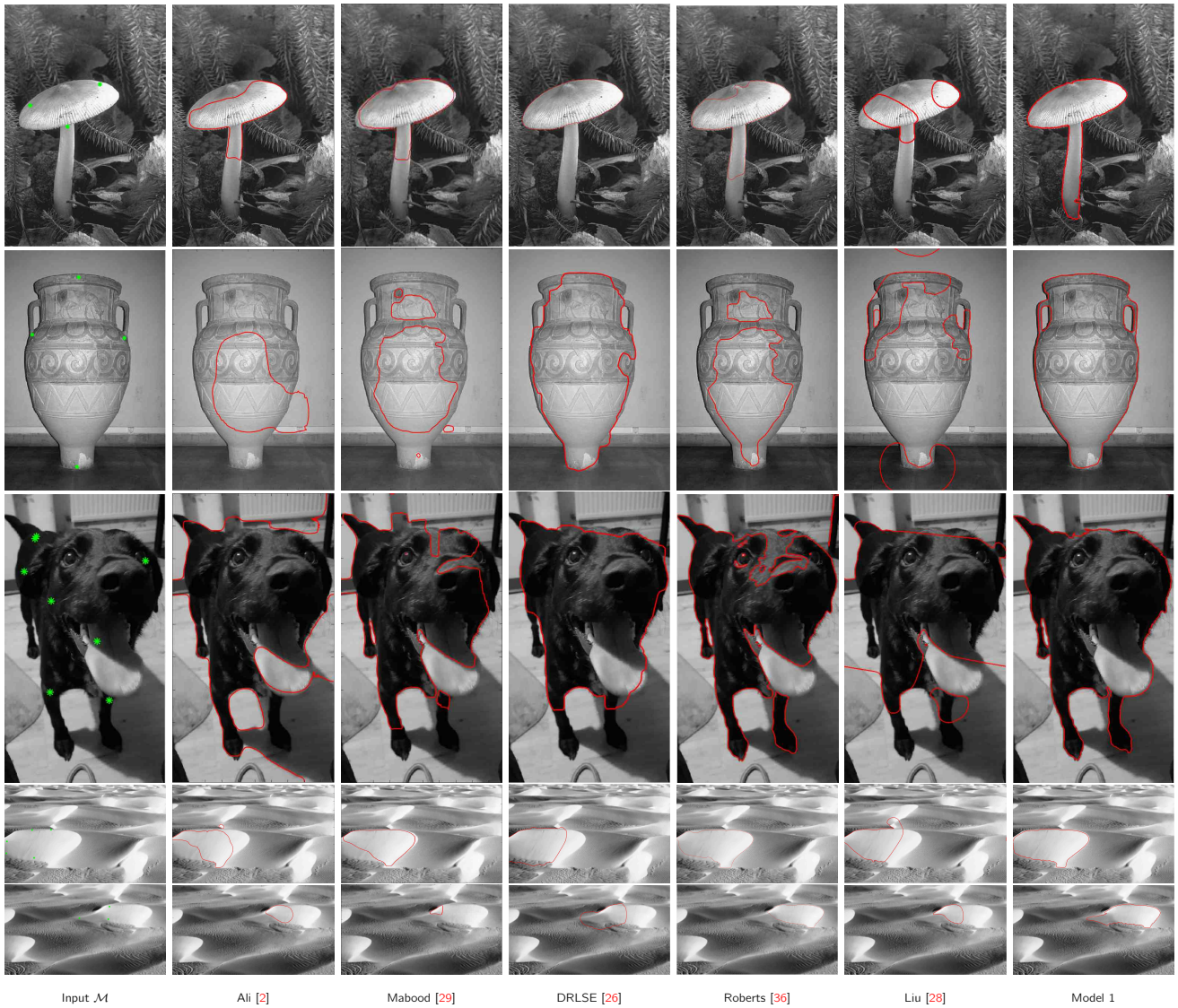


Figure 8. Natural Images. The sizes of each image are: Mushroom - 481 × 321, Vase - 481 × 321, Dog - 882 × 490, Sand Dune 1 - 573 × 942, Sand Dune 2 - 406 × 730

	Ali	Mabood	DRLSE	Roberts	Liu	**Model 1**
BraTS 1	0.955	0.943	0.946	0.961	0.967	**0.969**
BraTS 2	0.913	0.903	0.839	0.858	0.932	**0.955**
BraTS 3	0.804	0.865	0.819	0.815	0.821	**0.937**
BraTS 4	0.883	0.740	0.838	0.898	0.873	**0.904**
BraTS 5	0.853	0.840	0.662	0.878	0.849	**0.951**

Table 1. DICE scores of all the BraTS images found in Figure 9

can be found in Table 1, and a full quantitative display of all metrics applied to the 51 images can be found in Table 2, where we show the mean and standard deviation scores for each model, and for each metric. Moreover we display the mean time took for each model, where all images are of size 240 × 240.

Similarly, we also run a comparison on 41 images from the LiTS dataset, showing a few select results in Figure 11. A full quantitative display of all metrics applied to the 41 images can be found in Table 3, where we show the mean and standard deviation scores for each model for each metric. We again display the mean time took for each model, in this dataset all images are of size 512 × 512.

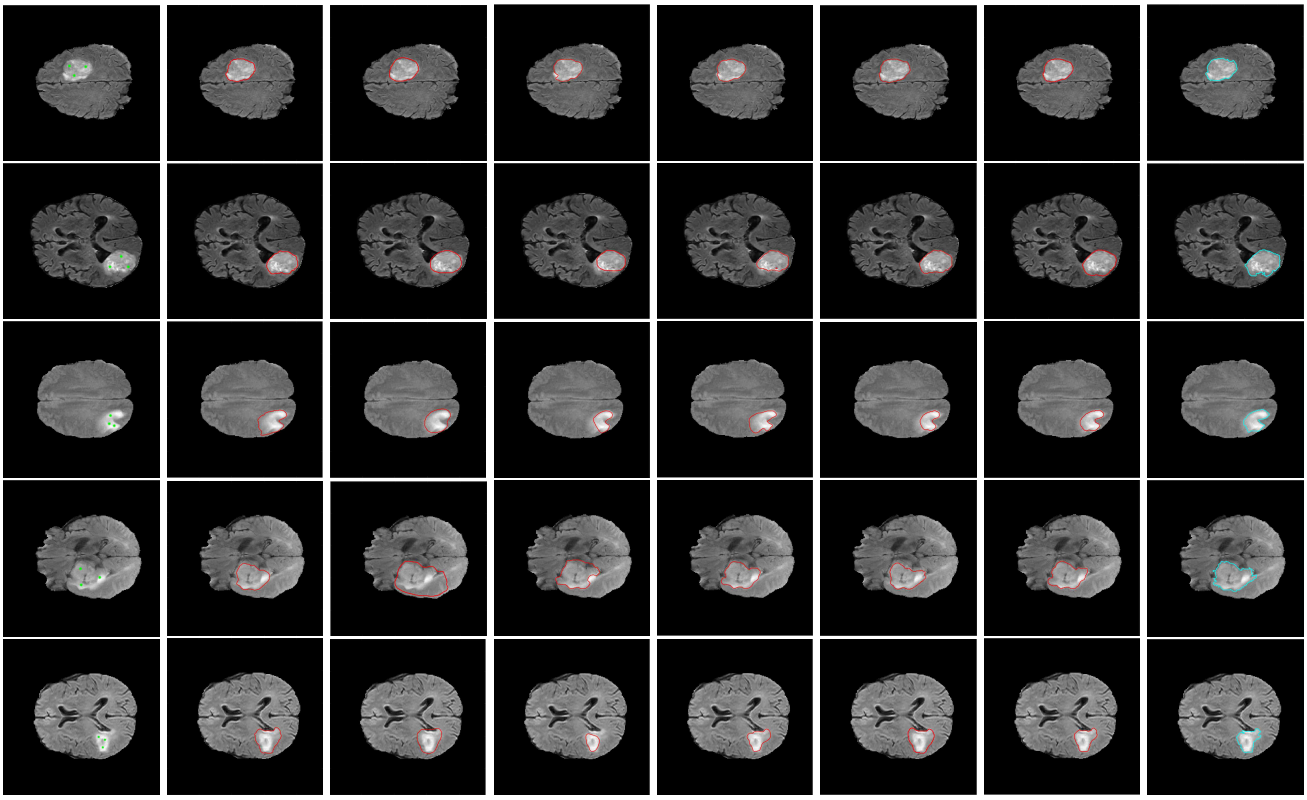


Figure 9. BraTS images: Each line corresponds to a new example, and we show in the first column the user input $\mathcal{M}$, in the middle columns the performance of the models, and the ground truth in the final column. Quantitative results can be found in Table 1

	time	DICE		Jaccard		"Accuracy"		"MCC"	
	(s)	mean	std	mean	std	mean	std	mean	std
Ali	6.43	0.891	0.094	0.814	0.133	0.993	0.008	0.895	0.083
Mabood	1.83	0.842	0.067	0.732	0.100	0.991	0.006	0.846	0.059
DRLSE	10.9	0.818	0.112	0.706	0.150	0.991	0.006	0.827	0.100
Roberts	28.5	0.808	0.130	0.695	0.162	0.990	0.008	0.821	0.110
Liu	10.9	0.939	0.045	0.889	0.050	0.996	0.003	0.939	0.046
Model 1	0.435	0.951	0.051	0.910	0.078	0.997	0.004	0.949	0.052

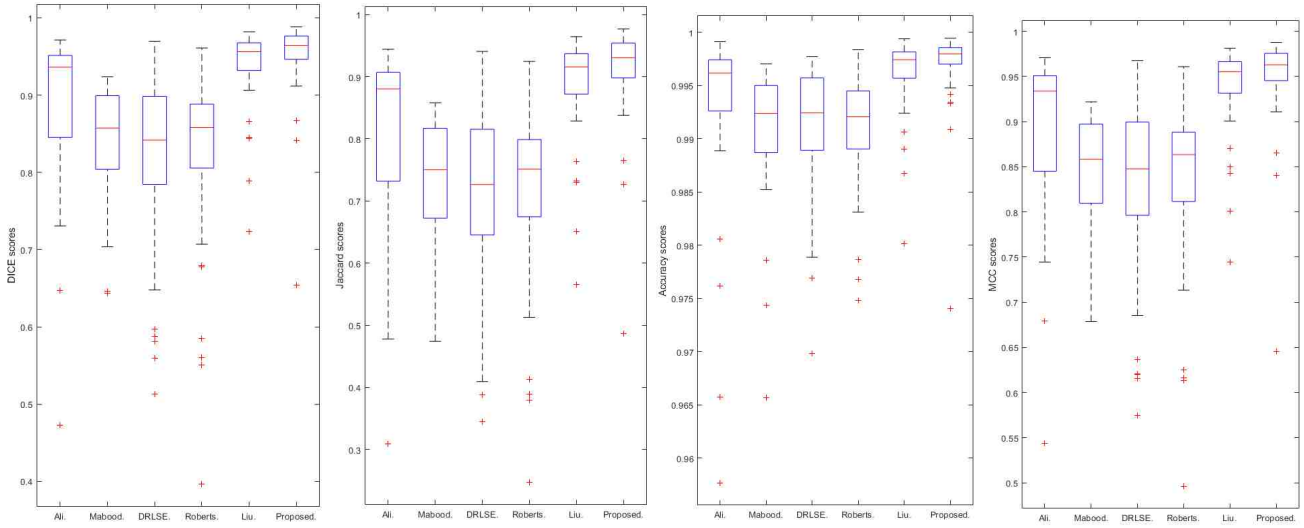
Table 2. Quantitative scores of the models evaluated on 51 images from the BraTS dataset. Here we display the mean DICE, Jaccard, Accuracy and MCC scores for all models, as well as displaying the standard deviations.

	time	DICE		Jaccard		"Accuracy"		"MCC"	
	(s)	mean	std	mean	std	mean	std	mean	std
Ali	29.27	0.564	0.176	0.414	0.170	0.996	0.004	0.613	0.147
Mabood	20.8	0.727	0.246	0.617	0.252	0.997	0.005	0.743	0.232
DRLSE	145.5	0.803	0.100	0.682	0.128	0.999	0.002	0.817	0.083
Roberts	12.3	0.832	0.131	0.729	0.152	0.999	0.001	0.839	0.112
Liu	135.6	0.715	0.225	0.599	0.253	0.996	0.009	0.740	0.200
Model 1	0.807	0.901	0.055	0.824	0.084	0.999	0.001	0.902	0.052

Table 3. Quantitative results on the LiTS dataset, shown for each algorithm the mean time taken to converge, and average and standard deviation values of DICE, Jaccard, Accuracy and MCC of the outputs of each respective model, compared with the ground truth.

4.4. Parameter selection

There are a few parameters that need to be selected in order to achieve a good selective segmentation result. For **Model 1**, we typically fix $\mu = 1$ and vary $\lambda \in [3, 10]$ depending on the noise level. The parameter θ, which controls the degree of selectiveness,



(a) DICE. (b) Jaccard. (c) Accuracy. (d) MCC.

Figure 10. Boxplot of the quantitative figures in Table 2, for the BraTS dataset.

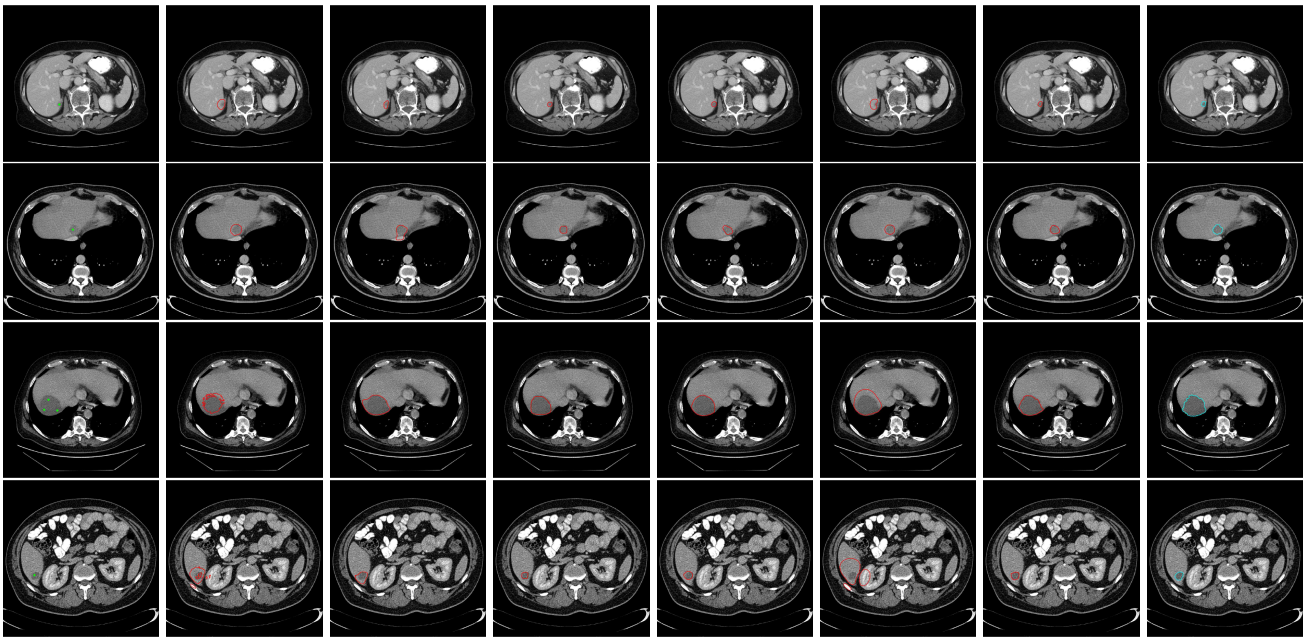
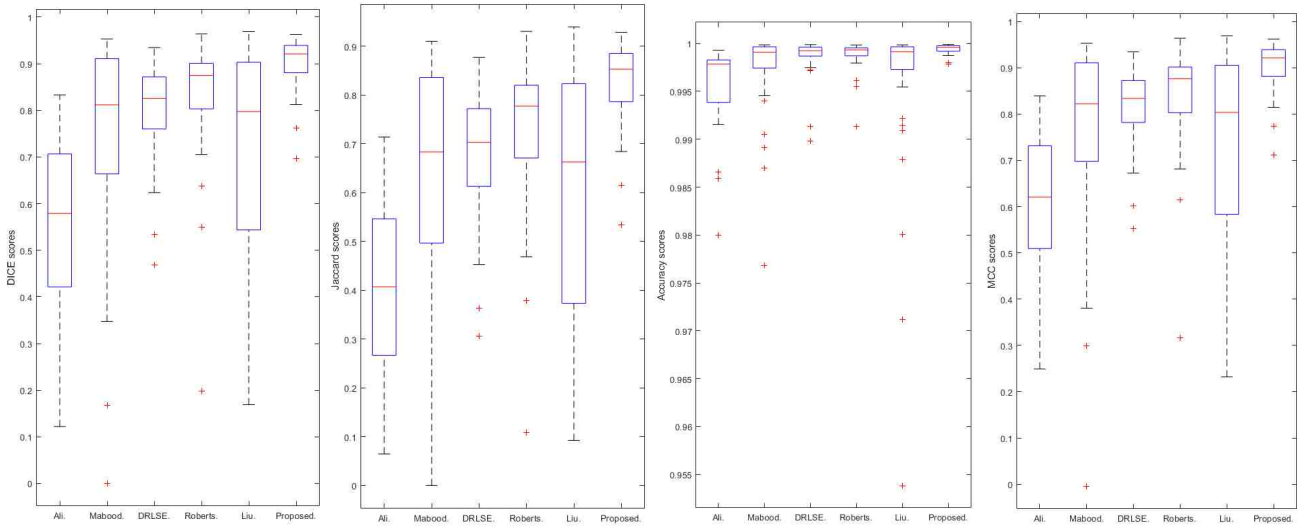


Figure 11. A few select images from the LiTS dataset. We show the input in the first column, the performance of the models in the middle columns, and the ground truth in ithe final column. Quantitative results on the entire LiTS dataset can be found in Table 3

is dependent on how complicated the image is, and how well the edges of the object are well defined. We find best results come from tuning empirically but in most cases the choice is fairly robust. We find $\theta \in [20, 50]$ all give satisfactory and similar results in the vast majority of cases.

A threshold parameter $\rho \in [0, 1]$ needs to be selected to acquire a binary segmentation result. We typically find the best results come from fine tuning manually, however this isn't too much trouble as the most expensive part of the model (optimisation) has already been performed, and a wide range values can easily be checked and the best selected.

In Figure 13, we show some heatmaps to demonstrate the parameter sensitivity on some of the examples already shown earlier in the section. We show the ground truth segmentation in the first and third columns, and in the second and fourth column respectively show a heatmap. The heatmap displays a range of input parameters to our proposed model on the x and y axis, and the colour corresponds to the DICE score. Here we look at the relationship of λ and θ, as μ is fixed as $\mu = 1$. The x-axis corresponds to θ varied from $\theta \in [1, 200]$ in increments of 0.5, the y-axis corresponds to λ varied from $\lambda \in [1, 30]$ in increments of 0.5, and the corresponding colour is the DICE of the output of our proposed model with those parameters; a colour of green



(a) DICE. (b) Jaccard. (c) Accuracy. (d) MCC.

Figure 12. Boxplots of quantitative results on the LiTS dataset. Values can be found in Table 3

is a DICE of 1, whereas red is 0. We see that for a wide range of combinations of λ and θ the DICE score is high, showing that our proposed model is not so sensitive to parameter choice.

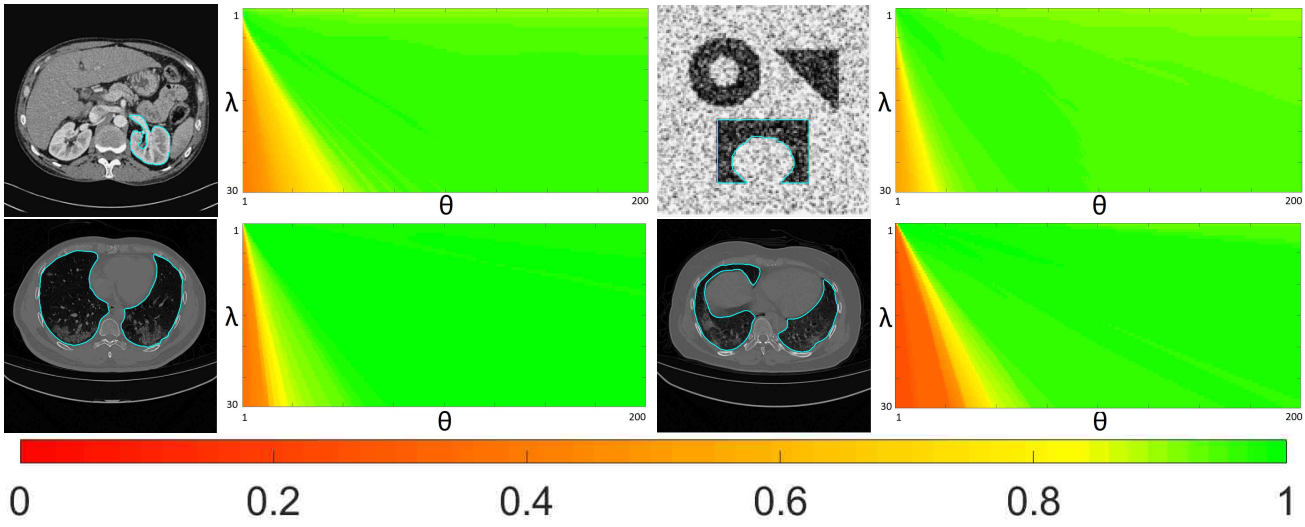


Figure 13. This figure displays parameter analysis for four examples shown earlier in this section.

5. Conclusions

We have proposed two convex and selective segmentation models based on approximation of the Mumford-Shah model. Different from previous works on selective segmentation, the new models requiring much less marker points (often just one point) can segment an image with intensity inhomogeneity. The recommended method is Model 1, a two-stage approach, consisting of first finding the minimiser to the model and then suitably thresholding the minimiser to achieve the segmentation result. For Model 1, we have proved that a unique minimiser exists. Numerical results by adopting a primal-dual algorithm for both models have shown that they are both effective and the new Model 1 outperforms recent novel works in a major way. Future works will generalise such models to a deep learning framework following the ideas of [8].

Acknowledgments

The first author is grateful to the UK EPSRC award EP/W522399/1. The second author is thankful for the UK EPSRC grant EP/N014499/1. The third author thanks the RLBUHT for their approval to participate this work and for partial funding.

References

1. G. Alberti, G. Bouchitté, and G. Dal Maso. The calibration method for the mumford-shah functional and free-discontinuity problems. *Calculus of Variations and Partial Differential Equations*, 16(3):299–333, 2003.
2. H. Ali, N. Badshah, K. Chen, and G. A. Khan. A variational model with hybrid images data fitting energies for segmentation of images with intensity inhomogeneity. *Pattern Recognition*, 51:27–42, 2016.
3. L. Ambrosio and V. M. Tortorelli. Approximation of functional depending on jumps by elliptic functional via γ-convergence. *Communications on Pure and Applied Mathematics*, 43(8):999–1036, 1990.
4. L. Ambrosio and V. M. Tortorelli. On the approximation of free discontinuity problems. *Boll. Un. Mat. Ital.*, 6-B:105–123, 1992.
5. N. Badshah and K. Chen. Image selective segmentation under geometrical constraints using an active contour approach. *Communications in Computational Physics*, 7(4):759, 2010.
6. P. Bilic et al. The liver tumor segmentation benchmark (lits). *arXiv preprint arXiv:1901.04056*, 2019.
7. J. P. Boyle and R. L. Dykstra. A method for finding projections onto the intersection of convex sets in hilbert spaces. In *Advances in order restricted statistical inference*, pages 28–47. Springer, 1986.
8. L. Burrows, K. Chen, and F. Torella. On new convolutional neural network based algorithms for selective segmentation of images. In *23rd Annual Conference on MIUA*, pages 93–104. Springer, 2020.
9. L. Burrows, W. Guo, K. Chen, and F. Torella. Reproducible kernel hilbert space based global and local image segmentation. *Inverse Problems & Imaging*, 2020.
10. X. Cai, R. Chan, C.-B. Schonlieb, G. Steidl, and T. Zeng. Linkage between piecewise constant mumford–shah model and rudin–osher–fatemi model and its virtue in image segmentation. *SIAM Journal on Scientific Computing*, 41(6):B1310–B1340, 2019.
11. X. Cai, R. Chan, and T. Zeng. A two-stage image segmentation method using a convex variant of the mumford–shah model and thresholding. *SIAM Journal on Imaging Sciences*, 6(1):368–390, 2013.
12. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79, 1997.
13. F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29(1):182–193, 1992.
14. A. Chambolle and G. Dal Maso. Discrete approximation of the mumford-shah functional in dimension two. *ESAIM: Mathematical Modelling and Numerical Analysis*, 33(4):651–672, 1999.
15. A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging and Vision*, 40(1):120–145, 2011.
16. T. F. Chan, S. Esedoglu, and M. Nikolova. Algorithms for finding global minimizers of image segmentation and denoising models. *SIAM journal on Applied Mathematics*, 66(5):1632–1648, 2006.
17. T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Transactions on Image Processing*, 10(2):266–277, 2001.
18. K. Chen, N. Debroux, and C. L. Guyader. A survey of topology and geometry-constrained segmentation methods in weakly supervised settings. In *Handbook of Mathematical Models and Algorithms in Computer Vision and Imaging (see [19])*, pages 1437–1482, 2023.
19. K. Chen, C.-B. Schonlieb, X.-C. Tai, and L. Younces. *Handbook of Mathematical Models and Algorithms in Computer Vision and Imaging*. Springer, 2023.
20. L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.
21. J. A. Geppert and G. Plonka. Frame soft shrinkage operators are proximity operators. *Applied and Computational Harmonic Analysis*, 57:185–200, 2022.
22. M. Gobbino. Finite difference approximation of the Mumford-Shah functional. *Commu. Pure and Appl. Math.*, 51(2):197–228, 1998.
23. C. Gout, C. Le Guyader, and L. Vese. Segmentation under geometrical conditions using geodesic active contours and interpolation using level set methods. *Numerical Algorithms*, 39(1-3):155–173, 2005.
24. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.
25. C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding. Minimization of region-scalable fitting energy for image segmentation. *IEEE transactions on image processing*, 17(10):1940–1949, 2008.
26. C. Li, C. Xu, C. Gui, and M. D. Fox. Distance regularized level set evolution and its application to image segmentation. *IEEE Trans. Image Process.*, 19(12):3243–3254, December 2010.
27. C. Li et al. A level set method for image segmentation in the presence of intensity inhomogeneities with application to mri. *IEEE transactions on image processing*, 20(7):2007–2016, 2011.
28. C. Liu, M. K.-P. Ng, and T. Zeng. Weighted variational model for selective image segmentation with application to medical images. *Pattern Recognition*, 76:367–379, 2018.
29. L. Mabood, H. Ali, N. Badshah, K. Chen, and G. A. Khan. Active contours textural and inhomogeneous object extraction. *Pattern Recognition*, 55:87–99, 2016.
30. B. H. Menze et al. The multimodal brain tumor image segmentation benchmark. *IEEE Trans. Medical Imaging*, 34(10):1993–2024, 2014.
31. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, 42(5):577–685, 1989.
32. S. Osher, R. Fedkiw, and K. Piechor. Level set methods and dynamic implicit surfaces. *Appl. Mech. Rev.*, 57(3):B15–B15, 2004.
33. P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pat. Anal. Mach. Intell.*, 12(7):629–639, 1990.

34. T. Pock, D. Cremers, H. Bischof, and A. Chambolle. An algorithm for minimizing the mumford-shah functional. In *2009 IEEE 12th International Conference on Computer Vision*, pages 1133–1140. IEEE, 2009.
35. L. Rada and K. Chen. Improved selective segmentation model using one level-set. *Journal of Algorithms & Computational Technology*, 7(4):509–540, 2013.
36. M. Roberts, K. Chen, and K. L. Irion. A convex geodesic selective model for image segmentation. *Journal of Mathematical Imaging and Vision*, 61(4):482–503, 2019.
37. M. Roberts and J. Spencer. Chan-Vese reformulation for selective image segmentation. *J. Math. Imaging and Vision*, 61(8):1173–1196, 2019.
38. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1-4):259–268, 1992.
39. O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. Springer, 2019.
40. J. A. Sethian. Fast marching methods. *SIAM rev.*, 41(2):199–235, 1999.
41. J. Spencer and K. Chen. A convex and selective variational model for image segmentation. *Commu. Math. Sciences*, 13(6):1453–1472, 2015.
42. X.-F. Wang, D.-S. Huang, and H. Xu. An efficient local Chan–Vese model for image segmentation. *Pattern Recognition*, 43(3):603–618, 2010.
43. H. K. Zhao. A fast sweeping method for eikonal equations. *Mathematics of Computation*, 74(250):603–627, 2005.

OrCiD: Liam Burrows at https://orcid.org/0000-0002-6910-6693
OrCiD: Ke Chen at https://orcid.org/0000-0002-6093-6623
OrCiD: Francesco Torella at https://orcid.org/0000-0003-0529-7387

