Chan-Vese Reformulation for Selective Image Segmentation

Michael Roberts · **Jack Spencer**

Abstract Selective segmentation involves incorporating user input to partition an image into foreground and background, by discriminating between objects of a similar type. Typically, such methods involve introducing additional constraints to generic segmentation approaches. However, we show that this is often inconsistent with respect to common assumptions about the image. The proposed method introduces a new fitting term that is more useful in practice than the Chan-Vese framework. In particular, the idea is to define a term that allows for the background to consist of multiple regions of inhomogeneity. We provide comparitive experimental results to alternative approaches to demonstrate the advantages of the proposed method, broadening the possible application of these methods.

1 Introduction

Image segmentation is an important application of image processing techniques in which some, or all, objects in an image are isolated from the background. In other words, for an image $z(\boldsymbol{x}) \in \mathbb{R}^2$, we find the partitioning of the image domain $\Omega \subset \mathbb{R}^2$ into subregions of interest. In the case of two-phase approaches this consists of the foreground domain Ω_F and background domain Ω_B, such that $\Omega = \Omega_F \cup \Omega_B$. In this work we concentrate on approaching this problem with variational methods, particularly in cases where

M. Roberts
Centre for Mathematical Imaging Techniques, Department of Mathematical Sciences, University of Liverpool
E-mail: michael.roberts@liverpool.ac.uk

J. Spencer
Translational Research Exchange @ Exeter, Living Systems Institute, University of Exeter
E-mail: j.a.spencer@exeter.ac.uk

user input is incorporated. Specifically, we consider the convex relaxation approach of [14,8] and many others. This consists of a binary labelling problem where the aim is to compute a function $u(x) \in \{0, 1\}$ indicating regions belonging to Ω_F and Ω_B, respectively. This is obtained by imposing a relaxed constraint on the function, $u \in [0, 1]$, and minimising a functional that fits the solution to the data with certain conditions on the regularity of the boundary of the foreground regions.

We will first introduce the seminal work of Chan and Vese [15], a segmentation model that uses the level set framework of Osher and Sethian [30]. This approach assumes that the image z is approximately piecewise-constant, but is dependent on the initialisation of the level set function as the minimisation problem is nonconvex. The Chan-Vese model was reformulated to avoid this by Chan et al. [14], using convex relaxation methods, that has the following data fitting functional

$$f_{CV}(u) = \int_{\Omega} (\lambda_1 f_1(\boldsymbol{x}) - \lambda_2 f_2(\boldsymbol{x})) u(\boldsymbol{x}) \, \mathrm{d}\Omega, \tag{1}$$

where $f_1(\boldsymbol{x})$ and $f_2(\boldsymbol{x})$ are data fitting terms indicating the foreground and background regions, respectively. In particular, in [15] and [14] these are given by

$$f_1(\boldsymbol{x}) = |z - c_1|^2, \quad f_2(\boldsymbol{x}) = |z - c_2|^2. \tag{2}$$

It should be noted that it is common to fix $\lambda = \lambda_1 = \lambda_2$. The data fitting functional is balanced against a regularisation term. Typically, this penalises the length of the contour. This is represented by the total variation (TV) of the function [15,36], and is sometimes weighted by an edge detection function $g(s) = 1/(1 + \beta s^2)$ [8,32,34,38]. Therefore, the regularisation term is given as

$$TV_g(u) := \int_{\Omega} g(|\nabla z(\boldsymbol{x})|)|\nabla u| \, \mathrm{d}\Omega. \tag{3}$$

Fig. 1: CT image with ground truth segmentation shown (green) and associated average intensity values (c_1 and c_2).
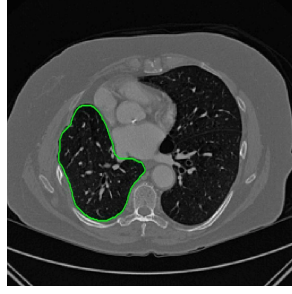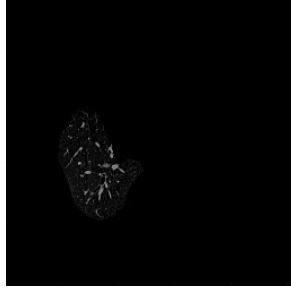
(i) Image with ground truth   (ii) Foreground, $c_1 = 0.08$ (iii) Background, $c_2 = 0.19$

The convex segmentation problem, assuming fixed constants c_1 and c_2, is then defined by

$$\min_{u \in [0,1]} \left\{ F_{CV}(u, c_1, c_2) = TV_g(u) + f_{CV}(u, c_1, c_2) \right\}. \tag{4}$$

In the case where the intensity constants are unknown it is also possible to minimise F_{CV} alternately with respect to u, c_1, and c_2, however, this would make the problem non-convex and hence dependent on the initialisation of u. Functionals of this type have been widely studied with respect to two-phase segmentation [8, 14, 15], which is our main interest. Alternative choices of data fitting terms can be used when different assumptions are made on the image, z. Examples include [1, 2, 16, 25, 39, 41]. We note that multiphase approaches [9, 40] are also closely related to this formulation although in this paper we focus on the two-phase problem due to associated applications of interest. It is also important to acknowledge analogous methods in the discrete setting such as [4, 18, 22, 35]. However, we do not go into detail about such methods here, although we introduce the work of [17] in §3 and compare corresponding results in §7.

In selective segmentation the idea is to apply additional constraints such that user input is incorporated to isolate specific objects of interest. It is common for the user to input marker points to form a set $\mathcal{M}$, where $\mathcal{M} = \{(x_i, y_i) \in \Omega, 1 \le i \le k\}$ and from this we can form a foreground region $\mathscr{P}$ whose interior points are inside the object to be segmented. In the case that $\mathcal{M}$ is provided $\mathscr{P}$ will be a polygon, but any user-defined region in the foreground is consistent with the proposed method. Some examples of selective or interactive methods include [10, 17, 21, 26, 29, 34, 22, 37, 38, 45]. A particular application of this in medical imaging is organ contouring in computed tomography (CT) images. This is often done manually which can be laborious and inefficient and it is often not possible to enhance existing methods with training data. In cases where learning based methods are applicable, the work of Xu et al. [43] and Bernard and Gygli [5] are state of the art approaches. At this stage we define the additional constraints in selective segmentation as follows:

$$f_S(u) = \theta \int_\Omega \mathscr{D}(\boldsymbol{x}) u \, d\Omega, \tag{5}$$

where $\mathscr{D}(\boldsymbol{x})$ is some distance penalty term, such as [33, 34, 38], and θ is a selection parameter. Essentially, the idea is that the selection term $\mathscr{D}(\boldsymbol{x})$ (based on the region $\mathscr{P}$ formed by the user input marker set) should penalise regions of the background (as defined by the data fitting term $f_2(\boldsymbol{x})$) and also pixels far from $\mathscr{P}$. In this paper we choose $\mathscr{D}(\boldsymbol{x})$ to be the geodesic distance penalty proposed in [34]. Explicitly, the geodesic distance from the region $\mathscr{P}$ formed from the marker set is given by:

$$\mathscr{D}_M(\boldsymbol{x}) = 0 \text{ for } \boldsymbol{x} \in \mathscr{P},$$

$$\mathscr{D}_M(\boldsymbol{x}) = \frac{\mathscr{D}_M^0(\boldsymbol{x})}{||\mathscr{D}_M^0(\boldsymbol{x})||_{L^\infty}} \text{ for } \boldsymbol{x} \notin \mathscr{P},$$

where $\mathscr{D}_M^0(\boldsymbol{x})$ is the solution of the following PDE:

$$|\nabla \mathscr{D}_M^0(\boldsymbol{x})| = q(\boldsymbol{x}), \qquad \mathscr{D}_M^0(\boldsymbol{x}_0) = 0, (\boldsymbol{x}_0) \in \mathscr{P}. \tag{6}$$

The function $q(\boldsymbol{x})$ is image dependent and controls the rate of increase in the distance. It is defined as a function similar to

$$q(\boldsymbol{x}) = \varepsilon_{\mathscr{D}} + \beta_G |\nabla z(\boldsymbol{x})|^2, \tag{7}$$

where $\varepsilon_{\mathscr{D}}$ is a small non-zero parameter and β_G is a non-negative tuning parameter. We set the value of $\beta_G = 1000$ and $\varepsilon_{\mathscr{D}} = 10^{-3}$ throughout. Note that if $q(\boldsymbol{x}) \equiv 1$ then the distance penalty $\mathscr{D}_M(\boldsymbol{x})$ is simply the normalised Euclidean distance, as used in [38].

A general selective segmentation functional, assuming homogeneous target regions, is therefore given by:

$$F_S(u, c_1, c_2) = TV_g(u) + f_{CV}(u, c_1, c_2) + f_S(u). \tag{8}$$

Assuming that the optimal intensity constants c_1 and c_2 are fixed, the minimisation problem is then:

$$\min_{u \in [0,1]} F_S(u, c_1, c_2). \tag{9}$$

Again, it is possible to alternately minimise $F_S(u, c_1, c_2)$ with respect to the constants c_1 and c_2 to obtain the average intensity in Ω_F and Ω_B, respectively. However, in selective segmentation it is often sufficient to fix these according to the

user input. In the framework of (9) the Chan-Vese terms [14, 15, 28] have limitations due to the dependence on c_2. In conventional two-phase segmentation problems it makes sense to penalise deviances from c_2 outside the contour, however for selective segmentation we need not consider the intensities outside of the object we have segmented. Regardless of whether the intensity of regions outside the object is above or below c_1, it should be penalised positively. The Chan-Vese terms cannot ensure this as they work based on a fixed "exterior" intensity c_2 and can lead to negative penalties on regions which are outside the object of interest. It is our aim in this paper to address this problem.

The motivation for this work comes from observing contradictions in using piecewise-constant intensity fitting terms in selective segmentation. Whilst good results are possible with this approach, the exceptional cases lead to severe limitations in practice. This is quite common in medical imaging as demonstrated in Fig. 1, where the target foreground has a low intensity. Given that the corresponding background includes large regions of low intensity, the optimal average intensities for this segmentation problem are $c_1 = 0.1534$ and $c_2 = 0.1878$. For cases where $c_1 \approx c_2$, we see that by (1), $f_1 - f_2 \approx 0$ almost everywhere in the domain Ω. This means that it is very difficult to achieve an adequate result, without an over-reliance on the user input or parameter selection.

The central premise for applying Chan-Vese type methods is the assumption that the image approximately consists of

$$z(\boldsymbol{x}) = c_1 \chi_F + c_2 \chi_B + \eta, \qquad (10)$$

where η is noise, χ_i is the characteristic function of the region Ω_i, for $i = F, B$ respectively. The idea of selective segmentation is to incorporate user input to apply constraints that exclude regions classified as foreground, based on their location in the image. We use a distance constraint which penalises the distance from the user input markers. However, a key problem for selective segmentation is that for cases where the optimal intensity values c_1 and c_2 are similar, the intensity fitting term will become obsolete as the contour evolves. This is illustrated in Fig. 3. The purpose of our approach is to construct a model that is based on assumptions that are consistent with the observed image and any homogeneous target region of interest. A common approach in selective segmentation is to discriminate between objects of a similar intensity [33,34,38]. However, the fitting terms in previous formulations [24,33,34,38] aren't applicable in many cases as there are contradictions in the formulation in this context. We will address this in detail in the following section.

In this paper our main contribution is to highlight a crucial flaw in the assumptions behind many current selective segmentation approaches and propose a new fitting term in relation to such methods. We demonstrate how our reformulation is capable of achieving superior results and is more robust to parameter choices than existing approaches, allowing for more consistency in practice. In §2 we give a brief review of alternative intensity fitting terms proposed in the literature, and detail them in relation to selective segmentation. We then briefly detail alternative selective segmentation approaches to compare our method against in §3. In §4 we introduce the proposed model, focussing on a fitting term that allows for significant intensity variation in the background domain. In §5 we discuss the implementation of each approach in a convex relaxation framework, provide the algorithm in §6, and detail some experimental results in §7. Finally, in §8 we give some concluding remarks.

2 Related Approaches

Here, we introduce and discuss work that has introduced alternative data fitting terms closely related to Chan-Vese [15]. In order to make direct comparisons, we convert each approach to the unified framework of convex relaxation [14]. It is worth noting that this alternative implementation is equivalent in some respects, but that the results might differ slightly if using the original methods. We are considering these models in the terms of selective segmentation, so all formulations have the following structure:

$$\min_{u \in [0,1]} \left\{ F(u) = TV_g(u) + f_S(u) + f(u) \right\}. \qquad (11)$$

We are interested in the effectiveness of $f(u)$ in this context, which we will focus on next. In particular, we detail various choices of $f(u)$ from the literature that are generalisations of the Chan-Vese approach. In the following we refer to minimisers of convex formulations, such as (11), by u_γ. Here, the minimiser of $F(u)$ is thresholded for $\gamma \in (0,1)$ in a conventional way [14].

2.1 Region-Scalable Fitting (RSF) [25]

The data fitting term from the work of Li et al. [25], known as Region-Scalable Fitting (RSF), consistent with the convex relaxation technique of [14] is given by

$$f_{RSF}(u) = \int_\Omega \left(\lambda_1 f_1(\boldsymbol{x}) - \lambda_2 f_2(\boldsymbol{x}) \right) u \, \mathrm{d}\Omega, \qquad (12)$$

where

$$f_1(\boldsymbol{x}) = \int_\Omega K_\sigma(\boldsymbol{x} - \boldsymbol{y}) |z - h_1(\boldsymbol{x})|^2 \, \mathrm{d}\Omega,$$

$$f_2(\boldsymbol{x}) = \int_\Omega K_\sigma(\boldsymbol{x} - \boldsymbol{y}) |z - h_2(\boldsymbol{x})|^2 \, \mathrm{d}\Omega, \qquad (13)$$

and $K_\sigma(x)$ is chosen as a Gaussian kernel with scale parameter $\sigma > 0$. The RSF selective formulation is then given as follows:

$$F_{RSF}(u) = TV_g(u) + f_S(u) + f_{RSF}(u). \tag{14}$$

The functions $h_1(x)$ and $h_2(x)$, which are generalisations of c_1 and c_2 from Chan-Vese, are updated iteratively by

$$h_1(x) = \frac{K_\sigma(x) * (u_\gamma z)}{K_\sigma(x) * u_\gamma},$$
$$h_2(x) = \frac{K_\sigma(x) * ((1 - u_\gamma) z)}{K_\sigma(x) * (1 - u_\gamma)}. \tag{15}$$

Using the RSF fitting term, any deviations of z from h_1 and h_2 are smoothed by the convolution operator, K_σ. This allows for intensity inhomogeneity in the foreground and background of target objects.

2.2 Local Chan-Vese (LCV) Fitting [41]

Wang et al. [41] proposed the Local Chan-Vese (LCV) model. In terms of the equivalent convex formulation, the data fitting term is given by

$$f_{LCV}(u) = \int_\Omega (f_1(x) - f_2(x)) u \, d\Omega \tag{16}$$

where

$$f_1(x) = \alpha |z - c_1|^2 + \beta |z^* - z - d_1|^2,$$
$$f_2(x) = \alpha |z - c_2|^2 + \beta |z^* - z - d_2|^2, \tag{17}$$

and $z^* = M_k * z$. Here, M_k is an averaging convolution with $k \times k$ window. The LCV selective formulation is then given as

$$F_{LCV}(u) = TV_g(u) + f_S(u) + f_{LCV}(u). \tag{18}$$

The values c_1, c_2, d_1, d_2 which minimise this functional for u_γ are given by

$$c_1 = \frac{\int_\Omega z u_\gamma \, d\Omega}{\int_\Omega u_\gamma \, d\Omega}, \quad c_2 = \frac{\int_\Omega z(1 - u_\gamma) \, d\Omega}{\int_\Omega (1 - u_\gamma) \, d\Omega},$$
$$d_1 = \frac{\int_\Omega (z^* - z) u_\gamma \, d\Omega}{\int_\Omega u_\gamma \, d\Omega}, \quad d_2 = \frac{\int_\Omega (z^* - z)(1 - u_\gamma) \, d\Omega}{\int_\Omega (1 - u_\gamma) \, d\Omega}. \tag{19}$$

The formulation is minimised iteratively. The LCV fitting term that $f_1(x)$ and $f_2(x)$ includes an additional term weighted by the parameters α and β. The principle for the LCV model is that the difference image $z^* - z$ is a higher contrast image than z and a two-phase segmentation on this image can be computed.

2.3 Hybrid (HYB) Fitting [1]

Based on extending the LCV model, Ali et al. [1] proposed the following data fitting term,

$$f_{HYB}(u, c_1, c_2, d_1, d_2) = \int_\Omega (f_1(x) - f_2(x)) u \, d\Omega \tag{20}$$

where

$$f_1(x) = \alpha |w - c_1|^2 + \beta |w^* - w - d_1|^2,$$
$$f_2(x) = \alpha |w - c_2|^2 + \beta |w^* - w - d_2|^2. \tag{21}$$

Here, $z^* = M_k * z$, $w = z^* z$, and $w^* = M_k * w$, with M_k the averaging convolution as used in the LCV model. The values c_1, c_2, d_1, d_2 are updated in a similar way to [41], with further details found in [1]. The authors refer to this approach as the Hybrid (HYB) Model. The HYB selective formulation is then given as

$$F_{HYB}(u) = TV_g(u) + f_S(u) + f_{HYB}(u). \tag{22}$$

The key aim of the HYB model is to account for intensity inhomogeneity in the foreground and background of the image through the product image w. In LCV, the presence of the blurred image z^* in the data fitting term deals with intensity inhomogeneity, whilst including z helps identify contrast between regions. The authors found that the product image $w = z^* z$ can improve the data fitting in both respects. Therefore they construct a LCV-type function with w rather than the original z. Their results suggest that this approach is more robust.

2.4 Generalised Averages (GAV) Fitting [2]

Recently, Ali et al. [2] proposed using the data fitting terms of Chan-Vese in a signed pressure force function framework [45]. They refer to this approach as Generalised Averages (GAV) as they update the intensity constants in an alternative way, detailed below. In the convex framework, we consider the selective GAV functional:

$$F_{GAV}(u) = TV_g(u) + f_S(u) + f_{GAV}(u), \tag{23}$$

where $f_{GAV}(u) = f_{CV}(u)$. This is identical to the CV selective formulation (8). However, the authors propose an alternative update for the fitting constants c_1 and c_2, given as follows:

$$c_1 = \frac{\int_\Omega z^\beta u_\gamma \, d\Omega}{\int_\Omega z^{\beta-1} u_\gamma \, d\Omega}, \qquad c_2 = \frac{\int_\Omega z^\beta (1 - u_\gamma) \, d\Omega}{\int_\Omega z^{\beta-1} (1 - u_\gamma) \, d\Omega}, \tag{24}$$

with $\beta \in \mathbb{R}$. If $\beta = 1$, the approach is identical to CV. In [2] the authors assert that the proposed adjustments have the following properties. As $\beta \to \infty$, c_1 and c_2 approach the maximum and minimum intensity in the foreground and

background of the image, respectively. Also, as $\beta \to -\infty$, c_1 and c_2 approach the minimum intensity in the foreground and background of the image, respectively. For example, if a high value of β is set, c_1 will take a larger value than in CV which can be useful for selective segmentation. For example, if we consider the image in Fig. 1 we can achieve a larger c_2 value by setting $\beta > 1$ and a smaller value by setting $\beta < 1$. Therefore, there is more flexibility when using this data fitting term in selective formulations. However, it should be noted that it involves the selection of the parameter β, which can be difficult to optimise.

3 Alternative Selective Segmentation Models

We now introduce two recent methods that incorporate user input to perform selective segmentation. Each involves input in the form of foreground/background regions to indicate relevant structures of interest. An example of this can be seen in Fig. 18, where red regions indicate foreground and blue regions indicate background. We compare against the work of Nguyen et al. [29], which uses a similar convex relaxation framework to the proposed approach, and Dong et al. [17], which uses a variation of the random walk approach. We summarise the essential aspects of each approach in the following.

3.1 Constrained Active Contours (CAC) [29]

The authors use a probability map, $P(\boldsymbol{x})$, from Bai and Sapiro [4] where the geodesic distances to the foreground/background regions are denoted by $D_F(\boldsymbol{x})$ and $D_B(\boldsymbol{x})$, respectively. An approximation of the probability that a point $\boldsymbol{x}$ belongs to the foreground is then given by

$$P(\boldsymbol{x}) = \frac{D_B(\boldsymbol{x})}{D_F(\boldsymbol{x}) + D_B(\boldsymbol{x})}. \tag{25}$$

Foreground/background Gaussian mixture models (GMM) are estimated from the user input. The terms $Pr(\boldsymbol{x}|F)$ and $Pr(\boldsymbol{x}|B)$ denote the probability that a point, $\boldsymbol{x}$, belongs to the the foreground and background, respectively. The normalised log likelihood for each is then given by

$$P_F(\boldsymbol{x}) = -\log Pr(\boldsymbol{x}|F)/(-\log Pr(\boldsymbol{x}|F) - \log Pr(\boldsymbol{x}|B)),$$
$$P_B(\boldsymbol{x}) = -\log Pr(\boldsymbol{x}|B)/(-\log Pr(\boldsymbol{x}|F) - \log Pr(\boldsymbol{x}|B)). \tag{26}$$

GMMs are widely used in selective segmentation [18,35,4, 22,17] and the authors in [29] incorporate this idea into the framework we consider with the following data fitting term:

$$h_c(\boldsymbol{x}) = \alpha_0 \left(P_B(\boldsymbol{x}) - P_F(\boldsymbol{x}) \right) + (1 - \alpha_0) \left(1 - 2P(\boldsymbol{x}) \right), \tag{27}$$

for a weighting parameter $\alpha_0 \in [0, 1]$. It is proposed that α_0 is selected automatically as follows:

$$\alpha_0 = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{\log Pr(x_i|F) - \log Pr(x_i|B)}{\log Pr(x_i|F) + \log Pr(x_i|B)} \right|, \tag{28}$$

where N is the total number of pixels in the image. Defining g_0 as the function $g(s)$ applied to the image $z(\boldsymbol{x})$ and g_p applied to the GMM probability map $P_F(\boldsymbol{x})$, an enhanced edge function is defined as

$$g_c(\boldsymbol{x}) = \beta_0 g_p + (1 - \beta_0) g_0, \tag{29}$$

for a weighting parameter $\beta_0 \in [0, 1]$, which can be set automatically in a similar way to (28). Thus, Nguyen et al. [29] define the Constrained Active Contours (CAC) Model as

$$\min_{u \in [0,1]} \left\{ \int_\Omega g_c(\boldsymbol{x}) |\nabla u(\boldsymbol{x})| \, \mathrm{d}\Omega + \lambda \int_\Omega h_c(\boldsymbol{x}) u(\boldsymbol{x}) \, \mathrm{d}\Omega \right\}. \tag{30}$$

They obtain a solution using the split Bregman method of Goldstein et al. [19], although other methods are applicable and will yield similar results. However, that is not the focus of this paper so we omit the details here. In the results section, §7, we will compare our method against CAC to see our data fitting term compares against a GMM-based approach.

3.2 Submarkov Random Walks (SRW) [17]

We now introduce a recent selective segmentation method by Dong et al. [17] known as Submarkov Random Walks (SRW). Rather than using the continuous framework of [14], this approach is based in the discrete setting where each pixel in the image is treated as a node in a weighted graph. Random walks (RW) have been widely used for segmentation since the work of Grady [22]. SRW is capable of achieving impressive results with user-defined foreground and background regions. The selective segmentation result can be obtained by assigning a label to each pixel based on the computed probabilities of the random walk approach. For brevity, we do not provide the full details of the method here, however, further details can be found in [17]. We compare SRW to our proposed approach on a CT data set in §7.4.

We now introduce essential notation to understand the approach of [17]. In RW an image is formulated as a weighted undirected graph $G = (V, E)$ with nodes $v \in V$ and edges $e \in E \subseteq V \times V$. Each node v_i represents an image pixel x_i. An edge e_{ij} connects two nodes v_i and v_j and a weight $w_{ij} \in W$ of edge e_{ij} measures the likelihood that a random walker will cross this edge:

$$w_{ij} = \exp\left(-\frac{||I_i - I_j||^2}{\sigma_0} \right) + \varepsilon_0, \tag{31}$$

where I_i and I_j are pixel intensities, with $\sigma_0, \varepsilon_0 \in \mathbb{R}$. In SRW a user indicates foreground/background regions in a similar way to CAC, as shown in Fig. 18, and can be viewed

as a traditional random walker with added auxiliary nodes. In [17], these are defined as a set of labelled nodes $V_M = \{V^{l_1}, V^{l_2}, ..., V^{l_K}\}$. A set of labels is defined, $LS = \{l_1, l_2, ..., l_K\}$, with K the number of labels $V^{l_k} = \{V_1^{l_1}, V_2^{l_1}, ..., V_{M_K}^{l_K}\}$, and M_k the number of seeds labelled l_k. The prior is then constructed from the seeded nodes (defined by the user). Assuming a label l_k has an intensity distribution H_k (based on GMM learning), a set of auxiliary nodes $H_k = \{h_1, h_2, \cdots, h_K\}$ is added into an expanded graph G_e to define a graph with prior $\bar{G}$. Each prior node is connected with all nodes in V and the weight, w_{ih_k}, of an edge between a prior node h_k and a node $v_i \in V$ is proportional to u_i^k, the probability density belonging to H_k at v_i.

The authors define the probabilities of each node $v_i \in V$ belonging to label l_k as the average reaching probability, denoted $\bar{r}_i^{l_k}$. This term incorporates the auxiliary nodes introduced above and is dependent on multiple variables and parameters, including w_{ij} (31). Further details can be found in [17]. The segmentation result is then found by solving the following discrete optimisation problem:

$$\bar{R}_i = \arg\max_{l_k} \bar{r}_i^{l_k}, \qquad (32)$$

where $\bar{R}_i$ represents the final label for each node. In other words, for a two-phase segmentation problem, $\bar{R}_i$ is analogous to the discretised solution of a convex relaxation problem in the continuous setting. Comparisons in terms of accuracy can therefore be made directly, which we elaborate on further in §7. The authors also detail the optimisation procedure and aspects of dealing with noise reduction.

4 Proposed Model

In this section we introduce the proposed data fitting term for selective segmentation. We consider objects that are approximately homogeneous in the target region. Intrinsically, it is then assumed that the region $\mathscr{P}$, provided by the user, is likely to provide a reasonable approximation of the optimal c_1 value and therefore an appropriate foreground fitting function, f_1, is given by CV (2). For this reason, it makes sense to retain this term in the proposed approach. The contradiction is in how the background fitting function f_2 is defined. Considering piecewise-constant assumptions of the image, and many of the related approaches, the background is expected to be defined by a single constant value, c_2. If $c_1 \approx c_2$ then $f_2 \approx f_1$ everywhere, and therefore the fitting term can't accurately separate background regions from the foreground. It is not practical to rely on $f_S(u)$ to overcome this difficulty as it will produce an over-dependence on the choice of $\mathscr{M}$ and $\mathscr{P}$. This is prohibitive in practice. An alternative function f_2 must therefore be defined which is compatible with f_1 and $f_S(u)$. Here, we define a new data fitting term that penalises background objects in such a way that

avoids these problems by allowing intensity variation above and below the value c_1. In order to design a new functional, we first look at the original CV background fitting function

$$f_2 = (z(\boldsymbol{x}) - c_2)^2.$$

It is clear that in an approximately piecewise-constant image this function will be small outside the target region (i.e. where the image takes values near c_2) and positive inside the target region. Our aim in a new fitting term is to mimic this in such a way that is consistent with selective segmentation, where regions with a 'foreground intensity' are forced to be in the background. It is beneficial to introduce two parameters, γ_1 and γ_2, to enforce the penalty on regions of intensity in the range $[c_1 - \gamma_1, c_1 + \gamma_2]$, i.e. enforce the penalty asymmetrically around c_1. We propose the following function to achieve this:

$$\tilde{f}_2(\boldsymbol{x}) = \begin{cases} 1 + \frac{z(\boldsymbol{x}) - c_1}{\gamma_1}, & c_1 - \gamma_1 \leq z(\boldsymbol{x}) \leq c_1 \\ 1 - \frac{z(\boldsymbol{x}) - c_1}{\gamma_2}, & c_1 < z(\boldsymbol{x}) \leq c_1 + \gamma_2 \\ 0, & else. \end{cases} \qquad (33)$$

This function takes its maximum value where $z(\boldsymbol{x}) = c_1$ and is 0 for $z(\boldsymbol{x}) > c_1 - \gamma_1$ and $z(\boldsymbol{x}) < c_1 + \gamma_2$. In Fig. 2 we provide a 1D representation of $\tilde{f}_2(\boldsymbol{x})$ for various choices of γ_1 and γ_2, with $z(\boldsymbol{x}) \in [0, 1]$ and $c_1 = 0.5$. Here, it can be seen how the proposed data fitting term acts as a penalty in relation to a fixed constant c_1. It is analogous to CV, whilst accounting for the idea of selective segmentation with a data fitting term. The main advantage of this term is that it replaces the dependence on c_2 in the formulation, which has no meaningful relation to the solution of a selective segmentation problem. Even when the foreground is relatively homogeneous, the background may have intensities of a similar value to c_1 which will cause difficulties in obtaining an accurate solution. We detail the proposed fitting term in the following section.

4.1 New Fitting Term

We define the proposed data fitting functional as follows:

$$f_{PM}(u) := \int_{\Omega} (\lambda_1 f_1(\boldsymbol{x}) - \lambda_2 \tilde{f}_2(\boldsymbol{x})) u \ \mathrm{d}\Omega, \qquad (34)$$

for $f_1(\boldsymbol{x}) = (z - c_1)^2$ and $\tilde{f}_2(\boldsymbol{x})$ as defined in (33). This is consistent with respect to the intensities of the observed object and the concept of selective segmentation. In Fig. 3 we see the difference between CV and the proposed fitting terms for given user input on a CT image. For the CT image, the CV fitting terms are near 0 within the target region. This is despite there being a distinct homogeneous area with good contrast on the boundary. This illustrates the problem we are aiming to overcome. With the proposed fitting term this phenomenon should be avoided in cases like this. By defining

Fig. 2: Three 1D plots of $\tilde{f}_2(x)$ whilst varying γ_1 and γ_2 (with $c_1 = 0.5$).

(i) $\gamma_1 = 0.1, \gamma_2 = 0.2$ (ii) $\gamma_1 = 0.3, \gamma_2 = 0.2$ (iii) $\gamma_1 = 0.3, \gamma_2 = 0.4$
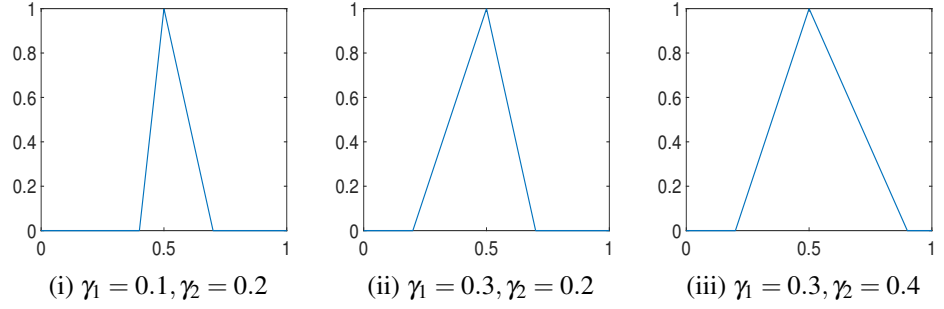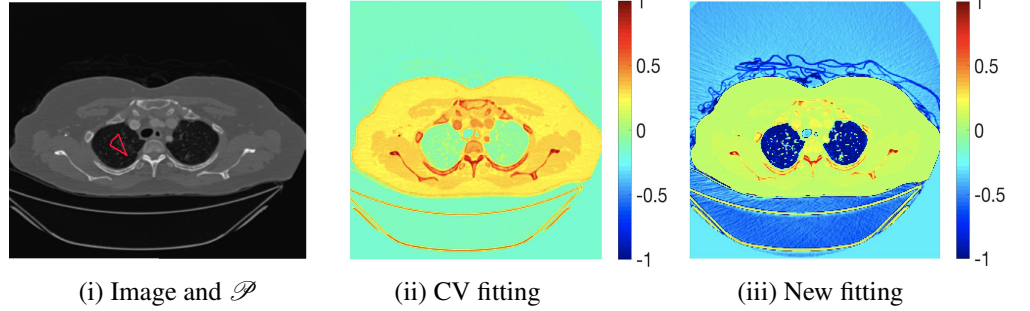
Fig. 3: An image with user input $\mathscr{P}$ shown in red ($c_1 = 0.152, c_2 = 0.188$). Here, we show the difference between the CV fitting function and the proposed approach. The target region is clearly defined by negative values in (iii).

(i) Image and $\mathscr{P}$ (ii) CV fitting   (iii) New fitting

$\tilde{f}_2$ as in (33) there is no contradiction if the foreground and background intensities of the target region are similar.

For images where we assume that the target foreground is approximately homogeneous, we have generally found that fixing c_1 according to the user input is preferable. We compute c_1 as the average intensity inside the region $\mathscr{P}$ formed from the user input marker point set. We therefore propose to minimise the following functional with respect to $u \in [0, 1]$, given a fixed c_1 :

$$F_{PM}(u) = TV_g(u) + f_{PM}(u) + f_S(u). \tag{35}$$

where f_S is the geodesic distance computed as described earlier using (6). The minimisation problem is given as

$$\min_{u \in [0,1]} F_{PM}(u) \tag{36}$$

The model consists of weighted TV regularisation with a geodesic distance constraint as in [34]. However, alternative constraints are possible, such as Euclidean [38], or moments [24]. It is important to note that we have defined the model in a similar framework to the related approaches discussed previously. The main idea is to establish how the proposed fitting term, $f_{PM}(u)$, performs compared to alternative methods. Next we describe how we determine the values of γ_1 and γ_2 in the function $\tilde{f}_2(x)$ automatically. This is important in practice as it avoids any additional user input or parameter dependence to achieve an accurate result. In subsequent sections we provide details of how we obtain a solution for the proposed model.

4.2 Parameter Selection

For a particular problem it is quite straightforward to optimise the choice of γ_1 and γ_2 experimentally, but we would like a method which is not sensitive to the choice of γ_1 and γ_2 and would also prefer that the user need not choose these values manually. Therefore, in this section we explain how to choose these values automatically based on justifiable assumptions about general selective segmentation problems. To select the parameters γ_1 and γ_2 we use Otsu's method [31] to divide the histogram of image intensities into N partitions. Otsu's thresholding is an automatic clustering method which chooses optimal threshold values to minimise the intra-class variance. This has been implemented very efficiently in MATLAB in the function `multithresh` for dividing a histogram such that there are $N - 1$ thresholds T_i.

We use the thresholds from Otsu's method to find γ_1 and γ_2 as follows. There are three cases to consider, based on the value of c_1 computed from the user input: i) $T_{i-1} \leq c_1 \leq T_i$ for some $i > 1$, ii) $0 \leq c_1 \leq T_1$, iii) $T_{N-1} \leq c_1 \leq 1$. For each case we set the parameters as follows:

(i) $\gamma_1 = c_1 - T_{i-1}, \quad \gamma_2 = T_i - c_1$
(ii) $\gamma_1 = c_1, \quad \gamma_2 = T_1 - c_1$
(iii) $\gamma_1 = c_1 - T_{N-1}, \quad \gamma_2 = 1 - c_1$

Choosing N too large could mean γ_1 and γ_2 are too small as the histogram would be partitioned too precisely. Generally we only ever need to consider a maximum of 3 phases for selective segmentation. If there is a large number of pixels in the image with intensity above or below c_1 the image can

Fig. 4: The histograms of intensities for some example images. The red lines are the automatic thresholds T_i obtained by Otsu's thresholding with $N = 3$.

(i) Test Image 1   (ii) Test Image 7   (iii) Test Image 9
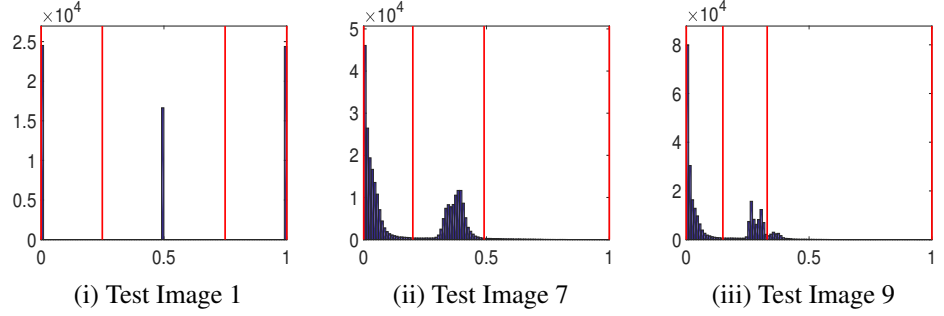
be considered two-phase in practice. Conversely, if a large number of pixels in the image have intensity above and below c_1 the image can essentially be considered three-phase in the context of selective segmentation. This is due to the way $\tilde{f}_2$ has been defined. Therefore, we set $N = 3$ for all tests. In Fig. 4 we can see the Otsu thresholds chosen for various images given in this paper. They divide the peaks in the histogram well and once we know the value of c_1 (the approximation of the intensity of the object we would like to segment) we can automatically choose γ_1 and γ_2 according to this criteria.

5 Numerical Implementation

We now introduce the framework in which we compute a solution to the minimisation of the proposed model, as well the related models introduced in Sections 1 and 2. All consist of the minimisation problem

$$\min_{u \in [0,1]} \left\{ F_X(u) = TV_g(u) + f_X(u) + f_S(u) \right\}, \tag{37}$$

for $X =$ CV, RSF, LCV, HYB, GAV, PM respectively. Minimisation problems of this type (37) have been widely studied in terms of continuous optimisation in imaging, including two-phase segmentation. A summary of such methods in recent years is given by Chambolle and Pock [13]. Our numerical scheme follows the original approach in [14]: enforcing the constraint in (37) with a penalty function, and deriving the Euler-Lagrange of the regularised functional. We then solve the corresponding PDE by following a splitting scheme first applied to this kind of problem by Spencer and Chen [38]. Whilst the numerical details are not the focus of the work, it is important to note widely used alternative methods. It has proved very effective to exploit the duality in the functional and avoid smoothing the TV term. A prominent example is the split Bregman approach for segmentation by Goldstein et al. [19]. This is closely related to augmented lagrangian methods, a matter further discussed by Boyd et al. [7]. Analogous approaches also consist of the first-order primal dual algorithm of Chambolle and Pock [12] and the max-flow/min-cut framework detailed by Yuan

et al. [44]. There are practical advantages in implementing such a numerical scheme for our problem, primarily in terms of computational speed. However, in the numerical tests we include we're mainly interested in accuracy comparisons. For this purpose the convex splitting algorithm of [38] is sufficient, and the extension of splitting schemes for convex segmentation problems may be of interest. Further details can be found in [38] and [34]. In the following, we first discuss the minimisation of (37) in a general sense and then mention some important aspects in relation to the alternative fitting terms discussed in §2.

5.1 Finding the Global Minimiser

To solve this constrained convex minimisation problem (38) we use the Additive Operator Splitting (AOS) scheme from Gordeziani et al. [20], Lu et al. [27] and Weickert et al. [42]. This is used extensively for image segmentation models [33, 34, 38]. It allows the 2D problem to be split into two 1D problems, each solved separately, with the results combined in an efficient manner. We address some aspects of AOS in §6, with further details provided in [34, 38].

A challenge with the functional (35), particularly with respect to AOS, is that this is a constrained minimisation problem. Consequently, it is reformulated by introducing an exact penalty function, $\nu(u)$, given in [14]. To simplify the formulation we define

$$r(\boldsymbol{x}) = \theta \mathscr{D}(\boldsymbol{x}) + f(\boldsymbol{x}),$$

$f(\boldsymbol{x})$ is the function associated with $f_X(u)$. We introduce a new parameter, $\tilde{\lambda}$, which allows us to balance the data fitting terms to the regularisation term more reliably. To be clear, we still only have two main tuning parameters (θ and $\tilde{\lambda}$) as we fix any variable parameters in $f(\boldsymbol{x})$ according to the choices in the corresponding papers. The unconstrained minimisation problem is then given as:

$$\min_{u} \left\{ TV_g(u) + \tilde{\lambda} \int_{\Omega} r(\boldsymbol{x}) u \, \mathrm{d}\Omega + \alpha \int_{\Omega} \nu(u) \, \mathrm{d}\Omega \right\}. \tag{38}$$

We rescale the data term with $\mathscr{F}(\boldsymbol{x}) = r(\boldsymbol{x})/||r(\boldsymbol{x})||_{\infty}$. In effect this change is simply a rescaling of the parameters. This

allows for the parameter choices between different models to be more consistent, as the fitting terms are similar in value. The problem (38) has the corresponding Euler-Lagrange equation (for fixed c_1):

$$\nabla \cdot \left(g(|\nabla z|) \frac{\nabla u}{|\nabla u|_{\varepsilon_1}} \right) - \tilde{\lambda} \mathscr{F}(\boldsymbol{x}) - \alpha v'_{\varepsilon_2}(u) = 0. \tag{39}$$

in Ω and $\frac{\partial u}{\partial \boldsymbol{n}} = 0$ where $\boldsymbol{n}$ is the outward unit normal. The constraint is enforced for $\alpha > \frac{\tilde{\lambda}}{2}||r(\boldsymbol{x})||$ by [14]. Two parameters, ε_1 and ε_2, are introduced here. The former is to avoid singularities in the TV term and the latter is associated with the regularised penalty function $v_{\varepsilon_2}(u)$ from [38]:

$$v_{\varepsilon_2}(u) = H_{\varepsilon_2}(b_{\varepsilon_2}(u))[b_{\varepsilon_2}(u)], \tag{40}$$

with $b_{\varepsilon_2}(u) = \sqrt{(2u-1)^2 + \varepsilon_2} - 1$ and regularised Heaviside function

$$H_{\varepsilon_2}(u) = \frac{1}{2}\left(1 + \frac{2}{\pi}\arctan\left(\frac{u}{\varepsilon_2}\right)\right). \tag{41}$$

The viscosity solution of the parabolic formulation of (39), obtained by multiplying the PDE by $|\nabla u|$, exists and is unique. The general proof for a class of PDEs to which (39) belongs, is included in [34] and we refer the reader there for the details. Once the solution to (39) is found, denoted u^*, we define the computed foreground region as follows:

$$u_\gamma = \{x \in \Omega \,|\, u^*(x) > \gamma\}. \tag{42}$$

We select $\gamma = 0.5$ (although other values $\gamma \in (0,1)$ would yield a similar result according to Chan et al. [14]). In the following we use the binary form of the solution, u^*, denoted u_γ. This partitions the domain into Ω_F and Ω_B according to the labelling function u_γ.

5.2 Implementation for Related Models

The discussion in this section so far has used the function $f(\boldsymbol{x})$ associated with the data fitting functional $f_X(u)$. This corresponding equations for the RSF, LCV, HYB and GAV models are detailed in §2, CV is discussed in §1, and our approach is given by eqn. (34). We use this implementation to obtain selective segmentation versions of each of those models, given by (37). When these terms contain parameter choices we follow the advice in the corresponding papers as far as possible, unless we have found that alternatives will improve results. In the next section we will give the results of these models and compare them to our proposed approach.

Note. We now discuss details behind tuning parameters for the GAV model. It is noted in §2 that the GAV model requires a parameter β to adapt the c_1 and c_2 calculation. We find that it is actually better to consider c_1 and c_2 separately

to achieve improved results, as sometimes we wish to tune the values to have a higher c_1 and lower c_2 (or vice-versa) simultaneously. Therefore we introduce parameters β_1 and β_2 to tune c_1 and c_2 as follows:

$$c_1 = \frac{\int_\Omega z^{\beta_1} u \, \mathrm{d}\Omega}{\int_\Omega z^{\beta_1 - 1} u}, \qquad c_2 = \frac{\int_\Omega z^{\beta_2}(1-u)}{\int_\Omega z^{\beta_2 - 1}(1-u)} \, \mathrm{d}\Omega, \tag{43}$$

In all experiments, we tested the following combinations of (β_1, β_2): $(1.5, 0.5)$, $(2, 0)$, $(3, -1)$, $(4, -2)$, $(0.5, 1.5)$, $(0, 2)$, $(-1, 3)$ and $(-2, 4)$. For each choice, we optimised the values of $\tilde{\lambda}$ and θ according to the procedure described in §7.1. This allowed us to select the optimal combination of (β_1, β_2) for each image.

6 Algorithm

Here, we will discuss the algorithm that we use to minimise the selective segmentation model (37). We utilise additive operator splitting techniques to solve the minimisation problem efficiently.

6.1 An Additive Operator Splitting (AOS) Scheme

Additive Operator Splitting (AOS) [20, 27, 42] is a widely used method for solving PDEs with linear and non-linear diffusion terms [33, 34, 38] such as

$$\frac{\partial u}{\partial t} = \mu \nabla \cdot (G(u)\nabla u) - f_0. \tag{44}$$

AOS allows us to split the two-dimensional problem into two one-dimensional problems, which we solve separately and then combine. Each one-dimensional problem gives rise to a tridiagonal system of equations which can be solved efficiently by Thomas' algorithm, hence AOS is a very efficient method for solving PDEs of this type. AOS is a semi-implicit method and permits far larger time-steps than the corresponding explicit schemes would. Hence AOS is more stable than an explicit method [42]. Note here that

$$G(u) = \frac{g(|\nabla z|)}{|\nabla u|_{\varepsilon_1}}, \qquad f_0 = \tilde{\lambda}\mathscr{F}(\boldsymbol{x}) + \alpha v'_{\varepsilon_2}(u), \tag{45}$$

and $\mu = 1$. The standard AOS scheme assumes f_0 does not depend on u, however in this instance that is not the case. This requires a modification to be used for convex segmentation problems, first introduced by [38]. This non-standard formulation incorporates the regularised penalty term, $v_{\varepsilon_2}(u)$, into the AOS scheme which we briefly detail next.

The authors consider the Taylor expansions of $v'_{\varepsilon_2}(u)$ around $u = 0$ and $u = 1$. They find that the coefficient b of the linear term in u is the same for both expansions. Therefore, for a change in u of δu around $u = 0$ and $u = 1$ the change

in $v'_{\varepsilon_2}(u)$ can be approximated by $b \cdot \delta(u)$. To address this, the relevant interval is defined as

$$I_\zeta := [0 - \zeta, 0 + \zeta] \cup [1 - \zeta, 1 + \zeta]$$

and a corresponding update function is given as

$$\tilde{b}(\boldsymbol{x}) = \begin{cases} b, & x \in \Omega, \ u(\boldsymbol{x}) \in I_\zeta \\ 0, & else. \end{cases}$$

The solution for (44) is then obtained by discretising the equation as follows:

$$\frac{u^{(k+1)} - u^{(k)}}{\tau} = \mu \sum_{\ell=1,2} A_\ell(u^{(k)}) u^{(k+1)} + \alpha \tilde{b}^{(k)}(u^{(k)} - u^{(k+1)}) - f_0^{(k)}.$$

where A_1 and A_2 are discrete forms of $\partial_x(G(u)\partial_x)$ and $\partial_y(G(u)\partial_y)$, respectively (given in [38,34]). The modified AOS update is then given by

$$u^{(k+1)} = \frac{1}{2} \sum_{\ell=1}^{2} \left(I - 2\tau\mu(I + \tilde{B}^{(k)})^{-1} A_\ell(u^{(k)}) \right)^{-1} \tilde{u}^{(k)}, \quad (46)$$

where $\tilde{B}^{(k)} = \mathrm{diag}(\tau\alpha\tilde{b}^{(k)})$ and $\tilde{u}^{(k)} = u^{(k)} + \tau(I + \tilde{B}^{(k)})^{-1} f_0^{(k)}$. This scheme allows for more control on the changes in f_0 between iterations due to the function $\tilde{b}$ and parameter ζ, and therefore leads to a more stable convergence. We refer the reader to [38] for full details of the numerical method.

6.2 The Proposed Algorithm

In Algorithm 1 we provide details of how we find the minimiser of the various selective segmentation models detailed above, defined by (37). The algorithm is in a general form to be applied to any of the approaches discussed so far. It is important to reiterate that alternative solvers to AOS are available, such as the dual formulation [3,8,11], split-Bregman [19], augmented Lagrange [6], primal dual [12], and max-flow/min-cut [44]. In all experiments we use the tolerance of 10^{-4} for the stopping criteria and set $\varepsilon_1 = 10^{-4}$, $\varepsilon_2 = 10^{-1}$ and $\tau = 10^{-2}$.

Algorithm 1 Selective Segmentation Algorithm

Provide user input region $\mathscr{P}$ and compute $\mathscr{D}$, according to (6).
Define $f(\boldsymbol{x})$ appropriately for the model (CV, RSF, LCV, HYB, GAV, or the proposed approach).
Compute $r(\boldsymbol{x}) = \theta\mathscr{D}(\boldsymbol{x}) + f(\boldsymbol{x})$. and $\mathscr{F}(\boldsymbol{x}) = r(\boldsymbol{x})/||r(\boldsymbol{x})||_\infty$.
Initialise u (arbitrary for $u \in [0,1]$).
while $\delta > tolerance$ **do**
    $u_{old} := u$.
    Update u according to the AOS iteration (46).
    $\delta = ||u - u_{old}||/||u_{old}||$.
end while
return $u^* = u$ and binary labelling function, u_γ.

---

7 Results

In this section we will present results obtained using the proposed model and compare them to using fitting terms from similar models (CV [15], RSF [25], LCV [41], HYB [1], GAV [2]), detailed in §2. We intend to provide an overview of how effective each model is in a number of key respects and analyse their potential for practical use in a reliable and consistent manner. Our focus is on how each model can be extended to a consistent selective segmentation framework. The key questions we consider are:

(i) How sensitive are the results to variations of the parameters $\tilde{\lambda}$ and θ?
(ii) Is the model capable of achieving accurate results?
(iii) To what extent is the proposed model dependent on the user input?

Test Images. We will perform all tests on the images shown in Figs. 5–7. We have provided the ground truth and initialisation used for each image. Test Images 1–3 are synthetic, Test Image 4 is an MRI scan of a knee, Test Images 5–6 are abdominal CT scans, and Test Images 7–9 are lung CT scans. They have been selected to present challenges relevant to the discussion in §2. We focus on medical images as this is the application of most interest to our work. In the following we will discuss the results in terms of synthetic images (1–3) and real images (4–9).

Measuring Segmentation Accuracy. In our tests we use the Jaccard Coefficient [23], often referred to as the Tanimoto Coefficient (TC), to measure the quality of the segmentation. We define accuracy with respect to a ground truth, GT, given by a manual segmentation:

$$GT = \{x \in \Omega \,|\, x \in \text{foreground}\}.$$

The Tanimoto Coefficient is then calculated as

$$\mathrm{TC} = \frac{|N(u_\gamma \cap GT)|}{|N(u_\gamma \cup GT)|},$$

where $N(\cdot)$ refers to the number of points in the enclosed region. This takes values in the range $[0, 1]$, with higher TC values indicating a more accurate segmentation. In the following we will represent accuracy visually from red (TC = 0) to green (TC = 1), with the intermediate scaling of colours used shown in Fig. 8. This will be particularly relevant in §7.2.

Note. In §2.4 we mentioned the tuning of parameters in the GAV model. To be explicit the optimal (β_1, β_2) pairs used in the following tests were (4,-2) for Test Images 1 and 2, (1.5,0.5) for Test Images 3,4, and 6, (2,0) for Test Image 5, and (-2,4) for Test Images 7,8, and 9. Results vary significantly as (β_1, β_2) are varied, but we found these to be the best choices for each image.

Fig. 5: Test Images 1–3; the ground truth contours are defined in the first row and the corresponding user input marker set is shown in the second row. These are synthetic images with homogeneous foregrounds selected to highlight the benefits of the proposed model.
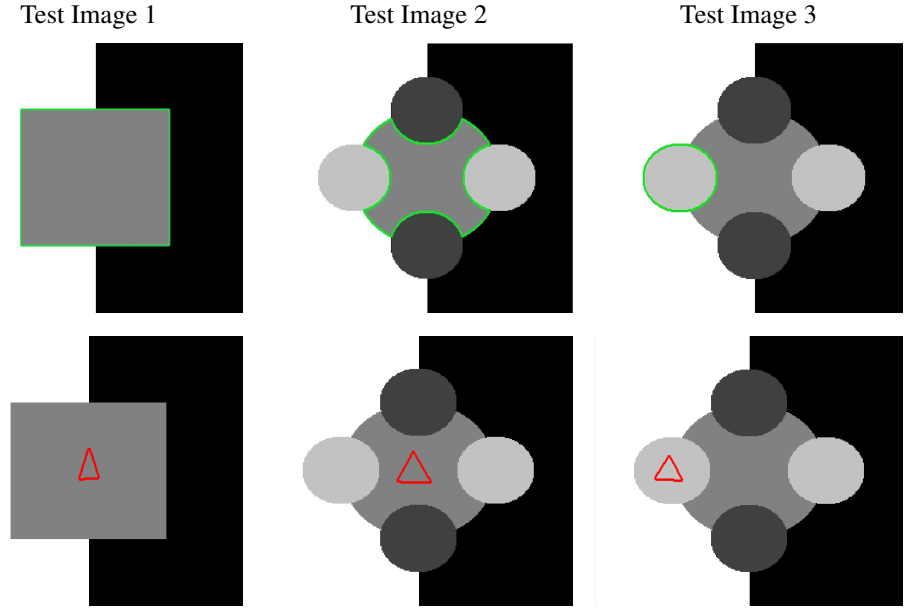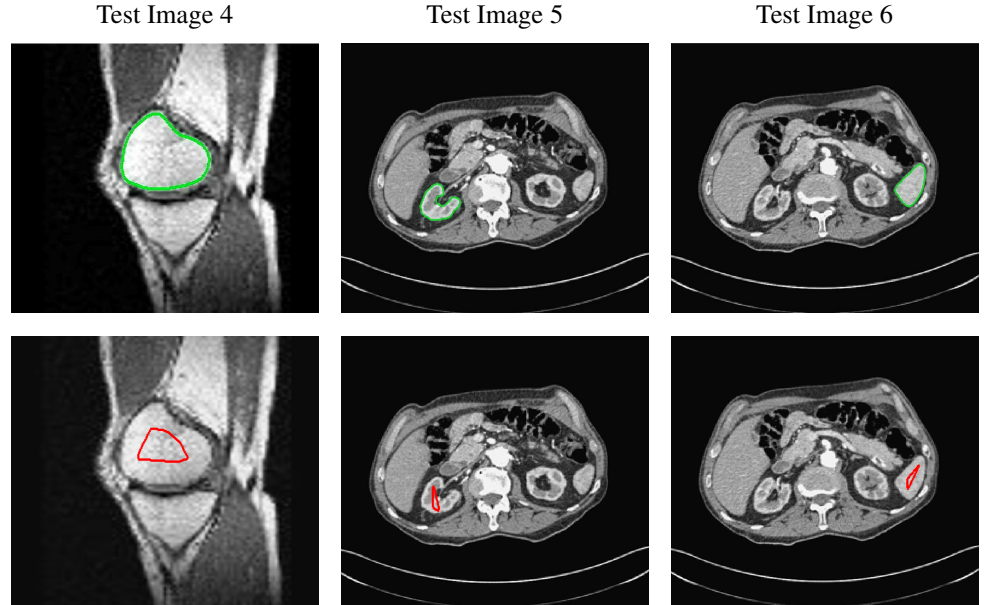
Test Image 1   Test Image 2   Test Image 3

Fig. 6: Test Images 4–6; the ground truth contours are defined in the first row and the corresponding user input marker set is shown in the second row. These are real images with some degree of intensity inhomogeneity in the foreground, with potential medical applications in mind.

Test Image 4   Test Image 5   Test Image 6

The discussion of results is split into four sections, addressing the questions introduced above. First, in Section 7.1, we will examine the robustness to the parameters $\tilde{\lambda}$ and θ for each model. Then, in Section 7.2, we will compare the optimal accuracy achieved by each method to determine what each data fitting term is capable of in the context of selective segmentation for these examples. In Section 7.3, we will test the proposed model with respect to the user input. By randomising the input we will determine to what extent the proposed model is suitable for use in practice. Finally, we will compare the proposed approach to the methods introduced in §3 on an additional CT data set. This will help establish how the algorithm performs against competitive approaches in the literature.

7.1 Parameter Robustness

In these tests we aim to demonstrate how sensitive to parameter choices each choice of fitting term is. To accomplish this we perform the segmentations for each of the models discussed (CV, RSF, LCV, HYB, GAV) and the proposed model for a wide range of parameters and compute the TC value. The parameter range used is $\tilde{\lambda}, \theta \in [1, 50]$. Due to computational constraints, we run for each integer $\tilde{\lambda}, \theta$ between 1 and 10, and every fifth from 15 to 50. This aspect of a model's performance is vital when used in practice. The less sensitive to parameter choices a model is the more relevant it is in relation to potential applications.

Fig. 7: Test Images 7–9;
the ground truth contours
are defined in the first row
and the corresponding user
input marker set is shown
in the second row. These
are real images with ap-
proximately homogeneous
foregrounds. The challenge
is that the background con-
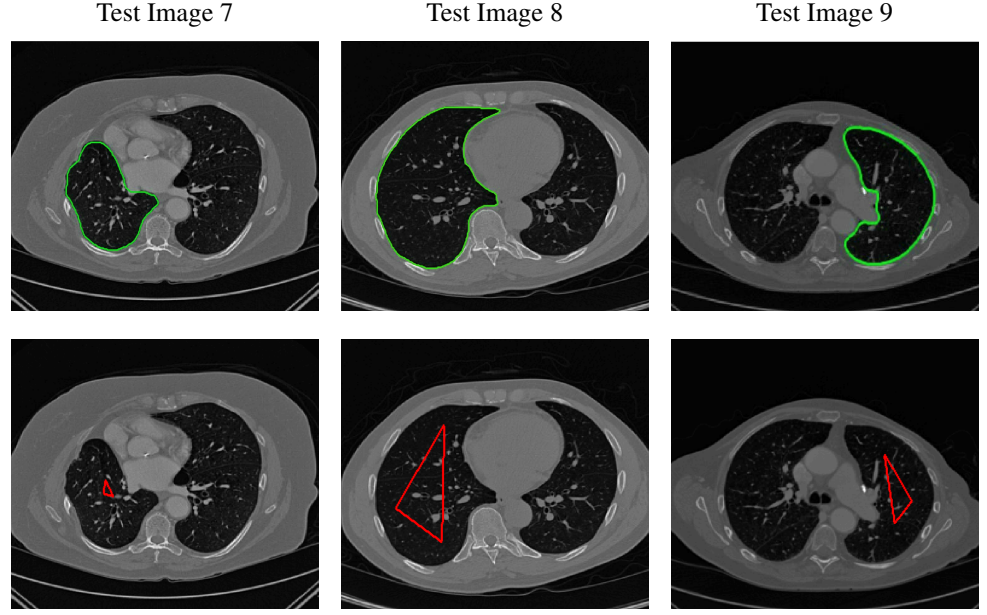tains substantial regions of
a similar intensity.

Test Image 7 | Test Image 8 | Test Image 9

Table 1: Optimal TC values
for Test Images 1–9, from
the heatmaps in Figs. 11,
12, and 13. The best result
for each image is given in
bold.

Model	Test Image								
	1	2	3	4	5	6	7	8	9
CV	0.000	0.000	0.970	0.969	0.933	0.988	0.889	0.931	0.180
RSF	**1.000**	0.997	0.993	0.924	0.884	0.956	0.785	0.950	0.782
LCV	0.313	0.142	0.970	0.970	0.941	0.988	0.911	0.960	0.828
HYB	0.184	0.091	0.988	0.960	0.870	0.988	0.000	0.000	0.000
GAV	0.984	0.960	0.988	0.967	0.965	0.988	0.950	0.954	0.919
Proposed	**1.000**	**1.000**	**1.000**	**0.973**	**0.989**	**0.990**	**0.965**	**0.961**	**0.971**
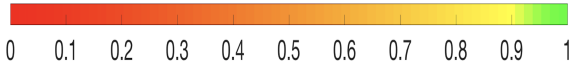
Fig. 8: Colour scaling corresponding to TC values, repre-
senting the accuracy of the result. This scale is used in sub-
sequent figures.

The TC values for the parameter sets $(\tilde{\lambda}, \theta)$ are pre-
sented as heatmaps in Figs. 11–13. A heatmap is a conve-
nient way to display accuracy results for hundreds of tests
concisely. In Fig. 9 we give an example heatmap with the
same axes used for those in Figs. 11–13. For each of the
combinations of parameter values $(\tilde{\lambda}, \theta)$ we give the TC
value of the segmentation result and represent it by the ap-
propriate colour. The corresponding colour scale is shown
in Fig. 8. Qualitatively, the more green areas of the heatmap
the more accurate the model is for a wider set of parame-
ters. Example results for Test Image 5 when varying $\tilde{\lambda}$ (with
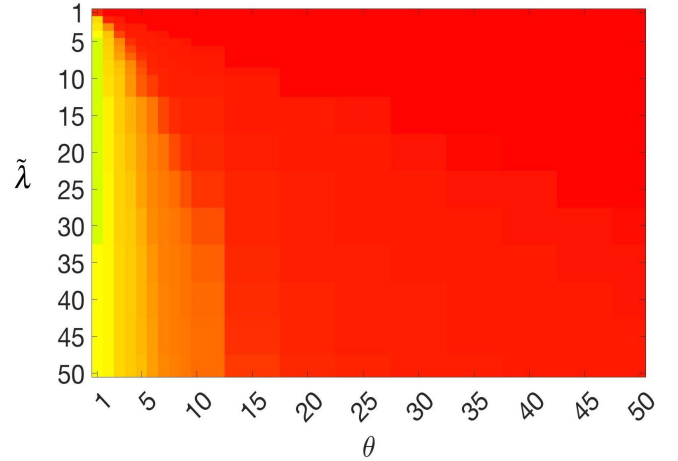$\theta = 4$) for the proposed model are given in Fig. 10. Here it

Fig. 9: Example heatmap of TC values to display segmenta-
tion accuracy for parameters $(\tilde{\lambda}, \theta)$.
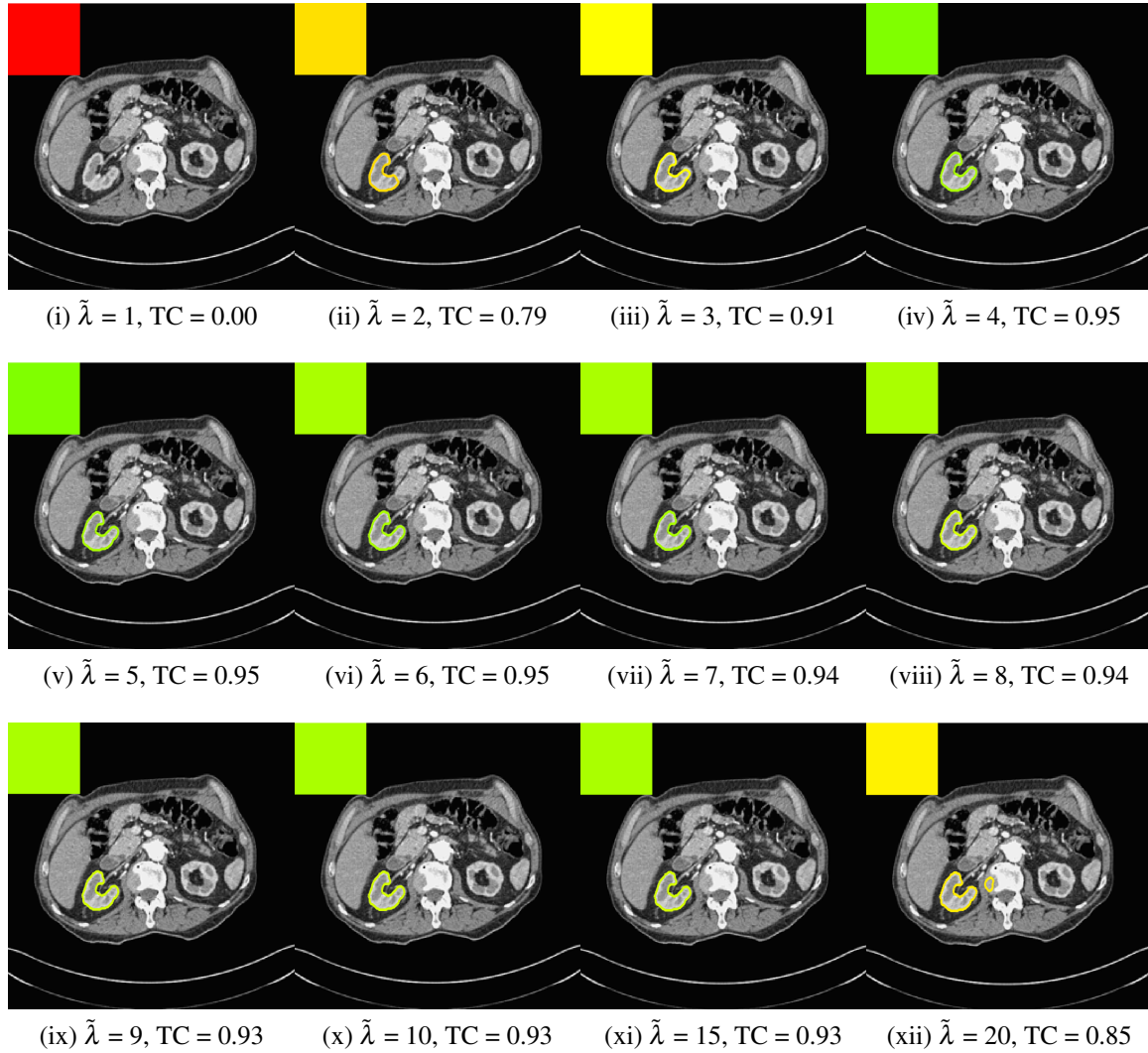
can be seen what each accuracy result corresponds to visu-
ally.

(i) $\tilde{\lambda} = 1$, TC = 0.00   (ii) $\tilde{\lambda} = 2$, TC = 0.79   (iii) $\tilde{\lambda} = 3$, TC = 0.91   (iv) $\tilde{\lambda} = 4$, TC = 0.95

(v) $\tilde{\lambda} = 5$, TC = 0.95   (vi) $\tilde{\lambda} = 6$, TC = 0.95   (vii) $\tilde{\lambda} = 7$, TC = 0.94   (viii) $\tilde{\lambda} = 8$, TC = 0.94

(ix) $\tilde{\lambda} = 9$, TC = 0.93   (x) $\tilde{\lambda} = 10$, TC = 0.93   (xi) $\tilde{\lambda} = 15$, TC = 0.93   (xii) $\tilde{\lambda} = 20$, TC = 0.85

Fig. 10: Segmentation results and TC values for the proposed model whilst varying $\tilde{\lambda}$ (with $\theta = 4$). The colours correspond to the TC value (green is TC = 1, red is TC = 0), consistent with the scale in Fig. 8. This is for Test Image 5, with the corresponding heatmap provided in Fig. 12.

Note. The axes have been removed from the heatmaps in Figs. 11–13 for presentational clarity. However, to be explicit, the axes used in all heatmaps are the same as those in Fig. 9.

Synthetic Images. These results are presented in Fig. 11. For Test Images 1–2 we see poor parameter robustness from all competing models, except for GAV which performs reasonably well. However, the proposed model has minimal parameter sensitivity for these images, with good results achieved for almost every combination of values tested. For Test Image 3 all models have a reasonable parameter range (except for RSF), however the proposed model gives better quality results for a wider parameter range. The other models achieve reasonable results here as the foreground intensity of the ground truth is greater than the background

$(c_1 = 0.75, c_2 = 0.49)$, whereas for Test Images 1–2 they are equal $(c_1 = c_2 = 0.50)$. These results highlight the key advantage of the proposed model.

Real Images. In Fig 12 we present results for Test Images 4–6. Here, the proposed model performs in a similar way to its competitors because these images are more typical selective segmentation problems in the sense that there is a clear distinction between the foreground and background intensities. In particular, the values in each case are: Test Image 4 $(c_1 = 0.85, c_2 = 0.25)$, Test Image 5 $(c_1 = 0.70, c_2 = 0.19)$, and Test Image 6 $(c_1 = 0.73, c_2 = 0.20)$. It can be seen that the proposed model is competitive compared to previous approaches. The performance is quite poor for Test Image 5, but is arguably still the best for this challenging case. In Fig. 13 we present results for Test Images 7–9. Here

Fig. 11: Heatmaps of TC values for permutations of $\tilde{\lambda}$ and θ. Each row and column is labelled according to the model used and the image tested. The colour is consistent with the scale in Fig. 8. Here, we present Test Images 1 – 3.
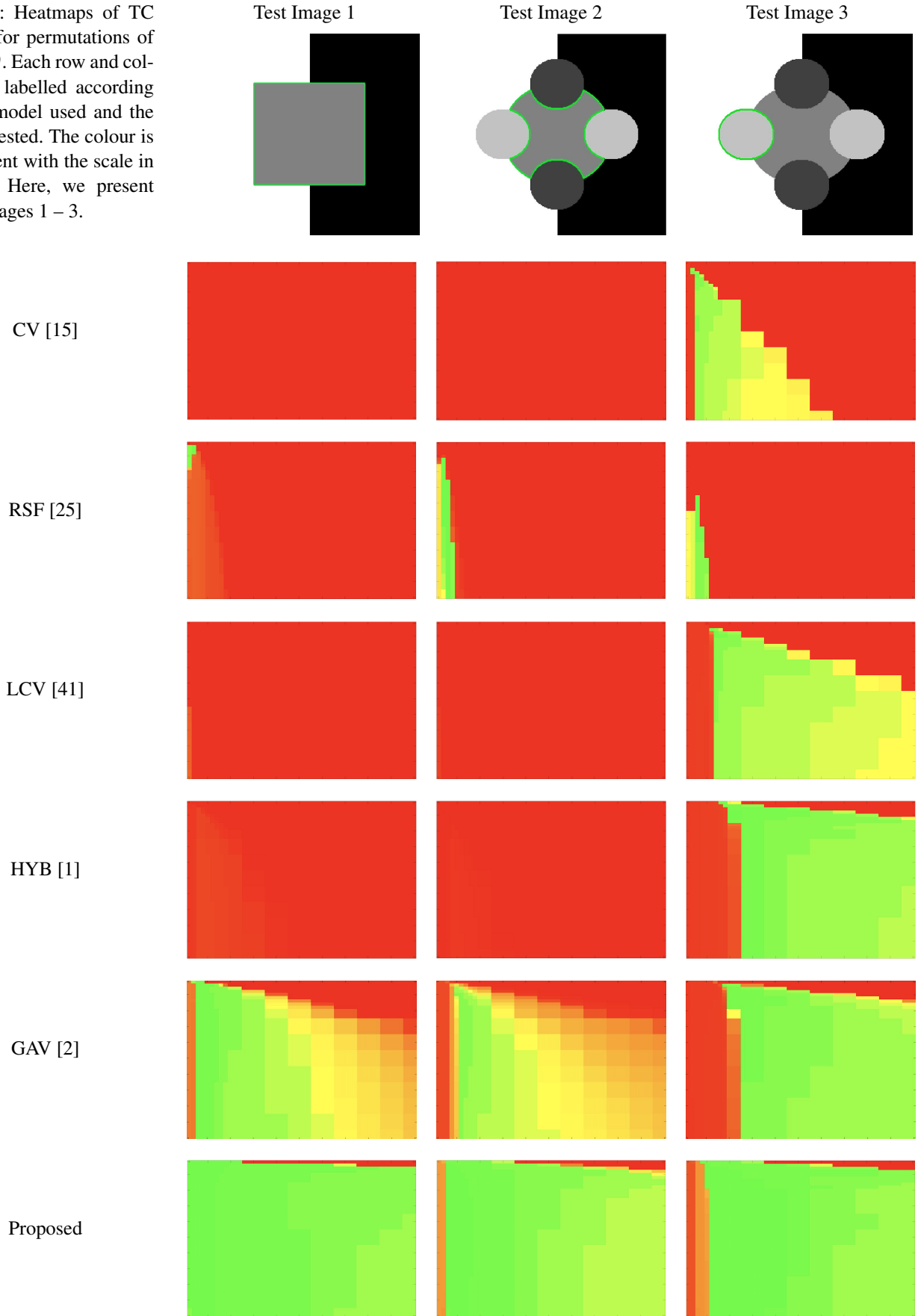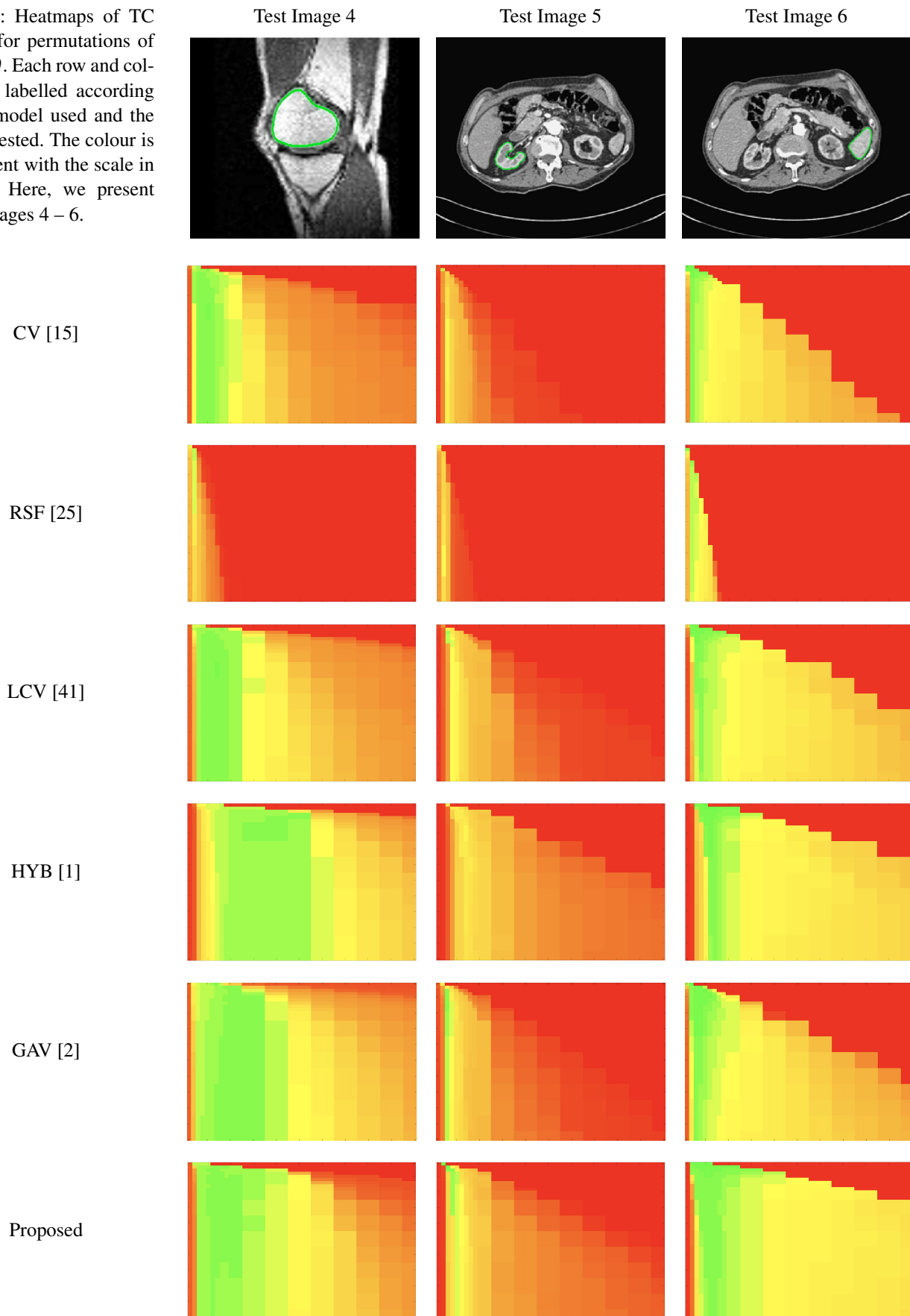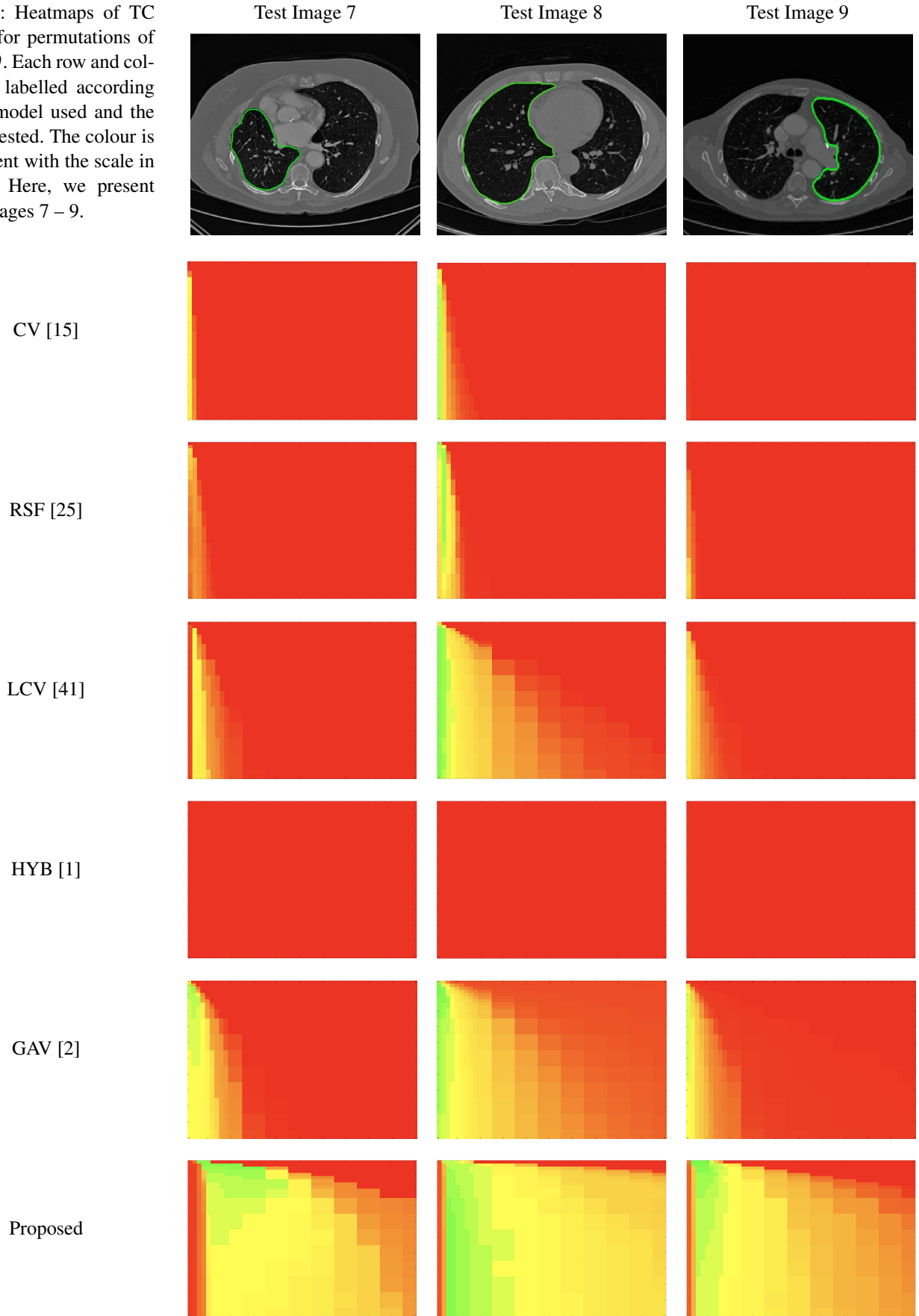
Test Image 1

Test Image 2

Test Image 3

CV [15]

RSF [25]

LCV [41]

HYB [1]

GAV [2]

Proposed

Fig. 12: Heatmaps of TC values for permutations of $\tilde{\lambda}$ and θ. Each row and column is labelled according to the model used and the image tested. The colour is consistent with the scale in Fig. 8. Here, we present Test Images 4 – 6.

Test Image 4   Test Image 5   Test Image 6

CV [15]

RSF [25]

LCV [41]

HYB [1]

GAV [2]

Proposed

Fig. 13: Heatmaps of TC values for permutations of $\tilde{\lambda}$ and θ. Each row and column is labelled according to the model used and the image tested. The colour is consistent with the scale in Fig. 8. Here, we present Test Images 7 – 9.

the proposed model outperforms previous approaches significantly for each image. This is mainly due to the type of image considered. Specifically, the true intensities are: Test Image 7 ($c_1 = 0.12, c_2 = 0.24$), Test Image 8 ($c_1 = 0.10, c_2 = 0.23$), and Test Image 9 ($c_1 = 0.08, c_2 = 0.14$). The proposed model is capable of achieving results where $c_1 \approx c_2$, with other models failing completely in these cases.

7.2 Accuracy Comparisons

Here we aim to address the question of whether each model is capable of achieving an accurate result. In other words, assuming that factors such as parameter and user input sensitivity are ignored, how successful is each approach. In Table 1 we present the optimal TC values for each model found from the tests described in the previous section, with the highest value in bold. Immediately we can see that the proposed model outperforms all other models in terms of accuracy for all test images (RSF equals it for Test Image 1). Below we will discuss some relevant details of the results, again by splitting the test images into synthetic and real.

Synthetic Images. We observe that for Test Images 1 and 2 (where $c_1 = c_2$, CV, LCV, and HYB fail completely. RSF and GAV perform well, with the proposed model being the most accurate with perfect results. For Test Image 3, all models are capable of achieving a good result. It should be noted that in this case $c_1 = 0.75$ and $c_2 = 0.49$. This difference enables the other models to perform well, although the proposed model is slightly superior with a perfect result.

Real Images. In Table 1 we can see that the proposed model is the most successful in terms of optimal accuracy. It is worth noting some inconsistency in the other models, with all but GAV having results that fall below TC = 0.9 for at least one image. GAV performs well for Test Images 4–9, with the proposed model slightly outperforming it in each case. It is worth reminding the reader that for GAV the parameters (β_1, β_2) have been refined for each example. Fixing this results in more variability in the quality of results. The proposed model has no such parameter optimisation between examples. We present the optimal results for Test Image 9 in Fig. 14. Here we can see how much variation there is in the quality of results for this lung CT image. GAV is the most competitive (TC = 0.919), but is visually inadequate. Two other models (CV, HYB) fail completely. In this case, the problem looks quite straightforward and yet other fitting terms are insufficient to produce a good result. Again, the proposed model tends to be superior in cases where $c_1 \approx c_2$ and is capable of achieving very good results for all the images considered. This highlight the advantages of the proposed fitting term.

7.3 User Input Randomisation

One key consideration for the practical use of selective segmentation models is that the result is not too reliant on user input. With intricate user input accurate results are almost guaranteed. However, the benefit of this kind of approach is that accuracy should be attainable with minimal, intuitive user input. One challenge in this setting is how to ascertain to what extent a method is dependent on the user input. In this section we will generalise the user input for the proposed model in order to determine how sensitive it is in this respect. By generalising in this way we will make two assumptions about the markers, $\mathscr{M}$, consistent with the above considerations:

(i) All points are within the target object.
(ii) Only 3 markers are selected.

We regard neither of these assumptions to be too onerous on a user, and are quite consistent with practical use. To perform this test, we randomly choose 1000 sets of 3 marker points and run each algorithm using them. The parameters $\tilde{\lambda}$ and θ are fixed at those which gave the optimal TC values in Table 1. For each set of marker points we compute the corresponding TC value of applying the proposed model with this input. The results for each image are summarised by boxplots in Fig. 15 with examples of the worst results, excluding outliers, shown in Fig. 16. Here, it can be seen that the worst result outperforms the optimal results of the alternative models considered, which is impressive. Below we discuss the results for the test images, by again splitting them into synthetic and real images.

Synthetic Images. For the Test Images 1–3 we achieve near perfect segmentations in all cases, shown by the mean TC being between 0.99 and 1.00 in all cases (for Test Image 1, the mean is precisely 1.00) and a small variance around the mean. Therefore, we can conclude that for images of this type, where the foreground is homogeneous, our method is very robust to user input. Essentially, any reasonable set of markers should produce excellent results. It should be noted that the optimal results from comparable approaches are less than the mean result of 1000 random tests for our method. This can be observed in Table 1. Furthermore, these methods often fail completely. This is a key result highlighting the advantages of our method. In visually simple cases (Test Images 1–3) our new data fitting term is an improvement on existing approaches by modifying the underlying assumptions involved.

Real Images. In all cases for Test Images 4–9 the mean values show that the segmentation results are highly accurate. Also, we notice that the variances are very reasonable demonstrating the robustness of varying the user input. This is an important aspect of selective segmentation, and highlights the advantages of the proposed fitting term. For Test Images 4–6 we observe more variability in the accuracy due

Fig. 14: We present the optimal result for each model for Test Image 9. The accuracy is represented by colour, consistent with the scale in Fig. 8. The proposed model significantly outperforms previous approaches in this case.
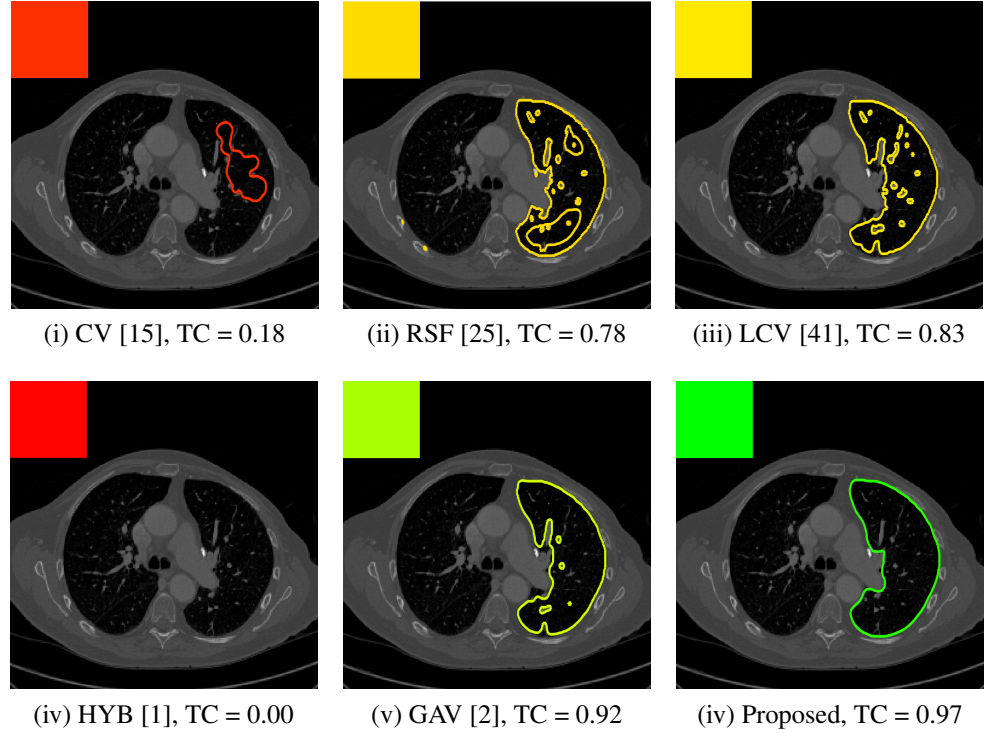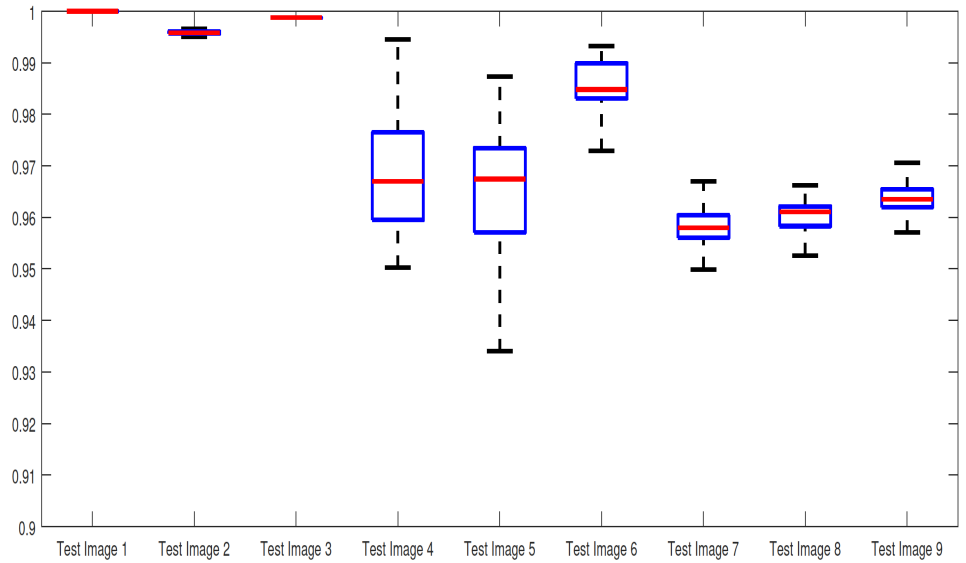

(i) CV [15], TC = 0.18   (ii) RSF [25], TC = 0.78   (iii) LCV [41], TC = 0.83

(iv) HYB [1], TC = 0.00   (v) GAV [2], TC = 0.92   (iv) Proposed, TC = 0.97

Fig. 15: Boxplots of the TC values for 1000 random user inputs using the proposed model. We observe that the method is remarkably consistent. Even the worst results, excluding outliers, are competitive with the optimal results of the existing approaches shown in Table 1.

to minor intensity inhomogeneity in the foreground. This means randomising the user input will be more sensitive. However, we can see that the results are very good with the mean accuracy being competitive with the optimal accuracy of comparable methods. In the case of the lung CT images (Test Images 7–9) the variance in TC values is very small, due to the homogeneity of the foreground. Again, it is important to compare the results of 1000 random results using our proposed model to the optimal result of comparable methods. For these images all of the methods except GAV

have at least one TC value below 0.9. However, GAV requires the tuning of additional parameters (β_1, β_2) whilst the proposed model does not. Compared to GAV, we can see that the mean of our tests is similar to the optimal value of GAV. One exception is for Test Image 9 (shown in Fig. 14), where there is a significant gap in favour of our model. Again, from Fig. 16, we can see that the worst result of randomising the user input for the proposed model is competitive with the optimal results of the alternatives. This is one of the most encouraging aspects of the tests; the proposed model is re-

markably robust to varying user input. This proves that successful results with minimal, intuitive user input is possible for a range of examples.

7.4 Additional Tests

In order to further establish the robustness of our method, we now introduce the results of testing our approach against competing interactive segmentation methods. Specifically, we compare against the work of Nguyen et al. [29] and Dong et al. [17], referred to as CAC and SRW respectively and detailed in §3. The results are presented in Fig. 17, showing a boxplot of accuracy in terms of TC on a set of 30 CT images (excluding outliers). The target structure we consider is the spleen, as this consists of a relatively homogeneous foreground, appropriate for the approach considered. The data has been manually contoured providing ground truth data for the image set. We compare CAC [29] and SRW [17] against our method with five variations of user input for each image. A representative example for three images is shown in Fig. 18. This shows foreground (red) and background (blue) user input regions. For our method, we define the red region as $\mathscr{P}$ as discussed in §1 and enforce hard constraints on the blue region. We refer to the results of the proposed approach using this input as Ours (i). We also include results of randomising the user input in an identical way to §7.3. For each image we generate 1000 simulated user input choices, which we present as Ours (ii). It is important to note that the difference between Ours (i) and (ii) is only the definition of $\mathscr{P}$. The method and parameters are fixed between each.

The performance of CAC [29] is very good, as shown in Fig. 17. We have included an additional figure to highlight the difference between CAC and Ours (i) and (ii) more precisely. This is shown in Fig. 19 (this is the same as Fig. 17 with TC restricted to [0.8,1]). Here we can see that the proposed approach has a slightly better median (0.96 compared to 0.94) and is generally more consistent than CAC. This is particularly evident when considering the worst TC results of CAC (0.19) against ours (0.87).

In Fig. 17 it can be seen that our method exceeds the performance of SRW by a large margin (0.66 compared to 0.95). One possible reason for this is that the input used, as displayed in Fig. 18, is restricted to be as intuitive as possible. SRW is capable of achieving improved results with more elaborate foreground/background input. However, it is generally reliant on a trial and error approach which is not ideal in practice. This highlights an important advantage of our method. It is able to achieve a high standard of results with simple user input. This is reinforced by considering Ours (ii), where the results of 30000 random variations of the user input does not cause a drop off in accuracy compared to the 150 manual user input selections. Again, this can be seen more clearly in Fig. 19. In fact, the results for

the proposed approach with the random input are slightly better than with the manual input. This underlines the robustness to user input in the model, which is a vital aspect of selective segmentation.

8 Conclusion

In this paper we have proposed a new intensity fitting term, for use in selective segmentation. We have compared it to fitting terms from comparable approaches (CV, RSF, LCV, HYB, GAV), in order to address an underlying problem in selective segmentation: if the foreground is approximately homogeneous what is the best way to define the intensity fitting term? Previous methods [33, 34, 38] involve contradictions in the formulation, which we attempt to address.

We have evaluated the success of the proposed model in four respects: parameter robustness, optimal accuracy, dependence on user input, and comparisons to competing selective models. Our focus is on medical applications, where the target object has approximately homogeneous intensity. In each way, the proposed model performs very well, particularly in cases where the true foreground and background intensities are similar. We have shown that our method is remarkably insensitive to varying user input, highlighting its potential for use in practice, and also outperforms competitive algorithms in the literature.

References

1. Ali, H., Badshah, N., Chen, K., Khan, G.: A variational model with hybrid images data fitting energies for segmentation of images with intensity inhomogeneity. Pattern Recognition **51**, 27–42 (2016)
2. Ali, H., Badshah, N., Chen, K., Khan, G.A., Zikria, N.: Multiphase segmentation based on new signed pressure force functions and one level set function. Turkish Journal of Electrical Engineering & Computer Sciences **25**, 2943–2955 (2017)
3. Aujol, J.F., Gilboa, G., Chan, T., Osher, S.: Structure-texture decomposition–modeling, algorithms, and parameter selection. International Journal of Computer Vision **67**(1), 111–136 (2006)
4. Bai, X., Sapiro, G.: A geodesic framework for fast interactive image and video segmentation and matting. IEEE International Conference on Computer Vision pp. 1–8 (2007)
5. Benard, A., Gygli, M.: Interactive video object segmentation in the wild. CoRR **abs/1801.00269** (2017)

Fig. 16: Results for the proposed model for each image, including TC values. The worst result, excluding outliers, of 1000 random user inputs for each example is presented. This demonstrates that the model is robust to user input, with poor results being competitive with the optimal result of competitors.
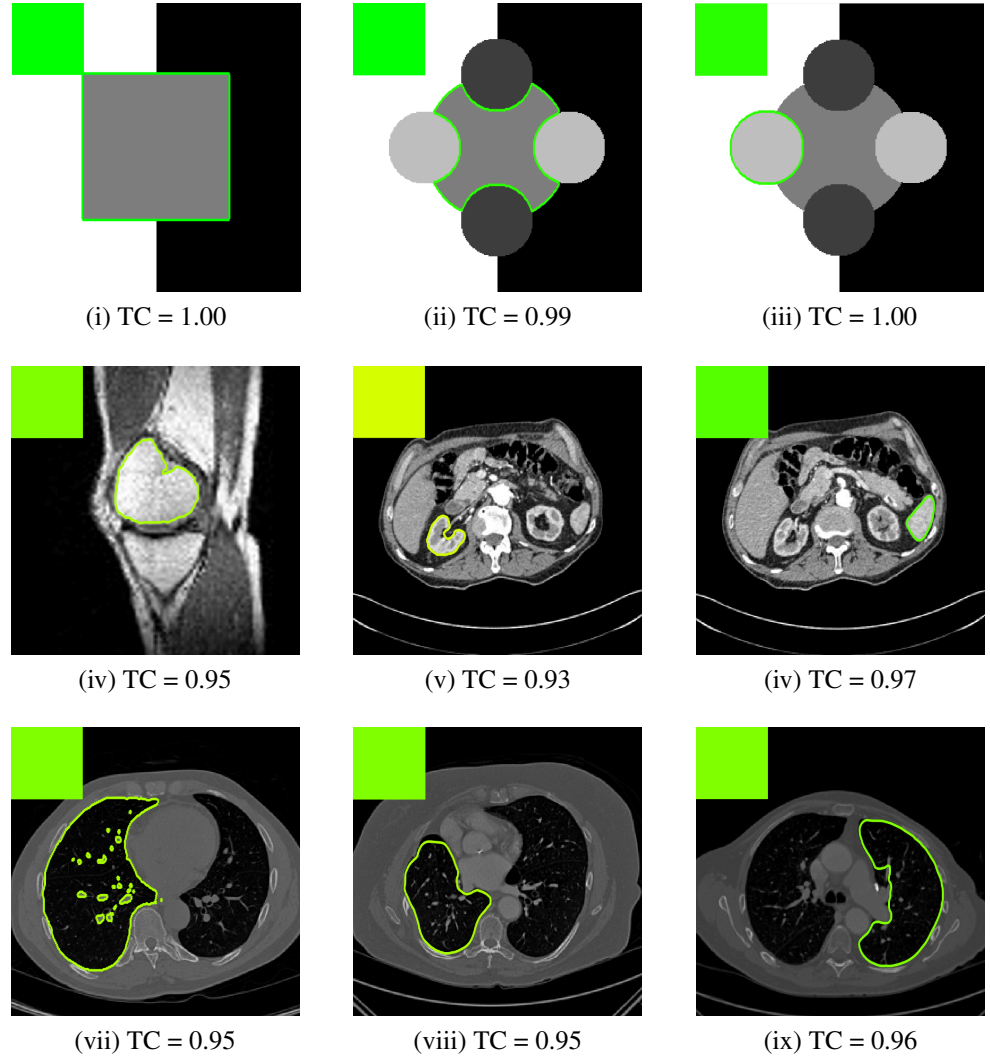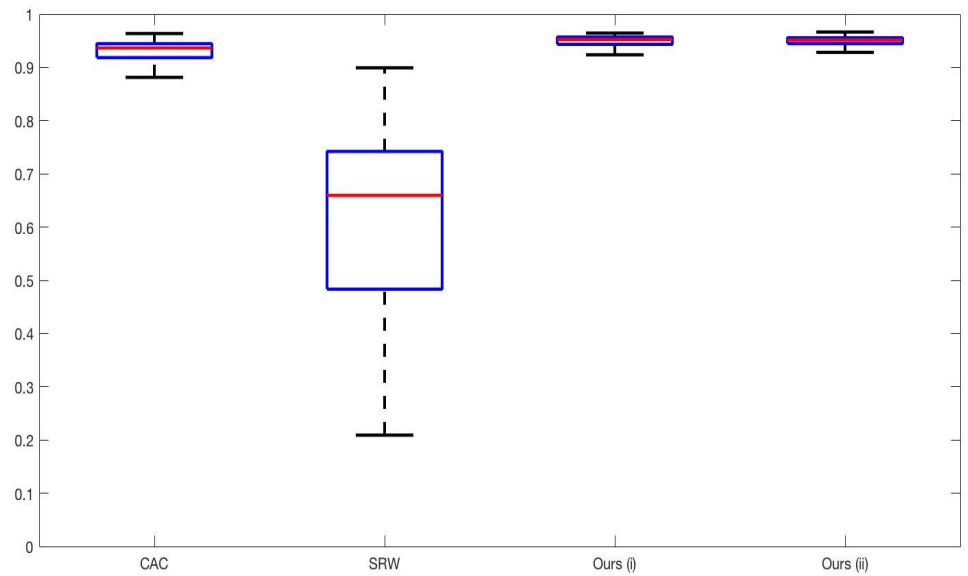

(i) TC = 1.00 (ii) TC = 0.99 (iii) TC = 1.00

(iv) TC = 0.95 (v) TC = 0.93 (iv) TC = 0.97

(vii) TC = 0.95 (viii) TC = 0.95 (ix) TC = 0.96

Fig. 17: Boxplots of the TC values comparing our method to CAC [29] and SRW [17] for 30 test images. Ours (i) refers to using identical user input to CAC and SRW, with a sample shown in Fig. 18. Ours (ii) refers to 1000 random variations of the user input for each image.
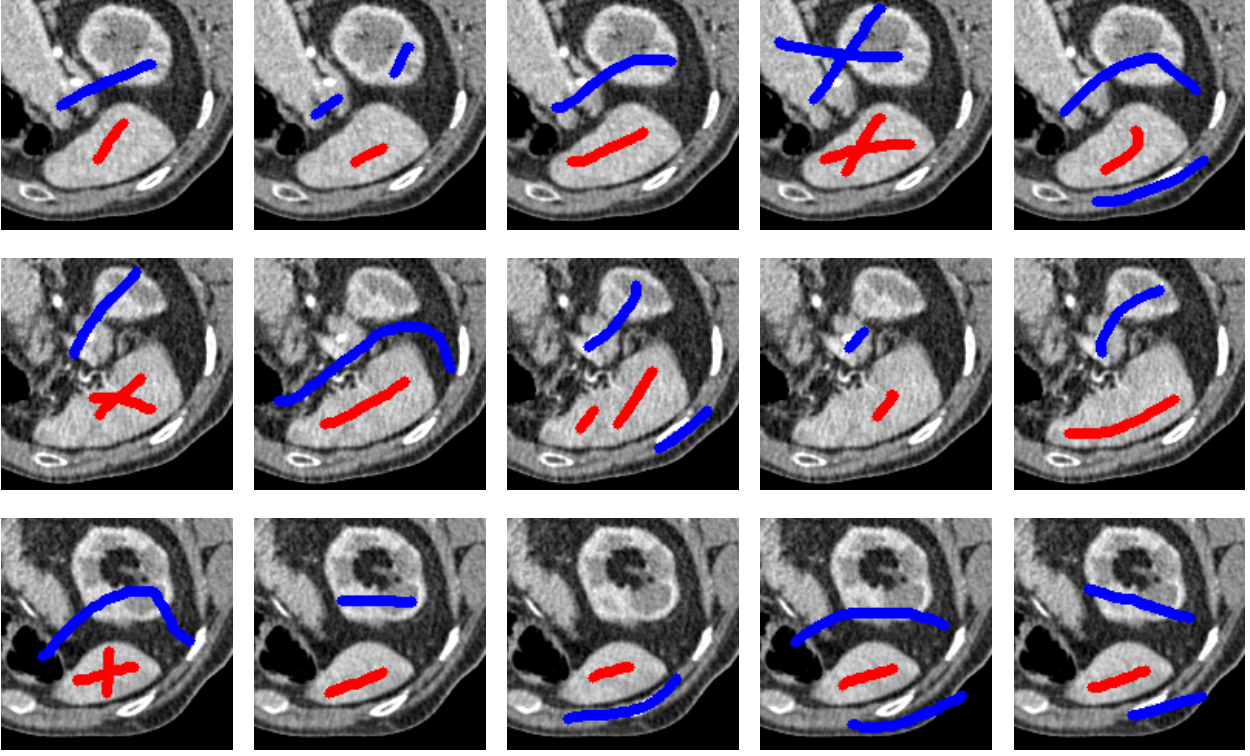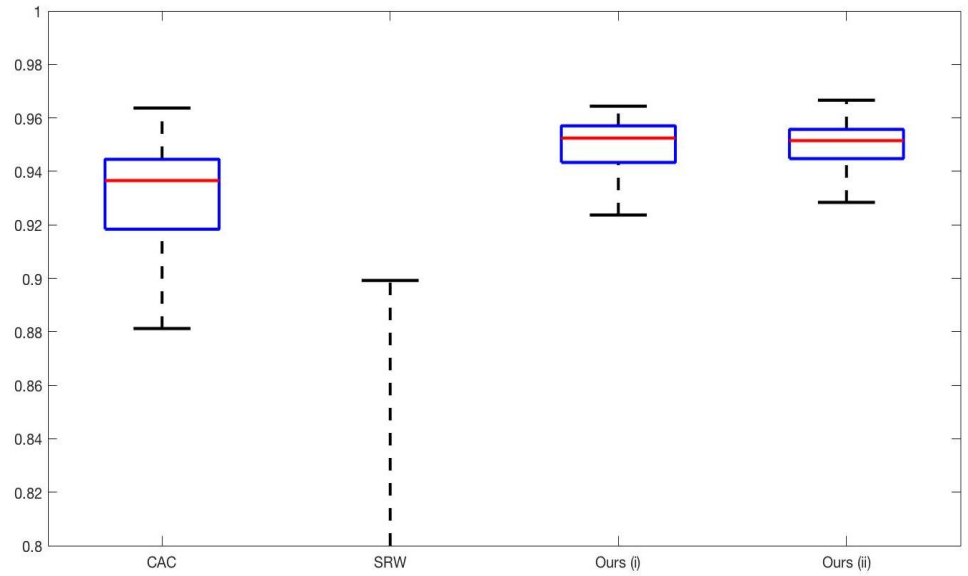
Fig. 18: Examples of the input used to compare our method to CAC [29] and SRW [17]. Each row represents an image in the dataset and we present five variations of the input used in the tests described in §7.4.

Fig. 19: Boxplots of the TC values from Fig. 17 for TC ∈ [0.8, 1]. Here, the extent to which the proposed method outperforms CAC [29] is clearer for both types of input.

6. Bertsekas, D.P.: Constrained optimization and Lagrange multiplier methods. Academic press (2014)
7. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning **3**(1), 1–122 (2011)
8. Bresson, X., Esedoglu, S., Vandergheynst, P., Thiran, J.P., Osher, S.: Fast global minimization of the active contour/snake model. Journal of Mathematical Imaging and Vision **28**(2), 151–167

(2007)
9. Brox, T., Weickert, J.: Level set segmentation with multiple regions. IEEE Transactions on Image Processing **15**(10), 3213–3218 (2006)
10. Cai, X., Chan, R., Zeng, T.: 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)

11. Chambolle, A.: An algorithm for total variation minimization and applications. Journal of Mathematical Imaging and Vision **20**, 89–97 (2004)

12. Chambolle, A., Pock, T.: A first-order primal-dual algorithm for convex problems with applications to imaging. Journal of Mathematical Imaging and Vision **40**, 120–145 (2011)

13. Chambolle, A., Pock, T.: An introduction to continuous optimization for imaging. Acta Numerica **25**, 161–319 (2016)

14. Chan, T., Esedoḡlu, S., Nikolova, M.: Algorithms for finding global minimizers of image segmentation and denoising models. SIAM Journal on Applied Mathematics **66**(5), 1632–1648 (2006)

15. Chan, T., Vese, L.: Active contours without edges. IEEE Transactions on Image Processing **10**(2), 266–277 (2001)

16. Chen, D., Yang, M., Cohen, L.: Global minimum for a variant Mumford-Shah model with application to medical image segmentation. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization **1**(1), 48–60 (2013)

17. Dong, X., Shen, J., Shao, L.: Submarkov random walk for image segmentation. IEEE Transactions on Image Processing **25**(2), 516–527 (2016)

18. Falcao, A., Udupa, J., Migazawa, F.: An ultrafast user-steered image segmentation paradigm: live wire on the fly. IEEE Transactions on Medical Imaging **19**(1), 55–62 (2002)

19. Goldstein, T., Bresson, X., Osher, S.: Geometric applications of the split bregman method. Journal of Scientific Computing **45**(1-3), 272–293 (2010)

20. Gordeziani, D., Meladze, G.V.: The simulation of the third boundary value problem for multidimensional parabolic equations in an arbitrary domain by one-dimensional equations. Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki **14**(1), 246–250 (1974)

21. Gout, C., Guyader, C.L., Vese, L.: Segmentation under geometrical conditions with geodesic active contours and interpolation using level set methods. Numerical Algorithms **39**, 155–173 (2005)

22. Grady, L.: Random walks for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **28**(11), 1768–1783 (2006)

23. Jaccard, P.: The distribution of the flora in the alpine zone.1. New Phytologist **11**(2), 37–50 (1912)

24. Klodt, M., Steinbrücker, F., Cremers, D.: Moment constraints in convex optimization for segmentation and tracking. In: Advanced Topics in Computer Vision, pp. 215–242. Springer (2013)

25. Li, C., Kao, C., Gore, J., Ding, Z.: Minimization of region-scalable fitting energy for image segmentation. IEEE Transactions on Image Processing **17**(10), 1940–1949 (2008)

26. Liu, C., Ng, M.K.P., Zeng, T.: Weighted variational model for selective image segmentation with application to medical images. Pattern Recognition **76**, 367–379 (2018)

27. Lu, T., Neittaanmäki, P., Tai, X.C.: A parallel splitting up method and its application to navier-stokes equations. Applied Mathematics Letters **4**(2), 25 – 29 (1991)

28. Mumford, D., Shah, J.: Optimal approximation by piecewise smooth functions and associated variational problems. Communications on Pure and Applied Mathematics **42**, 577–685 (1989)

29. Nguyen, T., Cai, J., Zhang, J., Zheng, J.: Robust interactive image segmentation using convex active contours. IEEE Transactions on Image Processing **21**, 3734–3743 (2012)

30. Osher, S., Sethian, J.: Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. Journal of Computational Physics **79**(1), 12–49 (1988)

31. Otsu, N.: A threshold selection method from gray-level histograms. IEEE Transactions on Systems, Man, and Cybernetics **9**(1), 62–66 (1979)

32. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence **12**(7), 629–639 (1990)

33. Rada, L., Chen, K.: Improved Selective Segmentation Model Using One Level-Set. Journal of Algorithms & Computational Technology **7**(4), 509–540 (2013)

34. Roberts, M., Chen, K., Irion, K.L.: A convex geodesic selective model for image segmentation. Journal of Mathematical Imaging and Vision (2018). DOI 10.1007/s10851-018-0857-2. URL https://doi.org/10.1007/s10851-018-0857-2

35. Rother, C., Kolmogorov, V., Blake, A.: Grabcut: Interactive foreground extraction using iterated graph cuts. ACM SIGGRAPH **23**(3), 1–6 (2004)

36. Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. Physica D: nonlinear phenomena **60**(1-4), 259–268 (1992)

37. Shen, J., Du, Y., Wang, W., Li, X.: Lazy random walks for superpixel segmentation. IEEE Transactions on Image Processing **23**(4), 1451–1462 (2014)

38. Spencer, J., Chen, K.: A convex and selective variational model for image segmentation. Communications in Mathematical Sciences **13**(6), 1453–1472 (2015)

39. Spencer, J., Chen, K.: Stabilised bias field: Segmentation with intensity inhomogeneity. Journal of Algorithms and Computational Technology **10**(4), 302–313 (2016)

40. Vese, L.A., Chan, T.F.: A multiphase level set framework for image segmentation using the mumford and shah model. International Journal of Computer Vision **50**(3), 271–293 (2002)

41. Wang, X., Huang, D., Xu, H.: An efficient local Chan-Vese model for image segmentation. Pattern Recognition **43**(3), 603–618 (2010)

42. Weickert, J., Ter Haar Romeny, B.M., Viergever, M.A.: Efficient and reliable schemes for nonlinear diffusion filtering. IEEE Transactions on Image Processing **7**(3), 398–410 (1998)

43. Xu, N., Price, B., Cohen, S., Yang, J., Huang, T.S.: Deep interactive object selection. In: IEEE Conference on Computer Vision and Pattern Recognition (2016)

44. Yuan, J., Bae, E., Tai, X., Boykov, Y.: A spatially continuous max-flow and min-cut framework for binary labeling problems. Numerische Mathematik **126**(3), 559–587 (2013)

45. Zhang, K., Zhang, L., Song, H., Zhou, W.: Active contours with selective local or global segmentation: a new formulation and level set method. Image and Vision Computing **28**(4), 668–676 (2010)