A survey on image segmentation methods

Yuri Iliash [1] Alexandra Yurova [2]

[1] Institut für Mathematik
Universität Augsburg

[2] Sechenov University

October 7, 2019

Image segmentation problem

Image segmentation problem:

find a partition of an image into semantically important parts.

Approaches to image segmentation:

- Variational approach
- Stochastic approach
- Neural networks / Deep learning

Segmentation types

Segmentation problem may be loosely divided into two main types:

- Detecting contours of the objects (edges detection)
- Detecting distinct homogeneous regions (Mumford–Shah model)

Mumford–Shah model

Idea: constructing a new image, close to original, made up of distinct homogeneus regions (“cartoon approximation”).

Original image and its cartoon approximation (from Strekalovskiy/Cremers)

Problem: Find a decomposition $\Omega = \Omega_1 \cup \Omega_2 \cup \ldots \Omega_n \cup K$, $K \subset \mathbb{R}^{d-1}$:

- Image intensity varies slowly and smoothly inside each Ω_i
- Image intensity may vary rapidly or discontinuously across K
- Set of internal boundaries K should be “small”.

Mumford–Shah functional

Given an image function $f \in L^2(\Omega)$, find $(u, K) = \arg\min E_{MS}(u, K)$, where

$$E_{MS}(u, K) = \underbrace{\int_{\Omega} (u - f)^2 \, dx}_{\text{fidelity to image}} + \underbrace{\alpha \int_{\Omega \setminus K} |\nabla u|^2 \, dx}_{\text{smoothness within segments}} + \underbrace{\beta \int_{K} d\sigma}_{\text{internal edges measure}}, \qquad \text{(MS)}$$

where α and β are positive parameters, $u \in H^1(\Omega \setminus K)$.

The functional is "minimal":

- without fidelity term, trivial solution: $u(x) = \text{const}$, $K = \emptyset$;
- without smoothness term, trivial solution: $u(x) = f(x)$, $K = \emptyset$;
- without internal edges term: $\inf E_{MS}(u, K) = 0$
  (for example, $u(x)$ piecewise constant, approximating $f(x)$).

Solution properties

Conjecture (Mumford–Shah)

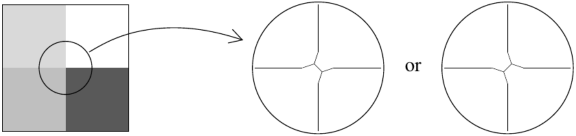
There exists a minimizer of E_{MS} such that K is the union of a finite set of $C^{1,1}$ simple curves: $K = \bigcup_{i=1}^{N} \gamma_i$.

Theorem

Let Mumford–Shah conjecture holds true, $d = 2$, and (u, K) is a minimizer of E_{MS}. Then the only vertices of K are:

- *points on $\partial\Omega$ where γ_i meets $\partial\Omega$ perpendiculary;*
- *triple junctions where three γ_i meet with angle $2\pi/3$ between each pair;*
- *crack-tips where γ_i ends.*

Solution is not unique.



Euler–Lagrange equations for Mumford–Shah functional

Theorem

Let (u, K) be a solution of (MS)*, satisfying the Mumford–Shah conjecture. Then*

$$\begin{aligned} \alpha\Delta u &= u - f \quad \text{on } \Omega \setminus K, \\ \frac{\partial u}{\partial \boldsymbol{n}} &= 0 \qquad\quad \text{on } \partial\Omega \cup \gamma_i, \end{aligned} \tag{1}$$

$$e(u^+) - e(u^-) + \beta\kappa(\gamma_i) = 0 \quad \text{on } \gamma_i, \tag{2}$$

where $e(u) = (u - f)^2 + \alpha|\nabla u|^2$*,* u^+ *and* u^- *are the traces of* u *on each side of* γ_i*,* $\kappa(\gamma_i)$ *is the (mean) curvature of* γ_i*.*

Discretization of the unknown set $K \subset \mathbb{R}^{d-1}$ is a complex task.

Ambrosio–Tortorelli approximation

Phase-field approach: “diffuse interface” instead of “sharp interface”:

- Introduce a function $z \approx \mathbb{1}_{\Omega\setminus K}$, $0 \leq z(x) \leq 1$.
- $z \approx 0$ if x close to the internal boundary γ_i, othewise $z \approx 1$.
- Approximate $E_{MS}(u, K)$ by a sequence $E_\varepsilon \to E_{MS}$, $\varepsilon \to 0$ in some sense.

Ambrosio–Tortorelli functional:

$$E_\varepsilon(u, z) = \int\limits_\Omega (u - f)^2\, dx + \alpha \int\limits_\Omega z^2 |\nabla u|^2\, dx + \beta \int\limits_\Omega \left(\varepsilon |\nabla z|^2 + \frac{(1-z)^2}{4\varepsilon} \right) dx,$$

- Integrand in the second term vanishes at the vicinity of K.
- Third term approximates the internal boundary length.
- $\Gamma\text{-}\lim\limits_{\varepsilon\to 0} E_\varepsilon(u, z) = \begin{cases} E_{MS} & \text{if } z = 1 \text{ a.e. in } \Omega \\ +\infty & \text{overwise} \end{cases}$ w.r.t. $(L^1(\Omega))^2$ topology.

Euler–Lagrange equations for Ambrosio–Tortorelli functional

$$
\begin{aligned}
(u - f) - \alpha \nabla (z^2 \nabla u) &= 0 \quad \text{in } \Omega, \\
\alpha z |\nabla u|^2 - \beta \varepsilon \Delta z + \frac{\beta}{4\varepsilon}(z - 1) &= 0 \quad \text{in } \Omega, \\
\frac{\partial u}{\partial \boldsymbol{n}} &= 0 \quad \text{on } \partial\Omega, \\
\frac{\partial z}{\partial \boldsymbol{n}} &= 0 \quad \text{on } \partial\Omega.
\end{aligned}
$$

- System of two nonlinear elliptic equations
- Finite element of finite difference approximation
- Newton method
- Nonconvex

Convex relaxation [Kee, Kim]

$E_\varepsilon(u, z)$ nonconvex due to the term $F(u, z) = \int\limits_\Omega z^2 |\nabla u|^2 \, dx$.

Idea: convex relaxation.

Theorem (McCormick relaxation)

Let u be Lipschitz continuous with constant L. Then

$$F_{rel}(u, z) = \int\limits_\Omega \max(0, L^2 z^2 + |\nabla u|^2 - L^2) \, dx \le F(u, z),$$

where $F_{rel}(u, z)$ is convex in $\left(W^{1,2}(\Omega) \cap W^{1,\infty}(\Omega)\right) \times W^{1,2}(\Omega)$.

Detecting object boundaries

Sharp variation of the image intensity on the boundary of an object
$\implies |\nabla f|$ is high on the boundary

Edge-detector function:

- $g : [0, +\infty[\to]0, 1]$
- g is strictly decreasing
- $g(0) = 1, \lim_{s \to +\infty} g(s) = 0$

A typical choice: $g(s) = \dfrac{1}{1+s^2}$,

f $|\nabla f|$ $g(|\nabla f|)$

Kass–Witkin–Terzopoulos model I

- Let $K = \bigcup_i \gamma_i$, where γ_i is a piecewise $C^{1,1}$ closed simple curve: $\gamma(q) : [a, b] \to \Omega, \gamma(a) = \gamma(b)$.
- Energy of the curve:

$$J(\gamma) = \underbrace{\int_a^b |\gamma'(q)|^2 \, dq + \alpha \int_a^b |\gamma''(q)|^2 \, dq}_{\text{spline energy}} + \lambda \underbrace{\int_a^b g^2(|\nabla f(\gamma(q))|) \, dq}_{\text{external energy}}. \quad \text{(KWT)}$$

- Spline energy: imposes smoothness constraint (membrane + thin plate).
- External energy: attracts the curve toward the object edges.
- Euler–Lagrange equation:

$$\frac{\delta J(\gamma)}{\delta \gamma} = 0 \quad + \quad \text{b.c.}$$

Kass–Witkin–Terzopoulos model II

Numerical solution: introduce an artificial parameter (time) and solve evolution equation (gradient descent):

$$\begin{aligned} \frac{\partial \gamma}{\partial t}(q,t) &= -\frac{\delta J(\gamma)}{\delta \gamma}(q,t), \\ \gamma(q,\,0) &= \gamma_0(q), \\ &+ \text{boundary conditions.} \end{aligned}$$

Drawbacks:

- $J(\gamma)$ is non-convex.
- $J(\gamma)$ is not intrinsic, it depens on the parametrization of γ.
- Method does not handle changes of topology: it is not possible to detect more than one object.
- It is not possible to detect nonconvex object.
- Numerical problems in solving.

Geodesic Active Contours model

Set $\alpha = 0$ in (KWT) (the curvature term is redundant):

$$J_1(\gamma) = \int_a^b |\gamma'(q)|^2 \, dq + \lambda \int_a^b g^2(|\nabla f(\gamma(q))|) \, dq.$$

Introduce an intrinsic functional, which in some sense equivalent to J_1:

$$J_2(\gamma) = 2\sqrt{\lambda} \int_a^b g(|\nabla f(\gamma(q))|)|\gamma'(q)| \, dq.$$

The curve energy, defined by functional J_2, does not depend on parametrisation. Weighted Euclidean length:

$$J_{GAC}(\gamma) = \int_0^{L(\gamma)} g(|\nabla f(\gamma(s))|) \, ds. \qquad \text{(GAC)}$$

Contour flow equation

Variational derivative of (GAC):

$$\frac{\delta J_{GAC}(\gamma)}{\delta \gamma} = \left(\langle \nabla g,\, \boldsymbol{n} \rangle - \kappa g\right) \boldsymbol{n}.$$

Gradient descent flow:

$$\frac{\partial \gamma}{\partial t} = \left(\kappa g - \langle \nabla g,\, \boldsymbol{n} \rangle\right) \boldsymbol{n},$$
$$\gamma(q,\, 0) = \gamma_0(q).$$

If $g \equiv 1$: mean curvature motion

$$\frac{\partial \gamma}{\partial t} = \kappa \boldsymbol{n}.$$

To make the detection of nonconvex objects easier and increase the convergence speed, add an additional term:

$$\frac{\partial \gamma}{\partial t} = \left(\kappa g - \langle \nabla g,\, \boldsymbol{n} \rangle + \beta g\right) \boldsymbol{n},$$

such that $\kappa + \beta > 0$.

Snakes in action

Problems with snakes

General contour normal flow equation:

$$\frac{\partial \gamma}{\partial t} = F(t,\, \gamma,\, \gamma',\, \gamma'')\,\boldsymbol{n},$$
$$\gamma(q,\, 0) = \gamma_0(q).$$

In the Geodesic Active Contours model

$$F(\gamma',\, \gamma'') = (\kappa + \beta)g - \langle \nabla g,\, \boldsymbol{n} \rangle\,.$$

Main difficulties in the numerical treatment:

- Very hard to tackle topology changes of γ
- Nontrivial choice of marker points for discretizing the curve: in the evolution process concentration and void regions can be created

Remedy: the Level-Set method.

Level-Set method [Oscher, Sethian]

Idea: γ can be seen as the zero-level of some function $u : \mathbb{R}^+ \times \mathbb{R}^2 \to \mathbb{R}$

$$u(t,\, \gamma(q,t)) = 0, \quad \forall q,\ \forall t \geq 0.$$

Differentiating w.r.t. t and taking into account contour flow equation leads to

$$\frac{\partial u}{\partial t} + \langle \nabla u,\, F\,\boldsymbol{n} \rangle = 0.$$

Substituting $\boldsymbol{n} = -\nabla u / |\nabla u|$ ($u < 0$ inside the contour), $\kappa = -\operatorname{div} \boldsymbol{n}$, obtain:

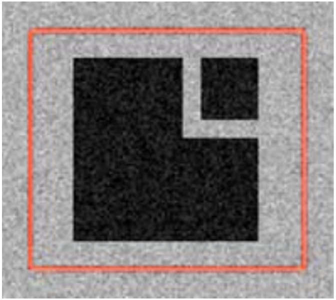
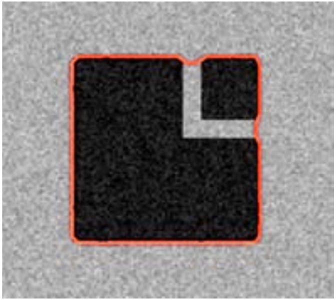
$$\frac{\partial u}{\partial t} = g(|\nabla f|) \left(\operatorname{div} \frac{\nabla u}{|\nabla u|} + \beta \right) |\nabla u| + \langle \nabla g,\, \nabla u \rangle\,,$$

$$\frac{\partial u}{\partial \boldsymbol{n}} = 0, \qquad u(0,\, x) = d(x, \gamma_0),$$

where $d(x, \gamma_0)$ is the signed distance to the initial contour γ_0.

Bad snakes

Non-convex problem: bad initial contour can give unsatisfactory results.

Global minimization of GAC model [Bresson et al.]

Introduce the weighted TV-norm:

$$TV_g(u) = \int_\Omega g(x)|\nabla u|\,dx, \quad u \in BV(\Omega).$$

This norm is closely related to (GAC), namely

$$TV_g(\mathbb{1}_{\Omega_\gamma}) = \int_\Omega g(x)|\nabla \mathbb{1}_{\Omega_\gamma}|\,dx = \int_\gamma g(s)ds = J_{GAC}(\gamma), \quad \gamma = \partial\Omega_\gamma.$$

Concider the following convex functional:

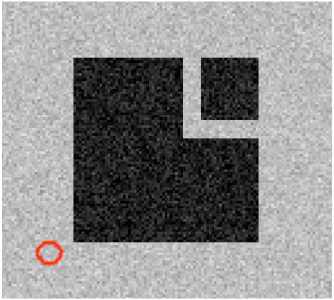
$$J_{TV\text{-}L^1}(u) = TV_g(u) + \lambda \int_\Omega |u - f|\,dx.$$

If the class of characteristic functions minimizing $J_{TV\text{-}L^1}$ is equivalent to minimization of J_{GAC}, while approximating f in L^1.

Works good for binary images with noise.

Good snakes

Good result even with bad initial contour.



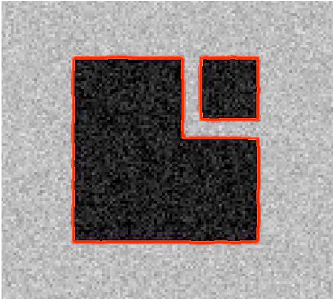


Image classification

Image classification problem is the segmentation problem with a labeling:

- partition image domain into homogenious regions:

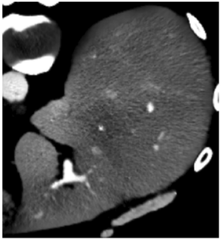
$$\bar{\Omega} = \bigcup_{i=1}^{N} (\Omega_i \cup \Gamma_i), \quad \Omega_i \cap \Omega_j = \emptyset, \ i \neq j$$

- each region belongs to one of several predefined classes

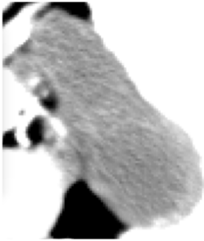
Discriminating criterion is based on certain image features, for example:

- Gaussian intensity distribution with known parameters $N(\mu_i, \sigma_i)$
- texture parameter
- entropy, contrast, etc…

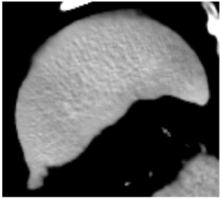
Textures



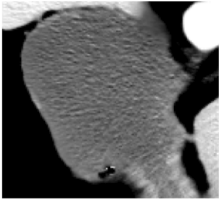
Liver



Gallbladder

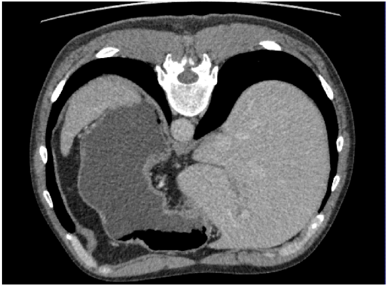


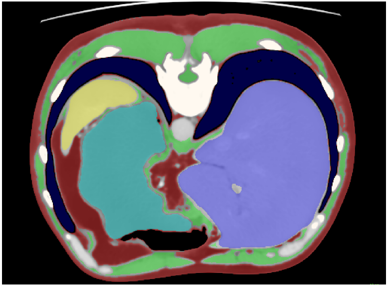
Spleen



Stomach

Example: abdominal cavity organs classification



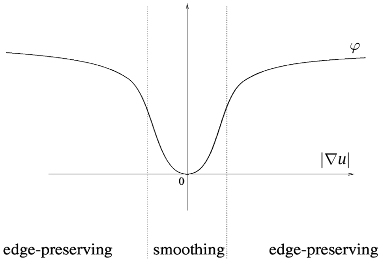


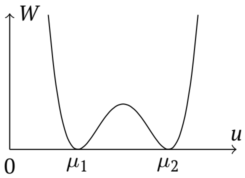
CT-scan cross-section of abdominal cavity organs, and classification: spleen, stomach, liver, lungs, bones

Variational model for image classification

The model, suggested by C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zerubia

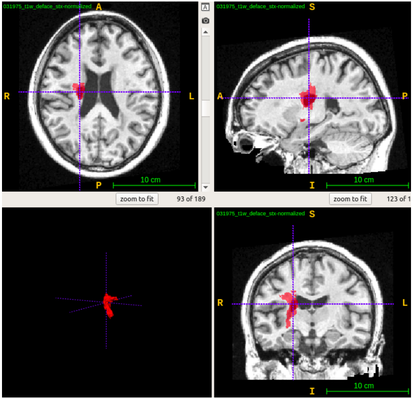
$$J_\varepsilon(u) = \underbrace{\int_\Omega (u-f)^2\,dx}_{\text{fidelity term}} + \underbrace{\varepsilon\alpha \int_\Omega \varphi(|\nabla u|)\,dx}_{\text{smoothness/restoration term}} + \underbrace{\frac{\beta}{\varepsilon}\int_\Omega W(u;\ \mu,\sigma)dx.}_{\text{classification term}} \qquad \text{(CH)}$$





Data source

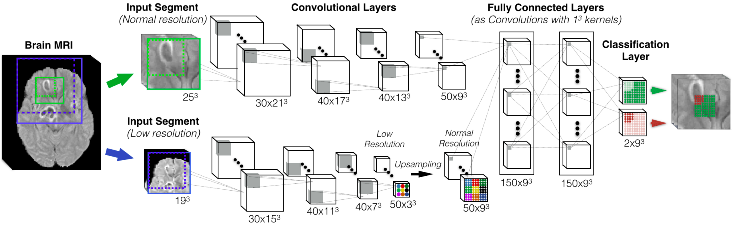
- ATLAS (Anatomical Tracings of Lesions After Stroke)
  T1-weighted MRI scans (n=220) with manually segmented lesions
- $197 \times 233 \times 189$



Convolutional Neural Network

- C_l channels (feature maps) on each layer
- kernels size $3 \times 3 \times 3$

$$\boldsymbol{y}_l^m = f\left(\sum_{n=1}^{C_{l-1}} \boldsymbol{k}_l^{m,n} * \boldsymbol{y}_{l-1}^n + b_l^m\right)$$



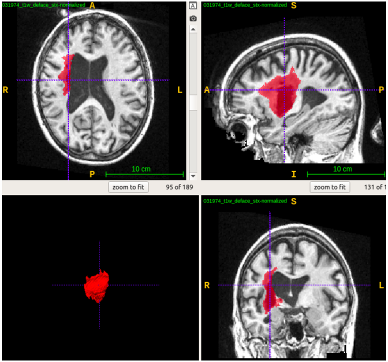
Experiments on ATLAS datasets

- Training (206 datasets)
- Testing (14 datasets)
- Dice Similarity Coefficient:

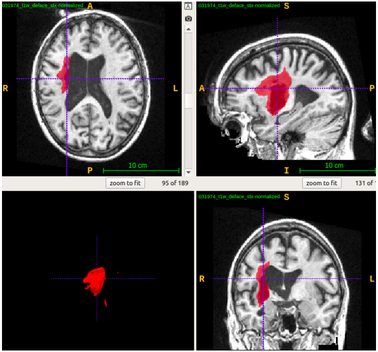
$$DSC = \frac{2|X \cap Y|}{|X| + |Y|}$$

Experiments on ATLAS datasets

Ground truth



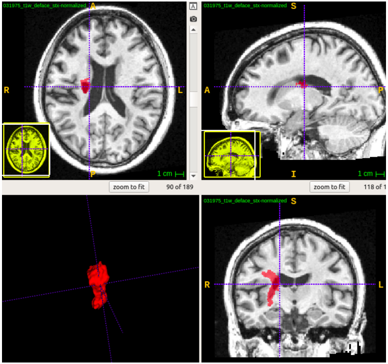
Result



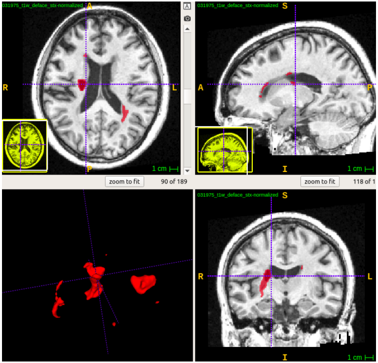
$$DSC = 0.8052$$

Experiments on ATLAS datasets

Ground truth



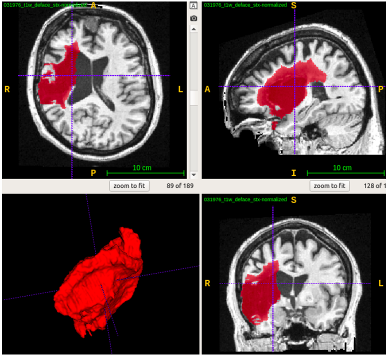
Result



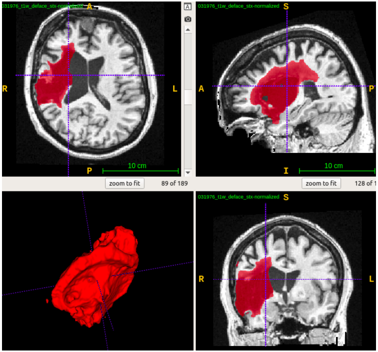
$$DSC = 0.4752$$

Experiments on ATLAS datasets

Ground truth



Result



$$DSC = 0.9304$$