Segmentation and Classification using Deep Learning Technologies

Workshop DIPOpt, Lyon, 2023

Xiaohao Cai
VLC / ECS / UoS
AI - ML – CV – IP – MI

Xiaohao Cai
VLC / ECS / UoS

- Computer vision
- Image processing
- Machine/deep learning
- Optimisation

Interdisciplinary

- Computer science
- Mathematics/statistics
- Physics, life science, etc.

Browse State-of-the-Art:
https://paperswithcode.com/sota

State-of-the-Art
Computer Vision

Data Size

1 100 1K 10K 100K 1M 10M > 10M

Methods

- Model-driven
- Data-driven

Interpretability

- Strong
- Weak
- Black-box
Conclusions

Current evidence on the use of AI systems in breast cancer screening is a long way from having the quality and quantity required for its implementation into clinical practice. Well designed comparative test accuracy studies, randomised controlled trials, and cohort studies in large screening populations are needed which evaluate commercially available AI systems in combination with radiologists. Such studies will enable an understanding of how AI can support radiologists.
It never happened—but not for lack of effort. Research teams around the world stepped up to help. The AI community, in particular, rushed to develop software that many believed would allow hospitals to diagnose or triage patients faster, bringing much-needed support to the front lines—in theory.

In the end, many hundreds of predictive tools were developed. None of them made a real difference, and some were potentially harmful.

That’s the damning conclusion of multiple studies published in the last few months. In June, the Turing Institute, the UK’s national center for data science and
In 2018, the Lords Committee on AI called for the development of AI systems that are “intelligible to developers, users and regulators”. It recommended that an AI system that could have a substantial impact on an individual’s life should not be used unless it can produce an explanation of its decisions. In a
What it means to look inside the black box

Explainability
Understanding reasoning behind each decision

Transparency
Understanding of AI model decision making

Provability
Mathematical certainty behind decisions

Source: PwC
“DATA IS THE NEW GOLD”
Methods

Data-driven vs. Model-driven

Interpretability

Strong vs. Weak vs. Black-box

Dataset Size

State of the art

Computer Vision

Medical Imaging Problems

10-20
100-1000
1 - 10 K
100 - 200 K
> 1.0 M (ImageNet)

Transfer learning
Few-shot learning

Train models from scratch

Browse State-of-the-Art:
https://paperswithcode.com/sota

[Credit Niranjan]
Few-Shot Learning Framework and Subspaces

[Data-driven part]

14 datasets covering 11 distinct diseases, with number of classes ranging from 2 to 11.

[Model-driven part]

Subspaces Representations
- SVD
- DA
- NMF

K-Nearest Neighbour Classifier

[1. Liu, K. Fan, X. Cai, and M. Niranjan, under review]
Classification

K-Nearest Neighbor (KNN)

Category A: 3 neighbors
Category B: 2 neighbors
New Data point

Support Vector Machine (SVM)

Optimal hyperplane
Maximum margin

https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning

https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning
Classification

Hyperspectral

LiDAR

Observed data

Ours
Active Contours Without Edges

Tony F. Chan, Member, IEEE, and Luminita A. Vese

Abstract—In this paper, we propose a new model for active contours to detect objects in a given image, based on techniques of curve evolution, Mumford–Shah functional for segmentation and level sets. Our model can detect objects whose boundaries are not necessarily defined by gradient. We minimize an energy which can be seen as a particular case of the minimal partition problem. In the level set formulation, the problem becomes a “mean-curvature flow”–like evolving the active contour, which will stop on the desired boundary. However, the stopping term does not depend on the gradient of the image, as in the classical active contour models, but is instead related to a particular segmentation of the image. We will give a numerical algorithm using finite differences. Finally, we will present various experimental results and in particular some the image (the external energy). Observe that, by minimizing the energy (1), we are trying to locate the curve at the points of maxima \( \nabla u_0 \), acting as an edge-detector, while keeping a smoothness in the curve (object boundary).

A general edge-detector can be defined by a positive and decreasing function \( g \), depending on the gradient of the image \( u_0 \), such that

\[
\lim_{x \to u_0} g(x) = 0.
\]

For instance
**T-ROF (Thresholded-ROF)**

**Image Restoration**

**ROF model**

\[ \min_{u \in BV(\Omega)} \left\{ TV(u) + \frac{\mu}{2} \int_{\Omega} (f - u)^2 dx \right\}, \]

**TV(u): total variation of u**

**Theorem**

(Relation between ROF and Chan-Vese model) Let \( \Omega \) be a simply connected bounded domain. For given \( 0 < m_0 < m_1 \leq 1 \), let \( \tilde{\Sigma} := \{ x \in \Omega : 0 < |\tilde{\Sigma}| \} \). Then \( \tilde{\Sigma} \) is a minimizer of the Chan-Vese model if \( m_0 = \text{mean}_f(\Omega \setminus \tilde{\Sigma}) \) and \( m_1 = \text{mean}_f(\tilde{\Sigma}) \).

<table>
<thead>
<tr>
<th>Method</th>
<th>Acc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MBO ('14)</td>
<td>99.1</td>
</tr>
<tr>
<td>GC-1 ('14)</td>
<td>98.4</td>
</tr>
<tr>
<td>GC-2 ('17)</td>
<td>98.7</td>
</tr>
<tr>
<td>GC-3 ('18)</td>
<td>98.6</td>
</tr>
<tr>
<td>GC-4 ('18)</td>
<td>98.4</td>
</tr>
<tr>
<td><strong>Ours</strong></td>
<td>99.4</td>
</tr>
</tbody>
</table>
Models

Models proposed in our work, e.g.:

\[ \min_x \left\{ \frac{\lambda}{2} \| y - Ax \|_2^2 + \| Wx \|_1 \right\} \]

\[ \min_g \left\{ \frac{\lambda}{2} \| f - Ag \|_2^2 + \frac{\mu}{2} \| \nabla g \|_2^2 + \| \nabla g \|_1 \right\} \]

\[ \mu \Phi(f, Ag) + \lambda \psi(g, u_i, c_i) + \sum_{i=1}^{K} \int_{\Omega} |\nabla u_i| \]

s.t. \quad \sum_{i=1}^{K} u_i(x) = 1, u_i(x) \in \{0, 1\}

\[ \min_{u \in S} \left\{ \frac{1}{2} \| f - Bu \|_2^2 + \lambda \| \nabla u \|_0 \right\} \]

\[ \min_{\psi} \left\{ D[T(\psi), R] + \alpha \| \Delta \psi \|_2^2 \right\} \]

Convex optimisation algorithms

- ADMM
- Primal-dual
- Split-Bregman
- Augmented Lagrangian

Sparse regularizations

- \( \| \cdot \|_0, \| \cdot \|_1, \| \cdot \|_2 \)
- with \( \nabla, \Delta, W \)
- \( W \): Wavelet transform
Proximal nested sampling for high-dimensional Bayesian model selection

Xiaohao Cai$^{1,2}$ · Jason D. McEwen$^{1,3}$ · Marcelo Pereyra$^4$

Received: 15 April 2022 / Accepted: 9 September 2022 / Published online: 5 October 2022
© The Author(s) 2022
Proximal nested sampling with data-driven priors for physical scientists

Jason D. McEwen 1,2*, Tobias I. Liaudat 1,3, Matthew A. Price 1, Xiaohao Cai 4 and Marcelo Pereyra 5

1 Mullard Space Science Laboratory, University College London (UCL), Dorking, RH5 6NT, UK;
2 Alan Turing Institute, London, NW1 2DB, UK;
3 Department of Computer Science, University College London (UCL), London, WC1E 6BT, UK;
4 School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK;
5 School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, EH14 4AS, UK;
* Correspondence: jason.mcewen@gmail.com

Abstract: Proximal nested sampling was introduced recently to open up Bayesian model selection for high-dimensional problems such as computational imaging. The framework is suitable for models with a log-convex likelihood, which are ubiquitous in the imaging sciences. The purpose of this article is two-fold. First, we review proximal nested sampling in a pedagogical manner in an attempt to elucidate the framework for physical scientists. Second, we show how proximal nested sampling can be extended in an empirical Bayes setting to support data-driven priors, such as deep neural networks learned from training data.

Keywords: Bayesian model selection; nested sampling; proximal calculus.
Model Selection

(a) Ground truth  (b) Dirty  (c) Hand-crafted prior  (d) Data-driven prior

\[
\text{Log evidence} = -2.96 \times 10^3 \quad \text{and} \quad -1.35 \times 10^3
\]
Subspace Representation / Embedding

Socher et al. (2013b)
Subspace feature representations (SVD vs DA)
DA (Linear discriminant analysis)

Maximise: \( R(d) = \frac{d^T S_B d}{d^T S_W d} \)

\[
S_B = \sum_{j=1}^{C} (\bar{y}_j - \bar{y})(\bar{y}_j - \bar{y})^T, \quad S_W = \sum_{j=1}^{C} S_{Wj}^j,
\]

Inter-class scatter  \quad Mean of the whole data  \quad Mean of class \( j \)  \quad No. of classes  \quad Intra-class scatter

\[
S_{Wj}^j = \sum_{k=1}^{N_j} (y_k^j - \bar{y}_j)(y_k^j - \bar{y}_j)^T,
\]
DA (Linear discriminant analysis)

Maximise: \[ \mathcal{R}(d) = \]

**GO-LDA:**

D Jiah

Abstract—Linear discriminant analysis. While linearity of class boundaries cannot be used to map complex data onto features which is variance preserving, LDA maximizes the class on a subspace. The solution to bin. It is well known that the multiclass LDA is:
NMF (Non-negative matrix factorization)

\[
\min_{U,V} \|X - UV^T\|_F^2, \text{ s.t. } U \geq 0, V \geq 0
\]
Supervised NMF (DNMF, SCNMFS)

DNMF (Discriminative NMF) [M. Babaee, 2016]:

$$\min_{U,V,A} \|X - UV^\top\|_F^2 + \alpha \|Q - AV^\top\|_F^2,$$

s.t. \(U \geq 0, V \geq 0\)

SCNMFS (Supervised and Constrained NMF with Sparseness) [Xibiao Cai, 2018]:

$$\min_{U,Z} \|X - UZ^\top Q\|_F^2 + \beta \|U\|_F^2,$$

s.t. \(U \geq 0, Z \geq 0\)
Subspace Representation / Tensor Decomposition

\[ Y = A^{(1)} \in \mathbb{R}^{I_1 \times J_1} + A^{(2)} \in \mathbb{R}^{I_2 \times J_2} + A^{(3)} \in \mathbb{R}^{I_3 \times J_3} + E \]

[Credit: Guoxu Zhou and Andrzej Cichocki]
Practical Sketching Algorithms for Low-Rank Tucker Approximation of Large Tensors

Wandi Dong¹ · Gaohang Yu¹ · Liqun Qi¹,² · Xiaohao Cai³

Received: 24 October 2022 / Revised: 27 January 2023 / Accepted: 28 February 2023 / Published online: 29 March 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023
A Randomized Block Krylov Method for Tensor Train Approximation

Gaohang Yu\textsuperscript{1*}, Jinhong Feng\textsuperscript{1}, Zhongming Chen\textsuperscript{1}, Xiaohao Cai\textsuperscript{2}, Liqun Qi\textsuperscript{1,3}

\textsuperscript{1}Department of Mathematics, Hangzhou Dianzi University, China. 
\textsuperscript{2}School of Electronics and Computer Science, University of Southampton, Southampton, UK. 
\textsuperscript{3}Huawei Theory Research Lab, Hong Kong, China.

*Corresponding author(s). E-mail(s): maghyu@163.com; Contributing authors: fengjinhong0502@163.com; zmchen@hdu.edu.cn; x.cai@soton.ac.uk; liqun.qi@polyu.edu.hk;
Subspace Representation / Tensor Decomposition

(a) Original gray video frame
(b) Noisy gray video frame
   PSNR: 7.6099
   CPU: 0.29; PSNR: 22.9434

(d) TT-rSVD
   CPU: 0.03; PSNR: 15.4260
(e) TT-rSI
   CPU: 0.07; PSNR: 20.4777
(f) TT-rBKI
   CPU: 0.10; PSNR: 22.0651
Few-Shot Learning Framework and Subspaces

Data-driven part

Model-driven part

[J. Liu, K. Fan, X. Cai, and M. Niranjan, under review]
Explainable AI (XAI) methods – Grad-CAM

Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization, IJCV, '19
Using computer vision to identify limpets from their shells: a case study using four species from the California peninsula

[Authors listed]
Multilevel Explainable Artificial Intelligence: Visual and Linguistic Bonded Explanations

Halil Ibrahim Aysel, Xiaohao Cai, and Adam Prugel-Bennett
Multilevel Explainable Artificial Intelligence: Visual and Linguistic Bonded Explanations

Halil Ibrahim Aysel, Xiaohao Cai, and Adam Prugel-Bennett
Multilevel XAI

The top five classes that maximally activate the given individual attribute on the left.
Data Augmentation

Use cGAN for data augmentation

D. Mallios and X. Cai, EUSIPCO, 2021
Data Augmentation

2018 Data Science Bowl
Find the nuclei in divergent images to advance medical discovery

Data Augmentation in Classification and Segmentation: A Survey and New Strategies
Khaled Alomar *, Halil Ibrahim Aysel © and Xiaohao Cai ©
Semantic Segmentation by Semantic Proportions

Halil Ibrahim Aysel¹*  Xiaohao Cai¹  Adam Prügel-Bennett¹
¹School of Electronics and Computer Science, University of Southampton, UK

Input: $X_i$  
Feature extraction  
$CNN$  
Features: $Y_i$  
$GAP$  
SP computation  
Predictions  
Ground-truth  
Predicted segmentation maps

[arXiv:2305.15608]
Balanced Neural Architecture Search and Its Application in Specific Emitter Identification

Mingyang Du, Xikai He, Xiaohao Cai, and Daping Bi

---

**Balanced Neural Architecture Search**

Sample architectures with probability $p$

Train the child network to convergence

The controller (RNN)

Validation accuracy

Network efficiency

Scale gradient of $p$ by $R$ to update the controller

---

Adversarial examples

NAS searching

Adversarial training

Clean data

Balanced NAS

Optimized models

Robust models

---

Block

Cell

[Diagram showing block and cell structures with operations like 'sep 1x7-7x1', 'sep 7x7', 'Softmax', 'Avg pooling', etc.]

---

[Diagram showing flow of adversarial training and network efficiency with control over gradients.]
Bayesian Attention Belief Network

Better robustness and generalisation

More efficient and effective

Structure of the proposed BABNet
Denoising and Classification

DNCNet: Deep Radar Signal Denoising and Recognition

Better robustness and generalisation
Recap:

- **Medical Imaging Problems**
  - 10-20
  - 100-1000
  - 1 – 10 K
  - 100 – 200 K
  - > 1.0 M (ImageNet)

- **State of the art Computer Vision**

- **Dataset Size**
  - Train models from scratch

- **Methods**
  - Transfer learning
  - Few-shot learning

- **Data augmentation**

- **Advanced DL models**

- **XAI**

- **Model-driven**
  - Strong

- **Data-driven**
  - Weak

- **Interpretability**
  - Black-box