

基于非监督深度学习的医学影像特征提取研究

贾开

清华大学

June 24, 2015

指导老师 宋亦旭
报告人 贾开
学号 2011011275

目录

研究背景与意义

研究内容

基于层叠卷积 ISA 的特征提取

基于深度卷积神经网络的特征提取

评测方法

实验结果

ISA

卷积神经网络

小结

总结

参考文献

目录

研究背景与意义

研究内容

 基于层叠卷积 ISA 的特征提取

 基于深度卷积神经网络的特征提取

 评测方法

实验结果

 ISA

 卷积神经网络

 小结

总结

参考文献

医学影像处理

意义重大

- ▶ 辅助医务人员进行疾病诊断
- ▶ 医疗自动化、标准化
- ▶ 医学研究
- ▶ 分割、识别、配准等经典问题，以图像特征为基础

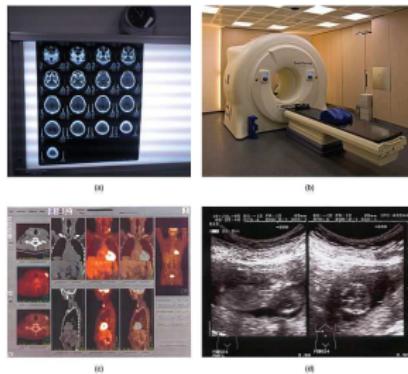


image source: http://en.wikipedia.org/wiki/Medical_imaging

医学影像处理

困难重重

- ▶ 分辨率不高，如 CT 的图像矩阵往往只有 512×512 ；
- ▶ 动态范围大，如 CT 值范围在 $[-1000, 1000]$ Hu 之间；
- ▶ 来源单一而封闭，需要大型扫描设备，而且涉及患者隐私；
- ▶ 由于人体构造、扫描时体位等不同而呈现很大差异性，病变部位则更是千差万别；
- ▶ 维度高，一般为 3D 体数据，或者带时间信息的 4D 图像。

医学影像处理

困难重重

- ▶ 分辨率不高，如 CT 的图像矩阵往往只有 512×512 ；
- ▶ 动态范围大，如 CT 值范围在 $[-1000, 1000]$ Hu 之间；
- ▶ 来源单一而封闭，需要大型扫描设备，而且涉及患者隐私；
- ▶ 由于人体构造、扫描时体位等不同而呈现很大差异性，病变部位则更是千差万别；
- ▶ 维度高，一般为 3D 体数据，或者带时间信息的 4D 图像。

现实：临幊上大部分软件基于人机交互，自动化程度低，无法将医生从一些重复工作中解放出来

特征提取

好的特征：强**鲁棒性**和高**区分度**

1. 人工设计：LBP, HOG, SURF, SIFT 等，性能有限
2. 监督学习：大部分依赖于对匹配点的标注
3. 非监督学习：大部分基于概率假设，对特征应具有的鲁棒性和区分力没有显式描述

深度学习

成功要素：

- ▶ 数据
- ▶ GPU
- ▶ Tricks

深度学习

成功要素：

- ▶ 数据
- ▶ GPU
- ▶ Tricks

现状：

- ▶ 基于深度学习的方法已在机器视觉、机器翻译、自然语言处理、语音识别、文字识别等诸多领域取得领先结果
- ▶ 在医学影像方面有所尝试，但未到理想程度，胸腹腔数据上未见类似工作
- ▶ 本文主要使用卷积神经网络

本研究所受课题资助

- ▶ 十二五国家科技支撑计划课题 影像引导肿瘤微创治疗技术体系研究
- ▶ 国家自然科学基金重点项目 多模态影像融合肿瘤热消融手术规划和导航研究

目录

研究背景与意义

研究内容

 基于层叠卷积 ISA 的特征提取

 基于深度卷积神经网络的特征提取

 评测方法

实验结果

 ISA

 卷积神经网络

 小结

总结

参考文献

ISA:概览

- ▶ 原始形式
- ▶ 层叠卷积 ISA
- ▶ 实验
- ▶ 局限性

原始 ISA 的形式

划分子空间结构，仅要求子空间之间的独立性。 \mathbf{V} 为待求解的 ISA 矩阵， $\mathbf{z}_1, \dots, \mathbf{z}_T$ 为白化的输入数据， $S(k)$ 为第 k 个子空间包含的下标，则优化目标为：

$$\begin{aligned} \mathbf{V}^* &= \arg \min_{\mathbf{V}} \sum_{t=1}^T \sum_{k=1}^K \sqrt{\sum_{i \in S(k)} (\mathbf{v}_i^\top \mathbf{z}_t)^2} \\ \text{subject to } \quad \mathbf{V}^\top \mathbf{V} &= \mathbf{I} \end{aligned} \tag{2.1}$$

使用 ISA 从输入 \mathbf{x} 提取特征 \mathbf{e} :

$$\begin{aligned} e_k &= \sqrt{\sum_{i \in S(k)} s_i^2} \\ \mathbf{s} &= \mathbf{Vx} \end{aligned} \tag{2.2}$$

原始 ISA 的形式

划分子空间结构，仅要求子空间之间的独立性。 \mathbf{V} 为待求解的 ISA 矩阵， $\mathbf{z}_1, \dots, \mathbf{z}_T$ 为白化的输入数据， $S(k)$ 为第 k 个子空间包含的下标，则优化目标为：

$$\begin{aligned} \mathbf{V}^* &= \arg \min_{\mathbf{V}} \sum_{t=1}^T \sum_{k=1}^K \sqrt{\sum_{i \in S(k)} (\mathbf{v}_i^T \mathbf{z}_t)^2} \\ \text{subject to} \quad \mathbf{V}^T \mathbf{V} &= \mathbf{I} \end{aligned} \tag{2.1}$$

在 T 上拆分，多 GPU 并行实现。

层叠卷积 ISA

将 4D 图像作为输入，矩阵乘转化为卷积，成为两层神经网络

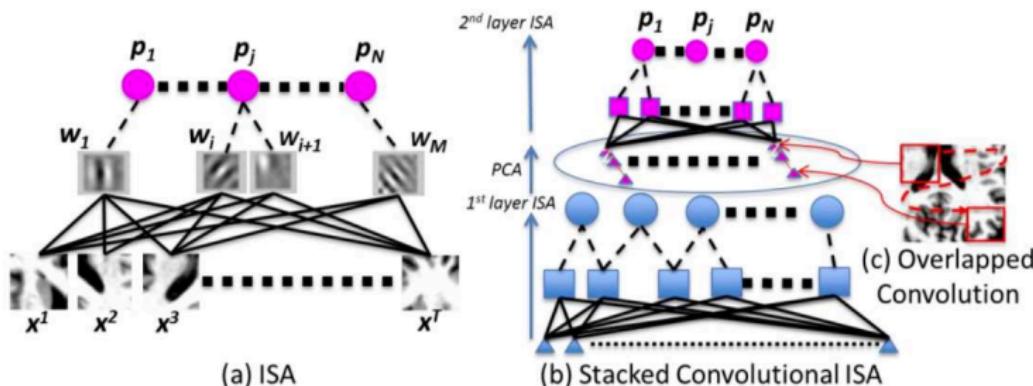


Fig. 2. Graphical depiction of ISA and the stacked convolutional ISA network

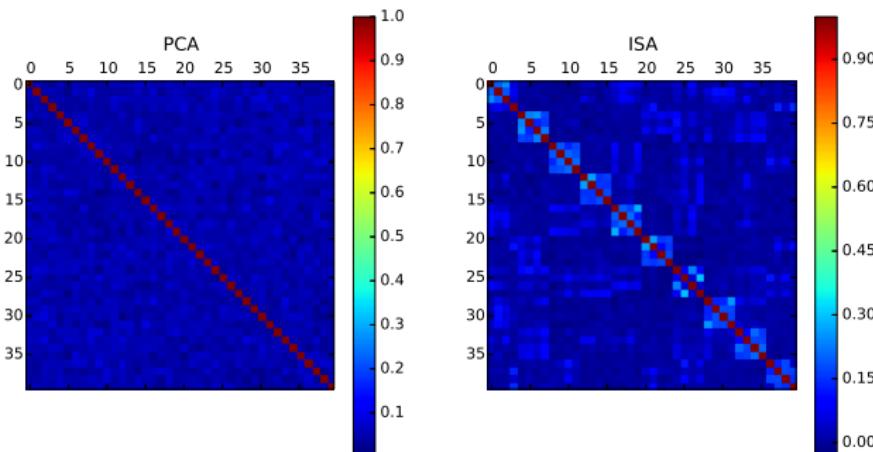
image source: Wu et al. 2013

实验: toy example

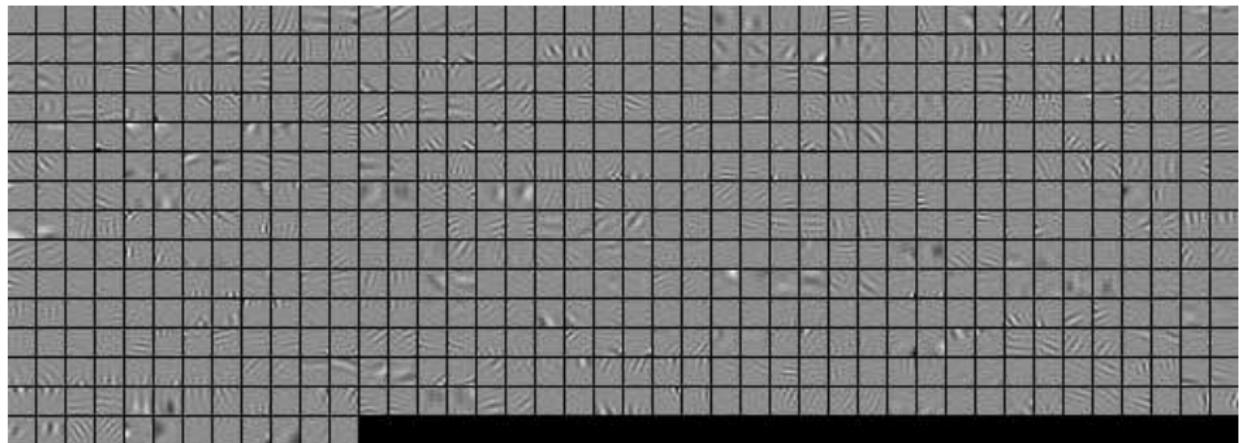
40 维独立高斯分布: $\mathbf{x} = (x_1, \dots, x_{40})^\top$

10 维独立均匀分布: $\mathbf{y} = (y_1, \dots, y_{10})^\top$

$$\begin{aligned}\mathbf{z} &= (z_1, \dots, z_{40})^\top \\ z_{4i+j} &= x_{4i+j+1} y_{i+1}\end{aligned}\tag{2.2}$$



实验: SLIVER07 上习得的滤波器



局限性

1. 卷积核大
2. 贪心逐层训练
3. 层数少
4. 完全非监督

CNN:概览

- ▶ 鲁棒性与区分力的形式化描述
- ▶ 两种损失函数
- ▶ 数据增广方法

鲁棒性与区分力

对于变换集合 \mathcal{T} , 定义

$$\begin{aligned} D^{\mathcal{T}}(\mathbf{X}) &= \{T(\mathbf{X}) : T \in \mathcal{T}\} \\ S_{\mathbf{W}}^{\mathcal{T}}(\mathbf{X}) &= \{f(\mathbf{X}; \mathbf{W}) : \mathbf{X} \in D^{\mathcal{T}}(\mathbf{X})\} \end{aligned} \quad (2.3)$$

要求:

鲁棒性 同一位置的 \mathbf{X}_1 和 \mathbf{X}_2 , $S_{\mathbf{W}}^{\mathcal{T}}(\mathbf{X}_1)$ 和 $S_{\mathbf{W}}^{\mathcal{T}}(\mathbf{X}_2)$ 类间变化小; $\mathbf{X} \in \mathcal{I}$, $S_{\mathbf{W}}^{\mathcal{T}}(\mathbf{X})$ 类内变化小

区分力 不同位置的 \mathbf{X}_1 和 \mathbf{X}_2 , $S_{\mathbf{W}}^{\mathcal{T}}(\mathbf{X}_1)$ 和 $S_{\mathbf{W}}^{\mathcal{T}}(\mathbf{X}_2)$ 类间变化大

损失函数

多分类 训练数据: $\mathbf{X}_1, \dots, \mathbf{X}_N$

监督信号: $D^T(\mathbf{X}_i)$ 中的所有图像都被分类到第 i 类

损失函数

多分类 训练数据: $\mathbf{X}_1, \dots, \mathbf{X}_N$

监督信号: $D^T(\mathbf{X}_i)$ 中的所有图像都被分类到第 i 类

度量学习 不同位置的图像 $\mathbf{X}_1, \mathbf{X}_2$

增广: $\mathbf{X}'_1 \in D^T(\mathbf{X}_1), \mathbf{X}'_2 \in D^T(\mathbf{X}_1), \mathbf{X}'_3 \in D^T(\mathbf{X}_2)$

记 $\mathbf{y}_i = f(\mathbf{X}'_i; \mathcal{W})$

$$\begin{aligned} L(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3) &= \max(0, d(\mathbf{y}_1, \mathbf{y}_2) + \delta - d(\mathbf{y}_1, \mathbf{y}_3)) \\ d(\mathbf{x}, \mathbf{y}) &= 1 - \frac{\mathbf{x}^\top \mathbf{y}}{|\mathbf{x}| |\mathbf{y}|} \end{aligned} \quad (2.4)$$

损失函数

多分类 训练数据: $\mathbf{X}_1, \dots, \mathbf{X}_N$

监督信号: $D^T(\mathbf{X}_i)$ 中的所有图像都被分类到第 i 类

全局优化

度量学习 不同位置的图像 $\mathbf{X}_1, \mathbf{X}_2$

增广: $\mathbf{X}'_1 \in D^T(\mathbf{X}_1), \mathbf{X}'_2 \in D^T(\mathbf{X}_1), \mathbf{X}'_3 \in D^T(\mathbf{X}_2)$

记 $\mathbf{y}_i = f(\mathbf{X}'_i; \mathcal{W})$

$$\begin{aligned} L(\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3) &= \max(0, d(\mathbf{y}_1, \mathbf{y}_2) + \delta - d(\mathbf{y}_1, \mathbf{y}_3)) \\ d(\mathbf{x}, \mathbf{y}) &= 1 - \frac{\mathbf{x}^\top \mathbf{y}}{|\mathbf{x}| |\mathbf{y}|} \end{aligned} \quad (2.4)$$

对数据量无苛刻要求

数据增广

数据增广本质上是定义 \mathcal{T} 的过程

- ▶ Gamma 校正:

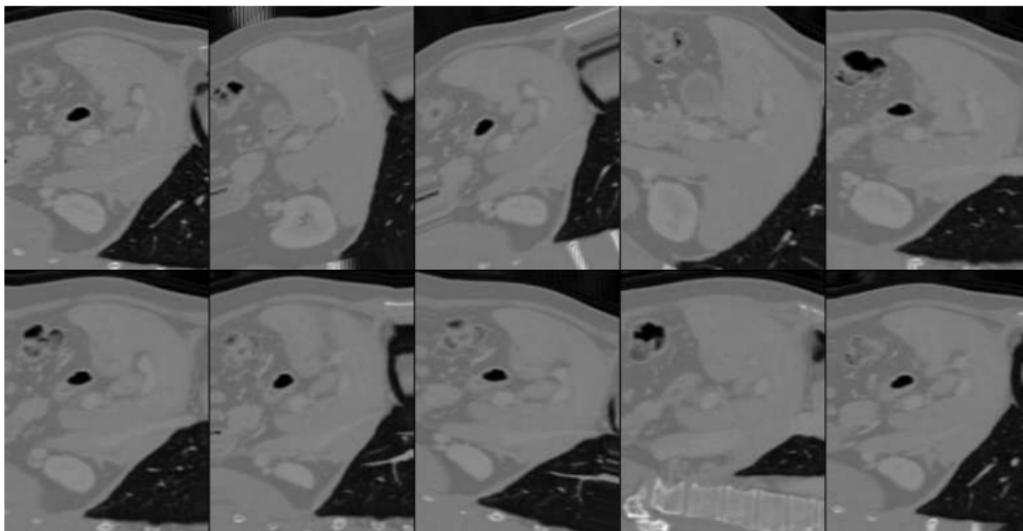
$$\mathbf{y} = L + (U - L) \left(\frac{\mathbf{x} - L}{U - L} \right)^\gamma \quad (2.5)$$

- ▶ 仿射变换

数据增广：仿射变换

对变换矩阵 A_0 做 SVD 分解：

$$\begin{aligned} A_0 &= U \Sigma V^T \\ &= U \Sigma U^T R \end{aligned} \quad (2.6)$$



评测方法:概览

- ▶ 单点匹配
- ▶ ROC 曲线

单点匹配的判定

要求以前景掩膜形式提供对某内脏的分割标注

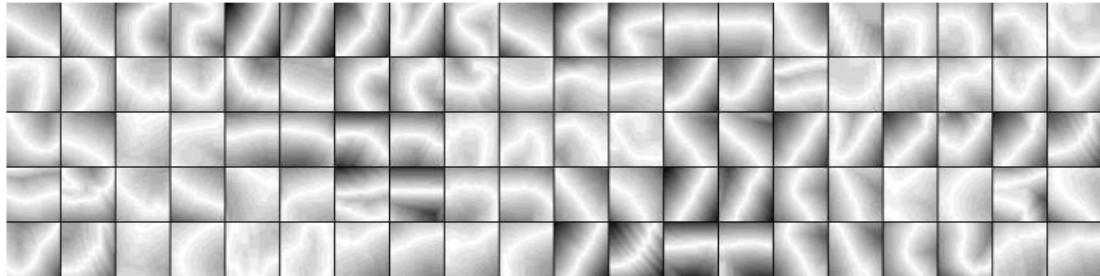
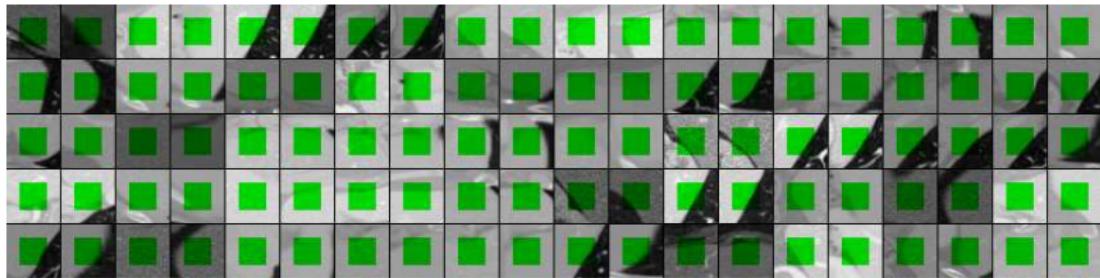
边界距离 某点到标注上脏器边界的距离，脏器内部为正，外部为负

参考曲面 边界距离为 2 的点构成

匹配点 训练图像参考曲面上某点，到某测试图像所有点中特征距离最近的点

成功匹配 匹配点在测试图像上的边界距离在 $[1, 3]$ 范围内

成功匹配点对的抽样观察



绘制 ROC 曲线

设有 N 个训练图像，每个的参考曲面上采样 T 个点；对测试图像 i ，对所有匹配点按 $2 \times 2 \times 2$ 方块去重，然后限制特征距离不超过 θ ：

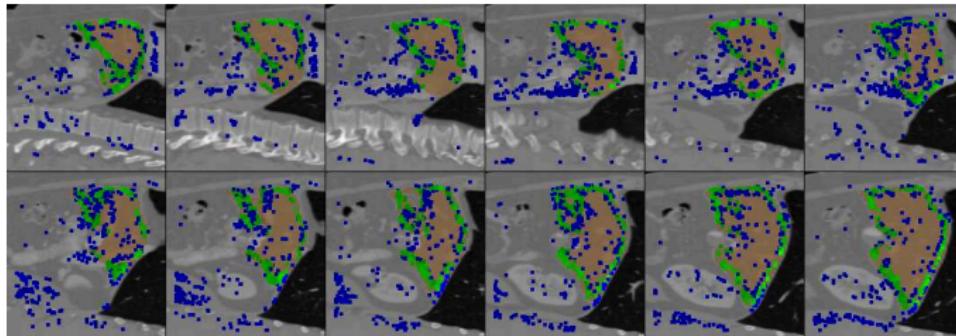
顶峰比 $t_i^{(\theta)} = \frac{\text{去重后点数}}{NT}$

精确度 $a_i^{(\theta)} = \frac{\text{去重后成功匹配点数}}{\text{去重后点数}}$

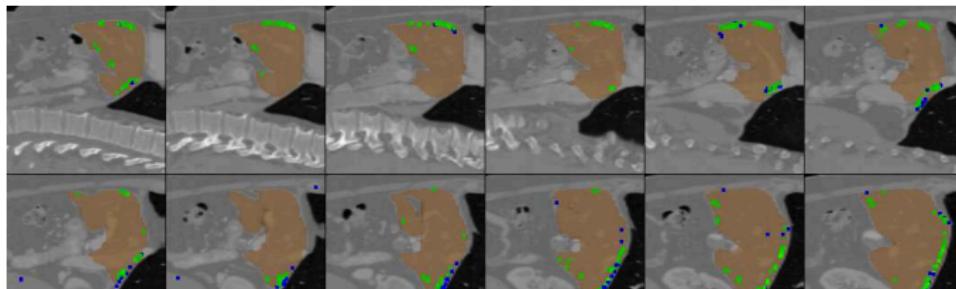
$t^{(\theta)}$ 、 $a^{(\theta)}$ 分别为 $\{t_i^{(\theta)}\}$ 、 $\{a_i^{(\theta)}\}$ 的均值，遍历 θ ，绘制 $(t^{(\theta)}, a^{(\theta)})$ 曲线。

阈值限制对点分布的影响

所有匹配点，顶峰比为 0.568，匹配精确度为 48.2%



限制阈值，顶峰比为 0.078，匹配精确度为 80.1%。



目录

研究背景与意义

研究内容

基于层叠卷积 ISA 的特征提取

基于深度卷积神经网络的特征提取

评测方法

实验结果

ISA

卷积神经网络

小结

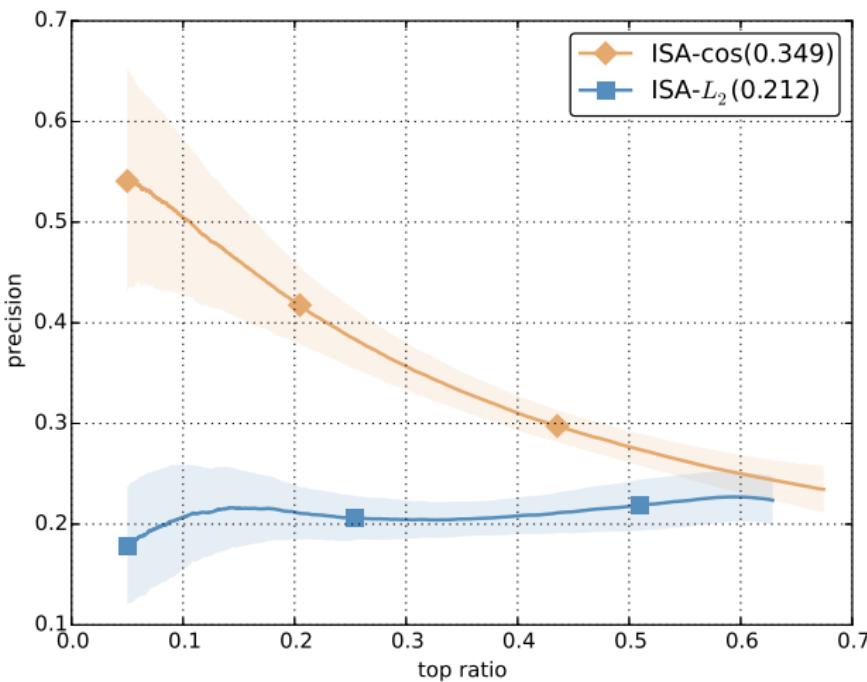
总结

参考文献

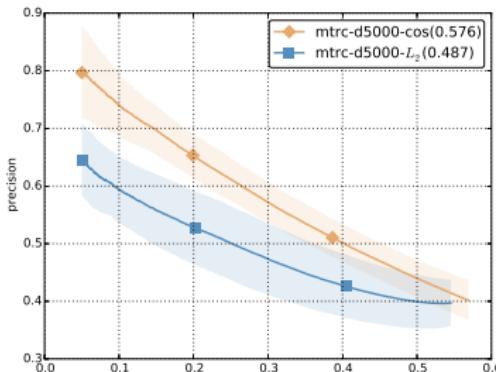
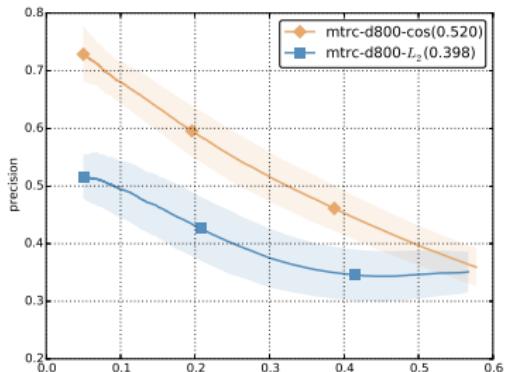
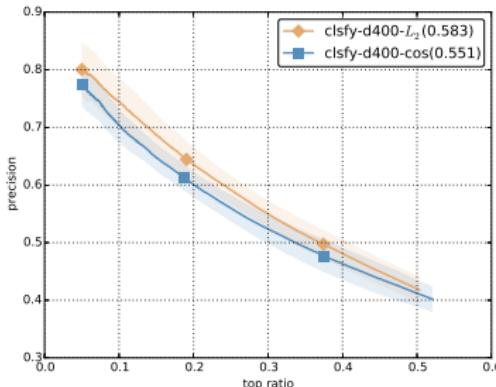
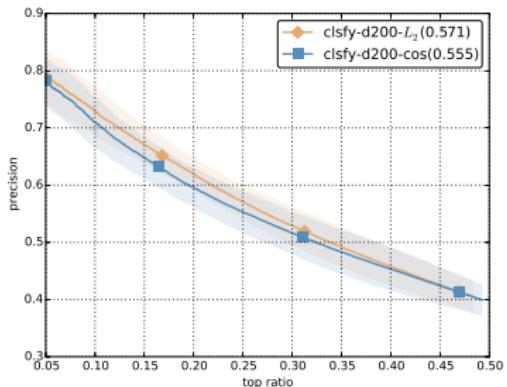
实验内容概览

- ▶ 距离度量: \cos vs L_2
- ▶ 训练数据量
- ▶ δ in $\max(0, d(\mathbf{y}_1, \mathbf{y}_2) + \delta - d(\mathbf{y}_1, \mathbf{y}_3))$
- ▶ PCA
- ▶ 边界距离与特征距离的关系反映特征性能

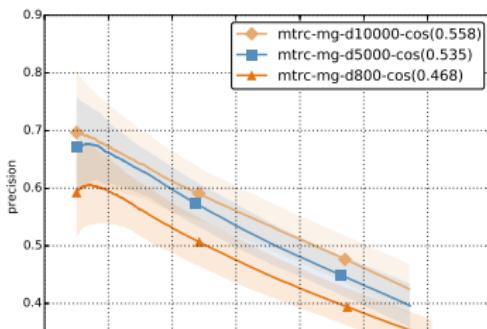
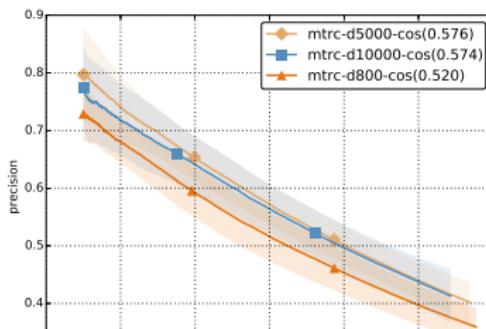
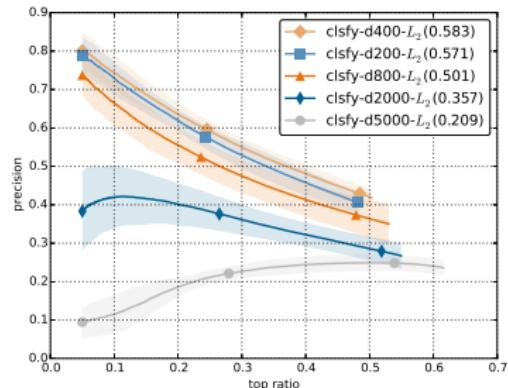
ISA



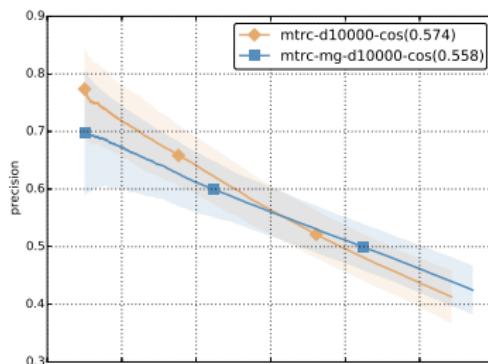
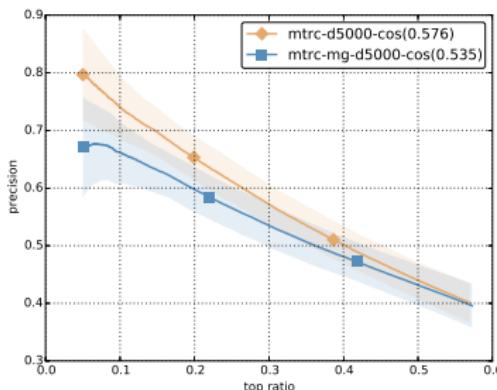
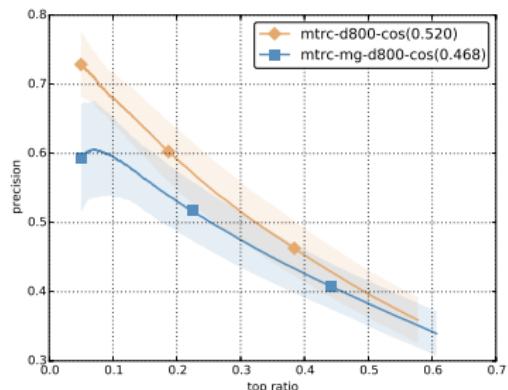
距离度量



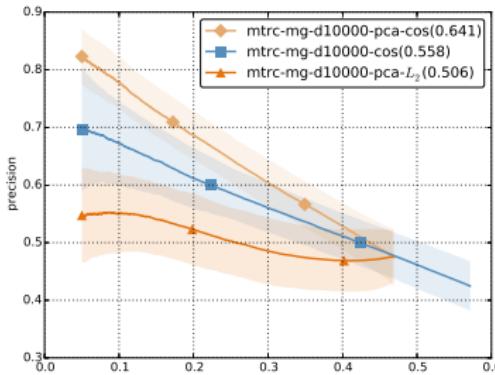
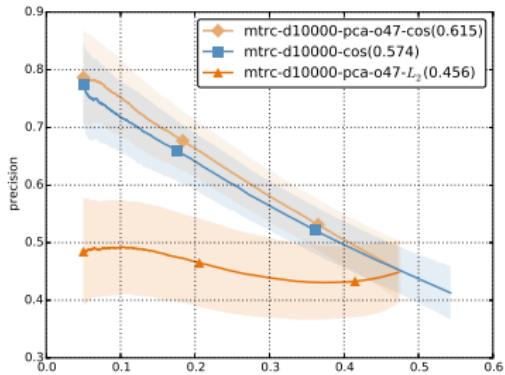
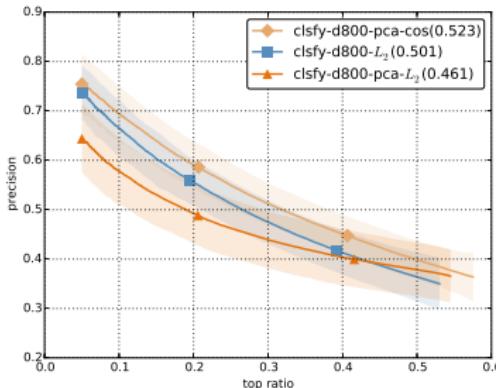
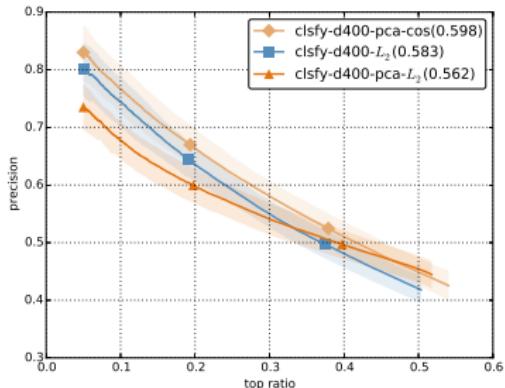
训练数据量



度量学习中的 δ

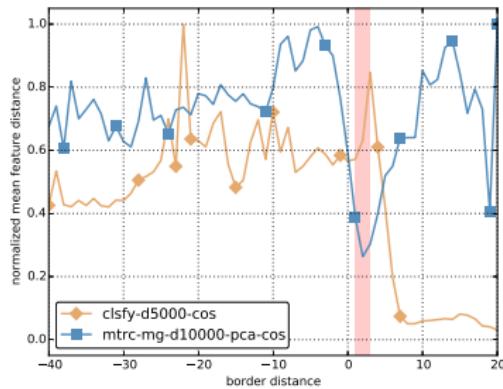
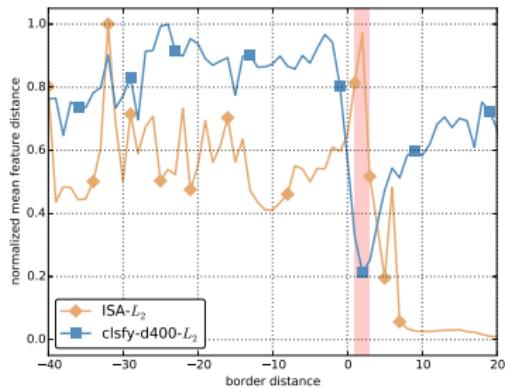


PCA



边界距离、特征距离与特征性能

当小的边界距离不对应小的特征距离时，特征距离不反映匹配配置信度，此时往往性能较差。如 ISA 的 L_2 、多分类时 5000 数据量等。



小结

1. 监督信号加入对鲁棒性和区分力要求的 CNN，比完全非监督的 ISA 好
2. 每种方法有对应的恰当的距离度量
3. 多分类训练的网络要求恰到好处的训练数据量
4. 度量学习的损失函数中需要合适的 δ
5. 结合 PCA 一般可提高性能

最好结果：度量学习 + 大量数据 +PCA+ 余弦距离

目录

研究背景与意义

研究内容

 基于层叠卷积 ISA 的特征提取

 基于深度卷积神经网络的特征提取

 评测方法

实验结果

 ISA

 卷积神经网络

 小结

总结

参考文献

主要工作及创新点

1. 层叠卷积 ISA: GPU 的数据并行实现, 极大提高了训练速度¹
2. 深度卷积神经网络: 将对鲁棒性和区分力的要求显式整合进损失函数, 使用多分类和度量学习两种损失函数, 辅以详细实验
3. 用深度神经网络分析胸腹腔医学影像, 并实验证明其在特征提取方面的优势
4. 特征评测标准: 无需依赖其它任务、无需对应点的标注

¹相关代码已开源<https://github.com/jia-kai/bachelor-thesis>

Thanks!

目录

研究背景与意义

研究内容

 基于层叠卷积 ISA 的特征提取

 基于深度卷积神经网络的特征提取

 评测方法

实验结果

 ISA

 卷积神经网络

 小结

总结

参考文献

参考文献 |

- [1] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. "SURF: Speeded up robust features". In: *Computer vision–ECCV 2006*. Springer, 2006, pp. 404–417.
- [2] Yoshua Bengio. "Learning deep architectures for AI". In: *Foundations and trends® in Machine Learning* 2.1 (2009), pp. 1–127.
- [3] Yoshua Bengio, Aaron Courville, and Pierre Vincent. "Representation learning: A review and new perspectives". In: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35.8 (2013), pp. 1798–1828.
- [4] James Bergstra et al. "Theano: a CPU and GPU Math Expression Compiler". In: *Proceedings of the Python for Scientific Computing Conference (SciPy)*. Oral Presentation. Austin, TX, June 2010.
- [5] Navneet Dalal and Bill Triggs. "Histograms of oriented gradients for human detection". In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. Vol. 1. IEEE. 2005, pp. 886–893.

参考文献 II

- [6] Alexey Dosovitskiy et al. "Discriminative unsupervised feature learning with convolutional neural networks". In: *Advances in Neural Information Processing Systems*. 2014, pp. 766–774.
- [7] Kaiming He et al. "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification". In: *arXiv preprint arXiv: 1502.01852* (2015).
- [8] Tobias Heimann et al. "Comparison and evaluation of methods for liver segmentation from CT datasets". In: *Medical Imaging, IEEE Transactions on* 28.8 (2009), pp. 1251–1265.
- [9] Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. "Distilling the knowledge in a neural network". In: *arXiv preprint arXiv:1503.02531* (2015).
- [10] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. "Multilayer feedforward networks are universal approximators". In: *Neural networks* 2.5 (1989), pp. 359–366.

参考文献 III

- [11] Aapo Hyvärinen and Patrik Hoyer. "Emergence of phase-and shift-invariant features by decomposition of natural images into independent feature subspaces". In: *Neural computation* 12.7 (2000), pp. 1705–1720.
- [12] Aapo Hyvärinen, Jarmo Hurri, and Patrick O Hoyer. *Natural Image Statistics: A Probabilistic Approach to Early Computational Vision*. Vol. 39. Springer Science & Business Media, 2009.
- [13] Diederik Kingma and Jimmy Ba. "Adam: A method for stochastic optimization". In: *arXiv preprint arXiv:1412.6980* (2014).
- [14] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks". In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.
- [15] Jack B Kuipers. *Quaternions and rotation sequences*. Vol. 66. Princeton university press Princeton, 1999.

参考文献 IV

- [16] Quoc V Le et al. "Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis". In: *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on.* IEEE. 2011, pp. 3361–3368.
- [17] Yann LeCun et al. "Gradient-based learning applied to document recognition". In: *Proceedings of the IEEE 86.11 (1998)*, pp. 2278–2324.
- [18] David G Lowe. "Object recognition from local scale-invariant features". In: *Computer vision, 1999. The proceedings of the seventh IEEE international conference on.* Vol. 2. IEEE. 1999, pp. 1150–1157.
- [19] Timo Ojala, Matti Pietikainen, and David Harwood. "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions". In: *Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on.* 1. 1994, pp. 582–585.

参考文献 V

- [20] Jean-Marc Peyrat et al. "Registration of 4D cardiac CT sequences under trajectory constraints with multichannel diffeomorphic demons". In: *Medical Imaging, IEEE Transactions on* 29.7 (2010), pp. 1351–1368.
- [21] Florian Schroff, Dmitry Kalenichenko, and James Philbin. "Facenet: A unified embedding for face recognition and clustering". In: *arXiv preprint arXiv:1503.03832* (2015).
- [22] Paul Scovanner, Saad Ali, and Mubarak Shah. "A 3-dimensional sift descriptor and its application to action recognition". In: *Proceedings of the 15th international conference on Multimedia*. ACM. 2007, pp. 357–360.
- [23] Dinggang Shen and Christos Davatzikos. "HAMMER: hierarchical attribute matching mechanism for elastic registration". In: *Medical Imaging, IEEE Transactions on* 21.11 (2002), pp. 1421–1439.
- [24] Ilya Sutskever et al. "On the importance of initialization and momentum in deep learning". In: *Proceedings of the 30th international conference on machine learning (ICML-13)*. 2013, pp. 1139–1147.

参考文献 VI

- [25] Guorong Wu et al. “Unsupervised deep feature learning for deformable registration of MR brain images”. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer, 2013, pp. 649–656.
- [26] 金征宇. 医学影像学. 第 2 版. 人民卫生出版社, 2010.