

主要内容

1 思路摘要 2 算法流程图 3 相关文章工作 4 附件

思路摘要

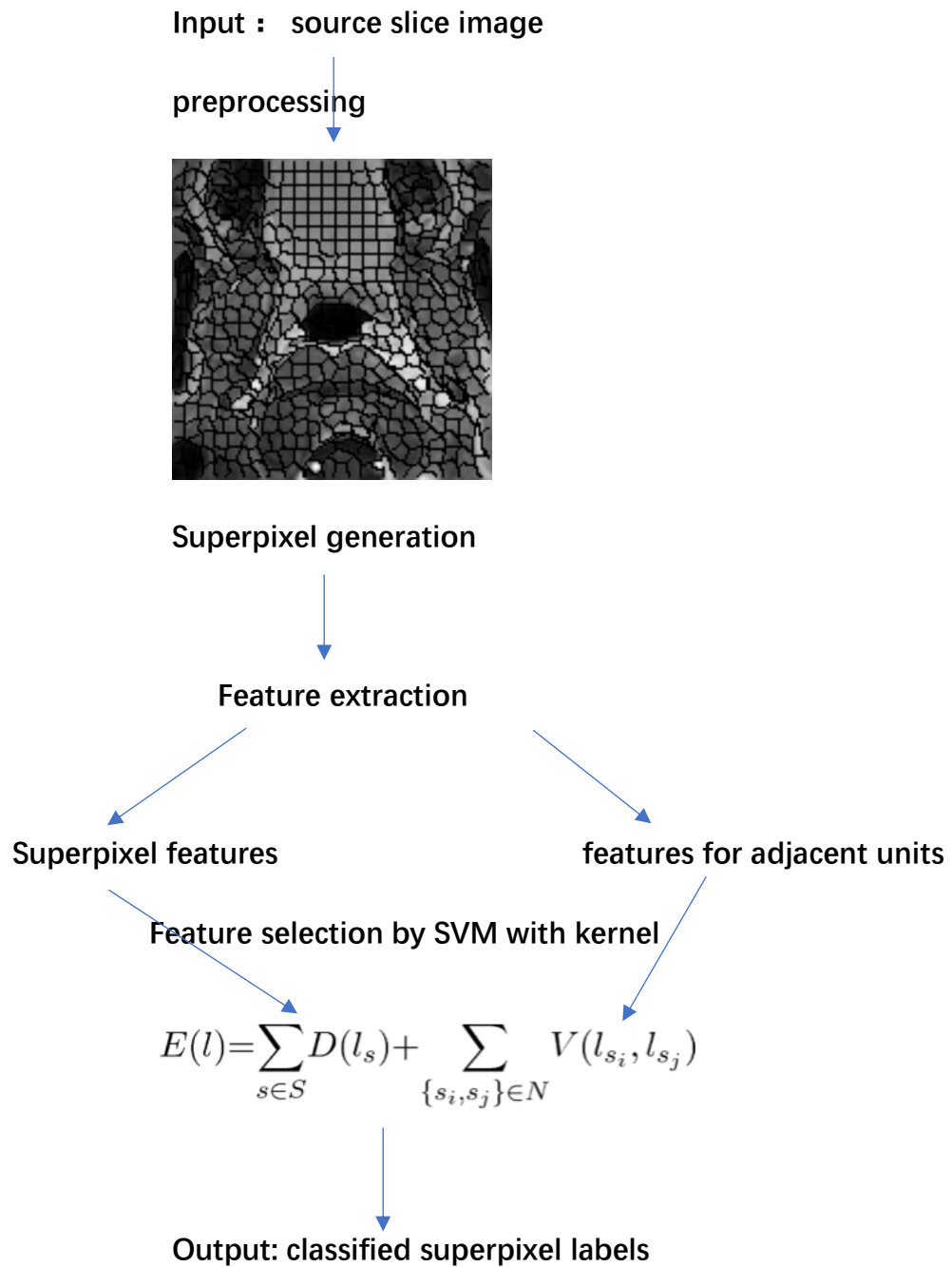
在阅读了相关文章后，考虑到鼻咽癌肿瘤的复杂程度和目前的数据集的规模，觉得通过学习的方法，得到准确的特征组合表达是比较合理的想法。我们打算采用 superpixel-based 的 Markov random field (graph cut) 的算法来做。

过分割产生的 Superpixel 携带很多有意义的特征信息，比如纹理，形状特征。同时 superpixel 边界能很好的与图像的原始边界保持一致。用 superpixel 替代传统的 pixel 作为基本的图像处理单元能方便于特征提取，同时大大减少计算量。

这样问题就转换为了对超像素块的二分类问题。相比于直接提取超像素块特征分类，Graph cut 综合考虑了像素本身的特征和相邻像素的相似性，通过构建能量函数最小化求解。能量函数简单结合灰度信息和位置信息做出推断，这样的信息并不能准确地描述像素块的特征。我们打算采用学习的方法来找到准确描述像素块肿瘤性质的特征参数组合。同时也采用学习的方法来找到准确描述相邻超像素块相似性的特征参数组合，SVM 和随机森林是供考虑的分类器。将学习到的结果带入 graph cut 框架（或者说 MRF）计算得到最终结果。

另外，考虑到在不同层面上不同特征可能有不一样的重要程度，计划按照结构特点把 slice image 分成二到三个部分来学习不同特征组合参数。

流程图



相关工作

主要阅读了三篇 superpixel-based 的文章

Paper1, Automated brain tumor detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI

用 SLIC 生成超像素, 对超像素块提取了灰度 纹理 分形不同特征, 然后用 mRMR 选择特征 最后用 ERT 对超像素做二分类。这样的方法没有考虑相邻像素块的影响, 不能保证分割结果的一致性。

Paper2, Superpixel-based Segmentation for 3D Prostate MR Images

用 SLIC 生成超像素后, 采用 3d 的 graph cut 求解, 像素本身特征考虑了灰度信息和位置信息, 用相邻像素的灰度差来衡量相似性, 考虑了三维相邻的超像素块的情况, 缺点是特征直接构建缺少训练过程, 不能确保选取的特征能否准确描述。

Paper3 , Class Segmentation and Object Localization with Superpixel Neighborhoods

用 Quick Shift 生成超像素, 用 sift descriptor 来描述超像素块特征, 用 SVM 训练分类超像素, 之后引入 CRF 来控制结果一致性, pairwise 对综合考虑了相邻像素块的灰度差和共同边界长短。

阅读文章之后思考, 为什么不把能量函数的第二项也通过学习的方法表达呢? 查阅资料发现存在很多描述相邻区域相似性的特征, 我们把训练集中有 相同标签的相邻像素对 label 为 0, 有不同标签的相邻像素对 label 为 1。这样也可以通过学习的办法来表达能量函数的第二项了。

附件 1 features for adjacent region

Table 1: Features for adjacent regions

Feature names	Dimension	Notation
differences in each channel of CIELAB	3	d_l, d_a, d_b
Euler distance of CIELAB values	1	d_c
average gradient maximum along boundary	1	d_g
χ^2 distance between RGB histograms	1	χ_h^2
χ^2 distance between gradient histograms	1	χ_H^2
variances of RGB values	3	s_r, s_g, s_b
variances of CIELAB values	3	s'_l, s'_a, s'_b
average HED maximum along boundary	1	d_h

Brightness and Colors

三个通道灰度特征

Euler distance of CIELAB values

CIELAB 空间上的距离

Average gradient maximum along boundary

不直接利用梯度，遍历相邻边界上所有的点，对于一个点画一个圈，得到这个圈内部最大的梯度值，最后求平均

χ^2 distance between RGB histograms

To make use of the details of color information, we employ the color histogram that has $8 \times 8 \times 8$ dimensions in the RGB color space. For histograms belonging to adjacent segments, we use χ^2 distance to measure their difference.

χ^2 distance between gradient histograms

The χ^2 distance of two segments when computing histograms of oriented gradient for each segment is also an attractive choice.

Variances

计算两块共同的颜色空间的方差，方差越小 表示越相似

Variance is a good measure for the fluctuation of a piece of data. We apply variances in the RGB (s_r, s_g, s_b) and CIELAB (s'_l, s'_a, s'_b) color spaces to R_i and R_j , where R_i and R_j are adjacent segments. The magnitude of the variances reflects the similarity between the two segments.

Average HED maximum

along boundary The HED feature is computed similarly to gradient feature above. However, because of the extra overhead of HED, we make this choice optional.

附件 2，三篇文章详细总结

paper1:

Supervoxel-based Segmentation for 3D Prostate MR Images

文章选用了 SLIC 生成了超像素

之后选取灰度信息和位置信息作为 supervoxel 的特征

灰度信息：高斯混合模型去拟合 MR 图的灰度分布，

位置信息：计算像素块到手动初始轮廓距离

结合这两个信息得到对 $p(1|s)$ 和 $p(0|s)$ 的概率估计（这里没有训练集做支撑，作者自己设计的，凭感觉设计？？）

pairwise: 考虑了三维范围内的相邻像素块，

相 同 层 面 相 邻

$$V^{intra}(l_{s_i}, l_{s_j}) = (1 - \delta(l_{s_i}, l_{s_j})) e^{-\beta \|G_i - G_j\|^2}$$

不同层面相邻（重合面积）

$$V^{inter}(l_{s_i}, l_{s_j}) = (1 - \delta(l_{s_i}, l_{s_j})) \frac{A(s_i \cap s_j)}{A(s_i \cup s_j)}$$

设置好了 $P(x1, x2)$ 和 $p(y1|x)$ 后 为标准 graph cut 结构 用 max-flow/min-cut 求解

Paper2 总结

Class Segmentation and Object Localization with Supervoxel Neighborhoods

首先使用 **quick shift** 生成 supervoxel，这种方法生成的 supervoxel 有一个特点：每张图片 supervoxel 的大小和数目都是不定的。如果图片越复杂，那么 supervoxel 的数目就越多，从而 supervoxel 就越小。

然后对生成的 supervoxel 进行分类。具体做法如下：首先对所有的像素求 SIFT descriptor，然后使用 K-mean quantize 到 K 类。最后对每个 supervoxel，将其中的像素 SIFT descriptor normalize 11 histogram，将每个 supervoxel 赋予其中像素标签最多的那个标签，使用 one-vs-rest SVM 进行分类。其判别函数为

$$C(h^0) = \sum_{j=1}^L c_i \exp(-\gamma d_{\chi^2}^2(h^0, h_i^0))$$

其中 \hat{h}^0 表示 histogram，在后面的改进中，会用另外一个值替代。其中，

$$d_{\chi^2}^2(h^0, h_i^0) = \sum_{k=1}^K \frac{(h^0(k) - h_i^0(k))^2}{h^0(k) + h_i^0(k)}.$$

但是直接这样分类会有两个缺点：1) 直接分类没有考虑周围 superpixel 的信息；2) superpixel 的 histogram 值比较稀疏，因为在一个 superpixel 内部，像素的 SIFT descriptor 都比较类似，所以它们的值都比较接近。

基于这两点，作者做了一个改进，即将图像的 superpixel 构造成一个图 $G(S, E)$ ，其中 superpixel $s_i \in S$ 表示全体 superpixel， E 表示两个 superpixel (s_1, s_2) 的边。

$D(s_i, s_j)$ 表示两个 superpixel 之间最短的距离。 H_i^0 表示第 i 个 superpixel 的 unnormalized histogram。 H_i^N 表示所有与第 i 个 superpixel 距离 N 个 node 之内的

histogram，其计算公式如下：

$$H_i^N = \sum_{s_j | D(s_i, s_j) \leq N} H_j^0$$

所有的学习算法没有改变，其实就是做了一个平均。

使用这种 histogram，解决了我们前面的两个问题：1) 结合了临近 superpixels 的 histogram，其分布更加广泛，因为使用了很多不一致的特征，这样就可以防止 SVM 过拟合。2) 当我们增加 N ，那我们就可以从周围的 superpixel 学到更多，从而慢慢从『part』level 转变成『object』level。

当然，如果 N 增加的一定程度，那么我们会慢慢模糊物体的边界，所以，我们需要使用 CRF 来减少这种效果。用 CRF 进行优化为了使得在提高 N 的情况下仍然能够有准确的边界，必须要增加一些新的限制来减少边界位置的错分类问题。

CRF 通过加入一些 pairwise potential 来实现这些限制。用 $p(c|G; w)$ 来表示已知和的条件概率，能量函数为

$$-\log(P(c|G; w)) = \sum_{s_i \in S} \Psi(c_i | s_i) + w \sum_{(s_i, s_j) \in E} \Phi(c_i, c_j | s_i, s_j)$$

单点势能函数 $\Psi(c_i | s_i) = -\log(P(c_i | s_i))$ 表示每个 superpixel 的 SVM 的概率输出，具体为：而边势能函数可以表示为：

$$\Phi(c_i, c_j | s_i, s_j) = \left(\frac{L(s_i, s_j)}{1 + \|s_i - s_j\|} \right) [c_i \neq c_j]$$

其中, $[\cdot]$ 是 0-1 指示函数, $\|s_i - s_j\|$ 表示两个 superpixel 之前颜色差异的 norm 值 (在 LUV 颜色空间内)。 $L(s_i, s_j)$ 表示两个 superpixel 的边界长度, 用来作为减少小的分散区域的正则项。一旦模型训练完之后, 我们使用 alpha-expansion 来 inference。

当改变 K-Means 中 K 的值时, 结果也会有很大的变化, 当 K=400 会比较理想的效果。

另外, 还有你读的 paper, 我认为价值在于提供了新的特征提取和筛选的思路。

最后补充一下 MRF 和 CRF 很相似, graph cut 是他们的一种特殊情况。

To mao: 大体思路就是这样了, 从选取的 3 篇文章来看, 有很多操作空间给我们去设计, 感觉还是很值得一做的, 另外很多新的知识, 我看了几天文章也有好多不懂, 不过没有什么关系, 能学到新知识才是最重要的, 我们制定好计划, 分工明确一步一步来。

Paper3: Automated brain tumor detection and segmentation using superpixel-based extremely randomized trees in FLAIR MRI

文章介绍的方法主要分五步: 预处理、超像素、提取特征、特征选择、二分类超像素

1、预处理:

去掉颅骨 (FSL[38])

直方图匹配: itk 实现。直方图匹配算法

http://www.idlcoyote.com/ip_tips/histomatch.html

2、超像素: SLIC 方法

3、提取特征:

(1) 灰度统计特征: 基于超像素, 对每个超像素包含的所有像素的灰度做统计

假设一个超像素包含 N 个像素, 灰度值分别记为: $\{x_1, x_2, \dots, x_N\}$

➤ 平均值: $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$

➤ 最大值

➤ 最小值

➤ 中值

➤ 众数

- 标准差: $\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2}$
- 方差: $s = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$
- 平均绝对偏差: $\frac{1}{N} \sum_{i=1}^N |x_i - \bar{x}|$
- 绝对偏差的中值
- 差异系数: $CV = \frac{\sigma}{\bar{x}} \times 100\%$ (系数越大, 表示数据的离散程度越大)
- 偏度: $Skewness = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^3 / \sigma^3$ (统计数据分布非对称程度的数字特征)
- 峰度: $Kurtosis = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^4 / \sigma^4 - 3$ (描述总体数据中所有取值分布形态陡缓程度的统计量, 与正态分布相比较)
- 中心矩: k 阶中心矩 $E\{[X - E(X)]^k\}$
- 极差: $\max\{x_i\} - \min\{x_i\}$ (最大值与最小值之间的差距, 也叫全距)
- 四分位差: 上四分位数 (Q3, 位于 75%) 与下四分位数 (Q1, 位于 25%) 之差,
 $Q = Q_3 - Q_1$
- 熵: $-\sum_i \sum_j P(i, j) \log(P(i, j))$, $P(i, j)$ 表示第 i 个和第 j 个数据同时出现的几率 (? 公式存疑)

(2) 纹理特征:

Gabor 滤波, 取不同的滤波尺寸 (filter size, σ) 和不同方向 (θ) 产生多个滤波 (kernel), 与图像做卷积。

$$G(x, y; \theta, \sigma, \lambda, \psi, \gamma) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi \frac{x'}{\lambda} + \psi\right)\right)$$

其中:

$$x' = x \cos \theta + y \sin \theta$$

$$y' = -x \sin \theta + y \cos \theta$$

σ : 滤波尺寸

λ : 正弦函数波长

θ : 滤波方向

假设设置不同滤波尺寸和不同滤波方向, 产生 N_{FB} 个滤波, 与图像卷积后, 图像上每个像素点与一个滤波产生一个卷积响应值, 则对于每个像素点有 N_{FB} 个响应值, 记为 N_{FB} 维向量。根据每个像素的 N_{FB} 维响应值做 k-means 聚类, 聚类之后得到一副带标号的图像, 记作 texton map。Texton map 上每个超像素的 (标号?) 分布直方图即为该超像素的 texton 特征。

(3) 分形特征

(4) 曲率特征

图像中像素 (x, y) 的二维曲率计算公式:
$$\text{Curv} = \frac{f_{xx}f_y^2 + f_{yy}f_x^2 - 2f_{xx}f_xf_y}{(f_x^2 + f_y^2)^{3/2}}$$

其中 f_{xx} 和 f_{yy} 是图像灰度 $I(x, y)$ 的二阶导。

每个超像素的曲率特征是超像素中所有像素的曲率的平均值。

4、特征选择

应用 mRMR (the Minimum Redundancy Maximum Relevance) 方法选择与问题相关性高的特征, 去除不相关的特征。

用互信息来度量相关性, 对于特征集合 S 中的特征 f_i , 与 class c 的最大相关性通过计

算: $\max D(S, c), \quad D = \frac{1}{|S|} \sum_{f_i \in S} I_M(f_i; c)$

其中 I_M 表示特征 f_i 和 class c 之间的互信息。

计算最小冗余: $\min R(s), \quad R = \frac{1}{|S|^2} \sum_{f_i, f_j \in S} I_M(f_i, f_j).$

运用留一法交叉检验、投票机制, 对所有特征向量进行特征选择。

5、二分类超像素: tumor、non-tumor

ERT (Extremely randomized tree), 由多个决策树构成, 与随机森林算法相似。

<http://blog.csdn.net/zhaocj/article/details/51648966>