

图像分割概述

姓名：彭瑶 学号：1120172773 班级：07111707

摘要：图像分割是计算机视觉的基础，是理解图像的语义内容的首要任务，也是当前的一个研究难点和热点。为了更好的把握和理解图像分割的相关算法，本文综述了图像分割的主要算法，包括经典的分割方法以及基于机器学习的方法，并比较和分析了这些方法的优缺点和可行的改进方式。经典的图像分割算法往往具有较好的执行效率，但面对越来越多的图像分割任务，普适性较差，而近些年飞速发展的机器学习算法为图像分割提供了新的思路。在文末，笔者还进一步总结了图像分割的挑战和应用。

关键词：阈值分割、区域分割、特征编码

图像分割的目的是根据图像的颜色、空间结构乃至语义将图像分割成若干具备语义信息的区域。由于分割的粒度大小没有明确定义，分割边界不够明晰而且很多情况下语义信息不够明确，图像分割仍然是一个非常棘手的问题。不同层次的图像分割所涉及到的技术和困难程度也不一而同。图像的内容分割的目的是从图像中分割出前景的部分和背景部分，图像的实体分割指的是从图像中找到并注明实体，而图像的语义分割需要从图像中识别出具体的语义信息。

早期的图像分割技术主要依赖于图像的颜色和空间，通过局部的颜色一致性和空间结构协同性得到局部特征，再通过局部特征的聚合得到分割区域。随着近些年来机器学习技术的发展，图像分割的技术也不断更新迭代，深度学习和语义分析等技术不断运用在图像分割中。通过训练神经网络不断提取和抽象出语义信息，最终实现分割的目的。

一、经典图像分割算法

传统的图像分割方法主要依赖图像的颜色信息和拓扑结构，并用数学中的理论方法加以分析。这些算法具备高效的运行效率，实现起来也比较简单，但往往只适用于某种类型的图像，对一些特定图像处理效果不佳。

1. 阈值的分割

一幅图像的不同部分往往具有不同的颜色，而

属于同一区域的像素点往往具有相近的颜色，因而可以通过简单的阈值来分割图像，比如选取白色调的某个区间作为阈值，然后选取蓝天白云中包含在阈值中的像素点作为一个分割，从而可以将白云从整幅图中分割出来。这个方法非常简单，然后缺点也显而易见，首先这种分割方法完全忽略了图像中像素点的空间结构，其次阈值的选取比较困难，往往需要多次尝试，而且当某个事物的颜色跨度较大时无法正确的将物体分割出来。

2. 区域的分割

基于区域的分割方式建立在基于阈值分割的基础之上，能够很好的克服以上提到的缺点，根据分割的次序不同，区域分割可以大致分为从一个像素点不断生长到整个区域的区域生长模式和从整个图片出发不断将区域细分的分裂模式。

2.1 区域生长

区域生长的主要思想是每个像素点都必须属于一个区域，属于同一个区域的像素点都必须连续相邻，因此该分割算法可以为：从图像中不断选择未分类的像素点作为当前像素点，然后从该像素点不断拓展，直到不能再拓展为止。这里的拓展需要考虑到像素点的颜色差异以及局部结构信息。除了从像素点出发，类似的算法还可以从像素点和像素点之间的边出发。先将为所有相邻像素点建立边，边的权重可以由连接的像素点颜色差异决定，然后对边进行从小到大的排序，再对每条边一次进行处理，

判断连接的两个像素点是否属于同一分割区域。

基于该方法的分割算法有很多，比如 K-Means 聚类算法、Mean-Shift 算法以及基于图的分割算法。这些算法都过度依赖于颜色差异评估函数的选取，不同的函数取得的效果往往有很大差异，不同图像所适合的评估函数也会不同，常见的颜色差异评估函数为颜色空间的距离函数，如欧式距离、汉明距离、哈密顿距离、余弦距离等。

这里以基于图的分割算法为例简单介绍下算法的流程：

先为所有相邻像素点生成边，这些边具有所连接的像素点位置信息和权重信息，其中的权重指的是相连像素点的差异程度，这里采用了颜色空间的欧氏距离来衡量，即 $w = \sqrt{\text{sum}((R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2)}$ 。例子中 R、G、B 分别表示三维颜色空间的红、绿、蓝三通道。除了用于衡量像素点之间的相似度的权重，还需要用一定的方式衡量已经聚好类的像素点簇的内聚性，以及能够用算法评估两个簇之间的相似性，簇与边权重之间的相似性。为此定义了簇的内聚性 $c = \frac{k}{\text{size}}$ ，其中 k 表示一个常量（通常对不同类型的图片有不同取值），size 表示簇的大小。同时为了表示簇和类的权重之间的关系用 $w + c$ 表示簇的权重。接着将所有边按权重大小排好序，并对所有边依次进行如下操作：

1、判断该边连接的两个像素点的权重是否都小于他们所在簇的权重，如果不是，跳过下面步骤直接分析下一条边，否则按如下方式将两个像素点所在簇合并为一个：

2、为了方便操作并查集来表示所有像素点所属的簇，当需要合并两个并查集的时候，只需要调用相应的 merge 函数，同时为了能够查到某个簇的大小以及权重，在并查集的基础上加了 sizes 和 values 两个变量。

完成所有像素点的聚类操作后，接着需要将大小较小的簇进一步合并（这里通过 min_pixels 变量来控制最小的簇大小）。通过合并能够很好的减少噪声对结果的影响。以上就完成了分割的过程，接下来需要从并查集中生成分割好的图片，具体做法如下：

完成所有像素点的聚类操作后，接着需要将大小较小的簇进一步合并（这里通过 min_pixels 变量来控制最小的簇大小）。通过合并能够很好的减少噪声对结果的影响。以上就完成了分割的过程，接下来需要从并查集中生成分割好的图片，具体做法如下：

从并查集中提取出所有的像素点集合，将集合的均值作为该簇的像素值，并还原到相应的图片位置上。

为了能够更好的过滤掉噪点对分割效果的影响，

对图像进行了一定的预处理过程，主要体现在实现了高斯模糊函数，并通过调整高斯模糊的半径大小和高斯函数sigma的取值来改变模糊的程度。

2.2 区域分裂合并

区域分裂的思想和区域合并的思想恰恰相反，区域分裂通过将图像不断分裂成小的组成部分来实现图片的分割，直到所有组成部分都无法继续分割下去，分割完成之后需要进行的是图像合并，即将分割好的小部分合并成适合的整体。

根据一次分裂的数量不同可以四分裂、九分裂、十六分裂等。分裂的终止条件是无法继续分裂下去，因此如何评估一个区域能否继续分裂是一个关键问题，通常用一个区域的内聚性来衡量是否可再分，内聚函数有很多种，如用最大颜色差异的倒数、平均颜色差异的倒数、最大颜色梯度的倒数等。分割后区域的合并也同样重要，对于比较普遍的规则形状的分割，区域的合并可以借鉴区域生长的方式，只不过这里用小的区域代替上面的像素。对不规则形状的分割，则需要结合具体区域的纹理走向进行分析。

区域分裂合并方法分割出的图像往往非常清晰且规则化，这是它的优点也是一个缺点，对结构化的图像分割效果较好，但对具有复杂纹理、颜色对比度高的图像分割的效果较差。

2.3 分水岭算法

分水岭算法源自拓扑理论的形态学的分割方法，它将图像看成是具有海拔高度的平面图，具体的高度值由颜色确定。图片的分割过程进而转变成为找到图像中的最低点和围绕最低点的高峰，进而形成一个盆地，也就是一个分割。它的基本思想借鉴了地理里面的形态学分析，和阈值法有一定的类似，但区别在于该方法还注重了像素点之间的位置关系。

该方法对边缘明显，区域内部颜色变化平缓图像具有很好的分割效果，而且分割出来区域的形状和实际非常拟合，同时这个方法对一些具备复杂纹理的图像效果不佳。

3. 遗传算法

遗传算法源自自然选择，这种思想如此广泛存

在以至于各各学科都能看到它的身影，在图像分割中也不例外。遗传算法由美国教授 Holland 在 1973 年提出，它的基本思想是将问题的解编码成字符串表示的基因，而自然界中基因的变异和重组看成问题的搜索过程，最后自然选择变为人为的筛选问题的解的过程。传统的问题求解算法往往存在计算复杂度过高以及局部最优解的问题，而基因算法在这方面能够很好的避免这个问题。由于基因的重组和变异过程具有很大的随机性，因此能够跳过局部最优解，而自然选择的过程就是淘汰不适应个体，选择优秀个体并留下后代的过程。

在图像分割领域，图像的分割方式被编码成一个字符串，字符串的排列组合和交替看成基因的重组和变异。通过衡量程序分割的效果得到更优的解。也正因为如此，遗传算法适用于全局搜索，但局部搜索能力不足，所以常把遗传算法和其他算法结合起来应用。将遗传算法运用到图像处理主要是考虑到遗传算法具有与问题领域无关且快速随机的搜索能力。其搜索从群体出发，具有潜在的并行性，可以进行多个个体的同时比较，能有效的加快图像处理的速度。但是遗传算法也有其缺点：搜索所使用的评价函数的设计、初始种群的选择有一定的依赖性。要是能够结合一些启发算法进行改进且遗传算法的并行机制的潜力得到充分的利用，这是当前遗传算法在图像处理中的一个研究热点。

4. 主动轮廓模型

主动轮廓模型是图像分割的一类非常经典的方法，这类方法具有很多的变体，但所有变体都有一个固定的模式，研究者可以在这个模式之下自主的开发新的算法。因此主动轮廓的图像分割模型具有很高的灵活性和实用性，通过定义不同的约束和初始轮廓以及轮廓的演变方法，可以得到更好的分割效果。

该方法的一个重要特点是在给定图像中利用曲线演化来检测目标，因此可以得到精确的边缘信息，这为准确提取图像中的信息提供了便利。它的基本思想是，首先定义一条初始曲线作为最初的分割面，然后利用图像数据得到能量函数，再通过最小化能量函数的方法来得到曲线不断向最优结果靠近的方式，从而最终和目标边缘重合。由于这是一种动态的逼近方式，在很多地方都具有很大的优势，比如求得的边缘来自于曲线的动态变化，因此是封闭的

且连续的。

主动轮廓模型分为参数主动轮廓模型和几何主动轮廓模型。其中参数主动轮廓模型将曲线或曲面的形变以参数化形式表达，而几何主动模型用几何的形式进行曲线演化。在参数主动模型中，有著名的 Snake 模型，该模型最初由 Kass 等人提出，Snake 的定义源自能量极小化的样条曲线，在曲线自身的内力和来自图像数据的外力的共同作用下自动向更优的边缘移动。

由于内力和外力的相互弥补和合作，因此不需要人为设定曲线的具体演化方式，它的优点是能够与模型直接进行交互，且模型表达紧凑，实现速度快，但缺点也很明显，就是对一些棘手的拓扑结构变化难以胜任，这一点恰好可以用几何活动轮廓方法来进行弥补。

5. 边缘检测

边缘检测方法基于这样的一个事实，即每个图像的区域和周围图像之间必然存在明显或暗含的边缘，这非常符合我们的直观认识。通过对图像边缘的检查，我们可以直接得到图像的分割线。常见的边缘检测方法为傅里叶变换，通过将空间域上的图片信息转化为频域上，从而高频的部分对应于原图像的边缘，低频对应于图像分割区域的内部。检测出图像的边缘后如何进一步对图像进行分割也是一个关键问题，因为得到的边缘信息往往非常复杂多变且缺乏规则性，需要拓扑结构的分析才能正确的将图像分割出来，因此这类边缘检测方法经常和其他方法结合起来使用。

边缘检测的优点是分割效率高，缺点是需要进一步的边缘组合，而且对于某些地方的分割不能得到闭合的分割边缘从而无法得到正确的分割结果。一个普适的启发式规则是先将大的明显的边缘先组合，然后再用分割的思想用已经识别好的边缘进一步分割。

二、应用和挑战

1. 图像分割的应用

图像分割是机器视觉的基础，它的应用非常广

泛，一个好的分割模型往往能够为上层应用带来巨大的效益。

1.1 医学影像识别

医学的病情诊断、治疗方案制定都需要对病人进行组织扫描，然后通过扫描得到的影像确定各内部器官的情况。过去这个工作都是靠有经验的医生来做，但是医生的判断受自身情绪的影响较大，判断结果也因人而异，而图像分割技术可以很好的应用到这个领域。一个优秀的图像分割识别系统能够快速高效的对病人的扫描影像进行分析，为了减少误判的可能，可以采用多个分割模型同时进行分割，并将它们作为医生最终诊断的依据。

由于机器影像识别具有以上的多种优点，很多地方都有了类似的产品投入使用，为医疗体系减轻了压力，为病人带来了福音。

1.2 图像优化

如何拍出好的图片，过去可能是先用相机拍摄，后期靠人工修补。现在很多软件都有美图的功能，比如背景虚化、高对比度、美颜、抠图。这些功能里面都离不开图像分割的支撑。图像分割能够将我们 Luc P , Couprie C , Chintala S , et al. Semantic Segmentation using Adversarial Networks[J]. 2016. 需要的部分分割出来单独优化，从而让图像优化变得非常简单，一个普通的用户只靠点击操作就可以完成绝大部分功能。

1.3 机器智能

人有五大感知能力，视觉是其中的关键一种，90%以上的外部都是通过视觉输入的，因此让机器拥有智能的第一步就是让机器具备人类的感知能力（尤其是视觉）。图像分割能够将视觉捕获到的信息分割成有意义的实体，从而为后面的语义理解和推理做好铺垫。像人脸识别、手写识别、物体检测等方向都取得了巨大的成功。

2. 现阶段的挑战

传统图像分割算法在各个领域取得了巨大的成果，但也存在很多的挑战。比如算法的普适性较差，算法的稳定性不够好。随着机器学习近些年来

的飞速发展，产生了很多的基于机器学习的图像分割算法，这些方法给图像分割提供了新的思路，同时也带来了更多的挑战。机器学习的方法非常通用，但需要巨大的特定数据进行训练，如何获取这些已经分割好的图片数据是一个问题，此外机器学习算法的可解释性较差，对一幅图像的分割缺乏理论支撑。

参考文献

- [1] Long, Jonathan, Shelhamer, Evan, Darrell, Trevor. Fully Convolutional Networks for Semantic Segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39 (4):640-651.
- [2] Yu, Fisher, Koltun, Vladlen. Multi-Scale Context Aggregation by Dilated Convolutions[J].
- [3] Chen L C , Papandreou G , Kokkinos I , et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, PP (99):834-848.
- [4] Pinheiro P O , Collobert R , Dollar P . Learning to Segment Object Candidates[J]. 2015.
- [5] Dai J , He K , Sun J . Instance-aware Semantic Segmentation via Multi-task Network Cascades[J]. 2015.
- [6] Luc P , Couprie C , Chintala S , et al. Semantic Segmentation using Adversarial Networks[J]. 2016.
- [7] Souly N , Spampinato C , Shah M . Semi Supervised Semantic Segmentation Using Generative Adversarial Network[C]// 2017 IEEE International Conference on Computer Vision (ICCV). IEEE, 2017.
- [8] He Kaiming, Gkioxari Georgia, Dollar Piotr, 等. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence:1-1.
- [9] He Kaiming, Gkioxari Georgia, Dollar Piotr, 等. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence:1-1. Learning to Segment Every Thing
- [10] Chen L C , Papandreou G , Kokkinos I , et al. Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs[J]. Computer Science, 2014 (4):357-361.