



多媒体通信技术

基于聚类的图像分割算法

作者： 罗雁天

学号： 2018310742

日期： 2018 年 12 月 18 日

目录

1	问题描述	2
2	算法描述	2
2.1	Kmeans 聚类算法	2
2.2	DeepLab v3 算法	2
3	实验结果	3
3.1	基于聚类的图像分割实验结果	3
3.2	DeepLab v3 实验结果	3
4	总结	4

1 问题描述

在计算机视觉领域, 图像分割 (Segmentation) 指的是将数字图像细分为多个图像子区域 (像素的集合)(也被称作超像素) 的过程。图像分割的目的是简化或改变图像的表示形式, 使得图像更容易理解和分析。图像分割通常用于定位图像中的物体和边界 (线, 曲线等)。更精确的, 图像分割是对图像中的每个像素加标签的一个过程, 这一过程使得具有相同标签的像素具有某种共同视觉特性。

简单来说, 图像分割可以看做是像素级别的分类, 其在医疗领域、自动驾驶等方面有着重要的应用, 在目前的算法研究中, 图像分类可以分为 Semantic Segmentation(语义分割) 和 Instance Segmentation(实例分割)。

本次大作业采用首先采用传统机器学习中聚类的方式来进行图像的语义分割, 然后使用近几年较好的算法与其进行对比。

2 算法描述

在本节首先介绍传统机器学习中的聚类算法, 然后介绍一种基于条件随机场与循环神经网络的算法。

2.1 Kmeans 聚类算法

聚类算法是一种无监督学习的算法, 训练样本的标记信息是未知的, 目标通过对无标记训练样本的学习来揭示数据的内在性质及规律, 为进一步的数据分析提供基础。基于聚类的图像分割不需要很大的训练集来进行监督训练, 相比于当前基于深度学习的图像分割算法训练时间更短。

给定样本集 $D = \{x_1, x_2, \dots, x_m\}$, “k 均值”(kmeans) 算法针对聚类所得簇划分 $C = \{C_1, C_2, \dots, C_k\}$ 最小化平方误差:

$$E = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|_2^2 \quad (2.1.1)$$

其中, $\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x$ 是簇 C_i 的均值向量。直观来看, 式 (2.1.1) 在一定程度上刻画了簇内样本围绕簇均值向量的紧密程度, E 值越小则簇内样本相似度越高。

KMeans 算法流程如算法1所示, 其中第 1 行对均值向量进行初始化, 在第 4-8 行与第 9-16 行依次对当前簇划分及均值向量迭代更新, 若迭代更新后聚类结果保持不变, 则将当前簇划分结果返回。

2.2 DeepLab v3 算法

DeepLabV3 算法是一种基于深度学习的图像分割算法, 是 DeepLab 系列的第三篇文章, 由 Liang-Chieh Chen, George Papandreou, Florian Schroff, Hartwig Adam 于 2017 年提出, 使用了基于 ResNet 的全卷积网络已经空洞卷积、ASPP(Atrous Spatial Pyramid Pooling) 模块等技术来做图像分割, 具体的算法细节在展示的时候已经讲解, 在此不再赘述, 可以 pre 文件夹下的文件。本

算法 1 KMeans 聚类算法

输入: 样本集 $D = \{x_1, x_2, \dots, x_m\}$, 聚类簇数 k

输出: 簇划分 $\mathcal{C} = \{C_1, C_2, \dots, C_k\}$

```
1: 从  $D$  中随机选择  $k$  个样本作为初始均值向量  $\{\mu_1, \mu_2, \dots, \mu_k\}$ 
2: repeat
3:   令  $C_i = \emptyset (1 \leq i \leq k)$ 
4:   for  $j = 1, 2, \dots, m$  do
5:     计算样本  $x_j$  与各均值向量  $\mu_i (1 \leq i \leq k)$  的距离:  $d_{ji} = \|x_j - \mu_i\|_2^2$ ;
6:     根据距离最近的均值向量确定  $x_j$  的簇标记:  $\lambda_j = \operatorname{argmin}_{i \in \{1, 2, \dots, k\}} d_{ji}$ ;
7:     将样本  $x_j$  划入相应的簇:  $C_{\lambda_j} = C_{\lambda_j} \cup \{x_j\}$ ;
8:   end for
9:   for  $i = 1, 2, \dots, k$  do
10:    计算新的均值向量:  $\mu'_i = \frac{1}{|C_i|} \sum_{x \in C_i} x$ 
11:    if  $\mu'_i \neq \mu_i$  then
12:      将当前均值向量  $\mu_i$  更新为  $\mu'_i$ 
13:    else
14:      保持当前均值向量不变
15:    end if
16:  end for
17: until 当前均值向量均未更新
```

次大作业使用此算法来对照传统机器学习中的聚类算法，分析和对比传统机器学习用于图像分割的效果和现在基于深度学习的图像分割的优缺点。

3 实验结果

3.1 基于聚类的图像分割实验结果

在本节，我们将上一节提到的聚类算法进行实现，使用了 python3.6 和 sklearn 包进行实验，由于使用聚类的方法对较复杂的图像进行分类得到的效果较差，因此我们选择了数据集 PASCAL VOC 2012 数据集上的较简单的图像进行分割。使用 Kmeans 聚类方法进行分割的实验结果与原始图像对比如图4.1、4.2、4.3、4.4所示。从图中可以看出，对于这种比较简答的图像，kmeans 分割的效果较好，能够较清晰的分割出物体，并且边缘部分分割较准确。但是对于较复杂的图像，kmeans 的分割效果就一般了，如图4.5、4.6所示，分割效果就很一般了。

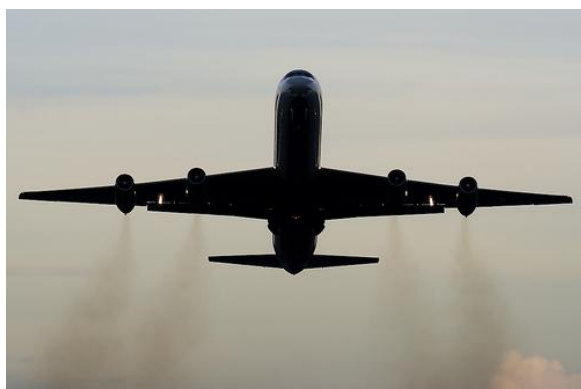
3.2 DeepLab v3 实验结果

此部分使用了 DeepLab v3 的开源 Demo 程序：DeepLabV3 Demo，使用此 demo 程序对上一节中分割效果较差的图像进行分割，分割结果如图4.7、4.8所示。从图4.7、4.8和图4.5、4.6的对比

中可以看出，基于监督训练的深度学习算法 DeepLabV3 相比于传统的基于聚类的图像分割算法的效果要好很多，这也是为什么当今使用深度学习来做图像分割很火热的原因。

4 总结

本次大作业主要使用了基于聚类的非监督学习算法 KMeans 来对简单图像进行分割，并且在复杂图像上将分割结果与现在很火热的基于深度学习的 DeepLabV3 算法进行对比，可以看出 DeepLabV3 算法在分割准确性上要高很多，因此现在多用深度学习的方式来进行图像分割。但是由于深度学习需要大量的数据进行监督训练，需要的数据量较高并且训练时间花费也很长，因此对于简单图像，我们依然可以用传统的机器学习的方式来进行图像分割。



(a) 原图像

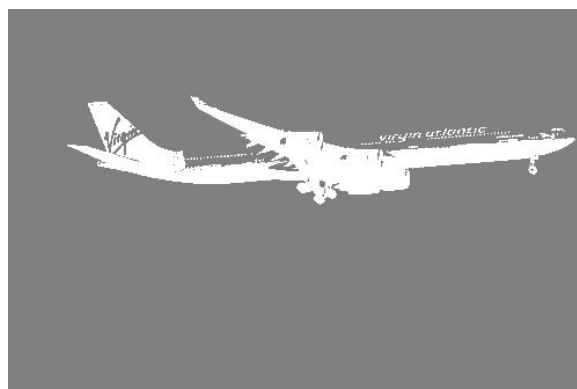


(b) 分割图

图 4.1: 使用 Kmeans 聚类算法进行图像分割



(a) 原图像



(b) 分割图

图 4.2: 使用 Kmeans 聚类算法进行图像分割



(a) 原图像



(b) 分割图

图 4.3: 使用 Kmeans 聚类算法进行图像分割



(a) 原图像

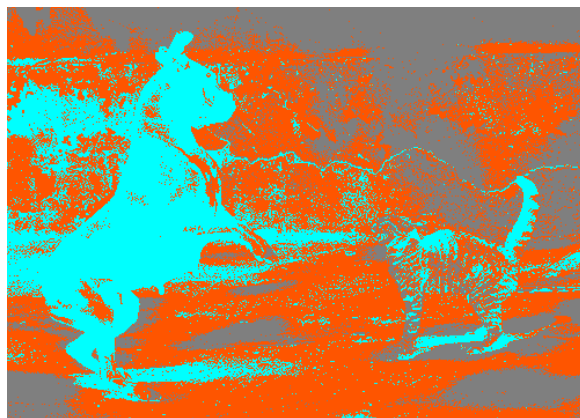


(b) 分割图

图 4.4: 使用 Kmeans 聚类算法进行图像分割



(a) 原图像

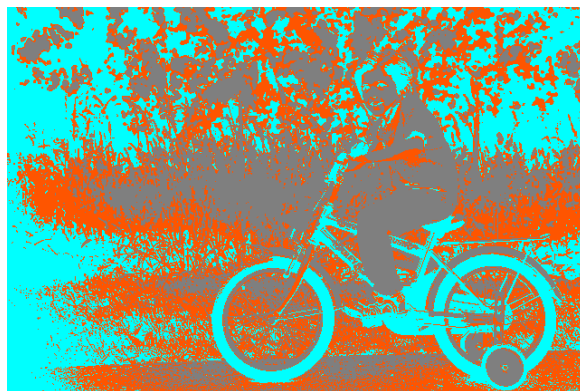


(b) 分割图

图 4.5: 使用 Kmeans 聚类算法进行图像分割



(a) 原图像



(b) 分割图

图 4.6: 使用 Kmeans 聚类算法进行图像分割

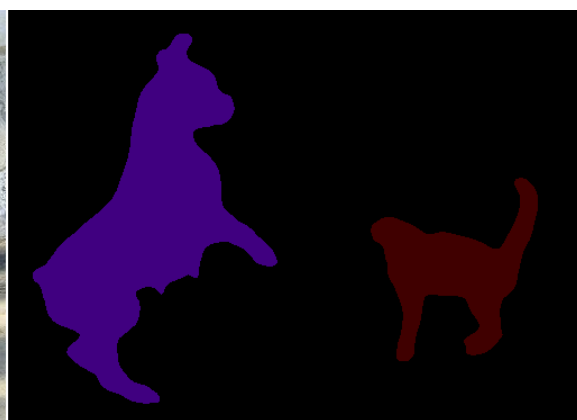


图 4.7: 使用 DeepLabV3 算法进行图像分割



图 4.8: 使用 DeepLabV3 算法进行图像分割