

8-14-2020

Object Detection in RGB-D Image Based on ANNet

Cai Qiang

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China;

Liwei Wei

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China;

Haisheng Li

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China;

Cao Jian

Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Object Detection in RGB-D Image Based on ANNet

Abstract

Abstract: The wide spread of depth images acquisition devices makes object detection in RGB-D images a hotspot in the field of computer vision. In order to make the features extracted by CNN more robust and to improve the detection accuracy, an improved CNN called ANNet was designed. *To enhance the model discriminability of local patches within the receptive field, some linear convolutional layers in the AlexNet with nonlinear convolutional layers were replaced which contained multilayer perceptron against the linear feature between convolution filter and underlying data patch.* The experiment result shows that the detection accuracy is improved by 3% in the RGB images and 4% in the RGB-D images on the NYUD2 datasets using the improved network.

Keywords

object detection, convolutional neural network, AlexNet, RGB-D images

Recommended Citation

Cai Qiang, Wei Liwei, Li Haisheng, Cao Jian. Object Detection in RGB-D Image Based on ANNet[J]. Journal of System Simulation, 2016, 28(9): 2260-2266.

基于 ANNet 网络的 RGB-D 图像的目标检测

蔡强, 魏立伟, 李海生, 曹健

(北京工商大学计算机与信息工程学院食品安全大数据技术北京市重点实验室, 北京 100048)

摘要: 由于深度图像采集设备的广泛使用, 使得利用 RGB-D 图像进行目标检测成为计算机视觉领域研究热点。为了使得利用卷积神经网络所提取的特征更具有鲁棒性, 设计了一种改进的卷积神经网络(本文称为 ANNet), 以提高检测准确率。为了提高卷积层中局部感受区域的模型分辨能力, 针对 AlexNet 网络中卷积层中卷积核与下层数据块的线性特性, 将部分卷积层改进为带有多层感知机的非线性卷积层。在 NYUD2 数据集上实验, 结果表明, 使用改进后的网络结构, 在彩色图像上的检测结果提升了 3%, 在 RGB-D 图像上的检测结果提升了 4%。

关键词: 目标检测; 卷积神经网络; AlexNet 网络; RGB-D 图像

中图分类号: TP391 文献标识码: A 文章编号: 1004-731X (2016) 09-2260-07

Object Detection in RGB-D Image Based on ANNet

Cai Qiang, Wei Liwei, Li Haisheng, Cao Jian

(Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China)

Abstract: The wide spread of depth images acquisition devices makes object detection in RGB-D images a hotspot in the field of computer vision. In order to make the features extracted by CNN more robust and to improve the detection accuracy, an improved CNN called ANNet was designed. To enhance the model discriminability of local patches within the receptive field, some linear convolutional layers in the AlexNet with nonlinear convolutional layers were replaced which contained multilayer perceptron against the linear feature between convolution filter and underlying data patch. The experiment result shows that the detection accuracy is improved by 3% in the RGB images and 4% in the RGB-D images on the NYUD2 datasets using the improved network.

Keywords: object detection; convolutional neural network; AlexNet; RGB-D images

引言

目标检测是计算机视觉、模式识别与机器学习领域非常活跃的研究方向^[1], 其任务在于从图像中检测并定位目标物体。随着深度图像采集工具的普及, 在 RGB-D 图像上进行目标检测成为目标识别

与检测领域的研究热点。目标检测的关键技术是如何从目标物体当中提取特征, 目标特征提取的质量, 直接关系到检测的结果。

在 RGB-D 图像上进行特征提取的方法主要分为两种, 一种方法是手工特征提取, 例如在彩色图像和深度图像上利用 SIFT^[2]、纹理直方图(texture histograms)^[3]等方法提取特征。考虑到深度图像的特性, 还有应用在深度图像上的深度核描述符(Depth Kernel Descriptors)^[4]、卷积 Fisher 核(Convolutional Fisher Kernels)^[5]等特征提取方法; 人工选取特征仅仅采集了对象单一方面的信息而



收稿日期: 2016-05-10 修回日期: 2016-07-11;
基金项目: 北京市自然科学基金(4162019);
作者简介: 蔡强(1969-), 男, 重庆, 博士, 教授, 硕
导, 研究方向为计算机图形学, 计算几何, 科学可视
化, 智能信息处理; 魏立伟(1987-), 女, 河北, 硕士
生, 研究方向为图像识别、机器学习。

<http://www.china-simulation.com>

• 2260 •

忽略了其他方面的信息, 例如 SIFT 特征采集的是像素点的水平和垂直方向的边缘信息, 而忽略了颜色信息, 并且有时不同的数据集需要提取的特征不一样, 从而特征的提取工作需要较强的专业背景, 因此出现了另一种利用深度学习模型进行特征提取的方法^[6-7]。

目前对于利用深度图像的目标检测方法分为 2D 物体检测和 3D 物体检测两种类型。在 3D 物体检测研究中, 贡献较为突出的是普林斯顿大学的 Song 等人提出一种利用深度卷积神经网络在 3 维空间进行类物体区域采样以及目标检测的方法^[6], 这种方法是目前利用深度图形进行 3D 物体检测的最新研究成果。

在 2D 物体检测研究中, 目前表现最好的是 RCNN^[8], Fast RCNN^[9]和 Faster RCNN^[10]三个迭代版本, 这三种检测架构引领了目标检测的潮流。伯克利大学的 Gupta 等人将 RCNN 扩展到了深度图像上^[7], 并且后续又改进了其部分缺陷, 将检测的目标扩展到包括识别非可见部分的全部目标物体。

卷积神经网络(Convolutional Neural Network)的概念最初是由 LeCun 等人^[11]提出, 他们提出了一种名为 LeNet-5 的多层人工神经网络结构, 这种网络能够成功识别手写体数字。由于网络的训练依赖于大量的训练样本和较高的计算代价, 当时的条件达不到相应的要求, 这种网络结构并没有引起足够的重视。

直到 2006 年, 网络训练的各种条件得以改善, 卷积神经网络重新被广泛研究。Krizhevsky 等人提出了经典的 AlexNet^[12]卷积神经网络, 这种网络相比其他的方法在图像分类上的准确率有了明显提升。AlexNet 的网络结构和 LeNet-5 的网络结构相似, 只是网络结构加深到 8 层。自此之后, 有许多工作集中到 AlexNet 网络结构的改进上, 用以提高模型分类的准确率。其中几种网络的表现尤其突出:

ZFNet^[13]将 AlexNet 网络中第一个卷积层的卷积核大小从 11*11 减小到 7*7 并且步长从 4 减小到 2, 给下层的卷积提供了更丰富的信息, 从而能够

提取到更为细致的特征;

VGGNet^[14]将网络结构加深到 19 层并且将每层的卷积核设定为 3*3 大小, 其实验的结果表明了网络的深度对于分类的准确率有较大的影响, 网络结构加深, 分类准确率提高; 由于该网络模型参数较多所以收敛速度较慢, 即使将计算分布在 4 块 GPU 上, 训练时长达到 20 d 之多;

GoogLeNet^[15]不仅拓宽了网络结构, 采用了 1*1、3*3 和 5*5 三种不同大小的卷积核, 而且将网络的层数增加到 22 层; 该网络结构使用了两个辅助分类器进行训练进而解决超深度模型中误差反向传播困难问题; 由于模型的深度和宽度都比较复杂, 所以该网络模型的训练需要强大的计算资源的支撑;

DeepID-Net^[16]网络是一个用于人脸检测的网络结构, 目前该网络已经有三个迭代的版本, 该网络由 4 层卷积层和 3 层池化层构成, 目前最新版本的 DeepID-Net 网络在人脸识别领域的表现已经达到最高水平^[16];

Network In Network^[17]网络将传统的卷积层改进为带有多层感知器的卷积层, 提高了提取特征的抽象程度, 从而增加了目标检测的准确率, 该网络模型是本文模型改进的想法来源。

另外还有很多其他的改进方法, 例如网络结构的改造、激活函数的改进或者过拟合问题的改进等等。本文从目标检测框架流程、ANNet 网络模型分析以及实验和总结几部分来进行介绍。

1 目标检测框架

我们采用的目标检测框架为 RCNN^[10], 其流程如图 1 所示。首先在图像上进行似物性采样, 每个样本提取大约 2K 个目标候选区域, 之后在数据集上进行模型训练, 用来提取特征, 之后, 利用训练好的模型在每个候选区域中提取特征, 将提取好的特征送入分类器中进行分类器的训练, 最后输入带有目标候选区域的测试数据集到分类器中获取检测结果。

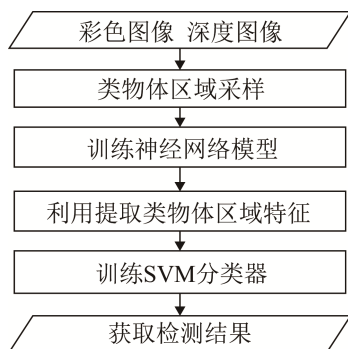


图 1 算法流程图

本文主要的研究工作集中在用于提取特征的网络结构的研究, 由于我们是在彩色图像和深度图像两种图像上进行特征提取, 所以分别训练两种网络。

由于卷积神经网络在较少的数据集上并不能现出很好的性能, 所以首先使用在较大数据集上例如 *Imagenet* 上预训练的模型, 在这个模型的基础上进行进一步调优训练(fine-tuning), 调优训练实际上就是在自己的数据集上继续训练, 与直接训练的差异是网络训练的初始化方式, 直接训练是按照网络定义指定的方式初始化(例如高斯随机初始化), 而调优训练是用预训练的模型文件训练好的参数来初始化, 用较少的迭代次数就可以达到较高的准确率, 因此我们采用在 *Imagenet* 数据集上训练过的网络模型作为预训练的模型。

对于作为用来提取特征的卷积神经网络, 我们做了很多思考, 在 *AlexNet* 网络结构中, 包含卷积层和池化层等, 其中卷积层是特征提取层, 池化层是特征映射层, 卷积层对于特征提取的质量起着至关重要的作用。我们认为卷积层还有改进的空间: 在卷积层中, 由于网络的输入维度较高, 如果下层中每个神经元都与上层所有输入相连, 也即全连接方式, 如图 2 所示, 这种方式使得网络参数会很多。

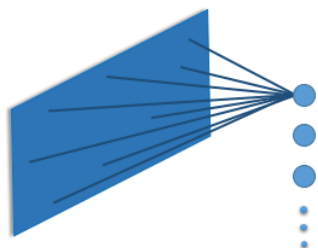


图 2 卷积层神经元全局连接方式

为了加快训练速度, 可以考虑将下层每个神经元与上层的局部相连, 即局部连接, 如图 3 所示。

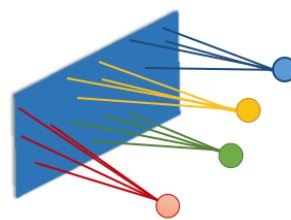


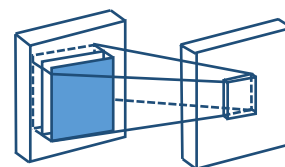
图 3 卷积层神经元局部连接方式

通常, 即使通过局部连接的方法之后, 参数依然很多, 所以目前的卷积都采用权值共享的手段来进一步降低参数数量, 也即同一层的卷积核使用相同的权值进行卷积。但是实际上同一图像当中不同位置重要程度不同, 采用相同的权重无疑降低了特征的抽象程度, 使得提取的特征鲁棒性降低。

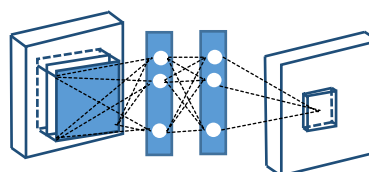
另外, 卷积层中的卷积核相对于的下层数据块来说, 是一个广义线性模型(GLM generalized linear model), 广义的线性模型抽象程度并不高, 当样本线性可分的时候, 线性模型表现良好, 当样本线性不可分时, 线性模型表现有待改进。

2 改进的卷积网络模型 ANNet

受到 *NIN* 网络模型的启发, 为了提高网络模型提取特征的抽象程度, 因此在卷积层中增加多层感知机, 将线性卷积层改进为包含多层感知机的卷积层, 形成非线性卷积结构。图 4 展示了两种卷积层的结构。



(a) 线性卷积



(b) 包含多层感知机的卷积层

图 4 线性卷积层和包含多层感知机卷积层对比图

在线性卷积层中, 通过线性滤波器加上非线性激活函数(例如 rectifier, sigmoid, tanh 等)形成特征图, 特征图的计算公式如下:

$$f_{i,j,k} = \max(w_k^T x_{i,j}, 0) \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (1)$$

通过增加多层感知机, 卷积层与池化层之间的关系不再是线性关系, 这两层之间通过多个线性滤波器的组合形成非线性关系。特征图的计算公式为:

$$\begin{aligned} f_{i,j,k_1}^1 &= \max(w_{k_1}^1 T x_{i,j} + b_{k_1}, 0) \\ &\vdots \\ f_{i,j,k_n}^n &= \max(w_{k_n}^n T f_{i,j}^{n-1} + b_{k_n}, 0) \end{aligned} \quad (2)$$

由于网络结构的改进是受益于 NIN 网络结构的启发, 所以本文取两个网络 AlexNet 和 NIN 的首字母得名 ANNet。如图 5 所示, 图中虚线框中为改进的带有多个感知机的非线性卷积层, 这种方法提高了特征提取的抽象程度, 进而提高物体分类的准确程度。需要说明的是, 在 AlexNet 网络结构中, 第 1 层和第 3 层的卷积层未被分组, 为了避免修改完的网络结构参数过多, 所以选取这两个卷积层进行改进。

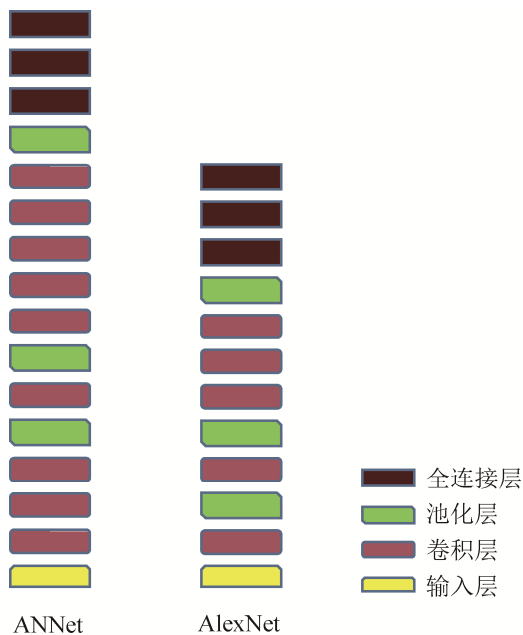


图 5 AlexNet 与 ANNet 网络结构对比图

由于原 AlexNet 网络中其中虚线框中为包含多层感知机的卷积层, 即网络的主要改进之处, 为本文的主要贡献。

ANNet 网络结构为 8 层网络, 其中前 5 层是卷积层, 后 3 层是全连接层, 最后一个全连接层的输出是分类的类别:

在第一层卷积层中, 采用 96 个 $11 \times 11 \times 3$ 的卷积核在步长为 4 的情况下对输入图像进行卷积操作, 之后紧跟两个卷积层 Cccp1 和 Cccp2, 其卷积核都为 1×1 ;

第二层卷积层中, 采用 256 个 5×5 的卷积核进行卷积操作;

第三层卷积层是 384 个 3×3 的卷积核进行卷积操作, 之后紧跟两个卷积层 Cccp3 和 Cccp4, 其卷积核都为 1×1 ;

第四层卷积层是 384 个 3×3 的卷积核进行操作;

第五层卷积层是 256 个 3×3 的卷积操作;

第六层和第七层是全连接层, 输出节点个数为 4096 个;

最后一层全连接层输出最终分类个数。

其中网络的各层参数如表 1 所示。

表 1 ANNet 网络各层参数

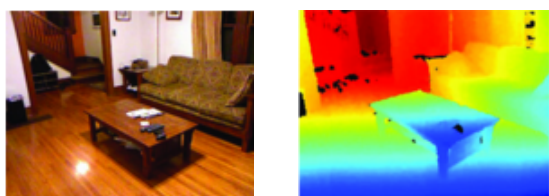
N	T	OD	PS	S
Conv1	C	$55 \times 55 \times 96$	11	4
Cccp1	C	$55 \times 55 \times 96$	1	1
Cccp2	C	$55 \times 55 \times 96$	1	1
Pool1	MP	$27 \times 27 \times 96$	3	2
Conv2	C	$27 \times 27 \times 256$	5	1
Pool2	MP	$13 \times 13 \times 256$	3	2
Conv3	C	$13 \times 13 \times 384$	3	1
Cccp3	C	$13 \times 13 \times 384$	1	1
Cccp4	C	$13 \times 13 \times 384$	1	1
Conv4	C	$13 \times 13 \times 384$	3	1
Conv5	C	$13 \times 13 \times 256$	3	1
Pool5	MP	$6 \times 6 \times 256$	3	2
Fc6	FC	4096	-	-
Fc7	FC	4096	-	-
Fc8	FC	21	-	-

注: N 代表网络层名称, T 代表网络层类别, OD 代表输出维度, PS 代表卷积核大小, S 代表步长。类别 C 代表卷积, MP 代表最大池化, FC 代表全连接层。

3 实验

3.1 实验数据集及评价指标

数据集：采用纽约大学公布的 NYUD2^[18]，该数据集共包含 407 024 对室内图，物体类别达到 894 多种，选取其中 1 499 对图像共 19 类物品，训练集为 795 对图像，测试集为 654 对图像。图 6 展示了其中的一对图像对。其中深度图像中不同的颜色代表距离的远近。



(a) 彩色图像 (b) 深度图像

图 6 深度图像和彩色图像

评价标准：评价标准选取 PASCAL VOC 数据集检测评价标准，平均准确率(average precision AP)。平均准确率计算查全率从 0 到 1 之间的查准率的均值，其定义为：

$$AP = \int_0^1 P(R)dR \quad (3)$$

其中：P 为查准率率，R 为查全率。

3.2 网络训练结果

本实验采用如图 5 所示的 ANNet 网络结构，使用在 Imagenet 数据集上预训练的模型在彩色图像和深度图像上分别进行调优训练，迭代次数为 30 K，训练时常约为 20 小时，比 AlexNet 网络训练时长增加，得到的模型大小分别为 218 MB 和 217 MB。彩色图像和深度图像迭代次数为 10 K、20 K、30 K 时得到的准确率和误差值如表 2 所示。

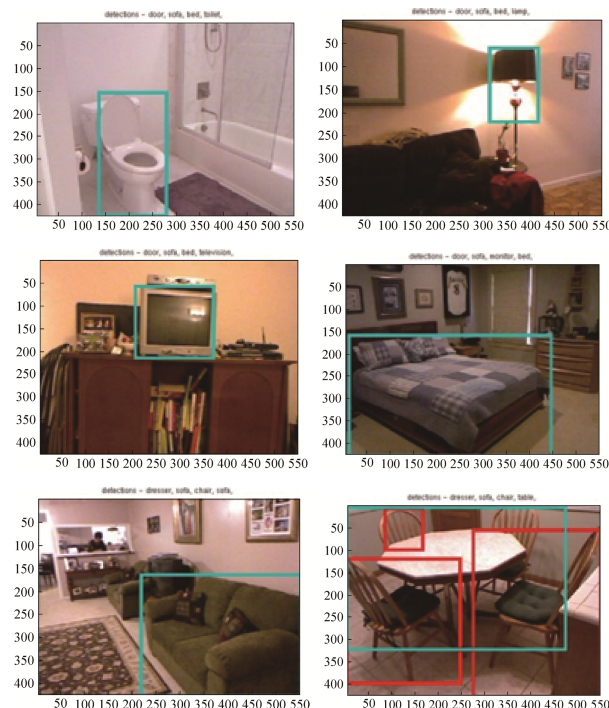
从表中可以看出,通过迭代次数不断增加，误差值逐渐减小，准确率逐渐上升。使用最大迭代次数 30 K 的模型提取第 6 层卷积层的特征，该层为全连接层，所提取的特征为全局特征。

表 2 网络模型训练结果

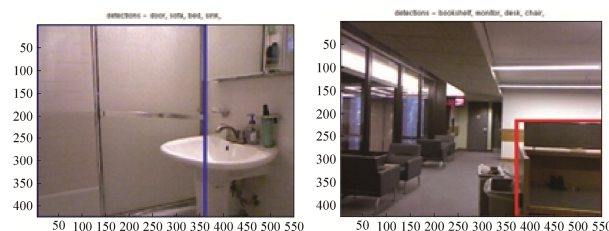
彩色图像模型训练结果			
对比指标	10K	20K	30K
准确率	0.750 625	0.752 266	0.759 92
误差值	3.156 1	2.961 034	2.565 18
深度图像模型训练结果			
对比指标	10K	20K	30K
准确率	0.807 187	0.809 375	0.815 938
误差值	1.042 5	0.961 034	0.822 918

3.3 物体检测结果及分析

物体检测结果：图 7 所示为一些类别的检测结果示例。其中图 7(a)展示了结果比较好的检测效果，图 7(b)中出现检测错误和检测不全的情况。



(a) 检测结果良好情况



(b) 检测结果错误和缺失情况

图 7 不同类别的检测结果

虽然我们一直不断追求更好的识别效果,但是检测的准确率仍没有达到非常令人满意的结果,从图 7(b)中可以看出,没有检测到图中沙发。表 3 展示了 19 类物体的平均准确率。可以看出准确率最高的类别是 bed 类,我们推测可能由于相同大小的图像当中,床这类物体的大小相比其他类别物体大的多,所以训练时候正样本数量相比负样本数量较多,所以这个类别的检测准确率相对于其他类别要高。

表 3 19 类物体检测平均准确率 /%

输入通道	RGB	RGB-D	RGB	RGB-
Bathtub	5.5	22.9	18.2	39.6
Bed	52.6	66.5	40.6	72.3
Bookshelf	19.5	21.8	27.2	36.0
Box	1.0	3.0	0.7	1.2
Chair	24.6	40.8	27.0	41.3
Counter	20.3	37.6	34.1	46.6
Desk	6.7	10.2	8.9	18.1
Door	14.1	20.5	14.7	21.5
Dresser	16.2	26.2	16.9	30.4
Garbage-bin	17.8	37.6	16.1	32.5
Lamp	12.0	39.3	25.9	34.3
Monitor	32.6	43.4	37.5	50.0
Night-stand	18.1	39.5	12.4	33.4
Pillow	10.7	37.4	14.5	37.6
Sink	6.8	24.2	25.8	40.3
Sofa	21.6	42.8	28.5	51.4
Table	10.0	24.3	11.0	21.1
Television	31.6	37.2	26.2	34.6
Toilet	52.0	53.0	44.8	47.1
Mean	19.7	32.5	22.7	36.3

从表 3 中可以看出,通过使用包含多层感知机卷积层的神经网络所提取的特征,相比传统的卷积神经网络,在牺牲一些时间效率的基础上,使得检测的平均准确率在 RGB-D 通道上提高了近 4 个百分点,在彩色图像上的平均准确率提高了 3 个百分点。另外,通过增加深度信息以后,检测的准确率相比单纯使用彩色信息有所提高。从整体上看,检测的平均准确率效果还是普遍偏低,在 19 个类别当中,达到 50%准确率的类别只有两类,检测的准确率有待进一步提升

4 结论

卷积神经网络目前已经被广泛应用在识别和检测的研究中,本文在原始的 AlexNet 网络结构基础上进行了改进,使用新的网络 ANNet 在彩色图像和深度图像上进行特征提取使得检测准确率稍有提升。实验的基础架构是基于 RCNN 的架构,今后会尝试其后期迭代版本 Fast-RCNN 和 Faster-RCNN。从结果可以看出利用深度信息和彩色信息相结合的方法比单独利用彩色信息得到的准确率要高一些。深度学习虽然相比传统方法有准确率方面的优势,但其弊端为网络参数过多,训练时间过长,如何减少网络的训练参数进而降低网络的训练时长是将来研究的方向之一。

参考文献:

- [1] 黄凯奇,任伟强,谭铁牛. 图像物体分类与检测算法综述 [J]. 计算机学报, 2014, 37(6): 1225-1240.
- [2] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (S0162-8828), 2002, 24(4): 509-522.
- [3] Leung T, Malik J. Representing and recognizing the visual appearance of materials using three-dimensional textures [J]. International Journal of Computer Vision (S0920-5691), 2001, 43(1): 29-44.
- [4] Bo L, Ren X, Fox D. Depth kernel descriptors for object recognition [C]// Proceedings of Intelligent Robots and Systems (IROS), 2011 IEEE. France: IEEE, 2011: 821-826.
- [5] Cheng Y, Cai R, Zhao X, et al. Convolutional fisher kernels for rgb-d object recognition [C]// Proceedings of 2015 International Conference on 3D Vision (3DV). France: IEEE, 2015: 135-143.
- [6] Song S, Xiao J. Deep Sliding Shapes for Amodal 3D Object Detection in RGB-D Images [J]. arXiv preprint arXiv: 1511.02300, 2015.
- [7] Gupta S, Girshick R, Arbeláez P, et al. Learning rich features from RGB-D images for object detection and segmentation [M]// Computer Vision—ECCV 2014. Zurich, Switzerland: Springer, 2014: 345-360.
- [8] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// Proceedings of the IEEE Conference

- on Computer Vision and Pattern Recognition. Columbus, USA: IEEE, 2014: 580-587.
- [9] Girshick R. Fast r-cnn [C]// Proceedings of the IEEE International Conference on Computer Vision. 2015: 1440-1448, Santiago, Chile. USA: IEEE, 2015: 91-99.
- [10] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]// Advances in Neural Information Processing Systems. Quebec, Canada: NIPS, 2015: 91-99.
- [11] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE (S0018-9219), 1998, 86(11): 2278-2324.
- [12] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [C]// Advances in Neural Information Processing Systems. Nevada, USA: NIPS, 2012: 1097-1105.
- [13] Zeiler M D, Fergus R. Visualizing and understanding convolutional networks [M]// Computer vision-ECCV 2014. Germany: Springer International Publishing, 2014: 818-833.
- [14] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv: 1409.1556, 2014.
- [15] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA: IEEE, 2015: 1-9
- [16] Ouyang W, Luo P, Zeng X, et al. Deepid-net: multi-stage and deformable deep convolutional neural networks for object detection [J]. arXiv preprint arXiv: 1409.3505, 2014
- [17] Lin M, Chen Q, Yan S. Network in network [J]. arXiv preprint arXiv: 1312.4400, 2013.
- [18] Silberman N, Hoiem D, Kohli P, et al. Indoor Segmentation and Support Inference from RGB-D Images [M]// Computer Vision-ECCV 2012. Berlin, Heidelberg, Germany: Springer, 2012: 746-76.

(上接第 2259 页)

方案性能的影响, 近距离范围单距离图像训练和远距离范围混合多距离图像训练自适应动态调整以优化识别性能和计算复杂度, 以及考虑扭曲翻转人脸图像的识别问题。

参考文献:

- [1] 宗文杰, 余青松, 张敏, 等. 基于 GIS 的视频监控系统研究 [J]. 计算工程与设计, 2011, 32(2): 745-748.
- [2] 罗翌陈, 杨辉华, 李灵巧, 等. SIFT 算法在 CUDA 加速下的实时人物识别与定位 [J]. 计算机科学, 2012, 39(11): 391-394.
- [3] Tsai H C, Wang W C, Wang J C, et al. Long distance person identification using height measurement and face recognition[C]//Region 10 Conference TENCON. USA IEEE, 2009: 1-4.
- [4] Y Yi, B Abidi, N D Kalka, et al. High magnification and long distance face recognition[C]//Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference, Baltimore, MD, 2006: 1-6.
- [5] 王婧瑶, 许勇, 曹本希, 等. 基于主成分分析的超声人脸识别识别算法研究 [J]. 计算机工程与设计, 2013, 34(8): 2868-2871.
- [6] 刘忠宝. 一种改进的线性判别分析算法在人脸识别中的应用 [J]. 计算机工程与科学, 2011, 33(7): 89-93.
- [7] 宋大伟. 一种改进的远距离视频图像性别辨识方法 [J]. 科技通报, 2013, 29(8): 121-124.
- [8] 李轶, 明东, 王璐, 等. 融合步态和人脸特征的远距离身份识别研究 [J]. 仪器仪表学报, 2011, 32(2): 264-270.
- [9] Wang Qunming, Shi Wenzhong. Utilizing multiple subpixel shifted images in subpixel mapping with image interpolation [J]. IEEE Geoscience and remote sensing letter (S1545-598X), 2014, 11(4): 798-802.
- [10] Chai Zhenhua, Sun Zhenan, Heydi Vazquez, et al. Gabor ordinal measures for face recognition [J]. IEEE Transactions on information forensics and security (S1556-6013), 2014, 9(1): 14-26.
- [11] Shantanu Rane, Petros Boufounos. Privacy-preserving nearest neighbor methods [J]. IEEE Signal Processing Magazine (S1053-5888), 2013, 30(2): 20-29.
- [12] 周慧, 陈澎. 利用 Mahalanobis 距离的 SAR 图像溢油识别 [J]. 计算机工程与应用, 2011, 47(18): 195-197, 203.