


博客 学





官方公众号



区块链大本营

原

## 视频人员

2017年09月19日 21:2

个人分类：深度学习

版权声明：本文为博主原创

### 一. 提出背景

目标：给定一

问题：可以定

这里提出的问

### 联系我们



官方公众号



区块链大本营

kefu@csdn.net

400-660-0108

QQ客服

客服论坛

### 关于我们 招聘 广告服务 网站地图

 百度提供站内搜索 京ICP证09002463号  
 ©2018 CSDN版权所有

经营性网站备案信息  
 网络110报警服务  
 北京互联网违法和不良信息举报中心  
 中国互联网举报中心

CSDN向您免费提供服务与产品的重要支持，希望您将csdn.net加入AdBlock

[答](#)
[商城](#)
[VIP会员](#)
[活动](#)
[招聘](#)
[ITeye](#)
[GitChat](#)

搜博主文章

👍 3名单，感谢支持！

2

写博客 发C

3



>

tion)

视频

行为识别

## Action Recognition

CNN

C3D

更多

[nolzhang/article/details/78034823](#)

为。

分类训练，解决一个输入视频的多分类问题。

的前提条件是：场景目标为单人，并且占据图片比较大的比例，如下图所示：



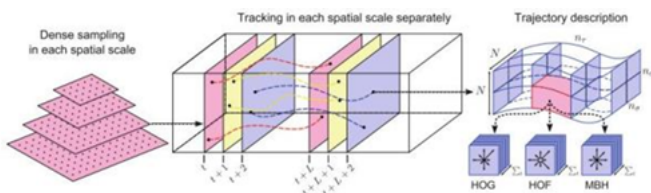
还有一类问题是基于行人检测，去估计行人的姿态和动作，暂时不在本篇讨论范围内。

## 二. 行为识别的发展

和其他领域一样，我们还是先从未被深度学习攻占的传统方法讲起，我们标记的里程碑算法是 iDT。

**论文:** Action Recognition with Improved Trajectories

iDT 方法是基于 DT (Dense Trajectories) 方法, 第一印象可以理解为 稠密光流 的轨迹。



如图所示，我们将算法描述为以下步骤：

- 1) 在原始图像多尺度上进行密集特征点采样，采样间隔为 $W$ （上图左）；
- 2) 进行有效的特征点筛选（只保留有用的），这里选用的方式是基于自相关矩阵的特征值；

和直接通过surf去选择特征点的思路基本上一样。

该 Step 形成空域信息。

3) 跟踪特征点, 在时间轴形成特征点上的轨迹序列 (上图左);

该 Step 形成

4) 对应每个时间

对序列进行编

5) 采用分类器

具体方法不再展

基于这个假设,

Detector 来消除人

联系我们

官方公众号

区块链大本营

kefu@csdn.net

400-660-0108

QQ客服

客服论坛

关于我们

招聘

广告服务

网站地图

百度提供站内搜索 京ICP证09002463号

©2018 CSDN版权所有

经营性网站备案信息

网络110报警服务

北京互联网违法和不良信息举报中心

中国互联网举报中心

进行特征采样 (HOG、HOF、MBH) (上图右),

可能会对光流有很大的影响。

来消除背景影响, 即通过 Surf 特征匹配来估算相邻帧的投影变换矩阵。另外, 论文设置

与估算)。

2

3

三. 深度学习方法

深度学习方法攻占该领域的时间是2014年, 开山之作也是具有代表性的 Two Stream 方法。

论文下载: [Two-Stream Convolutional Networks for Action Recognition in Videos](#)

来看其框架图:

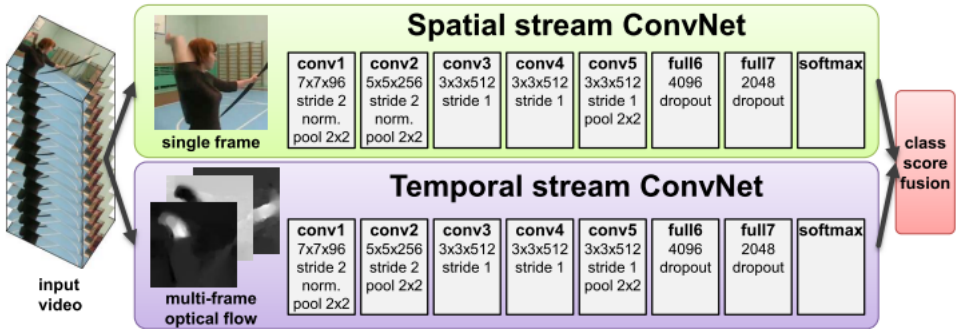


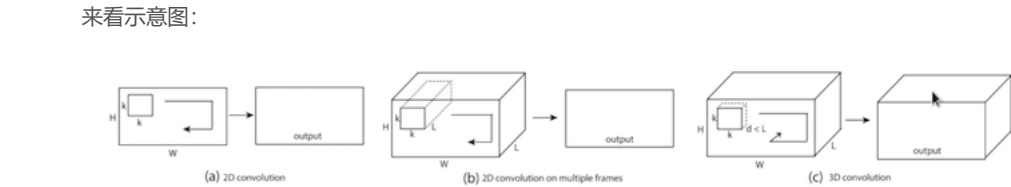
Figure 1: Two-stream architecture for video classification.

描述非常清晰, 通过CNN网络对 single frame 提取图像的特征信息, 下面通过多帧间的密集光流 (与iDT类似) 提出时域信息, 后面通过 fusion + 果。

针对 Two Stream 的改进比较多, 主要思路包括 网络的改进、Fusion 方法、结合RNN (LSTM)、选择Key Frame等, 这里没有太多创新的东西, refer一下。

这里重点提一下 C3D Network, 也就是3D卷积。

论文下载: [Learning Spatiotemporal Features with 3D Convolutional Networks](#)



与传统卷积的区别就在于将平面特征的提取扩展到3维, 将空域特征和时域特征同时提取, 该方法相比传统的2D方法, 效率有明显的提高, 基于VC 率达到了300FPS +。

虽然精度并不高, 但是C3D是该方向上的一个创新, 同样的基于视频的Task也将C3D看作是一个比较好的方法:

Code: <http://vlg.cs.dartmouth.edu/c3d/>

四. 参考数据集

Action Recognition 相关数据库比较多, 这里仅列出几个常用的供参考:

UCF101: [http://www.csfail.com/UCF101.html](#)

HMDB51: [http://www.hmdb51.net/](#)

Sports-1M: [http://www.sports1m.com/](#)

YouTube-8M: [http://www.youtube.com/watch?v=...](#)

ActivityNet: [http://activitynet.cs.cmu.edu/](#)

联系我们

官方公众号

区块链大本营

kefu@csdn.net

400-660-0108

QQ客服

客服论坛

关于我们

招聘

广告服务

网站地图

想对作者说点什么

百度提供站内搜索 京ICP证09002463号

©2018 CSDN版权所有

王金0803: 谢谢呢? (4个月前)

dashan8608: ...

large-human-motion-database

eo

load.html

2

3

频数据后, 是先需要对数据进行分成一帧帧图像的预处理, 还是说基于python的第三方库中有可以处

车子了, 好开心, 点赞! (11个月前 #1楼)

深度学习在视频动作识别应用

深度学习在最近十来年特别火, 几乎是带动AI浪潮的最大贡献者。互联网视频在最近几年也特别火, ...

来自: Touch\_Dream的博客

深度学习在视频行为识别中应用

深度学习在最近十来年特别火, 几乎是带动AI浪潮的最大贡献者。互联网视频在最近几年也特别火, ...

来自: dashan8608的博客

行为识别特征提取综述 ——2014年6月

行为识别特征提取综述 转自: [http://www.cnblogs.com/tornadomeet/archive/2012/06/22/2558548.htm...](#)

来自: xiaopihaierletian的博客

基于深度学习的视频检测 (一)

一、简介图像目标检测任务在过去几年深度学习的发展背景下取得了巨大的进展, 检测性能得到明显...

来自: fruit的博客

视频行为识别检测综述 IDT TSN CNN-LSTM C3D CDC R-C3D

本文github地址 Video Analysis之Action Recognition(行为识别) 行为识别就是对时域预先分割好的序...

来自: 万有文的博文

【神经网络与深度学习】【计算机视觉】视频中行为识别公开数据库汇总

1. Weizmann 人体行为库 此数据库一共包括90段视频, 这些视频分别是由9个人执行了10个不同的动...

来自: ZhangPY的专栏

SSH面试和笔试指导

适合人群: 所有人,章节: 视频介绍

视频中行为识别公开数据库汇总

1. Weizmann 人体行为库 此数据库一共包括90段视频, 这些视频分别是由9个人执行了10个不同的动...

来自: qq\_38080117的博客

视频行为识别第一讲: iDT算法

转发请注明出处: [http://blog.csdn.net/wzmsltw/article/details/53023363](#)iDT算法是行为识别领域中非...

来自: lin5885533的博客

视频行为识别年度进展

[https://mp.weixin.qq.com/s/?\\_\\_biz=MzI1NTE4NTUwOQ==&mid=2650326555&idx=1&sn=ffb945f2781...](#)

来自: java、c++、机器学习...

相关热词

bs视频教学视频 腾讯视频视频接口 c视频教学视频 视频视频纯音乐 视频地视频地址址

博主推荐

山水之间2018

关注

104篇文章

上园村蜻蜓队长

关注

0篇文章

lgy\_keira

关注

93篇文章

行为识别笔记: improved dense trajectories算法 (iDT算法)

iDT算法是行为识别领域中非常经典的一种算法, 在深度学习应用于该领域前也是效果最好的算法。由...

来自: Will Lin的博客

Video Analysis相关领域介绍之Action Recognition(行为识别)

这篇文章投稿在极视角公众号, 微信链接 随着深度学习技术的发展, 以及计算能力的进步 (GPU等) ...

来自: Will Lin的博客

<div><div>基于3D关节点的</div><div>文基于3D关节点的</div><div>C3D：视频动作分类</div><div>基于视频深度学习的</div><div>视频中行为识别公开数据集</div><div>基于深度学习的行人检测</div><div>python+opencv实现</div><div>图像识别训练样本集</div></div>	<div><div>联系我们</div><div><div><div><div></div><div>官方公众号</div></div><div><div></div><div>区块链大本营</div></div></div><div><div><div></div><div>kefu@csdn.net</div></div><div><div></div><div>400-660-0108</div></div><div><div></div><div>QQ客服</div></div><div><div></div><div>客服论坛</div></div></div><div><div>关于我们</div><div>招聘</div><div>广告服务</div><div>网站地图</div></div><div><div></div><div>百度提供站内搜索</div></div><div><div>京ICP证09002463号</div><div>©2018 CSDN版权所有</div></div><div><div>经营性网站备案信息</div><div>网络110报警服务</div><div>北京互联网违法和不良信息举报中心</div><div>中国互联网举报中心</div></div></div></div>	<div><div>in： 3D skeleton-based human... 来自： <a href="#">yeluohanchan的博客</a></div><div><div>temporal feature with 3DConvol... 来自： <a href="#">AUTO1993的博客</a></div><div><div>hCV学习 怎么使用OpenCV进... 来自： <a href="#">切磋砥砺，振衣起行</a></div><div><div>AI浪潮的最大贡献者。互联网... 来自： <a href="#">张天师的专栏</a></div><div><div>须车辆检测与车牌检测（一） 是其中的一个组成部分，我选... 来自： <a href="#">TripleLift`的博客</a></div><div><div>图像识别训练样本集ImageNetImageNet是一个计算机视觉系统识别项目，是目前世界上图像识别最大... 来自： <a href="#">liujiandu101的博客</a></div><div><div>人体动作识别（一） 最近几个月来，一直在研究人体动作识别和手势识别的课题。对于人体动作和手势的识别，本质上就... 来自： <a href="#">丫头的博客</a></div><div><div>视频行为识别[1]Flow Guided Feature: A Fast and Robust Motion Representation for Vid... 2018年论文 [ 1 ] Flow Guided Feature: A Fast and Robust Motion Representation for Video Acti... 来自： <a href="#">DavidZh</a></div></div></div></div></div></div></div></div></div>	<div><div>1.6万</div><div>2</div></div> <div><div>8332</div><div>3</div></div> <div><div>2860</div><div></div></div> <div><div>1.6万</div><div></div></div> <div><div>1.1万</div><div></div></div> <div><div>438</div><div></div></div> <div><div>1.2万</div><div></div></div> <div><div>90</div><div></div></div>
--	--	---	--

linolzhang

博客专家

关注

原创170

粉丝1119

喜欢629

评论191

等级： 博客 5

访问： 61万+

积分： 7312

排名： 4422

勋章：

最新文章

OCR文字识别（1）

基础网络架构探究之DiracNets

Facebook开源检测工具包 Detectron （by Caffe2）

TensorRT深度学习推理框架介绍

人群密度估计-Crowd Density

博主专栏

浅入浅出TensorFlow

阅读量： 69716

10 篇

深度学习基础

阅读量： 253288

25 篇

深度学习进阶

阅读量： 88122

8 篇

归档

2018年9月

1篇

2018年3月

2018年2月

2018年1月

2017年12月

展开

热门文章

Mask-RCNN技术解析  
阅读量：56925

目标检测-RCNN系列  
阅读量：33817

迁移学习：经典算法解  
阅读量：30473

目标检测 - Tensorflow API  
阅读量：24386

浅入浅出TensorFlow 7 - 行人检测之Faster-RCNN  
阅读量：20209

最新评论

GB28181技术基础之6 - T...  
lty1118： sb

浅入浅出TensorFlow 7 ...  
jixinpu： [reply]qq\_27535293[/reply] 注意一下，make.sh先后进入两个文件夹， ...

深度学习算法之YOLOv2  
qq\_36122764： 谢谢博主总结，请问YOLO V2里面的分类机制是怎样的呀？ 怎么还有softmax？ 我看源程序输出最...

基于视频的目标检测  
weixin\_40662331： 博主介绍的这两个模型代码都是基于Caffe框架下的， 请问有没有在Tensorflow框架下实现的"...

OCR文字识别（1）  
Maru\_Lim： 非常喜欢作者的写作风格， 支持

联系我们



官方公众号



区块链大本营

 kefu@csdn.net

 400-660-0108

 QQ客服

 客服论坛

关于我们 招聘 广告服务 网站地图

 百度提供站内搜索 京ICP证09002463号

©2018 CSDN版权所有

经营性网站备案信息 网络110报警服务

北京互联网违法和不良信息举报中心

中国互联网举报中心

2

3