

文章编号: 2095-2163(2019)01-0262-02

中图分类号: TP391.4

文献标志码: A

基于 EasyDL 的超新星自动搜寻系统设计

孙睿康

(淄博实验中学, 山东 淄博 255000)

摘要: 文章通过总结国内外现有的超新星搜寻方法, 基于百度 EasyDL 定制化图像识别平台, 提出了超新星搜寻系统设计。采用新疆星明天文台拍摄的可见光图像作为数据, 对超新星自动搜寻系统的设计进行了研究, 着重探讨了超新星特征提取、自动识别等问题。最终测试结果表明, 此系统能迅速地缩小需要人工确认的可见光图像数量, 从而准确地确定超新星。

关键词: 图像识别; 深度学习; 超新星; EasyDL

Design of automatic supernova recognition system based on EasyDL

SUN Ruikang

(Zibo Experimental High School, Zibo Shandong 255000, China)

[Abstract] This paper summarizes domestic and foreign methods of searching supernovas and presents the design of automatic searching system of supernovas based on EasyDL customized image recognition system. Visible light images are used photoed by Xinjiang Xingming Observatory as data, and the principle of design of automatic search of supernovas is studied, focusing on the problems of supernova feature extraction and automatic identification. The final results show that the system can rapidly reduce the number of images that require manual confirmation and precisely determine the supernovas.

[Key words] image identification; deep learning; supernova; EasyDL

0 引言

超新星作为恒星在演化接近末期时经历一次剧烈爆炸而形成的天体, 形成时会发出强烈的电磁波, 其中可见光波段是超新星搜寻的主要途径^[1]。超新星的可见光搜寻工作主要分为 2 类: 一类是专业天文学家所做的巡天工作, 使用大型望远镜自动搜寻大面积天空, 如美国加州大学 Berkeley 分校的 CCD 巡天和北京天文台兴隆观测站的 60cm 望远镜巡天^[2-3]。另一类为天文爱好者使用小型望远镜人工在较亮星系搜寻的工作, 如新疆星明天文台的公共超新星搜寻项目^[4]。前者的搜寻工作主要依靠计算机自动处理, 具有速度快、自动化程度高的特点, 但同样具有准确性不高、容易遗漏目标的缺点, 而且其图像识别系统与相关程序不公开。后者的搜寻工作准确性高, 但效率低下, 因为超新星爆发持续时间、视星等有限。可见光图像数据量庞大, 对超新星的人工搜寻效率低下, 会浪费大量人力。近年来, 信息技术迅速发展, 计算机的数据处理能力迅速上升, 使得人工智能应用更加广泛。机器学习作为人工智能领域的重要分支, 也迅速发展。计算机处理可见光图像具有速度快、准确性高的特点, 能有效弥

补上述 2 种超新星搜寻工作的不足, 故通过计算机模式识别来搜寻超新星的相关研究十分有必要。本文总结了前人的研究成果, 通过“EasyDL”定制化图像识别平台, 建立了可见光图像搜寻超新星的模型, 从而完成超新星搜寻。

1 图像数据的获取

本文选取了新疆星明天文台在 2015 年 7 月至 2018 年 7 月间拍摄的部分可见光静态图像。此静态图像是将新拍摄的图像与历史图像(以前拍摄的同一目标位置的图像)进行减法运算的产物。新拍摄的图像与历史图像均由 Celestron C14 望远镜拍摄, 视场 19 分 (Dec) * 25 分 (RA), 曝光时间 60s。理想状况下, 在静态图像中的超新星以圆形白色亮斑的形式存在, 对超新星的搜寻可以通过搜寻可见光图像上的白色亮斑的方式进行。获得的可见光静态图像总共 572 张, 从中人工选取了拍摄质量合格、拍摄时间距离超新星爆发时间接近的典型图像 335 张。此典型图像上的超新星亮度高、所占面积大、与图像背景的边缘明锐, 具有超新星的典型特征。

作者简介: 孙睿康(2000-), 男, 学生, 主要研究方向: 瞬变天体光变测量、天体自动识别系统。

通讯邮箱: 孙睿康 sunruikang2000@gmail.com

收稿日期: 2018-09-20

2 EasyDL 平台的使用

本文采用百度 EasyDL 定制化训练平台对图像进行深度学习。该平台支持定制图像分类和物体检测 2 类图像识别模型。采用物体检测模型能在一张图片上定制识别出每个物体的位置、数量、名称。平台用户创建模型后可上传并标注数据,然后训练模型并检验其效果,最终上线模型,获取应用程序编程接口(API)或离线软件开发工具包(SDK)。具体步骤如下:

(1)将 335 张典型静态图像上传 EasyDL 平台后,使用 EasyDL 平台的标注工具,拖动画框,圈出图像中的超新星,人工标注超新星在可见光图像的位置,使超新星与可见光照片背景间有一条分明的界线。

(2)标注完成后,训练该模型,令其进行机器学习。EasyDL 平台自动将图像的 70%用于训练,将 30%的模型用于检验模型效果。约 2 h 后第一次训练完成。

(3)重复训练操作 5 次。为了评估系统的效果,本文计算了 3 次模型训练结果的平均准确率(mAP)、精确率、召回率^[5]。

mAP 是衡量模型效果的指标。对于物体检测任务,每一类物体都可以计算出其精确率和召回率,在不同阈值下多次计算/试验,每个类都可以得到一条 P-R 曲线,曲线下的面积就是 AP 的值。“m”的意思是对每个类的 AP 再求平均,得到的就是 mAP 的值。 mAP 越接近 1,模型效果越好。精确率为正确识别的物体数与识别物体总数之比,召回率为正确识别的物体数与真实物体数之比。

3 测试结果分析

5 次模型训练结果的 mAP 、精确率、召回率见表 1。由于在检验模型效果时部分“supernova”标签被 EasyDL 平台错误地认为“标签检测有误”或“标签未被检测到”,训练结果的 mAP 、精确率、召回率的真实值应优于实验值。

实验结果表明,此超新星自动搜寻系统有一定

效果,能在较快的时间内识别超新星,能准确地反馈超新星在可见光图像的位置坐标,大大减少了搜寻超新星的工作量。系统采用 EasyDL 平台,可获取系统的 API 和 SDK,不仅能实现超新星的自动搜寻,而且通过适当扩充训练图像可以实现系统的进化,进而提高系统的识别效果,具有较强的可开发性。

表 1 测试结果

Tab. 1 Test results				%
训练次数	mAP	精确率	召回率	
1	63.51	61.39	75.61	
2	64.13	62.77	71.95	
3	62.92	74.24	59.76	
4	57.84	67.50	65.85	
5	67.07	67.86	69.51	

4 结束语

超新星自动搜寻系统经过长时间的发展,已经取得了较多的进展,其应用与相关应用(如近地小行星自动搜寻系统)前景广泛^[6-7]。本文依托百度 EasyDL 定制化图像识别系统平台,设计出了超新星搜寻系统。在未来的超新星搜寻工作中,可以使用“自动搜寻系统为主,人工确认为辅”的工作模式,提高搜寻超新星的效率。但由于时间与技术有限,本文设计的超新星自动搜寻系统在识别暗弱超新星、分辨噪点等方面有待进一步改进。

参考文献

- [1] 李宗伟, 王晓峰. 超新星研究进展[J]. 物理学进展, 2010, 30(3): 247-279.
- [2] 李卫东, 赵昭旺, 裴予雷, 等. 北京天文台超新星巡天系统: I. 巡天设备及软件[J]. 中国科学(A 辑), 1998, 28(3): 283-288.
- [3] TU Liangping, LUO Ali, WU Fuchao, et al. A method of searching for supernova candidates from massive galaxy spectra[J]. Science China Physics Mechanics and Astronomy, 2010, 53(10): 1928-1938.
- [4] 高兴. 新疆的明星天文台[J]. 中国国家天文, 2013(2): 66-75.
- [5] 秦莹华, 李菲菲, 陈虬. 基于迁移学习的多标签图像标注[J]. 电子科技, 2018, 31(8): 21-24.
- [6] 汪珍如. 超新星遗迹[J]. 天文学进展, 2001, 19(2): 238-240.
- [7] 林乎荣. 基于两晚各两帧 CCD 图像资料准确测量主带小行星地心距离[D]. 广州: 暨南大学, 2017.