



内蒙古农业大学学报(自然科学版)

Journal of Inner Mongolia Agricultural University(Natural Science Edition)

ISSN 1009-3575,CN 15-1209/S

《内蒙古农业大学学报(自然科学版)》网络首发论文

题目: 基于 Faster R-CNN 的西藏牧区牦牛检测算法
作者: 王菽裕, 李春国, 宋俊芳, 江英华
网络首发日期: 2020-09-28
引用格式: 王菽裕, 李春国, 宋俊芳, 江英华. 基于 Faster R-CNN 的西藏牧区牦牛检测算法. 内蒙古农业大学学报(自然科学版).
<https://kns.cnki.net/kcms/detail/15.1209.S.20200927.1703.002.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

基于 Faster R-CNN 的西藏牧区牦牛检测算法

王菽裕^{1,3}，李春国²，宋俊芳¹，江英华¹

(1 西藏民族大学信息工程学院，咸阳 712082; 2 东南大学信息科学与工程学院，南京 211189; 东南大学网络空间安全学院，南京 211189)

摘要：目前西藏牧区的牦牛养殖在我国畜牧业当中尚处于发展阶段，它的发展很大程度的影响我国牧区畜牧业的经济水平，尤其是对我国西藏牧区为主畜牧养殖业，彻底改革牧区的传统养殖方式非常必要，利用基于图像处理的牦牛目标检测就可以解决传统方式的耗时耗力问题。针对实际的牦牛放养场景，运用深度学习目标分类算法中具有代表性的 Faster R-CNN 网络结构，联合 ImageNet 中的牦牛数据集及采集的牦牛样本数据，把场景目标检测转换为目标二分类区别问题，进行牦牛目标检测。通过实际场景的实验结果及数据分析，牧区基于图像的牦牛目标检测方法在检测精度和执行效率上具有良好的检测效果。本文为解决西藏牧区牦牛目标检测提供了新技术，为改革牧区传统养殖方式上提出了新思路。

关键词：牧区；图像处理；深度学习；牦牛目标检测；Faster R-CNN；二分类

中图分类号 TP391

YAK DETECTION ALGORITHM BASED ON FASTER R-CNN IN TIBETAN PASTORAL AREA

WANG Shuyu^{1,3}, LI Chunguo², SONG Junfang¹, JIANG Yinghua¹

(1.School of Information Engineering, Xizang Minzu University, Xianyang 712082,China;

2.School of Information Science and Engineering, Southeast University, Nanjing 211189,China;

3.School of Cyber Science and Engineering, Southeast University, Nanjing 211189,China)

Abstract : At present, yak farming in pastoral areas of Tibet is still in the stage of development in China's animal husbandry. Its development has a great impact on the economic level of animal husbandry in pastoral areas of China. Especially in Tibet, the main animal husbandry in China, the traditional farming methods in pastoral areas is necessary to thoroughly reform. The time-consuming and labor-consuming problems of traditional methods can be solved by using image processing to detect yak target. Aiming at the actual yak grazing scene, the representative Faster R-CNN network structure is applied, combining the yak data set in ImageNet with the yak sample data collected, converted to the target two-class discrimination problem. Through the experimental results and data analysis of the actual scene, the image-based yak target detection method in pastoral areas has better detection accuracy and efficiency. This paper provides a new technology for yak detection in pastoral areas of Tibet, and puts forward a new idea for reforming traditional farming methods in pastoral areas.

Keywords :pastoral area; image processing; deep learning; yak target detection; faster R-CNN; two-category

西藏是我国重要的组成部分，其中绝大部分地区属于牧区，对于牧区的普通百姓，绝大部分家庭都有好几百的牦牛，牧区草场面积非常巨大，如何高效又准确的统计得到放牧结束的牦牛数目，是一个对于牧

基金项目：国家自然科学基金（61671144）；西藏自然科学基金（XZ2017ZRG-53(Z)）；校内重大培育（19MDZ03）；校内青年（20MDQ08）；National Natural Science Foundation of China (61671144); Tibet Natural Science Foundation (XZ2017ZRG-53 (Z)); Major Cultivation in School (19MDZ03); Youth in School(20MDQ08)

作者简介：王菽裕（1991—）男，助教，主要从事机器学习与模式识别方面的研究。

民非常必需解决的问题。我国的牧区发展还是非常传统的、人力投入为主的养殖方式，针对解决准确统计牦牛数目这一问题，现有的解决方式依旧是人为的统计牦牛数目，这种方式一方面不能完全准确的统计其数目，另一方面就是耗费人力；如何高效而有符合我国牧区环境的特点，提出一种准确统计牦牛数目的方法就显得尤为必要。由于我国西藏经济发展相对滞后，现代的技术手段在牧区的应用研究往往是很少。四川省草原科学研究院的赵洪文等人提出了一种基于视频数据的牦牛计数方法（申请号 201710524645.9）^[1]和动物进圈轨迹提取、统计模型构建、统计方法及装置（申请号 201910379411.9）^[2]；他们一方面是依赖于高分辨率的牦牛视频图像，一方面是依赖图像中人工设计的牦牛目标的表观属性特征进行检测，这样的方法泛化能力不够好。

本文提出牧区基于图像的牦牛目标检测属于目标检测和机器视觉的一个重要应用研究，目标检测可以运用传统的机器学习方法，其主要是通过对目标属性特征^[3]进行提取，属性特征提取的经典方法有 SIFT(scale invariant feature transform)^[4]，SURF(Speeded-Up Robust Features)^[5]，HOG(histogram of oriented gradient)^[6]等，再将以上方法得到的属性特征作为分类器的输入，进行类别的识别处理，其中分类器获取的方法有支持向量机 SVM(support vector machine)^[7]、迭代器 Adaboost^[8]等，以上的属性特征实际上是一种人为手工设计特征，手工提取特征的好坏直接影响着识别结果，同时针对某一个特定识别任务，提取属性特征的样本数据规模不够大，导致在实际应用中问题重重。

近几年深度卷积神经网络在理论和实践两个方面得到了迅猛发展，使得目标检测方法也有了质的发展规模；2014 年 R-CNN^[9]作为深度卷积神经网络目标检测框架的先导者，在 PASCAL VOC 2007 测试集的实验数据 mAP(mean average precision) 是 66%，微软亚洲研究院的 He 等人^[10]设计了一种新的网路结构 SPP-Net，然后 Ross B.Girshick 创新性的融合了 R-CNN 和 SPP-Net 的各自优势，提出了更加简单且有效的 Fast R-CNN^[11]，其依据 PASCAL VOC 2007 数据集作为测试数据，其实验数据 mAP 是 66.9%，微软亚洲研究院为了进一步提升和改善检测效率，就有了 Faster R-CNN^[12]算法的出现，此方法的核心其实是提出了一种全新的 RPN 网路结构，后来又相继出现了 YOLO^[13]类深度回归网路、依据 YOLO 的劣势点，再出现了 SSD^[14]网路等目标检测的优化网路结构，其中 Faster R-CNN 具备诸多的优势点，检测框架相当优

异、检测精度相对很高、扩展性优势明显、可改进性容易实现等，使之迅速成为目标检测算法中的佼佼者，本文就是采用 Faster R-CNN 作为牦牛目标检测算法。该算法的提出，为牧区牦牛目标检测提供了新思路，后期具有极大的应用价值。

本文采用基于视频图像结合 Faster R-CNN 网络结构作为牦牛目标检测的核心方法，应用在西藏牧区的牦牛数目统计和跟踪，与现有牦牛检测和计数方法的主要区别体现在：

1) 该方法首先是基于视频图像，可以认为是机器视觉技术在传统畜牧业领域中的应用，现有的牦牛管理靠人工安装在牦牛身上的传感器或者电子耳标，作为管理牦牛的便捷手段，基于视频图像可以克服人工耗时费力的问题。

2) 该方法并不是采用手工提取牦牛特征，手工提取特征往往利用目标的外观属性，诸如在专利（申请号 201910379411.9）^[2]中采用的外观特征有牦牛体长，牦牛体宽度，牦牛行走的轨迹等等，手工特征的弊端在于不能解决几何畸变、光照、颜色、角度的问题。

3) 该方法采用 Faster R-CNN 深度学习网路可以降低对几何畸变、光照、颜色、角度的依赖性，可以有效地克服牦牛外观属性特征的差异性，最重要的一点是可以根据训练样本数据自适应地构建牦牛目标的属性描述，具有更强的泛化和灵活性。

1 牦牛检测 Faster R-CNN 构建过程

本文牦牛目标检测方法的整个过程划分为两个步骤，分别定义为训练检测网路步骤和牦牛目标检测步骤。整个牦牛目标检测步骤的流程示意图如图 1 所示。

训练检测网路阶段，训练样本来源于两个部分，一个部分是刷选 ImageNet 数据集中牦牛图像样本，另外一部分是实际牧区场景下采集的牦牛图片，为了采用 VOC 2007 数据集的格式和算法评价机制，将所有训练数据样本的格式进行转换为统一标准格式。

牦牛目标检测阶段，将待测图像数据输入到已经训练好的检测网路模型中，得到最后的牦牛检测结果。

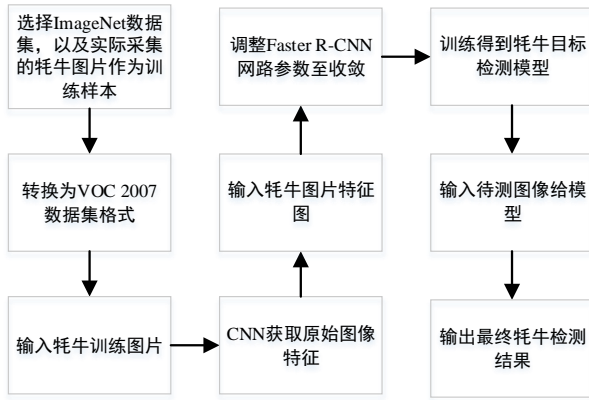


图 1 算法流程图

Fig.1 Algorithm flow chart

1.1 卷积神经网络 CNN

机器视觉中，卷积神经网络通过运用卷积层来实现自主学习训练牦牛目标的属性特征，这样做的好处就是可以避免人工定义牦牛目标本质特征刻画力度不够的影响，从而使得通过卷积神经网络提取的牦牛目标的属性特征更加准确，更加符合客观事实，更有利于得到精确的检测结果；图 2 为 CNN 的结构图，其中包括了 CNN^[15] 网络结构的主要组成部分：卷积层，卷积核，池化层，池化核，全连接层等。

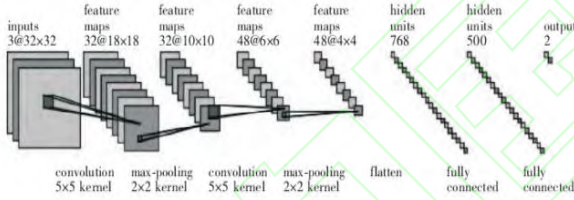


图 2 CNN 网络结构

Fig.2 CNN network architecture

1.2 Faster R-CNN 网络构建

Faster R-CNN 网络结构具有检测准确率高的优异特点，而且具有强扩展性、可改进性的优点。其主要是结合了 RPN 网络结构和 Fast R-CNN 网络结构的优势，重新构建的新型深度学习网络

1.2.1 区域建议网路 RPN

区域建议网路 RPN^[16] 定义描述如下 3 点：

1) anchor 就是特征图上的一个像素，以该 anchor 为中心，可以生成 k 种 anchor boxes (region proposal)，每个 anchor boxes 有着不同的宽高比和缩放比，论文中使用的 3 种宽高比是 (1:2, 1:1, 2:1)，3 种缩放比是 (64, 128, 256)，因此每个 anchor 产生 9 个 anchor boxes。

2) RPN 网络结构由 3 部分组成：

①头部：生成 anchor，使用的 3 种宽高比是 (1:2, 1:1, 2:1)，3 种缩放比是 (64, 128, 256)，生成 anchor；

②中部：分类层分支和回归层分支分别对这堆 anchors 进行各种计算

③尾部：通过对 2 个分支的结果进行汇总，来实现对 anchor 的初步筛选（先剔除越界的 anchor，再根据分类层分支结果通过 NMS 算法去重）和初步偏移（根据回归层分支结果），此时输出的都改头换面叫 region proposal 了。

3) 执行牦牛检测：最终依据 region proposal 的得分进行排序操作，选取在排序表中的前 300 个的 region proposal 为结果，将这 300 个结果输入到 Fast R-CNN 的网络结构，执行牦牛检测。RPN 结构如图 3 所示：

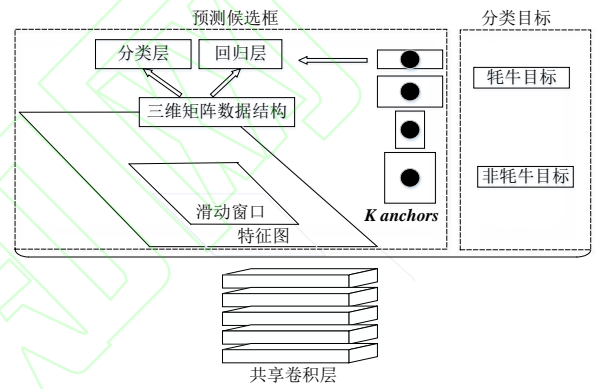


图 3 RPN 网络结构

Fig.3 RPN network architecture

1.2.2 区域建议损失函数

为了实现训练 RPN 网络的效果，本文算法的实现是为每一个 anchor 分配与之对应的一个二进制标签（是不是牦牛目标），具体的实施方式是牦牛目标的正标签的 anchor 是：与任意 GT (Ground-Truth) 包围有大于 0.7 的 IoU (intersection-over-union) 的 anchor；非牦牛目标的负标签的 anchor 是：与所有 GT 包围的小于 0.3 的 IoU 的 anchor。

其中 IoU 的定义就是真实边框和检测算法得到的边框之间的重叠程度，表达式如下式(1)：

$$IoU = \frac{Detection\ Result \cap GroundTruth}{Detection\ Result \cup GroundTruth} \quad (1)$$

经过上面内容的定义描述后，文章中的算法运用了 Fast R-CNN 中的多任务损失最小化目标函数，对于一副二维图像来说，将损失函数定义如下式(2)：

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (2)$$

上式中用 i 表示一个 mini-batch 中的 anchor 的编号；用 p_i 表示第 i 个 anchor 目标的预测概率。如果是文中的待检测目标牦牛，则为正标签， $p_i^*=1$ ；如果是非牦牛目标，则为负标签， $p_i^*=0$ 。 t_i^* 表示的是与正

anchor 对应起来的 GroundTruth 包围框坐标的向量。其中将用于计算分类的损失函数 如式(3)所示:

$$L_{cls}(p_i, \hat{p}_i) = -\log [p_i \hat{p}_i + (1 - p_i)(1 - \hat{p}_i)] \quad (3)$$

将用于计算回归的损失函数 定义为(4)表示:

$$L_{reg}(t_i, \hat{t}_i) = R(t_i - \hat{t}_i) \quad (4)$$

其中 R 表示的是 Fast R-CNN 中规定的用于计算鲁棒性的损失函数，如式(5)所示:

$$smooth_L(x) = \begin{cases} 0.5x^2 & |x| \leq 0.5 \\ |x| - 0.5 & |x| > 0.5 \end{cases} \quad (5)$$

1.2.3 Fast R-CNN 网路

Fast R-CNN 的输入包括有 3 个内容，具体是正负样本图像、正负样本图像的标定值(GroundTruth)以及目标建议 object proposals。经过真实边框和检测算法得到的边框之间的重叠程度，就可以得到每一个样本图像对应的一组感兴趣区域RoI (region of interests)。每一副图像的感兴趣区域RoI 组成来源如下表 1 所示。

表 1 RoI 的来源
Table 1 Sources of RoI

类别	比例	方法
1	30%	IoU 在[0.5,1]的区域，作为图像背景
2	70%	IoU 在[0.1,0.5]中较大区域，作为前景

Fast R-CNN 网路结构如图 4 所示，首先经过一些卷积层和池化层取得样本的卷积特征，然后再利用 RoI 池化层的运算操作对每一个感兴趣区域从卷积特征中进行特征提取，就可以获取到一个与之相对应的并且规格化的特征向量，接下来将所有的特征向量作为输入，进入到全连接层，把全连接层输出的结果进行共享，通过产生的两条支路，进入两个不同的网路层中，第一个网路层的责任在于运用 softmax 回归的方法计算两类(一类是牦牛，一类是非牦牛)目标外加一个“背景”类别的概率估计值；第二个网路层的责任在于输出检测框的位置，也就是用来表示每副图像上是不是牦牛目标的检测框坐标的 4 个值，输出的检测框坐标值用四元坐标表示为(r,c,h,w)，其中(r, c)为检测框左上角坐标，以及高度 h 和宽度 w。

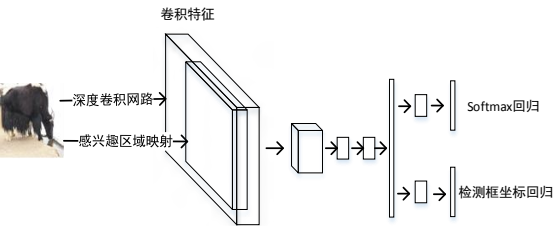


图 4 Fast R-CNN 网路结构
Fig.4 Fast R-CNN network architecture

1.2.4 Faster R-CNN 网路的构建

在 1.2.1 和 1.2.2 阐述了怎么为生成 RPN 网路结构进行学习训练，1.2.3 阐述了怎么样为生成 Fast R-CNN 网路进行学习训练以及训练过程中依赖于固定的待检测目标建议框，由于在上述叙述中 RPN 网路结构和 Fast R-CNN 网路结构都是进行独立学习训练的网路，所以需要想办法融合构建一种允许两个网路之间共享卷积层的方法，于是 Faster R-CNN 网路结构得以出现。经过在区域建议网路和 Fast R-CNN 两个网路进行交替学习训练得以改善优化，来学习得到可以进行共享使用的特征，从而产生 Faster R-CNN 网路，本文最后采用 Faster R-CNN 网路进行牦牛目标检测。

Faster R-CNN 模型激活的过程，大致如下步骤:

- 1)运用 ImageNet 数据集上预先学习训练好的初始化 RPN 网路指标参数，对 RPN 指标参数进行稍微改善;
 - 2) 运用 RPN 网路提取的候选框作为输入，训练 Fast R-CNN 网路各项指标参数;
 - 3) 运用训练得到的 Fast R-CNN 网路各项指标参数，从新初始化 RPN 网路指标参数，使得 RPN 网路卷积层固定不变，再次对 RPN 指标参数进行稍微改善;
 - 4)固定不变 Fast R-CNN 的卷积层，运用 RPN 提取的候选框作为输入，从新开始对 Fast R-CNN 的网路指标参数进行稍微改善;
 - 5)重复步骤 2)~4)，使得最后的网路模型收敛。
- 以上过程如图 5 所示:

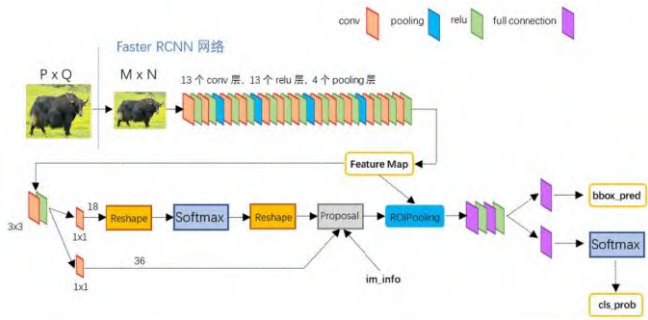


图 5 Fast R-CNN 网路结构
Fig.5 Faster R-CNN network architecture

2 实验结果及分析

2.1 数据集介绍

本文方法是简化了 Faster R-CNN 网路的经典的对 1 000 类物体进行的分类和检测，而是将牦牛检测问题分解为二分类区别的问题(是不是牦牛目标)。学习训练检测网路所采用的数据来源于 ImageNet^[17]数据集以及在西藏牧区采集的牦牛图像样本，本文还需要将数据集样本格式转换为 VOC 2007 的格式，从而适用 VOC 2007 的算法评价机制。

本文牦牛样本数据集包括有：牦牛样本图片 3 203

张/jpg, 对应的标签数据 3 203 个/xml, Bounding-Boxes 9 531 个/矩形框, 为了增大训练样本的数量, 我们对原有牦牛样本图片实施水平翻转操作, 数据集如下表 2 所示:

表 2 数据集
Table 2 Data sets

数据集	数量	水平翻转	备注
训练集	3 203	6 406	总数据集
测试集 A	1 676	/	部分
测试集 B	6 406	/	总数据集
测试集 C	3 278	/	部分

不同视角采集的部分牦牛图片样本如下图 6 所示:



图 6 部分不同视角下的牦牛样本图片

Fig.6 Yak sample pictures from different perspectives

2.2 实验结果及分析

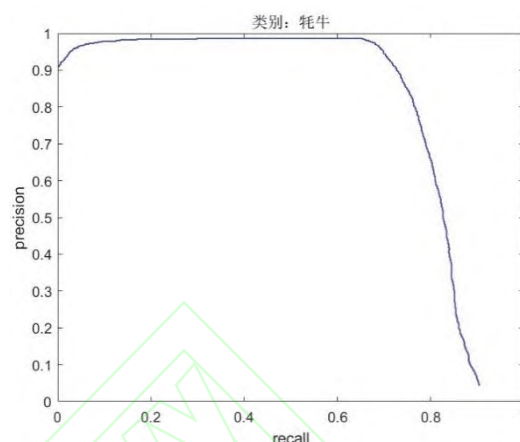
2.2.1 模型分析

Faster R-CNN 采用 ZF5 和 VGG16 模型对网路参数进行微调, 其过程大致如下步骤:

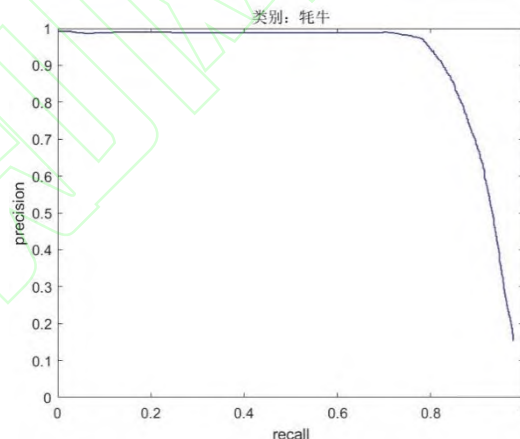
- 1) 运用 ImageNet 数据集上预先学习训练好的初始化 RPN 网路指标参数, 对 RPN 指标参数进行稍微改善;
- 2) 运用 RPN 网路提取的候选框作为输入, 训练 Fast R-CNN 网路各项指标参数;
- 3) 运用训练得到的 Fast R-CNN 网路各项指标参数, 从新初始化 RPN 网路指标参数, 使得 RPN 网路卷积层固定不变, 再次对 RPN 指标参数进行稍微改善;
- 4) 固定不变 Fast R-CNN 的卷积层, 运用 RPN 提取的候选框作为输入, 从新开始对 Fast R-CNN 的网路指标参数进行稍微改善;
- 5) 重复步骤 2)~4), 使得最后的网路模型收敛。

在训练集上利用 ZF5 模型和 VGG16 模型对检测网路训练完成之后, 在测试集 A,B,C 上进行实验, 实

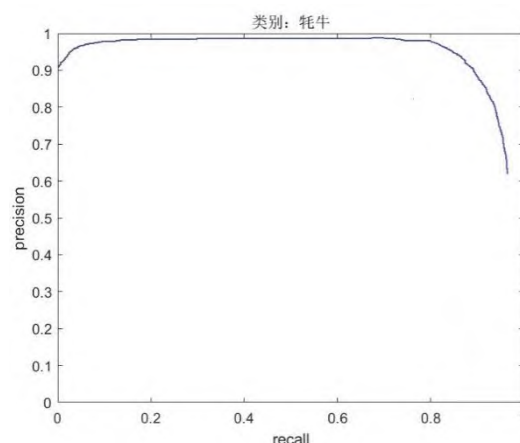
验结果如图 7 所示, 其中 AP 是平均准确率, 准确率召回率曲线(PR 曲线)与坐标轴包围起来的面积就是 AP 值; mAP 是不同种类 AP 值的平均。



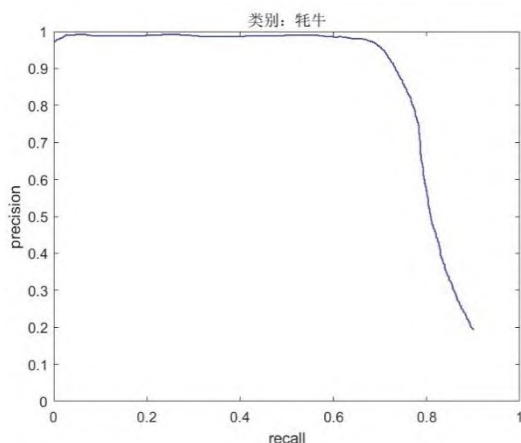
(a) 测试集 A(ZF5 模型, AP=0.742, 2 类)
图中 Precision 是准确率, recall 是召回率



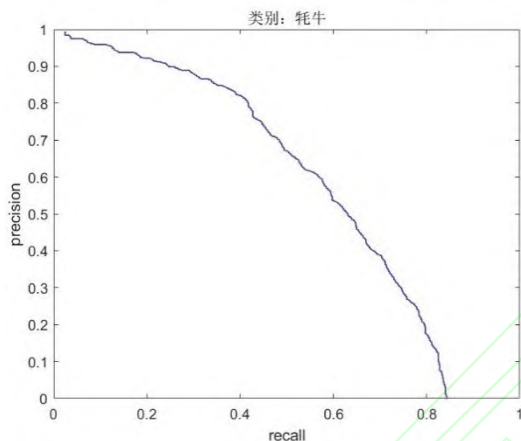
(b) 测试集 B(ZF5 模型, AP=0.808, 2 类)
图中 Precision 是准确率, recall 是召回率



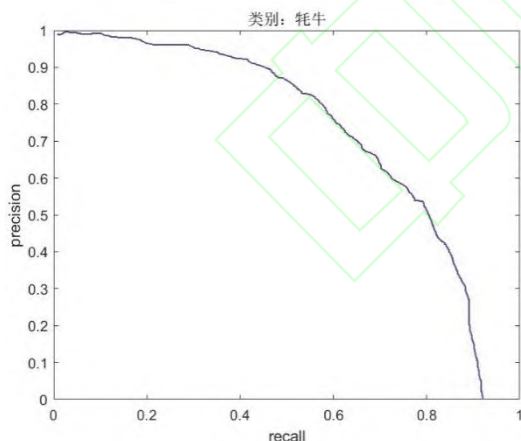
(c) 测试集 B(VGG16 模型, AP=0.873, 2 类)
图中 Precision 是准确率, recall 是召回率



(d) 测试集 C(ZF5 模型,AP=0.775,2 类)
图中 Precision 是准确率, recall 是召回率



(e) 测试集 B(ZF5 模型,AP=0.688,10 类)
图中 Precision 是准确率, recall 是召回率



(f) 测试集 B(VGG16 模型,AP=0.732,10 类)
图中 Precision 是准确率, recall 是召回率

图 7 不同网路模型、测试集下PR曲线

Fig.7 PR curves under different models and test sets

比较图 7 中的实验结果我们可以发现以下 3 条结论:

1)比较(a)和(b)与(a)和(d)图,随着学习样本的数量增加, mAP 值提高;

2)比较(b)和(c)图,随着网路模型深度的增加, mAP 值提高,深层模型的 mAP 值较于浅层模型提高很明显;

3)比较(b)和(e)图与(c)和(f)图,检测二分类相较于多分类问题, mAP 值更高。

图 8 为同一个测试集,不同模型下的牦牛目标检测效果。



(a)ZF5 模型,测试集 B



(b)VGG16 模型,测试集 B

图 8 牦牛检测效果

Fig.8 Detection effect of yak

表 3 为统计不同模型、样本数量下牦牛目标检测情况汇总;实验表明适度增大样本数量,可以提高深度检测网路的检测准确率;深度检测网路的卷积层增加,对于待检测牦牛目标的属性特征描述更加细致准确,也可以提高检测准确率;通过降低不必要的分类类别,从而可以减少检测的类别数目,这样也可以提高检测准确度。图 9 所示为不同模型,不同测试集下牦牛目标检测精度的对比。

表 3 不同模型、测试集下牦牛目标检测情况/%

Table 3 Yak detection under different models and test sets /%

实验描述	测试集 A (1 676 个)	测试集 C (3 278 个)	测试集 B (6 406 个)
ZF5,2 分类	74.2	77.5	80.8
VGG16,2 分类	79.1	83.5	87.3
ZF5,10 分类	63.7	66.1	68.8
VGG16,10 分类	66.9	70.2	73.2

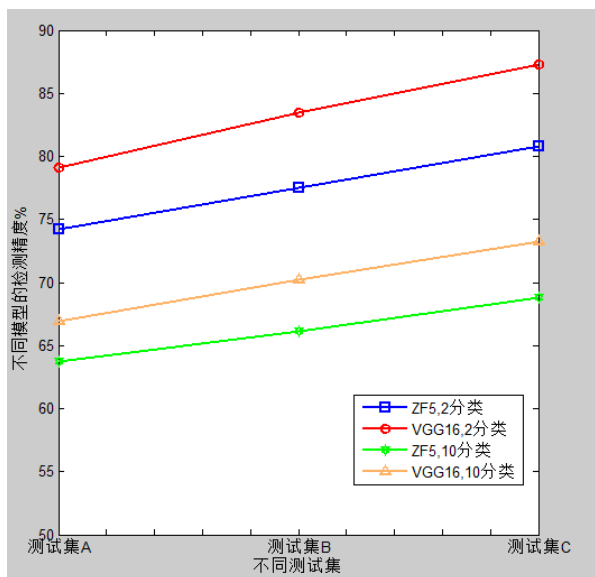


图9 不同模型,不同测试集下牦牛目标检测精度的对比
Fig.9 Comparison of Yak target detection accuracy under the different model and test sets

2.2.2 模型的测试实验

经过 2.2.1 验证模型在各个数据集上的有效性,以及对于特定的应用场景,如何选择合适的模型;我们新的测试集中,基于牦牛所在牧场环境的单一性,以及在训练模型过程中的时间复杂度和空间复杂度;最后选取二分类和 ZF5 模型作为检测模型,在特定测试集上的平均检测精度为 94.6%,存在漏检的现象主要是小牦牛被遮挡在大牦牛身下,以及牦牛目标在镜头画面不足 1/3。具体实验结果统计如下表 4 所示:

表 4 二分类和 ZF5 模型牦牛目标检测情况

Table 4 Yak detection under ZF5 models

视频	实际牦牛 (头)	测出牦牛 (头)	精度 (%)	平均精度 (%)
视频 1	83	76	91.6	94.6
视频 2	65	60	92.3	
视频 3	32	32	100	

3 结语

本文作者在西藏牧区考察之后,看到传统畜牧养殖业耗时耗力,利用基于图像的牦牛目标检测方法,将现代技术和传统畜牧业结合起来,利用深度学习下的 Faster R-CNN 进行牦牛目标检测。通过实验结果及数据分析发现,该方法平均检测牦牛的精度在测试集可以到达 83.3%,图像处理速率为毫秒级别,相较传统的人工牦牛计数的方式,无论在准确率和处理速率上得到了巨大的提高,同时,本文算法中训练检测网路阶段使用的牦牛目标数据是采集不同场景,不同视角的牦牛图片以及 ImageNet 数据集中的部分图片,使得对于该算法的通用性得到了提高,但是对于远场景下的微小牦牛目标无法检测,后续我们将继续开展更加深入细致的研究,为牧区畜牧养殖业的现代化智能化发展,贡献我们的力量。

参考文献

- [1] 赵洪文, 罗晓林, 安添午.一种基于视频数据的牦牛计数方法:中国 201710524645.9[P], 2017-6-30.
- [2] 王菽裕,李春国,宋俊芳等.动物进圈轨迹提取、统计模型构建、统计方法及装置:中国 201910379411.9[P], 2019-9-24
- [3] 白明月, 薛河儒, 姜新华,等.基于拐点的羊体测量点提取及体尺计算[J].内蒙古农业大学学报(自然科学版), 2017, 38(6):73-78.
- [4] Wang Y, Luo X, Ding L, et al. Object tracking via dense SIFT features and low-rank representation[J]. Soft Computing, 2019, 23(20):10173-10186..
- [5] Buch N, Velastin S A, Orwell J.A Review of Computer Vision Techniques for the Analysis of Urban Traffic[J].IEEE Transactions on Intelligent Transportation Systems, 2011, 12(3):920-939
- [6] Taigman Y, YANG Ming, Ranzato M A, et al. Deepface: closing the gap to human-level performance in face verification[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE Press, 2014 :1701-1708.
- [7] 胡学刚, 吴开元.基于 SVM 的显著性目标自动分割方法 [J]. 计算机工程与设计 ,2019,40(9), 2572-2577.
- [8] Freund Y,Schapire R E.A decision-theoretic generalization of on-line learning and an application to boosting [C] //Proc of Computational Learning Theory. [S.l.]: IEEE Press, 2014:23-27.
- [9] Girshick R, Donahue J, Darrel T, et al.Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]. 2014 :580-587.
- [10] He K, Zhang X, Ren S, et al.Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[C]//European Conference on Computer Vision. Springer, Cham, 2014:346-361.
- [11] Grishick R.Fast R-CNN [C] //IEEE International Conference on Computer Vision. IEEE Computer Society, 2015:1440-1448.
- [12] Ren S, He K, Girshick R, et al.Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Network[J].IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 39(6):1137-1149.
- [13] Redmon J, Divvala S, Girshick R, et al.You Only Look Once: Unified, Real-Time Object

-
- Detection[C]//IEEE Computer Society, 2016: 119-188.
- [14] Liu W, Anguelov D, Erhan D, et al.SSD: Single Shot MultiBox Detector[C] // European Conference on Computer Vision. Springer International Publishing, 2015:21-37.
- [15] Long J, Shelhamer E, Darrel T.Fully convolutional networks for semantic Segmentation[C] //Proc of IEEE Conference on Computer Vision and Pattern Recognition. [S.1.]: CVPR Press, 2015:1109-1123.
- [16] Felzenszwalb P F, Girshick R B, McAllester D, et al.Object detection with discriminatively trained part-based models[J].IEEE Trans on Pattern Analysis and Machine Intelligence, 2010, 32(9):1627-1645.
- [17] Russakovsky O, Deng Jia, Su Hao, et al. ImageNet large scale visual recognition challenge [J].International Journal of Computer Vision, 2014, 115(3):211-252.
- 