

基于深度学习和图像处理的马体尺测量设计

张婧婧

(新疆农业大学 计算机与信息工程学院, 新疆 乌鲁木齐 830052)

摘要:以伊犁马为研究对象,通过马场图像采集、实地测量,完成了马体尺测量中关键技术的研究。基于 YOLACT 实例分割技术,在 MS-COCO 数据集完成马体与背景的快速、高性能分割;采用边缘检测 Canny 算子对分割后的图像进行轮廓提取;在获取的马体轮廓上,对比动物特征点的 Harris 角点检测算法,提出动态网格的测点标定方法,完成马体尺特征点的数据标定,同时部分解决了马体站姿与摄像头不平行带来的体长修正问题;比较 Regress 及 Polynomial 的多元线性回归方式,量化、完成马体尺数据中胸围、管围的数据拟合及三维预测,并以像素为 640×480 两匹伊犁马体图像为例,获得了体尺测量结果。结果表明,基于深度学习和图像测量技术,可有效进行伊犁马体尺的自动测量并将其误差控制在较小范围之间,就大体型动物的体尺测量技术而言,该研究具备范例参考意义。

关键词:马体尺;YOLACT;实例分割;动态网格;数据标定;Regress 拟合

中图分类号:TP391.41

文献标识码:A

文章编号:1673-629X(2020)11-0180-05

doi:10.3969/j.issn.1673-629X.2020.11.033

Measurement and Design of Horse's Body Size Based on Deep Learning and Image Processing Technology

ZHANG Jing-jing

(School of Computer and Information Engineering, Xinjiang Agricultural University, Urumqi 830052, China)

Abstract: Taking the Yili horse as the research object, the key technologies in the measurement of horse body size are studied by image acquisition and field survey. Based on the example segmentation technology of YOLACT, the fast and high performance segmentation of horse body and background is completed in MS-COCO data set. The Canny operator of edge detection is used to extract the contour of the segmented image. Compared with Harris corner detection algorithm of animal feature points, the method of measuring points calibration of dynamic grid is proposed to complete the data calibration of horse scale feature points. At the same time, the length correction problem caused by the non parallel of the horse body and the camera is partly solved. By comparing the multiple linear regression methods of Regress and Polynomial, the data fitting and three-dimensional prediction of chest circumference and tube circumference in horse body size data are quantified and completed. Then taking the image of two Yili horses with the pixel of 640×480 as an example, the body size measurement results are obtained. It is shown that the automatic measurement of Yili horse's body size based on depth learning and image measurement technology can be effectively carried out and its error can be controlled in a small range. As far as the measurement technology of large animal's body size is concerned, this study has a reference value for example.

Key words: horse size; YOLACT; instance segmentation; dynamic grid; data calibration; Regress fitting

0 引言

马体尺数据是衡量马的生长发育及科学饲养、育种的重要依据。基于深度学习的马体测量技术重点解决了马体与背景的有效分割问题。YOLACT 是一种较为简单的全卷积网络实现实例分割的方法,该方法以其快速、易泛化且产生高质量的 mask 等技术优势适用于不同实例分割场景。马体的目标检测与背景分割过程采用该方法首先看中其在 MS-COCO 数据集^[1]

上约 30 mAP 的分割表现;其次马场目标检测对象及环境相对单一,目标图像的动态稳定性不作要求,且检测的精度相对不高,就 YOLACT 进行实例分割的图像低损失率而言,符合马体目标检测与背景分割的基本要求。

在马体与背景分割的基础上,体测图像中特征点的标定也是马体尺测量技术的重点。不同于传统的角点检测方法,设计中利用边缘检测 Canny 算子^[2]提取

收稿日期:2019-12-12

修回日期:2020-04-15

基金项目:新疆维吾尔自治区高校科研计划项目(XJEDU2020Y020)

作者简介:张婧婧(1981-),女,硕士,副教授,研究方向为农业信息化技术。

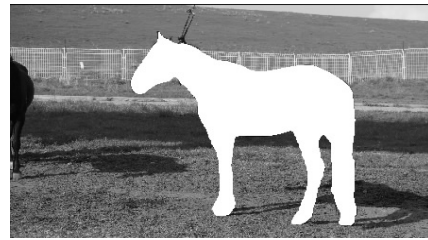
马体分割后的轮廓,在其轮廓线上采用动态网格完成体尺关键点的标定。该方法的研究旨在为相关动物体测技术提供范例参考。

1 马体的实例分割

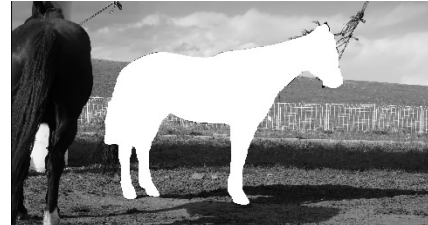
目前,大型马场的马养殖方式以散养居多,采集马体的图像信息时不可避免地纳入草场的背景、房屋、其他建筑甚至养殖人员等,为此采用深度学习技术实现马体目标与背景的实例分割,旨在为马体尺测量模型的建立准备条件。设计中采用的 YOLACT 实例分割方法属于目前深度学习的单级目标检测方法之一,在处理图像分割算法中,其以性能损失小、速度快而备受推崇^[3-5]。

1.1 YOLACT 的分割实例

YOLACT 将分割问题分解为两个并行的过程,利用全连接 fc 层和卷积 conv 层分别产生“掩模系数”(mask coefficients)和“原型掩模”(prototype masks),将掩模和预测的相应系数进行线性组合,并通过预测的 b-box 进行裁剪,实现掩模合成(assembly),计算中通过单个矩阵乘法实现。由此,该方法在马体特征空间中保持了空域一致性,且满足了马体目标检测的快速分割要求。以伊犁马的图像分割为例,文中采用 YOLACT 方法在训练好的 MS-COCO 数据集上完成马体目标检测与背景分割,如图 1 所示。



(a)马面水平朝左



(b)马面水平朝右

图 1 马体的实例分割

1.2 YOLACT 的网络结构

在 YOLACT 目标检测的前端网络中,低层的特征图信息量较少,但其特征图较大,适于检测目标位置准确且容易识别的一些小的物体;高层特征图信息量比较丰富,但是目标位置较为粗略,因此小物体的检测性能急剧下降。为了提高检测精度,YOLACT 的网络前端借鉴 Mask RCNN^[6]中采用的 FPN^[7]的网络结构,构建一个自底向上的线路,如 C_1 、 C_2 、 C_3 、 C_4 、 C_5 ;一个自顶向下的线路,如 P_7 、 P_6 、 P_5 、 P_4 、 P_3 以及对应层的链接,如图 2 所示。

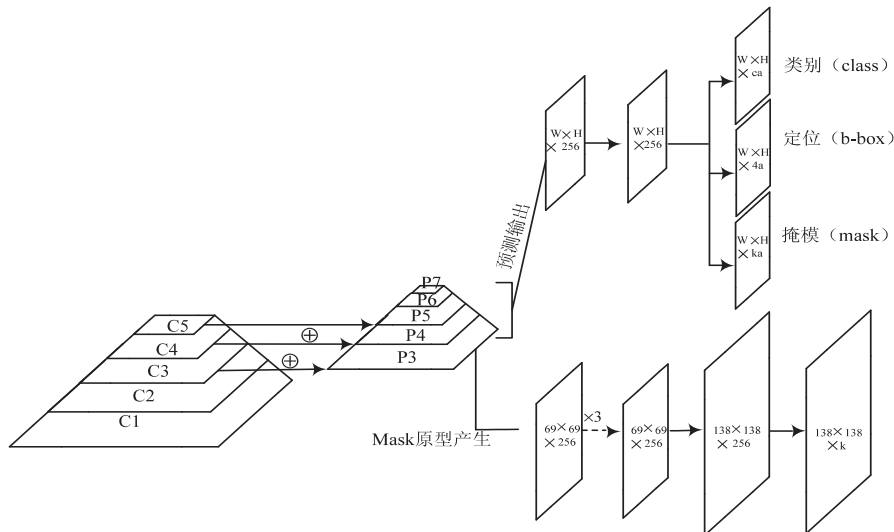


图 2 YOLACT 的网络结构

其中底层的特征图用于检测较小的目标,而顶层的特征图则用于检测较大的目标,对单目标的伊犁马体图像检测而言,其顶层特征图的检测性能要求相对较高。

Mask 原型的产生结构为 protonet,由全卷积网络 FCN 实现^[8],实验中预测马体图像时将生成一组(K 个)原型掩模,其最后一层含 K 个 channels(每个原型

一个),并将它们附加到 backbone 特征层中预测,如图 2 右下部所示。

作为 YOLACT 检测器的主干网络,其预测结构采用 RetinaNet 中带 FPN 的 ResNet-101^[9-10] 网络结构,其中头部结构设计如图 2 右上部所示。基于 Anchor 的典型目标检测算法中,预测头通常有两个分支,一支用于获得类别置信度,另一支用于进行边界框回归。

YOLACT 中添加第三个分支,预测 K 个掩模系数,用于对应每一个原型。马体图像分割实验中,从最终掩模中减去原型即可获得图 1 的分割效果^[8],并在非线性计算中产生稳定的输出。

1.3 马体分割图像的预处理

基于 YOLACT 目标检测与实例分割方法,文中对站姿较为标准、背景相对简单的马体图像进行了分割后的预处理。基于 YOLACT 分割后形成的马体掩膜,实验中分别将图 1 中透明掩膜换成白色和黑色,并通过异或运算去除背景,再与原图点乘,即可得到分割后的图像,由此完成马体测量模型的预处理。

2 马体尺测量点的标定

依据新疆马产业协会发布的地方标准,马体尺主要测量指标包括体高、体长、胸围、管围等,其中涉及的马体相关部位包括臀端点、胸骨前缘点、髻甲顶点、肩脚骨后缘、左前管部上 1/3 部等^[11]。具体而言,体高是指从髻甲顶点到地面的垂直距离;体长是指从胸骨前缘点到臀端点的直线距离;胸围是指肩脚骨后缘垂直绕胸一周的长度;管围是指左前管部上 1/3 部的下端最细处,水平绕其一周的长度。

2.1 马体边缘检测

据观察,马体尺标定所需的测点均分布在马体边缘轮廓上,为了得到马体尺测点,需要进行马体的边缘检测。实验中使用 Canny 算子对经过 YOLACT 实例分割的伊犁马体图像进行轮廓提取。

2.2 马体测点的 Harris 角点检测

在马体边缘检测的基础上,设计中尝试使用常见的 Harris 角点检测^[12-13]方法筛选出候选的测量点。通过调整角点检测的 blockSize 参数及 ksize 参数获得较多角点时,测量所需的测点与非测点并不能很好地区分;反之角点较少时,部分测点被遗失。

此外,马体尺测量点虽然大部分集中于被检测的角点之中,而鉴于马体体型、姿态的多样化,针对角点中测点的筛选以及设计更为通用的筛选模板仍较为困难。

2.3 动态网格的测点标定

综上,采用角点检测方法寻找马体尺的关键节点,其操作性不强且复杂度较高,本设计则采用了动态网格方法部分解决了马体测量中寻找测点的难题。

(1) 体高的获取。

对站姿各异、体型不同的马体而言,动态网格中各测点的标定均建立在其轮廓图像中平均像素值变化的基础之上。如图 3 中,水平方向的中部横线 H_m 表示马体像素点纵坐标的动态均值,计算方式如下:

$$y_m = \bar{y} \quad (1)$$

进一步地, H_u 表示 H_m 均值线以上像素纵坐标的动态均值,计算见式(2);而 H_l 则为 H_u 均值线以上像素纵坐标的动态均值,计算见式(3)。

$$y_u = \overline{y_m - \max(0, y_m - y)} \quad (2)$$

$$y_l = \overline{y_u - \max(0, y_u - y)} \quad (3)$$

其中,用于标定体高的髻甲顶点 A 恰好由 H_l 与马体轮廓的相交点得到。与此同时,根据髻甲顶点到前后足最低点所在直线 H_b 的距离,即为马体的体高,如式(4)所示。

$$ax + by + c = 0, \text{ height} = \frac{|ax + by + c|}{\sqrt{a^2 + b^2}} \quad (4)$$

其中, $ax + by + c = 0$ 表示马体前后足底部的连线。

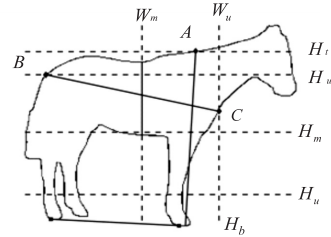


图 3 体尺测点的标定

(2) 体长的获取。

如图 3 所示,动态网格中 W_m 线表示马体像素点横坐标的均值,计算如式(5)所示:

$$x_m = \bar{x} \quad (5)$$

进一步地,网格中 W_u 线表示水平像素均值线 W_m 与其右边界线的等分线,计算如式(6)所示:

$$x_u = \frac{x_m + \frac{|\min(\text{flag} * x)|}{2}}{2} \quad (6)$$

其中,马面水平朝左时 $\text{flag} = 1$,朝右时 $\text{flag} = -1$ 。

如图 3 所示,在马体的体长计算中,臀端点 B 正是由 H_u 与马体轮廓的尾部交点得到;胸骨前缘点 C 则定为 W_u 与其轮廓交点中纵坐标较大的点,而体长的计算即为臀端点与胸骨前缘点的欧氏距离,如式(7)所示。

$$\text{length} = \sqrt{(x_B - x_C)^2 + (y_B - y_C)^2} \quad (7)$$

其中, (x_B, y_B) 为臀端点坐标, (x_C, y_C) 为胸骨前缘点坐标。

(3) 体长的修正。

在图 3 中,容易看出马体站姿与摄像头不平行,体长的像素值较实际值偏小,需进行数据修正,其修正方法如图 4 所示。

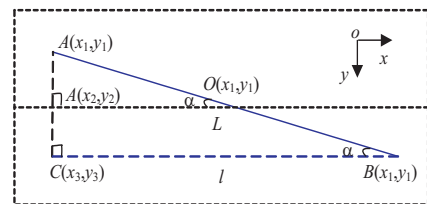


图 4 体长的数据修正

在图 4 中,由 $A(x_1, y_1)$ 、 $B(x_1, y_1)$ 确定的直线 L 表示马体中心线,其与马体前后足底部的连线 $ax + by + c = 0$ 平行,且与摄像头平面 1 形成角度为 α 的夹角,其中各参数间的关系如式(8)表示。

$$\cos\alpha = \frac{1}{\sqrt{1 + \left(-\frac{a}{b}\right)^2}} \quad (8)$$

将式(7)得到的体长数值 length 带入公式(9),即可得到体长的修正值 len_adjust 。

$$\text{len_adjust} = \frac{\text{length}}{\cos\alpha} \quad (9)$$

(4)胸围直径的获取。

与体高、体长的二维数据相比,基于平面图像获取马体的胸围和管围指标更为困难,为此引入两项相关性较强的平面指标:胸围直径和管围直径,用于后期马体胸围、管围的预测。

设计中,引入胸围直径旨在利用其与胸围的强相关性,即二者标定的测点在马体平面图像中重合。如图 3 所示,直线 W_m 与轮廓交点的纵坐标间距即为胸围直径 d_1 ,如式(10)所示,实际测量中其值与胸围数据相关性最强。

$$d_1 = |y_{1x=\bar{x}}^1 - y_{1x=\bar{x}}^2| \quad (10)$$

(5)管围直径的获取。

同样的,设计中引入管围直径在于其与管围的强相关性,即二者在马体管围测量中的测点一致。如图 3 所示,管围直径计算中,需要寻找动态网格线 H_g ,即为足底向上三分之一处与前后足轮廓的交点,从中获取最小值即为管围 d_2 ,如式(11)所示。

$$d_2 = \min(\Delta x_{|y=\bar{y}}^i), Y = \frac{2\max(\bar{y}) + \bar{y}}{3} \quad (11)$$

表 1 马体测点数据 (单位:像素)

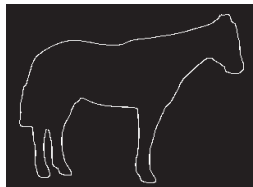
序号	flag	臀端点	胸骨前缘点	髻甲顶点	体高	体长 (修正)	胸围 直径	管围 直径
1	-1	(208,188)	(382,228)	(359,162)	191.64	183.76	80	11.00
2	1	(441,198)	(282,234)	(321,171)	179.21	169.35	74	13.00

以表 1 的像素数据为基础,根据人工测试与马体图像像素的多次匹配,选择比例为 1 : 1.21 进行像素

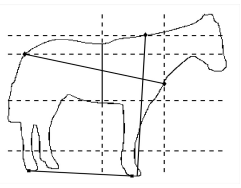
3 马体尺的测量实验

3.1 获取体尺数据

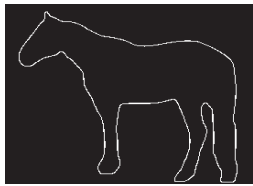
基于上文中马体图像分割、轮廓提取与特征测点的标定技术,对图像像素为 $640 * 480$ 的两匹伊犁马进行了体尺的测量,获取其特征测点,如图 5 所示。



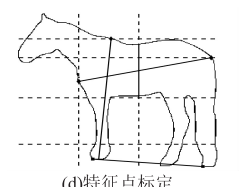
(a)flag 为-1 马体轮廓提取



(b)特征点标定



(c)flag 为 1 马体轮廓提取



(d)特征点标定

图 5 体尺特征测点的标定

据图 5 获取的体尺特征测点,分别得到两匹马的测点数据,如表 1 所示。

表 2 马体尺测量的对比(比例 1 : 1.21)

马体序号	测量方式	体高	体长(修正)	胸围直径	管围直径
1	图像	158.38	151.87	66.12	9.09
	人工	157	149	63	9
	误差	0.88%	1.93%	4.95%	1%
2	图像	148.11	139.96	61.16	10.74
	人工	147	143	59	10
	误差	0.75%	-2.1%	3.66%	7.4%

3.2 数据拟合与模型预测

依据马体尺数据的测量标准,三维体尺指标如胸围、管围的测量无平面图像的测量依据。前文设计中引入的胸围直径、管围直径指标,旨在利用其数据的相关性建立胸围与管围的预测模型^[14]。以 100 匹伊犁马体的人工测量数据为样本,完成体尺的胸围、管围预

测模型如下:

(1) 胸围的预测模型。

为了获得较好的预测效果,实验中采用了 Regress 及 Polynomial 两种多元回归方式完成马体胸围数据的预测,其预测结果如表 3 所示。

表 3 胸围预测结果

马匹序号	胸围预测	体高	体长	胸径	预测值	实际值	误差
1	Regress	158.38	151.87	66.12	173.618 3	172	0.94%
	Polynomial				174.466 5		1.43%
2	Regress	148.11	139.96	61.16	163.533 3	165	-0.90%
	Polynomial				166.930 7		1.17%

根据表 3 所示,选择误差较小的 Regress 回归模型进行马体胸围预测较为理想。其胸围的回归方程为:

$$Y_1 = 30.613\ 8 + 0.536\ 7 * h + 0.429\ 2 * w - 0.108\ 6 * d \quad (12)$$

其中, Y_1 为胸围, h 为体高, w 为体长, d 为胸围直径。

(2) 管围预测模型。

同样,实验采用了 Regress 及 Polynomial 回归方式完成马体管围数据的预测,其预测的结果如表 4 所示。

表 4 管围预测结果

马匹序号	管围预测	体高	体长	管径	预测值	实际值	误差
1	Regress	158.38	151.87	9.09	19.7	19	3.70%
	Polynomial				20.08		5.68%
2	Regress	148.11	139.96	10.74	18.52	18	2.92%
	Polynomial				17.71		-1.63%

根据表 4 可知,为了将测量误差有效控制在 5% 之内,仍选择 Regress 回归模型预测管围较为理想,其管围回归方程为:

$$Y_2 = 8.377\ 5 + 0.000\ 9 * h + 0.081 * w - 0.123\ 1 * c \quad (13)$$

其中, Y_2 为管围, h 为体高, w 为体长, c 为管围直径。

4 结束语

以伊犁马体为研究对象,通过采集马体的图像、马体尺的人工测量数据,完成了马体尺测量中四项指标的关键技术研究。

实验以深度学习的 YOLACT 实例分割技术为基础,首先实现了复杂背景下马体图像的快速、高质量分割;其次提出动态网格方式完成马体特征测点的数据标定,实时预测了马体在图像中的位置及身体方向,并部分解决了因其站姿不同带来的体长修正问题;然后采用 Regress 及 Polynomial 的多元线性及非线性预测方式,比较并实现了马体胸围及管围的预测,最后对两匹伊犁马体样本进行了体尺数据测量及误差计算,其各项误差均处于 0.75% ~ 3.7% 之间。就大体型动物体尺测量而言,该技术较为完整且具备范例参考意义。

设计的不足之处在于对马体的姿态仍有一定要

求,在马体的站姿过度倾斜或头部高昂的情况下,动态网格的测点将产生较大误差,使得测量不能准确完成。这些还有待进一步完善。

参考文献:

- [1] LIN T, MAIRE M, BELONGIE S J, et al. Microsoft COCO: common objects in context [C]//European conference on computer vision. Zurich, Switzerland: [s. n.], 2014.
- [2] 王文豪, 姜明新, 赵文东. 基于 Canny 算子改进的边缘检测算法[J]. 中国科技论文, 2017, 12(8): 910-915.
- [3] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(12): 2481-2495.
- [4] TREML M, ARJONA-MEDINA J, UNTERTHINER T, et al. Speeding up semantic segmentation for autonomous driving [C]//NIPS workshop on deep learning for action and interaction. Barcelona, Spain: [s. n.], 2016.
- [5] ZHAO H, QI X, SHEN X, et al. ICNet for real-time semantic segmentation on high-resolution images [C]//European conference on computer vision. Munich, Germany: [s. n.], 2018.
- [6] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN

(下转第 189 页)