

评估莫斯科市人口稠密地区的人流量和交通流量 使用固定电话的视频相机

R.M.沃罗比约夫

注解—在本文中，我们考虑了使用基于计算机视觉和深度神经网络的视频分析方法获取信息数据的问题，以及它们的进一步分析评估，以便自动化莫斯科人口稠密地区的交通统计过程，这与莫斯科政府有关。提出了一种解决此问题的方法，该方法可以计算所选位置的两种不同交通量：人和汽车。创建神经网络训练过程所必需的数据集。通过测量检测器质量的量度来评估算法的准确性。

过程 通过 完全的 例外
人来解决这个问题。

迄今为止，计算机视觉领域正在积极发展，包括朝着解决此类问题的方向发展。使用神经网络方法的方法也越来越多地用于各个领域，并成功地展示了良好的效果。

基于以上所述，我们可以得出结论，所考虑的问题尚未解决，但有发展方向，可以加快和改进分析过程。

工作考虑 **任务** **定义**
(识别) 每个单独的物体(人和车辆)并进一步分析其运动。它在于这样一个事实，即对于给定的视频文件，需要确定在几个给定方向之一上移动的人数、车辆数量。未来，将对这些数据进行分析(与时间间隔相关联，有助于识别“热门目的地”、“热区”(兴趣增加的区域))，并在此基础上构建图表和热图。这个任务涉及到图像中物体检测的区域。

二。关于本领域方法概述

由于此问题的主要问题是检测输入图像中的对象，因此搜索方法正是针对此主题进行的，以进一步编译基本算法。

为了检测，通常使用
两种主要方法(参见[18]、[19])：1) 基于检测对象的定性特征(颜色、大小、形状等)的方法；2) 基于神经网络的方法

关键词—对象检测、视频分析、计算机视觉、数据集创建。

使用定性特征的方法的本质是预先知道所需对象的一些重要和独特的特征。也就是说，搜索算法本身将取决于已知的内容。

我。在管理

目前，专门聘请人员解决交通评估问题，他们观看大量视频。现有的流量预估问题解决方案的弊端可以说明：1) 需要人工参与(1.1：需要雇佣大量人员；1.2：巨大流量计算误差大)；2) 巨大的时间成本(视频素材很多，每天都需要对其进行处理)。因此，迫切需要将此自动化

最常用的算法基于查找图像中对象的边缘，并使用负责所需对象参数的数据对其进行进一步分析。例如，如果已知检测到的物体大致尺寸，那么如果找到图像中所有物体的边缘，就可以猜测它们有多大，并找到最合适的搜索准则。

边缘搜索是通过评估亮度信号来执行的。也就是说，在物体的边界处，极有可能存在亮度差异。在基本方法中分析的正是这些差异。为了找到差异，使用了一种卷积算法(通常由 Sobel 算子)，即整个图像矩阵通过一个“掩码”，它

是一个矩阵，其元素代表不同符号的权重系数。通过“掩码”后，输出是一个矩阵，由沿其中一个轴的近似导数组成（取决于要通过图像的掩码）。接下来，对得到的矩阵的元素的平方和求平方根，形成最终的矩阵，有待进一步分析，即在物体的边缘会出现明显的高落差。这些方法在算法 [3]、[4]、[5]、[8]、[9]、[10] 中有更详细的描述。

选择图像矩阵 算法 是 分析 在 基础 颜色 成分。也就是说，如果预先知道目标物体的颜色，并且知道它会与背景不同，则可以搜索位于某个范围内的像素，如果该像素属于范围，如果不是则为零。这样就得到了一个矩阵，其中只剩下给定颜色的物体，然后通过剩余物体大小的近似估计，就可以找到想要的那个。

作为上述方法的替代方案，应考虑基于神经网络的方法。在现代方法中，由于技术组件的积极发展，基于深度学习的方法越来越受欢迎。这种方法的独特之处在于，神经网络可以针对各种同时特定的任务进行“锐化”。

随着这一领域的逐渐发展，构建神经网络的复杂性被降低为它们找到正确的层和系数序列的任务。也就是说，实际上，现代神经网络是一个“正确组装的构造函数”，需要

“提交”一个他可以学习的数据集，然后发布预测作为所提出问题的答案。

在寻找适合作为基础的算法的过程中，发现了一篇文章，其中编译了现代神经网络方法的顶部，这些方法可以根据它们的工作速度、有效性和初始训练数据的相关性来解决特定问题。[23]

在物体检测领域最引人注目的竟然是R-CNN，其意义在于将输入图像分割成大量的区域，每个区域都进行分类。此外，最适合期望对象的区域被认为是检测到的对象。

文章重点介绍了两个网络——YOLO 和 SSD，并对它们进行了比较。

A) 固态硬盘

使用单个深度检测图像中对象的方法之一

神经网络是SSD。在这种方法中，作者建议将具有边界框的空间离散化为一组用于不同比例和比例的对象的字段。每次调用时，网络都会生成每个存在的百分比

目的 里面 每个人 字段 和 产生 调整此框的大小以更好地适应对象的形状。此外，网络

结合了几个不同分辨率的特征图的预测，以更好地识别不同大小的物体。

SSD，由于这样的事实 他 充分地 消除猜测和后续像素（特征）重采样步骤，并且

将所有计算封装在一个网络中，易于组装、集成到系统中，要求 成分 检测 对象， 相对地 其他的 方法， 要求 对象的假设。

实验结果 在 各种各样的 数据集（PASCAL VOC、COCO）清楚地表明，对于使用额外步骤来猜测对象的方法，SSD 具有具有竞争力的准确性，并且在这样做时，执行

数数 很多 快点， 提供 单身的 结构 如何 为了 学习， 所以 和 为了 测试。此外，SSD 适用于原始尺寸较小的图像。

您可以在文章中了解有关 SSD 的更多信息。 [24]。

二) 优洛

YOLO 网络 [15] 可以作为 SSD 网络的替代方案。根据文章 [14] 的作者所述，YOLO v2 是当时最现代（最先进）的模型，用于检测图像中的物体问题，并在已知数据集（PASCAL VOC，可可），显示

这 伟大的 结果 经过 在准确性方面，超越替代方法（Faster R-CNN 和 SSD），同时在处理时间方面也胜出。

该网络的主要优点之一是它的多尺度训练方法，它允许对不同尺寸的原始图像使用 YOLO 模型，而不会损失处理速度和准确性。

YOLO 的第二个也是同样重要的优势是该网络实现了一种用于对象检测及其分类的联合学习方法。也就是说，该网络可以预测图像中对象的位置，而该对象不在用于检测器的数据集的标记中。

该网络的第三个重要优势是它的处理速度。在源图像中查找和分类对象的整个过程非常庞大，以至于可以实时处理 25 FPS 的视频文件。

在研究文献的基础上，决定开发和实施包含允许解决问题的基本方法的算法。这是因为问题远非微不足道，需要某种方法。它还需要开发能够彻底解决问题的成熟功能。

在较早考虑的网络中，基于比较特性，决定选择 YOLO 神经网络进行考虑，因为与其他网络相关的主要优势之一是其处理速度（支持实时视频处理）。此外，它的优点是可扩展性（即不绑定训练样本中的图像大小）。最后，由于使用了锚点，识别小物体的能力。

因此，为了解决这个问题，有必要开发一种算法，该算法将使用所研究的基本方法并为此任务重新制作考虑过的 Darknet Yolo 网络（选择系数，形成数据集，重新训练神经网络，在测试样本）。

三、和创建数据集

为了执行神经网络训练程序，有必要创建我们自己的数据集。发生这种情况的原因是现有数据集的对象检测质量存在问题，这些数据集取自互联网并用作神经网络的训练样本。在这方面，决定创建一个新的数据集。

从需要进行计数的摄像机的视频序列中拍摄（剪辑）的照片被用作图像。图像是从不同角度和相机分辨率拍摄的，以获得更多种类的数据。

为了创建必要的基地，有必要在初始图像上标记所有单独的车辆和人员的特定类别。也就是说，最终期望从两个类别中形成一组图像。

Labellmg 脚本，取自网络，后续修改为基于更方便快捷的标记

特定任务的要求。

该脚本允许您加载从图像目录顺序，产生

他们的通过突出显示标记矩形框并将标记的坐标保存为xml文件。由于 Yolo 神经网络的输入是带有标记框中心坐标的文本文件，

花了 写作 脚本， 哪个 会 执行将信息从 xml 文件转换为所需格式的操作。

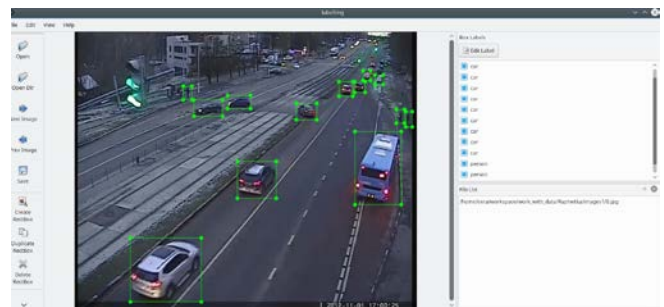


图1. 在 Labellmg 应用程序中标记数据集。

于是，对Darknet神经网络所需的数据集进行了标注。

四、P提出的算法

所提出的算法包含四个主要步骤：

- 步1：定位所有单个对象（交通工具、人）。
- 步2：估计运动物体的数量使用连续帧和位置数据的分析指定区域（跟踪所有对象的移动）。
- 步3：每个流量计数指示的方向（区域的连接“运动的开始-结束”）。
- 步4：第3步数据评估和整形“热点区域”和交通时刻表。

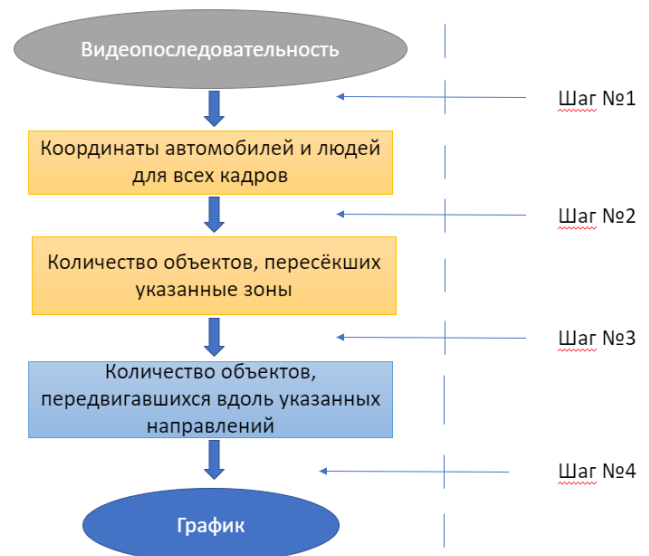


图2. 算法的简要说明。

五、关于经文步骤

步骤1:

曾是 尝试过 一堆 方法 到 对图像中的对象进行分割，但其中最好的结果是分割方法

使用渐变。这是因为物体通常具有相同的颜色（包括其上像素的亮度）和形状，因此，要搜索人，您需要找到宽度小而长度长的一种颜色图像中的高度数字，这将是 people。从这些条件来看，在人的“边缘”处，沿水平和垂直坐标的梯度应该变化很大。对于车辆来说，一切都一样，只是形式会比人类大得多。这将分割图像。所有符合这些条件的对象都将被进一步检查是否存在误报（如果对象的宽度比图像宽度和高度类别中的每一个都小得多，或者相反，大于一定百分比，则这些对象将被忽略）。

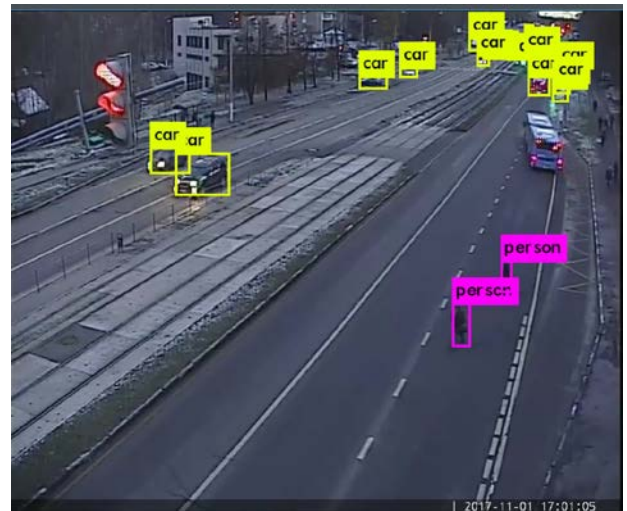


图3. 物体检测。

在 给予 步 有一个很大 由于他们可以彼此站得非常近（人群情况），因此与计数人员有关的错误，因此将被分割为一个对象，这是不正确的。

为了避免此类问题，尝试了一种方法，将分割后的对象“进行细化”发送到检测神经网络，而检测神经网络又必须确定每个单独的对象（人群中的人或巨大交通流中的汽车）。

但由于对象的不正确分割，它也被证明是低效的。比如有的时候一个人分到两个不同的段，导致弄清楚后，结果翻了一番，就造成了很强的错误。

到 摆脱 从 问题 错误 分割，提出了一种替代方法，在原始图像上使用 YOLO 检测神经网络来确定单个对象的位置。即，将检测物体的任务完全转移给神经网络。它应该在原始图像上找到所有预训练的类（对象）。

为了正确解决任务，需要先前收集的数据集（人员、交通）。

第2步:

在这个阶段，找到的两个类（车辆，人）的对象都有坐标。

现在有必要使用从连续帧中获得的数据分析来跟踪这些对象。也就是把属于一个移动的坐标连接起来

里面的物体 每个 从 给予 区域，V 后续。

这些连接是通过逐帧估计检测到的对象之间的欧氏距离形成的。也就是说，对于在当前帧上找到的每个对象，使用限制条件搜索对象在前一帧上的位置附近的坐标，如果找到任何对象，则将这些对象坐标连接成一个向量。

因此，对于每个对象，其自己的运动矢量在指定区域（即初始坐标和最终坐标）内形成。

此外，这些向量是为所有计算的 这些区域，这将最终给出的概念 穿过给定区域的对象数。



图 4。将两个位置标记为区域 曲目被计算在内。

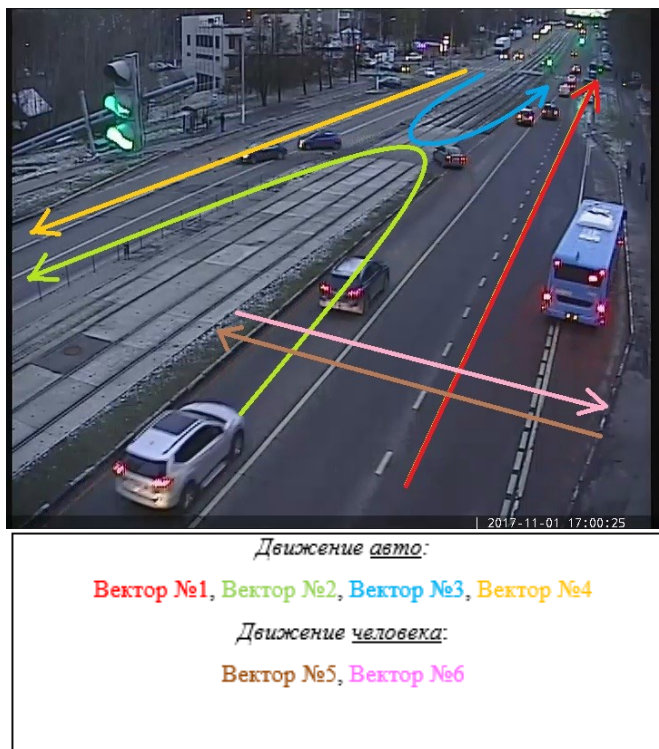
步骤#3:

图5. 物体的移动方向由于有关初始和最终区域数量及其连接的先验信息，区域之间的流量分布发生了。即构造一个方程组，其中沿给定方向运动的物体的数量特征作为未知变量。

例如，考虑汽车运动的4个方向（图5 - v1、v2、v3、v4）、2个初始区域（图4，位置编号1 - start_1、start_2）、2个结束区域（图4，位置1号 - end_1、end_2）和1个中间体（图4，位置1 - 中间）。

从所有可能的运动方向考虑一个方程组。我们得到沿每个给定方向移动的汽车数量：

$$\begin{cases} v1+v2 = start_1 \\ v1+v3 = end_1 \\ v3+v4 = start_2 \\ v2+v4 = end_2 \\ v2+v3 = middle \end{cases} \Rightarrow \begin{matrix} 1100 & start_1 \\ 1010 & end_1 \\ 0011 & start_2 \\ 0101 & end_2 \\ 0110 & middle \end{matrix} \Rightarrow \begin{matrix} 1000 & (start_1+end_1+middle)/2 \\ 0010 & (end_1-start_1+middle)/2 \\ 0001 & (start_2+end_2+middle)/2 \\ 0100 & (end_2-start_2+middle)/2 \end{matrix} \Rightarrow \begin{cases} v1 = (start_1+end_1+middle)/2 \\ v2 = (end_2-start_2+middle)/2 \\ v3 = (end_1-start_1+middle)/2 \\ v4 = (start_2+end_2+middle)/2 \end{cases}$$

如果只有一个初始区域和一个最终区域对应于移动方向（图4，位置2），则只应标记一个区域，物体将通过该区域移动。

因此，穿过该区域的对象数量将对应于沿该方向移动的对象数量。

步骤4:

在最后阶段，有一组向量描述所有检测到的物体（人、车）的移动方向。有了这些信息，就可以构建两个单独的图（对于

people and for transport），它将给定时间段内的数据结合起来，在一个轴上显示初始阶段指定的每个方向上通过的人（过往车辆）的数量特征，另一个轴将负责每个给定的方向，分别。

此外，使用第3步中的信息，您可以确定该位置的“热区”，该位置正在通过摄像机进行监控。也就是说，可以分析指定时间段内各种运动方向的数据，并确定其中最受欢迎的运动方向。



图6. 5分钟交通图

因此，汽车最流行的驾驶方向是#1和#4，而对于人来说最流行的驾驶方向是#5。

六. 乙探测器的实验评估

为了评估检测器的完整性和准确性，我们需要在测试样本 N (test_N, 其中 N 是该文本文件所属的图像的编号) 的图像实验过程中获得的文本文件以及带有属于已标记类的对象的预标记坐标 (true_N, 其中 N 是此文本文件所属的图像的编号)。

为了评估所实施算法的准确性，编写了一个 Python 脚本。

1. 脚本的本质是在一个循环中从1到N读取文本文件 test_N 和 true_N。
2. 此外，在循环中，执行 true_N 文件的行 - 所有框的坐标都转换为该框中心的坐标，并且该中心的坐标由键“类号”。这样就得到了一个字典，由原始图像中包含的所有对象框的中心坐标组成，我们的神经网络/基本算法必须确定它。
3. 接下来，您需要检查文件 test_N 的行。如果到达文件末尾，剩余的数量

坐标 V 列表 和 固定的
每个坐标列表作为缺失对象 (++)FN)。过渡到下一步。否则，以相同的方式转换坐标，并以类号为关键字，在上一步创建的坐标列表中进行搜索。如果存在这样的键，则比较坐标。如果预先知道的坐标与算法生成的坐标之间的差异小于 20 像素，则认为对象已正确定义 (++)TP)，这些坐标将从列表中删除。如果差异超过 20 个像素，或者没有这样的键，则这被视为误报 (++)FP)。随后返回到此步骤的开头。

移动物体，将根据指定的链接方向分发数据并构建流量图。

4. 对检测器操作进行了实验评估。

作为进一步发展的方向
建议方法：

1. 改变 Yolo 神经网络的配置以提高算法的速度。
2. 分段网络的使用，也将在两个类别（人员、运输）上进行训练。使用这种方法，可以准确地检测对象组（人群和汽车群）。然后就可以使用检测神经网络 Yolo。

4. 在这一步，一张图片有多个 positives，分别是：TruePositive, FalsePositive, FalseNegative，可以让你评估一张特定图片结果的准确性 (precision) 和完整性 (recall)：

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN}$$

5. 接下来是 F-measure：

$$F = 2 * \frac{Precision * Recall}{Precision + Recall}$$

6. 这个值被添加到列表的末尾，循环返回到开头。

因此，获得了 N F-measures 的列表。接下来，您应该计算平均值：

$$F_{avg} = \frac{F[0] + \dots + F[N-1]}{N}$$

该值是评估检测质量的关键值。

Идентификатор локации	Время видео (мин)	Средняя F-мера (Точность)
1	1	0.96
1	5	0.95
2	5	0.98

图7. 检测器对测试样品的评估结果。

七. Z 结论和进一步研究

在这项工作的过程中，解决了以下任务：

1. 对图像中对象的选择和分类的现有方法进行了审查。其中借鉴并改造了 Darknet Yolo 神经网络。
2. 形成并标记了一个数据集，其中包含来自摄像头的图像，上面有车辆和人员。
3. 开发并实施了一种基本算法，该算法针对指定区域的输入视频序列，将计算

乙 参考书目

- [1] Haykin S., Network N. 综合基础 // 神经网络。- 2004。- T. 2。- 没有。2004。- 第 41 章。
- [2] Krizhevsky A., Sutskever I., Hinton GE Imagenet classification with deep convolutional neural networks // 神经信息处理系统的进展。- 2012。- P.1097-1105。
- [3] Girshick R. 等人。Rich feature hierarchies for accurate object detection and semantic segmentation // IEEE 计算机视觉和模式识别会议论文集。- 2014 年。- P.580-587。
- [4] Viola P., Jones M. Rapid object detection using a boosted cascade of simple features // Computer Vision and Pattern Recognition, 2001. CVPR 2001. 2001 年 IEEE 计算机协会会议论文集。- IEEE, 2001。- T. 1。- P. II。
- [5] Viola P., Jones M. 稳健的实时对象检测 // 国际计算机视觉杂志。- 2001。- T. 4。- 没有。34-47。
- [6] Szegedy C. 等人。深入了解卷积 // IEEE 计算机视觉和模式识别会议论文集。- 2015 年 - 第 1-9 节。
- [7] Simonyan K., Zisserman A. 用于大规模图像识别的非常深的卷积网络 // arXiv 预印本 arXiv:1409.1556。- 2014 年。
- [8] Opelt A. 等。通用对象检测和识别的弱假设和提升 // 计算机视觉-ECCV 2004。- 2004。- P.71-84。
- [9] Lienhart R., Maydt J. 用于快速对象检测的一组类似 haar 的扩展功能 // 图像处理。2002. 诉讼。2002 年国际会议。- IEEE, 2002。- 第 1 卷 - P. II。
- [10] Dalal N., Triggs B. 用于人体检测的定向梯度直方图 // 计算机视觉和模式识别，2005 年。CVPR 2005. IEEE 计算机协会会议。- IEEE, 2005。- V. 1。- 第 886-893 号。
- [十一] Suard F. 等。使用定向梯度直方图的行人检测 // 智能车辆研讨会，2006 IEEE。- IEEE, 2006 年。- P.206-212。
- [12] Lee H., Center S. 使用深度学习 (CNN) 进行多对象类检测和定位。- 2015 年。
- [13] Buhler K., Lambert J., Vilim M. 视频 CS 229 课程项目中的实时对象跟踪。
- [14] Redmon J., Farhadi A. YOLO9000: 更好、更快、更强 // arXiv preprint arXiv:1612.08242。- 2016 年。
- [15] Redmon J. 等人。你只看一次：统一的实时对象检测 // IEEE 计算机视觉和模式识别会议论文集。- 2016。- P.779-788。
- [16] Girshick R. 等人。探测器。- 2018 年。
- [17] Redmon J., Farhadi A. YOLOv3: 增量改进 // arXiv preprint arXiv:1804.02767。- 2018 年。
- [18] 简历技巧。从零到英雄：使用深度学习进行目标检测指南：Faster-RCNN、YOLO、SSD。cv-tricks.com/object-detection/faster-r-cnn-yolo-ssd/ 2018
- [19] 深度学习目标检测：权威指南。tryolabs.com/blog/2017/08/30/object-detection-an-overview-in-the-age-of-deep-learning/ 2018

- [20] Girshick R. 等人。Rich feature hierarchies for accurate object detection and semantic segmentation // IEEE 计算机视觉和模式识别会议论文集。- 2014 年。- P.580-587。
- [21] Uijlings JRR 等人。Selective search for object recognition // 国际计算机视觉杂志。- 2013。- T. 104。- 没有。2。- 第 154-171 条。
- [22] 任 S. 等。Faster r-cnn: 使用区域建议网络实现实时目标检测 // 神经信息处理系统的进展。- 2015 年 - 第 91-99 页。
- [23] 一次性目标检测。machinethink.net/blog/object-detection/ 2018
- [24] 刘 W. 等。Ssd: Single shot multibox detector // 欧洲计算机视觉会议。- Springer, Cham, 2016 年。- 第 21-37 页。

使用视频评估莫斯科人口稠密地区的人流量和交通流量 从固定相机

R.沃罗比耶夫

摘要的我在本文中，任务是使用基于计算机视觉和深度神经网络的视频分析方法获取信息数据，并对其进行进一步的分析评估，以便自动计算莫斯科人口稠密地区的交通量。提出了一种解决此问题的方法，允许在选定位置计算两种不同的交通：人和汽车。神经网络学习过程需要一个数据集。通过测量检测器的质量度量来评估算法的准确性。

本文涉及确定（识别）每个单独对象（人和车辆）并进一步分析其运动的任务。它在于，对于给定的视频文件，需要确定在几个预定方向之一上移动的人数、车辆数量。将来，将分析这些数据（与时间间隔相关，有助于识别“热门目的地”、“热区”（兴趣增加的区域）），并在此基础上构建图表和热图。该任务涉及图像中物体检测领域。

关键词—对象检测、视频分析、计算机视觉、数据集创建。

R参考文献

- [1] Haykin S., Network N. 综合基础 //神经网络。- 2004。- T. 2。- 没有。2004。- 第 41 章。
- [2] Krizhevsky A., Sutskever I., Hinton GE Imagenet classification with deep convolutional neural networks //神经信息处理系统的进展。- 2012。- P.1097-1105。
- [3] Girshick R.等人。Rich feature hierarchies for accurate object detection and semantic segmentation //IEEE 计算机视觉和模式识别会议论文集。- 2014 年。- P.580-587。
- [4] Viola P., Jones M. Rapid object detection using a boosted cascade of simple features // Computer Vision and Pattern Recognition, 2001. CVPR 2001. 2001 年 IEEE 计算机协会会议论文集。- IEEE, 2001。- T. 1。- P. II。
- [5] Viola P., Jones M. 稳健的实时对象检测 // 国际计算机视觉杂志。- 2001。- T. 4。- 没有。34-47。
- [6] Szegedy C. 等人。深入了解卷积 //IEEE 计算机视觉和模式识别会议论文集。- 2015 年 - 第 1-9 节。
- [7] Simonyan K., Zisserman A. 用于大规模图像识别的非常深的卷积网络 //arXiv 预印本 arXiv:1409.1556。- 2014 年。
- [8] Opelt A. 等人。通用对象检测和识别的弱假设和提升//计算机视觉-ECCV 2004。- 2004。- P.71-84。
- [9] Lienhart R., Maydt J. 用于快速对象检测的一组类似 haar 的扩展功能//图像处理。2002.诉讼。2002年国际会议。- IEEE, 2002。- 第 1 卷 - P. II
- [10] Dalal N., Triggs B. Histograms of oriented gradients for human detection // Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on。- IEEE, 2005。- V. 1。- 第 886-893 号。
- [11] Suard F. 等人。使用定向梯度直方图的行人检测 // 智能车辆研讨会, 2006 IEEE。- IEEE, 2006 年。- P.206-212。
- [12] Lee H., Center S. 使用深度学习 (CNN) 进行多对象类检测和定位。- 2015 年。
- [13] Buhler K., Lambert J., Vilim M. 视频 CS 229 课程项目中的实时对象跟踪。
- [14] Redmon J., Farhadi A. YOLO9000: 更好、更快、更强 //arXiv 预印本 arXiv:1612.08242。- 2016 年。
- [15] Redmon J. 等人。你只看一次: 统一的实时对象检测//IEEE 计算机视觉和模式识别会议论文集。- 2016。- P.779-788。
- [16] Girshick R.等人。探测器。- 2018 年。
- [17] Redmon J., Farhadi A. YOLOv3: 增量改进 //arXiv preprint arXiv:1804.02767。- 2018 年。
- [18] 简历技巧。从零到英雄: 使用深度学习进行对象检测指南: Faster R-CNN, YOLO, SSD. cv-tricks.com/objectdetection/faster-r-cnn-yolo-ssd/ 2018
- [19] 深度学习目标检测: 权威指南. tryolabs.com/blog/2017/08/30/object-detection-an-overview-in-the-age-of-deep-learning/ 2018
- [20] Girshick R. 等人。Rich feature hierarchies for accurate object detection and semantic segmentation //IEEE 计算机视觉和模式识别会议论文集。- 2014 年。- P.580-587。
- [21] Uijlings JRR 等人。Selective search for object recognition//国际计算机视觉杂志。- 2013。- T. 104。- 没有。2。- 第 154-171 条。
- [22] 任 S.等。Faster r-cnn: 使用区域建议网络实现实时目标检测//神经信息处理系统的进展。- 2015 年 - 第 91-99 页。
- [23] 一次性目标检测. machinethink.net/blog/object-detection/ 2018
- [24] 刘W.等。Ssd: Single shot multibox detector //欧洲计算机视觉会议。- Springer, Cham, 2016 年。- 第 21-37 页。