

文章编号: 2095-2163(2022)11-0117-05

中图分类号: TP391.41

文献标志码: A

基于机器视觉的运动目标识别与定位

赵佳冉¹, 张爱华¹, 姜超群¹, 蒋群², 李建军¹

(1 中原工学院 电子信息学院, 郑州 450007; 2 新郑市恒凯电子科技有限公司, 郑州 450007)

摘要: 随着机器视觉技术的持续发展,分拣技术在工业生产中得到了广泛应用。本文提出了一种基于机器视觉的目标识别与定位装置的设计方案,通过识别运动状态物体的形状对其进行分类,同时估计运动物体轨迹,实现物体抓取。本方案中,由工业摄像机采集物体图像,利用全局阈值图像分割大津法(Otsu)进行阈值分割,轮廓逼近确定物体的形状,实现物体形状分类。采用灰度重心法确定物体中心像素坐标,然后利用旋转矩阵和偏移向量之间的矩阵关系将物体中心像素坐标转换为相机坐标,通过刚体变换把相机坐标转换为传送带坐标,获取运动物体坐标信息,实现目标定位。实验结果表明,智能分拣系统对不同运动目标物体抓取率达90%以上,分拣效果较好。

关键词: 机器视觉; 目标识别; 灰度重心法; 目标定位; 物体分拣

Object recognition and positioning based on machine vision

ZHAO Jiaran¹, ZHANG Aihua¹, JIANG Chaoqun¹, JIANG Qun², LI Jianjun¹

(1 School of Electronic and Information, Zhongyuan University of Technology, Zhengzhou 450007, China;

2 Xinzheng Hengkai Electronic Technology Co., Ltd., Zhengzhou 450007, China)

[Abstract] With the continuous development of machine vision technology, sorting technology has been widely used in industrial production. This paper introduces a design scheme of a target recognition and positioning device based on machine vision, which can classify objects by recognizing the shape of moving objects, at the same time estimate the trajectory of moving objects to achieve object grasping. In this scheme, the image of the object is collected by an industrial camera, the global threshold image segmentation Otsu method is used for threshold segmentation, and the shape of the object is determined by contour approximation to realize the shape classification of the object. The gray-scale barycenter method is used to determine the center pixel coordinates of the object, and then the center pixel coordinates of the object are converted into camera coordinates by using the matrix relationship between the rotation matrix and the offset vector, and the camera coordinates are converted into conveyor belt coordinates through rigid body transformation to obtain the coordinate information of the moving object, to achieve target positioning. The experimental results show that the intelligent sorting system has a grasping rate of more than 90% for different moving objects, and the sorting effect is good.

[Key words] machine vision; target recognition; gray barycentric method; target positioning; object sorting

0 引言

机器视觉技术和数字图像理论的快速发展,使得目标识别与分拣技术日渐成熟,且性能持续提升。在工业生产领域应用机器视觉技术^[1]和图像识别原理实现工业智能分拣。该类方法具有较强的灵活性和适应能力,使工业生产速度得到很大提升,同时减少了人力资源的消耗^[2-6]。

张大为等人^[3]基于机器视觉设计了可实现不同运行速度下的物料抓取的机器人自动上料系统。徐青青^[4]利用智能分拣系统能够实现不同物体识别分类并获取位姿,但均存在算法不易移植的问题。彭杰等人^[5]对不同位姿下的铝件和有机玻璃通过

条形码进行分拣,但应用场景受限。王诗宇等人^[6]利用并联机械手机器人和康耐视相机搭建智能分拣系统,克服光照因素影响的同时解决因为环境改变程序不适应问题,成本较高。

针对以上问题,设计了基于机器人视觉的运动物体形状识别与定位方案,该方案可应用于实验教学,所以成本较低、应用场景广泛且算法易于移植。

1 目标物体识别分拣系统

物体识别分拣系统主要由传送模块、机器视觉模块、运动控制模块和抓取模块四部分组成。对此可做流程阐述如下。

传送模块负责运动目标物体的运动数据;机器

基金项目: 河南省高等学校重点科研项目计划; 基于稀疏毫米波大规模 MIMO 信道信息技术研究(21A510015)。

作者简介: 赵佳冉(1996-),女,硕士研究生,主要研究方向:机器视觉、机器人;张爱华(1976-),女,博士,教授,主要研究方向:大规模 MIMO 传输技术、稀疏信道估计、通信信号处理。

通讯作者: 张爱华 Email: zhah@zut.edu.cn

收稿日期: 2022-03-19

视觉模块对采集到的图像做预处理,物体轮廓提取判断物体形状、获取物体重心坐标,为实现物体分拣提供必要信息;运动控制模块在调控X轴、Y轴移动的同时,控制Z轴运动到目标物体坐标进行吸取;抓取模块负责物体的吸放操作。

当物体进入检测区,触发光电传感器,相机启动,上位机界面计时开始,得到物体到达抓取点时间,再一并返回物体此时坐标值。桁架轴此时开始运动到抓取位置,完成目标抓取。系统工作流程如图1所示。

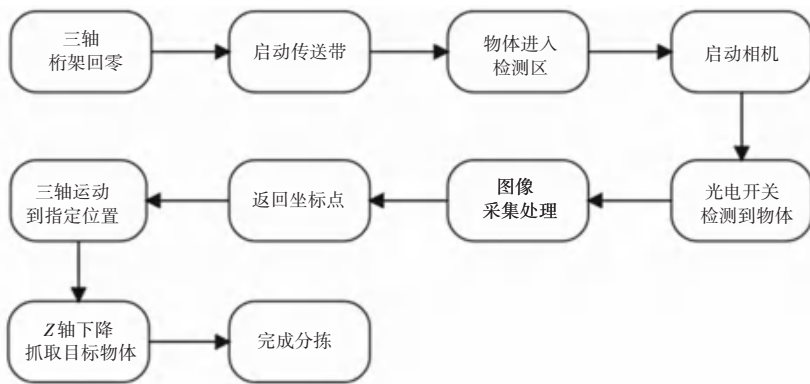


图1 系统工作流程

Fig. 1 System workflow

2 相机标定

由于相机镜头透镜的存在,使得光线投影到成像平面的过程中会产生畸变,此时就需要对相机进行标定。通过标定,获取相机的内参矩阵和外参矩阵,得到相机坐标系在世界坐标中的位置信息。相机坐标系与世界坐标系的关系^[7]可表示为:

$$\begin{pmatrix} \hat{e}_c^X \\ \hat{e}_c^Y \\ \hat{e}_c^Z \\ \hat{e}_c^1 \end{pmatrix} = \begin{pmatrix} \hat{e}_w^X \\ \hat{e}_w^Y \\ \hat{e}_w^Z \\ \hat{e}_w^1 \end{pmatrix} \begin{matrix} \hat{R} \\ T \\ \mathbf{1} \\ \hat{0}^T \end{matrix} \quad (1)$$

其中, \mathbf{R} 表示该相机的旋转矩阵; \mathbf{T} 表示偏移向量;矩阵 $[X_w \ Y_w \ Z_w \ 1]^T$ 表示空间点在世界坐标系下的向量; $[X_c \ Y_c \ Z_c \ 1]^T$ 表示空间点在相机坐标系下的向量。

将相机坐标系 (X_c, Y_c, Z_c) 转换至理想投影平面坐标系 (X_u, Y_u) ,可表示为:

$$\begin{cases} X_u = f \frac{X_c}{Y_c} \\ Y_u = f \frac{X_c}{Z_c} \end{cases} \quad (2)$$

其中, f 为摄像头焦距。由理想投影平面坐标系 (x_u, y_u) 到实际的投影平面坐标系 (x, y) 之间的转换关系为:

$$\begin{pmatrix} \hat{e}_u^x \\ \hat{e}_u^y \\ \hat{e}_u^1 \end{pmatrix} = (1 + k_1 r^2 + k_2 r^4) \begin{pmatrix} \hat{e}_w^x \\ \hat{e}_w^y \\ \hat{e}_w^1 \end{pmatrix} + \begin{pmatrix} 2p_1 x_u y_u (r^2 + 2x_u^2) \\ p_1 (r^2 + 2y_u^2) + 2p_2 x_u y_u \\ r = x_u^2 + y_u^2 \end{pmatrix} \quad (3)$$

其中, k_1, k_2 为径向畸变系数; p_1, p_2 为切向畸变系数; r 为点距成像中心的距离。

实际投影平面坐标 (x, y) 和相机图像坐标 (u, v) 的线性转换关系可分别表示为:

$$\begin{cases} u = f_x x + u_0 \\ v = f_y y + v_0 \end{cases} \quad (4)$$

$$\begin{cases} f_x = f/\Delta x \\ f_y = f/\Delta y \end{cases} \quad (5)$$

其中, (u_0, v_0) 为基准坐标,表示图像的中心坐标; Δx 和 Δy 分别表示坐标轴中水平和垂直方向相邻的2个像素之间的距离(mm/pixel); f_x 和 f_y 表示单个像素的焦距。结合式(1)、式(4)~(5)可得到具体的表达式为:

$$s \hat{e}_u = \hat{e}_0 \begin{pmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{pmatrix} [\mathbf{R} \ \mathbf{T}] \begin{pmatrix} \hat{e}_w^X \\ \hat{e}_w^Y \\ \hat{e}_w^Z \\ \hat{e}_w^1 \end{pmatrix} \quad (6)$$

其中, s 表示比例因子。

由此就可以得到图像中像素坐标系和世界坐标

系的对应关系式。

3 图像处理

3.1 图像预处理

图像预处理是物体形状识别中不可或缺的一步,实际图像采集时,外部环境干扰较大,为了消除无关信息的影响,需要对图像进行预处理。本文采用高斯滤波方法对图像进行滤波。利用二进制阈值分割方法中的 Otsu 算法^[8]将物体和背景分离,增强物体形状识别可靠性。

Otsu 算法在实际图像分割中是最佳的阈值选取算法,该算法计算简单且不受图像亮度和对比度的影响,能将物体和背景更好地分离出来,使后期的物体形状识别更加准确。

其它类型的阈值分割算法是用户自定义阈值,而 Otsu 阈值类型通过大津算法,自动获取合适阈值进行 Binary 二值化,所以该方法更适合应用在工程实践中。图 2 是利用 Otsu 阈值分割方法对目标物体分割后的图,可以看出该方法可以很好地分离出物体和背景。

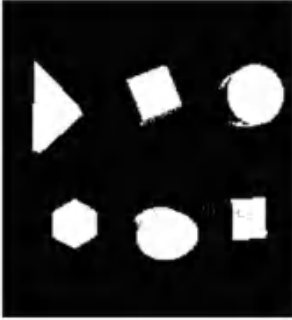


图 2 Otsu 阈值分割

Fig. 2 Otsu threshold segmentation

Otsu 阈值类型通过大津算法,自动获取合适的阈值进行 Binary 二值化^[9]。Otsu 算法原理为:令 $\{0, 1, 2, \dots, L-1\}$ 表示一幅大小为 $M \times N$ 像素的数字图像中的 L 个不同的灰度级,用 n_i 表示灰度级为 i 的像素数。图像中的像素总数为 $MN = \sum_{i=0}^{L-1} n_i$, 归一化直方图具有分量 $p_i = n_i/MN$, 选择一个阈值 T , 并用其把输入图像阈值化处理为 2 类: C_0 和 C_1 。这里, C_0 由图像中灰度值在 $[0, T]$ 范围内的所有像素组成, C_1 由灰度值在 $[T+1, L-1]$ 范围内所有像素组成。由 $P_0(T)$ 、 $P_1(T)$ 、 $\mu_0(T)$ 、 $\mu_1(T)$ 和 $\sigma_0^2(T)$ 、 $\sigma_1^2(T)$ 分别表示 2 类的累计概率、平均灰度和方差,各值间的关系列出公式见如下:

$$P_0(T) = \sum_{i=0}^T p_i \quad (7)$$

$$P_1(T) = \sum_{i=T+1}^{L-1} p_i = 1 - P_0(T) \quad (8)$$

$$\mu_0(T) = \frac{1}{P_0(T)} \sum_{i=0}^T ip_i \quad (9)$$

$$\mu_1(T) = \frac{1}{P_1(T)} \sum_{i=T+1}^{L-1} ip_i \quad (10)$$

$$\sigma_0^2(T) = \frac{1}{P_0(T)} \sum_{i=0}^T (i - \mu_0(T))^2 p_i \quad (11)$$

$$\sigma_1^2(T) = \frac{1}{P_1(T)} \sum_{i=T+1}^{L-1} (i - \mu_1(T))^2 p_i \quad (12)$$

令 μ 表示全局均值, $\sigma_b^2(T)$ 、 $\sigma_w^2(T)$ 分别表示类间方差及类内方差,可得:

$$\mu = \sum_{i=0}^{L-1} ip_i = P_0(T)\mu_0(T) + P_1(T)\mu_1(T) \quad (13)$$

$$\sigma_b^2(T) = P_0(T)(\mu_0(T) - \mu)^2 + P_1(T)(\mu_1(T) - \mu)^2 \quad (14)$$

$$\sigma_w^2(T) = P_0(T)\sigma_0^2(T) + P_1(T)\sigma_1^2(T) \quad (15)$$

由此,推得最佳阈值为:

$$T^* = \arg \max_{0 \leq T < L-1} \sigma_b^2(T) \quad (16)$$

3.2 形状识别

图像特征提取采用了多边形逼近的方法,获取目标基本轮廓。多边形逼近主要应用 *approxPolyDP* 多边形拟合函数识别物体形状,该函数采用道格拉斯-普客算法 (Douglas-Peucker) 实现拟合函数的运算^[10]。

算法应用步骤为:首先,做目标曲线的一条弦,假设目标曲线的首点为 A ,尾点为 B ,首尾相接得到一条直线 AB ,将该直线视为目标曲线的弦。然后寻找目标曲线与弦间的最大距离点 C ,使曲线和弦之间的最大距离为 d ,将 d 与预先设置好的阈值进行比较,若小于指定阈值,则这一段直线可近似为目标曲线,此段曲线处理完毕。如果距离 d 大于阈值,则用点 C 将线段分为 AC 和 BC 两段,继续上述操作,直到所有曲线处理结束。在本文中通过分别定义输出的多边形点集,将其作为多边形的边长,由此确定多边形的形状,形状示意如图 3 所示。

3.3 运动目标定位

利用灰度重心法计算物体的重心^[11]来定位运动目标位置。其重心坐标通过求图像的矩的方法得到,如果将图像的像素记为密度函数 $f(x, y)$,对像素点求期望即为图像的原点矩。原点矩的公式为:

$$m_{pq} = \sum_{x=1}^M \sum_{y=1}^N x^p y^q f(x, y) \quad (17)$$

其中, m_{pq} 表示图像中所有像素的总和,每个像

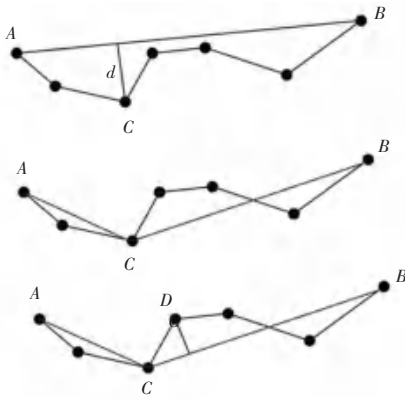


图3 Douglas-Peucker 逼近曲线示意图

Fig. 3 Schematic diagram of Douglas - Peucker approximation curve

素的 x, y 都要和相应的比例因子 x^p, y^q 相乘。

由于图形的一阶矩是与图形的形状有关的, 因此在本方案中采用一阶矩的方法求取物体的重心, 重心的求取公式为:

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (18)$$

其中, m_{00} 表示图像上所有非零区域, 处理轮廓的长度为:

$$m_{00} = \sum_{x=-r, y=-r}^r x^0 y^0 f(x, y) = \sum_{x=-r, y=-r}^r f(x, y) \quad (19)$$

m_{10} 和 m_{01} 表示轮廓在 x, y 方向上的累加, 累加过程为:

$$m_{10} = \sum_{x=-r, y=-r}^r x^1 y^0 f(x, y) = \sum_{x=-r, y=-r}^r x * f(x, y) \quad (20)$$

$$m_{01} = \sum_{x=-r, y=-r}^r x^0 y^1 f(x, y) = \sum_{x=-r, y=-r}^r y * f(x, y) \quad (21)$$

三角形重心位置如图4所示, 可通过上述一阶矩的运算求解得到。采用相似方法可求解其它图形的重心坐标。

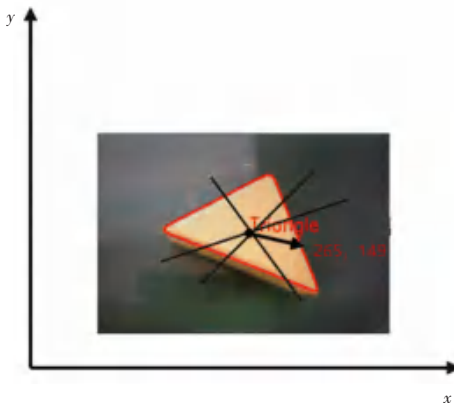


图4 三角形重心示意图

Fig. 4 Center of gravity of the triangle

4 系统实验验证

为了验证算法的有效性, 通过搭建的智能分拣装置对运动目标进行抓取实验。

当获取运动物体的传送带位置信息后, 对运动控制模块发送信号, 驱动三轴桁架机械臂运动到指定位置对物体进行吸取。

上位机界面如图5所示, 左框为摄像机采集到的原图, 右框显示结果为物体形状和此时重心位置坐标, 同时可通过速度窗口设置传送带传送速度。抓取实物图如图6所示。



图5 系统上位机界面

Fig. 5 Upper computer interface of the system

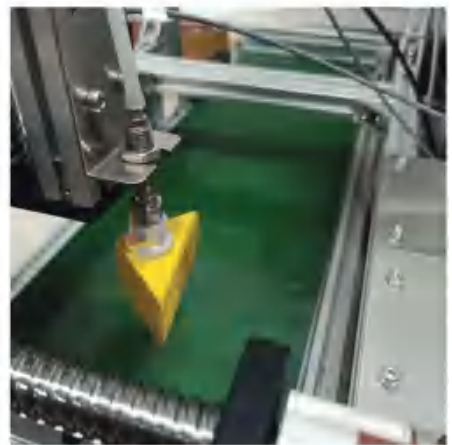


图6 目标物体抓取图

Fig. 6 Grasping diagram of target objects

通过程序调试进行多次实验, 物体的识别成功率在速度为 20 mm/s 时见表1, 30 mm/s 时见表2。

表1 速度为 20 mm/s 识别成功率

Tab. 1 Recognition success rate at a speed of 20 mm/s

	黄四边形 积木	橙方形 积木	粉圆形 积木	绿椭圆 积木	黄三角 积木
抓取成功率/%	100	97	100	100	93

表 2 速度为 30 mm/s 识别成功率

Tab. 2 Recognition success rate at a speed of 30 mm/s

	黄四边形 积木	橙方形 积木	粉圆形 积木	绿椭圆 积木	黄三角 积木
抓取成功率/%	100	96	100	97	90

从表 1、表 2 可以看出,不同形状物体抓取成功率不同,随着传送速度的提升,物体抓取成功率开始下降。

5 结束语

本文利用 .NET Framework 作为实验装置的开发平台,在 VS2015 环境下,使用 C#语言作为程序开发语言,采用 EmguCV 中的算法库对运动物体进行识别。利用 Otsu 阈值分割算法自动获取合适阈值,结合轮廓逼近方法获取目标物体形状,实现目标物体识别。通过摄像机坐标和传送带坐标之间的转换关系,确定运动物体重心坐标的运动轨迹,实现目标物体定位,最终实现分拣。

本系统主要是实验室演示装置,实验中因为对光源处进行遮光处理,排除外界光照影响,在实际工业应用时可能会出现因受到外界光照影响导致物体识别错误的情况,所以在实际应用中由于光照产生

(上接第 116 页)

4 结束语

本文分析讨论了 CBAM 注意力模块,通过实验选择出扩张率最合适的扩张解码,在此基础上,提出了基于注意力机制的改进方案,进行实验,满足了当下在实际环境中道路场景的分割准确率和实时性的需求,证明改进后网络的有效性和可行性。但仍有改进的空间,未来会在轻量化网络的改进上继续研究。

参考文献

- [1] CHENG Guo, ZHENG Jiangyu, KILICARSLAN M. Semantic segmentation of road profiles for efficient sensing in autonomous driving[C]// 2019 IEEE Intelligent Vehicles Symposium (IV). Paris, France; IEEE, 2019: 564-569.
- [2] LI Hanchao, XIONG Pengfei, AN Jie, et al. Pyramid attention network for semantic segmentation[J]. arXiv preprint arXiv: 1805.10180, 2018.
- [3] HUANG G, LIU Z, Van MAATEN D L, et al. Densely connected convolutional networks [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ; IEEE, 2017: 4700-4708.
- [4] LONG J, SHELHAMER E, DARRELL T. Fully Convolution networks for semantic segmentation[C]// Proceedings of IEEE

物体误检仍亟待进一步深入研究。

参考文献

- [1] 张冬梅,武杰,李丕丁. 基于机器视觉的运动目标检测算法综述[J]. 智能计算机与应用, 2020, 10(03): 192-195, 201.
- [2] 景卓,陈超波,曹凯,等. 基于机器视觉的物体识别分拣装置研究[J]. 计算机与数字工程, 2019, 47(03): 657-661.
- [3] 张大为,沈勇. 基于机器视觉的机器人自动上料系统设计[J]. 单片机与嵌入式系统应用, 2021, 21(03): 55-58.
- [4] 徐青青. 基于机器视觉的工业机器人智能分拣系统设计[J]. 仪表技术与传感器, 2019(08): 92-95, 100.
- [5] 彭杰,孟祥印,李晟尧,等. 基于机器视觉的工件分拣及上下料系统[J]. 机床与液压, 2021, 49(21): 38-42.
- [6] 王诗宇,林浒,孙一兰,等. 基于机器视觉的机器人分拣系统的设计与实现[J]. 组合机床与自动化加工技术, 2017(03): 125-129, 133.
- [7] 高智伟,谭晓东,刘客. 基于双目视觉的物体识别定位与抓取[J]. 科学技术与工程, 2020, 20(20): 8285-8291.
- [8] OTSU N. A threshold selection method from gray-level histograms [J]. IEEE Transactions on Systems Man & Cybernetics, 2007, 9(1): 62-66.
- [9] 陈思. 基于 Otsu 算法的车牌图像二值化及其 Matlab 实现[J]. 长春师范学院学报, 2012, 31(03): 33-35.
- [10] 李帅,方源敏,喜文飞. 基于无拓扑丈量曲线的快速压缩算法[J]. 科学技术与工程, 2011, 11(18): 4324-4327.
- [11] 任守峰,唐新明,祝小勇,等. 一种高分七号卫星激光足印光斑质心的高精度提取算法[J]. 光学学报, 2021, 41(10): 125-132.
- [12] International Conference on Computer Vision and Pattern Recognition (CVPR). Boston; IEEE Press, 2015: 3431-3440.
- [13] HYEONWOO N, HONG S, HAN B. Learning Deconvolution Network for Semantic Segmentation [C]// 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile; IEEE, 2015: 1520-1528.
- [14] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(12): 2481-2495.
- [15] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks [J]. arXiv preprint arXiv: 1609.02907, 2016.
- [16] FU Jun, LIU Jing, TIAN Haijie, et al. Dual attention network for scene segmentation [C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA; IEEE, 2019: 3141-3149.
- [17] POUDEL R P, LIWICHI S, CIPOLLA R. Fast-SCNN: Fast semantic segmentation Network [J]. arXiv preprint arXiv: 1902.04502, 2019.
- [18] SZEGEDY C, LIU Wei, JIA Yangqing, et al. Going deeper with convolution [C]// Computer Vision and Pattern Recognition. Boston, MA; IEEE, 2015: 1-9.
- [19] WANG H R, FAN Y, WANG Z X, et al. Parameter-free spatial attention network for person Re-identification [EB/OL]. [2021-06-21]. <https://arxiv.org/abs/1811.12150>.