迈向个人防护装备检测的边缘-云端架构

Jarosław Legierski
jaroslaw.legierski@orange.com Orange Innovation，Orange Polska S.A. 波兰华沙

Kajetan Rachwał Piotr Sowinski
kajetan.rachwal@ibspan.waw.pl
piotr.sowinski@ibspan.waw.pl
波兰科学院系统研究所 波兰华沙理工大学

Wojciech Niewolski
wojciech.niewolski@orange.com Orange Innovation，Orange Polska S.A. 波兰华沙理工大学

Przemysław Ratuszek Zbigniew Kopertowski
przemyslaw.ratuszek@orange.com
zbigniew.kopertowski@orange.com
Orange Innovation，Orange Polska S.A. 波兰华沙

Marcin Paprzycki Maria Ganzha
marcin.paprzycki@ibspan.waw.pl
maria.ganzha@ibspan.waw.pl 波兰科学院系统研究所 波兰华沙

arXiv:2301.01501v1 [cs.CV] 4 Jan 2023

摘要

在图像和视频流中检测个人防护装备，是保障施工工人安全的一项重要工作。本文提出了一种架构，可对这类装备进行实时图像识别。该解决方案可部署于边缘-云端和仅边缘两种环境。作为更大场景的一部分，该系统在一个正在施工的现场进行了测试，测试属于 ASSIST-IoT H2020 项目的范围。为确定仅边缘变体的可行性，我们采用 YOLOX 方法开发了一个用于统计佩戴安全头盔人员数量的模型。结果发现，鉴于现场可用的硬件基础设施，此用例可以采用纯边缘部署。在初步评估中，我们得出了若干重要观察结果，这些结果对于系统的后续开发和部署至关重要。未来的工作将深入研究两种架构变体的性能表现。

关键词：边缘-云连续体架构、个人防护装备检测、图像识别、工人安全

1 引言

如今，智能视频分析在广泛的应用领域中日益受到重视 [17]。此类系统的核心通常是图像识别（IR）组件。然而，迄今为止，IR 子系统最常见的部署方式是在云端。这种方式具有多项优势，例如能够使用大量且可扩展的计算资源、获得可靠的应用程序接口，并将系统维护的负担转移给云服务提供商。不过，这也需要付出代价。将数据发送到云端会引发安全性和隐私方面的担忧。此外，与云端通信必然会产生网络延迟，在时间要求严格的应用中，这种延迟可能相当显著。为解决以云端为中心的方案带来的问题，人们提出了边缘计算。该方法的核心是尽可能靠近数据源处理数据。这样既能降低延迟，也有助于确保数据的安全性和隐私，因为数据会保留在本地网络内。不过，边缘计算也有自身的一系列问题。通常，边缘端可用的计算资源要少得多。结合这两种方案的优势、弥补其不足的一种可行方法是采用边缘-云端连续体：数据的一部分在边缘端处理，另一部分则在云端处理。然而，这也引出了一个显而易见的问题：系统的各个部分应部署在这一连续体中的哪些位置?

在此，我们以监控正在施工的施工现场入口这一真实场景来探讨这个问题。具体而言，系统需要确保：（1）未经授权的人员不得进入工地；（2）所有人都佩戴适当的个人防护装备（PPE），即安全帽和安全背心。该场景作为 ASSIST-IoT 项目的一部分进行评估，地点位于波兰华沙，由建筑公司 Mostostal Warszawa 管理的施工现场。在此，边缘端与云端的取舍尤为重要。一方面，工人的隐私至关重要，同时必须尽可能降低延迟，以确保快速响应，这表明采用

边缘部署。另一方面，考虑到边缘端可用的硬件资源有限，而且施工现场的环境极其恶劣，采用云端部署似乎颇具吸引力。

鉴于这两种方案各自可能带来的优势，本文提出了一种面向边缘-云连续体的视频分析架构解决方案。该架构可部署为两种变体（仅边缘和边缘-云端），具体描述见架构部分。此外，为了确定该方案的可行性，研究开展了一项初步实验。研究人员开发了一个 IR 模型，并将其集成到该架构的仅边缘变体中。随后，该模型被用于检测佩戴个人防护装备的人员何时进入和离开工作现场。

2 背景

为本研究提供背景，本文概述了（1）IR 系统架构以及（2）用于个人防护装备检测的机器学习模型的最新进展。

系统架构。在边缘端部署 IR 系统最显而易见的优势是延迟更低。这一点已在 [20], 中得到证实，研究人员将人脸识别模型部署在边缘端。作者发现，与部署在云端相比，将模型部署在边缘端能显著提高响应速度。在其他研究中， [8, 9], 研究人员探讨了在仅边缘场景中部署深度卷积神经网络（CNN）的可行性。卷积神经网络具有较高的资源利用率，因此通常部署在云端。研究发现，当部分计算任务可以卸载到其他边缘设备时，在移动设备上部署卷积神经网络是可行的。边缘部署使系统在使用卷积神经网络进行实时增强现实目标跟踪时，能够持续实现 2.24 ms 的低延迟 [8]。两项研究均表明，通过边缘部署分配工作负载可以增强系统的推理能力，因为仅靠未连接的移动设备无法运行这些模型。

一项研究 [5] 探讨了一种边缘-云端架构，其中数据预处理服务器部署在靠近数据源的位置。预处理后的数据随后被发送到基于云端的深度学习平台。这降低了网络延迟和流量，同时提高了原始数据的安全性和隐私保护。

另一方面，边缘部署在可用硬件方面受到更多限制。计算资源有限，自然会限制模型规模和推理速度。一项研究比较了同一视频处理模型的不同实现方式（基于 TensorFlow、TensorRT 和 TFLite） [6], , 发现它们的资源利用率各不相同。实现方式的选择会影响模型的能耗及其推理速度。有趣的是，速度最慢的实现方式（TFLite）能效最高。研究还发现，TFLite 在处理低分辨率视频时，其速度能够与其他实现方式持平。对于高分辨率视频，则需要资源消耗更大的模型才能维持速度，这表明在资源有限的环境中，云端部署可能更有优势。尽管如此，只要边缘端有足够的可用资源，某些资源消耗较大的模型仍可部署在边缘端。另一项研究提出的部署方案要求所有节点都配备 GPU [9], 从而使作者能够在边缘端使用卷积神经网络。在 [13], 中，研究人员将 IR 模型部署在配备摄像机的 Raspberry Pi 4B 和 Intel Neural Compute Stick 2（一种用于边缘端深度学习推理的 USB 设备）上进行了研究。选择这些设备是因为它们功耗低且计算能力出色。总体而言，一个用于检测安全帽和安全背心等个人防护装备（PPE）的模型，准确率达到了 99.5% 左右。

PPE检测模型。 鉴于工人佩戴防护头盔对健康与安全至关重要，基于视频分析、能够有效检测防护头盔是否存在的方法，目前是一个热门研究课题。

近期一项研究 [19]证明，使用现有且未经修改的机器学习模型检测防护头部覆盖物，无法达到足够的检测准确率。该文比较了流行的 YOLO 算法 [1] 的多个版本。研究表明，用于头盔检测时，效果最好的 YOLO 版本是 v4。通过改进损失函数，该版本在测试中的准确率超过了 93%。另一项类似研究 [10]致力于改进 YOLOv5 算法。得益于神经网络结构的改进，该系统的准确率接近 97%。另一项研究 [18], 也探讨了如何改进 YOLOv5。不过，该研究并未改进算法本身，而是通过对输入图像应用滤波器来处理输入数据，从而将准确率提高到 95 % 以上。还有一项研究 [15]提出了一种基于 ConCaNet 注意力机制设计多层金字塔特征融合网络的方法，以提高检测速度和准确率。该研究采用 YOLOv3，并使用了一个包含 6000 张图像的数据集。结果表明，这种方法行之有效，能够减少所需的参数数量。

也可以使用 SSD-MobileNet 算法进行头盔检测 [4], , 该算法基于 CNN 的另一种变体。 [7], 中报告了对该方法的分析，测试了其有效性，测试期间的准确率达到 80%。在对多种算法类型进行更广泛的比较后， [11], 作者提出了一种基于动态变化神经网络的头盔检测方法——SHDDM（安全帽检测动态模型）。该模型会分析人体姿态，并确定安全帽应位于的区域，从而避免在头部区域之外检测到安全帽，进而降低假阳性率。还有

其他头盔检测方法，例如利用颜色和形状定位面部并判断安全帽是否正确佩戴的方法 [16]。另一种方案使用从视频流中获取的低分辨率图像，结合局部二值模式（LBP）、灰度共生矩阵（GLCM）和反向传播神经网络 [14]。

另一项研究 [2] 探讨了人工生成的图像在训练用于个人防护装备检测的卷积神经网络时的效用。论文介绍了使用 YOLOv3 所取得的结果；该模型在由电子游戏《侠盗猎车手V》中的 Rockstar 高级游戏引擎（RAGE）生成的人工图像上进行训练。在由真实场景图像组成的测试数据集上，该方法的平均精度均值（mAP）仅为 55.11%。而在合成图像上的 mAP 则高得多，达到 87.24%。值得注意的是，真实场景图像上的结果不佳，很可能是因为 RAGE 引擎无法生成足量的头部、焊接面罩、耳部防护装备及胸部物体变体。

由此可见，检测防护头盔的方法有很多。值得特别关注的是 SHDDM，因为它有一项重要功能：能够检查头盔是否佩戴正确，而不仅仅是检测是否存在头盔。这一点在真实场景的应用中尤为重要。

3 所提架构

所提出的视频分析系统可采用两种架构变体部署：边缘-云端（图1）和仅边缘（图 2）。如上所述，有理由认为这两种变体都适用于所考虑的场景。两种架构都包含部署在边缘端的共同核心，由摄像机、图像处理器（IP）组件以及职业安全与健康（OSH）经理的移动设备组成。

摄像机（在所报告的实验中使用的是大华 IPC-HFW5449T-ASE-LED）提供实时 RTSP 视频流，并将其传输至图像处理器。IP 是一项用 Python 编写的服务，可选择执行初步图像分析。通过运动检测和亮度阈值处理等可配置方法，IP 能够丢弃不包含移动人员的图像帧，从而减少参与实际图像分析的组件所需处理的网络流量。它还负责与 ASSIST-IoT 部署进行通信

系统的其余部分按照 ASSIST-IoT 参考架构 [3] 进行设计。IP 通过向 MQTT 主题发布警报，与系统的其余部分通信。此设计使 ASSIST-IoT 部署中的其他组件和设备能够以流式方式获知任何职业安全与健康违规行为，例如工人未佩戴防护头盔。

在架构的第一个版本——边缘-云端部署中，IP 配置为使用基于云端的 AWS Rekognition 平台及其个人防护装备检测服务。

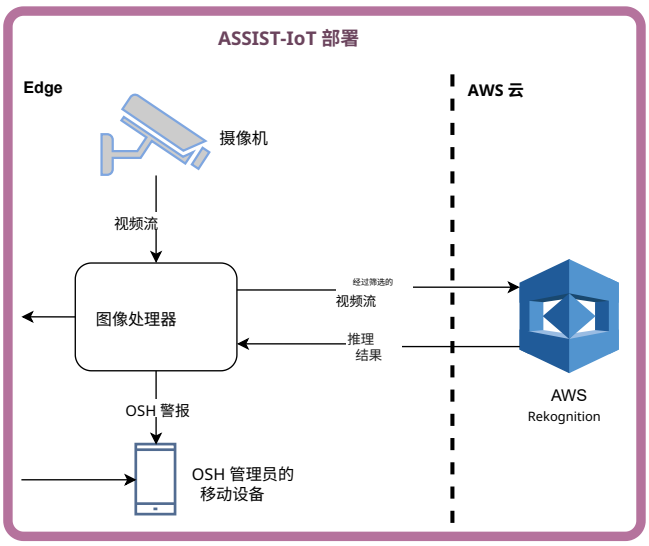


图 1. 边缘-云端部署

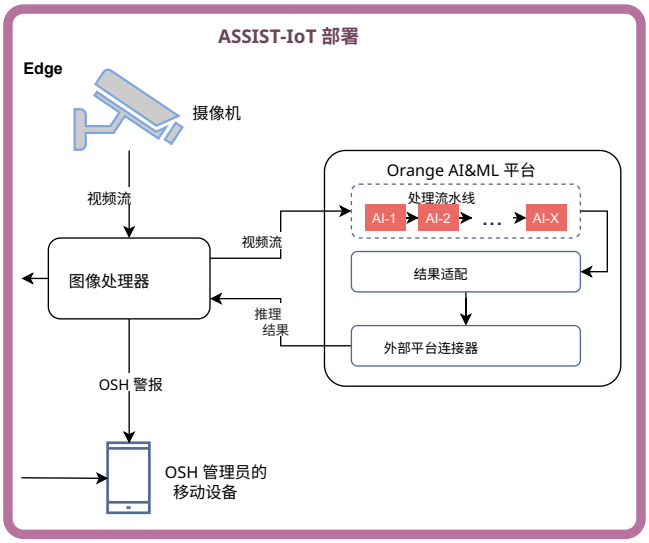


图 2. 边缘部署

在仅边缘变体中，视频分析由部署在施工现场服务器上的 Orange AI&ML 平台执行。这种边缘部署有助于降低网络延迟，并确保工人数据的隐私。AI&ML 平台的服务以可运行的 Python 模块编写，各自提供应用程序接口和图形用户界面。这些服务可以复用平台提供的应用程序接口和图形用户界面，也可以从头构建。服务会从视频源采集视频帧，在该服务专用的机器学习流水线中进行处理，并对结果进行调整或解读。平台生成的推理结果会通过提供的连接器转发给外部服务。由于 Orange AI&ML 平台在边缘端运行，所有视频

处理都在客户现场进行，从而确保客户数据（视频）的安全，并符合通用数据保护条例等相关法规。

4 方法

作为本研究的一部分，该架构的仅边缘变体初版已部署在一个正在施工的工地。利用 Orange AI&ML 平台，训练了一个用于统计进出特定区域且佩戴安全帽人员数量的模型。系统在定义的识别区域（边界框）内，对穿过图中可见的黄色线和绿色线的安全帽佩戴人员进行计数。3 和 4。进入施工现场的人员在穿过绿色线后计数，离开人员则在穿过黄色线后计数。机器学习流程由一个 YOLOX 目标检测模型和一个用于检测佩戴安全帽的头部、以及一个 Deep-SORT [12] 多目标跟踪算法组成。YOLOX 模型使用中国东北大学提供的数据集进行训练（https://public.roboflow.com/object-detection/hard-hat-workers）。

系统的结果与大华摄像机内置算法获得的结果进行了比较。需要注意的是，该摄像机统计了所有 进出人员，包括未佩戴防护头盔的人员。不过，这对结果的影响应该不大，因为该工地的安全规定禁止未佩戴安全帽进入，且在工人到达计数位置之前，这一规定会得到严格执行。

测量分为两个系列进行，每个系列采用不同的边界框定义。每个系列覆盖施工现场的一个完整工作日。按小时统计进入和离开的人数（上午5点至晚上7点之间）。

5 结果

表1和表2展示了所进行实验的结果。表1包含2022年11月22日进行的测量，边界框设置如图3所示。摄像机测得的进入人数与模型测得的人数之间的平均差值为 -6.21，标准差为 $\sigma = 5.08$；离开人数方面，平均差值和标准差分别为1.35和 $\sigma = 2.73$ 。摄像机检测到的入口人数与部署在人工智能与机器学习平台上的模型检测结果之间的相关性，以皮尔逊系数表示为0.988；出口人数方面为0.995。相关性在统计学上显著（$p \leq 0.05$）。表2包含2022年11月24日的测量结果（采用修改后的检测区域，如图4所示）。当天，进入人数的平均差值为 -4.93 ，标准差为 $\sigma = 4.25$；离开人数的平均差值为3.92，标准差为 $\sigma = 4.92$。对于这些测量，进入人数的皮尔逊系数为0.993，离开人数为0.989。相关性在统计学上显著（$p \leq 0.05$）。



图3. 11月22日边界框的位置（修改前）

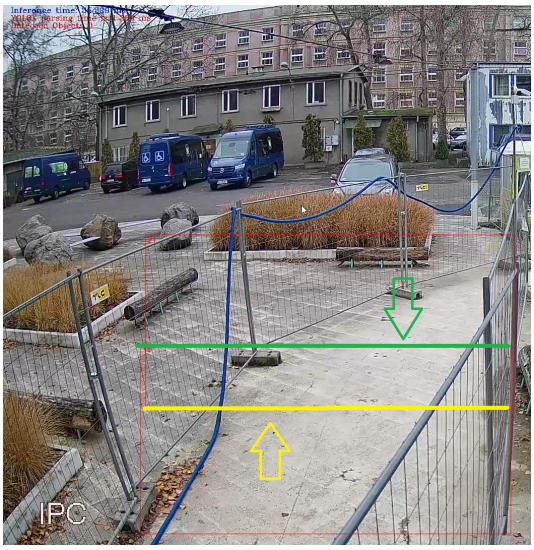


图4. 11月24日边界框的位置（修改后）

表格还展示了摄像机与人工智能与机器学习平台检测到的人数差异，以及分别对两个移动方向（进入和离开）计算出的差异总和。

实验期间发生了几起意外情况，对报告的结果产生了重大影响。观察到工人的行为出乎意料——他们在检测区域逗留或四处走动（图 5）。还注意到，有时工人通过检测区域后才戴上安全帽（图6）。这些行为给未来的系统带来了挑战，因为它们会显著影响系统的准确性。

Hour	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	总计
大华进入人数	12	65	84	47	26	84	50	51	28	70	28	8	9	0	562
大华离开人数	2	15	21	35	81	44	61	31	63	32	59	66	26	8	544
AI&ML进入人数	11	78	87	52	28	96	60	58	43	75	33	18	10	0	649
AI&ML离开人数	2	13	23	33	73	44	63	31	58	31	59	62	25	8	525
进入人数差异	1	-13	-3	-5	-2	-12	-10	-7	-15	-5	-5	-10	-1	0	-87
离开人数差异	0	2	-2	2	8	0	-2	0	5	1	0	4	1	0	19

表1。2022年11月22日施工现场的进入和离开人数。

Hour	05:00	06:00	07:00	08:00	09:00	10:00	11:00	12:00	13:00	14:00	15:00	16:00	17:00	18:00	总计
大华进入人数	4	57	113	62	34	73	75	65	56	93	27	10	9	0	678
大华离开人数	0	10	29	57	80	53	74	41	82	52	49	84	20	8	639
AI&ML进入人数	3	61	113	68	43	76	79	73	69	98	33	21	10	0	747
AI&ML离开人数	0	11	26	46	64	48	72	38	73	49	47	82	22	6	584
进入人数差异	1	-4	0	-6	-9	-3	-4	-8	-13	-5	-6	-11	-1	0	-69
离开人数差异	0	-1	3	11	16	5	2	3	9	3	2	2	-2	2	55

表 2。进入人数 施工现场的出入口，2022 年 11 月 24 日　日。

6 结论

测试模型在所研究的场景中表现相对良好。经验证，该模型在统计佩戴防护头盔人员数量时准确度足够，并通过与另一套系统进行对比得到了验证。模型与摄像机的计数结果存在一些差异，这可能归因于意外情况（图 5和 6），以及大华摄像机无法区分佩戴和未佩戴安全帽的人员这一事实。摄像机与 Orange AI&ML 平台模型之间较高的相关系数表明，这两种方案的表现相当出色。

需要指出的是，实验天数与边界框之间的相关性发生了变化。这些差异可由边界框的变化解释。这是一个有待进一步研究的参数。

所提出架构的两种变体均可用于所研究的施工现场个人防护装备检测场景。研究证实，采用边缘部署是可行的：服务器的计算能力足以维持令人满意的推理准确率。因此，可以得出结论，该施工现场配备了足够的硬件，可支持进一步开展部署实验。

未来将从网络延迟、资源利用率和准确度等方面比较所提出的两种架构变体。还将进一步测试所展示的模型，包括对视频进行人工标注，以获取用于比较的真实标注。这将有助于确定所开发模型的实际准确度。此外，还计划进一步优化边界框位置。

致谢

本研究得到 ASSIST-IoT 项目的支持，该项目由欧盟 H2020 RIA计划资助，资助协议编号为 957258。

参考文献

[1] Alexey Bochkovskiy、Chien-Yao Wang 和 Hong-Yuan Mark Liao。2020。YOLOv4：目标检测的最佳速度与精度。*arXiv* 预印本 *arXiv:2004.10934*（2020）。[2] Marco di Benedetto、Enrico Meloni、Giuseppe Amato、Fabrizio Falchi 和 Claudio Gennaro。2019。利用虚拟世界学习安全装备检测。载于*2019*年内容型多媒体索引国际会议（*CBMI*）。1–6。https://doi.org/10.1109/ CBMI.2019.8877466[3] Alejandro Fornés-Leal、Ignacio Lacalle、Carlos E Palau、Paweł Szmeja、Maria Ganzha、Marcin Paprzycki、Eduardo Garro 和 Francisco Blan-quer。2022。ASSIST-IoT：面向下一代物联网的参考架构。载于智能软件方法、工具与技术新趋势。IOS Press，109–128。[4] Andrew G. Howard、Menglong Zhu、Bo Chen、Dmitry Kalenichenko、Weijun Wang、Tobias Weyand、Marco Andreetto 和 Hartwig Adam。2017。MobileNets：面向移动视觉应用的高效卷积神经网络。https://doi.org/10.48550/ARXIV.1704.04861[5] Yutao Huang、Xiaoqiang Ma、Xiaoyi Fan、Jiangchuan Liu 和 Wei Gong。2017。深度学习与边缘计算的结合。载于*2017*年*IEEE*第*25*届网络协议国际会议（*ICNP*）。 1 2 10 1109 2017 8117585 – 。https://doi.org/ 。[6] /ICNP. 。Anis Koubaa 、 Adel Ammar 、 Anas Kanhouch 和 Yasser AlHabashi。2022。实时人脸识别推理的云端与边缘部署策略。*IEEE*网络科学与工程汇刊9，1（2022），143–160。https://doi.org/10.1109/TNSE.2021.3055835[7] Yange Li、Han Wei、Zheng Han、Jianling Huang 和 Wei-Dong Wang。2020。基于卷积神经网络的工程管理安全帽深度学习检测。土木工程进展2020（09 2020），1–10。https://doi.org/10.1155/2020/9703560[8] Luyang Liu、Hongyu Li 和 Marco Gruteser。2019。面向移动增强现实的边缘辅助实时目标检测。载于第*25*届移动计算年度国际会议和



图 5. 异常工人行为——在检测区域停留过久



图 6. 异常工人行为——在检测线后方戴上安全帽

网络（洛斯卡沃斯，墨西哥）（*MobiCom '19*）。计算机协会，纽约州纽约市，美国，文章25，共16页。https://doi.org/ 10 1145 3300061 3300116 。 / 。Peng Liu、Bozhao Qi 和 [9] Suman Banerjee。2018。EdgeEye：面向实时智能视频分析的边缘服务框架。1–6。https://doi.org/ 10 1145 3213344 3213345 。 / 。 Weiran Liu、Yi Hu 和 [10] Dawei Fan。2022。基于改进YOLOv5的安全帽佩戴识别。载于*2022*年第*11*届国际信息与通信技术会议（*ICTech*））。 466 470 10 1109 55460 2022 00099 – 。 https://doi.org/ 。 /ICTech 。 。Yao Nan、Qin Jian-Hua、Wang Zhen 和 [11] Wang Hong-Chang。2022。基于关键区域注意力机制的安全帽检测动态模型。载于*2022*年第7届亚洲电力与电气工程会议（*ACPEE*）。1296–1303。https://doi.org/10.1109/ ACPEE53904.2022.9783764

[12] Nicolai Wojke、Alex Bewley 和 Dietrich Paulus。2017。基于深度关联度量的简单在线实时跟踪。https://doi. org/ 10 48550 1703 07402 . /ARXIV. . Pei-Shao Wu、Chun-Yi [13] Lin、Tang-Yu Cheng 和 Wu-Sung Yao。2021。基于 Tiny-YOLOv3 的工业安全自动识别系统分析与设计。载于 2021 年 IEEE 国际未来能源电子会议（IFEEC）。1–6。https://doi.org/10.1109/ IFEEC53238.2021.9661778[14] JIANG Xinhua、XUE Heru、ZHANG Lina 和 ZHOU Yanqing。2021。结合统计特征与人工神经网络的低分辨率安全帽图像识别研究。载于《国际仿真：系统、科学与技术》（ IJSSST）。1–6。https://doi.org/ 10 5013 17 37 11 . /IJSSST.a. . . Chang Xu、 [15] Jinyu Tian 和 Zhiqiang Zeng。2022。用于头盔识别的轻量级融合通道注意力卷积神经网络。载于 WSPC会议论文集。1–8。[16] Geng Zhang、Lei Lv、Li Dan 和 Min Zhu。2017。基于颜色和形状的安全帽识别方法。载于第五届国际机械、材料与计算技术会议（ICMMCT）。1–5。[17] Qingyang Zhang、Hui Sun、Xiaopei Wu 和 Hong Zhong。2019。面向公共安全的边缘视频分析：综述。《IEEE会刊》107，8（2019），1675–1696。[18] Fangbo Zhou、Huailin Zhao 和 Zhen Nie。2021。基于 YOLOv5 的安全帽检测。载于 2021 年 IEEE 国际电力电子与计算机应用会议（ICPECA）。6–11。https: //doi.org/ 10 1109 51329 2021 9362711 . /ICPECA . . Mudi [19] Zhou、Zhuli Fang、Bin Zhao 和 Pengfei Li。2021。基于 YOLOv4 的安全帽佩戴检测与识别。载于 2021 年第三届科学技术创新国际学术交流会议（IAECST）。https://doi.org/10.1109/IAECST54258.2021. 9695790[20] Xihao Zhou 和 Sye Loong Keoh。2020。边缘端人脸识别模型部署：可行性研究。载于 2020 年第 21 届亚太网络运营与管理研讨会（APNOMS）。214–219。 https://doi.org/10.23919/APNOMS50412.2020.9236972