基于注意力机制的光线昏暗条件下口罩佩戴检测

【摘要】

1、针对在光线昏暗条件下口罩佩戴检测准确率较低的问题，提出了将注意力机制与YOLOv5网络模型相结合的口罩佩戴检测方法。首先对训练集图片使用图像增强算法进行预处理，然后将图片送入到引入了注意力机制的YOLOv5网络中进行迭代训练，完成训练后，将最优权重模型保存并在测试集上测试。在注意力的加持下，该模型能有效的增强人脸和口罩等关键点信息的提取，提高模型的鲁棒性，在光线昏暗条件下对口罩佩戴的检测准确率能达到92%，能够有效满足实际需求。

2、本文主要针对在可见度不高、光线昏暗的场景下，利用图像增强算法对图片进行预处理，将通道注意力和空间注意力结合，充分挖掘人脸口罩等关键特征点，同时对YOLOv5网络的损失函数进行相应的改进，提高模型在昏暗条件下的鲁棒性。

【背景】新冠肺炎疫情暴发后，人类健康受到巨大威胁，人们正常的工作与生活也受到了极大影响。为避免新冠肺炎疫情继续传播，人们在外出时规范佩戴口罩便成为了一项有效的防控手段。仅靠人工方式对流动人员进行口罩佩戴检测不仅效率较低，且会耗费大量人力物力，同时由于新冠肺炎传染性极强，近距离接触待检测人员更是存在较大的安全风险。因此，构建口罩自动检测系统检测活动人员在各类车站、大型商场等公共场合的口罩佩戴情况，对疫情防控具有重要的现实意义。

【其他技术的不足】对可见度不高、光照强度不强的昏暗条件下，其检测精度仍有待提高。

【改进后的网络模型】

1、模型损失函数(Loss)由分类损失(Classification Loss)、定位损失(Localization Loss)和目标置信度损失(Confidence Loss)组成，如公式(2)所示。本文采用二元交叉熵损失函数来计算分类损失和目标置信度损失，分别如公式(3)和公式(4)所示。

2、在原始YOLOv5的网络基础上，引入卷积注意力模块CBAM[19](Convolutional Block Attention Module Network)。CBAM包含两个子模块，分别是通道注意力模块CAM(Channel Attention Module)和空间注意力模块SAM(Spatial Attention Module)。CAM汇总通道注意力信息，SAM汇总空间注意力信息。与通道注意力不同，SAM主要关注于目标在图像上的位置信息，它将CAM的输出特征图作为本模块的输出特征图。由于CBAM模型在通道注意力模块中加入了全局最大池化操作，它能在一定程度上弥补全局平均池化所丢失的信息。其次，生成的二维空间注意力图使用卷积核大小为7的卷积层进行编码，较大的卷积核对于保留重要大的卷积核对于保留重要的空间区域有良好的帮助。添加了CBAM的YOLOv5网络不仅能更为准确的对目标进行分类识别，而且能更为精准的定位目标所在的位置。

【数据集】

1、通过网络爬取与自行拍摄相结合制作数据集，数据集中80%来源源于网络，20%来源于实际拍摄。在昏暗条件下的口罩佩戴图片相对缺乏，实际拍摄主要获取的是这类图片，实验过程中从楼道、室内等光线昏暗的场所进行了图片采集，同时也在傍晚和清晨等光线较弱的环境下进行了拍摄。

2、实验数据集共包含9000张图片，对其进行手工标注。该数据集分为两种类别，分别是bad和good，bad表示人员未佩戴或未按规范佩戴口罩，good表示正确佩戴口罩。数据集采用YOLO格式，图片标注使用LabelImg，标注后的文件以.txt作为后缀，文件名称和图片名称一致。训练集和测试集所占的比例为8：1

【模型训练】网络模型训练阶段，迭代批量大小设置为32，总迭代次数为600次。初始学习率设置为0.001，采用小批量梯度下降法，并使用Adam优化器计算每个参数的自适应学习率。大约在350次迭代后，模型开始逐渐收敛。

【结论】

1、实验结果表明，在可见度不高、光照强度不大的昏暗条件下，与参考文献[5]和AIZOO的方法相比，本文方法使用图像增强能改善图片的质量，然后通过注意力机制能更加准确的提取人脸口罩关键点特征，从而使得检测的准确率更高，具有较强的鲁棒性和扩展性，基本能够达到视频图像实时性的要求。

2、本文提出的基于注意力机制的光线昏暗条件下的口罩检测方法可应用在口罩佩戴识别系统中，具有较强的鲁棒性和可扩展性，对推进口罩佩戴检测的自动化、智能化，实现疫情防控和公共卫生安全具有重要的现实意义。