

基于语义分割的防侵入系统^①

罗耀俊^{1,2}, 向海^{1,2}, 胡晓兵^{1,2}, 牛洪超^{1,2}, 魏上云^{1,2}

¹(四川大学机械工程学院, 成都 610065)

²(四川大学宜宾园区, 宜宾 644000)

通信作者: 胡晓兵, E-mail: huxb@scu.edu.cn



摘要: 对现有门式起重机的防侵入系统进行了详细分析后, 针对其工作环境复杂、智能化水平较低的特点, 建立了基于机器视觉和深度学习技术的防侵入监测模型. 针对各目标检测算法和语义分割算法的优缺点, 采用了语义分割算法作为防侵入模型, ICNet 作为主要的语义分割网络. 对比其他网络, ICNet 网络具有 99.37% 的训练精度和 1.81% 的训练损失, 都展现出了最优的精度. 整体证明了基于语义分割的防侵入系统的智能性、可行性.

关键词: 语义分割; 防侵入; 门式起重机; ICNet

引用格式: 罗耀俊, 向海, 胡晓兵, 牛洪超, 魏上云. 基于语义分割的防侵入系统. 计算机系统应用, 2022, 31(1): 65-72. <http://www.c-s-a.org.cn/1003-3254/8276.html>

Anti-intrusion System Based on Semantic Segmentation

LUO Yao-Jun^{1,2}, XIANG Hai^{1,2}, HU Xiao-Bing^{1,2}, NIU Hong-Chao^{1,2}, WEI Shang-Yun^{1,2}

¹(School of Mechanical Engineering, Sichuan University, Chengdu 610065, China)

²(Yibin R&D Park of Sichuan University, Yibin 644000, China)

Abstract: After a detailed analysis of the existing anti-intrusion system of gantry cranes, an anti-intrusion monitoring model based on machine vision and deep learning is built in view of the complex working environments and low intelligence levels of the system. In consideration of the advantages and disadvantages of each target detection algorithm and semantic segmentation algorithm, the semantic segmentation algorithm is adopted as the anti-intrusion model, and the ICNet is used as the main semantic segmentation network. Compared with other networks, ICNet displays the best accuracy, with a training accuracy of 99.37% and a training loss of 1.81%. The results prove the intelligence and feasibility of the anti-intrusion system based on semantic segmentation.

Key words: semantic segmentation; anti-intrusion system; gantry crane; ICNet

现如今, 随着微电子、机器视觉、5G 通讯、人工智能技术的快速发展, 市场对门式起重机的自动化、信息化和智能化水平要求越来越高^[1]. 对于开放式工作的门式起重机来讲, 安全是其面临的一个重大的挑战, 尤其是门机轨道运行区域, 直接影响其能否畅通运行. 所以, 对于研究一套安全、可靠、智能的防侵入系统, 是具有十分重要的工业实用意义的^[2-5].

对于门式起重机的防侵入监测任务而言, 研究的

重点在于如何在运行过程中对监测区域中的障碍目标以及可能进入特定区域的目标进行检测、跟踪和预警. 障碍物检测研究主要集中在机器视觉, 向荣等人^[6]使用帧差法进行障碍物检测, 但容易受环境光线的影响; 战强等人^[7]采用 HIS 颜色空间对障碍物进行检测分割, 通过建立空间位置坐标关系, 使用栅格搜索算法对障碍物进行定位; 王荣本等人^[8]对智能车辆的障碍物检测进行了综合性分析, 并提出多传感器融合方法

^① 基金项目: 中国制造 2025 四川行动计划 (2018ZZ011); 川大-宜宾校地合作项目 (2019CDYB-13)

收稿时间: 2021-04-03; 修改时间: 2021-04-30; 采用时间: 2021-05-17; csa 在线出版时间: 2021-12-17

进行车辆前方障碍物检测;辛煜^[9]采用激光雷达实时对无人车周围环境进行三维重建,并提出了时空障碍物栅格图的障碍物检测方法;刘聪美等人^[10]采用Hough直线检测对煤矿铁轨进行识别,并采用支持向量机算法对行人进行监测,根据行人与铁轨的相对位置关系,发出不同的预警信号。

本文针对实际工作环境和工业需要,结合语义分割在各方向的应用^[11-17],对门式起重机防侵入问题进行分析,将其转换为语义分割问题。标定用于防侵入研究的语义分割数据集GCAID,并利用多个语义分割网络进行训练对比,最后决定选用ICNet网络运用于门式起重机的防侵入监测任务。门式起重机防侵入监测模型首先对图像进行语义分割,得到轨道和入侵目标的分割结果。然后根据轨道的分割结果,将监测视场划分为三个互不重叠的区域,分别是禁止区域、过渡区域和安全区域。接着计算检测的入侵目标所处的入侵区域,并根据入侵目标所处的区域,向控制系统发送不同的预警信号。

1 问题建模

如图1所示,将环境数据划分为4部分,分别是禁止区域、过渡区域、安全区域以及障碍目标。禁止区域属于轨道区域,禁止出现行人等危险源,当出现行人等危险源时应及时对门机进行减速和制动;过渡区域是沿轨道区域外扩0.5 m的区域,分布在禁止区域两侧,当出现行人等危险源时应发出警报信息并减速;安全区域是指除禁止区域和过渡区域的其他区域,对于这些区域存在的潜在危险目标可以按照正常速度运行;障碍目标是指可能影响门机安全运行的目标,结合本文的实际研究情况,暂定只考虑行人目标。在门机的防侵入检测问题中,只需重点检测禁止区域和过渡区域即可,过渡区域的划分是基于禁止区域的划分,因此需要首先检测出禁止区域。如图1左图所示,禁止区域包括轨道以及固定螺栓等目标,整个禁止区域与过渡区域存在明显的边缘、颜色等特征的差异,因此可以将整个禁止区域作为其中的一个检测对象。过渡区域的划分是在禁止区域的基础上进行的,且过渡区域与安全区域的特征基本是一致的,因此不做单独检测。障碍目标以行人检测为主,其可以在任何位置出现,针对不同位置对门机的运行状态进行不同的控制,分别是禁止区域减速停车、过渡区域警报减速、安全区域正常

行驶。如前所述,如果采用目标检测算法来对问题进行建模,存在的问题是采用边界框来划分区域容易使得处于安全区域的行人被划入禁止区域,从而使得门机总处于预警和停车状态。此外,对于处于禁止区域、但距离门机较远位置的障碍目标,在提前预警的情况下,如果在门机到达前障碍目标能够提前移除过渡区域,那么仍然可以不用停车。对此,可以将该问题转换为语义分割问题,前景分别是轨道区域(即禁止区域)和障碍目标(即行人),背景为过渡区域和安全区域。通过轨道区域来确定禁止区域、过渡区域和安全区域,根据行人的位置所属区域来判断门机的执行命令。

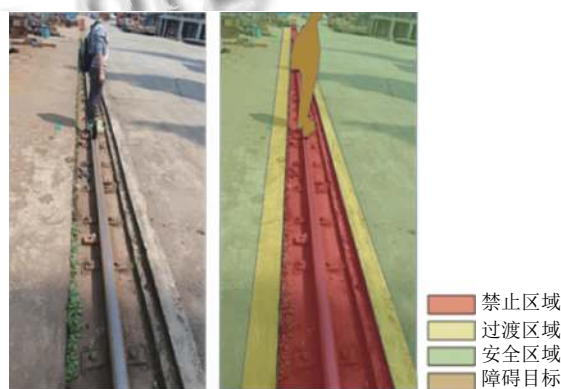


图1 门机防侵入建模示意图

基于以上分析,门机的防侵入检测系统的关键是对轨道区域和行人进行准确地检测和分割。经典的语义分割算法,如灰度分割、条件随机场等,均难以满足复杂自然环境下的在线检测;基于深度学习的语义分割算法近年来取得大量研究成果,在医学图像分割、缺陷检测与分割、街景分割、自动驾驶、卫星图像分割等领域已经成功应用。因此,本文采用基于深度学习的语义分割模型,以实现轨道区域和行人的精确分割。

2 ICNet 网络网络架构

ICNet (image cascade network) 网络^[18]是由香港中文大学联合腾讯优图、商汤科技于2018年在欧洲计算机视觉国际会议 (European Conference on Computer Vision, ECCV) 上首次公开,相比于与UNet、SegNet、FCN等网络,ICNet综合利用了不同分辨率的图像来提高网络的检测精度和速度,其核心思想是采用图像级联框架来逐步细化分段预测,首先使用低分辨率图像来获取粗略的预测掩膜,然后通过级联融合来源以引入高分辨率图像特征,并逐步改进已获得的粗略预

测. ICNet 通过将图像语义分割任务划分为粗预测和精调整, 在单块英伟达 1080Ti 图像处理器的情况下, 可以以 30.3 帧/s 的速度处理 1024×2048 分辨率的图像, 并且平均 IOU 精度可以达到 71%.

降采样输入图像、减少特征图数量以及模型压缩是最常用的检测加速策略, 但是无论哪一种策略, 都难以平衡网络推断速度和精度. 尽管这 3 种策略可以显著地减少运行时间, 但同时也只能得到比较粗糙的预测结果. 而如果直接使用高分辨率图像进行检测, 计算量通常是难以接受的. 如图 2 所示, ICNet 并不是简单地选择任何一种策略来加快推断速度, 相反, 它是通过采用不同分辨率大小的输入, 并采用级联特征融合单元来进行快速的语义分割. 训练时采用级联标签进行

引导, 而在测试阶段直接输出最终预测结果. 输入图像尺寸为全尺寸, 如 1024×2048 分辨率的 Cityscape 数据集图像, 并将输入图像进行下采样至输入的 1/2 和 1/4, 从而得到低、中、高 3 种不同分辨率的图像, 并将它们分别输入到 3 个不同的网络分支中. 将 1/4 输入图像送入 PSPNet 进行 8 倍下采样至输入分辨率的 1/32, 以进行粗分割; 将 1/2 输入图像和原图像分别送入两个更浅的卷积网络进行特征提取和下采样; 将最上分支的 1/32 特征图进行上采样并与第 2 分支的 1/16 特征图通过级联特征融合 (cascade feature fusion, CFF) 单元进行求和, 经过激活后再进行上采样, 并与第 3 分支的 1/8 特征图通过 CFF 模块进行融合, 并通过两次上采样以及卷积操作后得到与原图相同分辨率的预测图.

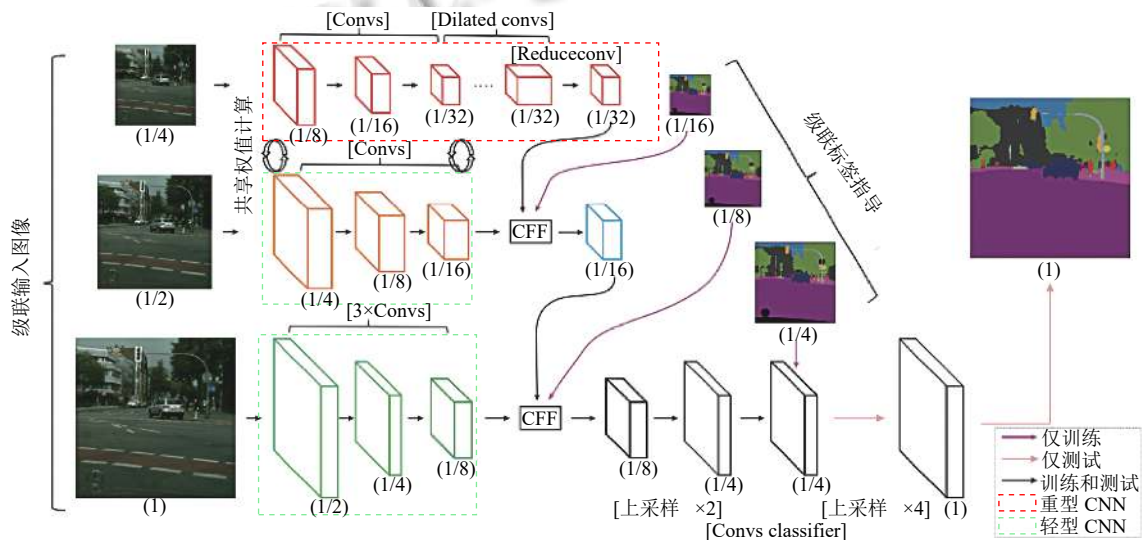


图 2 ICNet 网络架构图

虽然第 1 分支的卷积层数量比较多, 采用的是完整的 PSPNet 分割网络, 但由于输入图像的分辨率非常低, 因此其计算量依然比较有限, 推断时间仅需 18 ms, 内存占用也只有 0.6 GB. 由于第 2 层的卷积核共享了第 1 层的卷积核参数, 因此第 2 层的级联特征融合也仅仅只耗费了 6 ms. 第 3 层虽然输入图像的分辨率非常高, 但是由于卷积层数量比较少, 网络深度比较浅, 整个推断过程也仅需 9 ms. 因此, ICNet 通过不同分支进行不同分辨率输入的计算以及特征图的级联融合, 在保证分割精度的前提下, 可以有效地提高检测速度和减少内存占用.

图 3 为 CFF 单元示意图, 该单元地输入包括 3 个

部分: 两个特征图 F_1 、 F_2 以及真实标签, 尺寸分别为 $C_1 \times H_1 \times W_1$ 、 $C_2 \times H_2 \times W_2$ 和 $1 \times H_2 \times W_2$, 其中 F_2 的尺寸为 F_1 的两倍. 在特征融合前, 先对 F_1 使用双线性插值进行上采样 2 倍, 从而得到与 F_2 具有相同高度和宽度的特征图 F_1' . 然后再使用卷积核为 $C_3 \times 3 \times 3$ 、空洞率为 2 的空洞卷积, 对 F_1' 进行调整, 得到尺寸为 $C_3 \times H_2 \times W_2$ 的特征图 F_1'' . 空洞卷积可以融合相邻像素的特征信息, 相比于反卷积而言, 使用空洞卷积只需要更小的卷积核, 就可以获得与较大卷积核的反卷积一样大的感受野. 在保持感受野不变的条件下, 反卷积操作比空洞卷积需要更大的卷积核, 如反卷积的卷积核为 7×7 , 而空洞卷积可以使用 3×3 、空洞率为 2 的卷积核实现, 因

此空洞卷积需要的计算量更少。

为了使得 F_2 的通道数与 F_1 的通道数保持一致, 使用尺寸为 $C_3 \times 1 \times 1$ 卷积核对 F_2 进行卷积操作, 得到 F_2' , 其尺寸为 $C_3 \times H_2 \times W_2$ 。然后分别对 F_1' 和 F_2' 进行批归一化操作, 使得两个特征图可以保持在同一尺度空间。最后对归一化后的 F_1' 和 F_2' 进行逐元素求和以及 ReLU 激活操作, 从而得到融合后的特征图 F_2'' , 其尺寸为 $C_3 \times H_2 \times W_2$ 。为了增强网络对 F_1 的学习, 使用降采样后的真实标签来引导学习。如图 3 所示, 在第 1 层与第 2 层融合后, 使用 1/16 真实标签进行学习, 从而得到融合后的损失值; 同理, 使用 1/8 真实标签学习第 2 层与第 3 层的融合特征。对第 3 层融合后的特征进行上采样后, 计算与 1/4 真实标签的损失值。因此在训练期间, 损失值包括 3 部分, 第 1 层和第 2 层融合后的特征与 1/16 真实标签间的损失 $L_{1/16}$ 、第 2 层和第 3 层融合后的特征与 1/8 真实标签间的损失 $L_{1/8}$ 、第 3 层融合后的特征上采样后与 1/4 真实标签间的损失 $L_{1/4}$ 。

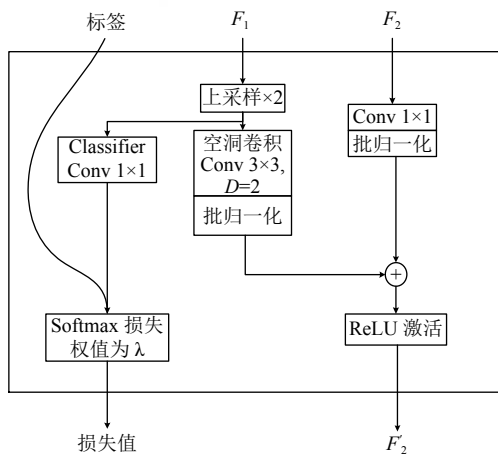


图 3 级联特征融合图

训练时, 为了增加每个分支的学习率, 采用级联标签引导策略, 利用 1/16、1/8 和 1/4 真实标签来引导学习低、中、高 3 种不同分辨率的学习, 对于 T 个分支和 N 个类别, 在分支 $t(t=1, 2, \dots, T)$, 预测特征图 F_t 的空间尺寸为 $Y_t \times X_t$, 在位置 (n, y, x) (n 表示特征图 F_t 的第 n 通道) 的值为 $F_{n,y,x}^t$, 真实标签在二维坐标中 (y, x) 对应的值为 \hat{n} 。对于分支使用权值为 λ_t 的带权 Softmax 交叉熵损失函数, 其数学表达式为:

$$\mathcal{L} = - \sum_{t=1}^T \lambda_t \frac{1}{Y_t X_t} \sum_{y=1}^{Y_t} \sum_{x=1}^{X_t} \log \frac{e^{F_{\hat{n},y,x}^t}}{\sum_{n=1}^N e^{F_{n,y,x}^t}} \quad (1)$$

测试时, 仅保留高分辨率的引导操作, 这样既可以保持训练过程中的稳定性, 同时又能防止网络受单个分支支配。

3 数据集的处理

由于目前尚无公开的门机轨道数据, 为了更好地匹配实际问题, 通过摄像头采集不同光线和天气条件下的现场视频。采集视频的分辨率为 1024×720 , 帧率为 30 帧/s, 颜色通道为 RGB 三通道。对所采集的视频间隔 15 帧取一张图片, 共计 1 680 张图片, 其中晴天为 1 420 张、阴天为 140 张、傍晚为 120 张。

考虑到只检测轨道和人, 因此标记类别有 3 类, 分别是轨道、人和背景。由于行人的数据在现场非常有限, 本系统在网络上下载了 12 个包含人的视频, 并按照 45 帧的间隔进行数据提取, 然后筛选不包含人的图片, 共得到 540 张分辨率为 1024×720 的数据集。最后选择将两个子数据集合并为最终的 GCAID 数据集。

Labelme 是一个开源的可视化图像标注工具, 它采用 Python+Qt 进行开发。Labelme 是目前常用的标注工具之一, 它可以标注用于语义分割、实例分割、场景分割等视觉任务的数据。Labelme 可以根据任务的需要对图像进行多边形、矩形、圆形、多线段、线段、点等形状的标注, 也可以对图像直接进行类别标注用于图像分类。此外, Labelme 还可以对视频进行标注。Labelme 标注的数据格式包括.xml 和.json 格式, 分别对应 VOC 类型和 COCO 类型的数据集。Labelme 兼容所有操作系统, 其操作界面如图 4 所示。

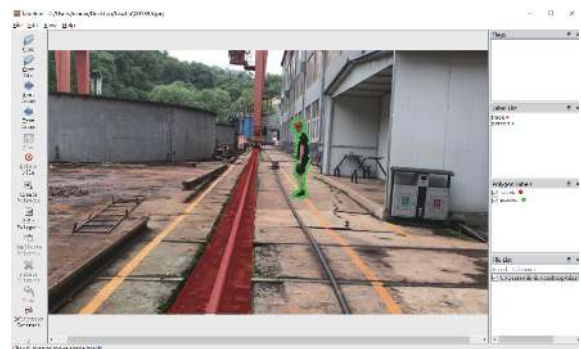


图 4 Labelme 标记工具操作界面

ICNet 对于图像的认识精度、数据集的标注精度及面积计算精度, 都会影响分割精度。由于数据集的标注精度会影响到网络对图像的认识精度, 网络对图像

的识别精度通过直接影响最后的类别预测, 从而影响面积计算精度, 最后影响分割. 所以, 本系统首先采用通过只标定轨道 (包括整个轨道安装区域) 与人, 背景区域自动生成的策略, 通过减少分割数量来减少 ICNet 对于图像识别精度的影响. 然后, 由于系统是对于门式起重机轨道的防侵入监测, 只需轨道与人处在分割区域内. 所以在数据集标记过程中, 对于轨道和人, 采用了不影响整体识别的前提下, 扩大标注区域, 来减小标注精度和图像识别精度对面积计算、类别分割的影响.

基于 Labelme 工具箱, 对包含 2 220 张图片的 GCAID 数据集按 json 格式进行标注. 如图 4 所示, 标签包含 2 类, 分别是 track 和 human, 代表图片中的轨道区域和人. 本文所标记的 track 并非仅仅表示轨道, 而是包含整个轨道安装区域, 因此其特征更加复杂, 难以通过经典的图像处理算法进行检测. 实际的语义标签包含 3 个类, 分别是背景、轨道和人, 背景类在标定的时候自动生成. 如图 5 所示, 图 5(a) 列表示原始数据, 包含了轨道、人以及各种复杂的背景, 其中 74% 的数据只包含轨道或人中的一种, 26% 的数据同时包含轨道和人; 图 5(b) 列表示真实语义分割图, 这部分数据是由人工标定, 作为训练时的监督数据; 图 5(c) 列表示掩膜可视化的结果, 通过图 5 可以清晰地看出, 数据集包含 3 个类别, 分别为 _background_、track 和 human, 对应地标签分别为 0、1 和 2. 在模型训练过程时, 通常只需要原始图片和真实语义分割图即可, 通常记为 (x, y) , 分别表示模型输入数据和训练目标.

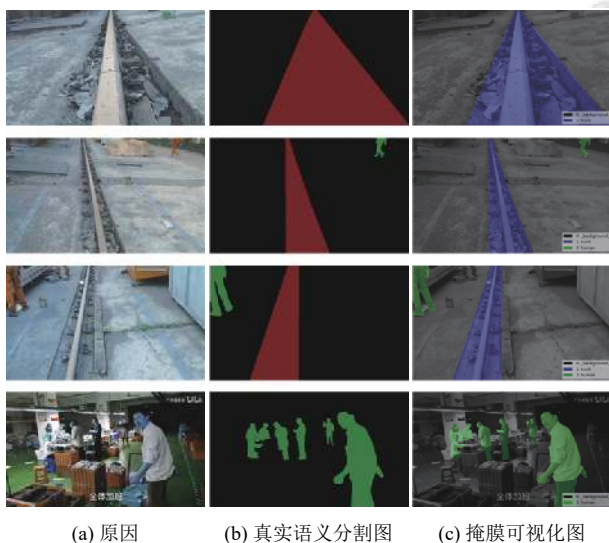


图 5 GCAID 数据集部分样本

4 实验分析

4.1 训练参数设置

采用 Keras 来实现 ICNet 网络, 训练时采用 4 块技嘉 2080Ti 图形处理显卡, 并使用 CUDA10.0 和 CUDNN v7.5 来加速训练过程. 将 GCAID 数据集划分为训练集、验证集和测试集共 3 个互不重叠的子集, 各子集包含的样本数分别为 1 698、189 和 333.

在训练超参数设置方面, 每个批次的样本数为 2, 训练轮数为 50, 每一轮迭代 375 次, 共迭代 37 500 次. 初始学习率为 0.01, 在第 30 轮、40 轮将学习率分别降低 10 倍. 训练过程中采用早停技术, 以防止出现过拟合. 在训练过程中, 每迭代 1 轮就进行 1 次验证. 为了增加样本的多样性, 采用包括随机缩放 (比例在 0.5~2)、随机镜像等数据增强策略. 训练时, $L_{1/16}$ 与 $L_{1/8}$ 的权值 λ_1 、 λ_2 均为 0.4, $L_{1/4}$ 的权值 λ_3 设置为 1. 网络的分割精度评估采用逐类交并比平均值 (mean of class-wise intersection over union, mIoU). mIoU 是通过计算每一类的 IoU 值后再求平均值, 计算公式为:

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (2)$$

其中, k 表示类别数, p_{ii} 表示像素预测类别 i 与真实标定类别 i 一致的像素点总数, $\sum_{j=0}^k p_{ij}$ 表示实际标定类别为 i 却预测类别为 j 的其他类的像素点数 (包含 p_{ii}), $\sum_{j=0}^k p_{ji}$ 表示实际标定类别为 j 却预测为类别 i 的像素点数 (包含 p_{ii}). 由于分母重复计算了 p_{ii} , 因此需要减去一个 p_{ii} .

4.2 数据分析

如图 6(a) 所示, 经过 50 轮的训练, 训练精度均达到稳定状态, 尽管 ICNet 的初始精度较低, 但是经过快速的上升后, 获得最高训练精度; 而 SegNet 模型的训练精度则明显低于其他模型. 如图 6(b) 所示, 所有模型的训练损失值都处于向下收敛, 其中, ICNet 的损失值下降速度最快.

图 7(a) 为常用语义分割模型验证精度曲线图, 从图中可以看出, SegNet 在前 20 轮训练中, 验证损失的波动幅度较大, 然后就趋于平缓; FCN-8 的验证精度持续上升, 在第 10 轮后一致保持最高的验证精度; UNet 的验证精度值最低, 始终处于 80% 以下. 如图 7(b) 所示, 除了 SegNet 的损失值前 20 轮波动幅度较大以外, 其他的模型的下落过程较为平稳.

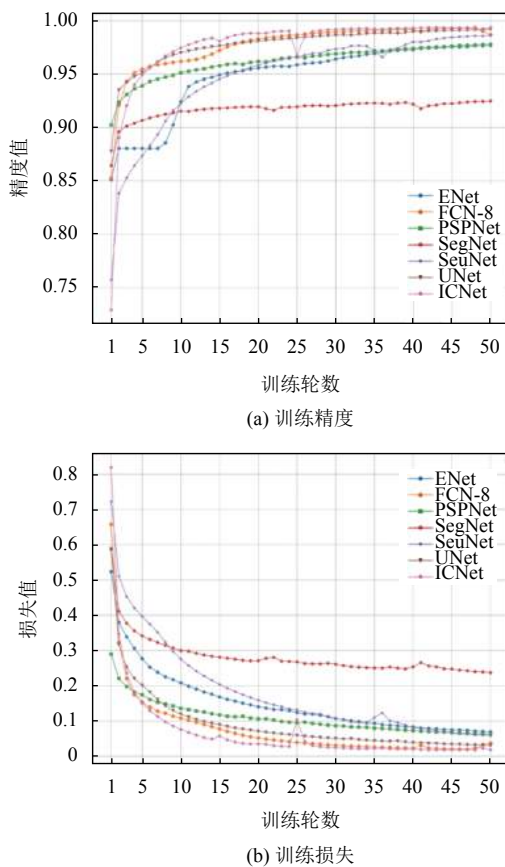


图6 常用模型训练精度曲线图与损失曲线图

如表1所示, ICNet 的训练精度和训练损失均优于其他模型, 而在验证精度与损失方面, FCN-8 则取得了最好的成绩, 且 ENet 也优于 ICNet. 通过对 333 张测试集进行测试分析, 并选用每个网络保存的最佳模型, 得到测试 mIoU 如表2所示, ICNet 取得 63.72% 的测试精度, 而 SegNet 的测试 mIoU 最低, 仅为 30.25%.

其次, 对各模型在测试集上的表现进行了分析, 评估指标包括 mIoU 和各类的 IoU. 测试结果如表2所示. 从表中可以看出, ICNet 对于 GCAID 数据集的平均分割精度最高, 为 63.72%.

其他多数框架针对高分辨率输入都具有相对密集的计算过程, 而 ICNet 由于采用了独特的级联特征融合网络, 通过在训练时使用了级联的标签监督, 有效地利用了低分辨率图和高分辨率图信息结合. 其中低分辨率分支被送入了 Heavy CNN 中, 得到了粗糙的语义预测. 中分辨率和高分辨率的分支只被轻量级的 CNN 处理, 进行恢复和完善模糊边界和丢失的细节, 可以减少中部分支和下部分支的参数数目. 虽然低分辨率分支导致了细节缺失和边界模糊, 但它已经获得了大部

分语义信息, 细节再由其他分支补充. 正是这种独特的级联网络使得测试结果中 ICNet 在所选网络中都呈现出较好的精度.

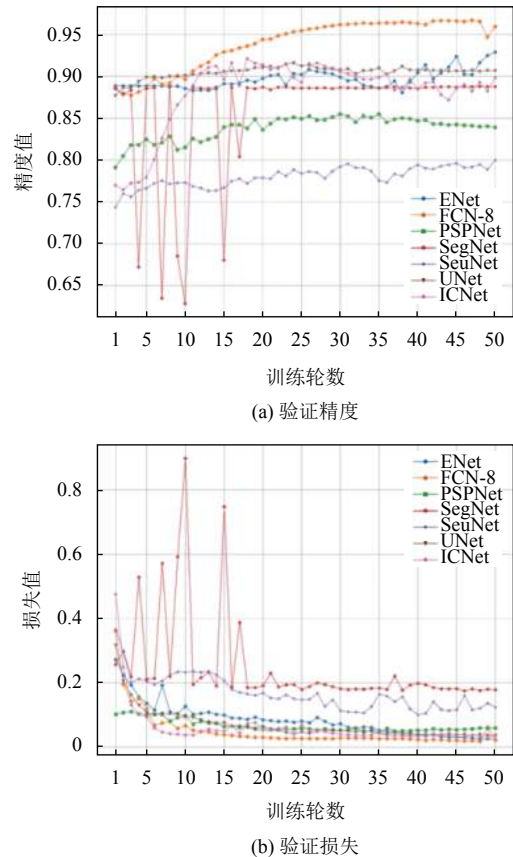


图7 常用模型训练精度曲线图与损失曲线图

表1 实验中相关参数设置对照

训练轮数	模型名称	训练精度	训练损失	验证精度	验证损失
50	UNet	0.992 0	0.0305	0.9065	0.0358
50	ENet	0.977 0	0.069 4	0.929 0	0.022 8
50	FCN-8	0.987 0	0.037 3	0.959 4	0.022 0
50	SegNet	0.924 9	0.237 1	0.882 7	0.177 9
50	ICNet	0.993 7	0.018 1	0.898 3	0.035 1
50	PSPNet	0.978 1	0.062 5	0.839 0	0.059 2
50	SEUNet	0.986 2	0.058 4	0.799 8	0.123 7

表2 实验中相关数据对照

模型名称	mIOU	IoU_{track}	IoU_{human}	$IoU_{background}$
UNet	0.391 5	0.087 0	0.291 6	0.795 9
ENet	0.485 8	0.476 7	0.157 8	0.835 2
FCN-8	0.610 4	0.605 3	0.348 2	0.877 8
SegNet	0.302 5	0.011 2	0.109 1	0.786 7
ICNet	0.637 2	0.654 4	0.402 1	0.855 3
PSPNet	0.481 7	0.233 7	0.382 4	0.828 8
SEUNet	0.403 9	0.190 0	0.208 8	0.812 9

最后,本文对 ICNet 网络的训练轮数进行了对比分析,如表 3 所示,随着迭代次数的增加,测试 mIoU 先出现上升,然后呈现下降,在 70 轮取得最佳 mIoU 值.可以看出,随着训练轮数增加,模型出现了过拟合情况.此外,本文展示了部分 ICNet 分割结果,如图 8 所示,可以看出,ICNet 对轨道的分割比较准确,但是对于人的分割精度并不是很理想,且容易将部分背景错误分类为人.针对人的分割精度较差的原因主要有两点:第一是人的姿态更加复杂,且在整个图像中的占比并不是很高,因此在训练过程中即使错误分类,依然可以得到较高的精度;第二是因为数据集中人的数量并不是很多,从而导致样本失衡.

表 3 ICNet 在不同迭代轮数下的测试结果

训练轮数	测试mIoU	提升幅度(%)
50	0.637 2	—
60	0.641 2	0.63
70	0.645 3	1.27
80	0.637 8	0.09
90	0.628 7	-1.33
100	0.622 1	-2.37

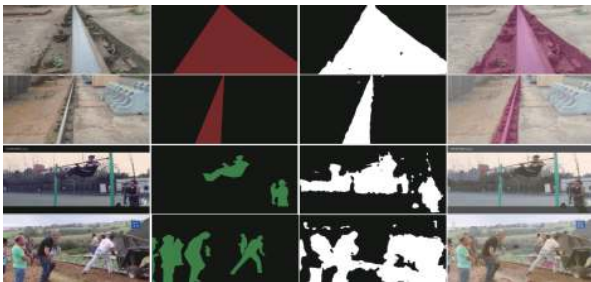


图 8 部分测试结果图

4.3 防侵入测试

在对防侵入进行测试中,本文基于轨道的分割结果将图片分为禁止区、过渡区和安全区,然后根据人的分割结果及其所处的区域发出不同的预警信号.测试时,将帧率设置为 20 帧/s,测试结果如表 4 所示,正确预警样本包含 48 张不包含人的样本和 5 张包含人但处于警告区的样本,错误预警的 3 张样本是将人划入到禁止区.

表 4 防侵入系统测试结果

样本数量	包含人的样本	不包含人的样本	正确预警样本	预警精度(%)
56	8	48	53	94.6

图 9 是部分防侵入测试结果,从图中可以看出,模型能够很好的对轨道各区域进行正确划分,效果良好.



图 9 防侵入分割测试结果

5 结论

本文针对门式起重机实际工作环境和工业需要,结合深度学习技术,对门式起重机防侵入问题进行了分析,并对门机的防侵入监测任务进行了详细的建模,确定了门式起重机智能防侵入系统的方案.采用了语义分割算法(ICNet 网络)对视场中的特定目标进行了检测和分割,并对选用的 ICNet 的网络结构做了详细的介绍.针对门机的防侵入监测需求,使用 Labelme 图片标注工具制作了 GCAID 数据集,用于训练 ICNet 和其他语义分割模型,也对训练网络参数设置进行了详细说明,并基于 Keras 测试了包括 ICNet、UNet、SegNet 等语义分割模型.对比其他网络,ICNet 网络具有 99.37% 的训练精度和 1.81% 的训练损失,都展现出了最优的精度,并且在实际防侵入测试中,也证明了基于语义分割的防侵入系统的智能性、可行性,为门式起重机或类似场景,提供了智能防侵入系统方案.

参考文献

- 1 许光辉,孙勇娜,叶浪.桥门式起重机智能化设计现状及研究.中国科技纵横,2016,(14):36.[doi:10.3969/j.issn.1671-2064.2016.14.030]
- 2 秦长海.基于视觉定位的起重机智能化控制系统.建筑机械化,2009,30(5):50-52.[doi:10.3969/j.issn.1001-1366.2009.05.014]
- 3 白雪,黄廷磊.运动目标精确检测算法.计算机系统应用,2009,18(12):41-44.[doi:10.3969/j.issn.1003-3254.2009.12.010]
- 4 朱良.户外周界期传感技术——周界防入侵系统的眼睛和耳朵.中国安防,2008,(3):39-41.[doi:10.3969/j.issn.1673-7873.2008.03.013]
- 5 乔传标,王素玉,卓力,等.智能视觉监控中的目标检测与

- 跟踪技术. 测控技术, 2008, 27(5): 22–24. [doi: [10.3969/j.issn.1000-8829.2008.05.007](https://doi.org/10.3969/j.issn.1000-8829.2008.05.007)]
- 6 向荣, 蒋荣欣. 铁路机车快速超视距障碍物识别算法. 湘潭大学自然科学学报, 2013, 35(2): 103–108. [doi: [10.3969/j.issn.1000-5900.2013.02.021](https://doi.org/10.3969/j.issn.1000-5900.2013.02.021)]
- 7 战强, 吴佳. 未知环境下移动机器人单目视觉导航算法. 北京航空航天大学学报, 2008, 34(6): 613–617.
- 8 王荣本, 赵一兵, 李琳辉, 等. 智能车辆的障碍物检测研究方法综述. 公路交通科技, 2007, 24(11): 109–113, 124. [doi: [10.3969/j.issn.1002-0268.2007.11.025](https://doi.org/10.3969/j.issn.1002-0268.2007.11.025)]
- 9 辛煜. 无人驾驶车辆运动障碍物检测、预测和避撞方法研究 [博士学位论文]. 合肥: 中国科学技术大学, 2014.
- 10 刘聪美. 基于图像处理的矿用电机车行人监测预警系统设计. 煤炭工程, 2016, 48(5): 17–19, 23. [doi: [10.11799/ce201605006](https://doi.org/10.11799/ce201605006)]
- 11 Geng QC, Zhou Z, Cao XC. Survey of recent progress in semantic image segmentation with CNNs. *Science China Information Sciences*, 2018, 61(5): 051101. [doi: [10.1007/s11432-017-9189-6](https://doi.org/10.1007/s11432-017-9189-6)]
- 12 Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich: Springer, 2015. 234–241.
- 13 Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(4): 640–651. [doi: [10.1109/TPAMI.2016.2572683](https://doi.org/10.1109/TPAMI.2016.2572683)]
- 14 He KM, Gkioxari G, Dollár P, *et al.* Mask R-CNN. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 386–397. [doi: [10.1109/TPAMI.2018.2844175](https://doi.org/10.1109/TPAMI.2018.2844175)]
- 15 Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481–2495. [doi: [10.1109/TPAMI.2016.2644615](https://doi.org/10.1109/TPAMI.2016.2644615)]
- 16 柳锋, 李俊峰, 戴文战. 基于深度学习语义分割的导光板缺陷检测方法. 计算机系统应用, 2020, 29(6): 29–38. [doi: [10.15888/j.cnki.csa.007411](https://doi.org/10.15888/j.cnki.csa.007411)]
- 17 韩利丽, 孟朝晖. 基于深度学习的视频语义分割综述. 计算机系统应用, 2019, 28(12): 1–8. [doi: [10.15888/j.cnki.csa.007195](https://doi.org/10.15888/j.cnki.csa.007195)]
- 18 Zhao HS, Qi XJ, Shen XY, *et al.* ICNet for real-time semantic segmentation on high-resolution images. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 418–434.