第 43 卷　第 9 期
2022 年 9 月

仪　器　仪　表　学　报
Chinese Journal of Scientific Instrument

Vol. 43 No. 9
Sep. 2022

DOI: 10.19650/j.cnki.cjsi.J2209185

基于 LAM-Net 的轨道侵入界异物自主检测系统*

叶　涛,赵宗扬,郑志康
(中国矿业大学(北京)机电与信息工程学院　北京　100083)

摘　要:针对轨道入侵异物对行车安全造成的极大威胁,而现有的轨道目标检测算法难以平衡检测精度和速度、易受复杂环境影响以及难以部署于嵌入式设备等问题,提出了一种轻量型自适应多尺度卷积神经网络,其通过特征图线性变换简化特征提取过程,使用自适应多尺度特征融合优化特征表达能力,并通过设计轻量型注意力进一步提升异物检测精度;同时,结合 NVIDIA Jetson TX2 嵌入式平台,研制了轨道入侵异物自主检测系统。实验结果表明,本文提出的模型很好地平衡了检测速度和精度,在 NVIDIA GeForce GTX1080Ti 的 GPU 平台上对轨道数据集的检测速度为 297 FPS,检测精度为 92.96%,比 YOLOv4-tiny 高 7.72%,实现了在轨道交通复杂场景下高精度、高速度以及高鲁棒性的检测入侵异物。

关键词: 目标检测算法;轻量型卷积神经网络;深度学习;轨道入侵异物;自适应特征融合;检测系统
中图分类号: U491.2　TH39　　**文献标识码**: A　　**国家标准学科分类代码**: 580.70

Research on the autonomous detection system for railway intrusion obstacles based on LAM-Net

Ye Tao, Zhao Zongyang, Zheng Zhikang
(*School of Mechanical Electronic & Information Engineering, China University of Mining & Technology, Beijing 100083, China*)

Abstract: The railway obstacles in front of the train have great threat to traffic safety. The existing railway object detection algorithms are difficult to balance the detection accuracy and speed, which are susceptible to complex environment and difficult to deploy in embedded equipment. To address these issues, the lightweight and adaptive multiscale convolutional neural network is proposed in this article. The model simplifies the computation of redundant feature maps in feature extraction process by means of feature map linear transformation, and the adaptive multi-scale feature fusion is used to optimize the ability and further improve the accuracy of foreign obstacles detection. In addition, combined with NVIDIA Jetson TX2, an autonomous intrusion detection system for railway traffic scenes is developed. Experimental results show that the proposed model performs a great compromise between detection speed and accuracy. The detection speed of LAM-NET on the NVIDIA GeForce GTX1080Ti is 297 FPS, and the detection accuracy is 92.96% (7.72% higher than that of YOLOv4-tiny), which can well realize the high precision, real-time and high robustness detection for railway obstacles.

Keywords: object detection algorithm; convolutional neural network; deep learning; railway foreign obstacles; adaptive feature fusion; detection system

0　引　　言

铁路运输业是国民经济的基础产业,在交通运输体系中居骨干地位,对国民经济发展起着强有力的支持作用。然而,随着我国铁路运输能力的不断提高,铁路行驶的速度提升、行车的密度增加,列车线路上的异物将对列车安全运行造成极大的威胁[1-2]。高铁列车之间的碰撞追尾主要发生在列车调车[3]过程中。在该模式下,目前主要依赖列车驾驶员来判断列车前方是否有异物进而采取紧急措施,而在铁路弯曲的情况下司机的视线更容易被堵塞,这已经很难保证高铁列车的安全,要想让列车在

收稿日期:2022-01-13　　Received Date: 2022-01-13
*基金项目:煤炭资源高效开采与洁净利用国家重点实验室开放基金(2021-CMCU-KF012)、中央高校基本科研业务费专项基金(2022YQJD04,2022YJSJD01)项目资助

雾天、阴天和夜间等弱光线环境和复杂背景环境下安全运行更是存在着更大的挑战。因此研究一种能够高精度实时检测轨道入侵异物的智能避障系统的理论和方法非常有必要,对于保证列车的行驶安全进而保护公众的生命财产安全具有很现实的意义。

许多学者对轨道入侵异物检测进行了大量的研究。Garcia 等[4]在道路两侧安装不同的传感装置形成红外屏障,利用数据融合生成图像特征计算物体的大小和方向并得到异物图像。Catalano 等[5]在轨道下安装了光纤光栅传感器检测轨道上行人产生的声波,实现了集成光纤传感器的轨道入侵异物检测系统。上述方法利用固定的监测点进行入侵异物检测,对体积较大、高度一定的异物检测效果较好,但很容易漏检小目标异物。随着计算机视觉技术的发展,基于视觉的检测方法在轨道异物检测领域得到了广泛的应用。Nanasone 等[6]提出一种单目相机和图像处理算法相结合的方法进行轨道异物检测;Teng 等[7]提出一种基于超像素的轨道目标检测方法,结合支持向量机[8-9]以实现异物分类,该算法计算简单、易于移植。但上述采用图像处理或手工特征提取来实现轨道入侵异物检测的方法过于依赖设计者的经验且无法自适应地提取特征,对于复杂场景中目标的高级语义信息难以充分的挖掘,场景迁移能力也差。因此,传统的目标检测算法往往无法在实际复杂多变的轨道环境中高精度实时检测轨道入侵异物。

随着深度学习的快速发展,基于卷积神经网络的自适应特征提取算法的检测速度和精度都较高且突破了传统手工特征构建的局限性。该目标检测算法大致可以分为两类:两阶段算法和单阶段算法。代表性的两阶段算法如 R-CNN[10-12]系列先产生候选框,再通过卷积神经网络进行目标检测。虽然这些方法精度较高,但由于过于复杂的计算量导致其检测速度很慢,无法实现在实际应用中实时检测目标。单阶段算法(you only look once, YOLO [13-17]及 single shot multibox detector, SSD[18])将检测问题视为回归问题,直接预测目标的类别和位置。相比于两阶段算法,它们的算法结构更加简单,实时性提升巨大。李晖晖等[19]采用改进 SSD 算法主干网络中修正线性单元 CReLU 和特征金字塔网络的方式加强浅层特征的传递效率,并提升了网络的定位精度和分类精度;Zhang 等[20]增加了 YOLOv3 特征融合中的卷积层,以增加计算量为代价使得目标定位更加准确。但二者的检测速度仍旧无法满足实际场景中实时检测目标的需求,并且难以部署在嵌入式设备中。由于实际场景下的目标检测对算法的检测精度和实时性的要求越来越高,近年来涌现出了许多旨在平衡检测速度和检测精度神经网络模型。Li 等[21]将标准卷积替换为深度可分离卷积降低计算量提高检测速度,易于部署,但检测精度却有待提高。文献[22]将候选区域的生成与检测任务结合在一起,采用跳层提取特征的方式既能获取高层语义信息又可以得到低层位置信息,有效提升了检测精度但实时性差。Womg 等[23]大量减少 SSD 主干网络中 3×3 滤波器的数量后再进行下采样,大幅度提升了算法的检测速度,但是其检测精度低仍旧无法保证在实际场景中准确高效检测目标。Wei 等[24]增加了 YOLOv3 中特征图拼接的次数提高检测,但由于数据集包含的目标类别不够全面的限制,该模型对于复杂和多目标场景下的识别能力仍有待验证。

针对上述问题,为在平衡算法的检测速度和精度的同时提高对复杂背景下异物的检测精度,本文提出了一种自适应多尺度轻量型卷积神经网络检测模型(lightweight and adaptive multiscale convolutional neural network, LAM-Net),检测速度为 297 FPS,检测精度为 92.96%,比 YOLOv4-tiny 高 7.72%。LAM-Net 的特征提取网络采用线性变换生成相似的卷积核特征,降低了模型计算复杂性,提升了检测速度;特征融合网络在融合前赋予不同尺度特征图独立的权重参数,实现了不同尺度特征图的自适应融合,提高了目标检测精度;使用注意力机制对通道权值进行优化,在保证实时检测的前提下使得检测精度进一步提高。并结合 NVIDIA JETSON TX2 计算平台研制了小型化的智能轨道入侵异物检测系统,满足了在轨道交通复杂场景下高精度、高速度以及高鲁棒性自主检测轨道异物的需求。

1　轨道入侵异物检测模型的设计原理及结构

1.1　模型设计原理和总体结构

轨道交通列车运行过程中由于复杂光线和多尺度目标形态的变化将加剧模型的漏检、错检风险,严重影响了轨道异物自主检测系统的可靠性。为实现轨道交通复杂场景下异物的精确检测并满足嵌入式端的计算需求,本文在单阶段的目标检测模型 YOLOv4-tiny 的基础上进行深入研究,从图像特征提取、特征融合等方面进行优化,提出了一种端到端的自适应多尺度特征融合目标检测网络 LAM-Net,实现了复杂的轨道场景中列车运行前方异物的全天候检测。模型结构如图 1 所示,主要包括轻量型特征提取网络和自适应多尺度特征融合网络两部分。

首先,将在真实轨道场景中采集的不同天气和背景的图像缩放至固定分辨率(416×416);其次,通过由线性特征运算和深度可分离卷积组成的轻量型特征提取网络构建了高鲁棒性的多尺度图像金字塔;然后,在特征融合部分利用可学习的权重参数来自适应调整不同尺度通道间融合的权重参数,实现自适应融合语义丰富的低分辨率特征图和位置信息丰富的高分辨率特征图,为进一步

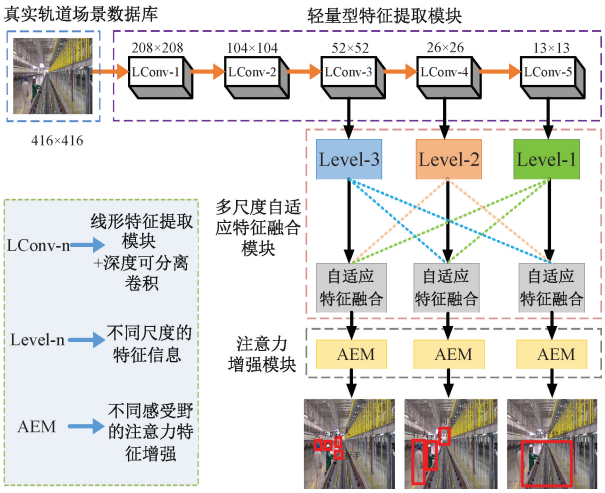


图 1 LAM-Net 总体框架

Fig. 1 LAM-Net overall framework

目标分类和位置回归提供可靠的信息来源,提高模型的目标检测能力;再次,通过不同感受野的注意力特征增强重新分配通道权重,在保证检测实时性的同时进一步提高检测精度;最后在融合后的不同尺度特征图上生成一系列预设候选框,将预设候选框与目标真实的标注框按照交并比匹配和非极大值抑制方法获得最终不同尺度异物检测的结果。

1.2 轻量型特征提取模块

本文通过对 YOLOv4-Tiny 目标检测模型结构进行了深入分析,其特征提取网络中存在大量密集连接运算,导致了模型内存的占用较大且检测速度较慢等问题。为简化特征提取方式以及提升模型的计算效率,本文提出了 LConv 结构以构建一种轻量高效的线性特征提取网络,与传统的深度卷积神经网络相比,该特征提取网络通过简化特征图冗余信息的生成过程来提高特征提取的效率,有利于大幅减少模型内存占用并提高模型检测的检测速度。1 个 LConv 结构由 2 个高效线性特征提取模块和深度可分离卷积组成,如图 2 所示。首先,第 1 个线性特征提取模块用来对输入特征图的通道数进行升维,实现高效鲁棒的目标特征提取;然后,使用深度可分离卷积对特征进行下采样,极大地降低了模型计算量;接着第 2 个卷积模块实现输出特征图的通道数的降维,在降低计算量的同时保证提取特征的丰富性;最后,添加捷径层将输入特征图和其他特征图进行相加,与传统的残差网络相加操作相比,该方式可以在降低内存访问成本提高检测速度的同时实现高效的特征整合。受 MobileNet 系列非线性激活函数对高维空间影响小、低维空间影响大的启发,本文在第 1 个线性特征提取模块对输入特征图进行降维后,采用 Leaky ReLU 对输入特征图进行激活,目的是在几乎不影响高维数据的前提下提升模型的特征表达能力;在第 2 个线性特征提取模块进行特征整合并降维后,本文使用 Linear activation 代替 Leaky ReLU,目的是避免低维的特征信息被干扰和破坏。

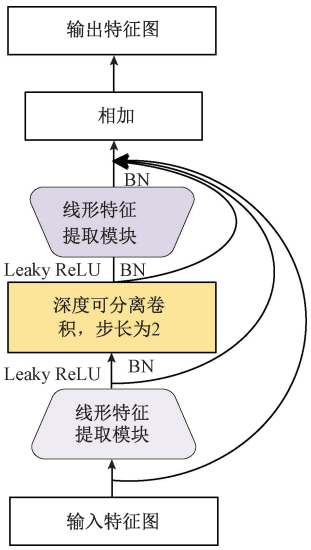


图 2 LConv 结构图

Fig. 2 LConv structure

LConv 中的线性特征提取模块用于解决传统卷积提取的中间特征图存在过多冗余的特征信息无法被充分利用且占用着巨大的计算资源的问题。其核心是将非线性卷积的计算转换为线性变换过程,将其中一个相似的特征图通过线性变换来近似地生成另一个相似特征映射,使用更少的计算参数量来获取更多特征图的冗余信息,从而保证在利用较小的计算资源的同时提高目标识别和检测结果的可靠性。如图 3 所示,线性特征提取模块的实现流程主要分为普通卷积层和线性特征变换层 2 个步骤。第 1 步,原始 c 个通道的输入特征图 $\boldsymbol{X}$ 经过普通卷积后生成 m 个通道的特征图 $\boldsymbol{Y}'$, 如式(1)所示:

$$\boldsymbol{Y}' = \boldsymbol{X} * \boldsymbol{f}' \tag{1}$$

其中,$\boldsymbol{f}' \in \boldsymbol{R}^{c\times k\times k\times m}$ 表示卷积运算中使用的卷积核维度信息,且 $m \leqslant n$(仅在所有特征图无相似时取等),同时为简化卷积计算过程,相比于普通卷积式(1) 省略了偏置项,其他卷积超参数与普通卷积保持一致,目的是使得模型输出特征图的空间尺寸大小与普通卷积相一致。

在第 2 步线性特征变换层的计算过程中,为获得与普通卷积相似的 n 个特征图,本文对第一步普通卷积输出 $\boldsymbol{Y}'$ 中的每个原始特征图应用简单的线性变换关系生成 s 个线性特征图,新的特征图生成原理如式(2)所示:

$$\boldsymbol{y}_{ij} = \boldsymbol{\Phi}_{i,j}(\boldsymbol{y}'_i), \quad \forall i = 1,\cdots,m, \quad j = 1,\cdots,s \tag{2}$$

其中,$\boldsymbol{y}'_i$ 表示第 1 步生成的原始特征图 $\boldsymbol{Y}'$ 中第 i 个通道的特征图;$\boldsymbol{\Phi}_{i,j}$ 表示第 j 个通道特征图的线性变换过程,用于生成第 j 个通道的线性特征图 $\boldsymbol{y}_{ij}$,即:特征图 $\boldsymbol{y}'_i$ 可

以有一个或多个相似的线性特征图 $\{\boldsymbol{y}_{ij}\}_{j=1}^{s}$,$s$ 为进行线性变换的次数;$\boldsymbol{\Phi}_{i,s}$ 表示表示原始输入特征图的恒等映射,最终将恒等映射的特征图和经 $s-1$ 次线性变换后的特征图进行拼接。通过应用简单的线性变换操作,生成 $n=m\times s$ 个通道特征图作为线性特征提取模块的输出特征图,用 $\boldsymbol{Y}=[\boldsymbol{y}_{11},\boldsymbol{y}_{12},\cdots,\boldsymbol{y}_{ms}]$ 表示。由于线性变换运算在每个单独通道特征图上进行,故相比于普通卷积,线性变换操作的计算量会大幅降低。此外,本文所提出的线性特征提取模块首先使用传统卷积在整合特征信息的同时进行降维以降低计算量,且在逐通道卷积提取特征之后并未使用逐点卷积进行计算,而是恒等变换等线性特征变换所生成的特征图进行拼接,故该模型的计算量相比于 MobileNet 系列得到了进一步降低。

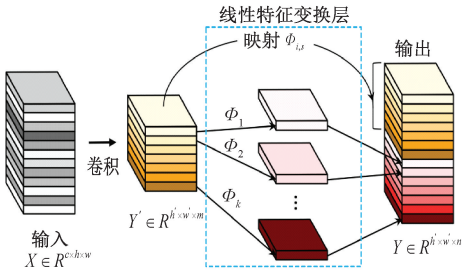


图 3　线性特征提取模块计算流程图

Fig. 3　Linear feature extraction module calculation flow chart

为进一步突显本文所设计的轻量型特征提取网络在处理特征冗余信息时的有效性,对 LConv1 中线性特征提取模块中的相似特征图进行了可视化,特征提取网络第一层输出特征图的可视化结果如图 4 所示,左上方展示了输入的原始图像,左框展示了由准卷积操作生成的卷积核特征可视化结果,右框展示由第 1 个线性特征提取模块变换的卷积核特征可视化结果。可视化结果表明,线性变换生成的卷积核特征足够灵活,能够在高鲁棒特征提取的同时减少模型的计算参数量和提升模型的计算效率,满足在实际轨道环境中检测轨道异物的需求。

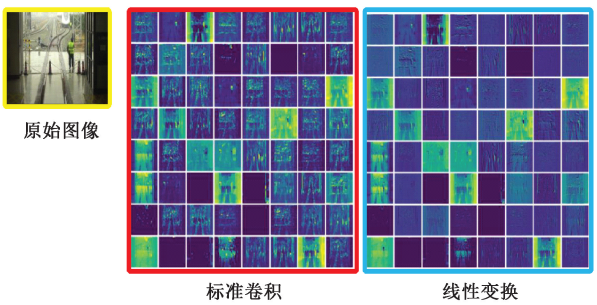


图 4　轻量型特征提取模块第 1 层卷积特征图可视化

Fig. 4　Lightweight feature extraction module for the first layer of convolution characteristics graph visualization results

1.3　多尺度自适应特征融合模块

图像特征金字塔常用于解决场景中目标尺度变化较大等问题,而不同特征图之间的尺度不一致是造成特征融合的主要限制。基于此问题,本文在 ASFF 融合策略的基础上,提出了一种自适应感受野融合结构嵌入至模型的特征融合模块预测层中,以实现自适应地学习不同感受野分支的融合权重,降低因目标尺寸不一致而造成的定位分类损失,增强模型的特征表示能力,提高模型对轨道异物的检测性能。融合模型结构如图 5 所示,对于融合某一层特征,首先将其他层次的特征进行尺度调整,然后进行特征融合并对其进行训练,以获得融合权重对不同特征尺度对于预测特征图的贡献进行学习。在多尺度检测的通道数调整中,采用 1×1 卷积操作将特征的通道数进行压缩,然后利用插值的方法来提高调整尺度分辨率;针对 2 倍下采样,利用步长为 2 的 3×3 卷积来调整通道数和分辨率,对于 4 倍下采样,在卷积前添加步长为 2 的最大池化层来实现。

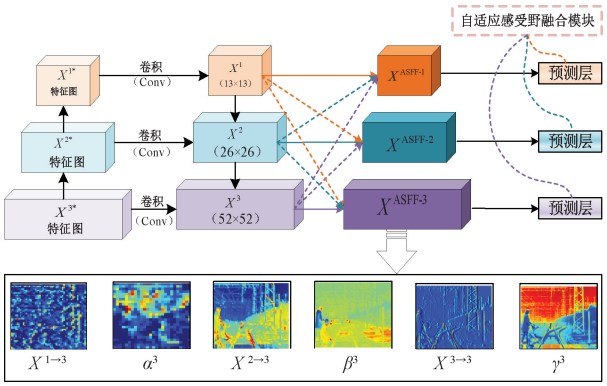


图 5　自适应特征融合方式

Fig. 5　Adaptive feature fusion method

模型多尺度特征融合的过程如式(3)所示,其中第 n 层特征图缩放后所得到的特征向量 $\boldsymbol{X}_{ij}^{n\to l}$ 与第 l 层特征图中(i,j) 位置处的特征向量大小相同;$\boldsymbol{\alpha}_{ij}^{l}$、$\boldsymbol{\beta}_{ij}^{l}$、$\boldsymbol{\gamma}_{ij}^{l}$ 分别表示融合后的第 l 层特征图自适应学习得到的 3 个融合权重参数值,三者的关系如式(4) 所示:

$$\boldsymbol{Y}_{ij}^{l}=\boldsymbol{\alpha}_{ij}^{l}\otimes\boldsymbol{X}_{ij}^{1\to l}+\boldsymbol{\beta}_{ij}^{l}\otimes\boldsymbol{X}_{ij}^{2\to l}+\boldsymbol{\gamma}_{ij}^{l}\otimes\boldsymbol{X}_{ij}^{3\to l} \tag{3}$$

$$\boldsymbol{\alpha}_{ij}^{l}+\boldsymbol{\beta}_{ij}^{l}+\boldsymbol{\gamma}_{ij}^{l}=1,\boldsymbol{\alpha}_{ij}^{l},\boldsymbol{\beta}_{ij}^{l},\boldsymbol{\gamma}_{ij}^{l}\in[0,1] \tag{4}$$

其中,嵌入至预测层中的自适应感受野融合模块如图 6 所示,该结构主要利用多分支感受野生成不同大小感受野的特征图,并通过模型自适应学习特征图的融合权重,提高模型的特征表示能力。

该模块可分为瓶颈结构、多分支感受野结构和感受野融合结构。自适应性感受野融合模块使用 1×1 卷积层来构建瓶颈结构,首先通过使用 1×1 卷积将前一层的通道数减少到原来的 1/4,大幅减少模型参数量并提升训

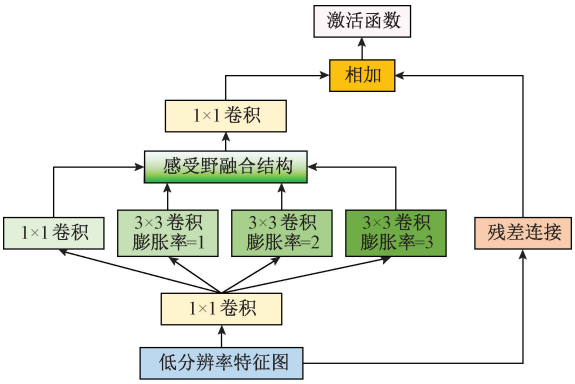
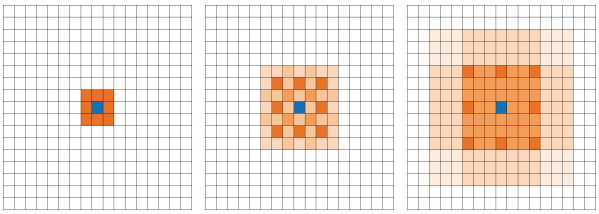


图 6 自适应感受野融合模块

Fig. 6 Adaptive fusion module of the reception field



(a) 膨胀率=1
(a) Dilation rate=1

(b) 膨胀率=2
(b) Dilation rate=1

(c) 膨胀率=3
(c) Dilation rate=3

图 7 不同膨胀率的空洞卷积示意图

Fig. 7 Schematic diagram of dilated convolution with various dilation rates

练效率,然后使用 1×1 卷积来调整特征图的通道数并对特征信息进行整合;如图 7 所示,在多分支感受野结构中,使用 4 个不同大小的卷积分支即 1 个 1×1 卷积和 3 个不同扩张率的空洞卷积在不增加计算参数的前提下扩展模型的感受野,实现了在更为充分利用特征图的上下文信息的同时增强模型的特征表示能力;在感受野融合结构中不同感受野的特征图被充分利用,该结构通过过滤掉特征图中的冗余信息并结合重要目标信息自适应性提高模型的特征表达能力,计算过程如式(3)所示。

$$\boldsymbol{Y}_{ij} = \boldsymbol{X}_{ij}^{0} \times \boldsymbol{A}_{ij} + \boldsymbol{X}_{ij}^{1} \times \boldsymbol{B}_{ij} + \boldsymbol{X}_{ij}^{2} \times \boldsymbol{C}_{ij} + \boldsymbol{X}_{ij}^{3} \times \boldsymbol{D}_{ij} \quad (5)$$

其中,$\boldsymbol{F}^{n}$ 代表多分支感受野结构各个分支的输出特征图,$\boldsymbol{X}_{ij}^{n}$ 代表特征图 $\boldsymbol{F}^{n}$ 在位置(i,j) 处所对应的特征向量。$\boldsymbol{W}_{ij}$ 代表特征图 $\boldsymbol{F}^{n}$ 的融合权重且 $\boldsymbol{W}_{ij} \in [\boldsymbol{A}_{ij},\boldsymbol{B}_{ij},\boldsymbol{C}_{ij},\boldsymbol{D}_{ij}]$,融合权重通过网络的自适应学习获得且在所有通道中共享。$\boldsymbol{Y}_{ij}$ 表示输出特征图在位置(i,j) 处的特征向量,征向量,融合权重定义和约束如式(6) ~ (8) 所示。

$$\boldsymbol{A}_{ij} + \boldsymbol{B}_{ij} + \boldsymbol{C}_{ij} + \boldsymbol{D}_{ij} = 1 \quad (6)$$

$$\boldsymbol{A}_{ij},\boldsymbol{B}_{ij},\boldsymbol{C}_{ij},\boldsymbol{D}_{ij} \in [0,1] \quad (7)$$

$$\boldsymbol{X}_{ij} = \frac{e^{\boldsymbol{\theta}_{M_{ij}}}}{e^{\boldsymbol{\theta}_{A_{ij}}} + e^{\boldsymbol{\theta}_{B_{ij}}} + e^{\boldsymbol{\theta}_{C_{ij}}} + e^{\boldsymbol{\theta}_{D_{ij}}}}, \boldsymbol{M} \in [\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\boldsymbol{D}] \quad (8)$$

其中,融合权重使用 Softmax 函数方式进行定义,加权标量映射 $\boldsymbol{\theta}_{A_{ij}}$、$\boldsymbol{\theta}_{B_{ij}}$、$\boldsymbol{\theta}_{C_{ij}}$ 和 $\boldsymbol{\theta}_{D_{ij}}$ 可通过反向传播自适应学习。

本文所提出的多尺度自适应特征融合模块通过特征融合网络的基础上添加可学习的模型参数并将自适应感受野融合模块嵌入至预测层中,实现了自适应地学习不同尺度上的特征信息,有效地降低了在各个尺度中因目标尺寸不一致而导致的错检率。相较于传统的 FPN 特征融合,该融合方法的特征融合的更加充分,特征表达能力更强且对于不同尺度目标的分类和回归更为精准高效。

1.4 注意力增强模块

实时性目标检测算法对于图像的处理效率较高,但在进行特征提取和目标框的回归时易受到复杂背景的影响导致检测性能下降,在复杂多变的轨道环境中该错检漏检异物等问题更为普遍。针对上述问题,LAM-Net 中加入了 AEM 轻量型通道注意力模块,通过重新分配输出通道的权值,降低了复杂背景对检测结果的影响和漏检率,在基本不影响检测速度的同时提升了轨道小目标异物的识别能力。

本文绘制了 AEM 轻量型通道注意力模块并对特征图进行可视化如图 8 所示,输入通道被平均分为两个分支,分支一使用 3×3 的卷积进行特征提取,分支二使用 5×5 的卷积提取特征获得更大的感受野,最后将取得不同感受野的特征图进行融合,在建立全局上下文关系的同时建模通道之间的相互依赖关系,实现自适应地重新校准通道的特征响应,使得各通道的权值更具代表性,以增加极少计算量为代价大幅度提升了轨道异物检测精度。AEM 注意力机制模块的计算过程如下所示:

$$\boldsymbol{s}_1 = g(\boldsymbol{F}) \quad (9)$$

$$\boldsymbol{s}_2 = \mathrm{ReLU}(\boldsymbol{W}_1\boldsymbol{s}_1) \quad (10)$$

$$\boldsymbol{s}_3 = Sigmoid(\boldsymbol{W}_2\boldsymbol{s}_2) \quad (11)$$

$$\boldsymbol{F} = \boldsymbol{s}_3\boldsymbol{F} \quad (12)$$

其中,$\boldsymbol{F} \in \boldsymbol{R}^{C\times H\times W}$ 表示输入特征图。函数 g 表示平均池化操作,$\boldsymbol{s}_1 \in \boldsymbol{R}^{C\times 1\times 1}$ 表示平均池化后的输出,$\boldsymbol{W}_1 \in \boldsymbol{R}^{C/16\times C}$ 表示全连接层 FC1 的参数,$\boldsymbol{s}_2 \in \boldsymbol{R}^{C/16\times 1\times 1}$ 表示经过 FC1 和 ReLU 函数处理后的输出,$\boldsymbol{W}_2 \in \boldsymbol{R}^{C\times C/16}$ 表示经过全连接层 FC2 的参数,$\boldsymbol{s}_3 \in \boldsymbol{R}^{C\times 1\times 1}$ 表示 FC2 和 *Sigmoid* 函数处理后的输出。特征图可视化结果表明,AEM 轻量型通道注意力模块在自适应特征融合的基础上有效地优化了神经网络模型的特征表达,能够使输出特征图的注意力更为集中,输出的通道信息更具代表性,进而大幅度提升轨道异物检测精度。

1.5 小型智能轨道异物检测样机

NVIDIA Jetson TX2 是一种基于 NVIDIA Pascal™ 架

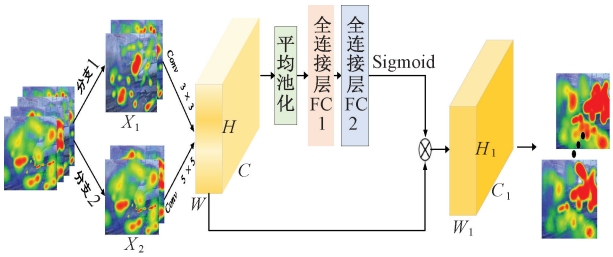
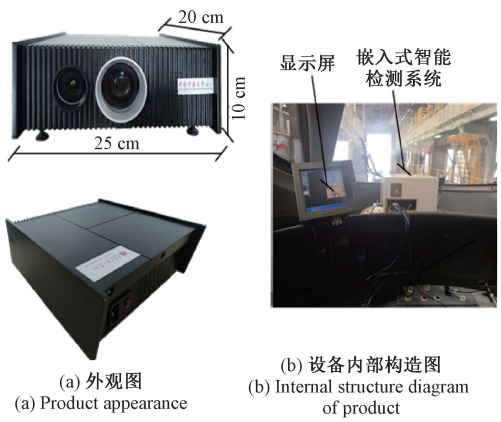


图 8　AEM 轻量型通道注意力模块结构图

Fig. 8　AEM lightweight channel attention module structure

构的 AI 单模块超级计算机,具有性能强大、外形小巧且功耗较低等优点,对于集成小型化智能终端设备非常有利。由于列车在实际轨道环境下运行时头部位空间较为有限,故能否将高性能算法嵌入至 NVIDIA Jetson TX2 中并设计出小型智能轨道异物检测样机对于实际轨道场景下异物的精准高效检测尤为重要。针对上述问题,本文研制了基于嵌入式的轨道入侵异物自主检测系统硬件平台并搭建了小型智能轨道异物检测样机如图 9 所示,主要包括 NVIDIA Jetson TX2 嵌入式处理器、近红外激光灯、可见光相机、电源、图像成像控制器以及语音报警器等,其设备长宽高分别为 25、20、10 cm,最大功耗仅为 20 W,能够在实际场景中低功耗、精准高效地检测轨道入侵异物。



(a) 外观图
(a) Product appearance

(b) 设备内部构造图
(b) Internal structure diagram of product

图 9　小型智能列车入侵异物检测样机

Fig. 9　Small intelligent train invasion of obstacles detection prototype

2　轨道入侵异物自主检测系统的组成

针对在弱光线轨道环境中捕获图像质量差难以满足模型训练需要、模型检测精度和速度在嵌入式部署过程中难以平衡无法满足实际应用需求等难点问题,本文设计了轨道异物自主检测系统,如图 10 所示。

图像采集层主要利用可见光相机和近红外激光灯采集列车前方的环境信息,其中近红外激光灯在弱光

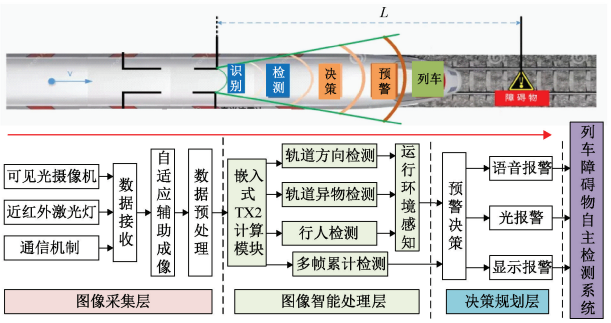


图 10　轨道异物自主检测系统组成

Fig. 10　Orbital foreign bodies independent of the detection system

线环境中主动开启进行补光,提高所捕获图像的质量;图像智能处理层主要以高性能嵌入式计算平台 TX2 为依托,在保证模型实时性的同时将 LAM-Net 嵌入式移植至 TX2 中有利于在实际轨道环境中高精度实时检测入侵异物;决策规划层主要对图像智能处理层输出的信息进行处理,并结合预警策略实现列车辅助驾驶功能。本文主要针对该系统图像智能处理层中模型的检测性能为研究重点,提出了一种轻量型线性自适应多尺度特征融合目标检测模型(LAM-Net),实现了在实际轨道环境中高精度、高鲁棒性实时检测轨道入侵异物,并通过大量的实验对比和结果分析验证了该异物检测模型的有效性。

3　实验结果和分析

为了评估该目标检测模型的有效性,本文在所采集的真实轨道数据集上进行了实验及测试。训练用服务器使用 4 块 NVIDIA GTX 1080TI 型号的高性能 GPU 和 Intel i7-6950X 型号的处理器,同时配备有 512 GB 固态硬盘、4 TB 机械硬盘和 64 GB 内存。设备软件运行环境主要采用 Pytorch 1.7.0 深度学习框架,系统环境 Ubuntu 18.04.5,编程语言采用 Python3.7,计算机视觉库 Opencv4.5.1。除上述服务器硬件外,本文还在小型智能轨道异物检测样机的嵌入式平台 NVIDIA Jeston TX2 中测试了模型的检测性能。

3.1　数据采集

为评估算法对铁路异物的检测性能,本文建立了真实环境下的轨道数据集,数据集示例如图 11 所示。该数据集是通过采集真实场景中的视频和图像来获取的,利用视频抽帧方法生成了 11 330 张图片,包含了不同天气、光照和复杂背景的轨道图像,能够满足模型训练需要并使得模型在真实轨道交通场景下更精准地检测异物。本文将收集到的图像分为 7 类进行标注:左转轨道、右转轨道、直行轨道、列车、行人、扳手、安全

帽。数据集中 70% 图像用来进行训练和验证,30% 图像用于测试。

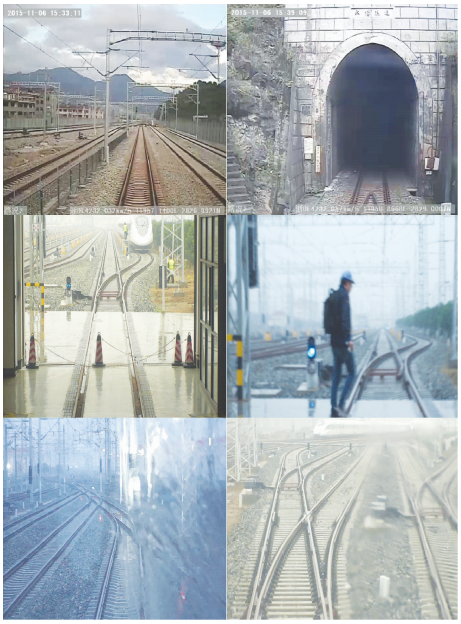


图 11 轨道数据集示例

Fig. 11 Samples of railway dataset

3.2 实验结果分析

基于卷积神经网络的模型训练和测试过程中包含大量的超参数设置,本文采用随机梯度下降法对模型进行优化,将初始学习率设置为 0.01,非学习率随着迭代次数的增长而降低,最大迭代次数为 100 轮,动量参数为 0.937,权重衰减为 0.000 5,批量大小为 16,输入分辨率为 416×416,非极大值抑制值域为 0.6,目标检测置信度值域为 0.45。为验证本文提出的目标检测方法对轨道场景下不同大小、不同天气、不同类别目标的检测效果,本文在轨道交通数据集上将所提出的自适应多尺度目标检测方法(LAM-Net)与鲁棒高精度的目标检测方法(YOLOv4、YOLOv5x、DSSD、DFF-Net[24]、ThunderNet-YOLOV3 等),以及实时轻量目标检测方法(YOLOXs[25]、YOLOv4-tiny、YOLOv5s、FR-Net-320[26]、LFD-Net[27]等)分别做了对比,实验结果展示于表 1 和 2 中。

表 1 的实验结果表明,LAM-Net 取得了 92.96% 的检测精度和 297 FPS 的检测速度,对于各类别入侵异物的检测精度较高,且检测速度明显优于其他模型,能够实现在实际轨道场景中精准高效地检测轨道异物。相比之下,Faster-RCNN 的检测速度极慢,低于表中展示的所有模型且小目标识别性能差,不能在实际轨道场景下精准检测小目标异物。SSD 的检测精度尤其是小目标识别能力相对 LAM-Net 也有较大差距。DSSD 相比于 SSD 模型的小目标识别能力得到了提升,但安全帽和扳手类检测精度仍比 LAM-Net 分别低 5.29% 和 4.03% 且实时性差。而且,LAM-Net 的检测精度和速度皆优于 DFF-Net、YOLOv3 和 ThunderNet-YOLOv3,能够在复杂多变的轨道环境中实现更为精准高效地目标检测。虽然 YOLOv4 和 YOLOv5x 的检测精度优于 LAM-Net,但 LAM-Net 更好地平衡了检测速度和精度,在保证精准检测异物的同时,检测速度是 YOLOv4 和 YOLOv5x 的 5 倍有余。除此之外,LAM-Net 表现出了更优的小目标检测能力,对于扳手类的检测精度可达 89.42%,较 YOLOv4 和 YOLOv5x 的 AP 值分别提升了 2.92% 和 1.51%。总之,LAM-Net 很好的平衡了检测速度和精度,即便对于安全帽和扳手这两类小目标也可以实现准确高效的实时检测,更符合实际轨道场景的实际应用需求。

表 1 鲁棒高精度模型结果比较

Table 1 Robust and high precision model results comparison

类别	SSD	DSSD	YOLOv3	YOLOv4	YOLOv5x	DFF-Net	ThunderNet YOLOV3	Faster-RCNN	LAM-Net
列车/%	90.20	91.85	92.28	94.01	**95.26**	89.96	87.01	90.43	93.22
行人/%	87.08	80.42	86.45	**94.41**	94.15	90.77	79.30	79.37	92.05
直行轨道/%	90.13	92.25	95.07	98.56	**98.95**	90.79	95.62	90.24	95.58
左转轨道/%	86.11	87.54	95.04	**95.90**	95.56	90.08	83.50	87.19	92.43
右转轨道/%	90.79	93.82	93.64	97.40	**97.67**	90.44	93.69	90.55	94.49
安全帽/%	86.99	88.24	86.95	93.18	**93.61**	90.83	86.53	70.62	93.53
扳手/%	83.68	85.39	86.17	86.50	87.91	87.99	78.84	34.17	**89.42**
平均精度/%	87.85	89.93	90.08	94.28	**94.73**	90.12	86.35	77.57	92.96
检测速度/FPS	47	13	67	55	42	54	20	10	**297**

表 2　实时轻量模型结果比较

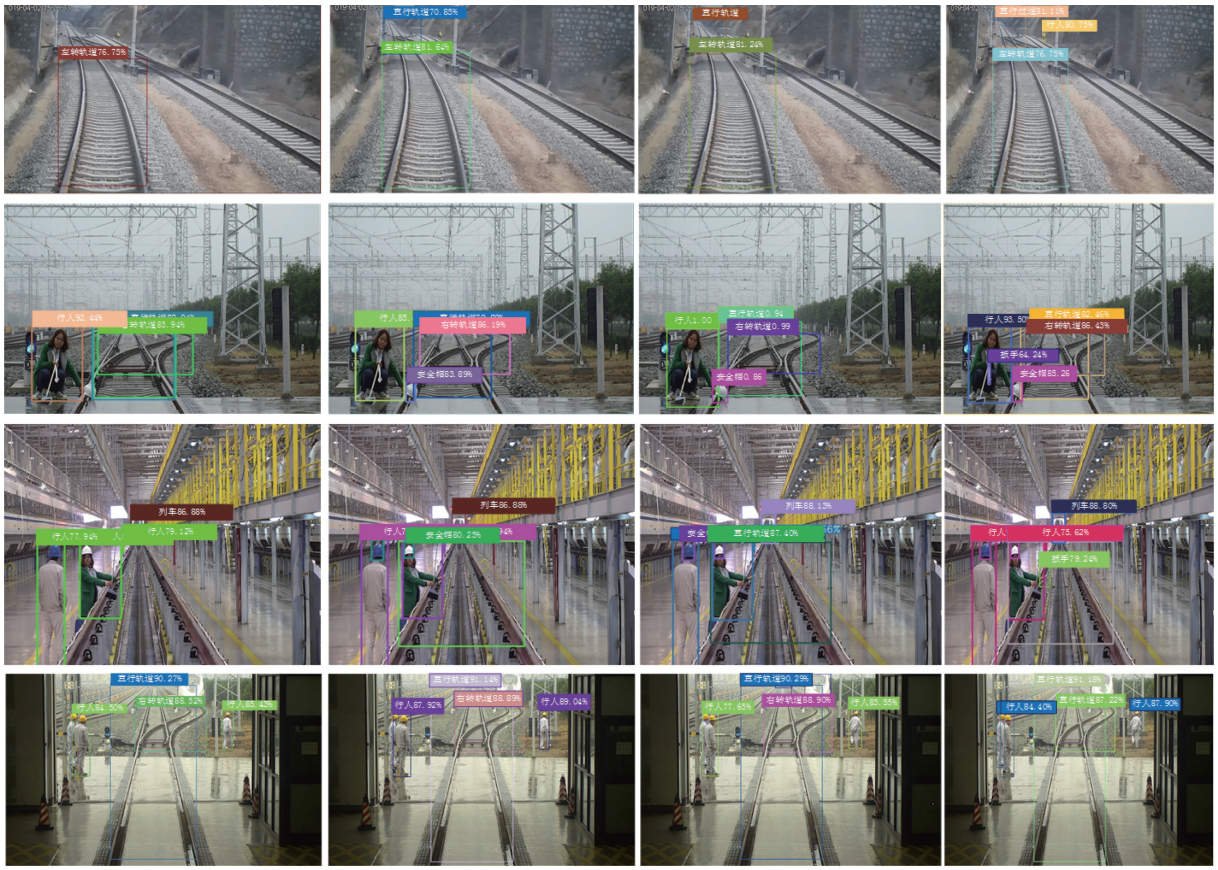
Table 2　Real-time lightweight model results

类别	YOLOv3-tiny	YOLOXs	YOLOv4-tiny	YOLOv5s	Mobilenetv2-YOLOv3	FR-Net	FB-Net	LFD-Net	LAM-Net
列车/%	84.08	91.82	88.97	92.24	87.93	90.50	90.62	90.92	93.22
行人/%	70.80	90.35	72.17	85.82	81.03	88.14	82.94	80.66	**92.05**
直行轨道/%	95.46	92.09	91.69	95.13	96.42	90.32	90.57	**97.04**	95.58
左转轨道/%	85.86	91.26	89.34	89.02	85.12	87.63	88.25	91.31	**92.43**
右转轨道/%	90.58	**96.58**	88.90	95.54	93.42	90.55	90.53	96.26	94.49
安全帽/%	88.30	92.72	88.27	88.98	79.14	86.67	87.85	87.58	**93.53**
扳手/%	76.60	89.04	77.37	77.46	56.75	81.33	82.60	88.91	**89.42**
平均精度/%	84.52	91.98	85.24	89.88	82.83	87.88	87.62	90.24	**92.96**
检测速度/FPS	312	113	**330**	102	75	72.1	82	209	297

表 2 的实验结果表明,LAM-Net 很好地平衡了轨道异物的检测速度和精度,在保证实时检测的前提下以 92.96% 的平均检测精度在实时轻量模型中取得最优。LAM-Net 的检测精度和速度皆优于 YOLOv5s、Mobilenetv2-YOLOv3、FR-Net 和 FB-Net,能够在轨道环境中更精准高效地识别出异物目标。虽然 YOLOv3-tiny 和 YOLOv4-tiny 的检测速度较快,但是它们的检测精度过低,尤其是对于小目标的检测精度差。LAM-Net 对于小目标安全帽和扳手的检测精度分别为 93.53% 和 89.42%,比 YOLOv3-tiny 高了 9.82% 和 8.32%,比 YOLOv4-tiny 高了 9.06% 和 8.25%,能够在复杂的轨道环境中精准定位并检测异物目标。而且,LAM-Net 能够鲁棒地检测行人目标并取得了 92.05% 的最优检测效果,此类目标虽然属于大尺度目标,但存在大面积遮挡且形态差异较大,为目标识别率和召回率等技术指标带来了巨大的挑战,实验结果表明其他模型在此类目标上的检测精度与 LAM-Net 差距较大。YOLOXs 和 LFD-Net 分别在右转轨道和直行轨道上取得了最优结果,但其他异物类别的检测精度与 LAM-Net 仍有存在差距,尤其是对于小目标的识别能力差距较大。LAM-Net 在多数目标中的检测精度都取得最优,尤其是小目标识别能力较其他模型提升巨大,在复杂背景下也可以保证极强的目标识别能力,可用于实际铁路场景下的实时入侵异物检测。

图 12 展示了本文所提出的方法与不同模型所检测的对比结果。所有模型均在轨道交通数据集上进行训练和测试。每一行代表一类典型场景的检测效果,每一列代表一类模型的检测效果,从左至右每一列依次为 FR-Net 检测结果、FB-Net 检测结果、YOLOv4-tiny 检测结果以及本文所提出的 LAM-Net 检测结果。其中,第 1 行展示的是列车实际运行过程中远距离动态目标检测效果,图像中的目标具有明显的尺度变化,FR-Net 仅能检测出近距离的大目标,而 FB-Net 和 YOLOv4-tiny 对远距离的小目标(行人)的检测存在漏检且检测精度较低,本文提出的 LAM-Net 对于不同尺度和距离的目标具有更好的检测效果。第 2 行展示的为静态场景下异物的识别效果,虽然目标距离采集装置较近但目标尺度相差较大,最终检测结果表明这几种算法都能准确地检测出大部分的目标,但前三者模型均存在小尺度目标(扳手、安全帽)的漏检情况,而本文提出的 LAM-Net 模型实现了所有异物的检测,具有较高的检测精度和召回情况,在多尺度目标位置回归方面更具优势。第 3 行选取室内列车运检环境作为模型输入,FR-Net 能够检测出特征明显的目标,但对于远处的行人和近处较大的轨道方向存在漏检问题,FB-Net 和 YOLOv4-tiny 能够检测出大多数目标,但由于远处的行人目标特征与背景较为相似,上述算法难以进行有效检测,而 LAM-Net 可以在不同场景下有效的对多尺度目标进行检测,展现出了更好的泛化性能。第 4 行选取具有遮挡情况下小目标的检测场景,场景中行人目标具有明显的遮挡关系且安全帽目标属于小目标范围,实验结果表明 FR-Net、FB-Net 和 YOLOv4-tiny 对于行人遮挡、安全帽等小目标的检测结果皆较差,而 LAM-Net 能检测出潜在的所有目标,特别是遮挡目标和小目标。上述检测结果可视化分析表明,LAM-Net 能够精确检测出目标的位置并预测其类别,特别是对于目标尺度变化较大的场景、遮挡目标以及小目标较多的场景具有较好的检测效果。

为验证 LAM-Net 的泛化性,本文使用该模型对其他场景进行了检测。如图 13 所示,LAM-Net 在多目标场景中仍旧保持着良好的目标检测性能,对检测目标较小的场景、昏暗光线下的场景、目标较密集且存在遮挡的场景也有着较高的检测精度。总之,泛化实验结果表明,LAM-Net 可以有效提高目标检测精度,能够在各种环境



(a) FR-Net测试结果
(a) FR-Net test result

(b) FB-Net测试结果
(b) FB-Net test result

(c) YOLOv4-tiny测试结果
(c) YOLOv4-tiny test result

(d) LAM-Net测试结果
(d) LAM-Net test result

图 12 不同模型测试结果对比

Fig. 12 Different model test results comparison

图 13 模型鲁棒性检验

Fig. 13 Model robustness test

下精准的定位并检测目标,并且其强大的泛化性能允许它被推广到其他场景中进行目标检测。

3.3 不同环境下的鲁棒性检验

本文分析了该模型在不同环境下的鲁棒性,如图 14 所示。虽然存在光照较暗、背景复杂和图像成像质量较差等问题,但是 LAM-Net 仍可以很好的检测不同类型的异物,以确保列车在雾天、黑夜和雨天等能见度较低的环境下安全运行(图 14(a)、(c)、(e)),特别是 LAM-Net 在复杂天气下保持着很强的多目标识别能力(图 14(a));在较为复杂的轨道交错的环境中仍可以很好的检测目标(图 14(b)),保证行车安全;在行人出现的场景中避免了安全帽这一小目标的漏检(图 14(d));当列车进入和驶出隧道时,尽管光线条件会发生突变,但是该网络依旧可以很好地检测到铁路方向从而确保了列车安全运行(图 14(f)),即便是在昏暗弱光线的环境中 LAM-Net 的铁路方向检测能力仍旧很强。鲁棒性实验结果表明,LAM-Net 能够有效地检测出目标的位置,能够实现不同场景下的目标检测并具有较高的鲁棒性,即便在昏暗环境和存在背景干扰的轨道环境中仍有着较好的检测性能,满足恶劣环境下列车辅助驾驶的要求,能够保证列车的安全运行。

(a) 雾天
(a) Foggy days

(b) 复杂交错轨道
(b) Complex staggered Railway tracks

(c) 夜晚
(c) Nights

(d) 小目标存在的场景
(d) Scenarios with small objects

(e) 雨天
(e) Rainy days

(f) 隧道
(f) Tunnel

图 14　不同场景的鲁棒性测试

Fig. 14　Model robustness test of different scenarios

3.4　消融实验结果对比

为更好地理解本文提出的各结构的有效性,本文设计了多组消融实验来对比不同结构对目标检测性能的影响,分别从检测精度和检测速度两个方面来对比模型性能,如表 3 所示。YOLOv4-tiny 作为基准测试模型,检测结果用来作为对比基准;AM-Net 由自适应特征融合模块和注意力机制模块组成,MA-Net 由 Moblienetv3、自适应特征融合模块和注意力机制模块组成,目的是验证特征提取模块对提升检测速度的有效性;LM-Net 由轻量型特征提取模块和注意力机制模块组成,以验证自适应特征融合模块对提高检测精度的有效性;LA-Net 由轻量型特征提取模块和自适应特征融合模块组成,以验证注意力机制模块对进一步提升目标检测精度的有效性。

表 3　消融实验结果对比

Table 3　Comparison of ablation results

子模型名称	YOLOv4-tiny	AM-Net	LM-Net	LA-Net	MA-Net	LAM-Net
轻量型特征提取模块	–	–	√	√	–	√
Moblienetv3	–	–	–	–	√	–
自适应特征融合模块	–	√	–	√	√	√
注意力机制模块	–	√	√	–	√	√
列车/%	88.97	92.38	89.12	88.72	92.73	**93.22**
行人/%	72.17	**93.17**	77.06	90.61	91.60	92.05
直行轨道/%	91.69	95.85	96.02	**96.42**	92.51	95.58
左转轨道/%	89.34	92.27	91.10	90.29	92.19	**92.43**
右转轨道/%	88.90	**96.01**	93.98	93.17	93.05	94.49
安全帽/%	88.27	93.36	81.27	92.25	92.81	**93.53**
扳手/%	77.37	89.15	85.16	88.34	87.36	**89.42**
检测精度/%	85.24	93.17	87.53	91.40	91.75	92.96
检测速度/FPS	330	261	**332**	310	283	297

首先,将 AM-Net 中的特征提取模块替换为本文所设计的轻量型特征提取模块,与 AM-Net 相比,LAM-Net 以牺牲极小精度为代价使得检测速度提升了 36 FPS。此外,本文将 MA-Net 中的 Moblienetv3 替换为轻量型特征提取模块。与 MA-Net 相比,LAM-Net 的平均检测精度和检测速度分别提升了 1.21% 和 14 FPS,尤其是对于小目标轨道入侵异物安全帽和扳手的检测精度分别提升了 0.72% 和 2.06%,更利于在复杂多变的轨道环境中实现高精度实时轨道入侵异物检测。实验结果证明轻量型特征提取模块可以在几乎不降低检测精度的前提下大幅度提升模型的检测速度,实现在复杂的轨道环境中快速高效地检测轨道异物。其次,将 LM-Net 中的特征融合模块替换为自适应特征融合模块,与 LM-Net 相比,检测精度提高了 5.43%,对于安全帽和扳手类小目标的检测精度提升了 12.26% 和 4.26%。这些数据表明,该模块采用分享权重的方式自适应融合各个尺度的特征并将多感受野模块嵌入至预测层中的方式有利于提升模型的检测精度。最后,将注意力机制模块加入至 LA-Net,使得平均检测精度提升了 1.56%,且提升了除直行轨道外的所有异物类别的检测精度。最后,实验数据表明,注意力机制模块更好地分配了不同通道的权值,突出目标更重要的特省信息,在基本不增加计算负荷的前提下使得检测精度进一步提高。总之,LAM-Net 在 6 种类别中取得了最好的成绩。尤其是

对于安全帽、扳手等小物体检测的 AP 值较 YOLOv4-tiny 分别提高了 9.06% 和 8.25%，而且所提 3 个模块的组合使用使得 LAM-Net 在检测速度和精度之间取得了良好的平衡，能够满足实际轨道场景下检测异物的应用需求。

3.5 模型在嵌入式设备中的性能测试

为测试轨道交通列车前方典型目标检测系统在真实运行场景中的性能，本文将研制的图像采集装置、目标检测算法、嵌入式计算平台以及预警控制策略相结合搭建基于嵌入式的轨道异物识别系统。本文选取的嵌入式平台为 NVIDIA Jetson TX2，配有 NVIDIA Pascal™ GPU、高达 8 GB 内存、59.7 GB/s 的显存带宽，提供各种标准硬件接口，其核心模组如图 15 所示。在上述嵌入式平台中配置了目标检测所需的深度学习环境，实现典型目标检测系统实时运行。所采用的软件包版本为 Jetpack4.4，系统版本为 Ubuntu 18.04，Cuda 软件版本为 10.2，Cudnn 软件版本为 8.0.0，opencv 软件版本为 4.1.1，Tensorrt 软件版本为 7.1.3.0，Deepstream 软件版本为 5.0，Python 软件版本为 3.6。

本文将多个轻量型目标检测模型部署至嵌入式设备中，并采用检测精度和检测速度对不同模型进行评估，如表 4 所示。实验结果表明，LAM-Net 实现了 92.96% 的检测精度和 32 FPS 的检测速度，检测精度和速度皆优于 YOLOv5s、YOLOXs、FB-Net 和 LFD-Net，能够在复杂的轨道环境中更精准高效地检测异物目标。YOLOv4-tiny 的检测速度较快，但是其检测精度过低，无法保证在复杂的轨道环境中精准地检测异物。与 YOLOv4-tiny 相比，LAM-Net 在降低 3 FPS 的检测速度前提下，使得检测精度获得 7.72% 的大幅提升。YOLOv5x 和 YOLOv4 对异物的检测精度较高，但检测速度过慢，无法实时检测异物目标。综上分析，LAM-Net 在嵌入式平台中仍旧能很好地平衡检测速度和精度，满足在复杂多变的轨道环境中精准高效检测异物的需求。

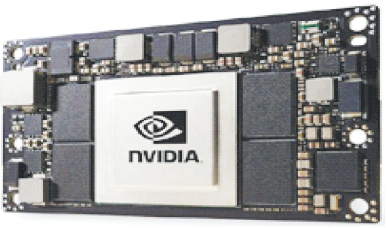


图 15 NVIDIA Jetson TX2 嵌入式平台

Fig. 15 NVIDIA Jetson TX2 embedded platform

表 4 不同模型嵌入式测试结果

Table 4 Embedded test results of different models

评价指标	YOLOv5s	YOLOv5x	YOLOv4	YOLOv4-tiny	YOLOXs	LFD-Net	FB-Net	LAM-Net
检测精度/%	89.88	94.73	94.28	85.24	91.98	90.24	87.62	92.96
检测速度/FPS	14	4	6	35	16	23	21	32

4 结 论

本文提出了一种轻量型自适应多尺度卷积神经网络目标检测模型 LAM-Net，并将其应用于复杂环境下的轨道异物检测中。特征提取模块使用线性变换和深度可分离卷积降低了模型计算复杂性大幅度提升检测速度；特征融合模块可以自适应地学习不同感受野分支的融合权重，实现了不同尺度特征图的自适应融合，有效地提升了检测精度；通道注意力机制模块在不影响检测速度的使得轨道异物的检测精度进一步提升。

在真实场景下的轨道数据集上的实验结果表明，LAM-Net 的检测速度和精度分别 297 FPS 的检和 92.96%，实现了检测速度和精度的良好平衡。LAM-Net 在牺牲少量检测精度的前提下，检测速度分别达到了 YOLOv5s 和 YOLOv4 的 5.4 倍和 7.1 倍。在检测精度方面，LAM-Net 的平均检测精度比 YOLOv4-tiny 高 7.72%，尤其是对于小目标安全帽和扳手的检测精度分别比 YOLOv4-tiny 高了 9.06% 和 8.25%，能够更好地在复杂的轨道环境中精准高效地检测小目标异物，确保行车安全。

针对于 LAM-Net 训练出的模型进行了泛化性实验和鲁棒性检验。泛化性实验结果表明，该模型具有很强的泛化能力，即便在除轨道外的其他复杂环境中也能精准的检测目标；鲁棒性检验结果表明，该模型在昏暗、雨天、雾天等复杂轨道环境中仍旧保持着较高的检测精度，实现了在复杂轨道环境中快速鲁棒检测异物。

结合 NVIDIA JETSON TX2 计算平台研制了小型智能轨道异物自主检测系统。结果表明，LAM-Net 在系统中的检测精度和检测速度分别达到了 92.96% 和 32 FPS，优于现有的目标检测模型，验证了异物自主检测的有效性和实时预警系统的可行性，能够满足在轨道环境下对列车前方异物自主检测的应用需求。

本文将模型部署在 NVIDIA JETSON TX2 中进行轨道入侵异物检测,若结合 TensorRt 加速将进一步提升算法的检测性能,这将是未来继续深入的研究内容。

参考文献

[1] GUAN M Y, WEN C Y, ZOU C L, et al. Real-time event-triggered object tracking in the presence of model drift and occlusion[C]. IEEE Transactions on Industrial Electronics, 2018: 2054-2065.

[2] 王洋,余祖俊,朱力强,等. 基于 CNN 的高速铁路侵限异物特征快速提取算法[J]. 仪器仪表学报,2017,38(5):1267-1275.
WANG Y, YU Z J, ZHU L Q, et al. Fast feature extraction algorithm for high speed railway clearance intruding objects based on CNN [J]. Chinese Journal of Scientific Instrument, 2017,38(5): 1267-1275.

[3] 钟明轩,乐逸祥,周磊山. 高铁站列车和调车作业计划一体化编制方法[J]. 交通运输系统工程与信息, 2021, 21(3): 176-186.
ZHONG M X, YUE Y X, ZHOU L SH. Scheduling method for integrated plan of train operation and shunting operation in high-speed railway station[J]. Journal of Transportation Systems Engineering and Information Technology, 2021, 21(3): 176-186.

[4] GARCIA J, URENA J, HERNANDEZ A, et al. Efficient multisensory barrier for obstacle detection on railways[J]. IEEE Transactions on Intelligent Transportation Systems, 2010, 3(11): 702-713.

[5] CATALANO A, BRUNO F, PISCO M, et al. Intrusion detection system for the protection of railway assets by using fiber Bragg grating sensors: A case study[C]. Third Mediterranean Photonics Conference, Trani, 2014: 1-3.

[6] NANASONE R, NAGAMINE N, UKAI M, et al. Frontal obstacle detection using background subtraction and frame registration [J]. Quarterly Report of RTRI, 2017, 58(4): 298-302.

[7] TENG Z, LIU F, ZHANG B P. Visual railway detection by superpixel based intracellular decisions [J]. Multimedia Tools and Applications, 2016, 75 (5): 2473-2486.

[8] LIU H C, LI S T. Decision fusion of sparse representation and support vector machine for SAR image target recognition[J]. Neurocomputing, 2013, 113 (1): 97-104.

[9] 傅成红,杨书敏,张阳. 改进支持向量回归机的短时交通流预测[J]. 交通运输系统工程与信息, 2019, 19(4): 130-134.
FU CH H, YANG SH M, ZHANG Y. Promoted short-term traffic flow prediction model based on deep learning and support vector regression [J]. Journal of Transportation Systems Engineering and Information Technology, 2019, 19(4): 130-134.

[10] GIRSHICK R, DONAHUE J, DARELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition, Columbus, 2014: 580-587.

[11] GIRSHICK R, REN S, HE K M, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[12] 石杰,周亚丽,张奇志. 基于改进 Mask RCNN 和 Kinect 的服务机器人物品识别系统[J]. 仪器仪表学报, 2019,40(4):216-228.
SHI J, ZHOU Y L, ZHANG Q ZH. Service robot item recognition system based on improved Mask RCNN and Kinect [J]. Chinese Journal of Scientific Instrument, 2019,40(4):216-228.

[13] GE Z, LIU S, WANG F, et al. Yolox: Exceeding YOLO series in 2021[J]. ArXiv Preprint, 2021, ArXiv: 2107. 08430.

[14] WANG C, BOCHKOVSKIY A, LIAO H M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J]. ArXiv Preprint ,2022, ArXiv: 2207. 02696.

[15] RDEMON J, FARHADI A. YOLOv3: An incremental improvement [J]. ArXiv Preprint, 2018, ArXiv: 1804. 02767.

[16] BOCHKOVSKIY A, WANG CY, LIAO M Y. YOLOv4: Optimal speed and accuracy of object detection [J]. ArXiv Preprint, 2020, ArXiv: 2004. 10934.

[17] 尹安东,张万兴,赵韩,等. 基于神经网络的磷酸铁锂电池 SOC 预测研究[J]. 电子测量与仪器学报, 2011, 25(5):433-437.
YIN AN D, ZHANG W X, ZHAO H, et al. Research on estimation for SOC of LiFePO_4 Li-ion battery based on neural network [J]. Journal of Electronic Measurement and Instrumentation, 2011,25(5):433-437.

[18] LIU W, ANGUELOV D, ERHAN D. SSD: Single shot multibox detector [C]. European Conference on Computer Vision (ECCV), Germany, 2016: 21-37.

[19] 李晖晖,周康鹏,韩太初,等. 基于 CReLU 和 FPN 改进的 SSD 舰船目标检测[J]. 仪器仪表学报,2020,41(4):183-190.
LI H H, ZHOU K P, HAN T CH, et al. Ship object detection based on SSD improved with CReLU and FPN[J]. Chinese Journal of Scientific Instrument, 2020,41(4):183-190.

[20] ZHANG H. Real-time detection method for small traffic signs based on YOLOV3 [J]. IEEE Access, 2020, 12(8): 64145-64156.

[21] LI J, ZHOU F Q, YE T. Real-world railway traffic detection based on faster better network [J]. IEEE Access, 2018, 17(6): 68730-68739.

[22] KONG T, YAO A, CHEN Y, et al. HyperNet: Towards accurate region proposal generation and joint object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 845-853.

[23] WOMG A, LI F, CHWYL B, et al. Tiny SSD: A tiny single-shot detection deep convolutional neural network for real-time embedded object detection[C]. Conference on Computer and Robot Vision (CRV), Toronto, 2018: 95-101.

[24] WEI X, SUO D, LI Y. Multi-target defect identification for railway track line based on image processing and improved YOLOv3 model [J]. IEEE Access, 2020, 11(8): 61973-61988.

[25] YE T, ZHANG X, ZHANG Y, et al. Railway traffic object detection using differential feature fusion convolution neural network [J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 22 (3): 1375-1387.

[26] WANG R J, LI X, LING C X. Pelee: A real-time object detection system on mobile devices[J]. ArXiv Preprint, 2020, ArXiv: 1804.06882.

[27] YE T, WANG B C, PING S, et al. Automatic railway traffic object detection system using feature fusion refine neural network under shunting mode[J]. Sensors, 2018, 18(6):1916-1935.

[28] YE T, REN C, ZHANG X, et al. Application of lightweight railway transit object detector [J]. IEEE Transactions on Industrial Electronics, 2021, 68(10): 10269-10280.

作者简介

叶涛(通信作者),2009 年于中国矿业大学获得学士学位,2012 年于中国矿业大学(北京)获得硕士学位,2016 年于北京航空航天大学获得博士学位,现为矿业大学(北京)机械电子与信息工程学院的高级工程师。主要研究方向为深度学习、目标检测和轨道交通等。
E-mail: ayetao198715@163.com

Ye Tao (Corresponding author) received his B. Sc. degree from China University of Mining and Technology in 2009, received his M. Sc. degree from China University of Mining and Technology (Beijing) in 2012, and received his Ph. D. degree from Beijing University of Aeronautics and Astronautics in 2016. He is currently a senior engineer in the School of Mechanical Electronics and Information Engineering at University of Mining and Technology (Beijing). His main research interests include deep learning, objection detection and rail transit.

赵宗扬,2020 年于中国矿业大学(北京)获得学士学位,现为中国矿业大学(北京)机械与电子工程学院在读研究生。主要研究方向为深度学习、目标检测和轨道交通等。
E-mail:303616426@qq.com

Zhao Zongyang received his B. Sc. degree from China University of Mining and Technology (Beijing) in 2020. He is currently a graduate student in the School of Mechanical and Electronic Engineering at China University of Mining and Technology (Beijing). His main research interests include deep learning, object detection and rail transit.

郑志康, 2021 于武汉科技大学获得学士学位,现为中国矿业大学(北京)机械与电子工程学院在读研究生。主要研究方向为深度学习、目标检测和轨道交通等。
E-mail:2980572453@qq.com

Zheng Zhikang received his B. Sc. degree from Wuhan University of Science and Technology in 2021. He is currently a graduate student in the School of Mechanical and Electronic Engineering at China University of Mining and Technology. His main research interests include deep learning, object detection and rail transit.