

### 3.6.3 卷积神经网络训练

车载信息识别方面，通过加载经大型自然数据集 ImageNet 预训练过的深层 AlexNet 模型开展相关训练工作。这里，同样在 Lane-1 车道与 Lane-2 车道试验所测得数据库中随机选取 70% 的样本作为训练集，并将其余 30% 的样本作为测试集进行准确率评估。参数设置方面，则与前述相关迁移学习训练试验保持一致。测试过程中的函数变化曲线如图 3.27 所示，可发现经训练 4 个周期后，模型在训练集上的准确率开始在由 50% 逐步上升至 90% 左右，且相应的损失函数值则由 0.9 逐步降低至 0.1 左右；而在迭代约 30 个周期后，该准确率曲线开始逐渐收敛至 100% 左右，相应的损失函数值则稳定在 0.0002 附近。最终，将所设置的图像测试集输入至该模型中进行精度评估测试，发现试验 1~5 的平均识别精度约为 100%，表明由大型自然图像集 ImageNet 预训练过的 AlexNet 模型，不仅可以通过迁移学习的训练方式实现对数字信号图像的分类识别，而且对车辆行驶轨道识别问题可达到较为理想的辨识效果。

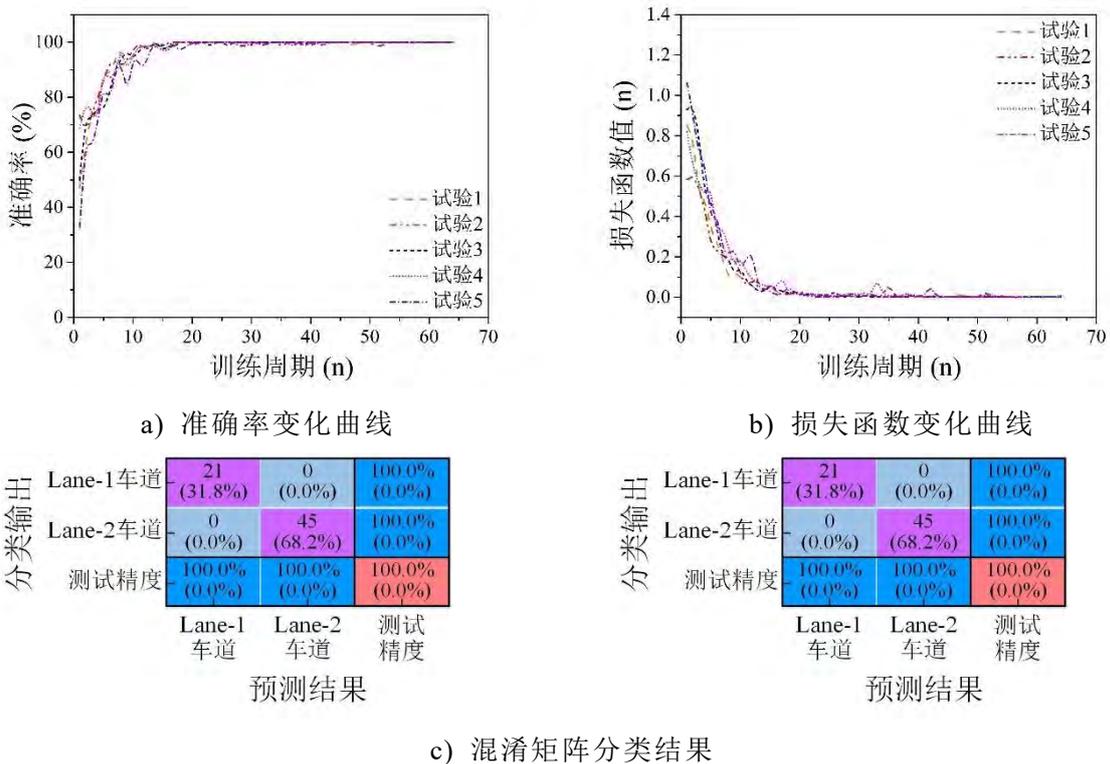


图 3.27 DCNN 网络模型训练结果

为进一步提高 DCNN 网络模型的性能，针对所建立形成的各有效测试样本，对其分别采用如表 3.3 所示的图像变换技术对数据库规模进行扩充，并利用扩充后的样本库对该神经网络模型进行训练。这里，试验所用训练/测试集及相关模型训练参数均与前述保持一致。试验过程中的函数变化曲线如图 3.28 所示，可发现经迭代约 30 个训练周期后，该模型的准确率曲线开始由 50% 逐渐上升至 90% 左

右，相应的损失函数值则由 1.5 逐渐下降至 0.2 附近；在迭代约 200 个周期后，其准确率曲线开始逐渐收敛至 100%左右，相应的损失函数值也稳定在 0.001 附近。最终，将所设置的测试数据集导入至该 DCNN 网络模型中，用以对训练结果进行精度评估。这里，可发现五次随机试验的平均识别精度达到了 99.82%，表明可利用图像变换技术对数据库样本容量进行扩充的方式提高对车辆行驶轨道的判断识别效果。

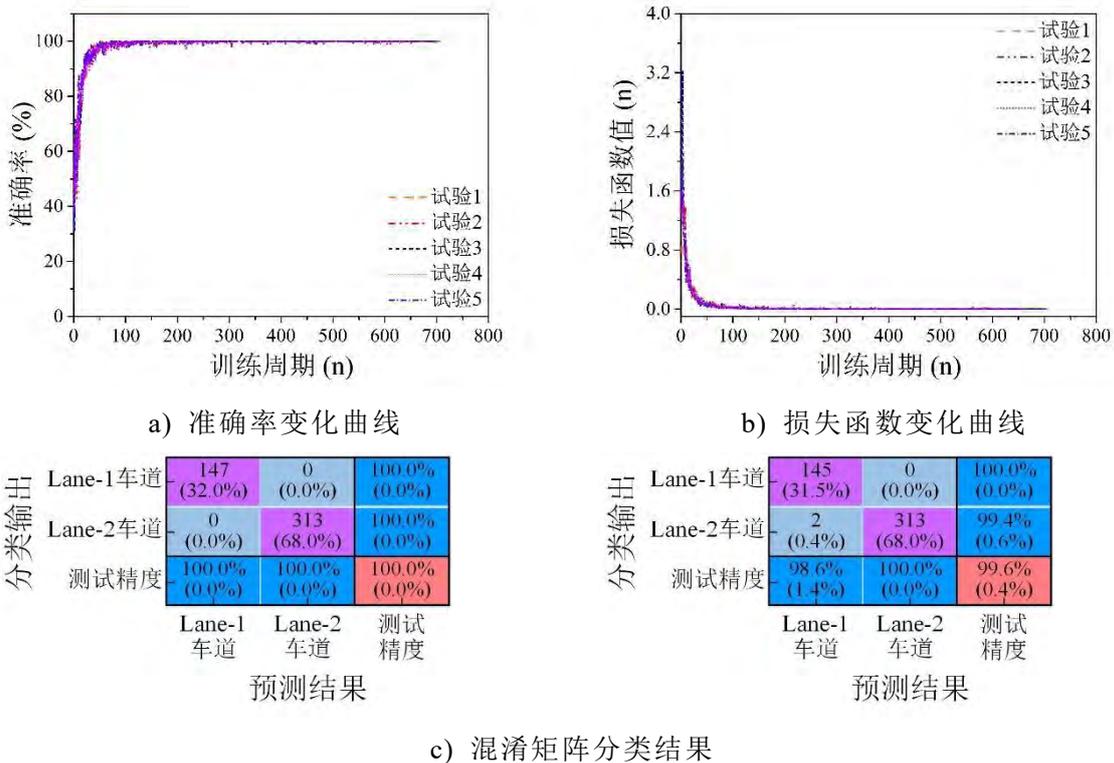


图 3.28 DCNN 网络模型优化训练结果

### 3.7 车辆荷载信息综合检测

#### 3.7.1 试验测试方案设计

经前述试验测试结果可知，所提出的基于深度学习理论及结构特征响应信息提取的车辆荷载信息识别技术具有较好的可行性。这里，考虑到作用在结构上的单个外界荷载本质上包含有一系列特定的参数(如特定的车载、车速、车型等)，如何由单条结构响应实现对外部荷载信息的综合识别是一个值得进行探讨的科学问题。这里，通过将经前述单独训练过的车载分类器进行串联整合，提出了一种对车辆荷载信息进行综合检测的技术手段。具体而言，针对由车辆荷载作用引起的单条结构响应，在对其完成时-频分析的基础上，分别利用前述训练所得的车载 DCNN 分类模型、车速 DCNN 分类模型、车型 DCNN 分类模型及车道 DCNN 分类模型，对相应的车辆信息进行依次检测识别，如图 3.29所示。

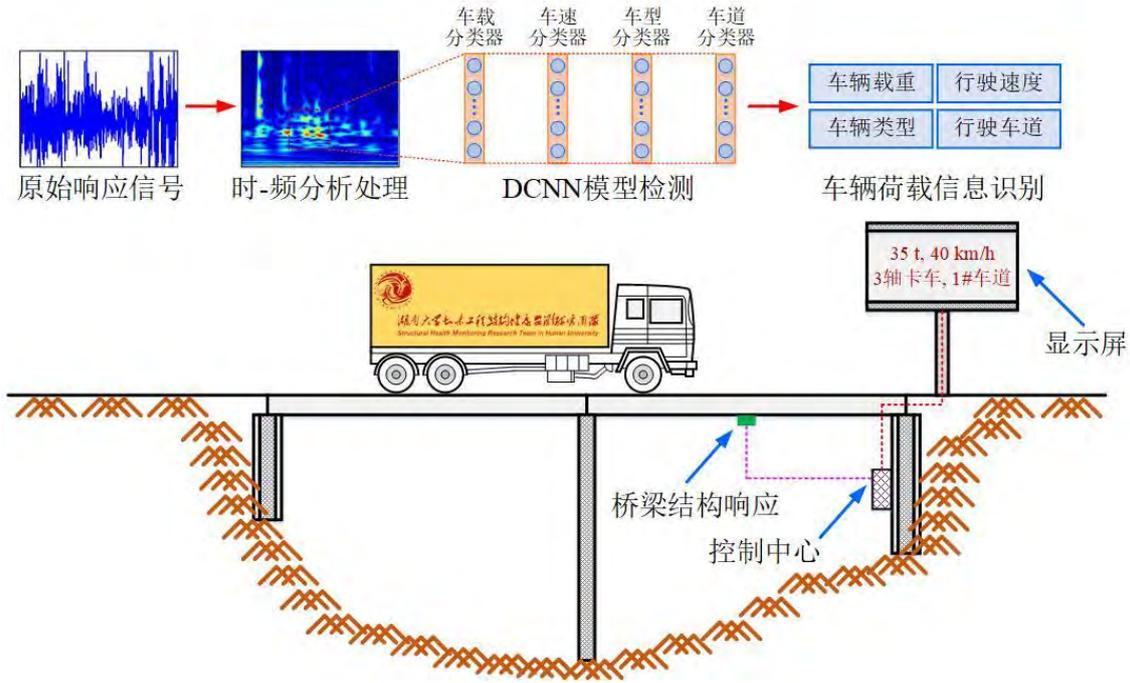


图 3.29 车辆荷载信息综合检测系统示意图

测试样本选取方面,针对前述章节 3.3~3.6 所建立的各结构特征响应数据库,随机抽取 80 条结构响应信息进行分析,将其依次编号为工况 1~80。各样本所对应的相关车辆荷载信息如图 3.30 所示,其中测试工况 1~60 及 71~80 所使用的车辆模型为三轴车,工况 61~70 则为五轴车模型;同时,测试工况 1~70 均行驶在 Lane-1 车道上,工况 71~80 则行驶在 Lane-2 车道上。

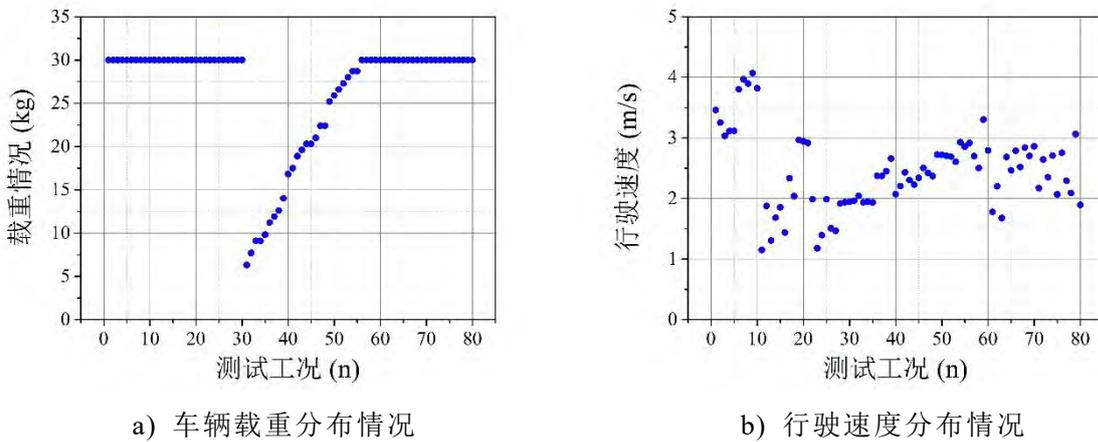


图 3.30 测试样本车辆荷载信息分布情况

### 3.7.2 车辆荷载信息识别结果

针对由时域加速度信号转换所得的各二维小波时-频功率图,分别将其输入至前述经训练过的车载、车速、车型及车道检测 DCNN 模型中,用以实现对车辆荷载信息的综合检测识别。相应的试验测试结果如图 3.31 所示,可发现神经网络模型对这四类车辆荷载信息的识别精度分别约为 58.80%、90.00%、97.50%和 90.00%,

表明所提方法能较好地实现对车辆荷载信息的全面检测识别,且具有方便快捷、成本低廉等优势。但值得注意的是,当车辆行驶速度较高时,该方法对车辆实际载重情况的检测效果有待提高。在后续研究过程中,可通过增加训练数据库样本规模的方式进行改善。除此之外,在实际应用中还可通过制定相关行驶规则(如限速等)等方式,对该技术的应用效果进行辅助提高。

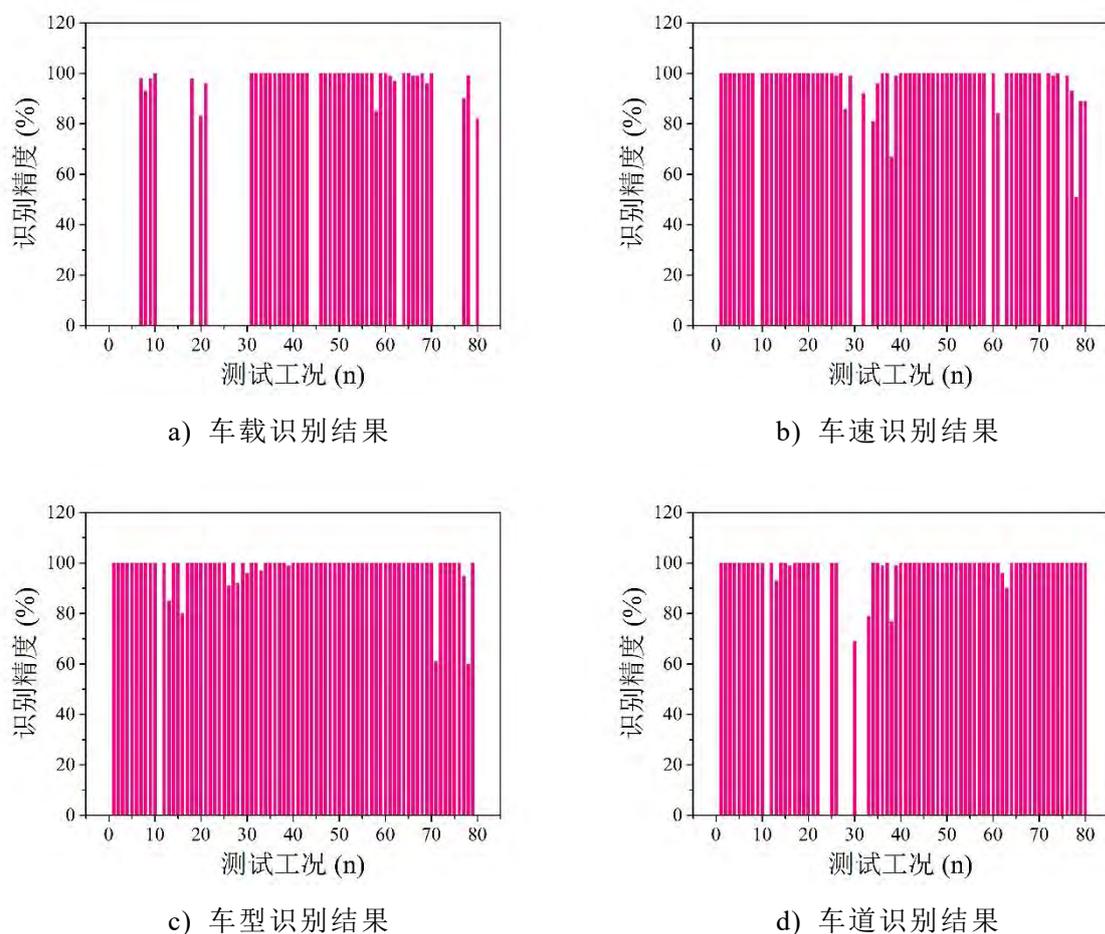


图 3.31 车辆荷载信息检测识别结果

## 3.8 野外实桥测试

### 3.8.1 工程概况

在完成实验室缩尺车-桥振动测试、获得较为理想测试效果的基础上,拟计划通过开展实际情况更为复杂、不确定性干扰因素更多的野外试验测试,对所提基于深度学习理论和结构特征响应提取的车载信息识别方法的有效性进行评估。

通过选取某野外桥梁进行了现场测试。试验过程中,所测试桥梁结构为钢筋混凝土连续梁式桥,其采用 C40 强度等级混凝土浇筑而成,如图 3.32 所示。桥梁全长约 140m,桥宽为 8m。结构采用等截面钢筋混凝土箱梁;同时为防止弯扭失稳,设置了小跨径箱梁,其对应的梁高约为 1.6m。当桥墩高度大于 30m 时,为

防止高墩失稳，采用薄壁箱型空心墩，其截面尺寸为 4.0m×2.2m；相应的基础则为桩基承台形式，且桩径设置为 2.0m。当桥墩高度小于 30m 时，则采用实心花瓶墩。承台处桥墩截面尺寸为 2.5m×1.6m，墩顶变化段桥墩截面尺寸为 4.0m×1.6m。基础为桩基承台形式，其中桩径为 2.0m。该桥梁结构采用“桩柱式”桥台，桥台桩基为端承桩，其桩径设置为 1.5m。该桥梁结构的运营车辆限制速度规定不超过 20km/h。

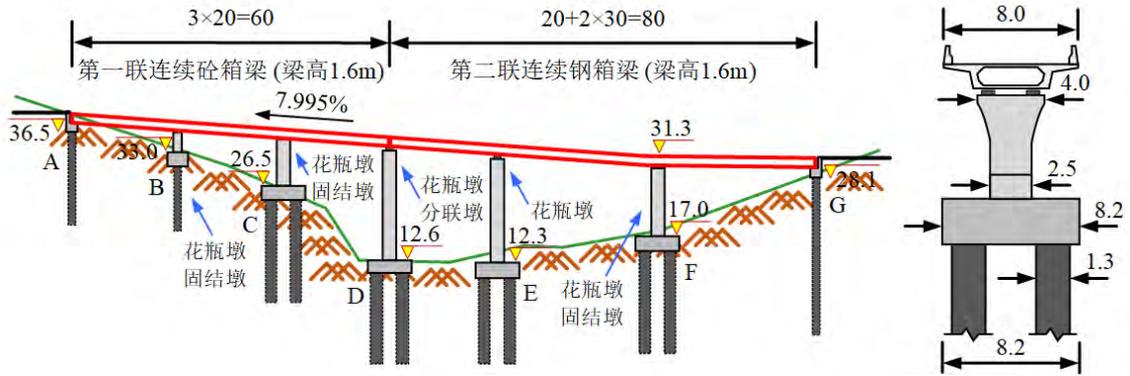


图 3.32 野外测试所用桥梁结构示意图

### 3.8.2 测试方案

针对该结构特性及现场测试环境条件，拟计划通过设置不同的车辆荷载工况，通过开展基于深度学习理论及结构特征响应提取的车载识别方法，实现对车辆荷载的反问题识别。由于该桥梁车道全长为一段 S 型曲线，为避免试验过程中加载卡车在弯道处发生碰撞等意外事故发生，故仅针对直线段 DE 部分开展了系列试验测试。其测试桥梁段结构全长约 21.60m，车道宽度约 7.07m。试验过程中，设置了两种不同重量的载重卡车进行跑车试验，如图 3.33 所示。所用测试车辆均为 3 轴 10 轮重型卡车，车辆全长约 8.14m，宽度约为 2.42m。其中，车辆前轴为单轴两车轮，后两轴均为单轴四车轮。同时，车辆前轴距车头外侧最边缘的水平距离约为 1.5m，最后一轴距车尾最边缘约为 1.5m。其相关参数如表 3.6 所示。



a) 侧视图

b) 主视图

图 3.33 试验测试所用加载卡车

表 3.6 试验所用卡车参数取值

车辆编号	总重 $W/t$	轴距 $AS_1/m$	轴距 $AS_2/m$
卡车 A	17.02	3.80	1.36
卡车 B	35.06	3.80	1.36

试验过程中采用 Wilcoxon-799LF 型加速度传感器进行信号采集,其量程为 $\pm 10g$ ,灵敏度为 $500mV/g$ ,频响范围为 $0.1\sim 2500Hz$ ,如图 3.34(a)所示。仪器设备方面,利用 LMS-Cadax-8 便携式动态信号采集仪,对试验过程中桥梁结构所产生的加速度信号进行实时采集,如图 3.34(b)所示。



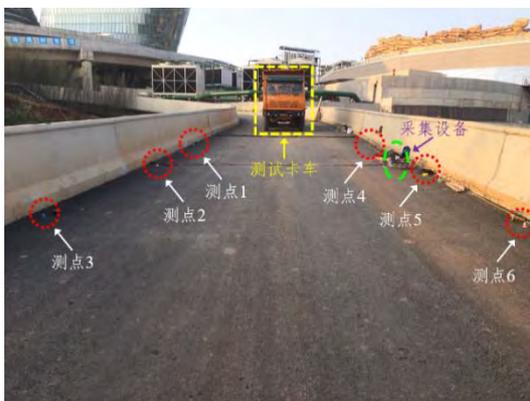
a) 加速度传感器



b) LMS 采集仪

图 3.34 试验测试所用仪器设备

监测传感器布设方面,对测试所用桥梁结构 DE 段部分的 $1/4$ 跨、 $1/2$ 跨及 $3/4$ 跨等位置处,分别在桥面板两端布置有一个加速度传感器,并依次编号为测点 1~6。其中,传感器通过耦合剂稳定地沿垂直方向布置于桥面上,且其安装位置应尽可能地贴近桥面板端部边缘。试验测试时的信号采样频率设置为 $1024Hz$ 。具体在试验测试过程中,规定载重卡车由桥梁结构 D 段匀速驶向 E 段。这里,由于该桥面限速为 $20km/h$ ,故设置载重卡车在 $10\sim 20 km/h$ 速度范围内随机进行跑车测试。测试过程中,同一时刻仅允许桥梁结构上行驶有一辆载重卡车。同时,对于每辆载重卡车,依次规定其进行 40 次跑车测试。现场测试情况如图 3.35 所示。



a) 测点布置情况



b) 测点布置详情



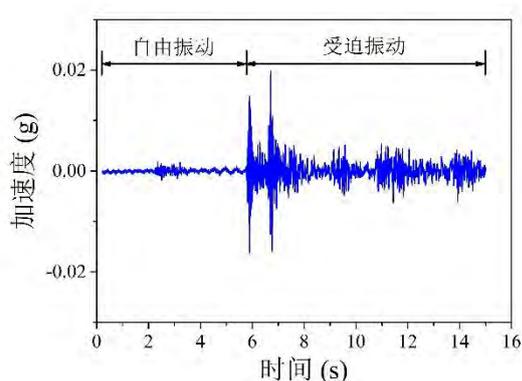
c) 现场测试全景(视角一)

d) 现场试验全景(视角二)

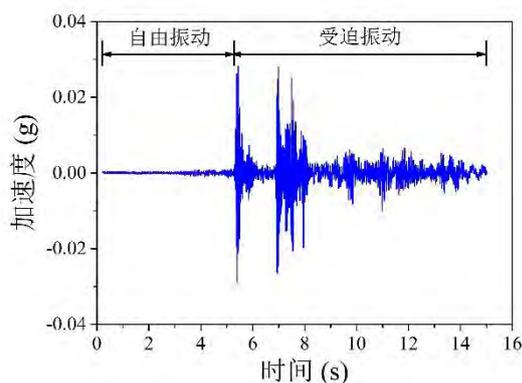
图 3.35 现场试验测试概况

### 3.8.3 试验测试结果

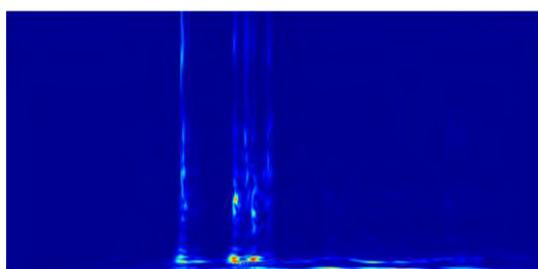
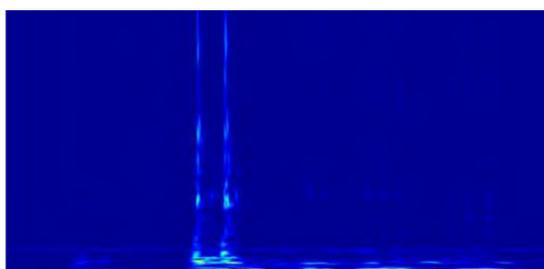
试验测试过程中,针对每种车型分别进行了 40 组随机跑车试验。由于桥面板上共设置有 6 个加速度传感器,故经一系列跑车测试后,理论上对于每类车载工况可形成 240 组动态加速度时域信号。然而,由于实际测试过程中受个别传感器接触不良、电缆线短路等不可避免因素的影响,最终针对车辆 A 建立了 148 条响应数据库,针对车辆 B 则建立了 160 条数据库。为尽可能地提高神经网络的测试精度,依据数值模拟及实验室试验研究所得经验结论,选用基于 Morlet 基函数的 CWT 变换的时-频分析技术对采集的时域加速度响应进行信号处理。对于本试验中所设置的测试工况,不同载重下运营车辆对桥梁结构所产生的典型响应及特征提取情况如图 3.36 所示。



a) 车致加速度响应(车辆 A)



b) 车致加速度响应(车辆 B)

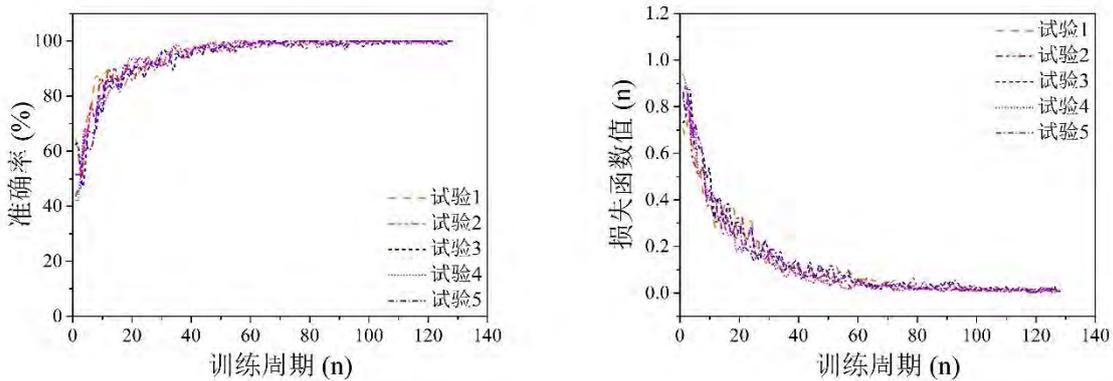


c) CWT 分析(车辆 A)

d) CWT 分析(车辆 B)

图 3.36 时-频分析技术处理下图像特征信息提取

深度学习网络训练方面,利用经 ImageNet 预训练后的 AlexNet 模型进行迁移学习训练;且通过设置五次随机测试的方式,尽可能地避免偶然性因素对试验结果造成的影响。其中,在车辆 A、B 所形成的数据库中随机选取 70%的样本作为训练集,其余 30%的样本则作为测试集进行准确率评估。参数设置方面,将每次训练迭代的最小批次设置为 64,最大迭代次数设置为 64,学习率设置为 0.0001。测试过程中的模型性能水平变化曲线如图 3.37 所示,可发现经训练 14 个周期后,该模型的准确率曲线开始由 50%逐步上升至 90%附近,而损失函数值则由 0.8 逐步降低至 0.2 左右;在迭代约 70 个周期后,其准确率曲线开始逐渐收敛至 99.68%左右,相应的损失函数值则稳定在 0.015 附近。最终,将设置好的测试数据集输入至该模型中进行精度评估,可发现五次随机试验的平均识别精度约为 94.13%,表明所提基于深度学习理论及结构特征响应提取的车辆荷载识别方法,在实际野外桥梁现场测试方面仍具有较好的可行性。



a) 准确率变化曲线

b) 损失函数变化曲线

分类输出	卡车A	44 (47.8%)	4 (4.3%)	91.7% (8.3%)
	卡车B	0 (0.0%)	44 (47.8%)	100.0% (0.0%)
	测试精度	100.0% (0.0%)	91.7% (8.3%)	95.7% (4.3%)
		卡车A	卡车B	测试精度
		预测结果		

分类输出	卡车A	44 (47.8%)	6 (6.5%)	88.0% (12.0%)
	卡车B	0 (0.0%)	42 (45.7%)	100.0% (0.0%)
	测试精度	100.0% (0.0%)	87.5% (12.5%)	93.5% (6.5%)
		卡车A	卡车B	测试精度
		预测结果		

c) 混淆矩阵分类结果

图 3.37 DCNN 网络模型训练结果

为进一步提高 DCNN 神经网络模型的性能,针对所建立形成的各有效测试样本,对其分别采用如表 3.3 所示的图像变换技术对数据库规模进行扩充。利用扩充后的样本库对该 DCNN 网络模型进行训练。这里,试验所用训练/测试集及相关模型训练参数均与前述保持一致。试验过程中的函数变化曲线如图 3.38 所示,

可发现经迭代约 60 个训练周期后,相应的准确率曲线开始由 50%逐渐上升至 90%左右,相应的损失函数值则由 1.1 逐渐下降至 0.2 附近;在迭代约 700 个周期后,其准确率曲线开始逐渐收敛于 100%左右,相应的损失函数值也稳定在 0.002 附近。最终,将所设置的测试数据集导入至该神经网络模型中,用以对训练结果进行精度评估。这里,可发现五次随机试验的平均识别精度达到了 97.21%,表明可利用图像变换技术对数据库样本容量进行扩充的方式,有效提高对野外车辆荷载信息的判断识别效果。

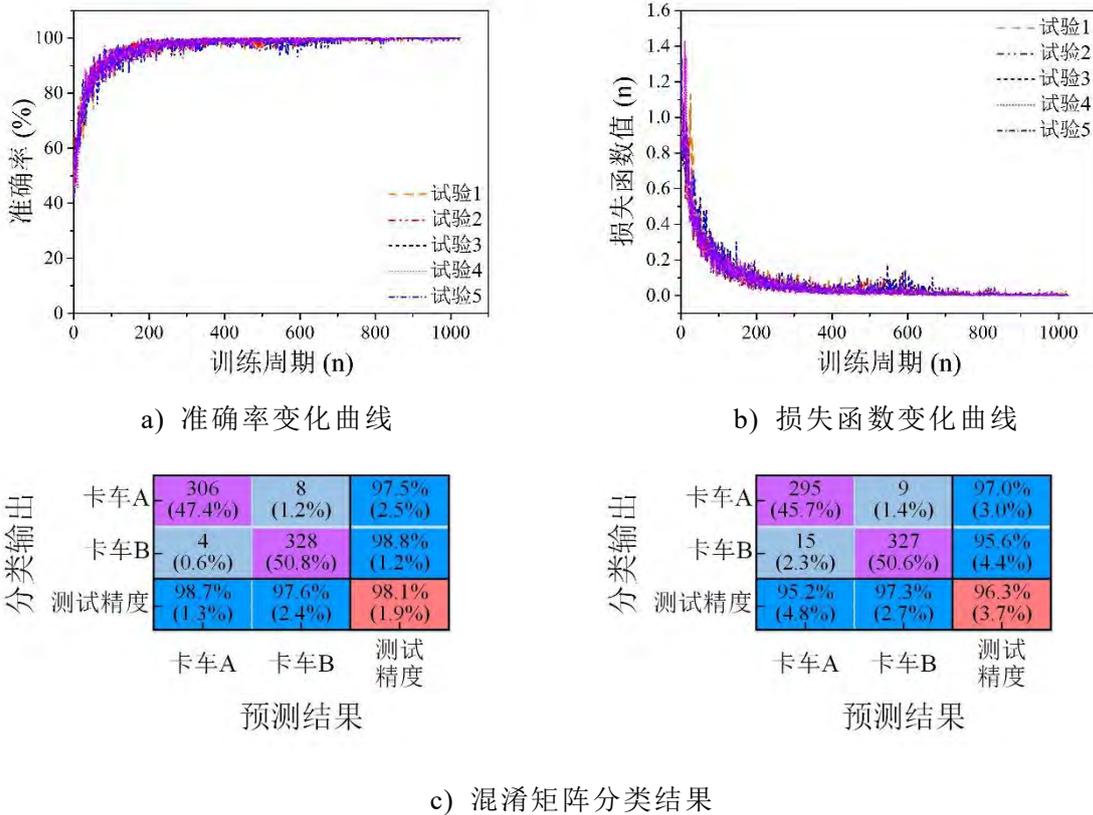


图 3.38 DCNN 网络模型优化训练结果

### 3.9 本章小结

为进一步验证所提基于深度学习理论及结构特征响应提取的桥梁上车辆荷载信息识别方法的有效性,本章分别开展了实验室缩尺模型振动测试及野外现场试验测试。在建立不同工况下结构动态响应样本的基础上,采用时-频分析技术对结构响应中的特征信息进行提取,并通过 DCNN 网络模型对运营状态下车辆荷载的相关信息依次进行了辨识。相关结论具体如下:

- (1) 利用时-频分析技术对测试所得桥梁结构的加速度响应进行分析。测试结果表明,基于 Morlet 基函数的 CWT 变换技术具有计算效率高、能够较为充分地提取结构特征响应信息等优势,可适用于本章所提出的车载荷载信息识别研究。
- (2) 通过采用迁移学习训练方式,将经大型自然物体图像集 ImageNet 训练后

的 AlexNet 模型引入至对缩尺车辆模型荷载信息的分类识别任务当中，并获得了 96.37% 的平均测试精度。同时，通过以图像处理技术扩充样本库数量的方式进行了优化训练，并获得了 98.87% 的测试精度，有效提高了模型的泛化能力。

(3) 将经单独训练过的车载分类器进行串联整合，提出了一种对车辆荷载信息进行综合检测的技术手段。测试结果表明，针对由车辆荷载所引起的单条结构响应，神经网络模型能够较好地实现对外界输入车辆荷载的载重情况、行驶速度、车辆类型及所在车道等信息的判断识别任务，且具有方便快捷、成本低廉等优势。

(4) 通过开展野外现场试验测试，对所提车辆荷载识别方法的实用性进行验证。测试过程中对足尺卡车荷载信息的识别精度达到了 97.21%，表明该方法可应用至对实际桥梁上车辆荷载的监测应用当中。

## 第4章 基于深度学习理论及压缩感知技术的车辆检测跟踪研究

### 4.1 引言

为实现对桥面上车辆荷载信息的动态监测,采用非接触式机器视觉技术对桥面上的运营车辆进行实时检测跟踪具有重要意义。针对传统的目标检测及跟踪方法所存在的测试精度易受干扰、实时性效果有待提高等不足之处,本章提出了一种基于深度学习理论及压缩感知技术的运营车辆目标检测跟踪方法。首先,在经统计建立车辆外观图像数据库的基础上,利用 Faster-RCNN 神经网络模型实现了对监控视频中车辆目标的定位检测;其次,将网络模型的目标检测输出结果用以更新压缩感知跟踪器中的相关参数,从而实现对车辆目标的实时跟踪。在此基础上,通过统计多种车辆的物理参数信息,实现了不同车辆类型与相应载重区间之间的映射关系,有效避免了传统视频监测技术仍需依赖 BWIM 系统进行车辆载重信息获取的局限性。通过分别开展实验室试验和野外试验的方式,以摄像机及无人机作为媒介对所提非接触式视频监测方法的可行性进行测试。除此之外,通过设置不同的环境测试工况,对该方法的稳定性进行评估分析。

### 4.2 车辆检测跟踪算法

#### 4.2.1 目标检测理论概述

在前述第二章及第三章中,所涉及的目标分类(Object classification)任务主要是指通过一系列算法将物体图像转化为某特定类别的数据信息,并通过预先设置好的标签类别对计算输出结果进行依次分类。然而,目标检测(Object detection)任务却主要侧重于关注图像中特定的目标信息,即要求通过一系列算法对相应目标的类别及位置等有效信息进行输出。对于传统的目标检测算法而言,其理论计算过程主要包含三个步骤:首先,利用不同规格尺寸的滑动窗口选取图像中的目标部分作为候选区域;其次,利用相关特征算法对候选区域内的特征信息进行提取(如利用 Haar 特征进行面部检测,利用 HOG 特征进行行人检测等);最后,通过分类器(如支持向量机等)将目标对象与背景信息进行分离输出<sup>[113,114]</sup>。值得注意的是,传统方法是指通过相关算法对一系列视频帧中的运动目标进行检测,即要求视频中的目标对象必须时刻处于运动变化状态。典型的代表算法有边缘检测法、背景差分法及光流法等。

近年来,基于深度学习理论的目标检测方法可实现对图像中目标物体的直接

检测，即与检测对象是否处于静止或运动状态并无关系。通常情况下，该技术可以主要分为两阶段(Two-stages)方法和一阶段(One-stage)方法等两大类<sup>[113,114]</sup>。其中，两阶段方法是指在利用算法生成一系列样本候选区域后，再利用深度学习网络对其进行归纳分类，其中典型的代表算法主要有 R-CNN、Fast-RCNN 及 Faster-RCNN 等；一阶段方法则无需产生候选区域，通过采用端至端的计算方式，直接将目标区域定位问题转化为回归问题进行处理，典型的代表算法主要有 YOLO、SSD 等。总体而言，基于候选区域方式的目标检测算法在识别准确率及定位精度上具有优势，而基于端至端方式的检测算法则在计算效率上表现突出。

## 4.2.2 基于 Faster-RCNN 的目标检测理论

网络结构设置方面，基于深度学习理论的 Faster-RCNN 目标检测算法主要由区域生成网络(Region proposal network, RPN)和 Fast-RCNN 网络两部分组成<sup>[115]</sup>，如图 4.1 所示。其在依次完成确定目标候选区域、提取对象深度特征、辨识特定对象及获取相应位置信息等步骤的基础上，通过共享卷积层相关参数训练结果的方式，在降低神经网络计算工作量的前提下，较好地提高了网络模型的计算效率。相关科学问题研究方面，2018 年，Cheng 等<sup>[116]</sup>利用 Faster-RCNN 神经网络，对污水管道所存在着的植物根部侵入、裂缝损伤、开裂漏水及变形下沉等受损位置实现了自动化检测识别；2018 年，Luo 等<sup>[117]</sup>为实现建筑工程项目现场施工的智能化，利用 Faster-RCNN 对监控相机拍摄所得的施工现场静态图像进行检测，并有效识别出了 22 类与建筑工程施工现场有关的物体对象及 17 类与生产活动相关的施工行为；2019 年，Wang 等<sup>[118]</sup>则利用 Faster-RCNN 神经网络实现对长城上砖石损伤情况的定位检测任务。然而，现有土木工程领域内将 Faster-RCNN 应用至车辆荷载检测识别方面的科学研究问题开展地相对较少。

### 4.2.2.1 Faster-RCNN 模型结构

网络模型架构设立方面，Faster-RCNN 在卷积网络层中设置有 13 个卷积层、13 个 ReLU 层和 4 个池化层，如图 4.1 所示。网络模型通过卷积层、ReLU 层及池化层等对所输入信息进行计算后，即可实现对物体目标的特征提取工作；在此基础上，将所得特征图输出至 RPN 网络中以生成候选区域，并通过叠加 Softmax 分类器来判断所提取到的特征信息属于前景目标还是背景目标；同时，通过利用边框回归技术进行特征信息修正，以此获得更加精确的目标区域检测结果<sup>[119]</sup>。除此之外，在获得多个候选区域的基础上，通过进行 ROI 池化操作，将不同尺寸的特征输入转换为具有统一维度的结果输出；并通过在全连接层中对特征信息进行分类回归，在判断输出候选区域物体特征的所属类别后，便可最终实现对检测框最终精确位置的确认。具体在候选区域建议网络设置方面，为避免传统选择搜索法、边缘提取法等技术所可能导致的计算冗余度过高现象，Faster-RCNN 通过采

用基于共享卷积层的 RPN 网络，对最后一级共享卷积层的特征图上采用  $n \times n$  滑动窗口进行扫描，使得每个窗口都能够产生单个特征向量；同时，在进行边界回归等操作后，从而可生成相对应的区域建议<sup>[120]</sup>。

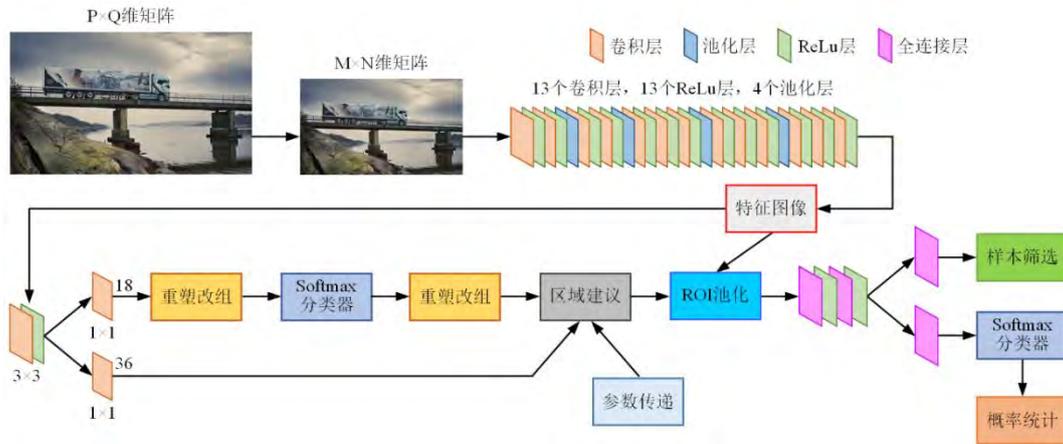


图 4.1 Faster-RCNN 神经网络结构示意图

#### 4.2.2.2 Faster-RCNN 模型检测原理

在 Faster-RCNN 神经网络中，所输入的图像样本在经卷积网络层进行传递后，通过获取物体共享特征图等信息，将所得结果输出至候选区间建议网络中进行处理。同时，在利用所设置的具有特定尺度的滑动窗口进行一系列卷积处理后，即可得到所输入对象的特征向量，从而实现对各候选区域中特征属性的判断识别。其中，每个滑动窗口均设置有 9 个不同大小的候选框。针对所提取到的图像特征信息，ROI 池化层通过在后续网络层中对特征信息开展下采样操作，从而将所更新的特征输入至 Softmax 分类器中；在完成辨识、回归及非极大值抑制处理等步骤后，即可得到特定目标的相关信息<sup>[115]</sup>。需要说明的是，在 Faster-RCNN 模型中，RPN 网络及 Fast-RCNN 网络可通过使用共享卷积层的方式，使得同一特征图能够在两个网络中进行并行计算，从而极大地提高神经网络模型的计算效率，如图 4.2 所示。根据现有文献建议，该共享网络层通常采用 AlexNet、VGG-16 等卷积神经网络模型<sup>[115]</sup>。

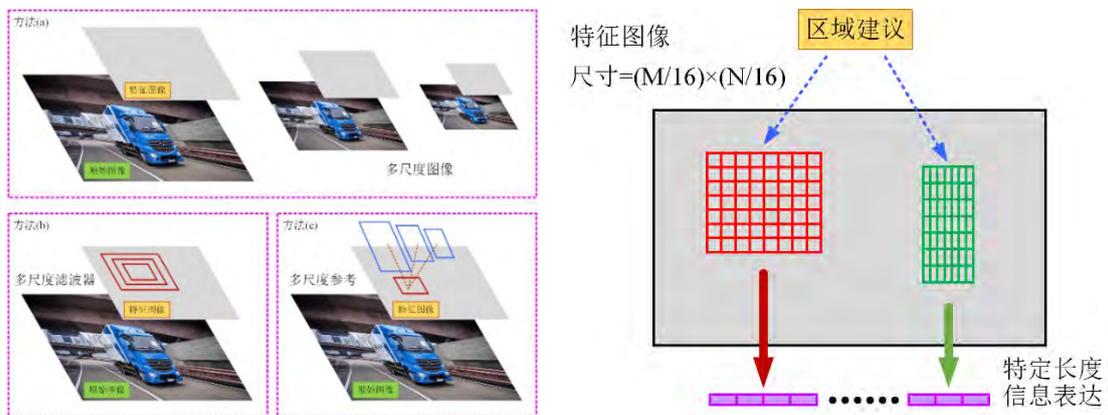


图 4.2 RPN 网络区间建议选取

### 4.2.2.3 Faster-RCNN 模型计算过程

值得注意的是，RPN 网络在利用卷积层对目标对象中的特征信息进行提取、并生成具有特定维度图像特征的基础上，额外定义了具有两个分支的全连接层。其中，reg-layer 层将主要负责对目标区域的框选尺寸信息进行确定，主要参数包括角点坐标  $(x,y)$  和宽度高度尺寸  $(w,h)$ ；而 cls-layer 层则主要负责对检测对象的特征属性进行判断，即辨识其属于前景目标还是背景目标<sup>[121]</sup>。除此之外，RPN 网络中所具备的特定滑动窗口处理方式，将保证 reg-layer 和 cls-layer 网络层实现对卷积层中所提取特征信息的有效关联。其中，每个滑动窗口将生成  $k$  个区域建议(即对应于  $k$  个候选框)。各候选框在经全连接层进行回归处理后，将依次输出  $2k$  个样本信息和  $4k$  个具有  $(x,y,w,h)$  表达形式的坐标信息。对于每个候选框，其性能表达程度将由滑动窗口纵横比及窗口中心点位置共同确定。一般情况下，候选框的纵横比包括 1:1, 1:2 及 2:1 三种情况，区域面积则分为  $128^2$ ,  $256^2$  及  $512^2$  等情况。其经排列组合后，可共生成 9 种不同组合。因此，对于维度尺寸为  $W \times H$  的目标特征图，其所可能产生的总候选框数量为  $W \times H \times k$  个。除此之外，由于 RPN 网络具有平移不变性，可使得即使对图像中的目标对象进行平移处理后，神经网络模型也能够正确地对所建议区域进行准确候选，如图 4.3 所示。

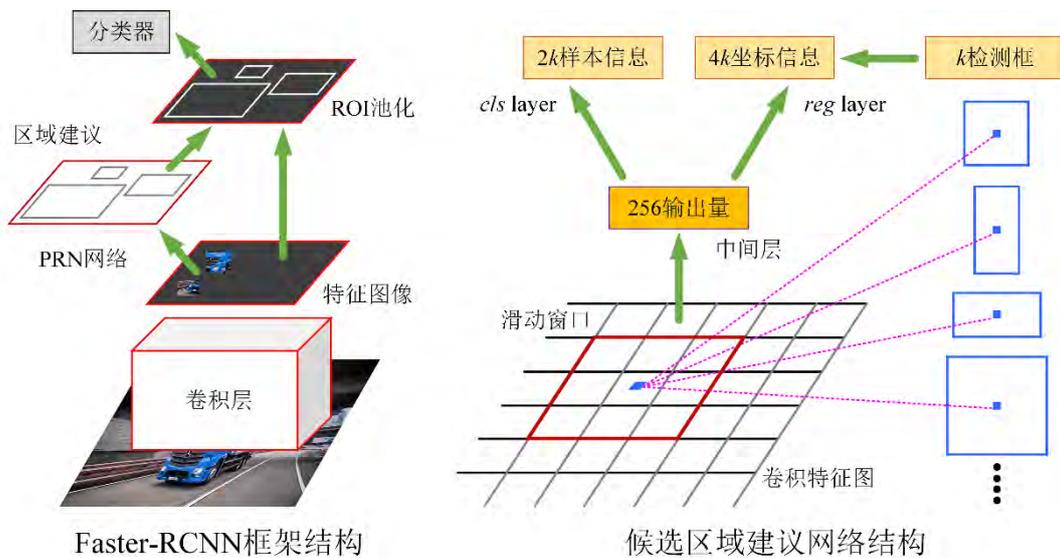


图 4.3 Faster-RCNN 网络特征目标选取<sup>[115]</sup>

在进行候选区域建议网络训练时，通过将候选框中对象位置与真实目标所在位置之间的重叠比例(IOU)作为衡量标准，从而判定所候选区域中是否准确存在所预期的目标对象。对此，Faster-RCNN 中定义有两种判定方式，其中一种是将与真实目标之间重叠度最高的区域认定为正确预测结果，另一种则是将重叠比例大于 0.7 的区域对象判定为前景目标，否则便判定为背景目标。一般情况下，多采用第二种方式进行检测目标判定识别。这里，重叠比函数被定义为<sup>[122]</sup>：

$$IOU = \frac{S_a \cap S_{gt}}{S_a \cup S_{gt}} \quad (4.1)$$

其中,  $S_a$ 表示候选框所在的区域;  $S_{gt}$ 表示真实目标所在的区域。

针对具有特定对象的目标识别任务, 在经候选区域建议网络迭代训练后, 其损失函数值需达至最小才可满足收敛要求。其中, 目标对象的损失函数值可表达为:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (4.2)$$

其中,  $i$ 表示经最小训练批次计算后所对应的候选区域;  $p_i$ 表示当前候选区域 $i$ 属于检测目标的概率值(当为前景目标时,  $p_i^*$ 取值为 1; 否则为 0);  $L_{cls}$ 表示特定目标分类任务所对应的损失函数值;  $L_{reg}$ 表示候选区域边框经回归处理后所对应的损失函数值;  $\lambda$ 作为经验系数, 被用于平衡计算过程中 $L_{cls}$ 和 $L_{reg}$ 之间的数值差距, 其通常取为 10; 集合 $t_i$ 为预测区域 $i$ 所在位置的坐标向量, 可表示为 $\{t_x^*, t_y^*, t_w^*, t_h^*\}$ 。其中,  $t_i^*$ 代表真实目标所在区域的坐标向量。

损失函数值计算方面, 为避免传统均方误差理论等方法所可能导致的局部最优解混沌现象, Faster-RCNN 中通过设计逻辑回归方法对损失函数进行了有效计算。该过程可表示为:

$$L(y^*, y) = -(y \log y^* + (1 - y) \log (1 - y^*)) \quad (4.3)$$

其中,  $y$ 表示各样本的类别归属情况;  $y^*$ 表示经辨识后分类器的特征输出结果。对于 RPN 网络层, 该结果可表示为:

$$L_{cls}(p_i, p_i^*) = -(p_i^* \log p_i + (1 - p_i^*) \log (1 - p_i)) \quad (4.4)$$

针对候选区域内的采样对象均属于前景信息(即 $p_i^* = 1$ )这一特殊情况, 上式(4.4)可表示为:

$$L_{cls}(p_i, 1) = -\log p_i \quad (4.5)$$

其中, 需要特别进行说明的是, 当 $p_i$ 为前景信息的概率值越收敛于 1 时, 其在训练过程中所对应的损失函数值将相应越小。

在 Faster-RCNN 神经网络中, 通过设置 smooth-L1 函数, 对边框回归的损失函数情况进行考虑<sup>[123]</sup>。同时, 当预测-候选区域之间的相对位置与真实-候选区域之间的相对位置相比较为相近时, 候选区域的回归效果将随之增强。该过程可表示为:

$$L_{reg}(t_i, t_i^*) = \text{smooth}_{L1}(t_i - t_i^*) = \begin{cases} 0.5(t_i - t_i^*)^2 & \text{if } |t_i - t_i^*| < 1 \\ |t_i - t_i^*| - 0.5 & \text{otherwise} \end{cases} \quad (4.6)$$

若将上式(4.6)中的参数 $t_i$ 及 $t_i^*$ 分别以向量 $(t_x^i, t_y^i, t_w^i, t_h^i)$ 、 $(t_x^{i*}, t_y^{i*}, t_w^{i*}, t_h^{i*})$ 的形式进行表示, 则可相应得到:

$$\begin{cases} t_x = (x - x_a)/w_a, t_y = (y - y_a)/h_a, t_w = \log(w/w_a), t_h = \log(h/h_a) \\ t_x^* = (x^* - x_a)/w_a, t_y^* = (y^* - y_a)/h_a, t_w^* = \log(w^*/w_a), t_h^* = \log(h^*/h_a) \end{cases} \quad (4.7)$$

其中,  $(x, y, w, h)$  表示矩形候选框的中心坐标值及宽(高)度信息;  $x, x_a, x^*$  依次表示预测框、候选框及真实框所对应的中心横坐标值;  $y, y_a, y^*$  则依次表示相应预测框、候选框及真实框所对应的中心纵坐标值。

在经过上述操作处理后, 单个检测目标对象将可能产生多个候选区域。此时, 可通过采用非极大值抑制算法从中选择与真实情况重叠度最高的区域, 并将其作为最优区域进行输出。

### 4.2.3 目标跟踪理论概述

目标跟踪方法研究方面, 现有相关技术对运动目标进行实时跟踪的本质是通过图像分割技术, 从运动场景中选择出用户所关注的目标对象; 在此基础上, 通过算法计算等一系列步骤最终完成对目标对象分类及后续场景跟踪的任务。因此, 对特定目标实施初始判断分割的准确程度将极大地影响后续识别跟踪效果。相关目标跟踪技术的计算流程可大致分为以下步骤: 首先, 定义初始化目标框, 并提取下一帧图像中所产生候选框的特征; 其次, 依据相关方法对该候选框进行权重估计; 最后, 选取期望最高的候选框作为预测结果, 并以此循环从而完成目标跟踪任务<sup>[113,114]</sup>。

根据计算方法与原理的不同, 目标跟踪方法可分为基于相关滤波理论的跟踪算法、基于稀疏表示理论的跟踪算法和基于深度学习理论的跟踪算法。其中, 基于相关滤波理论的跟踪算法可利用相关性对两个信号之间的相似程度进行描述, 通过搜寻特定的滤波模板完成前后帧图像间的卷积操作, 并将最大响应所在的区域位置作为预测目标, 从而完成特定目标的追踪任务。所对应典型的代表算法有 Mosse、CSK、KCF 及 SR-DCF 等。对于基于稀疏表示理论的目标跟踪算法而言, 其通过将所输入信号以稀疏性约束形式进行表示, 从而可将目标追踪问题转化为稀疏逼近问题进行求解。所对应典型的代表算法有 L1-track、ALSA 及 L1-APG 等。除此之外, 基于深度学习理论的目标追踪算法通过将在经分类任务训练过的神经网络模型迁移至目标跟踪领域, 使得其在近年来得到了一定发展。其中典型的代表算法有 DLT、Deep-sort 及 MDNet 等。

### 4.2.4 基于压缩感知技术的目标跟踪理论

作为近年来国际上信息理论研究领域内所取得的重要突破, 压缩感知理论 (Compressive sensing, CS) 已开始逐渐被各行业的科学研究工作者所关注。该技术最早由 Candes 及 Romberg 等首先提出, 其在保证高维信号在某变换域上具有稀疏性的基础上, 通过利用特定随机测量矩阵将其投影至一低维空间上, 并通过进行最优化问题求解方式对原始信号进行重构<sup>[124]</sup>。总体而言, 压缩感知理论能够

较好地解决传统 Shannon 定理对信号采样频率的限制问题，可通过较少的采样资源实现对原始信号的有效重构。目前，该技术已被应用至医学成像、遥感监测、多媒体混合编码、数字通讯及结构健康监测等领域的研究工作当中<sup>[125]</sup>。其中，在土木工程行业的健康监测研究领域当中，由于建筑结构的动力响应信号通常仅由前几阶模态参与，因此其在频域上整体表现为具有稀疏性的窄带信号。若当受到噪声等干扰信号污染时，经试验测试结果表明，结构的动力信号在其他变换域上仍可表现出近似稀疏的特性<sup>[126]</sup>，从而有效保证了其卓越的稳定性。

#### 4.2.4.1 压缩感知模型结构

近年来，受工程结构建设规模不断扩大的影响，传统结构健康监测研究所需的传感器数量及监测所需成本也随之增多，将十分不利于监测数据的分析处理。然而，压缩感知理论的出现极大地改善了这一问题。其在减小获取数据所需成本的基础上，对于新型无线传感器所可能存在的数据丢失情况也有较好的解决途径，因此已开始逐渐被相关学者应用至土木工程行业内的结构健康监测领域研究当中。例如在缺失数据重构研究方面，2013 年，Bao 等<sup>[127]</sup>为解决无线传感器在信号传输过程中所存在的部分数据丢失问题，提出了基于压缩感知理论的丢失数据重构方法；2014 年，Bao 等<sup>[128]</sup>为解决测点快速移动情况下无线传感所可能导致的数据丢失问题，通过在桥梁结构及行驶车辆上分别布置无线传感器和信号接收基站，实现了一种基于压缩感知的数据重构技术。对于桥梁上车载时空分布等动力学反问题研究，由于该类问题本身也具有稀疏性的特征表现形式，故也可将其归纳于稀疏约束重构问题。

作为近年来所发展的一种新型方法，基于压缩感知理论的运动目标跟踪理论针对具有可压缩性的原始信号，可通过利用随机测量矩阵在对其进行压缩处理的基础上，采用分类器完成对目标特征的分类辨识，最终实现对目标对象的跟踪效果。整个计算理论的流程示意图如图 4.4 所示。具体实施方面，在将原始信息输入至所建立的系统架构中后，首先对目标对象的特征模板进行确定，并在特征区域内进行前(背)景特征信息提取；其次，利用随机测量矩阵降维处理原始输入信号，同时对分类器中的权值参数进行迭代更新；最后，按照最大相似原理对初始目标的所在位置进行确定，并经优化处理后即可实现对目标最终位置的确定<sup>[129]</sup>。



图 4.4 基于压缩感知理论的目标跟踪原理

#### 4.2.4.2 压缩感知模型跟踪原理

针对现有基于深度学习理论的目标跟踪算法所存在的计算复杂度高、跟踪效

率有限等不足之处，基于压缩感知的目标跟踪算法通过从图像特征空间直接进行信息采样、并得到相应的压缩特征等措施，有效避免了对冗余信息的计算处理过程，从而确保其具有计算量小、跟踪效率高等优点。该算法的具体流程设置方面，在经人工手动确定目标对象的所在位置后，通过在目标周围进行前(背)景样本信息采样，实现对所得区域内图像特征信息的提取工作，如图 4.5 所示。其中，首帧图像中所提取到的目标特征将被用于初始化定义分类器的相关参数。然而，传统的压缩感知追踪算法主要集中在对运动状态下人体面部的实时捕获<sup>[130]</sup>，对于高速移动行驶状态下车辆目标的跟踪效果并不明确。因此，通过开展基于压缩感知理论的运动车辆目标跟踪问题研究具有重要意义。

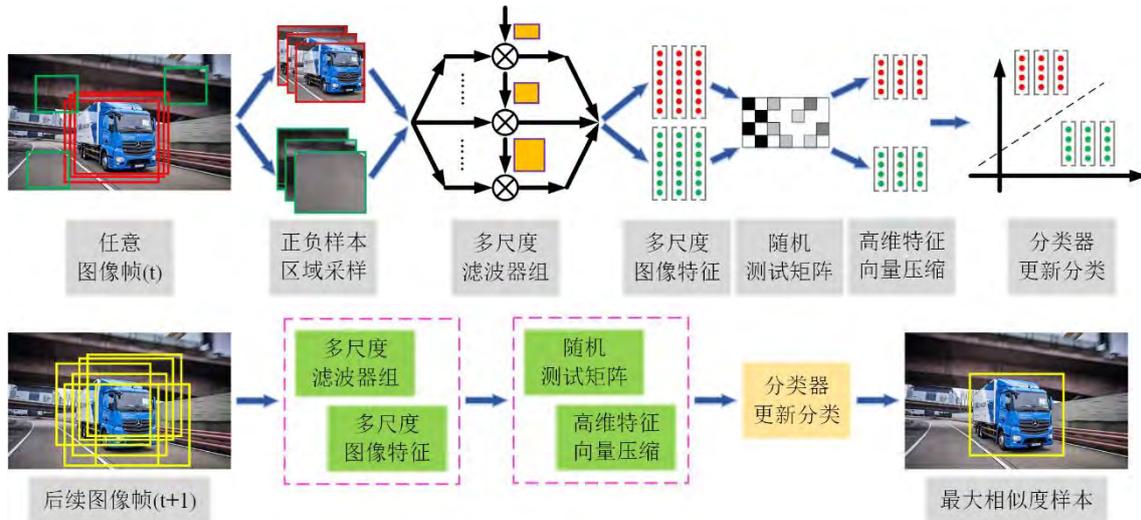


图 4.5 基于压缩感知理论的目标跟踪流程

#### 4.2.4.3 压缩感知模型计算过程

在压缩感知理论的采样过程中，首先需要对目标区域进行确定。即通过人工手动框选特征区域的方式，在该区域内部生成一系列采样网格，并结合所设置的采样条件，依次选择符合要求的目标样本。其中，在前景目标特征信息提取阶段，样本区域外边界的搜索半径将相应变小，以便于算法在目标区域附近进行正样本提取；在背景目标特征提取阶段，将主要围绕内、外边界半径之间的环形区域进行特征采样。除此之外，通过设置概率满足条件用以实现对样本数量的有效控制<sup>[131]</sup>。其中，当概率条件大于 1 时，所在区域内的采样样本均被视为前景目标；否则，将在所在环形区域内随机进行背景样本的筛选。这里，该采样过程可表示为<sup>[123]</sup>：

$$ind = (rd < prob) \wedge (dist < inrad) \wedge (dist < outrad) \quad (4.8)$$

其中， $rd$ 表示各元素均服从 0-1 均匀分布的随机矩阵； $prob$ 表示目标判定过程中所设置的概率条件； $dist$ 表示采样对象与目标对象之间的坐标距离； $inrad$ 表示采样过程中所设置的区域外边界半径； $outrad$ 表示所对应设置的采样区域内边界半径。针对采样阶段所设置的概率条件，可将其表示为：

$$prob = \begin{cases} \max num \\ (\max row - \min row) \times (\max col - \min col) \end{cases} \quad (4.9)$$

其中， $\max num$ 表示特征区域内所能接受的样本容量上限； $\max row$ 及 $\min row$ 分别表示在采样区域内所设置的上、下边界条件； $\max col$ 及 $\min col$ 则分别表示采样在区域内所设置的左、右边界条件。

依据压缩感知理论的相关计算流程可知，在对目标区域完成采样操作后，首先需利用随机测量矩阵将待压缩信号投影至某低维空间进行处理；其次，在特征提取过程中，通过将各样本与多个具有不同尺度的滤波器进行卷积运算，并在得到高维特征向量后，利用随机测量矩阵对其进行降维处理；最后，利用分类器对经降维后的图像特征进行更新或分类。具体在多尺度滤波器选取方面，针对目标跟踪任务中所存在的尺度变换问题，压缩感知理论通过选取具有尺度不变性的矩形滤波器 $\{h_{1,1}, \dots, h_{w,h}\}$ ，用以实现对目标区域内的各样本进行卷积操作。其中，所涉及到的矩形滤波器可表示为：

$$h_{i,j}(x,y) = \begin{cases} 1, & 1 \leq x \leq i, 1 \leq y \leq j \\ 0, & otherwise \end{cases} \quad (4.10)$$

其中， $i, j$ 分别表示矩形滤波器的宽度和高度信息。值得注意的是，对于经卷积操作后维度尺寸为 $w \times h$ 的特征图像，可通过对应设置 $w \times h$ 个滤波器，将所得的全部图像特征进行重构，从而得到维度尺寸为 $(w \times h)^2$ 的列向量。在此基础上，通过将所有列向量进行拼接，在得到对应的高维多尺度图像特征后，利用随机测量矩阵对相乘运算过程中的非零项进行考虑，从而实现降维运算处理。

随机测量矩阵选取方面，不仅要考虑计算复杂度，也需要依据目标特性进行合理选取。根据相关文献建议可知<sup>[132]</sup>，压缩感知理论中所常用的随机测量矩阵为一随机高斯矩阵，其中各元素均满足 $N(0,1)$ 分布。该理论可表达为：

$$r_{ij} = \sqrt{s} \times \begin{cases} 1, & prob. = 1/2s \\ 0, & prob. = 1 - 1/s \\ -1, & prob. = 1/2s \end{cases} \quad (4.11)$$

其中， $s$ 表示计算复杂度，通常将其数值大小设置为4。

高维特征信息压缩方面，采样样本在与矩形滤波器进行卷积运算后，会得到维度尺寸处于 $10^6 \sim 10^{10}$ 范围之间的高维特征向量 $\mathbf{X}$ 。此时，通过利用随机测量矩阵 $\mathbf{R}$ ，将其投影至低维特征向量 $\mathbf{V}$ 。其中，该投影过程无需考虑原始信号所可能产生的特征丢失问题。这一过程可表示为：

$$\mathbf{V} = \mathbf{R}\mathbf{X} \quad (4.12)$$

至于分类器更新设置方面，针对特征区域采样所得的不同样本，通过在目标对象特征提取阶段对其使用同一组滤波器，从而便可得到样本容量保持一致、且各样本特征值具有多样性的数据集。依据所设置样本类型的不同，可将所提取到

的图像特征分为两组。其中，第 I 组为前一帧中所提取到的前(背)景图像特征值，第 II 组则为当前帧所提取到的图像特征值。那么依据压缩感知计算理论可知，第 I 组样本将用于对分类器进行权值更新，第 II 组则将通过分类器进行对象分类。这里，通过将目标跟踪问题间接转化为二进制分类问题进行考虑，并利用 Naive-Bayes 分类器对所得低维特征中的所有信息做处理。该过程可表示为<sup>[123]</sup>：

$$H(v) = \log \left( \frac{\prod_{i=1}^n p(v_i|y=1)p(y=1)}{\prod_{i=1}^n p(v_i|y=0)p(y=0)} \right) = \sum_{i=1}^n \log \left( \frac{p(v_i|y=1)}{p(v_i|y=0)} \right) \quad (4.13)$$

依据相关数学理论可知，若假设参数  $v$  服从独立分布统计，则高维变量随机矩阵的投影情况应相应满足高斯概率分布。因此，分类器  $H(v)$  中的条件分布概率应满足以下分布情况：

$$p(v_i|y=1) \sim N(\mu_i^1, \sigma_i^1), \quad p(v_i|y=0) \sim N(\mu_i^0, \sigma_i^0) \quad (4.14)$$

其中， $\mu_i^1$ 、 $\sigma_i^1$  分别表示前景目标中第  $i$  个特征所对应的均值和标准差； $\mu_i^0$ 、 $\sigma_i^0$  分别表示背景目标中第  $i$  个特征所对应的均值和标准差。在实际场景中的目标跟踪过程中，通过对每一帧图像中的分类器进行参数重置，即可实现对 Naive-Bayes 分类器的权值更新。其中，该参数重置过程可表示为：

$$\lambda \mu_i^1 + (1-\lambda)\mu^1 \rightarrow \mu_i^1, \quad \sqrt{\lambda(\sigma_i^1)^2 + (1-\lambda)(\sigma^1)^2 + \lambda(1-\lambda)(\mu_i^1 - \mu^1)^2} \rightarrow \sigma_i^1 \quad (4.15)$$

其中， $\lambda$  表示算法对于目标特征信息的学习效率。另一方面，上式(4.15)中所涉及的参数均值和方差的初始化过程可表示为：

$$\mu^1 = \frac{1}{n} \sum_{k=0|y=1}^{n-1} v_i(k), \quad \sigma^1 = \sqrt{\frac{1}{n} \sum_{k=0|y=1}^{n-1} (v_i(k) - \mu^1)^2} \quad (4.16)$$

综上所述，基于压缩感知理论的目标跟踪方法具有计算效率高、实时性好等特点，在相关理论技术中具有一定特有优势。然而，现有的压缩感知追踪技术研究对象主要为人体面部特征，对高速移动状态下运营车辆的跟踪效果并不明确，因此有必要对其开展相关试验测试。除此之外，该方法在初始目标特征提取阶段也需要借助人工经验进行对象框选，其自动化效率有待提高。因此，研发一种无需人类先验知识辅助的自动化目标检测追踪技术具有重要意义。

## 4.3 实验室试验测试

### 4.3.1 神经网络训练

针对本章所提出的基于深度学习及压缩感知理论的运动车辆目标检测跟踪方法，首先通过开展实验室试验对其有效性进行测试。值得注意的是，在基于非接触式机器视觉测量技术的科学问题研究当中，为使测试结果尽可能满足实际需求，要求缩尺模型与真实对象在视觉外观上具有一定可比性。因此在本试验过程中，选用了外观构造相似度较高的缩尺仿真车辆模型对实际车辆进行模拟。其