**硕士学位论文**

基于视觉的地下管道缺陷检测方法研究

**RESEARCH ON DETECTING METHOD OF SEWER DEFECTS BASED ON VISION**

**陈克凡**

**哈尔滨工业大学**

**2018年12月**

国内图书分类号：TP23621 学校代码：10213

国际图书分类号：621 密级：公开

**工学硕士学位论文**

基于视觉的地下管道缺陷检测方法研究

|  |  |  |
| --- | --- | --- |
| 硕士研究生 | ： | 陈克凡 |
| 导 师 | ： | 胡泓教授 |
| 申请学位 | ： | 工学硕士 |
| 学科 | ： | 机械电子工程 |
| 所 在 单 位 | ： | 哈尔滨工业大学（深圳） |
| 答 辩 日 期 | ： | 2018年12月 |
| 授予学位单位 | ： | 哈尔滨工业大学 |

Classified Index: TP23621

U.D.C: 621

A dissertation submitted in partial fulfillment of

the requirements for the academic degree of

Master of Engineering

**RESEARCH ON DETECTING METHOD OF SEWER DEFECTS BASED ON COMPUTER VISION**

|  |  |
| --- | --- |
| **Candidate：** | Chen Kefan |
| **Supervisor：** | Prof. Hu Hong |
| **Academic Degree Applied for：** | Master of Engineering |
| **Speciality：** | Mechatronic Engineering |
| **Affiliation：** | Harbin Institute of Technology, Shenzhen |
| **Date of Defence：** | December, 2018 |
| **Degree-Conferring-Institution：** | Harbin Institute of Technology |

# 摘 要

随着城市化的深入，地下管道越来越密集，对地下管道进行定期检查维护可在很大程度上减少管道缺陷带来的社会经济损害。目前，CCTV（闭路电视）检测是主流的地下管道检测方法，通过地下管道机器人对管道进行实况拍摄，而后交由专业检查人员进行判定。大量的管道视频由人工进行检测存在易疲劳、主观性强、效率低下等缺点。为解决上述缺点，本课题利用计算机视觉方法构建了针对管道视频的一体化智能检测系统，对管道有无缺陷、缺陷类别及缺陷发生位置进行自动检测。

本课题将管道缺陷检测划分为视频帧正异常检测、异常分类、伪异常帧过滤和缺陷信息识别等四个步骤。对比了基于方向梯度直方图特征的支持向量机模型、基于多类别Gabor滤波器特征的极端随机树模型以及卷积神经网络模型的性能，最终选用卷积神经网络作为检测模型。针对大型深层网络算法实时性差的问题，采用了深度分离卷积和分组Shuffle卷积替代传统卷积的方式，在保持准确率的同时大大提高了模型的检测效率；针对由于管道地形变化及人为偏摆控制等原因产生的过检率较高的问题，利用改进的双向光流法对摄像机的位姿变化进行监测，过滤位姿变化幅度过大的伪异常帧以降低过检率；针对实际检测中准确率不高的问题，采用了迁移学习和模型快照集成策略提高了实际场景的准确率。同时，为满足发现缺陷同时自动记录缺陷地段、距离等信息的功能需求，本课题设计了视频字符检测和识别算法。针对管道场景视频字符特点利用最大极值稳定区域算法和基于级联分类器的非字符区域筛除算法进行字符定位；利用颜色信息融合尺度信息的层次聚类算法进行字符分割；分别利用模板匹配和OCR软件进行数字和文字识别。

本课题构建的针对管道视频的一体化智能检测系统，在准确性和实际场景适应能力上相对传统检测方式有了大幅的提高。同时，仅利用正异常检测进行视频缺陷段提取，再结合人工复检的检测方式可以在保证较高缺陷检出率的同时缩短专业人员约70%的检测时间。

关键词：地下管道；缺陷检测；卷积神经网络；图像处理；字符识别

# Abstract

With the deepening of urbanization, the scale of urban underground pipelines is expanding. Regular inspection and maintenance of underground pipelines can greatly reduce the social and economic damage caused by pipeline defects. At present, CCTV (Closed Circuit Television) detection is the mainstream underground pipeline detection method. The underground pipeline robot is used for live shooting of the pipeline, and then the videos are judged by professional inspectors. A large number of pipeline videos are detected manually， which has the disadvantages of fatigue, subjectivity and inefficiency. In order to solve the above shortcomings, an integrated intelligent detection method for pipeline video based on computer vision is proposed. This method can automatically determine whether the pipeline has defects, what defects exist and where the defects occur.

Pipeline defect detection is divided into four steps: video frame anomaly detection, abnormal frame defect classification, pseudo anomaly frame filtering and defect information recognition. Compared the performance of support vector machine model based on HOG (Histogram of oriented gradients) feature, extremely random tree model based on GIST (Gabor filter features of multi-direction and multi-scale) feature and convolutional neural network, and finally chose convolutional neural network as the detection model. Aiming at the problem of poor real-time performance of large-scale deep network algorithms, depthwise separable convolution and group shuffle convolution are used to replace traditional convolution, which greatly improves the detection efficiency of the model while maintaining accuracy. Aiming at the problem of high over-detection rate due to pipeline topography and artificial yaw control, the improved bidirectional optical flow method is used to filter pseudo anomaly frame. Aiming at the problem of low accuracy in actual detection, the transfer learning and model snapshot integration strategy is adopted to improve the accuracy. At the same time, video character detection and recognition algorithm is designed to meet the functional requirements of automatically recording information such as defect location and distance while finding the defect. According to the video character characteristics of pipeline scenes, the maximum stable extremum region algorithm and the non-character region filtering algorithm based on cascaded classifier are used for character localization. The hierarchical clustering algorithm based on color and scale information is used for character segmentation. Template matching and OCR software are used for number and text recognition respectively.

Compared with the traditional detection method, the proposed integrated intelligent detection method for pipeline video has greatly improved the accuracy and actual scene adaptability. At the same time, the use of anomaly detection for video defect segment extraction, combined with the manual re-inspection can shorten the detection time of professionals by about 70% while ensuring a higher defect detection rate.

**Keywords:** underground pipeline, defect detection, convolutional neural network, image processing, character recognition

目 录

[摘 要 I](#_Toc530236981)

[Abstract II](#_Toc530236982)

[第1章 绪 论 1](#_Toc530236983)

[1.1 课题背景及研究的目的和意义 1](#_Toc530236984)

[1.1.1 课题来源 1](#_Toc530236985)

[1.1.2 研究目的及意义 1](#_Toc530236986)

[1.2 国内外研究现状 2](#_Toc530236987)

[1.2.1 视觉管道检测技术研究现状 2](#_Toc530236988)

[1.2.2 图像识别算法研究现状 4](#_Toc530236989)

[1.3 本文主要研究内容 6](#_Toc530236990)

[第2章 管道缺陷检测方案设计及方法研究 8](#_Toc530236991)

[2.1 引言 8](#_Toc530236992)

[2.2 管道缺陷检测方案设计 8](#_Toc530236993)

[2.2.1 地下管道检测缺陷类别及功能需求分析 8](#_Toc530236994)

[2.2.2 管道缺陷检测方案制定 10](#_Toc530236995)

[2.2.3 管道缺陷检测性能评估指标分析 11](#_Toc530236996)

[2.3 基于传统机器学习的管道缺陷检测方法 13](#_Toc530236997)

[2.3.1 支持向量机管道检测方法 13](#_Toc530236998)

[2.3.2 极端随机树管道检测方法 16](#_Toc530236999)

[2.4 基于卷积神经网络的管道缺陷检测方法 18](#_Toc530237000)

[2.4.1 卷积神经网络分析 18](#_Toc530237001)

[2.4.2 正异常检测方法 20](#_Toc530237002)

[2.4.3 异常分类方法 22](#_Toc530237003)

[2.5 卷积神经网络模型训练 26](#_Toc530237004)

[2.5.1 数据增强 26](#_Toc530237005)

[2.5.2 训练策略 28](#_Toc530237006)

[2.6 本章小结 29](#_Toc530237007)

[第3章 管道缺陷检测性能优化 30](#_Toc530237008)

[3.1 引言 30](#_Toc530237009)

[3.2 管道正异常检测方法时间效率优化 30](#_Toc530237010)

[3.2.1 深度分离卷积 30](#_Toc530237011)

[3.2.2 分组Shuffle卷积 31](#_Toc530237012)

[3.2.3 优化结果分析 32](#_Toc530237013)

[3.3 管道检测过检率优化 33](#_Toc530237014)

[3.3.1 单帧特征点检测 33](#_Toc530237015)

[3.3.2 帧间特征点跟踪 34](#_Toc530237016)

[3.4 管道异常分类准确率优化 36](#_Toc530237017)

[3.4.1 迁移学习 36](#_Toc530237018)

[3.4.2 模型集成策略 37](#_Toc530237019)

[3.4.3 优化结果分析 38](#_Toc530237020)

[3.5 本章小结 39](#_Toc530237021)

[第4章 管道信息及缺陷位置信息识别方法 40](#_Toc530237022)

[4.1 引言 40](#_Toc530237023)

[4.2 信息字符定位 40](#_Toc530237024)

[4.2.1 字符候选区域提取 40](#_Toc530237025)

[4.2.2 基于Adaboost算法的非字符区域筛除 41](#_Toc530237026)

[4.3 信息字符分割 43](#_Toc530237027)

[4.3.1 基于K-Means的颜色分割 43](#_Toc530237028)

[4.3.2 基于层次聚类的尺度信息融合分割 46](#_Toc530237029)

[4.4 信息字符识别 48](#_Toc530237030)

[4.5 本章小结 50](#_Toc530237031)

[第5章 管道缺陷检测方法的实验与分析 51](#_Toc530237032)

[5.1 引言 51](#_Toc530237033)

[5.2 管道图像数据集构建 51](#_Toc530237034)

[5.3 管道检测性能实验分析 52](#_Toc530237035)

[5.4 视频时间压缩实验分析 54](#_Toc530237036)

[5.5 信息识别精度实验分析 55](#_Toc530237037)

[结 论 58](#_Toc530237038)

[参考文献 59](#_Toc530237039)

[攻读硕士学位期间发表的论文及其他成果 63](#_Toc530237040)

[哈尔滨工业大学学位论文原创性声明和使用权限 64](#_Toc530237041)

[致 谢 65](#_Toc530237042)

# 绪 论

## 课题背景及研究的目的和意义

### 课题来源

本课题来源于深圳某公司项目——基于视觉的地下管道缺陷检测系统，本文主要针对该系统中的视觉算法部分，包括视频帧正异常检测、异常帧分类、伪异常帧过滤及缺陷位置信息识别记录。

### 研究目的及意义

随着我国城镇化建设不断深入，地下管道越来越密集，大批管道因为年限较为久远，老化而发生泄漏、爆破等事故，对环境、社会经济及人身安全造成损害，所以对地下管道进行定期检查维护是十分有必要的[1]。

目前较为成熟的管道检测系统有声纳检测系统、超声导波检测系统、激光扫描三维重建检测系统以及管道闭路电视检测系统（CCTV）等[2]。超声导波检测系统[3]通过仪器给探头施加激励信号，根据反射回来的导波进行缺陷快速定位，但仅适用于缺陷种类较少、结构形态变化性小的运输管道，其原理结构如图 1‑1所示。管道声纳检测系统只能检测液面以下管道状况，且无法直观判断缺陷种类[4]。激光扫描三维重建检测通过管道机器人上安装的激光测距仪获取管道内部的三维点云数据，利用配准算法和重构技术，重建出管道内部三维模型[5]。该方法能够对缺陷定量检测，且能够检测出较细微缺陷，但主要针对管道的结构性缺陷，缺乏语义信息无法对管道功能性缺陷进行分类检测且检测费用过高，其原理结构如图 1‑2所示。CCTV检测是目前地下管道应用最普遍的检测系统，其由操作人员控制地下管道机器人对管道进行实况拍摄，而后交由专业检查人员进行判定，其结构原理如图 1‑3所示。CCTV巡检方式成本低、效率高，能直观的检测管道结构性和功能性缺陷[6]。

由于地下管道缺陷种类多样、背景复杂多变，传统图像处理算法远无法满足实际场景检测的需求，目前世界范围内没有成熟的管道CCTV视频自动检测技术。在学术领域，也有大量的学者致力于地下管道自动检测算法的研究，但目前所能达到的水准与实际场景检测仍存在较大的差距[7]，而人工检测存在低智能化、主观性强，易疲劳等缺点[8]。近年来，深度学习迅猛发展以及图片数据的易获取性使得管道视频的自动检测可行性大大增加。基于大量数据训练的卷积神经网络模型结合图像处理算法能够实现一定程度的自动检测，不仅能减少操作人员的劳动强度和人工经验的依赖程度，更提高了检测效率和信息描述的一致性，具有很高的研究和应用价值。

|  |
| --- |
| 图 1‑1超声导波检测原理[3] |
| 图 1‑2激光扫描三维重建检测[5] |
| 图 1‑3 CCTV检测 |

## 国内外研究现状

### 视觉管道检测技术研究现状

由于近年来机器视觉技术的高速发展，其在许多工业自动化场景有了广泛的应用[9]。目前，也有大量的学者、工程师们致力于管道缺陷检测视觉算法的研究。工业场景一般具有较好的打光条件，较小的场景变动性，而管道检测由于其背景复杂、缺陷种类多样，检测难度相比于模式比较固定的工业化场景大大增加，单纯地使用传统的图像处理算法远不能达到高准确率的识别要求。现今已有的中外文献中的算法大多只能针对小场景变动范围内的特定缺陷进行识别或对已知存在缺陷的图像进行缺陷分类，还远不能应用于实际管道场景检测。

Sunil等人通过一系列以腐蚀和膨胀为基础的形态学操作对管道缺陷部分进行分割，提取分割后缺陷的形态学特征，釆用模糊神经网络对图像进行识别[10]，分割效果如图 1‑4所示。但该算法在复杂管道场景下容易受到背景干扰导致分割成功率低，同时模糊神经网络算法单帧处理时间长导致检测效率差。北京信息科技大学王鹏等人对管道图像进行预处理后，提取管道缺陷图像的边缘特征，然后釆用人工神经网络对缺陷进行分类[11]。但该方法只能对人工识别出来的缺陷帧进行缺陷分类而无法从大量视频帧提取缺陷帧，整体智能化程度低。同时人工神经网络基于反向传播原理进行训练，无法保证能够收敛到全局最优解[12]。Hengmeechai等人提出了一种树枝侵入缺陷检测算法，通过与正常管道图像差值提取疑似缺陷区域，而后通过基于方向梯度直方图特征的支持向量机算法对疑似缺陷区域进行识别，能够达到86%的准确率和缺陷召回率[13]，但该文献仅针对一种缺陷进行实验，不具备通用性。

|  |
| --- |
| 图 1‑4形态学分割效果图[10] |

Halfawy等人[14]提出了完整的地下管道视频检测框架，是目前学术领域所见智能化程度最高的检测方法。首先通过光流法[15]分析管道机器人及携带摄像机的运动状态和位姿变化，提取行进速度低于一定阈值的视频段为疑似异常段，具体如图 1‑5所示。而后通过haar-like特征加上级联分类器[16]对疑似缺陷帧进行视角分类。不同视角会存在不同种类的缺陷，针对所判定的视角设计对应缺陷的检测算法。同时，利用光流向量结合管道总长度判断缺陷发生地和管口之间的距离。该方法实现了端到端的检测方式，即输入管道检测视频即可输出该段管道是否存在缺陷、存在何种缺陷及缺陷发生位置等信息。然而，该方法成立的条件是操作人员在碰到疑似缺陷发生处会放慢管道机器人行进的速度进行细节拍摄，但这一假设在很多管道检测场景是不成立的。同时，该方法针对错口、浮渣等缺陷设计的图像处理算法在场景变化比较大的情况下失效的概率很高。

|  |
| --- |
| 图 1‑5光流分析提取异常视频段[14] |

上述检测方法大多通过图像处理算法对缺陷进行分割，之后通过特征描述子对特征进行量化描述，最后通过机器学习的分类器进行缺陷分类。这类方法在管道检测中存在很大的局限，一是只能实现半自动化，缺少从大量正常视频中寻找缺陷段的过程；二是管道场景变化复杂，使用图像处理算法进行缺陷分割会很大程度的受到背景干扰，导致分割成功率不高；三是管道缺陷种类多样且同种缺陷在颜色、纹理、形状等特征上也会产生很大的变化，传统的机器学习模型的识别能力不能满足要求。而Halfawy等人依据摄像机行进速度进行缺陷视频段提取的方式在大部分实际检测中并不适用。因此，目前已有的算法无法适应真实的管道场景检测。

### 图像识别算法研究现状

图像识别是管道检测的关键技术，无论是正异常检测还是异常分类都很依赖于检测系统在复杂场景下对当前视频帧的识别能力。目前，机器学习算法被大量的应用于图像识别中。近年来，深度学习技术的的日渐成熟使得其在一些方面达到了优于传统机器学习算法的性能。然而，深度学习还远远没有达到取代传统机器学习方法的程度，后者在许多方面也存在着前者无法具备的优势。因此，对一项视觉任务而言，针对其特性设计最合适的算法才能达到最优的综合性能。

Epaillard等人[17]考虑到管道缺陷具有不确定的模式和构成方式以及异常出现概率低的特点，将缺陷定义为与排污管道CCTV视频中的正常模式的背离。同时考虑到管道正常样本的数量的远大于缺陷样本，他提出了以隐马尔可夫模型为基础的管道检测异常模型，用正常帧对模型演练，从而能够提供其常态。低于预定阈值较大可能性的任何其他视频序列将被分为异常和潜在缺陷。该算法充分利用了管道缺陷样本占总样本比例较小的特点，但是仅依赖正常帧信息而未充分利用异常信息，会导致在实际检测中存在较大的误检率。Joshua Myrans 等人[18]利用Gabor滤波器能够分析信号局部时频特性的特点，引入时间局部化的窗函数对信号进行时频分析。采用不同尺度、不同方向的Gabor滤波器对分块图像提取特征对图片进行描述，具体如图 1‑6所示。将提取到的特征通过分类器进行识别分类，采用的是一种由决策树按一定策略集成生成的分类器——极端随机树。这种算法是随机森林算法[19]的一种变种，在其基础上引入了更强的随机性，在变动性强的管道场景具有不错的性能。

|  |
| --- |
| 图 1‑6多尺度、多方向Gabor滤波器特征提取[18] |

深度学习算法中的卷积神经网络在视觉任务中表现出不俗的性能，也越来越多的被应用在图像识别领域。卷积神经网络能够通过一层层的卷积池化层提取到相比手工特征更具泛化性、抽象性的特征，从而实现更好的识别分类效果[20]。2012年Krizhevsky等人[21]在其提出的AlexNet模型上引入了Relu非线性激活函数和Dropout概念大大提高了模型在ImageNet数据集上的准确率。Simonyan和Zisserman采用多个小卷积核代替大卷积核，同时增加了模型的深度进一步提高了模型的识别能力和识别准确率[22]。2014年，Christian Szegedy等人[23]提出了GoogleNet，在之前的基础上提出了inception结构，使用不同大小的卷积核进行特征提取并进行多尺度拼接，进一步提升了模型的理解能力。另外，在之后的改进中研究者们引入了批规范化[24]的方法，对中间层输出进行归一化加快了模型的收敛使模型可训练深度进一步提高。2015年，微软亚洲研究院的何凯明等人提出了ResNet，除了正常的非线性映射之外，还设置了一条跳跃连接在一定程度上解决了随着网络层数增加导致的梯度消失问题，使模型深度可达到一百多层，大大增加了模型的理解能力和识别准确率[25]。

在大部分学者将目光聚焦在提升网络识别能力的同时，部分学者致力于深度学习网络的轻量化的研究。Song Han等人[26]通过剪枝、量化、霍夫曼编码等技术大大压缩了模型的参数量，提高了深度学习算法在内存受限设备上的可用性。FN Iandola等人[27]提出了SqueezeNet，将一层卷积层拆分成挤压层和扩张层，挤压层通过11卷积减少参数量，扩张层通过11卷积和33卷积结合多尺度特征，能够在保证与AlexNet相近的准确率的同时将模型大小压缩为原来的。Andrew G等人提出了MobileNet，利用深度分离卷积技术获取模型的稀疏化表达，能够将计算量压缩至正常卷积的八到九分之一，大大减小了前向传播所需要的时间[28]。

深度学习的出现不断地提高了人工智能算法在计算机视觉任务中的基线，也为管道复杂场景的识别带来了更大的可行性。但即使经过大量模型压缩和加速的研究，深度学习算法的计算量依然制约着其在许多领域的落地与应用，而这也会直接影响到管道的检测效率。当然，深度学习目前也处在高速的发展之中，相信深度学习在未来会更多地落地到生活场景中去。

## 本文主要研究内容

本课题的目的是利用计算机视觉技术构建一个一体化的智能管道缺陷检测系统，自动地检测输入视频中是否存在缺陷、存在何种缺陷以及缺陷发生位置。同时，要保证在实际场景检测中有实时性好、过检率低、准确率高等特性。其主要研究内容有以下几点：

（1）设计合理的检测方案，以满足管道缺陷检测的功能需求和性能需求。

根据管道数据正异常样本分布不均，真实管道检测中存在特殊视角干扰等特点以及缺陷检测和缺陷信息识别等功能需求，设计合理的整体方案及检测流程。

（2）设计缺陷检测算法，并优化其时间效率和准确性。针对实际管道变化性强，缺陷种类多样的特点，设计具有较强识别能力以及泛化能力的检测算法。同时为适应实际管道场景检测需求，需要对算法的准确性以及实时性进行优化。

（3）设计伪异常帧过滤算法，提高实际检测流畅性以及特殊情况处理能力。由于管道地形变化及机器人本体的位姿变化，在运行检测过程中会出现大量的伪异常帧（由于摄像机大幅偏摆及抖动而非管道本身缺陷产生的异常帧），从而导致较高的过检率。因此，需要设计一定的算法将伪异常帧从检测序列中剔除，降低在实际场景中的过检率。

（4）设计视频文字检测识别算法，实现缺陷信息的自动识别和记录。为了进一步提高缺陷检测系统的智能性，当某视频帧存在缺陷时会自动识别并记录缺陷管道所在地段及缺陷处与管口之间的距离。由于这些信息以视频文字的形式记录在视频中，本系统需要针对视频文字设计适应本场景的字符定位、分割、识别算法来自动记录缺陷位置信息。

# 管道缺陷检测方案设计及方法研究

## 引言

管道缺陷检测系统是针对现实场景管道视频的一体化智能检测系统，包括缺陷检测、伪异常帧过滤、缺陷信息识别记录及缺陷报告生成等功能。本章根据检测缺陷类别及功能需求设计整体方案，包括算法处理流程、异常处理策略、方法性能对比、模型构建及训练。同时，由于管道检测领域缺乏公共数据集，为综合评估本文检测方法的优劣，本章复现了管道检测领域论文中检测效果较好的两种方法，分别是基于方向梯度直方图特征的支持向量机模型和基于多尺度、方向Gabor滤波器[29]特征的极端随机树模型。利用本文所构建的训练集进行训练，并在测试集上对三种方法进行直观性能对比。

## 管道缺陷检测方案设计

### 地下管道检测缺陷类别及功能需求分析

地下管道因为用途复杂、材料各异会产生很多缺陷种类，根据其性质大体可分为结构性缺陷和功能性缺陷[30]。典型的功能性缺陷为以下几种：（1）障碍：管道内存在大型障碍物影响管道流通，如图 2‑1 a) 所示；（2）沉积：杂质在管道底部沉淀淤积，如图 2‑1 b)所示；（3）气体：管道内存在某种弥散气体，如图 2‑1 c)所示；（4）树根：大量树根生长进入管道，如图 2‑1 d)所示。典型的结构性缺陷为以下几种：（1）破裂：外部压力过大导致管道产生裂缝、破损，如图 2‑1 e)所示；（2）接口破损：管道连接处发生破损，如图 2‑1 f)所示；（3）异物穿入：尖锐的物体（非管道自身材料）穿入管道内，如图 2‑1 g)所示；（4）变形：管道所受挤压力过大而产生形变，如图 2‑1 h)所示。

根据某管道机器人公司提供的检测视频数据发现，障碍、沉积、异物穿入、气体最为常见，考虑到缺陷数据的获取情况，本课题所构建的检测系统只针对上述几种常见缺陷。另外，本方法为端到端的检测方法，即输入管道视频即可输出该段管道的缺陷分析报告，包括管道是否存在缺陷、存在缺陷种类、缺陷发生地段、缺陷距离管口位置等信息。真实管道场景存在背景复杂多变缺陷种类多样等特点，这导致目前存在的视觉检测算法远无法达到现实场景落地的水平，这也对本课题的检测方法提出了很大的挑战。因此，为适应实际需求，本文提出的检测方法必须满足如下条件：

（1）具有较高的准确性和泛化能力，以适应真实管道场景；

（2）具有较好的实时性以提高检测效率；

（3）具有特殊视角处理能力以提高检测流畅性；

（4）具有自动管道地段信息及缺陷位置信息识别能力以提高检测智能性。

|  |
| --- |
| a) 障碍物  b) 沉积  c) 气体  d) 树根  e) 破裂  f) 接口破损  g) 异物穿入  h) 变形  图 2‑1管道缺陷样本 |

### 管道缺陷检测方案制定

本节根据管道检测特点及功能需求，合理地制定功能模块和整体方案。针对高准确性和泛化能力要求，仅依靠图像处理算法无法适应场景变化性，因此本方法核心检测任务应由基于统计特性的模式识别和图像处理算法的融合算法完成。针对检测方法实时性要求，考虑到视频帧的连续性可采取相隔固定帧采样检测方法，同时需要针对算法时间效率进行优化。针对特殊视角处理能力及流畅性要求，由于管道地形变化和人为控制，机器人在行进过程中会产生大幅振动和视角偏摆，这会导致出现一些正向和侧向以外的非常规视角，这一类视角通常以一种运动过渡状态短暂地出现在视频段中，但通常会被常规检测方法误判为异常。本文将这种由于镜头振动和大幅偏摆而非管道本身缺陷造成的异常帧称之为伪异常帧，如果不设计算法应对这种特殊情况会造成大量的误判，影响到整体检测的流畅性。因此，本文需要设计一种伪异常帧过滤算法，防止这类视频帧造成整体过检率偏高的问题。针对缺陷信息自动识别记录要求，由于每一段视频所检测管道属于特定地段，且每一处缺陷发生位置距离管口有特定的距离，如果能够在检测到缺陷的同时自动地识别并记录缺陷管道地段信息、缺陷发生距离信息，这将提高检测方法的效率和智能性。而这些信息通常以字符的形式出现在视频当中，因此针对管道场景设计合适的视频字符检测识别算法是本文不可或缺的一部分。

另外，获取的管道数据集中同一类型异常样本远少于总体正常样本，考虑到数据分布不均衡会影响检测准确率[31]，本文将管道检测划分为视频帧正异常检测和异常分类两个阶段。上述划分存在两个优势，一是将正常样本分为一类，所有类型的异常样本分为一类，这能改善数据的均衡性，使训练模型有更好的性能。二是正异常检测面向所有视频帧，而异常分类仅面向疑似存在异常的视频帧，两者无论从面向对象的出现频率还是检测任务的难度来看都是存在差异的，分离之后可对这两级检测进行分别处理。准确性是缺陷检测方法的首要指标，但正异常检测方法因其面向所有视频帧，考虑到整体检测效率需要兼顾实时性。而缺陷分类任务因其面向对象占总体比例小，实时性要求不高。同时，两阶段划分之后可以在标注数据时带有一定的“偏向性”。相比于准确率这个指标，在缺陷检测方法中，缺陷的召回率更为重要，即找出缺陷的数量占缺陷总数的比例。因此，在正异常检测阶段数据标定中只要样本与正常管道形态特征存在微小偏差即标定为疑似异常样本，这会提高缺陷召回率，但同时也会带来大量的正常样本被判定为异常，导致整体过检率较高。为解决上述问题，在第二阶段异常分类中同样设立正常类别，考虑到数据均衡问题正常类别样本数量需与其他缺陷类别相近，故选取被第一阶段错判为异常的正常样本作为本阶段训练样本。这能对被上一级检测器认定为异常样本的正常样本进行结果更正，在一定程度上降低过判率。总而言之，通过上述方法将一级检测划分成二级检测，能够针对两阶段特性进行分别处理，同时在构建两阶段训练样本时加入一定的偏向性能在保证缺陷检出率的同时大幅挺高系统的检测效率。

同时，本系统还提供了一种现实场景可落地性更强的检测方式，先由正异常检测模型联合光流分析法对视频进行缺陷帧筛选，筛除正常帧，而后由人工对异常帧进行复验及缺陷分析。经过前一阶段处理，视频时间能被压缩至原视频的30%左右，这能在保证很高缺陷检出率的同时大幅减小专业检测人员的干预时间，提高检测效率。综合上述分析，得到本方案的整体流程见图 2‑2。

|  |
| --- |
| 图 2‑2 检测流程图 |

### 管道缺陷检测性能评估指标分析

准确率是评估分类效果的最常用的指标，但不同场景有不同的需求，当正负样本错分时产生的代价不同时，则不能一味的通过准确率来衡量模型的性能。根据场景的不同，精确率、召回率以及F1系数有时也会被作为首要评估指标。统计测试集中真值为正的样本被预测为正样本的数量为TP，真值为负的样本被预测为负样本的数量为TN，真值为负的样本被预测为正样本的数量为FP，真值为正的样本被预测为负样本的数量为FN，则准确率公式见( 2‑1 )，精确率公式见( 2‑2 )，召回率公式见( 2‑3 )、F1系数公式见( 2‑4 )。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |

本文将异常样本设为正类，将正常样本设为负类，因此将正类错判为负类会导致缺陷漏检，将负类错判为正类会造成过判，相比之下前者造成的损失更大，且后者可通过二次检测进行排除。正类召回率的直观意义是检测出的缺陷占缺陷总数的比例。因此，召回率和准确率对于管道缺陷检测方法来说都是重要的衡量指标。然而，当两种检测方法分别在召回率和准确率上具有优势时，则需要一个更为综合的指标去判断优劣。 ）可以综合地评估一个分类器性能的优劣，该曲线以假阳率（）为横坐标，真阳率()为纵坐标[38]，计算公式分别见( 2‑5 )、( 2‑6 )所示：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |

将所有样本按照预测为正类的概率由高到低进行排列，分别取不同的阈值作为划分正负样本的边界，并求出对应的真阳率和假阳率即可得到坐标系中的点集，由得到的点集构成曲线。曲线只能定性的描述分类模型的综合性能优劣无法给出量化结果，故采用曲线下的面积作为模型的定量综合评估指标。

## 基于传统机器学习的管道缺陷检测方法

### 支持向量机管道检测方法

能否对分类图像进行针对性的特征提取是影响支持向量机检测方法性能的重要因素。本文采用特征，又称方向梯度直方图特征，其统计图像局部区域内的边缘梯度和方向，并按照一定的规律将其构建成梯度方向直方图，最终形成特征描述子[32]。特征提取流程和原理示意分别如图 2‑3、图 2‑4所示。

|  |
| --- |
| 图 2‑3 HOG特征构造流程图 |
| 图 2‑4 HOG特征提取原理图 |

（1）图像归一化 不同管道图像场景不同、光照条件不同，首先利用图像归一化操作来消除亮度差异,减少纹理失真，抑制噪声干扰，利用Gamma校正算法[33]，如公式( 2‑7 )所示，其中为处理前图像，为处理后图像。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

（2）计算图像梯度 图像归一化之后需要求取每一个像素的梯度幅值和方向，利用算子[34]与图像进行卷积得到得到像素水平梯度和竖直梯度，其计算公式分别见( 2‑8 )、( 2‑9 )，在通过公式( 2‑10 )、( 2‑11 )计算其梯度幅值和方向。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |

（3）统计的特征 将之间的梯度方向以为间隔划分成个，将梯度方向落在某一个中的所有像素以其像素梯度幅值为权值求和作为此的特征值，每个可以得到一个维特征向量。

（4）统计的特征 个组成一个，将内所有的梯度方向直方图归一化后进行整合即可得到的梯度方向直方图。

（5）统计全图特征 以为单位以固定步长在图像上滑动，将每一个滑动状态的特征进行组合即可得到全图的梯度方向直方图特征向量。本文取大小为个像素，大小为个像素，即个，滑动步长为，则对于的图像，可以得到维的特征向量。分别对正常、异常样本进行HOG特征可视化操作即可得到图 2‑5、图 2‑6。

|  |
| --- |
| 图 2‑5正常样本HOG特征可视化 |
| 图 2‑6 异常样本HOG特征可视化 |

提取了管道图像的HOG特征后，需要对这些特征进行分类训练，以判断该管道图片是否存在缺陷。本文采用（，支持向量机）进行分类，该算法本质就是用一个最大间隔的超平面将不同类别数据点分隔开来[35]，如图 2‑7所示。

|  |
| --- |
| 图 2‑7最大间隔超平面示意图 |

假定最大间隔为，则的表达式见公式( 2‑12 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中 ——正向支持向量；

——负向支持向量；

——分隔平面范数。

为了使最大，需要最小化，可通过公式( 2‑13 )的拉格朗日方程求最小化得到最优解，其中是拉格朗日乘子。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

很多分类场景数据存在线性不可分的情况，即在原始数据空间中无法利用线性超平面将不同类别的数据分隔，此时需要利用核函数将原有低维数据映射到高维空间，在高维空间中寻找最大间隔分隔平面。高维映射后SVM算法的通用表达式为( 2‑14 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

常用的核函数有线性核函数、多项式核函数、径向基核函数、核函数，其公式分别为( 2‑15 )、( 2‑16 )、( 2‑17 )、( 2‑18 )所示。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |

其中线性核函数在线性可分的情况下准确率和效率较高；多项式核函数参数较多寻优过程复杂，同时如果较大会带来较大的计算量；核函数使得具有类似神经网络的分类特性；径向基核函数只有一个可调节参数，可将原参数空间映射到无穷维，在大部分情况下效果较好。本文分别对四种核函数进行网格参数寻优，采用交叉验证法确定最佳参数，得到调优之后的四种核函数性能如表2-1所示。

表 2‑1 支持向量机测试结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型 | 核类型 | 核参数 | 召回率 | 准确率 |
| SVM-1 | 线性 | — | 0.79 | 0.714 |
| SVM-2 | Sigmoid | =1e-2, =0.0 | 0.75 | 0.707 |
| SVM-3 | 多项式 | =3,=0.0 | 0.86 | 0.714 |
| SVM-4 | 径向基 | =1e-4 | 0.81 | 0.737 |

由数据可知，采用径向基核函数的模型具有最佳准确率，采用多项式核函数模型具有最佳召回率，但总体来说支持向量机模型在正异常检测任务上的效果并不可观。

### 极端随机树管道检测方法

滤波器相对傅里叶变换能够提供更好的方向选择和尺度选择性,抽取空间局部频度特征，引入时间局部化的窗函数对信号进行时频分析，具体表达见公式( 2‑19 )。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中为滤波器波长，为滤波器方向，为相位偏移，空间纵横比，为滤波器带宽。对识别图像进行分块，对每一块区域采用多个尺度、多个方向的滤波器进行特征提取，将这些特征进行拼接得到该识别图像的特征描述。其中的值分别取3，5，7，9，的取值为之间增量为的值，这样就能对图像的各个时频率特征进行提取，对一张分块图像，得到的特征的维数为，其中每一个滤波器对分块图像滤波后得到的是一个二维数组，取这个二维数组的平均值（平均模式）或最大值（最大模式）作为特征的一个值，具体分块方式及滤波器形态如图 2‑8所示。

|  |
| --- |
| 图 2‑8 GIST特征提取示意图 |

极端随机树是随机森林的一种变体，随机森林的本质是机器学习中的集成学习方法，其基本决策单元是决策树[36]。决策树通过给定的训练数据集学得一种决策机制，而后利用这种决策机制对未知数据进行分类，算法流程见表2-2。

表 2-2 决策树算法流程

|  |
| --- |
| **算法1**：决策树构建 |
| **输入**：训练集； |
| 属性集 |
| **过程**：函数  生成节点；  分别以中每一个属性作为划分特征，利用公式( 2‑20 )计算信息增益，并以此为依据选取最优划分特征；  的每一个值  为生成一个分支，用表示中在上取值为的样本集；  为空  将该节点标记为叶子节点，属于中样本最多的类别；      以为分支节点； |
| **输出**：针对训练数据的决策树 |

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中 ——样本集合；

——样本集合中属性的值为的集合；

——样本类别总数；

——第类样本占总样本比例。

随机森林通过引入样本随机扰动和属性随机扰动来构建许多决策树，而后选取投票数最大的决策作为分类结果。而极端随机树在随机森林的基础上又引入了一个随机性，在每次进行最优分裂特征选择时不选择信息增益最大的特征，而是随机选取一个特征作为分裂特征，在管道检测这种变化性很大的场景能够获得优于随机森林的检测性能。

对滤波器特征和极端随机树分类器参数进行调优，确定决策树数量为，滤波器波长为，滤波器空间纵横比为，得到测试结果见表2-3。

表 2‑3 极端随机树测试结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 模型 | 图像尺寸 | 模式 | 滤波器尺寸 | 召回率 | 准确率 |
| ET-1 | 112 | 平均 | 11 | 0.70 | 0.693 |
| ET-2 | 112 | 最大 | 11 | 0.82 | 0.776 |
| ET-3 | 224 | 最大 | 19 | 0.81 | 0.765 |

由上述信息可知，采用最大模式相对平均模式模型性能有较大的提升，

同时在图像尺寸达到112之后继续增大图像尺寸增加GIST特征维数对性能并无提升且会降低计算效率。由于管道场景的变动性较强，采用极端随机的树模型较支持向量机模型有较大的提升，但仍然不太可观。

## 基于卷积神经网络的管道缺陷检测方法

### 卷积神经网络分析

不同于传统的神经网络结构，卷积神经网络是一种权值共享模型，其结构如图2-1所示。它由多个相邻的卷积层和池化层组成（图 2‑9中仅画了两层），每一层都是由多个特征图组成。其中，卷积层是由多个不同的卷积核与前一层每一个特征图卷积得到。假设第层有个特征图，卷积核的数量为，则卷积操作之后得到的第层卷积层有个特征图，其中第个特征图计算公式见( 2‑21 )，卷积神经网络示意见图 2‑9。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中 ——激活函数；

——卷积滤波器的偏置量；

——用于生成，针对前一级第i个特征图卷积核。

|  |
| --- |
| 图 2‑9卷积神经网络结构示意图 |

池化层通过一定规则的降采样能够缩小特征图尺寸，降低模型计算量，使提取的特征具有平移、尺度、旋转不变性，使得模型在预测阶段具有更强的鲁棒性。典型的池化方式有最大池化和平均池化，对于来自前一卷积层的某特征图，以表示其第个大小的局部下采样区域，在池化层得到输出特征图，的第个节点处的输出值的最大池化、平均池化计算方式分别见公式( 2‑22 )、( 2‑23 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |

卷积神经网络利用极小化实际输出与理想输出的差以反向调整卷积网络的参数，其局部权值共享性不仅降低了神经网络的复杂度，同时使其对尺度缩放、平行移动、旋转等各种形式变化具有较高的不变性。卷积神经网络的诞生掀起了计算机视觉领域的浪潮，大大提高了计算机在图像识别、目标检测、语义分割等多项视觉任务中的性能。

### 正异常检测方法

从上述实验中可以看到支持向量机模型和极端随机树模型在实际场景中的准确性都不太可观。因此必须构建识别性能更好的检测模型，卷积神经网络便具备这种特性。对地下管道视频进行采样，将采样得到的图片作为训练样本，部分正常和异常样本如图 2‑10、图 2‑11所示。

|  |
| --- |
| 图 2‑10正常样本 |
| 图 2‑11 异常样本 |

目前，深度学习算法大多运行在专用GPU上，对于一般工业场景工控机而言，复杂度过高的模型会导致单帧计算时间过长，导致系统实时性差。由于正异常检测面向所有视频帧，因此需要综合考虑准确性和实时性。本文分别使用SqueezeNet[27]和InceptionV3[37]在构建的数据集上进行训练，根据两者在测试集上的综合性能初步确定正异常检测模型结构。

SqueezeNet是一个轻量化卷积神经网络，不同于传统的卷积神经网络，该网络提出了Fire module结构，将一层卷积拆分为挤压层和扩张层。挤压层采用11的卷积核降低上一层的特征图数起到了参数压缩的作用，扩张层分别使用11和33的卷积核进行卷积同时在特征图维度进行融合，使模型获取不同感受野的信息。Fire module的结构如图 2‑12所示，SqueezeNet网络参数见表2-4，其中S代表挤压层，E代表扩张层。而InceptionV3是一个大型深度卷积神经网络，其采用不同大小的卷积核获取不同大小的感受野，而后通过拼接将不同尺度特征融合。同时其采用两个33的卷积层代替一个55的卷积层，使用不规则的13卷积和31卷积代替33卷积。上述策略使得该模型具备很强的学习能力，但在加深网络的同时也增加了其前向传播计算量，其具体结构参数及模型构建原理见2.4.3节，本节不做详细阐述。

|  |
| --- |
| 图 2‑12 Fire module结构 |

表 2‑4 SqueezeNet网络结构参数

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 网络层 | 输出尺寸 | 卷积核大小/步长 | S11通道 | E11通道 | E33通道 |
| Input | 224×224×3 | — | — | — | — |
| Conv1 | 111×111×96 | 7×7/2 | — | — | — |
| Maxpool1 | 55×55×96 | 3×3/2 | — | — | — |
| Fire2 | 55×55×128 | — | 16 | 64 | 64 |
| Fire3 | 55×55×128 | — | 16 | 64 | 64 |
| Fire4 | 55×55×256 | — | 32 | 128 | 128 |
| Maxpool4 | 27×27×256 | 3×3/2 | — | — | — |
| Fire5 | 27×27×256 | — | 32 | 128 | 128 |
| Fire6 | 27×27×384 | — | 48 | 192 | 192 |
| Fire7 | 27×27×384 | — | 48 | 192 | 192 |
| Fire8 | 27×27×512 | — | 64 | 256 | 256 |
| Maxpool8 | 13×13×512 | 3×3/2 | — | — | — |
| Fire9 | 13×13×512 | — | 64 | 256 | 256 |
| Con10 | 13×13×2 | 1×1/1 | — | — | — |
| Avgpool | 1×1×2 | 13×13/1 | — | — | — |

经过训练和调优之后，在测试集上对SqueezeNet、InceptionV3、支持向量机以及极端随机树模型进行性能测试，得到对比数据如表2-5所示，绘制得到ROC曲线如图 2‑13所示，测试电脑的CPU型号为Intel（R） Core （TM）i5-540M。可以看出，卷积神经网络SqueezeNet和InceptionV3相比于支持向量机和极端随机树在准确性上有较大幅度的提升，绘制的ROC曲线也能够完全包络后两者方法，整体性能较优。而相较两种卷积神经网络，InceptionV3准确性较高但单帧处理时间过长，而SqueezeNet准确性相对较低但具备时间优势。正异常检测面向所有视频帧，检测效率和准确性都是重要指标，可根据功能需求侧重进行选择，但显然两者都不是最优检测方法，后续需要对该阶段检测方法进行优化。

表 2‑5 模型性能对比数据

|  |  |  |  |
| --- | --- | --- | --- |
| 模型 | 召回率 | 准确率 | 时间（s） |
| 支持向量机 | 0.81 | 0.74 | 0.112 |
| 极端随机树 | 0.82 | 0.78 | 0.165 |
| SquezeNet | 0.91 | 0.87 | 0.265 |
| InceptionV3 | 0.94 | 0.91 | 0.876 |

|  |
| --- |
| I:\MyFolder\大论文\Image\roc1-1000.jpg  图 2‑13 各模型ROC曲线 |

### 异常分类方法

在一段检测视频中，并非所有的视频帧都属于正向视角，因为检测需要操作人员会控制摄像机镜头偏摆对侧向管壁进行拍摄。同时，大部分管道检测公司会在视频的开始段和结束段插入信息页面显示该管道地段、材质、检测人员等相关信息。另外，由于上一级检测模型更加注重异常样本召回率而对过检率相对宽松，这会导致较多的正常样本被前一级模型认定为异常。为了对错检样本进行二次纠正，需要在本级检测中设立正常样本类别。但考虑到训练数据的均衡性，需要对正常样本进行欠采样。因此，在本级检测中将训练样本划分为三个大类别：（1）正常样本，见图 2‑14 a)；（2）异常非缺陷样本：①侧向样本，见图 2‑14 b)；②信息样本，见图 2‑14 c)；③井口样本，见图 2‑14 d)；（3）缺陷样本：①障碍样本，见图 2‑14 e)；②沉积样本，见图 2‑14 f)；③异物穿入样本，见图 2‑14 g)；④气体样本，见图 2‑14 h)。

|  |
| --- |
| a) 正常  b) 侧向  c) 信息  d) 井口  e) 障碍  f) 沉积  g) 异物穿入  h) 气体  图 2‑14 异常分类样本图 |

由于异常分类检测难度高于正异常检测，且异常分类仅针对上一级判定为异常的视频帧而正异常检测针对每一帧视频，故前者出现频率远小于后者。基于上述特性，本级检测对模型的理解能力要求更高而对模型时间效率要求放宽，故使用InceptionV3模型作为基础网络。InceptionV3的核心在于三种Inception模块，其中InceptionA结构如图 2‑15所示。

|  |
| --- |
| 图 2‑15 InceptionA结构图 |

结构有四个分支，第一个分支对输入进行11的卷积，它可以跨通道组织信息，提高网络的表达能力，同时可以对输出通道升维和降维，可以看到四个分支都用到了11卷积来进行低成本跨通道的特征变换；第二个分支使用了33的最大池化后加上11卷积，能够增强提取特征的平移、旋转、尺度不变性；第三个分支在11卷积之后使用33卷积来获取更大感受野特征；第四个分支采用两个33卷积代替55卷积，在获取大感受野特征的同时降低参数数量，增强了特征的非线性程度。四个分支在最后在特征图数量维度进行聚合，使得模型获得了多尺度融合信息。另外InceptionB和InceptionC结构分别如图 2‑16、图 2‑17所示。

|  |  |
| --- | --- |
| 图 2‑16 InceptionB结构图 | 图 2‑17 InceptionC结构图 |

在另外两种inception结构将nn的卷积拆分成1n的卷积和n1的卷积来进一步减少参数增加非线性，完整网络结构参数见如表2-6。

表 2‑6 InceptionV3网络结构

|  |  |  |
| --- | --- | --- |
| 网络层 | 输入尺寸 | 卷积核大小/步长 |
| Conv | 299×299×3 | 3×3/2 |
| Conv | 149×149×32 | 3×3/1 |
| Conv padded | 147×147×32 | 3×3/1 |
| Pool | 147×147×64 | 3×3/2 |
| Conv | 73×73×64 | 3×3/1 |
| Conv | 71×71×80 | 3×3/2 |
| Conv | 35×35×192 | 3×3/1 |
| 3×InceptionA | 35×35×288 | 见图2-15 |
| 5×InceptionB | 17×17×768 | 见图2-16 |
| 2×InceptionC | 8×8×1280 | 见图2-17 |
| Pool | 8×8×2048 | 8×8 |
| Linear | 1×1×2048 | 1×1/1 |
| Softmax | 1×1×8 | — |

卷积神经网络的训练依赖于反向传播原理和梯度下降法[39]，而由于卷积神经网络反向传播算法的特性，随着网络的加深模型训练难度加大。卷积神经网络前向传播公式如下：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中为激活函数，为第层激活输出，为第层卷积核，为第层偏置系数。假定为损失函数，则对和的分别为偏导为：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |

由于取决于使用的损失函数，取决于使用的激活函数，使用不同的损失函数和激活函数会有不同的结果，为方便推导统一记为。由上述公式可知只要求出了某一层的就可以求解出和的梯度，而则由第层的递推而来：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

由于网络初始化时采用均值为0，方差很小的高斯分布初始化，因此都是一些接近0的小数值，随着网络向前传播的值会越来越小，出现梯度消失导致网络无法训练。为改善上述状况，对网络结构进行调整，加入跳跃连接，具体结构如图 2‑18所示。

|  |
| --- |
| 图 2‑18反向传播原理示意图 |

正常卷积模块梯度通过上方Inception多重卷积层反向传播时梯度要乘以的多次方造成梯度大幅衰减，随网络加深这一现象愈加显著导致无法正常训练。而加入了33卷积的跳跃连接之后，可通过下方跳跃连接层经少量衰减后直接反向传播给上一层，很大程度缓解了梯度消失现象，使网络更好地收敛。经过训练之后，将得到的固化模型在测试数据集上进行测试，并与其他论文方法进行性能对比。其中，极端随机树在正异常检测任务中性能优于支持向量机，但在异常分类任务中性能较差，故不将数据列出。支持向量机模型和本文模型测试数据分别见表2-7、2-8。

表 2‑7 异常分类SVM测试结果

|  |  |  |  |
| --- | --- | --- | --- |
| 缺陷类型 | 精确率 | 召回率 | F1系数 |
| 沉积 | 0.62 | 0.67 | 0.65 |
| 气体 | 0.90 | 0.78 | 0.83 |
| 异物穿入 | 0.67 | 0.34 | 0.45 |
| 障碍物 | 0.54 | 0.49 | 0.52 |
| 合计 | 0.67 | 0.67 | 0.66 |

表 2‑8 异常分类InceptionV3测试结果

|  |  |  |  |
| --- | --- | --- | --- |
| 缺陷类型 | 精确率 | 召回率 | F1系数 |
| 沉积 | 0.79 | 0.68 | 0.73 |
| 气体 | 0.81 | 0.95 | 0.87 |
| 异物穿入 | 0.94 | 0.92 | 0.93 |
| 障碍物 | 0.71 | 0.82 | 0.76 |
| 合计 | 0.83 | 0.83 | 0.83 |

由于管道场景变化复杂，同种缺陷在形态、颜色、位置等特征上变动性也较强，故检测难度较大。本文模型相对于支持向量机模型在性能上有了较大的提升，在管道实际场景自动化检测中向前迈出了一步，但距离实际使用仍有一定差距。

## 卷积神经网络模型训练

### 数据增强

深度学习算法性能优越的同时也存在一定的弊端，弊端在于其对训练样本的依赖，如果不能获得大规模、多场景的训练数据，算法的性能就会受到一定的制约。根据视觉任务所属领域的不同，有些领域的数据较难获取且不具备可重构性，这时有效的数据增强手段就能在一定程度上提升训练模型的性能[40]。管道检测领域不存在大型公共数据集且数据获取的成本较大，因此本文针对管道检测场景使用了如下数据增强方法：

（1） 水平翻转 能够在增加样本数量的同时引入模型对水平视角转换的适应能力，具体如图 2‑19 b)所示。

（2） 旋转变换 管道机器人在管道行进过程中无法一直保证水平视角，随着管道地形的变化会发生一定的偏转。而引入小幅度的正负向旋转增强既能够增加样本的多样性也能提高模型对偏转视角的适应能力，具体如图 2‑19 c)所示。

（3） 尺度变换 由于管道管径存在差异，管道与检测机器人之间的相对位置也会发生变化，这就导致视角中管壁和缺陷的尺度也会随着场景的不同而发生变化。对原图进行随机缩放，然后在缩放图上随机裁剪固定尺寸的子区域。若缩放图的尺寸小于固定裁剪尺寸，则进行零填充，具体如图 2‑19 d)、e)所示。

（4） 亮度变换 不同管道场景，不同检测设备打光能力存在差异，为了使模型拥有更强的亮度变化适应能力需要对原样本进行亮度变换增强，具体如图 2‑19 f）、g)所示。

（5） 颜色增强 管道场景变化复杂是检测的一大难点之一，在原图的基础上加入对比度变换、HSV颜色空间中的饱和度、色度扰动能够增强训练样本的场景多样性，增强训练模型的场景泛化能力，具体如图 2‑19 h)所示。

|  |
| --- |
| a) 正常  b) 水平翻转  c) 旋转  d) 尺度增大  e) 尺度缩小  g) 亮度增强  h) 颜色变换  f) 亮度减弱  图 2‑19数据增强样本 |

### 训练策略

卷积神经网络利用极小化模型预测输出与训练样本真值之间的的差异以反向调整卷积神经网络的参数。预测输出与真值之间的差异由损失函数来衡量，本文选用交叉熵作为模型的损失函数，具体公式见( 2‑28 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中 ——样本真实值；

——网络预测值。

交叉熵损失函数值可以理解为模型预测数据分布与训练数据分布的差异程度，在检测任务中我们无法获得数据的真实分布，只能退而求其次让模型学到的分布尽可能接近训练数据的分布，以达到最优预测效果。损失函数是模型优化的目标而优化算法的选择决定模型的训练时间以及最终收敛效果。梯度下降法是常用的神经网络训练方法，其公式见( 2‑29 )，其中是学习率，是梯度。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

在检测任务的损失函数等高线并不是那么“圆“的情况下，从某一点开始的梯度下降过程是极其曲折的，这样会使下降速度变慢，具体如图 2‑20所示。

|  |
| --- |
| 图 2‑20梯度下降图 |

本课题引入指数加权平均，采用动量梯度下降法[41]来更新参数以减少在训练过程中来回折返的情况，具体公式见( 2‑30 )所示：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

动量的引入使得在时刻下降方向不仅仅由当前时刻的梯度方向所决定，还受前次累积动量的影响。它不仅能在前期起到加速的作用，而且能在中后期更容易跳过局部最优值。上述公式中学习率仍然是一个可变量，为了得到更好的训练效果，本文采用动态学习率方法，学习率更新公式见( 2‑31 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中 ——学习率；

——学习率衰减系数；

——当前迭代次数；

——衰减迭代周期。

采用指数衰减策略使训练初期时学习率的值较大，能够将使参数更快接近最优解，起到加速训练的作用。而在后期接近最优解时使得学习率减小，避免了在最优解附近来回动荡的情况，使参数能够收敛到更优解。

## 本章小结

本章对常见地下管道缺陷做了简单的介绍，根据检测系统的功能需求设计了一体化智能检测流程，并在本文构建的数据集上对本文算法和其他算法进行了性能对比。在视频帧正异常检测任务中使用支持向量机模型准确率74%，召回率81%，单帧处理时间0.112s。极端随机树模型准确率78%，召回率82%，单帧处理时间0.165s。本文基于卷积神经网络SqueezeNet模型的准确率87%，召回率91%，单帧处理时间0.265s，InceptionV3模型的准确率91%，召回率94%，单帧处理时间0.876s，总体性能有较大幅度的提升，绘制的ROC曲线能完全包络其他文献算法。在缺陷分类任务中，基于InceptionV3的模型准确率83%，相较其他文献算法准确率优势明显，但时间效率较低。另外，为增强模型泛化能力和最终性能，在数据方面根据管道场景特性加入数据增强，在模型训练中使用动量梯度下降和指数衰减的动态学习率。

# 管道缺陷检测性能优化

## 引言

检测模型的构建只完成了检测系统最基本的需求，在实际场景检测过程中仍存在许多挑战。本章针对系统的实时性、准确性、过检率等性能进行优化，使得检测系统能够有更好的综合性能和实际场景适应能力。

## 管道正异常检测方法时间效率优化

深度学习算法相比于传统机器学习算法在性能提升的同时也带来了更大的计算复杂度，一般管道检测工控机无法满足计算需求。考虑到视频帧的连续性和关联性，只需对视频进行相隔固定帧采样检测，这在一定程度上可缓解计算量压力，但当使用类似InceptionV3这种大型深层神经网络时检测效率仍然无法达到实际场景检测需求。本节通过对模型结构进行优化，使得在保证准确率相近的情况下大幅压缩网络前向传播时间，达到最优综合性能。

### 深度分离卷积

随着深度学习的发展，学者们在不断优化加深模型深度提升模型准确率的同时，将大量研究投入模型轻量化中，降低模型的大小和计算时间，使得深度学习算法得以在移动端等计算资源受限的条件下运行。使用深度分离卷积能在保证模型准确性相差不大的情况下大幅度提高算法的时间效率。深度分离卷积将正常卷积拆分成深度层面卷积以及点对点卷积，具体原理如图 3‑1所示。

|  |
| --- |
| 图 3‑1深度分离卷积图 |

假设上一层卷积层有个通道，当前卷积层有个通道，正常卷积使用个宽度为的3D卷积核进行卷积，每一个卷积核对上一层所有通道进行卷积得到当前卷积层的一个通道。深度分离卷积则先进行深度层面卷积，使用个宽度为1的2D卷积核，对上一层的每一个通道逐个卷积之后得到一个通道数同样为的卷积层，而后进行点对点卷积，使用个11的卷积核进行信息关联得到个通道的当前卷积层。深度分离卷积可以得到冗余信息更少的稀疏化表达，在大幅减少计算量的同时并不会带来太多的信息损失，计算量比率见公式( 3‑1 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

其中是特征图的尺寸，由此可见在使用常规33卷积的情况下，深度分离卷积能够将计算量压缩至正常卷积的八到九分之一。

### 分组Shuffle卷积

深度分离卷积只能对卷积核大小在33以上的卷积进行时间效率优化，随着11小卷积核的大规模使用，在一个模型中11卷积计算量占总比重增加。为了进一步优化时间效率，需要对11卷积运算进行优化。本节使用分组的思想来降低计算量，假设第层卷积层通道数为，第层卷积层通道数为，卷积核大小为11，由于第层每一个通道都需要接受上一层所有个通道的信息，若特征图大小为，则总计算量为。如果对卷积通道进行分组，则同样情况下第层每一个通道只需要接受上一层个通道的信息，总计算为，计算量相当于原先的，具体原理如图 3‑2所示。

|  |
| --- |
| 图 3‑2分组Shuffle卷积 |

但是上述方法存在一个弊端，如果第层的前通道只能看到上一层前的信息会使模型的理解存在偏差，导致准确率下降。为保证在减少计算量的同时准确率与原模型相近，还需要对卷积通道进行Shuffle操作，将通过分组卷积得到的卷积通道进行重组。上述操作能够在使计算量下降的同时保证每一个卷积通道仍然能够看到上一层卷积层的所有信息。

### 优化结果分析

利用深度分离卷积和分组Shuffle卷积分别对33卷积核和11卷积核进行优化，同时对模型进行拓宽和加深，使得在获得接近InceptionV3准确率的同时，运行时间大大减小，优化前后模型的准确率、召回率和精确率以及单帧处理时间如表3-1所示，ROC曲线如图 3‑3所示.

表 3-1优化前后测试结果对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 模型 | 准确率 | 召回率 | 精确率 | 时间（s） |
| SqueezeNet | 0.873 | 0.908 | 0.886 | 0.265 |
| InceptionV3 | 0.915 | 0.946 | 0.935 | 0.876 |
| Ours | 0.908 | 0.931 | 0.928 | 0.359 |

|  |
| --- |
| I:\MyFolder\大论文\Image\roc2-1000.jpg  图 3‑3 优化前后ROC曲线对比 |

由上述数据可知，通过模型结构优化和网络拓宽加深提升了检测准确率。同时，相比于直接使用大型深层网络，本章优化方法在单帧处理时间上面增加的幅度较小，处于可接受范围内。

## 管道检测过检率优化

漏检是将异常样本认定为正常样本，而过检则是将正常样本认定为异常样本。通常，准确率和召回率是衡量视频帧正异常检测的一组指标，实际检测场景中我们更注重是否尽可能多的异常样本被识别出来，更直观的讲是宁可错检不可漏检，因此我们的模型更加注重召回率指标。这就导致在实际检测中允许一部分正常样本被认定为异常样本，因为异常样本还将进行二次检查，也就是说检测系统允许存在一定的过检率。然而，在一个系统中一定的“随机误差”是允许存在的，但“系统误差”是需要尽力去避免的。由于管道地形变化产生的摄像机剧烈抖动以及人为控制的大幅度偏摆等原因出现非正常视角，这一类视角通常以一种运动过渡状态短暂地出现在视频段中，有很大的概率被认定为异常帧。这种非管道本身缺陷导致的伪异常帧属于“系统误差”，应该尽可能的去避免。本节利用改进的Lucas-Kanade算法监测摄像机的位姿变化，对偏摆幅度大于一定阈值的视频帧从预测序列中剔除，这就能有效的降低过检率。同时，由于视频本身的连贯性这一策略并不会导致缺陷漏检。L-K算法的原理是对当前视频帧进行特征点检测，而后利用光流原理对检测出的特征点进行跟踪，利用特征点移动的距离和方向来监测摄像机的位姿变化。

### 单帧特征点检测

本文选用角点作为待跟踪特征点，定义角点为多个方向的平均强度变化值都比较高的像素点。在假定的兴趣点周围放置一个小窗口，并观察窗口内某个方向上强度值的平均变化。假定位移向量为，那么平均强度值变化就是：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中 ——图像像素灰度；

——兴趣点邻域。

将式( 3‑2 )泰勒展开得到式( 3‑3 )，将式( 3‑3 )转换为矩阵形式得到式( 3‑4 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |

记中间矩阵为，该矩阵的两个特征值分别表示最大平均强度值变化和其垂直方向的平均强度值变化。如果两者都很大，那么就是在角点的位置。定义角点响应函数S为式( 3‑5 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中 ，——矩阵的两个特征值；

det——矩阵行列式；

trace——矩阵的迹。

将响应函数大于一定阈值的窗口中心认定为角点。由于管道场景整体灰度趋于平滑，本文将设定为0.1，窗口大小设为15。同时，在算法中加入非极大值抑制，滤除极值相邻的角点，使检测角点较为分散地分布在视角中，具体检测效果如图 3‑4所示。

|  |
| --- |
| 图 3‑4角点检测图 |

### 帧间特征点跟踪

根据视频帧的连续性，利用Lucas-kanade算法跟踪检测到的角点，判断当前帧角点在后续帧中的新位置，根据位置变化幅度来判断摄像机的偏摆程度。假定相邻帧中的同一特征点强度值是不变的，这个过程就是寻找如下的位移：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中和分别代表当前帧和下一个瞬间的帧。强度值不变的假设适用于相邻图像上的微小位移。将式( 3‑6 )泰勒展开得到式( 3‑7 )，去掉两个表示强度值的项可得到式( 3‑8 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |

单个方程无法解出两个未知量，算法假设光流向量在像素点的邻域保持不变，然后使用最小二乘法对邻域中的所有像素点求解光流方程。然而，为了提高跟踪结果的精度，应该选择较小的邻域窗口，而为了处理快而长的运动应该选择大的邻域窗口。为了解决这一矛盾引入了金字塔算法，先在图像金字塔的最顶层计算光流，用上一层估计到的运动结果作为下一层金字塔的初始值，下一层在上一层基础上计算光流向量，重复这样的估计直到金字塔的最底层，将最底层光流向量作为最终结果。上述策略能够使得算法适用于速度快、位移大的运动，从而提高算法在真实检测场景中的准确性和鲁棒性，原理如图 3‑5所示。

|  |
| --- |
| 图 3‑5金字塔Lucas-Kanad光流 |

然而该特征点跟踪算法适用于背景固定前景运动的场景，本课题涉及到的场景背景前景都在运动容易产生误跟踪。为了解决这一问题，本文采用了双向光流法。即在当前帧利用L-K算法检测出当前帧特征点在下一帧中对应的特征点，同时进行反向计算将作为当前帧，将作为下一帧，计算出在中对应的特征点，若和的偏差小于一定的阈值才认定为跟踪成功。

经实际验证，双向光流法有效的提高了本场景特征点的跟踪效果。获取到了相邻帧的特征点位置信息后，根据下式：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中n为当前帧特征点数量，为特征点i前后帧x，y坐标差之和，为所设定的阈值，如果式( 3‑9 )成立，则认定该帧为伪异常帧并将其从缺陷检测序列中移除。光流检测效果如图 3‑6所示。

|  |
| --- |
| 图 3‑6光流检测效果图 |

## 管道异常分类准确率优化

### 迁移学习

本文算法相比于其他文献算法在性能上已经具备比较大的提升，但与实际检测需求相比仍存在一定的差距，为此必须进一步对模型准确率进行优化。本文所构建的管道图像数据集在规模和场景多样性上相对其他文献已经较为可观，但和目前图像识别的主流数据集ImageNet数据集[42]相比仍然存在较大的局限性。同时，管道数据与常见物体图像相比，其在网络上几乎没有高质量数据资源，整体获取难度和成本较大。在训练数据受限且进一步扩大数据集成本较大的情况下，利用其它已有数据集并在此基础上进行迁移学习能够实现低成本性能提升。

管道图像识别和大规模日常物体识别存在相关性，两者在低层次的视觉信息如颜色、纹理、形状等特征参数的学习上具有很强的一致性。因此，本节将构建的模型先在ImageNet数据集上进行训练，将训练得到的参数模型固化并用以初始化管道图像识别模型。由于ImageNet数据集的图像类别是1000，而正异常检测是2分类任务，缺陷分类是8分类任务，故需根据任务类别数量对管道识别模型的后几层迁移层进行结构调整。同时，由于视觉任务在低级特征上具备一定的共通性，故将在ImageNet上训练得到的前端参数锁死，在管道数据集上微调时不再参与训练。最后，将模型在管道图像数据集上进行微调，使后端网络能够学习到更适应管道检测任务的参数。经过实验发现，利用上述迁移学习方式获得的模型性能相比于仅利用管道图像数据集训练得到的模型有小幅的性能提升，具体流程如图 3‑7所示。

|  |
| --- |
| 图 3‑7 迁移学习流程 |

### 模型集成策略

如同随机森林算法在性能上相比于单棵决策树有较大的提升，集成策略在深度学习模型中也同样适用，常用的集成策略如下：

（1）不同初始化集成 由于神经网络的训练机制是基于随机梯度下降， 故不同的网络参数初始化会导致不同的训练结果，尤其是对于小样本的学习场景， 首先对同一模型进行不同初始化，之后得到的网络模型进行结果集成会大幅度缓解其随机性，提升最终任务的预测结果。

（2）不同网络结构集成 可以在使用如VGGNet，深度残差等不同网络架构的网络进行训练，最后将不同架构网络得到的结果做集成。

（3）不同目标函数集成 采用不同的损失函数作为目标函数分别训练模型。在预测阶段,对不同模型预测结果做“置信度级别”的平均或投票完成预测任务。

（4）网络快照集成 深度神经网络模型复杂的解空间存在非常多的局部最优解， 但经典的随机梯度下降方法只能让网络模型收敛到其中一个局部最优解。 网络快照便利用了网络解空间中这些局部最优解来对单个网络做模型集成[43]。 通过循环调整学习率可使网络依次收敛到不同的局部最优解，收敛示意图如图 3‑8所示。

|  |
| --- |
| 图 3‑8模型快照集成法[43] |

利用余弦函数对学习率进行“循环退火”处理，即将学习率从初始设定值随训练次数增加逐渐减缓到0，之后再重新放大而跳出局部最优解，学习率变化公式( 3‑10 )如下：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中为初始学习率，为模型迭代次数,为总训练次数，为学习率“循环退火”的次数，具体学习率变化曲线如图 3‑9所示。学习率“循环余弦退火”的过程中，每个循环结束可使模型收敛到一个不同的局部最优解，称之为模型快照。在预测阶段，对各个模型快照预测结果取平均，得到最终预测结果。

|  |
| --- |
| 学习率  图 3‑9余弦退火学习率下降 |

本文采用模型快照集成法，其保留一次训练过程中的不同状态，只需对模型进行一次完整的训练，而上述集成策略的前三种需要采用不同网络结构或不同初始化状态多次训练网络。

### 优化结果分析

利用迁移学习和模型集成的方法能够在原基础上带来一定的性能提升，在异常分类任务的测试数据上进行实验，得到数据如表3-2所示。

表 3‑2优化前后测试结果对比

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 参数类型 | 原模型 | 迁移学习 | 快照集成 | 迁移+快照 |
| 准确率 | 0.831 | 0.852 | 0.843 | 0.870 |
| 精确率 | 0.830 | 0.854 | 0.845 | 0.868 |
| 召回率 | 0.830 | 0.853 | 0.848 | 0.867 |
| F1系数 | 0.830 | 0.853 | 0.849 | 0.869 |

由上述数据可知，迁移学习和快照集成都能够给模型带来一定的性能提升。迁移学习需要在两个数据集上进行训练，该策略增加了一定的训练时间成本，收益略大于模型快照集成。模型快照集成属于单模型集成，只需要一次训练没有增加训练时间成本，但在预测阶段需要计算各个“快照”的输出进行平均，增加了预测阶段时间成本。训练时间成本增加带来的影响是可忽略的，但预测阶段时间增加会直接影响到检测效率，但考虑到异常情况在实际检测中占比较小，相对整体来说带来的影响在可接受范围内。因此，两种策略的结合能够在实际检测中带来正向收益。

## 本章小结

本章根据正异常检测和异常分类任务的特性，分别对上述两阶段检测算法进行了优化。针对正异常检测任务，利用深度分离卷积和分组Shuffle卷积优化了模型结构，在保证准确性的同时大大减小了单帧处理时间。针对异常分类任务，利用迁移学习和模型快照集成法提升了检测准确率。同时，也针对实际检测过程中的出现伪异常帧导致过检率过高的问题设计了过滤算法，对摄像机的位姿变化进行监测，滤除偏摆或振动幅度过大的视频帧，降低了过检率。上述优化措施提高了模型实际场景检测适应能力，增加了非缺陷异常帧的处理能力，使检测系统更加智能化。

# 管道信息及缺陷位置信息识别方法

## 引言

检测视频的开头和结尾会显示该段视频检测管道所处地段、井号、检测员、管道类型、管道材质等信息，检测视频中段会实时显示管道机器人与井口之间的距离信息。根据功能需求，系统需要在检测每一段视频时自动识别并记录管道地段、材质等信息，如果发现缺陷需要记录缺陷发生处与井口之间的距离。这就需要设计一种视频字符识别方法，当异常分类模型识别到当前视频帧为信息帧时自动调用，对管道信息进行识别并记录，从而实现一体化检测。管道信息字符相比于自然场景字符较为有序且变化较小，但不同的厂家有不同的规格且管道背景较为复杂，因此仍需要设计相适应的字符识别算法。字符检测识别的基本流程包括字符定位、字符分割和字符识别，本章结合管道场景视频字符特点对上述各个流程进行算法设计。

## 信息字符定位

### 字符候选区域提取

字符定位是字符分割和识别的前提，其性能直接影响到整个流程结果。相比于毫无规律的自然场景字符检测视频字符检测在字体大小、颜色等方面变化性较小，但由于不同检测单位采用不同的规格，信息字符在视频中的位置会发生变化导致无法使用固定ROI（Region of Interest，感兴趣区域）提取字符位置。针对上述特性，本文采用MSER算法，即最大稳定极值区域算法对文字进行粗定位[44]。MSER通过使用不同的阈值对图像进行二值化，二值化阈值范围为0-255，如果设置阈值增量为，定义阈值变化前后区域的面积变化率为公式( 4‑1 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中为区域灰度值在以下的面积，MSER算法原理见图 4‑1。图中高度代表灰度值的高低，区域A属于稳定极值区域而区域C则不是。字符区域具有相近的灰度值，算法能够对字符区域进行粗定位，检测效果如图 4‑2所示。可以发现，该算法的检测结果存在大量的重叠区域，为提高后续算法时间效率，需对重叠区域进行筛除。利用包含从属关系以及文字区域的形状特征筛除之后得到图 4‑3，经过这一步骤候选框的数量从2203减少到200。

|  |
| --- |
| 图 4‑1 MSER算法原理图 |

|  |  |
| --- | --- |
| 图 4‑2 MSER检测图 | 图 4‑3 包含关系滤除图 |

### 基于Adaboost算法的非字符区域筛除

从图片中可以看出，筛选之后的候选区域仍存在着一些背景区域，这会对后续区域融合造成极大的干扰。本文利用Adaboost算法[16]对字符区域和背景区域进行识别，该算法以决策树为基分类器进行串行级联，决策树的构建在第二章中已经进行阐述。相比于随机森林的并行集成策略，Adaboost算法是基于串行集成策略，先利用原始数据训练弱分类器，而后将数据中的错分样本权重提高，同时根据该弱分类器的错误率给出该弱分类器的权重，之后利用新的数据分布训练下一个分类器，串行重复上述过程，直达准确率达到设定标准或弱分类器数量超过设定上限，最终得到的强分类器为所有弱分类器的加权和，过程示意如图 4‑4所示。

|  |
| --- |
| 图 4‑4 Adaboost算法示意图 |

单分类器权重计算公式如下，其中为弱分类器错误率：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

样本权重分布更新公式( 4‑3 )如下，其中为训练所得分类器，为数据真实值：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

最终集成分类器如式( 4‑4 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

利用MSER算法截取候选区域构建训练数据集，其中部分负样本和正样本分别如图 4‑5、图 4‑6所示。

|  |  |
| --- | --- |
| 图 4‑5 字符负样本 | 图 4‑6 字符正样本 |

利用训练得到的模型去除候选区域中的背景区域，得到图 4‑7。有些字符由于字体形态原因，同一个字符分为多个部分，比如“川”、“始”。这类字符的每一个部分都会被认定为一个候选区域，而不是一个候选区域对应整个字符，如果以候选区域为单位进行分割识别会造成很多区域内并不是完整文字而导致无法识别。因此需要对候选区域进行融合，以字段的形式进行分割和识别操作。首先，将候选区域内像素值置为，其他区域像素值置为，得到二值化图 4‑8。而后使用形态学闭运算进行区域融合，闭运算以形态学腐蚀和膨胀操作为基础。如果集合为与的并集，为原图像，为形态学结构元素，结构元素的中心点在中滑动，将的中心点滑动到时所覆盖的区域记为，则膨胀和腐蚀操作公式分别见( 4‑5 )、( 4‑6 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |

闭运算时用相同的结构元素对先膨胀，后腐蚀，能够填充整个字段区域内部的背景间隙，又不使字段区域的外边界扩张，其定义如下：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

经过闭运算之后进行水平方向和垂直方向的直方图投影，其中水平直方图投影如图 4‑9所示，根据两个方向的直方图投影对字段区域进行切割，得到图 4‑10。

|  |  |
| --- | --- |
| 图 4‑7 Adaboost滤除图 | 图 4‑8 二值化 |
| 图 4‑9 水平直方图投影 | 图 4‑10 字符区域融合 |

## 信息字符分割

### 基于K-Means的颜色分割

为了能更好的对视频字符进行识别，在定位出字符区域之后需要对字符进行前背景分割，分割的效果直接决定字符识别的准确率。管道信息字符的分割相对于一般视频字幕分割难度较大，原因在于字幕要么处于背景相对单一的位置，要么在字幕显示过程中背景是不断流动的。若字幕文字处于简单背景位置则只需简单的阈值操作就能起到很好的分割效果，若字幕处于流动的复杂背景下，则可通过视频多帧融合算法进行分割。后者中字符信息是固定不变的，而流动的背景是不断变化的，只需对多帧视频取像素均值即可弱化背景像素对字符像素的干扰。然而，本文获取的管道视频字符处于复杂多变的背景之下同时在管道信息字符显示过程中背景是固定不变的，无法进行多帧融合增强处理。由于视频字符灰度值会随场景亮度的变化而变化，所以无法使用固定阈值分割。

常用的自适应阈值分割法有大津法[45]和局部自适应阈值分割[46]。局部自适应阈值分割算法适用于图像不同区域亮度情况不同的场景，其对每一个像素点采用不同的阈值，该阈值由像素点所在的邻域内的平均值或高斯加权和决定，利用局部自适应阈值分割算法对图 4‑11进行字分割可得到图 4‑12。大津法是一种基于全局的二值化算法，其计算出使得图像背景和前景有最大类间方差的阈值并以此进行二值化，最大类间方差计算公式见( 4‑8 )。对范围内的灰度值迭代计算最大类间方差，将使得最大的灰度作为阈值，利用大津法对图 4‑11进行分割可得到图 4‑13。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中 ——前景像素点数占图像比例；

——前景平均灰度；

——背景像素点数占图像比例；

——背景平均灰度；

——图像总灰度

由分割效果可见上述两种自适应阈值分割算法在管道场景取得的效果并不可观，本文采用自适应能力更强的无监督机器学习方法——均值聚类，其算法流程见表4-1。本课题利用图像的三通道颜色信息作为输入特征，选取聚类个数为5，最终得到分割结果如图 4‑14所示。该分割方法在管道字符分割中起到较优的效果，但仍存在一些弊端。第一，均值聚类算法的损失函数是非凸函数，意味着无法保证取得的最小是是全局最小，其对质心初始位置的选取比较敏感。第二，部分场景背景颜色与字符区域近乎一致，仅根据颜色信息无法有效的将前景文字分割出来，难例的原图和分割图分别如图 4‑15图 4‑16所示。

表 4-1 均值聚类算法流程

|  |
| --- |
| **算法2**：均值聚类 |
| **输入**：训练集； |
| 聚类数目. |
| **过程**： |
| 从中随机选取个聚类中心点为； |
|  |
| 令 |
|  |
| 计算样本所属的类： |
| 将样本划入相应的簇： |
|  |
| 计算新均值向量： |
|  |
|  |
| 当前均值向量均未更新 |
| **输出**：簇划分 |

|  |
| --- |
| 图 4‑11原图 |
| 图 4‑12局部自适应分割图 |
| 图 4‑13 Ostu分割 |

|  |
| --- |
| 图 4‑14 K均值聚类分割 |
| 图 4‑15 背景干扰图 |
| 图 4‑16 干扰分割图 |

### 基于层次聚类的尺度信息融合分割

由上节末尾可知，在像素颜色信息近乎相同的情况下，如果仅根据颜色信息是无法对字符进行有效分割的。除颜色特征之外，字符还存在一个特有特征—字宽特征[47]，即字符笔画具有特定的宽度，这是一些其他区域所不具备的，具体示意如图 4‑17所示。

|  |
| --- |
| 图 4‑17 字宽尺度示意图[47] |

假定存在连通区域，该区域的宽度特征可由其中每一个像素点的尺度取均值得到，具体见公式( 4‑9 )、( 4‑10 )。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |

式中，，，分别为0°，45°，90°，135°的直线与连通区域交集部分的长度，为连通区域的像素个数。本文利用种子填充算法对待分割图像进行连通区域分析，其算法的基本流程见表4-2。

表 4-2 种子填充算法流程

|  |
| --- |
| **算法3**：种子填充算法 |
| **输入**：待分割图像，像素初始值为0的遍历标记图像； |
| 生长规则：像素值相似、 |
| **过程**：  中的每一个像素点 |
| **if**  将当前像素点压入栈中，令，  **:**  S不为空  弹出栈顶元素，令  的每一个符合生长规则的八连通像素  将压入中，令， |
| **输出**：中具有相同的像素为同一连通区域 |

对图 4‑15进行连通区域分析，并对每一个连通区域赋予不同的颜色，得到图 4‑18。然后，提取并记录每一个区域的字宽尺度特征。同时，针对均值聚类算法存在的局限性，本文使用层次聚类算法[48]进行分割，该算法通过计算任意两个数据点（组）间的相似性，对其中最为相似的两个数据点（组）进行合并，并反复迭代这一过程，直到类别数目缩小到设定的阈值，具体示意图和树状图分别如图 4‑19、图 4‑20所示。

|  |  |
| --- | --- |
| 图 4‑18连通区域分析 | |
| 图 4‑19 自下而上层次聚类示意图 | 图 4‑20 层次聚类树状图 |

不同于上一节仅将颜色信息作为度量数据点间相似度的标准，本节算法融合了字符的宽度信息，相似度计算方法见公式( 4‑11 )。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

式中 ——颜色信息；

——尺度信息；

——比例权重。

迭代初期为数据点的相似度计算，迭代一定步数之后需要对数据簇进行相似度度量，计算公式见( 4‑12 )：

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |

经过上述优化之后能够对一些颜色相近的难例取得更好的分割效果，这直接提高了下一步字符识别的精度，分割对比效果见图 4‑21。

|  |
| --- |
| 1. 原图 |
| 1. 局部自适应分割 |
| c) K均值聚类分割 |
| d) 本文优化方法  图 4‑21 分割效果对比图 |

## 信息字符识别

在进行完字符位置检测和字符分割之后，需要对字符信息进行识别。由于管道地段、材质等信息为汉字，而汉字类别繁多，训练和识别是一个很大的工程。目前，商用和开源OCR识别软件已经较为成熟，对分割良好的二值化输入图片能够保证很高的准确率。因此，本文在进行字符定位和字符分割之后，识别部分由Google开源文字识别工具Pytesseract完成，具体检测效果如图 4‑22所示。



图 4‑22 管道信息识别效果

除了记录检测管道所属地段、材质等信息外，本文所提出的检测系统在发现缺陷之后需要记录缺陷与管口之间的距离。距离信息仅包含数字和少量符号，利用模板匹配[49]技术就能对数字进行精准而快速的识别，距离信息识别样例和模板图像分别如图 4‑23、图 4‑24所示。在有效分割之后对数字进行切分，将切分的图像块做零值边缘扩充，分别利用各个模板在待识别数字图像上滑动，计算模板在不同位移时两者的标准相关匹配值，而在计算相关系数之前需要对模板和待匹配图像进行标准化操作，具体见公式( 4‑13 )、( 4‑14 )、( 4‑15 )。选取所有滑动状态中的最大匹配值作为该模板和待识别数字的相似度，认定待识别数字等于最大相似度模板图像中的数字。

|  |  |  |
| --- | --- | --- |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |
|  |  | ( ‑ ) |
| 图 4‑23 距离信息识别样例 | | | | |
| 图 4‑24 数字模板 | | | | |

## 本章小结

为增加检测系统智能化程度，实现自动管道信息、缺陷位置信息识别和记录的功能，本章设计了管道场景视频字符识别算法。在字符定位阶段，利用最大稳定极值区域算法和Adaboost算法的级联实现了精准的定位；在字符分割阶段，对原有的基于单纯基于颜色信息的聚类分割算法进行改进，使用融合字宽尺度信息的层次聚类算法能对前背景颜色相近的难例进行分割；在字符识别阶段，利用开源文字识别工具Pytesseract进行文字识别，利用基于标准相关匹配算子的模板匹配进行距离数字识别。最后，将识别到的信息记录并整合，输出对应管道的缺陷分析报告。

# 管道缺陷检测方法的实验与分析

## 引言

本章介绍了管道图像的采集设备和采集环境，基于真实管道场景的图像和视频进行实验，对算法可行性和准确性进行分析。另外，由于部分实验结果已在前文算法性能对比中给出，本章就未分析部分进行完善，同时对检测方法进行整体性评估。

## 管道图像数据集构建

管道视频由进入地下管道的机器人携带的摄像机拍摄得到，实际采集过程中使用的深圳施罗德公司的型号为SINGA-300的履带机器人，实物见图 5‑1，具体参数见表5-1。

|  |
| --- |
| 图 5‑1管道检测机器人实物图 |

表5-1 SINGA-300机器人主要参数

|  |  |  |  |
| --- | --- | --- | --- |
| 参数类型 | 参数值 | 参数类型 | 参数值 |
| 适用管径 | DN300-200mm | 拍摄像素 | 400万 |
| 控制距离 | 40m | 爬坡角度 |  |

本文数据的主要来源是施罗德公司提供的管道检测视频，采用相隔固定帧进行采样构建训练和测试集。正异常检测阶段数据分布及数量见表5-2，异常分类阶段数据分布及数量见表5-3。其中异常分类阶段也设立了正常类别，包含了正异常检测阶段错分概率较大的部分正常样本，以对上一级错分样本进行更正。管道检测相对人脸、行人、物体检测相对小众，缺乏具有代表性的公共数据集，故仅能通过在本文构建的数据集上复现其他文献算法来进行间接性能对比。同时，管道数据网络资源较少且质量较差，自行获取需要专门的管道机器人检测设备导致数据获取成本大。通过与其他管道检测领域文献对比发现，本文构建的数据集分辨率更高，数据规模更大，覆盖场景更广，对算法性能的评估具有一定的代表性。

表5-2 正异常检测阶段数据分布

|  |  |  |
| --- | --- | --- |
| 数据集 | 正常样本 | 异常样本 |
| 训练集 | 9440 | 7657 |
| 测试集 | 1788 | 1676 |

表5-3 异常分类阶段数据分布

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 数据集 | 沉积 | 障碍 | 异物穿入 | 气体 | 正常 |
| 训练集 | 1202 | 1229 | 1019 | 1336 | 1215 |
| 测试集 | 400 | 404 | 335 | 443 | 400 |

## 管道检测性能实验分析

管道检测中准确性和检测效率是两个重要指标。在正异常检测中，使用在SqueezeNet基础上经过深度分离卷积和分组Shuffle卷积结构优化后的模型，可达到准确性和检测效率综合性能最优，具体指标见表5-4，多模型ROC曲线见图 5‑2，正常识别样例如图 5‑3所示，异常识别样例如图 5‑4所示。

表5-4 正异常检测各模型测试结果

|  |  |  |  |
| --- | --- | --- | --- |
| 模型 | 召回率 | 准确率 | 时间（s） |
| HOG+SVM | 0.81 | 0.74 | 0.112 |
| GIST+ET | 0.82 | 0.78 | 0.165 |
| Ours | 0.93 | 0.91 | 0.359 |

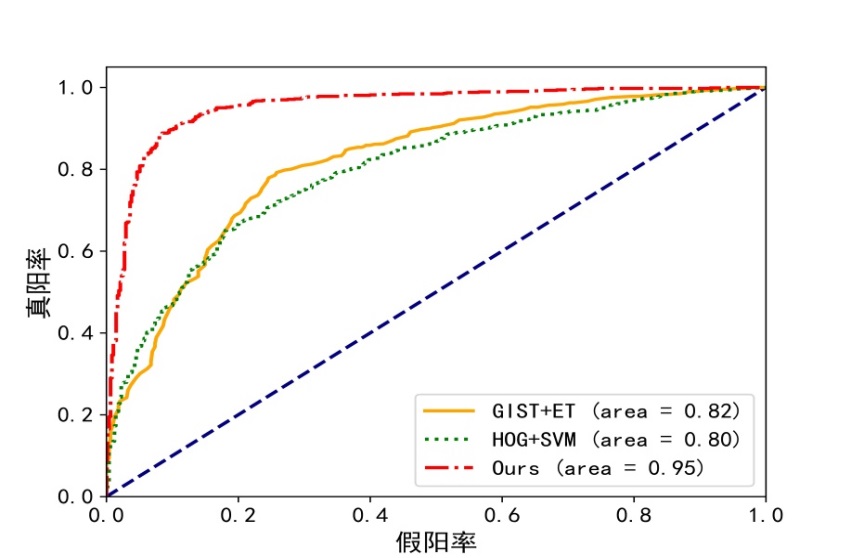


图 5‑2 多模型ROC曲线

|  |
| --- |
| 图 5‑3 正常识别样例 |
| 图 5‑4 异常识别样例 |

在异常分类中，其面向对象为第一级分类器判定为异常的视频帧，其出现频率较低，该任务单帧处理时间对整体检测效率影响不大，故把准确性指标放在第一位。该阶段使用基于InceptionV3模型的残差网络，并在此基础上使用迁移学习和模块快照集成进行优化，最终检测结果见表5-5，沉积识别样例见图 5‑5，气体识别样例见图 5‑6，异物穿入样例见图 5‑7，障碍样例见图 5‑8。

表5-5 异常分类测试结果

|  |  |  |  |
| --- | --- | --- | --- |
| 缺陷类型 | 精确率 | 召回率 | F1系数 |
| 沉积 | 0.84 | 0.85 | 0.85 |
| 气体 | 0.95 | 0.95 | 0.95 |
| 异物穿入 | 0.97 | 0.88 | 0.92 |
| 障碍物 | 0.81 | 0.84 | 0.83 |
| 合计 | 0.87 | 0.87 | 0.87 |

|  |
| --- |
| 图 5‑5 沉积识别样例 |
| 图 5‑6 气体识别样例 |
| 图 5‑7 异物穿入识别样例 |
| 图 5‑8 障碍识别样例 |

由上述数据可知，本文检测方法在准确性方面距离代替人工检测还存在一些差距，但相较其他文献算法有了大幅度的提升。本文完全基于真实管道场景进行实验，整体而言在管道自动检测领域向前迈进了一大步。

## 视频时间压缩实验分析

由上述实验结果可知，本检测方法在准确性方面相较其他文献算法有了较大幅度的提升，在实际场景检测中的泛化和适应能力也明显增强，但距离完全代替人工检测还存在差距。因此，本文提出另一种检测方案，利用正异常检测模型和双向光流算法对视频时间进行压缩，即排除模型认定为正常的和摄像机保持静止视频段，从所有视频帧中筛选出异常视频段交由人工进行复检。本节采用多段复合视频进行实验，压缩前后对比结果如图 5‑9所示，缺陷统计结果见表5-6。

|  |
| --- |
| C:\Users\Administrator\Desktop\zhifangtu.jpg  图 5‑9视频时间压缩前后对比 |

表5-6 复合视频缺陷统计结果

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 视频序号 | 原始时间 | 压缩后时间 | 缺陷总数 | 发现缺陷数 | 缺陷检出率 | 压缩率 |
| 1 | 1280s | 388s | 14 | 14 | 100% | 0.696 |
| 2 | 775s | 353s | 13 | 13 | 100% | 0.545 |
| 3 | 770s | 209s | 9 | 8 | 88.9% | 0.729 |
| 4 | 1243s | 260s | 15 | 14 | 93.3% | 0.791 |
| 5 | 734s | 111s | 3 | 3 | 100% | 0.849 |
| 6 | 663s | 76s | 4 | 4 | 100% | 0.885 |
| 总计 | 5465s | 1397s | 58 | 56 | 96.5% | 0.745 |

所有视频段的总时间压缩率为74.5%，总体缺陷召回率为96.5%，由此可见利用正异常检测和双向光流进行时间压缩结合人工复检能够在保证很高的缺陷检测率的同时，大大减少专业检测人员的干预时间，从而大幅提高检测效率。而由上节得到单帧正异常检测中缺陷召回率为93%，由于视频帧的连续性，缺陷会出现在连续的多帧视频中，对其中一帧或几帧漏判有时并不会导致遗漏这一处缺陷。因此，在本节实验中总体缺陷召回率略高于单帧检测缺陷召回率。

## 信息识别精度实验分析

本节使用200张信息视频帧对管道信息识别方法进行性能评估，其中以字段为单位对字符定位算法进行准确性统计，具体数据见表5-7，字符定位样例见图 5‑10。

表5-7 字符定位检测结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 总字段数 | 漏检数 | 过检数 | 漏检率 | 过检率 |
| 1771 | 3 | 16 | 0.17% | 0.9% |

|  |
| --- |
| 漏检  过检  图 5‑10 字符定位样例 |

由上述数据可知，本文字符定位算法的漏检率在相当低的一个水平，只有在背景色彩和亮度极端接近信息字符的时候会发生漏检。而过检率相对较高，但仍然在1%以下且部分过检情况可通过后期识别算法滤除，并不会对整体信息识别造成影响。完成定位之后需要进行分割和识别，由于分割没有一个直观定量的评估指标，故将分割和识别进行联合准确率评估，具体数据见表5-8，而针对分割算法的评估仅给出部分难例分割效果，如图 5‑11所示。

表5-8 字符识别检测结果

|  |  |  |
| --- | --- | --- |
| 总字符数 | 正确识别字符数 | 准确率 |
| 15939 | 15767 | 98.92% |

|  |
| --- |
| 图 5‑11 难例分割图 |

本文使用谷歌公司的开源字符识别工具Pytesseract进行识别，准确率较为可观。另外，在发现缺陷时还需要对缺陷距离信息进行识别记录，本文使用基于标准相关匹配值的模板匹配算法进行识别，采用800张视频帧进行实验得到数据如表5-9所示。可以看出，本文算法在距离信息识别上能够达到很高的准确率，能够符合实际检测要求。最终将所有模块整合形成端到端的一体化检测系统，其图形用户界面如图 5‑12所示。

表5-9 距离信息识别检测结果

|  |  |  |
| --- | --- | --- |
| 总帧数 | 正确识别帧数 | 准确率 |
| 800 | 795 | 99.37% |

|  |
| --- |
| C:\Users\ADMINI~1\AppData\Local\Temp\WeChat Files\5979c3eb473daea02c5ff153b5efffc.png  图 5‑12 管道检测系统图形用户界面 |

# 结 论

目前地下管道检测普遍采用CCTV检测方式，其拍摄回来的视频需由专业人员进行缺陷分析。大量的管道视频由人工进行检测存在易疲劳、主观性强、效率低下等缺点。针对这一现状，本课题利用计算机视觉方法构建了针对管道视频的一体化智能检测系统，对管道有无缺陷、缺陷类别及缺陷发生位置进行自动检测，取得的成果如下：

（1）根据功能需求及实际管道场景检测特性，提出了一体化管道检测方案，整体流程包括伪异常帧过滤、正异常检测、异常分类、管道信息及缺陷信息识别以及缺陷报告生成。通过多种算法融合处理提高了管道检测的抗干扰性、准确性以及智能性，完成端到端的检测。

（2）以卷积神经网络为基础，构建了正异常检测和异常分类模型。通过深度分离卷积和分组Shuffle卷积对网络进行优化用于正异常检测任务，可达到91%的准确率和93%的缺陷召回率。通过迁移学习和模型快照集成对InceptionV3残差网络进行优化用于异常分类任务，可实现87%的F1系数。在缺陷检测准确性上方面，相较其他文献方法有了大幅度的提升，虽仍无法实际落地，但向着真实管道场景缺陷检测自动化迈进了一大步。

（3）利用改进的光流算法对摄像机位姿进行监测，滤除了伪异常帧，增加了检测系统在真实场景中的特殊视角处理能力。同时基于缺陷检测现状，提出正异常检测和双向光流算法结合人工复检的检测方式，能够在保证96.5%的缺陷检出率的同时将视频时间压缩约74%，从而大大减少人工干预时间。

（4）针对管道场景特性设计了视频字符识别算法，实现了管道综合信息和缺陷距离信息的自动识别记录功能。使用最大稳定极值算法和Adaboost算法级联进行字符定位，可将漏检率和过检率分别控制在0.17%和0.9%。使用融合尺度信息的层次聚类算法进行字符分割再结合Pytesseract识别工具，可实现98.92%的管道综合信息字符识别准确率。使用基于标准相关匹配值的模板匹配算法，可实现99.37%的缺陷距离信息识别准确率。

虽然本课题取得了一定的研究成果，但在以下方面仍然具备改进空间：

（1）构建规模更大、场景和缺陷多样性更强的管道数据集。

（2）持续优化检测算法，进一步提升检测方法的准确性和实时性。

（3）优化字符定位和分割算法，提高背景极端复杂情况下的字符定位和分割能力，进一步提高字符识别准确率。

# 参考文献

1. 李波锋. 基于机器视觉的排水管道缺陷检测算法研究[D]. 广州：广东工业大学，2015：15-21.
2. Wirahadikusumah R, Abraham D M, Iseley T, et al. Assessment Technologies for Sewer System Rehabilitation[J]. Automation in Construction，1998，7（4）：259-270.
3. 宋志东. 超声导波技术在管道缺陷检测中的研究[D]. 天津：天津大学，2006：55-71.
4. 李田, 郑瑞东, 朱军. 排水管道检测技术的发展现状[J]. 中国给水排水，2006，22（12）：11-13.
5. 闵令伟. 管道三维重建与缺陷检测研究[D]. 哈尔滨：哈尔滨工业大学，2015：66-84.
6. 郑瑞东. 上海市排水管道CCTV检测评价技术研究[D]. 上海：同济大学，2006：11-19.
7. Guo W, Soibelman L. Automated Defect Detection for Sewer Pipeline Inspection and Condition Assessment[J]. Automation in Construction，2009，18（5）：587-596.
8. Wu W, Liu Z, He Y. Classification of Defects with Ensemble Methods in the Automated Visual Inspection of Sewer Pipes[J]. Pattern Analysis & Applications，2015，18（2）：263-276.
9. 卞正岗. 机器视觉技术的发展[J]. 中国仪器仪表，2015（6）：40-42.
10. Sinha S K. Automated Underground Pipe Inspection Using a Unified Image Processing and Artificial Intelligence Methodology[M]. University of Waterloo，2000.
11. 王鹏. 基于图像处理和人工智能的排水管道病害自动检测方法[D]. 北京：北京信息科技大学，2008：22-29.
12. Goh A T C. Back-propagation Neural Networks for Modeling Complex Systems[J]. Artificial Intelligence in Engineering，1995，9（3）：143-151.
13. Halfawy M R, Hengmeechai J. Automated Defect Detection in Sewer Closed Circuit Television Images Using Histograms of Oriented Gradients and Support Vector Machine[J]. Automation in Construction，2014，38（5）：1-13.
14. Hengmeechai J, Halfawy M R, Integrated Vision-Based System for Automated Defect Detection in Sewer Closed Circuit Television Inspection Videos[J].American Society of Civil Engineers，2015，29（1）：1-9.
15. Lucas B D, Kanade T. Optical Navigation by the Method of Differences[C]// International Joint Conference on Artificial Intelligence，Nevada：Morgan Kaufmann Publishers，1985：981-984.
16. Yoav F, Robert E S. A Decision-Theoretic Generalization of On-Line Learning and An Application to Boosting[M]. Computational Learning Theory. Springer Berlin Heidelberg，1995：119-139.
17. Epaillard E, Bouguila N. Proportional Data Modeling with Hidden Markov Models Based on Generalized Dirichlet and Beta-Liouville Mixtures Applied to Anomaly Detection in Public Areas[J]. Pattern Recognition，2016，55：125-136.
18. Myrans J, Kapelan Z, Everson R. Automated Detection of Faults in Wastewater Pipes from CCTV Footage by Using Random Forests[J]. Procedia Engineering，2016，154：36-41.
19. Svetnik V, Liaw A, Tong C, et al. Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling[J]. Journal of Chemical Information & Computer Sciences，2003，43（6）：1947.
20. 周飞燕, 金林鹏, 董军. 卷积神经网络研究综述[J]. 计算机学报，2017，40（6）：1229-1251.
21. Krizhevsky A, Sutskever I, Hinton G E. ImageNet Classification with Deep Convolutional Neural Networks[C]// International Conference on Neural Information Processing Systems，California：Curran Associates，2012：1097-1105.
22. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv：1409.1556.
23. Szegedy C, Liu W , Jia Y , et al. Going Deeper with Convolutions[C]// IEEE Conference on Computer Vision and Pattern Recognition，Boston：[Computer Vision Foundation](http://www.cv-foundation.org/)，2015：1-9.
24. Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[C]// International Conference on International Conference on Machine Learning，Washington：IEEE Computer Society，2015：448-456.
25. He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]// The IEEE Conference on Computer Vision and Pattern Recognition，Lasvegas：IEEE Computer Society，2016：770-778.
26. Han S, Mao H, Dally W J. Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding[J]. Fiber， 2015，56（4）：3-7.
27. Iandola F N, Han S, Moskewicz M W, et al. SqueezeNet: AlexNet-Level Accuracy with 50x Fewer Parameters and <0.5MB Model Size. arXiv： 1602.07360.
28. Howard A G, Zhu M, Chen B, et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv：1704.04861.
29. Lyons M, Akamatsu S, Kamachi M, et al. Coding Facial Expressions with Gabor Wavelets[C]// International Conference on Automatic Face and Gesture Recognition，California：IEEE Computer Society，2002：200-205.
30. 王太勇, 蒋奇. 管道缺陷定量识别技术的研究[J]. 天津大学学报:自然科学与工程技术版, 2003，36（1）：55-58.
31. Searle S R. Linear Models for Unbalanced Data[M]. Wiley-Interscience, 2006.
32. Navneet D, Bill T, et al. Histograms of Oriented Gradients for Human Detection[C]// Computer Society Conference on Computer Vision and Pattern Recognition，California：IEEE Computer Society，2005：886-893.
33. Bar-Hillel A, Hertz T, Weinshall D. Object Class Recognition by Boosting a Part-Based Model[C]// Computer Vision and Pattern Recognition，California：IEEE Computer Society，2005：702-709.
34. Kanopoulos N, Vasanthavada N, Baker R L. Design of An Image Edge Detection Filter Using the Sobel Operator[J]. IEEE Journal of Solid-State Circuits，2002，23（2）：358-367.
35. Cortes C, Vapnik V. Support-Vector Networks[J]. Machine Learning，1995，20（2）：273-297.
36. Geurts P, Ernst D, Wehenkel L. Extremely Randomized Trees[M]. Kluwer Academic Publishers，2006.
37. Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the Inception Architecture for Computer Vision[C]// Computer Vision and Pattern Recognition，Lasvegas：IEEE Computer Society，2016：2818-2826.
38. Fawcett T. An Introduction to ROC Analysis[J]. Pattern Recognition Letters，2005，27（8）：861-874.
39. Ruder S. An Overview of Gradient Descent Optimization Algorithms. arXiv： 1609.04747.
40. Sylvia F. Data Augmentation and Dynamic Linear Models[J]. Journal of Time Series Analysis，2010，15（2）：183-202.
41. Qian N. On the Momentum Term in Gradient Descent Learning Algorithms[J]. Neural Netw，1999，12（1）：145-151.
42. Olga Russakovsky, Jia Deng, Hao Su, et al. ImageNet Large Scale Visual Recognition Challenge[J]. International Journal of Computer Vision，2015，115（3）：211-252.
43. Gao H, Li Y X, Geoff P, et al. Snapshot Ensembles: Train 1, Get M for Free[J]. Journal of Time Series Analysis，2017，12（5）：167-173..
44. Chen H, Tsai S S, Schroth G, et al. Robust Text Detection in Natural Images with Edge-Enhanced Maximally Stable Extremal Regions[C]// IEEE International Conference on Image Processing，Washington：IEEE Computer Society，2011：2609-2612.
45. Jiao S, Li X, Lu X. An Improved Ostu Method for Image Segmentation[C]// International Conference on Signal Processing，Nevada：IEEE Computer Society，2007：164-166.
46. 王序哲. 局部自适应二值化方法研究[J]. 软件导刊，2011，10（11）：13-14.
47. 董越. 视频中的文字提取技术[D]. 天津：天津大学，2014：59-73.
48. Kaufman L, Rousseeuw P J. Finding Groups in Data: An Introduction to Cluster Analysis[M]. DBLP，1990.
49. 唐琎, 李青. 一种快速的模板匹配算法[J]. 计算机应用，2010，30（6）：1559-1561.

# 攻读硕士学位期间发表的论文及其他成果

**（一）撰写的学术论文**

1. Kefan Chen, Hong Hu, Long Chen, Caiying He and Chaozhan Chen. An Intelligent Sewer Defect Detection Method Based on Convolutional Neural Network [C]// Proceedings of the 2018 IEEE International Conference on Information and Automation(ICIA), Wuyi Mountain，Fujian, China，August 2018: 1311-1316.

# 哈尔滨工业大学学位论文原创性声明和使用权限

学位论文原创性声明

本人郑重声明：此处所提交的学位论文《基于视觉的地下管道缺陷检测方法研究》，是本人在导师指导下，在哈尔滨工业大学攻读学位期间独立进行研究工作所取得的成果，且学位论文中除已标注引用文献的部分外不包含他人完成或已发表的研究成果。对本学位论文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。

作者签名： 日期： 年 月 日

学位论文使用权限

学位论文是研究生在哈尔滨工业大学攻读学位期间完成的成果，知识产权归属哈尔滨工业大学。学位论文的使用权限如下：

（1）学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文，并向国家图书馆报送学位论文；（2）学校可以将学位论文部分或全部内容编入有关数据库进行检索和提供相应阅览服务；（3）研究生毕业后发表与此学位论文研究成果相关的学术论文和其他成果时，应征得导师同意，且第一署名单位为哈尔滨工业大学。

保密论文在保密期内遵守有关保密规定，解密后适用于此使用权限规定。

本人知悉学位论文的使用权限，并将遵守有关规定。

作者签名： 日期： 年 月 日

导师签名： 日期： 年 月 日

# 致 谢

时光飞逝，转眼间两年半的硕士生涯即将接近尾声。这短暂而又宝贵的时光是我人生中一段不可或缺的旅程，在这个过程中我不仅学到了丰富的专业理论知识，学会了严谨务实的科研态度，也懂得了许多为人处世的道理。这是一段苦痛和艰辛的旅程，也是一段收获和成长的旅程。在这里，我要真心地感谢每一个关心过、帮助过我的人。

首先，我要感谢我的研究生导师胡泓教授，胡老师严谨的治学风格、刻苦的专研精神、真诚的待人之道和积极的人生态度是我一生都要学习的榜样。在课题选择方面，胡老师在给予我充分自由空间的同时也能准确地把握课题研究大方向，给予悉心的指导和中肯的建议。在研究进度推进方面，胡老师会定期举行例会，汇报各阶段课题进展，就课题中碰到的难题进行深度的探讨，也会毫不保留地利用自己的资源消除我们课题中理论和设备方面的阻塞点。在生活上，胡老师也对我们关怀备至，鼓励我们在学习之余多多参加体育运动，养成健康的生活方式，也会定期组织旅游团建活动促进实验室同门之间的交流。在此，谨对恩师致以最崇高的敬意和真挚的感谢！

感谢实验室的师兄师姐，在学习和生活上给了我很大的帮助，让我在专业课学习和课题研究上少走了许多弯路。感谢袁晓建、许江长、钱荣辉、肖洋师兄、张雯、左丹师姐在课题选择和学习方向以及课题难点上给予我的指导和建议。感谢陈建师兄、邓红红师姐在企业实习的过程中帮助我更快的融入到工作环境中。衷心的希望各位师兄师姐能够工作顺利，生活圆满。

感谢实验室同级的同学陈超湛、陈建楷、王照华、薄纯强、柴培林、赵小贺、郭毅强等人，以及我的舍友吴啟菡和汪欣，感谢你们给予我的不论是学习上的还是生活上的帮助和关心，真心的祝愿你们学业有成，前程似锦。

也要感谢一直在我背后给予关怀和支持的父母，你们无私的付出一直激励着我不断前行，这一路走来你们一直是我心灵的港湾，感谢你们！

最后还要感谢我的母校哈尔滨工业大学（深圳），为我们提供了良好的学习生活环境，我深深的为作为母校的一份子而感到骄傲和自豪，真心的祝愿母校越来越好！