油气管道线路感知技术

调研报告

# 摘 要

考虑油气管道线路感知技术研究项目要求，并结合管线的实际应用场景，我们将相调研报告分为以下三部分：  
 针对工程机械的行为态势，研判其对管线的威胁态势，提出一种多尺度时序建模算法FPN-X的框架设计，直接分析视频数据，对目标视频的行为动作检测与预测。首先调研时序动作检测任务的的技术发展，并针对工程机械的行为态势识别，提出的多尺度时序建模算法FPN-X的框架设计。FPN-X算法把时序动作检测分为：前置视频特征提取网络、特征金字塔FPN、时序动作提议生成和动作分类四部分。将时序和空间两个通道的多个分类结果融合，识别目标动作。

针对管道沿线大规模监控产生的海量视频数据存储及查询问题，提出一种基于视频图像资源中用户关注的活动目标数据结构化框架，形成有价值的信息积累，能够快速准确实现特征检索、以物搜人、轨迹回放等应用功能。首先调研目标视频结构化的主要识别方法及提取框架，并针对管道智能监控感知场景，提出将FF-YOLO算法与动作序列算法相结合，解决低质量视频的目标识别问题并研判其运动态势，通过目标检测、场景识别、动作识别提取视频帧集合中的目标、场景、动作信息，生成相应的一段文字描述。

对计算能力不足问题，我们提出采用算力感知网（CAN）进行解决。针对计算设备的不稳定性，我们提出一种使用GSM手动控制的断电开关器实现设备断电重启的方法，提高监控系统运行效率，降低风险漏报率。

# 目 录

[**摘要**](#_Toc69290648)

[第一章 基于时序建模的视频动作检测与预测方法 1](#_Toc92822285)

[1.1 研究背景及意义 1](#_Toc92822286)

[1.2 国内外研究现状 1](#_Toc92822287)

[1.3相关算法 2](#_Toc92822288)

[1.3.1 动作识别 2](#_Toc92822289)

[1.3.2 时序动作定位 3](#_Toc92822290)

[1.4需求分析 5](#_Toc92822291)

[1.5解决方案 5](#_Toc92822292)

[1.4.1 视频特征提取 6](#_Toc92822293)

[1.4.2 FPN模块 8](#_Toc92822294)

[1.4.3 多尺度提议生成 8](#_Toc92822295)

[1.4.4 提议动作分类 9](#_Toc92822296)

[1.6本章小结 10](#_Toc92822297)

[1.7参考文献 10](#_Toc92822298)

[第二章 基于视频数据的场景结构化建模方法 11](#_Toc92822299)

[2.1 研究背景及意义 11](#_Toc92822300)

[2.2 国内外研究现状 12](#_Toc92822301)

[2.3 需求分析 14](#_Toc92822302)

[2.4 视频结构化方法 14](#_Toc92822303)

[2.5 本章小结 17](#_Toc92822304)

[2.6 参考文献 17](#_Toc92822305)

[第三章 面向大规模智能监控系统的算力共享网络解决方案 18](#_Toc92822306)

[3.1 研究背景及意义 19](#_Toc92822307)

[3.2 国内外研究现状 21](#_Toc92822308)

[3.3 需求分析 22](#_Toc92822309)

[3.4 解决方案 23](#_Toc92822310)

[3.5 本章小结 25](#_Toc92822311)

[3.6 参考文献 25](#_Toc92822312)

[第四章 总结 27](#_Toc92822313)

# 第一章 基于时序建模的视频动作检测与预测方法

## 1.1 研究背景及意义

时序动作检测和视频描述是智能分析研究领域中的一个重要方向，也是一项具有十分挑战性的研究。时序动作检测由视频特征提取、候选提议生成评估和动作提议分类三部分组成，其输出的信息同时包含时序动作提议输出的目标移动位置信息和行为识别信息；视频描述中密集事件描述任务是视频特征提取、候选提议生成评估和基础上，增加了基于LSTM网络的描述生成语言模型，给出了目标位置信息的同时用语言模型生成自然语言描述。时序动作检测的研究严重依赖对目标动作的时序提议效果，视频数据具有结构复杂、目标变化多样且持续时间长短不一等问题，使得对应时序提议方法存在着对视频特征利用不充分。目标动作时序区域检测困难等问题。视频自动翻译图像的研究室2014年前后图片的描述生成开始，2015年图像翻译推广到视频领域，视频描述成为新兴的研究课题，而视频描述本身就有内容处理复杂，语法解析难等挑战，尤其在2017年视频理解竞赛ActivityNet提出密集事件描述的概念后，把问题印象了包含多个目标事件、且目标时间间存在时序的于一相关性的视频描述生成，这给视频描述提出更大挑战。同时利用机器自动翻译视频内容是机器理解视频的最终目的，存在着巨大的研究价值。

## 1.2 国内外研究现状

时序动作检测任务的目标是从未知输入的视频中定位目标动作的开始和结束时间并给出目标动作的分类结果。在符合真实场景下的视频具有无组织、语义信息丰富的特点。规范的时序动作检测工作开始于Ke Y.等人、Rohrbach M.等人分别使用hand-centric和object-centric特征来检测固定摄像机厨房hengren视频中的具体动作。更广泛领域的时序动作检测工作是在经典的视频理解数据集THUMOS-14出现后展开，其中Oneata D.等人、Wang L等人和Karaman S.等人分别用DT特征、单帧级CNN特征或者融合声音特征等方式，通过滑动窗产生候选提议方式设计时序动作检测框架。同时开始出现基于时空信息的时序动作检测方法，在时间维度仍然使用基于滑动窗的动作提议生成方法。2016年，Yuan J.等人提出基于传统的iDT特征的时序动作检测算法，该算法基于iDT特征对视频提取了一种分数分布金字塔特征，之后再使用LSTM对PSDF特征序列进行处理，并根据输出的帧级的行为类别置信度分数处理得到行为片段的预测，再THUMOS-14数据集上达到了18.8%。虽然该方法证明了传统的特征提取iDT的有效性，但是iDT特征的提取实在太耗时和耗存储空间了，iDT特征的大小甚至远远小于视频 的本身。受图片领域目标检测算法RCNN的启发，Shou Z.等人提出多片段网络SCNN，先用滑窗和提议生成网络的方式产生并判断候选视频判断为动作的概率，然后使用分类网络和定位网络分别给出动作分类概率和重叠度相关的损失函数，使网络能更好的估计视频片段和准确位置，最后采用NMS来去除重叠片段完成预测。为了获得更准确的动作边界，Shou Z.等人又基于C3D网络设计了一个卷积逆卷积网络，通过帧级动作类别概率预测，对时序动作检测中的动作边界进行微调，使动作更加准确，从而提高mAP。2018年CVPR上密歇根大学和谷歌研究院Chao Y.等人提出了基于R-C3D进行了多项优化的时序动作检测网络TAL-Net，主要优化思路是提出一个应用多尺度结构时应时间片段变化的感受野对齐方法，并扩展感受野提取和融合光流信息以获得更好的提取时序信息。

2018年ActivityNet竞赛的时序动作检测冠军方案是来自上海交大的边界敏感网络BSN。BSN算法主要由时序评估模块TEM和提议评估模块PEM两个部分，TEM的输入是通过预训练的双流法模型TSN得到输入视频的特征，利用FCN输出和输入特征等长的“动作”、“开始和“结束三条概率分布曲线，然后根据概率分布曲线得到候选提议集合，之后通过PEM对每个提议的置信度进行分析，再通过NMS去除重叠度高的候选提议给出的结果。这个方案主要存在两点问题，一是对不同时序长度的视频通过采样方式统一为固定长度输入到TEM，这让对尝试品种的多尺度动作检测的召回率不高。例如假设视频的特征位F=,其中包含目标动作检测A = {(, ), (, ),…}，若采用BSN算法，输入TEM模块的特征位F的采样=，当存在目标动作满足-<-时，则导致该动作难召回，而当- > /2时，在（, ）之间往往出现大于阈值的开始和截至时间，可能出现（，）和(,)的错误检测结果。二是输入特征采用的是TSN行为识别结构的MLP输出，然后输入时序评估模块用一维卷积提取时序特征，忽略了不同通道对结果的影响。

## 1.3相关算法

### 1.3.1 动作识别

动作识别的目标是识别出视频中出现的动作，通常是视频中物体的动作。视频可以看作是由一组图像帧按时间顺序排列而成的数据结构，比图像多了一个时间维度。动作识别不仅要分析视频中每帧图像的内容，还要从视频帧之间的时序信息中挖掘线索，动作识别是视频理解的核心领域。

如图1.1所示，视频是由一系列的图像帧组成的，每帧图像各自前馈（Feedforward）一个图像分类模型，不同帧的图像分类模型之间相互共享参数，得到每帧图像的特征之后，对各帧图像进行汇合（Pooling），例如采用平均汇合，得到固定维度的视频特征，最后经过一个全连接层和Softmax激活函数进行分类以得到视频的类别特征，最后再进行预测。

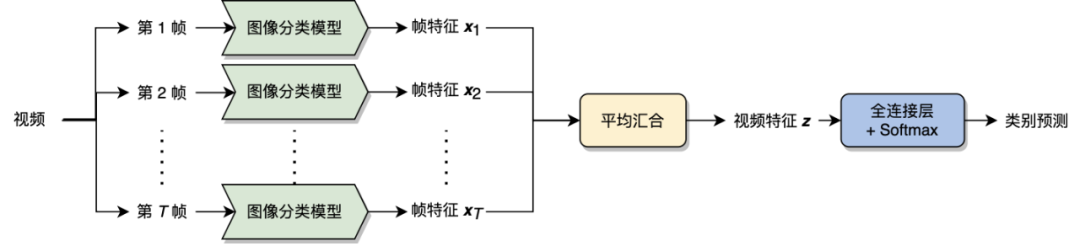


图1.1 视频类别预测流程图

汇合的方法有很多，主要包括：平均汇合、NetVLAD/NeXtVLAD、NetFV、RNN等。另外，可以借助一些传统的算法来补充时序关系，例如，双流法利用光流显示地计算帧之间的运动关系，TDD利用iDT计算的轨迹进行汇合等。基于2D卷积的动作识别方法的一个优点是可以快速吸收图像分类领域的最新成果，改变骨骼网络，新的图像分类模型可以十分方便地迁移到基于2D卷积的动作识别方法中。

3D卷积可以从视频片段中同时学习图像特征和相邻帧之间复杂的时序特征，最后利用学到的高层级特征进行分类。相比于2D卷积、3D卷积可以学习到视频帧之间的时序关系。由于3D卷积神经网络的参数量和计算量比2D卷积神经网络大的很多，不少研究工作专注于对3D卷积进行低秩近似，如FSTCN、P3D、S3D等。

### 1.3.2 时序动作定位

时序动作定位也称为时序动作检测，是视频理解的另一个重要领域。动作识别可以看作是一个纯分类问题，其中要识别的视频基本上已经过剪辑(Trimmed)，即每个视频包含一段明确的动作，视频时长较短，且有唯一确定的动作类别。而在时序动作定位领域，视频通常没有被剪辑，视频时长较长，动作通常只发生在视频中的一小段时间内，视频可能包含多个动作，也可能不包含动作，即为背景类。时序动作定位不仅要预测视频中包含了什么动作，还要预测动作的起始和终止时刻。相比于动作识别，时序动作定位更接近现实场景。

时序动作定位可以看作由两个子任务组成，一个子任务是预测动作的起止时序区间，另一个子任务是预测动作的类别。由于动作识别领域经过近年来的发展，预测动作类别的算法逐渐成熟，因此时序动作定位的关键是预测动作的起止时序区间，有不少研究工作专注于该子任务，ActivityNet 竞赛除了每年举办时序动作定位竞赛，还专门组织候选时序区间生成竞赛(也称为时序动作区间提名)。

既然要预测动作的起止区间，一种最朴素的想法是穷举所有可能的区间，然后逐一判断该区间内是否包含动作。对于一个 T 帧的视频，所有可能的区间为，穷举所有的区间会带来非常庞大的计算量。有关时序动作定位的算法如下：

（1）基于滑动窗口的算法

这类算法的基本思路是预先定义一系列不同时长的滑动窗，之后滑动窗在视频上沿着时间维度上进行滑动，逐一判断每个滑动窗对应的时间区间内具体是什么动作类型。

（2）基于候选时序区间的算法

目标检测算法中的两阶段算法将目标检测分为两个阶段：第一阶段产生图像中可能存在目标的候选区域，一般一张图片可以产生成百上千个候选区域，这一阶段和具体的类别无关；第二阶段逐一判断每个候选区域的类别并对候选区域的边界进行修正。类似于两阶段的目标检测算法，基于候选时序区间的时序定位算法也将整个过程分为两个阶段：第一阶段产生视频中动作可能发生的候选时序区间：第二阶段逐一判断每个候选时序区间的类别并对候选时序区间的边界进行修正。最终将两个阶段的预测结果结合起来，得到未剪辑视频中动作的类别和起止时刻预测。

（3）自底向上的时序定位算法

自底向上的时序动作定位算法，这类算法首先局部预测视频动作开始和动作结束的时刻，之后将开始和结束时刻组合成候选时序区间，最后对每个候选时序区间进行类别预测。相比于自顶向下的算法，自底向上的算法预测的候选时序区间边界更加灵活。

除了上述算法之外，还有对时序结构信息建模的方法、逐帧预测的算法和单阶段算法等算法。

## 1.4需求分析

近些年来，随着城市的不断发展和建设，经常会遇见工地附近有各种管线的情况，而且十分复杂。在施工过程中，周围土体会被扰动，带着附近的管线产生位移，而且在施工过程中，对附近管线产生附加应力或者变形超出管线的应力时，管线就会被破坏，导致施工进程受到影响，而且对于管线的修复也会带来一定的人力和物力的开销。西气东输作为我国重要的运输管线之一，如何避免大型机械对管线所造成的影响已成为迫在解决的问题。针对本文所解决的问题，对以下需求进行分析：

（1）数据获取

本文所提出的解决方案主要是对视频进行处理，需要在大型机械出现的场景中获取大量的视频。在目前技术的支持下，可以通过监控的方式对管线周围场景中的大型机械进行视频捕捉。

（2）视频处理

在获取完视频数据之后需要对视频数据进行标注处理，对视频中的大型机械的类别和动作进行标注。

（3）对视频中的大型机械的动作进行识别

对视频中大型机械动作的识别是本项工作的核心任务，对视频中大型机械进行定位并对其进行动作识别，方可判断该大型机械是否对该区域的管线造成影响。

## 1.5解决方案

为了针对目前边界敏感网络BSN在固定大小特征维度上产生候选提议造成的多尺度groundtruth目标动作检测困难问题，借鉴了FPN在多尺度特征图上进行预测的思想，先通过FPN结构得到多个尺度的特征图，然后借鉴BSN的思路在不同尺度的特征图上得到了每个尺度上的目标动作、动作开始和动作结束的三类概率分布曲线。同时为了更好的连续视频上下文信息和注意到不同通道的影响，采用TSN双流法特征提取网络中最后一个池化层的输出作为FPN-TAD算法的输入，采用2D时序-通道特征建模，得到了更好的视频特征表示。

如图X位提出的多尺度时序建模算法FPN-X的框架设计。FPN-X算法把时序动作检测分为四个部分，前置视频特征提取网络、特征金字塔FPN、时序动作提议生成和动作分类四部分。第一部分为基础的视频特征提取，主要基于双流法的代表结构TSN来提取输入视频的高位语义特征信息。第二部分为FPN模块，把输入的视频特征表示通过2D卷积操作得到多个尺度的特征金字塔，采用top-down的方式融合其中的部分小尺度特征图作为第三部分时序动作提议模块的输入。第三部分为时序动作提议生成模块，包括时序评估模块TEM、提议评估模块PEM和NMS后处理等组成部分。其中时序评估模块在输入的多尺度特征图的时序维度上分别进行一维卷积时序评估，生成动作区域、开始位置和结束位置三条概率曲线，分别表示当前特征对应的视频区域为动作、动作开始和动作结束的概率，不同尺度的特征图上均会得到这样的动作概率分布曲线；提议评估模块以及NMS后处理则主要针对TEM产生的候选提议给出置信度评分，并通过soft-NMS对候选提议进行re-ranking。第四部分位动作分类模块，根据第三部分得到的提议位置，取第一部分产生的候选特征用TSN模型采用的single-shot的分类预测方式，得到候选提议包含的多个snippet的分类结果，最后融合时序和空间两个通道的多个snippet分类结果最终的输入候选提议的动作分类结果。

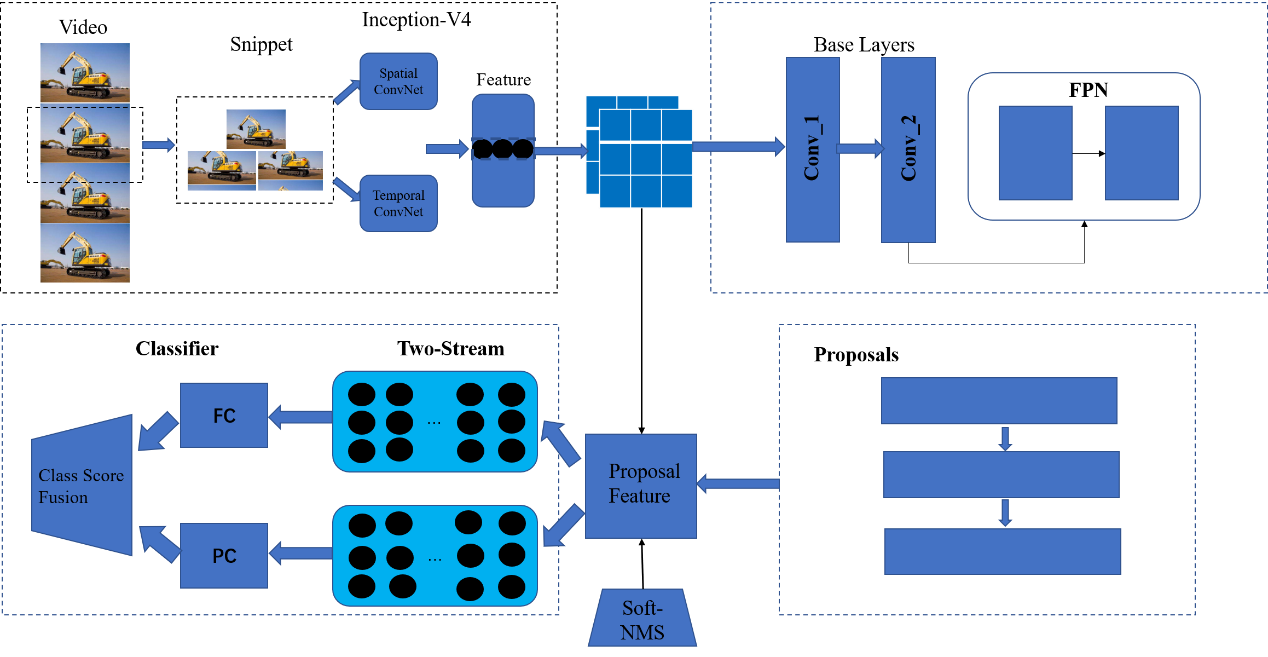


图1.2 双流法特征学习框架图

在视频的特征提取网络结构中，相比于用C3D去提取特征，双流法有者更明显的性能优势，本次使用双流法去学习视频里丰富的特征表示。具体过程如图1.2表示，特征提取分为空间特征卷积和时序特征卷积的两个通道进行，其中空间特征的输入为视频帧RGB图像，时序特征的输入为帧间密集光流，两个通道均采用Inception-v4作为基础的2D卷积。

### 1.4.1 视频特征提取

特征提取的具体过程如图1.3所示：1，输入含有T帧的输入视频V,V = ,把视频分为T/个snippet作为特征提取的最小单元，其中代表snippet的长度，建议采用=16。2，对每个snippet，从共有帧的序列种随机选择一张RGB的图像计算其光流，光流计算结果包含x方向上的flow\_x和y方向的flow\_y两张光流图像。3，把RGB输入空间卷积网络，把光流图像输入到时序卷积网络，取最后Inception-v4的最后一个pooling层输出作为视频特征的提取的最终结果。

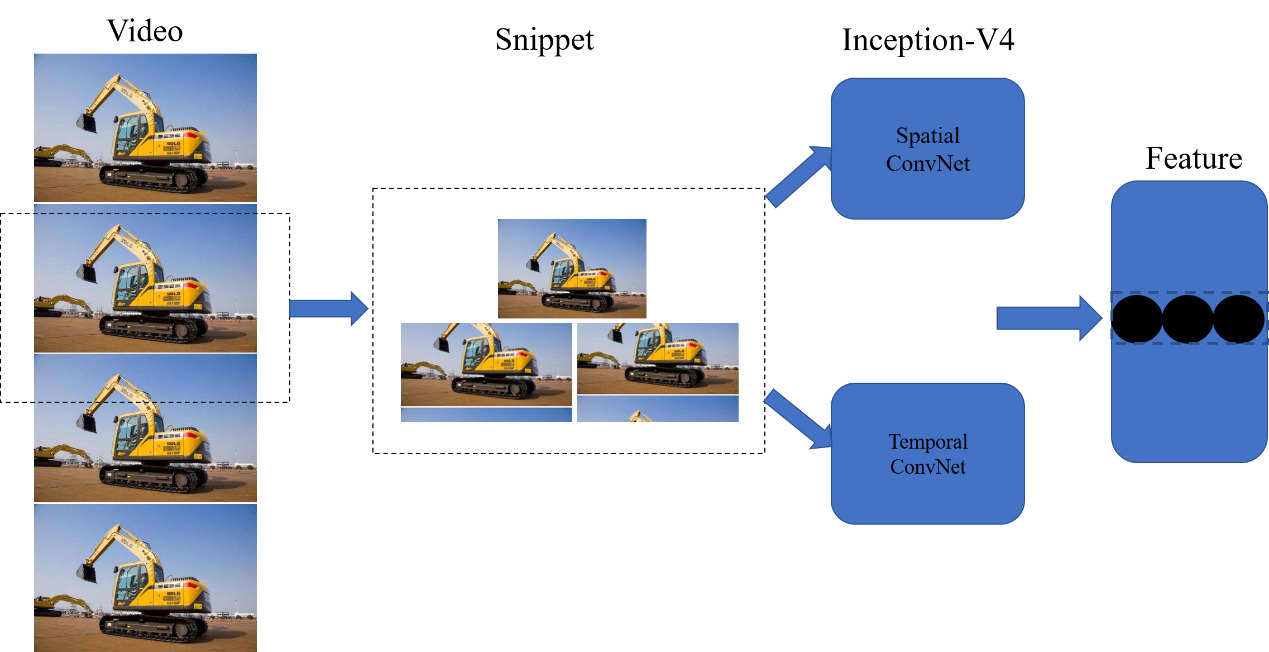


图1.3 特征提取流程图

卷积神经网络Inception-v4的整体网络结构和输入图像如图1.4所示，视频特征的提取过程以RGB图像特征提取为例，Inception-v4的输入为从snippet中抽取视频帧，被Resize到299×299×3的大小，其中299×299为输入图片的尺寸，3表示RGB三个通道。Inception-v4网络首先使用Stem模块完成数据的预处理，Stem模块采用多次卷积+2次pooling组成，得到35×35×384的特征图，其中pooling采用了Inception-v3论文中提到的卷积+pooling并行的结构，用来防止bottleneck问题。Stem之后采用了3种结构共14个Inception模块，3种Inception模块之间的Reduction模块的作用类似pooling，同样使用了并行结构来防止bottleeneck问题。在全部卷积层完成后得到了8×8×1536的特征图（8×8大小，共1536个通道），经过平均池化即可得到1536维的最终图像特征。挂光流图像的特征提取过程类似，网络同样为Inception-v4，最终得到的图像特征维度同样是1536个通道。因此每个snippet可以得到1536×2的特征表示，考虑到不同长度的视频影响，将得到的视频特征采用线性插值的方式固定为512。即，每个视频的特征表示为512×1536×2的特征图。

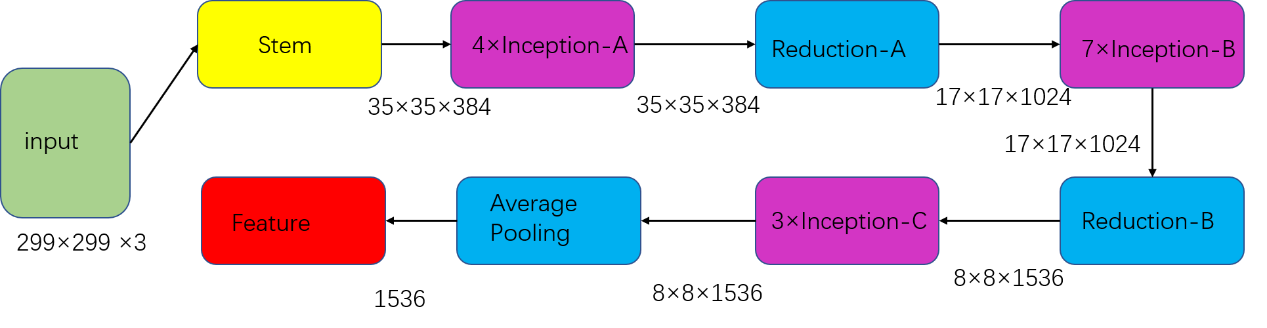


图1.4 Inception-v4的结构图

### 1.4.2 FPN模块

特征金字塔FPN设计结构由基础层、FPN和特征图宽度压缩三部分组成，分别完成输入特征的预处理、特征金字塔和输出特征宽度压缩的功能。该部分的输入为前置视频提取网络的输出，在本模型中对应512×1536×2大小的特征图。为了更好的联系时序上下文信息和通道特征信息，把512×1536×2的特征图视为时序和空间两个维度的512×1536大小的图像处理，采用图像特征提取常用的2D卷积核对时序和通道联合建模。其中，BaseLayers包含Conv\_1和Conv\_2两层卷积神经网络，Conv\_1的卷积核大小kernel=(1,9)，步长stride=(1,3)，使得输出的特征图为512×512×2；Conv\_2的设置为kernel=(3,3),stride=(2,2)，得到256×256×64的特征图，作为标准2D卷积FPN的输入。FPN部分包含5层卷积Conv\_3~Conv\_7，卷积大小和Conv\_2一直，通道数都是256。今次FPN得到了256、128、64、32、16五种特征图，考虑太小的特征图会造成开始结束位置的不准确，对16，32，64三种尺寸采用top-down方式叠加在64尺寸上进行融合，最终得到(256,128,64)三个尺寸的特征图。在宽度压缩部分采用Conv\_8和池化层把宽度压缩为1.最终，FPN模块产生的特征图为（256,128,64）×256，其中×256表示通道数，分别输入到TEM生成三种尺度的动作概率分布曲线。

### 1.4.3 多尺度提议生成

多尺度时序动作提议参考边界敏感网络BSN，由时序评估模块TEM和提议评估模块PEM两部分组成，分别完成生成动作概率分布曲线和给出候选提议的置信度评分功能。

（1）时序评估模块TEM

时序评估模块TEM接受FPN模块的输出作为输入，在输入特征的时序维度上进行一维卷积。 对于输入特征尺度,用和分别视频的特征序列和标注(annotation),其中表示输入FPN输出的特征图尺寸数量。对于包含的正确(ground truth)的动作实例,将包含的区域作为动作区域开始区域可表示为，相应的结束区域可以表示为， 其中表示动作实例的持续时间。

（2）提议评估模块PEM

FPN的引入，使得TEM部分会产生= {256,128,64}三种尺度的时序概率序列。对于每一个时序概率序列，认为的为候选开始位置得到集合,同样的方式可以得到结束位置集合，将开始位置和结束位置组合可获得Np个候选提议其中表示候选提议对应的特征。提议评估模块PEM对于给定的候选提议;给出置信度评分然后计算其与标签的,保存最大重叠度分数。此时，候选提议可以表示为.根据可以将所有的候选提议π=Ap划分为正提议A书和负提议A股R两个类别，划分标准是如果> 0.7,则; .如果 <0.3,则分。考虑到数据平衡的问题，实验中在划分结束后采用随机抽样的方式以保证:= 1:2.

（3）后处理

在预测中，可采用如.上所述的FPN、TEM和PEM模型处理输入特征得到候选提议集合Api = {{an = (ts, terPeomn,pE,p;)1=1}h[n", Ap,x对应lz = 256，以此类推。为了得到最终的提议集合，需要首先对不同尺度的预测结果进行汇总，然后对提议评估得分进行融合，最后根据得分抑制冗余候选提议。

### 1.4.4 提议动作分类

PEM模块输出的是时序动作提议的结果，包含动作的开始结束和结束时间。提议动作分类则需要根据PEM得到的proposal结果，给出提议动作的类别。对于给定的提议，取对应的前置视频特征提取模块的提议特征，其中和分别表示空间特征和时序特征，长度为。

FPN-TAD网络的训练分为四个阶段，第一阶段通过动作识别任务在Activitynet-1.3数据集对前置视频特征提取网络进行预训练，得到输入snippet得到1536维特征的Inception-v4模型；第二阶段训练时序动作提议生成网络，将前置特征提取网络的输出作为输入，训练得到多尺度时序动作概率分布曲线；第三阶段基于时序动作提议的结果训练时序评分网络筛选候选提议；最后通过候选提议特征训练动作分类器。

## 1.6本章小结

本文所提出的解决方案主要是对视频进行处理，对视频中的大型机械的类别和动作进行标注。对视频中大型机械动作的识别是本项工作的核心任务，对视频中大型机械进行定位并对其进行动作识别，判断该大型机械是否对该区域的管线造成影响。

## 1.7参考文献

# 第二章 基于视频数据的场景结构化建模方法

长输油气管道安防主要依靠人力巡检巡查，西气东输每年投入的人力巡护成本较高，为了响应国家油气体制改革要求，西气东输公司率先开展了“智能管道 智慧管网”的探索与建设，并在管道智能监控感知威胁态势方面取得了实质性的突破，本报告主要集中在基于视频数据的场景结构化建模方法。将视频图像资源中用户关注的活动目标数据进行结构化解析,形成有价值的信息积累,可以提供特征检索、以物搜人、轨迹回放等应用功能，本主题首先调研目标视频数据结构化的主流实现方法，并针对管道智能监控感知场景进行建模，识别主要目标的结构化信息进行讨论。

## 2.1 研究背景及意义

在智能管道、智慧管网建设深入开展的背景下，由于大规模视频产生的海量数据让监控中心的工作变得相对繁重且效率低，监控视频除了时间和空间的属性外，并没有其他的标签。除了按照时间和地点查找相应的视频外，大多的视频只能靠人慢慢甄别。并且由于监控视频受自身硬件设施及光照条件的限制，在这些低质量的视频数据中进行比对工作不仅困难且准确度低。

视频结构化，即视频数据的结构化处理，就是通过原始视频进行智能分析，提取出关键信息，并进行文本的语义描述。视频结构化是一种视频内容信息提取的技术，它对视频内容按照语义关系，采用时空分割、特征提取、对象识别等处理手段，组织成可供计算机和人理解的文本信息的技术。从数据处理的流程看，视频结构化描述技术能够将监控视频转化为人和机器可理解的信息，并进一步转化为可以被用户所用的情报，实现视频数据向信息、情报的转化。

目前视频结构化分析服务器采用目标检测、特征提取、对象识别、深度学习等分析手段提取视频图像中的目标对象及运动轨迹；对目标对象进行分类为人、车、物、行为、事件等不同类别；根据目标轨迹信息可生成浓缩视频；进一步提取目标的高层特征包括目标颜色特征，分类特征，速度特征等，并把目标的轨迹信息以及高层特征形成高效的索引数据。

通过视频图像结构化技术，实现视频智能分析和数据挖掘，让视频从人工抽检向智能提取转变，实现智慧化的信息提取、分析、应用，让人从观看视频监控劳动中解脱出来，进行高效、准确的视频图像实战应用。具有以下三个优点：

（1）视频查找速度得到极大的提升。视频结构化之后，从百万级的目标库中(对应数百到一千小时的高清视频)查找某张截图上的工程机械目标，数秒即可完成;千万级目标的库中查找，几分钟即可完成(如果实现云化，速度会更快)。在结构化基础上进行检索查询，可以解决快速目标查找问题。

（2）存储容量极大的降低，经过结构化后的视频，存储人的结构化检索信息和目标数据不到视频数据量的2%;对于车辆，不到1%;对于行为降得更多。存储容量极大地降低，可以解决视频长期存储的问题。

（3）视频结构化可以盘活视频数据，可作为数据挖掘基础，视频经过结构化处理后，存入相应的结构化数据仓库，对各类的数据仓库可以进行深度的数据挖掘，充分发挥大数据作用，提升视频数据的应用价值，提高视频数据的分析和预测功能。

## 2.2 国内外研究现状

近年来，随着视频数据的巨增，视频结构化技术也逐渐成为计算机视觉热点问题，业内学者也高度关注这一问题，在影视视频、短视频等海量视频各式各样的呈现背景下，需要的是后台的服务器能够对视频数据进行更安全的存储传输和更便捷快速的检索推荐，尤其是近年来的推荐技术，根据用户对不同类型、不同风格的喜爱度来推送相关的视频。这些技术更需要我们利用视频结构化技术来获取更准确的视频描述来确定视频的标签，而不是通过人工主观判断设定标签，导致效率与准确率较低。在监控安防领域，据相关部门统计,我国的摄像头总数达到了 1.76 亿个，遍布我国每一个地方。每天产生的数据是巨大的。我们也迫切需要视频结构化系统来判断视频画面中人员、车辆等感兴趣的目标，并对其实现信息自动化语义生成和标注，便于存储和检索。总之，由于视频数据的爆炸式增长以及视频数据推的业务的巨大价值，推动着视频结构化成为视频处理、视频分析领域的研究热点，目前国内外已经有不少的研究机构和公司进行了研究并取得不俗成果。

国外学者对视频结构化技术研究相对国内起步较早，早在 1970 年英国出现ADVISOR 项目的研究，研究目的是监控视频中异常行为的识别和自动报警。哥伦比亚大学研发了一种自动检索系统 VideoQ[1]，可以通过视频内容的一些视觉特性来进行视频检索，完成对剪辑过的复杂视频的检索、标注任务。Almaden 实验室研究的 Cue Video 也是一种视频检索系统[2]，该系统可以自动实现镜头的边界检测，提取关键帧，并利用关键帧作为图片信息作为视频的可视化描述。Mecocci[3]等在文献中介绍了关于检测异常事件的监控系统的设计，目标是在不同场景下根据自学习特性在特定环境下的特定行为与异常行为相对比而触发警报。Duque 等人提出了基于目标追踪的视频异常监控系统 OBSERVER[4]，采用背景消除法分割目标，对目标进行跟踪，通过追踪目标轨迹的匹配程度来判断异常情况。随着机器学习、深度学习技术不断引入图像与视频处理技术中，卷积神经网络来实现视频结构化描述模型的方法不断被提出，Ibrahim 等提出了一种新的基于深度学习的视频特征表示方法[5]，利用深度学习技术从视频中的帧序列中提取的基于图像的特征。然后通过映射方法将提取的特征转化为一个深度特征视频矩阵，通过特征矩阵对视频进行结构化描述。Hori 等人提出基于注意力机制的视频结构化模型[6]，该方法提取图像特征、运动特征、音频特征融合的多模态特征，再利用循环神经网络来实现语义描述从而实现视频结构化描述。

此外，国内对视频结构化技术的研究起步相对较晚，但是发展十分迅速，海康威视、旷世科技、商汤科技等企业和清华大学、中科大等高校极大的推动视频结构化技术在我国的发展。例如在智能安防领域，公安部第三研究所研发的视频结构化系统，主要功能为分析道路交通事件视频内容，转化为数字及其文字索引事件。针对事件索引出视频内容，反映事件发生的情况，实现道路监控视频的语义感知和服务。另外，海康威视公司推出的视频结构化系统，采用目前业界主流技术，在各类复杂场景下具有良好的适用性，可以准确实现在海量高清视频中提取多种有价值的信息，并对其进行存储并服务于用户。

通过对国内外视频结构化技术的历程的了解，可见视频结构化技术随着近来多媒体技术的发展不断发展，同时视频结构化也广泛运用到监控、影视、医疗等行业，从而也对视频结构化技术提出了更高的要求。总之，视频数据的安全存储和准确检索，依然是未来视频结构化技术的发展方向，依托于深度学习技术等来实现视频结构化描述也依旧是世界学者研究的热门问题之一。

## 2.3 需求分析

（1）将工程机械的视频数据生成结构化信息，便于视频存储和检索。

（2）低质量的视频数据中，能准确、快速识别分析工程机械的威胁态势。

（3）结构化信息的数据库存储问题，实现快速存储、分析视频信息。

## 2.4 视频结构化方法

目前视频结构化算法多用于安防与智能监控领域，其核心思想是对视频图像中关注的人、车、非机动车等目标提供更深层次的结构化解析，如出现在视频中人的样貌、衣着，车辆的类型、颜色等。而根据我们的应用场景，主要是对于挖掘机等工程机械的识别，需要考虑工程机械的种类、品牌、动作、运动趋势等信息，在海量的视频数据中能快速识别潜在威胁场景（如管道附近施工工地的挖掘机），研判其对管线的威胁态势。因此，在该场景进行视频结构化从任务实现的角度是适用的，并且在保证视频结构化稳定准确输出的前提下，可以利用输出结果，为后续决策控制模块提供支撑。

本节主要是将基于目标识别的算法与基于时序建模的视频动作检测与预测方法相结合，来对监控视频的中的数据进行特征提取，以下为视频结构化步骤和流程图如图2.1所示：

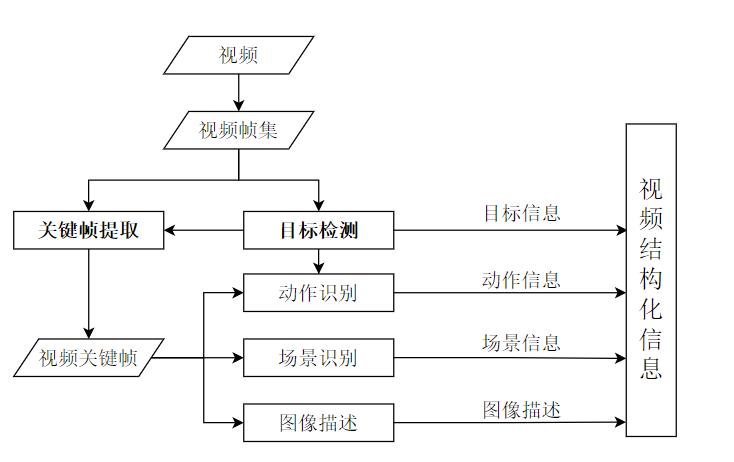


图2.1视频结构化流程图

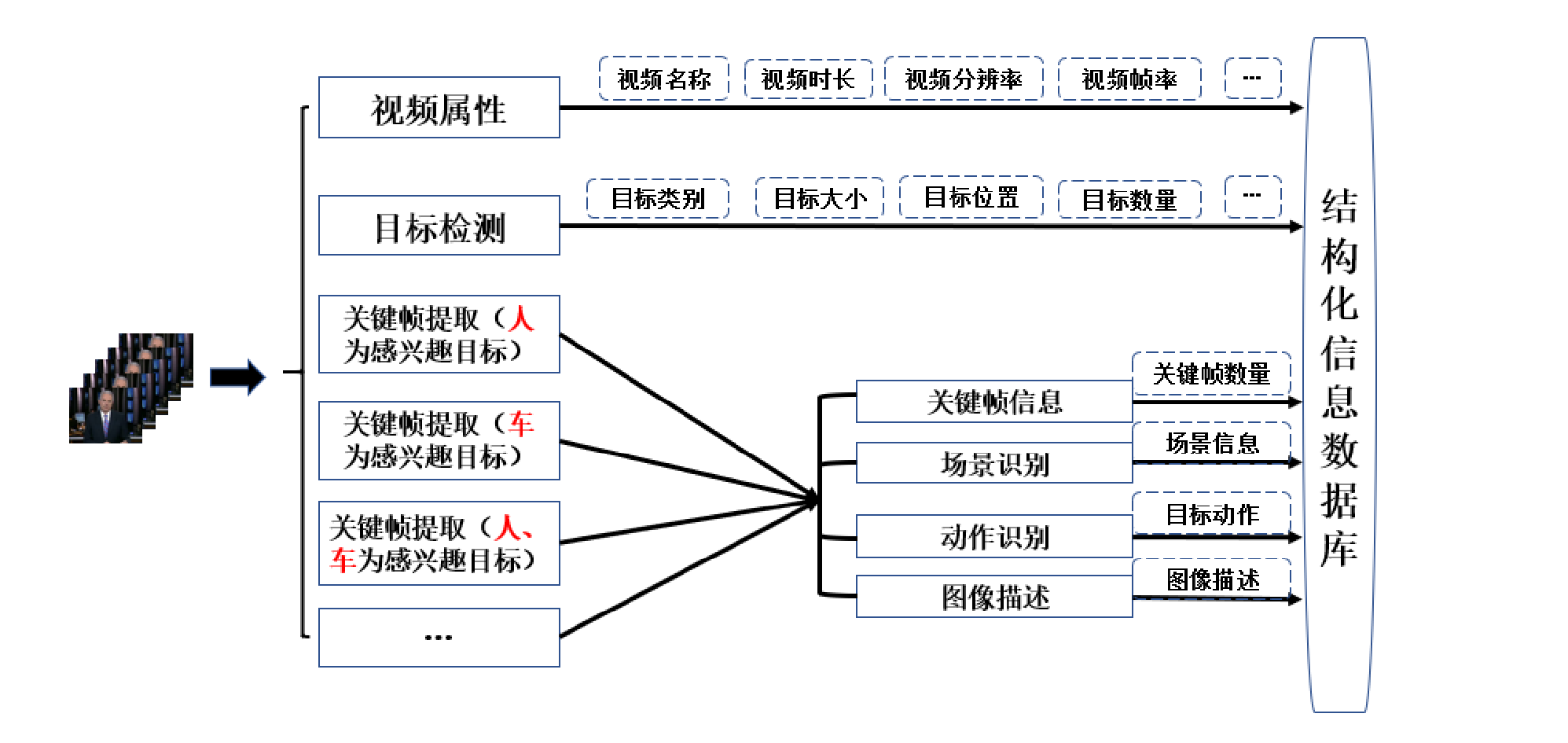
（1）通过目标检测网络算法实现复杂场景视频帧的目标信息的获取，目标信息主要包括目标的车辆类型、作业位置等信息，作为第一层结构化信息，便于对视频中出现的不同目标进行统计与检索。

（2）由于视频数据相邻帧之间的高度相似性，导致大量冗余，而目标信息也不能直观反映该帧表达的具体内容，不能生成直观的语义描述，还需要对视频关键帧提取，获取视频关键帧集。通过基于感兴趣目标分布的关键帧提取方法，该方法结合深度学习网络来进行特征提取和处理。首先对视频帧进行特征提取，获取颜色、纹理、运动、深度特征，然后将特征根据不同权值相融合获取帧间距离，通过自适应阈值算法获取镜头初检边界，再对边界帧复检获取最终镜头边界。最后采取目标检测网络计算镜头内视频帧的目标信息丰富度来选取关键帧以及过滤相似关键帧。

（3）通过基于时序建模的视频动作检测与预测方法，识别工程机械的行为状态、运动轨迹等信息，作为第二次结构化信息，研判威胁态势其对于管道的威胁。

（4）通过目标检测、场景识别、场景识别提取视频帧集合中的目标、场景、动作信息，目标信息主要反映图片内容主体是“谁”，场景信息反映图片内容发生在“什么地方”，动作信息主要针对目标是人的图片，反映人在“干什么”，通过这些信息可以反映出关键帧描述的主要内容，有利于基于场景、动作的关键词检索。

此外，为了对图片中的内容进一步语义描述，采取图像描述技术，图像描述可以对关键帧中的内容生成相应的一段文字描述，不仅用于提取高层语义信息，一定程度填补人与机器之间语义鸿沟，而且可以与获取的其他结构化信息互补，保证视频结构化信息的完整性。传统的视频结构化提取视频中的目标相关信息作为结构化信息[7]，而本文提出一种基于显著内容的视频结构化方法，方法实现框架如图2.2 所示。

图2.2视频结构化整体框架图

基于显著内容的结构化方法将视频信息以分为三个层次，视频层次是从对视频本身属性信息的提取。视频帧层次是以视频中每一帧图像的目标检测结果来分析，对目标检测结果的目标类别、目标大小、目标数量、目标分布进行存储[8]。对于关键帧帧层次，可以通过设置不同的感兴趣目标，选取不同的关键帧，例如在监控中工程机械等目标，对于不同场景提取不同的关键帧，然后基于关键帧进行场景识别、动作识别等，充分提取关键帧中的信息。因此，本文提出的关键帧是基于镜头内部再划分子镜头而筛选的，各个关键帧代表着每一个子镜头的内容。所以，提取关键帧的场景、动作、目标信息，再生成图像描述，可以代表视频不同时间段的内容信息，可以更好反映视频的内容，便于视频内容的理解。

## 2.5 本章小结

本章针对管道沿线大规模监控产生的海量视频数据存储及查询问题，提出一种基于视频图像资源中用户关注的活动目标数据结构化框架，形成有价值的信息积累，能够快速准确实现特征检索、记录获取等应用功能。

## 2.6 参考文献

1. S. F. Chang, W. Chen, H. Sundaram. VideoQ: a fully automated video retrieval system using motion sketches[C]. Proceedings Fourth IEEE Workshop on Applications of Computer Vision. WACV'98 (Cat. No. 98EX201). IEEE, 1998: 270-271
2. D. Ponceleon, A. Amir, S. Srinivasan, et al. Cuevideo: automated multimedia indexing and retrieval[C]. Proceedings of the seventh ACM international conference on Multimedia (Part 2). 1999: 199
3. A. Mecocci, M. Pannozzo, A. Fumarola. Automatic detection of anomalous behavioural events for advanced real-time video surveillance[C]. The 3rd International Workshop on Scientific Use of Submarine Cables and Related Technologies, 2003, 2003: 187-192
4. D. Duque, H. Santos, P. Cortez. The OBSERVER: An intelligent and automated video surveillance system[C]. International Conference Image Analysis and Recognition, 2006: 898-909
5. Z. A. A. Ibrahim, M. Saab, I. Sbeity. VideoToVecs: a new video representation based on deep learning techniques for video classification and clustering[J]. SN applied sciences, 2019, 1(6): 1-7
6. C. Hori, T. Hori, T. Y. Lee, et al. Attention-based multimodal fusion for video description[C]. Proceedings of the IEEE international conference on computer vision, 2017: 4193-4202
7. 瞿明军. 基于视频结构化的目标识别与分析[D]. 山东大学,2020
8. 蒲博建. 基于显著内容的视频结构化技术研究[D].电子科技大学,2021. DOI:10.27005/d.cnki.gdzku.2021.002109.

# 第三章 面向大规模智能监控系统的算力共享网络解决方案

面向大规模智能监控系统的算力共享网络是面向大规模智能监控系统发展的主旋律，基于边缘计算的算力共享网络在未来大规模视频监控预警市场占绝对优势。随着“智能管道 智慧管网”的迅速推进，将安防监控与运营商的网络、边缘计算、云技术相结合，为视屏监控云网融合、云边协同带来可能。

智能监控系统作为实现“智能管道 智慧管网”系统的重要组成部分受到了社会各界的重视。为提高监控系统效率，降低风险漏报率，面向大规模智能监控系统的算力共享网络已经成为迫切需求。本文聚焦于大规模智能监控系统所面临的设备出现异常导致的计算能力不足、计算核心故障导致的风险漏报等问题。

针对设备出现异常（电力不足、过热降频以及计算任务激增）导致的计算能力不足，提出一种算力感知网络的架构，围绕着周围算力盈余的节点分配计算资源，克服传统部署方式在设备出现异常导致的计算核心无法发挥全部性能而产生的计算能力不足的问题。通过服务器向周围区域计算性能正常的节点发送区域内因设备异常导致无法正常工作的节点的视频数据，实现区域计算节点任务的统一调度和管理，实现基于区域和视频源的视频计算资源部署和分配模式。受影响的计算节点只需将摄像机记录的信息片段发送到服务器，有效解决了因设备计算能力不足导致的风险漏报问题。

针对边缘节点计算能力不足问题，提出采用算力感知网络（CAN，Computing-aware Networking）。通过考虑网络和算力两个维度地性能指标，将分布的边缘计算资源进行动态融合，并结合网络路径的度量，将资源与网络进行统一协同调度，使能力不足的边缘节点能够按需实时地调用分布计算资源。

针对计算核心故障（死锁、过热死机以及硬件故障）导致的风险漏报问题，提出两种重启电源恢复设备的策略。

（1）由于计算核心出现的死锁、过热死机等故障具有可自恢复特点，提出一种使用GSM手动控制的断电开关器实现设备断电重启的方法。当设备计算核心故障无法响应时，GSM手动控制的断电开关器会自动通知后勤维修人员，并能执行后勤维修人员的远程操作，即时对故障设备进行恢复，减少后勤维修成本。

（2）除了使用GSM手动控制的断电开关器实现设备断电重启的方法，我们还提出了给电源开关加装蓝牙模块，由蓝牙模块来控制断电开关器的方法。该策略同样也能实现第一种方案中所实现的对故障设备进行恢复，减少后勤维修成本需求。

以上两种方案都能在保证整个监控系统稳定的同时最大减少风险漏报的问题。本文基于以上提出的计算架构与方法，探讨面向大规模智能监控系统的算力共享网络的可行性、有效性和实用性。

## 3.1 研究背景及意义

长输油气管道安防主要依靠人力巡检巡查，西气东输每年投入的人力巡护成本较高，为了响应国家油气体制改革要求，西气东输公司率先开展了“智能管道 智慧管网”的探索与建设，并在管道智能监控感知威胁态势方面取得了实质性的突破。视频监控技术作为实现“智能管道 智慧管网”的重要技术手段也越来越受到社会各界的重视，同时在各个行业得到了广泛的应用。2010年，为了顺利、安全地举办世博会，上海市大规模增加监控网络的范围和密度，仅在浦东新区就建设了包含一万多个监控点位的高清监控系统；截至2011年，黑龙江投入8亿多建设资金用于监控系统建设，共建设近千个监控中心，二十多万个监控点；武汉市为推进“平安城市”建设，建设了覆盖跨江大桥、主干道和主城区城道口等重点部位的监控系统。

长输油气管道安防主要依靠人力巡检巡查，西气东输每年投入的人力巡护成本较高，为了响应国家油气体制改革要求，西气东输公司率先开展了“智能管道 智慧管网”的探索与建设，并在管道智能监控感知威胁态势方面取得了实质性的突破。视频监控技术作为实现“智能管道 智慧管网”的重要技术手段也越来越受到社会各界的重视，同时在各个行业得到了广泛的应用。2010年，为了顺利、安全地举办世博会，上海市大规模增加监控网络的范围和密度，仅在浦东新区就建设了包含一万多个监控点位的高清监控系统；截至2011年，黑龙江投入8亿多建设资金用于监控系统建设，共建设近千个监控中心，二十多万个监控点；武汉市为推进“平安城市”建设，建设了覆盖跨江大桥、主干道和主城区城道口等重点部位的监控系统。

视频监控系统的发展主要分为以下三个阶段：

第一阶段：以模拟摄像机及模拟录像机为主的模拟视频监控系统，这种监控系统多是以摄像机、录像机、分割器为核心，通过模拟信号输出视频，传输距离短，图像质量差，目前已逐步退出市场。

第二阶段：基于模拟摄像机和数字录像机的视频监控系统，采用模拟信号传输视频，数字方式处理与贮存，属于过渡阶段。

第三阶段：基于网络摄像机的全数字视频监控系统，以数字方式传输视频，传输距离长，图像质量好，目前正得到广泛应用。

以上三代视频监控系统，主要解决视频捕捉、传输、存储等问题，并不具备视频职能分析的功能，真正起到安全防范的作用则需要人为观察视频从而做出安全评估。一种方法是实时观看，建设一个监控中心，将监控点所有的视频流导入监控中心，同时配备多块显示屏来播放视频画面，由工作人员通过观看屏幕上的监控点的视频，事后进行调阅，本质上还是以人观看的方式来查找感兴趣的信息。

然而人工监控具有本身固有的缺陷，例如：消耗大量人力、易使人产生疲劳、可扩展差等等。尤其随着“智能管道 智慧管网”监控系统规模不断扩大，面对数以万计的监控点和海量视频数据，以人工为核心的监控系统越来越无法满足需求。如何解决人工监测带来的种种问题，加强监控系统的事前预警和事后的检索，成为当前摆在“智能管道 智慧管网”发展前的重要问题。

监控界提出了智能视频监控(Intelligent Video Surveillance, IVS)的概念。智能监控即使用计算机强大的计算能力以及图像分析技术对传统的监控视频进行处理、分析和理解，过滤无用信息。目前智能视频监控在计算机视觉领域中备受关注。

相比于传统监控系统，智能监控系统具有以下优势：

（1）全天候的监控能力：使用计算机代替人来监控视频画面，实现对监控画面的不间断分析，彻底改变以往由监控人员对监控画面进行监视和分析的方式。

（2）快速的响应能力：通过设置某些可疑行为的识别规则，在其发生危害之前通知相关人员采取相应措施，为潜在威胁做好准备工作。

（3）扩展视频资源的应用领域：传统视频监控系统只能获取单一的视频数据，而智能监控系统可以提取视频中的各种信息，为其他应用提供支持。

随着“智能管道 智慧管网”监控系统规模的不断扩大，面向大规模监控系统的视频计算对智能监控系统提出了新的挑战：

（1）监控设备出现电力不足、过热降频以及计算任务激增等状况所导致的算力不足问题。视频监控节点通常部署在野外，由太阳能供电，而监控视频流数据量大，计算核心需要在稳定的工作电压下才能高效运算，如果出现电量不足、散热不良导致过热降频或者短时间内计算任务激增等情况，计算核心将无法发挥全部性能以负荷繁重的计算任务，这将导致设备反应缓慢，遇到风险时无法及时汇报。如何解决监控设备出现故障导致算力不足的问题，是智能监控系统需要解决的首要问题；

（2）计算核心故障导致的风险漏报问题。视频监控摄像头每分每秒都在向计算核心输出大量的视频流数据，计算核心任务十分繁重。当计算核心出现故障（如死锁、过热和硬件故障）时，该区域的风险监控将瘫痪，此时若存在风险，势必会出现风险漏报问题，后果不堪设想。如何保证智能监控系统的稳定性，杜绝因计算核心故障而导致的风险漏报也是亟待解决的问题；

本文针对以上两个问题进行研究，有利于实现智能监控，推进“智能管道 智慧管网”建设， 对公共安全、社会稳定等具有重要意义。

## 3.2 国内外研究现状

基于边缘计算的智能城市实时监控视频分析目前引起了从业者和研究人员的广泛关注，大量的视频数据被提交到互联网。根据原始视频内容，如何删除冗余视频帧、正确分割视频序列并减少不必要的计算资源消耗是一项具有挑战性的工作。X. Zhang等人提出了一种检测和定位具有运动场形状的视频异常的方法，S. Ding等人利用改进的 Harris-Laplace 时空兴趣点从大视频中识别有趣的片段，然后从中选择关键帧。时空兴趣点可以准确地附着在检测到的目标周围，兴趣区域构建算法可以快速准确地定位包含目标的候选区域。由于图像分类和目标识别的巨大成功，特别是视频监控、目标计数和目标检测，深度学习受到了广泛关注。相比之下，Y. Fang等人提出了空间和时间特征图的计算，然后检测视频显著性。由于资源限制（能源消耗、计算和内存），在智能终端设备上部署这些具有低延迟要求的大型、强大的视频任务仍然是一个挑战。

因此，考虑将这些计算任务转移到更强大的边缘服务器或云。然而，云计算模型不适用于需要短期执行的边缘服务，因为将任务卸载到云中心增加了网络往返传输延迟，应用服务请求不会很快得到响应，但是充分利用云中强大的计算和内存资源将减少总响应时间。由于边缘节点和边缘服务器离用户近，可以快速响应用户请求，成为首选的帮手。在边缘服务器上运行计算密集型任务时，需要对多个终端设备资源进行有效管理。目标是平衡精度、能耗、延迟和负载平衡的性能参数。 VideoStorm引入了这些权衡，以在满足准确性和延迟目标的前提下为每个请求选择正确的配置。例如，在 Chameleon中，配置在流视频输入期间在线更新。然而，基于任务划分的不同模型分割点会导致不同的计算延迟。因此，有必要选择合理的切割策略，以最大限度地发挥端边甚至云协同的优势。在将任务卸载到边缘服务器时，我们可以在边缘进行数据预处理，减少冗余、带宽、延迟和对云中心的依赖，同时提高视频分析的效率。为了减少带宽消耗，有学者提出了端边云协同架构和模型压缩，以消除不同环境下的数据传输。例如M. Song等人提出只有从边缘设备推断出的数据传输到云端进行再训练，以减少数据传输，建议在不影响精度的情况下去除冗余数据以减少数据传输。Glimpse 将所有 DL 计算任务迁移到最近的边缘服务器，同时使用更新检测来删除应该卸载的相机帧。如果检测没有变化，将在本地进行帧跟踪。这种过滤增强了系统的处理能力并使移动设备上的实时目标检测成为可能。Vigil提出了一种分布式架构，它巧妙地利用边缘和云之间的处理任务来减少视频监控中的带宽消耗。类似地，VideoEdge提出了边缘和云的分层架构来处理摄像机流，以便在多个目标和约束之间实现更好的权衡，公平地分配资源。我们在智能视频监测系统中提出采用CAN，该网络通过对多个边缘节点算力资源的合理分配和调度， 综合考虑实时的网络资源状况和计算资源状况，通过网络灵活匹配、动态调度，将动态调度能力调至最优节点，让网络支持提供动态的服务来保证边缘计算节点正常工作。

## 3.3 需求分析

在大规模智能视频监控系统的背景下，虽然现存网络利用边缘计算技术，缓解了云中心计算延迟高、带宽不足、安全性低的问题，实现了“就地、就近”提供服务的功能，但是当智能边缘计算设备存在算力不稳定情况，依然会对整体网络造成巨大影响。如果出现以上情况，很有可能会造成该终端节点所连接摄像机中视频数据的丢失，那么对于该监控区域来说，必定存在一定风险，因为智能边缘计算设备无法进行实时的检测和报警工作。

除此以外，考虑到一些智能边缘计算设备因为部署在野外，采用太阳能板进行供电，其面临各种由环境所带来的问题。如连续阴雨天气导致无法为太阳能板持续提供能量、夏季长时间阳光暴晒导致计算核心过热降频以及遇上风沙或大雾天气导致计算任务激增等问题，那么边缘计算节点将面临着计算能力不足的问题。在这种情况下，无法对所收集到的视频数据提供正常的处理，往往会出现对异常情况的漏报。针对以上智能边缘计算设备的不稳定性和计算能力不足问题，我们提出以下需求分析：

（1）规避因计算能力不足而导致的风险漏报问题：当边缘计算节点的电压没有达到计算核心的工作电压或者计算核心因为天气或者计算任务激增导致过热降频时，将导致计算能力下降，可以借助算力共享方法，使得当前节点只需执行少量本地任务，将部分视频数据任务传输给距离当前节点最近且具有足够计算能力的边缘设备帮助计算。除此之外，当边缘计算节点的性能不足以支撑其计算任务时，可能会产生风险漏报的情况，那么此时，只需要减轻该边缘计算节点的计算压力，将所有的视频数据信息传输给最近的稳定节点进行计算，以避免漏报风险。

（2）减少后勤工作人员的修理成本：从物理安全方面来说，因为边缘计算节点往往部署在室外环境，缺少一定的保护措施。当物理设备受到由外部环境所带来的不可预知的干扰或破坏时，通常会导致设备死机，故障等问题。通常在这种情况下，后勤工作人员会来到现场进行维修，维修的耗时以及人工的耗费不可避免的增加了维护成本。但由于部分问题（例如：死锁，过热等）可以通过重启电源来解决，那么此时我们就需要一个能够进行远程开关电源的处理方案，用以最快恢复故障节点和最大节省后勤成本。

## 3.4 解决方案

针对计算能力不足问题，我们提出采用算力感知网（CAN）进行解决。为了实现泛在计算和服务的感知、互联和协同调度，CAN从三个角度改变了现存边缘计算能力不足问题：（1）针对当前将网络和计算资源分别管理，CAN能够将网络和计算能力相融合，进行统一控制与管理；（2）针对当前网络设备通过路由表信息选择最佳路径，CAN同时考虑网络设备路由信息及边缘节点的计算资源，旨在选择最佳路径与最佳服务节点；（3）针对当前利用带宽、时延和抖动等指标度量网络性能，CAN还会考虑CPU、GPU和ARM处理器等终端设备性能。算力感知网整体框架图如图3.1所示。

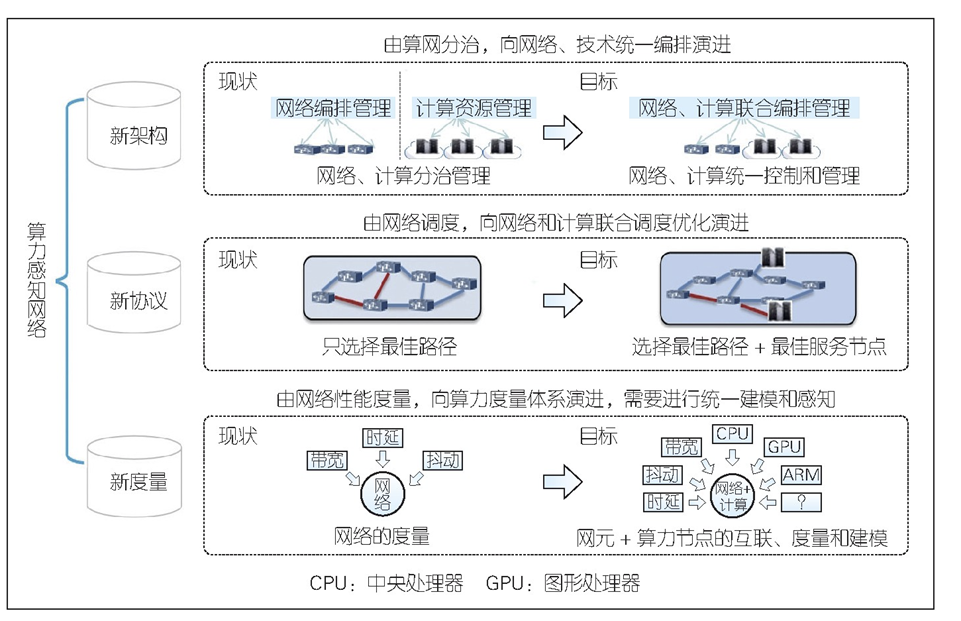


图3.1 算力感知网整体框架图

利用CAN，我们将无处不在的网络连接为基础，通过服务的自动化部署、最优路由和负载均衡，构建算力感知的全新的网络基础设施，最终实现网络的无所不达及算力无处不在。CAN的实施主要分为三个模块：

（1） 算力量化

对于各种异构的计算资源，将采用统一的度量标准进行量化，通过对设备算力进行抽象描述，形成算力度量模板，对算力路由和算力设备管理等提供标准的算力度量规则。首先，对于不同的异构设备进行算力度量和建模，将设备的硬件资源，如：CPU或GPU等物理资源的算力进行统一描述，当边缘计算节点进行自身计算或协调调度时，能够提供有力帮助。其次，考虑到边缘计算节点需要使用时序建模的方法，对摄像机所记录的视频数据进行处理，因此需要量化该任务所消耗的算力大小，以便更好为边缘节点的自身应用提供服务。最后，由于部分边缘计算节点会因为计算能力不足的原因，无法满足自身节点的应用所需服务，因此在发出算力请求的过程中，需要将其算力需求进行量化，映射为实际所需的算力资源，从而可以充分感知边缘设备的需求，提高协同调度效率。

（2）算力路由

通过算力路由，实现对“网络+计算”的联合考虑，对算力不足的边缘计算节点提供拥有足够算力并且距离最近的节点提供服务。算力路由分为算力路由控制和算力路由转发两种技术，通过将网络、计算和存储等多维度资源进行整合，及时对边缘计算节点的服务进行感知与通告，帮助实现按需调度。算力路由控制会根据边缘计算节点的算力信息，生成具有算力感知的新型路由表，以此表示边缘计算节点的剩余计算能力；另外，还能根据边缘计算节点的请求，动态生成按需的算力调度策略，实现对算力网络的协同调度，帮助完成节点任务。算力路由转发利用网络路由协议对全网环境进行感知，将网络路径和算力相结合，选择最优节点提供服务。

（3）算力管理

边缘计算节点需要进行统一管理，以实现对网络和算力的监测和控制。因为边缘计算节点是灵活地存在于网络之中，会动态地增加、改变和减少，因此通过算力管理可以对节点进行注册、更新和注销；算力度量实现了对不同类型资源地统一抽象描述，而算力管理在此基础上形成算力能力模板，对算力设备的操作维护管理等提供标准的算力度量规则；算力管理通过对算力性能的实时监测，充分了解网络和算力的整体情况，以便对相应的路由通告策略进行管理等。

针对计算设备的不稳定性，我们提出一种使用GSM手动控制的断电开关器实现设备断电重启的方法。当节点出现死机等故障时，远程控制系统中该节点的颜色会标记为灰色，那么我们可以利用后端程序告知后勤维修人员，然后该人员采用远程控制方式进行设备开关机，而不必到达故障节点所在现场。当某边缘计算节点频繁出现故障并且需要进行重启时，我们需要将每次重启的设备节点以及故障时间进行记录。如果在一个月的时间内，被记录的次数过多，那么说明该节点可能受到硬件性能或其他因素的影响，此时需要管理人员到边缘计算节点所在现场进行调研，从根源上解决该节点的问题，以减少重启的次数，帮助保护设备硬件并且让其更加稳定地进行计算。

## 3.5 本章小结

本章针对计算能力不足问题，我们提出采用算力感知网（CAN）进行解决。利用CAN，我们将无处不在的网络连接为基础，通过服务的自动化部署、最优路由和负载均衡，构建算力感知的全新的网络基础设施，最终实现网络的无所不达及算力无处不在。针对计算设备的不稳定性，我们提出一种使用GSM手动控制的断电开关器实现设备断电重启的方法。

## 3.6 参考文献

1. 雷波,陈运清.边缘计算与算力网络——5G+AI 时代的新型算力平台与网络连接[M].北京:电子工业出版社,2020
2. 中国移动.算力感知网络技术白皮书[R].2019

# 第四章 总结

长输油气管道安防主要依靠人力巡检巡查，西气东输每年投入的人力巡护成本较高，为了响应国家油气体制改革要求，西气东输公司率先开展了“智能管道 智慧管网”的探索与建设，并在管道智能监控感知威胁态势方面取得了实质性的突破，本咨询项目主要聚焦以下三方面进行调研研究：

1、基于时序建模的视频动作检测与预测方法

调研基于视频流的动作检测和演化技术，并讨论针对识别管道智能监控视频中捕捉到的潜在威胁场景（如管道附近施工工地）进行动态时序建模来识别并预测其演化方向的可能性。

2、基于视频数据的场景结构化建模方法

将视频图像资源中用户关注的活动目标数据进行结构化解析,形成有价值的信息积累,可以提供特征检索、以物搜人、轨迹回放等应用功能，本主题首先调研目标视频数据结构化的主流实现方法，并针对管道智能监控感知场景进行建模，识别主要目标的结构化信息进行讨论。

3、面向大规模视频数据的智能边缘计算技术

在管道智能监控感知场景下，集中式云计算无法提供低时延、高效率的视频监控服务。基于此,考察分布式边缘计算模型在大规模视频流数据处理中的研究现状，并针对管道智能监控感知环境下，采用智能边缘计算技术，充分发挥终端设备算力，实现大规模视频数据处理的可能性。