

# 基于元学习的 少样本水闸图像识别方法研究

薛凌峰<sup>1</sup>, 宋 炜<sup>1</sup>, 鲍建腾<sup>1</sup>, 焦 野<sup>1</sup>, 戚荣志<sup>2</sup>

(1. 江苏省水旱灾害防御调度指挥中心, 江苏 南京 210029; 2. 河海大学 计算机与软件学院, 江苏 南京 211100)

**摘要:** 针对实际工程环境中采集的水闸图像样本不均衡、前后景混融导致的识别效果不好的问题, 提出一种基于元学习的少样本水闸图像识别方法。首先构建水闸图像数据集, 并使用图像增强和预处理对数据集进行优化; 再使用多头注意力, 提升网络准确捕捉多种与任务相关的关键特征信息的能力, 更好地与时序卷积协作, 进一步提高水闸图像的识别效果。在构建的 sluice-ImageNet 数据集上进行实验, 实验结果表明, 相比其他方法, 所提方法在水闸启闭状态图像识别任务上更有效性和优越性。该方法部署于重点水利工程视频监测平台, 辅助人工监管, 可实现对水闸异常运行情况的实时监测, 为防汛决策提供智能化支持。

**关键词:** 图像识别; 少样本; 元学习; 多头注意力

中图分类号: TV211

文献标识码: B

文章编号: 1007-7839(2024)03-0020-0005

## Research of few-sample sluice image recognition method based on meta-learning

XUE Lingfeng<sup>1</sup>, SONG Wei<sup>1</sup>, BAO Jianteng<sup>1</sup>, JIAO Ye<sup>1</sup>, QI Rongzhi<sup>2</sup>

(1. Flood and Drought Disaster Prevention and Control Center of Jiangsu Province, Nanjing 210029, China;

2. College of Computer Science and Software Engineering, Hohai University, Nanjing 2111000, China)

**Abstract:** Aiming at the problem of poor recognition effect caused by uneven samples of sluice gate images collected in the actual engineering environment and the mixing of front and rear views, a few-sample sluice image recognition method based on meta-learning was proposed. First, the sluice image dataset was constructed, and the dataset was optimized using image enhancement and preprocessing. Then, multi-head attention was used to improve the ability of the network to accurately capture a variety of key feature information related to the task, and better cooperate with the temporal convolution to further improve the recognition effect of the sluice image. Experiments were carried out on the constructed sluice-ImageNet dataset. The experimental results show that the proposed method is more effective and superior than other methods in the sluice opening and closing recognition task. The method was deployed on the video monitoring platform of key water conservancy projects to assist manual supervision and realize real-time monitoring of abnormal operation of sluices, and provide intelligent support for flood control decision-making.

**Key words:** image recognition; few-sample; meta-learning; multi-head attention

收稿日期: 2024-01-15

基金项目: 江苏省水利科技项目(2018057)

作者简介: 薛凌峰(1985—), 男, 工程师, 主要从事防汛抗旱工作。E-mail: 84713687@qq.com

通信作者: 戚荣志(1980—), 男, 副教授, 博士, 主要从事水利信息化、智能软件工程等工作。E-mail: rzqi@hhu.edu.cn

随着智慧水利工作的推进,国内各级水利防汛视频监控系统相继建成<sup>[1-3]</sup>,可以更及时、便捷地获得水利影像数据,有助于更好开展水利监管工作。其中,水闸启闭状态监控是防汛调度工作中的一项重要内容,应用水闸场景图像自动识别技术辅助或代替人工监管,可以提高异常情况监测的实时性<sup>[4]</sup>。但不同类型水闸的启闭状态判断依据相似而不尽相同,而且一些低调度频率的水闸的各种不同运行状态的影像资料难以收集,在历史上也没有足够的水闸影像资料积累,于是许多水闸的开启或关闭状态的样本数量很不均衡,部分状态样本很少。针对该问题,本研究将不同类型的水闸图像识别工作视为多个具有相似性的子任务,基于元学习理论<sup>[5]</sup>捕捉各类型水闸启闭状态识别任务的共性,实现跨水闸类型的闸门启闭状态图像识别。

## 1 水闸图像数据集构建

水闸图像识别领域内缺乏权威可靠的开源数据集,因此本研究需自行构建数据集。所有采集到的原始数据需要进行一系列的预处理才能成为可以进行学习的样本。

### 1.1 数据来源

本文水闸启闭状态图像数据主要采集自江苏省防汛视频监控系统,从系统中导出水闸闸口监控视频片段作为初始影像数据,再从视频中提取各种状态的水闸场景图像。本文共收集了石梁河水库南北泄洪闸、皂河闸、海口闸、黄沙港闸等16个闸门的影像数据,由于各个闸门在每个年周期内的调度频率不同,因此尽量收集了各水闸的不同状态数据。

### 1.2 图像数据预处理

省级防汛视频监控系统中不同水闸监控点的视频质量各异,需首先借助视频格式转换工具将AVI、MOV等格式统一转换为MP4格式,再人工筛选记录闸门场景时间段以及对应标签,并处理视频损坏、跳帧等。在此基础上,利用OpenCV计算机视觉模块提供的CV2.VideoCapture函数从视频流中截取图像。为了避免异常干扰,确保正确标记闸门启闭状态,对每一连续场景每隔25帧(约2s)截取一张图像,再人工筛选不合格图片。

视频在生成和传输过程中容易受到各种噪声的干扰,导致图像质量下降,对后续图像的处理和图像视觉效应将产生不利影响。非局部均值算法<sup>[4]</sup>充分利用图像中的冗余信息,在去除噪声的同时最大程度地保持图像的细节特征,其基本思想是将图

像中与当前像素具有相似领域结构的像素加权平均作为当前像素的估计值,其权值由像素欧拉距离确定。对于每一个像素点的权值,采用以该像素点为中心的图像子块与当前像素点为中心的子块之间的高斯加权欧氏距离来计算。一般认为,图像中的其他像素与当前像素的欧拉距离越小,则两者越相似,因此为此像素设置较高的权重。

视频监控设备通常安置在拍摄水闸工程整体的位置,因此每张图像上有多个闸口。为便于做标记和后续算法分析,以单个闸口图像区域作为感兴趣区域(region of interest, ROI),将各ROI截取为单独图片。具体操作为:利用OpenCV2模块的鼠标事件函数实现交互,得到并记录感兴趣矩形区域的左上、右下顶点坐标值,将其作为Pillow图型库Image.crop函数的输入值,从而完成图片裁剪,得到ROI区域。预处理后效果示例如图1所示。

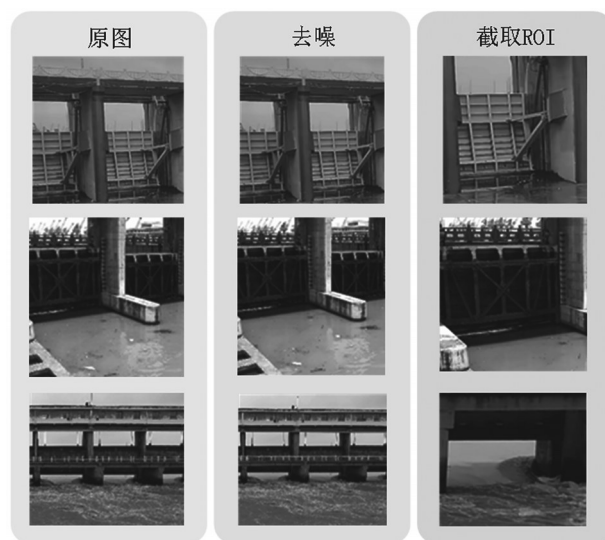


图1 图像预处理效果

### 1.3 图像增强

通过图像数据预处理后得到的样本存在类别不均衡的问题,部分水闸图像数据稀缺、启闭状态图像数量差距大,不利于模型的训练。数据增强使用平移变换、缩放和剪裁等图像处理方法对原始数据进行变换,生成相同对象不同形态的更多数据,用于补充欠缺的类型样本,促使模型学习到具有分辨性的特征,增强方法的鲁棒性。其中,水平翻转方法操作简单,并且在水闸启闭状态识别场景中可以保证标签安全。在实际环境中,水利防汛监控摄像设备的安置位置受限,在同一水闸站点往往仅从某一单侧视角进行拍摄记录。采用水平翻转进行增强,尽量

补充了数据集中样本量少的开启或关闭状态下的某些闸门图像,并能在确保色彩、线条等其他特征不丢失的同时丰富了水闸的不同视角特征。

## 2 水闸图像识别方法

水闸数据集中的图像大都没有明确的前后景区分,而且关注区域不是简单的某个对象,而是以

图中的闸口、浪花、水波纹等要素综合作为分类结果判断的依据,这与在标准数据集上开展的图像分类研究存在较大差异。因此,为了适应少样本水闸启闭状态图像识别的应用场景,本文基于SNAIL元学习模型,提出了多头自注意力水闸图像识别方法(Multi-Head Self-Attention SNAIL, MH-SNAIL)。MH-SNAIL网络框架如图2所示。

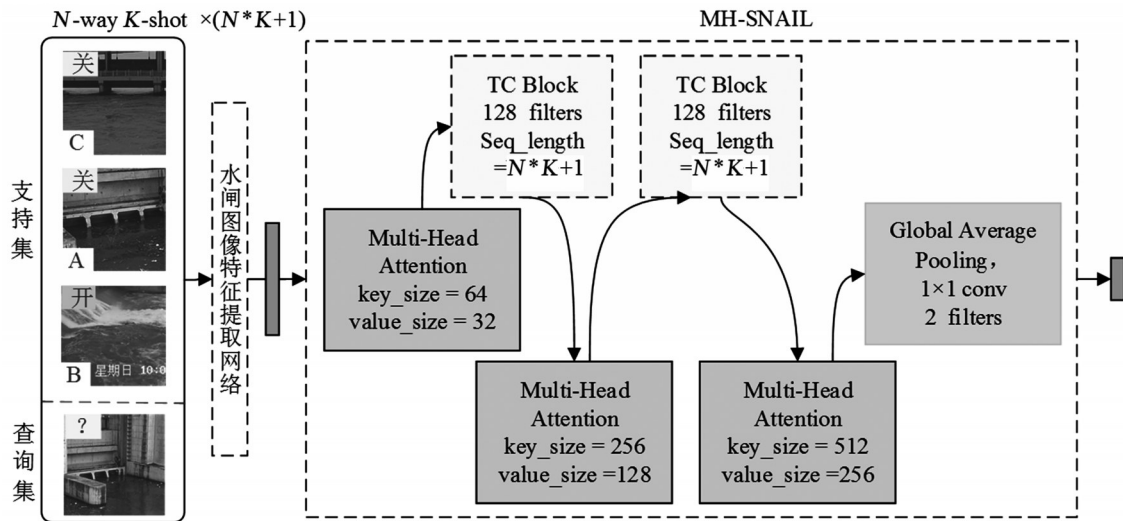


图2 MH-SNAIL网络框架

MH-SNAIL接收水闸图像特征向量组作为输入,通过3个多头注意力层和2个时序卷积层交替,后接一个全局平均池化层计算输出表征预测结果的二维概率向量。其中,3个多头注意力层的键-值嵌入向量的维数依次为64-32、256-128、512-256,2个时序卷积层的通道数均为128,每个通道包含 $N \times K + 1$ 个稠密块。

### 2.1 多头注意力

多头注意力机制根据查询嵌入与键嵌入的相似性分布计算得到一组关于值嵌入元素的权重,再基于这组权重更新值嵌入中的元素,能够很好地学习长序列信息。多头注意力首先对查询嵌入、键嵌入与值嵌入进行线性变换映射到多个不同子空间中,在子空间中分别执行普通的注意力运算后再合并结果,过程示意如图3所示。

多头注意力首先将查询嵌入 $Q$ 、键嵌入 $K$ 与值嵌入 $V$ 线性映射到多个各异的子空间中(设子空间的个数为 $h$ ,取输入特征向量维数不大于6的最大公因数),再在各个子空间中并行执行 $h$ 次注意力函数,从而得到 $h$ 个不同的注意力结果,即 $h$ 个头(head),最后将这些结果拼接起来并进行线性变换,

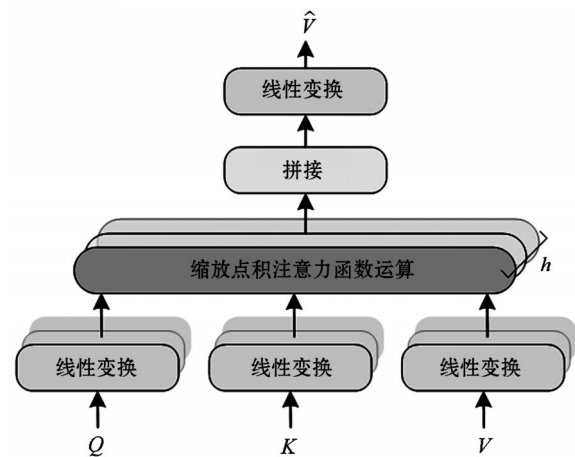


图3 多头注意力示意

生成最后的注意力结果。

多头的注意力运算模式能够从不同的特征表示子空间中学习到更多有价值的信息。显然,在少样本水闸图像分类任务中,相比普通的注意力机制,多头注意力机制更能达到综合关注多种特征的要求,有利于网络更加全面地捕捉水闸图像中的关键信息,也可以更好地弥补时序卷积在长距离依赖中的不足。



## 2.2 全局平均池化

MH-SNAIL网络的输出层使用全连接层,将学习到的特征表示综合起来并映射到样本类别空间中,但全连接层存在参数多、训练速度慢的弊端,并且会将一定的特征存储在这些参数中。本文考虑使用全局平均池化层来代替全连接层,先采用一个GAP将 $N$ 个特征图降维成一个 $1 \times N$ 像素尺寸的特征图,再用 $C$ 个 $1 \times 1$ 卷积核将 $1 \times N$ 的特征图卷成 $1 \times C$ 的向量,其中 $C$ 为类别数量,在本文的水闸启闭状态图像分类问题中取2。通过这种替换方案,整个过程在维度上来看相当于一层全连接层,但GAP能够大大减少特征的维度,有利于提升训练速度。

## 3 实验验证

### 3.1 实验数据及评价指标

实验使用本文第2小节所构建的水闸图像数据集sluice-ImageNet,共包含来自16类水闸的15 000幅 $84 \times 84$ 像素的彩色闸门启闭状态图像样本。为了方便实验,采用水平翻转处理数据集的所有图像,使得样本数量扩增为原数据集的两倍。此时,为了保证每一个元任务中都能学习到多于一种水闸类型的样本,必须设置类别数 $N$ 不小于5。根据数据采样规则,将数据集划分为不同闸门且不同启闭状态的25个子批,每个子批包含1 200幅图像。其中,9类闸门包含开启态和关闭态的两个子批,挑选其中8个闸门的16个子批作为训练集,将剩余1个闸门的2个子批连同其他7个子批随机划分为测试集和验证集,具体数据集结构如表1所示。

表1 数据集结构

数据类	闸门类别	分类子批	样本(幅)
训练集	8	16	19 200
验证集	4	4	4 800
测试集	4	5	6 000

实验使用训练阶段的平均验证准确率(Validate Accuracy)和测试阶段的测试准确率(Test Accuracy)来评判所提方法的效果。在训练阶段,每个epoch训练结束就进行验证评估和统计每次验证的结果,取平均值作为一个epoch的验证准确率。在测试阶段,在测试集上构建 $10^3$ 个随机任务,用训练阶段每次验证准确率最高的权重模型进行测试,所得结果的平均值作为测试准确率。其中,每次任务的验证与测试结果将会取95%的置信区间。

### 3.2 实验环境及配置

实验的所有网络模型均运行于相同的软硬件环境,其详细软硬件环境配置如表2所示。

表2 实验软硬件环境配置

配置名称		配置参数
硬件环境	CPU	Intel(R) Xeon(R) Silver 4116 CPU @ 2.10GHz
	GPU	GeForce RTX 3090 TURBO, 24G 显存
	内存	96G
软件环境	操作系统	Ubuntu 16.04, 64位
	编程环境	Python 3.6, Pytorch 1.8

实验采用了5-way 1shot和5-way 5-shot这两种任务的方式,使用批处理大小(batch size)为16、初始学习率为 $1 \times 10^{-4}$ 的Adam优化算法调整模型的参数,设置训练次数(epoch)为100,每次训练的元任务数为 $10^3$ ,学习率每训练20次迭代减半,权重衰减为 $10^{-6}$ 。训练阶段,每完成一次训练就进行验证评估,并统计每一次验证结果,将其中验证准确率最高的权重模型保存下来,作为测试阶段的模型。

### 3.3 实验结果与分析

为了评估MH-SNAIL在水闸启闭状态图像识别任务上的性能,选择匹配网络、MAML、原型网络和关系网络这4种经典的元学习方法与本文的MH-SNAIL方法进行比较,通过与少样本图像分类领域内的其他常用方法比较,进一步验证本文的MH-SNAIL方法的有效性。

表3中展示了各方法在sluice-ImageNet数据集上以5-way 1-shot和5-way 5-shot的训练方式进行训练后得到的测试准确率。从表3可以看出,MH-SNAIL方法在sluice-ImageNet数据集上的表现明显优于这些方法。在1-shot场景下,MH-SNAIL相较于匹配网络、MAML、原型网络和关系网络分别具有11.78%, 8.01%, 7.46%和5.18%的分类性能提升;在5-shot场景下,则分别具有11.94%, 5.57%, 4.48%和

表3 在sluice-ImageNet上与其他方法的分类准确率对比

方法	5-way 1-shot / %	5-way 5-shot / %
匹配网络	45.07	58.40
MAML	48.84	64.77
原型网络	49.39	65.86
关系网络	51.67	63.92
MH-SNAIL	56.85	70.34

6.42%的分类性能提升。这有力地验证了本文提出的MH-SNAIL方法的有效性和优越性。另外,在5-way 5-shot情景下,MH-SNAIL方法的分类准确率为70.34%,初步达到了跨水闸类型闸门启闭状态图像识别应用的要求。

本文采用Python编程语言,基于TensorFlow深度学习框架实现了MH-SNAIL模型。模型将应用在重点水利工程视频监控平台中。模型部署时,首先将训练好的模型导出为ONNX格式,然后在GPU服务器上加载ONNX模型,以加速模型推理的速度。监测平台通过模型提供的Restful接口进行交互。模型接收用户通过平台Web界面上传的JPG或PNG格式的水闸图像,或者对接平台获取的实时视频,收到图像信息后在服务器上进行推理,并将水闸实时状态识别结果以JSON格式的报文发送给用户平台。

## 4 结 语

提出了基于元学习的少样本水闸图像识别方法,采用多头注意力,提升网络准确捕捉多种与任务相关的关键特征信息的能力,更好地与时序卷积协作实现水闸启闭状态图像识别。在 sluice-

ImageNet数据集上的实验验证了MH-SNAIL方法的有效性和优越性,初步达到了跨水闸类型闸门启闭状态图像识别应用的要求,并且,通过微服务架构实现水闸图像识别方法的服务化封装,部署于重点水利工程视频监控平台中,具有高度灵活性和拓展性。该方法可作为防汛视频监控系统的延伸,辅助或代替人工监管,提高异常情况监测的实时性,实现更智能化的水利工程管理的运行。

### 参考文献:

- [1] 牟舵,刘斌.大藤峡水利枢纽施工过程智能视频监控与应用[J].水利信息化,2021(2):67-69,90.
- [2] 杨凯,郭振霆,杨甜,等.视频监控在偏远水利工程管理的应用[J].海河水利,2021(1):118-120,126.
- [3] 陈述,纪勤,陈云,等.基于知识图谱的智慧水利研究进展[J].河海大学学报(自然科学版),2023,51(3):143-153.
- [4] 余再康,程井.贵州省水库运行及安全监测监管系统关键技术[J].河海大学学报(自然科学版),2023,51(4):46-54.
- [5] VILALTA R, DRISSI Y. A Perspective View and Survey of Meta-learning[J]. Artificial Intelligence Review, 2002, 18(2):77-95.

(上接第19页)

传输。建立洮湍湖治理分析系统,强化对洮湍地区河湖水量、水质、水生态、出入湖污染物通量的分析评价,研究区域水量交换、污染物输移规律。定期系统评估洮湍地区河湖的水安全、水生态、水资源、水环境和岸线等水域状况,为河湖监管提供支撑。

### 2.4.4 强化数字赋能治理

加快数字孪生流域和数字孪生工程建设,依托监测体系完善,强化信息基础设施建设,为模型平台提供精准物理参数和现实约束条件。基于全省水利一张图,汇聚基础数据、监测数据、业务管理数据、跨行业共享数据和地理空间数据,搭建数字孪生场景<sup>[5]</sup>。优化现有水量水质模型,开发突发性水污染、蓝藻AI识别和预测等模型,满足不同场景预演模拟需要。根据运行管护及应用场景需求强化业务平台系统开发,建立水工程调度、出入湖水水质监测、河湖健康评价、水域岸线监管、蓝藻水华预警等业务平台系统,提升全要素预报、预警、预演、预案能力。

## 3 结 语

研究主要以“两核两区一带”(洮、湍两湖生态

绿核、西南水源涵养区、腹部生态水网区、入太湖生态缓冲带)建设为抓手,以陆域控源截污为前提,以加强水系连通为基础,以着力河湖生态修复为关键,以强化河湖综合管理为支撑,系统推进洮湍地区河湖综合治理,全面提升河湖生态环境质量,实现全域幸福河湖,促进太湖水生态环境改善,满足人民群众对优美生态环境的需要,为区域经济社会高质量发展提供有力支撑。

### 参考文献:

- [1] 胡惠良,谈俊益.江苏太湖流域水环境综合治理回顾与思考[J].中国工程咨询,2019(3):92-96.
- [2] 杨春瑞,查亮.淮河流域高塘湖健康评价分析研究[J].工程与建设,2023,37(1):35-39.
- [3] 张卫.安徽省城市水生态治理模式探讨[J].水利发展研究,2020,20(6):32-35,43.
- [4] 谢三桃,朱慧雯,唐晓先,等.巢湖流域山水林田湖草生态保护和修复[J].水资源保护,2023,39(2):252-258.
- [5] 刘国庆,范子武,廖铁鹏,等.江苏数字孪生水网建设与预报调度一体化应用初探[J].中国水利,2023(3):60-65.