

基于深度学习的水位线检测算法

廖 赞¹, 段 清^{1*}, 刘俊晖¹, 周 豪²

(1. 云南大学 软件学院, 昆明 650091; 2. 云南省览易网络科技有限公司, 昆明 650091)

(* 通信作者电子邮箱 676295641@qq.com)

摘 要: 目前对河道、水库等开放水域的水位进行监控, 通常需要在监控地点部署用于测量液位的特殊部件, 如水位标尺、压力传感器等设备, 无法仅使用摄像头完成水位的监控。为解决以上问题, 提出一种基于卷积神经网络(CNN)的水位检测算法, 该算法输入为一张包含水位线的静态图像, 输出为水位线与检测图像左侧边界交点的预测坐标及水位线与水平方向的夹角预测值, 最后根据网络输出的交点坐标及夹角绘制预测水位线。测试结果表明, 该方法对检测环境的适应能力极强, 即便在夜间下雨, 且只由红外光源照明的情况下也能对水位线进行有效预测。使用该方法可以对河道湖泊等开放水域实现全天候非接触式水位连续监控。

关键词: 深度学习; 卷积神经网络; 残差网络; 水位线检测; 水位线识别

中图分类号: TP391.41 **文献标志码:** A

Water line detection algorithm based on deep learning

LIAO Yun¹, DUAN Qing^{1*}, LIU Junhui¹, ZHOU Hao²

(1. School of Software, Yunnan University, Kunming Yunnan 650091, China;

2. Yunnan Province Lanyi Network Technology Company Limited, Kunming Yunnan 650091, China)

Abstract: Water level monitoring in rivers, reservoirs and other open waters, usually requires special components for water level measurement at the monitoring site, such as water level measuring sticks, pressure sensors and other equipment. It is not possible to monitor the water level by only using a camera. To solve the above problems, a water line detection algorithm based on Convolutional Neural Network (CNN) was proposed. The input of this algorithm was a static image containing water line. The output were the predicted coordinates of the intersection of the water line and the boundary on the left side of the detected image, and the predicted angle between the water line and the horizontal direction. Finally, the predicted water line was drawn according to the intersection coordinates and included angle of the CNN network output. The test results show that the proposed algorithm has a strong adaptability to the detection environment. Even if it rains at night and is only illuminated by infrared light source, the water line can be predicted effectively by this algorithm. The algorithm can realize all-weather non-contact continuous monitoring of water line in open water areas such as rivers and lakes.

Key words: deep learning; Convolutional Neural Network (CNN); Residual Network (ResNet); water line detection; water line identification

0 引言

对河道、水库等水域的水位实施全天候的监控对于环境治理、水资源管理、防洪防灾、农业生产等领域来说是十分重要的。根据测量特征的不同, 常见的水位检测可分为5种类型: 基于浮子的方式、基于压力的方式、基于超声波的方式、基于雷达的方式和基于图像的方式。

目前, 基于压力的水位检测方法及基于超声波的水位检测方法应用最多。基于压力的方法将压力传感器放置于待测液位底部, 当液位高度变化时, 底部液体压力会发生变化, 通过压力及液位深度之间的关系可以将压力直接转换为水位值。基于压力的检测方法虽然易于使用, 但是在开放水域环境下, 由于水下环境不确定性, 为了保证测量的准确, 需要经

常对传感器进行校准; 同时, 由于传感器长期承受水压, 容易造成损害。基于超声波的水位测量方法利用超声波在同种介质中传播速度相对恒定以及碰到障碍物能反射的原理制成, 因为不需要将传感器置于水中, 避免了基于压力的水位检测方法的相关问题; 然而, 由于声速易受温度变化影响, 因此, 在大雨或大雪期间, 环境温度变化较大, 其测量精度无法保证。

基于图像的水位检测方法是近年来水位检测研究的热点。早期的基于图像的水位检测方法^[1-5]主要使用图像的空间信息, 即基于图像边缘信息进行水位检测。这类方法的主要思想是, 首先对拍摄图像进行校正, 然后通过边缘检测算法检测特定区域的边缘信息, 最后在已检测出的众多边缘中寻找可能的水位线。这类方法比较容易受到水岸特征的影响, 因此, 这类方法的应用范围比较有限。目前较新的基于图像

收稿日期: 2019-08-06; 修回日期: 2019-10-10; 录用日期: 2019-10-21。

基金项目: 国家自然科学基金资助项目(61762089); 云南省软件工程重点实验室开放基金资助项目(2017SE101)。

作者简介: 廖赞(1980—), 男, 云南昆明人, 讲师, 博士, 主要研究方向: 计算机视觉、机器学习; 段清(1975—), 女, 云南昆明人, 副研究员, 博士, CCF会员, 主要研究方向: 数据挖掘、智能信息处理、软件工程; 刘俊晖(1980—), 男, 云南昆明人, 讲师, 博士, 主要研究方向: 数据挖掘、机器学习、软件工程; 周豪(1995—), 男, 云南昆明人, 硕士, 主要研究方向: 计算机视觉、机器学习。

的水位检测方法^[6-9]则同时使用图像的空间信息以及序列图像间的时间运动信息来克服由于错误的边缘检测导致的水位误检问题。这类方法的基本思想是:首先通过滤波器去除镜头上的灰尘、雨或雪产生的图像噪声;然后通过边缘检测提取图像的空间信息,并通过光流提取图像的时间信息;在估计了时空特征之后,将每个时空特征投影到 y 轴上以创建累积直方图;最后基于提取到的图像空间特征及时间特征预测水位线位置。这类方法比较依赖光流提取图像的时间信息,如果水域比较平静,则时间信息特征将失效;此外,这类算法需要保证水位线为水平方向,因此,对摄像机拍摄角度有一定要求,如果摄像机无法水平拍摄水位线,就需要对水岸进行标注,后期通过图像校正进行处理,所以应用范围也比较有限。

深度学习^[10-11]是近年来快速发展的机器学习技术,它模拟人脑的学习过程^[12],从输入图像的低级别信息中逐步抽象、提取高级特征^[13]。本文提出了一种基于深度学习的单帧图像水位线检测方法,设计了一种基于 ResNet50 的卷积神经网络(Convolutional Neural Network, CNN)架构,并在此架构基础上给出了对网络进行有效训练的损失函数。实验结果表明,即使在复杂的环境中本文算法也可以有效地提取相关图像中的水位线。

本文的主要工作是:

- 1) 针对普通摄像头抓拍的单帧图像,提出了一种基于深度学习的水位线检测方法;
- 2) 基于 ResNet50,设计了一种 CNN 架构,可以从输入数据中提取水位线特征;
- 3) 给出了基于斜率及截距的损失函数,基于该损失函数可对上述 CNN 进行有效训练;
- 4) 给出了创建水位检测训练数据集的具体方法及相关工具,并基于该工具创建了水位检测数据集及验证集。

1 数据来源及数据处理

1.1 数据来源

深度学习属于监督学习方式,需要大量的数据对算法进行训练。从昆明 5 条主要河流的共 20 个监控摄像头,通过变换摄像头的视角及缩放参数的方式,总共采集包含水位线的图片 3 955 张,如表 1 所示。其中 3 486 张照片用于训练,469 张照片用于测试,相关图片示例如图 1 所示。

表 1 水位线图像数量统计

天气	夜间红外	白天自然光	夜间自然光
	(00:00—6:30)	(7:00—19:00)	(20:00—23:30)
晴朗	518	542	445
多云	0	225	0
小到中雨	359	446	424
大雨	280	370	346

1.2 数据增强

由于实际获得的图像总量十分有限,不足以支撑深度学习的训练,因此在本项目中对训练数据进行了增强,图像增强包括以下方式:左右翻转、切变、旋转、HSV 颜色变化、高斯模糊、随机噪声,如图 2 所示。

1.3 数据标注

水位线在自然环境下是非常多变的,可能是任意折线。为了使问题相对简单,本文假设带检测水位区域的水位线均

为近似直线。在此假设基础上,本文提出的基于深度学习的水位检测方法实际上是一种直线段的回归方法,即从输入的图像中找到最接近真实水位线的一条直线。基于此考虑,对数据的标注也是基于直线段进行的。该标注方法非常简单,即在给定图像上近似直线的水位线区域的两边确定两个端点,端点间的连线即为标注水位线。如图 3 所示。其中:在图像标注过程中,如果整条水岸比较平直,如图(a)所示,可以使用一条直线段对整个水岸线进行整体标注;如果水岸出现折线,如图(b)所示,可以使用两条分段折线进行标注;如果水岸不完整,如图(c)、(d)所示,可以使用两条分段直线段进行部分标注。

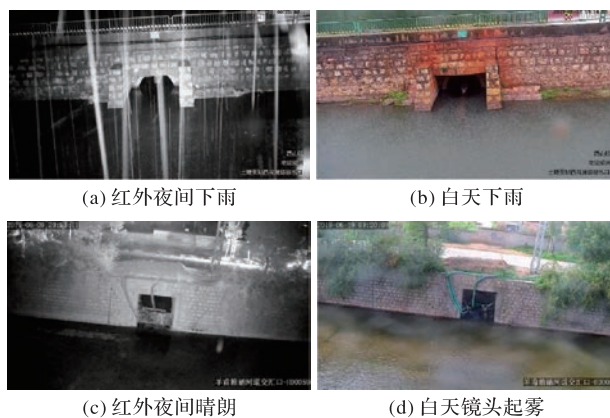


图 1 水位线图像示例

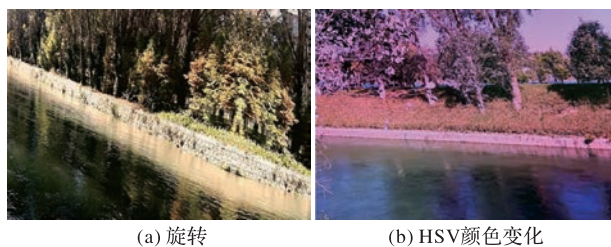


图 2 部分图像增强结果示例

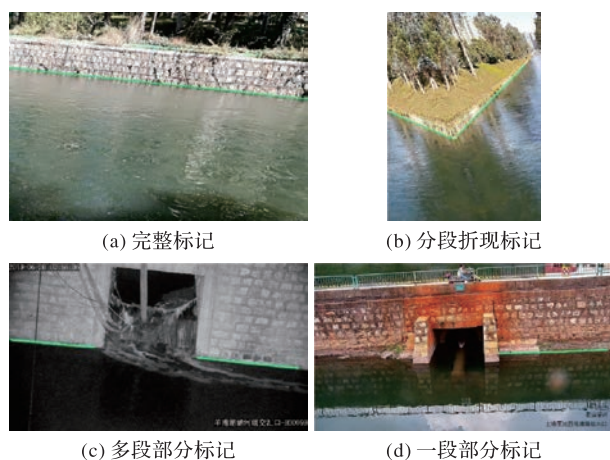


图 3 水线标注结果示例

图像标注完成后的标签数据存储格式如下所示。其中,lines 表示当前图像中的所有已标注水位线段;line 表示其中的一条线段; x_bgn 表示当前线段起始点的 x 坐标轴; y_bgn 表示当前线段起始点的 y 坐标轴; x_end 表示当前线段结束点的 x 坐标轴; y_end 表示当前线段结束点的 y 坐标轴; $x1\sim x5$ 、 $y1\sim y5$ 表示实际坐标值。


```

{“lines”: [
  {“line”: {“x_bgn”:x1,“y_bgn”:y1,“x_end”:x2,“y_end”:y2}},
  {“line”: {“x_bgn”:x3,“y_bgn”:y3,“x_end”:x4,“y_end”:y4}},
  {“line”: {“x_bgn”:x5,“y_bgn”:y5,“x_end”:x6,“y_end”:y6}}
]}

```

2 水位预测卷积神经网络架构

2.1 标准 CNN 架构

CNN 是一种模仿生物视觉识别过程的人工神经网络^[14],主要用于图像分类、物体检测等与计算机视觉相关的任务。与浅层神经网络算法比较,它具有权重共享、分层特征提取、端到端训练、易于扩展等优点^[15]。CNN 已经成为计算机视觉任务的研究热点。CNN 通常包括输入层、卷积层、池化层、全连接层及输出层。输入层接收输入图像,图像在输入 CNN 之前通常需要对图像进行预处理(如归一化),并将图像缩放或填充成为输入层的大小。卷积层通过卷积核及卷积运算提取图像的特征,低层次的卷积层提取图像较低级别的特征,如边、线、角等信息;较高层次的卷积核提取图像较高级别的抽象信息,如眼、嘴等高层次抽象信息。卷积核是一组包含数值的矩阵,称为权重,CNN 通过训练数据调节权重参数,并使用经过训练的权重提取图像的特征。卷积运算后通常需要跟随一个激活函数,激活函数的输出结果称为特征图,它是下一个卷积层的输入数据。卷积层的运算过程可以用如下公式表达:

$$G_i = f(G_{i-1} \otimes W_i + b_i) \quad (1)$$

其中: G_{i-1} 表示第*i-1*层输出的特征图; W_i 表示第*i*层卷积核权重; b_i 表示第*i*层的偏置项; f 表示激活函数; \otimes 表示卷积运算; G_i 表示第*i*层卷积运算后的输出特征图。池化层对输出特征图进行下采样,其目的是压缩特征,降低网络的计算量,然而在压缩特征的同时也损失了部分特征。常见的池化操作有最大池化及平均池化两种。卷积层及池化层的作用是对输入图像上的特征进行提取,提取到的特征最终需要通过一个传统的神经网络进行分类或是回归运算,该任务由全连接层完成。

2.2 水位预测 CNN 架构的设计

CNN 的权重共享机制极大地减少了整个网络的参数量,网络过拟合的风险降低,因此,网络的深度大大增加。VGG 网络^[16]同时也证明,随着网络层次的增加,网络提取特征的能力也逐步增加。然而随着网络层次的不断增加,梯度消失问题逐渐明显^[17],使得网络也越来越难以训练,此外,参数量的增加也使得网络过拟合的风险增加^[18]。因此,基于卷积层加池化层的标准 CNN 架构一般不超过 19 层。

2015 年,微软亚洲研究院,He 等^[19]提出了基于残差学习的卷积神经网络(ResNet)。ResNet 基于残差模块构建。152 层的 ResNet 在 ImageNet 上的 Top-5 准确率已经超过人类,并且文献^[19]也证实,ResNet 的识别率不会随着网络层次的增加而降低。近年来,以 ResNet 作为基础结构的新计算机视觉算法层出不穷。此外,基于 ImageNet 训练的 ResNet 权重在各种深度学习计算平台上都极易获得。基于以上原因,本文选用 ResNet50 作为图像特征提取的基础网络。本文给出的水位预测网络的总体架构如表 2 所示。

表 2 水位检测 CNN 网络架构

层名称	层信息	BN	激活	输出大小
Conv1	7×7, 64	yes	ReLU	112×112
3×3 max pool				
Conv2_x	$\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$	yes	ReLU	56×56
Conv3_x	$\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 4$	yes	ReLU	28×28
Conv4_x	$\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 6$	yes	ReLU	14×14
Conv5_x	$\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$	yes	ReLU	7×7
平均池化层				2

3 网络训练

3.1 二次数据增强算法

由于标注数据结果比较复杂,为了充分利用标注数据,图像在送入网络进行训练前,对数据进行了二次增强。二次增强算法如下:

步骤 1 将原始图像按照图 4(a)的标注结果,在线段端点处沿 y 轴方向裁剪图像,得到子图像 slice_image,如图 4(b)、(c)所示。

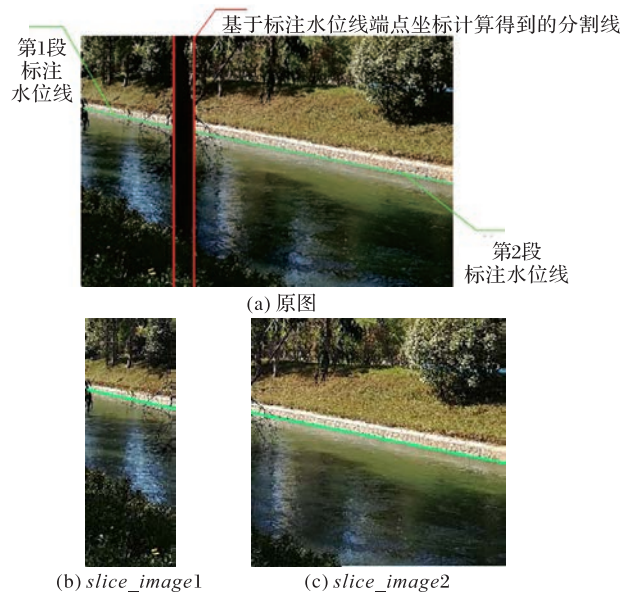


图 4 图像二次增强切片示意图

步骤 2 在 0~1 取随机参数 t ,使用以下公式计算水位线上随机坐标点:

$$\begin{aligned} point_random &= point_bgn + \\ & t \times (point_end - point_bgn) \end{aligned} \quad (2)$$

步骤 3 在 slice_image 图像中随机选取两个点 tl_point 及 br_point , tl_point 及 br_point 需满足如下条件:

$$\begin{cases} tl_point.x < random_point.x \\ tl_point.y < random_point.y \\ br_point.x > random_point.x \\ br_point.y > random_point.y \end{cases} \quad (3)$$

步骤4 以 tl_point 为矩形左上角顶点, br_point 为矩形右下角顶点, 构成一矩形 $crop_rect$ 。

步骤5 以 $crop_rect$ 为裁剪区域, 对 $slice_image$ 进行裁剪, 得到子图像 $patch_image$, 如图5所示。



图5 $patch_image$ 图例

步骤6 对 $patch_image$ 进行随机水平翻转, 随机调整HSV颜色, 随机模糊。

步骤7 基于图6(a)及如下公式所示计算 $patch_image$ 的标签值。

$$label = [\cos \theta, \sin \theta, d] \quad (4)$$

步骤8 完成。

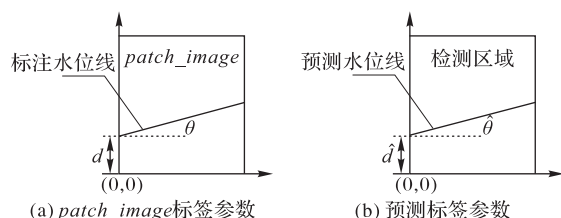


图6 标签计算参数示意图

3.2 损失函数

基于ResNet50对图像特征进行提取后, 最终网络输出层为2个节点, 预测结果为二维向量, 如下所示:

$$f(x) = [\hat{\theta}, \hat{d}] \quad (5)$$

其中: $\hat{\theta}$ 为预测水线与水平方向的夹角, 以下称为预测夹角, 如图6(b)中的 $\hat{\theta}$ 所示; \hat{d} 为预测水线与图像左侧边界交点的截距值, 以下称为预测截距, 如图6(b)中 \hat{d} 所示; $f(\cdot)$ 为本文给出的CNN; x 为输入图像。基于CNN的输出结果, 使用如下公式计算预测标签:

$$\logit_label = [\cos \hat{\theta}, \sin \hat{\theta}, \hat{d}] \quad (6)$$

网络损失函数定义如下:

$$interception_loss = |\hat{d} - d| \quad (7)$$

$$slop_loss = 1 - \cos \hat{\theta} \cos \theta - \sin \hat{\theta} \sin \theta \quad (8)$$

$$total_loss = slop_loss + interception_loss + 0.2 \times reg_loss \quad (9)$$

其中: $interception_loss$ 为预测截距与标签截距的差异损失, $slop_loss$ 为预测水位线与标注水位线斜率的差异损失, $total_loss$ 为总体损失, reg_loss 为L2正则化惩罚。

3.3 训练参数及过程

CNN非常适合于计算机视觉任务, 然而需要大量的高质量的数据标注。本文使用了训练前数据增强及训练中二次数据增强两种策略, 然而, 对于ResNet50这样规模的CNN网络而言, 训练数据依然是远远不够的。迁移学习^[20]是解决深度学习训练过程中数据饥饿问题的常用方法。本文研究以ResNet50在ImageNet上的训练结果作为预训练权重, 在网络初始化时加载ResNet50预训练权重中除分类网络部分的相关权重参数。使用该预训练权重可以极大地加快网络的收敛

并提高网络的泛化能力。训练使用参数及损失下降曲线如表3和图7。

表3 训练参数

参数	值
训练数据量	13 944(包含数据增强)
测试数据量	1 407(包含数据增强)
初始学习率	1e-3
学习率衰减	每2 000 batch, 衰减10%
batch size	32
训练步数	50 000
优化器	momentum

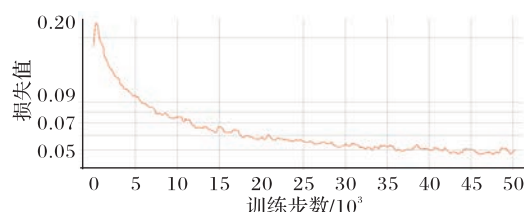


图7 平均损失下降曲线

4 测试和讨论

通过观察训练曲线, 在45 000步迭代后, 模型的损失下降趋势已趋于平稳, 故选择50 000步的训练结果作为模型的最终权重参数。对样本预测误差的评价方式使用如下公式计算:

$$erro = \frac{1}{2} |d_1 - \hat{d}_1| + |d_2 - \hat{d}_2| \quad (10)$$

其中: d_1 为样本标注水线与样本图像左侧边界交点的Y坐标值, \hat{d}_1 为样本预测水线与样本图像左侧边界交点的Y值; d_2 为样本标注水线与样本图像右侧边界交点的Y坐标值, \hat{d}_2 为样本预测水线与样本图像右侧边界交点的Y值。

测试数据集共1 407张图像(含数据增强), 测试时从以上测试集中顺序提取图像, 并从原图像中随机裁剪5个包含水位线且大小为224的图块进行预测。测试结果如表4和表5所示。

表4 测试结果统计

测试项	值
测试图像数量	1 407张(包含数据增强)
总测试样本数量	1 407×5 = 7 035 张(随机裁剪)
总像素误差	77 436.3 555 像素
平均像素误差	5.503 65 像素

表5 误差区间样本量统计

误差区间/像素	样本数量	误差区间/像素	样本数量
[0, 2]	949	[10, 12]	251
[2, 4]	1 934	[12, 14]	124
[4, 6]	1 826	[14, 16]	81
[6, 8]	1 162	[16, 18]	22
[8, 10]	681	[18, 20]	5

由表5可知: 0~8个像素的预测误差最为集中, 占总样本量的84.45%; 误差超过10个像素的样本占总样本量的16.55%, 共483个样本。通过对预测误差超过10像素的测试样本进行分析, 可分为图8中所示4种情况。如图8(a)、(b)所示情况, 是由于测试时从样本图像中随机裁剪出的区域中, 水线十分难以辨认(即便通过人眼都难以确认), 这类情况是导

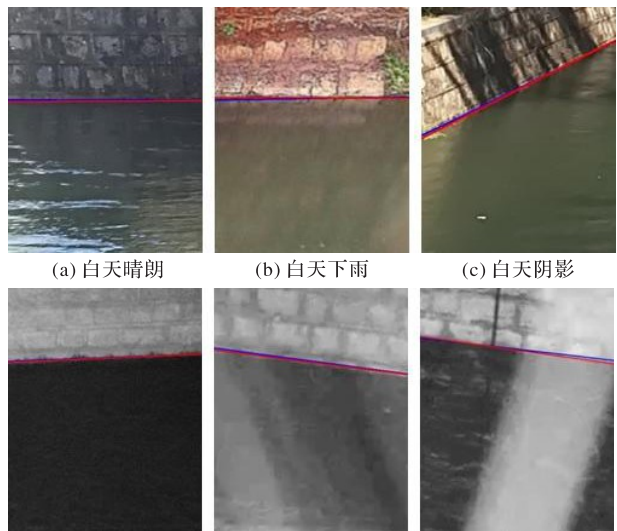
致预测误差过大的主要原因,共199个样本,占49.90%;如图8(c)所示情况,是由于人工标注数据时产生偏差,预测结果实际更加精确,这类情况在大误差预测结果中共42个样本,占7.66%。如图8(d)所示情况,是水线肉眼清晰可见的情形,但是预测结果有较大误差,这类情况在大误差预测结果中共205个样本,占42.4%。



(a) 水线难分辨 (b) 水线存遮挡 (c) 标记误差 (d) 预测误差

图8 大误差预测样本分析图例

算法在各种环境下的水位预测结果示例如图9所示。图9中,因为标注结果与预测结果几乎一致,所以多数图像只能看到一条预测直线;图9(f)为下雨情况下存在明显由下雨产生的自右上斜向左下贯穿图像的灰白色条块状遮挡,这种情况对识别产生了一定的影响,可以看到两条直线,其中一条直线为预测结果,另一条直线为标记结果。



(a) 白天晴朗 (b) 白天下雨 (c) 白天阴影 (d) 夜间晴朗 (e) 夜间下雨1 (f) 夜间下雨2

图9 预测结果示例

5 结语

本文提出了一种基于深度学习的水位线检测算法,该算法以ResNet50为特征提取网络,输出特征向量为预测水位线截距及预测水位线与水平方向的夹角。实验结果表明,本文算法稳定可靠,完全可以替代传统的基于图像的水位检测方法。

参考文献 (References)

- [1] HIES T, PARASURAMAN S B, WANG Y, et al. Enhanced water-level detection by image processing[C]// Proceedings of the 10th International Conference on Hydroinformatics. Hamburg, Germany: [s. n.], 2012: 563-571.
- [2] 鲍江,陶青川,张鹏. 基于图像处理的水位线检测算法[J]. 水电能源科学, 2015, 33(4): 96-99.
- [3] UDOMSIRI S, IWAHASHI U M. Design of FIR filter for water level

detection[J]. World Academy of Science, Engineering and Technology, 2008, 48: 47-52.

- [4] 高晓亮,王志良,王馨,等. 基于HSV空间的视频实时水位检测算法[J]. 郑州大学学报(理学版), 2010, 42(3): 75-79.
- [5] IWAHASHI M, UDOMSIRI S, IMAI Y, et al. Water level detection for functionally layered video coding[C]// Proceedings of the 2007 IEEE International Conference on Image Processing. Piscataway: IEEE, 2007: II-321-II-324.
- [6] KWAKB J Y, KOJ C, NAM Y. Automatic water-level detection using CCD camera and k-means clustering[J]. Korean Society of Image Science and Technology, 2011, 17(9): 1-8.
- [7] YU H, HAHN H. Remote detection and monitoring of a water level using narrow band channel[J]. Journal of Information Science & Engineering, 2010, 26(1): 71-82.
- [8] SAKAINO H. Camera vision based water level estimation[J]. IEEE Sensors Journal, 2016, 16(21): 7564-7565.
- [9] BRUININK M, CHANDARR A, RUDINAC M, et al. Portable automatic water level estimation using mobile phone cameras[C]// Proceedings of the 2015 14th IAPR International Conference on Machine Vision Applications. Tokyo: [s. n.], 2015: 426-429.
- [10] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521: 436-444.
- [11] SCHMIDHUBER J. Deep learning in neural networks: an overview[J]. Neural Networks, 2015, 61: 85-117.
- [12] VIEIRA S, PINAYA W H L, MECHELLI A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications[J]. Neuroscience Biobehavioral Reviews, 2017, 74(A): 58-75.
- [13] GU J, WANG Z, KUEN J, et al. Recent advances in convolutional neural networks[EB/OL]. [2019-04-08]. <https://arxiv.org/pdf/1512.07108v3.pdf>.
- [14] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]// Proceedings of the 25th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2012, 1: 1097-1105.
- [15] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[C]// Proceedings of the 2014 European Conference on Computer Vision. Cham: Springer, 2014: 818-833.
- [16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. arXiv, 2014: 1409.1556.
- [17] LOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift[EB/OL]. [2019-03-07]. <https://arxiv.org/pdf/1502.03167.pdf>.
- [18] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. Journal of Machine Learning Research, 2014, 15: 1929-1958.
- [19] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[EB/OL]. [2019-04-05]. <https://arxiv.org/pdf/1512.03385.pdf>.
- [20] YOSINSKI J, CLUNE J, BENGIO Y, et al. How transferable are features in deep neural networks?[C]// Proceedings of the 27th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2014: 3320-3328.