

无人机航拍图像中裸露地表的识别

钟映春¹, 谢林烽¹, 郑海阳¹, 罗志勇²

(1. 广东工业大学 自动化学院, 广州 510006;

2. 广州市优飞信息科技有限公司, 广州 510630)

摘要: 电力输电线路下方或附近的无覆盖物的裸露地表, 是引起输电线路事故的主要隐患之一; 从无人机电力巡检航拍图像中识别裸露地表可以预防类似事故的发生; 由于 Mask RCNN 识别无人机电力巡检航拍图像中裸露地表的精度较低, 提出一种图像特征融合的方法, 即人工提取 HOG 和 LBP 两种不同的图像特征, 经过不同权重的融合共同表征图像中裸露地表区域的特征, 再对 SVM 进行训练并用于识别; 实验结果表明, 采用该方法识别率可以达到 80% 以上, 识别时间少于 60 ms; HOG 和 LBP 两种特征在进行融合时, 当两种特征的数量级相当时, 得到的识别率最高; 可见, 该方法在具有较高识别率的同时, 具有比较好的实时性, 适合于无人机机载平台对航拍图像的初筛, 且训练时间较少, 权重参数规模小, 为无人机航拍图像中目标物的识别提供一种新思路。

关键词: 无人机; 航拍图像; 裸露地表; 特征融合

Recognizing Bared Earth Region of Aerial Images from UAV

ZHONG Yingchun¹, XIE Linfeng¹, ZHENG Haiyang¹, LUO Zhiyong²

(1. School of Automation, Guangdong University of Technology, Guangzhou 510006, China;

2. Guangzhou Ufly Technology Co., Ltd., Guangzhou 510630, China)

Abstract: The bared earth under or near power transmission lines is one of the main hidden trouble. It is one of the important tasks to recognize the bared earth region from aerial images of unmanned aerial vehicle (UAV) power inspection. After failing to employ the Mask RCNN to recognize the bared earth, this paper proposed an approach of image feature fusion. The HOG and LBP features of aerial images were extracted and fused by different weights. The experiments show that, (1) the average precision of recognizing the bared earth reaches 80%; (2) the average recognizing precision is best when the weights of two features make the order of magnitude of the two features equal. The approach proposed in this paper is suitable for the first image filtering by the UAV airborne platform because it has not only quite good recognizing precision but also quite good real-time performance, which provides a new idea for target recognition in UAV aerial images.

Keywords: unmanned aerial vehicle (UAV); power inspection; bared earth; features fusion

0 引言

采用多旋翼无人机进行电力输电线路巡检是当前电网巡检的主要发展趋势之一^[1]。无人机在巡检过程中不断对输电线路进行航拍, 而后通过算法对航拍图像进行线路缺陷或故障、事故隐患的初次筛选, 再把结果图像发送到远程服务器端进行二次甄别。这样可以显著减少无人机上图像数据的传送量, 提高效率。

输电线路下方或附近若存在无覆盖物的裸露地表, 是输电线路隐患的主要因素之一。其原因在于: 首先, 输电线路下方或附近存在裸露地表, 说明该处很可能会进行工程施工。在施工过程中, 工程机械在举升挖斗或者起重臂时, 极易进入输电线路的安全范围, 导致击穿事故的发生。其次, 裸露地表存在泥土流失或塌方的隐患, 严重时会导致

致输电线的支撑铁塔失稳甚至倾覆。因此, 在输电线路巡检过程中, 发现并及时处置裸露地表是电网巡检的重要工作内容之一。

采用无人机进行电力输电线路巡检, 实现无人机航拍图像中的裸露地表的自动识别, 可以显著提高巡检效率和人员安全性, 是当前电网巡检的主要趋势之一。Zhengrong Li 等^[2]利用脉冲耦合神经滤波器消除背景噪声和改进的 Hough 变换实现无人机航拍检测电力线。何思远等^[3]通过对无人机输电线路图像进行中值滤波、膨胀和腐蚀的预处理, 提取预处理图像的小波特征值, 利用基于近邻传播的聚类方法实现对目标图像中绝缘子、防震锤和输电塔设备的分类与识别。Carlos Sampedro 等^[4]采用梯度直方图 (HOG, histogram of grey) 特征训练两个多层感知神经网络

收稿日期: 2020-12-10; 修回日期: 2021-01-08。

基金项目: 国家自然科学基金(61975248); 广东省自然科学基金(2018A0303130137); 广州市科技计划项目(202007040004)。

作者简介: 钟映春(1973-), 男, 江西人, 博士, 副教授, 主要从事模式识别与图像理解, 无人机自动驾驶方向的研究。

通讯作者: 谢林烽(1994-), 男, 广州人, 硕士, 主要从事模式识别与图像理解方向的研究。

引用格式: 钟映春, 谢林烽, 郑海阳, 等. 无人机航拍图像中裸露地表的识别[J]. 计算机测量与控制, 2021, 29(7): 190-196.

络, 第一个用于背景-前景分割, 第二个用于分类 4 种不同类型的电塔取得不错的效果。Zhenbing Zhao 等^[5]采用一种深度卷积神经网络检测无人机航拍图像绝缘子的新颖方法, 应用具有多面体特征提取方法的卷积神经网络 (CNN, convolution neural network) 模型来表示绝缘子的状态, 并基于这些特征训练支持向量机 (SVM, support vector machine)。Xiao Wu 等^[6]提出采用条纹方向直方图和条纹长度直方图描述鸟巢的特征, 并将其用于高铁架输电线接触网系统中鸟巢的检测识别。Meng Lan 等^[7]提出一种替换掉 RoI 池化层的基于区域的 Faster R-CNN 模型, 用于无人机巡检过程中对电气设备的缺陷检测。钟映春等^[8]采用 YOLO (you only look once) V3 算法对无人机航拍图像中铁塔上的鸟巢进行识别。邹捷等^[9]提出一种优化的 R-FCN 深度学习图像识别算法, 该算法提高了验收图像缺陷识别的效率, 提升了无人机在巡检输电线路工程验收的工作效率。这些研究表明, 从图像中甄别电力输电线路各个元器件的缺陷、故障、隐患等是电网巡检的主要趋势之一。但是, 针对无人机电力巡检图像中裸露地表的识别研究尚未见诸报道。

针对无人机电力巡检图像中裸露地表的识别问题, 本文在采用 Mask R-CNN (mask region-convolution neural network) 识别效果不佳的基础上, 提出了一种图像特征融合的方法, 即人工提取两种不同的图像特征, 经过不同权重的融合共同表征图像中裸露地表的特征。

1 研究思路

本文研究的思路如图 1 所示。

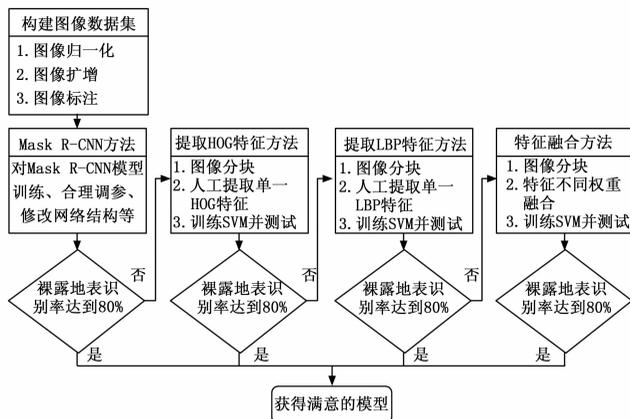


图1 研究思路

从图 1 可见, 在构建图像数据集的基础上, 本文首先尝试了深度学习方法里的 Mask R-CNN 算法, 发现该方法在识别裸露地表时的识别率难以达到使用的最低要求。在 Mask R-CNN 方法难以取得突破的情况下, 本文尝试采用人工提取单一特征和 SVM 识别方法, 结果表明识别率同样难以达到使用的最低要求。在尝试了两种单一特征后, 采用了特征融合的方法进行裸露地表识别, 使识别率

达 80% 以上, 且实时性比较高, 基本满足裸露地表的识别要求。

2 图像数据集的构建

2.1 图像获取

本文采用大疆“御” Mavic 2 无人机搭载 LID-20c 哈苏航拍相机采集图像, 航拍图像分辨率为 $5\,472 \times 3\,684$ pixels, 共采集包含不同场景的图像 1 493 张。对采集的图像进行归一化处理为 $1\,280 \times 1\,040$ pixels。

2.2 构建图像数据集

由无人机采集的原始图像总体数量不多。因此为了增加图像数据的多样性、减少模型训练过程中产生的过拟合、提高模型泛化能力, 本文采用了图像扩增技术^[10], 主要包括:

- 1) 随机角度旋转原始图像。
- 2) 在原始图像中增加高斯噪声等。

经过图像扩增后, 得到图像数据集共计 4 479 张, 随机选取 75% 的图像 3 359 张图像作为训练集图像, 其余的 1 120 张图像作为测试集图像。任意挑选一张无人机航拍图像及其经过图像扩增后得到的图像如图 2 所示。

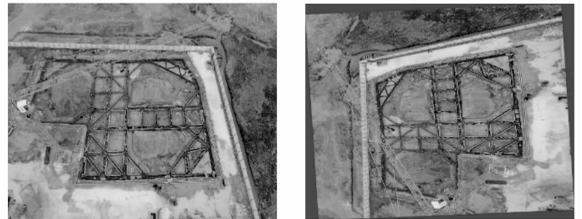


图2 图像扩增示例

2.3 图像标注

本文使用 LabelImg 软件作为标注工具, 对图像数据集中裸露地表的这一类别和位置进行标注^[11]。

3 构建识别模型

3.1 裸露地表的 Mask R-CNN 识别模型

Mask R-CNN^[12]是在延续了 Faster R-CNN 的思想上发展而来的, 模型结构图如图 3 所示。其模型框架由两个阶段组成, 第一阶段扫描图像并生成可能包含一个目标的 Region Proposal, 第二阶段对提议进行分类并生成边界框和 Mask 掩码, Mask R-CNN 算法在 Faster R-CNN 基础上增加 RoIAlign 以及全卷积网络 (fully convolutional network, FCN), Mask R-CNN 将分类预测和掩码 (mask) 预测拆分为网络的两个分支, 分类预测分支与 Faster R-CNN 相同, 对感兴趣区域给出预测, 产生类别标签以及矩形框坐标输出, 而掩码预测分支产生的每个二值掩码依赖分类预测结果, 基于此分隔出物体。Mask R-CNN 对每个类别均独立地预测一个二值掩码, 避开类间的竞争。在特征提取网络上, 它采用 Resnet101 残差网络与 FPN 特征金字塔网络结合, 能有效减少因网络过于深导致的网络退化问题。过往的经验表明 Mask R-CNN 具有优秀的目标

识别精确度，这是本文采用该方法的主要原因。

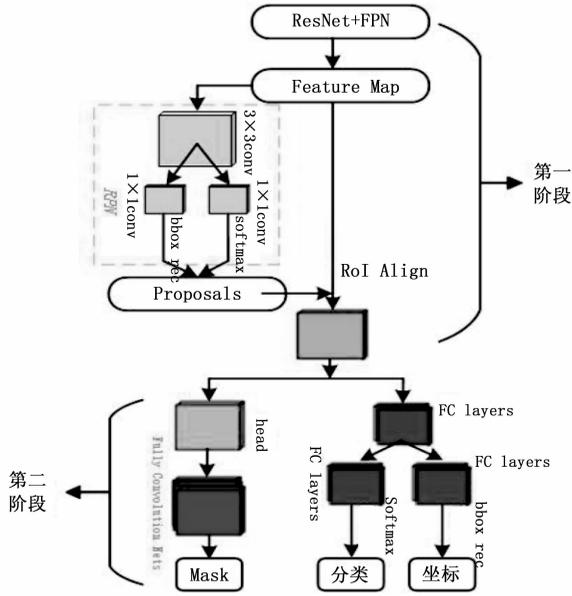


图 3 Mask R-CNN 模型结构

3.2 采用人工提取特征的裸露地表识别模型

采用人工提取特征的裸露地表识别模型总体框架如图 4 所示。

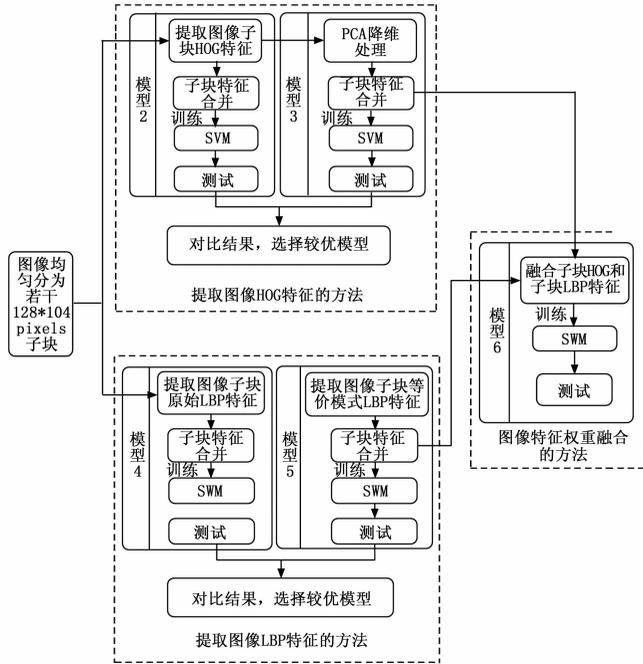


图 4 人工提取特征的裸露地表识别模型框架

对标注好裸露地表类别的每张图像裁切出若干个 128×104 pixels 的子块，然后对图像子块分别采用以下 5 种方法构建识别模型。

方法一：采用原始 HOG 特征的识别方法。首先，对图像划分为若干个图像子块，而后提取图像子块 HOG 特征；

其次，将子块 HOG 特征合并为单张图像的 HOG 特征；再次，输入到 SVM 分类器中训练获取得到模型 2。

方法二：采用主成分分析 (PCA, principal component analysis) 降维 HOG 特征的识别方法。首先，提取图像子块 HOG 特征；其次，采用 PCA 降维方法对特征矩阵进行降维处理；第三，将降维后的子块 HOG 特征合并为单张图像的 HOG 特征；第四，输入到 SVM 分类器中训练得到模型 3。

方法三：采用原始局部二值模式 (LBP, local binary patterns) 特征的识别方法。首先，提取图像子块原始 LBP 特征；其次，将子块原始 LBP 特征合并为单张图像的原始 LBP 特征；再次，输入到 SVM 分类器中训练获取得到模型 4。

方法四：采用等价模式 LBP 特征的识别方法。首先，提取图像子块等价模式 LBP 特征；其次，将子块等价模式 LBP 特征合并为单张图像的等价模式 LBP 特征；再次，输入到 SVM 分类器中训练获取得到模型 5。

方法五：采用降维后的 HOG 特征与等价模式 LBP 特征融合的识别方法。首先，分别提取图像 HOG 特征和等价模式 LBP 特征；其次，对 HOG 特征进行 PCA 降维处理；第三，将降维后的子块 HOG 特征合并为单张图像的 HOG 特征，将子块等价模式的 LBP 特征合并为单张图像的等价模式 LBP 特征；第四，用不同权重的融合降维后的 HOG 特征和等价模式 LBP 特征；第五，输入到 SVM 分类器中训练得到模型 6。

3.2.1 提取图像 HOG 特征的裸露地表识别模型

1) HOG 特征：

梯度方向直方图 HOG 特征是一种描述图像边缘和形状特征的算子^[13]。它通过计算和统计图像局部区域的梯度方向直方图来构成特征。对于一副图像，梯度或者边缘的方向密度分布能够较好地描述局部目标的表象和形状。HOG 特征提取算法的实现过程如下：

- (1) 对样本图像进行标准化 Gamma 空间和颜色空间，将图像进行 Gamma 标准化，这种处理能有效降低图像局部的阴影和光照变化等因素的影响。
- (2) 计算图像中横坐标和纵坐标方向的梯度。
- (3) 统计单元内梯度直方图。
- (4) 获取最终特征。本方法采用滑动窗口法，滑动步长为 8×8 ，把每个归一化好的块内的特征向量串联起来，即可得到最终的 HOG 特征向量。

2) 特征降维：

本方法实验采用的图像子块为 128×104 pixels，提取图像子块 HOG 特征所得特征向量维数为 6 480 维。由于该特征向量维度较高，包含信息较多，有些信息可能会对识别产生干扰，而且高维度特征所需训练时间也较长，所以尝试对特征向量进行降维。主成分分析 PCA 降维是较为常用的线性降维方法^[13]，主要是通过将原来高维度的变量做线

性投影到低维度的空间上。一般通过式 (1) 来计算特征的主成分:

$$y = U^T(x_i - \bar{x}) \quad (1)$$

其中: y 表示主成分特征, \bar{x} 为训练样本的特征均值, x_i 为需要降维的特征, 为协方差矩阵计算公式, 如式 (2) 所示:

$$U^T = \frac{1}{N} \sum_{n=1}^N (x_i - \bar{x})(x_i - \bar{x})^T \quad (2)$$

由人工提取图像 HOG 特征后, 对特征进行 PCA 降维处理, 处理后获得的特征输入到 SVM 分类器进行训练, 获取裸露地表图像识别模型。

3.2.2 提取图像 LBP 特征的裸露地表识别模型

1) LBP 特征:

局部二值模式 LBP 特征是一种描述图像局部纹理特征的算子。它通过利用中心像素点的灰度值作为阈值, 其相邻像素点的灰度值与中心像素点进行大小比较得到的二进制编码来表示局部的纹理特征。原始的 LBP 算子定义为: 在 3×3 的窗口内, 以窗口中心像素点为阈值, 周围的 8 个像素点的灰度值与中心点进行比较, 若周围像素值大于等于中心像素值, 则该像素点的位置标记为 1, 小于中心像素值标记为 0。如此, 按顺时针方向的 8 个像素点经过比较后可产生 8 位二进制数, 即可得到该窗口中心像素点的 LBP 值。如图 5 所示。

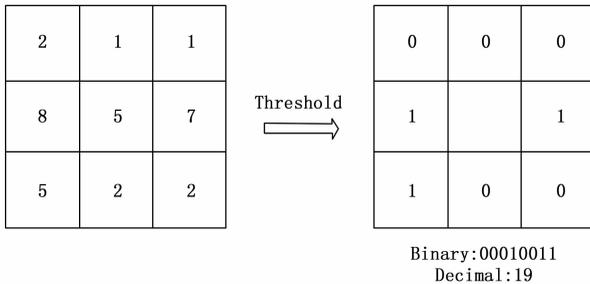


图 5 原始 LBP 算子

2) 等价模式 LBP 特征:

原始 LBP 算子仅能够处理局部 3×3 小范围内的信息, 难以满足不同尺寸的特征提取的需求, 具有局限性。Timo Ojala 等改进了原始的 LBP 算子^[14], 由原来的 3×3 范围扩展为任意范围, 但与此同时, 邻域的扩展使得二进制模式类别显著增加, 产生过多模式, 降低了效率, 不利于统计和分类。由此, Ojala 提出一种“等价模式”(Uniform Pattern) 的 LBP 算子, 实现了对 LBP 算子的降维处理。

当 LBP 算子的循环二进制数从 0 跳到 1 或从 1 跳到 0 最多 2 次时, 则属于等价模式, 例如 00011000。由以下公式 (3)、(4) 可计算是否为等价模式:

$$s(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (3)$$

$$U(G_p) = |s(g_{p-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{p-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \quad (4)$$

其中: g_c 对应局部领域中心像素的灰度值, g_p 对应于半径为 R 的圆周上等间隔像素 p 的灰度值, 若计算所得 $U(G_p)$ 不大于 2, 则为等价模式。等价模式能使原来二进制模式类别由种降低到种, 通过此种方式, 不但可以简化模式类别, 还可以降低高频噪声干扰, 在不损失有效的图像信息同时降低了特征数据的维度。实验证明采用等价模式的 LBP 特征比采用别的模式 LBP 特征对裸露地表分类识别的效果更好。

由人工提取图像等价模式的 LBP 特征后, 输入到 SVM 分类器进行训练, 获取裸露地表图像识别模型。

3.2.3 融合 HOG 与 LBP 的裸露地表识别模型

1) 特征融合的原因:

LBP 特征是一种用来描述图像局部纹理特征的算子, 具有灰度不变性和旋转不变性、计算较为简单等显著的优点, 对图像局部纹理特征描述较好, 但单独使用时容易对复杂图像特征提取不完整。HOG 特征是一种描述图像边缘和形状特征的算子, 对图像几何和光学的形状保持良好的不变性, 但 HOG 特征对边缘主体部分以及一些容易混淆的边缘部分很难表征, 且当背景有较多噪声边缘时也会削弱 HOG 特征的描述能力。针对单一 LBP 特征以及单一 HOG 特征在表征图像特征上的不足, 本方法将图像的 LBP 特征和 HOG 特征融合, 该融合方法既可以描述图像的局部纹理信息又可以描述图像的边缘轮廓信息, 提供了更加丰富的图像特征信息, 能更好提升裸露地表识别的效果。

2) 融合方式:

本方法采取的特征融合方式为两个特征向量以不同的权重系数和后进行串联融合, 由于实验所得到的 HOG 特征与 LBP 特征数值上不在同一个数量级上, 若 HOG 特征与 LBP 特征直接串联融合容易出现其中一方特征数据远远大于另一方特征数据, 使得另一方的特征表征特性被削弱, 降低识别分类精度的情况。所以在尝试两个特征以不同权重进行融合时, 一般选取的两个特征的权重数量级上的差距与原始的两个特征数量级的差距相当, 以此消除因两个特征在数量级上的差距而造成的一方特征表征特性被削弱的问题。融合后的特征 F 如式 (5) 所示:

$$F = \alpha HOG + \beta LBP \quad (5)$$

由人工分别提取图像的 HOG 特征、等价模式的 LBP 特征, 对 HOG 特征进行 PCA 处理, 然后经 PCA 处理后的 HOG 特征和等价模式的 LBP 特征以不同权重进行串联融合, 通过反复评价不同权重的 LBP 特征和 HOG 特征融合后的识别效果, 选取最优的权重组合, 获取裸露地表的最佳识别模型。

4 实验与结果分析

4.1 实验环境配置

4.1.1 深度学习实验环境

实验硬件环境: CPU 为 i7-6700K, 内存 32 G, GPU

为英伟达公司的 GTX1070，显存 8 G，硬盘 1 T。

实验软件环境：操作系统为 linux-ubuntu16.04，深度学习专用编程框架 CUDA，加速库 cuDNN，高层神经网络 API 库 Keras，TensorFlow 和 Python 3.6。

4.1.2 人工提取特征的实验环境

实验硬件平台：CPU 为 Core i5-8250U，内存 8 G，GPU 为英伟达公司的 MX150，显存 2 G，硬盘为 1 T。

实验软件平台：操作系统为 windows10，图像处理库 skimage，常用的机器学习库 sklearn，Python 3.6。

4.2 实验方案设计

本文设计以下实验。

实验 1：采用模型 1 即 Mask R-CNN 模型进行裸露地表识别，获得识别精确度和 F_1 值。

实验 2：采用模型 2 即原始的 HOG 特征方式进行裸露地表识别，获得识别精确度和 F_1 值。

实验 3：采用模型 3 即经过 PCA 降维的 HOG 特征进行裸露地表识别，获得识别精确度和 F_1 值。

实验 4：采用模型 4 即原始的 LBP 特征方式进行裸露地表识别，获得识别精确度和 F_1 值。

实验 5：采用模型 5 即等价模式的 LBP 特征方式进行裸露地表识别，获得识别精确度和 F_1 值。

实验 6：在对比上述各个实验的识别精确度和 F_1 值的基础上，采用模型 6 即特征串联融合的方式进行裸露地表识别，并探索最佳的融合权重。

4.3 评价指标

为了评估本文所采用的方法的效果，本文拟使用以下两个评价指标进行评价：精确度 (AP, average precision)^[15]、 F_1 指标^[16]。

其中，AP 值为识别模型在测试集上识别每个类别的平均精确度，它由精确度 (Precision) 和召回率 (Recall) 一同决定，计算式子如下：

$$P = \frac{TP}{TP + FP} \quad (6)$$

$$r = \frac{TP}{TP + FN} \quad (7)$$

$$AP = \int_0^1 p(r)dr \quad (8)$$

上述式子中，TP 为正确预测为正类的正样本数；FP 为错误预测为正类的负样本数；FN 为错误预测为负类的正样本数。AP 值为在 0 到 1 之间的所有召回值的精确平均值，即 P-R 曲线 (Precision-recall curves) 下的面积。AP 值越接近 1，说明模型效果越好。

F_1 指标综合表达了召回率 Recall 和精确度 Precision 相互促进又相互制约的关系，可以在二者之间取得平衡。 F_1 值越高，模型效果越好。 F_1 值计算式子如下：

$$F_1 = \frac{2 * p * r}{p + r} \quad (9)$$

4.4 实验结果与分析

4.4.1 实验 1 的结果与分析

实验 1 采用原始的 Mask R-CNN 模型，以及调整学习率、batch size 等参数的 Mask R-CNN 模型。实验结果如表 1 所示。

表 1 实验 1 结果

模型	AP/%	F_1 /%	MTD/ms
原始模型 1	52.37	51.25	153
调参后的模型 1	57.66	55.61	160

由表 1 可以看出，采用 Mask R-CNN 模型在测试集上识别裸露地表的效果一般，调参后测试的 AP 值只有不到 58%， F_1 值也只有 55.61%，说明在无人机采集的原始图像数据集不多的情况下，采用 Mask R-CNN 模型的表现一般。

4.4.2 实验 2 与实验 3 的结果与分析

实验 2 和实验 3 的实验结果如表 2 所示。

表 2 实验 2 与实验 3 结果

模型	AP/%	F_1 /%	MTD/ms
模型 2	72.96	70.58	73
模型 3	74.05	71.63	57

由表 2 可以看出，经过特征 PCA 降维处理的模型 3，在测试集上的表现 AP 值和 F_1 值分别相对于模型 2 提高了 1.09% 和 1.05%。同时，由于降维后的特征维度更低、计算量更少，识别时间也比未降维前快了 16 ms。实验说明采用 PCA 对原始 HOG 特征进行降维对裸露地表的识别效果和效率均有一定提升。

4.4.3 实验 4 与实验 5 的结果与分析

实验 4 和实验 5 的实验结果如表 3 所示。

表 3 实验 2 与实验 3 结果

模型	AP/%	F_1 /%	MTD/ms
模型 4	59.49	57.55	58
模型 5	65.82	62.67	55

由表 3 可以看出，采用等价模式 LBP 特征的模型 5 比模型 4 在测试集上的 AP 值高出 6.33%， F_1 值高出 5.12%。这说明等价模式的 LBP 特征有效降低了噪声干扰，提升了裸露地表的识别效果。同时，等价模式的 LBP 特征相比原始 LBP 特征简化了模式类别、降低了特征维度，所以在识别时间上也有了 3 ms 的提升。

4.4.4 实验 6 的结果与分析

通过实验 2~5 结果可以看出，无论是经过降维处理的 HOG 特征还是等价模式的 LBP 特征，单一特征的识别方法对裸露地表的识别精确度始终无法达到 80% 以上，无法达到实际应用的基本要求。

实验 6 对经过降维处理的 HOG 特征和等价模式的 LBP 特征进行不同权重的融合, 最后输入到 SVM 分类器训练。不同权重组合经训练的模型 6 在测试集上的表现如图 6 所示。

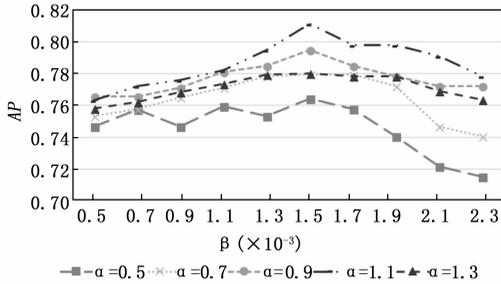


图 6 不同权重特征融合测试结果

由图 6 可以看出, 当 LBP 特征的权重系数为 0.001 5, HOG 特征权重系数为 1.1 时, 两特征融合后训练所获取的模型对裸露地表识别效果最佳, AP 值达到了 81.13%, 同时该权重组合的测试 F_1 值也达到了 78.37%。由此可见, 相比于提取单一特征训练获取的模型, 提取 HOG 特征和 LBP 特征经融合后再训练的模型对识别裸露地表效果提升明显。此外, 不同的权重系数组合会对识别效果产生影响, 通过 AP 值的高低反复评价不同权重对裸露地表识别的效果, 多次实验, 寻找出最优的权重组合。其本质就是, 选取的两个特征的权重数量级上的差距与原始的两个特征数量级的差距相当。

4.4.5 实验结果汇总与分析

实验 1 至实验 6 获得的实验结果如表 4 所示。

表 4 实验结果汇总

实验	AP/%	F_1 /%	MTD/ms	训练时间/min	权重参数规模/m
实验 1	57.66	55.61	160	1560	255
实验 2	72.96	70.58	73	44	8
实验 3	74.05	71.63	57	52	0.8
实验 4	59.49	57.55	58	32	2
实验 5	65.82	62.67	55	27	0.4
实验 6	81.13	78.37	62	60	1

由表 4 可知, 采用 Mask R-CNN 模型的实验 1 无论是训练时间还是训练出来的权重参数规模都大于其余实验, 且识别时间也比其余实验多将近 100 ms, 对于小数据集训练样本应用到无人机平台来说, 训练时间和权重参数规模都不宜过大。实验 6 采用的特征按照权重系数融合的方法, 既保证了识别的精确度, 且运行的效率比较高, 训练出来的权重参数规模占用空间也比较小。

任意挑选 2 张无人机航拍采集的图像, 模型 6 对该图像裸露地表识别结果如图 7 所示, 基本满足要求。

5 讨论

电力输电线路附近的裸露地表是导致电力输电线路发



图 7 模型 6 的识别案例

生相关事故的重要隐患之一。在无人机巡检电力输电线路过程中, 识别发现航拍图像中的裸露地表, 可以预防相关事故的发生, 是当前无人机电力巡检过程中现实而急迫的需要。然而, 目前尚未有从航拍图像中识别裸露地表的相关研究见诸报道。

针对航拍图像中裸露地表的识别问题, 我们分别采用深度学习中的 Mask R-CNN 模型、人工提取单一 HOG/LBP 特征和特征融合等 3 种方法进行了研究和探索。在规模为 4 000 张图像左右的数据集上进行的实验结果表明, 1) 将 HOG 和 LBP 特征按照权重进行特征融合后用于识别裸露地表是有效方法之一。2) 在 HOG 和 LBP 特征按照权重进行特征融合的方法中, 选取的权重系数在使得两种特征的数量级相当时, 得到的 AP 值和 F_1 值最佳。

当前, 广泛采用深度学习的方法识别无人机航拍图像中电力输电线路的金具、绝缘子、输电线、铁塔等各个具体环节^[5,7-9,17-18]。鉴于此, 本研究也将深度学习方法作为首选, 采用 Mask R-CNN 模型识别裸露地表。但是, 实验结果表明 Mask R-CNN 模型的识别率和难以令人满意。对比文献和本文的研究发现, 其中的原因很可能有: (1) 图像数据集规模不够大; (2) 文献中的识别目标多数都具有显著的细节特征。在深度学习中以 CNN 为基础的模型较擅长提取目标的细节特征, 但是在识别裸露地表这类具有宏观特征的目标时, 其局限性比较明显。

在采用 Mask R-CNN 模型难以取得满意结果后, 我们尝试采用了人工提取特征的方法。虽然也有一些研究采用人工提取特征方法进行输电线路相关的研究^[3-4,6], 但是这些特征提取方法均未涉及到识别裸露地表这样的目标。我们的实验表明, 提取图像中裸露地表的单一特征的识别率同样难以令人满意。从航拍图像可见, 从宏观上看裸露地表与周边地块有显著差异, 同时具有一定的纹理特性。HOG 特征更多的表达了目标的灰度宏观特征^[19], 而 LBP 特征主要表达目标的纹理特征^[20]。为此, 我们采用权重的方式将两者融合在一起, 并研究了权重对识别率的影响。实验结果表明, 在权重使得两个特征的数量级相当时, 识别率基本达到实际使用的要求。

相对于深度学习的 Mask R-CNN 模型而言, 特征融合的方法不仅在训练时间上显著更少, 而且在小规模图像数据集的情况下识别精度也显著比 Mask R-CNN 模型更高。

相对于单一特征的方法而言,采用权重融合特征的方法兼具了宏观特征和微观特征表达,能够更加准确的反映图像区域的本质。由此可见,针对无人机航拍图像中的裸露地表识别问题,在图像数据集规模不大的情况下,我们提出权重融合特征的识别方法,比较准确高效的识别出图像中的裸露地表区域。这个方法不仅是无人机机载平台对图像进行实时初筛的关键基础之一,而且为解决其它具有宏观特征的目标识别提供了一种良好的解决方案。

需要指出的是,尽管特征融合的方法能够使得识别率基本达到要求,但是距离实际应用相去甚远。后续,我们将探索融合更多的特征,将裸露地表的识别率提高到 95% 以上,真正能够进入实际应用阶段。此外,尽管目前 Mask R-CNN 模型对裸露地表的识别率不高,但是相比于人工提取特征,深度学习方法显然具有更大的提升空间。因此,后续我们将在增大图像数据集规模的基础上,探索改进深度学习中相关模型在宏观表达上的能力。

6 结束语

针对无人机巡检航拍图像的裸露地表识别,本文分别对 Mask R-CNN 模型、人工提取单一 HOG/LBP 特征和特征融合等 3 种方法进行了研究和探索。在小规模数据集下,将 HOG 和 LBP 特征按照权重进行特征融合后用于识别裸露地表相较于深度学习模型和人工提取单一特征的模型是更为有效的方法,在保证识别裸露地表一定的精确度同时,也具备了实时性的特点,同时训练时间以及权重参数规模都更低。本文的探索为无人机巡检过程中实时识别航拍图像中的目标物提供了一种新思路。

参考文献:

- [1] DONG G, CHEN X G, WANG B H, et al. Inspecting transmission lines with an unmanned fixed-wings aircraft [C] // 2012 2nd International Conference on Applied Robotics for the Power Industry (CARPI), IEEE, 2012: 173-174.
- [2] LI Z R, LIU YUEE, RODNEY WALKER, et al. Towards automatic power line detection for a UAV surveillance system using pulse coupled neural filter and an improved Hough transform [J]. Machine Vision & Applications, 2010, 21 (5): 677-686.
- [3] 何思远, 刘刚, 王玲, 等. 基于无人机的输电线路设备识别方法研究 [J]. 红外与激光工程, 2013, 42 (7): 1940-1944.
- [4] CARLOS SAMPEDRO, CAROL MARTINEZ, ANEESH CHAUHAN, et al. A supervised approach to electric tower detection and classification for power line inspection [C] // 2014 International Joint Conference on Neural Networks (IJCNN), IEEE, 2014: 1970-1977.
- [5] ZHAO Z B, XU G Z, QI Y C, et al. Multi-patch deep features for power line insulator status classification from aerial images [C] // 2016 International Joint Conference on Neural Networks (IJCNN), IEEE, 2016: 3187-3194.
- [6] WU X, YUAN P, PENG Q, et al. Detection of bird nests in overhead catenary system images for high-speed rail [J]. Pattern Recognition, 2016, 51: 242-254.
- [7] LAN M, ZHANG Y P, ZHANG L F, et al. Defect detection from UAV images based on region-based CNNs [C] // 2018 IEEE International Conference on Data Mining Workshops (ICDMW), IEEE, 2018: 385-390.
- [8] 钟映春, 孙思语, 吕帅, 等. 铁塔航拍图像中鸟巢的 YOLOv3 识别研究 [J]. 广东工业大学学报, 2020, 37 (3): 42-48.
- [9] 邹捷, 许瑞庆, 习雨同. 深度学习在输电线路工程验收智能缺陷识别中的应用 [J]. 江西电力, 2020, 44 (2): 5-9.
- [10] 魏伟航, 马乾力, 高治良, 等. 改进卷积神经网络在图像烟雾检测的应用 [J]. 智能计算机与应用, 2019, 9 (3): 47-51.
- [11] CAROLINE GEVAERT, CLAUDIO PERSELLO, SANDER OUDE ELBERINK, et al. An automated technique for basemap updating using UAV data [C] // 2017 Joint Urban Remote Sensing Event (JURSE), IEEE, 2017: 1-4.
- [12] He K M, GEORGIA GKIOXARI, PIOTR DOLLÁR, et al. Mask RCNN [C] // Proceedings of the IEEE international conference on computer vision, 2017: 2961-2969.
- [13] PAN S J, SUN S T, YANG L F, et al. Content retrieval algorithm based on improved HOG [C] // 2015 3rd International Conference on Applied Computing and Information Technology/2nd International Conference on Computational Science and Intelligence, IEEE, 2015: 438-441.
- [14] TIMO OJALA, MATTI PIETIKALNEN, DAVID HARWOOD. A Comparative Study of Texture Measures with Classification Based on Feature Distributions [J]. Pattern Recognition, 1996, 29 (1): 51-59.
- [15] CHRISTOPH FEICHTENHOFER, AXEL PINZ, ANDREW ZISSERMAN. Detect to track and track to detect [C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 3038-3046.
- [16] REN S Q, HE K M, ROSS GIRSHICK, et al. Object detection networks on convolutional feature maps [J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 39 (7): 1476-1481.
- [17] 宁柏锋. 基于深度学习的无人机巡检图像销钉故障检测 [J]. 计算机测量与控制, 2019, 27 (11): 25-29.
- [18] 欧家祥, 史文彬, 张俊玮, 等. 基于深度学习的高效电力部件识别 [J]. 贵州电力技术, 2018, 021 (9): 1-8.
- [19] NAVNEET DALAL, BILL TRIGGS. Histograms of oriented gradients for human detection [C] // 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), IEEE, 2005, 1: 886-893.
- [20] TIMO OJALA, MATTI PIETIKAINEN, TOPI MAENPAA. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on pattern analysis and machine intelligence, 2002, 24 (7): 971-987.