

基于机器学习的草坪湿度自动检测喷洒系统

付予泽¹, 林浩岚²

(1. 复旦大学第二附属中学 2023 届 12 班; 2. 复旦大学附属中学 2023 届 12 班)

指导教师 钱隆 (复旦大学附属中学)

摘要 本文工作针对目前草坪浇灌不足和水资源利用率较低的现象, 在原有固定喷灌的基础上, 运用图像识别和机器学习技术, 对已有草坪图像数据进行训练, 产生训练集。然后将待喷洒的真实草坪图像分割为多个区域后导入训练集, 实现对草坪各区域湿度状况的自动分类。根据分类结果向喷头控制器发送指令, 执行喷洒操作, 从而避免定时定量喷灌所导致的水资源浪费, 达到节水并提升草坪质量的目的。

关键词 图像识别; 机器学习; 草坪自动喷灌系统

一、课题的背景和意义

随着我国足球文化的快速普及, 真草球场越来越多地出现在我们的生活中。同时, “美丽中国”的倡导也使得大草坪在城市中遍地开花, 成为滋养城市居民身心的必需品。无论是真草球场还是大草坪都有体感真实、减少受伤、美化环境等效用, 但同时也存在铺设成本高、养护工序复杂、需水量大等问题。由于我国节水灌溉技术起步较晚, 而草坪养护的需水量不断扩大, 产生大量浪费, 也极大提高了草坪的养护费用。因此, 真草操场在校园中的普及度一直不高, 且大多养护较差。图 1 为我校真草操场, 可以看到养护情况并不理想。另一方面, 城市公园中的大草坪往往采用定时定量浇灌的方式, 节省人力的同时也造成水资源的浪费。



图 1. 高中操场实拍图

■ 相关领域现状

草坪对湿度温度要求较高：适宜草坪生长的温度在 $15^{\circ}\text{C}\sim 20^{\circ}\text{C}$ 之间；一个面积为 7140 平方米的标准足球场^[1]平均每天需要 8000 升水，在温度可达 40 摄氏度的夏季，需水量更多；另一方面，黄梅天等雨季降水量已经过大，再洒水会导致土地养分流失、草地萎蔫且不宜使用。

经资料查询和实地考察，上海大多数草坪都是采用定时定量喷洒技术。在铺设草坪时事先安装好喷头，在特定的时间点，如：清晨、夜晚，喷头从地底伸出，开始进行圆圈型喷洒，一段时间后结束。该技术存在三个问题：1) 固定的绕圈喷洒方式重复性强，容易喷至场外（如图 2 所示），造成水资源的浪费；2) 重叠范围大，但依然存在喷洒不到的盲区，导致草坪旱涝不一；3) 由于每天洒水量一定，对天气的适应能力差，晴天“旱”，雨天“涝”。经过论文查询，大部分新型草坪浇灌技术，也都是使用 PLC 等控制器对喷洒进行调节，缺

少对草坪各区域缺水情况的智能分析^[2-4]。



图 2. 操场的绕圈喷洒

二、 本文研究内容和结构

我们的课题意在开发一种新型的智能草坪喷洒系统，以达到更好地养护草坪、节水节能的目的。该系统由摄像头、计算机、控制器、喷头等模块组成，后文将详细介绍。

课题的核心是我们设计的一个“图像识别+机器学习”程序：搜集大量草坪图像作为数据集，运用图像识别和机器学习算法对数据集进行训练，产生训练集，然后获取待喷洒的草坪图像进行区域分割，导入训练集后对草坪各区域的缺水状况加以计算和分类，根据分类结果控制喷头的喷洒量，达到按需喷洒的效果。

在报告的第二部分，将描述程序中所使用算法的原理；第三部分讨论我们课题的解决方案；第四部分详细论述方案的具体实现；第五部分将对全文进行总结，并提出改进方案和未来展望。

三、 算法基础和解决方案

在本节，我们先论述课题所使用算法的思路和原理，然后给出解决方案及系统部署。

1. 算法基础

机器学习(Machine Learning, ML)是一门多领域交叉学科，涉及概

率论、统计学、逼近论、凸分析、算法复杂度理论等多门学科^[5]。它专门研究计算机怎样模拟或实现人类的学习行为，用数据或以往的经验获取新的知识或技能，使自身的性能不断改善。我们的课题就是利用机器学习算法对以往的草坪数据进行训练产生训练集，然后将需要自动喷洒的草坪图片导入训练集进行计算分类，根据分类结果控制喷洒量，使得喷洒系统具有一定“智能”。

经过学习和查询比较，我们选择机器学习的经典算法 KNN 算法对草坪数据集进行训练和分类，并采用了有丰富的第三方库（numpy, os, matplotlib, cv2, pillow）支持的 python 语言加以实现。

■ KNN 算法

K 最近邻（K-Nearest Neighbor）算法，即 KNN 算法，是一种基于实例的学习，或者是局部近似和将所有计算推迟到分类之后的惰性学习^[6]。用最近的邻居（k）来预测未知数据点。

在我们的程序中，使用了 KNN 算法来识别草坪需水量的不同区域，即，将相同类别的样本在特征空间中聚在一起。首先，我们将草坪各区域分为绿草、枯草、裸露土壤等三种类型，先搜集大量的各类型的图片数据，然后利用 KNN 算法对这些数据进行学习，运用 cv2 和 os 库对图片文件进行批量操作，取各类图片的 RGB 平均值为特征值。这样，在特征空间中就有了三类草坪数据的特征值，即，训练集。接着，将待喷洒草坪的真实图像分割为若干区域，计算草坪各区域与特征值的抽象“距离”（利用 numpy 库中 linalg.norm 函数），选出 k 个最近数据实现分类（classification）。k 值是预测精度的一个关键因素，需要通过实验加以确定。

■ 梯形/桶形畸变校正

大多数实拍照片中的草坪由于透视关系呈梯形而不是真实情况

中的矩形，不方便处理，所以需进行梯形畸变校正，将梯形转换为原始的矩形。有时图片使用了广角镜头拍摄，还可能产生桶形畸变，也需要采用相应校正方法进行转换。

■ 边框标识算法

OpenCV(Open Source Computer Vision Library)是一个基于(开源)发行的跨平台计算机视觉库，它轻量级而且高效，同时提供了 Python 接口，实现了图像处理和计算机视觉方面的很多通用算法。CV2 是 OpenCV 官方的一个扩展库^[7]。在我们的程序中，导入 cv2 库进行待喷洒草坪各区域分类结果的展示，通过以绿黄红三种颜色标记草坪的各区域边框，达到直观显示各区域湿度的目的。

2. 解决方案

我们设计的智能喷洒系统(图 3)由五部分组成：摄像头、计算机、控制器、喷头和草地。

摄像头：安装在操场附近的高处，如教学楼楼顶，要求能够将操场的全景拍摄下来。另外，摄像头需要与操场中线对齐，拍摄到的草地图像呈梯形。摄像机像素要求较高，下文会有提及。每日一次将拍摄的图像传递给位于机房的计算机。

计算机：得到摄像头传递的图像后，将图像导入程序，计算出草地不同位置需要喷洒的强度(该步骤下文将重点阐述)，然后将其转化为指令传递给位于操场附近的控制器。

控制器：收到计算机传递的指令，将其转化为电信号，传输到操场不同位置的喷头。

喷头：收到控制器的电信号，根据指示喷洒特定区域，以满足各处草地不同的需水量。

应用部署

摄像头：SXGA(1280*1024)

喷头：购买利用改进起落点精确控制的草坪喷头专利和控制水压水流的FS200系列流量开关。

场地：在足球场地下铺设电缆、水管。

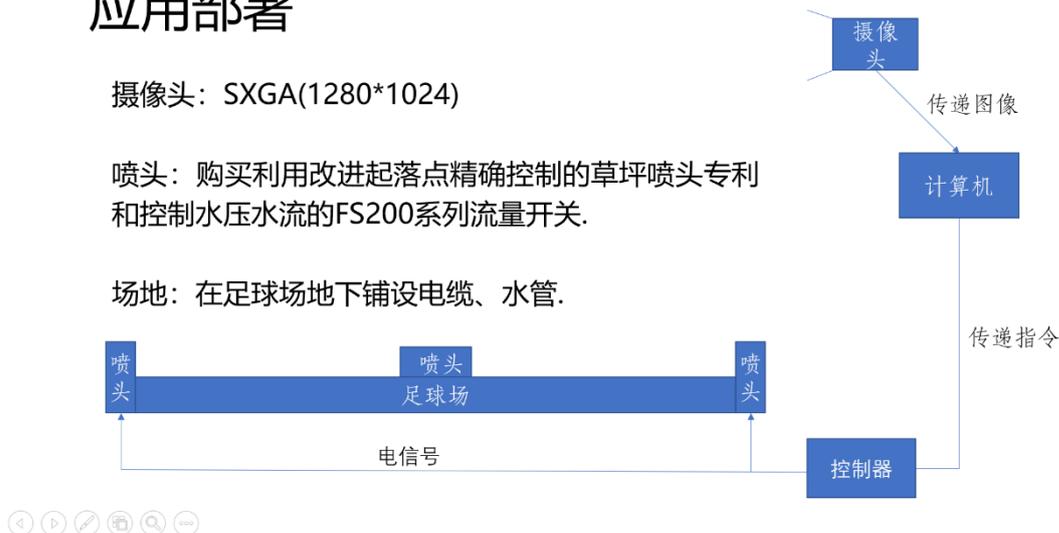


图 3. 草地的智能化喷洒系统部署

四、 具体实现

解决方案的具体实现包括：数据采集，数据处理，建模，评估等多个步骤，最后是实验结果和分析。

1. 数据采集

计划通过上网搜索和实地考察等多种方式，收集绿草、枯草、裸露土壤等三类图片各 1000 张(图 4)，学校操场全景俯视图 100 张(图 5)。由于时间关系，未完全搜集到所计划的图片数量。



图 4. 收集的草地图片示例



图 5. 拍摄的草地照片示例

2. 算法实现

■ 数据处理

对绿草、枯草、裸露土壤等三种类型各使用 800 张图片运用 KNN 算法进行训练，用卷积的方法提取三种类型草坪的颜色属性（RGB 值），允许一定的上下浮动，完成训练集创建。然后，各使用 200 张照片进行测试，由上一步获取的训练集自动判断照片属性，正确率达到 90%左右。至此，测试集训练完成。

【数据处理程序】

```
import cv2
import numpy as np
import os

# Define a function to extract our feature (average color)
def averagecolor(image):
    return np.mean(image, axis=(0, 1))

trainY = ['green', 'dry', 'dirt']

# 训练
trainX2 = []
trainY2 = []

# Let's loop through the training images in the 4 folders in the image
subdirectory
for label in trainY:
    print ("Loading training images for the label: "+label)

    #Load all images inside the subfolder
    for filename in os.listdir(label+"/"):
        img = cv2.imread(label+"/"+filename)
        img_features = averagecolor(img)
        trainX2.append(img_features)
        trainY2.append(label)

# 读取 jpg
new_card = cv2.imread("test dry.jpg")
new_card_features = averagecolor(new_card)
calculated_distances = []
for card in (trainX2):
    calculated_distances.append(np.linalg.norm(new_card_features-card))
cd = np.array(calculated_distances)

top3 = cd.argsort()[:3][: :-1]
```

```

top = []
print(top3)
for i in top3:
    top.append(i)
print(trainY2[max(top, key = top.count)])

```

■ 建模

首先使用梯形/桶形畸变校正算法将草坪的实拍图像转换为矩形，然后将矩形图像切割为 32*32 像素的小正方形，进行网格化处理。接着，将切割出的小正方形依次输入第一步建立的训练集进行计算，根据计算结果将绿草标记为绿色、枯草标记为黄色、裸露土壤标记为红色，分别对应三种不同的湿度和喷洒强度。

■ 评估

将实地拍摄获取的 100 张不同草地照片逐一导入上述模型，网格化后进行区域分析，并利用 cv2.rectangle 标记程序的计算结果(图 6)。然后人工评判其准确率。若准确率小于 80%，则对训练集进行改进：删除个体差异较大的草地照片；调整 KNN 算法的近邻个数 k 等。

【分类和标记程序】

```

import cv2
from PIL import Image
import numpy as np

k = 3

img = cv2.imread("test2.jpg")
canvas = img.copy()
dic = {'green':(0, 255, 0), 'dry':(0,255,255), 'dirt':(0,0,255)}

def averagecolor(image):
    return np.mean(image, axis=(0, 1))

def draw(type,x1,y1,x2,y2):
    cv2.rectangle(canvas,(x1,y1),(x2,y2),dic[type],2)

def check(image):
    new_card_features = averagecolor(image)
    calculated_distances = []
    for card in (trainX2):
        calculated_distances.append(np.linalg.norm(new_card_features-
card))

```

```

cd = np.array(calculated_distances)

topnp = cd.argsort()[0:1][::-1]
top = []
for i in topnp:
    top.append(i)
return trainY2[max(top, key = top.count)]

ss = 30
for i in range(ss, (img.shape[0] // ss) * ss+1, ss):
    for j in range(ss, (img.shape[1] // ss) * ss+1, ss):
        cut = img[i - ss : i, j - ss : j]
        draw(check(cut), j - ss, i - ss, j, i)
cv2.imshow('image', canvas)
cv2.waitKey(0)
cv2.destroyAllWindows()

```

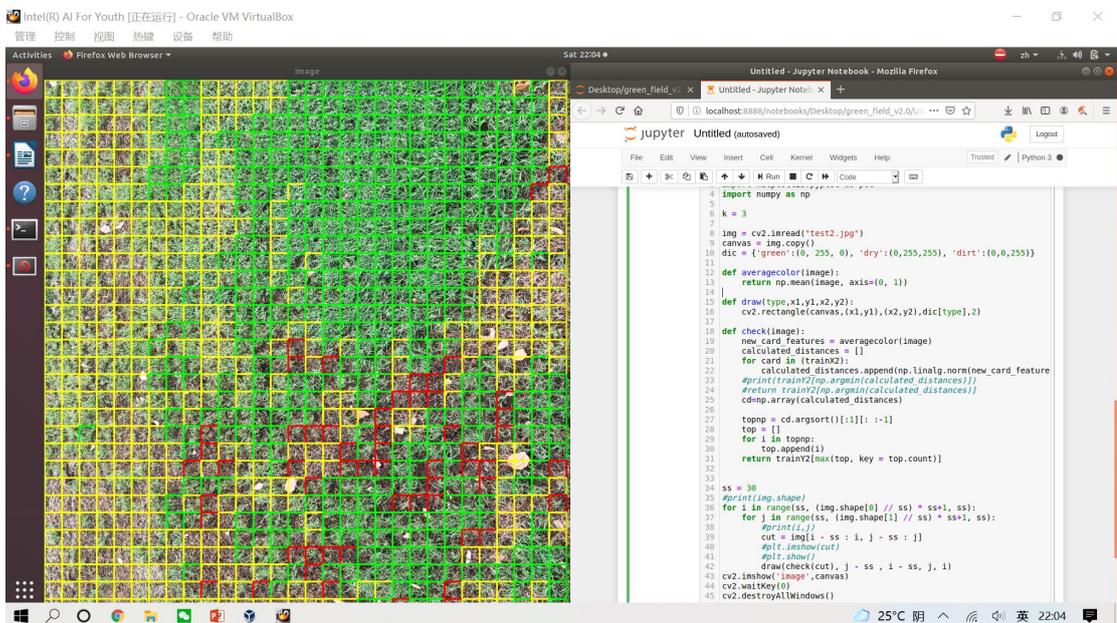


图 6. 利用训练集对草地的分类和标

图 7. 程序界面

3. 实验结果和分析

■ 实验结果

对实拍草坪图像的分析计算结果示例如图 6 所示，经人工评测，准确率达到 80%以上，实现了草坪全区域的有效分类。

实验刚开始时，模型准确率较低。原因在于：产生训练集的大部分草地图片通过网络获取，经常包括蓝天、花卉之类的干扰因素，影响平均色彩属性。因此我们调整了训练集，删除了非典型图片。通过

这个操作，模型准确率得到显著提升，达到 60%左右。接下来，我们寻找较为合适的 KNN 算法的不同近邻个数，在尝试了 2、3、4、5 等多个近邻数之后，最终确定近邻数为 3 时，结果较为准确。

■ 实验结果分析

由于只涉及二维图像，因此无法对是否有异物进行判定。若图像中出现了较深色的其它物品，如：垃圾袋，则程序将会判定其为裸露的土壤，而进行非必要的浇水。我们的解决方法是：在训练集中加入会出现在操场上的常见异物（如：禁区线、塑料瓶、衣物等），使程序能够进行识别，然后在网格化图像前先进行大范围搜索，找出异物并标记为不浇水区域。

由于数据原因，我们只设计了三个需水量档位，较为粗略，缺乏渐变，且喷头难以操作。解决方法是：在两种不同色块之间添加渐变色，通过颜色属性的差值确定需水量的差值。在不添加训练集分类的情况下，达到喷水渐变的效果。

天气也会影响草地实时图像的色调，晴天较亮，阴天较暗，色彩属性有一定平移。可以读取当天的天气情况，根据不同的天气对颜色参数进行校正，再对图像进行判断。另外，天气还关系到需水总量的大小，因此，通过当天降水量计算出减少的喷洒量，能够有效节省水资源。

五、 总结和展望

本文针对现有的草坪自动喷灌系统进行了改进，通过机器学习技术对草地的不同区域进行缺水量的自动判断，配合增设摄像头、提高喷头精度，可达到草坪全覆盖、按需分配水量的效果。

由于资金和时间的的原因，我们虽然设计了完整的草地自动喷洒系统，但核心部分是实现了草地的智能化区域计算和分析。后续若能将

整个系统投入实践，将极大的降低耗水量，提高草坪品质，节省养护费用。

1. 实践方案

第一步，我们将在自家的小草坪，进行安装和测试，确保系统的可行性和稳定性。

第二步，对学校原有的系统略微进行更改，加入摄像头，更换喷头，利用程序进行处理，测试系统实际效果。

第三步，总结实验数据，获得专家认证后，可向全市足球场和草坪绿地进行推广。



图 8. 理想的节水草坪

2. 推广扩展

扩展一：可以将本系统应用于草原牧场。由于本程序采用平面图像识别技术，只要是较为平整的草地都可以运用。城市公园中的草坪大多形状规整，没有高低起伏，面积偏小，正符合本系统的应用对象。草原牧场的草地虽然平整，但是不适合铺设地下管道。原因有三点：1) 面积较大，铺设成本高昂；2) 喷灌牧场对精度的要求较低，不需要大范围铺设喷头；3) 地下铺设大兴水土，会破坏生态环境。面对这种情况，可以将地下部分改为地上，使用无人机进行超低空人工喷洒。此时，自动喷洒系统可以借助蓝牙等无线通讯方式，将计算机程

序计算出来的不同区域需水量指示精确地发送至无人机。如果该方案能得以实施，将为防止西北荒漠化蔓延提供极大的助力。



图 9. 可应用于青青牧场

扩展二：可应用于花圃。目前本课题只采用了单色调提取技术，无法适用于花卉等颜色丰富的植物。但是像花圃这种花卉种类较规则，长势较平均的，仍然可以通过改进程序以达到自动浇灌的效果。可以根据不同的花卉品种，构建多个训练集，将原来的单组颜色参数替换成多组，即可实现自动湿度识别。



图 10. 可应用于花圃浇灌

参考文献

- [1] <https://baike.baidu.com/item/%E8%B6%B3%E7%90%83%E5%9C%BA%E9%9D%A2%E7%A7%AF/2537916>
- [2] 张淑红, 陈进. 草坪自动喷灌系统结构研究与应用[J]. 农业装备技术, 2007, 33(6): 25-27.
- [3] 蒋晓云. 基于 PLC 草坪自动喷灌系统设计[J]. 甘肃水利水电技术, 2015, 51(8): 36-40.
- [4] 方伟, 苏仲洋, 栾海. 万达足球训练基地慧云智能化管理系统简析[J]. 智能建筑, 2020, 234: 38-41.
- [5] <https://baike.baidu.com/item/%E6%9C%BA%E5%99%A8%E5%AD%A6%E4%B9%A0/217599?fr=aladdin>
- [6] https://blog.csdn.net/m0_60720471/article/details/119818963
- [7] https://blog.csdn.net/qq_40723205/article/details/100139047