# 机器学习组 报告

## 实验思路

利用python实现对遥感影像的林地识别。根据影像的波段信息，读取需要波段，分别为红、蓝、绿、近红外波段。此后，读入影像标注后的信息，作为一个新的波段，包含林地、非林地信息。

参考Distinguishing Planting Structures of Different Complexity from UAV Multispectral Images以及《植被及颜色相近地物分类的高光谱图像降维方法研究》两边论文思路，我们根据红、蓝、近红外波段计算植被指数，并对红、蓝、绿、近红外四个波段计算纹理指数，将计算得到的所有指数合并到一个特征向量中。根据各个指数之间的相关系数以及互信息，筛选得到10个指数。最后采用机器学习方法识别影像中的林地，并进行精度评价。

## 实验数据

选取Planet卫星以及哨兵2号卫星在巢湖附近的遥感影像，Planet影像分辨率为3M级，哨兵2号影像分辨率为10M级。各影像波段介绍如下：

### Planet 各波段介绍

|  |  |  |
| --- | --- | --- |
| 波段 | 波谱范围 | 描述 |
| B1 | 420 nm - 530 nm | 蓝色 |
| B2 | 500 nm - 590 nm | 绿色 |
| B3 | 610 nm - 700 nm | 红色 |
| B4 | 760 nm - 860 nm | 近红外 |

### 哨兵2号各波段介绍

|  |  |  |
| --- | --- | --- |
| 波段 | 中心波长 | 描述 |
| B1 | 443 nm | 超蓝(沿海和气溶胶) |
| B2 | 490 nm | 蓝色 |
| B3 | 560 nm | 绿色 |
| B4 | 665 nm | 红色 |
| B5 | 705 nm | 可见光和近红外(VNIR) |
| B6 | 740 nm | 可见光和近红外(VNIR) |
| B7 | 783 nm | 可见光和近红外(VNIR) |
| B8 | 842 nm | 可见光和近红外(VNIR) |
| B9 | 865 nm | 可见光和近红外(VNIR) |
| B10 | 940 nm | 短波红外线(SWIR) |
| B11 | 1375 nm | 短波红外线(SWIR) |
| B12 | 1610 nm | 短波红外线(SWIR) |
| B13 | 2190 nm | 短波红外线(SWIR) |



图1. 哨兵2号卫星影像

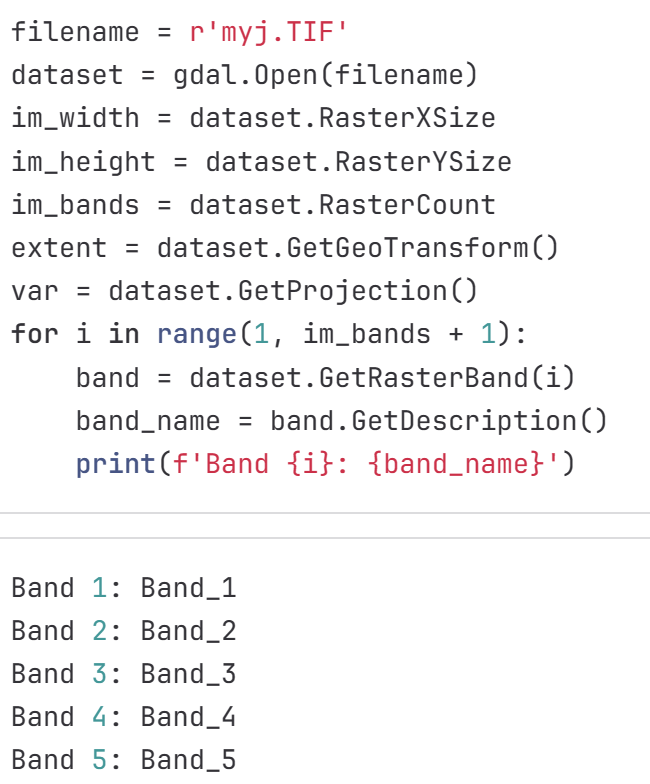


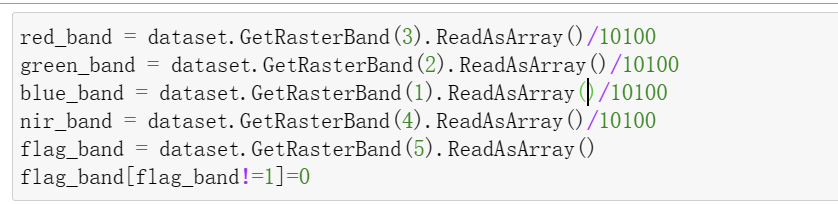
图2. Planet卫星影像

## 实验步骤与结果

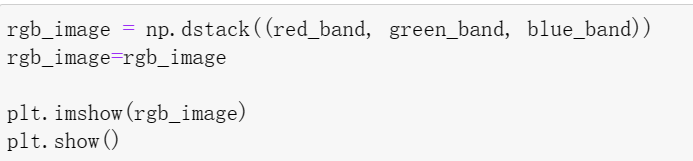
### 数据读取

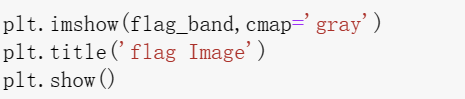
（1）以planet卫星数据为例,用GDAL 库打开 TIF 文件，检索图像的宽度和高度、波段数和范围。将波段数据存储在red\_band、green\_band、blue\_band、nir\_band 和 flag\_band 的变量中，其中flag\_band 是数据标注后的结果，为TIF文件的第 5 个波段，其中所有值不等于 1 的像素都被替换为了 0。

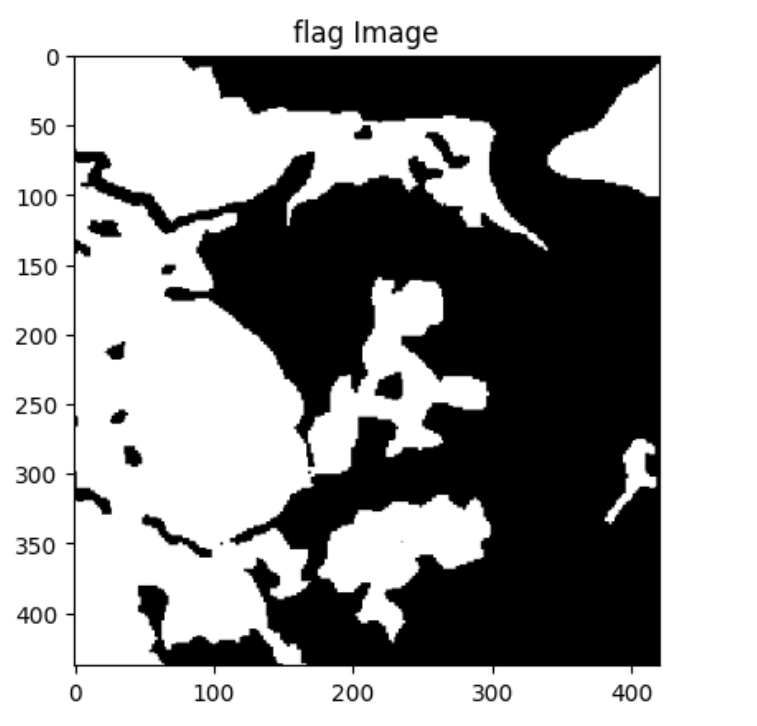




1. 影像可视化，Flag\_band图下图所示：

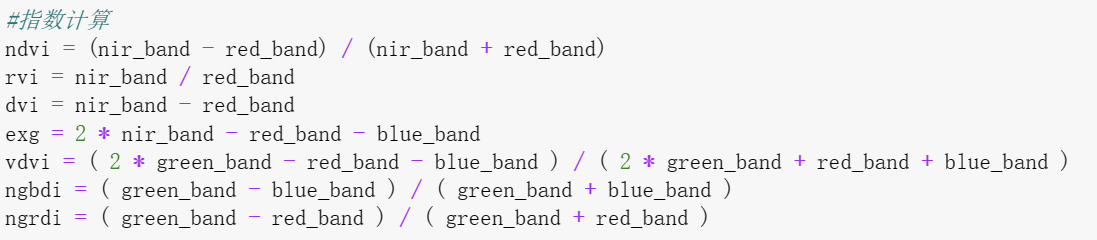


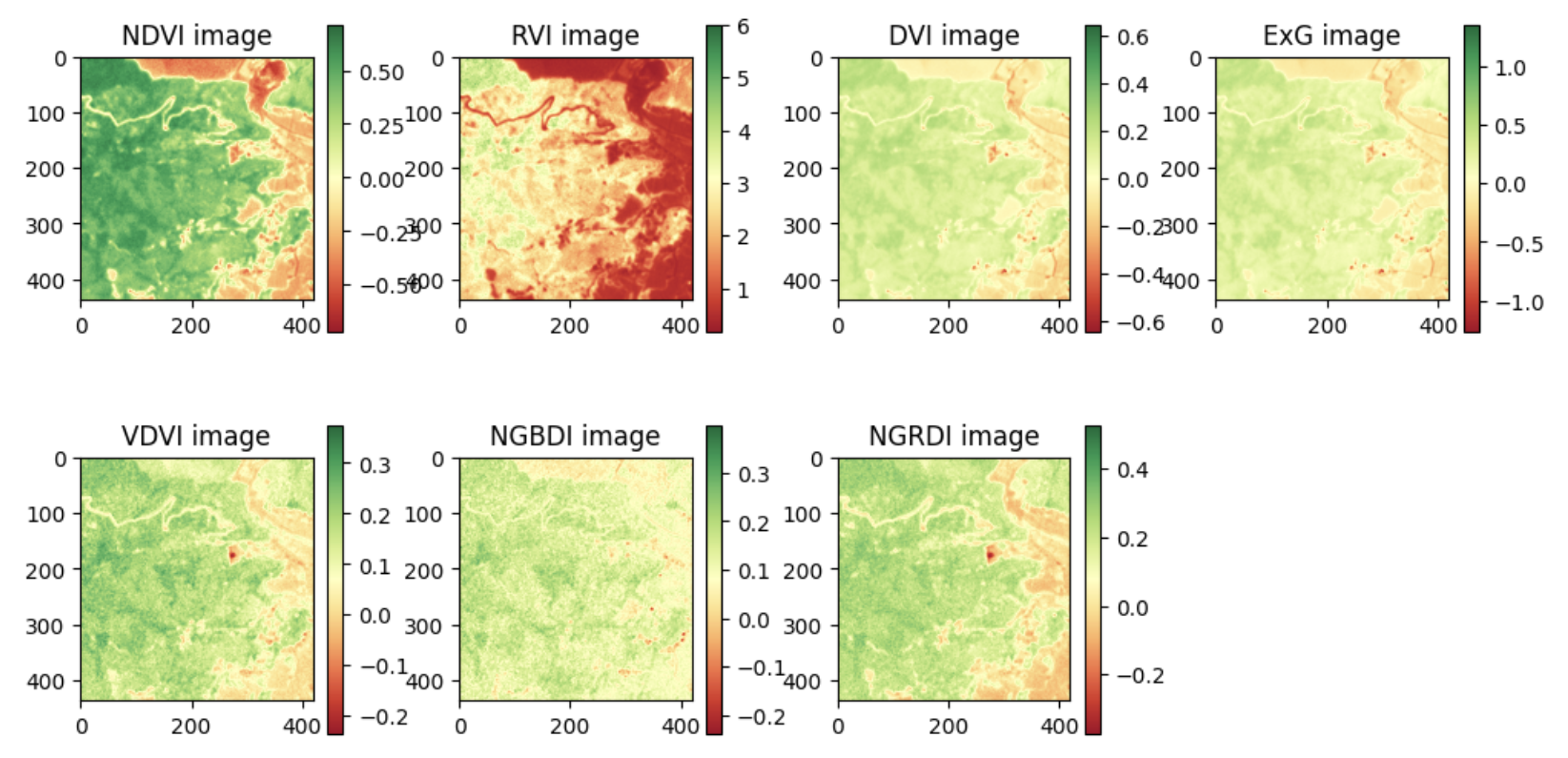




### 指数计算

1. 植被指数计算与可视化





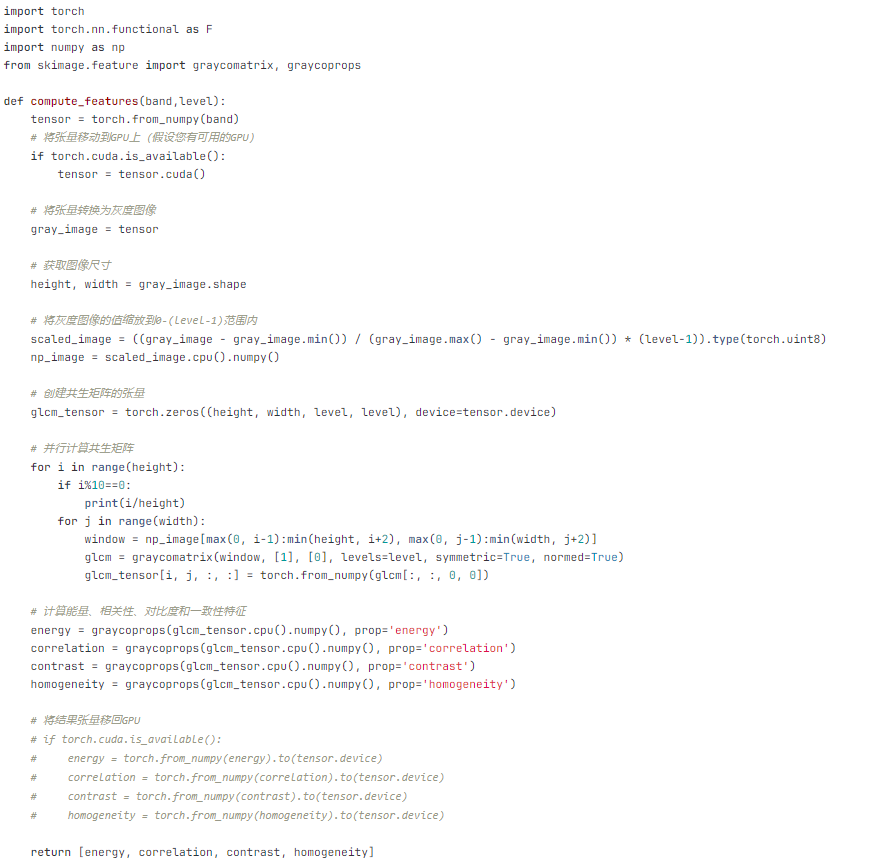
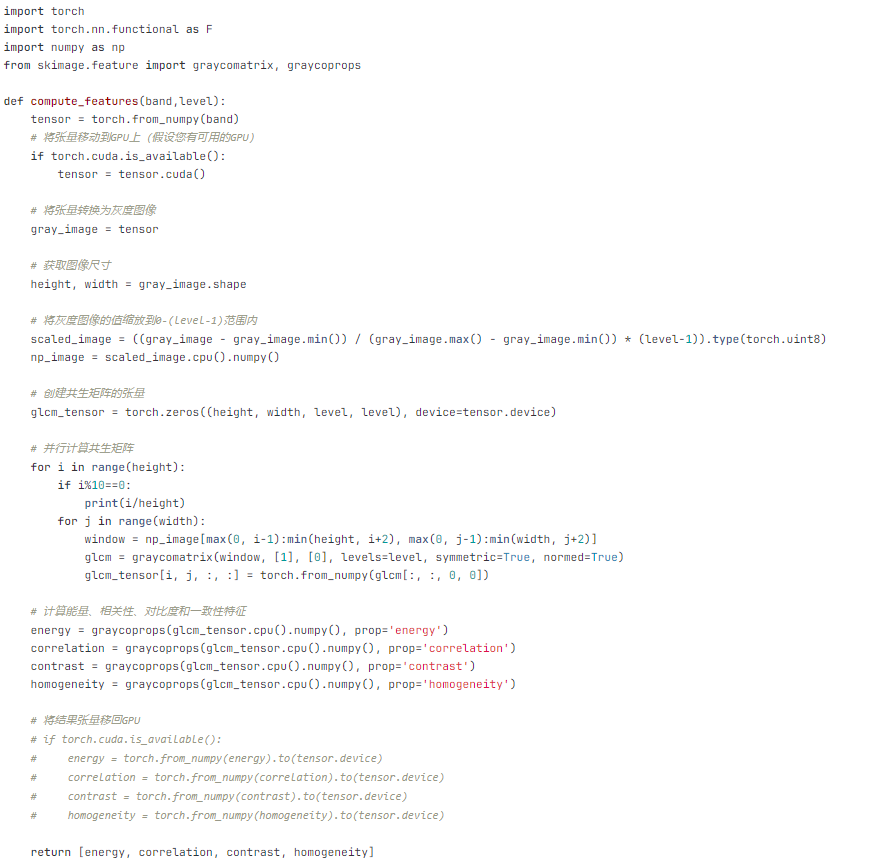
1. 纹理指数计算与可视化

遍历每个像素，计算灰度共生矩阵。提取energy, correlation, contrast, homogeneity 作为纹理特征。



1. 纹理计算的GPU加速

由于纹理计算量较大，参考GPU-Accelerated Features Extraction From Magnetic Resonance Images，利用GPU加速纹理计算。



### 基于互信息的波段筛选及归一化

1. 计算指数之间的相关系数，并筛选



1. 基于互信息筛选波段至10维



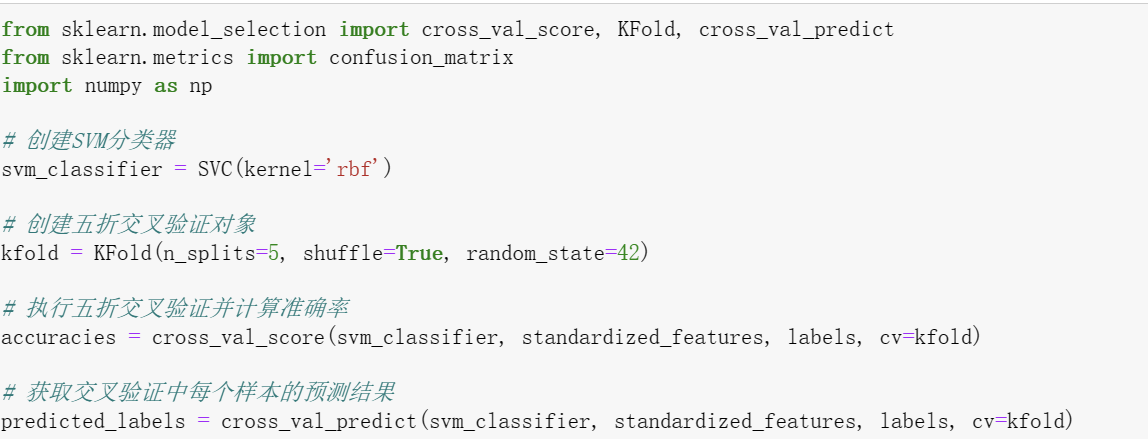
1. 指数归一化



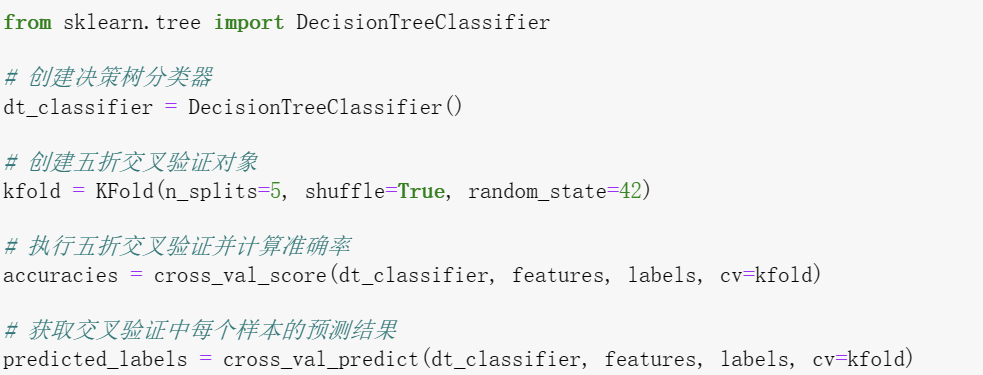
### 使用机器学习方法进行林地识别

我们分别采用了支持向量机、决策树、随机森林、逻辑回归四种机器学习方法。在训练时，对样本进行5折交叉验证，更好地体现模型泛化能力。

1. 支持向量机



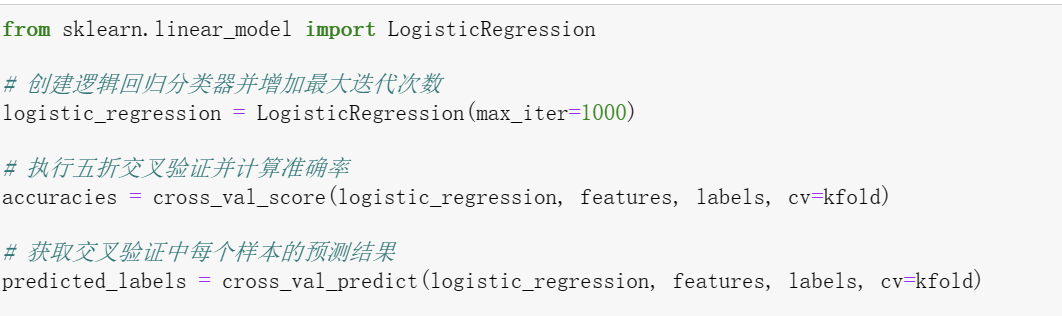
1. 决策树



1. 随机森林



1. 逻辑回归



## 实验结果与讨论

### 实验结果

1. 哨兵2号卫星影像精度评价结果：
2. 支持向量机

|  |  |  |
| --- | --- | --- |
|  | Predicted negative | Predicted positive |
| Negative | 27685 | 2394 |
| Positive | 5171 | 4749 |
| Accuracy | 0.810875 | |

1. 决策树

|  |  |  |
| --- | --- | --- |
|  | Predicted negative | Predicted positive |
| Negative | 27008 | 3072 |
| Positive | 2875 | 7045 |
| Accuracy | 0.850300 | |

1. 随机森林

|  |  |  |
| --- | --- | --- |
|  | Predicted negative | Predicted positive |
| Negative | 28838 | 1242 |
| Positive | 2686 | 7234 |
| Accuracy | 0.900950 | |

1. 逻辑回归

|  |  |  |
| --- | --- | --- |
|  | Predicted negative | Predicted positive |
| Negative | 27896 | 2184 |
| Positive | 2867 | 7053 |
| Accuracy | 0.873725 | |

1. Planet卫星影像精度评价结果：
2. 支持向量机

|  |  |  |
| --- | --- | --- |
|  | Predicted negative | Predicted positive |
| Negative | 11531 | 4433 |
| Positive | 2009 | 22027 |
| Accuracy | 0.83895 | |

1. 决策树

|  |  |  |
| --- | --- | --- |
|  | Predicted negative | Predicted positive |
| Negative | 11749 | 4331 |
| Positive | 4614 | 19306 |
| Accuracy | 0.775325 | |

1. 随机森林

|  |  |  |
| --- | --- | --- |
|  | Predicted negative | Predicted positive |
| Negative |  |  |
| Positive |  |  |
| Accuracy | 0.840825 | |

1. 逻辑回归

|  |  |  |
| --- | --- | --- |
|  | Predicted negative | Predicted positive |
| Negative | 9469 | 6611 |
| Positive | 1577 | 22343 |
| Accuracy | 0.793230 | |

### 结果分析

由于planet卫星影像分辨率较哨兵2号低，实验精度相对低。实验发现，经指数归一化后，实验精度有所上升。此外，两种影响均在随机森林方法下精度最高.......