特征分解

2022年5月13日 9:03

特征分解

项目背景：

这项工作是给定了两种不同模态的遥感图像，一种是高光谱图像，另一种是 SAR 图像，我们的任务是对这两种模态的数据进行有效融合，提高模型在高光谱遥感图像的像素级分类性能。

批注 [李1]: 两种数据的存储格式

数据介绍：

高光谱影像具有几十甚至上千个光谱波段，能够捕捉地物精细的光谱信息，光谱范围大，波段信息丰富，常用于地物的精细分类与识别。合成孔径雷达 SAR 可以获取全天候的振幅和相位信息，具有丰富的空间信息，因此，SAR 在融合的过程中可以为高光谱或者多光谱提供丰富的空间信息

明确任务：

这项工作分为三个部分，包括特征提取前的数据降维去冗余，特征提取模块，以及最后的特征融合及分类模块。我主要负责的第一个模块（数据降维去冗余）的算法设计与实现。

批注 [李2]: 数据降维方法

降维去冗余目的：

高光谱图像的波段很多，光谱分辨率比较高，SAR 图像的空间分辨率很大，通常都是几万像素。如果直接使用他们进行特征提取，那么对计算机算力的要求会很大，模型也难以训练。其次，他们是同一场景的不同模态的信息表征，虽然说他们的优势信息表征存在差异，但他们包含的共有信息（背景信息）基本一致，存在信息冗余的问题。
因此，我们需要对这两种数据先进行降维和去冗余操作，从大量数据中有效地筛选有效的数据。

痛点问题 1 解决方案：

针对信息高维且冗余问题，我们主要是借鉴了特征分解的思想，利用 NMF 非负矩阵分解方法将原始的特征分解为主特征矩阵和系数矩阵两部分。主特征矩阵中的每一行都代表了一种主成分，系数矩阵的每一列表示了每种主成分在原始特征中的占比。接着我们利用两种模态的主特征矩阵计算他们之间的相似性矩阵 S，相似性矩阵 S 衡量了两种主特征矩阵之间的整体相似性，对于 S 中的任意一个元素，它的值越大，表示这个主特征矩阵对应位置的主成分相似性越高。我们的目的是去冗余，因此首先对相似性矩阵取反（1-S），再将它作用于原本的主特征矩阵，我们就可以对两种模态之间的共有冗余信息进行抑制，最后再利用处理过的主特征矩阵与他们对应的系数矩阵相乘来重建两个模态的特征，实现对高维多模态数据的有效筛选，获得更加鲁棒的特征表征

批注 [李3]: NMF 非负矩阵分解原理以及 PyTorch 实现

批注 [李4]: PyTorch 矩阵相乘实现

痛点问题 2 解决方案：

针对高光谱和 SAR 影像的优势特征各异问题，从空间和通道两个维度出发，分别构建了空间特征重表征模块和通道特征重表征模块，用于强调各单模态的优势特征。

痛点问题 3 解决方案：

针对融合特征的决策不确定性问题（单模态特征对目标的表征能力是不同的，这就导致每个模态在决策阶段的结果可能存在差异性）。利用狄利克雷分布对各单模态特征的初级决策结果进行确定度评估，结合评估结果得到最终的决策结果。

最终结果：

最终我们取得了 84%的分类精度，较基线模型提高了 3%。

批注 [李5]: OA,AA,Kappa 计算方式

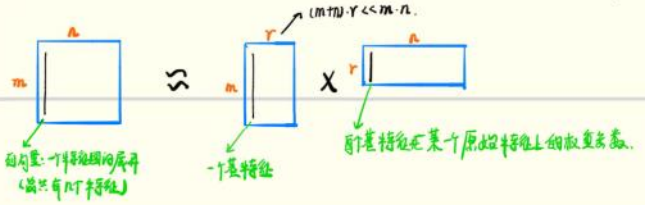
• NMF 非负矩阵分解.

• 利用矩阵分解降维的方法: PCA, LSA(潜在语义分析), SVD(奇异值分解), VQ(矢量化).

有效问题: 上述方法的共同点是将大矩阵 V 近似分解为形状的 $V=WH$ 乘积, 即 W, H 中的元素可正可负, 即使原始数据是全正的, 分解之后的矩阵中仍可能有负数, 但负值在实际问题中往往是没有意义的. (eg: 图像数据中不可能有负值的像素点)

• NMF 基本思想

思想: 对于任意给定的一个非负矩阵 V, NMF 能够找到一个非负矩阵 W 和一个非负矩阵 H, 使得 $V \approx WH$.



用基矩阵作为原始矩阵, 从而实现原始矩阵的降维, 同时有压缩空间, 减少计算资源.

上图解释: V中的每个列向量是 W中每个列向量的加权和, 加权重系数就是 H中对应的列向量.

NMF 非负矩阵分解是一个 NP 难问题, 可以转为优化问题来近似求解 W 和 H.

从图像方向: PCA 降维后得到的每个特征都是一个局部的脸部特征图像, 相当于是得几个局部人脸加在一起一张脸.
NMF 得到的是脸部基特征, 相当于取甲的眼睛, 乙的鼻子, 丙的嘴巴等, 然后拼接出一张脸.

- NMF非负矩阵分解原理

NMF问题可以看作是一个最优化问题，用梯度法有解，应用最广泛的是欧几里得距离和KL散度。

- 用梯度法求出解：

欧几里得距离：$\|V-WH\|^2=\sum_{ij}(V_{ij}-(WH)_{ij})^2$

散度：$D(V\|WH)=\sum_{ij}(V_{ij}\log\frac{V_{ij}}{(WH)_{ij}}-V_{ij}+(WH)_{ij})$

- 算法步骤：

①. 随机生成一个W矩阵

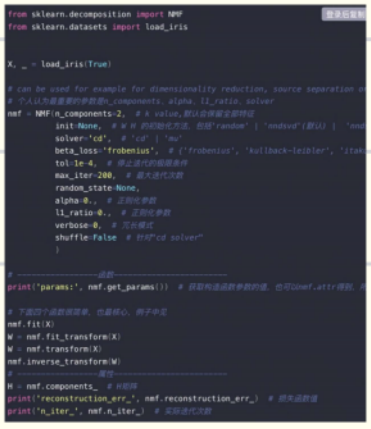
②. 固定W矩阵，按照 $H_{kj}=H_{kj}\cdot\frac{(W^TV)_{kj}}{(W^TWH)_{kj}}$（欧几里得）更新H，直到收敛（H不变or变化很小）

③. 固定H矩阵，按照 $W_{ik}=W_{ik}\cdot\frac{(VH)^T_{ik}}{(WHH^T)_{ik}}$（欧几里得）迭代更新W直到收敛。

④. 重复②，③步骤，直到损失函数趋于0。

- NMF实现

from sklearn.decomposition import NMF



- **多光谱图像及SAR图像的读取.**

已确定好的
- 多光谱数据格式：xxx.tif　　Shape：W×H×C（C：200左右.）
- SAR数据格式：xxx.tif

导入库和读取图像

```
1 import numpy as np
2 import cv2
3
4 img = cv2.imread('sar_example.tif',-1)
```

需要用到的库包括numpy和cv2，如果缺少cv2库可以conda install opencv进行安装，如果不是使用conda环境可以通过pip install opencv-python进行安装。通过cv2.imread的方式可以完整的读入16位的tif格式图片（注意替换函数调用的文件名）。

转换和存图

```
1 img_8 = (img / 256).astype('uint8')
2 cv2.imwrite('example.tif', 256-img_8)
```

登录后复制

这里使用的是numpy库中的astype进行转换。由于16位的和8位的二进制数值差倍数转换成10进制是256，所以将16位的原图片除以256就可以相应地收缩成8位下的对应值，但是这样会生成小数，于是再使用astype('uint8')进行转换即可以转成8位的。
最后再写出自己想要的格式就好了。如果是用256-img_8的话就可以得到背景较白的图片，如果直接存就可以得到背景较黑的图片，如果不想转换格式，只是想变成直接可视的8位图片，也可以直接存成tif格式。

- tif图像是16位的，可以转换为8位的可视灰度图.

- **Pytorch 矩阵乘法：**

torch.mm(mat1, mat2)：一般只用来计算二维矩阵的乘法.

mat1：(m×n)　　mat2：(n×d)　　out：(m×d)　　不支持broadcast操作.

torch.bmm(mat1, mat2)：三维带batch-size的矩阵乘法

mat1：(B×m×n)　　mat2：(B×n×d)　　out：(B×m×d)　　不支持 broadcast 操作.

torch.matmul(mat1, mat2)：多维矩阵乘法　　matmul()使用多参数的后两个维度进行矩阵乘法，其余维度可以认为是batch维度.

mat1：(1000×500×99×11)　　mat2：(500×11×99)

首先进行后两位矩阵乘法：(99×11)×(11×99)=(99×99)　　mat1的batch为(1000×500)　　mat2的batch为500

将500)广播为(1000×500)，则最终out为(1000×500×99×99)　　支持 broadcast 操作.

```python
# 读取图片——ms
ms4_tif = TIFF.open('./data/image10/ms4.tiff', mode='r')
ms4_np = ms4_tif.read_image()
print('原始ms4图的形状: ', np.shape(ms4_np))

# 读取PAN图
pan_tif = TIFF.open('./data/image10/pan.tiff', mode='r')
pan_np = pan_tif.read_image()
print('原始pan图的形状;', np.shape(pan_np))

# 读取标签
# label_np = np.load("./data/image6/label6.npy")
label_dict = h5py.File("./data/image10/label10.mat")
label_np =np.array(label_dict['label'][:]).T
label_dict.close()
print('label数组形状: ', np.shape(label_np))
```

torch.mul(mat1, mat2)：逐元素相乘

• OA、AA、kappa计算方法.

• 混淆矩阵：

	数量	预测 类1	预测 类2	预测 类3
真实	类1	43	25	76
真实	类2	...	...	...
真实	类3	...	39	46

混淆矩阵可以直观的看出每个地物被正确分类和错误分类的个数，但混淆矩阵并不能一眼看出分类精度的好坏.

• 整体分类精度OA：正确分类个数与总个数的比值　　易受到大数据量的类别的影响，不能很好的表征每个地物.

$$OA = \frac{TP + TN}{TP + FP + TN + FN}$$

• kappa系数：一致性检验.

$$kappa = \frac{OA - pe}{1 - pe}$$

$$pe = \frac{(TP+FN)\times(TP+FP)+(TN+FP)\times(TN+FN)}{N^2}$$

- 多光谱和SAR图像详细信息

  - 多光谱图像:

    ①.能在可见光到短波红外的光谱区间连续成像

    ②.以数十甚至数百个连续窄波段对目标区域同时成像

    ③.多光谱图像集图像信息与光谱信息于一身，图像信息可以反映物体的形状、大小甚外部结构。而不同物质对光谱的吸收程度是不同的，在某个特定波长下，图像会对某个物质有较好的反映，因此光谱信息更能反映物体的内部结构和物质差异。

  - SAR图像:

    - SAR图像上的信息是对雷达波束的反映，雷达发送连续的无线电脉冲照亮目标场景，并且接收并记录每个脉冲的回波。
    - SAR图像上色调的变化主要取决于目标的后向散射强度。每一个接收到的回波被转换为电信号，以特定的灰度色调记录在光学胶片上。

      目标的后向散射主要受目标的表面粗糙程度影响。