

文章编号: 2095-2163(2022)12-0128-06

中图分类号: TP751

文献标志码: A

基于 RNN 和多重注意力机制的高光谱图像分类

郑凯东, 李阿莹

(西安石油大学 计算机学院, 西安 710065)

摘要: 现今伴随着网络模型的愈加复杂, 高光谱图像分类的实时性较差。考虑到高光谱图像本质上是一类序列数据, 本文提出 LZ-RNN 模型, 通过引用双向循环神经网络(BRNN)来处理序列数据, 减少时间成本。引入多重注意力机制来提升模型的精度, 通过三分支捕获跨纬度交互来计算注意力权重。实验表明, 本文所提方法在 Pavia Centre 和 Botswana 两个数据集上取得优异的效果, 与其他主流算法相比, 在较短的时间内达到了较高的精度。

关键词: 高光谱图像分类; 双重循环神经网络; 注意力机制

Hyperspectral images classification based on RNN and multiple attention mechanisms

ZHENG Kaidong, LI Aying

(College of Computer Science, Xi'an Shiyou University, Xi'an 710065, China)

[Abstract] Nowadays, with the increasingly complex network models, the real-time performance of hyperspectral images classification is poor. Based on the fact that hyperspectral images are essentially a type of sequence data, this paper proposes the LZ-RNN model, which reduces the time cost by referring to the bidirectional recurrent neural network (BRNN) to process the sequence data. A multi-attention mechanism is introduced to improve the accuracy of the model, and the attention weights are calculated by capturing cross-latitude interactions through three branches. Experiments show that the method proposed in this paper achieves excellent results on the Pavia Centre and Botswana datasets, and achieves higher accuracy in a shorter time than other mainstream algorithms.

[Key words] hyperspectral images classification; bidirectional recurrent neural network; attention mechanism

0 引言

高光谱遥感是指用大量狭窄的电磁波通道获取地物的空间、辐射和光谱三重信息的技术^[1]。高光谱成像遥感技术可以获取地物的光谱、辐射和空间信息, 是通过电磁波谱中可见光、红外线的波段范围获取地物的数据信息, 构成图像数据立方体^[2]。

在研究早期, 科研人员主要把关注点放在光谱特征的提取上, 相关的方法包括主成分分析、独立成分分析、线性判别式分析等, 利用这些方法来提取高光谱图像的光谱特征, 并使用分类器进行分类, 比较常用的有支持向量机(Support Vector Machine, SVM)^[3]、随机森林^[4]、多项式逻辑回归分类^[5]等方法。

卷积神经网络(Convolutional Neural Network, CNN)是深度学习的代表算法之一, 已有众多学者在应用 CNN 及改良的网络结构对高光谱图像分类进行研究, 如 Chen 等人^[6]利用高光谱图像的像素信

息作为输入, 显著提高了结果精度。然而由于高光谱图像具有二维空间信息和一维光谱信息, 因此在利用 CNN 模型对高光谱图像分类时会因为这些冗余信息而对精度造成影响^[7]。在此基础上, Xu 等人^[8]提出了 3D Octave 卷积模型, 能够从高光谱图像中提取空间-光谱特征, 并引入注意力机制来加强重要区域的空间-光谱信息的提取, 但该模型在网络训练方面耗费了更多的时间成本。由于高光谱图像的光谱数据本质上是一类序列数据^[9], Venkatesan 等人^[10]首次把循环神经网络(Recurrent Neural Network, RNN)引入到高光谱图像分类中, 并提出了一种激活函数和参数校正函数, 用于分析高光谱图像中的数据序列。Zhou 等人^[11]提出了一种基于光谱-空间长短时记忆网络的高光谱图像分类方法, 大大减少了网络模型的训练时间, 但精度略逊于改良后的 CNN 模型。

基于以上研究, 本文提出了一种基于 RNN 及多重注意力机制融合网络(LZ-Bidirectional Recurrent

作者简介: 郑凯东(1964-), 男, 副教授, 主要研究方向: 图形学与虚拟现实、程序设计、计算机基础教育; 李阿莹(1996-), 女, 硕士研究生, 主要研究方向: 图形图像处理、智能计算与可视化。

收稿日期: 2022-03-21

哈尔滨工业大学主办 ◆ 系统开发与应用

Neural Network ,LZ-BRNN),用双向循环神经网络 BRNN 来处理高光谱图像的序列数据,减少时间成本,并引入三重注意力机制来提升网络模型的精确度。

1 本文算法

本文提出一种基于 RNN 及多重注意力机制融

合的网络模型(LZ-BRNN),网络框架流程如图 1 所示。图 1 中,模型用双向循环神经网络(BRNN)来处理高光谱图像的序列数据,减少时间成本,同时引入三重注意力机制,通过三支结构捕获跨维度交互来计算注意力权重,用来提高整个网络模型分类的精确度。

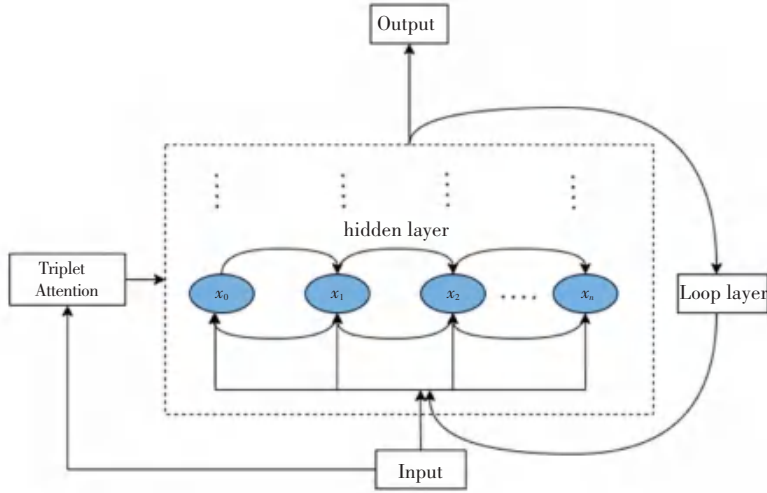


图 1 LZ-BRNN 网络框架流程图

Fig. 1 LZ-BRNN network framework flowchart

1.1 双向循环神经网络

BRNN 的基本思想是提出每一个训练序列向前和向后分别是 2 个循环神经网络(RNN),而且这 2 个都连接着一个输出层。该结构给输出层输入序列中提供每一个点的完整的过去和未来的上下文信息。文中给出 BRNN 隐藏层和输出层的计算公式分别如下:

$$S_t = f(U_1 \cdot X_t + W_1 \cdot X_{t-1}) \quad (1)$$

$$S'_t = f(U_2 \cdot X_t + W_2 \cdot X_{t+1}) \quad (2)$$

$$O_t = g(V_1 \cdot S_t + V_2 \cdot S'_t) \quad (3)$$

1.2 Triplet Attention 模型

深度学习中的注意力机制(Attention Mechanism, AM)是根据重要性给图像不同部分分配不同的权重,有助于忽略掉不重要的特征信息。本文引用三重注意力机制(Triplet Attention, TA),由 3 个分支组成。其中,2 个用于捕获通道 C 和空间 H 或 W 之间的跨维交互,第 3 个分支发挥着类似卷积注意力模块(CBAM)的作用。TA 模型设计原理如图 2 所示。

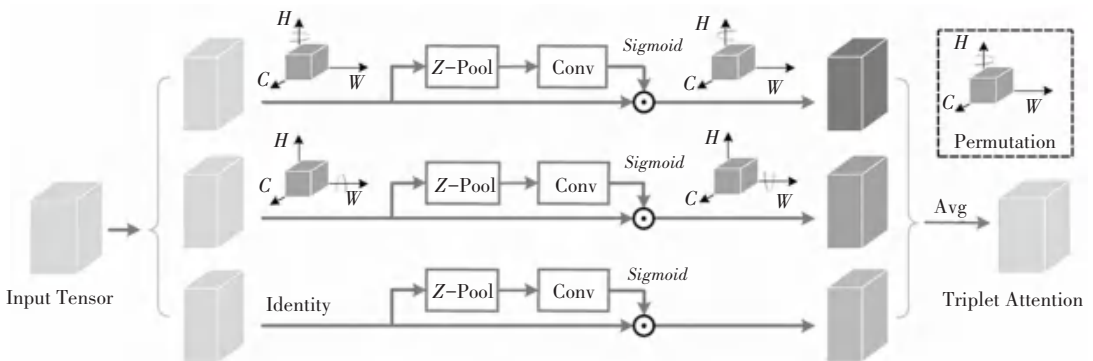


图 2 Triplet Attention 模型原理图

Fig. 2 Schematic diagram of the Triplet Attention model

2 实验结果与分析

所有实验的硬件设备均在 Intel(R) Core(TM) i7-10870H CPU @ 2.20 GHz 2.21 GHz 处理器, Nvidia GeForce RTX 2070MQ 显卡的系统上运行的。同时,主机选用 Windows10 操作系统,运行环境为 Python3.6, Torch1.10.0+CU102。

2.1 实验数据集描述

为了验证本文方法的有效性,在高光谱图像的 2 个公开数据集、即 Botswana 和 PaviaC 上进行实验。其中,Botswana 数据集像素分辨率为 30 m,包含 145 个可用于高光谱图像分类的波段,数据涵盖了 14 种地物类别,分别是 Water、Hippo grass、Floodplain grasses1 等。PaviaC 数据集经过处理后共有 102 个可用波段,数据包含了 9 种地物类别,分别是 Water、Trees、Asphalt 等。

2.2 实验参数设置以及对照实验

本次实验的主要评价指标有:迭代次数

(*epoch*)、训练集划分(*training_sample*)以及学习率(*learning_rate*),经过多次实验,最终将测试集和训练集各取样本数据的 50%,*epoch* 取 200,*learning_rate* 设置为 0.001,在实验中能取得更好的结果。

本文实验主要与传统分类方法、如 SVM^[12],以及近年来在高光谱图像分类问题中常用的算法进行比较,如 RNN^[13]、3D-FCN^[14]、3D-CNN^[15]。根据高光谱常用的衡量精度指标:总体分类精度(Overall Accuracy, *OA*)以及 *Kappa* 系数进行分类精度比较^[16],同时还要计算每次分类所用时间,用来验证分类的实时性。

2.3 实验结果与分析

在参数设定的基础上,分别在 Botswana、PaviaC 数据集上各进行 10 次实验,最终实验结果取平均值,表 1 为 Botswana 数据集实验分类结果。

表 1 Botswana 数据集分类结果

Tab. 1 Classification results of the Botswana dataset

class	SVM	RNN	3D-FCN	3D-CNN	LZ-RNN
Water	1.000	0.996	1.000	1.000	1.000
Hippo grass	0.971	0.990	1.000	0.970	0.990
Floodplain grasses1	0.967	0.992	0.992	1.000	1.000
Floodplain grasses2	0.918	0.968	0.982	1.000	1.000
Reeds	0.882	0.903	0.938	0.956	0.960
Riparian	0.691	0.870	0.933	0.945	0.930
Firescar	0.985	1.000	1.000	1.000	1.000
Island interior	0.771	0.990	0.969	1.000	0.994
Acacia woodlands	0.850	0.943	0.975	0.981	0.985
Acacia shrublands	0.702	0.963	0.956	0.992	0.995
Acacia grasslands	0.773	0.974	0.971	0.993	0.998
Short mopane	0.994	0.920	0.910	0.989	1.000
Mixed mopane	0.908	0.927	0.924	0.993	0.984
Exposed soils	0.854	0.979	0.989	1.000	1.000
<i>OA</i> / %	87.13	95.57	96.55	98.65	98.82
<i>Kappa</i>	0.860	0.952	0.963	0.985	0.986
<i>Time</i> / s	65.3	46.2	218.0	572.0	74.0

由表 1 可以得出,在 Botswana 数据集上,本文所提出的方法在精度上高于 SVM 和普通的循环神经网络以及 3D-FCN 网络模型,在 *OA* 和 *Kappa* 值上高于三者,比改进后的 3D-CNN 略高一点。在分类耗费时间成本上,RNN 用时最短,也表现出了 RNN

处理时序数据的优越性,随着网络层数的增加,3D-CNN 虽然在精度上要高于 3D-FCN、但用时更长,LZ-RNN 在 RNN 基础上进行改良,使得整体时间成本仍在可接受范围内。图 3 为不同模型在 Botswana 数据集上的分类结果。

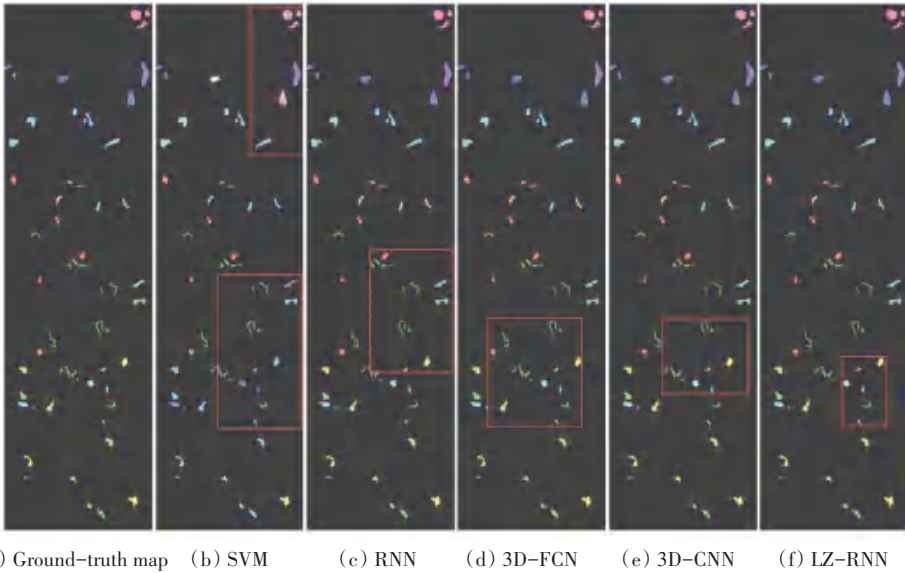


图 3 不同模型在 Botswana 数据集分类结果图

Fig. 3 Classification results graph of different models in the Botswana dataset

图 3(a) 是在 Botswana 数据集取波段为 (75, 33, 15) 得到的真值图。用红色选框标注的是通过模型分类与真值有偏差的区域, 随着模型不断优化, 出错区域明显越来越少, 这也与表 1 得出的数据相吻合。

第二组实验是在 Pavia Center 数据集上进行, 由于该数据集像素点远高于 Botswana 数据集, 分类所需时间也有所增加, 在本次实验中将迭代次数 *epoch* 设为 100, 其他参数不变, 表 2 为 PaviaC 数据集实验分类结果。

表 2 PaviaC 数据集分类结果

Tab. 2 Classification results of the PaviaC dataset

class	SVM	RNN	3D-FCN	3D-CNN	LZ-RNN
Water	0.985	0.999	1.000	0.996	0.999
Trees	0.863	0.915	1.000	0.970	0.975
Asphalt	0.368	0.823	0.992	0.915	0.930
Self-Blocking Bricks	0.302	0.502	0.982	0.800	0.852
Bitumen	0.439	0.841	0.938	0.935	0.959
Tiles	0.919	0.866	0.933	0.980	0.970
Shadows	0.887	0.832	1.000	0.970	0.953
Meadows	0.948	0.896	0.969	0.996	0.998
Bare Soil	0.714	0.888	0.975	0.999	0.998
<i>OA</i> / %	0.916	96.64	96.55	98.12	98.68
<i>Kappa</i>	0.879	0.953	0.963	0.973	0.981
<i>Time</i> / min	15.60	13.33	66.75	133.50	20.08

从表 2 可以得出, 在 PaviaC 数据集上, 本文提出的 LZ-RNN 模型在 *OA* 值和 *Kappa* 值上高于其他算法。由于 PaviaC 图像的像素点远高于 Botswana 数据, 但 RNN 模型仍能在较短的时间内进行分类, 且在精度上略高于 CNN 模型。而通过改良得到的 LZ-RNN 提升了原模型的精度。图 4 为不同模型在 PaviaC 数据集上的分类结果。

12) 得到的真值图。用红色选框标注的是通过模型分类与真值有偏差的区域。随着模型的不断优化, 出错区域明显越来越少, 图像上冗余斑点也越来越少。实验证明, 本文提出的 LZ-RNN 模型对于大部分地物类别的分类均取得了良好的效果, 有多个地物类别的效果达到 99.9%, 既保证了精度, 又将时间成本控制在可接受的范围内。

图 4(a) 是在 PaviaC 数据集取波段为 (55, 41,

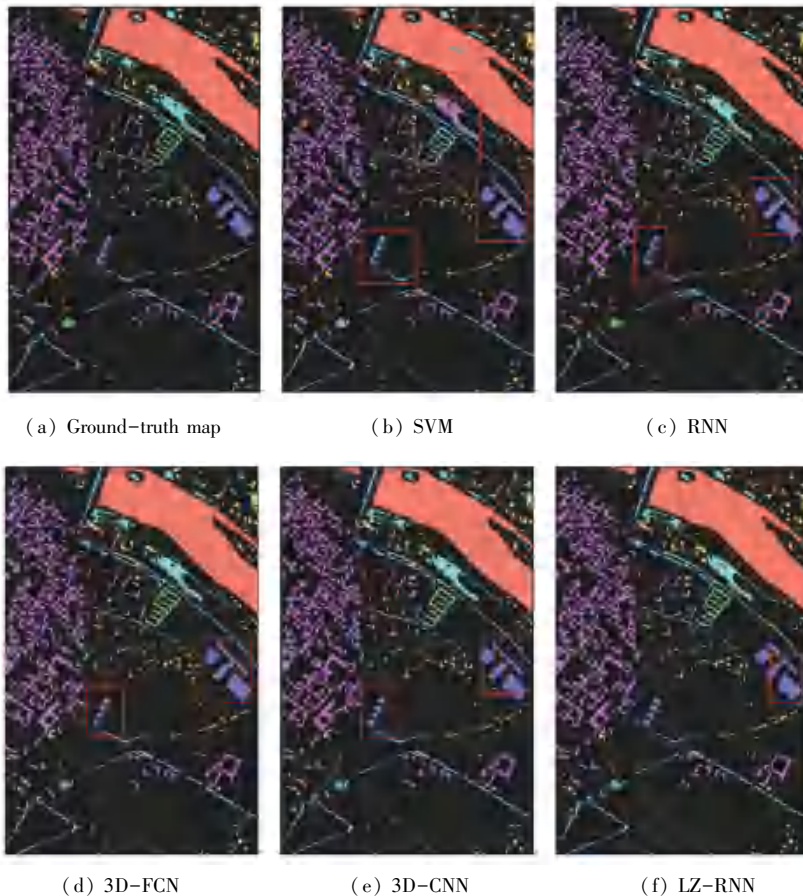


图4 不同模型在 PaviaC 数据集分类结果图

Fig. 4 Classification results graph of different models in the PaviaC dataset

3 结束语

本文提出了一种基于循环神经网络和多重注意力机制的高光谱图像分类模型,网络结构简单且有效,模型用双向循环神经网络(BRNN)来处理高光谱图像的序列数据,降低了时间成本,引入多重注意力机制来提升分类精度,通过实验对比,证明本文提出的方法能够在较短的时间内完成高光谱图像分类任务,且展现出优秀的分类性能。

参考文献

- [1] 童庆禧,张兵,郑芬兰. 高光谱遥感—原理、技术与应用[M]. 北京:高等教育出版社,2006.
- [2] 杨红艳,杜健民. 高光谱遥感图像波段选择研究进展综述[J]. 计算机工程与应用,2022,58(10):1-12.
- [3] 王立国,赵亮,刘丹凤. SVM在高光谱图像处理中的应用综述[J]. 哈尔滨工程大学学报,2018,39(06):973-983.
- [4] HAM J, CHEN Yangchi, CRAWFORD M M, et al. Investigation of the random forest framework for classification of hyperspectral data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2005,43(3):492-501.
- [5] LI Jun, BIOUCAS - DIAS J M, PLAZA A. Semisupervised hyperspectral image segmentation using multinomial logistic

- regression with active learning[J]. IEEE Transactions on Geoscience and Remote Sensing, 2010,48(11):4085-4098.
- [6] CHEN Yushi, LIN Zhouhan, ZHAO Xing, et al. Deep learning-based classification of hyperspectral data[J]. IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, 2017,7(6):2094-2107.
- [7] HE MINGYI, LI BO, CHEN HUAHUI. Multi-scale 3D deep convolutional neural network for hyperspectral image classification[C]//2017 IEEE International Conference on Image Processing (ICIP). Beijing:IEEE, 2017:3904-3908.
- [8] XU Tang, MENG Fanbo, ZHANG Xiangrong. Hyperspectral image classification based on 3-D octave convolution with spatial-spectral attention network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020,59(3):2430-2447.
- [9] LIANG Lianhui, LI Jun, ZHANG Shaoquan. Hyperspectral images classification method based on 3D octave convolution and Bi-RNN attention network[J]. Acta Photonica Sinica, 2021,50(9):0910001.
- [10] VENKATESAN R, PRABU S. Hyperspectral image features classification using deep learning recurrent neural networks[J]. Journal of Medical Systems, 2019,43(7):1-10.
- [11] ZHOU Feng, HANG Renlong, LIU Qingshan, et al. Hyperspectral image classification using spectral-spatial LSTMs[C]//CCF Chinese Conference on Computer Vision. Singapore: Springer, 2017:577-588.

(下转第137页)