



计算机应用  
*Journal of Computer Applications*  
ISSN 1001-9081, CN 51-1307/TP

## 《计算机应用》网络首发论文

题目：数据异常情况下遥感影像时间序列分类算法  
作者：任媛媛，汪传建  
收稿日期：2020-09-07  
网络首发日期：2020-10-26  
引用格式：任媛媛，汪传建. 数据异常情况下遥感影像时间序列分类算法[J/OL]. 计算机应用. <https://kns.cnki.net/kcms/detail/51.1307.TP.20201026.1335.002.html>



**网络首发：**在编辑部工作流程中，稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定，且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式（包括网络呈现版式）排版后的稿件，可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定；学术研究成果具有创新性、科学性和先进性，符合编辑部对刊文的录用要求，不存在学术不端行为及其他侵权行为；稿件内容应基本符合国家有关书刊编辑、出版的技术标准，正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性，录用定稿一经发布，不得修改论文题目、作者、机构名称和学术内容，只可基于编辑规范进行少量文字的修改。

**出版确认：**纸质期刊编辑部通过与《中国学术期刊（光盘版）》电子杂志社有限公司签约，在《中国学术期刊（网络版）》出版传播平台上创办与纸质期刊内容一致的网络版，以单篇或整期出版形式，在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊（网络版）》是国家新闻出版广电总局批准的网络连续型出版物（ISSN 2096-4188，CN 11-6037/Z），所以签约期刊的网络版上网络首发论文视为正式出版。

# 数据异常情况下遥感影像时间序列分类算法

任媛媛<sup>1</sup>, 汪传建<sup>2\*</sup>

(1.石河子大学 信息科学与技术学院, 新疆 石河子 832000;

2.安徽大学 互联网学院, 安徽 合肥 230039)

(\*通信作者电子邮箱 wej\_si@ahu.edu.cn)

**摘 要:** 针对时序遥感图像数据异常时卷积神经网络分类性能较差的问题, 提出了一种多模式与多单模相结合的网络结构。该网络首先通过多元时序模型和单变量时间序列模型对多维时间序列进行多尺度特征提取, 并基于像素空间坐标信息, 通过自动编码形式完成遥感图像的时空序列特征构建。最后通过全连接层和 softmax 函数实现分类。在数据异常(数据缺失和数据扭曲)的情况下, 提出的算法和一维卷积神经网络(1D-CNN)、多通道深度神经网络(MCDNN)、时序卷积神经网络(TSCNN)和长短期记忆(LSTM)等通用时间序列遥感影像分类算法进行分析比较。实验结果表明, 利用端到端的多模式与多单模式架构融合的网络在数据异常的情况下分类精度最高, F1 值达到了 93.40%。

**关键词:** 遥感影像; 时序数据; 卷积神经网络; 分类; 数据异常

**中图分类号:** TP391.4

**文献标志码:** A

## Remote sensing time-series images classification algorithm with abnormal data

REN Yuanyuan<sup>1</sup>, WANG Chuanjian<sup>2\*</sup>

(1. School of Information Science and Technology, Shihezi University, Shihezi Xinjiang 832000, China

2. School of Internet, Anhui University, Hefei Anhui 230039, China)

**Abstract:** Concerning the problem of convolutional neural networks making poor classification performance when confronted with abnormal time series remote sensing image data, a novel end to end network based on the integration of multi-mode and multi-single-mode architecture was introduced. Multi-scale features of the multi-dimensional time series could be extracted by network from multivariate time series model and the univariate time series model. Then spatio-temporal sequence feature was completely constructed by automatic coding based on the pixel space coordinate information. Finally, the classification was implemented by fully connected layer and the softmax function. In the case of data anomaly (data loss and data distortion), the proposed model algorithm was compared with other general time series remote sensing image classification algorithms such as 1D Convolutional Neural Networks (1D-CNN), Multi-Channels Deep Neural Network (MCDNN), Time Series Convolutional Neural Networks (TSCNN) and Long Short-Term Memory (LSTM) network simultaneously. The experimental results show that the novel fusion network using the end-to-end training method had higher classification accuracy to alleviate abnormal data problem, and the F1 value reached 93.40%.

**Keywords:** remote sensing image; time serials; convolutional neural networks; classification; data abnormal

### 0 引言

地球陆地是人类赖以生存和发展的物质基础, 获取陆地表面各种资源及其变化数据是人类认识和利用陆地表面的前提。遥感技术通过人造地球卫星、航空、无人机等平台上的遥测仪器对地球表面实施感应遥测, 产生了海量的不同空间分辨率和时间分辨率的遥感数据, 这些遥感数据蕴含着地球表面各类地物在空间和时间上的巨量信息。如何精确的区分各类地物是研究陆地表面的基础性工作。遥

感影像分类就是利用计算机通过对遥感影像中各类地物的光谱信息和空间信息进行分析, 选择特征, 将影像中各个像元按照某种规则或算法划分不同的类别, 然后获得遥感影像中与实际地物的对应信息<sup>[1]</sup>。在遥感影像分类方面, 对于常用的单时相遥感影像分类技术, 只能利用波谱反射率的异同来分类, 有效信息获取不足使得分类结果存在许多问题, 比如混合像元的判定、“同物异谱, 异物同谱”现象的存在、分类精度在时间和空间上难以保持稳定等。众所周知, 地表事件的发生和地物的演化, 是随着时间的推

收稿日期: 2020-09-07; 修回日期: 2020-10-21; 录用日期: 2020-10-22。

基金项目: 国家重点研发计划(2017YFB0504203)

**作者简介:** 任媛媛(1996—), 女, 陕西岐山人, 硕士研究生, CCF 会员, 主要研究方向: 机器学习、遥感数据处理; 汪传建 (1977—), 男, 安徽怀宁人, 教授, 博士, CCF 会员, 主要研究方向: 机器学习、遥感影像处理。

进而进行的,这使得地表覆被在时间上具有一定的变化规律。遥感影像反映出,大多数地物的波谱反射率经过长期观测可以重复获取,且由波谱反射率组成的时间序列表现出一定的周期性。遥感影像时间序列由于蕴含显著的物候信息,使得不同地物拥有不同的光谱反射率轨迹,利用这些数据特点能够更加有效地区分地物类型,时间序列在土地覆盖分类方面已经被证明优于单时相分类<sup>[2]</sup>。

随着深度学习的兴起,学者们提出了很多基于深度学习的遥感影像时间序列分类算法,常见的用于时间序列建模的深度学习框架包括长短期记忆网络(Long Short-Term Memory, LSTM)、门循环单元(Gate Recurrent Unit, GRU)、一维/三维卷积神经网络(1D/3D Convolutional Neural Networks, 1D/3D CNN)和时间变换网络(Temporal Transformer Networks, TTN)等。利用 LSTM 模型<sup>[3]</sup>的序列数据建模能力,文献[4]提出两种 LSTM 模型(深度栈式 LSTM 和深度双向 LSTM)用于 MODIS 时间序列分类,分类精度优于随机森林(Random Forest, RF)和支持向量机(Support Vector Method, SVM)分类方法。卷积神经网络自从被提出之后被广泛应用于各种遥感任务,包括陆地覆盖分类<sup>[5-6]</sup>、影像分割<sup>[7]</sup>、目标检测<sup>[8]</sup>、丢失数据重建<sup>[9]</sup>等。在这些遥感任务中,卷积神经网络通过对影像  $x$  维和  $y$  维进行卷积操作来充分影像数据的空间结构,也就是说卷积操作主要运用在影像的光谱域或者波段域,但没有考虑影像的时间特征。为了处理遥感影像时间序列数据,学者们又将卷积神经网络应用于遥感影像时间序列分类。文献[10]利用一维卷积神经网络<sup>[11-12]</sup>可有效描述一维序列数据的特性,提出基于一维卷积神经网络的农作物分类方法,试验结果表明该方法分类性能优于 LSTM 和传统分类方法。文献[13]多尺度卷积神经网络(Multi-scale Convolutional Neural Networks, MCNN)提取不同比例和频率的特征,将特征提取和分类合并在一个框架。文献[14]提出一种端到端的遥感影像时间序列分类深度学习算法,以一维卷积为核心构建三个组件分别描述遥感影像各像素的波段间依赖性、全波段全卷积特征以及邻居像素特征来实现像素级分类,并在 2017 年时间序列陆地覆盖分类挑战赛<sup>[15]</sup>(TiSeLaC 2017)获得第一名。由于深层 LSTM 形式的深层递归神经网络在这种像素较少的情况下,训练并不是很成功<sup>[14]</sup>。文献[16]利用多通道深度卷积神经网络(Multi-Channels Deep convolutional Neural Networks, MCDNN)体系结构通过在输入的每个维度上独立(即并行)应用卷积,来利用多模态时间序列数据的不同特征之间的假定独立性。但指标值与反射率的测量值相关,文献[17]提出的时序卷积神经网络(Time Series Convolutional Neural Networks, TSCNN)通过共同训练多元时间序列以进行更好的特征提取,从而获得了更好的分类结果。

现有的遥感影像分类算法严重依赖数据本身,而数据本身存在数据缺失、数据扭曲等数据异常问题。且现有分

类算法大多针对某一个具体研究区,同类作物物候期相对变化很小,同类作物的时间序列数据组成像素对应的归一化植被指数(Normalized Vegetation Index, NDVI)等特征曲线高度相似,直接输入到分类算法中。实际上,同一种作物的物候期大体相似,但随年份和区域的不同呈现出多样性,因此对应的 NDVI 等特征曲线也呈现出多样性特征。这种遥感影像中提取的与物候期相对应的特征曲线多样性可以通过扭曲操作来得到,即对特征曲线进行时间轴和特征值轴两个维度进行适度的平移、拉伸、压缩等操作,可以得到同一种作物的丰富多样的特征曲线。目前尚无学者关注数据扭曲情形下遥感影像时间序列分类算法的鲁棒性问题。文献[4]中提出遥感影像时间序列数据存在扭曲的现象,但采取了数据增强的方式进行解决,没有从模型和算法的角度来研究提高算法鲁棒性的方法。

为解决以上问题,本文的主要贡献有:(1)提出了一种端到端的多模式与多单模式架构融合的网络,解决数据异常情况下时间序列分类问题;(2)在基于卷积神经网络分类器中引入了能够提取单个波段和多波段特征的结构,在网络中结合多种特征进行时间序列分类,不仅仅是单靠其中一种;(3)将模型应用到遥感影像时间序列分类中,取得了较好的效果。通过 1D-CNN、MCDNN、TSCNN、LSTM 和本文的网络进行对比,对 TiSeLaC 遥感影像时间序列数据集进行分类,分别对数据缺失和数据扭曲两种数据异常情况进行分析。实验结果表明端到端的多模式与多单模式架构融合的网络在数据异常的情况下分类精度较高,F1 值达到了 93.40%。

## 1 相关工作

卷积神经网络随着深度学习的兴起,逐渐应用于遥感影像时间序列分类任务中。时间序列数据分类问题中,训练样本通常由一系列已标记的单序列时间序列样本构成,假设单变量时间序列  $\mathbf{X} = [x_1, x_2, \dots, x_T]$  是有序的一组实数值,其中  $\mathbf{X}$  的长度为  $T$ 。 $M$  维时间序列  $\mathbf{X}_M = [\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^M]$  是由  $M$  个不同的单变量时间序列构成的多元时间序列,其中  $\mathbf{X}^i \in R^T$ 。数据集  $D = \{(\mathbf{X}_1, \mathbf{Y}_1), (\mathbf{X}_2, \mathbf{Y}_2), \dots, (\mathbf{X}_N, \mathbf{Y}_N)\}$  是  $(\mathbf{X}_i, \mathbf{Y}_i)$  的集合,其中  $\mathbf{X}_i$  可以是单变量或多元时间序列,  $\mathbf{Y}_i$  是其对应的 one-hot 编码标记向量。对于包含  $K$  个类别的数据集,时间序列  $\mathbf{X}_i$  对应标记  $\mathbf{Y}_i \in \{0, 1\}^K$ ,表示训练样本中有  $K$  个类别,  $\mathbf{Y}_i$  中仅有一个元素取 1,取 1 元素的索引为该时间序列的类别,其余元素均为 0。多波段遥感影像提供了同一个位置上的多个不同光谱波段的强度,各个波段影像之间存在互补性和冗余性<sup>[18]</sup>。在遥感影像时间序列中,  $M$  维时间序列即是  $M$  个波段组成的多元时间序列。则数据集中地面特



征可以表示为  $\mathbf{X}_M \in R^{N \times T \times M}$ , 其中  $N$  为像素数,  $T$  为时间序列长度即天数,  $M$  为时间序列维度即波段数量。本文中数据异常情况包括数据缺失和数据扭曲两种情况, 其中数据缺失是指丢失部分像素, 即  $N$  减少, 以模拟遥感影像采集中的数据量少。数据扭曲则是针对时间序列长度, 又分为压缩时间序列长度, 拉伸时间序列长度和二者混合三种情况, 即  $T$  变化, 以模拟遥感影像采集中由于天气等原因, 导致缺失某天或某段时间的数据等情况。

同时, 地物的时间序列数据在遥感影像中呈点簇状分布, 整体数据信息较为离散。在此情况下, LSTM 分类效果不理想<sup>[10,14]</sup>。文献[10–12]应用 1D-CNN 算法可有效描述一维序列数据的特性, 解决遥感影像时序分类问题。使用 MCNN<sup>[13]</sup>对时序数据集进行分类。MCNN 模型包括三个阶段, 转换阶段, 局部卷积阶段和全卷积阶段。其中转换阶段对输入时间序列应用各种转换, 以应对现有时间序列分类方法中由于特征提起与分类过程分离, 所导致的无法提取不同时间尺度的不同特征的问题。MCNN 的局部卷积阶段, 通过多组时间序列上进行卷积, 提取不同时间尺度序列的特征。不同的分支彼此独立, 所有的输出将通过多个具有不同大小的最大池化层。全卷积阶段中将提取的所有特征连接起来, 并通过多个跟着最大池化的卷积层、全连接层和 softmax 层来生成最终的输出。MCNN 整体通过这个端到端的系统, 反向传播所有的参数进行训练。该方法可以一定程度上应对数据缺失的问题, 但该方法并未对数据的各个波段之间的关系进行考虑, 无法应对数据扭曲, 分类精确率较低。考虑到多模态时间序列的不同特征之间的假定独立性, 多通道深度卷积神经网络(MCDNN)<sup>[16]</sup>在输入的每个维度上独立并行应用卷积。MCDNN 对传统的 CNN 进行了改进, 将多变量时间序列分解为单变量时间序列, 即对每个  $\mathbf{X}^i$  分别进行处理, 然后在特征学习的最后连接一个多层感知机(Multi-Layer Perceptron, MLP)进行分类, 但其无法应对数据缺失的情况。由于 MCDNN 无法发掘出不同单变量之间的相互关系, TSCNN<sup>[17]</sup>联合训练多元时间序列进行特征提取。不同的单变量可能具有不同的价值, 有的单变量区分度较大更有价值。传统的 CNN 通常使用交叉熵的 softmax 输出层, TSCNN 的不同之处在于其使用了均方误差(Mean Squared Error, MSE) 损失函数的 Sigmoid 输出层。并且在池化层部分使用平均池化层代替通常使用的最大池化层。在最后一个卷积层之后不再使用池化层。该方法考虑到了变量之间的关系, 将整个数据作为一个整体进行训练, 但不能学习不相关的时间序列之间的信息,

无法应对数据扭曲的情况。通过分析各分类模型的优缺点, 本文提出一种端到端的多模式与多单模式架构融合的网络。

## 2 端到端模式融合网络

端到端的多模式与多单模式的架构融合的网络的整体架构如图 1 所示。该网络主要由两阶段构成: 第一部分是自动编码阶段; 第二部分为分类阶段。其中第一阶段又分为三个部分, 第一个部分为多元时序模型; 第二个部分为单变量时间序列模型组合; 第三个部分为像素坐标输入部分。

多元时序模型的输入为整个多维时序数据整体, 这样可以很好的衡量多个单变量之间的关系。通过三个卷积层进行特征抽取, 并且进行扁平化, 输出整体多维时间序列数据的特征。单变量时间序列组合则是对各个单变量时间序列单独进行多尺度的提取, 包括两个带着最大池化层的卷积层, 其中各个单变量在第一层卷积之后进行连接, 并再次卷积。第二层卷积之后的特征图也进行连接, 再进行卷积与扁平化。该特征与单个通道及第一层后的连接特征图共同进行最大池化与扁平化后连接, 送入自动编码阶段的最后一层, 输出一系列单变量时间序列的特征。像素坐标输入部分则将经过预处理的像素坐标, 按比例缩放后直接送入最后一层。自编码阶段最后一层连接以上三个部分的输出, 将其输入分类阶段。总的来说, 待分类的时间序列数据会以整体和单个变量的形式分别输入网络, 经过不同的卷积神经网络, 得到所有的特征集合, 其中包括不同尺度的特征, 最后与像素坐标一同连接, 得到时间序列特征。

分类阶段则是通过全连接层的叠加对自动编码阶段得到的时间序列特征进行分类。将五个全连接层间学习到的时序特征映射到样本标记空间。并在第二个全连接层与第三个全连接层使用 Dropout, 使得神经元与其他随机挑选神经元共同工作, 来避免模型过拟合, 增加了泛化能力。最后一层的全连接层通过 softmax 获得分类概率, 对像素进行分类, 得到地物的类别标签。最终整个网络以时序特征的分类损失为整体的网络损失进行反向传播, 得到端到端的时间序列分类模型。网络的整体学习过程将在下文进行具体的介绍。

由于时序分类网络对待时序数据角度不同, 一个假定时间序列不相关, 可以应对数据缺失情况, 另一个则认为这些时间序列是一整个实体, 可以应对数据扭曲的情况。本文提出的端到端的多模式与多单模式架构融合的网络则可以将二者结合起来, 以解决遥感影像时间序列数据异常的问题。

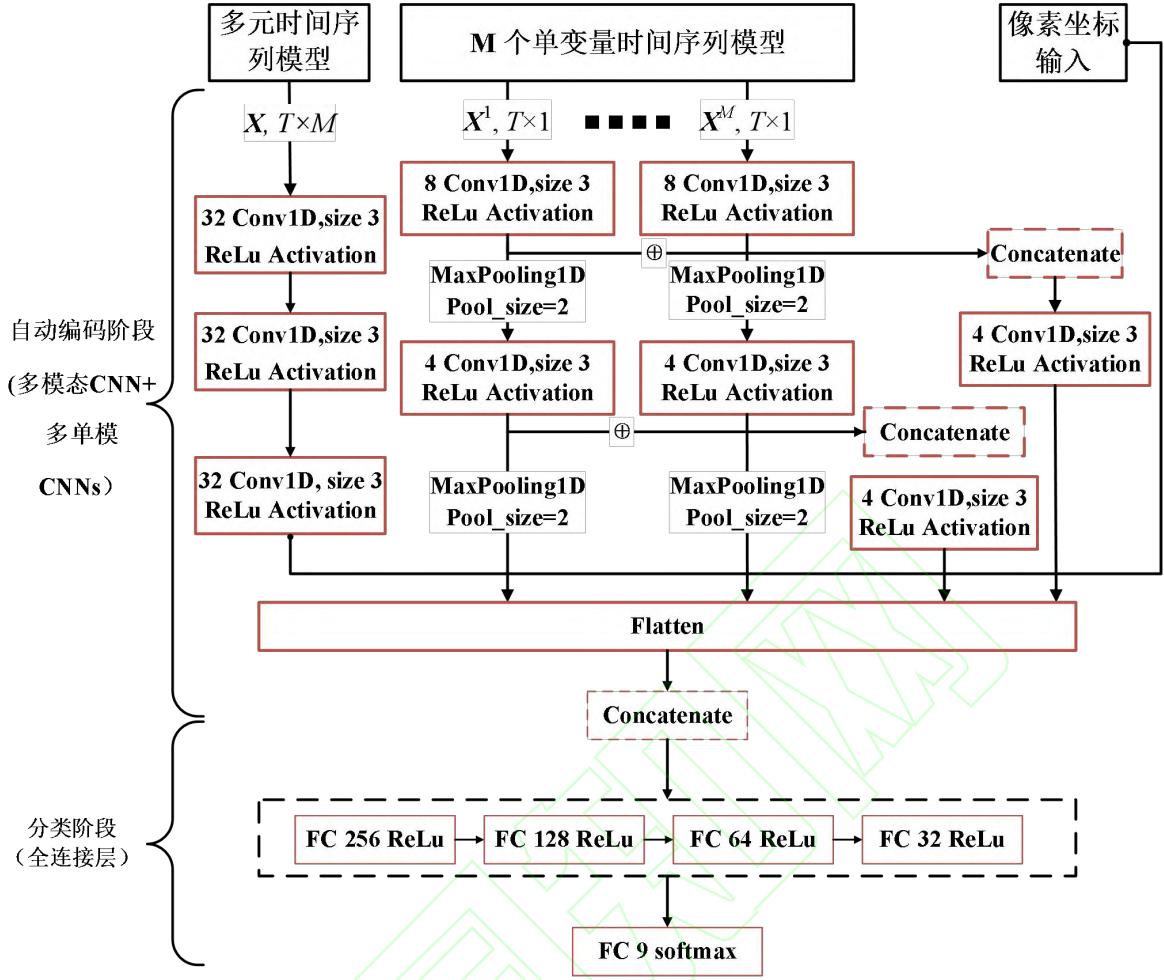


图1 端到端的多模式与多单模式架构融合的网络

Fig. 1 End-to-end multi-mode and multi-single-mode architecture fusion network

### 3 网络结构

时间序列分类问题中, 训练样本通常由一系列已标记的单变量时间序列样本构成, 假设单变量时间序列  $\mathbf{X} = [x_1, x_2, \dots, x_T]$  是有序的一组实数值, 其中  $\mathbf{X}$  的长度  $T$ 。 $M$  维时间序列  $\mathbf{X}_M = [\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^M]$  是由  $M$  个不同的单变量时间序列构成的多元时间序列, 其中  $\mathbf{X}^i \in \mathbb{R}^T$ 。

在整个体系结构中的自动编码阶段, 使用了3种不同的模型, 输入数据首先被用来训练一个多元时序模型卷积神经网络, 将  $M$  维遥感影像时间序列作为一个整体进行处理, 输入时序数据的维度是  $T \times M$ , 即对  $\mathbf{X}_M$  整体进行卷积。通过连续三层卷积, 其中每层包括为32个尺寸为  $3 \times 3$  的卷积核, 以获得整体的时间序列特征。

#### 3.1 单变量时间序列模型组合

各类地物对应的遥感影像时间序列变化不一, 单靠一种波段很难进行分类, 且各个波段信息之间存在一定的冗余。因此需要对多个单变量进行结合。单变量时间序列模型对

$M$  维时间序列数据  $\mathbf{X}_M = [\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^M]$  中每一维单独进行处理, 即对  $\mathbf{X} = [x_1, x_2, \dots, x_T]$  进行处理。并且将  $M$  个处理后的数据在不同的层级进行连接以获取不同尺度的时间序列, 如图2所示。每个波段的时间序列被输入到卷积层中, 第一个卷积层包括8个尺寸为  $3 \times 3$  的卷积核。每个波段的第二个卷积层包括4个尺寸为  $3 \times 3$  的卷积核。分别在第一个、第二个卷积层后对得到的特征图进行连接, 对第一层卷积后的连接特征图使用8个尺寸  $10 \times 10$  的卷积核进行卷积并卷积, 对第二层卷积后的连接特征图使用4个尺寸为  $10 \times 10$  的卷积和进行卷积, 并且将得到的结果扁平化与每个通道卷积后的结果扁平化后连接。这样既结合了单变量进行训练, 又使用了不同尺度的特征进行采样, 丰富了卷积网络所提取到的特征图, 有利于解决数据扭曲的问题。

自动编码阶段的第3部分模型负责处理像素的坐标, 预处理按比例缩放的像素坐标并与另外两个模型的结果连接, 交给全连接层一起处理。

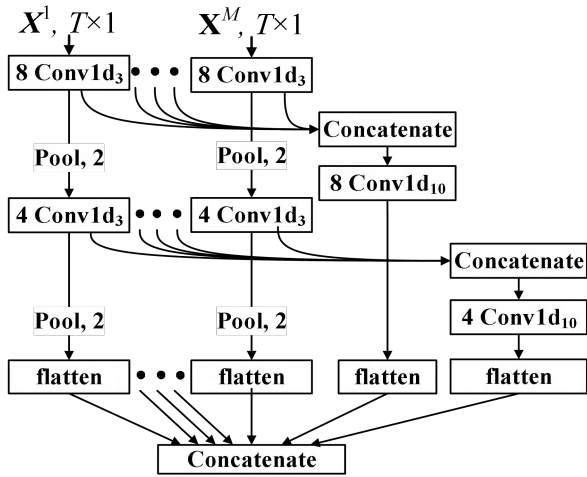


图2 多单模式的架构

Fig. 2 Unimodal architecture

### 3.2 分类阶段

本阶段使用具有 256 个神经元的全连接层, 激活函数为线性整流函数(Rectified Linear Unit, ReLU)。数据经过 128 个神经元的全连接层, 并且以 0.3 的概率神经元不工作, 再经过 64 个神经元的全连接层, 以 0.3 的概率神经元不工作, 避免过拟合。最后通过 32 个神经元的全连接层, 连接到 softmax 层对像素进行分类。分类阶段通过在 128 个神经元、64 个神经元和 32 个神经元这三个全连接层之间增加 Dropout 层, 来避免过拟合, 增加模型的泛化能力。

模型使用自适应矩估计 Adam 优化器, 学习率为 0.001, 每个批次的大小 Batch size 的大小为 256, 共计 50 个 Epoch。动态调整学习率, 每次调整时学习率变为原学习率的 0.2 倍。经过 3 个 Epoch 训练后, 模型性能不再提升, 学习率减少。学习率的下限为 0.00001。

## 4 试验与分析

### 4.1 试验环境和试验设置

采用 TensorFlow 2.2.0 开源框架, 搭建在操作系统为 Ubuntu16.04, Python 版本为 3.6 的平台上, 平台显卡型号为 GeForce GTX TITAN XP。本文设置模型学习率(Learning Rate, LR)为 0.001, 每个批次的大小 Batch size 为 256, 训练数据集次数 Epochs 为 100 次。对比模型的参数:

1D-CNN: LR=0.01, Batch size=256, Epochs=100, 降采样因子为 2, 即时间序列下采样为原先的二分之一。滑动平均窗口大小为 4, 即时间序列的值为本值与该位置前两个和后两个的平均值。

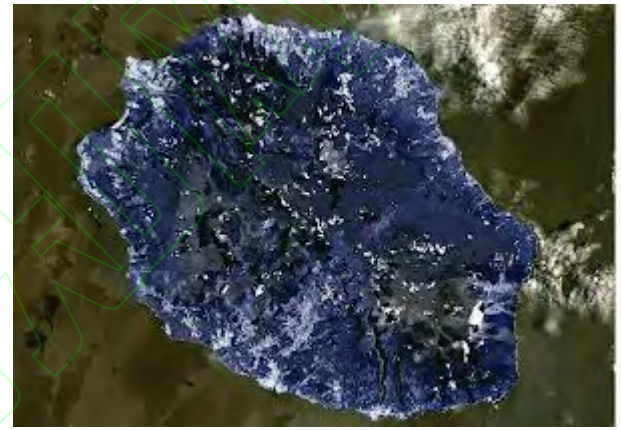
MCDNN: LR=0.01, Batch size=256, Epochs=100。

TSCNN: LR=0.001, Batch size=128, Epochs=100。

LSTM: LR=0.001, Batch size=256, Epochs=100。

### 4.2 试验数据和数据准备

所使用的数据集是来自于 The time series land cover classification challenge (TiSeLaC) [15], 2014 年在留尼汪岛的 23 幅 Landsat 8 影像生成的时间序列数据集。留尼汪岛影像如图 3(a)所示, 土地覆盖类型如图 3(b)所示。该遥感影像级别为 2A, 尺寸为 2866×2633 像素, 地面分辨率 30m。在每个时间戳, 每个像素有 10 个特征值, 包括 7 个表面反射率(海岸波段、蓝波段、绿波段、红波段、近红外波段、短波红外 1 和短波红外 2)和 3 个辐射指数: NDVI、归一化水指数(Normalized Difference Water Index, NDWI)和亮度指数(Brightness Index, BI)。而数据集的土地覆盖类型数据, 即像素类别标签则是由 2012 年 Corine 土地覆盖地图[19]和 2014 年 Graphical Land Parcel 注册信息联合标注而得, 以保证其真实可靠。训练集包括 81714 个像素, 测试集为 17973 个像素, 共计 9 个类别, 如表 1 所示。



(a) 留尼汪岛

(a) Reunion Island



(b) 土地覆盖类型

(b) Land Cover Classes

图3 留尼汪岛影像

Fig. 3 Remote sensing images of Reunion Island

数据集中地面特征可以表示为  $X_M \in R^{N \times T \times M}$ , 其中  $N = 81714$  个像素,  $T = 23$  天,  $M = 10$  为特征数量即波段数与辐射指数数量之和。



表1 数据集类别分布

Tab. 1 Data set class distribution

类别编号	类别名称	Train	Test
1	Urban areas	16000	4000
2	Other built-up surface	3236	647
3	Forests	16000	4000
4	Sparse vegetation	16000	3398
5	Rocks and bare soil	12942	2588
6	Grassland	5681	1136
7	Sugarcane crops	7656	1531
8	Other crops	1600	154
9	Water	2599	519

本文通过数据处理来模拟数据异常情况。首先是数据缺失情况，即存在部分像素点缺失。本文采用随机删减策略进行模拟，即通过一定概率随机删除部分像素点，并最终只保留固定数量的像素点进行数据训练。其次，针对数据扭曲，包括时序数据压缩操作、拉伸操作和两者混合操作。其中，图4(a)表示原始时间序列。压缩操作表现为基于时序数据按照固定步长进行抽取，如图4(b)所示。拉伸操作表现为对特定时序区间进行线性插值进行数据填充，如图4(c)所示。

#### 4.3 评价指标

为客观评价卷积神经网络的分类性能，本文使用精确率(Precision)、召回率(Recall)和F1值作为评价指标。其中真阳(True Positive, TP)、假阳(False Positive, FP)、假阴(False Negative, FN)和真阴(True Negative, TN)分别表示：

TP：实际为正，预测为正的样本数量

FP：实际为负，预测为正的样本数量

FN：实际为正，预测为负的样本数量

TN：实际为负，预测为负的样本数量

##### (1) 精确率

精确率指模型预测为正的样本中实际也为正的样本，占被预测为正的样本的比例，即

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

精确率体现了模型对负样本的区分能力，精确率越高，模型对负样本的区分能力越强。

##### (2) 召回率

召回率指实际为正的样本中，被预测为正的样本，占实际为正的样本的比例，即

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

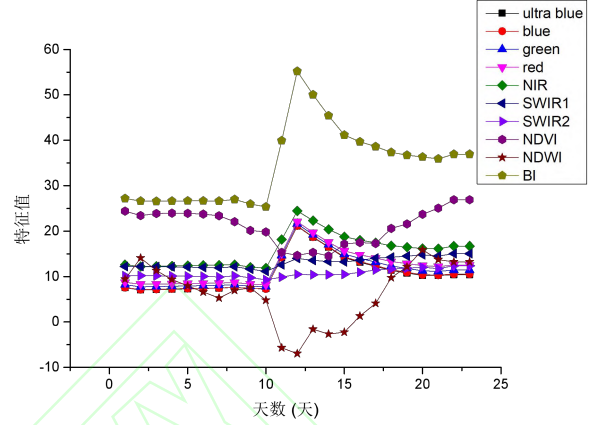
召回率体现了模型对正样本的识别能力，召回率越高，模型对正样本的识别能力越强。

##### (3) F1 值

F1 值是评价卷积神经网络性能的常见指标。F1 值通过精确率和召回率计算，公式如下：

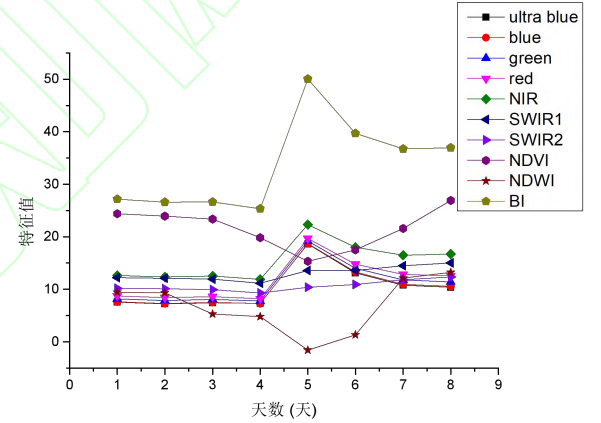
$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (3)$$

F1 值作为精确率和召回率的调和平均值，其值越高，说明模型越稳健。



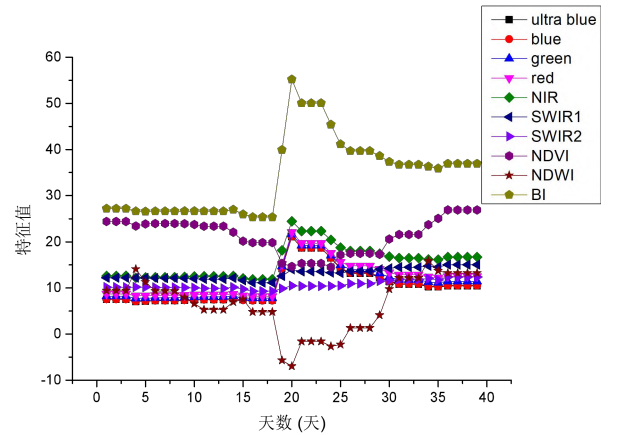
(a) 原始时间序列数据

(a) Original time series data



(b) 压缩时间序列数据

(b) Compressed time series data



(c) 拉伸时间序列数据

(c) Stretched time series data

图4 时间序列数据操作

Fig. 4 Operation for time series data

表 2 时序数据分类结果

Tab. 2 The results of time series data classification

数据变化情况	模型	精确率/%	召回率/%	F1 值/%
原始数据	1D-CNN	88.50	88.68	88.47
	MCDNN	86.28	86.75	86.34
	TSCNN	88.60	88.87	88.68
	LSTM	92.43	92.40	92.41
	端到端的多模式与多单模式架构融合的网络	93.49	93.33	<b>93.40</b>
数据缺失(缺失部分像素点)	1D-CNN	82.34	83.15	82.58
	MCDNN	73.31	75.00	73.50
	TSCNN	84.51	85.38	84.89
	LSTM	85.63	86.63	85.98
	端到端的多模式与多单模式架构融合的网络	87.83	87.87	<b>87.41</b>
数据扭曲(压缩部分时序数据)	1D-CNN	86.15	86.48	86.48
	MCDNN	85.20	85.74	85.27
	TSCNN	77.18	78.72	77.27
	LSTM	90.69	90.84	90.73
	端到端的多模式与多单模式架构融合的网络	92.71	92.62	<b>92.65</b>
数据扭曲(部分时序数据删除后进行线性插值)	1D-CNN	87.81	88.23	87.93
	MCDNN	86.63	86.91	86.46
	TSCNN	87.66	88.12	87.78
	LSTM	92.18	92.07	92.11
	端到端的多模式与多单模式架构融合的网络	93.27	93.16	<b>93.20</b>
数据扭曲(拉伸部分时序数据进行填充)	1D-CNN	87.65	88.01	87.71
	MCDNN	87.20	87.56	87.29
	TSCNN	88.85	89.17	88.93
	LSTM	93.29	93.20	93.23
	端到端的多模式与多单模式架构融合的网络	93.47	93.30	<b>93.35</b>

#### 4.4 实验结果分析

本文比较了不同模型在数据缺失与数据扭曲情况下对时序数据分类性能的影响。以 23 天时间序列数据为原始数据集, 1D-CNN、MCDNN、TSCNN、LSTM 和端到端的多模式与多单模式架构融合的网络这五种时序分类网络得到的结果如表 2 所示, F1 值分别为 88.47%, 86.34%, 88.68%, 92.41%和 93.40%。其中 1D-CNN 采用 Adam 优化器, 初始学习率为 0.001。MCDNN 采用随机梯度下降策略 SGD 与 0.01 的学习率。TSCNN 采用学习率不变化的 Adam 优化器, 学习率为 0.001。由于 LSTM 形式的深层递归神经网络在这种像素较少的情况下, 训练效果不佳<sup>[14]</sup>。本文所使用的 LSTM 网络为借助 LSTM 结构对端到端的多模式与多单模式架构融合的网络进行结合, 以提升 LSTM 网络性能。通过将本文网络中的多元时序模型, 替换为输出维度为 512 的 LSTM 网络, 对时序数据进行整体特征提取, 再结合网络中的多单变量网络与像素坐标, 最后在分类阶段对数据进行分类。

在这种原始情况下, 五种时序分类网络的精确率和召回率相仿, 说明其对正负样本的识别能力都很强。从中可以看出, 端到端的多模式与多单模式架构融合的网络对时序数据分类性能最好, 精确度达到了 93.5%。而数据缺失的情况, 即以 80%的概率删除训练集中的像素点, 剩余 16343 个像素点。当训练数据变少时, 端到端的多模式与多单模式架构融合的网络的 F1 值仍然最优。同时, MCDNN 在此时受影响较大, 精确度只有 73.31%, 且各个网络的 F1 值均有所降低, 说明训练数据减少会影响分类效果。其中对 TSCNN 影响最小, 精度下降约 4%。端到端的多模式与多单模式架构融合的网络的前后变化较小, 说明数据量变小对其影响较小。端到端的多模式与多单模式架构融合的网络由于使用三种不同模型来对原始时序数据进行分类, 故而较为分类效果稳定, 可以应对数据缺失的问题。

为模拟数据扭曲的情况, 本文压缩部分时序数据同时拼接剩余数据, 共剩余 6 天时序数据。此时在数据缺失情况下表现优异的 TSCNN, 分类精度却下降最大, 说明在时间维度较小的数据集中, 将所有波段特征作为一个整体,



区分度较低。本文模型性能受影响较小，仍能保持 92.65% 的 F1 值，分类性能优于其他模型。为模拟数据采集较少的情况，即只有 15 天时序数据的情况下，需要考虑对数据进行填充，通过线性插值的方法对其进行填充，将数据填充回原数据集的大小进行训练。在这种情况下，得到的分类结果与准线基本一致，说明这类时序数据中线性插值的方法是有效的，可以提高训练数据较少时的分类精度。而在数据拉伸的情况下，本文通过对相邻时间区间的数据进行之间线性插值，获得共 40 天时序数据，并对此进行分类。在这种情况下由于 1D-CNN 和 MCDNN 网络结构较为简单，对遥感影像时序数据描述能力较弱，可能会产生过拟合现象，导致分类效果变差 F1 值降低。但 TSCNN 和本文提出的端到端的多模式与多单模式架构融合的网络结构考虑到了将多维特征作为整体学习，模型复杂描述能力强，在数据较为复杂的情况下，仍能有效表示时序数据。可以考虑扩充时间维度来增加数据，提升模型分类能力。

综合考虑各种情况，端到端的多模式与多单模式架构融合的网络的分类结果均好于其他模型，说明了其能应对数据异常的问题。

在原始情况下，五种时序分类网络的时间性能如表 3 所示，由于 1D-CNN 需要对原始时间序列数据进行转换，对于单一时间序列输入进行降采样和滑动平均等变化，所以训练用时最长，性能最差。其余模型时间性能接近，MCDNN 用时最短，单个 Epoch 用时为 2s 6ms，效率最高。其次是 TSCNN。本文模型单 Epoch 用时与 LSTM 基本一致，分别为 5s 16ms 和 5s 18ms。综合精度与时间性能，本文模型效果更好。

表 3 时序数据分类用时

模型	单次 Epoch 用时
1D-CNN	123s 482ms
MCDNN	2s 6ms
TSCNN	3s 6ms
LSTM	5s 18ms
端到端的多模式与多单模式 架构融合的网络	5s 16ms

本文网络在 9 个地物类别的分类效果有所不同，图 4 展示了在 9 类不同地物分类任务上的混淆矩阵。从图中可以看出，9 个类别都可以被区分开，其中其他作物类相对区分度较低，其原因主要是其在训练样本中所占数量较少，训练较为不充分。城市地区与其他建筑物由于光谱反射率较为接近，也存在一定的误分现象。

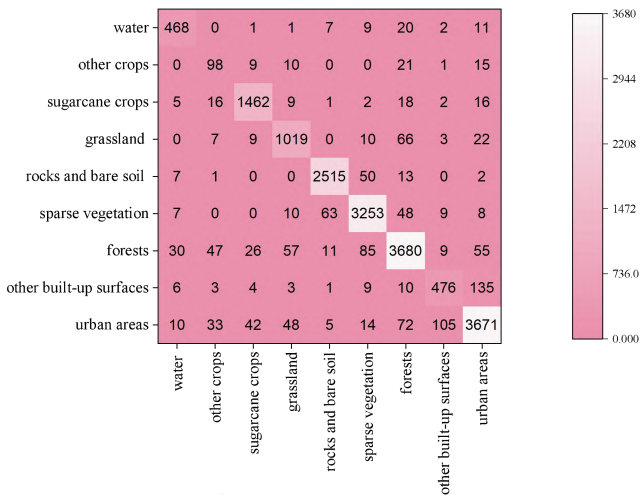


图 5 混淆矩阵

Fig. 5 Confusion Matrix

5 总结

遥感影像时间序列分类在遥感影像分类中十分重要，本文提出了一种端到端的多模式与多单模式架构融合的网络，可以通过结合多元时序模型和多尺度单变量模型进行特征提取，与像素坐标汇总后进行分类，并通过使用公开遥感影像时序数据集对 1D-CNN、MCDNN、TSCNN、LSTM 和本文模型进行验证，实验表明：在数据异常的情况下，本文模型仍然保持了分类精度优于其他模型，且 F1 值变化较小，说明了本文模型具有鲁棒性。该模型在解决时序分类的问题上，尚有进一步研究的空间，后续工作将使用模型初始化方法提高模型分类精度，以及研究更极端条件下对模型的影响。

参考文献(References)

[1] 邓书斌. ENVI 遥感图像处理方法[M]. 高等教育出版社, 2014:407-424. (DENG S B. ENVI remote sensing image processing method [M], Higher Education Press, 2014:407-424)

[2] FRANKLIN S E, AHMED O S, WULDER M A, et al. Large area mapping of annual land cover dynamics using multitemporal change detection and classification of Landsat time series data [J]. Canadian Journal of Remote Sensing, 2015, 41(4): 293-314.

[3] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. neural computation, 1997, 9(8): 1735-1780.

[4] 林蕾. 基于循环神经网络模型的遥感影像时间序列分类及变化检测方法研究[D]. 中国科学院大学 (中国科学院遥感与数字地球研究所), 2018:45-74.(LIN L. Remote sensing image time series classification and change detection based on recurrent neural network model [D]. University of Chinese Academy of Sciences (Institute of remote sensing and digital earth, Chinese Academy of Sciences), 2018:45-74.)

[5] MAGGIORI E, TARABALKA Y, CHARPIAT G, et al. Convolutional neural networks for large-scale remote-sensing image classification [J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 55(2): 645-657.

[6] POSTADJIAN T, LE BRIS A, SAHBI H, et al. Investigating the potential of deep neural networks for large-scale classification of very high resolution satellite images [J]. ISPRS Annals of

- Photogrammetry, Remote Sensing & Spatial Information Sciences, 2017, 4: 183–190.
- [7] VOLPI M, TUIA D. Dense semantic labeling of subdecimeter resolution images with convolutional neural networks [J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 55(2): 881–893.
  - [8] AUDEBERT N, LE SAUX B, LEFÈVRE S. Segment-before-detect: vehicle detection and classification through semantic segmentation of aerial images [J]. Remote Sensing, 2017, 9(4): 368.
  - [9] ZHANG Q, YUAN Q, ZENG C, et al. Missing data reconstruction in remote sensing image with a unified spatial-temporal-spectral deep convolutional neural network [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(8): 4274–4288.
  - [10] ZHONG L, HU L, ZHOU H. Deep learning based multi-temporal crop classification [J]. Remote sensing of environment, 2019, 221: 430–443.
  - [11] HU W, HUANG Y, WEI L, et al. Deep convolutional neural networks for hyperspectral image classification [J]. Journal of Sensors, 2015(2):1-12.
  - [12] WANG Z, YAN W, OATES T. Time series classification from scratch with deep neural networks: A strong baseline [C]//2017 International joint conference on neural networks (IJCNN). 2017: 1578–1585.
  - [13] CUI Z, CHEN W, CHEN Y. Multi-scale convolutional neural networks for time series classification[EB/OL]. [2016-05-11]. <https://arxiv.org/pdf/1603.06995.pdf>
  - [14] DI MAURO N, VERGARI A, BASILE T M A, et al. End-to-end learning of deep spatio-temporal representations for satellite image time series classification [C]// PKDD/ECML 2017: Proceedings of the 2017 European Conference on Machine Learning & Principles and Practice of Knowledge Discovery in Databases. Skopje: Springer, 2017:18-22.
  - [15] TiSeLaC: Time Series Land Cover Classification Challenge - Dino Ienco[EB/OL].[2020-05-24]. <https://g0.njuu.cf/extdomains/sites.google.com/site/dinoienco/tisela>.
  - [16] ZHENG Y, LIU Q, CHEN E, et al. Exploiting multi-channels deep convolutional neural networks for multivariate time series classification [J]. Frontiers of Computer Science, 2016, 10(1): 96–112.
  - [17] ZHAO B, LU H, CHEN S, et al. Convolutional neural networks for time series classification [J]. Journal of Systems Engineering and Electronics, 2017, 28(1): 162–169.
  - [18] 蔡涛,王润生. 一个从多波段遥感图像提取道路网的算法[J]. 软件学报, 2001, (6): 943–948.(CAI T, WANG R S. An algorithm for road network extraction from multi band remote sensing images [J]. Journal of software, 2001, (6): 943–948.)
  - [19] CORINE Land Cover — European Environment Agency[EB/OL].[2020-05-22].<https://www.eea.europa.eu/publications/COR0-landcover>.

**This work is partially supported by** the National Key Research and Development Program of China under Grant (2017YFB0504203).

**REN Yuanyuan**, born in 1996, M. S. candidate. Her research interests include machine learning, remote sensing data processing.

**WANG Chuanjian**, born in 1977, Ph. D., professor. His research interests include machine learning, remote sensing data processing.