

# 基于 Stacked ConvLSTM 的时间序列 森林火烧迹地检测

李淑君<sup>1,2</sup>, 郑柯<sup>1</sup>, 唐娉<sup>1</sup>, 霍连志<sup>1</sup>, 袁媛<sup>3</sup>

1. 中国科学院空天信息创新研究院, 北京 100094;

2. 中国科学院大学, 北京 100049;

3. 南京邮电大学, 南京 210042

**摘要:** 确定森林火烧迹地的准确时间点以及空间范围对于森林的受损评价、管理、碳核算以及森林恢复的管理有重要意义。由于森林火烧迹地在空间分布上具有一定的连续性, 现有的森林火烧迹地提取方法大都采用先分类再后处理的两步处理策略来抑制虚警像素的影响。本文提出将时空检测方法 Stacked ConvLSTM 用于时间序列森林火烧迹地的检测, 在保持结果具有较好空间连续性的基础上避免了具有主观性的后处理操作, 实现端到端提取森林火烧迹地信息, 提升了森林火烧迹地的提取精度。采用 MODIS 时间序列数据, 基于 2001 年—2008 年以及 2001 年—2016 年的黑龙江沾河林业局伊南河林场和内蒙古自治区毕拉河林业局北大河林场两个区域的历史时间序列, 分别对这两个区域 2009 年以及 2017 年发生的特大火灾区域进行火烧迹地检测, 利用 Stacked ConvLSTM、Stacked LSTM 以及 bfast 算法在两个区域的 MODIS 时间序列中提取森林火烧迹地, 并将火烧迹地检测结果与 ESA 发布的 Fire\_CCI 5.1 火烧迹地产品进行对比分析。结果表明: 首先, 从目视效果来看, 在研究区域 I, Stacked ConvLSTM 检测的结果比 Stacked LSTM 和 bfast 算法错误检测点少, 并且在空间分布也保持较高连续性; 在研究区域 II, Stacked ConvLSTM 检测到了较完整的火烧迹地区域。其次, 在定量的精度评价指标上, 在研究区域 I, Stacked ConvLSTM 的精确度比 Stacked LSTM 和 bfast 算法分别高出 0.120 和 0.405, 并且召回率、准确度和 F1-score 也更高, Fire\_CCI 5.1 召回率虽更高, 由于错检区域较大, 其他精度指标远低于 Stacked ConvLSTM; 在研究区域 II, Stacked ConvLSTM 精确度达 0.924, 召回率、准确度和 F1-score 相比 Stacked LSTM 和 bfast 算法以及 Fire\_CCI 5.1 更高。

**关键词:** Stacked ConvLSTM, 时间序列, 时空预测, 火烧迹地

**引用格式:** 李淑君, 郑柯, 唐娉, 霍连志, 袁媛. 2022. 基于 Stacked ConvLSTM 的时间序列森林火烧迹地检测. 遥感学报, 26(10): 1976-1987

Li S J, Zheng K, Tang P, Huo L Z and Yuan Y. 2022. Forest burned area detection with time series data based on Stacked ConvLSTM. National Remote Sensing Bulletin, 26(10): 1976-1987 [DOI: 10.11834/jrs.20210471]

## 1 引言

森林火灾是一种突发性强、破坏性高、处置困难的自然灾害 (Cao 等, 2020)。近年来, 受全球气候变化与人类活动的影响, 森林防火行业面临着愈发严峻的考验, 对森林火灾识别技术提出了新的需求与挑战 (魏书精 等, 2020; 王振师 等, 2019; 孙红斌 等, 2018)。火烧迹地是描述森林燃烧最重要的信息之一, 是经火灾烧毁后尚未

长成新林的土地, 能够提供森林火灾发生时间、频度、位置、面积以及空间分布等重要信息 (Ruiz 等, 2014), 是火灾形成机制不可或缺的研究对象 (Chuvieco 等, 2016), 也是影响森林资源保护、植被恢复、碳排放等至关重要的变量 (Chuvieco 等, 2018), 以及碳循环扰动和全球变化研究的重要参数 (孙桂芬 等, 2019)。遥感技术具有观测覆盖范围大、可动态观测等优势, 使其成为最为有效的森林监测手段, 如何利用时间序列遥感图像提取

收稿日期: 2020-10-26; 预印本: 2021-03-28

基金项目: 国家自然科学基金(编号:41971396, 41901356); 中国科学院(A类)战略性先导科技专项(编号:XDA19080301)

第一作者简介: 李淑君, 研究方向为遥感图像处理。E-mail: lishujun@aircas.ac.cn

通信作者简介: 霍连志, 研究方向为遥感图像处理和分类。E-mail: huolz@aircas.ac.cn

森林变化信息是当前研究的热点和难点问题(杨辰, 2013)。

根据数学统计方法划分, 传统的时间序列变化检测算法可分为6大类, 包括阈值法、差分法、分段法、轨迹分类法、统计边界法和回归法(Zhu, 2017)。阈值法采用了一个预定的阈值确定时间序列中的森林, 当与阈值存在显著偏差时检测到变化, Hilker等(2009)基于Landsat和Modis数据的NDVI指数利用融合模型对森林实现变化检测, Huang等(2009)利用森林综合指数完成率对美国西部森林区域动态变化的评定(IFZ); 差分法通过比较不同时间获取的影像, 并且将变化很大的差异定义为变化区域, Bolton等(2015)使用差分法对火灾后的残留结构和森林恢复进行分析; 分段法将完整的历史时间序列数据根据残差或角度标准分为一系列直线段, 检测突变和渐变信息(Chance等, 2016); 轨迹分类方法从时间序列中提取某种变化(训练目的)信息, 用该信息对图像中的每个时间序列进行分类, Kennedy等(2007)基于轨迹对森林动态特征进行自动匹配实现变化检测; 统计边界法将明显偏离边界的检测为变化, Zhu等(2016)利用Landsat数据对绿度趋势分析土地覆盖变化; 回归法假定预测变量和响应变量之间存在线性关系, 使用回归获得结果(Markham等, 2012)。此外, 还有一些其他常用的经典时间序列变化检测算法, 如干扰和趋势监测(LandTrendr)方法(Kennedy等, 2010)、植被追踪算法(VCT)(Huang等, 2010)、bfast算法(Verbesselt等, 2012)、连续变化检测和分类算法(CCDC)(Zhu和Woodcock, 2014)等。

近年来, 深度学习得以迅速发展, 并越来越多地用于时间序列分析(Långkvist等, 2014), 在时间序列森林变化监测方面的应用也越来越广泛, Reddy和Prasad(2018)将LSTM用于海岛植被动态监测, 预测岛上植被未来时间序列以获取植被变化信息; Kong等(2018)基于LSTM模型对长时间序列森林火灾进行监测, 检测率高于传统的bfast算法; Yuan等(2020)进一步在LSTM模型中加入Attention机制并用于森林火灾监测, 检测精度相比LSTM有很大提升。然而, LSTM模型主要是对图像中单个像元的时间序列进行分析, 没有考虑像素的空间邻域信息, 而空间邻域的相关性对于遥感图像分析具有重要意义(Huo等, 2015)。

为在时间序列预测中同时考虑空间信息, Shi等(2015)提出了ConvLSTM, 将原来LSTM中的全连接改进为卷积连接, 可以较好地刻画时间序列的时空特征, 并用于降雨预测。Stacked ConvLSTM堆叠多个ConvLSTM增加网络的表达能力, 提升了对时间序列数据时空结构信息的提取能力, 因此, Kim等(2017)用两层ConvLSTM基于雷达数据进行降雨的多步预测, Zhao等(2019)基于三层的ConvLSTM对超声波时间序列影像进行单幅影像预测。

由于Stacked ConvLSTM尚未用于森林火烧迹地检测, 其在森林火烧迹地检测的有效性和效果有待验证, 本文提出将时空检测方法Stacked ConvLSTM用于时间序列森林火烧迹地的检测, 实现端到端提取森林火烧迹地信息。由于火烧迹地在空间分布上具有连续性(武晋雯等, 2020), 因而在时间序列变化检测中加入邻域空间信息具有重要意义。一方面, 能够在一定程度上降低因数据缺失(云、阴影、地形遮挡等所致)造成的影响(Khan等, 2017); 另一方面, 相比现有方法大多采用先分类再后处理抑制零散虚警信息的两阶段策略(张兆明等, 2020), 端到端数据驱动的方法能够避免后处理方法选择的主观性。

## 2 研究区域、数据与方法

### 2.1 研究区域及数据

两个研究区域分布如图1所示, 研究区域I是黑龙江沾河林业局伊南河林场, 地理位置在 $48^{\circ}39'N$ ,  $128^{\circ}14'E$ 附近区域, 于2009年4月27日发生草甸森林火灾, 过火总面积 $1557.562\text{ km}^2$ , 其中有林地 $1072.746\text{ km}^2$ 。

研究区域II在内蒙古自治区毕拉河林业局北大河林场, 地理位置在 $49^{\circ}31'N$ ,  $123^{\circ}06'E$ 附近, 2017年5月2日发生森林火灾, 过火面积 $115\text{ km}^2$ , 有林地占60%, 受害森林面积达 $82.816\text{ km}^2$ 。

选择这两个研究区域的原因有以下几点: (1)两个研究区域的火灾面积较大, 在遥感影像上有明显的过火区域轮廓, 便于研究实验的进行; (2)两个区域的历史时间序列都比较平稳, 除了已知的火灾外没有受到其他明显的森林干扰发生, 便于模型算法的验证; (3)两个研究区分别包含草甸森林和森林区域, 为在不同的森林覆盖类型验证模型的有效性提供了条件。

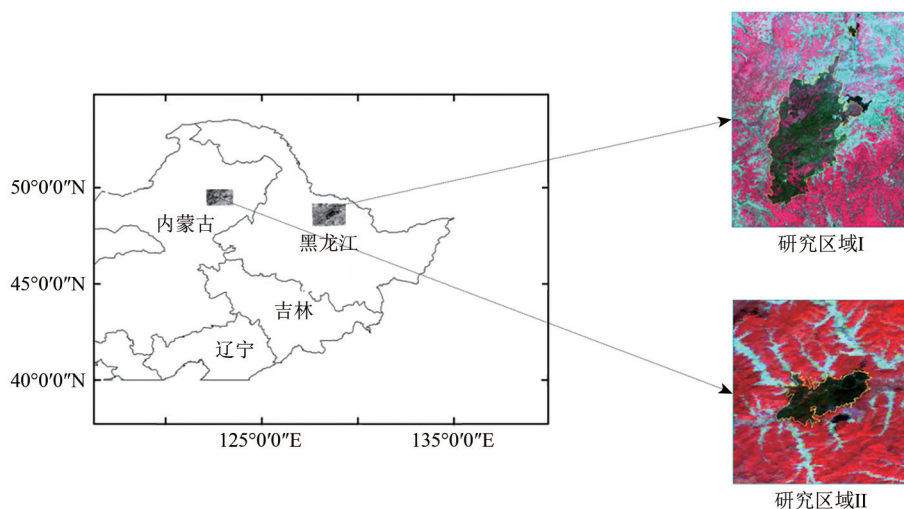


图1 研究区域(研究区域 I :伊南河林场火灾区域,研究区域 II :北大河林场火灾区域)

Fig. 1 Study area (study area I : Yinanhe Forest Farm fire area, study area II : Beidahe Forest Farm fire area)

本文采用的研究数据是中分辨率成像仪 (MODIS) 陆地产品 MOD13Q1。MOD13Q1 空间分辨率 250 m, 是 16 天合成产品, 每年有 23 幅影像。分别获取了研究区域 I 的 2001 年—2009 年以及研究区域 II 的 2001 年—2017 年间的时间序列数据。实验剔除了序列中受云、雪影响大的春冬季影像, 采用一年中季相变化比较明显的 14 幅影像作为研究的时间窗口 (DOY97—DOY305)。

MODIS 数据的 MOD13Q1 产品中增强型植被指数 (EVI) 在植被监测中具有时相多、覆盖面广、且不易产生过饱和等优势 (Huete 等, 2002; 王正兴等, 2003), 被广泛应用于大区域尺度的植被物候变化监测等研究领域 (Peng 等, 2017; Zhang 等, 2003)。因此本文选用增强型植被指数 EVI 来研究其时间序列上的动态变化, EVI 的计算公式可表示为

$$EVI = 2.5 \frac{\rho_{NIR} - \rho_R}{\rho_{NIR} - 6\rho_R - 7.5\rho_B + 1} \quad (1)$$

式中,  $\rho_R$ 、 $\rho_B$ 、 $\rho_{NIR}$  分别表示 MODIS 数据的红、蓝和近红外波段的反射率。

选择 MODIS 数据的原因主要有两点: (1) MODIS 数据的时间跨度长, 可以提供足够长的时间序列; (2) MODIS 16 天合成产品, 时间间隔均匀, 方便用于时间序列的统计建模。

## 2.2 Stacked ConvLSTM

时间序列数据预测是指学习过去的时间序列并预测未来的变化。传统的神经网络无法解决随

时间轴变化的问题, 随之诞生了 RNN (Recurrent Neural Network) (Jordan, 1997); 但由于经典的 RNN 在提取较长时间序列信息的表现较差, 提取的时序信息有限, Hochreiter 和 Schmidhuber (1997) 提出了 LSTM 网络模型, 在经典 RNN 中加入门 (gates) 的结构来选择性的增加和删除过去的时序信息, 即加入输入门、输出门和遗忘门, 分别控制本单元 (一个 LSTM 结构为一个基本单元) 数据的输入、输出以及上一步单元的输出信息增减。LSTM 公式表示如下 (Hochreiter 和 Schmidhuber, 1997):

$$I_t = \sigma(W_{xi}X_t + W_{hi}H_{t-1} + W_{ci}oC_{t-1} + b_i) \quad (2)$$

$$F_t = \sigma(W_{xf}X_t + W_{hf}H_{t-1} + W_{cf}oC_{t-1} + b_f) \quad (3)$$

$$C_t = F_t o C_{t-1} + I_t o \tanh(W_{xc}X_t + W_{hc}H_{t-1} + b_c) \quad (4)$$

$$O_t = \sigma(W_{xout}X_t + W_{hout}H_{t-1} + W_{cout}oC_{t-1} + b_o) \quad (5)$$

式中,  $I$ 、 $F$ 、 $O$  分别表示输入门、遗忘门和输出门,  $C$  和  $H$  分别表示细胞状态 (经过门控输出的信息) 和隐藏状态 (每个时间点的输出值),  $W$  表示对应数据的权重,  $X$  表示输入数据,  $b$  表示偏置值,  $\sigma$  表示激活函数,  $o$  表示哈达玛积, 下标  $t$  表示  $t$  时刻; 下标  $i$ 、 $f$ 、 $out$  分别表示 3 个控制门对应的权重和偏置值, 下标  $c$  表示细胞状态  $C$  对应的权重和偏置;  $W$  的下标  $x$  表示对应输入数据  $X$  的权重,  $h$  表示对应隐藏状态  $H$  的权重。

ConvLSTM 是在 LSTM 基础上提出的一种 LSTM 变体, 将 LSTM 的输入层到隐藏层和隐藏层到隐藏层之间的全连接状态替换为卷积连接, 对

LSTM 无法充分利用空间信息进行了改进。LSTM 在处理图像数据时需要将图像数据转为一维向量, 无法处理原图像数据的空间结构信息。相比 LSTM 模型, ConvLSTM 能够更好地提取时间序列图像中的时空结构信息。ConvLSTM 模型公式表示如下:

$$I_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i) \quad (6)$$

$$F_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f) \quad (7)$$

$$C_t = F_t \circ C_{t-1} + I_t \circ \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \quad (8)$$

$$O_t = \sigma(W_{xout} * X_t + W_{hout} * H_{t-1} + W_{cout} \circ C_{t-1} + b_{out}) \quad (9)$$

式 (6) — (9) 中字母表达含义和式 (2) — (5)

相同, 输入变量的全连接换成了卷积操作, \* 表示卷积操作, 由图 2 的 ConvLSTM 内部结构可以看出, 输入门、输出门和遗忘门对于输入和隐藏层都进行了卷积操作; 其中  $W_{ci} \circ C_{t-1}$ 、 $W_{cf} \circ C_{t-1}$  和  $W_{cout} \circ C_{t-1}$  表示输入门、遗忘门和输出门与上一细胞状态的 peephole (Gers 和 Schmidhuber, 2000) 连接。如图 2 所示, peephole 连接在每个门都加入细胞状态的信息, 由于网络单元可能会有门状态为 0 的情况, 导致缺乏重要信息, 添加 peephole 操作可以改进这个缺点。

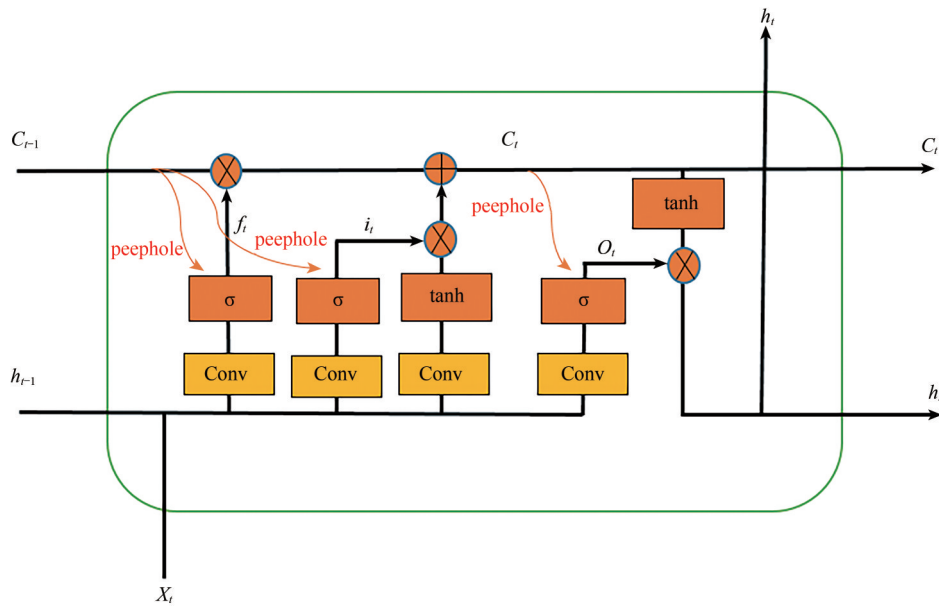


图2 ConvLSTM 内部结构

Fig. 2 Internal structure of ConvLSTM

本文的 Stacked ConvLSTM 网络结构如图 3 所示, Stacked ConvLSTM 在 ConvLSTM 基础上进行改进, 由多个 ConvLSTM 堆叠而成。假设网络输入时间序列数据为  $(X_1, X_2, \dots, X_n)$ , 每次输入一个  $X_i$  经过  $n$  次数据输入提取输入的时序特征, 经过  $k$  层 ConvLSTM 网络单元前向传递网络学习到的时空信息, 最后一层三维卷积层输出最后结果  $(X_{n+1}, X_{n+2}, \dots, X_{n+m})$ , 将输出结果与目标序列  $(\hat{X}_{n+1}, \hat{X}_{n+2}, \dots, \hat{X}_{n+m})$  计算损失函数, 并反向传递进行网络训练。

实验中预测序列  $(\hat{X}_{n+1}, \hat{X}_{n+2}, \dots, \hat{X}_{n+m})$  和目标序列  $(X_{n+1}, X_{n+2}, \dots, X_{n+m})$  采用的损失函数是均方误差 MSE 函数, 损失函数的计算可表示为

$$Loss = \sum_{i=n+1}^{n+m} \|X_i - \hat{X}_i\|^2 \quad (10)$$

式中,  $X_i$  表示第  $i$  时间点目标序列值,  $\hat{X}_i$  表示第  $i$  时间点网络模型的预测值。

### 2.3 火烧迹地检测精度评价指标

混淆矩阵也称作误差矩阵, 是表示分类精度的一个  $n \times n$  矩阵 (Liu 等, 2007)。在本文中火烧迹地和非火烧迹地属于二分类问题, 混淆矩阵如表 1 所示。

基于混淆矩阵, 进一步计算以下常用的结果评价指标, 包括: 精确度 ( $P$ )、召回率 ( $R$ )、准确度 ( $Acc$ )、F1-score 值 ( $F1$ )。以上 4 个评估指标公式如下:

$$P = \frac{TP}{TP + FP} \quad (11)$$

$$R = \frac{TP}{TP + FN} \quad (12)$$

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (13)$$

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \quad (14)$$

式中,  $TP$  表示正确分为火烧迹地的像素个数,  $FP$  表示错误分为火烧迹地的像素个数,  $TN$  表示正确

分为非火烧迹地的像素个数,  $FN$  表示错误分为非火烧迹地的像素个数;  $P'$  是分为火烧迹地的总像元个数,  $N'$  是分为非火烧迹地的总像元个数;  $P$  是真实火烧迹地像元个数,  $N$  是真实非火烧迹地像元个数。

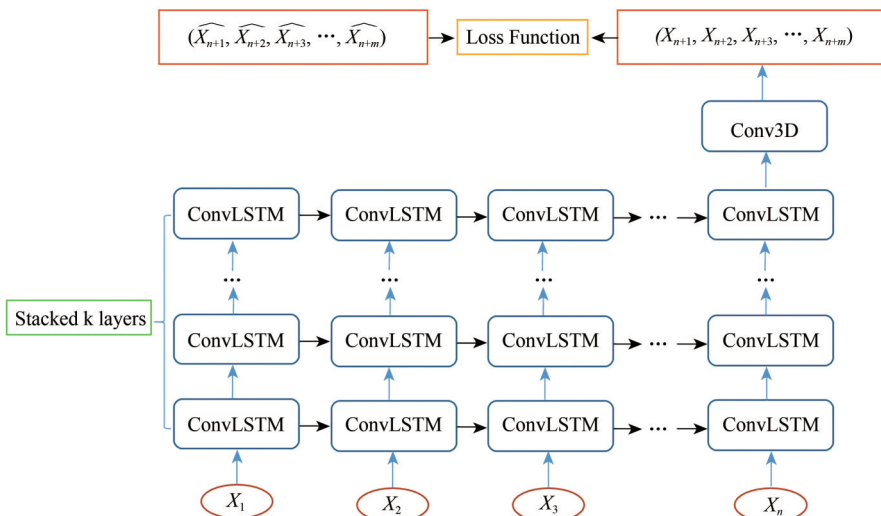


图3 堆叠 ConvLSTM  
Fig. 3 Stacked ConvLSTM

表1 火烧迹地检测混淆矩阵

**Table 1 Confusion matrix of fire detection classification**

		分类类别		
		火烧迹地	非火烧迹地	
实际类别	火烧迹地	$TP$	$FN$	$P$
	非火烧迹地	$FP$	$TN$	$N$
		$P'$	$N'$	$P+N$

注:  $TP$  表示正确分为火烧迹地的像素个数,  $FP$  表示错误分为火烧迹地的像素个数,  $TN$  表示正确分为非火烧迹地的像素个数,  $FN$  表示错误分为非火烧迹地的像素个数;  $P'$  是分为火烧迹地的总像元个数,  $N'$  是分为非火烧迹地的总像元个数;  $P$  是真实火烧迹地像元个数,  $N$  是真实非火烧迹地像元个数。

### 3 数据处理与实验

#### 3.1 数据处理

基于深度学习框架 Keras 完成了 Stacked ConvLSTM 的构建 (Chollet 等, 2015), 实验设备环境是 NVIDIA GeForce GTX1080 GPU。研究区域 I 的 MODIS 时间序列数据集图像像素区域大小为  $512 \times 512$ , 研究区域 II 图像像素区域大小为  $400 \times 400$ 。在这两个研究区域, MODIS 时间序列中观测到森林火烧迹地信息的变化时间点分别为 2009 年和 2017 年的第 126 天, 对应本研究的时间序列中

第 3 个时间点影像。如图 4 是在两个研究区分别取像素区域大小为  $10 \times 10$  的平均值在时间轴的变化, 发生火灾的时间点在最后一年 (图 4 中红色显示), 是需要进行预测的年份。

本文使用滑动窗口法 (Cheng 等, 2016) 将研究所用 MODIS 数据划分为监督训练的输入输出集。在时间上, 输入序列长度为 1 年时序数据 14 幅影像, 输出序列长度为后 1 年时序数据 14 幅像, 滑动的步长为一年即 14 幅影像。例如, 2001 年数据作为输入, 2002 年数据作为输出, 下一个滑动窗口的输入为 2002 年数据, 输出为 2003 年数据, 依次类推至时序最后一年。最后一次滑动作为测试数据, 研究区域 I 输入 2008 年数据预测 2009 年, 研究区域 II 输入 2016 年数据预测 2017 年。在空间上, 将研究区域 I 的像素区域大小  $512 \times 512$  切分成像素区域大小  $64 \times 64$  的子块, 研究区域 II 的像素区域大小  $400 \times 400$  分成像素区域大小  $50 \times 50$  的子块。对处理后的数据集, 除了最后一年的数据作为测试数据, 剩下的数据按照 8 : 2 划分为训练集和验证集。

#### 3.2 时序预测与火烧迹地检测

本文对 Stacked ConvLSTM 模型中的超参数设置了不同的值进行实验, 得到的最优参数分别为

卷积核大小 $3\times 3$ , Batch Size 大小设置为 16, 采用的优化器为 RMSprop (Zhong 等, 2012), 学习率

大小设置为 0.001, 梯度衰减值为 0.9, 网络的循环训练次数为 100 次。

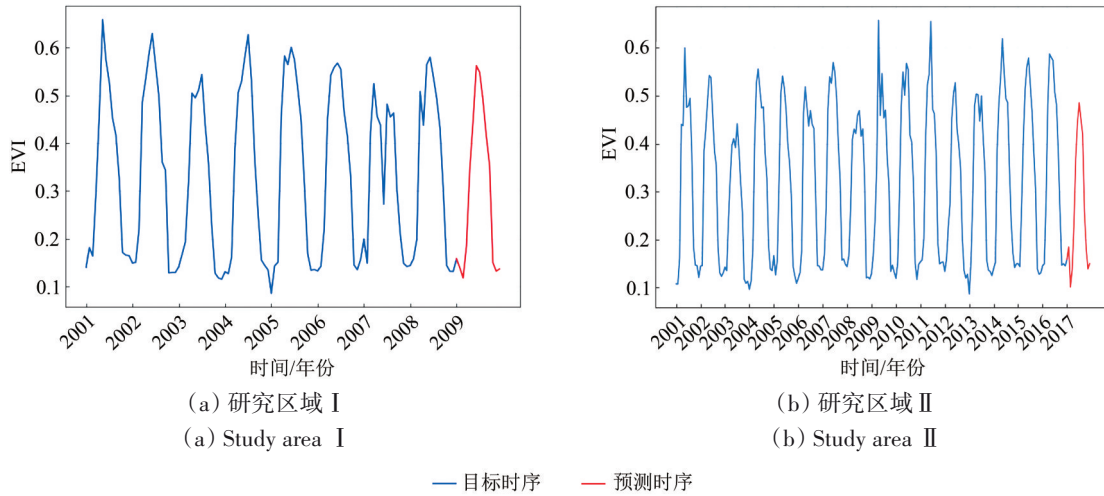


图4 研究区域 EVI 在时间轴上变化

Fig. 4 EVI varies along the time axis in the study area

Stacked ConvLSTM 的 Stacked Layers 及每层的单元个数对于网络模型的学习能力影响很大, 因此对不同的网络结构进行测试。网络所测试的层数以及单元数是在 Stacked ConvLSTM 的应用以及实验中得出的经验值, 根据预测精度取在本文数据表现最佳的网络层数和单元数, 作为的森林火烧迹地检测网络结构。

研究区域 I 测试结果如表 2 所示, 由平均绝对误差 MAE (Mean Absolute Error) 和均方误差 MSE (Mean Squared Error) 指标可以看出, 研究区域 I 的预测效果最好的是 32-32-32-batch16 结构; 研究区域 II 的结构如表 3 所示, 效果最好的结构是 64-32-batch16。

表 2 研究区域 I 预测精度

Table 2 Predict precision of study area I

Stacked Layers-Batch	MAE	MSE
32-32-32-batch16	0.0375	0.0188
32-32-32-batch32	0.0666	0.0234
64-32-batch16	0.0409	0.0193
128-batch16	0.0450	0.0193

表 3 研究区域 II 预测精度

Table 3 Predict precision of study area II

Stacked Layers-Batch	MAE	MSE
32-32-32-batch16	0.0554	0.0213
32-32-32-batch32	0.0517	0.0202
64-32-batch16	0.0428	0.0185
128-batch16	0.0460	0.0196

将两个研究区域的 Stacked ConvLSTM 最佳网络结构用于 MODIS 数据的森林火烧迹地检测。取预测的时间序列以及验证的目标序列中像素区域大小  $10\times 10$  求平均值作图结果如图 5 所示, 可以直观地看出, 未变化区域的预测值和目标序列值很接近, 而变化区域在第 3 个时间点 (MODIS 时间序列的第 126 天) 变化之后一段时间 EVI 值下降很大, 与预测值相距悬殊, 植被在后半年有所恢复, EVI 值回升至和预测值相近。

对测试序列中每一个像元采用经验阈值的方法确定是否为变化像元 (Kong 等, 2018)。根据检测精度分别调整两个研究区域的阈值, 获得最佳阈值为 -0.05。如图 6 所示, 阈值分割线表示预测序列每个像素的 EVI 值减去 0.05。并提取目标序列首次低于阈值分割线的像元作为变化像元, 在测试时间序列的第 3 个时间点检测到变化。

## 4 结果分析评价

本文利用 Stacked ConvLSTM、Stacked LSTM 以及 bfast 算法在两个区域的 MODIS 时间序列中提取森林火烧迹地, 并将火烧迹地检测结果与 ESA 发布的 Fire\_CCI 5.1 火烧迹地产品进行了对比分析。地面真值数据分别由 2009 年 5 月 23 日 Landsat-5 TM 影像和 2017 年 5 月 25 日 MODIS 影像目视解译生成。

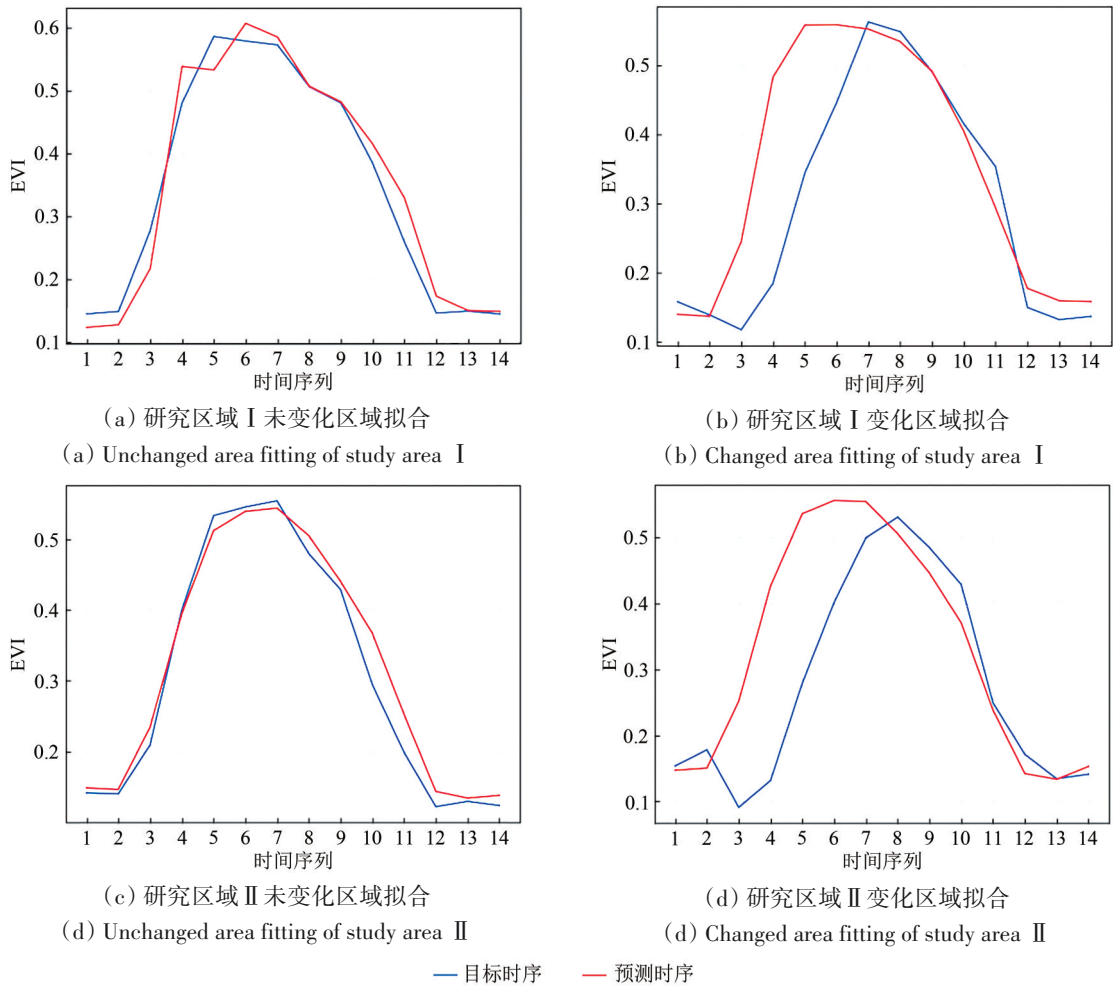


图5 未变化和变化区域时序拟合  
Fig. 5 Unchanged and changed area fitting

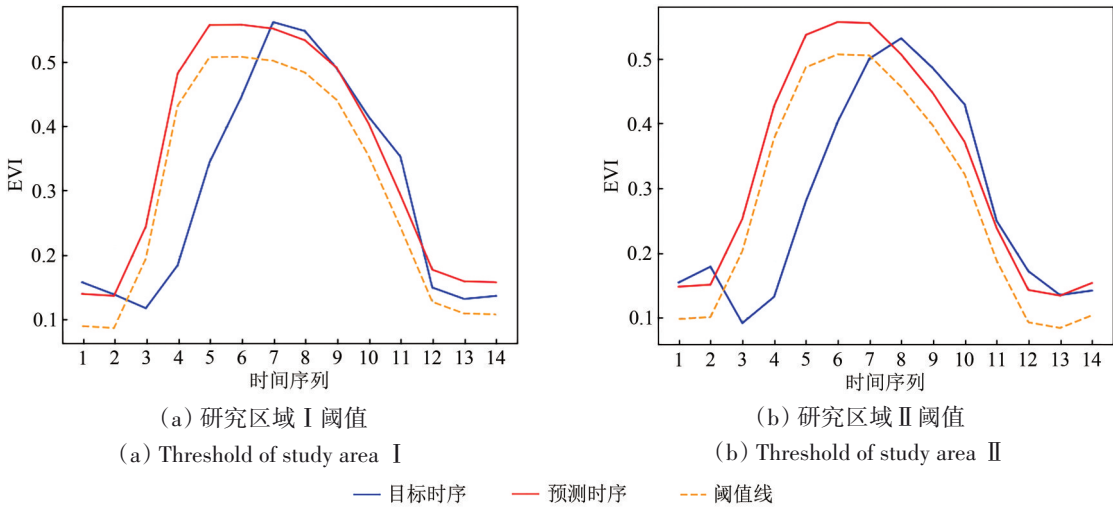


图6 变化区域阈值  
Fig. 6 Threshold of changed area

火烧迹地检测结果如图7所示,从目视效果来看,在研究区域 I, Stacked ConvLSTM 检测的结果比 Stacked LSTM 和 bfast 算法错误检测点少,并

且在空间分布也保持较高连续性;在研究区域 II, Stacked ConvLSTM 检测到了较完整的火烧迹地区域。在研究区域 I, Stacked ConvLSTM 检测的结

果相对 Stacked LSTM 以及 bfast 算法更少细碎的像素, 错误检测点少并且在空间分布也保持相对高的连续性, 检测的火烧迹地精度相对于另外两种方法也更高, Fire\_CCI 5.1 的错检区域比 Stacked ConvLSTM 更大。在研究区域 II, Stacked ConvLSTM

检测到了较完整的火烧迹地区域, Stacked LSTM 以及 bfast 算法的错误检测像元相对更多, 并且在空间连续性上表现更差, Fire\_CCI 5.1 相比 Stacked ConvLSTM 漏检像元更多。

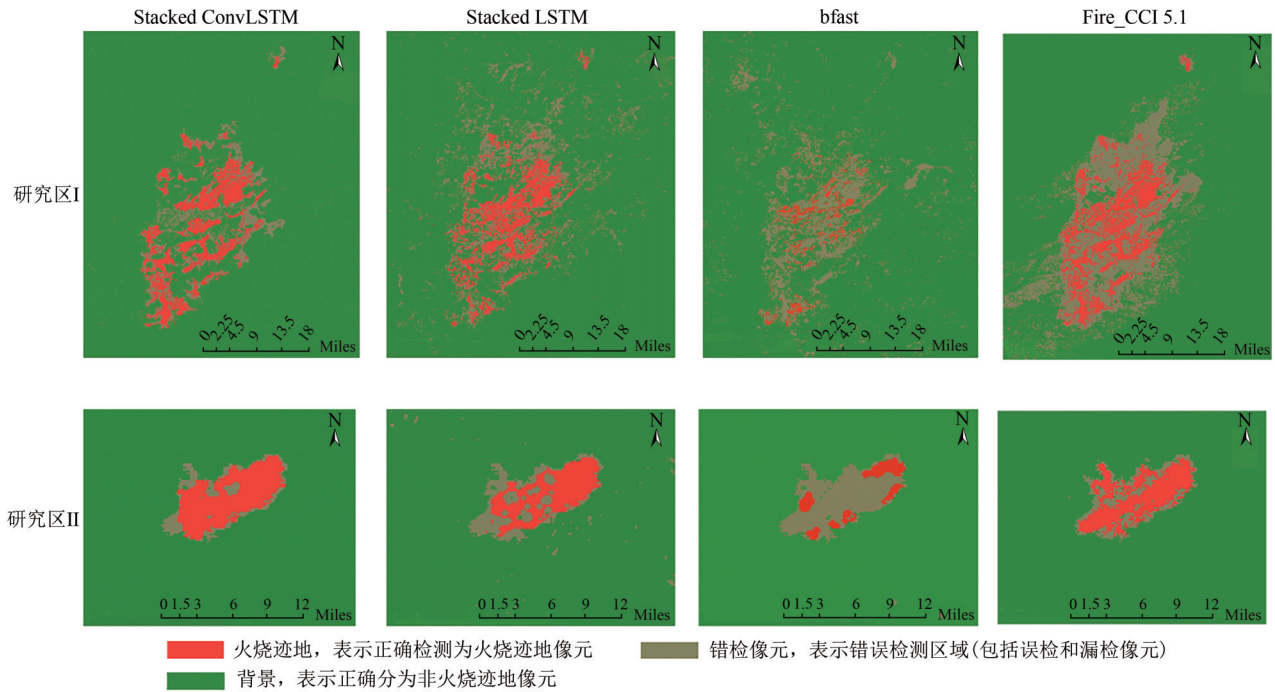


图7 森林火烧迹地检测结果图

Fig. 7 Forest burned areas detection results

由表4和表5可以看出, 在两个研究区域 Stacked ConvLSTM 的检测精度相比 Stacked LSTM 和 bfast 算法更好。在研究区域 I, Stacked ConvLSTM 的精确度比 Stacked LSTM 和 bfast 算法分别高出 0.120 和 0.405, 并且召回率、准确度和 F1-score 也更高, Fire\_CCI 5.1 召回率虽更高, 由于错检区域较大, 其他精度指标远低于 Stacked ConvLSTM; 在研究区域 II, Stacked ConvLSTM 精确度达 0.924, 召回率、准确度和 F1-score 相比 Stacked LSTM 和 bfast 算法以及 Fire\_CCI 5.1 更高。

表4 研究区域 I 火烧迹地检测精度表

研究方法	精确度	召回率	准确度	F1-score
Stacked ConvLSTM	0.827	0.802	0.975	0.814
Stacked LSTM	0.720	0.707	0.956	0.713
bfast	0.422	0.418	0.921	0.420
Fire_CCI 5.1	0.275	0.952	0.825	0.427

表5 研究区域 II 火烧迹地检测精度表

Table 5 Precision table of burned areas detection in study area II

研究方法	精确度	召回率	准确度	F1-score
Stacked ConvLSTM	0.924	0.867	0.982	0.895
Stacked LSTM	0.907	0.785	0.978	0.842
bfast	0.597	0.268	0.929	0.370
Fire_CCI 5.1	0.922	0.674	0.964	0.779

将两个区域的精度表绘制成柱状统计图, 见图8所示, 可以更直观地看出, 在研究区域 I, Stacked ConvLSTM 比 Stacked LSTM 的检测精度要好, 并远高于 bfast 算法; 在研究区域 II, Stacked ConvLSTM 的召回率以及 F1-score 比另两种方法高很多。总体来看, Stacked ConvLSTM 的检测表现比 Stacked LSTM 和 bfast 算法表现更好。

Stacked ConvLSTM 在两个区域检测的火灾发生的时间点分别在 MODIS 时间序列中的 2009 年和 2017 年的第 126 天, 而实际发生的火灾时间点分



别在2009年的4月27日和2017年的5月2日,即第116天和第122天。以上3种方法在两个区域的检测时间点如表6和表7所示, Stacked ConvLSTM和Stacked LSTM比实际发生的时间点分别迟10天和4天, bfast算法的检测时间点相对另两种方法要

迟一步。检测时间点的推迟和MOD13Q1数据的时间分辨率有关, MOD13Q1是16天合成产品, 发生的火灾点和获取的数据点正好相差了一个时间周期, 导致了检测时间点的延迟。

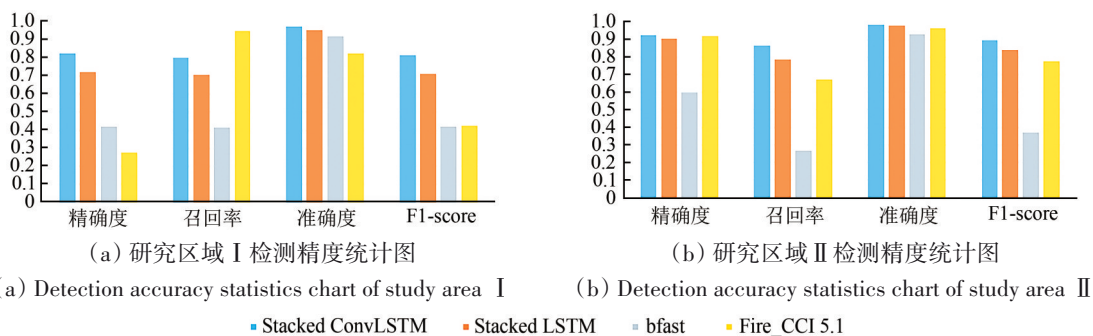


图8 森林火烧迹地检测精度柱状统计图

Fig. 8 Precision of forest burned areas column chart

表6 研究区域 I 变化时间点

Table 6 Change detected time of study area I

	检测点(DOY)	变化点(DOY)
Stacked ConvLSTM	126	116
Stacked LSTM	126	116
bfast	145	116

注: DOY表示Day of Year,一年中的第多少天,常用于时间序列数据中的表述。

表7 研究区域 II 变化时间点

Table 7 Change detected time of study area II

	检测点(DOY)	变化点(DOY)
Stacked ConvLSTM	126	122
Stacked LSTM	126	122
bfast	145	122

注: DOY表示Day of Year,一年中的第多少天,常用于时间序列数据中的表述。

## 5 结论

本文提出将Stacked ConvLSTM用于时间序列森林火烧迹地检测,利用Stacked ConvLSTM学习时空特征的优势,在保持结果具有较好空间连续性的基础上避免了具有主观性的后处理操作,实现端到端提取森林火烧迹地信息,提升了森林火烧迹地的提取精度。基于两个发生过特大火灾区域的MODIS时间序列数据,定性和定量分析Stacked ConvLSTM、Stacked LSTM以及经典的bfast算法火烧迹地提取结果,并将3种方法的火烧迹地

提取结果与Fire\_CCI 5.1火烧迹地产品进行了对比。研究结果表明,在检测精度和空间分布的连续性上,Stacked ConvLSTM检测的火烧迹地结果都有很大提高;此外,在火灾发生时间点检测效果上,由于MODIS时间分辨率(MOD13Q1数据为16天合成数据)为16天,Stacked ConvLSTM和Stacked LSTM均落后真实火灾时间点一步,但均比bfast算法的在时间上更准确一步。

Stacked ConvLSTM在森林火烧迹地检测存在一些不足之处。首先,本文基于单个EVI指数开展Stacked ConvLSTM在森林火烧迹地检测的有效性研究,更多植被相关指数及其组合的检测效果值得进一步研究;其次,确定火烧迹地区域的方法是基于经验阈值,有待研究自动的阈值计算方法;再次,由于MODIS数据的分辨率属于中低分辨率,Stacked ConvLSTM模型的卷积核设置太大反而会模糊空间信息,因此提取的空间信息有限,在更高空间分辨率和时间分辨率的影像数据上,Stacked ConvLSTM模型的时空预测表现可能会更好;最后,网络模型训练中是固定时间长度的输入输出,只能学习有限的时空序列信息,这大大限制了网络在时空预测上的表现。因此,在后续的研究中,考虑在Stacked ConvLSTM基础上加入一些可以输入输出不定长时序数据和更高效的措施如Encoder-Decoder结构以及Attention机制等,提高网络在时空序列上的预测效果。

志 谢 此次实验过程中, 在袁媛老师的帮助下获取了本文的实验数据, 在此表示由衷感谢!

## 参考文献(References)

- Bolton D K, Coops N C and Wulder M A. 2015. Characterizing residual structure and forest recovery following high-severity fire in the western boreal of Canada using Landsat time-series and airborne Lidar data. *Remote Sensing of Environment*, 163: 48-60 [DOI: 10.1016/j.rse.2015.03.004]
- Cao Y C, Wu Z P, Zhou Y F, Wei S J and Feng X Q. 2020. Forest fire recognition method based on recurrent neural network. *Forestry and Environmental Science*, 36(5): 34-40 (曹毅超, 吴泽鹏, 周宇飞, 魏书精, 封晓强. 2020. 基于循环神经网络的森林火灾识别研究. *林业与环境科学*, 36(5): 34-40) [DOI: 10.3969/j.issn.1006-4427.2020.05.006]
- Chance C M, Hermosilla T, Coops N C, Wulder M A, White J C. 2016. Effect of topographic correction on forest change detection using spectral trend analysis of Landsat pixel-based composites. *International Journal of Applied Earth Observation and Geoinformation*, 44:186 - 194[DOI:10.1016/j.jag.2015.09.003]
- Cheng M, Xu Q, Lv J M, Liu W Y, Li Q and Wang J P. 2016. MS-LSTM: a multi-scale LSTM model for BGP anomaly detection// 2016 IEEE 24th International Conference on Network Protocols (ICNP). Singapore: IEEE: 1-6 [DOI: 10.1109/ICNP.2016.7785326]
- Chollet F, others. 2015. Keras. GitHub. [https://github.com/fchollet/keras]
- Chuvieco E, Lizundia-Loiola J, Pettinari M L, Ramo R, Padilla M, Tansey K, Mouillot F, Laurent P, Storm T, Heil A and Plummer S. 2018. Generation and analysis of a new global burned area product based on MODIS 250 m reflectance bands and thermal anomalies. *Earth System Science Data*, 10(4): 2015-2031 [DOI: 10.5194/essd-10-2015-2018]
- Chuvieco E, Yue C, Heil A, Mouillot F, Alonso-Canas I, Padilla M, Pereira J M, Oom D and Tansey K. 2016. A new global burned area product for climate assessment of fire impacts. *Global Ecology and Biogeography*, 25(5): 619-629 [DOI: 10.1111/geb.12440]
- Gers F A and Schmidhuber J. 2000. Recurrent nets that time and count.// *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks/IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*. Como: IEEE: 189-194 [DOI: 10.1109/IJCNN.2000.861302]
- Hilker T, Wulder M A, Coops N C, Linke J, McDermid G, Masek J G, Gao F and White J C. 2009. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sensing of Environment*, 113(8): 1613-1627 [DOI: 10.1016/j.rse.2009.03.007]
- Hochreiter S and Schmidhuber J. 1997. Long short-term memory. *Neural Computation*, 9(8): 1735-1780 [DOI: 10.1162/neco.1997.9.8.1735]
- Huang C Q, Goward S N, Masek J G, Thomas N, Zhu Z L and Vogelmann J E. 2010. An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, 114(1): 183-198 [DOI: 10.1016/j.rse.2009.08.017]
- Huang C Q, Goward S N, Schleeuwis K, Thomas N, Masek J G and Zhu Z L. 2009. Dynamics of national forests assessed using the Landsat record: case studies in eastern United States. *Remote Sensing of Environment*, 113(7): 1430-1442 [DOI: 10.1016/j.rse.2008.06.016]
- Huete A, Didan K, Miura T, Rodriguez E P, Gao X and Ferreira L G. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1/2): 195-213 [DOI: 10.1016/S0034-4257(02)00096-2]
- Huo L Z, Tang P, Zhang Z and Tuia D. 2015. Semisupervised classification of remote sensing images with hierarchical spatial similarity. *IEEE Geoscience and Remote Sensing Letters*, 12(1): 150-154 [DOI: 10.1109/LGRS.2014.2329713]
- Jordan M I. 1997. Serial order: a parallel distributed processing approach. *Advances in Psychology*, 121: 471-495 [DOI: 10.1016/S0166-4115(97)80111-2]
- Kennedy R E, Cohen W B and Schroeder T A. 2007. Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment*, 110(3): 370-386 [DOI: 10.1016/j.rse.2007.03.010]
- Kennedy R E, Yang Z Q and Cohen W B. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr-temporal segmentation algorithms. *Remote Sensing of Environment*, 114(12): 2897-2910 [DOI: 10.1016/j.rse.2010.07.008]
- Khan S H, He X M, Porikli F and Bennamoun M. 2017. Forest change detection in incomplete satellite images with deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 55(9): 5407-5423 [DOI: 10.1109/TGRS.2017.2707528]
- Kim S, Hong S, Joh M and Song S K. 2017. DeepRain: ConvLSTM network for precipitation prediction using multichannel radar data. *arXiv Preprint arXiv: 1711.02316*[DOI: 10.48550/arXiv.1711.02316]
- Kong Y L, Huang Q Q, Wang C Y, Chen J B, Chen J S and He D X. 2018. Long short-term memory neural networks for online disturbance detection in satellite image time series. *Remote Sensing*, 10(3): 452 [DOI: 10.3390/rs1003 0452]
- Långkvist M, Karlsson L and Loutfi A. 2014. A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recognition Letters*, 42: 11-24 [DOI: 10.1016/j.patrec.2014.01.008]
- Liu C, Frazier P, Kumar L. 2007. Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment*, 107(4): 606 - 616[DOI:10.1016/j.rse.2006.10.010]
- Markham B L, Helder D L. 2012. Forty-year calibrated record of earth-reflected radiance from Landsat: a review. *Remote Sensing of Environment*, 122:30 - 40[DOI:10.1016/j.rse.2011.06.026]
- Peng D L, Wu C Y, Li C J, Zhang X Y, Liu Z J, Ye H C, Luo S Z, Liu X J, Hu Y and Fang B. 2017. Spring green-up phenology products

- derived from MODIS NDVI and EVI: intercomparison, interpretation and validation using National Phenology Network and AmeriFlux observations. *Ecological Indicators*, 77: 323-336 [DOI: 10.1016/j.ecolind.2017.02.024]
- Reddy D S and Prasad P R C. 2018. Prediction of vegetation dynamics using NDVI time series data and LSTM. *Modeling Earth Systems and Environment*, 4(1): 409-419 [DOI: 10.1007/s40808-018-0431-3]
- Ruiz J A M, Lázaro J R G, Cano I D Á and Leal P H. 2014. Burned area mapping in the North American boreal forest using Terra-MODIS LTDR (2001—2011): a comparison with the MCD45A1, MCD64A1 and BAGEOLAND-2 products. *Remote Sensing*, 6(1): 815-840 [DOI: 10.3390/rs6010815]
- Shi X J, Chen Z R, Wang H, Yeung D Y, Wong W K and Woo W C. 2015. Convolutional LSTM network: a machine learning approach for precipitation nowcasting//Proceedings of the 28th International Conference on Neural Information Processing Systems. Montreal: MIT Press: 802-810 [DOI:10.48550/arXiv.1506.04214]
- Sun H B, Wang Z L, Zhu Z W, Huang F F, Gan X H and Zhang W Q. 2018. Stoichiometry characteristics of dominant tree species and soil in *Heritiera littoralis* community in Baguang Wetland Park in Shenzhen. *Forestry and Environmental Science*, 34(6): 15-20 (孙红斌, 王佐霖, 朱子维, 黄芳芳, 甘先华, 张卫强. 2018. 深圳市坝光湿地园银叶树群落优势树种与土壤生态化学计量特征分析. *林业与环境科学*, 34(6): 15-20 [DOI: 10.3969/j.issn.1006-4427.2018.06.003])
- Sun G F, Qin X L, Liu S C, et al. Potential analysis of typical vegetation index for identifying burned area[J]. *Remote Sensing for Land and Resources*, 2019, 31( 1) : 204-211 (孙桂芬, 覃先林, 刘树超, 等. 2019. 典型植被指数识别火烧迹地潜力分析. *国土资源遥感*, 31(1):204-211) [DOI:10.6046 /gtzyyg.2019.01.27]
- Verbesselt J, Zeileis A and Herold M. 2012. Near real-time disturbance detection using satellite image time series. *Remote Sensing of Environment*, 123: 98-108 [DOI: 10.1016/j.rse.2012.02.022]
- Wang Z S, Wei S J, Xie J H, Zhong Y X, Li X C, Tang H H, Wu Z P, Zhou Y F and Li Q. 2019. Effect of forest fire combustion environment on forest fire suppression. *Forestry and Environmental Science*, 35(2): 84-88 (王振师, 魏书精, 谢继红, 钟映霞, 李小川, 唐洪辉, 吴泽鹏, 周宇飞, 李强. 2019. 林火燃烧环境对灭火效果的影响研究. *林业与环境科学*, 35(2): 84-88) [DOI: 10.3969/j.issn.1006-4427.2019.02.014]
- Wang Z X, Liu C and Huete A. 2003. From AVHRR-NDVI to MODIS-EVI: advances in vegetation index research. *Acta Ecologica Sinica*, 23(5): 979-987 (王正兴, 刘闯, Huete A. 2003. 植被指数研究进展: 从 AVHRR-NDVI 到 MODIS-EVI. *生态学报*, 23(5): 979-987) [DOI: 10.3321/j.issn:1000-0933.2003.05.020]
- Wei S J, Luo S S, Luo B Z, Li X C, Wang Z S, Wu Z P, Zhou Y F, Zhong Y X and Li Q. 2020. Occurrence regularity of forest fire under the background of climate change. *Forestry and Environmental Science*, 36(2): 133-143 (魏书精, 罗斯生, 罗碧珍, 李小川, 王振师, 吴泽鹏, 周宇飞, 钟映霞, 李强. 2020. 气候变化背景下森林火灾发生规律研究. *林业与环境科学*, 36(2): 133-143) [DOI: 10.3969/j.issn.1006-4427.2020.02.019]
- Wu J W, Sun L Y, Ji R P, Feng R, Yu W Y and Zhang Y S. 2020. Research progress and prospect of remote sensing on extracting burned areas information. *Journal of Catastrophology*, 35(4): 151-156 (武晋雯, 孙龙彧, 纪瑞鹏, 冯锐, 于文颖, 张玉书. 2020. 火烧迹地信息遥感提取研究进展与展望. *灾害学*, 35(4): 151-156) [DOI: 10.3969/j.issn.1000-811X.2020.04.028]
- Yang C. 2013. Monitoring Regional Forest Disturbance by Remote Sensing: A Case Study of Wuning County. Nanjing: Nanjing University of Information Science and Technology (杨辰. 2013. 区域森林植被扰动遥感监测研究——以武宁县为例. 南京: 南京信息工程大学)
- Yuan Y, Lin L, Huo L Z, Kong Y L, Zhou Z G, Wu B and Jia Y. 2020. Using an attention-based LSTM encoder-decoder network for near real-time disturbance detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13: 1819-1832 [DOI: 10.1109/JSTARS.2020.2988324]
- Zhang X Y, Friedl M A, Schaaf C B, Strahler A H, Hodges J C F, Gao F, Reed B C and Huete A. 2003. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84(3): 471-475 [DOI: 10.1016/S0034-4257(02)00135-9]
- Zhang Z M, Tang D, He G J, Long T F and Wei M Y. 2020. Global 30-meter resolution burned area products over the globe based on Landsat 8 images. *China Scientific Data*, 5(4): 91-96 (张兆明, 唐朝, 何国金, 龙腾飞, 魏明月. 2020. 全球30米分辨率火烧迹地产品. *中国科学数据(中英文网络版)*, 5(4): 91-96) [DOI: 10.11922/sciencedb.976]
- Zhao C J, Zhang P, Zhu J, Wu C R, Wang H M and Xu K L. 2019. Predicting tongue motion in unlabeled ultrasound videos using convolutional LSTM neural networks//2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP): 5926-5930. IEEE, Piscataway, NJ. [DOI: 10.1109/ICASSP.2019.8683081]
- Zhong X G, Mohamed A S and Florian M W. 2012. Testing the accuracy of query optimizers//Proceedings of the Fifth International Workshop on Testing Database Systems (DBTest '12. Association for Computing Machinery, 11:1-6 [DOI:10.1145/2304510.2304525]
- Zhu Z. 2017. Change detection using Landsat time series: a review of frequencies, preprocessing, algorithms, and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, 130: 370-384 [DOI: 10.1016/j.isprsjprs.2017.06.013]
- Zhu Z, Fu Y C, Woodcock C E, Olofsson P, Vogelmann J E, Holden C, Wang M, Dai S and Yu Y. 2016. Including land cover change in analysis of greenness trends using all available Landsat 5, 7, and 8 images: a case study from Guangzhou, China (2000—2014). *Remote Sensing of Environment*, 185: 243-257 [DOI: 10.1016/j.rse.2016.03.036]
- Zhu Z and Woodcock C E. 2014. Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 144: 152-171 [DOI: 10.1016/j.rse.2014.01.011]

# Forest burned area detection with time series data based on Stacked ConvLSTM

LI Shujun<sup>1,2</sup>, ZHENG Ke<sup>1</sup>, TANG Ping<sup>1</sup>, HUO Lianzhi<sup>1</sup>, YUAN Yuan<sup>3</sup>

1. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China;

3. Nanjing University of Posts and Telecommunications, Nanjing 210042, China

**Abstract:** As the largest land cover, forests play an important role in human living environment, biological habitat, and global carbon cycle. Forest health is directly related to global ecological security and sustainable development of human society. In recent years, urban construction, disasters, forest management and deforestation, and other factors have caused different degrees of disturbance to forests. It is important to determine the exact time point and spatial range of forest burned area for forest damage assessment, management, carbon accounting, and forest restoration management. Owing to the continuity of spatial distribution of forest burned areas, most of the existing methods of forest burned area extraction use the two-step treatment strategy of first classification and then post-processing to suppress the effect of false alarm pixels. In this paper, a spatiotemporal detection method, Stacked ConvLSTM, is proposed for the detection of forest fire tracks in time series. This method avoids subjective post-processing operations on the basis of maintaining better spatial continuity of the results, and achieves end-to-end extraction of forest burned area information, which improves the extraction accuracy of forest fire-burning land. This paper proposes to use Stacked ConvLSTM to detect forest disturbance in time and space. Combined with the characteristics of ConvLSTM in extracting temporal and spatial characteristics from long-term historical series, it can predict the change trend of vegetation in a period of time in the future, and accurately determine the time point and spatial range of forest disturbance. ConvLSTM is an LSTM variant proposed on the basis of LSTM. The full connection state from input layer to hidden layer and from hidden layer to hidden layer of LSTM is replaced by convolution connection, which can make full use of spatial information. Compared with single-pixel-based methods, ConvLSTM can extract the spatiotemporal structure information of time series images at the same time, which is better for spatiotemporal analysis. In this paper, Stacked ConvLSTM is used to detect the temporal and spatial distribution of forest burned areas, predict the change trend of vegetation in a period of time in the future, and determine the presence of forest burned areas by comparing with the newest time-series images. With MODIS long time series data, based on the historical time series of Yinanhe Forest Farm of Zhanhe Forestry Bureau in Heilongjiang Province and Beidahe Forest Farm of Bilaha Forestry Bureau in Inner Mongolia from 2001—2008 and 2001—2016, the extraction results of burned areas were compared with Stacked LSTM and bfast algorithm. The Stacked ConvLSTM, Stacked LSTM, and bfast algorithms were used to extract forest burned areas from MODIS time series in both regions, and to compare the detection results with the Fire\_CCI 5.1 burned areas products released by ESA. Results show that, firstly, from the visual effect, in study area I, the error detection of Stacked ConvLSTM is fewer than that of Stacked LSTM and bfast algorithm and maintains high continuity in spatial distribution. In study Area II, Stacked ConvLSTM detected a more complete area of fire. Secondly, in study area I, Stacked ConvLSTM was 0.120 and 0.405 more accurate than Stacked LSTM and bfast algorithms, respectively. Moreover, the recall rate, accuracy, and Fire\_CCI 5.1 F1-score were higher. In study area I, the accuracy of Stacked ConvLSTM is 0.924 had a higher recall rate, accuracy, and F1-score than Stacked LSTM, bfast algorithms, and Fire\_CCI 5.1. The detection accuracy of ConvLSTM model in space is higher than that of the other two methods, and its continuity of detection results in space is better. The detection effect of ConvLSTM model is equivalent to that of Stacked LSTM in time, but both of them are closer to the real fire time point than bfast algorithm. Results show that Stacked ConvLSTM has advantages in obtaining the change trend of forest long-term historical series for spatiotemporal prediction, and improves the detection accuracy of forest fire to a certain extent.

**Key words:** Stacked ConvLSTM, time series, spatiotemporal prediction, forest burned area

**Supported by** National Natural Science Foundation of China (No. 41971396, 41901356); Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDA19080301)