

doi:10.12196/j.issn.1000-3274.2022.02.006

张翔, 孙宪坤, 胡峻, 等. 结合数据扩增与残差收缩网络的地震短临预测[J]. 地震, 2022, 42(2): 74-88.
ZHANG Xiang, SUN Xian-kun, HU Jun, et al. Short-term and imminent earthquake prediction combined with data augmentation and residual shrinkage network[J]. *Earthquake*, 2022, 42(2): 74-88.

结合数据扩增与残差收缩网络 的地震短临预测^{*}

张翔¹, 孙宪坤¹, 胡峻², 尹京苑², 熊玉洁¹

(1. 上海工程技术大学电子电气工程学院, 上海 201620; 2. 上海市地震局, 上海 200062)

摘要:中强地震发生前存在地表异常增温现象, 热红外信息可以成为地震短临预测的途径之一。然而, 地震预测研究常存在可供分析的震例样本不足的问题。本文基于 MODIS 地表温度数据, 先通过深度卷积生成对抗网络(DCGAN)对地震前的地温数据做扩增处理, 再将扩增后的地温数据输入深度残差收缩网络进行特征提取并预测未来短期内是否存在发生 5 级及以上地震事件的可能性。实验针对中国地震较多的中西部地区, 将地表温度数据依照地震实际发生情况标记为发生地震数据和未发生地震数据, 样本比为 3:1, 分不同的预测时间段进行比较。结果显示, 5 日预测的准确率最高为 73.86%, 正确预测发生占实际发生的比例为 68.09%。多次实验准确率曲线趋向稳定, 证明该预测方法有很好的实用性。基于 MODIS 数据结合数据扩增与残差收缩网络的预测方法为短临中强地震预测研究提供了一种新思路。

关键词:地震预测; MODIS; 数据扩增; 深度卷积生成对抗网络; 深度残差收缩网络

中图分类号:P315.75

文献标识码:A

文章编号:1000-3274(2022)02-0074-15

引言

地震作为自然灾害之一, 具有非常大的破坏性, 严重威胁着人民群众生命财产安全和社会经济发展。1975 年海城 7.3 级地震, 鉴于当时地震系统对地震事件的预测, 政府及时下达了疏散命令, 有效降低了人员的伤亡^[1]。由此可见, 地震预测研究工作对于保障人民群众生命和财产安全具有重大的意义。

近年来, 各国学者开展了通过地震前兆信息分析来预测地震的相关研究工作, 虽然取得了一定的进展, 但还不成熟, 地震预测至今仍是一个世界性难题^[2]。目前, 地震预测方

^{*} 收稿日期: 2021-10-25; 修改回日期: 2021-12-01

基金项目: 国家重点研发计划项目(2019YFC1509202); 国家自然科学基金项目(62006150); 上海青年科技英才扬帆计划项目(19YF1418400)

作者简介: 张翔(1996-), 男, 江苏扬州人, 硕士研究生, 主要从事模式识别与智能系统研究。

通讯作者: 孙宪坤, 副教授。E-mail: xksun@sues.edu.cn

法大致有三种: 运用地震前兆信息、识别地质活动和利用统计学的方法。地震预测的时间阶段通常分为长期、中期、短期以及临震四种, 短临地震预测往往与地震前兆密切相关, 地震前兆指标大致涉及重力、地形变、地电、地下水化学、地磁、地下水物理、地温和气象等^[3]。

大量研究表明, 中强地震发生前地表温度存在异常增温的现象, 地表温度异常是地震前兆之一。卫星红外遥感技术具有范围广、全天候、信息量大、精度高和观测条件限制少的优点。遥感技术的快速发展使利用热红外遥感数据监测地震发生前后的热信息变化成为可能。同时大量的实验表明, 热红外信息在地震短临预测中具有独特优势, 是地震预报的可能途径之一^[4]。早期利用卫星热红外数据提取地震前兆信息大多是通过大量原始遥感图像的直观分析与推断来分析地震热信息异常的基本特征。由于受到构造活动、日照、大气及地表植被状况等因素影响, 研究地表热异常与地震相关性往往需要较强的图像识别能力及遥感理论知识, 且大多是经验性的依据, 大大增加了研究的复杂性与困难。因此, 分析方法的不断尝试和改进是深入研究热异常与地震相关性的关键。李勇等^[5]利用分类回归树法找出热红外异常数据与地震震级的关系, 结果显示, 影响地震震级的主要因素是最大异常亮温面积。熊攀^[6]利用小波分析法结合 2008 年四川汶川 8.0 级地震和 2007 年云南普洱 6.4 级地震资料研究发现, NOAA 卫星长波辐射数据在汶川地震发生前震中 300 km 范围内均出现明显的异常。陈顺云等^[7]基于 2006 年 3 月至 2008 年 2 月遥感影像和实测地表温度, 从差值、相关性及时频特性等几个方面进行对比, 结果显示夜间热红外图像与实测地表温度之间的相关性优于白天; 而分析地表温度的长期变化时, 选用白天或夜间的数据没有实质性区别。该研究初步明确了图像与实测地温的相关性, 有助于更合适地选取遥感资料开展地震预测研究。解涛等^[8]基于中国静止气象卫星 FY-2E 的热红外遥感亮温数据利用小波能谱法对 2012 年云南彝良 5.7 级地震进行分析, 结果表明地震发生前震中附近出现了热红外异常, 同时异常区域出现扩大、集中并形成条带的过程。上述研究通过提取与地震的相关性遥感热红外信息, 探索利用热红外信息开展地震预报。

随着计算机技术的快速发展, 机器学习方法在不断地进步并得到了广泛的应用。同样地, 机器学习也被应用到了地震学研究领域。近年来, 基于机器学习的地震预测研究也取得了一定的效果。Moustra 等^[9]使用震电信号 (Seismic Electric Signals, SES) 作为神经网络的输入, 将大地震震级预测的准确率提升至 83.56%。Asencio-Cortés 等^[10]在日本东京及周围 200 km 以内的区域试验人工神经网络对于地震震级预测的可行性与准确性并取得了良好的效果。Li 等^[11]提出了基于深度学习模型 (Deep Learning Earthquake Discrimination, DLEP) 的地震震级范围预测, 以 8 个地震活动性参数作为模型的输入, 采用动态损失函数作为输出, 针对中国新疆、中国四川、中国山东—江苏, 青藏高原, 菲律宾, 日本, 美国芝加哥和洛杉矶等 8 个具有不同特征的区域进行算法性能测试, 结果显示, DLEP 相比其他经典机器学习模型具有更准确的性能。Huang 等^[12]将地震事件投影到地形图中并生成图像数据集, 通过卷积神经网络 (Convolutional Neural Network, CNN) 预测未来 30 天内是否会发生大地震。Nicolis 等^[13]使用深度神经网络长期短期记忆 (Long Short-Term Memory, LSTM) 和卷积神经网络预测智利地震的发生率。与经典模型 CNN 相比, LSTM 既能识别大地震的特征模式, 又能近似预测地震最大强度。

本文利用热红外遥感范围广、全天候、信息量大、精度高和观测条件限制少的优点,

直接将热红外图像 MODIS 地表温度数据输入到深度卷积网络中进行特征提取。由于深度学习网络需要大量的数据,而地震震例样本有限,因此,先将热红外图像通过深度卷积生成对抗网络(Deep Convolutional Generative Adversarial Networks, DCGAN)进行扩增,扩大样本数据量。为检验扩增后数据的有效性,本文利用 RST 算法进行热异常提取,对比真实图像与生成图像热异常的分布范围以及分布位置是否吻合。残差网络是对卷积神经网络的一种改进,采用了恒等路径来优化参数。在深度残差网络中,梯度不仅逐层反向传播,而且通过恒等路径直接传递到之前的层。由于其出色的参数优化能力,近几年深度残差网络多被应用于故障诊断。深度残差收缩网络是对残差网络的改进,在残差网络基础上增加了软阈值化、注意力机制,在提取特征时关注重点信息而忽略次要信息,是一种很好的降噪方法。中强地震前短时间的热红外图像在震源附近会出现异常,云层等影响可以通过注意力机制和软阈值化方法进行去除,从而有效提取热异常信号。

1 实验数据

1.1 震例选取

采用 2006—2016 年中国中西部地区四川、新疆、青海等地发生的 120 个 5 级及以上地震,包括 2014 年 2 月 12 日新疆于田 7.3 级地震(36.1°N , 82.5°E)、2013 年 4 月 20 日四川芦山 7.0 级地震(30.3°N , 103°E)、2010 年 4 月 14 日青海玉树 7.1 级地震(33.2°N , 96.6°E)、2008 年 5 月 12 日四川汶川 8 级地震(31°N , 103.4°E)、2008 年 3 月 21 日新疆于田 7.3 级地震(35.6°N , 81.6°E)等震例。无震样本则随机选取 40 个。

1.2 数据预处理

MODIS 是中分辨率成像光谱辐射计(Moderate-Resolution Imaging Spectroradiometer)的简称,是搭载在 TERRA 和 AQUA 卫星上的传感器。两颗卫星均为极地轨道太阳同步卫星,分别为上午星和下午星,从而可以使 MODIS 每 1~2 d 获得一次全球观测数据(包括白天的可见光图像以及白天/夜晚的红外图像)。研究所用的热红外遥感数据是 1 km 分辨率夜间 MODIS 地表温度数据,即 MYD11A1 1 km $\times$ 1 km 地表温度数据,原始数据为 .hdf 文件,通过 ERDAS 9.2 软件对数据进行预处理,从 12 个数据集中挑选出 LST_

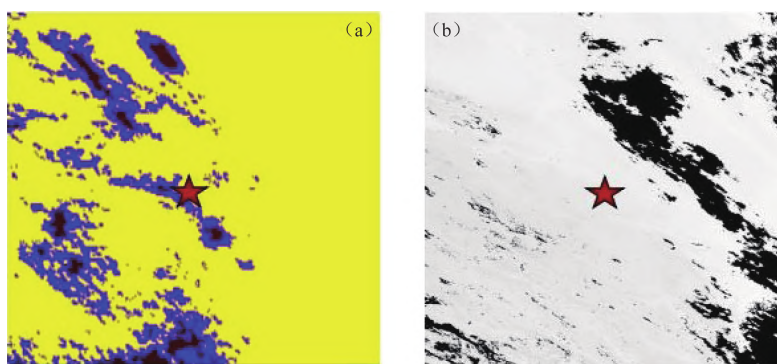


图 1 MODIS 地表温度图像转换

(a) MODIS 地表温度图像; (b) 格式转换后地表温度图像

Fig.1 MODIS surface temperature image conversion

(a) MODIS surface temperature image; (b) Surface temperature image after format conversion

Night_1 km 夜间温度数据, 重投影为 Geographic Lat/Lon 投影, 该投影具有角度和方向不变的特性。

以地震震中位置为中心, 往南北和东西 256 km 的范围裁剪成尺寸为 256×256 大小的图像, 使用最邻近内插法进行重采样, 将数据中被大气干扰的低质量数据直接设置为 0。最后, 通过 IDL 编程将 .hdf 数据中的夜间温度数据 LST_Night_1 km 输出成 .jpg 图片, 图 1 是震中为四川汶川 (31°N , 103.4°E) 的 MODIS LST 地表温度数据在 ERDAS 9.2 中的成像以及转换后的 jpg 格式图像, 红色五角星为地震震中位置。

2 研究方法

2.1 数据扩增

深度学习需要大量的数据来训练。基于深度学习的热红外遥感地震前兆研究同样需要大量震例图像, 但这一需求往往难以满足。因此, 本文提出对热红外图像做数据扩增处理, 扩大训练的数据量。数据扩增是扩大数据规模, 提高数据质量常用的一种方法, 该方法不仅能够有效地解决数据集规模小、质量差、难以获取等问题, 而且也能够有效地减少深度学习网络模型的过拟合问题, 提高网络模型的泛化性能。目前, 数据扩增技术可分为监督数据扩增和无监督数据扩增。监督数据扩增包括单样本数据的几何变换, 如翻转、旋转、缩放、仿射等在图像形态上做变换的简单方法; 在图像内容上, 有颜色变换、添加噪声、随机擦除、模糊图像等方法。无监督数据扩增的方法包括两种: 一是通过模型学习适合当前任务的数据扩充方法; 二是通过学习源数据生成新数据。

本文采用的 MODIS 热红外图像, 图像上每个像素对应地表温度值, 在一定的区域范围内, 由于地形的多样性, 地表温度数据呈现不规则分布, 但是同一地区同一段时间内温度还是存在一定年变化规律, 该数据属于单样本数据同时覆盖全图且分布不规则。利用温度数据预测地震发生, 关键在于分析提取地震前温度的异常信息, 由于历史震例样本有限, 对应的发生地震的热红外图像数量也有限, 如果采用简单的翻转、旋转、缩放等监督数据增强方法增强地震前的温度图像, 并不能很好地进行温度数据的扩增^[14]。

2.1.1 生成对抗网络

生成对抗网络 (Generative Adversarial Network, GAN) 是一种生成数据的神经网络方法, GAN 网络由生成器和判别器组成, 通过两者不断对抗学习的方式来进行训练。生成器网络将随机采样的样本作为输入, 生成与训练集中数据分布相似的新数据样本; 判别器网络则判别生成器生成样本的真实性。生成器的目的是生成更真实的数据欺骗判别器, 而判别器的目的是将生成的数据和真实数据区分开来。该网络的思想受到了博弈论中二人零和博弈的启发, 通过两者不断对抗互相提升^[15]。图 2 为 GAN 网络模型结构示意图。

2.1.2 深度卷积生成对抗网络

深度卷积生成对抗网络 (Deep Convolutional Generative Adversarial Networks, DCGAN) 是一种将卷积神经网络与原始 GAN 网络相结合的无监督学习模型, 相较于 GAN 模型结合了深度神经网络概念, 优化了网络结构, 提高了样本生成质量和收敛速度。DCGAN 的优势是使用了卷积层, DCGAN 的生成器 G 接收一个随机噪声 z , 通过噪声生成图像 $G(z)$ 。判别器 D 接收一张图像 x , 输出图像 x 是真实图像的概率为 $D(x)$, DCGAN 的数学模型为

$$V(D, G) = E_x - p_{\text{data}}(x) [\ln D(x)] + E_z - p_z(z) \{ \ln [1 - D(G(z))] \} \quad (1)$$

其中, $p_{\text{data}}(x)$ 是真实图像的数据分布, $p_z(z)$ 是生成器 G 产生的图像数据分布, DCGAN 的任务则是通过训练调整生成器参数增大 $D(G(z))$ 的值, 同时调整判别器增大 $D(x)$ 使得 $D(G(z))$ 减小, 最终使两个模型达到一个均衡点。

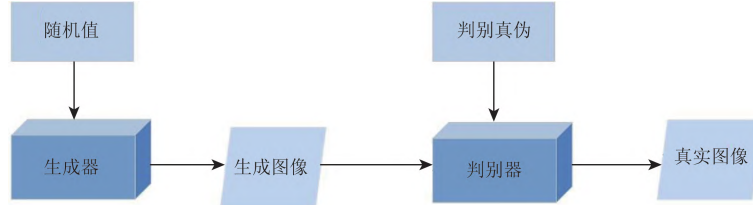


图 2 GAN 网络模型结构

Fig. 2 GAN network model structure

2.1.3 基于改进 DCGAN 的热红外数据扩增

本文采用改进 DCGAN 网络结构, 在 GAN 的基础上加入特征融合技术。图 3 为网络结构示意图, 其中生成器网络由 5 层反卷积层组成, 利用反卷积学习图像信息进行上采样, 反卷积核大小为 5×5 , 卷积核个数由 1024 逐次缩小一半, 生成图像为 256×256 。判别器网络采用多特征融合技术以提高生成图像的质量, 由 5 层卷积层和 1 层特征融合层组成。卷积层的卷积核大小为 5×5 , 特征融合层融合最后 3 层卷积层的特征并归一化到 8×8 尺寸, 倒数第三层进行最大池化, 将 64 张 32×32 大小的特征图归一化到 64 张 8×8 大小。倒数第二层进行最大池化, 将 128 张 16×16 大小的特征图归一化到 128 张 8×8 大小。然后, 融合到倒数第一层 256 张 8×8 大小的特征图, 融合后 8×8 特征图为 224 张, 进行重塑, 变成 $8 \times 8 \times 224$ 大小的特征向量。最后, 通过判别器对其进行识别, 输出真假图像的概率值, 概率值 $D(x)$ 为“0”时表示鉴别为生成图像; 为“1”时, 则表示是真实图像。训练过程中, 生成器的目标就是尽量生成真实的图片去欺骗判别器。而判别器的目标就是尽量辨别出生成的假图像和真实的图像。这样就构成了一个动态的“博弈过程”, 最终的平衡点为

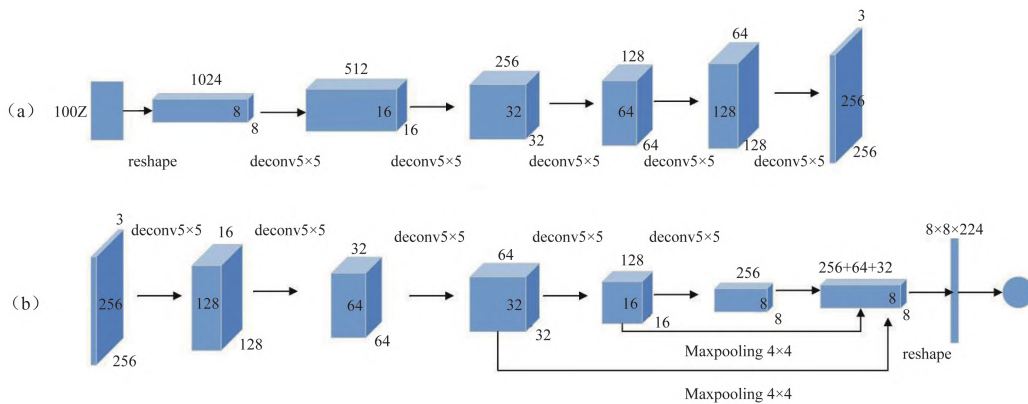


图 3 改进 DCGAN 网络结构

(a) 生成器; (b) 判别器

Fig. 3 Improved DCGAN network structure

(a) Generator; (b) Discriminator

纳什均衡点。不断重复训练直到判别器输出概率接近 0.5 时,说明此时判别器无法判别生成图像的真假,训练结束。

2.1.4 热红外异常提取算法

针对地震历史样本不足而导致深度学习样本有限的问题,提出利用深度卷积生成对抗网络(DCGAN)对热红外数据进行扩增处理,对于扩增后的数据是否可作为样本提取热异常特征进行地震预测,本文利用 RST 算法对扩增后的生成图像进行有效性检验。

RST 算法是一种基于多时相数据的分析方法,目的在于突出时空域上相对空间域的异常信息,RST 模型由 RAT(Robust AVHRR Techniques)模型演化而来,最初用于云层、火山喷发、森林火灾等方面的监测^[16]。Tramutoli 等^[17]改进了 RST 算法并将其应用于地震前热红外监测。该方法将地震引起的热红外异常作为信号,其他因素引起的异常作为噪声,引入了地震热异常信噪比的概念。近年来,研究者利用该方法开展了大量震例研究,并在多个震例研究中发现了明显持续的热红外异常现象,该算法以多年同期的遥感卫星数据为基础,提取出相对于背景场的异常值,公式为

$$R = \frac{\Delta T(x, y, t) - \mu_{\Delta T}(x, y)}{\sigma_{\Delta T}(x, y)} \quad (2)$$

其中, R 为热异常稳健估计值, $\Delta T(x, y, t)$ 是相对于区域均值的差值 [$\Delta T(x, y, t) = T(x, y, t) - T(t)$], $\mu_{\Delta T}(x, y)$ 是 $\Delta T(x, y, t)$ 多年相同位置、相同时刻的平均值, $\sigma_{\Delta T}(x, y)$ 为对应的标准差, $T(t)$ 为 t 时刻影像均值。统计学中,大于两倍标准差的数值属于异常范畴,已有研究都将 R 阈值设置为 2,即当 $R \geq 2$ 时属于异常范畴^[18]。

2.2 深度残差收缩网络模型

地物本身的热过程是复杂的,能使地表温度发生改变的因素还包括地球基本温度场、云层、寒暖气流以及人类社会活动等,从这些复杂背景中提取与地震相关联的弱信号是利用热异常预测地震的关键。深度残差收缩网络是在深度残差网络基础上,加入注意力机制和软阈值化函数。该网络的优势即从复杂信号背景中提取有用的弱信号,残差收缩网络里面增加的注意力机制和软阈值化,实际上是一种很好的降噪方法。中强地震前短时间的热红外图像在震源附近会出现异常,存在的包括云层等影响可以通过注意力机制和软阈值化方法进行去除,从而有效提取热异常信号。因此,本文选用深度残差收缩网络模型训练热红外遥感数据,提取与地震相关联的热红外信号特征,并预测未来发生 5 级以上地震事件的概率。

2.2.1 深度残差网络

传统的神经网络可能会存在信息丢失、梯度消失、爆炸等问题。深度残差网络是改进的深度卷积神经网络,该网络的特点是容易优化,能够通过增加相当的深度来提高准确率,其内部的残差块使用了跳跃连接,缓解了在深度神经网络中增加深度带来的梯度消失问题。卷积层训练后会产生新的特征图,此时模型会自发地去学习不同特征图之间的权重,通常一个输入 a 经过卷积层形成一个新的输出,即普通映射 $H(a)$,而在残差模块中,经过卷积后的映射就为 $Y(a) + a$ 。使用多层卷积去接近 $H(a)$,不如去近似接近残差 $Y(a)$, $Y(a)$ 表示为

$$Y(a) = H(a) - a \quad (3)$$

残差网络将每一层卷积层产生的新特征进行叠加,增加网络深度的同时减少网络的计算参

数量,防止梯度消失。本文所搭建的网络输入层尺寸大小为 11×11 ,卷积层为 Conv2d_1、Conv2d_2、Conv2d_3、Conv2d_4,大小分别为 9×9 、 6×6 、 5×5 、 3×3 ,同时 Conv2d_1 使用两层残差学习单元,输出的 8 通道的 3×3 和 8 通道的 1×1 的特征残差单元通过最大池化融合到 Conv2d_3 卷积层,Conv2d_2 卷积层也使用两层残差学习单元,输出的 16 通道的 3×3 和 16 通道的 1×1 特征残差单元通过最大池化融合到 conv2d_3 卷积层。特征融合技术可以让网络层更加全面地提取特征,防止梯度消失。在网络的输出端包括一个全局的均值池化层和全连接层。为了避免过拟合现象,模型采用了包括 BN(Batch Normalization)^[19]、Dropout^[20]和激活函数^[21]在内的正则化策略。图 4 为本文深度残差网络结构示意图。

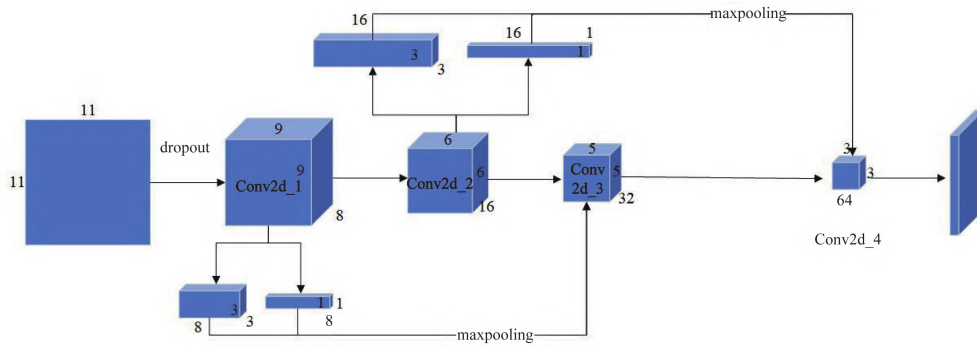


图 4 深度残差网络结构

Fig. 4 Deep residual network structure

2.2.2 注意力机制

在日常生活中,人们对于所看到的事物会进行特征检索,着重观察重要的信息而忽略次要的信息,注意力机制则利用人们这一视觉机制,提取图像中有用的信息而忽略次要的信息。近几年,在注意力机制应用上效果较好的是 SENet(Squeeze-and-Excitation Networks),该算法通过一个小型网络自动学习每个特征的重要度,提取重要特征而抑制相关性较小的特征。算法主要通过 Squeeze 和 Excitation 两个操作实现,Squeeze 操作对特征进行空间维度上的压缩,将二维特征压缩成一个实数,使该实数在某种程度上具有全局感官野。接着,通过 Excitation 给每个特征生成权重,最后通过 Reweight 操作将 Excitation 输出的权重加到先前的特征上,对原始特征重新标定,筛选重要特征^[22]。

2.2.3 软阈值化函数

热红外图像中由于复杂地形、大气层等因素的影响,图像中会存在噪声,软阈值化是深度残差收缩网络降噪的关键步骤,软阈值函数表达式和对 a 的偏导数为

$$y = \text{soft}(a, \tau) = \begin{cases} a - \tau, & a \leq -\tau \\ 0, & |a| \leq \tau \\ a + \tau, & a \geq \tau \end{cases} \quad (4)$$

$$f(a) = \frac{\partial y}{\partial a} = \begin{cases} 1, & a \leq -\tau \\ 0, & |a| \leq \tau \\ 1, & a \geq \tau \end{cases} \quad (5)$$

假设 τ 的值为 1, 通过式(4)和式(5)可以看出软阈值化输入与输出关系, 如图 5 所示。

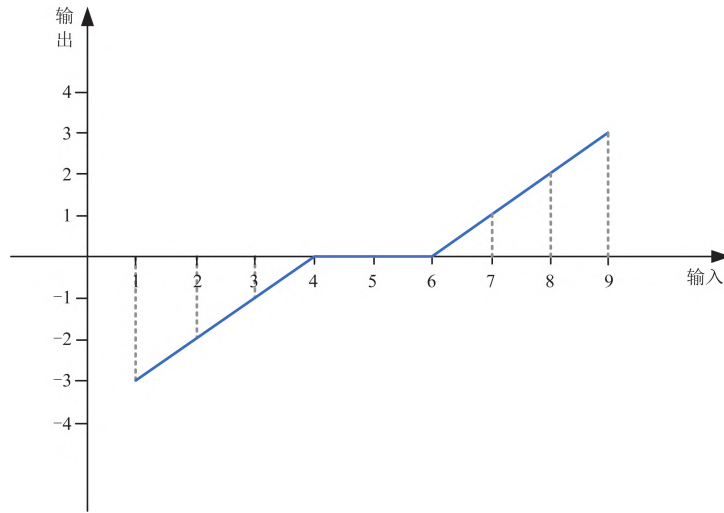


图 5 软阈值化输入与输出关系

Fig. 5 The relationship between soft thresholding input and output

软阈值化是一种非线性变换, 有着和 ReLU 激活函数相似的性质——梯度为 0 或者 1。将部分区域特征值设为“0”, 有利于抑制次要的特征信息。

2.2.4 基于深度残差收缩网络的地震预测方法

本文对不同时间段进行短临地震预测。例如, 5 日预测是指将地震发生前 4 d 的图像标记为“1”, 即发生地震; 未发生地震前 4 d 的图像标记为“0”, 即未发生地震。将提取特征后的深度残差收缩网络模型输入测试数据集, 如果在此时间段内出现预测结果为“1”, 说明在此 5 日以内有发生地震事件的可能。以此类推, 本文分为 9 日预测、7 日预测、5 日预测和 3 日预测。

3 实验结果与分析

3.1 数据增强结果与有效性分析

通过改进 DCGAN 模型对预处理后的已有真实热红外遥感图像进行训练并生成图像。对 2008 年 5 月 12 日四川汶川(31°N, 103.4°E)大地震当天及震前的热红外图像进行扩增, 当训练到 25 个步数时, 实验结果如图 6 所示。可以发现, 生成的热红外图像上温度的灰度分布与真实图像相似度较高。再利用 RST 算法分别提取真实图像和生成图像的热异常信息, 图 6 中分布比较集中的红点为地震发生前热异常信息, 绿色标记为地震震中位置(31°N, 103.4°E), 通过对比发现, 2008 年汶川地震前, 真实图像上热异常分布位置与生成图像上的热异常分布位置大致相同, 热异常的范围大小有所差别但影响不大。同时, 为体现改进 DCGAN 方法对于扩增热红外图像的有效性, 实验增加了传统的 DCGAN 模型进行对比, 在参数不变的情况下, 取消特征融合技术, 即利用传统的 CNN 训练并生成热红外图像。

根据 RST 算法求出生成图像的标准差, MODIS_LST 数据为反射的亮温值, 需要将原始影像的像素值乘以 0.02 得到开尔文温度, 再减去 273.15 得到摄氏度。本文将原始图像

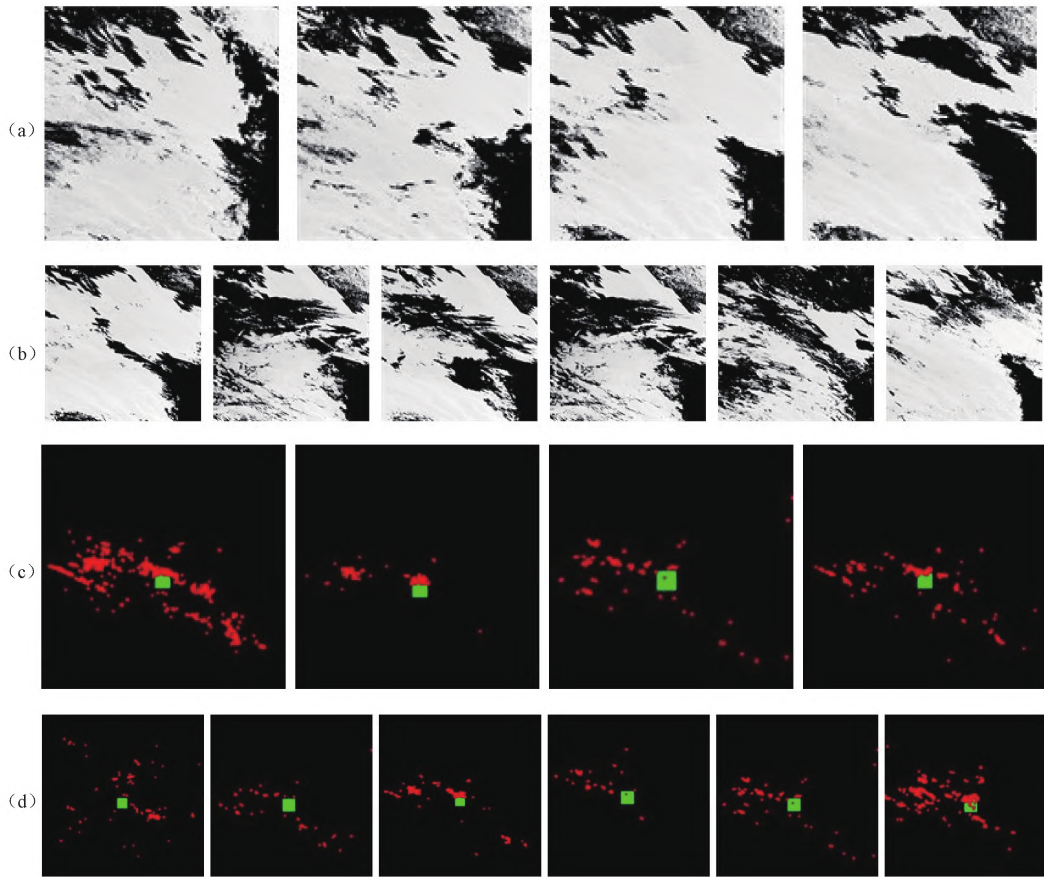


图 6 DCGAN 生成图像及有效性检验

(a) 真实地表温度图像; (b) 生成地表温度图像; (c) 真实图像 RST 算法提取热异常结果;
(d) 生成图像 RST 算法提取热异常结果

Fig. 6 DCGAN generated image and validity check

(a) Real surface temperature image; (b) Generate surface temperature image; (c) Thermal anomaly results extracted by RST algorithm by real image; (d) Thermal anomaly results extracted by RST algorithm by generated image

转化成灰度图像, 灰度值范围为 $0 \sim 255$, 计算真实图像的均值, 将其代入同一时刻生成图像中计算生成图像的标准差, 如式(6)所示。

$$\sigma_{\Delta T}(x, y) = \sqrt{\frac{\sum_{t \in T} [\Delta T(x, y, t) \cdot A(x, y, t) - \mu_{\Delta T}(x, y)]^2}{\sum_{t \in T} A(x, y, t)}} \quad (6)$$

其中, $A(x, y, t)$ 为每张图像中排除像素值为“0”的像元个数, 得到生成图像的标准差, 如表 1 所示。

改进 DCGAN 同一时刻生成图像的标准差均小于 1, 说明生成图像的温度值和真实图像的温度均值差异较小, 传统 DCGAN 方法生成图像的标准差少部分大于 1, 但总体上小于 1, 与真实图像的差异较小。改进 DCGAN 方法生成图像与真实图像差异相比传统 DCGAN 方法更小。结合图 6 热异常分布, 说明通过 DCGAN 生成的震前热红外图像温度值

近似真实图像温度值，增强后的图像可以作为最后预测模型的输入数据。

表 1 生成图像标准差

Table 1 Standard deviation of generated image

	生成图像标准差					
	生成图像 1	生成图像 2	生成图像 3	生成图像 4	生成图像 5	生成图像 6
DCGAN	0.8719	1.2333	0.9480	0.8166	1.0475	0.8397
改进 DCGAN	0.5332	0.7056	0.6183	0.7211	0.8394	0.7792

3.2 不同时间段预测结果对比

将经过数据扩增后的热红外图像直接输入深度残差收缩网络，本文样本为 120 个地震样本与 40 个随机的无地震样本，每个样本根据不同的预测时段选取不同天数的图像，经过扩增后，数据量变为原来的 3 倍。网络的训练集与测试集之比为 3 : 1。网络提取热红外图像特征后，最终进入判别器，判断为“1”，则为发生地震；判断为“0”，则未发生地震。实验对不同预测时间段的预测结果进行了对比，分为 9 日预测、7 日预测、5 日预测和 3 日预测。

本文设置预测指标分别为预测准确率 A，预测发生结果 A_1 和预测未发生结果 A_2 ，预测准确率 A 为正确预测地震是否发生的天数/预测的天数，即

$$A = (\text{正确预测发生天数} + \text{正确预测未发生天数}) / \text{预测的天数} \times 100\% \quad (7)$$

$$A_1 = \text{正确预测发生天数} / \text{实际发生的天数} \times 100\% \quad (8)$$

$$A_2 = \text{正确预测未发生天数} / \text{实际未发生的天数} \times 100\% \quad (9)$$

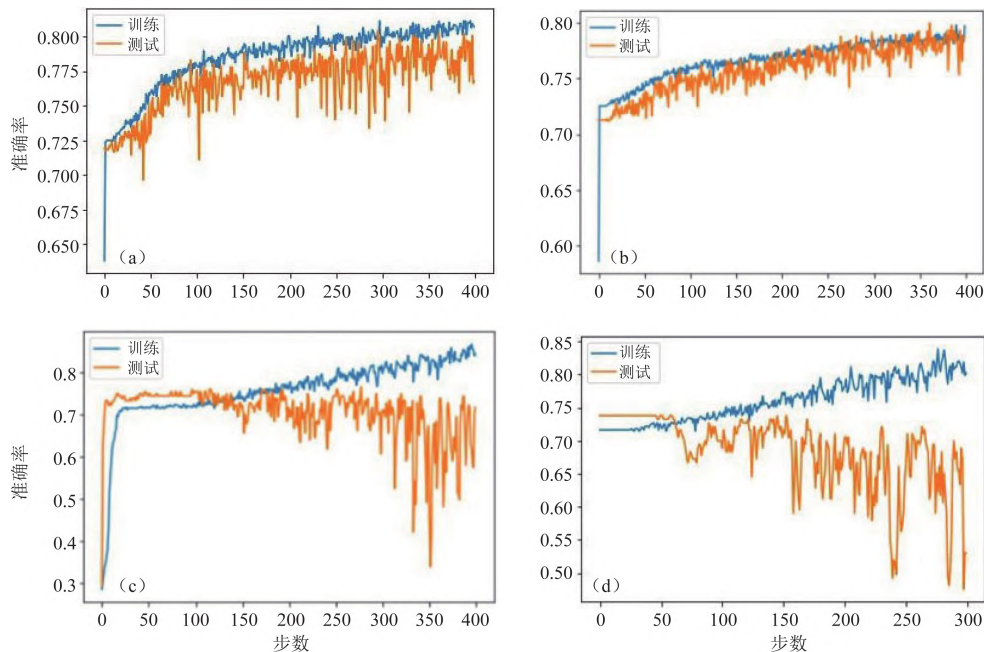


图 7 预测准确率曲线图

(a) 3 日预测准确率曲线；(b) 5 日预测准确率曲线；(c) 7 日预测准确率曲线；(d) 9 日预测准确率曲线

Fig. 7 Curve of forecast accuracy

(a) 3-day forecast accuracy curve; (b) 5-day forecast accuracy curve; (c) 7-day forecast accuracy curve; (d) 9-day forecast accuracy curve

由图 7 可见, 3 日预测和 5 日预测的准确率在训练达到 100 个步数时, 准确率值逐渐稳定在 0.725~0.780, 而 7 日预测和 9 日预测的准确率值在训练达到 200 个步数时, 准确率开始出现大幅度波动, 随着步数的增加, 准确率越来越不稳定。同时对比 3 日预测和 5 日预测可知, 5 日预测的准确率更加稳定, 测试集与训练集的曲线也更加吻合。经多次实验, 结果依旧是 5 日预测的准确率最高并更加稳定。此外, 通过实验对比数据有无扩增对于最终预测结果的影响。实验先将原始数据输入深度残差收缩网络模型进行特征提取并进行预测, 后将扩增后的数据输入网络模型进行预测, 多次实验取平均值, 每个时段分别得出两种结果, 如表 2 所示。经过数据扩增后, 每个预测时段的准确率相比扩增前都有了提高, 其中 5 日预测扩增前准确率为 69.11%, 扩增后准确率提高到 73.86%。正确预测地震发生的天数占实际发生天数的比例也有所提高, 由扩增前的 63.99% 提高到 68.09%。随着预测时间段的增加, 预测准确率也随之降低。7 日预测与 9 日预测的正确预测地震发生比例相比 3 日预测和 5 日预测结果较低。通过数据扩增后的 5 日预测准确率混淆矩阵, 如图 8 所示。

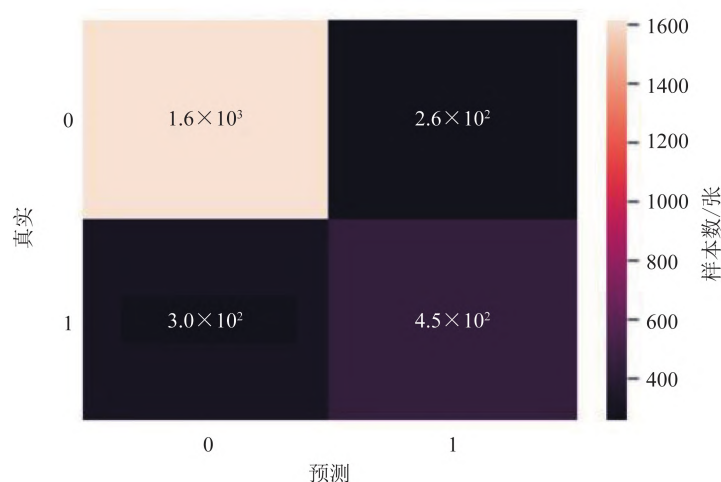


图 8 5 日预测混淆矩阵

Fig. 8 5-day forecast confusion matrix

表 2 不同时段预测结果对比

Table 2 Comparison of forecast results in different periods

3 日预测结果/%			5 日预测结果/%			7 日预测结果/%			9 日预测结果/%		
扩增前		扩增后	扩增前		扩增后	扩增前		扩增后	扩增前		扩增后
A	64.74	72.93	A	69.11	73.86	A	54.83	61.33	A	43.97	57.81
A ₁	59.25	64.71	A ₁	63.99	68.09	A ₁	57.04	63.12	A ₁	49.41	60.99
A ₂	60.09	76.33	A ₂	63.12	79.32	A ₂	50.79	59.67	A ₂	48.57	56.44

由图 8 能够看出, 真实发生地震的样本中, 预测正确的样本量比预测错误的样本量多; 未发生地震的样本中, 预测正确的样本也比预测错误的样本量多。说明预测效果较好。

3.3 不同模型预测结果对比

为体现深度卷积神经网络应用于短临地震预测结果的有效性, 在上述网络参数不变的

情况下,增加 CNN 和深度残差网络两种方法进行地震预测。去掉本文中的残差模块,直接利用 2D-CNN 进行训练并对测试集进行有无地震发生判断。深度残差网络是本文深度残差收缩网络的基础模型,在本文的基础上去掉软阈值化和注意力机制。多次实验得出两种方法的预测准确率,与本文方法进行对比,如表 3 所示。

表 3 不同方法预测结果对比
Table 3 Comparison of forecast results of different methods

	3 日预测结果/%			5 日预测结果/%			7 日预测结果/%			9 日预测结果/%		
	CNN	ResNet	本文方法	CNN	ResNet	本文方法	CNN	ResNet	本文方法	CNN	ResNet	本文方法
A	64.73	70.84	72.93	62.71	70.97	73.86	51.97	58.64	61.33	50.92	55.63	57.81
A ₁	49.11	65.87	64.71	57.91	72.23	68.09	50.01	59.73	63.12	48.75	56.91	60.99
A ₂	73.33	75.90	76.33	64.23	67.33	79.32	52.69	57.67	59.67	51.69	53.45	56.44

由表 3 能够看出,ResNet 和本文方法在 3 日和 5 日的预测准确率都能达到 70%以上,随着预测时间段的增加,预测准确率逐渐降低并达到 60%左右。通过三种方法得出的结果进行对比得出本文方法的预测精度相对更高。因此,通过深度卷积神经网络对热红外图像进行特征提取并研究其与地震的相关性是可行的,结果表明 3 日和 5 日短临预测的准确率均能达到 70%以上。

4 结论与展望

本文基于 MODIS 热红外遥感图像,提出用深度卷积网络的方法针对中国中西部地区大地震进行短临预测。由于深度学习网络在训练时需要大量的样本数据,而地震历史样本有限,本文提出先利用深度生成对抗网络 DCGAN 生成大量热红外图像进行数据扩增,增大数据量,再将扩增后的图像输入到深度残差收缩网络中进行特征提取,经判别器判别是否会发生地震。本文选取 120 个地震样本以及随机的 40 个无地震样本,采用 MODIS LST 地表温度数据。以地震震中为中心,256 km×256 km 为范围作为图像区域,图像尺寸大小为 256×256。实验分不同的时间段进行预测,分别为 3 日预测、5 日预测、7 日预测、9 日预测。结果显示,5 日预测的准确率最高,即正确预测地震是否发生的比例最高。随着预测时间段的增加,准确率值逐渐降低。同时,随着训练步数的增长,3 日和 5 日预测的准确率曲线拟合度更高且更加稳定,7 日和 9 日预测的曲线波动较大。最后,比较得出 5 日的预测准确率最高,为 73.86%,预测发生占实际发生的比例为 68.09%。数据扩增方法也使得地震预测的准确率有所提高,数据量充足避免了网络的过于拟合。本文研究表明,通过结合数据扩增与深度残差收缩网络对预测短期内大地震事件是否发生的方法是可行的。本次试验主要针对中国中西部(四川、青海、甘肃、新疆等地区),研究的震例震中位置与地震时间为已知,试验结果(预测准确性)实际上是热异常信息与地震的非线性相关性,未来实际预测地震并不是已知位置,需要结合专家知识与活动断裂带进一步研究。

参考文献:

- [1] 陈棋福,王克林. 2008 年汶川地震与中国的地震预报[J]. 世界地震译丛,2010,3(5): 34-54.
CHEN Qi-fu, WANG Ke-lin. The 2008 Wenchuan earthquake and earthquake prediction in China

- [J]. *World Earthquake Translation Collection*, 2010, 3(5): 34-54 (in Chinese).
- [2] 徐程, 邹金慧. BP 神经网络在地震预测中的应用研究[J]. *自动化仪表*, 2012, 33(6): 12-14.
XU Cheng, ZOU Jin-hui. Research and application of BP neural network in earthquake prediction [J]. *Automation Instrumentation*, 2012, 33(6): 12-14 (in Chinese).
- [3] 强祖基, 赁常恭, 李玲芝, 等. 卫星热红外图像亮温异常——短临震兆[J]. *中国科学(D 辑: 地球科学)*, 1998, 28(6): 446-455.
QIANG Zu-ji, LIN Chang-gong, LI Ling-zhi, et al. Satellitic thermal infrared brightness temperature anomaly image: short-term and impending earthquake precursors[J]. *Science in China (Series D: Earth Science)*, 1998, 28(6): 446-455 (in Chinese).
- [4] 闫丽莉, 屈春燕, 温少妍, 等. 卫星热红外亮温、气温及地温观测的年变变化对比研究[J]. *地震学报*, 2012, 34(2): 257-266.
YAN Li-li, QU Chun-yan, WEN Shao-yan, et al. A comparison study on annual variation of thermal infrared brightness temperature and land temperature from meteorological stations[J]. *Acta Seismologica Sinica*, 2012, 34(2): 257-266 (in Chinese).
- [5] 李勇, 何洋波, 马丽, 等. 用分类回归树方法研究年亮温异常变[J]. *北京师范大学学报(自然科学版)*, 2003, 39(2): 143-146.
LI Yong, HE Yang-bo, MA Li, et al. A study on abnormality brightness by classification and regression tree methods[J]. *Journal of Beijing Normal University (Natural Science)*, 2003, 39(2): 143-146 (in Chinese).
- [6] 熊攀. 小波方法在地震遥感信息提取中的应用[D]. 北京: 中国地震局地震预测研究所, 2009.
XIONG Pan. Wavelet-based methods for detecting earthquake information in remote sensing data [D]. Beijing: Institute of Earthquake Prediction, China Earthquake Administration, 2009 (in Chinese).
- [7] 陈顺云, 刘培洵, 刘力强, 等. 遥感与实测地表温度的对比分析及在地震研究中的意义[J]. *地球物理学报*, 2011, 54(3): 747-755.
CHEN Shun-yun, LIU Pei-xun, LIU Li-qiang, et al. Comparative analysis between land surface temperatures obtained by field measurement and satellite remote sensing and its implication in earthquake research[J]. *Chinese Journal of Geophysics*, 2011, 54(3): 747-755 (in Chinese).
- [8] 解涛, 康春丽, 卢军, 等. 2012 年彝良 $M_s5.7$ 和 $M_s5.6$ 级地震热辐射亮温异常分析[J]. *地球物理学进展*, 2013, 28(5): 2322-2327.
XIE Tao, KANG Chun-li, LU Jun, et al. Study of brightness temperature anomalies of Yiliang (China) $M_s5.7$ and $M_s5.6$ earthquakes on 7 Sep., 2012[J]. *Progress in Geophysics*, 2013, 28(5): 2322-2327 (in Chinese).
- [9] Moustra M, Avraamides M, Christodoulou C. Artificial neural networks for earthquake prediction using time series magnitude data or Seismic Electric Signals[J]. *Expert Systems with Applications*, 2011, 38(12): 15032-15039.
- [10] Asencio-Cortés G, Martínez-álvarez F, Troncoso A, et al. Medium-large earthquake magnitude prediction in Tokyo with artificial neural networks[J]. *Neural Computing and Applications*, 2017, 28(3): 1043-1055.
- [11] Li R, Lu X, Li S, et al. DEP: A deep learning model for earthquake prediction[C]. In Proceedings of the 2020 IEEE International Joint Conference on Neural Networks, 2020, IJCNN'20: 1-8.
- [12] Huang J P, Wang X A, Zhao Y, et al. Large earthquake magnitude prediction in Taiwan based on deep learning neural network[J]. *Neural Network World*, 2018, 28(2): 149-160.

- [13] Nicolis O, Plaza F, Salas R. Prediction of intensity and location of seismic events using deep learning[J]. *Spatial Statistics*, 2021, 42: 100442.
- [14] 宋冬梅, 臧琳, 单新建, 等. 基于 LST 年趋势背景场的地震热异常提取算法[J]. 地震地质, 2016, 38(3): 680-695.
SONG Dong-mei, ZANG Lin, SHAN Xin-jian, et al. A study on the algorithm for extracting earthquake thermal infrared anomalies based on the yearly trend of LST[J]. *Seismology and Geology*, 2016, 38(3): 680-695 (in Chinese).
- [15] 李冠东, 张春菊, 高飞, 等. 双卷积池化结构的 3D-CNN 高光谱遥感影像分类方法[J]. 中国图象图形学报, 2019, 24(4): 639-654.
LI Guan-dong, ZHANG Chun-ju, GAO Fei, et al. Doubleconvpool-structured 3D-CNN for hyperspectral remote sensing image classification[J]. *Journal of Image and Graphics*, 2019, 24(4): 639-654 (in Chinese).
- [16] Chen Y S, Jiang H L, Li C Y. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2016, 54(10): 6232-6251.
- [17] Tramutoli V, Aliano C, Corrado R, et al. On the possible origin of thermal infrared radiation (TIR) anomalies in earthquake-prone areas observed using robust satellite techniques (RST)[J]. *Chemical Geology*, 2013, 339: 157-168.
- [18] 李雪松, 李劲华, 吕智涵. 基于改进深度残差收缩网络的轴承故障诊断[J]. 青岛大学学报(自然科学版), 2022, 35(2): 38-43+50.
LI Xue-song, LI Jin-hua, LÜ Zhi-han. Bearing fault diagnosis based on improved deep residual shrinkage network[J]. *Journal of Qingdao University (Natural Science Edition)*, 2022, 35(2): 38-43+50 (in Chinese).
- [19] Peng Y H. Super-resolution reconstruction using multiconnection deep residual network combined an improved loss function for single-frame image[J]. *Multimedia Tools and Applications*, 2020, 79(3): 9351-9362.
- [20] Ni J, Shi P, Zhao Y, et al. Fixed-time output consensus tracking for high-order multi-agent systems with directed network topology and packet dropout[J]. *IEEE/CAA Journal of Automatica Sinica*, 2021, 8(4): 817-836.
- [21] 杜圣杰, 贾晓芬, 黄友锐, 等. 面向 CNN 模型图像分类任务的高效激活函数设计[J]. 红外与激光工程, 2022, 51(3): 493-501.
DU Sheng-jie, JIA Xiao-fen, HUANG You-rui, et al. High efficient activation function design for CNN model image classification task[J]. *Infrared and Laser Engineering*, 2022, 51(3): 493-501 (in Chinese).
- [22] 刘成勇, 乔文杰, 陈蜀喆, 等. 基于 LSTM 与注意力机制的船舶航迹预测模型研究[J]. 中国航海, 2021, 44(4): 94-100.
LIU Cheng-yong, QIAO Wen-jie, CHEN Shu-zhe, et al. On ship track prediction with LSTM and attention mechanism[J]. *Navigation of China*, 2021, 44(4): 94-100 (in Chinese).

Short-term and Imminent Earthquake Prediction Combined with Data Augmentation and Residual Shrinkage Network

ZHANG Xiang¹, SUN Xian-kun¹, HU Jun², YIN Jing-yuan², XIONG Yu-jie¹

(1. School of Electronic and Electrical Engineering, Shanghai University of Engineering and Science, Shanghai 201620, China; 2. Shanghai Earthquake Agency, Shanghai 200062, China)

Abstract: Anomalies of temperature increase will occur before moderate and strong earthquakes, and thermal infrared information can be one of the ways to predict the short-term and imminent earthquakes. However, earthquake prediction research often has the problem of insufficient earthquake samples for analysis. Based on the MODIS surface temperature data, the deep convolutional generative adversarial network (DCGAN) is used to amplify the pre-earthquake ground temperature data, and then the amplified ground temperature data is input into the deep residual shrinkage network for feature extraction, and predict whether there is the possibility of $M \geq 5$ earthquake in the short-term future. The experiment is aimed at the central and western regions of China where earthquakes often occur, and the surface temperature data is marked according to the actual occurrence of the earthquake as the data of the occurrence of earthquakes and the data of no earthquakes. The sample ratio is 3 : 1, and the comparison is divided into different prediction time periods. The results show that the highest accuracy rate of 5-day prediction is 73.86%, and the proportion of correct predictions to actual occurrences is 68.09%. The accuracy rate curve of multiple experiments tends to be stable, which proves that the prediction method has good practicability. Based on MODIS data, the prediction method combining data augmentation and residual shrinkage network provides a new idea for short-term, moderate-strong earthquake prediction research.

Key words: Earthquake prediction; MODIS; Data augmentation; Deep convolution generation adversarial network; Deep residual shrinkage network