

文章编号: 2096-1618(2020)06-0589-05

基于 ConvLSTM 的雷达回波外推

赖 灿¹, 王海江¹, 李 静², 徐自励², 刘 涛¹

(1. 成都信息工程大学电子工程学院, 四川 成都 610225; 2. 中国民航局第二研究所, 四川 成都 610041)

摘要:为对雷达回波产品进行预测,研究一种基于 ConvLSTM 的雷达回波外推神经网络模型,采用结构为长短时记忆网络,对其结构做出一些改变后可以使用多维数据作为输入。且采纳单极化 SA 雷达的反射率数据作为数据集。结果证明:(1)采用相当多或者足够多的数据集,并进行数据预处理,包括数据筛选、排序、拟合、滤波。之后进行标准化,将数据按比例缩放,使其反射率值以 0 为中心并分布在一个小的区间内。(2)使用卷积层对反射率数据提取特征,并在隐藏层后面接入全连接层和回归层后,可以得到更优化和更准确地预测结果。(3)利用适当的激活函数和一定的正则化方法可以减少模型的过度拟合,提高模型的训练精度。实验结果表明,对 18 min 内的雷达回波有较好的预测效果,对反射率小于 30 的回波有较好的预测能力。

关键词:ConvLSTM; 雷达回波外推; 数据预处理; 优化; 正则化

中图分类号:TN957.51⁺³

文献标志码:A

doi:10.16836/j.cnki.jcuit.2020.06.001

0 引言

临近天气预报指的是在未来短时间内对一些强度较大、变化较快的,如雷暴、强降水、冰雹之类的天气现象进行预测,这些天气过程通常持续时间短但是破坏性强。一般来说,临近天气预报的可用预报时效性为 2 h 内^[1]。而进行临近天气预报的主要手段是雷达回波外推技术,对回波运行的速度、方向、强度和形态变化等特征进行跟踪和预测。传统的雷达回波外推主要包括单体质心法和交叉相关法。

单体质心法通过计算雷达连续体扫中对流单体的回波参数,识别和跟踪对流单体和对流单体群,处理对流单体的合并和分裂,提供对流单体的移动和演变信息^[2]。Michael Dixon 等^[3]将雷暴视为三维单体加以识别、追踪和外推。Johnson 等^[4]设计的风暴单体识别和跟踪算法(SCIT)通过设定一系列从小到大的阈值识别出三维雷暴单体,并根据雷暴质心位置线性加权外推出雷暴质心以后的位置。乔春贵等^[5]使用质心法对整幅雷达回波图进行线性外推,结果普遍显示单体质心法对强度大且稳定的对流单体追踪效果比较好,而对于强度小和结构比较复杂的回波的识别较为困难。

Hilst 等^[6]先后将交叉相关法应用于跟踪风暴簇的整体运动。交叉相关法首先被 Austin 等^[7]运用一个平均的矢量对降水进行预测。Rinehart 等^[8]将雷达回波区域进行划分,得到不同的回波运动矢量,并在此基础上发明了使用交叉相关法跟踪回波运动的 TREC

(tracking radar echo by correlation)法。该方法可以在一定程度上反演风暴内部的运动,并在之后提出使用 TREC 方法的时间间隔应该小于 5 min,否则产生无序矢量。Li 等^[9]在此基础上提出了 COTREC 方法,该方法很好地减少了杂波的影响,减少了无序矢量的产生。

目前为止,单体质心法和交叉相关法都是气象业务部门重要的预报方式,但是由于交叉相关法应用面更广且精度较高,所以应用更为频繁。

1 方法研究

1.1 数据预处理

本次神经网络模型的输入,直接选择了反射率数据。数据的质量非常重要,特别是具有明显天气现象的雷达回波数据对实验有重要意义,雷达回波数据的总体反射率强度较高,而且范围也较大。选用 12 d 的具有明显降水特征的雷达回波数据,每 1 天的雷达回波数据有接近 240 个雷达体扫二进制文件,这些雷达体扫二进制文件以 6 min 为一个间隔,体现了时序之间的关联性。

每 1 次的雷达体扫有 11 个仰角层。每 1 个仰角层有 360 ~ 372 条径向(360 ~ 372 个方位角),而对于反射率数据来说,每 1 条径向有 460 个距离库,所以说每 1 个雷达体扫二进制文件能够提供 11 个(360 ~ 372)×460 的反射率数据。为尽量排除地物杂波的影响,又使反射率值不会偏小,选择雷达体扫第 5 层的反

射率数据。

数据集的详细信息如表 1 所示。

表 1 数据集详细信息

	训练集	测试集
年份	2019、2018、2013、2007	2019、2018
天数/d	21	2
数据总量	4000	300

1.1.1 雷达回波数据径向排序和拟合

由于每部雷达每层开始扫描的方位角是不固定的,为提取雷达回波数据在每条径向上的时序信息,必须将雷达回波数据的径向进行排序,即按照方位角进行排序。将径向数按照方位角从小到大排列后,固定雷达回波数据的空间位置信息。由于每部雷达的体扫径向数不尽相同,为了统一径向数,将所有雷达回波反射率数据的径向数统一为 360 个。为尽量反映出雷达回波反射率数据的特征,过滤掉无效数据,只选用前 200 个距离库的反射率数据。这样,每一个第 5 层的雷达回波数据就是一个形状为 360×200 大小的矩阵。统一了形状的输入对于训练是十分便利的。

1.1.2 雷达回波数据滤波

雷达回波数据上一般都会有地物杂波或者其他不相关的杂波。为滤除杂波,尽量减少其对于训练的影响,必须对其进行滤波处理。采用的滤波方法有 2 种^[10]。第一种是针对孤立点的处理,需要遍历每一个数据库,在每一个非零的有效数据上建立一个 $N \times N$ 的矩形窗口。通过判断窗口内所有数据中的有效数据总数 M ,得到该窗口有效数据的比例 S ,如式 1 所示。设置判断为孤立点的阈值为 M_0 (缺省值为 0.7)。若 S 小于 M_0 ,则判断该点为孤立数据点,将该点数据置为无效数据。

$$S=M/N^2 \tag{1}$$

第二种方法是补足缺测点针对斑秃问题,与去除孤立点的处理类似,遍历整个数据,在空白数据上建立 $M \times M$ 的矩形窗口(文中窗口尺寸 $M=5$)。统计窗口内 M^2 个数据中有效数据的个数,设置阈值 M_1 (缺省值为 12)。若有效数据个数大于此阈值,则有目标格点的值替换为

$$V=\frac{1}{M^2}\sum_{i=1}^{M^2}V_i \tag{2}$$

其中 V_i 为窗口内第 i 个数据的值。

1.1.3 雷达回波数据标准化

对于神经网络的输入来说,对数据进行归一化或者标准化是十分必要的,标准化的公式为

$$x=\frac{x-\text{mean}(x)}{\sigma} \tag{3}$$

其中, \bar{x} 为 x 的均值, σ 为 x 的方差。

标准化将数据变换为均值为 0,标准差为 1 的分布。这种分布并非一定是正态的。

最大-最小值归一化的公式为

$$x'=\frac{x-\min(x)}{\max(x)-\min(x)} \tag{4}$$

其中, $\min(x)$ 为 x 的最小值, $\max(x)$ 为 x 的最大值, x 为当前目标数据。

没有采用最大-最小值归一化而采用标准化的原因是,雷达回波数据里有较多的无效数值,将这些无效数值赋值为某一个较小的反射率强度值。如果采用最大-最小值归一化,那么数据的缩放比例仅仅和极值有关。所以这使每一列的反射率无效数据的归一化值不尽相同,而且对损失值计算和反归一化产生较大的影响,所以对于雷达回波数据用最大-最小归一化值是不现实的。

所以采用了标准化,因为标准化的结果跟每一列的每个数据都有关系,所以标准化后的数据不会改变其分布,从而使损失值计算和反归一化都可以正常进行。同时,标准化的好处可以使不同量纲的特征处于同一数量级,减少方差大的特征的影响,让数据尽量满足 sigmoid 和 tanh 的梯度下降区间,从而使梯度不会趋于 0,使模型更加准确;可以加快学习算法的收敛速度^[11]。

1.2 LSTM 神经元

传统的 RNN 在处理 long term memory 的时候存在缺陷,容易出现梯度爆炸或者梯度消失。因为在这种结构中,其只考虑最近的状态,而不考虑以前的状态。传统的 RNN 总是用“覆写”的方式计算状态:

$$S_t=f(S_{t-1},x_t) \tag{5}$$

其中 $f(\cdot)$ 表示仿射变换后在套上一个 tanh 或者 sigmoid 激活函数, x_t 表示输入序列在时刻 t 的值。根据求导的链式法则,这种形式直接导致梯度被表示为连成积的形式,以至于造成梯度消失,粗略得说,很多个小于 1 的项连乘后梯度就很快逼近零^[12]。

原始的 RNN 的隐藏层只有一个状态,即 h ,它对于短期的输入非常敏感。而对于 LSTM 来说,多了一个状态,即 c ,让它来保存长期的状态,称为单元状态(cell state)。LSTM 结构是 Jürgen Schmidhuber 在 1997 年发明的一种记忆形式,或者是一种神经网络的环境。可以说,LSTM 是 RNN(recurrent neural network)的升级版。LSTM 被证明在处理长序列问题上有出色的能力^[14-15]。LSTM 的公式如下:

$$\begin{aligned} i_t &= \text{sigmoid}(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i) \\ f_t &= \text{sigmoid}(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f) \\ c_t &= c_{t-1} \circ f_t + i_t \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \\ o_t &= \text{sigmoid}(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o) \\ h_t &= o_t \tanh(c_t) \end{aligned} \tag{6}$$

其中 \circ 代表哈达玛乘积, i_t 代表当前时刻的输入, W_* 代表权重矩阵, b_* 代表偏置, c_{t-1} 代表上一个时刻的状态。 f_t 代表遗忘门的值, 意味着哪些细胞状态应该被遗忘, c_t 代表着当前时刻的状态, o_t 代表着当前时刻的输出, h_t 代表当前输出和当前状态做哈达玛乘积后的最终输出。

经典的 LSTM 结构将数据展开成一维进行预测, 可以较好地解决时间相关性, 但是 FC-LSTM 只能提取时间序列信息, 无法提取空间信息。对于空间数据特别是雷达回波数据来说, 蕴含了大量的冗余信息, FC-LSTM 无法处理。

为了解决这个问题, 一种在 input-to-state 和 state-to-state 之间的卷积结构应运而生。ConvLSTM 使用卷积替换全连接, 从而可以提取序列的空间信息。即 FC-LSTM 和 ConvLSTM 的主要差别在于 ConvLSTM 将矩阵乘法操作替换成了卷积操作。即 input-to-state 和 state-to-state 的转换中使用的是卷积而不是全连接。其公式如下:

$$\begin{aligned} i_t &= \text{sigmoid}(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} \circ c_{t-1} + b_i) \\ f_t &= \text{sigmoid}(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} \circ c_{t-1} + b_f) \\ c_t &= c_{t-1} \circ f_t + i_t \tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c) \\ o_t &= \text{sigmoid}(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} \circ c_t + b_o) \\ h_t &= o_t \tanh(c_t) \end{aligned} \tag{7}$$

其中 $*$ 代表卷积, \circ 代表哈达玛乘积。

1.3 模型构造和参数

对于 LSTM(或者 ConvLSTM)来说神经网络模型一般最多三层, 采用了两层的隐藏神经元。该模型的总体结构如图 1 所示。

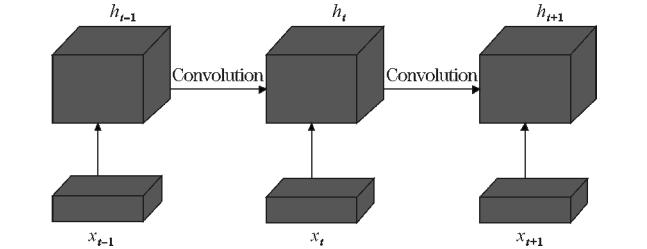


图 1 模型总体结构图

其中, h_{t-1} 代表上一个时刻的输出, x_{t-1} 代表上一个时刻的输入, h_t 代表当前时刻的输出, x_t 代表当前时刻的输入, h_{t+1} 代表下一个时刻的输出, x_{t+1} 代表下一个时刻的输入。

该模型的超参数如表 2 所示。

表 2 模型超参数

超参数	数值大小
输入大小	[360,200]
学习率	0.001
隐藏层 1 神经元	64
隐藏层 2 神经元	32
每一层是否采用 normalization	是
每一层是否采用 dropout	是
全连接层 1 神经元	64
全连接层 2 神经元	32
回归层神经元	1

2 实验结果

用 5 张反射率图预测未来 3 张反射率图, 对于评价其预测精确度使用雷达回波外推的 POD(命中率)和 CSI(临界成功指数), 其中 POD 和 CSI 数值越高代表预测精确度越高。其中第一批测试集真实雷达回波反射率图如图 2 所示。

预测雷达回波反射率图如图 3 所示。

该测试集 POD 和 CSI 如表 3 所示。

第二批测试集真实雷达回波反射率图如图 4 所示。

预测的雷达回波反射率图如图 5 所示。

该测试集 POD 和 CSI 如表 4 所示。

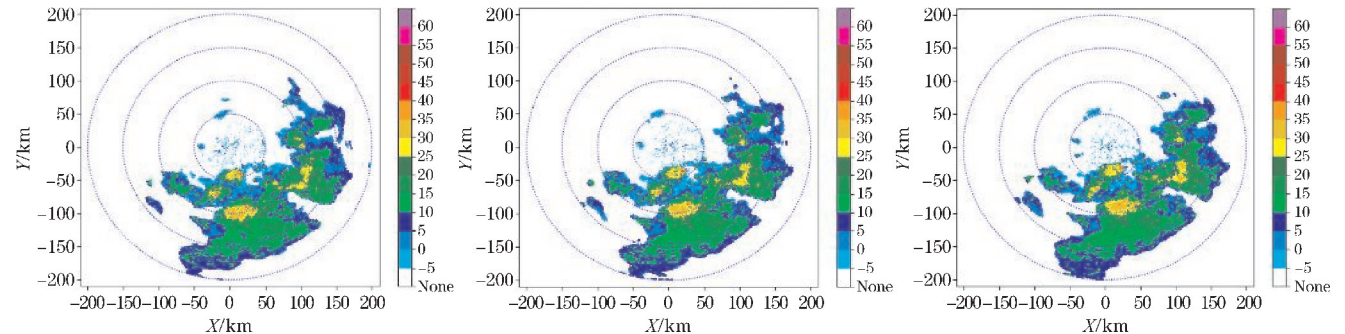


图 2 第一批测试集真实雷达回波反射率图

表 3 第一批测试集 POD 及 CSI		
回波阈值	POD	CSI
10	0.70	0.61
20	0.54	0.35
30	0.32	0.22

表 4 第二批测试集 POD 及 CSI		
回波阈值	POD	CSI
10	0.96	0.89
20	0.97	0.80
30	0.57	0.45

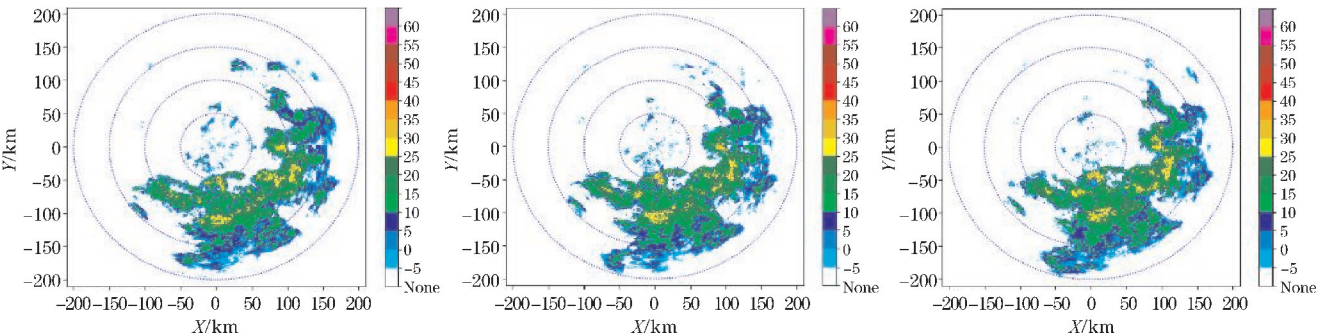


图 3 第一批测试集预测雷达回波反射率图

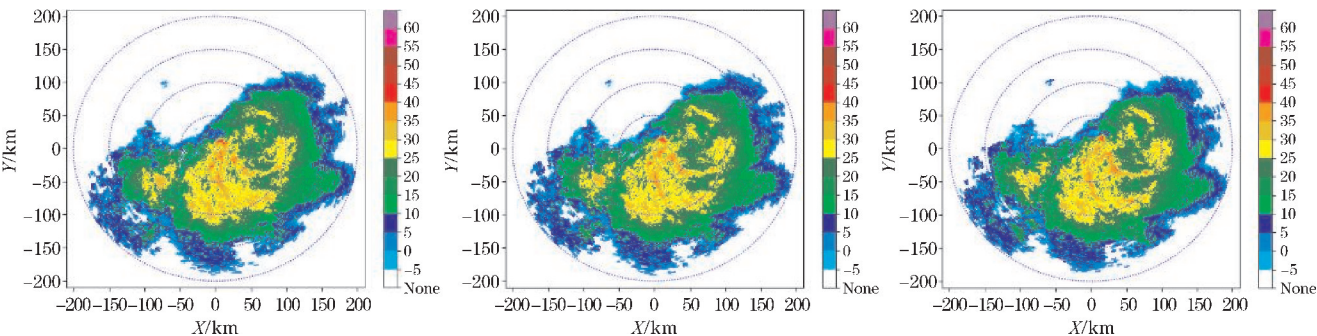


图 4 第二批测试集真实雷达回波反射率图

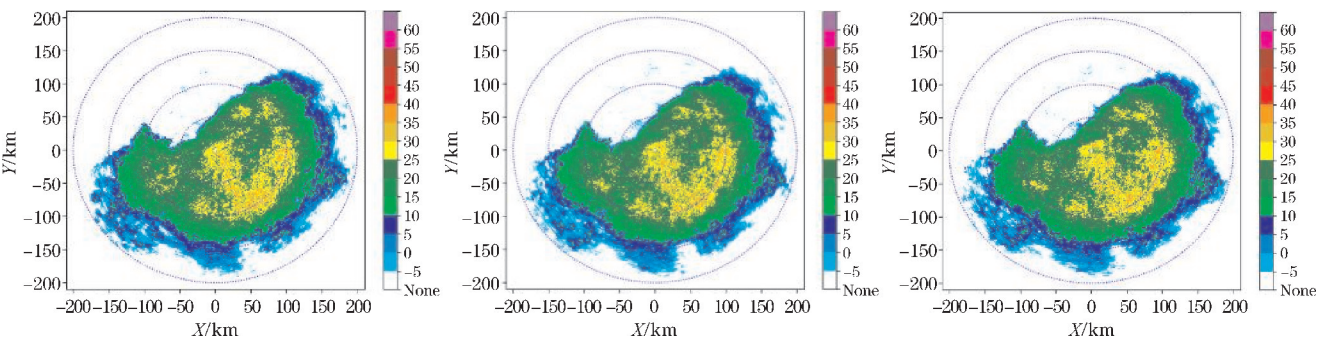


图 5 第二批测试集预测雷达回波反射率图

3 结束语

模型可以较好地预测强度为 30 以下的反射率值, 因为其命中率和临界成功指数都比较高, 但是如果提高回波阈值, 预测 30 以上强度的反射率值, 其命中率和临界成功指数就会急剧下降, 第一是因为训练的数据集小, 没有学习到过多的特征。第二是因为模型比

较简易, 其学习特征变化和预测未来数据的能力比较弱, 且对于强度大于 30 的反射率数据不易预测。为解决上述问题, 提高预测的精确度, 可以采用以下方法: 一是利用更多的数据集训练, 使模型能充分学习数据的特征, 并且提高数据的筛选性, 尽量减少无效数据的影响。二是提高模型的复杂度, 比如说利用含 attention 机制的 encoder-decode 模型, 其可以训练和学习较复杂的多维时序数据。可以预测强度大于 30 的反射率数据及 0.5 ~ 1 h 的雷暴和强对流天气变化。

参考文献:

- [1] 俞小鼎,周小刚,王秀明. 雷暴与强对流临近天气预报技术进展[J]. 气象学报,2012,70(3):311-337.
- [2] 张蕾,魏鸣,李南,等. 改进的光流法在回波外推预报中的应用[J]. 科学技术与工程,2014,14(32):133-148.
- [3] Dixon M, Wiener G. TITAN: Thunderstorm Identification, Tracking, Analysis, and Nowcasting-A Radar-based Methodology[J]. Journal of Atmospheric and Oceanic Technology, 1993, 10(6):785.
- [4] Johnson J T, Mackeen P L, Witt A, et al. The Storm Cell Identification and Tracking Algorithm: An Enhanced WSR-88D Algorithm [J]. Weather & Forecasting, 1998, 13(2):263-276.
- [5] 乔春贵,郑世林,杨立志,等. 质心法雷达回波外推的原理及应用[J]. 河南气象,2006(3):29-30.
- [6] Hilst G R, Russo J A Jr. An objective extrapolation technique for semi-conservative fields with an application to radar patterns[R]. Tech Memo No 3, Travelers Weather Research Center, Harford, CT, 1960:34.
- [7] Austin, G L Bellon A. The use of digital weather radar records for short-term precipitation forecasting[J]. Quarterly Journal of the Royal Meteorological Society, 1974.
- [8] Rinehart R E, Garvey E T. Three-dimensional storm motion detection by conventional weather radar[J]. Nature, 1978, 273:287-289.
- [9] LI L. Nowcasting of Motion and Growth of Precipitation with Radar Over Complex Orography [J]. j. appl. meteor, 1995:34.
- [10] 李若楠,张鸿,白志娜,等. 气象雷达基数据质量控制及其在回波识别方面的简单应用[J]. 现代农业科技,2015(5):245-247.
- [11] Shanker M, Hu M Y, Hung M S. Effect of data standardization on neural network training[J]. Omega, 1996, 24(4):385-397.
- [12] Jozefowicz, Rafal, Zaremba, et al. An Empirical Exploration of Recurrent Network Architectures[C]. International Conference on International Conference on Machine Learning. JMLR. org, 2015.
- [13] Hochreiter S, Schmidhuber J. Long Short-Term Memory[J]. Neural computation, 1997, 9(8):1735-1780.
- [14] Pascanu R, Mikolov T, Bengio Y. On the difficulty of training Recurrent Neural Networks[J]. 2012.
- [15] Gers F A, Schmidhuber, Jürgen, et al. Learning to Forget: Continual Prediction with LSTM[J]. Neural Computation, 2000, 12(10):2451-2471.

The Radar Echo Extrapolation based on ConvLSTM

LAI Can¹, WANG Haijiang¹, LI Jing², XU Zili², LIU Tao¹

(1. College of Electronic Engineering, Chengdu University of Information Technology, Chengdu 610225, China; 2. The Second Research Institute of CAAC, Chengdu 610041, China)

Abstract: In order to make predictions for radar echo productions, this paper has researched a radar echo extrapolation neural network model based on ConvLSTM, and the construction adopted by this model is long-short term memory network. I made changes for the structure so multiple dimensions data can be used as input. Single polarization SA radar reflectivity data is regarded as data sets. To achieve ideal prediction effect, the following approaches have been approved effectively: (1) There should be enough or relatively enough data sets, then we preprocess the data sets, including data selection, sort, fitting and filter. Afterwards do standardization, zoom data according to scale, and the reflectivity value is centered on 0 and distributed in a small interval. (2) The convolutional layer is used to extract characteristics of reflectivity data. Placing full connection and regression layers after hidden layers, we can obtain more optimal and accurate prediction results. (3) We take advantage of some activation functions and regularization methods, these can reduce the over-fitting of the model, improve the training accuracy of model. The experimental results have shown that pretty good prediction effect has obtained within 18 minutes, and prediction ability is strong while reflectivity value is small than 30.

Keywords: ConvLSTM; radar echo extrapolation; data preprocessing; optimization; regularization