

3DuA-Net:融合 3D 卷积和注意力的 雷达回波外推预报*

包广斌 罗 瞳 彭 璐 赵怀森

(兰州理工大学计算机与通信学院 兰州 730050)

摘 要: 针对短临降雨预测模型对历史雷达数据的建模结果存在视觉性能模糊和低估高回波值的问题,提出一种融合 3D 卷积和双端注意力机制的短临降雨雷达回波外推模型 3DuA-Net。以 ST-LSTM 时空长短期记忆网络为循环单元,将普通卷积替换为 3D 卷积,增强模型从全局视角强化对短期运动特征信息的捕获能力。并提出 DuAtt 高效双端注意力机制,提高模型对长期雷达图像序列局部和全局重要特征信息的保存及结合能力。采用深圳气象局公开的多普勒雷达数据集进行实验,结果表明:在 10、20、40 dBz 阈值下,该模型相比 Conv-LSTM 基线模型的 CSI 指标平均提升 7.74%,HSS 指标平均提升 5.54%,MAE 指标下降 3.8%,SSIM 指标提升 8.86%。

关键词: 短临降雨;三维卷积;注意力机制;深度学习;雷达回波外推

中图分类号: TP391.41;TN959.4 **文献标识码:** A **国家标准学科分类代码:** 520.2040

3DuA-Net: Fusion of 3D convolution and attention for radar echo extrapolation forecasting

Bao Guangbin Luo Tong Peng Lu Zhao Huaisen

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: In response to the issues of visual performance blurring and underestimation of high echo values in the modeling results of traditional short-term rainfall prediction models for historical radar data, we propose a short-term rainfall radar echo extrapolation model that integrates 3D convolution and dual-end attention mechanism 3DuA-Net. Using ST-LSTM space-time long short-term memory networks as the recurrent units, replacing ordinary convolution with 3D convolution enhances the model's capability to capture short-term motion features from a global perspective. Additionally, an efficient dual self-attention module DuAtt is proposed to improve the model's ability to preserve and integrate important local and global features in long-term radar image sequences. Experimentation conducted using publicly available Doppler radar datasets from the Shenzhen Meteorological Bureau shows that, at 10, 20, 40 dBz thresholds, the model exhibits an average improvement of 7.74% in the CSI metric compared to the Conv-LSTM model, an average improvement of 5.54% in the HSS metric, a decrease of 3.8% in the MAE metric, and an improvement of 8.86% in the SSIM metric.

Keywords: precipitation nowcasting; 3D convolution; attention mechanism; deep learning; radar echo extrapolation

0 引 言

联合国政府间气候变化专门委员会的评估报告强调,全球气候变化正导致中高纬度和热带地区的降水量显著增加,进而引发极端降雨事件的频率上升。以近年为例,2021 年郑州“7·20”暴雨^[1]和 2023 年 7 月的华北特大暴雨都导致严重内涝和山洪灾害。因此,实现实时效的降雨预测尤

为重要。短临降雨预报专注于未来 0~2 小时内的精细化预测^[2],力求实现“定时、定点、定量”的精准度。由于降雨受季节、地理高度和大气环流等多因素影响,变化趋势复杂,因此常利用高时空分辨率和高时效性天气雷达回波图像数据来训练预测模型。

雷达回波外推是短临降雨预报的核心方法^[3]。其原理是利用天气雷达探测到的回波数据,精确分析回波的强度

分布及回波体的移动速度和方向,随后对回波体进行外推,以预测未来雷达回波状态。

基于深度学习的雷达回波外推方法能够自适应地学习雷达回波序列中降雨的变化规律,直接建立历史数据到未来预测数据的映射,相较传统方法避免了不必要的假设^[4]。目前,它面临的主要挑战是如何设计一个合适的深度网络架构以有效捕捉时空特征。为解决该问题,Shi 等^[5]将卷积神经网络(convolutional neural network,CNN)和长短期记忆网络(long short-term memory,LSTM)结合形成 Conv-LSTM 模型,建立端到端的短临降雨预报模型。但该模型准确性较低,由此衍生出多种改进方法。Wang 等^[6]提出时空记忆单元级联方法在层次不变的情况下,增加更多的非线性操作。Lin 等^[7]提出将注意力机制引入到 Conv-LSTM 单元中,增强了时空表征能力。Trebing 等^[8]提出增加注意力模块的小型 U-net 模型,提升预测效果的同时减小了网络规模。Zhang 等^[9]提出结合回溯注意力机制的局部注意力预测模型,解决了因 Conv-LSTM 网络不能同时捕获局部和全局的空间特征的问题。此外,Zheng 等^[10]开发了一个时间交互块和一个空间提取块并将模块合并得到 TISE-LSTM 模型,该模型有效解决了遗忘门信息丢失的问题。尹传豪等^[11]提出基于非对称双路特征提取结构 GLNet 模型,该模型融合不同尺度下雷达降雨图片的局部特征和全局特征,在降雨区域具有更高的预测准确性。

上述模型在预测方面都表现尚佳,但存在两大显著问题:一是视觉性能略显模糊,二是对于中高回波值的预估偏低。中高回波值区域常用与暴雨紧密相连,可能会诱发自然灾害。所以需要针对这些问题进行预测模型的改进。实验证明,这些问题的根本原因有:1)模型未能捕捉到长期序列数据流中相邻时间步之间的短期运动信息;2)模型无法有效保留长期序列中局部和全局空间的重要特征。

因此,本文提出了一种新的 3DuA-Net 模型,旨在通过学习历史多普勒雷达回波图序列中云层的空间分布和强度演变规律,实现对雷达回波图像的有效外推,以达到更高的短临降雨预测精度。

1 DuA-Net 模型设计

1.1 短临降雨预测问题定义

在机器学习的视角,短临降雨预测本质上是一个时空序列预测问题,其核心在于精准预测随时间演变且存在显著空间关联性的数据序列。在实际应用中,每 6 min 获取一次雷达图,若要预测某区域 1 h 内的降雨情况,通过先前训练得出的算法模型外推未来 6~15 帧雷达图,然后结合雷达图和 Z-R 关系^[12]完成天气预测任务。Z-R 公式如式(1)所示。

$$dBZ = pixel_value \times 95/255 - 10 \quad (1)$$

式中: $pixel_value$ 表示像素值, dBZ 表示反射率值。

时空序列预测问题的形式化定义为:假设某一动态系

统每隔固定的时间产生 P 个测量值,且每个测量值均为 $M \times N$ 大小的具有空间相关性的格点数据,那么每个测量值均可表示为相应大小的 P 通道图片,记为 $X_t \in R^{P \times M \times N}$,其中 R 表示观测特征的域。令 t 为最后一个观测时刻, $X_{t-J+1:t}$ 为过去 J 个历史时刻的观测数据, $\hat{X}_{t+1:t+K}$ 为该动态系统未来 K 个时刻的状态,则定义一般形式的时空序列预测问题如式(2)所示。

$$\hat{X}_{t+1}, \dots, \hat{X}_{t+K} = \operatorname{argmax}_p(X_{t+1}, \dots, X_{t+K} \mid X_{t-J+1}, \dots, X_t) \quad (2)$$

此外,如果动态系统在除确定性趋势外,在时间和空间的任意维度上还具有典型的季节性特征,那么需要对上述表达式进行调整。

1.2 3DuA-Net 外推模型框架

3DuA-Net 模型框架如图 1 所示,由多个三维双端注意力预测单元(3DuA-Unit)堆叠而成,生成时间戳 2 到 T 的预测。在任意时间戳 t ,模型用当前雷达图 X_t 和历史雷达序列 $X_{0:t}$ 来预测下一个时间戳 $t+1$ 的雷达图 X_{t+1} 。这一过程中,时间记忆信息 C 沿水平方向传输,而空间记忆信息 M 则以锯齿状方式传输。在 3DuA-Unit 中采用高效双端注意力机制(efficient dual self-attention module,DuAtt)保留长期历史特征信息,提高输出的时空表达能力。通过上述过程,模型建模运动模式的同时保持了空间外观的连贯性,且使得所保留的空间外观能够有效地与时间运动特征相结合。值得注意的是,大多数基于 Conv-LSTM 模型的预测网络采用了相似的体系框架,不同主要在于所使用的单元,而非框架本身。

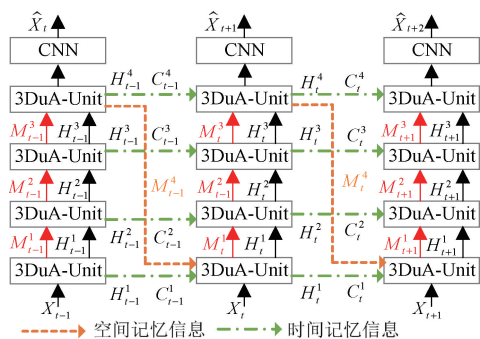


图 1 3DuA-Net 外推预测模型架构图

Fig. 1 Framework of 3DuA-Net model

1.3 外推单元网络结构

1) 外推单元 3DuA-Unit

3DuA-Unit 外推单元如图 2 所示。该单元有 4 个输入:当前输入 X_t 、空间记忆 M_{t-1}^i 、历史时间记忆 $C_{t-\tau,t-1}$ 和隐藏状态 H_{t-1}^i ,输出包括当前的隐藏状态 H_t^i 、空间记忆 M_t^i 、时间记忆 C_t^i 。与常规循环单元不同,3DuA-Unit 的输入门、遗忘门和输出门不依赖于同一层先前时间步的隐藏状态和细胞状态,而是依赖于当前输入和当前时间步上一

层更新的细胞状态。为了捕捉连续时空领域中的短期和局部运动特征,3DuA-Unit 采用 3D 卷积作为运动感知器。提取的局部和全局特征信息随后与多个先前的时间记忆信息一同输入到开发的 DuAtt 机制中,旨在恢复可能被遗忘的重要信息。最终,经过处理的信息被传递到下一时刻,作为下一时间步的输入。这种设计让 3DuA-Unit 能够有效处理时空序列,同时在分析中保留关键信息。

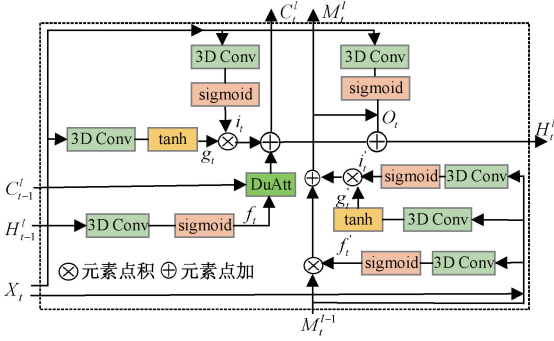


图2 3DuA-Unit 外推预测单元网络图

Fig. 2 Network of 3DuA-Unit prediction

3DuA-Unit 单元在计算形式上表示,如式(3)~(12)所示。

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1}^l + b_i) \quad (3)$$

$$g_t = \tanh(W_{xg} * X_t + W_{hg} * H_{t-1}^l + b_g) \quad (4)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1}^l + b_f) \quad (5)$$

$$C_t^l = i_t \circ g_t + \text{DuAtt}(f_t, C_{t-\tau, t-1}^l) \quad (6)$$

$$M_t^l = f_t' \circ M_{t-1}^{l-1} + i_t' \circ g_t' \quad (7)$$

$$H_t = O_t \circ \tanh(W_{1 \times 1} * [X_t^l, W_t^l]) \quad (8)$$

$$i_t' = \sigma(W_{xi}' * X_t + W_{mi}' * M_{t-1}^{l-1} + b_i') \quad (9)$$

$$f_t' = \sigma(W_{xf}' * X_t + W_{mf}' * M_{t-1}^{l-1} + b_f') \quad (10)$$

$$g_t' = \tanh(W_{xg}' * X_t + W_{mg}' * M_{t-1}^{l-1} + b_g') \quad (11)$$

$$O_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1}^l + W_{co} * C_{t-1}^l + W_{mo} * M_{t-1}^{l-1} + b_o) \quad (12)$$

式中: σ 是 sigmoid 函数, \circ 表示矩阵乘积即范德蒙运算, τ 是先前时间记忆的数量, $*$ 表示 3D 卷积操作。式(5)中的 DuAtt 表示本文改进的双端注意力机制。 C_t^l 是当前时间步的记忆状态。 i_t 、 f_t 和 g_t 分别表示输入门、遗忘门和输入模块门,用于控制纵向信息传递。 O_t 和 H_t^l 分别表示输出门和隐藏状态。

2) 3D 卷积编码

如何摆脱手动设计的时空信息特征是视频分析中的一项关键问题。黄金贵等^[13-14]研究表明,3D 卷积神经网络在时空特征的探索方面较 2D 卷积更为有效。如图 3(a)所示,二维卷积层只能从前一层的局部邻域中提取空间维度特征,生成一个特征图。如图 3(b)所示,二维卷积应用于视频体积时,多个相邻帧被视为多个通道,这导致多帧操作后被压缩成一个特征图,丧失了时序信息。如图 3(c)所

示,3D 卷积是通过将一个三维卷积核与多个相邻帧叠加形成的立方体卷积。卷积层中的特征映射与前一层中的多个相邻帧连接来获取运动信息,与 2D 卷积相比多生成了一个视频体积,且保留了时序信息。

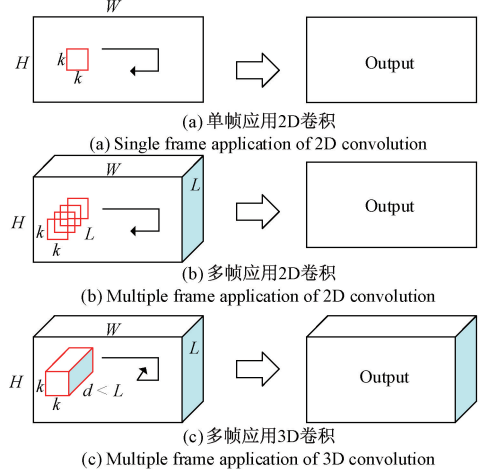


图3 2D和3D卷积操作对比图

Fig. 3 Comparison of 2D and 3D convolution operations

每个时间间隔的雷达回波图可以类比为视频帧,鉴于3D卷积在视频分析中的优势,本文将该技术引入短临降雨预测领域,旨在自动提取雷达回波数据中的时空特征,特别是云层运动等关键信息。

如图4所示,采用滑动窗口技术处理原始数据。这种方法通过将每个滑动窗口内的多帧数据视为一个时间步的输入,增加一个深度通道,将输入数据扩展为宽度、高度和时间(即多帧数)3个维度,使得模型能够捕获到雷达回波数据中的动态时空特征。

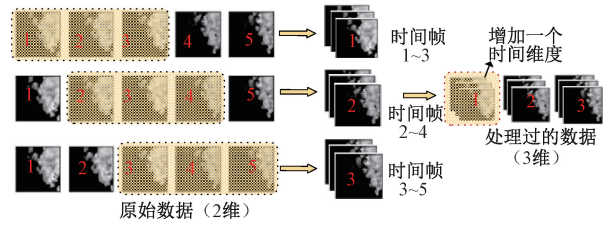


图4 使用滑动窗口处理的数据

Fig. 4 Data processed using sliding windows

为了与三维空间的输入数据兼容,对预测模型内部结构进行调整,即将预测单元中的普通卷积替换为3D卷积,以提取相邻数据帧之间的运动信息。

3) DuAtt 机制

Woo等^[15]提出的轻量级通用注意力模块在不增加网络参数的同时显著提升了网络的图像表征能力。Fu等^[16]在图像处理领域成功利用注意力机制捕捉并建模了长期依赖关系。借鉴上述经验,本文提出一个嵌入预测单元的双端注意力机制,如图5所示。在图2中标识为关键模块。该机制采用动态加权策略将注意力图与特征图相乘,自适应

应地学习并强调先前记忆中对于预测结果重要的特征信息。

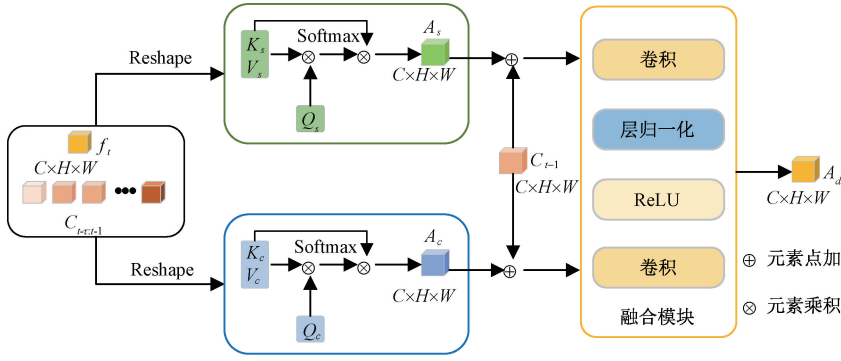


图 5 双端注意力机制图

Fig. 5 Dual attention mechanism

多普勒雷达广阔的成像范围导致回波图像成像内容繁杂,如何更好地识别显著性区域并忽略背景区域对于实现雷达图像特征表示至关重要。空间注意力模块的目的就是在 CNNs 提取的高层特征基础上,进一步捕获不同区域特征的上下文联系,并为各个区域生成重要性权重,识别出与任务目标高度相关的显著性区域。空间注意力模块的网络结构为图 5 中的绿色部分所示。该模块输入包括一系列特征图 $C_{t-\tau:t-1} \in R^{N \times C \times \tau \times H \times W}$ 和通过遗忘门的信息 $f_t \in R^{N \times C \times H \times W}$ 。

其中, N 、 C 、 H 和 W 分别表示批量大小、通道数、高度和宽度。将一系列特征图 $C_{t-\tau:t-1}$ 通过重塑处理后,得到键向量序列 $K_s \in R^{N \times (\tau \times H \times W) \times C}$ 和值向量序列 $V_s \in R^{N \times (\tau \times H \times W) \times C}$ 。 f_t 遗忘信息通过重塑处理后,得到查询向量序列 $Q_s \in R^{N \times (H \times W) \times C}$ 。其中, τ 表示系列的长度。空间注意力模块的输出表达式为:

$$A_s = \text{DuAtt}_s(f_t, C_{t-\tau:t-1}) = \text{norm}(C_{t-1} + \text{softmax}(Q_s \cdot K_s^T) \cdot V_s) \quad (13)$$

式中: A_s 表示给定特征图 f_t 和长期系列映射 $C_{t-\tau:t-1}$ 的空间相关程度,其中, $Q_s = f_t, K_s = V_s = C_{t-\tau:t-1}, \text{softmax}(Q_s \cdot K_s^T)$ 表示 Q_s 和 K_s 两者的相似性。

不同特征对雷达回波图像语义内容的表达起着不同的作用。通道注意力模块在 CNN 提取特征图的基础上,根据各通道特征的重要性程度,生成相应的权重,从而优化网络对图像语义内容的理解和表达。通道注意力模块的网络结构如图 5 中的蓝色部分所示。与空间注意力模块不同, $f_t \in R^{N \times C \times H \times W}$ 被转换为查询向量序列 $Q_c \in R^{N \times (\tau \times H \times W) \times C}$,特征图 $C_{t-\tau:t-1}$ 被重塑为键向量序列 $K_c \in R^{N \times (\tau \times H \times W) \times C}$ 和值向量序列 $V_c \in R^{N \times (\tau \times H \times W) \times C}$ 。通道注意力模块的输出表达式为:

$$A_c = \text{DuAtt}_c(f_t, C_{t-\tau:t-1}) = \text{norm}(C_{t-1} + \text{softmax}(Q_c \cdot K_c^T) \cdot V_c) \quad (14)$$

式中: A_c 表示给定特征图 $Q_c = f_t$ 和长期系列映射 $C_{t-\tau:t-1}$ 的通道相关程度。 $Q_c = f_t, K_c = V_c = C_{t-\tau:t-1}, \text{softmax}(Q_c \cdot K_c^T)$

表示查询向量序列 f_t 对键向量序列 $C_{t-\tau:t-1}$ 的通道方面影响。

时空信息融合模块将通道注意力和空间注意力的输出结果进行整合,该整合结果不仅涉及局部空间特征,还包括长序列 τ 周期中相应位置的信道关系。因此,融合模块的输出比每个单一模块更有效。融合计算的输出表达式为:

$$A_d = \text{DuAtt}_d(f_t, C_{t-\tau:t-1}) = \text{SumFusion}(\text{DuAtt}_s(f_t, C_{t-\tau:t-1}), \text{DuAtt}_c(f_t, C_{t-\tau:t-1})) \quad (15)$$

具体来说,对 A_c 和 A_s 独立应用两个卷积层。第 1 个卷积层包括一个内核尺寸为 3 的卷积、一个层归一化和 ReLU 激活函数。第 2 个卷积层使用了内核尺寸为 1 的卷积操作,对两个结果逐元素求和。

2 实验设置

2.1 实验参数与数据

3DuA-Net 模型使用适应性矩估计 (adaptive moment estimation, Adam) 优化器进行训练,初始学习率为 0.001,采用 L1 + L2 损失函数进行优化。批处理大小和最大迭代次数分别设置为 4 和 60 000 次。并且采用提前停止和定时采样策略进一步优化模型。

实验数据采用深圳气象局公开的多普勒雷达数据集。该数据集包含 10 000 个样本作为训练集,4 000 个样本作为测试集。训练集的数据来自连续两年的采样数据,测试集则是在下一年内采样。从训练集中随机选择 2 000 个样本作为验证集。每个样本包含 15 张 CAPPI 雷达图像,每张图像的空间分辨率为 101×101 像素,图片的时间间隔为 6 min。

实验专注于云层的演变、密度和结构特征,故采用灰度图像来捕捉这些信息,此举不仅避免了 RGB 图像的冗余信息,还加强了模型的泛化能力。

数据集链接: <https://tianchi.aliyun.com/dataset/1085>。

2.2 评估方法

选择降雨临近预报任务中广泛认可的 Heidke 技能分数 (heidke skill score, HSS) 和临界成功指数 (critical success index, CSI)。以及时空序列预测任务常用的测量指标,结构相似性 (structural similarity index, SSIM) 和平均绝对误差 (mean absolute error, MAE) 进行评估。其中, HSS、CSI 和 SSIM 的取值范围为 $[0, 1]$, MAE 的取值范围为 $[0, +\infty]$ 。

HSS 和 CSI 计算方法如式(16)~(17)所示。

$$HSS = \frac{2(FN \times TN - FN \times FP)}{(TP + FN)(FP + FN) + (TP + FP)(FP + TN)}$$

(16)

$$CSI = \frac{TP}{TP + FP + FN}$$

(17)

式中: TP 表示真阳性预测, FP 表示假阳性预测, TN 真阴性预测, FN 表示假阴性预测。

MAE 和 SSIM 计算方法如式(18)~(19)所示。

$$MAE(X_{k+1:T}, X_{k+1:T}) = \frac{\sum_i^H \sum_j^H |X_{k+1:T}(i, j) - X_{k+1:T}(i, j)|}{H \times W}$$

(18)

$$SSIM(X_{k+1:T}, X_{k+1:T}) = \frac{(2u_x u_y + C_1)(2\sigma_{xy} + C_2)}{(u_x^2 + u_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

(19)

式中: H 和 W 分别表示雷达回波图的高和宽, (i, j) 表示

雷达图像的位置, x 和 y 表示比较的两幅图片, u_x 和 u_y 分别为 x 和 y 的平均值, σ_x 和 σ_y 表示两者的方差, σ_{xy} 是 x 和 y 之间的协方差, C_1 和 C_2 是常数。

3 实验结果和模型对比

3.1 对比实验

本文的实验目标是用过去半小时的历史观测图像来预测未来一小时的观测结果, 因此每个示例中, 前 5 个回波图作为输入, 后 10 个回波图为预期输出。为了对比本文模型的性能效果, 将 3DuA-Net 模型与 Conv-LSTM、PredRNN++、SA-ConvLSTM、SmaAt-Unet、RPA-Net、TISE_LSTM 六种经典时空预测模型, 在 3 个不同阈值 (10、20、40 dBz) 情形下进行训练测试。结果如表 1 所示, 每一行记录了不同度量标准下每个模型的平均值, 最佳效果用加粗标注。比较发现, 除了 10 dBz 的 CSI 指标和 20 dBz 的 HSS 指标, 3DuA-Net 模型在 CSI 和 HSS 各阈值情形下均取得了最高的指标值。观察 MAE 和 SSIM 数值, 可看出 3DuA-Net 模型有最小的误差和最高的结构相似性, 即体现该模型在综合性能方面优于其他模型。PredRNN++ 和 SA-ConvLSTM 相比, SSIM 值相似, 但 3DuA-Net 模型的其他评估分数明显高于它们。这也表明了 DuA-Unit 的优点。虽然, TISE_LSTM 在最低阈值 (10 dBz) 下具有最佳的 CSI, 但在中高阈值 (20 dBz) (20 dBz、40 dBz) 下的性能较差, 这表明与传统的注意力机制相比, 双端注意力机制可以改善雷达回波中高区域的预测。

表 1 测试数据集在 HSS、CSI、SSIM 和 MAE 方面比较结果

Table 1 Comparison results of test data sets in terms of HSS、CSI、SSIM、MAE

方法	HSS↑				CSI↑				MAE↓	SSIM↑
	10 dBz	20 dBz	40 dBz	Agv.	10 dBz	20 dBz	40 dBz	Agv.		
Conv-LSTM	0.705 2	0.485 0	0.125 2	0.438 4	0.771 2	0.412 1	0.074 9	0.419 4	14.69	0.335 3
PredRNN++	0.706 9	0.496 1	0.130 8	0.444 6	0.773 4	0.420 2	0.079 3	0.424 3	14.54	0.334 1
SA-ConvLSTM	0.706 6	0.503 7	0.122 8	0.444 4	0.771 2	0.428 4	0.076 6	0.425 4	14.51	0.335 7
SmaAt-Unet	0.704 6	0.530 8	0.151 7	0.462 4	0.767 1	0.441 1	0.097 2	0.435 1	14.69	0.329 6
RPA-Net	0.705 0	0.509 1	0.089 3	0.434 5	0.780 4	0.431 9	0.052 1	0.421 5	14.52	0.334 8
TISE_LSTM	0.715 0	0.508 5	0.178 6	0.467 4	0.785 8	0.420 5	0.108 5	0.438 2	14.45	0.355 6
3DuA-Net	0.717 6	0.519 7	0.188 3	0.475 2	0.771 7	0.452 2	0.115 3	0.446 4	14.12	0.367 9

图 6 为模型在不同阈值下的预报表现。从 HSS 和 CSI 曲线可见, 随着预测时间的增加, 模型的预报准确性均呈现下降趋势, 这也凸显出长期预报的困难。如图 6(a) 和(b)所示, 当阈值为 10 dBz 时, 本文模型对未来 1 h 的回波预测效果超过其他模型。如图 6(c)和(d)所示, 当阈值为 20 dBz 时, 36 min 后该模型预测优于其他模型。如图 6(e)和(f)所示, 在 40 dBz 阈值下, 3DuA-Net 模型表现尤为突出。这一结果验证了新模型单元中的 3D 卷积和双端注意力机制, 对于提高高降雨区域预测的准确性是有

效的。

为了更好地比较和理解结果, 图 7、8 给出了不同模型的可视化预测结果。第 1 行图片表示真实回波图像, 最下面颜色表格代表不同的雷达回波强度, 回波强度越高表示该区域发生极端天气情况的几率越高。可以看出, Conv-LSTM 模型对中高回波强度的预测误差很大, PredRNN++ 从第 3 帧开始, 不能很好地预测高回波区域, SmaAt-Unet 和 TISE_LSTM 均在 6 帧之后无法很好地对中等回波强度进行预测, 只有 3DuA-Net 可以在第

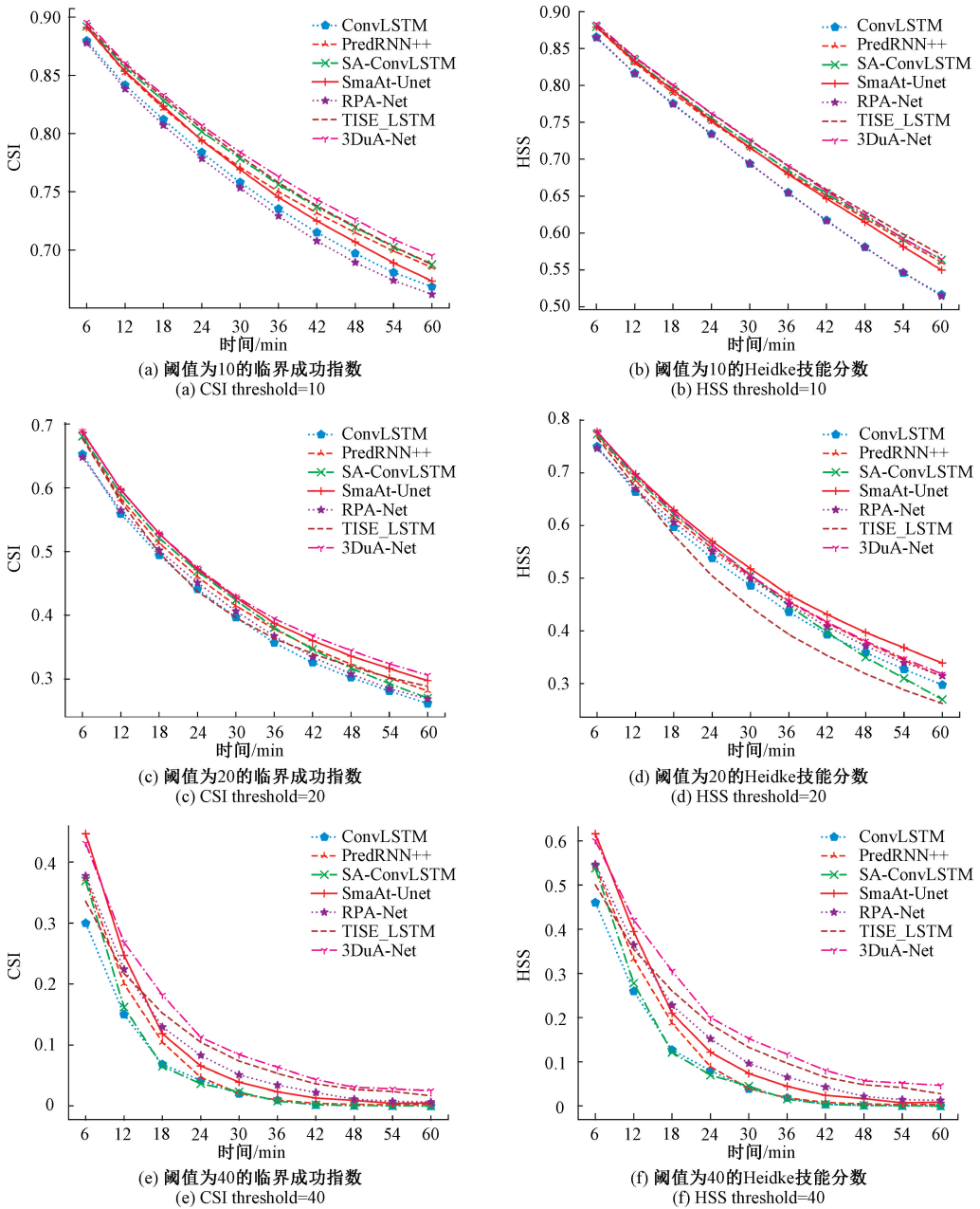


图 6 不同回波强度下的 CSI 和 HSS 曲线

Fig. 6 CSI and HSS curves under different echo intensities

10 个预测图像中保留高回声值区域(红色部分),且红色部分随着预报时间的推移逐渐消失,这是因为其同时应用了三维卷积和双端注意机制,能够更好地模拟短期和长期依赖关系。

3.2 消融实验

为验证 3DuA-Net 模型中 DuAtt 机制和 3D 卷积操作在预测模型中的影响和不同作用,进行了一系列消融实验。首先对消融模型进行相关说明,设计 3 个变种模型:3D-LSTM、CAAtt-Net 和 SAAtt-Net,分别代表在时空长短期记忆网络(spatiotemporal LSTM,ST-LSTM)基础上将普

通卷积替换为 3D 卷积的 3D-Net 模型、在 3D-Net 模型单元基础上仅添加通道注意力机制的 CAAtt-Net 模型、以及在 3D-Net 模型单元基础上仅添加空间注意力机制的 SAAtt-Net 模型。以上模型变种的实验结果在表 2 中展示。

从表 2 可以看出,DuAtt 机制和三维卷积操作对预测结果均有不同程度的提升。前两行数值对比得出,3D 卷积操作优于普通卷积操作,且能够获取更多的重要时空特征,但在高阈值下效果并不明显。2、3、4 行数值对比得出,CAAtt-Net 模型和 SAAtt-Net 模型相较于 3D-Net 模型的预测结果指标仅有不太明显的提升,其原因可能是通道注意

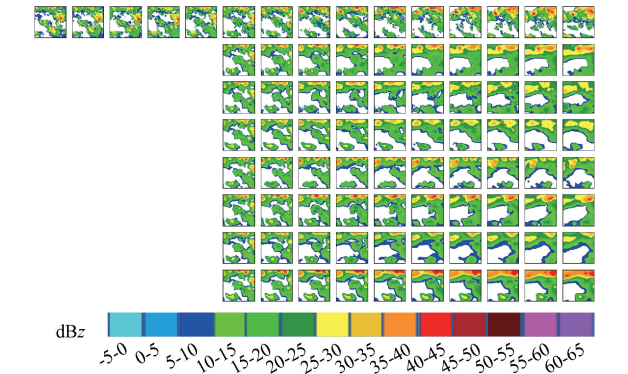


图 7 第 758 序列测试样本图

Fig. 7 Image of the 758th sequence test sample

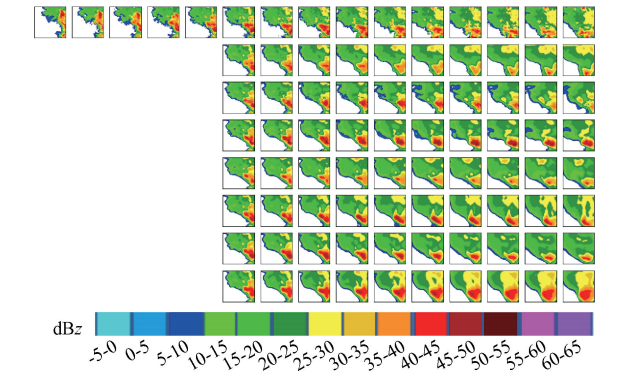


图 8 第 2 908 序列测试样本图

Fig. 8 Image of the 2 908th sequence test sample

表 2 消融实验(学习率为 10^{-4})

Table 2 Ablation experiment (learning rate 10^{-4})

方法	HSS \uparrow				CSI \uparrow				MAE \downarrow	SSIM \uparrow
	10 dBz	20 dBz	40 dBz	Agv.	10 dBz	20 dBz	40 dBz	Agv.		
PredRNN++	0.706 9	0.496 1	0.130 8	0.444 6	0.773 4	0.420 4	0.079 3	0.424 3	14.54	0.334 1
3D-Net	0.710 1	0.504 5	0.135 3	0.450 0	0.776 8	0.422 1	0.089 0	0.429 3	14.48	0.334 9
CAtt-Net	0.710 3	0.508 0	0.144 1	0.454 1	0.777 5	0.428 1	0.085 7	0.430 3	14.23	0.324 1
SAtt-Net	0.716 5	0.515 3	0.152 2	0.461 3	0.778 7	0.429 7	0.092 9	0.433 8	14.22	0.325 1
3DuA-Net	0.717 6	0.530 8	0.188 3	0.478 9	0.780 4	0.452 2	0.115 3	0.449 3	14.10	0.357 9

力机制和空间注意力机制未能得到充分的训练以及没有保留必要的时间和空间特征细节。而在集有 DuAtt 机制和三维卷积操作的 3DuA-Net 模型预测指标优于其他消融模型,尤其在较高阈值中预测更加明显。

4 结 论

本文通过构建基于雷达回波外推的 3DuA-Net 短临降雨预测模型,解决了短临降雨预测结果视觉性能模糊和低估中高回波值的问题。在公开的多普勒雷达数据集进行的一系列实验中,3DuA-Net 模型在综合性能上超越了现有的主流深度学习预测模型,包括 Conv-LSTM、PredRNN++、SA-ConvLSTM、SmaAt-Unet、RAP-Net、TISE_LSTM 等。特别是在预测强降雨事件和降雨强度持续上升的地区,3DuA-Net 展现出了卓越的性能,预测结果更为精确和清晰。

未来,在进一步优化模型时考虑加入地面温度、气压、大气湿度等其他气象要素的输入。同时,在保证现有精度的基础上,提高模型对预测值的长期依赖性,进一步解决多帧外推结可视化结果的模糊性问题以达到更高短临预报降雨准确率。

参考文献

[1] 郭换换,王坤. 基于双偏振雷达和降水现象仪的郑州“7·20”极端强降水微物理特征分析[J]. 大气科学学报, 2023,46(5):713-724.

GUO H H, WANG K. Analysis of microphysical characteristics of the “7·20” extreme heavy rainfall in Zhengzhou based on dual polarization radar and precipitation phenomenon instrument [J]. Chinese Journal of Atmospheric Sciences, 2023, 46 (5): 713-724.

[2] ZHANG Y CH, LONG M SH, CHEN K Y, et al. Skilful nowcasting of extreme precipitation with NowcastNet[J]. Nature, 2023, 619(7970): 526-532.

[3] 吴卓升,张巍,林艳,等. 动态概率卷积神经网络在雷达回波外推中的应用[J]. 计算机应用研究, 2021,38(7): 2125-2129.

WU ZH SH, ZHANG W, LIN Y, et al. Application of dynamic probabilistic convolutional neural network in radar echo extrapolation[J]. Journal of Computer Application Research, 2021, 38(7): 2125-2129.

[4] 方巍,庞林,易伟楠. 基于深度时空融合网络的雷达回波外推模型[J]. 电子学报, 2023,51(9):2526-2538.

FANG W, PANG L, YI W N. Radar echo extrapolation model based on deep spatiotemporal fusion network [J]. Journal of Electronics, 2023, 51(9): 2526-2538.

[5] SHI X J, CHEN ZH R, WANG H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting [J]. 28th

- International Conference on Neural Information Processing Systems, 2015, 1:802-810.
- [6] WANG Y B, GAO ZH F, LONG M SH, et al. PredRNN++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning [C]. International Conference on Machine Learning, 2018: 5123-5132.
- [7] LIN ZH H, LI M M, ZHENG ZH B, et al. Self-attention convlstm for spatiotemporal prediction[C]. AAAI Conference on Artificial Intelligence, 2020, 34(7): 11531-11538.
- [8] TREBING K, STANCZYK T, MEHRKANOON S. SmaAt-UNet: Precipitation nowcasting using a small attention-UNet architecture[J]. Pattern Recognition Letters, 2021, 145: 178-186.
- [9] ZHANG ZH, LUO CH, FENG SH SH, et al. RAP-Net: Region attention predictive network for precipitation nowcasting [J]. Geoscientific Model Development Discussions, 2022, 15(13): 5407-5419.
- [10] ZHENG CH Y, TAO Y F, ZHANG J J, et al. TISE-LSTM: A LSTM model for precipitation nowcasting with temporal interactions and spatial extract blocks [J]. Neurocomputing, 2024, 590: 127700.
- [11] 尹传豪,秦华旺,戴跃伟,等. GLnet:融合全局和局部信息的短临降雨预报网络[J]. 电子测量技术, 2023, 46(17):102-108.
- YIN CH H, QIN H W, DAI Y W, et al. GLnet: A short-term rainfall forecast network integrating global and local information [J]. Electronic Measurement Technology, 2023, 46(17): 102-108.
- [12] SINGH S, SARKAR S, MITRA P. A deep learning based approach with adversarial regularization for Doppler weather radar ECHO prediction [C]. 2017 IEEE International Geoscience and Remote Sensing Symposium(IGARSS), 2017: 5205-5208.
- [13] 黄金贵,黄一举. 基于注意力时空解耦 3D 卷积 LSTM 的视频预测[J]. 微电子学与计算机, 2022, 39(9): 63-72.
- HUANG J G, HUANG Y J. Video prediction based on attention spatiotemporal decoupling 3D convolutional LSTM [J]. Microelectronics and Computers, 2022, 39(9): 63-72.
- [14] KHANAN D, JINDAL N, RANA P S, et al. Enhanced spatio-temporal 3D CNN for facial expression classification in videos [J]. Multimedia Tools and Applications, 2024, 83(4): 9911-9928.
- [15] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]. European Conference on Computer Vision(ECCV), 2018: 3-19.
- [16] FU J, LIU J, TIAN H J, et al. Dual attention network for scene segmentation [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 3146-3154.

作者简介

包广斌,副教授,研究生导师,博士,主要研究方向为计算机网络体系结构、人工智能、计算机视觉和图像处理。

E-mail:bao.gb@qq.com

罗瞳(通信作者),硕士研究生,主要研究方向为人工智能、计算机视觉和图像处理。

E-mail:luotong999@qq.com