

杜先君, 郭航飞, 程生毅. 用于短临降水预报的多尺度注意力编码-动态解码网络[J]. 海洋学报, 2024, 46(12): 122-134, doi:10.12284/hyxb2024119

Du Xianjun, Guo Hangfei, Cheng Shengyi. Multiscale attentional coding-dynamic decoding network for short-range precipitation nowcasting[J]. Haiyang Xuebao, 2024, 46(12): 122-134, doi:10.12284/hyxb2024119

用于短临降水预报的多尺度注意力编码-动态解码网络

杜先君¹, 郭航飞¹, 程生毅¹

(1. 兰州理工大学 电气工程与信息工程学院, 甘肃 兰州 730050)

摘要: 短临降水预报是气象学和水文学中的重要任务之一,但在现有深度学习方法中,其预测结果模糊不清,并且累计误差大。为了克服这些预测方法中存在累计误差的局限性,以及预测序列结果模糊不清的问题,本文构建了一种基于多尺度注意力编码-动态解码网络(Multi-scale Attention Encoding-Dynamic Decoding Network, MAEDDN)的短临降水预报方法,通过学习输入数据的时空特征来预测未来的降水情况。为了得到更多输入序列的特征信息,在编码过程中,使用带有空间及通道注意力的卷积块进行编码,并增加多尺度融合模块解决降水分布中小尺度与大尺度信息无法同时捕获的问题;增强预测序列的清晰度,需要模型更好地理解降水过程,因此在解码过程中,针对短临降水过程伴随的生成与消散过程,提出了一种动态解码网络,通过学习输入过去数据的强度分布及变化趋势对解码过程进行灵活地筛选。使用公开数据集 SEVIR 的降水数据进行实验,并与现有最好模型进行对比,实验结果表明:(1) MAEDDN 提升了在高强度降水区域的预测能力;(2) MAEDDN 预测的图像序列清晰度显著优于其他模型。构建的多尺度注意力编码能够更好地捕捉气象数据的复杂关系;动态解码能够根据不同的情况自适应地选择解码过程,提供更准确的预测结果。

关键词: 短临预报; 动态解码网络; 时空特征; 多尺度融合

中图分类号: P457.6

文献标志码: A

文章编号: 0253-4193(2024)12-0122-13

1 引言

短临降水预报与海洋学研究密切相关。海洋通过水汽输送和海气相互作用显著影响沿海地区的降水模式,尤其在极端天气事件(如强对流降水和台风)中,海洋对大气的能量输送在降水形成中发挥了关键作用。短临降水预报主要针对中小尺度天气系统,尤其是强对流天气,实现 0 到 2 h 内的精准预报。与大尺度天气系统不同,短临降水具有突发性强、机理复杂、生命周期短的特点,使其预测成为一个极为复杂和富有挑战性的课题。基于海洋学研究的相关性,本研究结果在沿海气象应用中具有间接的参考价值。近年来,气候变化问题引起了全世界的关注,极端降

水事件的频发给国家和地区带来了巨大的损失。因此,针对强对流天气如此之强的破坏性,就必须提供及时可靠的短临降水预报手段^[1]。

目前,短临降水预报方法主要是雷达回波法,利用雷达发射的微波信号与大气中的降水粒子发生散射,根据接收到的回波信号强度和反射率,估计降水的强度和分布。这种传统的外推方法主要是质心跟踪、互相关和光流法。其中,质心跟踪法首先需要确定要跟踪的降水区域的质心位置,根据连续观测到的质心位置,推断降水区域的运动方向和速度,进而预测未来的位置和路径。然而该方法往往无法考虑到降水系统内部的微观结构和变化,导致在短临预报中可能无法捕捉到细节信息。互相关法将整个数据区

收稿日期: 2024-06-19; 修订日期: 2024-10-17。

基金项目: 国家自然科学基金(62241307); 甘肃省科技计划项目(22YF7FA166、24JRRA173、24CXGA050); 兰州市科技计划项目(2024-3-47)。

作者简介: 杜先君(1979—),男,浙江省杭州市人,副教授,博士生导师,主要研究方向为复杂系统建模与控制。E-mail: xdu@lut.edu.cn

域划分为几个小区域,然后计算相邻时间雷达回波图像小区域之间的相关系数。回波区域的平均运动由区域之间的相关性决定。但是在恶劣的对流天气下,错误跟踪将显著增加。光流法通过两个基本假设(亮度恒定性、小运动假设)计算像素在目标表面上运动的瞬时速度场,然后预测下一帧图像。光流法的优点是在不了解场景信息的情况下可以检测和识别运动目标的位置,但无法检测运动目标的旋转形变与像素点之间的生消反应。

近年来,随着人工智能(Artificial Intelligence, AI)的快速发展,在计算机视觉(Computer Vision, CV)、自然语言处理(Natural Language Processing, NLP)等领域已经获得了大量成果。很多学者将短临预报问题视为时空序列回归问题。因此循环神经网络(RNN)及其变体、长短期记忆(LSTM)和门控循环单元(GRU)已被广泛使用。在LSTM的变体中,Shi等^[2]提出的ConvLSTM是短临降水预报的代表性模型,后续加以改进的TrajGRU^[3]模型一度成为短临降水预报任务中的基准模型。Wang等人的PredRNN^[4]与PredRNN++^[5]模型中设计了一个新的时空LSTM(ST-LSTM)单元。对于上述模型来说,它们在原始形式下通常是用于预测下一个时间步的输出,而不是一次性预测多个时间步。这是因为这些模型都是基于循环结构的,每个时间步的输出都依赖于前一时刻的输入和模型的状态。常见的方法是使用循环方式进行多步预测。在每个时间步,模型生成一个时间步的输出,然后将这个输出作为下一个时间步的输入,迭代多次来预测多个时间步的输出。另一种方法是使用自回归方式进行多步预测。模型在每个时间步生成一个时间步的输出,并将其追加到输入序列中,然后用更新后的输入序列继续生成下一个时间步的输出。在这样的预测方式下,多步预测会增加预测的不确定性和累计误差,并且误差会随着预测步数的增加而逐渐累积放大的现象。

为了解决上述问题,本文设计了一种新的编码器-解码器结构来构建用于短临降水预报的多尺度注意力编码-动态解码网络(Multi-Scale Attention Encoding-Dynamic Decoding Network, MAEDDN)。所设计的网络训练和预测只需要输入所选地区过去降水量的分布图,便能预测未来时间的降水图。该方法只使用固定长度的向量特征进行预测,而不是直接使用前一步的预测结果,因此误差累积的问题得到有效的缓解。Xue等^[6]提出的Widenet,该模型的方向是更宽而不是更深,通过混合专家(MoE)代替前馈网络(FFN),使模型沿宽度缩放,使用单独LN用于转换各种语义表

示,而不是共享权重。受这项工作的启发,本文设计了动态解码路由块(dynamically decode routing blocks, DDRBs),具有像素级的选择机制,充分融合四层解码器,聚合了短临降水任务的生消机制。在研究工作中,设计并采用多尺度注意力模块(multi-scale attention module, MAM)来进一步从编码网络中提取高级语义特征,并提高全局上下文特征利用的效率。为了评估所提算法的性能,本文在SEVIR数据集上进行了实验,所提算法在不同的评估指标下均取得了有竞争力的性能。

本文的主要贡献可以总结如下:

(1)设计了多尺度注意力模块(MAM),区别于以往模型仅依赖单一尺度特征提取的方式,该模块通过构建多尺度感受野,提取降水序列的更多形状信息与细节信息,从而更好地预测不同尺度的降水序列。

(2)动态解码路由块(DDRBs),与只对模型预测过程中的中间特征做处理不同的是,该模块通过学习输入降水序列的模式得出动态解码权重,从而对模型提取出的特征进行加权融合与筛选,为模型补充降水过程生消机制。

2 相关工作

在用于短临降水预报的深度学习方法中,大致可以分成两个方法:基于卷积神经网络的编码-解码网络和基于循环神经网络预测模型。

2.1 卷积神经网络

这种方法通常使用编码器-解码器结构,其中编码器负责提取输入数据的特征,解码器则将这些特征转换为降水预报结果。编码器部分由多个卷积层和池化层组成,解码器部分通过反卷积层和上采样操作将提取的特征映射回原始数据的尺寸,在保留空间信息的同时,有效地学习输入数据的时空关系。

吴卓升等^[7]在卷积神经网络基础上搭建动态概率层,以增强预测序列与输入序列的关联性。SmaAt-UNet^[8]基于UNet架构开发,并配备注意力模块和深度可分离卷积,通过荷兰地区的降水图和法国云覆盖的二进制图像评估模型在真实数据集验证,在预测性能方面,SmaAt-UNet与其他预测模型相当,且仅使用了四分之一的可训练参数。Chen和Wang^[9]使用单帧3D(经度、纬度和高度)气象分析场(温度、位势高度、相对湿度、纬向和经向风速)作为输入进行一系列3D卷积,自动提取水平和垂直特征用来生成未来24h内的总降水量。Qiu等^[10]集成了图卷积网络(GCN)、门控循环单元(GRU)和经验模态分解(EEMD),有效

缓解了卷积神经网络 CNN 对于非欧式数据的局限性,且模型可以同时实现单步预测和多步预测。基于 UNet 和 SegNet 架构提出的 RainNet^[11],采用递归的方法预测下一时刻的降水分布,模型通过使用德国气象局提供的雷达合成图进行训练与测试,实验结果明显优于光流等传统模型。曹伟华等^[12]在 RainNet 基础上建立了不同的预报方式,以京津冀地区降水进行实验分析,讨论了不同预报方式对预报性能的影响。Yang 和 Yuan^[13]提出定制多尺度深度学习框架,并设计了一个捕捉尺度变换的模块,强调了多尺度时空建模的重要性。Wu 等^[14]为了缓解预测模型中局部灵活性差、长时间跨度预测性能差等缺陷,提出了局部卷积与基于全局傅立叶变换并行的框架,从而能够动态捕捉局部-全局空间相互作用和依赖关系。

2.2 循环神经网络

这种方法主要基于循环神经网络(如 LSTM)或其变种,用于捕捉时间序列数据中的时序信息。通过将过去的降水观测序列作为输入,模型能够学习到时间序列的规律,并预测未来的降水情况。

方巍等^[15]考虑到卷积运算在处理剧烈变化特征的局限性,采用光流法进行特征追踪,并引入对抗网络,拟合预测序列的分布。Yu 等^[16]通过将预训练的 TransGAN 生成器与 ConvLSTM 相结合,能够较好地提取时空序列特征且有效缓解了循环神经网络预测结果分辨率不足的缺陷。时间感知自适应特征融合网络 TAFFNet 等^[17],将基于物理学的数值天气预报与雷达回波数据自适应融合到统一的深度学习模型中,进而为模型提供更多的时空信息。庄潇然等^[18]将雷达反射率因子与降水观测资料融合至 PhyDNet 模型中,可以更好地再现降水区域的分布特征。基于风场与降水数据融合的模型 ASS-TGCN^[19],则使用了改进的图卷积神经网络结合风向及地理位置信息捕捉空间信息,并使用门控循环单元捕捉时间信息。Chen 等^[1]在 LSTM 基础上分析确定了降雨时间序列最佳滞后时间的最佳窗口大小,提升了降雨预测的确定性。Fang 等^[20]为了解决雷达回波外推模糊的困境,选择时空长短期记忆(ST-LSTM)作为模型的循环单元,并将 3D 卷积操作集成到模型中,以增强模型捕捉雷达回波运动趋势的能力。Xiong 等^[21]为了克服输入数据和网络输出之间的上下文关系缺失的问题,对 LSTM 网络的输出进行卷积操作以捕捉全局依赖关系。Wu 等^[22]提出的 ISA-PredRNN 基于 PredRNN-V2 模型,在模型中引入了自注意力机制和长期记忆状态,并设计了一组新的门控机制,能够更好

地捕捉不同强度的降水。Lin 等^[23]将自编码器作为特征提取层,ViT 作为预测层,实现了对于特征的高效学习。Küçük 等^[24]使用基于 transformer 的预测模型,可预测出不同天气现象下的雷达场,并且在衰减场和复杂场的预测上有良好的稳健性。

综上,与传统雷达回波外推方法相比,卷积神经网络在训练时能够利用大量的历史雷达回波资料,从中学习其变化规律,包括降雨强度演变的增长和衰减过程。但是,卷积结构的位置不变性使雷达图像具有相同的降雨场转换,存在一定的局限性。循环神经网络在训练时能够建立时序关系,但由于短期降水过程图像序列相邻帧变化剧烈,在递归预测过程中会造成巨大的累计误差。

3 方法

3.1 动态编解码网络

为了解决预测过程中存在累计误差与降水演变过程难以学习的问题,本文提出了多尺度注意力编码-动态解码网络,将降水序列按照时间顺序整合为一个输入,在编码过程,通过多尺度特征提取捕捉输入图像序列更多的空间信息,并融入注意力机制提取序列中空间与时间变化信息。在预测部分使用反卷积操作,并且考虑到降雨存在生成与消散等突发过程,加入了由输入决定的像素级路由模块对预测进行干预,在实验过程中,发现本文的方法能够较好地预测出未来降雨的分布及强度。

多尺度注意力编码-动态解码网络如图 1 所示。编码器将原始的气象数据作为输入,通过学习将降雨分布运动与时间进行耦合并且通过多尺度特征融合补充降雨分布的运动细节信息表示成中间向量,然后用解码器从中间向量中生成初步的降水预报图,并由输入所决定的动态解码路由模块对解码器的生成图进行加权融合形成最后的降水预报图。图 1a 为多尺度注意力编码模块,实现输入通过不同的尺度变换进行编码;图 1b 为动态路由解码模块,确定由输入序列所决定的选择张量,在预测阶段提供不同的权重信息。设任务为:输入给定过去 10 帧降水分布图(5 min/帧) $I_i \in R^{h \times w}$, $i = 1, 2, \dots, 10$,为了方便后续表示,将输入记为 I_{1-10} 。来预测未来 10 帧的雷达回波图 $\tilde{I}_{t+1}, \tilde{I}_{t+2}, \tilde{I}_{t+3}, \dots, \tilde{I}_{t+10}$,记为 \tilde{I}_{t-10} 。

将短临预测模型设为 $G^\theta(I_1, I_2, I_3, \dots, I_{10})$,其中 θ 是模型要学习的参数集合,学习目标为:使模型预测值: $\tilde{I}_{t+1}, \tilde{I}_{t+2}, \tilde{I}_{t+3}, \dots, \tilde{I}_{t+10} = G^\theta(I_1, I_2, I_3, \dots, I_{10})$ 与真实观测值: $I_1, I_2, I_3, \dots, I_{10}$ 之间差异值最小。在任务中可以使用

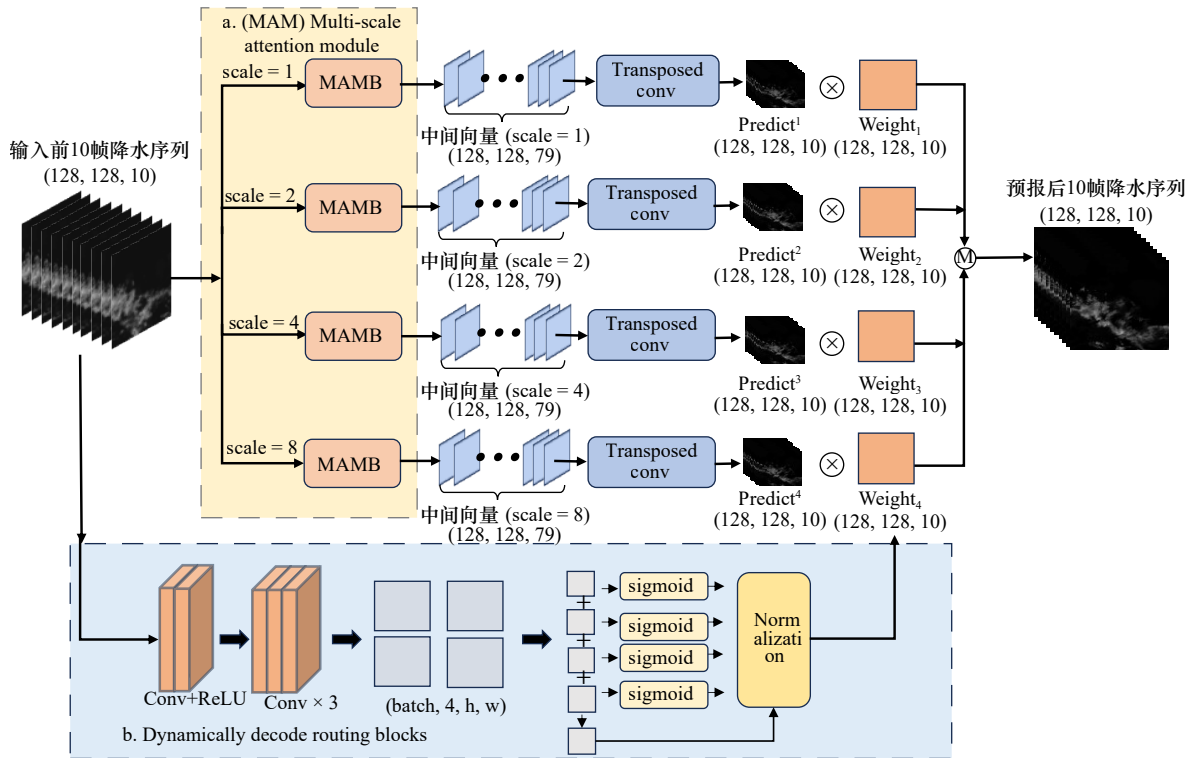


图1 多尺度注意力编码-动态解码网络结构

Fig. 1 Multi-scale attention encoding-dynamic decoding network structure

a. 多尺度注意力编码模块, b. 动态路由解码模块

a. Multi-scale attention encoding module, b. dynamic routing decoding module

下式表示:

$$I_{inter}^i = \text{MAMB}_{\text{scale}=i}(I_{1-10}), \quad (1)$$

$$\text{Pred}^i = \text{DeConv}_{\text{scale}=i}(I_{inter}^i), \quad (2)$$

$$\{\text{Weight}_i | i = 1, 2, 3, 4\} = \text{DDRDB}(I_{1-10}), \quad (3)$$

$$\text{Output} = \sum_i \text{Pred}^i \times \text{Weight}_i \quad i = 1, 2, 3, 4, \quad (4)$$

式中, MAMB为网络中的多尺度注意力编码块; DeConv为反卷积; DDRDB为动态解码路由模块; Pred^i 为各尺度所生成的预测结果; Weight_i 为路由模块生成的权重张量; Output为最终加权融合的预测结果。

3.2 多尺度注意力模块

降水是一个复杂的多尺度过程,涉及到从小尺度的湍流到大尺度的气旋系统等多个尺度的现象。编码网络提取特征时,无法同时考虑到不同尺度降水发生过程中的运动强度变化,通过多尺度机制能够捕捉并处理不同尺度的信息。并且在降水过程中流体的发展与移动路径是降水预报的关键因素,通过注意力机制帮助模型聚焦于降水发生的关键区域。为了缓解降水运动变化尺度不同的采集域问题,同时获取更有价值的特征,本文设计了多尺度注意力模块(MAM)。

MAM的主要目的是在存在大量不同尺度的运动变化的情况下进行多尺度编码融合,使网络能够有效捕捉各种尺度的细节信息。MAM使用4个并联的MAMB进行特征提取,通过输入尺度列表scale进行控制, MAMB的结构如图2所示。在图2中, MAMB由双线性插值、卷积运算、卷积注意力机制(CBAM)、张量联接运算组成。F1表示主干道特征提取,输入由前端的尺度信息决定; F2表示副干道特征提取,为主干道特征提供全局信息。该模块利用CBAM^[25]模块获得卷积运算后的更具价值的特征信息,并将尺度特征与全局特征联接为中间特征,为后续的解码做准备。CBAM模块根据人类视觉特性能够提取感兴趣的特征。在CBAM中沿着两个独立的维度(通道和空间)依次判断注意力图,然后将注意力图与输入特征图相乘以进行自适应特征优化。本文中将CBAM模块放在特征提取的最后一层。图2中卷积神经网络的详细配置如表1所示。

由于降水过程复杂多变,模型对于多种运动强度的编码困难,同样权重的编码网络无法同时提取输入图像中的大尺度与微小尺度变化,而将输入图像的尺寸缩小可以更好地检测出降水区域的运动变化。本文设计的MAM可以选择性地、充分地融合各尺度的

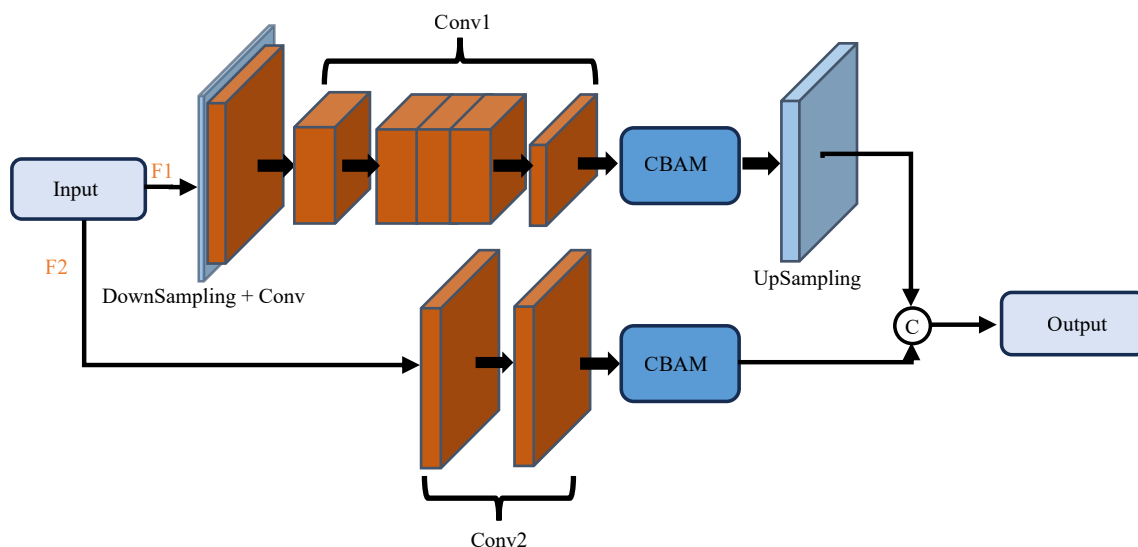


图 2 多尺度特征注意力模块 (MAMB) 架构的详细信息

Fig. 2 Detailed information on the Multi-scale Attention Feature Module (MAMB) architecture

表 1 MAMB 中卷积神经网络参数配置

Table 1 Parameter configuration for convolutional neural networks in MAMB

F1 主干道卷积神经网络		
类型	卷积核	步长
Conv	$3 \times 3 \times 150$	1×1
	PRReLU	
Conv	$3 \times 3 \times 300$	1×1
	PRReLU	
Conv	$3 \times 3 \times 300$	1×1
	PRReLU	
Conv	$3 \times 3 \times 300$	1×1
	PRReLU	
Conv	$3 \times 3 \times 75$	1×1
	PRReLU	
F2 副干道卷积神经网络		
类型	卷积核	步长
Conv	$3 \times 3 \times 32$	2×2
	PRReLU	
Conv	$3 \times 3 \times 32$	1×1
	PRReLU	

特征信息,使 MAEDDN 能够更好地对降水运动信息进行编码,最终实现更好的预测结果。在消融实验中将进一步证明 MAMB 的有效性。

3.3 动态解码路由模块

本工作设计的 DDRB 是为了处理在解码过程中图像的遮挡问题(生成与消散过程),在降水过程中不同时间点,某些特征可能比其他特征更为重要,例如,

在降水开始阶段,水汽含量和不稳定能量可能是决定性因素,而在降水结束阶段形成和扩散则变得更为关键。动态解码器通过学习过程中的动态权重分配,能够自动识别哪些特征在特定时间点更重要,并据此调整预测结果,通过多尺度注意力编码器提取的特征,动态解码器能够根据不同尺度特征之间的相互作用和关联性进行融合,从而提高预报的准确性。其中包括对原始输入图像的编码与对于中间向量的解码等操作。如图 1b 中所示,首先获取原始输入连续的运动特征信息并且经过 sigmoid 层形成解码器的选择张量,在主干网络 MAM 提取特征后使用反卷积神经网络将各尺度特征信息解码为图像序列,在最后通过选择张量融合预测结果。如下式所示:

$$M = \text{Conv}_2(\text{ReLU}(\text{Conv}_1(I_{1-10}))) \quad M \in \mathbb{R}^{b \times 4 \times b \times w}, \quad (5)$$

$$W'_i = \text{Sigmoid}(W_i) \quad W_i \in M, \quad (6)$$

$$\bar{W}'_i = \frac{W'_i}{\sum_j W'_j} \quad j = 1, 2, 3, 4, \quad (7)$$

式中, M 为输入图像经过卷积神经网络提取出的特征信息, W 为各尺度特征信息进行概率化后的结果; \bar{W}'_i 为将 W' 归一化后的结果。

在 DDRB 中的卷积神经网络具体参数配置如表 2 所示。

面对降水过程中的生成与消散、边界遮挡问题,使用单一解码器可能无法灵活处理像素点的生成与消散。本文设计的 DDRB 旨在能够通过特征学习去理解输入中不同类型的降水过程,对预测结果进行动态选择,从而更好地预测降水分布过程。消融实验中

表2 DDRB中卷积神经网络参数配置
Table 2 Parameter configuration for convolutional neural networks in DDRB

类型	卷积核	步长
Conv	$3 \times 3 \times 32$	1×1
	ReLU	
Conv	$3 \times 3 \times 32$	1×1
Conv	$3 \times 3 \times 32$	1×1
Conv	$3 \times 3 \times 4$	1×1

将进一步证明 DDRB 的实用性。

3.4 损失函数

本文中使用的损失函数如下式:

$$\text{MSEloss} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}, \quad (8)$$

$$\text{L1loss} = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n}, \quad (9)$$

$$\text{loss} = \text{MSEloss} + \alpha \times \text{VGGloss} + \text{L1loss}, \quad (10)$$

式中, MSEloss为均方误差,是反应估计量与被估计量之间差异程度的一种度量。L1loss是把目标值与模型输出(估计值)做绝对值得到的误差。为了使预测图像更符合人类观念,本文引入了通过系数 α 调节的感知损失VGGloss^[26],通过将输入图像与真实图像送入

预训练的VGG19网络中提取各层特征,然后对比各层特征之间的差异(在本文中 α 设置为0.8,其中详细说明在消融实验部分)。

4 实验

4.1 数据集

为了评估所提出的网络的性能,在全球公开气象数据集SEVIR^[27]上进行实验。该数据集是时间和空间对齐图像的集合,其中包含卫星和雷达捕获的天气事件。该数据集是使用NOAA分发的公开数据集创建的,包括GOES-16对地静止卫星和NEXRAD天气达的数据。

在本文中使用垂直累积液态含水量进行实验,液态含水量是利用反射率因子强度资料和含水量之间的关系反演出云层含水量,是判别强降水及其降水潜力、短历时强降水、冰雹等灾害性天气的有效工具之一。图3中的图像显示了SEVIR数据集中降水事件的数千个样本之一。事件中包含在美国陆地使用NEXRAD雷达采样的 $384 \text{ km} \times 384 \text{ km}$ 斑块上的含水量数据,以5 min为间隔进行采样,每帧图像以灰度图的形式储存,一次天气事件共储存49帧。图像中的每个像素值为 $1 \text{ km} \times 1 \text{ km}$ 内的垂直累积液态含水量,数值范围为 $[0 \sim 255]$,数值越大则对应含水量越高。

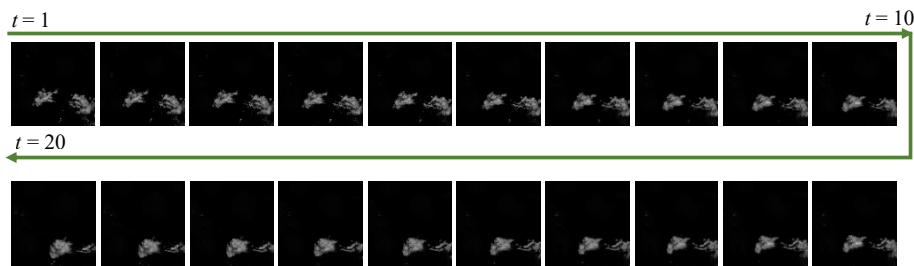


图3 SEVIR数据集中VIL示例

Fig. 3 Examples of VIL in the SEVIR dataset

4.2 训练

本文所提出的网络模型基于Pytorch框架,训练时,首先对网络进行50个epoch的预训练,优化器使用Adam^[28],预训练的学习率为 10^{-3} ,批量大小设置为18。然后将学习率设置为 10^{-4} ,使模型进行精细化学习,本文采用了提前停止标准,当测试损失在15个epoch未下降时,便停止训练。

本文实验环境基于Linux操作系统,系统内存为64GB,显卡采用NVIDIA 3060,显存为12GB。详细相关配置如表3所示。

在网络训练过程中,使用20个图像的滑动窗口

表3 实验环境配置

Table 3 Experimental environment configuration

名称	相关配置
操作系统	Linux-5.15.0
处理器	Intel Core i7-12700K
内存	64 GB
显卡	NVIDIA 3060
深度学习框架	PyTorch 1.8.1

对数据集进行划分采样,前 10 帧为输入数据,后 10 帧图像为真实数据。由于 GPU 显存有限,将图像裁剪为 128×128 的大小。在实验过程中数据集的详细配置如表 4 所示。

表 4 数据集详细配置

Table 4 Detailed configuration of the dataset

数据集	事件数	图像数	模式
SEVIR_VIL_2017	2 778	55 560	训练
SEVIR_VIL_2018	1 712	34 240	测试

4.3 评估指标

本文通过一些广泛使用的综合指标来评估所提出的算法,除了 MSE 之外还包括关键成功指数(CSI)、检测概率(POD)、误识率(FAR),这些指数常被用于评价临近降水预报的综合性能。CSI、POD、FAR 定义如下式:

$$CSI = \frac{Hits}{Hits + Misses + F.Alarms}, \quad (11)$$

$$POD = \frac{Hits}{Hits + Misses}, \quad (12)$$

$$FAR = \frac{F.Alarms}{Hits + F.Alarms}, \quad (13)$$

式中, Hits 真实值为 1, 预测值为 1 的总数; Misses 为真实值为 1, 预测值为 0 的总数; F.Alarms 为真实值为 0, 预测值为 1 的总数。

为了计算 Hits、Misses 和 F.Alarms, 预测值与真实值被重新调整到 0~255 范围并在阈值 [16,133,219] 处进行二值化。

4.4 对比实验

4.4.1 定量实验

为了评价短临降水预报的成功率,与性能最好的模型 SmaAt-UNet^[8]、ConvLSTM^[2]、PredRNN^[4]、PhyDnet^[29]、BEP^[30] 进行了定量比较实验。表 5 显示了在 SEVIR_VIL_2017 数据集上训练并在 SEVIR_VIL_2018 数据集上评估的结果。表中的评估指标包括不同阈值下的 CSI 和 MSE。指标的最佳性能使用粗体进行标注。比较结果表明,本文设计的模型从输入图像序列中捕获了更有价值的特征,并生成更接近未来真实的图像序列。定量比较结果表证明,所设计的算法 MAEDDN 在中、高强度降水区域(CSI-219、CSI-181、CSI-160)上取得了有竞争力的结果。并且为了客观衡量对于不同强度的降水运动的预报能力,进行了不同强度的检验评估分别以 16、133、219 作为阈值来进行二值化。在上述 3 个阈值条件下,进行了 10 个时间步的外推实验。如图 4 所示。

表 5 SEVIR 数据集上与 SOTA 模型对比实验结果(加粗数值为该指标性能最好)

Table 5 Comparison of experimental results with SOTA models on the SEVIR dataset (bold values indicate the best performance for that metric)

模型	评价指标							
	CSI-M \uparrow	CSI-219 \uparrow	CSI-181 \uparrow	CSI-160 \uparrow	CSI-133 \uparrow	CSI-74 \uparrow	CSI-16 \uparrow	MSE(10^{-3}) \downarrow
SmaAt-UNet	0.378 1	0.107 7	0.148 0	0.255 7	0.340 4	0.673 1	0.744 0	4.001 5
ConvLSTM	0.418 5	0.128 8	0.248 2	0.292 8	0.405 2	0.679 3	0.756 9	3.753 2
PredRNN	0.408 0	0.131 2	0.232 4	0.276 7	0.385 8	0.671 3	0.750 7	3.901 4
PhyDnet	0.394 0	0.128 8	0.230 9	0.270 8	0.372 0	0.655 6	0.705 9	4.816 5
BEP	0.384 3	0.110 5	0.162 9	0.265 0	0.350 6	0.673 0	0.743 9	3.986 8
MAEDDN(ours)	0.452 3	0.233 0	0.307 7	0.364 5	0.394 5	0.651 4	0.762 9	3.446 3

分析可知 MAEDDN 在所有强度上 CSI 值表现最佳,模型具有更强的鲁棒性,随着阈值增加和预报时效延长,对比的模型预报能力均显著下降,然而 MAEDDN 下降趋势相比更为缓慢,对于高强度的预报能力(133、219)优于其他模型。

4.4.2 定性实验

在本文中展示了 MAEDDN 与对比算法 SmaAt-UNet、ConvLSTM、PredRNN、PhyDnet、BEP 之间的定性比较结果,以便更直观的结果来验证所提出方法的

良好性能。其中 Ground truth 表示真实观测值, Input 表示模型输入, Ours 表示 MAEDDN 模型的预报结果。在预测出的降水图中根据像素值由大到小进行由深到浅的颜色标注,为了验证各个算法的预测有效性,在中小强度降水过程与中高强度降水过程数据集上进行对比。实验结果如图 5 和图 6 所示。图 5 为中小强度降水过程,图 6 为中高强度降水过程。

可以在各模型预测图像序列中观察到 ConvLSTM、PhyDnet、PredRNN 预测结果呈现消散趋势,虽能大概预

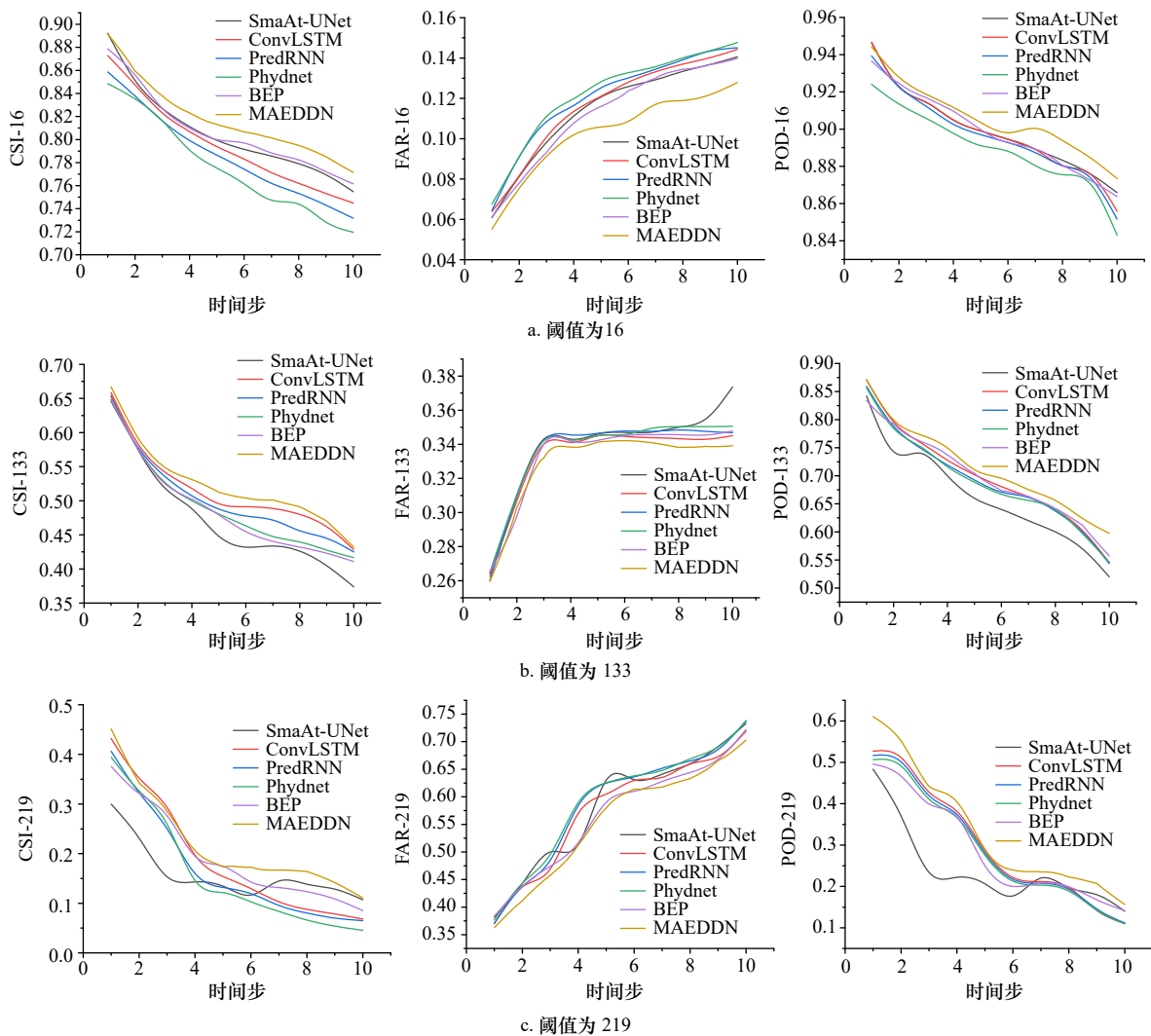


图4 反射率阈值在16(a)、133(b)、219(c)的10个时间步外推过程中的指标变化

Fig. 4 Variation of metrics during the extrapolation process over 10 time steps with reflectivity thresholds of 16 (a), 133 (b) and 219 (c)

测降水的主体位置,但降水强度预测不准确,并且降水区域的运动预测也不够准确。SmaAt-UNet虽然主体降水位置预测较为精确,然而降水强度(特别是高强度降水区域)预测错误。BEP虽然预测出的轮廓与强度值比上述模型有所改善,但是预测的不够精确。MAEDDN所预测的降水序列,不仅能够准确预测降水区域的位置,而且降水强度及运动预测得更加精准。定性实验结果表明MAEDDN能够更好地处理降水过程中的运动与强度并生成准确的未来时刻降水序列。基于LSTM的算法在面临降水强度变化时,存在严重强度错误预测,而本文设计的MAEDDN缓解了这一问题。主观实验结果图表明,本文设计的多尺度注意力模块(MAM)能够充分捕捉到输入图像序列中的运动强度变化信息,有效注意到微小尺度的强度特征,使编码网络提取到更强的运动时空编码信息。同时,路由动态解码模块(DDMR),进一步对输出结

果进行选择,以便使解码网络更加灵活处理降水区域的运动。

4.5 消融实验

MAEDDN以CNN为骨干网络进行测试,依次添加MAM和DDRM,并使用CSI指标对预测图像序列进行评估。在保证算法所有其他配置不变的情况下,SEVIR数据集上的测试性能如表6所示。通过将MAM模块加入模型后,可以观察到在CSI-16与CSI-74低强度降水阈值处指标有明显提升,而对于CSI-133等中高强度提升有限。这表明MAM能够有效得提取降水区域的变化特征,然而对于降水过程的生消反应不够敏感(更容易偏向均值),导致预报值不够精确。在加入DDRM模块后,CSI-133等中高强度得到了显著地提升,这说明DDRM通过对中间特征向量的加权融合与筛选,补充了模型对于局部细节信息变化的感知能力。

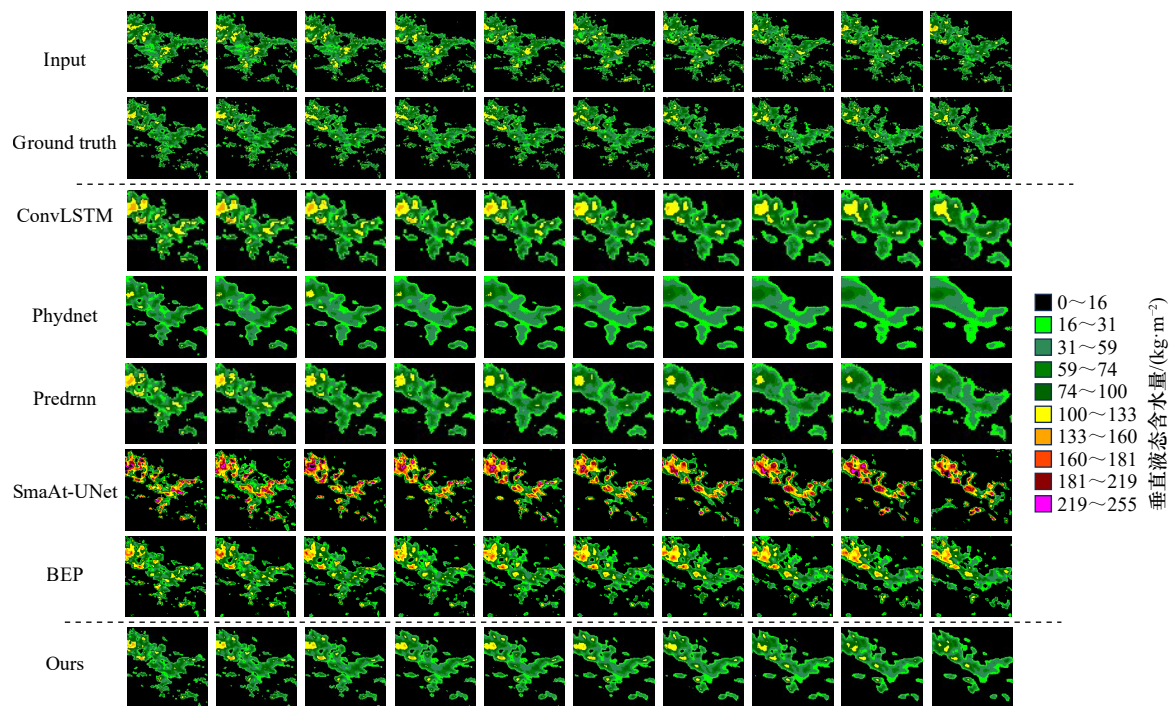


图 5 对比实验预测图像序列结果(生成)

Fig. 5 Results of the generated image sequences from the comparison experiments (generation)

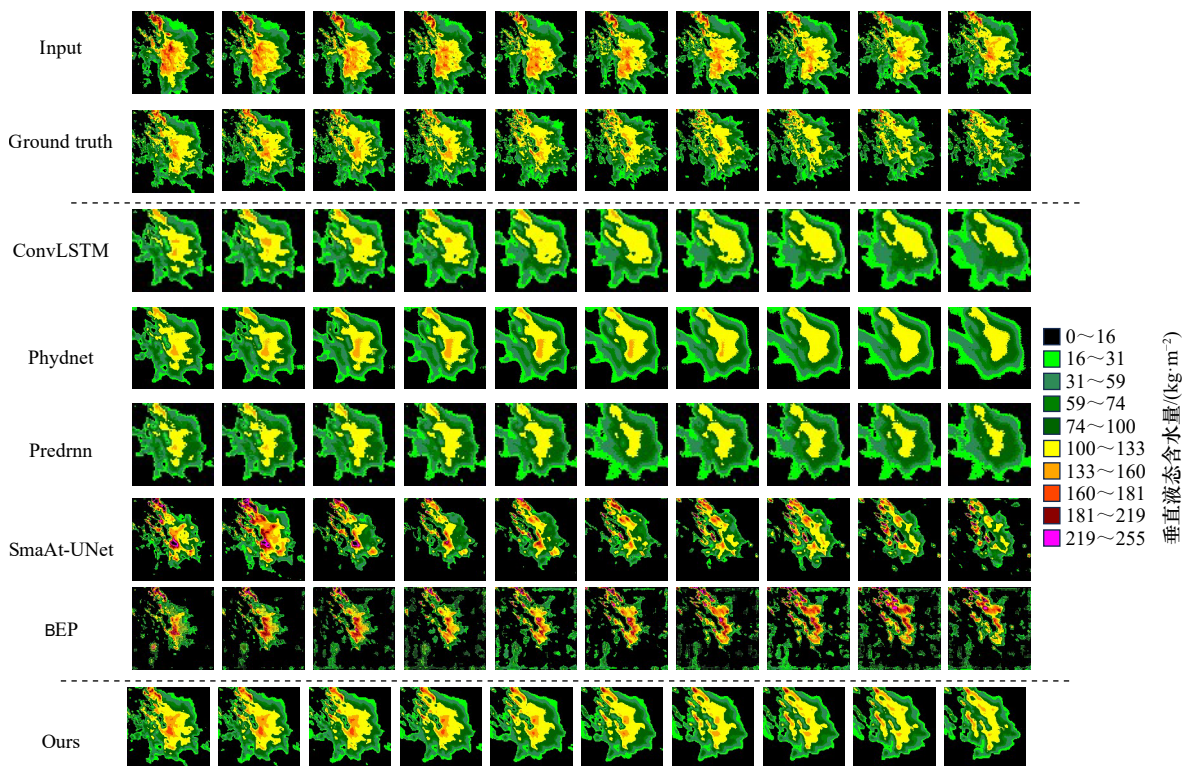


图 6 对比实验预测图像序列结果(消散)

Fig. 6 Results of the generated image sequences from the comparison experiments (dissipation)

为了更直观地展示和解释所提出的 MAM 和 DDRM 的效果,在本实验中引入了 MAM 和 DDRM 逐步添加的预测序列结果图。如图 7 所示,其中 Input 表示模型输入, GT 表示实际观测值, Backbone 表示基准网

络。可以看到 Backbone 所预测的图像序列只能获知未来序列的大致轮廓,并且降水强度严重失真并且降水区域的分布预测不准确;在加入 MAM 模块后可以观察到预测序列中的降水强度有非常大的改善,但对

表 6 提出网络的消融实验结果 (CSI) (加粗数值为该指标性能最好)

Table 6 Ablation experimental results of the proposed network (CSI) (bold values indicate the best performance for that metric)

模型设置	评价指标						
	CSI-M \uparrow	CSI-219 \uparrow	CSI-181 \uparrow	CSI-160 \uparrow	CSI-133 \uparrow	CSI-74 \uparrow	CSI-16 \uparrow
Backbone	0.356 3	0.070 5	0.229 5	0.258 7	0.274 1	0.604 7	0.700 6
+MAM	0.399 2	0.150 1	0.249 6	0.304 7	0.313 5	0.627 4	0.750 4
+MAM+DDRB	0.452 3	0.233 0	0.307 7	0.364 5	0.394 5	0.651 4	0.762 9

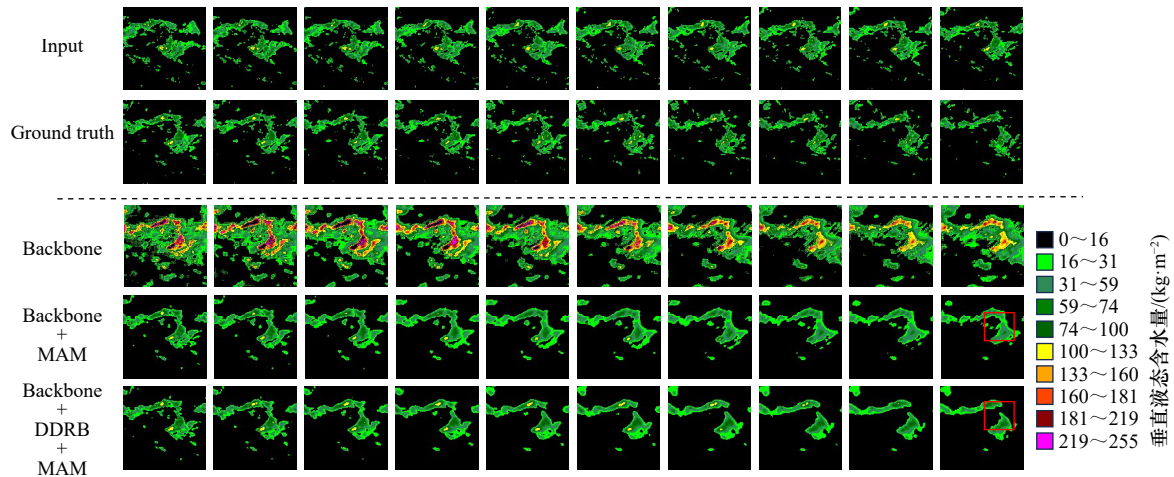


图 7 消融实验预测图像序列结果

Fig. 7 Results of the generated image sequences from the ablation experiments

于降水区域的运动预测有所欠缺,如图6中使用红色框框出的区域,无法预测未来时刻降水区域的分离。在加入 DDRB 模块后,有效解决了降水区域的运动变

化问题。通过添加 MAM 和 DDRB,可以更加有效分析序列中强度与分布运动编码信息,使网络对未来时刻降水图序列预测更加准确和精细。

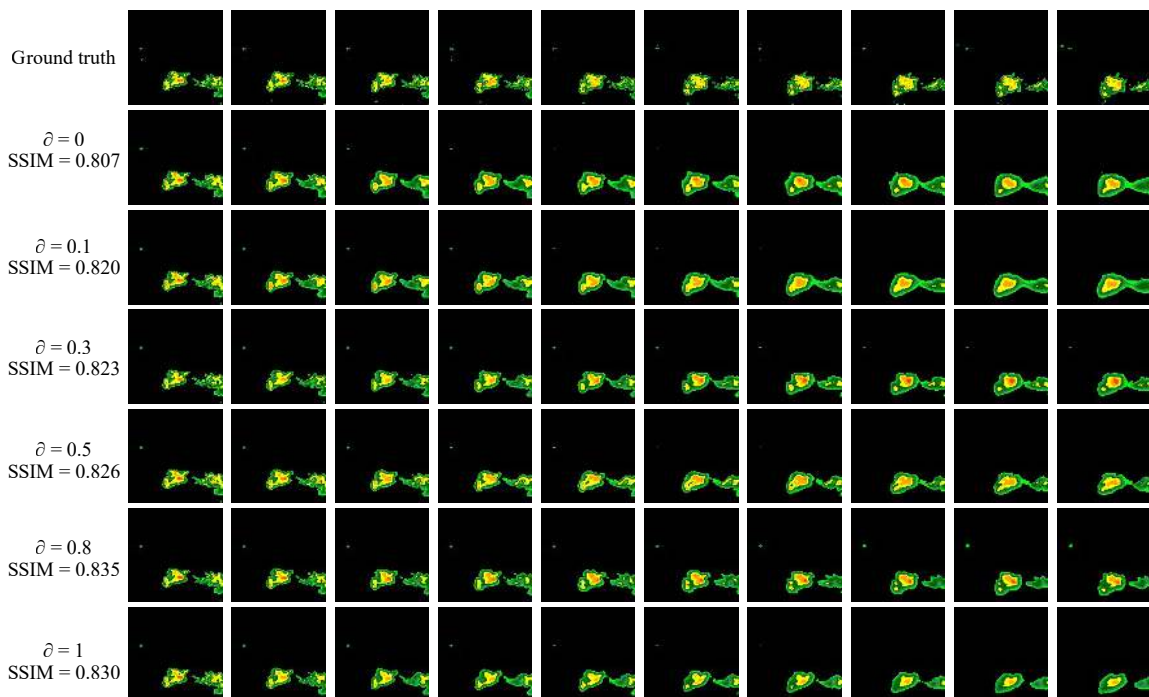


图 8 不同参数 δ 的对比实验

Fig. 8 Comparison experiments with different parameters δ

为了验证 Vgg 联合损失函数的有效性,通过配置参数 δ ,使 δ 进行递增,并使用 SSIM 指标对模型预测出的序列进行相似度检测,实验结果如图 8 所示。图中第 1 行为真实序列,第 2~6 行为参数 δ 递增后作为损失函数,并且模型以此训练并测试后的结果。可以发现随着参数 δ 的递增,模型预测序列的 SSIM 指标不断增长,并且预测序列的图像也更加偏向于真实。然而当 δ 递增至 0.8 时,SSIM 出现峰值, $\delta = 1$ 时,SSIM 指标开始下降,预测的图像也逐渐远离真实值。所以在模型训练过程中,联合损失函数中的 Vggloss 参数 δ 选择为 0.8。

5 结论

本文设计了一个用于短临降水预报的多尺度注意力编码-动态路由解码网络。该网络通过融合输入

图像序列的不同尺度的细节特征并通过注意力机制提取特征中的关键信息,最后使用由输入序列所决定的动态路由解码张量对预测结果进行筛选以指导降水区域的运动变化。实验结果表明:

(1)使用多尺度注意力编码架构可以更好捕获降水序列中的细节特征,能够同时预测出高强度与低强度的降水区域。

(2)对于在降水序列中的区域分布运动的情况,通过使用由输入序列指导的动态路由解码模块,有效提升了模型的预测能力。

MAEDDN 在短临降水预报任务中实现了有竞争力的性能。在未来的工作中可进一步对数据集进行改进,对数据集中的序列进一步标注:消散过程、生成过程、运动强度大小、降水强弱等标签为动态路由模块提供更可靠的支撑。

参考文献:

- [1] Chen Chengcheng, Zhang Qian, Kashani M H, et al. Forecast of rainfall distribution based on fixed sliding window long short-term memory[J]. *Engineering Applications of Computational Fluid Mechanics*, 2022, 16(1): 248–261.
- [2] Shi Xingjian, Chen Zhourong, Wang Hao, et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems-Volume 1. Montreal: MIT Press, 2015: 802–810.
- [3] Shi Xingjian, Gao Zhihan, Lausen L, et al. Deep learning for precipitation nowcasting: a benchmark and a new model[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017: 5622–5632.
- [4] Wang Yunbo, Long Mingsheng, Wang Jianmin, et al. PredRNN: recurrent neural networks for predictive learning using spatiotemporal LSTMs[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach: Curran Associates Inc., 2017: 879–888.
- [5] Wang Yunbo, Gao Zhifeng, Long Mingsheng, et al. PredRNN++: towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning[C]//Proceedings of the 35th International Conference on Machine Learning. Stockholm: PMLR, 2018: 5123–5132.
- [6] Xue Fuzhao, Shi Zijie, Wei Futao, et al. Go wider instead of deeper[C]//Proceedings of the 36th AAAI Conference on Artificial Intelligence. AAAI, 2022: 8779–8787.
- [7] 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 Zhuosheng, Zhang Wei, Lin Yan, et al. Application of dynamic probability convolutional neural network in radar echo extrapolation[J]. *Application Research of Computers*, 2021, 38(7): 2125–2129.
- [8] Trebing K, Stańczyk T, Mehrkanoun S. SmaAt-UNet: precipitation nowcasting using a small attention-UNet architecture[J]. *Pattern Recognition Letters*, 2021, 145: 178–186.
- [9] Chen Guoxing, Wang W C. Short-term precipitation prediction for contiguous United States using deep learning[J]. *Geophysical Research Letters*, 2022, 49(8): e2022GL097904.
- [10] Qiu Yunan, Lu Zhenyu, Fang Shanpu. A short-term precipitation prediction model based on spatiotemporal convolution network and ensemble empirical mode decomposition[J]. *IEEE/CAA Journal of Automatica Sinica*, 2022, 9(4): 738–740.
- [11] Ayzel G, Scheffer T, Heistermann M. RainNet v1.0: a convolutional neural network for radar-based precipitation nowcasting[J]. *Geoscientific Model Development*, 2020, 13(6): 2631–2644.
- [12] 曹伟华, 南刚强, 陈明轩, 等. 基于深度学习的京津冀地区精细尺度降水临近预报研究[J]. *气象学报*, 2022, 80(4): 546–564.
Cao Weihua, Nan Gangqiang, Chen Mingxuan, et al. A study on fine scale precipitation nowcasting in Beijing-Tianjin-Hebei region based on deep learning[J]. *Acta Meteorologica Sinica*, 2022, 80(4): 546–564.
- [13] Yang Shangshang, Yuan Huiling. A customized multi-scale deep learning framework for storm nowcasting[J]. *Geophysical Research Letters*, 2023, 50(13): e2023GL103979.
- [14] Wu Hao, Liang Yuxuan, Xiong Wei, et al. Earthfarsser: versatile spatio-temporal dynamical systems modeling in one model[C]//Proceedings of the 38th AAAI Conference on Artificial Intelligence. AAAI, 2024: 15906–15914.
- [15] 方巍, 庞林, 张飞鸿, 等. 对抗型长短期记忆网络的雷达回波外推算法[J]. *中国图象图形学报*, 2021, 26(5): 1067–1080.
Fang Wei, Pang Lin, Zhang Feihong, et al. Radar echo extrapolation algorithm based on adversarial long short-term memory network[J].

- Journal of Image and Graphics*, 2021, 26(5): 1067–1080
- [16] Yu Wenbin, Wang Suxun, Zhang Chengjun, et al. Integrating spatio-temporal and generative adversarial networks for enhanced nowcasting performance[J]. *Remote Sensing*, 2023, 15(15): 3720.
- [17] Wang Jingnan, Wang Xiaodong, Guan Jiping, et al. TAFNet: Time-aware adaptive feature fusion network for very short-term precipitation forecasts[J]. *Geophysical Research Letters*, 2023, 50(15): e2023GL104370.
- [18] 庄潇然, 郑玉, 王亚强, 等. 基于深度学习的融合降水临近预报方法及其在中国东部地区的应用研究 [J]. *气象学报*, 2023, 81(2): 286–303.
- Zhuang Xiaoran, Zheng Yu, Wang Yaqiang, et al. A deep learning-based precipitation nowcast model and its application over East China[J]. *Acta Meteorologica Sinica*, 2023, 81(2): 286–303.
- [19] Qiu Yunan, Lu Zhenyu, Tang Haibo. A short-term regional precipitation prediction model based on wind-improved spatiotemporal convolutional network[J]. *Earth and Space Science*, 2022, 9(9): e2022EA002411.
- [20] Fang Wei, Shen Liang, Sheng V S, et al. A novel method for precipitation nowcasting based on ST-LSTM[J]. *Computers, Materials & Continua*, 2022, 72(3): 4867–4877.
- [21] Xiong Taisong, He Jianxing, Wang Hao, et al. Contextual Sa-attention convolutional LSTM for precipitation nowcasting: a spatiotemporal sequence forecasting view[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 12479–12491.
- [22] Wu Dali, Wu Li, Zhang Tao, et al. Short-term rainfall prediction based on radar echo using an improved self-attention PredRNN deep learning model[J]. *Atmosphere*, 2022, 13(12): 1963.
- [23] Lin Fudong, Yuan Xu, Zhang Yihe, et al. Comprehensive transformer-based model architecture for real-world storm prediction[C]// *European Conference on Machine Learning and Knowledge Discovery in Databases: Applied Data Science and Demo Track*. Turin: Springer, 2023: 54–71.
- [24] Küçük Ç, Giannakos A, Schneider S, et al. Transformer-based nowcasting of radar composites from satellite images for severe weather[J]. *Artificial Intelligence for the Earth Systems*, 2024, 3(2): e230041.
- [25] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[C]// *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018: 3–19.
- [26] Kim B, Han M, Shim H, et al. A performance comparison of convolutional neural network-based image denoising methods: the effect of loss functions on low-dose CT images[J]. *Medical Physics*, 2019, 46(9): 3906–3923.
- [27] Veillette M S, Samsi S, Mattioli C J. SEVIR: a storm event imagery dataset for deep learning applications in radar and satellite meteorology[C]// *Proceedings of the 34th International Conference on Neural Information Processing Systems*. Vancouver: Curran Associates Inc., 2020: 22009–22019.
- [28] Barakat A, Bianchi P. Convergence and dynamical behavior of the ADAM algorithm for nonconvex stochastic optimization[J]. *SIAM Journal on Optimization*, 2021, 31(1): 244–274.
- [29] Le Guen V, Thome N. Disentangling physical dynamics from unknown factors for unsupervised video prediction[C]// *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 11471–11481.
- [30] Yang Nan, Li Xiaofeng. Lightweight AI-powered precipitation nowcasting[J]. *The Innovation Geoscience*, 2024, 2(2): 100066.

Multiscale attentional coding-dynamic decoding network for short-range precipitation nowcasting

Du Xianjun¹, Guo Hangfei¹, Cheng Shengyi¹

(1. College of Electrical and Information Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: Short-term precipitation nowcasting is a critical task in both meteorology and hydrology. However, current deep learning methods often yield ambiguous prediction results and exhibit significant cumulative errors. To address the limitations associated with these predictive methods, particularly the challenges of cumulative error and lack of clarity in prediction sequences, we propose a novel approach based on a Multi-scale Attention Encoding-Dynamic Decoding Network (MAEDDN) for short-term precipitation nowcasting. This method leverages the learning of spatiotemporal features from input data to accurately predict future precipitation scenarios. To obtain richer feature information from the input sequences, the encoding process employs convolutional blocks with spatial and

channel attention for encoding. And a multi-scale fusion module is introduced to address the challenge of capturing both small-scale and large-scale information in precipitation distribution simultaneously. To enhance the clarity of the predicted sequences, the model needs to better understand the precipitation process. Therefore, in the decoding process, a dynamic decoding network is proposed in response to the generation and dissipation processes accompanying short-term precipitation. This network flexibly filters the decoding process by learning the intensity distribution and change trends of past input data. Experiments are conducted by using the precipitation data from the open-source SEVIR dataset, and comparisons are made with the best methods reported so far. The experimental results reveal that: (1) MAEDDN enhances the forecasting capability in areas with high-intensity precipitation, and (2) The clarity of the predicted image sequences by MAEDDN is significantly better than that of other models. The constructed multi-scale attention encoding captures the complex relationships in meteorological data more effectively, while the dynamic decoding adapts the decoding process based on different scenarios, resulting in more accurate prediction outcomes.

Key words: precipitation nowcasting; dynamic decoding network; spatiotemporal features; Multi-scale fusion