

# 基于 Halo 注意力机制的双阶段临近降水预报网络<sup>①</sup>



周云龙<sup>1</sup>, 季繁繁<sup>2</sup>, 潘泽锋<sup>1</sup>

<sup>1</sup>(南京信息工程大学 计算机学院, 南京 210044)

<sup>2</sup>(南京信息工程大学 电子与信息工程学院, 南京 210044)

通信作者: 周云龙, E-mail: 20211249481@nuist.edu.cn

**摘要:** 先前基于深度学习进行临近降水预报的方法试图在统一架构中建模雷达回波的时空演变, 然而, 这些方法可能难以完全捕捉到这种复杂的时空关系. 本文提出了一种基于 Halo 注意力机制的双阶段临近降水预报网络, 该网络将降水预测的时空演变过程分为运动趋势预测和空间外观重建两个阶段. 首先, 可学习光流模块对雷达回波的运动趋势进行建模并生成粗略的预测结果. 其次, 特征重建模块对历史雷达回波序列的空间外观变化建模并对粗粒度预测结果的空间外观进行特征细化重建, 生成精细的雷达回波图. 通过在 CIKM 数据集上的实验表明, 本文所提出的方法与主流方法相比, 平均的海德克技能得分和关键成功指数分别提高了 4.60% 和 3.63%, 达到了 0.48 和 0.45; 结构相似性提高了 4.84%, 达 0.52; 均方误差降低了 6.13%, 达 70.23.

**关键词:** 深度学习; 临近降水预报; 光流; 注意力机制; 双阶段预测

引用格式: 周云龙, 季繁繁, 潘泽锋. 基于 Halo 注意力机制的双阶段临近降水预报网络. 计算机系统应用, 2024, 33(5): 67-75. <http://www.c-s-a.org.cn/1003-3254/9502.html>

## Two-stage Precipitation Nowcasting Network Based on Halo Attention Mechanism

ZHOU Yun-Long<sup>1</sup>, JI Fan-Fan<sup>2</sup>, PAN Ze-Feng<sup>1</sup>

<sup>1</sup>(School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China)

<sup>2</sup>(School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China)

**Abstract:** The previous methods for precipitation nowcasting based on deep learning try to model the spatiotemporal evolution of radar echoes in a unified architecture. However, these methods may face difficulty in capturing the complex spatiotemporal relationships completely. This study proposes a two-stage precipitation nowcasting network based on the Halo attention mechanism. This network divides the spatiotemporal evolution process of precipitation nowcasting into two stages: motion trend prediction and spatial appearance reconstruction. Firstly, a learnable optical flow module models the motion trend of radar echoes and generates coarse prediction results. Secondly, a feature reconstruction module models the spatial appearance changes in the historical radar echo sequences and refines the spatial appearance of the coarse-grained prediction results, generating fine-grained radar echo maps. The experimental results on the CIKM dataset demonstrate that the proposed method outperforms mainstream methods. The average Heidke skill score and critical success index are improved by 4.60% and 3.63%, reaching 0.48 and 0.45, respectively. The structural similarity index is improved by 4.84%, reaching 0.52, and the mean squared error is reduced by 6.13%, reaching 70.23.

**Key words:** deep learning; precipitation nowcasting; optical flow; attention mechanism; two-stage prediction

① 基金项目: 科技创新 2030—“新一代人工智能”重大项目 (2018AAA0100400); 国家自然科学基金 (U21B2049, 61936005)

收稿时间: 2023-11-28; 修改时间: 2023-12-29; 采用时间: 2024-01-12; csa 在线出版时间: 2024-03-22

CNKI 网络首发时间: 2024-03-27

临近降水预报是指某一地区在较短的时间内(一般来说是0-6 h)的降雨预报。准确的临近降水预报不仅可以为人们的日常生活提供便利,还可以减少降雨造成的损失。目前降水预报常用的业务系统是数值天气预报和光流法<sup>[1,2]</sup>,但由于数值天气预报基于流体力学和热力学方程,计算复杂度高,其精度严重依赖于初始条件。而光流法需要假设雷达回波的亮度是恒定的。实际上,雷达回波的亮度会因降水强度变化、衰减等各种因素而发生变化。因此在预测降水事件的起始和衰减方面仍然表现不佳。为了提高精度,深度学习已成为降水临近预报的重要发展方向。基于深度学习的临近降水预报一般通过雷达回波外推实现。首先利用过去雷达回波数据预测未来雷达回波数据,然后利用Z-R关系( $Z = 10 \log a + 10 b \log R$ )将预测雷达回波数据转化为降水强度数据,实现降水强度临近预报,其中R表示降雨量mm/h,Z表示雷达回波反射率dBZ,a和b是基于特定区域统计数据计算的两个常数。

目前,深度学习临近降水预报网络中主要使用卷积神经网络(convolutional neural network, CNN)和循环神经网络(recurrent neural network, RNN)。这些方法以端到端方式训练神经网络模型,通过学习大量的复合雷达观测数据的统计规律进行预测<sup>[3]</sup>。然而,雷达回波的演变涉及整体的平移变化和局部的空间形变<sup>[4]</sup>,如图1所示,橙色和红色区域明显的向右下角方向移动,同时局部区域也在时刻发生着变化,如,黑色框内红色雷达回波区域范围变得更大且强度增强。以前的方法尝试在一个统一的框架中对这两种变化进行建模,但是由于大气环境的多样性,雷达回波的变形、积累或消散非常复杂,一个统一的框架难以同时处理时间和空间上的复杂变化。

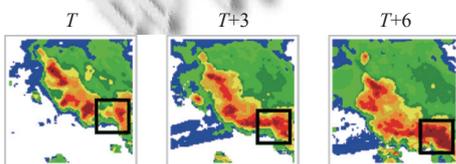


图1 雷达回波序列可视化结果

基于RNN的方法,如ConvLSTM等虽然在LSTM中集成了卷积操作来建模时空变化,但是在反向传播过程中容易出现梯度值逐渐减小并趋近于零或梯度值不断增大造成梯度消失或梯度爆炸的问题。除此之外,这些方法通常是一阶马尔可夫模型<sup>[5]</sup>,对于长序列的处

理能力相对较弱,容易出现短期记忆问题,即过早丢弃先前的信息,难以捕捉到长期的上下文信息。而基于CNN的方法在处理序列数据时对于时间关系的建模能力有一定的局限性。传统的卷积操作主要用于处理空间信息,而在处理时间序列数据时,卷积操作无法直接捕捉到时间上的依赖和顺序关系。

为了更好地处理复杂的雷达回波演化过程,本文提出了一种基于Halo注意力机制<sup>[6]</sup>的双阶段临近降水预报网络,该网络采用Halo注意力机制作为基本构建单元构建了可学习光流模块和特征重建模块,旨在分别对雷达回波的运动趋势和空间形变进行建模,提高深度学习方法的时空建模能力,提供准确且具有细粒度特征的预测结果。该网络将预测雷达回波演变分为两个阶段。第1阶段利用可学习光流模块学习雷达回波序列中降水云体的运动模式,提供对降水的位置和移动趋势的粗略预测。第2阶段利用特征重建模块对可学习光流模块的粗略预测结果进行特征细化重建。总体而言,本文提出的方法为临近降水方向提供了一种新的研究思路,具体的创新点总结如下。

(1) 将降水的演变过程分为两个阶段,并分别设计了可学习光流模块和特征重建模块,这种分阶段的设计可以充分利用时间和空间信息,提高降水预测的准确性和稳定性,并为未来的研究提供了新的思路和方法。

(2) 使用Halo注意力机制作为所提出网络的基本构建单元,使得网络能够自适应地分配注意力权重,更好地捕捉到雷达回波序列中的时间和空间相关性。

## 1 相关工作

### 1.1 光流法

光流法是计算机视觉领域中常用的方法,用于估计图像序列中像素的运动,其基本原理是通过比较相邻帧之间的像素亮度差异来计算光流场。光流场表示了每个像素在两帧之间的位移情况,即其运动方向和速度。通过对整个图像应用光流法,可以得到每个像素的光流向量场,从而揭示了整个图像中物体的运动情况。在气象领域,光流法<sup>[7-10]</sup>逐渐被应用于雷达回波的跟踪,并取得了良好的实际效果。例如,香港天文台开发了多尺度光流变分法,它结合了光流法和变分法的思想,并在不同尺度上进行计算,用于雷暴单体预报和定量降水预报。此外,香港天文台还进一步提出了实时雷达回波光流变分法,该方法采用了空间平滑法并且

允许调整光流变分参数<sup>[9]</sup>。另外, Ayzel 等人<sup>[10]</sup>开发了一种名为 rainymotion 的临近降水模型, 该模型基于多种光流算法计算降水的运动信息, 进而进行降水的预测。

## 1.2 深度学习方法

2015年, Shi 等人<sup>[11]</sup>将临近降水预报作为一个时空序列预测问题, 并提出了 ConvLSTM 网络, 该网络用卷积结构取代了 FC-LSTM 中的全连接结构进行预测, 这为临近降水预报开辟了新的方向。Wang 等人<sup>[12,13]</sup>提出了一种称为 PredRNN 的时空预测方法, PredRNN 额外增加了时空记忆单元, 并通过“之”字形结构将它们连接起来, 使特征在空间和时间上传播来帮助建模时空信息。Wang 等人<sup>[14]</sup>提出了一种名为 MIM 的模型, 该模型创新性地利用相邻循环状态之间的差分信号来模拟时空动力学的稳态和非稳态特征, 并在多个时空序列数据集上取得了最佳的预测结果。Niu 等人<sup>[15]</sup>提出了一种双阶段的时空细化网络 2S-STRef, 该模型第 1 个阶段生成粗粒度的预测结果, 第 2 个阶段对第 1 个阶段生成的预测结果进行细化。Ravuri 等人<sup>[3]</sup>提出了一种深度生成模型, 以消除较长预测时间内临近预报结果中的模糊问题。Wu 等人<sup>[16]</sup>基于物理世界运动可以分解为瞬态变化和运动趋势的原理提出了 MotionRNN 框架来捕捉运动中的复杂变化, 同时通过 MotionGRU 单元对瞬态变化和运动趋势进行建模。

最近一些研究<sup>[17-21]</sup>将 CNN 扩展到降水临近预报中, 其中具有代表性的 CNN 模型是 UNet 及其变体。Zhuang 等人<sup>[17]</sup>提出了 ST-CNN 来充分利用时空信息, 从大量的降水数据中自动学习降水的基本模式, 进行极端降水预测。Ayzel 等人<sup>[18]</sup>提出了 RainNet, 其采用 UNet 架构从大量的雷达数据中学习时空演变来提供精确的临近降水预测。Trebing 等人<sup>[19]</sup>通过为 UNet 配备卷积块注意力模块和深度可分离卷积提出了 SmaAt-UNet。Yang 等人<sup>[20]</sup>结合了 Transformer 和 UNet 模型提出了 AA-TransUNet, 并通过结合卷积块注意力模块和深度可分离卷积在降低模型参数的同时提升模型性能。

## 2 双阶段临近降水预报网络

### 2.1 网络架构

本文采用端到端的学习方法来生成预测雷达回波序列。如图 2 所示, 所提出的网络架构由可学习光流模块和特征重建模块组成, 这两个模块都是神经网络, 可以实现非线性建模。为了更好地捕捉到降水事件

的空间相关性和时序特征, 这两个模块采用 Halo 注意力机制作为基本的构建单元构成。此外, 为了更好地提取和保留图像特征, 两个网络都采用了“U”型结构并通过跳跃连接来拼接低级和高级雷达回波特征。

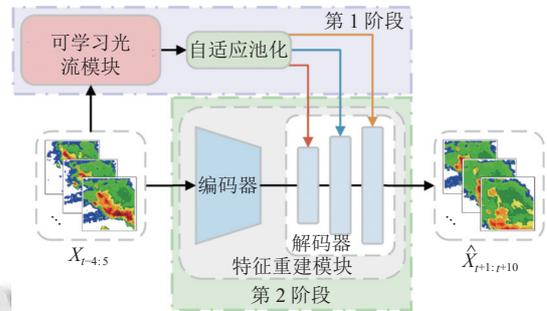


图2 双阶段临近降水预报网络架构图

每帧雷达回波图像可以表示为  $X_t \in R^{H \times W \times C}$ , 其中  $H$ 、 $W$  和  $C$  分别表示高度、宽度和通道数。第 1 个阶段, 可学习光流模块对 5 帧连续的雷达回波图像  $X_{t-4:t}$  进行时间相关性进行建模生成预测光流场  $\hat{V}_{t+1:t+10}$  并通过 warp 操作生成 10 帧预测的粗粒度雷达回波图像  $\hat{X}_{t+1:t+10}$ 。第 2 个阶段, 特征重建模块将从历史序列  $X_{t-4:t}$  学习到的空间变化模式对粗粒度预测结果  $\hat{X}_{t+1:t+10}$  进行细化重建, 生成具有细粒度特征的预测雷达回波序列  $\hat{X}_{t+1:t+10}$ 。这种在一个统一的网络中通过分别建模运动趋势和空间变化的方法, 可以充分地建模时空变化情况。

### 2.2 Halo 注意力机制

随着注意力机制的发展, 在计算机视觉领域, 注意力已是一种常用的算法, 并且已被应用于不同的任务中。注意力机制的公式可以表示为式 (1) 所示:

$$Attention(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

其中,  $Q, K, V \in R^{H \times W \times C}$  分别表示 Query、Key 和 Value,  $k$  为注意力头的数量,  $d_k = C/k$  是注意力头的维度。虽然注意力机制基于内容的交互和更强的捕获图像中远程依赖关系的能力使得重要信息获得了更多的关注。但是, 注意力机制通常需要大量计算, 导致计算复杂性增加。如果采用全局的注意力, 对于较大的图像就会导致计算量非常大。如果采用局部的注意力, 由于每次滑动窗口的结果在显存中得不到释放, 可能就会导致内存不足的问题。Halo 注意力将整张图片分为多个块, 并对每个块进行局部注意力, 从而解决了上述问题。

除此之外, 一个地点的降水受距离它较远地点的降

水的影响很小. 因此, 要预测给定地点的降水情况, 只需要关注附近区域的降水情况, 可以忽略那些太远地区的降水情况. 而 Halo 注意力恰恰遵循了局部处理的原则, 不仅可以提升降雨预测的效果, 而且解决了在高分辨率下注意力机制的计算成本过大的问题. 如图 3 所示, Halo 注意力机制主要分为 3 个步骤: 首先将整个图片分为多个块, 其次, 如果每次操作时只考虑每个块包含的信息, 则不可避免地会导致信息丢失. 因此, 在计算每个块之前, 在每个块之外, 利用原图像的信息进行填充 (简称 Halo), 使得每个块的感受野可以适当放大, 可以关注到更多的信息. 最后, 对每个块分别执行局部自注意力.

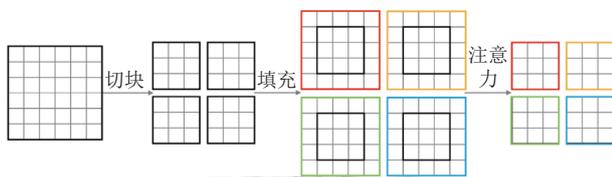


图 3 Halo 注意力机制的过程

另外, 传统的注意力模型在进行下采样的时候通常使用平均池化或最大池化. Halo 注意力机制的下采样层可以对每个块分别进行采样, 然后对这些采样后的信息进行 Halo 注意力操作, 从而达到下采样的效果. 因为这里是对每个块的信息进行采样后计算, 而不是对每个块内的信息进行计算, 所以这样的操作是不会影响模型精度的, 同时还将在下采样层中的 FLOPs 降低了 4 倍.

### 2.3 可学习光流模块

雷达回波图序列的演化是一个复杂的过程, 由于复

杂的大气环境, 雷达回波不仅会发生平移运动, 雷达回波的形状也可能会迅速膨胀、收缩或变化, 这与其他时空序列预测任务有极大的不同. 因此, 本文将降水预测任务分解为两个阶段, 并设计了可学习的光流模块和特征重建模块来处理这两个阶段的不同任务. 可学习光流模块旨在通过学习雷达回波序列中降水云体的运动模式, 粗略推测未来一段时间内降水云体的位置情况<sup>[22]</sup>. 与以往的基于 CNNs 的光流生成网络相比, 本文提出的可学习光流模块引入了 Halo 注意力机制, 这使得网络能够更好地关注重要的特征或动态地调整特征的权重, 以适应不同的运动场景. 这种动态调整特征权重的能力对于处理复杂的大气环境下的雷达回波图序列非常关键, 因为雷达回波可能会发生各种形状和运动模式的变化. 此外, 通过使用 Halo 注意力机制, 光流模块能够自适应地学习到不同区域和运动模式之间的相关性, 并将更多的注意力放在具有更显著运动的区域上.

如图 4 所示, 可学习光流模块由上侧编码器和下侧解码器构成. 编码器和解码器均有 5 层, 蓝色的 Halo 注意力机制不使用下采样层, 其输出的特征图尺寸不会改变, 而绿色的 Halo 注意力机制使用下采样层. 在本文中, 使用下采样层后, 其输出的特征图尺寸会缩小为原尺寸的 1/4. 对于编码器, 第 1 层将连续的 5 帧雷达回波图像  $X_{t-4:t}$  的通道维度扩充至 64 通道. 其余 4 层通过 Halo 注意力机制的下采样层减小空间维度来逐步提取更抽象的特征. 除编码器第 5 层不进行通道维度扩充外, 中间 3 层每一层都将通道维度翻倍.

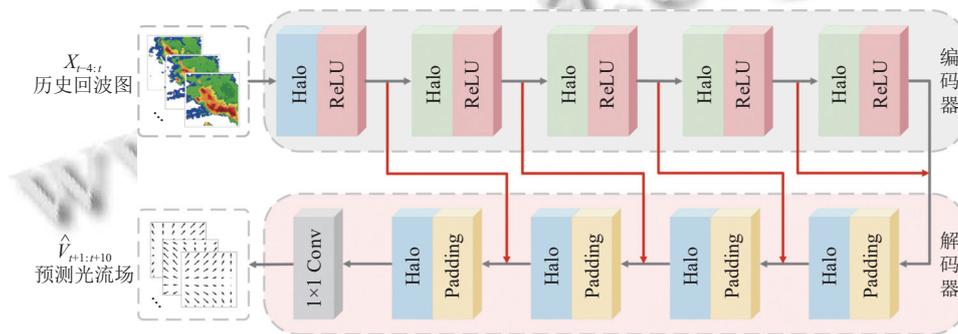


图 4 可学习光流模块架构

对于解码器, 除第 1 层不进行通道维度压缩, 其余各层均将通道维度压缩至输入维度的一半. 编码器和解码器通过跳跃连接将特征在通道维度上进行拼接, 如图 4 中红色箭头所示. 为了更好地建模运动趋势并改善光流估计的性能, 编码器和解码器之间拼接的是

不同尺度的特征信息, 由于拼接的特征图尺寸不一致, 因此在拼接前需要对尺寸较小的特征图进行边缘填充. 此外, 解码器最后一层使用  $1 \times 1$  卷积将特征维度压缩至 20, 从而最终生成连续的 10 帧未来雷达回波图像的光流场  $\hat{V}_{t+1:t+10}$ . 所生成的每个光流场具有两个通道,

表示了图像中每个像素点的运动方向和速度. 可学习光流模块每一阶段的具体特征尺寸如表 1 所示. 最后, 通过式 (2) 对输入雷达回波序列的最后一帧  $X_t$  进行 *warp* 操作得到下一时刻的雷达回波图  $\tilde{X}_{t+1}$ , *warp* 操作可以将雷达图像根据光流场信息进行变换, 重复此过程直至生成预测的未来 10 帧连续的雷达回波图像  $\tilde{X}_{t+1:t+10}$ .

$$\tilde{X}_{t+1} = \text{warp}(X_t, \hat{V}_{t+1}) \quad (2)$$

表 1 可学习光流模块各层输出尺寸

层级	长度	宽度	通道维度
编码器第1层	128	128	64
编码器第2层	64	64	128
编码器第3层	32	32	256
编码器第4层	16	16	512
编码器第5层	8	8	512
解码器第1层	16	16	256
解码器第2层	32	32	128
解码器第3层	64	64	64
解码器第4层	128	128	64
解码器第5层	128	128	20

### 2.4 特征重建模块

特征重建模块利用深度学习强大的表征学习能力学习雷达回波复杂的空间变化模式, 并基于可学习光流模块提供的粗粒度特征信息进行空间特征重建. 如图 5 所示, 本文提出的特征重建模块由上侧编码器和下侧解码器构成. 编码器的输入为在通道维度上拼接的 5 帧连续的历史雷达回波图像  $X_{t-4:t}$  和可学习光流模块生成的 10 帧粗粒度预测结果  $\tilde{X}_{t+1:t+10}$ . 历史雷达回波图像中包含了雷达回波的过去状态和空间外观变化. 而可学习光流模块的粗粒度预测结果提供了雷达回波

在时间上的运动趋势. 特征重建模块编码器通过结合运动趋势信息可以更好地理解和建模雷达回波的空间形态变化. 此外, 编码器通过逐渐抽象的过程使得编码器能够捕捉到不同尺度上的特征, 从而更好地描述雷达回波的演变过程, 包括膨胀、收缩和形状变化等. 解码器的主要任务是将编码器学习到的空间特征变化模式应用到多尺度的粗粒度预测序列上, 并通过非线性上采样逐步恢复特征的空间维度. 整个过程采用了循序渐进的方式, 通过逐渐对不同尺度的粗粒度特征进行空间细节重建, 并将这些不同尺度的信息融合在一起, 可以得到具有细粒度特征的预测序列  $\hat{X}_{t+1:t+10}$ .

与可学习光流模块结构图中的颜色意义相同, 蓝色的 Halo 注意力机制不使用下采样层, 其输出的特征图尺寸不会更改, 而绿色的 Halo 注意力机制使用下采样层. 编码器第 1 层保持图像尺寸不变, 通道维度扩充至 64, 随后 3 层均会将输入序列进行下采样, 图像尺寸缩减为缩小为原尺寸的 1/4, 通道维度增加一倍. 最后一层保持图像尺寸和通道维度不变. 在非线性和上采样前, 可学习光流模块的预测结果经过自适应池化化为解码器不同阶段特征的尺寸, 这种多尺度的预测结果在通道维度与编码器的特征拼接, 可以提供多尺度的雷达回波特征信息用于特征重建. 最后一层编码器的输出特征图经  $1 \times 1$  卷积操作进行维度压缩后得到最终的预测雷达回波序列. 特征重建模块每一阶段的具体特征尺寸如表 2 所示. 总体而言, 特征重建模块通过学习历史雷达回波序列的空间变化模式并对多尺度的雷达回波特征信息进行特征重建来生成具有细粒度细节的预测雷达回波图.

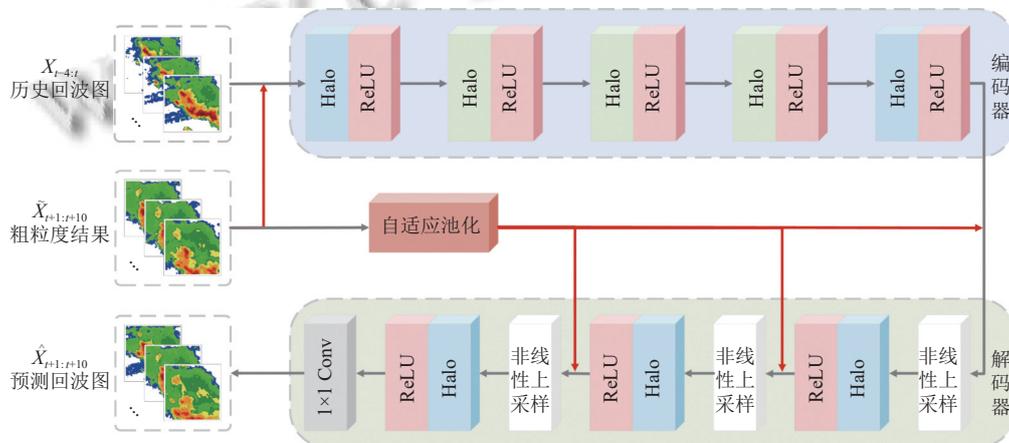


图 5 特征重建模块架构

表2 特征重建模块各层输出尺寸

层级	长度	宽度	通道维度
编码器第1层	128	128	64
编码器第2层	64	64	128
编码器第3层	32	32	256
编码器第4层	16	16	512
编码器第5层	16	16	512
解码器第1层	32	32	256
解码器第2层	32	32	128
解码器第3层	32	32	256
解码器第4层	64	64	256
解码器第5层	128	128	128
解码器第6层	128	128	64
解码器第7层	128	128	10

### 3 实验分析

#### 3.1 数据集

本文利用信息与知识管理国际会议 (CIKM AnalytiCup 2017) 挑战赛数据集来评估所提出的模型. 它包含 14 000 个序列, 其中 8 000 个序列作为训练集, 2 000 个序列作为验证集, 4 000 个序列作为测试集. 训练集和验证集中的数据是连续两年的, 而测试集中的数据是在下一年的<sup>[23]</sup>. 每个序列包含 15 帧雷达回波图像, 间隔 6 min, 仰角分别为 0.5 km、1.5 km、2.5 km 和 3.5 km. 在本文中, 选择 3.5 km 仰角的雷达回波图像来训练和测试所提出的模型. 在每个序列中, 前 5 帧雷达回波图被作为模型输入, 后 10 帧作为预测输出. 每张图像分辨率为 101×101. 每个像素代表 1 km×1 km 的面积, 所有像素的取值范围为 [0-255].

#### 3.2 评估标准

作为预处理步骤, 式 (3) 将每个像素的像素值转换为雷达反射率:

$$Z = \frac{\text{像素值}}{255} \times 90 - 10 \quad (3)$$

在评估指标方面, 采用阈值法将预测回波图与观察值进行转换. 如果该值大于给定阈值, 则将相应值设置为 1; 否则设置为 0. 本文计算了真阳性预测  $TP$  (预测=1, 真值=1)、假阳性预测  $FP$  (预测=1, 真值=0)、真阴性预测  $TN$  (预测=0, 真值=0) 和假阴性预测  $FN$  (预测=0, 真值=1) 的数量. 此外, 本文使用了 5 dBZ、20 dBZ 和 40 dBZ 这 3 个阈值, 并用式 (4) 和式 (5) 计算了关键成功指数 (critical success index,  $CSI$ ) 和海德克技能得分 (Heidke skill score,  $HSS$ ) 来评估结果:

$$CSI = \frac{TP}{TP + FN + FP} \quad (4)$$

$$HSS = \frac{2(TP \times TN - FN \times FP)}{(TP + FN)(FN + TN) + (TP + FP)(FP + TN)} \quad (5)$$

这两个指标广泛应用于降水预报的准确性测评中.  $CSI$  指标关注模型预测的降水事件与真实观测到的降水事件之间的一致性, 即成功预测的降水事件的比例, 其取值范围从 0 到 1, 值越接近 1 表示预测结果越准确.  $HSS$  评估的是排除随机预测后的正确预测比例, 它衡量模型的预测能力相对于随机猜测的改进程度, 其取值范围从 -1 到 1, 值越接近 1 表示预报结果越准确. 除此之外, 结构相似性 (structural similarity index,  $SSIM$ ) 和均方误差 (mean squared error,  $MSE$ ) 也作为本文的评估指标.  $SSIM$  考虑了图像的亮度、对比度和结构等方面的信息, 是基于人眼感知的图像质量评价方法之一, 可以用来评估图像的质量和清晰度.  $MSE$  用于衡量两幅图像之间差异的指标. 它计算了两幅图像对应像素之间的平均平方差, 从而反映了它们在像素级别上的差异程度.

#### 3.3 实验参数设置

所有实验均在 PyTorch 框架下实现, 并且在 NVIDIA GeForce RTX 3090 上运行. 所提出的网络使用 Adam 优化器在训练过程中进行优化, 学习率设置为 0.000 3, 批次大小设置为 16. 此外, 损失函数使用的是平均绝对误差 (mean absolute error,  $MAE$ ). 为防止模型过拟合, 该网络在训练过程中采用了早停机制. 原始输入  $X_i$  维度为 (101, 101, 1), 为方便数据操作, 本文使用 0 填充图像四周, 将输入数据维度改为 (128, 128, 1) 后输入到模型中进行预测. 预测结果去掉四周填充的部分后, 变为初始输入维度 (101, 101, 1).

#### 3.4 实验结果

为了验证所提出模型的优越性, 本文与 6 种主流模型进行了比较, 包括 ConvLSTM<sup>[11]</sup>、ConvGRU<sup>[24]</sup>、PredRNN<sup>[12]</sup>、PredRNN++<sup>[25]</sup>、MIM<sup>[14]</sup> 和 PredRNN\_V2<sup>[13]</sup>. 如表 3 所示, 最佳结果以粗体标出, 所提出的模型在所有指标上都表现出了优势. 为了更直观地展示本文网络的提升效果, 本文绘制了  $HSS$  和  $CSI$  指标的折线图, 展示了不同阈值下各模型随预测时间增加的表现. 如图 6 所示, 所提出的网络在所有阈值下均明显领先于先前的模型, 这意味着其具有更高的准确性, 进一步凸显了其卓越性能. 此外, 图 7 还展示了各模型的  $SSIM$

折线图,结果显示所提出的网络获得了最高的得分,这说明其预测结果更加逼真,与观测到的雷达回波图像更为接近.为了验证所提出网络的稳定性,本文还基于生成的4000个雷达回波序列计算了各项指标的方差.

为了更加直观地观察结果,我们将MSE指标的方差缩小了 $10^7$ 倍.如表4所示,可以明显看出本文所提出的方法在各项指标的方差方面表现最佳,这充分展示了所提出模型的稳定性.

表3 不同模型在CIKM数据集上的实验结果

模型	HSS↑				CSI↑				SSIM↑	MSE↓
	5	20	40	Avg	5	20	40	Avg		
ConvLSTM	0.7023	0.4913	0.0810	0.4249	0.7698	0.4227	0.0475	0.4133	0.4912	83.72
ConvGRU	0.6801	0.4847	0.0990	0.4213	0.7572	0.4072	0.0593	0.4079	0.4512	80.48
PredRNN	0.7069	0.4961	0.1307	0.4446	0.7734	0.4202	0.0793	0.4243	0.4996	84.85
PredRNN++	0.6905	0.5054	0.1347	0.4435	0.7497	0.4226	0.0850	0.4191	0.4946	77.76
MIM	0.7046	0.5197	0.1517	0.4587	0.7671	0.4411	0.0972	0.4351	0.4950	74.54
PredRNN_V2	0.7030	0.4823	0.1012	0.4288	0.7711	0.4111	0.0637	0.4153	0.4780	76.48
本文算法	<b>0.7277</b>	<b>0.5374</b>	<b>0.1742</b>	<b>0.4798</b>	<b>0.7922</b>	<b>0.4509</b>	<b>0.1096</b>	<b>0.4509</b>	<b>0.5238</b>	<b>70.23</b>

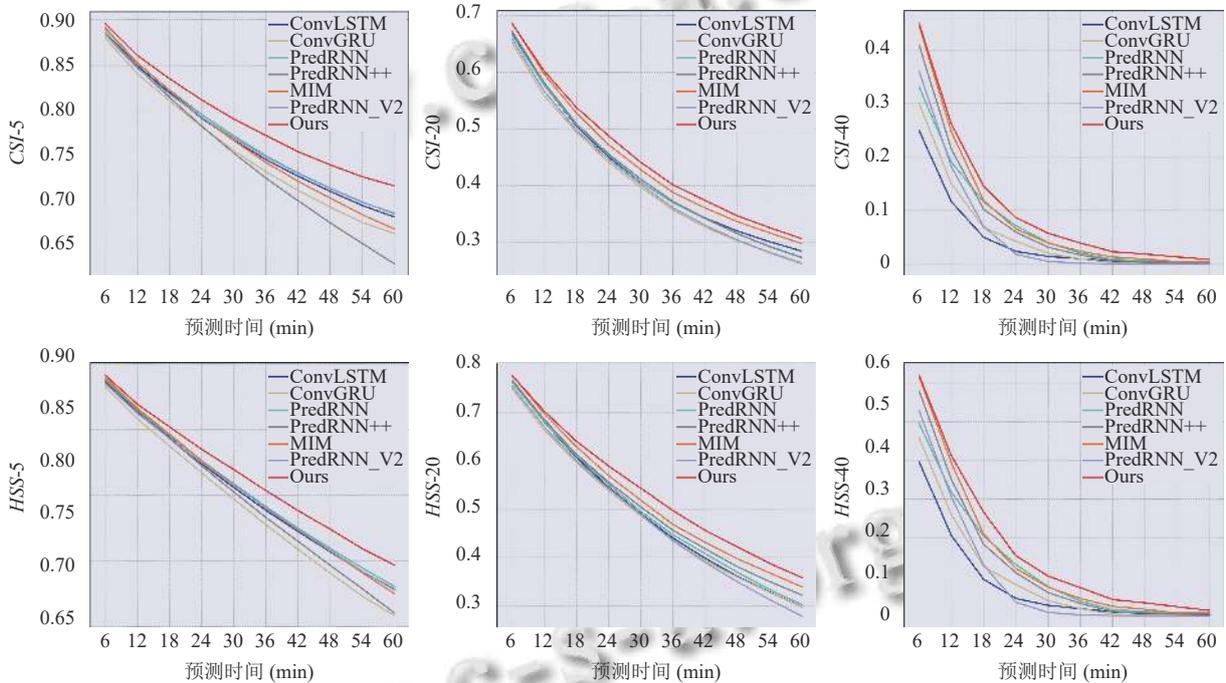


图6 不同模型在CIKM数据集上的CSI和HSS结果

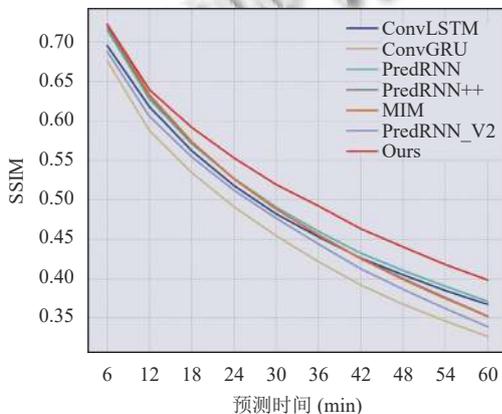


图7 不同模型在CIKM数据集上的SSIM结果

表4 在CIKM数据集上不同指标的方差

模型	HSS	CSI	SSIM	MSE
ConvLSTM	0.0113	0.0123	0.0193	1.49
ConvGRU	0.0123	0.0125	0.0226	1.19
PredRNN	0.0121	0.0119	0.0189	1.59
PredRNN++	0.0129	0.0135	0.0183	1.18
MIM	0.0128	0.0124	0.0197	1.11
PredRNN_V2	0.0110	0.0101	0.0205	1.44
本文算法	<b>0.0104</b>	<b>0.0098</b>	<b>0.0151</b>	<b>1.07</b>

为了进一步展示所提出模型的实际预测效果,本文随机选择了一个雷达回波序列,并绘制了所提出模型以及其他主流模型的预测效果图,如图8所示.通过对比可以明显看出,与先前的模型相比,所提出的模型

在预测雷达回波的位置和强度方面表现更为准确. 特别是在黄色区域, 先前的模型错误地预测了其位置, 甚至对雷达回波的强度进行了高估. 而所提出的网络则更加接近观测结果, 其预测结果在雷达回波的形状和强度方面更加精确. 综上所述, 通过对比各项指标的折线图、方差结果以及预测效果图的展示, 可以充分证明所提出模型相对于其他主流模型在准确性和可靠性方面的优越性, 为相关领域的研究和应用提供了有力的支持.

### 3.5 消融实验

为了评估所使用的 Halo 注意力机制在网络中的有效性以及本文所提出的双阶段预测网络的实际提升效果, 本文进行了消融实验, 分别对 Halo 注意力机制、单阶段模型进行了独立验证. 首先, 本文移除可学习光流模块, 用 3×3 卷积替代特征重建模块中的 Halo

注意力机制, 并以此作为基准 (Baseline) 来验证 Halo 注意力机制的效果. 其次, 本文使用应用了 Halo 注意力机制的特征重建模块作为单阶段预测模型. 根据表 5 的结果显示, 仅使用特征重建模块的 Baseline 网络在预测效果上表现最差, 这表明 Halo 注意力机制对于提高网络性能具有重要作用. 单阶段网络的效果明显提升, 进一步验证了 Halo 注意力机制的有效性. 值得注意的是, 当引入可学习光流模块, 构建了双阶段网络时, 预测效果进一步提升. 这说明可学习光流模块能够在时序上建模雷达回波的运动趋势, 其融合到预测过程中增强了模型对降水演变的理解和预测能力. 综上所述, 消融实验的结果验证了 Halo 注意力机制的有效性, 并证明了所提出的双阶段临近降水预报网络能够更好地利用空间和时间信息, 提高了对临近降水的预测能力, 为临近降水预报研究提供了有力的支持.

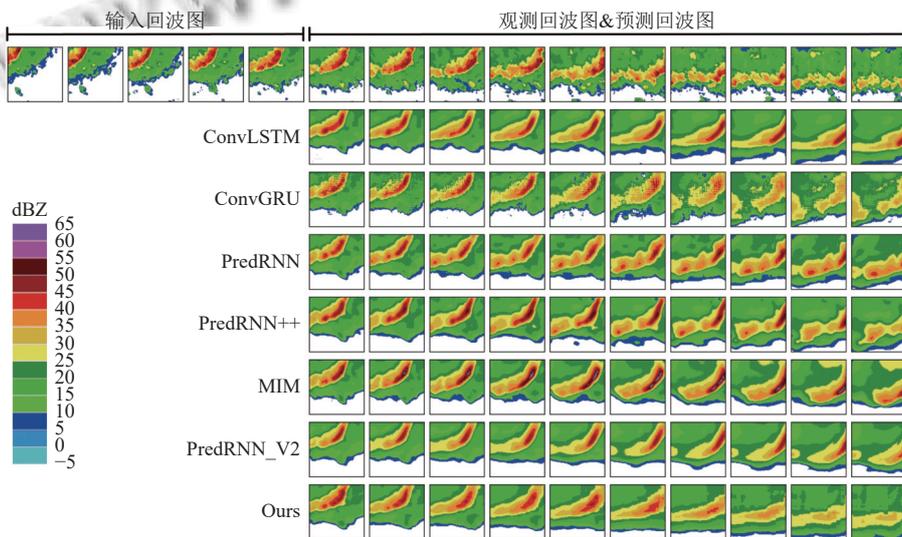


图 8 不同模型在 CIKM 数据集上的一个可视化结果

表 5 在 CIKM 数据集上的消融实验结果

模型	HSS↑				CSI↑				SSIM↑	MSE↓
	5	20	40	Avg	5	20	40	Avg		
Baseline	0.6925	0.5049	0.1387	0.4454	0.7711	0.4334	0.0895	0.4313	0.4860	84.90
单阶段	0.7168	0.5257	0.1451	0.4625	0.7829	0.4374	0.0973	0.4392	0.5116	73.62
双阶段	<b>0.7277</b>	<b>0.5374</b>	<b>0.1562</b>	<b>0.4738</b>	<b>0.7922</b>	<b>0.4509</b>	<b>0.1025</b>	<b>0.4485</b>	<b>0.5238</b>	<b>70.23</b>

## 4 结论与展望

本文提出了一种基于 Halo 注意力机制的双阶段临近降水预报网络, 该网络由可学习光流模块和特征重建模块组成. 这两个模块是由 Halo 注意力机制作为基本构建模块构建的, 并将降水预测的过程分为两个阶段. 在第 1 个阶段, 可学习光流模块通过估计雷达回波序列中的运动趋势生成粗粒度的预测结果. 在第

2 个阶段, 特征重建模块对历史雷达回波序列的空间外观进行建模并基于学习到的空间外观变化模式对粗粒度的预测结果进行特征重建, 从而实现对未来降水的准确预测. 这种双阶段的预测方法能够更好地捕捉雷达回波的时空演变特征, 提高降水预报的准确性和细节展示. 实验结果表明, 所提出的网络在应对复杂的雷达回波演变时具有更高的准确性和清晰度.

## 参考文献

- 1 Woo WC, Wong WK. Operational application of optical flow techniques to radar-based rainfall nowcasting. *Atmosphere*, 2017, 8(3): 48. [doi: [10.3390/atmos8030048](https://doi.org/10.3390/atmos8030048)]
- 2 Li L, He ZW, Chen S, *et al.* Subpixel-based precipitation nowcasting with the pyramid Lucas-Kanade optical flow technique. *Atmosphere*, 2018, 9(7): 260. [doi: [10.3390/atmos9070260](https://doi.org/10.3390/atmos9070260)]
- 3 Ravuri S, Lenc K, Willson M, *et al.* Skilful precipitation nowcasting using deep generative models of radar. *Nature*, 2021, 597(7878): 672–677. [doi: [10.1038/s41586-021-03854-z](https://doi.org/10.1038/s41586-021-03854-z)]
- 4 Yan BY, Yang C, Chen F, *et al.* FDNet: A deep learning approach with two parallel cross encoding pathways for precipitation nowcasting. *Journal of Computer Science and Technology*, 2023, 38(5): 1002–1020. [doi: [10.1007/s11390-021-1103-8](https://doi.org/10.1007/s11390-021-1103-8)]
- 5 Su JH, Byeon W, Kossaifi J, *et al.* Convolutional tensor-train LSTM for spatio-temporal learning. *Proceedings of the 34th International Conference on Neural Information Processing Systems*. Vancouver: Curran Associates Inc., 2020. 1150.
- 6 Vaswani A, Ramachandran P, Srinivas A, *et al.* Scaling local self-attention for parameter efficient visual backbones. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021. 12894–12904.
- 7 Li DW, Liu YD, Chen CH. MSDM v1.0: A machine learning model for precipitation nowcasting over eastern China using multisource data. *Geoscientific Model Development*, 2021, 14(6): 4019–4034. [doi: [10.5194/gmd-14-4019-2021](https://doi.org/10.5194/gmd-14-4019-2021)]
- 8 Sakaino H. Spatio-temporal image pattern prediction method based on a physical model with time-varying optical flow. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(5): 3023–3036. [doi: [10.1109/TGRS.2012.2212201](https://doi.org/10.1109/TGRS.2012.2212201)]
- 9 吴奕霄, 李劲, 项瑛, 等. Farneback 光流法在雷达回波追踪中的应用. *气象科学*, 2023, 43(4): 515–522. [doi: [10.12306/2023jms.0039](https://doi.org/10.12306/2023jms.0039)]
- 10 Ayzel G, Heistermann M, Winterrath T. Optical flow models as an open benchmark for radar-based precipitation nowcasting (rainymotion v0.1). *Geoscientific Model Development*, 2019, 12(4): 1387–1402. [doi: [10.5194/gmd-12-1387-2019](https://doi.org/10.5194/gmd-12-1387-2019)]
- 11 Shi XJ, Chen ZR, Wang H, *et al.* Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal: MIT Press, 2015. 802–810.
- 12 Wang YB, Long MS, Wang JM, *et al.* PredRNN: Recurrent neural networks for predictive learning using spatiotemporal LSTMs. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach: Curran Associates Inc., 2017. 879–888.
- 13 Wang YB, Wu HX, Zhang JJ, *et al.* PredRNN: A recurrent neural network for spatiotemporal predictive learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(2): 2208–2225. [doi: [10.1109/TPAMI.2022.3165153](https://doi.org/10.1109/TPAMI.2022.3165153)]
- 14 Wang YB, Zhang JJ, Zhu HY, *et al.* Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics. *Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE, 2019. 9154–9162.
- 15 Niu D, Huang JH, Zang ZL, *et al.* Two-stage spatiotemporal context refinement network for precipitation nowcasting. *Remote Sensing*, 2021, 13(21): 4285. [doi: [10.3390/rs13214285](https://doi.org/10.3390/rs13214285)]
- 16 Wu HX, Yao ZY, Wang JM, *et al.* MotionRNN: A flexible model for video prediction with spacetime-varying motions. *Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021. 15435–15444.
- 17 Zhuang Y, Ding W. Long-lead prediction of extreme precipitation cluster via a spatio-temporal convolutional neural network. *Proceedings of the 6th International Workshop on Climate Informatics*. San Francisco, 2016.
- 18 Ayzel G, Scheffer T, Heistermann M. RainNet v1.0: A convolutional neural network for radar-based precipitation nowcasting. *Geoscientific Model Development*, 2020, 13(6): 2631–2644. [doi: [10.5194/gmd-13-2631-2020](https://doi.org/10.5194/gmd-13-2631-2020)]
- 19 Trebing K, Stańczyk T, Mehrkanoon S. SmaAt-UNet: Precipitation nowcasting using a small attention-UNet architecture. *Pattern Recognition Letters*, 2021, 145: 178–186. [doi: [10.1016/j.patrec.2021.01.036](https://doi.org/10.1016/j.patrec.2021.01.036)]
- 20 Yang YM, Mehrkanoon S. AA-TransUNet: Attention augmented TransUNet for nowcasting tasks. *Proceedings of the 2022 International Joint Conference on Neural Networks*. Padua: IEEE, 2022. 1–8. [doi: [10.1109/IJCNN55064.2022.9892376](https://doi.org/10.1109/IJCNN55064.2022.9892376)]
- 21 Ayzel G, Heistermann M, Sorokin A, *et al.* All convolutional neural networks for radar-based precipitation nowcasting. *Procedia Computer Science*, 2019, 150: 186–192. [doi: [10.1016/j.procs.2019.02.036](https://doi.org/10.1016/j.procs.2019.02.036)]
- 22 Zhang YC, Long MS, Chen KY, *et al.* Skilful nowcasting of extreme precipitation with NowcastNet. *Nature*, 2023, 619(7970): 526–532. [doi: [10.1038/s41586-023-06184-4](https://doi.org/10.1038/s41586-023-06184-4)]
- 23 Luo CY, Li XT, Wen YL, *et al.* A novel LSTM model with interaction dual attention for radar echo extrapolation. *Remote Sensing*, 2021, 13(2): 164. [doi: [10.3390/rs13020164](https://doi.org/10.3390/rs13020164)]
- 24 Shi XJ, Gao ZH, Lausen L, *et al.* Deep learning for precipitation nowcasting: A benchmark and a new model. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach: Curran Associates Inc., 2017. 5622–5632.
- 25 Wang YB, Gao ZF, Long MS, *et al.* PredRNN++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning. *Proceedings of the 35th International Conference on Machine Learning*. Stockholm: PMLR, 2018. 5110–5119.

(校对责编: 孙君艳)