

MDA-MIM: 一种融合多尺度特征与双重注意力机制的雷达回波图预测模型

胡强¹, 高雅婷¹, 尹宾礼², 渠连恩^{1,3}

(1. 青岛科技大学信息科学技术学院, 山东 青岛 266000; 2. 青岛广播电视台, 山东 青岛 266000)

3. 新疆工程学院信息工程学院, 新疆 乌鲁木齐 830023;

摘要: 为提升雷达回波图中时空特征的提取质量, 提出了一种基于多尺度特征融合和双重注意力机制的MIM改进(MDA-MIM)模型。该模型基于空洞卷积实现多尺度特征提取与融合。通过在MIM模型中的非平稳模块集成自注意力机制, 调整不同时间步长和空间位置的权重, 更精确地捕捉雷达回波数据中的非平稳性特征。在平稳模块引入局部注意力机制, 以聚焦于局部区域内的特征关联, 增强对平稳性特征的捕捉能力。真实数据集上的实验结果表明, MDA-MIM具有优秀的预测性能, 在MSE、MAE、SSIM和PSNR等指标上均优于对比模型。

关键词: 雷达回波图; 时空预测; 注意力机制; 多尺度特征

中图分类号: TP391

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2025053

MDA-MIM: a radar echo map prediction model integrating multi-scale feature fusion and dual attention mechanism

HU Qiang¹, GAO Yating¹, YIN Binli², QU Lianen^{1,3}

1. School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266000, China

2. Qingdao Radio and Television Station, Qingdao 266000, China

3. College of Information Engineering, Xinjiang Institute of Engineering, Urumqi 830023, China

Abstract: To obtain high-quality spatiotemporal features from radar echo maps, an improved MIM (memory in memory) model, MDA-MIM (multi-scale feature fusion and dual attention mechanism MIM) was proposed. Multi-scale feature fusion and a dual attention mechanism were incorporated in MDA-MIM. Dilated convolution was used to extract and integrate multi-scale features. To better capture the non-stationary characteristics of radar echo data, a self-attention mechanism was introduced into the non-stationary module of the MIM model, dynamically adjusting the weights of different time steps and spatial positions. Meanwhile, a local attention mechanism was incorporated into the stationary module, enabling the model to focus on feature correlations within local regions and enhance its ability to extract stationary features. Experiments conducted on real-world datasets demonstrate that MDA-MIM achieves state-of-the-art predictive performance, consistently outperforming baseline models across multiple evaluation metrics, including MSE, MAE, SSIM, and PSNR.

Keywords: radar echo map, spatial-temporal prediction, attention mechanism, multi-scale feature

收稿日期: 2025-01-13; 修回日期: 2025-03-07

通信作者: 渠连恩, lianen.qu@qust.edu.cn

基金项目: 国家自然科学基金资助项目(No.61973180); 山东省重点研发计划(软科学)基金资助项目(No.2023RKY01009)

Foundation Items: The National Natural Science Foundation of China (No.61973180), Shandong Province Key Research and Development (Soft Science) Project (No.2023RKY01009)

0 引言

全球气候变化的不断加剧造成极端天气事件频发。准确、及时的降水预报对于保护公共安全、降低灾害损失具有重要意义。雷达回波数据作为降水预报的核心数据源,其动态变化和空间分布特征中蕴含着丰富的降水信息和时空变化特性^[1],可以有效地辅助降水预测,因此,雷达回波图的时序预测问题成为降水预测研究领域中的热点问题^[2]。

雷达回波图中包含的时空特征可以分为非平稳性特征和平稳性特征。非平稳性特征是指统计特性随时间变化显著的特征,在雷达回波图中表现为回波强度的快速变化,通常与快速变化的天气现象有关,如雷暴、龙卷风等。平稳性特征是指在一定时间内统计特性随时间变化较小的特征,在雷达回波图中表现为一定区域内回波强度的持续性,通常与稳定的降水现象相关,如层云状降水、锋面降水。这些特征对雷达回波图预测模型的性能提出了极高的要求。

传统的时空预测模型在处理雷达回波数据时,面临如何同时有效地捕捉非平稳性特征和平稳性特征这2种截然不同的数据的挑战^[3]。非平稳性特征要求模型具备识别数据中长期依赖关系和动态变化趋势的能力;而平稳性特征则需要模型精确地刻画局部空间结构和时间连续性。然而,现有的模型在平衡上述两方面需求时常存在局限性,影响了预测性能。

为了应对上述挑战,本文提出一种基于多尺度特征融合和双重注意力机制的MIM(memory in memory)改进(MDA-MIM, multi-scale feature fusion and dual attention mechanism MIM)模型。通过多尺度特征提取和注意力的引入提升雷达回波图的预测质量,主要贡献如下。

1) 设计了一个基于空洞卷积的多尺度特征提取与融合模块,使模型能够在不增加计算复杂度的情况下提取不同尺度的特征,弥补深层特征中细节信息的丢失问题。

2) 引入双重注意力机制改进了MIM模型,在非平稳特征模块MIM-N中加入自注意力机制,使模型能够自适应地调整不同时间步和空间位置上的权重,更精确地捕捉雷达回波数据中的非平稳性特征;在平稳特征模块MIM-S中加入局部注意力机制,使模型专注于局部区域内的特征关系,提高平

稳性特征的捕捉能力。

3) 实验结果表明,MDA-MIM模型在雷达回波图预测任务中展现出优异的性能,预测精度优于当前流行的预测模型。

1 相关工作

传统的基于雷达回波图的降水预报方法采用光流法^[4],通过计算连续帧之间的光流场捕捉每个像素的运动模式,并预测未来的运动状态。此外,雷达回波图预测方法还包括单体质心法^[5]、交叉相关法^[6]等,但大气运动涉及产生、消散和变化过程,限制了此类方法在雷达回波图降雨预测中的性能。

近年来,研究者将雷达回波图预测转化为时空数据预测问题,构建了一系列模型。例如,Shi等^[7]提出的卷积长短期记忆网络(ConvLSTM, convolutional LSTM)是在长短期记忆网络(LSTM, long short term memory)的基础上引入卷积结构,更好地捕捉空间相关性,但在处理长时间序列时容易出现过拟合现象。Wang等^[8]提出的PredRNN++设计了一个Gradient Highway单元解决PredRNN的梯度消失问题,该模型复杂度较高。Wang等^[9]提出的MIM使用差分运算逐层稳定高阶非平稳序列,以模拟长期运动趋势,取得了不错效果。Zhou等^[10]提出的SAMM-MIM(MIM with self-attention memory module)通过引入自注意力机制,增强对雷达回波图中重要时空信息的识别和记忆能力,但仍存在细节丢失问题。

Cheng等^[11]提出的AFR-LSTM(LSTM with attention fusion and recall)通过整合历史信息并在编码器与解码器之间构建信息回忆模块以保存细节信息。Fang等^[12]提出的STUNNER(spatiotemporal fusion neural network)通过集成时间差分网络和时空轨迹网络,实现了对数据的高效时空特征提取和融合。Zheng等^[13]提出的TISE-LSTM(LSTM with temporal interactions and spatial extract block)设计了时空交互模块和空间特征提取模块,以增强对数据的时间和空间特征的捕捉能力,但增加了模型的复杂性,短期预测性能不佳。

Naz等^[14]提出的ConvLSTM2D结合了时间分布式的二维卷积层和二维ConvLSTM,以及在不同层级上应用多尺度卷积核,从而有效捕获时空数据中的局部和全局特征。Song等^[15]提出的PPNet通

过泰勒冻结假设和特征约束机制,提升对时序信息的敏感度,但没有改变全卷积网络无法主动提取时序信息的缺陷,在长期预测中模型性能会下降。

为了提高雷达回波图的预测质量,一些研究者设计了基于生成对抗网络(GAN, generative adversarial network)的预测模型。例如, Fang 等^[16]提出的 VRNet (vivid radar network) 融合多尺度空间特征提取和 GAN 来增强时空信息表示,但在捕捉强度变化过程即非平稳性特征方面存在不足。基于 Transformer 架构的模型也受到了研究人员的欢迎。例如, Zhao 等^[17]提出的 MDTNet (multiscale deformable Transformer network) 通过结合图像多尺度块分割与融合策略和可变形多头自注意力机制,有效地捕捉不同空间尺度的降水特征,缩小预测和参考图像在傅里叶空间的频谱差异。Tang 等^[18]提出的 CMFPN (cross-modal fusion prediction network) 通过多尺度特征提取和时空对齐模块整合了雷达回波和数值天气预报数据,提高了降水预测的准确性。基于 Transformer 的模型通常需要大量的计算资源和数据进行训练,在小型数据集上的应用效果一般。

此外,一些研究人员使用其他模型进行雷达降水预报。Yan 等^[19]提出的 FNet 通过 2 个并行的交叉编码路径来分别捕捉雷达回波的整体运动趋势和局部形态变化,但短期降水预测结果仍有不足。Yao 等^[20]提出的 DSADNet (dual-source attention dynamic neural network) 利用动态卷积和自注意力机制预测未来降水,但在提取多尺度特征和细节信息方面存在不足,可能导致将弱降水区域预测为强降水区域。

上述工作从不同的模型角度研究了雷达回波图的时序预测问题,但多数工作主要考虑的是预测模型的改进,鲜有工作聚焦在雷达回波图特征提取。综合考虑各类模型的预测效果、复杂度,本文选择 MIM 模型作为预测基准模型,通过引入多尺度特征提取和注意力机制,构建了一个 MDA-MIM 模型,从雷达回波图的特征提取以及平稳性和非平稳性特征处理等多个角度提升预测精确度。

2 问题描述与研究框架

2.1 问题描述

雷达回波图预测将已有的雷达回波图序列作为

输入,预测未来的雷达回波图序列,如图 1 所示。

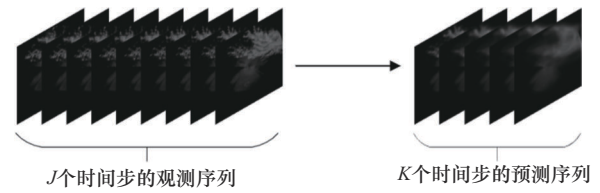


图 1 雷达回波图的预测

雷达回波图预测问题可以形式化描述如下。令雷达回波图序列在空间维度上由 $M \times N$ 个像素点表示,每个像素点都有 Q 个随时间变化的测量值,即每个像素点的特征数量为 Q 。因此,一帧雷达回波图的观测结果可以用张量 $X \in R^{Q \times M \times N}$ 表示,其中 R 表示观测特征的可能取值的范围。

每个时间步对应一帧二维雷达回波图^[21],按时间排列的雷达回波图序列对应于一组张量 X_1, X_2, \dots, X_t 。时空序列预测问题就是在包括当前观测数据在内的 J 个先前观测数据的基础上,预测未来最有可能出现的长度为 K 的序列,计算式为

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

其中, $p(\hat{X}_{t+1}, \dots, \hat{X}_{t+K} | X_{t-J+1}, X_{t-J+2}, \dots, X_t)$ 表示在给定过去 J 个时间步的观测序列(包括当前时间步),即 $X_{t-J+1}, X_{t-J+2}, \dots, X_t$ 的条件下,未来 K 个时间步的观测序列,即 $\hat{X}_{t+1}, \dots, \hat{X}_{t+K}$ 的概率; argmax 用于获得未来 K 个时间步的观测序列概率最大的预测; $\tilde{X}_{t+1}, \dots, \tilde{X}_{t+K}$ 表示预测的未来 K 个时间步的观测序列。

2.2 研究框架

本文提出的 MDA-MIM 模型架构如图 2 所示,模型由多尺度特征提取与融合(MEF, multi-scale feature extraction and fusion)模块、ST-LSTM 单元和 DA-MIM (dual attention-MIM) 单元组成。

模型的输入可以是真实帧或上一个时间步生成的帧,输出则则在每个时间步中生成帧。图 2 中水平方向的箭头表示长期状态和隐藏状态的传递;竖直方向的箭头表示时空记忆状态和隐藏状态的传递;斜向的箭头表示差分信息的传递。

在模型的输入端,MEF 模块首先对输入的雷达回波图像进行多尺度特征提取,通过应用不同空洞率的空洞卷积捕获多尺度特征^[22],这些特征随

后被融合,为后续的时空特征学习提供基础。在时间步 $t \neq 1$ 和网络层数 $l \neq 1$ 时,本文将隐藏状态 H_{t-1}^{l-1} 和 H_t^{l-1} 传递给 DA-MIM 模块以生成差分特征。在第一层中,由于没有前一层的状态,使用 ST-LSTM 生成隐藏状态,其他层的差分信息是通过隐藏状态 H_{t-1}^{l-1} 减去隐藏状态 H_t^{l-1} 得到的。

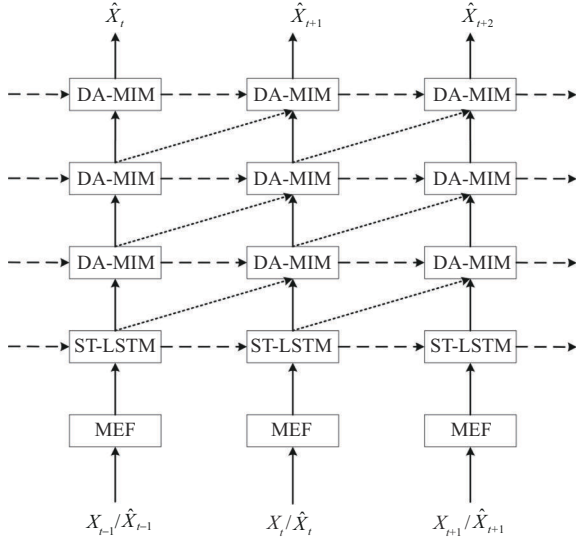


图2 MDA-MIM模型架构

DA-MIM为引入双重注意力机制的改进MIM模型,其中,在MIM-N模块加入自注意力机制,使模型能够自适应地调整不同时间步和空间位置上的权重,更精确地捕捉雷达回波数据中的非平稳性特征;在MIM-S模块中加入局部注意力机制,有助于模型专注于局部区域内的特征关系,提高平稳性特征的捕捉能力。上述特征在经过3层DA-MIM后,将输出下一时间步的预测帧。

3 MDA-MIM模型

3.1 多尺度特征提取与融合模块

在雷达回波图预测任务中,如果直接使用堆叠的采样网络层进行特征提取,则会降低图像分辨率,并逐层丢失重要的空间特征信息。鉴于不同尺度的图像均携带空间特征信息,因此,多尺度特征的提取有助于高质量地获取雷达回波图的空间分布和细节纹理信息。

为了提取不同尺度的特征表示,需要调整卷积核的感受野大小,然而过大的卷积核会降低模型的计算效率。为了克服这一局限,本文设计了一个多尺度特征提取与融合模块,通过引入空洞卷积,在

卷积核中相邻元素之间插入零值来扩大感受野。MEF结构如图3所示,其中,空洞率 $C_1=1, C_2=3, C_3=5$ 。

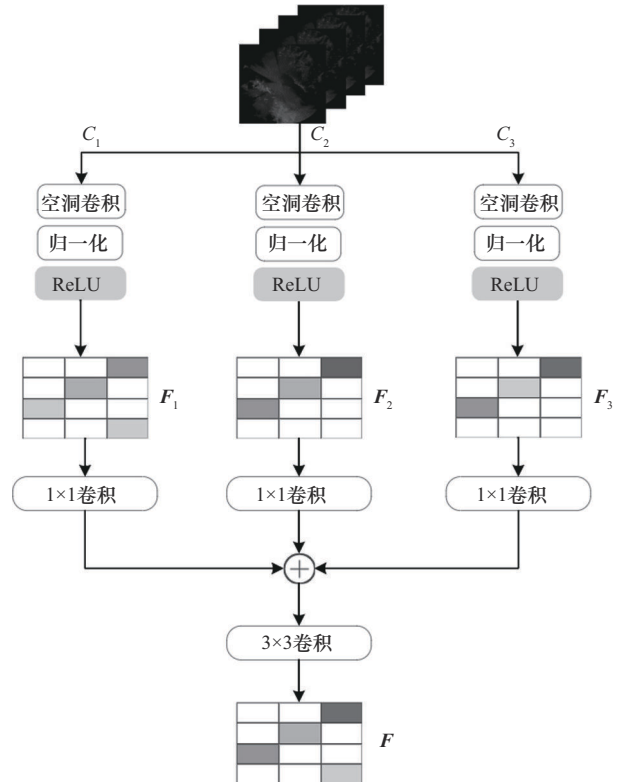


图3 MEF结构

MEF由3个具有不同空洞率的平行空洞卷积层组成。较大的空洞率能够获得较大的感受野,有助于更加全面地获取雷达回波整体轮廓的空间分布;空洞率越小,卷积核的感受野就越小,可以捕捉雷达回波图像的细节纹理特征。

在每个空洞卷积层之后,特征信息会依次经过归一化层和激活函数层,分别得到3个不同尺度感受野的特征图: $F_1 \sim F_3$ 。这些特征图随后被融合,通过 1×1 的卷积操作对3个尺度的特征图进行处理,并沿通道轴将它们合并。为了避免特征信息的冗余,更精确地实现特征融合,对融合后的特征图进行 3×3 的卷积操作。通过利用空洞卷积,能够有效地从雷达回波图中提取多尺度特征,不仅有助于在后续的预测过程中保留更多的特征细节,还能提高预测图像的清晰度和准确性。

3.2 基于自注意力机制的MIM-N模块

MIM-N模块将前一层的隐藏状态 H_{t-1}^{l-1} 作为输入,专注于捕捉基于2个连续隐藏状态之间差值

$(H_t^{l-1} - H_{t-1}^{l-1})$ 的非平稳变化。然而，过度的差分运算可能导致整体信息的丢失，尤其是在多次差分后，原始数据中的重要特征可能会被削弱，影响模型捕捉非平稳性特征的能力。此外，MIM-N 模块主要关注全局的非平稳性特征，可能会忽视序列中的局部特征和细节信息，在面对极端复杂或快速变化的雷达回波时，MIM-N 模块可能难以迅速适应并准确捕捉这些变化，进而影响预测的准确性。

为了克服 MIM-N 模块的这些局限性，本文引入了自注意力 (self-attention) 机制，该机制通过计算序列中每个元素与其他元素之间的注意力分数，捕捉输入序列中元素之间的局部依赖关系，这有助于模型在捕捉全局非平稳性特征的同时更好地保留和利用局部特征信息^[23]。基于自注意力机制的 MIM-N 模块简称 SA-MIMN 模块，其结构如图 4 所示。

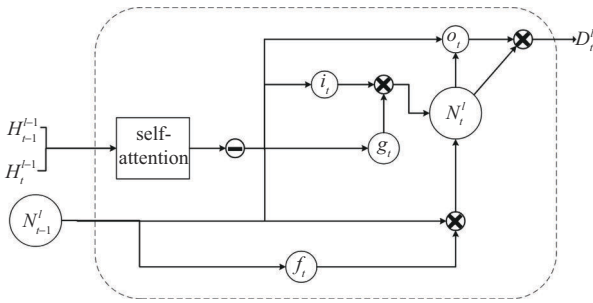


图 4 SA-MIMN 模块结构

隐藏状态 \mathbf{H} 包含了模型学习到的动态变化信息，SA-MIMN 模块中的自注意力机制能够区分 \mathbf{H} 中时空信息的重要性，计算式如下。

$$\tilde{\mathbf{H}} = \text{SA}(\mathbf{H}) \quad (2)$$

$$\mathbf{g}_t = \tanh(W_{xg} * (\tilde{H}_t^{l-1} - \tilde{H}_{t-1}^{l-1}) + W_{ng} * N_{t-1}^l + b_g) \quad (3)$$

$$i_t = \sigma(W_{xi} * (\tilde{H}_t^{l-1} - \tilde{H}_{t-1}^{l-1}) + W_{ni} * N_{t-1}^l + b_i) \quad (4)$$

$$f_t = \sigma(W_{xf} * (\tilde{H}_t^{l-1} - \tilde{H}_{t-1}^{l-1}) + W_{nf} * N_{t-1}^l + b_f) \quad (5)$$

$$N_t^l = f_t \odot N_{t-1}^l + i_t \odot \mathbf{g}_t \quad (6)$$

$$o_t = \sigma(W_{xo} * (\tilde{H}_t^{l-1} - \tilde{H}_{t-1}^{l-1}) + W_{no} * N_{t-1}^l + b_o) \quad (7)$$

$$D_t^l = \text{SA-MIMN}(H_t^{l-1}, H_{t-1}^{l-1}, H_{t-1}^l) = o_t \odot \tanh(N_t^l) \quad (8)$$

其中， \mathbf{N} 表示在非平稳模块中水平移动的存储单元，保存着非平稳性特征的信息； \mathbf{W} 表示权重矩阵； \mathbf{H} 表示隐藏状态，它包含模型学习到的动态变化信息； \mathbf{D} 表示由 SA-MIMN 模块生成的差分特征； $*$ 表示卷积， \odot 表示逐元素乘积 (即 Hadamard

乘积)。

首先，利用式(2)中的 SA，将隐藏状态 \mathbf{H} 进行线性变换，转换为向量 \mathbf{Q} 、 \mathbf{K} 和 \mathbf{V} ，计算每个元素与其他元素之间的相似度得分，并通过 softmax 函数得到权重向量。然后，将权重向量与值向量相乘，得到加权求和的结果。最后，将自注意力机制计算得到的特征与原始特征进行融合，通过门控机制控制信息的传递和更新。

具体来说，式(3)中的 \mathbf{g}_t 是输入调制门，主要功能是调节新信息的流入量，即如式(6)所示，输入调制门的输出 \mathbf{g}_t 和输入门的输出 i_t 逐元素相乘，决定新输入信息如何影响非平稳性特征 \mathbf{N} 。式(4)中的 i_t 为输入门，决定了隐藏状态和时空记忆状态中的信息有多少会被写入网络中的记忆单元。式(5)中的 f_t 是遗忘门，它控制了记忆单元中信息的遗忘程度，确保模型能够适应雷达回波图中的动态变化。式(7)中的 o_t 是输出门，负责决定有多少记忆单元中的信息会被输出到网络的下一层或作为最终预测。式(8)为该模块输出结果。由于自注意力机制可以直接计算序列中任意 2 个元素之间的关系，因此它不受序列长度的限制，这有助于模型在面对复杂或快速变化的动态系统时保持较高的预测准确性和适应性，降低预测误差的波动范围，使预测结果更加稳定可靠。

3.3 基于局部注意力机制的 MIM-S 模块

在 MIM 模型中，MIM-S 模块的核心职责是处理数据中的平稳性特征，该模块将 MIM-N 模块的输出差分特征 D_t^l 和上一时间步的时空状态 C_{t-1}^l 作为输入，以捕捉序列中近似静止的变化。

加入局部注意力 (local-attention) 机制的 MIM-S 模块简称 LA-MIMS 模块，其结构如图 5 所示。通过引入局部注意力机制，该模块能更加聚焦于局部区域的特征变化，这对于捕捉像素值的空间相关性和时间依赖性等低阶平稳性非常有效。局部注意力机制能够根据输入数据的不同自动调整关注的局部区域大小，使模型在处理不同空间尺度上的平稳性特征时更加灵活。

LA-MIMS 模块首先对细胞状态 \mathbf{C} 进行填充，然后通过卷积操作使用局部注意力权重进行局部注意力计算，如式(9)所示，接着与原始特征进行融合，通过式(10)~式(14)的门控机制控制信息的传递和更新。

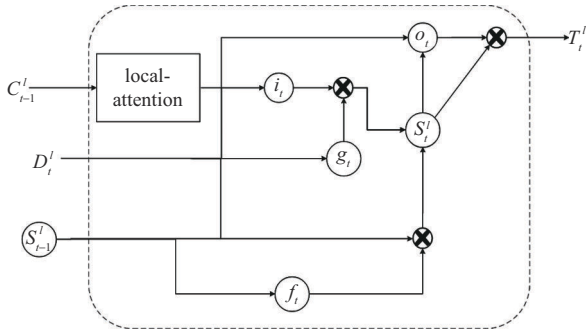


图5 LA-MIMS模块结构

$$\tilde{C} = \text{LA}(C) \quad (9)$$

$$g_t = \tanh(W_{dg} * D_t^l + W_{cg} * \tilde{C}_{t-1}^l + b_g) \quad (10)$$

$$i_t = \sigma(W_{di} * D_t^l + W_{ci} * \tilde{C}_{t-1}^l + b_i) \quad (11)$$

$$f_t = \sigma(W_{df} * D_t^l + W_{cf} * \tilde{C}_{t-1}^l + b_f) \quad (12)$$

$$S_t^l = f_t \odot S_{t-1}^l + i_t \odot g_t \quad (13)$$

$$o_t = \sigma(W_{do} * D_t^l + W_{co} * \tilde{C}_{t-1}^l + W_{so} * S_t^l + b_o) \quad (14)$$

$$T_t^l = \text{LA-MIMS}(D_t^l, \tilde{C}_{t-1}^l, S_{t-1}^l) = o_t \odot \tanh(S_t^l) \quad (15)$$

其中， S 表示在平稳模块中水平移动的存储单元，保存着平稳性特征的信息， D 是由SA-MIMN模块学习的差分特征， T 是一个存储单元，存储由LA-MIMS模块学习到的特征信息。

具体来说，式(10)中的 g_t 是输入调制门，主要功能是确定新输入的差分特征信息对平稳性特征 S 的影响程度，即如式(13)所示，输入调制门的输出 g_t 和输入门的输出 i_t 逐元素相乘，决定新输入信息如何影响平稳性特征 S 。式(11)中的 i_t 为输入门，决定了差分特征和时空记忆状态中的信息有多少会被写入网络中的记忆单元。式(12)中的 f_t 是遗忘门，它控制了差分特征和记忆单元中信息的遗忘程度。式(14)中的 o_t 是输出门，负责决定有多少差分特征、记忆单元和平稳模块中的信息会被输出到网络的下一层或作为最终预测。式(15)是该模块的输出结果。LA-MIMS模块引入局部注意力机制有助于捕捉数据中的平稳变化，通过聚焦于局部特征能够减少噪声和干扰对预测结果的影响，从而提高预测的精度和稳定性。

3.4 MDA-MIM模型

DA-MIM单元利用SA-MIMN模块和LA-MIMS模块分别处理数据中的非平稳性特征和平稳性特征，结构如图6所示。通过堆叠多个DA-MIM单元，能够捕捉更高阶的非平稳性特征，从而使时空

过程逐渐平稳，增强未来序列的可预测性。这种结构的关键是提供必要的隐藏状态以生成差异特征，并更好地促进非稳态建模。

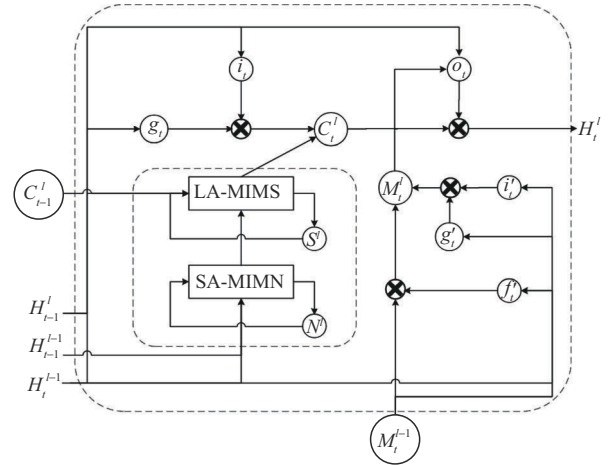


图6 DA-MIM单元结构

4 实验

实验中的采用CPU为64核128线程的AMD EPYC 7T83，GPU使用的是RTX 4090，操作系统为Ubuntu20.04，Python版本为3.8，深度学习框架为TensorFlow1.15.5，Adam为模型的优化器，模型在训练和测试阶段的初始学习率是0.001。

4.1 评估指标

本文采用均方误差（MSE）、平均绝对误差（MAE）、结构相似性指数（SSIM）、峰值信噪比（PSNR）作为评估模型的指标。

MSE和MAE的计算式分别为

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (16)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (17)$$

其中， n 为样本数量， y_i 为像素点的真实特征值； \hat{y}_i 为像素点的特征预测值。MSE、MAE的值越小代表误差越小。

SSIM的范围为[-1, 1]，计算式为

$$\text{SSIM}(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (18)$$

SSIM由3个部分组成， $\alpha, \beta, \gamma > 0$ ，用于调整3个部分的权重； $l(x, y)$ 用于比较2个图像的亮度； $c(x, y)$ 用于比较2个图像的对比度； $s(x, y)$ 用于比较2个图像的结构，SSIM的值越大代表相似度越高。

PSNR用于衡量2个图像之间的差异，最小值

为 0，其计算式为

$$\text{PSNR} = 10 \lg \left(\frac{\text{MaxValue}^2}{\text{MSE}} \right) \quad (19)$$

其中，MSE 为 2 个图像的均方误差；MaxValue 是图像像素可取到的最大值。对于彩色图像，可以分别计算 RGB 的 3 个通道的 PSNR，然后取平均值。PSNR 的值越大，2 个图像差异越小。

4.2 实验结果及分析

本文收集了山东省青岛市 2019 年 1 月至 2020 年 12 月的雷达回波图数据作为实验数据。首先对这些数据进行预处理，去除空白图片以消除不必要数据对模型预测能力的干扰，每张雷达图都被处理成为 64×64 的网格图像。训练数据集包含 96 240 个图片，测试数据集包含 20 200 个图片。使用一个 20 帧宽的窗口对连续图像进行切片，因此每个序列由 20 帧连续图像组成，其中前 10 帧为输入，后 10 帧为输出。在评估模型的预测能力时，以前 10 帧预测图像作为输入，迭代预测后 10 帧图像，以均值作为最终的实验结果。实验使用 MSE 和 MAE 作为评估模型预测精度的指标，同时采用 SSIM 和 PSNR 来衡量生成图像的质量。实验结果如表 1 所示。

表 1 不同模型的预测性能对比

模型	MSE↓	MAE↓	SSIM↑	PSNR↑
FRNN	50.07	270.80	0.570 1	19.39
PredRNN++	46.40	223.29	0.667 2	19.99
ConvLSTM	45.46	221.23	0.715 0	20.03
MIM	41.71	200.90	0.734 2	20.39
PPNet	41.19	190.16	0.737 5	20.45
AFR-LSTM	41.52	200.17	0.735 0	20.33
SAMM-MIM	41.29	201.34	0.734 9	20.42
FDNet	40.94	189.95	0.740 4	20.59
CMFPN	40.81	182.32	0.743 8	20.48
MDA-MIM	40.86	180.56	0.744 7	20.65

从表 1 可以看出，MDA-MIM 模型虽然在 MSE 方面略差于 CMFPN，但在其他指标上均表现最佳。与表 1 中所有的对比模型的各个指标均值对比，MDA-MIM 的 MSE 和 MAE 分别降低了 5.55% 和 13.57%，SSIM 和 PSNR 分别提升了 5.08% 和 2.08%，因此，MDA-MIM 在雷达回波图的预测质

量上优于对比模型。MDA-MIM 模型的 MSE 相较于 MIM 模型降低了 2.04%，MAE 相较于 MIM 模型降低了 10.12%，相较于 SAMM-MIM 模型降低了 10.32%。在图像质量评估方面，MDA-MIM 模型的 SSIM 和 PSNR 比 MIM 模型分别提高了 1.43% 和 1.28%，比 SAMM-MIM 模型分别提高了 1.33% 和 1.13%，这说明本文对 MIM 模型的改进是有效的。

为了更直观地比较模型的预测结果，本文将预测结果的第 12 帧、第 14 帧、第 16 帧、第 18 帧和第 20 帧可视化，如图 7 所示。

从图 7 可以看出，相比其他模型，MDA-MIM 模型预测的图像在形状和空间变化趋势上与真实图片更为相似，这说明本文的 MDA-MIM 预测精确度明显优于对比模型。此外，对于第 18 帧和 20 帧的图像，与真实图像对比，采用 MDA-MIM 模型预测的图像，不仅具有更高的预测精确度，而且具有清晰的边界，能够更好地保留细节信息，这说明 MDA-MIM 不仅预测质量高，而且具有优秀的长序列图像的预测能力。

4.3 消融实验

为了进一步研究多尺度特征提取与融合模块、SA-MIMN 模块和 LA-MIMS 模块的有效性，本文对比了去掉多尺度特征提取与融合模块的 DA-MIM 模型、去掉自注意力机制的 MLA-MIM 模型、去掉局部注意力机制的 MSA-MIM 模型、去掉双重注意力机制的 M-MIM 模型和本文提出的 MDA-MIM 模型在雷达回波数据集上的 MSE、MAE、SSIM 和 PSNR 指标。实验结果如表 2 所示，为了更直观地展示实验结果，分别给出不同指标下各模型的柱状图，如图 8~图 11 所示。

如图 8 所示，从 MDA-MIM 模型中去掉多尺度特征提取与融合模块、自注意力机制、局部注意力机制和双重注意力机制均会导致 MSE 增加，其中 DA-MIM 模型的 MSE 略有上升，表明多尺度特征对于提升预测精度具有积极作用。MLA-MIM 模型的 MSE 为 41.68，而 MSA-MIM 模型的 MSE 则为 41.57，说明注意力机制对于减少预测误差有显著贡献，其中自注意力机制的贡献最大。此外，M-MIM 模型的 MSE 为 41.38，尽管较 MDA-MIM 有所提升，但仍优于仅去除单一注意力机制的模型，这暗示了两者虽各有贡献，但共同作用时效果最佳。



图 7 不同模型下多步预测的雷达回波图

表 2 消融实验结果

模型	MSE↓	MAE↓	SSIM↑	PSNR↑
DA-MIM	41.15	189.84	0.738 8	20.57
MLA-MIM	41.68	197.56	0.733 4	20.45
MSA-MIM	41.57	193.61	0.734 3	20.43
M-MIM	41.38	191.20	0.736 7	20.63
MDA-MIM	40.86	180.56	0.744 7	20.65

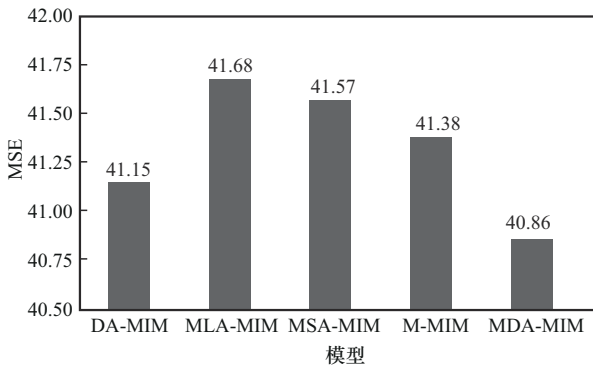


图 8 不同模型下的 MSE 对比

如图 9 所示, MAE 的结果与 MSE 类似, 去掉任意模块均会导致 MAE 增加, 其中自注意力机制的影响最大, 局部注意力机制次之, 最后是双重注意力机制和多尺度特征提取与融合模块。这说明双重注意力机制和多尺度特征提取与融合模块对于雷达回波图预测的准确性是不可或缺的。

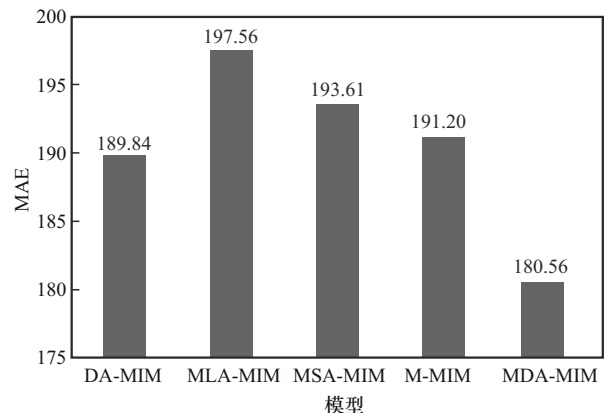


图 9 不同模型下 MAE 对比

从图 10 可以看到, DA-MIM 模型的 SSIM 略有下降, 说明多尺度特征提取与融合模块对保持图像结构信息有一定贡献。去掉自注意力机制的 MLA-MIM 模型和去掉局部注意力机制 MSA-MIM 模型的 SSIM 下降最多, 虽然去除双重注意力机制的 M-MIM 模型较任意去掉一个注意力机制的 SSIM 有所上升, 但它仍低于 DA-MIM 和 MDA-MIM, 进一步证实了这 2 种注意力机制在保持结构相似性方面的重要性。

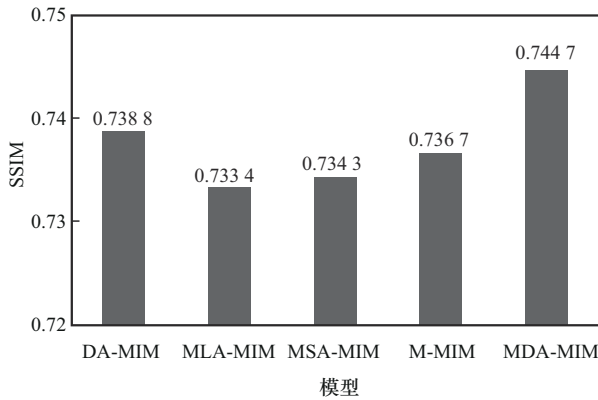


图 10 不同模型下的 SSIM 对比

如图 11 所示, 去除局部注意力机制的 MSA-MIM 模型的 PSNR 是最低的, 这说明局部注意力机制对图像质量的贡献最大, 其次是自注意力机制, 而同时去掉两者后, PSNR 值虽接近原始 MDA-MIM 但略低。去除多尺度特征提取与融合模块后, DA-MIM 模型的 PSNR 也略有降低, 表明多尺度特征对图像质量有一定贡献。这些结果表明, MDA-MIM 模型中的多尺度特征提取与融合模块以及注意力机制均对提升预测图像的峰值信噪比具有积极影响。

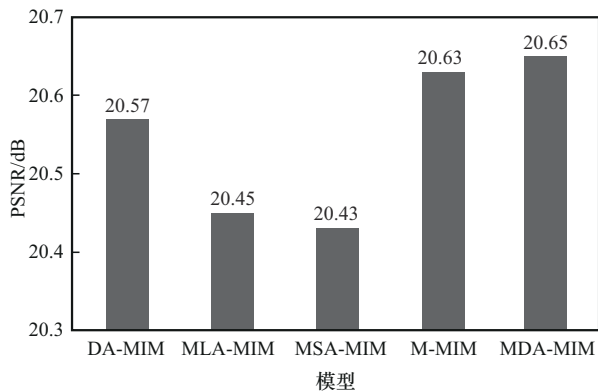


图 11 不同模型下的 PSNR 对比

综上所述, 虽然不同模块对不同指标的影响不同, 但是它们均不及 MDA-MIM 模型的效果好, 即 MDA-MIM 模型通过整合多尺度特征提取与融合模块和注意力机制, 有效地提升了雷达回波图预测的准确性。

4.4 复杂度分析

MDA-MIM 的复杂度由 MEF 和 DA-MIM 这 2 个模块的复杂度组成, 其时间复杂度为 $O(T \times (C_{out} \times C_{in} \times K_1^2 \times d_c^2 \times H \times W) + H \times W \times K_2^2 L + T d_a)$, 其中, T 为序列长度, H 和 W 分别为每一帧的高度和宽度, C_{in} 和 C_{out} 分别为空洞卷积的输入和输出通道数, K_1 和 K_2 分别为空间卷积网络和 MIM 的卷积核大小, L 为 MIM 块的数量, d_c 为空洞率, d_a 为注意力特征维度。MDA-MIM 空间复杂度为 $O(C_{out} \times H \times W + C_{out} \times C_{in} \times K_2 + T \times H \times W \times d_f + L \times P + T \times H \times W \times d_f \times L + T^2 + T \times d_a)$, 其中 P 为每个 MIM 块的参数量, d_f 为 MIN 中向量特征维度。具体地, 在本实验中部署的 MDA-MIM 模型的参数量为 137.1 MB, 训练时间为 7.36 h。

5 结束语

本文设计了一种雷达回波图预测模型——MDA-MIM, 该模型通过引入多尺度特征融合和双重注意力机制对 MIM 进行了改进。模型中的多尺度特征提取与融合模块利用空洞卷积技术实现了从不同尺度提取图像特征的能力, 增强了模型对细节的感知能力。通过在 MIM-N 模块中增加自注意力机制和在 MIM-S 模块中引入局部注意力机制, MDA-MIM 模型能够有效地捕捉雷达回波数据中的非平稳性特征和平稳性特征。实验结果表明, MDA-MIM 模型在雷达回波图预测任务中优于对比模型。消融实验表明本文中对 MIM 的 3 个模块的改进均能提高雷达回波图的预测精确度。下一步的研究工作是构建多模态的降水预测模型, 将雷达回波图的预测结果用于降水预测。

参考文献:

- [1] NIU D, LI Y H, WANG H B, et al. FsrGAN: a satellite and radar-based fusion prediction network for precipitation nowcasting[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2024, 17: 7002-7013.
- [2] HE W, XIONG T S, WANG H, et al. Radar echo spatiotemporal sequence prediction using an improved ConvGRU deep learning model[J].

- Atmosphere, 2022, 13(1): 88.
- [3] 张林, 李峰, 吴蕾, 等. CINRAD/SAD 双偏振雷达非降水回波识别技术[J]. 应用气象学报, 2022, 33(6): 724-735.
ZHANG L, LI F, WU L, et al. Non-precipitation identification technique for CINRAD/SAD dual polarimetric weather radar[J]. Journal of Applied Meteorological Science, 2022, 33(6): 724-735.
- [4] WOO W C, WONG W K. Operational application of optical flow techniques to radar-based rainfall nowcasting[J]. Atmosphere, 2017, 8(3): 48.
- [5] JOHNSON J T, MACKEN P L, WITT A, et al. The storm cell identification and tracking algorithm: an enhanced WSR-88D algorithm[J]. Weather and Forecasting, 1998, 13(2): 263-276.
- [6] 王改利, 刘黎平. 多普勒雷达资料在暴雨临近预报中的应用[J]. 气象, 2005, 31(10): 12-15, 97.
WANG G L, LIU L P. Application of Doppler radar observation to rainfall nowcasting[J]. Meteorological Monthly, 2005, 31(10): 12-15, 97.
- [7] SHI X J, CHEN Z R, WANG H, et al. Convolutional LSTM network[C]// Proceedings of the 29th International Conference on Neural Information Processing Systems - Volume 1. New York: ACM Press, 2015: 802-810.
- [8] WANG Y B, GAO Z F, LONG M S, 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. New York: ACM Press, 2018: 5123-5132.
- [9] WANG Y B, ZHANG J J, ZHU H Y, et al. Memory in memory: a predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics[C]// Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2019: 9146-9154.
- [10] ZHOU W, QU Z W, QU L N, et al. Radar echo maps prediction using an improved MIM with self-attention memory module[J]. Journal of Sensors, 2023, 2023(1): 8876971.
- [11] 程勇, 钱坤, 康志明, 等. 一种雷达回波外推的注意力融合和信息回忆的 LSTM 方法[J]. 热带气象学报, 2023, 39(5): 653-663.
CHENG Y, QIAN K, KANG Z M, et al. An attention fusion and information recall lstm method for radar echo extrapolation[J]. Journal of Tropical Meteorology, 2023, 39(5): 653-663.
- [12] FANG W, PANG L, SHENG V S, et al. STUNNER: radar echo extrapolation model based on spatiotemporal fusion neural network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 1-14.
- [13] ZHENG C 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.
- [14] NAZ F, SHE L, SINAN M, et al. Enhancing radar echo extrapolation by ConvLSTM2D for precipitation nowcasting[J]. Sensors, 2024, 24(2): 459.
- [15] 宋毅, 张晗奕, 孙丰, 等. PpNet: 基于预先预测的降雨短时预测模型[J]. 电子与信息学报, 2024, 46(2): 492-502.
SONG Y, ZHANG H Y, SUN F, et al. PpNet: a precipitation nowcasting model based on pre-prediction[J]. Journal of Electronics & Information Technology, 2024, 46(2): 492-502.
- [16] FANG W, SHEN L, SHENG V S. VRNet: a vivid radar network for precipitation nowcasting[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 4103811.
- [17] ZHAO Z W, DONG X C, WANG Y P, et al. MDTNet: multiscale deformable Transformer network with Fourier space losses toward fine-scale spatiotemporal precipitation nowcasting[J]. IEEE Transactions on Geoscience and Remote Sensing, 2024, 62: 4106417.
- [18] TANG Y Q, WANG H B, NIU D, et al. CMFPN: a cross-modal fusion prediction network for precipitation nowcasting[C]// Proceedings of the 2024 36th Chinese Control and Decision Conference (CCDC). Piscataway: IEEE Press, 2024: 2456-2461.
- [19] YAN B Y, YANG C, CHEN F, et al. FNet: a deep learning approach with two parallel cross encoding pathways for precipitation nowcasting[J]. Journal of Computer Science and Technology, 2023, 38(5): 1002-1020.
- [20] YAO J L, JI J W, WANG R B, et al. DSADNet: a dual-source attention dynamic neural network for precipitation nowcasting[J]. Sustainability, 2024, 16(9): 3696.
- [21] 薛丰昌, 章超钦, 王文硕, 等. 基于自注意力和门控循环神经网络的雷达回波外推算法研究[J]. 气象学报, 2024, 82(1): 127-135.
XUE F C, ZHANG C Q, WANG W S, et al. Improving radar echo extrapolation algorithms based on selfattention and gated recurrent neural networks[J]. Acta Meteorologica Sinica, 2024, 82(1): 127-135.
- [22] 高红民, 曹雪莹, 陈忠昊, 等. 基于多尺度近端特征拼接网络的高光谱图像分类方法[J]. 通信学报, 2021, 42(2): 92-102.
GAO H M, CAO X Y, CHEN Z H, et al. Hyperspectral image classification method based on multi-scale proximal feature concatenate network[J]. Journal on Communications, 2021, 42(2): 92-102.
- [23] 杨静, 吴成茂, 周流平. 基于全局-局部自注意力网络的视频异常检测方法[J]. 通信学报, 2023, 44(8): 241-250.
YANG J, WU C M, ZHOU L P. Novel video anomaly detection method based on global-local self-attention network[J]. Journal on Communications, 2023, 44(8): 241-250.

作者简介



胡强 (1980-), 男, 山东邹城人, 博士, 青岛科技大学教授、博士生导师, 主要研究方向为服务计算、人工智能。

高雅婷 (2001-), 女, 山东滨州人, 青岛科技大学硕士生, 主要研究方向为人工智能。

尹宾礼 (1970-), 男, 山东青岛人, 青岛广播电视台工程师, 主要研究方向为信号与系统、无线通信。

渠连恩 (1980-), 男, 黑龙江黑河人, 博士, 新疆工程学院副教授、硕士生导师, 主要研究方向为计算机视觉、时序预测、车辆非线性状态估计等。