

文章编号:1671-6833(2025)05-0137-08

# 融合时空重构单元和 Transformer 的雷达回波外推算法

方巍<sup>1,2,3,4</sup>, 王溟西<sup>1</sup>

(1. 南京信息工程大学 计算机学院, 江苏 南京 210044; 2. 南京信息工程大学 软件学院, 江苏 南京 210044; 3. 中国气象局武汉暴雨研究所 中国气象局流域强降水重点开放实验室/暴雨监测预警湖北省重点实验室, 湖北 武汉 430205; 4. 中国气象科学研究院 灾害天气国家重点实验室, 北京 100081)

**摘要:** 针对现有基于深度学习的雷达回波外推算法时空特征提取能力不足和建模能力有限的问题, 提出了一种融合时空重构单元(SRU)和 Transformer 的雷达回波外推模型 SRU-Former。首先, 在模型的编码器和解码器中引入新设计的时空重构单元, 通过分离、变换和重构的策略来提取雷达图像精细化时空特征; 其次, 在编码器和解码器之间引入 Transformer 的变体架构模型 Poolformer, 用全局平均池化操作代替自注意力机制, 帮助模型对高度动态变化的雷达序列进行建模; 最后, 在江苏省气象雷达数据集和上海市气象雷达数据集上训练和测试, 与目前主流的深度学习模型进行对比。实验结果表明: 在 2 h 外推任务中, *CSI*、*FAR*、*MSE* 和 *SSIM* 4 个指标均取得最优值, 在江苏省数据集上 *CSI* 提升了 0.020, 上海市中数据集上 *CSI* 提升了 0.048; SRU-Former 能够有效提升模型的预报准确率, 外推后期对强回波区域的捕捉更加精确, 细节纹理更加丰富清晰。

**关键词:** 深度学习; 雷达回波外推; Transformer; 时空重构单元; 全局平均池化

中图分类号: TP399

文献标志码: A

doi: 10.13705/j.issn.1671-6833.2025.02.007

多普勒天气雷达在短时降水预报中扮演着重要的角色, 提高雷达回波外推精度对强对流天气预警和经济稳定发展具有重要意义<sup>[1]</sup>。传统的雷达回波外推方法<sup>[2-4]</sup>大多是基于线性的方法, 忽略了中小尺度天气系统中雷达回波的非线性运动特征, 存在局限性。近年来, 深度学习和气象学之间的结合日益紧密, 可以从海量数据中揭示潜在的复杂关系, 对雷达回波的时空动态变化进行高效建模。

循环神经网络(recurrent neural network, RNN)是一种通过引入时间概念来处理序列数据的神经网络, 能够在雷达序列外推中发挥重要的作用。卷积长短期记忆网络(convolutional long short-term memory, ConvLSTM)<sup>[5]</sup>的提出极大地推动了 RNN 在雷达回波外推领域的应用, 其通过结合卷积神经网络(convolutional neural networks, CNN)和 RNN 的特性, 能够捕捉到雷达数据中的内在规律和模式。Wang 等<sup>[6-7]</sup>先后推出预测循环神经网络(predictive recurrent neural network, PredRNN)和预测循环神经

网络改进模型(predictive recurrent neural network plus plus, PredRNN++), 引入更高效的时空记忆单元来捕捉时空信息, 但随着预报时间的延长, 捕捉动态变化的能力逐渐降低。方巍等<sup>[8]</sup>在 ConvLSTM 的基础上提出了时空融合网络, 将瞬态变化与运动趋势同时融合。

尽管目前很多基于 RNN 的雷达回波外推模型被提出, 但其串行处理方式限制了计算效率, 导致预测结果的不稳定性和误差累积<sup>[9]</sup>。随后, 研究者们将经典卷积网络模型 U-Net<sup>[10]</sup>引入雷达回波外推领域观察它的适用性, 取得了不错的效果<sup>[11]</sup>。Trebing 等<sup>[12]</sup>在 U-Net 的编码器和解码器中嵌入注意力机制, 提出了 SmaAt-U-Net, 具有更高的运算效率和更低的资源需求。Gao 等<sup>[13]</sup>提出了 SimVP 模型, 它完全建立在 CNN 的基础上, 以较低的计算成本在雷达回波外推领域取得了出色的效果, 但 CNN 主要关注于图像的空间特征提取, 缺乏处理时间序列数据的天然机制, 无法充分建立雷达序列的时间依赖关系

收稿日期: 2024-11-02; 修订日期: 2024-12-28

基金项目: 国家自然科学基金资助项目(42075007, 42475149); 中国气象局流域强降水重点开放实验室开放研究基金(2023BHR-Y14); 江苏省研究生科研与实践创新计划项目(KYCX24\_1533, SJCX24\_0476)

作者简介: 方巍(1975—), 男, 安徽黄山人, 南京信息工程大学教授, 博士, 博士生导师, 主要从事人工智能和气象大数据处理研究, E-mail: hsfangwei@sina.com。

引用本文: 方巍, 王溟西. 融合时空重构单元和 Transformer 的雷达回波外推算法[J]. 郑州大学学报(工学版), 2025, 46(5): 137-144. (FANG W, WANG H X. Radar echo extrapolation algorithm integrating spatiotemporal reconstruction unit and Transformer [J]. Journal of Zhengzhou University (Engineering Science), 2025, 46(5): 137-144.)

模型<sup>[14]</sup>。随着 Transformer<sup>[15]</sup>在自然语言处理领域的出现,其提取长序列信息的潜力促使众多研究者将其应用于雷达回波外推。方巍等<sup>[16]</sup>将 Swin Transformer 模型<sup>[17]</sup>和 UNet 模型相结合来更好地应对雷达回波的非线性变化,提高预测结果的鲁棒性。Bi 等<sup>[18]</sup>设计出了盘古气象大模型,其主要思想就是使用 ViT(vision transformer)<sup>[19]</sup>的 3D 变体来处理各种复杂的气象要素,在 ERA5 再分析数据的中长期气象预报上首次超过了传统数值方法。基于 Transformer 的雷达回波外推模型领域虽然序列处理能力强,但也面临计算成本高和空间建模能力不足的问题。

为了克服现有深度学习雷达回波外推模型在时空特征提取上的局限性,本文提出了一种融合时空重构单元(spatiotemporal reconstruction unit,SRU)和 Transformer 的外推模型 SRU-Former,该模型融合 CNN 和 Transformer 的特性,充分发挥 CNN 在雷达图像特征提取和局部细节感知方面的优势,以及 Transformer 在捕捉长时雷达回波序列依赖关系的能力。本文应用 SRU-Former 网络模型在江苏省气象雷达数据集和上海市气象雷达数据集上与现有主流深度学习模型进行了实验对比,实验证明,SRU-Former 在多个评价指标上都取得很好的效果。

## 1 SRU-Former 模型

### 1.1 整体网络架构

本文提出的 SRU-Former 网络模型是一种编码器和解码器架构,编码器堆叠了 4 个模块,每个模块包括 1 个卷积核大小为  $3 \times 3$  的卷积层、1 个组归一化层、GELU 激活函数和 1 个 SRU。解码器同样也堆叠了 4 个模块,每个模块包括 1 个卷积核大小为  $3 \times 3$  的转置卷积层、1 个组归一化层、GELU 激活函数和 1 个时空重构单元 SRU。其中编码器和解码器的 4 个卷积层步长序列均为  $[2, 1, 2, 1]$ ,即分别包括 2 个下采样层和 2 个上采样层。SRU 用来加强模型对雷达图的时空特征提取能力。此外,在编码器和解码器之间,本文融合了 Transformer 的变体架构

模型 Poolformer<sup>[20]</sup>来捕捉雷达回波序列的时间依赖关系和动态变化信息,将 Poolformer 模块的个数设为 4。考虑到雷达图具有较高的分辨率,本文通过 Patch 操作将雷达图划分为若干个小块作为模型的输入单元,Patch 操作可以使得模型捕捉到更多的图像细节和局部变化,并且能够降低计算复杂度。图 1 为 SRU-Former 的网络模型结构图。

### 1.2 时空重构单元 SRU

雷达数据中存在着大量的非平稳信息,例如云层的堆积、变形和消散等,这就使得雷达回波序列的时空特征提取变得尤为困难。近年来,研究者们通常在模型中嵌入注意力机制<sup>[21]</sup>来自适应地关注较关键的雷达回波区域,但是注意力机制的效果在很大程度上依赖雷达数据的质量,且往往需要大量的计算资源来训练和优化模型。本文受 SCConv 模块<sup>[22]</sup>的启发,设计并提出了 SRU,能够很好地改善这个问题,在增强时空特征提取能力的同时,可以更高效地处理雷达数据。

特征提取时的信息冗余在很大程度上是导致雷达回波外推精度不高的重要因素。SRU 采用分离、变换和重构的策略来减少时空信息冗余并促进特征学习,图 2 为 SRU 模块结构。SRU-Former 模型的输入形状是五维张量  $B \times T \times C \times H \times W$ ,其中  $B$  为批次大小, $T$  为序列长度, $C$  为通道数, $H$  和  $W$  分别为输入图像的高和宽。本文将输入张量重新塑造成形状为  $B \times (T \times C) \times H \times W$  的四维张量,将时间和通道维度进行了合并,把这个新的维度叫作  $F$ ,SRU 的输入特征  $X$  维度为  $B \times F \times H \times W$ 。首先将  $X$  的  $F$  通道按照比例因子  $\alpha$  拆分为两部分  $\alpha \times F$  和  $(1 - \alpha) \times F$ ,本文将  $\alpha$  设为 0.5。随后对这两部分分别进行  $1 \times 1$  卷积,来压缩通道数从而降低计算复杂度。经过分离操作,得到上半部分特征  $X_{up}$  和下半部分特征  $X_{low}$ ,用分组卷积 GWC 和逐点卷积 PWC 分别对  $X_{up}$  进行特征提取,并将结果直接相加得到  $Y_1^w$ ,而对  $X_{low}$  则是直接用 PWC 进行特征提取,将其结果与  $X_{low}$  在  $F$  维度上进行拼接得到  $Y_2^w$ , $Y_1^w$  相比于  $Y_2^w$  特征信息要较丰富些,上述操作如式(1)和式(2)所示。GWC 和

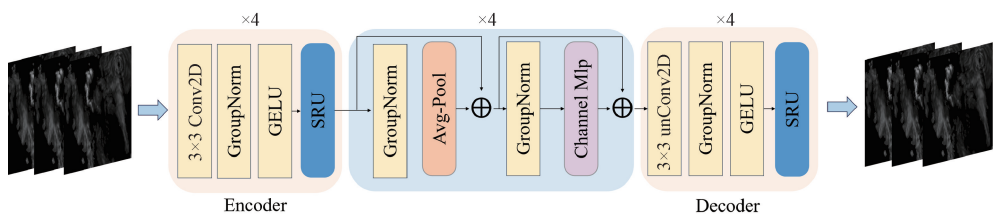


图 1 SRU-Former 网络模型结构图

Figure 1 Structure diagram of SRU-Former network model

PWC 是深度学习中两种不同的卷积方式,相比于标准卷积而言,可以更好地捕捉输入数据的特征。GWC 将输入特征图按照  $F$  维度分为若干个组,每个组内分别进行独立的卷积操作,PWC 就是  $1 \times 1$  卷积,主要用来减少特征图维度。

$$Y_1^w = M^C X_{up} + M^{P1} X_{up}; \quad (1)$$

$$Y_2^w = M^{P2} X_{low} \cup X_{low} \circ \quad (2)$$

式中:  $M^C$  为 PWC 卷积操作,本文将分组数设为 2;  $M^{P1}$  和  $M^{P2}$  两个  $1 \times 1$  卷积分别应用在上半部分特征  $X_{up}$  和下半部分特征  $X_{low}$  上,输入  $F$  维度数受分离操作中的比例因子  $\alpha$  影响,由于  $\alpha$  设置为 0.5,  $M^{P1}$  和  $M^{P2}$  完全相同;  $Y_1^w$  和  $Y_2^w$  为得到的上下部分新的特

征;  $+$  和  $\cup$  分别为变量求和与  $F$  维度叠加操作。

为了进一步减少时空特征冗余,本文并没有将特征信息较丰富的  $Y_1^w$  和特征信息较少的  $Y_2^w$  直接相加,而是采取交叉重构的操作,充分地融合这两部分特征,如式(3)~(5)所示:

$$Y_{11}^w \oplus Y_{22}^w = Y^{w1}; \quad (3)$$

$$Y_{21}^w \oplus Y_{12}^w = Y^{w2}; \quad (4)$$

$$Y^{w1} \cup Y^{w2} = Y^w \circ \quad (5)$$

具体而言,  $Y_1^w$  和  $Y_2^w$  在  $F$  维度上平均分割分别得到特征  $Y_{11}^w$  和  $Y_{12}^w$ ,  $Y_{21}^w$  和  $Y_{22}^w$ ,交叉地将  $Y_{11}^w$  与  $Y_{22}^w$  相加得到  $Y^{w1}$ ,  $Y_{21}^w$  与  $Y_{12}^w$  相加得到  $Y^{w2}$ ,最后将  $Y^{w1}$  和  $Y^{w2}$  在  $F$  维度上进行拼接得到时空精细化特征  $Y^w$ 。

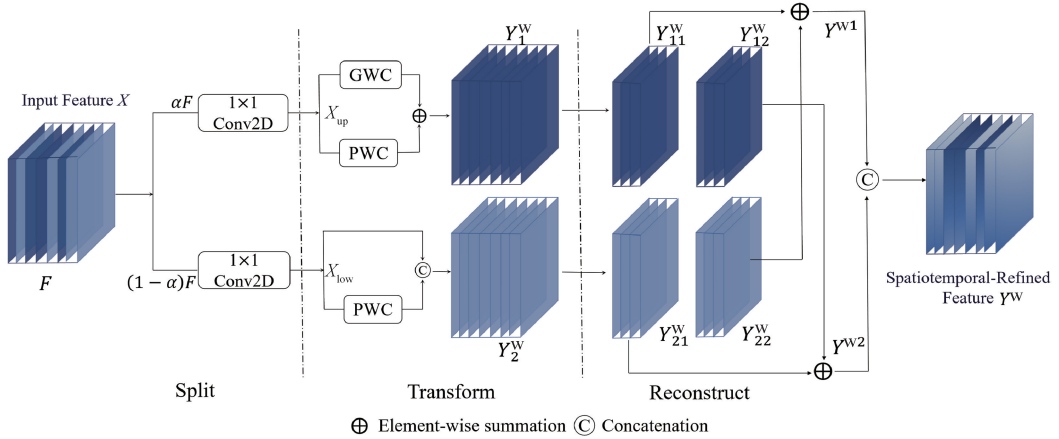


图 2 SRU 模块结构

Figure 2 SRU module structure

### 1.3 Poolformer 模块

雷达回波数据动态且复杂,RNN 虽能处理序列数据,但易梯度消失,而 Transformer 架构灵活,更适合时间序列处理。因此,本文在 SRU-Former 网络模型的编码器和解码器之间融合 Transformer 的变体模型 Poolformer<sup>[20]</sup>,帮助建立信息更加完整的长距离依赖关系。Poolformer 的提出证明了 Transformer 的成功主要归因于其通用架构,而非特定于自注意力机制的 Token Mixer 模块。在 Token Mixer 模块的选取上,自注意力机制和多层感知机都能够取得很好的效果。

图 3 为 Poolformer 模型结构。如图 3 所示,Poolformer 直接使用全局平均池化作为 Token Mixer 模块,更加高效地整合雷达图的时空信息。相较于自注意力机制用庞大的参数去存储和计算注意力权重,池化层不需要学习额外的参数,极大地减少了计算成本,提升了模型的鲁棒性和灵活性。具体流程如式(6)和式(7)所示:

$$Y = \text{TokenMixer}(\text{Norm}(X)) + X; \quad (6)$$

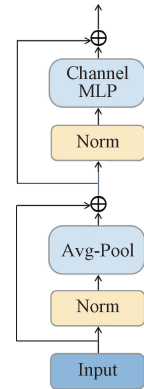


图 3 Poolformer 模型结构

Figure 3 Poolformer model structure

$$Z = \text{MLP}(\text{Norm}(Y)) + Y. \quad (7)$$

式中: TokenMixer 为对输入特征全局平均池化操作的结果再减去输入特征,这样有助于原始特征和池化特征的信息交互; Norm 为组归一化操作; MLP 为多层感知机,由两个全连接层和 GELU 激活函数构成;  $+$  为变量求和操作;  $X$  为输入特征,  $Y$  为经过第一个残差模块后输出的特征;  $Z$  为经过 Poolformer 整个模块后输出的特征;输入和输出张量形状保持

为  $B \times F \times H \times W$  不变。本文共堆叠 4 个 Poolformer 模块,有效地增强了模型在  $F$  维度(时空维度)上的特征表示能力。

将 Poolformer 的输出张量输入到解码器中,并将得到的输出张量重新塑造成形状为  $B \times T \times C \times H \times W$  的五维张量,保持模型初始输入和最终输出结果大小不变。

## 2 实验结果与分析

### 2.1 实验数据集

研究所使用的江苏省雷达回波数据集(见图 4(a<sub>1</sub>~a<sub>3</sub>))来源于江苏省气象台,共包含 31 892 个样本,265 个降水过程,每个样本是以 PNG 格式存储的灰度图。覆盖时间为 2019—2021 年 4—9 月份,覆盖空间范围为 116°18'E—121°57'E,30°45'N—35°20'N。数据样本的时间间隔为 6 min,图像分辨率为 480×560。使用步长为 3 的样本,长度为 30 的滑动窗口生成 10 000 个不重叠的雷达回波序列,能够降低样本重复率。顺序选取 9 000 个序列作为训练集,剩余 1 000 个序列作为测试集,10 帧长度用于输入,20 帧长度用于预测。

上海市雷达回波数据集(见图 4(b<sub>1</sub>~b<sub>3</sub>))由上海市气象台制作。资料共包含 40 000 个样本序列,243 个降水过程,每个样本是以 PNG 格式存储的灰度图。覆盖时间为 2017—2020 年 6—9 月份,覆盖空间范围为 120°51'E—122°12'E,30°40'N—31°53'N,图像分辨率为 460×460。顺序选取 32 000 个序列作为训练集,剩余 8 000 个序列作为测试集,10 帧长度用于输入,时间间隔为 6 min,后 10 帧用于预测,时间间隔为 12 min。

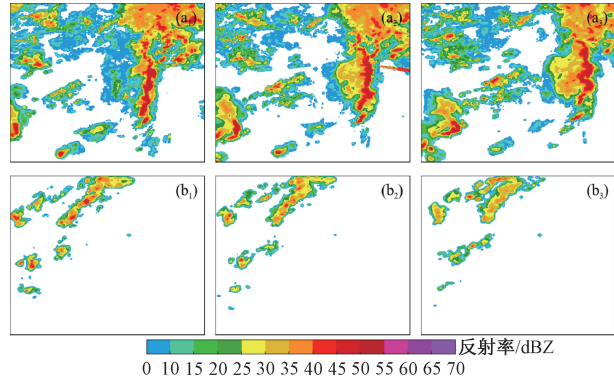


图 4 江苏省和上海市多普勒雷达反射率拼图示意

Figure 4 Example of the reflectivity mosaic of doppler radar in Jiangsu Province and Shanghai City

### 2.2 实验环境和训练策略

本研究使用 PyTorch 开源神经网络框架,所用服务器搭载 Linux 操作系统,运行内存为 125 G,显

卡为 GeForce RTX 4090 GPU。在模型训练过程中,本文将 *Batch Size* 设为 16, *Patch Size* 设为 4,即将原始图像分割为若干个大小为 4×4 的小块。使用 Adam 优化器,采用余弦退火策略<sup>[23]</sup>将学习率从  $10^{-3}$  降低到  $10^{-4}$ ,总共训练 200 个轮次。损失函数为均方误差 *MSE*,如式(8)所示:

$$MSE = \frac{1}{N} \sum_{n=1}^N \sum_{h=1}^H \sum_{w=1}^W (y_{n,h,w} - x_{n,h,w})^2. \quad (8)$$

式中:  $y_{n,h,w}$  为雷达回波图像的预测值;  $x_{n,h,w}$  为真实观测值;  $N$  为外推序列长度;  $H$  和  $W$  分别为图像的高和宽。

### 2.3 评价指标

实验采用气象业务中常用的临界成功指数 *CSI*、空报率 *FAR* 和命中率 *POD* 作为评估标准,如式(9)~(11)所示。雷达回波图中的反射率值表示单位体积内降水粒子的后向散射截面之和,根据反射率值的大小可以间接推断出降水粒子的密度、大小和分布情况。一般来说,当反射率值超过 15 dBZ 时,就可以认为与降水有关。因此,在本文中,计算阈值为 15 dBZ 下的 *CSI*、*FAR* 和 *POD*。在此之前要把模型预测的像素值转换成基本反射率,根据指定阈值将真实观测值和预测值分别进行 0-1 二值化处理。

$$CSI = \frac{n_{hits}}{n_{hits} + n_{misses} + n_{false\ alarms}}; \quad (9)$$

$$FAR = \frac{n_{false\ alarms}}{n_{hits} + n_{false\ alarms}}; \quad (10)$$

$$POD = \frac{n_{hits}}{n_{hits} + n_{misses}}. \quad (11)$$

式中:  $n_{hits}$ 、 $n_{misses}$  和  $n_{false\ alarms}$  分别表示命中点、漏报点和空报点的数量。*CSI* 和 *POD* 越高、*FAR* 越低,外推算法效果越好。此外,本文还采用结构相似性 *SSIM* 作为评价指标,如式(12)所示。*SSIM* 可以衡量预测雷达图像与标签雷达图像的相似程度。

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}. \quad (12)$$

式中:  $x$  和  $y$  表示真实观测值和预测值;  $\mu_x$  和  $\mu_y$  分别为  $x$  和  $y$  的平均值;  $\sigma_x^2$  和  $\sigma_y^2$  分别为  $x$  和  $y$  的方差;  $\sigma_{xy}$  为  $x$  和  $y$  的协方差;  $c_1$  和  $c_2$  为两个维持除式稳定的常数。

### 2.4 实验结果分析

本文应用 PredRNN<sup>[6]</sup>、PredRNN++<sup>[7]</sup>、SmaAtUNet<sup>[12]</sup>、SimVP<sup>[13]</sup>、ViT<sup>[19]</sup> 和 SRU-Former 共 6 个深度学习模型在两个数据集中分别进行训练和测试。表 1 为两个数据集上的定量指标结果对比。图 5 为

江苏省雷达数据集外推可视化结果图,测试样本时间为 2021 年 7 月 21 日 00 时 06 分—03 时 00 分。图 6 为上海市雷达数据集外推可视化结果图,测试样本时间为 2020 年 8 月 4 日 01 时 06 分—04 时 00 分。

由表 1 可知,本文提出的 SRU-Former 网络模型在两个数据集的所有指标中都取得了较好或者最好的结果。在两个数据集中,SRU-Former 的 *POD* 指标

均未达到最优,这主要是因为本文为了捕捉更多较弱的雷达回波信号,将反射率阈值设为 15 dBZ,在一定程度上弱化了 SRU-Former 在预报精度上的优势,*CSI* 作为一个综合考虑命中率 and 空报率的指标,更能有效地用于评估预报模型的准确性。从图 5 和图 6 中可以看出,SRU-Former 在对强回波区域的动态变化捕捉上优势显著,尤其在上海市雷达数据集中,精准地预测了强回波的演变趋势,在 2 h 外推后期(图 6(a<sub>3</sub>))

表 1 定量指标结果对比

Table 1 The comparison of quantitative indicator results

模型	江苏省数据集					上海市数据集				
	<i>CSI</i>	<i>POD</i>	<i>FAR</i>	<i>MSE</i>	<i>SSIM</i>	<i>CSI</i>	<i>POD</i>	<i>FAR</i>	<i>MSE</i>	<i>SSIM</i>
PredRNN	0.496	0.597	0.277	1.646	0.636	0.436	<b>0.901</b>	0.544	80.921	0.820
PredRNN++	0.505	0.610	0.268	1.633	0.649	0.755	0.815	0.088	8.396	0.973
SmaAt-UNet	0.523	0.673	0.322	1.753	0.649	0.688	0.773	0.140	13.956	0.955
SimVP	0.507	<b>0.678</b>	0.341	1.473	0.619	0.601	0.854	0.322	24.355	0.906
ViT	0.527	0.620	0.240	1.410	0.654	0.480	0.826	0.473	45.853	0.843
SRU-Former	<b>0.547</b>	0.632	<b>0.206</b>	<b>1.312</b>	<b>0.674</b>	<b>0.803</b>	0.870	<b>0.087</b>	<b>6.102</b>	<b>0.983</b>

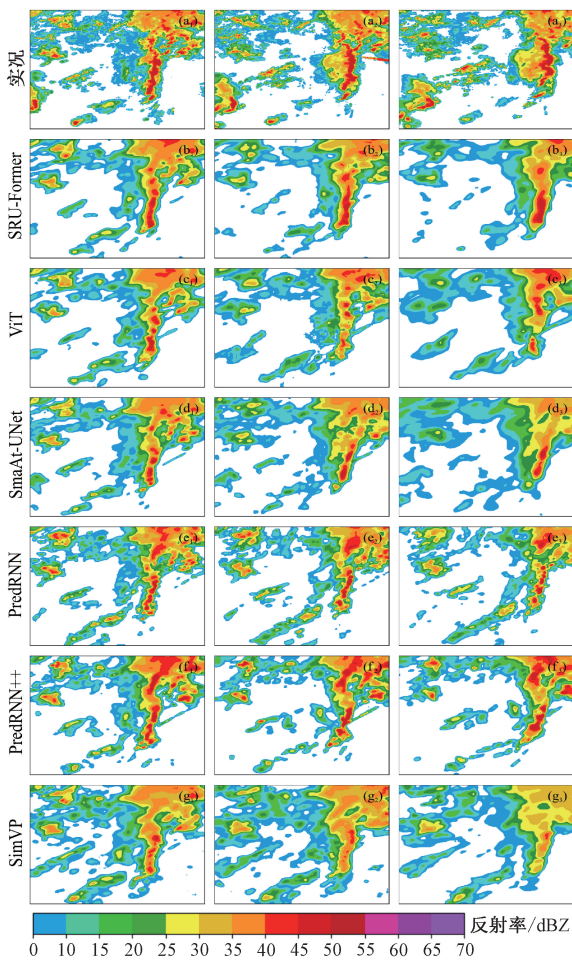


图 5 江苏省雷达反射率拼图外推结果对比

Figure 5 Comparison of extrapolation results of radar reflectivity puzzle in Jiangsu Province

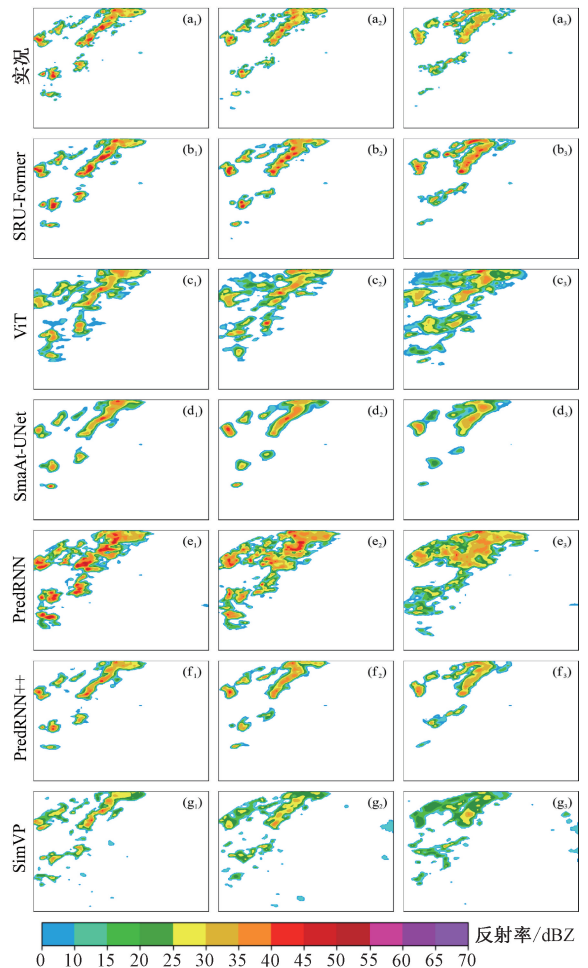


图 6 上海市雷达反射率拼图外推结果对比

Figure 6 Comparison of Extrapolation Results of Radar Reflectivity Puzzle in Shanghai City

依然能够保持对强回波区域的识别和预测。对比模型中,只有 SmaAt-UNet 和 PredRNN++在两个数据集中预报性能较为稳定,但仍落后于 SRU-Former; ViT 和 PredRNN 在江苏省雷达数据集中对回波区域尚能有效预测,但在上海市雷达数据集中外推图像细节丢失严重,空报率较高;SimVP 对于强回波区域的预测能力较弱,预测值过于保守,在两个数据集中都存在较为明显的回波削弱问题。

本文在实验中发现了不同雷达数据集之间的差异。江苏省雷达数据集质量相对较低,包含了大量的非均质性和不确定性。图 7 和图 8 展示了 SRU-Former 网络模型在两个数据集测试样本 10 个时间步上 CSI 评价指标的变化过程,每个时间步间隔 12 min。从图 7 可以看出,在数据质量较低的江苏省雷达数据集中,所有模型的预报能力均显著下降,SRU-Former 的下降趋势相较其他模型要稍缓慢些;从图 8 可以看出,在数据质量较高的上海市雷达数据集中,SRU-Former 表现得非常出色,CSI 有逐渐升高的趋势,表明 SRU-Former 网络模型在不断地学习和优化中可以更好地模拟雷达回波的变化规律,并且在更高质量的数据中可以实现更加精准的预报。

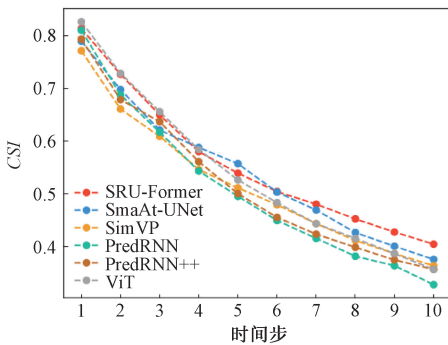


图 7 江苏省雷达数据集 CSI 变化

Figure 7 Changes in CSI of Jiangsu Province radar dataset

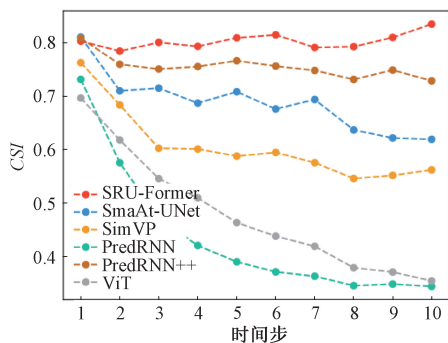


图 8 上海市雷达数据集 CSI 变化

Figure 8 Changes in CSI of Shanghai City radar dataset

## 2.5 消融实验

### 2.5.1 Poolformer 模块的计算效率分析

为了证明引入的 Poolformer 模块能够有效地提高模型的计算效率,本文在模型编码器和解码器之间融合其他 3 种时序依赖关系建模模型: ConvNext<sup>[24]</sup>、Swin Transformer<sup>[17]</sup>、Uniformer<sup>[25]</sup>, 比较计算量和参数量这两个指标。计算量和参数量越大,表示该模型的计算复杂度越高,而较低的计算量和参数量则表明模型具有较低的计算复杂度和较少的参数,不同时序依赖关系建模模型对比如表 2 所示。

表 2 不同时序依赖关系建模模型对比

模型	计算量/GFLOPs	参数量/ $10^6$
Poolformer	3.197	36.802
ConvNext	3.841	47.212
Swin Transformer	3.757	45.764
Uniformer	3.654	44.134

从表 2 中可以看出,相对其他 3 种模型而言,在编码器和解码器之间融合 Poolformer 模块能够取得最低的计算量和参数量,具有更快的训练速度和推理速度,计算资源消耗较少,便于模型的部署,在不同的预报任务中具有很好的鲁棒性。

### 2.5.2 比例因子 $\alpha$ 的影响

为了探讨时空重构单元中不同  $\alpha$  对模型效果的影响,本文选取 3 个不同的  $\alpha$  值进行对比。表 3 显示了不同  $\alpha$  在上海市雷达数据集中计算量和 CSI 指标上的对比。

表 3 不同  $\alpha$  因子值指标对比

Table 3 Comparison of indicators with different  $\alpha$  factor values

$\alpha$	计算量/GFLOPs	CSI
0.25	3.037	0.795
0.50	3.197	0.803
0.75	3.427	0.804

从表 3 中可以看出,随着  $\alpha$  值升高,计算量和 CSI 也逐渐增加,这主要是因为模型对输入张量在  $F$  通道按照比例因子  $\alpha$  拆分得到的上半部分特征应用分组卷积和逐点卷积操作,能够提取到更丰富的特征信息,但同时也会带来更高的计算复杂度,当  $\alpha$  取 0.50 时,模型在保持较高预报准确率的同时,计算效率也相对较高,二者达到一个很好的平衡,综合效果最佳。

## 3 结论

本文提出了一种融合时空重构单元和 Trans-

former 的雷达回波外推模型 SRU-Former。实验结果表明,本文在模型编码器和解码器中设计的时空重构单元模块 SRU,能够增强对雷达图像的时空特征提取能力,有效解决现有深度学习模型对复杂天气变化的特征表示不足的问题;在编码器和解码器之间融合的 Poolformer 模块,在捕捉序列依赖的同时,提升了模型的计算效率和鲁棒性。在未来工作中,将融合风场和温度等气象要素,进一步提高预报准确率。

## 参考文献:

- [1] 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.
- [2] 黄旋旋,朱科锋,赵坤. 改进后 TREC 外推方法在台风临近降雨预报中的应用[J]. 气象科学, 2017, 37(5): 610-618.  
HUANG X X, ZHU K F, ZHAO K. The improved TREC algorithm and its application to the precipitation nowcasting for landfall typhoon[J]. Journal of the Meteorological Sciences, 2017, 37(5): 610-618.
- [3] DEL MORAL A, RIGO T, LLASAT M C. A radar-based centroid tracking algorithm for severe weather surveillance: identifying split/merge processes in convective systems[J]. Atmospheric Research, 2018, 213: 110-120.
- [4] 吴剑坤,陈明轩. 基于雷达回波区域跟踪算法的临近预报技术进展[J]. 气象科技, 2018, 46(5): 899-909.  
WU J K, CHEN M X. Advances in nowcasting techniques based on radar echo area tracking algorithm[J]. Meteorological Science and Technology, 2018, 46(5): 899-909.
- [5] SHI X J, CHEN Z R, WANG H, et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting[J]. Advances in Neural Information Processing Systems, 2015, 2015: 802-810.
- [6] WANG Y B, WU H X, ZHANG J J, 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, 2017: 879-888.
- [7] 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[EB/OL]. (2018-04-17) [2024-10-16]. <https://doi.org/10.48550/arXiv.1804.06300>.
- [8] 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 spatio-temporal fusion neural network[J]. Acta Electronica Sinica, 2023, 51(9): 2526-2538.
- [9] 刘颖,杨鹏飞,张立军,等. 前馈神经网络和循环神经网络的鲁棒性验证综述[J]. 软件学报, 2023, 34(7): 3134-3166.  
LIU Y, YANG P F, ZHANG L J, et al. Survey on robustness verification of feedforward neural networks and recurrent neural networks[J]. Journal of Software, 2023, 34(7): 3134-3166.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]//Lecture Notes in Computer Science. Cham: Springer, 2015: 234-241.
- [11] 王善昊,胡志群,王福增,等. 基于 ConvLSTM 融合 RMAPS-NOW 数据的雷达回波外推研究[J]. 气象学报, 2024, 82(4): 554-567.  
WANG S H, HU Z Q, WANG F Z, et al. Extrapolation of radar echo based on ConvLSTM with fusion of RMAPS-NOW data[J]. Acta Meteorologica Sinica, 2024, 82(4): 554-567.
- [12] TREBING K, STAŃCZYK T, MEHRKANOON S. SmaAt-UNet: precipitation nowcasting using a small attention-UNet architecture[J]. Pattern Recognition Letters, 2021, 145: 178-186.
- [13] GAO Z Y, TAN C, WU L R, et al. SimVP: simpler yet better video prediction[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 3160-3170.
- [14] 李卫军,张新勇,高庚潇,等. 基于门控时空注意力的视频帧预测模型[J]. 郑州大学学报(工学版), 2024, 45(1): 70-77, 121.  
LI W J, ZHANG X Y, GAO Y X, et al. Video frame prediction model based on gated spatio-temporal attention[J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(1): 70-77, 121.
- [15] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[EB/OL]. (2017-07-12) [2024-10-16]. <https://doi.org/10.48550/arXiv.1706.03762>.
- [16] 方巍,齐媚涵. 基于深度学习的高时空分辨率降水临近预报方法[J]. 地球科学与环境学报, 2023, 45(3): 706-718.  
FANG W, QI M H. Precipitation nowcasting method with high spatio-temporal resolution based on deep learning[J]. Journal of Earth Sciences and Environment, 2023,

- 45(3): 706–718.
- [17] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: hierarchical vision Transformer using shifted windows [C]//2021 IEEE/CVF International Conference on Computer Vision. Piscataway: IEEE, 2021: 9992–10002.
- [18] BI K F, XIE L X, ZHANG H H, et al. Pangu-weather: a 3D high-resolution model for fast and accurate global weather forecast[EB/OL]. (2022–11–03) [2024–10–16]. <https://doi.org/10.48550/arXiv.2211.02556>.
- [19] DALMAZ O, YURT M, CUKUR T. ResViT: residual vision transformers for multimodal medical image synthesis [J]. IEEE Transactions on Medical Imaging, 2022, 41(10): 2598–2614.
- [20] YU W H, LUO M, ZHOU P, et al. MetaFormer is actually what you need for vision[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 10809–10819.
- [21] 魏明军, 王镛涵, 刘亚志, 等. 基于特征融合和混合注意力的小目标检测[J]. 郑州大学学报(工学版), 2024, 45(3): 72–79.
- WEI M J, WANG M H, LIU Y Z, et al. Small object detection based on feature fusion and mixed attention[J]. Journal of Zhengzhou University (Engineering Science), 2024, 45(3): 72–79.
- [22] LI J F, WEN Y, HE L H. SCConv: spatial and channel reconstruction convolution for feature redundancy [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2023: 6153–6162.
- [23] SMITH L N. Cyclical learning rates for training neural networks[C]//2017 IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE, 2017: 464–472.
- [24] LIU Z, MAO H Z, WU C Y, et al. A ConvNet for the 2020s[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2022: 11966–11976.
- [25] LI K C, WANG Y L, ZHANG J H, et al. UniFormer: unifying convolution and self-attention for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(10): 12581–1260.

## Radar Echo Extrapolation Algorithm Integrating Spatiotemporal Reconstruction Unit and Transformer

FANG Wei<sup>1,2,3,4</sup>, WANG Haoxi<sup>1</sup>

(1. School of Computer Science, Nanjing University of Information Science & Technology, Nanjing 210044, China; 2. School of Software, Nanjing University of Information Science & Technology, Nanjing 210044, China; 3. China Meteorological Administration Basin Heavy Rainfall Key Laboratory/Hubei Key Laboratory for Heavy Rain Monitoring and Warning Research, Wuhan Research Institute of Heavy Rain, China Meteorological Administration, Wuhan 430205, China; 4. State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing 100081, China)

**Abstract:** Addressing the limitations of existing deep learning-based radar echo extrapolation algorithms in spatiotemporal feature extraction and long-term dependency modeling, a novel SRU-Former model that integrated spatiotemporal reconstruction unit (SRU) and Transformer was proposed for radar echo extrapolation. Firstly, a newly designed SRU was introduced into the model's encoder and decoder to extract fine-grained spatiotemporal features from radar images via separation, transformation, and reconstruction. Secondly, a variant architecture model of Transformer, Poolformer, was introduced between the encoder and decoder, using global average pooling to replace the self-attention mechanism, thereby assisting the model in modeling highly dynamic radar sequences. Finally, SRU-Former was trained and tested on two meteorological radar datasets from Jiangsu Province and Shanghai City, respectively, and compared with current mainstream deep learning models. In the 2-hour extrapolation task, SRU-Former achieved optimal values in four metrics: *CSI*, *FAR*, *MSE*, and *SSIM*. Specifically, *CSI* was improved by 0.020 on the Jiangsu Province dataset and by 0.048 on the Shanghai City dataset. Experimental results showed that SRU-Former effectively improved model prediction accuracy, with more precise capture of strong echo regions and clearer detail textures in later extrapolation stages.

**Keywords:** deep learning; radar echo extrapolation; Transformer; spatiotemporal reconstruction unit; global average pooling