

# 基于时空分离的雷达降水图像时空降尺度研究

郑祥明 秦华旺 陈浩然

(南京信息工程大学电子与信息工程学院 南京 210044)

**摘要:**针对现有时空降尺度深度学习对雷达降水图像时空特征学习不充分的问题,提出了一种基于时空分离的三维深度学习模型 3DUST。该模型以 Unet3d 为核心架构,设计混合时空分离卷积单元增强降水图像局部时空特征的提取,并使用三维 Swin Transformer 补偿传统 Unet3d 编码器下采样造成的降水图像时空特征丢失问题,以提高时空降尺度预报的效果。通过法国气象局提供的公开数据集对模型进行检验和评估,结果表明:设计的混合时空分离单元具有较好的局部时空特征提取能力,基于时空分离的方法能够进一步提高时空降尺度预报效果。具体的,本文提出的 3DUST 模型使得 SSIM 和 PSNR 评价指标较对比模型分别提高了 5.2% 和 6.7%,且参数量减少了 3.2%。

**关键词:** 气象预报;深度学习;降尺度;降水图像;时空分离

**中图分类号:** P426.6 **文献标识码:** A **国家标准学科分类代码:** 170.15

## Spatio-temporal downscaling of radar precipitation images based on spatio-temporal separation

Zheng Xiangming Qin Huawang Chen Haoran

(School of Electronic and Information Engineering, Nanjing University of Information Science &amp; Technology, Nanjing 210044, China)

**Abstract:** To address the problem that the existing spatiotemporal downscaling deep learning methods are not enough to learn the spatiotemporal characteristics of radar precipitation images, a spatiotemporal separation based 3D deep learning model is proposed. The model takes Unet3d as the core architecture. A hybrid spatio-temporal separation convolution unit is designed to enhance the extraction of local spatio-temporal features of precipitation images, and a three-dimensional Swin Transformer is used to compensate for the loss of spatio-temporal features of precipitation images caused by traditional Unet3d encoder downsampling, so as to improve the effect of spatio-temporal downscaling forecast. The model was tested and evaluated through the open data set provided by METEO FRANCE. The results show that the designed hybrid spatio-temporal separation unit has a better ability to extract local spatio-temporal features, and the spatio-temporal separation based method can improve the spatio-temporal downscaling forecasting effect. Specifically, the 3DUST model proposed in this paper increased SSIM and PSNR evaluation indexes by 5.2% and 6.7%, respectively, and reduced the number of parameters by 3.2% compared with the comparison model.

**Keywords:** weather forecast; deep learning; precipitation image; downscaling; spatio-temporal separation

## 0 引言

雷达图像外推是目前降水预报的一个重要手段,但受雷达探测性能的限制,实际中往往难以获取高时空分辨率的雷达图像,此时需要对雷达图像进行降尺度<sup>[1]</sup>,即提高其时空分辨率。传统基于插值的降尺度方法容易滤除降水图像中的高频特征,极大制约了降水预报精度的提升。近年来,深度学习开始被引入降水图像的降尺度应用中。与传统插值方法相比,深度学习能够更加准确地捕获降水图像中的时空特征,从而有效提高各种复杂气象条件

下的降水预报精度。

基于深度学习的气象图像降尺度目前已成为气象预报领域中的一个重要研究方向,并在近年来取得了丰富的研究成果。Vandal 等<sup>[2]</sup>提出在卷积前将预报相关的变量与输入数据特征在信道维度上拼接,设计了相应的 DeepSD 网络,获得了较好的应用效果;Rodrigues 等<sup>[3]</sup>提出基于深度神经网络的降尺度模型,通过监督学习方法将多个不同天气模型的低分辨率预报数据作为输入,取得了优于对比标准的实验结果;Lu 等<sup>[4]</sup>针对传统插值方法处理图像时存在的问题,基于经典的 Unet<sup>[5]</sup>架构设计了优化的降尺度模

型。Höhlein 等<sup>[6]</sup>研究了基于 U-net 结构和残差结构的深度学习网络,应用于风速数据空间降尺度问题,并取得了优于对比模型的表现。Serifi 等<sup>[7]</sup>提出了基于多个深度卷积神经网络的深度学习架构,有效解决了多时空分辨率的数据重构问题。虽然近年来基于深度学习的气象图像降尺度研究已取得了较多研究成果,但现有成果仍然存在对气象图像时空特征学习不充分的问题,尤其对于时空特征极为丰富的雷达降水图像,特征提取不充分将会严重限制预报精度的提升。

气象图像降尺度与计算机视觉领域中的图像超分辨率十分相似,本质上都是在尽量保留图像特征的前提下提高其分辨率。Transformer<sup>[8]</sup>是一个利用注意力机制来提高模型训练速度的深度学习网络模型,并在图像超分辨率领域中得到广泛应用。Yang 等<sup>[9]</sup>提出了基于 Transformer 的图像超分辨率架构;Cao 等<sup>[10]</sup>提出了一种时空卷积注意力机制,用于对局部时空数据的信息挖掘,实现了 Transformer 在视频超分辨率应用中的改进;Swin Transformer<sup>[11]</sup>是 Transformer 的一种改进模型,早期被应用于图像识别任务中,后来通过使用时空因子分解计算全局自注意力来处理视频任务<sup>[12]</sup>。本文将基于 U-net3d<sup>[13]</sup>和 Swin Transformer 网络设计一种新型雷达降水图像降尺度模型。与视频超分辨率不同的是,降水区域的运动是非刚性的(即流体),而视频超分辨率主要关注刚体运动估计。因此,常规图像超分辨率的深度学习模型往往难以充分提取气象图像的时空特征。此外,气象数据的多源特性存在严重的时间偏差和时间稀疏问题,其特征提取比常规图像超分辨率更具挑战性。

针对深度学习模型在气象图像降尺度中时空特征学习不充分的问题,本文提出一种基于时空分离的三维深度学习模型 3DUST。该模型以 U-net3d 为基础架构,通过增加 Swin Transformer 和混合时空分离卷积等模块构建三维时空降尺度网络,以增强对气象数据时空特征的学习能力。本文中的 Swin Transformer 从常规的二维空间扩展到三维( $T$ 、 $W$  和  $H$ , 3D),用以捕获雷达降水图像的时空特征,并将其嵌入到时空分离的 3D U-net-2a 结构中作为解码单元的输入,然后与编码单元进行多尺度信息融合重构高分辨率降水图像。通过法国气象局提供的公开数据集对模型的验证结果表明,提出的时空分离降尺度网络可以有效增强雷达降水图像的时空学习和多尺度融合能力。本文的主要创新点如下:

1) 提出了一种基于时空分离的雷达降水图像时空降尺度深度学习模型 3DUST,该模型在 U-net3d 架构的基础上通过时空分离卷积模块和时空分离 Swin Transformer 模块实现对雷达降水图像空间和时间特征的同步提取,实现雷达降水图像的三维时空降尺度。

2) 针对雷达降水数据在 U 型编码器下采样中的时空特征丢失问题,通过使用时空分离的 Swin Transformer 模

块捕获降水数据的时间和空间特征,并嵌入到 3D U-net-2a 的解码器中作为解码器的输入,然后与编码单元进行多尺度信息融合重构高分辨率降水图像,增强对雷达降水数据时空特征的捕捉能力,有效提升降尺度效果。

## 1 深度学习降尺度模型

### 1.1 3DUST 网络结构

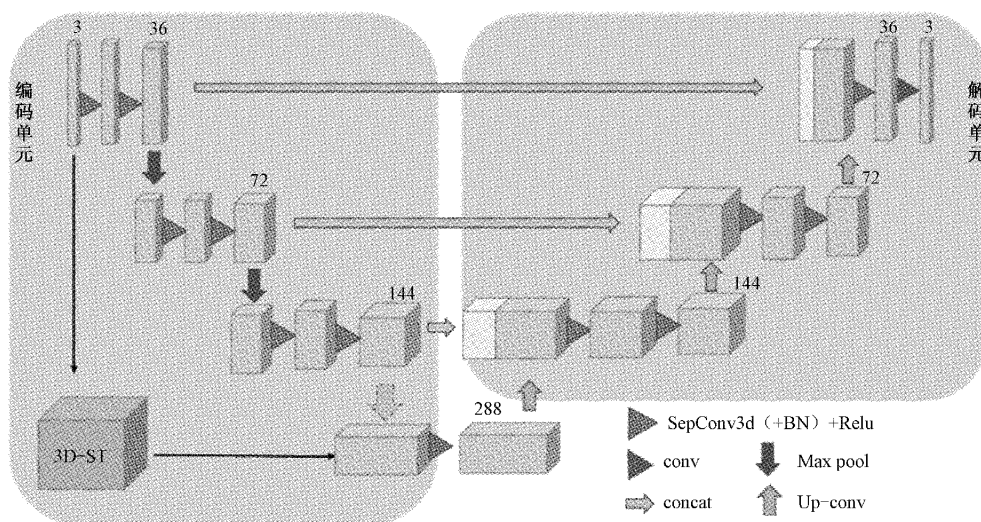
本文提出的 3DUST 深度学习降尺度网络模型由编码单元和解码单元两大部分组成,如图 1 所示。该网络以 U-net3d 为基础架构,通过设计时空分离的卷积模块替换标准的三维卷积模块构建时空分离的 3D U-net-2a 网络。此外,为优化 3D U-net-2a 编码器下采样造成的降水数据时空特征丢失问题,使用时空因子分解的三维 Swin Transformer 模块(如图 1 中 3D-ST 模块,详见 1.3 节)提取降水图像的时空特征,并将该特征嵌入到 3D U-net-2a 的解码器中作为解码器的输入。

在编码单元路径中,首先将间隔为 10 min 的数据作为输入,经三线性插值和上采样得到间隔为 5 min、尺寸为  $3 \times 400 \times 400$  的标签数据,然后分别传入 3D U-net-2a 和 3D-ST 两个路径进行时空特征的提取。在 3D U-net-2a 路径中,为增强模型的空间和时间感知能力,设计编码器从第 1 层 36 个特征图,经过 3 层卷积和下采样得到了 144 个特征映射,并且把每层的特征图残差连接到解码单元进行特征拼接和卷积融合。每一层采用三维时空分离卷积模块(如图 1 中 SepConv3d 模块,详见 1.2 节)和下采样单元( $2 \times 2 \times 2$  最大池化)组成。在 3D-ST 路径中,通过连续 4 个 3D-ST 层和下采样提取标签数据的时空特征,并将该特征嵌入到 3D U-net-2a 的解码器中作为解码器的输入,最大可能的将气象图像的内部特征和图像间的时空关系传递到解码单元,提高模型的时空捕获能力。

在解码单元路径中,首先将 3D-ST 提取的时空特征信息与编码器的多尺度残差信息融合后卷积再上采样,弥补 3D U-net-2a 下采样造成的时间和空间信息丢失。解码器的每一层由  $2 \times 2 \times 2$  的上卷积组成,在每个维度上跨 2 步,然后是 2 个三维时空分离卷积,每个卷积后面都有一个激活函数 ReLU。为了实现拼接统一性,本文在 3DUST 模型中根据 U-net3d 提取的特征图数量和分辨率配置 3D-ST 结构的超参数维度为 36,当补丁尺寸为 (1, 1, 1) 时,3D-ST 输出的特征数量为 288,分辨率为 (50, 50) 的特征图与 U-net3d 提取的特征图进行拼接融合和上采样处理,然后与前一层提取的特征图迭代直至重构出高分辨率雷达降水图像。最后一层使用  $1 \times 1 \times 1$  卷积将输出通道的数量减少到 3。

### 1.2 时空分离卷积模块

降水过程是一种流体运动,降水的位置和雨量与云有直接关系,云在流动过程中会在气温、风速和气压等多种气象因子作用下增长或消散,甚至形成新的云,具有较强的时空特性。多项研究表明,三维卷积神经网络具有较好的时



其中黄色三维立体柱表示特征映射,每层特征的通道数量在特征图上;3D-ST 表示 Swin Transformer 三维时空特征提取网络。

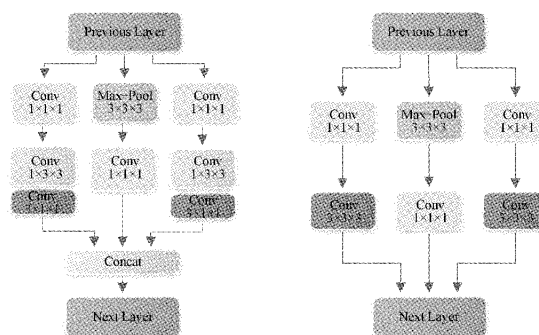
图 1 3DUST 的主要架构

空特征学习能力<sup>[14-15]</sup>。由于标准的 3D-CNN 计算成本相对较高,为解决深层次网络较难训练的问题,He 等<sup>[16]</sup>提出了一个残差学习模块,在此基础上 Tran 等<sup>[17]</sup>提出了基于 ResNet 架构的用于动作识别的时空卷积模型。在视频分类任务中,Xie 等<sup>[18]</sup>提出了时空分离的 S3D 网络,减少了参数量并且提高了模型的性能。在以上成果的启发下,本文根据降水时空特点设计三维混合时空分离卷积模块,如图 2 (a)所示。提出时空分离的卷积单元捕获降水雷达数据的时空特征,解决雷达降水产品的时空降尺度分析中时间和空间学习不充分的问题,以实现更精准的时空降尺度。

雷达降水产品具有很强的时空特征,为实现时间和空间特征的有效提取,同时减少模型的参数量,本文设计将空间和时间进行分离的方法。即将  $k_t \times k \times k$  结构的卷积模块替换为  $1 \times k \times k$  和  $k_t \times 1 \times 1$ ,其中  $k_t$  表示时间上的宽度, $k$  表示空间上的高度/宽度。首先,设计标准的 3D-CNN 多路径提取降水图像的时空特征,该模块分 3 条路径分别捕获雷达降水时空特征,如图 2(b)所示。然后,将标准的  $3 \times 3 \times 3$  三维卷积核替换为  $1 \times 3 \times 3$  和  $3 \times 1 \times 1$ ,在三维卷积中依次进行空间特征进行提取和时间卷积。同时,使用最大池化对当前层的时空特征进行最大池化,捕获雷达降水图像的细节边界特征。最后,将 3 条路径的特征拼接,实现 1 次时空特征的提取。通过将时空分离三维卷积模块替换标准的卷积模块,模型的参数量减少了 47.9%,同时性能指标仍有一定的提高。

### 1.3 3D Swin Transformer

通过利用移位窗口实现计算的层次化方案,可以解决将 Transformer 从自然语言应用到计算机视觉处理时遇到的多尺度和高分辨率问题,有效提高了效率<sup>[11]</sup>。Cao 等<sup>[10]</sup>基于 Swin Transformer 和 Unet 提出了 Swin-Unet 模型用



(a) 时空分离卷积 (b) 混合卷积模块  
图 2 3D 卷积单元结构

于医学图像分割任务中,并取得了较好表现。本文根据雷达降水图像的时空特征学习任务需要,在 Liu 等<sup>[12]</sup>研究的启发下,把 Swin Transformer 再扩展一个时间维度设计 3D-ST,如图 3 所示。与视频处理任务类似,本单元的输入图像尺寸为  $T \times H \times W \times 3$ ,其中  $T$  为 3,表示 3 个连续时间帧的气象图像。将图像按  $1 \times 4 \times 4 \times 3$  尺寸的 3D 补丁块处理,首先经过 Patch Partition 模块得到  $T \times \frac{H}{4} \times \frac{W}{4} \times 3$  48 个特征向量,然后 Linear Embedding 将这些特征线性投影到 48 个维度。在时间维度,已经实现了 10 min~5 min 的插值;在空间维度上,大小为 (3, 400, 400) 的标签数据应用 3D-ST 结构依次处理。

根据雷达降水时空降尺度需要,本文在基于非重叠 2D 窗口的多头自注意力模块基础上扩展一维时间变量,并把 3D-ST 块的多头自注意力模块设计为基于 3D 移动窗口的模块,详细结构如图 3 所示。为增强区域降水的时空联系,扩展了移位窗口以保证基于非重叠窗口自注意力的高效计算。对于连续两层,当输入的令牌数量为  $T' \times H' \times W'$  并且 3D 窗口的尺寸为  $P \times M \times M$  时,第 1



层会在规则模式下得到  $\frac{T'}{P} \times \frac{H'}{M} \times \frac{W'}{M}$  不重叠的 3D 窗口；第 2 层窗口的配置方式是在时间、高度和宽度等不同方向与前一层的自注意力模块通过  $\frac{P}{2} \times \frac{M}{2} \times \frac{M}{2}$  令牌处理，可以增强与前一层的非重叠 3D 窗口的联系，通过 3D 移位窗口的设计增加了窗口数量，在计算量不变的情况下提高了模型计算效率。将基于 3D 窗口的多头自注意力模块 (3DW\_MSA) 和移位窗口的多头自注意力模块 (3DSW\_MSA) 依次应用于图 3 中的两个连续 3D-ST 块，基于数学建模思想，该 3D-ST 单元可以用数学表达式表示为式 (1)~(4) 所示。

$$\hat{z}^l = 3DW\_MSA(LN(z^{l-1})) + z^{l-1} \quad (1)$$

$$z^l = MLP(LN(\hat{z}^l)) + \hat{z}^l \quad (2)$$

$$\hat{z}^{l+1} = 3DSW\_MSA(LN(z^l)) + z^l \quad (3)$$

$$z^{l+1} = MLP(LN(\hat{z}^{l+1})) + \hat{z}^{l+1} \quad (4)$$

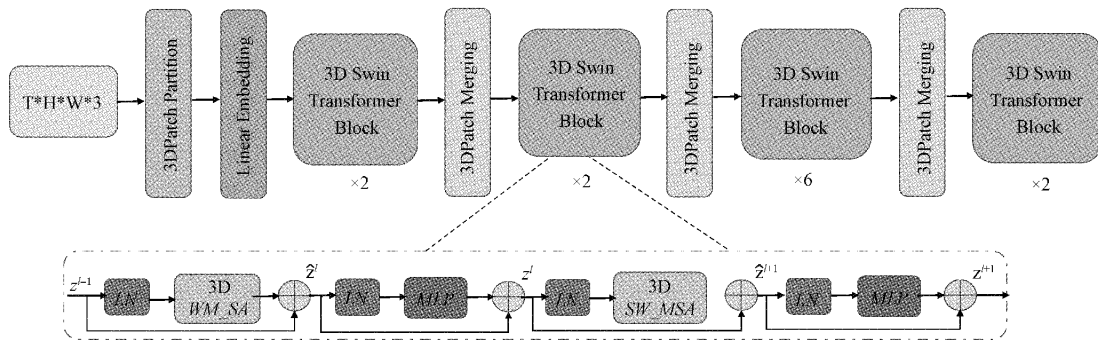


图 3 3D-ST 架构及 3D Swin Transformer 层的工作原理图

## 2 实验与分析

在法国气象局提供的公共降水反射率数据集上，测试了 3DUST 网络和 U-net3d 变体网络的模型性能及数据融合性能。同时，为了测试模型各模块的性能，在 3DUST 网络模型进行相关消融实验证明了该网络的先进性。

### 2.1 降水数据集

采用法国气象局<sup>[19]</sup>提供的公开数据集进行训练和测试，该数据集包含 2018 年 2 月~12 月每 5 min 的降水反射率数据、降水速率数据和反射测量高度数据，为了增强模型训练效率，对原始数据进行了处理<sup>[20]</sup>：

1) 排除降水产品中的缺失和未检测到的值，删除边缘缺失的数据；

2) 使用高通滤波器过滤降水范围低于 0.5 mm/h 的数据，表示低于 0.5 mm/h 的数据为没有降水<sup>[21]</sup>；

3) 由于降水不是日常现象，为了保证训练效果的质量，过滤掉了有效反射面积小于 20% 的数据。最后，随机选取了 7 000 组数据，其中 6 000 组作为训练集，1 000 组作为测试数据集。

其中，3DW\_MSA 和 3DSW\_MSA 分别表示基于 3D 窗口的规则和移位窗口划分的多头自注意力模块， $\hat{z}^l$  和  $\hat{z}^{l+1}$  分别表示第  $l$  层 3DW\_MSA 和 3DSW\_MSA 模块的输出，LN 和 MLP 表示层归一化和多层感知器模块。

大量的研究工作表明自注意力计算时考虑每个注意力头的相对位置偏差是有益的，因此为每个注意力头引入 3D 相对位置变量  $B (B \in \mathbb{R}^{P^2 M^2 M^2})$ ，自注意力的特征计算方法如式 (5) 所示，表示基于三维窗口的多头自注意力。

$$Attention(Q, K, V) = SoftMax\left(\frac{QK^T}{\sqrt{d}} + B\right)V \quad (5)$$

其中， $Q, K, V \in \mathbb{R}^{PM^2 \times d}$  分别表示查询、键和值矩阵， $d$  为查询维度和键的特征， $PM^2$  为 3D 窗口中的补丁数量，自相对位置沿各轴位于  $[-P+1, P-1]$  (时间) 或  $[-M+1, M-1]$  (高度或宽度) 范围内变量  $B$  的参数化的偏置矩阵  $\hat{B} \in \mathbb{R}^{(2P-1) \times (2M-1) \times (2M-1)}$ 。

### 2.2 评价标准

为了量化网络的效果，模型采用通用的图像质量评估指标，如：均方误差 (MSE)、峰值信噪比 (PSNR) 和结构相似性 (SSIM) 等。

在本文提出的模型测试中，设  $Y$  表示时空降尺度数据， $\bar{Y}$  表示原始高分辨率图像。当  $i \in \{1, \dots, n\}$  是一个时空区间的  $n$  个格点的索引，则 MSE 定义为：

$$MSE(Y, \bar{Y}) = \frac{1}{n} \sum_i |Y_i - \bar{Y}_i|^2 \quad (6)$$

通过 MSE 的大小来确定失真图像的失真程度，其中 MSE 越小表示时空降尺度的质量越高。将 MSE 转换为信号处理中常用的分贝 (dB) 形式，可得 PSNR (dB) 表达式：

$$PSNR = 10 \cdot \lg\left(\frac{MAX_Y^2}{MSE}\right) \quad (7)$$

其中， $MAX_Y^2$  为图像可能的最大像素值，由于将 MSE 放在了分母上，所以 PSNR 的数值越大表明图像质量越好。

MSE 和 PSNR 评价结果仅与当前位置的两个像素值相关，而忽略了图像所包含的一些视觉特征，尤其是图像

的局部结构信息。为综合评价模型的性能,选用 SSIM 对图像的局部信息进行评价,主要由亮度、对比度和结构 3 部分组成,分别为:

$$\begin{aligned}
 l(Y, \bar{Y}) &= \frac{2\mu_Y\mu_{\bar{Y}} + c_1}{\mu_Y^2 + \mu_{\bar{Y}}^2 + c_1} \\
 c(Y, \bar{Y}) &= \frac{2\sigma_Y\sigma_{\bar{Y}} + c_2}{\sigma_Y^2 + \sigma_{\bar{Y}}^2 + c_2} \\
 s(Y, \bar{Y}) &= \frac{\sigma_{Y\bar{Y}} + c_3}{\sigma_Y\sigma_{\bar{Y}} + c_3}
 \end{aligned} \tag{8}$$

其中,  $\mu_Y$  和  $\mu_{\bar{Y}}$  分别表示  $Y$  和  $\bar{Y}$  的平均值,  $\sigma_Y^2$  和  $\sigma_{\bar{Y}}^2$  分别表示  $Y$  和  $\bar{Y}$  的方差,  $\sigma_{Y\bar{Y}}$  表示  $Y$  和  $\bar{Y}$  的协方差,  $L$  表示像素值的取值范围,  $c_1 = (k_1L)^2$  和  $c_2 = (k_2L)^2$  为常数, 默认取  $k_1 = 0.01, k_2 = 0.03$ , 当  $c_3 = \frac{c_2}{2}$  时, SSIM 定义为:

$$\text{SSIM}(Y, \bar{Y}) = l(Y, \bar{Y}) \cdot c(Y, \bar{Y}) \cdot s(Y, \bar{Y}) \tag{9}$$

在实际训练中,系统设置参数  $k_1 = 5$  和  $k_2 = 10$ , 应用 Adam<sup>[22]</sup> 优化器,把学习率设置为  $1 \times 10^{-5}$ ,并在单个 NVIDIA Tesla V100 设备上训练。

### 2.3 3DUST 模型及其变体实验

为测试三维混合时空分离卷积模块对降水产品的时空特征提取和高分辨率图像重构的能力,本文将使用一个时空分离卷积模块的模型称为 3D UNet-a;使用两个时空分离深度卷积模块称为 3D UNet-2a;将 Unet3d 中  $3 \times 3 \times 3$  的卷积核直接替换为  $1 \times 3 \times 3$  和  $3 \times 1 \times 1$  的模型称为 3D UNet-b。在上述模型的基础上分别使用 3D-ST 单元替换原结构中的连接单元,称为: 3DUST-a、3DUST 和 3DUST-b。此外,本文根据 Unet3d 解码单元输入特征的维度和分辨率,当配置 3D-ST 的补丁尺寸为 (1, 2, 2) 时嵌入 3D UNet-b 中的模型称为 3DUST-b2。此时需要把 3D-ST 输出的特征数量为 288,分辨率为 (25, 25) 的特征图上采样为 (288, 50, 50) 再与 3D UNet-b 提取的特征图进行拼接融合和上采样处理,然后与前一层提取的特征图迭代直至重构出高分辨率降水图像。将训练数据集分别在 3DUST 模型及其变体模型上遍历 40 轮,得到了各模型变体的损失函数下降曲线,如图 4 所示。同时把各模型的数量、MSE 和 SSIM 等在测试集得到的评价指标记录在表 1 中。

从训练和测试结果中可以发现仅使用时空分离卷积模块的 Unet3d 模型变体的参数量均较少,其中 3DUnet-2a 效果最好,说明时空分离卷积具有良好的时空特征学习能力,并且适当增加卷积深度可以提高时空特征的提取能力。当进一步把 3D-ST 模块嵌入网络中,发现各模型的性能均有一定提升,表明 3D-ST 有助于弥补 U 型编码器下采样造成的特征丢失问题。综上,通过该对比实验结果可以说明以下问题:

1) 通过模型对比实验可以发现三维卷积网络的时空

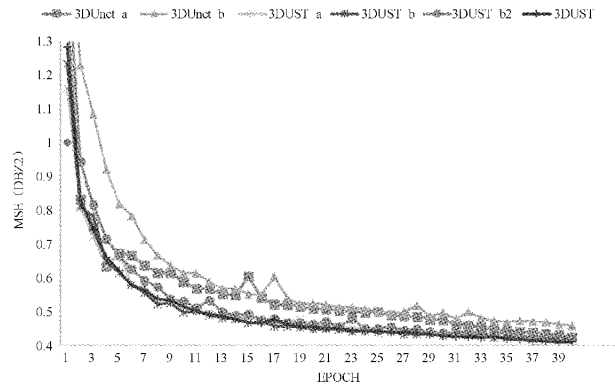


图 4 3DUST 模型及其变体的损失函数曲线

特征学习能力优于二维卷积网络,而时空分离的三维卷积神经网络可以在保障网络性能的前提下极大减少模型参数,且适当增加卷积深度可以提高时空特征的提取能力,即时空分离卷积网络更适合时空特征学习;

2) 3D-ST 具有良好的捕获时间和空间特征的能力,通过设计 3D-ST 作为编码器提取降水图像的时空特征并嵌入到 3DUnet-2a 模型中,可以有效补偿 3DUnet-2a 编码器在下采样过程造成的时空特征丢失问题,增强捕获降水图像上下文的联系,有利于模型重构高分辨率降水图像;

3) 通过变体模型的对比,可以发现基于时空分离的雷达降水图像时空降尺度网络模型 3DUST 的收敛更快、曲线平稳且在第 25 轮左右以后均方差表现最优。

表 1 测试集评价指标对比

模型	MSE	SSIM	参数量	PSNR
3DUnet-a	1.71	0.951	<b>3 224 197</b>	36.023
3DUnet-2a	1.639	<u>0.954</u>	<u>3 260 761</u>	<u>36.221</u>
3DUnet-b	1.708	0.952	3 535 669	36.091
3DUST-a	<u>1.536</u>	0.953	6 018 787	36.217
3DUST-b	1.568	0.949	6 330 259	36.103
3DUST-b2	1.728	0.953	6 330 367	36.020
<b>3DUST(ours)</b>	<b>1.483</b>	<b>0.958</b>	6 032 167	<b>36.325</b>

注:表中粗体表示性能最好,下划线部分表示性能次之。

### 2.4 模型对比分析

为了对比 3DUST 模型在时空降尺度问题中的性能指标,文章使用法国气象局提供的公开数据集分别在 Unet、Unet3d、TRU-NET<sup>[23]</sup>、CGAN<sup>[24]</sup>、时空分离模型 3D UNet-2a 以及 3DUST 模型上进行对比实验。将训练数据集分别训练 40 轮得到各模型训练损失函数曲线,如图 5 所示。在测试集上测试得到各模型的评价指标,如表 2 所示。

在图 5 中,从 Unet3d 与 3D UNet-2a 和 3DUST 的损失函数曲线对比中可以看出,Unet3d 损失曲线具有较低的拟合效果,但是表 2 展示的改进模型的 SSIM 和 PSNR 等指标均优于 Unet3d,表明改进模型具有更好的时空学习能

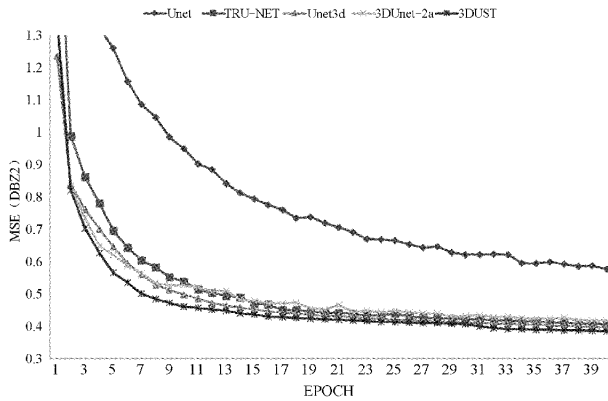


图 5 不同模型损失函数曲线

表 2 不同模型在测试集上的性能分析

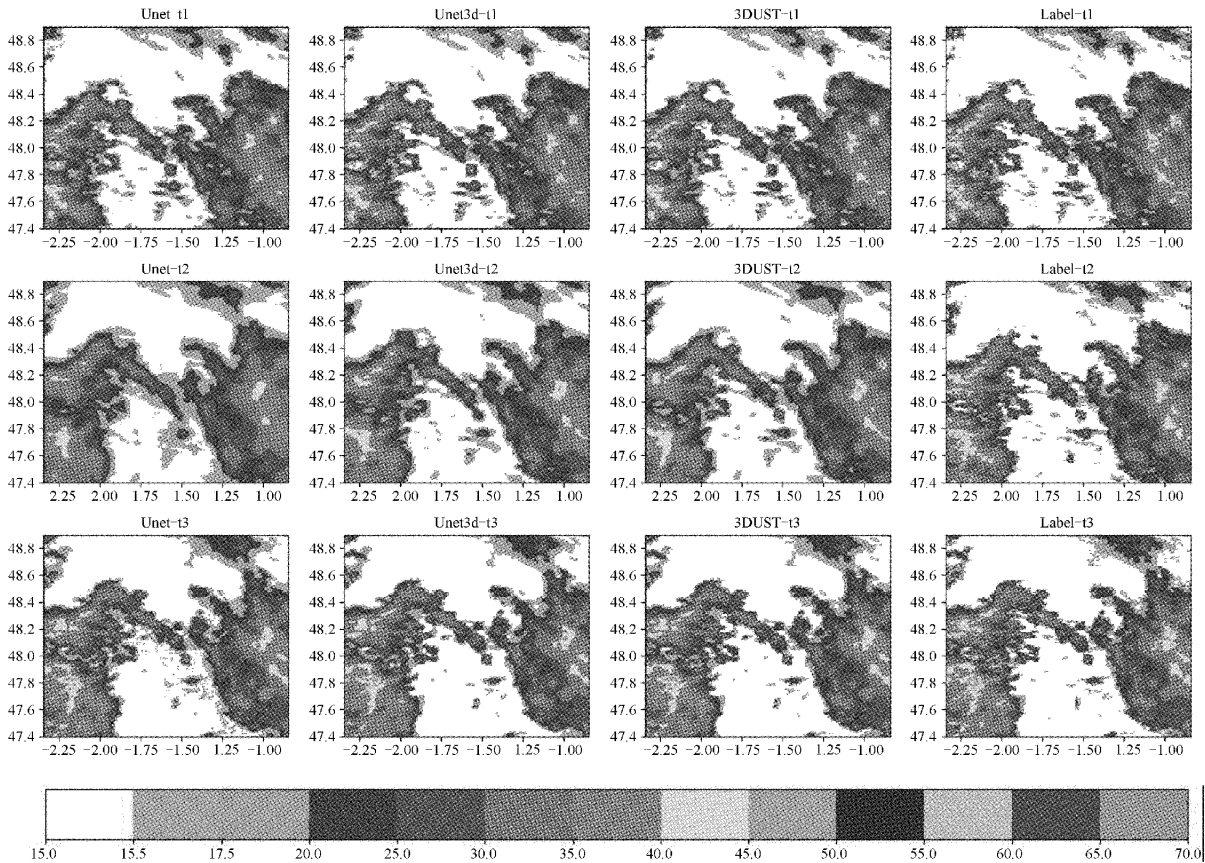
模型	MSE	SSIM	参数量	PSNR
Unet	2.054	0.931	17 273 676	35.281
TRU-NET	1.691	0.954	6 890 290	36.213
CGAN	1.836	0.951	6 960 321	36.017
Unet3d	1.769	0.953	6 256 189	36.084
3DUnet-2a	1.639	0.954	<b>3 260 761</b>	36.221
3DUST(ours)	<b>1.483</b>	<b>0.958</b>	6 032 167	<b>36.325</b>

注:表中粗体表示性能最好,下划线部分表示性能次之。

力。当观察对比模型及本文提出的改进模型时,可以看到采用混合时空分离的 3D UNet-2a 模型在参数量减少了 47.9%的情况下,SSIM 和 PSNR 指标仍优于 Unet3d 模型。在进一步优化的 3DUST 模型中,SSIM 和 PSNR 指标较 Unet3d 模型分别提高了 5.2% 和 6.7%,且参数量较少,表明 3D-ST 模块有助于优化 3D UNet-2a 在编码器下采样过程造成的特征丢失问题。

为了比较实际的降尺度效果,本文分别绘制了 Unet、Unet3d 及 3DUST 模型的降尺度雷达反射率图像及真实高分辨率图像,如图 6 所示。其中  $t_1$  和  $t_3$  时刻的图像为各模型输入的低分辨率图像经空间降尺度得到,而  $t_2$  时刻的图像为各模型时空降尺度重构的高分辨率图像,是课题的时空降尺度目标图像。从图中观察到 3 个模型的时空降尺度预报均取得了良好的表现,其中 3DUST 数据颜色分层更清晰,边缘细节的重构效果较对比模型更丰富,也更接近于真实高分辨率图像。

为更直观观察模型时空降尺度对比的效果,本文绘制了 Unet、Unet3d 和 3DUST 3 个模型  $t_2$  时刻时空降尺度雷达反射率图像与真实高分辨率图像的差值图,如图 7 所示。从预报的反射率图像和真实高分辨率图像的差值效果上可以更直观地观察到 3DUST 网络预报的数据散点较

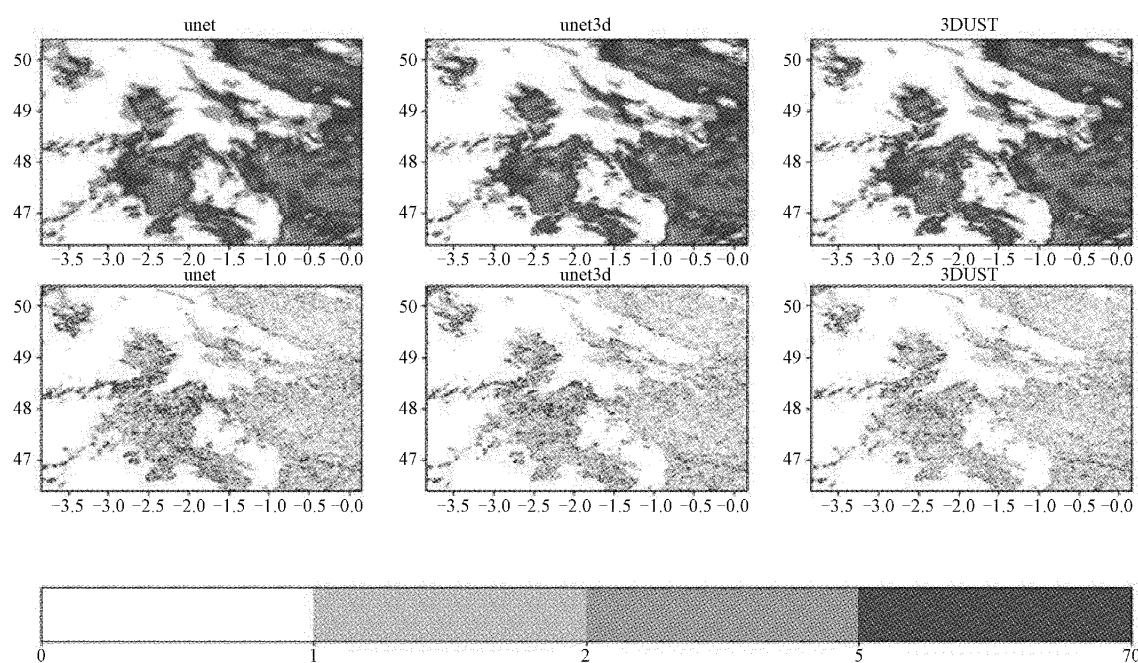


注:彩色最佳,黑白状态下可横向比较,即同一  $t$  时刻比较图像的纹理分布,以区分不同模型的时空降尺度预报效果。

其中,纹理分布与 Label 越接近,预报效果越好

图 6 3 个模型的雷达反射率降尺度图像及真实高分辨率图像对比





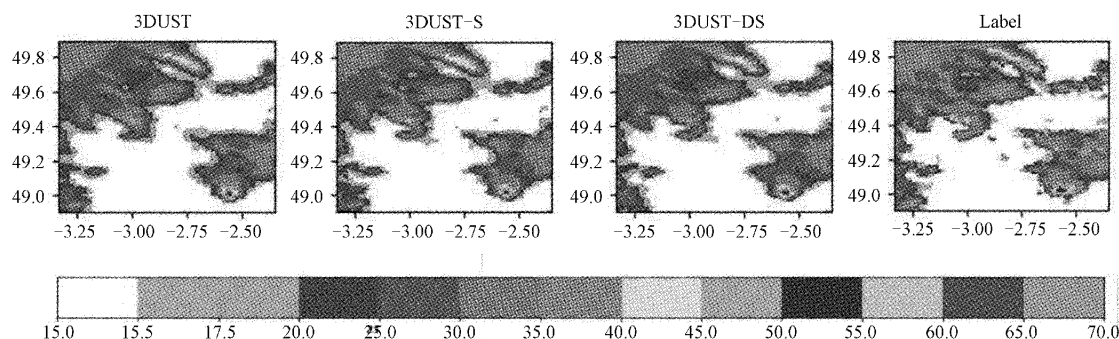
注：彩色最佳，黑白状态下颜色越深表示误差越大  
图 7 时空降尺度雷达反射率图像与真实图像的差值图

对比模型更稀疏，重构的效果更接近真实的数据效果。这个结果表明采用混合时空分离和 Swin Transformer 的三维网络模型对雷达降水图像时空特征具有更好的学习能力。

### 2.5 气象数据的融合验证

在上述 3DUST 模型消融实验中仅使用了降水反射率数据，而在气象降尺度领域中，地形、植被、经度、纬度等其他类型数据往往也会影响降尺度效果。本文采用的数据集也包含了降水速率和反射测量高度等数据，因此本文根据模型和数据集实际选择地形、海拔和测量高度等影响因子融合到降水反射率数据中进行降尺度预报，分别使用 Unet3d 变体模型和 3DUST 变体模型进行数据融合验证。为直观观察气象数据融合的效果，首先绘制了 3DUST、3DUST-S 和 3DUST-DS 在  $t_2$  时刻时空降尺度雷达反射率图像与真实高分辨率图像的对比，如图 8 所示。其实验步

骤是，首先，分别提取地形和海拔等静态影响因子的特征，合并后与降水反射率数据融合后输入模型进行验证，其中 3D U-Net-S 和 3DUST-S 分别为两个模型加入静态影响因子后的损失函数曲线。与静态数据不同，降水反射率数据本身就是动态数据。所以，本文把反射率和测量高度等随预报过程不断变化的动态变量以信道叠加的形式进行特征提取，与静态数据特征拼接后分别输入模型进行验证，得到损失函数曲线 3D U-Net-DS 和 3DUST-DS，如图 9 所示。从图 8 和图 9 中综合评价，当对比模型中融入静态影响因子后，损失函数收敛明显加快，曲线更加平稳，表明静态影响因子对降水的影响较大；当进一步融入动态变量后，对模型也产生了积极影响。因此，气象数据的融合实验也进一步验证了 3D-ST 模块的性能以及 3DUST 模型的有效性。



注：彩色最佳，黑白状态下通过比较图像的纹理分布，以区分不同模型的时空降尺度预报效果。  
其中，纹理分布与 Label 越接近，预报效果越好

图 8 数据融合效果图

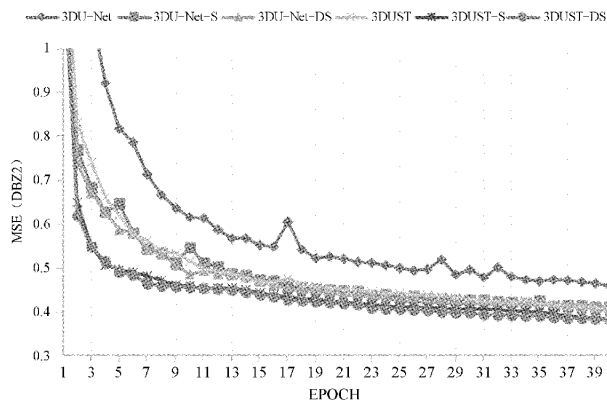


图 9 数据融合后的损失函数曲线

### 3 结 论

本文研究了深度学习在雷达降水图像时空降尺度中的应用问题。在二维卷积网络的基础上扩展时间维度设计三维卷积网络,并提出时空分离的三维卷积网络用于时空特征的捕获,在保证网络性能的前提下极大减少了模型参数。为优化 U 型编码器下采样造成的时空特征丢失问题,增加 3D-ST 模块进一步提高降水数据的时空特征捕获,并嵌入到 3D UNet-2a 的解码器中构建有利于雷达降水图像时空特征学习的深度学习网络。利用法国气象局提供的雷达降水反射率数据对 Unet、Unet3d、TRU-NET、CGAN 和 3DUST 及其变体网络结构进行测试,证明了采用时空分离的 3DUST 模型可以较好实现时空降尺度降水预报。最后,将影响降水的静态影响因子和动态变量分别融入输入数据中并在 Unet3d 模型和 3DUST 模型测试,进一步验证了模型的有效性。

### 参考文献

- [1] 董成,赵国良,王轶霞. 降雨强度对非饱和土降雨入渗的影响[J]. 中国测试, 2022, 48(7): 157-162.
- [2] VANDAL T, KODRA E, GANGULY S, et al. DeepSD: Generating high resolution climate change projections through single image super-resolution[C]. Proceedings of the 23rd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining, 2017: 1663-1672.
- [3] RODRIGUES E R, OLIVEIRA I, CUNHA R, et al. DeepDownscale: A deep learning strategy for high-resolution weather forecast [C]. 2018 IEEE 14th International Conference on e-Science, IEEE, 2018: 415-422.
- [4] LU Z, CHEN Y. Single image super-resolution based on a modified U-net with mixed gradient loss [J]. Signal, Image and Video Processing, 2022: 1-9.
- [5] ROSENBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, 2015: 234-241.
- [6] HOHLEIN K, KERN M, HEWSON T, et al. A comparative study of convolutional neural network models for wind field downscaling[J]. Meteorological Applications, 2020, 27(6): e1961.
- [7] SERIFI A, GUNTHER T, BAN N. Spatio-temporal downscaling of climate data using convolutional and error-predicting neural networks [J]. Frontiers in Climate, 2021, 3: 656479.
- [8] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. Advances in Neural Information Processing Systems, 2017, 30.
- [9] YANG F, YANG H, FU J, et al. Learning texture transformer network for image super-resolution [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 5791-5800.
- [10] CAO H, WANG Y, CHEN J, et al. Swin-unet: Unet-like pure transformer for medical image segmentation [C]. Computer Vision-ECCV 2022 Workshops, 2023: 205-218.
- [11] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 10012-10022.
- [12] LIU Z, NING J, CAO Y, et al. Video swin transformer [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 3202-3211.
- [13] ÇIÇEK Ö, ABDULKADIR A, LIENKAMP S, et al. 3D U-Net: learning dense volumetric segmentation from sparse annotation [C]. International Conference on Medical Image Computing and Computer-assisted Intervention, Springer, Cham, 2016: 424-432.
- [14] 晋良念,朱皓,王燃. 一种毫米波雷达快速高分辨三维点云成像方法 [J]. 仪器仪表学报, 2022, 43(12): 158-167.
- [15] 李哲,卢健,杨腾飞. 基于深度学习的三维点云重建方法 [J]. 国外电子测量技术, 2021, 40(3): 1-5.
- [16] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [17] TRAN D, WANG H, TORRESANI L, et al. A closer look at spatiotemporal convolutions for action recognition [C]. Proceedings of the IEEE conference on



- Computer Vision and Pattern Recognition, 2018: 6450-6459.
- [18] XIE S, SUN C, HUANG J, et al. Rethinking spatiotemporal feature learning: Speed-accuracy trade-offs in video classification [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 305-321.
- [19] LARVOR G, BERTHOMIER L, CHABOT V, et al. An open reference weather dataset by Météo-France [J]. MeteoNet, 2020.
- [20] CHEN H, QIN H, DAI Y. FC-ZSM: Spatiotemporal downscaling of rain radar data using feature constrained zooming slow-mo network[J]. MeteoNet, Frontiers in Earth Science, 2022: 756.
- [21] SHI X, CHEN Z, WANG H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting [J]. Advances in Neural Information Processing Systems, 2015:802-810.
- [22] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. ArXiv Preprint, 2014, ArXiv:1412.6980.
- [23] ADEWOYIN R A, DUEBEN P, WATSON P, et al. TRU-NET: A deep learning approach to high resolution prediction of rainfall[J]. Machine Learning, 2021, 110: 2035-2062.
- [24] WANG J, LIU Z, FOSTER I, et al. Fast and accurate learned multiresolution dynamical downscaling for precipitation[J]. Geoscientific Model Development, 2021, 14(10): 6355-6372.

### 作者简介

郑祥明, 博士研究生, 主要研究为深度学习, 气象图像处理。

E-mail: zymbzxy@163.com

秦华旺(通信作者), 博士, 教授, 博导, 主要研究方向为深度学习。

E-mail: qin\_h\_w@163.com

陈浩然, 硕士研究生, 主要研究方向为深度学习, 气象降尺度。

E-mail: 781241941@qq.com