

雷达回波图像预测

《大数据分析(B)》课程项目报告

龚傲凡 谢笛 梁婷

1 选题背景与文献调研

1.1 选题背景

临近预报的概念由 Browning 和 Collier 于 1989 年提出,是指通过时空序列外推方法实现高精度的短时定量天气预报,其核心是对强对流天气的预报^[1]。世界气象组织则将临近预报定义为“基于当前天气的详尽信息,通过外推方法实现 0-6 小时的天气预测”^[2]。临近预报涵盖降水、大风、冰雹、雷暴、雾霾等众多气象要素,降水临近预报是其中一个重要的研究方向。降水临近预报由于能提供降水强度和分布的短期预报,尤其是强降雨的预报,可以起到重要的紧急预警作用,其预报结果可以指导洪水预警、机场调度、市民出行等等,为人类生产生活带来显著的社会效益^[3]。

随着遥感技术的发展,天气雷达已经成为降水临近预报的重要工具。现代天气雷达基于多普勒效应,通过回波强度和反射率推算降水强度和分布信息,能为降水临近预报提供充分可靠的观测数据。基于天气雷达的外推法可以利用雷达回波时序数据,推算一定时段内降水天气系统的变化情况,并按此规律推测下一时刻的天气系统状态。现代天气雷达扫描间隔短,观测范围广,空间分辨率高,能获得高精度降水强度和时空分布信息,能够有效应用在降水临近预报任务中。然而实际情况中的雷达回波非常复杂,现有外推算法难以完全掌握规律,其准确性有待提高^[4];而且外推法的有效时间较短,通常仅为 1 h,最高不超过 3 h^[5]。

随着数据获取和存储技术的迅猛发展,人类已经进入数据量高速膨胀的大数据时代,如何高效利用数据获取信息成为各个行业的研究前沿,深度学习技术的蓬勃发展为降水临近预报问题的解决提供了无限可能。深度学习能够自动设计特征提取方法,捕捉数据中高度复杂的规律,并从中提取有效信息,能够减少人为的特征提取工作,降低对专业知识的依赖,从而实现预测能力的飞跃,这些优势

使深度学习在众多领域获得了广泛应用^{[6][7]}。卷积神经网络 (CNN) 和循环神经网络 (RNN) 是两类常用的深度学习方法,它们分别擅长处理图像/空间信息和语言/时间序列。对于同时集合了空间和时间信息的时空序列数据,已有研究采用 CNN 和 RNN 结合的方法并验证了其有效性^{[8][9]}。应用 CNN 和 RNN 等前沿深度学习技术,利用天气雷达提供的海量高分辨率数据,可以训练神经网络学习雷达回波的演变规律,进而推测天气系统的未来状态。如何运用深度学习技术实现准确有效的降水临近预报,具有重要的研究意义。

1.2 文献调研

人工神经网络是在现代生命科学研究的启发下提出的一种模拟神经系统处理和记忆信息的计算模型,其突出优势是强大的非线性自适应信息处理能力,是人工智能领域的有力工具^[10]。深度学习是基于深层人工神经网络的架构,对数据特征进行自动学习的研究方法,近年来已经在众多领域的大量任务中被广泛采用,如计算机与人工智能领域的人脸识别^[11]、语音识别^[12]、自动驾驶^[13]等,生命科学领域的医学影像诊断^[14]和分子结构推断^[15],以及金融、通讯、法律、制造业、自然资源等行业^{[7][10]}。

卷积神经网络 (CNN) 是一类以卷积运算和深度结构为核心的前馈神经网络,擅长处理具有网格结构的数据,如一维的信号数据、二维的图像数据和三维的视频数据等等,能够胜任分类、识别、回归等任务^[10]。其中,如何采用 CNN 挖掘具有时空结构的数据信息是近年来深度学习的热点之一。Ji 等提出 3D CNN 算法,可以从视频中捕捉人类动作信息^[16]; Tran 等设计出新的 C3D 网络,并将 3D CNN 的任务从视频分析推广到带有时空结构数据的特征提取问题^[17]; Qiu 等提出 P3DResNet 网络,采用精心设计的 2D 网络从 3D 视频信息中提取人类动作特征,缩短了神经网络的训练时间^[18]。在气象科学领域, Pan 等采用 CNN 实现区域降水量预测^[19], Klein 等采用动态 CNN 实现基于雷达数据的短期天气预报^[20], Ayzel 等采用全卷积网络实现雷达降水临近预报^[21]。

循环神经网络 (RNN) 是一类以循环结构为核心的前馈神经网络,擅长处理具有时序结构的数据,如自然语言序列和信号时间序列等等。单独的 RNN 只能提取一维时序信息,而 CNN 可以提取二维空间信息,将 CNN 和 RNN 结合便可以实现三维时空序列数据的特征提取和分析。如 Miao 等用 CNN 和 LSTM 的组合网络预测日降水量,预测效果与传统数值模式 ECMWF 相比有明显提高^[22]; Wu、Yang 等用 CNN 和 LSTM 建立了卫星和地面测站数据的融合模型,提高了中国 TRMM 原始降水资料的准确性^[23]。

CNN 和 RNN 组合网络的时空特征提取能力有限,为此 Shi 等提出了一种结

合方式——ConvLSTM，将 CNN 中的卷积运算纳入 LSTM 单元中，实现了 CNN 和 RNN 的底层结合，并证明了这种结构在降水临近预报任务的有效性^[24]。在 ConvLSTM 的启发下，研究者开发出一些融合 CNN 和 RNN 的底层结构及相应的网络组织结构，如基于 ConvLSTM、ConvGRU、TrajGRU 单元的 Encoder-Forecaster 结构^[25]，鲁棒性和计算效率更优的 dec-seq2seq 结构^[26]，具有动态卷积层和自循环网络的 RDCNN^[27]，兼具时间和空间记忆力机制的 PredRNN、PredRNN++结构^{[28],[29]}，以及融合多层高度雷达信息的 MLC-LSTM^[30]等等。

与传统雷达回波外推方法相比，基于深度学习的降水临近预报方法具有如下优势：第一，能够充分利用大量气象数据进行训练，从大规模数据集中学习天气系统的发展规律，数据利用率高；第二，能够学习到传统临近预报法难以捕捉的高维抽象特征和复杂非线性变化规律，特征提取能力强；第三，能够利用成熟的开源框架和 GPU 硬件显著加速训练和预报过程，计算效率高；第四，能够通过不同的网络结构、训练参数、数据集影响预测效果，这些参数结构以及数据处理方式都具有很大的优化和修改空间，优化潜力大。因此，基于深度学习的降水临近预报方法值得深入研究。

2 数据获取与数据处理

2.1 数据获取

实验采用 2017 年深圳气象局和阿里云联合举办的“天池大赛”提供的雷达回波数据集——CIKM 2017。该数据集包含的雷达回波反射率和中心格点的降水量，共有 14000 个样本，每个样本包含 0.5 km、1.5 km、2.5 km 和 3.5 km 四个高度的雷达反射率灰度图序列，并预先进行了线性变换处理。图像的像素值表示某个空间网格处的雷达反射率的线性变换值，取值范围为 0 至 255 的整数。图像尺寸为 101 km×101 km，空间分辨率为 1 km，序列时长为 90 min，时间分辨率为 6 min。数据格式为 txt 文本。如图 2.1 为一组样本在 3.5km 高度上的序列，2.1(a)为样本的第一张图像，2.1(b)为样本的所有图像（共 15 张）。为了方便展示，此处进行了色彩转换，实际数据为灰度图。

2.2 数据处理

为了在该数据集上训练神经网络，需要将数据处理成神经网络容易接受的数据形式。数据处理步骤如下：首先由于 0.5km 高度的数据中缺测值较多，扔掉这一高度的数据，将 txt 格式转换为方便 Python 处理也便于存储读取的 npz 高维数组。数组的维度为 $15 \times 14000 \times 3 \times (101 \times 101)$ ，各个维度分别表示序列长度、样

本数、通道数和空间尺寸。考虑到系统内存限制，将 14000 个样本平均切分为 140 份，并储存在不同的 npz 数组中。

在用神经网络处理图像数据时，需要将像素值转换到一个合适的区间内，这样可以使权重水平整体适中，防止训练过程中出现梯度消失和梯度爆炸现象。图 2.2 给出了数据集像素值的整体分布特征。像素值整体呈现偏态分布，并且稀疏程度较高（含有大量零值）。考虑到像素值单位相同，因此可以对像素值整体进行缩放。由于 180 以上的像素点很少，可以认为是噪音，以 180 为阈值进行归一化，将原始数据集像素值映射到[0, 1]区间中，如式 2-1 所示：

$$\tilde{Z} = \frac{Z}{180}, \tilde{Z} \in [0, 1] \quad (2-1)$$

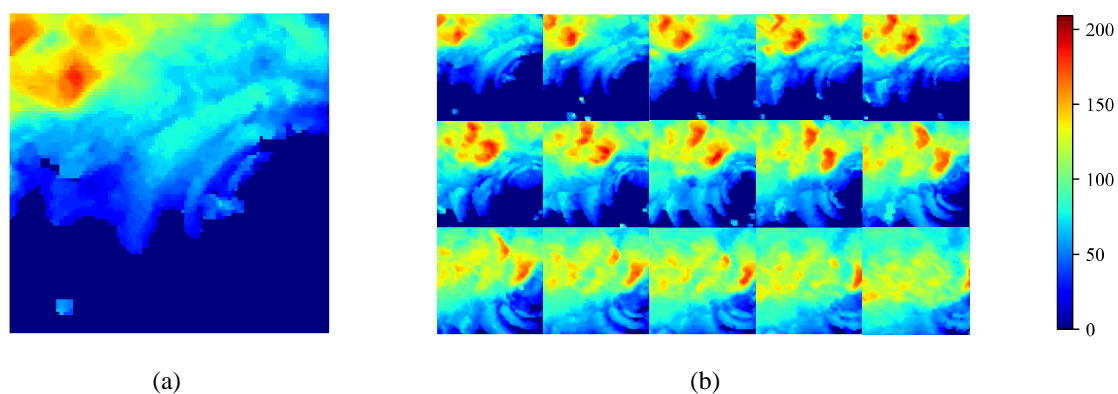


图 2.1 雷达回波图像序列样本 (3.5 km) 示例：(a)单张图像；(b)图像序列

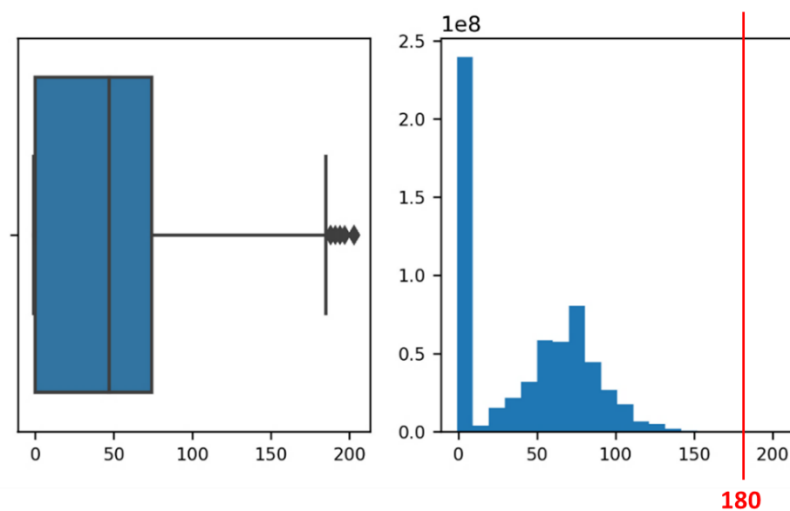


图 2.2 数据集像素值的整体分布特征

完成数据处理后,由于时空序列外推任务属于自监督学习,需要将数据集划分为训练集、验证集和测试集。网络在训练集上通过梯度的反向传播更新参数,完成训练过程,在验证集上调整超参数后,再在测试集上检验网络性能。将 14000 个样本按 5:1:1 的比例进行随机划分,即训练集、验证集、测试集分别包含 10000、2000、2000 个样本。

由于样本序列长度为 15,时间分辨率为 6 min。为了满足降水临近预报的业务需求,将前 5 帧作为输入,未来 10 帧作为输出,即用第-24 min 至第 0 min 的雷达回波图像预测第 6 min 至 60 min 的雷达回波图像。

3 网络构建与实验验证

3.1 网络构建

3.1.1 ConvLSTM

LSTM 单元擅长处理时间序列,也能处理空间信息,例如可以将图像的每一行看作时间序列在每个时间步的值。如果用 LSTM 处理时空序列,需要在空间维度和时间维度分别建立不同的 LSTM 模型,不仅网络参数十分冗杂,而且由于时空序列的点与点具有极强的时空关联性,维度单一的 LSTM 很难刻画这种复杂的时空特征。

为了提高 LSTM 对空间信息的捕捉能力,Shi 等提出了 ConvLSTM 单元^[14]。该单元是 CNN 与 LSTM 在底层结合的深度学习模块,用 CNN 中的卷积代替 LSTM 门控机制中的矩阵乘法,网络中隐藏状态之间的信息传递方式由全连接替换为卷积,从而在维持良好时序处理功能的同时,显著增强单元的空间特征刻画能力^[14]。图 3.1 展示了基于 ConvLSTM 的网络和基于 LSTM 的网络的差别,ConvLSTM 具有的卷积结构使得基于 ConvLSTM 的网络用卷积代替矩阵乘法,获得时空特征提取能力,可以接受具有时空结构的三维输入,而基于 LSTM 的网络只能接受一维输入或向量化的二维输入。如图 3.2 所示,ConvLSTM 可以处理三维张量,隐变量 C 和 H 由一维向量或二维矩阵变成了三维张量。

3.1.2 Encoder-Forecaster

Encoder-Decoder 可以处理序列长度不定的一维信息,图 3.3(a)展示了一种用 Encoder-Decoder 进行时空序列外推的方案。但是,该结构在时空序列外推任务中存在显著的局限性。一方面,Encoder-Decoder 的空间信息流向单一,在空间维度上缺少高层抽象信息向低层特征空间的传递过程,导致部分空间信息的丢失;另一方面,由于处理空间信息往往会引入卷积层和采样层,Encoder-Decoder 缺乏卷积层和采样层的参数共享机制,使得参数冗杂,计算效率低,而且梯度传播

路径长，训练不稳定。

为了适应时空序列外推任务，Shi 等提出 Encoder-Forecaster 结构，以改进 Encoder-Decoder 的不足^[25]。如 3.3(b)所示，Encoder-Forecaster 在 Encoder-Decoder 的基础上做出几点修改：第一，引入采样层，获取不同尺度的时空特征；第二，Forecaster 部分倒置了 Decoder 的信息流向，增强了高层特征空间对底层特征空间参数更新的影响和低层特征空间对预测的影响。

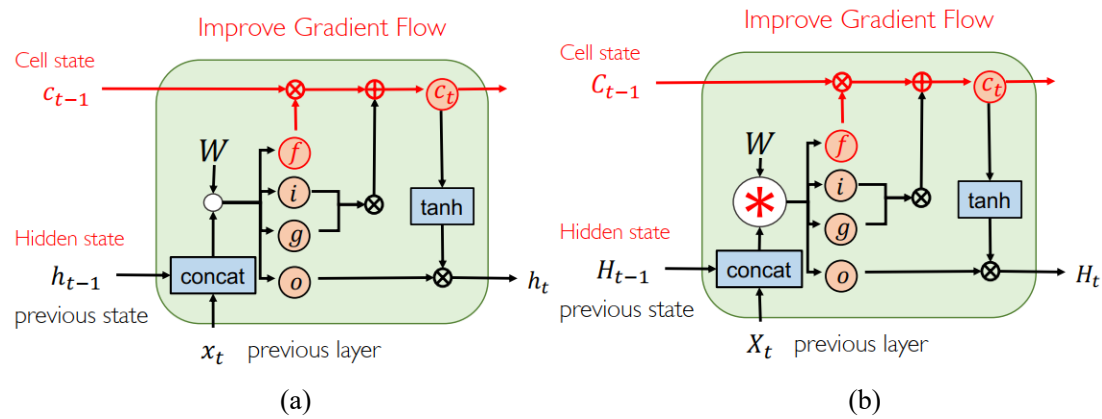


图 3.1 LSTM (a) 与 ConvLSTM (b) 对比

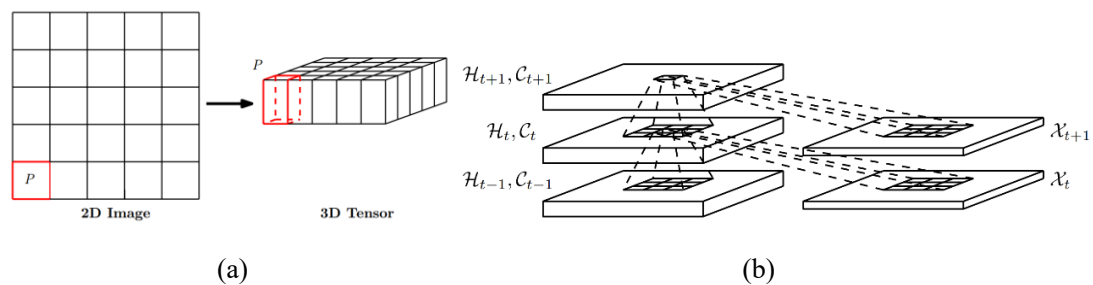


图 3.2 (a) 隐变量由 2D 向 3D 的转换, (b) ConvLSTM 序列

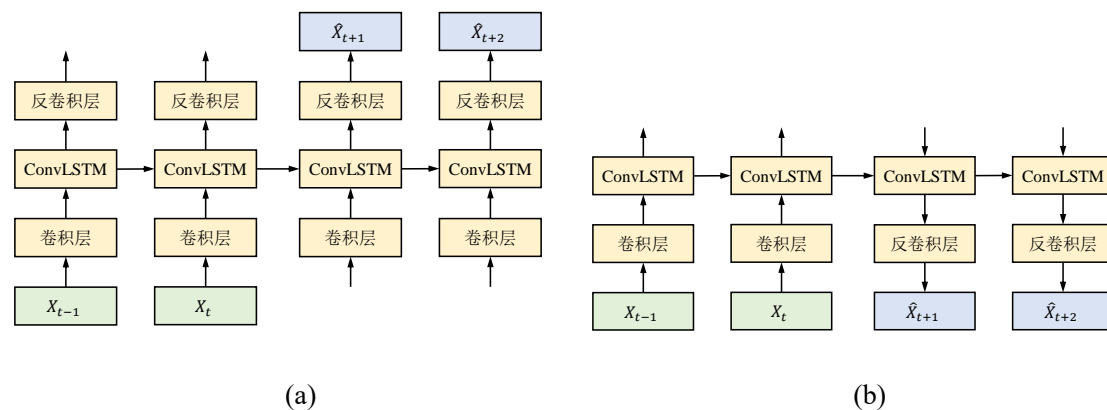


图 3.3 Encoder-Decoder (a) 和 Encoder-Forecaster (b) 对比

3.1.3 采样层

下采样层采用步长为 2 的卷积。对于上采样层，反卷积是常用的上采样方法。反卷积又叫转置卷积，其原理是先通过填充零值将小尺寸图像恢复到原始尺寸，转置卷积核后再进行卷积操作。实验中采用步长为 2 的反卷积，实验发现，用反卷积会出现 checkerboard artifacts (图 3.4(a))。Odena 等解释了使用反卷积产生此现象的原因，并建议用双线性插值卷积替代反卷积^[30]，我们经过尝试发现确实有效。将上采样层改为双线性插值卷积(图 3.5(b))，效果如图 3.4(b)，checkerboard artifacts 的问题得到解决。

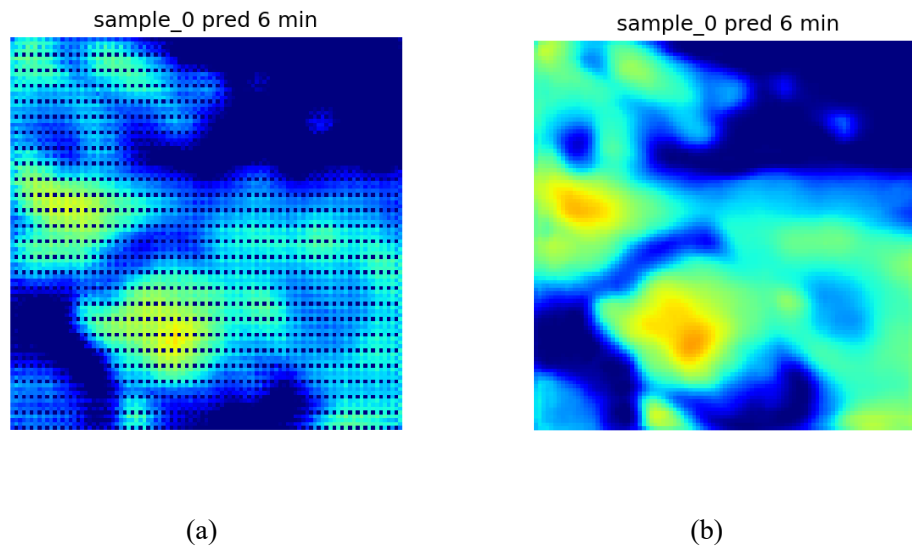


图 3.4 (a) Checkerboard artifacts; (b) No checkerboard artifacts

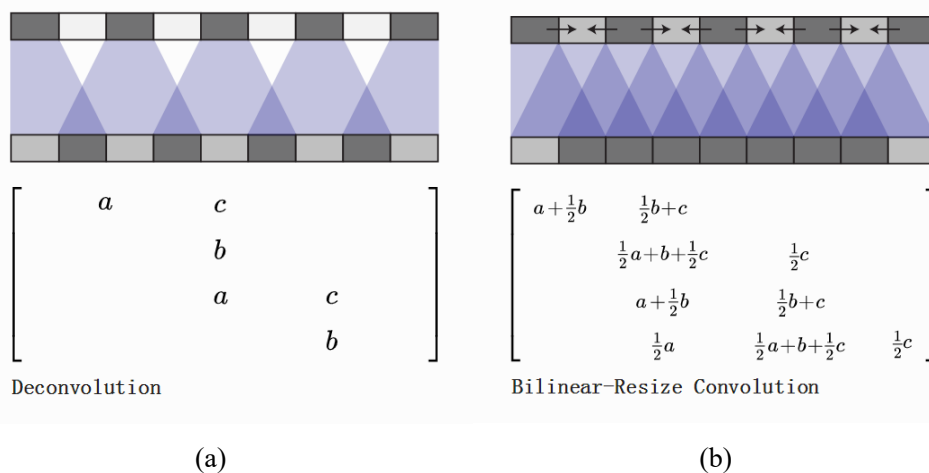


图 3.5 反卷积 (a) 与双线性插值卷积 (b)

3.1.4 网络整体结构

针对数据集的特征，本项目设计融合 ConvLSTM 单元和 Encoder-forecaster 结构的网络。网络结构如图 3.6 所示，为基于 ConvLSTM 的 3 层 Encoder-Forecaster，每层包括一个下采样层/上采样层和 ConvLSTM 层。Encoder 接收雷达回波图像序列，经过下采样层和 ConvLSTM 层输出隐藏状态信息，并分别传递到下一层和同一层下一时刻的 ConvLSTM 单元中；Forecaster 根据 Encoder 传递的隐藏状态信息，经过三层 ConvLSTM 和上采样输出预测序列。

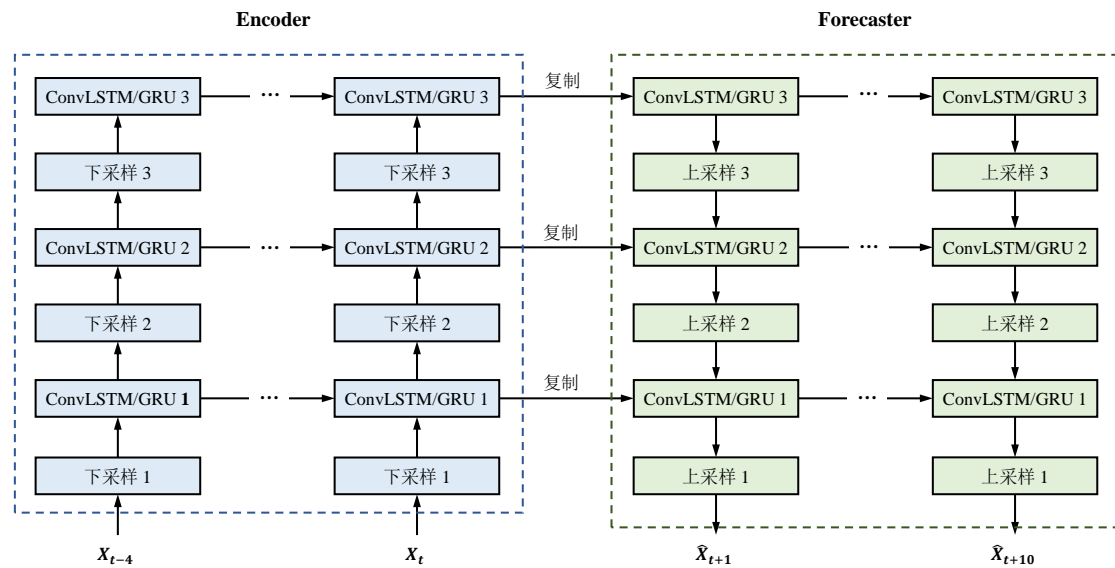


图 3.6 网络整体结构

表 3.1 网络参数

层名称	卷积核	步长	输入通道	输出通道	输入尺寸	输出尺寸
下采样 1	3×3	2×2	1/3	8/16	101×101	51×51
ConvLSTM/GRU 1	3×3	1×1	8/16	8/16	51×51	51×51
下采样 2	3×3	2×2	8/16	16/32	51×51	26×26
ConvLSTM/GRU 2	3×3	1×1	16/32	16/32	26×26	26×26
下采样 3	3×3	2×2	16/32	32/64	13×13	26×26
上采样 3	3×3	2×2	32/64	16/32	13×13	26×26
ConvLSTM/GRU 2	3×3	1×1	16/32	16/32	26×26	26×26
上采样 2	3×3	2×2	16/32	8/16	26×26	51×51
ConvLSTM/GRU 1	3×3	1×1	8/16	8/16	51×51	51×51
上采样 1	3×3	2×2	8/16	1	51×51	101×101

表 3.2 模型通道数和参数数量

序号	输入通道数	模型大小	参数数量
M1	1	小	791,043
M2	3	小	791,187
M3	1	大	3,158,019
M4	3	大	3,158,307

Jing 等在处理相似问题时将不同高度的雷达信息同时作为网络输入，而将中间高度作为预测目标^[33]。考虑网络卷积核数量和输入通道数量对网络性能的影响，设计两组组对比实验。表 3.1 和 3.2 列出了网络的详细参数，表 3.2 是模型通道数和参数数量。

3.2 实验验证

3.2.1 训练设置

深度学习模型由于参数众多、运算复杂，需要消耗大量计算资源，与传统 CPU 相比，GPU 可以大幅提高神经网络的训练和运算速度，而且业界已经开发出一些支持 GPU 运算的深度学习编程框架。本章全部实验都在以下实验环境中进行：编程语言为 Python 3.7.3，编程框架为 Pytorch 1.0.1，操作系统为 Ubuntu 14.04，GPU 为 NVIDIA TITAN X。

超参数是神经网络无法自动更新优化的参数，需要在训练时手动设置。经过反复调试，最终的超参数设置为：采用 Kaiming Uniform 的卷积核初始化方法^[31]；采用 Adam 优化方法^[32]，初始学习率为 0.0005，一阶动量和二阶动量系数分别为 0.9, 0.999；采用分批次训练方式，每批次输入 50 组样本；参数更新次数为 16000，数据集遍历次数为 80 轮。

3.2.2 损失函数

深度学习的损失函数决定了网络的梯度传播方式和优化方向。平均绝对误差损失 (MAE) 和均方误差损失 (MSE) 是常用的两种用来描述图像差异程度的损失函数，计算方法如式 (3-1) 和 (3-2)。

$$L_{MAE}(x, y) = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |x_{ij} - y_{ij}| \quad (3-1)$$

$$L_{MSE} = \frac{1}{MN} \sum_{i=1}^M (x_{ij} - y_{ij})^2 \quad (3-2)$$

其中 x, y 表示两张图像, M, N 分别表示图像的高度和宽度。网络可以通过极小化 MAE 或 MSE 实现参数优化。然而, 这类损失函数会引导网络趋于选择所有可能分布的均值作为优化方向, 这会导致预测结果产生模糊问题^[33]。本实验中, 预测图像损失采用 MAE+MSE 的形式 (式 3-3), 这是因为 MSE 损失处处可导, 收敛速度快, 但对离群点敏感; 而 MAE 损失具有更好的鲁棒性, 预测图像有更高的锐度, 但收敛速度慢^{[28][29][33]}。因此采取综合二者优点的组合损失形式。

$$\min J(X_i, \hat{X}_i) = L_{MAE} + L_{MSE} = \frac{1}{N} \sum_{i=1}^N [|X_i - \hat{X}_i| + (X_i - \hat{X}_i)^2] \quad (3-3)$$

3.2.3 训练过程

取验证集损失最小时的模型参数作为该模型的测试参数, 图 3.7 展示了小网络和大网络的训练过程。其中蓝色折线和红色折线分别代表模型的训练集损失和验证集损失。从整体上看, 迭代到 80 轮的时候, 各网络基本都已经收敛。

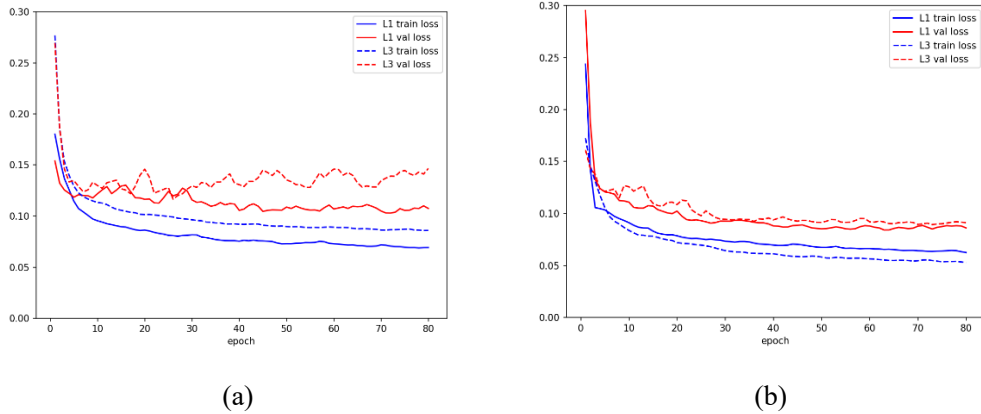


图 3.7 各网络的训练过程对比。(a) 小网络; (b) 大网络

4 结果分析与研究结论

4.1 样本预报结果分析

在测试集中抽取两个样本进行绘图分析, 对预报的准确性进行直观的评价。

(1) 样本 0

图 4.1 为四个网络对样本 0 雷达回波率的预报结果。第一行为输入的雷达回波图像序列 (此处只绘制了 2.5 km 高度序列), 对应时序为 -24min~0min (当前); 第二行为实际观测雷达回波图像序列, 对应时序为 6min~60min, 第 3 行至第 6 行为四个实验的预测结果, 分别对应 M1、M2、M3 和 M4。

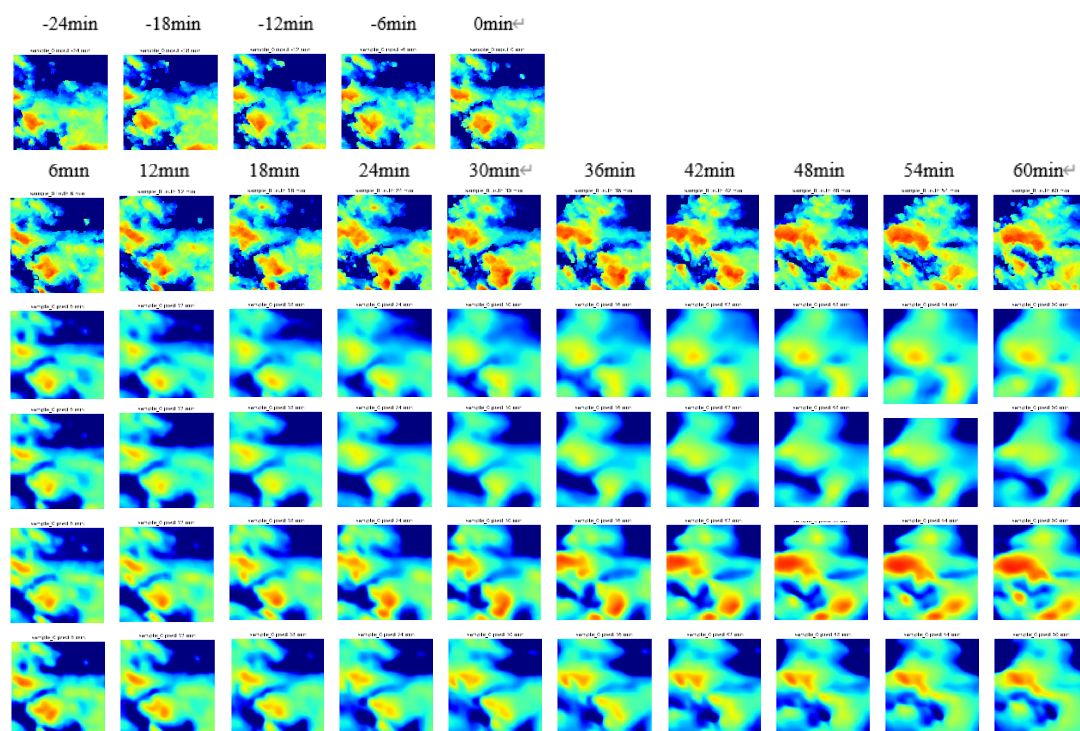


图 4.1 样本 0 预报结果

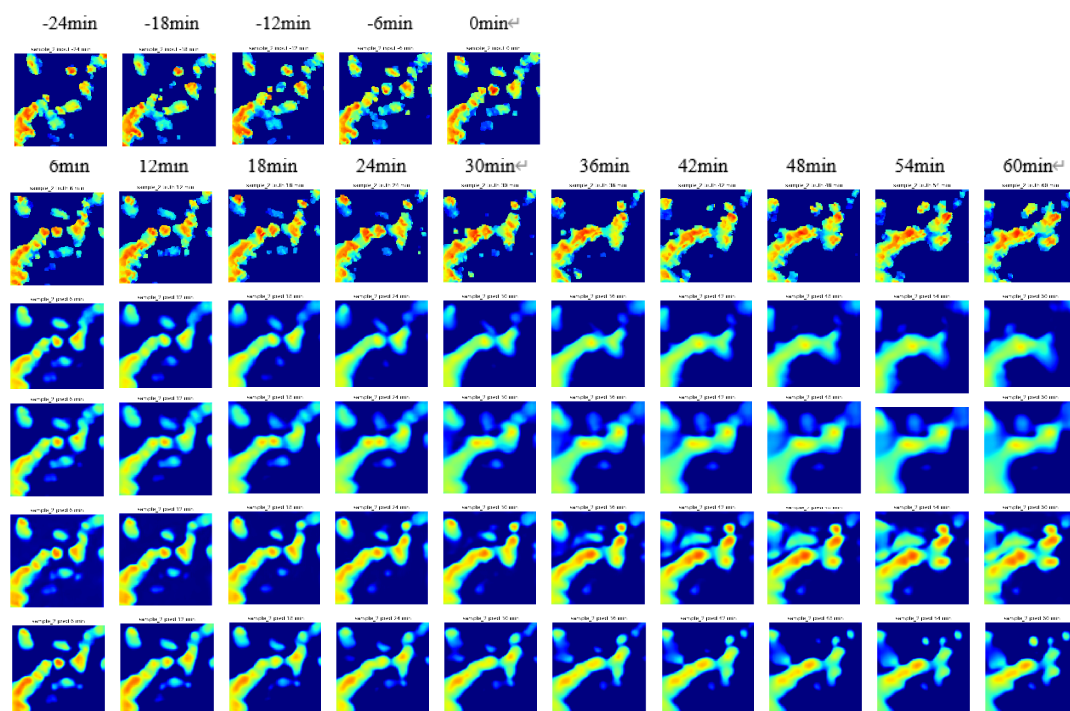


图 4.2 样本 1 预报结果

预测图像直观评价：

① 大网络结构和小网络结构均可以预测到回波的轮廓，但是大网络的预测结果细节更丰富，更接近于实际观测的结果，尤其是针对红色的高值区域，小网络结构基本无法预测。

② 与 3 通道结构相比，1 通道对整体强度和高值区域的预测值都更加接近实际观测值。

(2) 样本 1

图 4.2 为四个网络对样本 1 雷达回波率的预报结果。每张子图的含义与图 4.1 一致。

预测图像直观评价：

① 从样本 1 的预报结果中同样可以得到的和样本同样的结论，大网络的预测结果优于小网络，同时 1 通道对于高值区域的预测准确度优于 3 通道。

② 从 1 通道大网络模型的预测结果中可以看到，在 36min 后，模型对于部分稀疏区域的值出现了高估，与实测结果不符，这一现象将会在第 4.2 节进一步探讨。

4.2 预报准确性分析

为了定量评价深度学习网络的准确性，需要采取合适的预测指标进行分析。降水预报常用的预测指标为探测率 (POD)、误报率 (FAR)、临界成功指数 (CSI) 和 Heidke 技巧评分 (HSS)。首先按照某个阈值将预报值转换为二值矩阵（大于阈值为 1，小于阈值为 0），然后正确统计其中的正确探测（真实值=1，预测值=1）、漏报（真实值=1，预测值=0）、误报（真实值=0，预测值=1）和正确排除（真实值=0，预测值=0）的像素数量。POD、FAR、CSI 和 HSS 的计算方法如 (4-1)~(4-4) 式。

$$POD = \frac{n_h}{n_h + n_m} \quad (4-1)$$

$$FAR = \frac{n_f}{n_h + n_f} \quad (4-2)$$

$$CSI = \frac{n_h}{n_h + n_m + n_f} \quad (4-3)$$

$$HSS = 2 \frac{n_h n_c - n_m n_f}{(n_h + n_m)(n_m + n_c) + (n_h + n_f)(n_f + n_c)} \quad (4-4)$$

其中 n_h, n_m, n_f, n_c 分别代表正确探测、漏报、误报和正确排除的数量。POD、FAR、CSI 和 HSS 的取值范围在[0, 1]区间内，其中 POD、CSI 和 HSS 的取值越

接近 1 且 FAR 越接近 0，说明网络预报效果越好。实验中将阈值设为 0.2。

除此之外，为了评判预测图像序列与实际观测图像序列的整体和局部产剧，还采用均方误差 (MSE) 和结构相似性 (SSIM) 作为预测指标。MSE 在平均水平上描述图像整体误差，SSIE 则从亮度、对比度和结构方面评价图像整体和局部相似程度。MSE 和 SSIM 的计算方法如 (4-5)、(4-6) 式。

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (x_{ij} - y_{ij})^2 \quad (4-5)$$

$$SSIM(x, y) = \frac{(2u_x u_y + C_1)(2\sigma_{xy} + C_2)}{(u_x^2 + u_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (4-6)$$

其中 x, y 表示两张图像， μ_x, μ_y 表示 x, y 的均值， σ_x, σ_y 表示 x, y 的方差， σ_{xy} 表示 x, y 的协方差， C_1, C_2 为可调节参数，实验中分别设为 0.01 和 0.03。SSIM 的取值范围在 $[0, 1]$ 区间内，SSIM 越接近于 1 且 MSE 越接近与 0，则两张图像的相似程度越高。

分别计算 4 个网络的 6 个测试指标的值，测试结果如表 4.1 所示。

预测指标评价分析：

① 对于模型的大小而言，增加网络参数数量对预报效果的提升时显著的，M3 和 M1 相比几乎在所有的指标上都有了很大程度的提升。这个结果也和样本分析中直接观察到的结果相符合。

② 对于模型的输入通道数而言，与 1 通道相比，3 通道的 POD 和 FAR 的值都更低，这意味着网络变得更谨慎了，正确探测和误报的像素数量均出现了减少，这也对用这样本分析中样本 2 观察到的 1 通道的网络在某些稀疏区域雷达反射率存在高估的现象。

表 4.1 各网络在测试集上的预测指标比较

序号	输入通道	模型大小	参数数量	POD	FAR	CSI	HSS	SSIM	MSE
M1	1	小	791,043	0.8750	0.1567	0.7692	0.6035	0.4894	0.0171
M2	3	小	791,187	0.8419	0.1192	0.7642	0.6070	0.4323	0.0208
M3	1	大	3,158,019	0.8983	0.1275	0.8043	0.6447	0.5756	0.0138
M4	3	大	3,158,307	0.8755	0.1203	0.7923	0.6468	0.5455	0.0149

4.3 预报时效分析

预报时效指预报的有效期限,分析预报时效需要查看各时刻的预值与实际观测值。根据四组实验在各时刻的预报图像与实际观测图像,绘制不同预测指标随预报时间的变化曲线,如图 4.3 所示。

预报时效性分析:

- ① 随着时间的推移,四组实验的预报准确性都出现了明显的下滑,POD、CSI、HSS 和 SSIM 的值大幅度减小,而 FAR 和 MSE 的值明显增大。
- ② 大网络和小网络得时效性差异显著,大网络预报准确性的下降比小网络慢,在较长的时间段内都可以保持较好的准确性。
- ③ 3 通道的预测结果与 1 通道相比,在 POD 和 CSI 两个参数上表现更差, FAR 随时间的累积速率也更快,这一结果是反常的,与 Jing 等^[30]的研究结论相反,这一现象的原因还有待探讨。

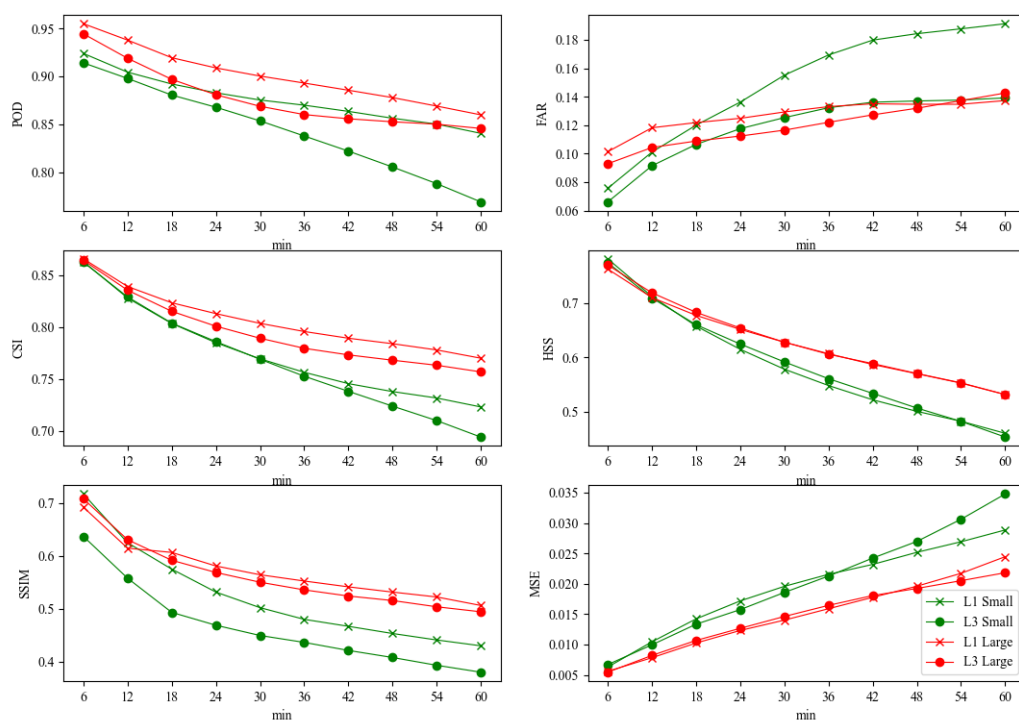


图 4.3 各实验的预测时效对比

4. 4 研究结论和未来研究计划

4. 4. 1 研究结论

(1) 网络参数数量（卷积核数量）对于预报效果的影响显著，说明需要较多的卷积核才能学到此数据集中雷达回波复杂的时空特征。

(2) 输入信息（雷达高度信息）对预报效果的影响显著，输入更多信息，预测结果的准确度反而降低了，说明按照通道见到堆叠高度信息不能使网络充分提取这些信息，可以尝试增加学习权重或者注意力机制。

4. 4. 2 未来研究计划

未来可以尝试更多方面因素对于预报效果的影响，例如卷积核尺寸、输入序列长度以及正则化方法等。

参考文献

- [1] Browning K, Collier C. Nowcasting of precipitation systems[J]. Reviews of Geophysics, 1989, 27(3): 345-370.
- [2] Nowcasting[EB/OL]. [May 9, 2016].
<https://www.wmo.int/pages/prog/amp/pwsp/Nowcasting.htm>.
- [3] 陈程. 卷积神经网络在气象短临预报的研究与应用[D]. 华南理工大学, 2018.
- [4] 施恩. 基于卷积神经网络的雷达回波外推方法研究[D]. 国防科技大学, 2017.
- [5] 俞小鼎, 周小刚, 王秀明. 雷暴与强对流临近天气预报技术进展[J]. 气象学报, 2012, 70(03): 311-337.
- [6] Baldassi C, Borgs C, Chayes J T, et al. Unreasonable effectiveness of learning neural networks: From accessible states and robust ensembles to basic algorithmic schemes[J]. Proceedings of the National Academy of Sciences, 2016, 113(48): E7655-E7662.
- [7] Shen C P. A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists[J]. Water Resources Research, 2018, 54(11): 8558-8593.
- [8] Fan Y, Lu X, Li D, et al. Video-based emotion recognition using CNN-RNN and C3D hybrid networks[C]. Proceedings of the 18th ACM International Conference on Multimodal Interaction, 2016: 445-450.
- [9] Zhao R, Ali H, Van Der Smagt P. Two-stream RNN/CNN for action recognition in 3D videos[C]. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017: 4260-4267.
- [10] Goodfellow I, Bengio Y, Courville A. Deep learning[M]. MIT press, 2016.
- [11] Parkhi O M, Vedaldi A, Zisserman A. Deep face recognition[J], 2015.
- [12] Deng L, Liu Y. Deep learning in natural language processing[M]. Springer, 2018.

- [13] Daily M, Medasani S, Behringer R, et al. Self-driving cars[J]. Computer, 2017, 50(12): 18-23.
- [14] Greenspan H, Van Ginneken B, Summers R M. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique[J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1153-1159.
- [15] Lyons J, Dehzangi A, Heffernan R, et al. Predicting backbone Ca angles and dihedrals from protein sequences by stacked sparse auto - encoder deep neural network[J]. Journal of computational chemistry, 2014, 35(28): 2040-2046.
- [16] Ji S, Xu W, Yang M, et al. 3D convolutional neural networks for human action recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 35(1): 221-231.
- [17] Tran D, Bourdev L, Fergus R, et al. Learning spatiotemporal features with 3d convolutional networks[C]. Proceedings of the IEEE international conference on computer vision, 2015: 4489-4497.
- [18] Qiu Z, Yao T, Mei T. Learning spatio-temporal representation with pseudo-3d residual networks[C]. proceedings of the IEEE International Conference on Computer Vision, 2017: 5533-5541.
- [19] Pan B X, Hsu K L, Aghakouchak A, et al. Improving Precipitation Estimation Using Convolutional Neural Network[J]. Water Resources Research, 2019, 55(3): 2301-2321.
- [20] Klein B, Wolf L, Afek Y. A dynamic convolutional layer for short range weather prediction[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 4840-4848.
- [21] Ayzel G, Heistermann M, Sorokin A, et al. All convolutional neural networks for radar-based precipitation nowcasting[J]. Procedia Computer Science, 2019, 150: 186-192.
- [22] Miao Q, Pan B, Wang H, et al. Improving monsoon precipitation prediction using combined convolutional and long short term memory neural network[J]. Water, 2019, 11(5): 977.
- [23] Wu H, Yang Q, Liu J, et al. A spatiotemporal deep fusion model for merging satellite and gauge precipitation in China[J]. Journal of Hydrology, 2020: 124664.
- [24] Shi X, Chen Z, Wang H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting[C]. Advances in neural information processing systems, 2015: 802-810.
- [25] Shi X, Gao Z, Lausen L, et al. Deep learning for precipitation nowcasting: A benchmark and a new model[C]. Advances in neural information processing systems, 2017: 5617-5627.
- [26] Tran Q-K, Song S-K. Computer Vision in Precipitation Nowcasting: Applying Image Quality Assessment Metrics for Training Deep Neural Networks[J]. Atmosphere, 2019, 10(5): 244.
- [27] Shi E, Li Q, Gu D, et al. A Method of Weather Radar Echo Extrapolation Based on Convolutional Neural Networks[C]. International Conference on Multimedia Modeling, 2018: 16-28.
- [28] Wang Y, Long M, Wang J, et al. Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms[C]. Advances in Neural Information Processing Systems, 2017: 879-888.
- [29] Wang Y, Gao Z, Long M, et al. Predrnn++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning[J]. arXiv preprint arXiv:1804.06300, 2018.

- [30] Odena. et al. (2016). Deconvolution and Checkerboard Artifacts.
- [31] He K, Zhang X, Ren S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]. Proceedings of the IEEE international conference on computer vision, 2015: 1026-1034.
- [32] Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.
- [33] Jing J, Li Q, Peng X. MLC-LSTM: Exploiting the Spatiotemporal Correlation between Multi-Level Weather Radar Echoes for Echo Sequence Extrapolation[J]. Sensors, 2019, 19(18): 3988.