

袁凯,李武阶,李明,等,2022. 四种机器深度学习算法对武汉地区雷达回波临近预报的检验和评估[J]. 气象,48(4):428-441.
Yuan K, Li W J, Li M, et al, 2022. Examination and evaluation of four machine deep learning algorithms for radar echo nowcasting in Wuhan Region[J]. Meteor Mon, 48(4):428-441(in Chinese).

四种机器深度学习算法对武汉地区雷达 回波临近预报的检验和评估^{*}

袁 凯 李武阶 李 明 庞 晶

武汉市气象台, 武汉 430040

提 要: 基于 PredRNN++、MIM、CrevNet 和 PhyDNet 四种机器深度学习算法, 利用武汉地区 2012—2019 年的雷达和降水资料, 开展了人工智能技术在武汉地区临近预报中的应用研究, 根据均方误差(MSE)、结构相似性指数(SSIM)、命中率(POD)、虚警率(FAR)和临界成功指数(CSI)等指标检验评估了四种机器学习算法对武汉地区雷达回波临近预报的预报性能, 并以半拉格朗日光流法进行了对比, 得到以下主要结论: MIM 算法的 MSE 和 FAR 最低, SSIM 最高; PredRNN++ 算法的 POD 和 CSI 最高。机器深度学习算法的 POD、CSI 和 SSIM 均高于光流法, FAR 和 MSE 则更低, 其中 SSIM、POD、CSI 三种指标的提升幅度在 3.2%~24.7%, MSE 和 FAR 两种指标的降幅在 13.1%~43.3%。30 min 以内, 除 CrevNet 外, 其余三种机器学习算法和光流法的预报能力较为接近; 30 min 以后, 深度学习算法和光流法都随着预报时效的延长, 预报能力均显著下降, 但机器学习算法下降得更缓慢, 尤其是 60 min 以后光流法的降幅进一步增加, 显示出机器学习长预报时效的优势。此外, 机器学习算法之间针对不同评分指标在不同预报时效的下降速度并不一致。PredRNN++ 算法在所有强度上 CSI 均表现最佳, MIM 和 PhyDNet 两种算法对 ≥ 40 dBz 的回波预报、CrevNet 算法对 ≥ 50 dBz 的回波预报均好于光流法。机器学习算法和光流法都随着回波强度的增加, CSI 和 POD 迅速降低, FAR 快速上升, 但机器学习算法的 FAR 上升得更慢。四个不同回波形态、不同发展趋势个例的分析结果表明, 机器学习算法不仅具备对一定回波强度变化的预报能力, 而且对回波强度和面积变化趋势的时间节点预报也与实况基本一致。此外, 机器学习算法对回波运动的预报能力明显强于光流法, 显示出机器学习算法良好的应用前景。

关键词: 机器深度学习, 雷达回波, 临近预报, 检验评估

中图分类号: P456, P409

文献标志码: A

DOI: 10.7519/j.issn.1000-0526.2021.102001

Examination and Evaluation of four Machine Deep Learning Algorithms for Radar Echo Nowcasting in Wuhan Region

YUAN Kai LI Wujie LI Ming PANG Jing

Wuhan Meteorological Observatory, Wuhan 430040

Abstract: Based on four machine deep learning algorithms (PredRNN++, MIM, CrevNet and PhyDNet), radar data and precipitation data in Wuhan from 2012 to 2019, this study investigates the possibility of application of artificial intelligence (AI) technology in the nowcasting of Wuhan Region. The forecasting skills of radar echo nowcasting are examined in terms of mean square error (MSE), structural similarity index (SSIM), probability of detection (POD), false alarm rate (FAR) and critical success index (CSI), then compared with the semi-Lagrangian optical flow method. The results are summarized as follows. The MSE and FAR are the lowest and SSIM is the highest in the MIM algorithm. The POD and CSI of

^{*} 2021 年 1 月 7 日收稿; 2021 年 8 月 3 日收修定稿

第一作者: 袁凯, 主要从事短时天气预报研究. E-mail: yuankai2009@126.com

通信作者: 李武阶, 主要从事天气预报分析和技术方法应用研究. E-mail: 1669625159@qq.com

PredRNN++ are the highest. The POD, CSI and SSIM of machine learning are higher than semi-Lagrangian optical flow, while the FAR and MSE of machine learning are much lower, of which the SSIM, POD and CSI of machine algorithms are improved by 3.2%—24.7% than semi-Lagrangian optical flow, but the MSE and FAR are reduced by 13.1%—43.3%. Within 30 minutes, except the CrevNet algorithm, the skills of other algorithms are similar to that of semi-Lagrangian optical flow. 30 minutes later, the skills of both machine algorithm and semi-Lagrangian optical flow decline significantly with the increase of forecast lead time. However, the skill of machine algorithms declines much more slowly. Especially after 60 minutes, the skill of semi-Lagrangian optical flow descends more quickly indicating the advantage of machine learning algorithms for long-term prediction. In addition, the descending rates at different forecast lead times for different score indexes are different among the machine algorithms. The CIS of PredRNN++ is the highest in any intensity, MIM and PhyDNet performance is better than semi-Lagrangian optical flow for radar echo intensity exceeding 40 dBz, but CrevNet shows better skill for radar echo intensity exceeding 50 dBz. The POD and CSI of machine algorithms and semi-Lagrangian optical flow decline significantly with the increase of forecast intensity of radar echo, while the FAR increases quickly, but the increase of FAR rate of machine learning algorithm is much slower. To sum up, the analysis of four different echo patterns and different development trends shows that the machine learning algorithm has the ability not only to predict the change of radar echo intensity in a certain content, but also to predict the time node of the evolution tendency of intensity and acreage, which are basically consistent with the observation. These results suggest that the ability of machine deep learning to predict the movement of radar echo is better than that of semi-Lagrangian optical flow, indicating its possible wide prospect for operational application.

Key words: machine deep learning, radar echo, nowcasting, examination and evaluation

引 言

在天气预报业务中,临近预报通常是指对某一区域未来0~2 h内的短时强降水、冰雹、雷暴大风、龙卷和雷电等灾害性天气进行预测(俞小鼎等,2012),在气象灾害防御中具有极为重要的地位。临近预报的方法主要有雷达回波外推预报、中尺度数值模式预报以及概念模型预报等(Wilson et al, 1998)。中尺度数值模式预报由于模式初始场协调和资料同化时效等问题的限制,在最初的几个小时内其预报效果较差,甚至无法直接用于临近预报(王丹等,2014;吴剑坤等,2019);概念模型预报带有预报员一定的个人主观性,且精细化程度较低。因此,目前业务上临近预报方法仍以雷达回波的外推预报为主(Wilson et al, 2010; Wang et al, 2016),主要包括单体质心法、交叉相关法和光流法。

目前业务上广泛使用的单体质心算法主要有TITAN(thunder identification tracking analysis and nowcasting)(Dixon and Wiener, 1993)和SCIT(storm cell identification and tracking)(Johnson

et al, 1998)。单体质心法最先由Austin and Bellon(1982)提出,然后Rosenfelds(1987)、Dixon and Wiener(1993)和徐月飞等(2011)对其进行了大量试验和改进,虽然取得了一定的效果,但其较大的计算量和只适用于强对流风暴的“先天性缺陷”,导致其预报准确率很难以较大幅度地提高。

交叉相关法由Rinehart and Garvey(1978)提出,其后国内外专家和学者从不同角度对其进行了改进,改进后的交叉相关法解决了回波波形的变化导致的运动无序(Rinehart, 1981)和反演矢量不连续等问题(Li et al, 1995),得到了较为均匀的风场(徐亚钦等, 2011),但基于交叉相关法的外推预报技术主要获得的是雷达回波的水平运动矢量,并没有考虑深对流系统通常伴随较强的垂直运动,因此该方法仅适用于缓慢变化的层状云降水系统,而对快速增长或消散的强对流系统的预报效果并不理想(韩雷等, 2008; 曹春燕等, 2015)。

为了进一步提高雷达回波的外推预报技术,许多专家和学者引入了计算机视觉领域的光流法,该方法最早由Gibson(1979)提出,众多专家学者对其进行了大量检验和优化升级,结果表明:光流法可以

较为准确地得到雷暴的整体运动趋势,对于移动型局地生成及强度和形状随时间变化较快的雷暴,光流法明显优于交叉相关法,而对于热带降水系统尤其是台风系统,光流法没有考虑雷达回波的旋转性,预报效果不如交叉相关法(曹春燕等,2015;Bechini and Chandrasekar,2017)。虽然光流法已较大幅度地提高了回波临近预报的准确率,但其基本模型较实况仍存在不少误差,比如要求图像遵循灰度不变形假设,即线性外推原则,此外求解光流场时,也只适用于回波运动较小的情况,对于快速移动的回波误差仍然较大。

近年来,以机器深度学习为代表的人工智能技术在图像识别和视频领域取得了突破性进展,由于其挖掘海量数据和非线性建模的优秀能力,众多学者将其引入到以雷达回波预报为主的短时临近预报业务中,并取得了良好的成果(许小峰,2018;黄骅文等,2021)。Shi et al(2015)提出了卷积长短期记忆单元网络算法(convolutional LSTM network, ConvLSTM),结果表明 ConvLSTM 优于普通的光流法;Wang et al(2017)提出了 PredRNN 算法,并在此之上,提出了 PredRNN++ 算法(Wang et al, 2018),结果显示 PredRNN++ 效果较 ConvLSTM 有明显的提升;Wang et al(2019)提出了 MIM (memory in memory)算法,结果表明 MIM 算法不仅具有良好的平稳信息的提取能力,还具备一定提取非平稳信息的能力;陈元昭等(2019)研究了基于生成对抗网络的临近预报方法,结果表明对中等强度的回波其预报效果较好。Yu et al(2020)提出了 CrevNet 算法,结果显示该算法对交通视频识别与处理精准度较高,具有较强的预测能力;Le Guen and Thone(2020)以 ConvLSTM 为内核,提出了 PhyDNet 算法,结果表明该算法在视频预测上具有良好的时空预测能力,尤其对长时间预测具有明显的优势。陈训来等(2021)利用卷积门控循环单元神经网络算法(convolutional gated recurrent unit, ConvGRU)对雷达回波进行临近预报,结果表明,基于 ConvGRU 方法对强对流天气具有较好的预报效果,业务中具有广泛的应用前景;顾建峰等(2020)采用 TrajGRU 算法,建立了重庆地区三维雷达回波智能预报系统,并将其业务化,结果显示机器学习在回波演变的临近预报方面具有明显优势,人工智能技术在重庆临近预警业务中已初见成效。

虽然基于深度学习的临近预报技术已经取得了

令人欣喜的成果,各种算法都表现出特有的预报性能,但雷达回波的发生、发展和移动是非常复杂的,它与站点所处的气候带、站点附近特殊的地形和地貌特征等因素相关,所以开展机器深度学习算法预报性能的检验评估,有利于各地选择适用于本地的深度学习方法,建立最优的深度学习方法(陈锦鹏等,2021;周康辉等,2021)。因此,本文选取在常规数据集中表现良好、计算成本较低且内部结构差异较大的四种深度学习算法,利用武汉地区较长时间序列的雷达和降水资料,开展基于机器深度学习算法的雷达回波临近预报应用研究和预报结果的检验评估,并与业务中常用的半拉格朗日光流法进行对比分析和检验,以期寻找到适合武汉地区的深度学习算法,为武汉地区雷达回波临近预报提供智能化的客观参考。

1 资料与方法

1.1 资料

本研究主要用到以下两种资料:逐日和逐小时的降水资料,主要用来选取训练和检验样本;新一代多普勒天气雷达组合反射率因子拼图产品,时间分辨率为 6 min,本文所选范围为覆盖湖北东部地区的以武汉雷达站为中心、边长为 256 km 的正方形区域,其经纬度范围分别为 $29.247^{\circ} \sim 31.797^{\circ}\text{N}$ 、 $113.098^{\circ} \sim 115.648^{\circ}\text{E}$,空间分辨率为 $1\text{ km} \times 1\text{ km}$ 。雷达回波数据进行了孤立噪音过滤和超折射回波抑制(吴涛等,2013),并利用中值滤波(赵悦等,2007)进行质量控制。

为尽可能地增多训练和检验样本个数,使算法能够学习到不同种类的回波特征,同时降低样本严重失调(弱降水过程和无降水的天气占全年的绝大多数时间)带来的不利影响,本文以武汉地区 5 个国家气象观测站和 14 个观测环境良好、数据质量较高且地理分布较为均匀的自动气象观测站为基础(具体站点和位置如图 1 所示),选取 2012 年 6 月 1 日至 2019 年 12 月 31 日,上述任一站点日降水量 $\geq 10\text{ mm}$ 、小时降水量 $\geq 0.6\text{ mm}$ 且组合反射率因子拼图连续 3 小时不断裂的连续雷达回波为 1 个样本(如有重叠则算为 1 个样本),然后对样本进行切片

亦如此,其中 PredRNN++、MIM 算法已经应用于雷达回波临近预报,而 CrevNet 和 PhyDNet 算法则是视频预测领域最新的算法,尚未应用于雷达回波临近预报。

2.1 PredRNN++ 算法

Shi et al(2015)将 RNN 中传统的 LSTM(long short-term memory)中的全连接层改为卷积层,提出了 ConvLSTM 算法,而后发展出的 ConvGRU、TrajGRU 等算法(Shi et al,2017)与 ConvLSTM 并无本质的区别,这些算法将 LSTM 或 GRU 中的记忆模块改造成 CNN 结构,增强了算法对空间特征的提取能力。Wang et al(2017)在 ConvLSTM 算法的基础之上,将可以记忆的单元放置在算法的堆叠结构中,提出了 PredRNN 算法,为了缓解该算法中梯度容易消失的问题以及提高其对短时非线性时空特征的提取能力,引入了 GHU(gradient highway unit),该结构使得梯度能够在第一层和第二层之间高速传递,有效抑制了梯度的消失,最终提出了 PredRNN++ 算法(Wang et al,2018)。

2.2 MIM 算法

Wang et al(2019)为了解决 PredRNN 中 LSTM 遗忘门的饱和问题,将图片的信息分为平稳信息和非平稳信息两部分,提出了 MIM 算法,该算法分两次对图片信息进行提取,首先由 MIM-N 结构提取出非平稳信息,而后传递给 MIM-S, MIM-S 则利用门控来选择记忆或忘记非平稳信息的多少,同时通过多层模块之间相互的差分运算,使得非平稳信息缓慢降低,从而提取出各种高阶的非平稳信息,最终将所提取的平稳信息和非平稳信息相结合,进行输出与预测。

2.3 CrevNet 算法

Yu et al(2020)提出的 CrevNet 算法,是一种全新的嵌套了三维卷积模块的双向可逆自编码结构,其在一系列正向和反向计算过程中使得输入和特征之间建立了一对一的双向映射关系,这种关系理论上保证了在特征提取过程中不丢失信息,因而保留更多信息进行预测,明显提高了预测图片的清晰度。此外,该算法的内存和计算开销都较小,对于硬件要求不高,易于训练和调试。

2.4 PhyDNet 算法

Le Guen and Thome(2020)提出了 PhyDNet 算法,该算法参考了 MIM 算法的基本假设,将图片信息分为已知的物理过程和未知因素(包括生消、发展等)两个部分,然后用深度网络结构来约束模型,以 ConvLSTM 为主要内核来提取未知因素,此外利用卷积过程模拟偏导,从而学习到新的物理信息,最后将物理信息和已有的未知因素结合,从而进行更好的预测。

2.5 半拉格朗日光流法

本文所指的半拉格朗日光流法均为 Lucas-Kanade 稀疏光流法(Lucas and Kanade,1981)。Lucas-Kanade 稀疏光流法处理的像素点较少,耗时短,该算法需满足以下 3 个基本假设:(1)亮度恒定,即同一点随着时间的变化,其亮度不变;(2)小运动,即随着时间的变化不会引起位置的剧烈变化;(3)空间一致,即场景中邻近点投影到图像上相邻位置不变且相邻点速度一致。具体的计算步骤:首先,根据每组样本中最后连续的 3 张雷达回波实际图像,利用 Lucas-Kanade 方法计算出光流场;然后,基于已经计算出的光流场,利用半拉格朗日方法进行外推预报(张蕾等,2014)。

3 结果分析

3.1 个例分析

3.1.1 2018 年 6 月 30 日片状回波个例

2018 年 6 月 30 日 11:30 的雷达回波上显示(图 2),武汉地区西部有强度超过 45 dBz 的片状回波,从回波的演变来看,预报时效内该片状回波不断东移进入湖北东部地区,回波面积变大,强度无明显变化。对比光流法和机器深度算法的预报结果:光流法对片状回波的位置把握较好,但对回波形态的预报较差,尤其是对后侧进入武汉地区的回波缺乏预报能力。而四种机器学习算法中,除 CrevNet 算法外均准确预报出了后期进入武汉地区的片状回波面积的增大,尤其是 PhyDNet 和 PredRNN++ 算法预报的回波形态与强回波位置与实况吻合较好。另外,从回波强度上看,深度学习算法预报的回波强

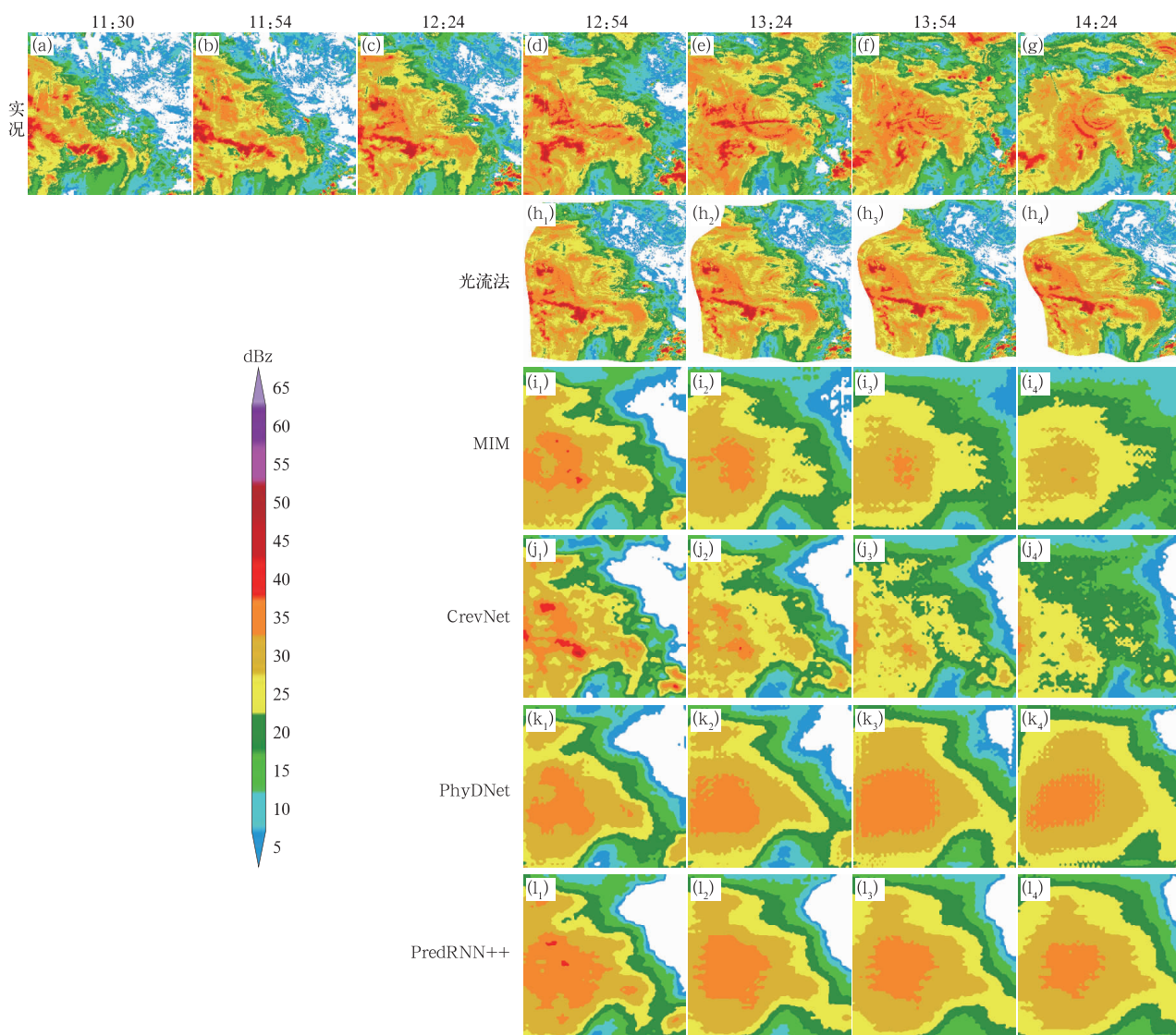


图 2 2018 年 6 月 30 日 11:30—14:24(a~g)回波实况和(h~l)预报回波对比

(a)11:30,(b)11:54,(c)12:24,(d)12:54,(e)13:24,(f)13:54,(g)14:24;

(h)光流法,(i)MIM,(j)CrevNet,(k)PhyDNet,(l)PredRNN++

Fig. 2 Comparison between (a—g) observation and (h—l) forecast of radar echo from 11:30 BT to 14:24 BT 30 June 2018

(a) 11:30 BT, (b) 11:54 BT, (c) 12:24 BT, (d) 12:54 BT, (e) 13:24 BT,

(f) 13:54 BT, (g) 14:24 BT; (h) optical flow, (i) MIM, (j) CrevNet,

(k) PhyDNet, (l) PredRNN++

度较实况都偏弱,PhyDNet 和 PredRNN++ 算法强度最强,最接近实况。

3.1.2 2014 年 7 月 4 日弓状回波个例

2014 年 7 月 4 日 08:30 的雷达回波上显示(图 3),武汉中北部和南部地区有多个强度达 45 dBz 以上的块状回波存在,从回波的演变来看,预报时效内这些块状回波在东移的过程中,逐渐合并和组织

化,强度几乎保持不变,最终形成一个经武汉东北部—东南部—西南部环绕湖北东部地区的大弓状回波。从光流法和机器深度算法的预报结果来看:光流法对东北侧回波位置的把握较好,但对整体的回波形态的预报较差,尤其是对块状回波移动过程中的合并缺乏预报能力。而除 CrevNet 以外的其他三种机器学习算法均成功预报出了块状回波在东

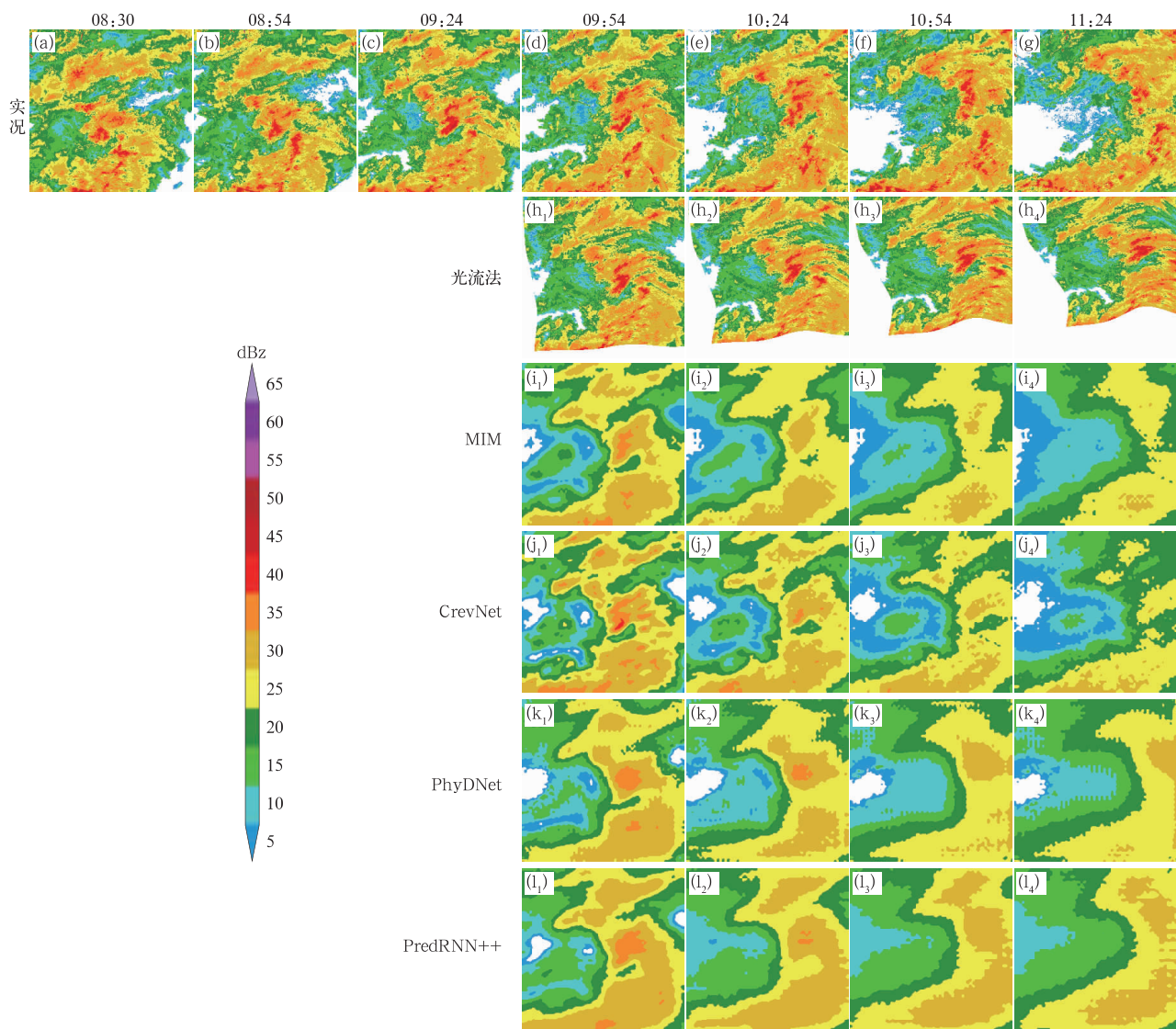


图 3 同图 2, 但为 2014 年 7 月 4 日 08:30—11:24(a~g) 雷达实况
(a)08:30, (b)08:54, (c)09:24, (d)09:54, (e)10:24, (f)10:54, (g)11:24

Fig. 3 Same as Fig. 2, but from 08:30 BT to 11:24 BT 4 July 2014

(a) 08:30 BT, (b) 08:54 BT, (c) 09:24 BT, (d) 09:54 BT, (e) 10:24 BT, (f) 10:54 BT, (g) 11:24 BT

移过程中的合并, 整体回波形状和回波中心位置与实况非常吻合。最后, 从回波强度上看, 所有深度学习算法预报的强度较实况都偏弱, 但 PredRNN++ 算法强度最强, 与实况最接近。

3.1.3 2018 年 7 月 30 日分散性回波个例

2018 年 7 月 30 日 20:30 的雷达回波上显示 (图 4), 武汉地区中部和南部有多个分散性块状回波存在, 其中中部块状回波强度超过 50 dBZ, 南部块状回波强度较弱, 从回波的演变来看, 预报时效内中部块状回波的面积和强度均快速减弱, 至 23:30

基本消散殆尽, 而南部的块状回波面积逐渐减小, 强度略有减弱。对比光流法和机器深度算法的预报结果: 光流法虚报出了中部地区多块的强回波以及过高估计了南部块状回波的强度和面积。而机器学习算法中, 除 PhyDNet 算法外, 均准确预报出了中部块状回波的强烈衰减, 而且对回波强度减弱和面积变小的时间节点的预报也与实况几乎一致, 都表现为 21:54—22:24 回波强度的稍有减弱和面积的稍微变小, 以及 22:24—23:24 回波强度的明显减弱和面积的显著缩小, 尤其是 CrevNet 算法还成功地预

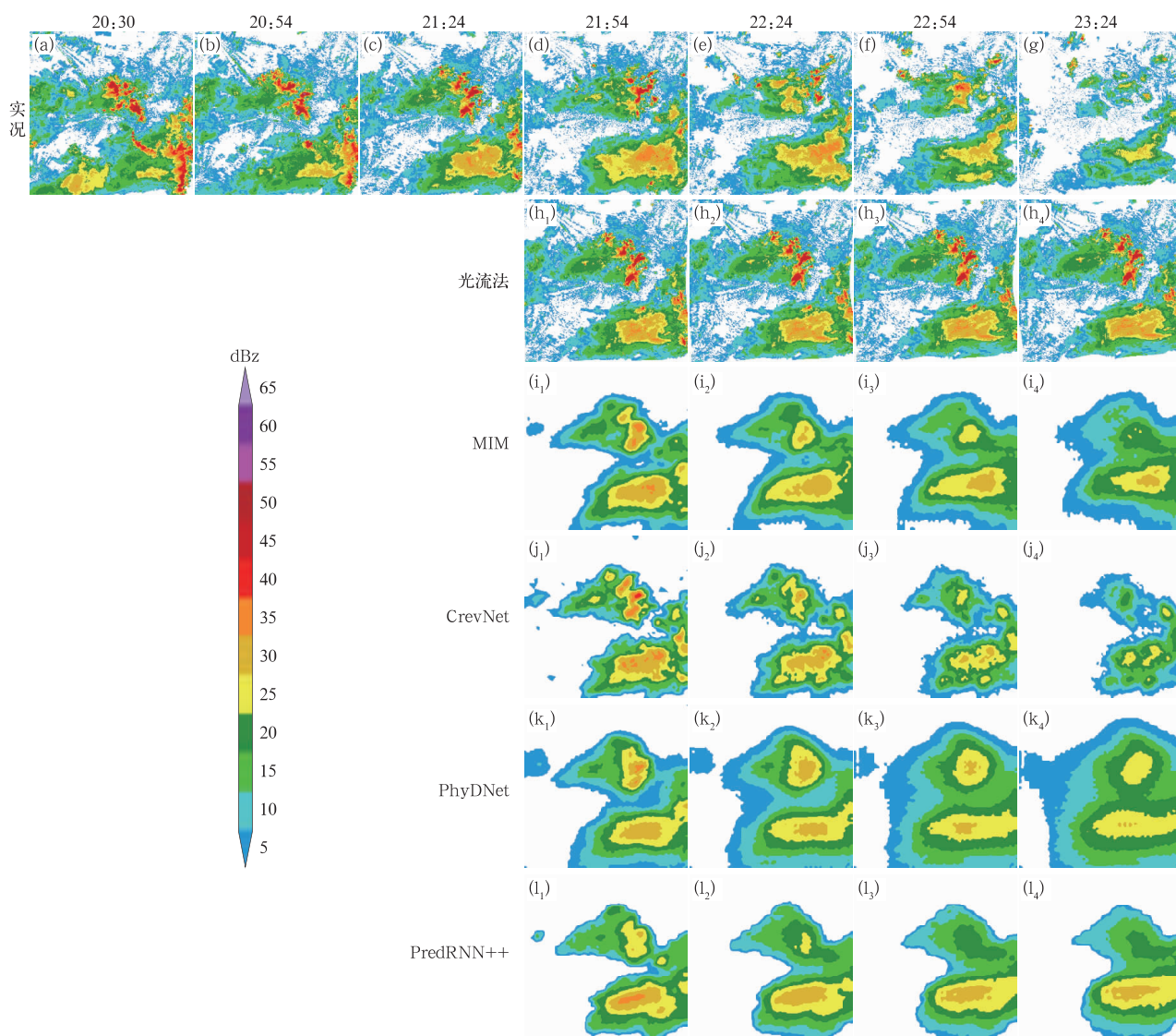


图4 同图2,但为2018年7月30日20:30—23:24

(a)20:30,(b)20:54,(c)21:24,(d)21:54,(e)22:24,(f)22:54,(g)23:24

Fig. 4 Same as Fig. 2, but from 20:30 BT to 23:24 BT 30 July 2018

(a) 20:30 BT, (b) 20:54 BT, (c) 21:24 BT, (d) 21:54 BT, (e) 22:24 BT, (f) 22:54 BT, (g) 23:24 BT

报出了南侧块状回波面积的变小,其预报的回波面积与回波位置与实况吻合较好,强度也非常接近实况回波强度。

3.1.4 2016年7月1日带状回波个例

2016年7月1日15:30的雷达回波上显示(图5),武汉地区有南北两块中心强度超过45 dBz的强回波,北部呈块状,南部则呈带状,从后期回波的演变来看,北部的块状回波缓慢东移,强度减弱,面积变小,南侧的带状回波强度和面积无太大变化,但位置略有南压。从光流法和机器深度算法的预报结果来看:光流法对北部回波移动的把握较好,预报

的强回波的位置与实况基本吻合,但其预报的回波强度较实况略偏强,而对南侧带状回波位置预报误差较大,尤其在30 min之后误差进一步增大。而四种机器学习算法均预报出了北部块状回波在东移过程中强度减弱和面积变小的变化趋势,而且对回波强度减弱和面积变小的时间节点与实况几乎一致,都表现为16:54—17:24回波强度的稍有减弱和面积的稍微变小以及17:24—18:24回波强度的明显减弱和面积的显著缩小。此外,机器学习算法对强回波位置的预报也与实况基本一致,其所预报的南北两块强回波位置与实况基本一致,但南侧

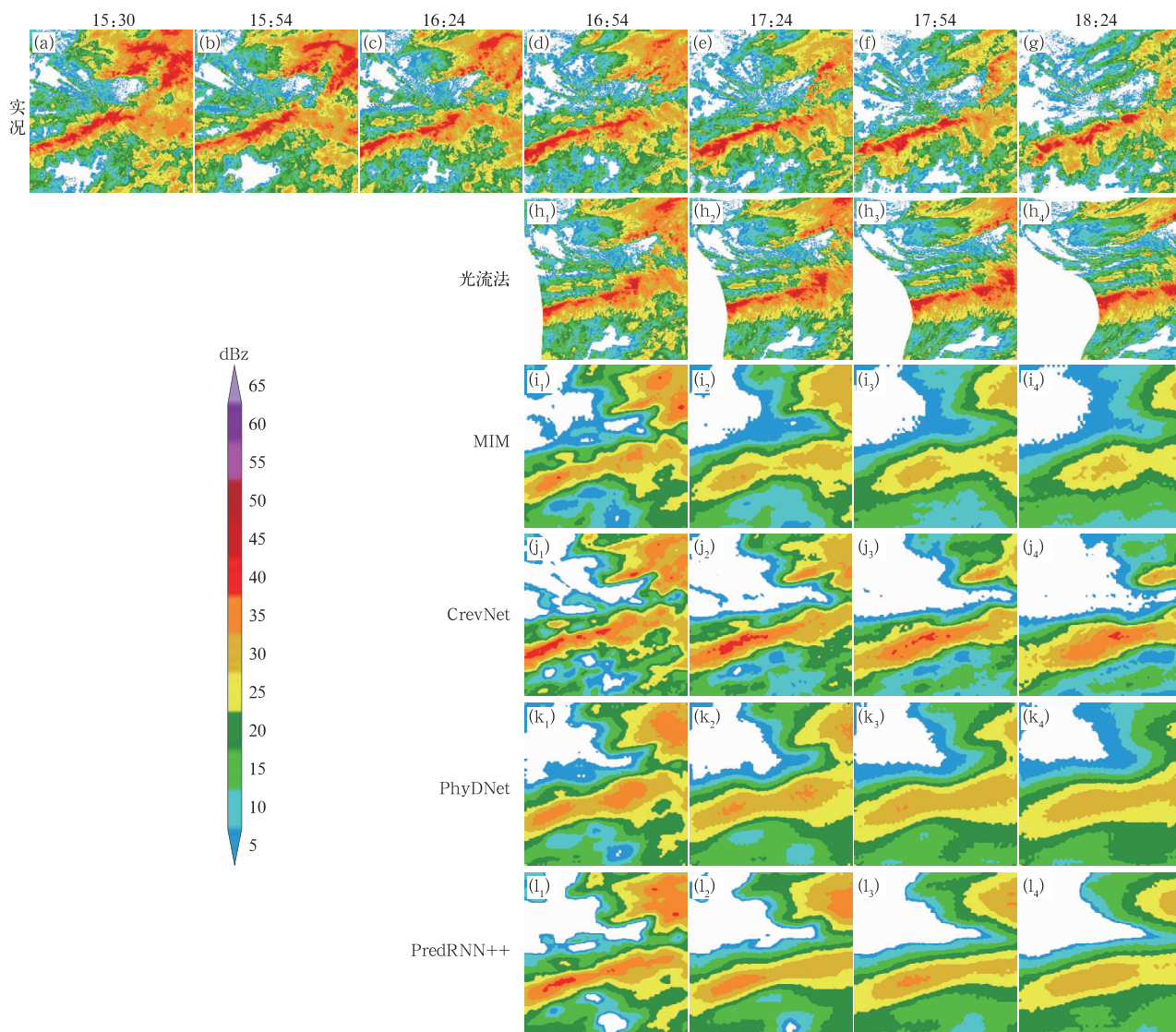


图 5 同图 2, 但为 2016 年 7 月 1 日 15:30—18:24

(a) 15:30, (b) 15:54, (c) 16:24, (d) 16:54, (e) 17:24, (f) 17:54, (g) 18:24

Fig. 5 Same as Fig. 2, but from 15:30 BT to 18:24 BT 1 July 2016

(a) 15:30 BT, (b) 15:54 BT, (c) 16:24 BT, (d) 16:54 BT,

(e) 17:24 BT, (f) 17:54 BT, (g) 18:24 BT

带状回波的面积较实况稍有偏大。另外,从强度上看,所有算法预报的回波强度较实况都偏弱,其中 MIM 算法偏弱的最明显,而 CrevNet 算法强度最强,与实况最为吻合。

上述四个不同回波形态、不同发展趋势个例的分析结果表明:光流法由于其“刚体边界”以及对非线性运动的拟合,随着外推时间的增长,光流场使得预报图片的形变就越明显,导致其对回波位置预报误差较大,尤其在 30 min 之后这种误差表现得更加明显。此外,由于亮度恒定假设的限制,使得光流法

缺乏回波强度变化的预报能力。而机器学习算法不仅对回波运动的把握能力明显强于光流法,而且还具备一定回波强度变化的预报能力,并且对回波强度和面积变化趋势时间节点的预报也与实况基本一致,但所有算法预报的回波强度较实况都偏弱。而造成各种机器学习算法预报的回波强度偏弱的原因主要有以下两个方面:一方面各算法所采用的网络结构在提取空间特征的过程中不可避免地存在信息损失,预报时效越长,信息损失得越明显,从而导致预报的回波强度较实况偏弱;另一方面还可能与损失

函数的设定有关,由于以 MSE 为损失函数平均了整幅图像的误差,使得在预报的后期时段像素值趋于平均,导致所预报的回波强度偏弱,在视觉上产生逐渐“模糊化”的过程(Zhao et al, 2017; 陈训来等, 2021)。后期将图像梯度差 (gradient difference loss, GDL) 引入损失函数中,进行多损失函数的加权试验(Nie et al, 2018),让算法更加注重图像整体结构而忽略局部噪音影响,同时探索不同预报方式的可行性,改进算法对空间信息提取的效果,让新算法预报的图像更接近实际的回波强度。

3.2 整体评分

为满足户外重大活动气象保障临近预报精细化服务的需求,以 10 dBz 为回波强度阈值,对 622 组检验样本 2 h 的平均得分(预报时间间隔为 6 min)进行再平均(结果见表 2),可以看到四种机器学习算法中, MIM 算法的 MSE 和 FAR 最低、SSIM 最高,表明其对大多数回波位置和整体回波形状的预报较其他算法更为准确;而 PredRNN++ 的 POD 和 CSI 最高,表明该算法对大多数强度回波预报最

表 2 平均的 MSE、SSIM、POD、FAR 和 CSI

Table 2 The average MSE, SSIM, POD, FAR and CSI of algorithms

算法	MSE/dBz	SSIM	POD	FAR	CSI
光流法	42.257	0.620	0.590	0.279	0.486
MIM	31.796	0.722	0.640	0.158	0.578
CrevNet	36.701	0.665	0.609	0.187	0.538
PhyDNet	32.592	0.716	0.677	0.188	0.592
PredRNN++	33.306	0.675	0.700	0.195	0.606

为准确。但四种机器学习算法无论是哪种检验指标都要明显好于光流法,其中 SSIM、POD、CSI 三种指标的提升幅度在 3.2%~24.7%, MSE 和 FAR 两种指标降幅在 13.1%~43.3%。

3.3 评分随时间的演变

从各种评分指标随时间的演变来看(图 6),在初始的 30 min 以内,除 CrevNet 算法外,其余三种算法和光流法的预报能力比较接近,都表现出 MSE、FAR 的缓慢上升以及 SSIM、POD、CSI 的缓慢下降,而 CrevNet 算法的 MSE 和 FAR 几乎持平,甚至略有降低,而且 SSIM、POD 甚至还略有上升的趋势,这种差异可能与其特有的结构有关,即 CrevNet 算法对预报的回波采取预测图像和实况回波图像相结合的策略,即对于靠前时刻的回波图像,它重点考虑输入的第 8~10 张实况回波图像,而后随着预报时刻的延长这种权重不断降低。换言之,在刚开始进行预测的半小时内,由于过分考虑了输入的第 8~10 张实况回波图像,导致其所预报的回波位置出现了较大偏差,因而 MSE 和 FAR 较其余算法明显偏大, POD、SSIM 和 CSI 则明显偏低,

但 30 min 之后这种权重减弱显著,与其他算法表现一致。

30 min 以后,各种机器学习算法和光流法都随着预报时效的延长,预报能力逐步下降,但所有指标都显示机器学习算法较光流法下降得更加缓慢,尤其是 60 min 以后,光流法的降幅进一步增加,而机器学习算法降幅相对平稳,两者之间的差距随着预报时效的延长而显著增加。以 MSE 为例,30 min 时光流法与 MIM 算法的 MSE 差值为 9.43 (dBz)²,到 120 min 时两者之间的 MSE 差值扩大到了 11.64 (dBz)²,这与顾建峰等(2020)和陈训来等(2021)的研究结果一致。此外,四种机器学习算法针对不同评分指标在不同预报时效内的变化趋势也不尽相同: MIM 算法的 MSE、SSIM、POD 和 FAR 在 30~120 min 内变化得最慢; PhyDNet 算法的 CSI 在 30~120 min 变化最慢; CrevNet 算法在 30 min 内的变化与其他算法相比有较大不同,预报效果最差,但在 30~120 min 内,其和 PredRNN++ 算法在 SSIM、POD 和 FAR 上的变化趋势几乎一致,只是 CrevNet 的 CSI 下降较 PredRNN++ 稍快。

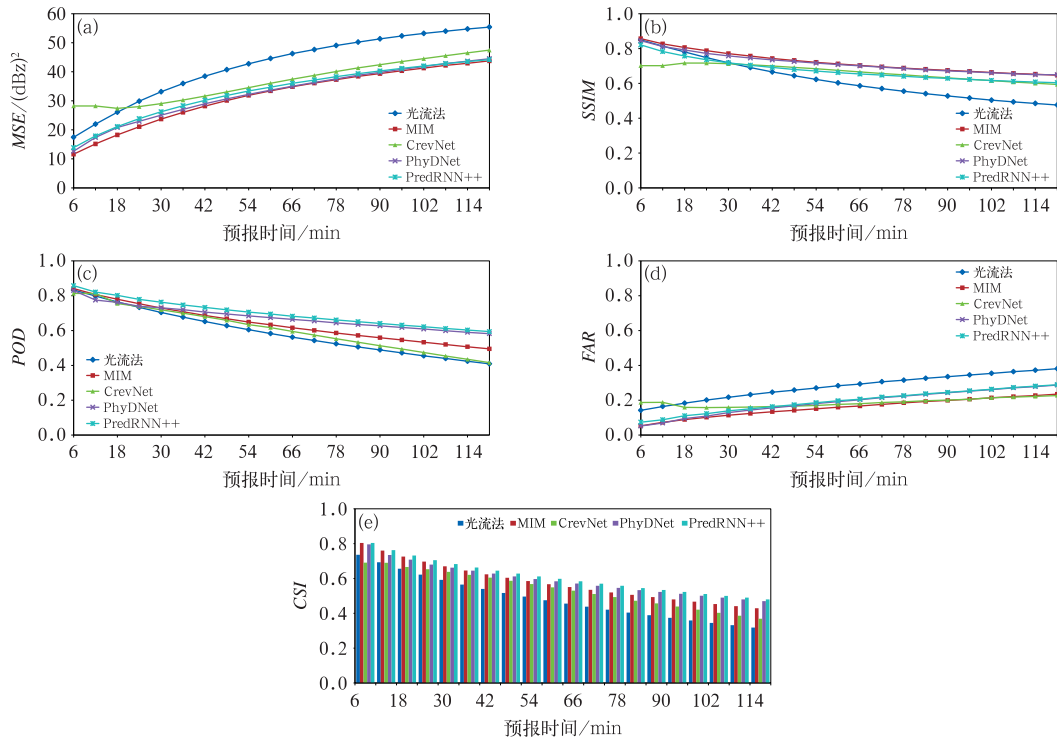


图 6 平均的(a)MSE、(b)SSIM、(c)POD、(d)FAR 和(e)CSI 随时间的演变
(横坐标为预报时间,预报时效为 0~120 min,时间间隔为 6 min)

Fig. 6 The average evaluation of (a) MSE, (b) SSIM, (c) POD, (d) FAR and (e) CSI
(The abscissa axis is forecast time, 0~120 min, the interval is 6 min)

3.4 不同阈值回波的检验

为了客观衡量各算法对于不同强度回波的预报能力,进行了不同阈值回波的检验评估(表 3)。整体来看,各算法随着回波强度的增加,CSI 和 POD 都迅速降低,FAR 则陡然上升,但各指标下降和上

升最快的区间有一定的差异:CSI 和 POD 在 20~30 dBz 的区间下降最快,而光流法和 CrevNet 算法的 FAR 在此区间上升最快,MIM、PredRNN++ 和 PhyDNet 算法则在 40~50 dBz 急剧上升。

逐个指标来看,PredRNN++ 算法在所有强度上 CSI 均表现最佳,MIM、PhyDNet 算法对 ≥ 40 dBz

表 3 各算法对不同阈值回波预报的平均评分表(预报时效 2 h,预报时间间隔为 6 min)

Table 3 The average score table of each algorithm for radar echo prediction with different thresholds (lead time: 2 h, forecast interval: 6 min)

指标	回波阈值/dBz	光流法	MIM	CrevNet	PhyDNet	PredRNN++
CSI	20	0.431	0.510	0.481	0.513	0.546
	30	0.236	0.297	0.244	0.279	0.323
	40	0.117	0.171	0.136	0.166	0.202
	50	0.168	0.154	0.185	0.132	0.195
POD	20	0.535	0.579	0.569	0.597	0.658
	30	0.339	0.338	0.312	0.317	0.397
	40	0.193	0.189	0.194	0.181	0.246
	50	0.262	0.163	0.284	0.161	0.212
FAR	20	0.344	0.221	0.274	0.243	0.270
	30	0.631	0.335	0.531	0.357	0.400
	40	0.889	0.490	0.778	0.458	0.482
	50	0.970	0.903	0.911	0.911	0.791

回波的预报较光流法好,而 CrevNet 算法则对 ≥ 50 dBz 回波的预报较光流法好;对 ≥ 20 dBz、 ≥ 30 dBz 和 ≥ 40 dBz 回波预报,POD 得分最高的均为 PredRNN++ 算法,而对 ≥ 50 dBz 回波预报,则是 CrevNet 算法的 POD 得分最高;从 FAR 指标检验来看, ≥ 20 dBz 和 ≥ 30 dBz 回波预报得分最低的都为 MIM 算法, ≥ 40 dBz 回波预报得分最低则为 PhyDNet 算法,而 PredRNN++ 算法则对 ≥ 50 dBz 回波预报的虚警率最低。

4 结论与讨论

本文基于 PredRNN++、MIM、CrevNet 和 PhyDNet 四种机器深度学习算法,利用 2012 年 6 月 1 日至 2019 年 12 月 31 日武汉地区的雷达和降水资料,以 MSE 为损失函数,对武汉地区雷达回波临近预报的性能进行了初步探讨,得到以下主要结论:

(1)通过 622 组随机样本的检验表明:MIM 算法的 MSE 和 FAR 最低、SSIM 最高,PredRNN++ 的 POD 和 CSI 最高。但这四种机器学习算法无论是哪种检验指标均要明显好于光流法,其中 SSIM、POD、CSI 三种指标的提升幅度约在 3.2%~24.7%,MSE 和 FAR 两种指标降幅约在 13.1%~43.3%。

(2)从各指标的时间演变来看:30 min 以内除 CrevNet 算法外,PredRNN++、MIM 和 PhyDNet 算法与光流法的预报能力较为接近,30 min 以后机器算法和光流法都随着预报时效的延长,预报能力均下降,但机器学习算法较光流法下降得更加缓慢,尤其是 60 min 以后光流法的降幅进一步增加,因此两者之间的差距进一步增大,显示出机器学习长预报时效的优势。此外,机器学习算法之间对不同评分指标在不同时段下降速度的反映不尽相同:MIM 算法的 MSE、SSIM、POD 和 FAR 在 30~120 min 内变化最慢,PhyDNet 的 CSI 在 30~120 min 变化最慢,CrevNet 和 PredRNN++ 算法在 SSIM、POD 和 FAR 上几乎一致,但 CrevNet 的 CSI 下降较 PredRNN++ 稍快。

(3)从不同阈值强度回波的评分来看:PredRNN++ 算法在所有强度上 CSI 均表现最佳,MIM、PhyDNet 算法对 ≥ 40 dBz 回波的预报较光流法好,而 CrevNet 则对 ≥ 50 dBz 回波的预报较光流法

好;POD 方面, ≥ 20 dBz、 ≥ 30 dBz 和 ≥ 40 dBz 回波得分最高的均为 PredRNN++ 算法, ≥ 50 dBz 以上则是 CrevNet 算法;FAR 方面, ≥ 20 dBz 和 ≥ 30 dBz 回波得分最低的都为 MIM 算法, ≥ 40 dBz 最低则为 PhyDNet, ≥ 50 dBz 最低则为 PredRNN++。

(4)四个不同回波形态、不同发展趋势个例的分析结果表明:光流法由于其固有缺陷,对回波强度的变化缺乏预报能力,并且 30 min 以后对回波移速和移向的估计也出现明显偏差,而且这种偏差随着预报时效的增长不断增加,最终导致其所预报的回波位置与实况的误差越来越大,而机器学习算法不仅对回波运动的把握能力明显强于光流法,而且还具备一定回波强度变化的预报能力,并且对回波强度和面积变化趋势时间节点的预报也与实况基本一致,但所有算法预报的回波强度较实况都偏弱。

(5)就四种机器学习算法而言:MIM 算法对大多数回波的位置和整体回波形状的把握最好,PredRNN++ 算法对大多数强度回波预报最为准确,CrevNet 算法对 50 dBz 以上强回波的预报能力最强,但 MIM、PhyDNet 和 PredRNN++ 算法所预报的回波强度较实况明显偏弱,尤其是 MIM 偏弱得非常显著,而 CrevNet 则最强与实况强度最为接近,但其强回波范围较实况略偏大。

虽然机器深度学习算法较光流法表现出一定程度的优势,但也存在一些问题:首先,从时间演变来看,机器学习算法预测的回波在 60 min 之后逐渐开始“雾化”,这种“模糊化”作用会让小面积的最强回波被平滑掉,并“泛化”出较大的次强回波的范围,因此在业务应用过程中,对于 60 min 后,尤其是 90 min 后,机器学习预报出回波面积显著增大的区域,需要特别警惕降水区域可能有显著的变化。此外,这种过度平滑的回波预报图像将会对机器算法的检验评分也产生一定程度的影响,比如对于一个实况上回波强度明显减弱的个例,将导致其评分偏高,而对于一个实况上回波强度明显增强的个例,又将导致其评分偏低,而具体定量的影响则可能与各算法对回波强度整体的平滑程度、检验样本中不同的回波强度演变趋势个例的占比等因素有关。其次,由于大多数雷暴的平均生命期仅为 30 min,对于生命期短于这个时间的雷暴,外推预报的意义不大(俞小鼎等,2012),只有对那些生命期较长的对流系统,外推预报才具有意义。最后,仅基于雷达回波的外推预报仍具有一定的局限性,快速更新循环同

化预报的高分辨率中尺度数值模式与不断发展的外推技术的融合将是未来临近预报的发展方向(王丹等,2014)。

参考文献

- 陈锦鹏,冯业荣,蒙伟光,等,2021. 基于卷积神经网络的逐时降水预报订正方法研究[J]. 气象,47(1):60-70. Chen J P, Feng Y R, Meng W G, et al, 2021. A correction method of hourly precipitation forecast based on convolutional neural network[J]. Meteor Mon, 47(1):60-70(in Chinese).
- 陈训来,刘军,郑群峰,等,2021. 基于卷积门控循环单元神经网络的临近预报方法研究[J]. 高原气象,40(2):411-423. Chen X L, Liu J, Zheng Q F, et al, 2021. A study on radar echo nowcasting based on convolutional gated recurrent unit neural network[J]. Plateau Meteor, 40(2):411-423(in Chinese).
- 陈元昭,林良勋,王蕊,等,2019. 基于生成对抗网络 GAN 的人工智能临近预报方法研究[J]. 大气科学学报,42(2):311-320. Chen Y Z, Lin L X, Wang R, et al, 2019. A study on the artificial intelligence nowcasting based on generative adversarial networks [J]. Trans Atmos Sci, 42(2):311-320(in Chinese).
- 曹春燕,陈元昭,刘东华,等,2015. 光流法及其在临近预报中的应用[J]. 气象学报,73(3):471-480. Cao C Y, Chen Y Z, Liu D H, et al, 2015. The optical flow method and its application to nowcasting[J]. Acta Meteor Sin, 73(3):471-480(in Chinese).
- 顾建峰,周国兵,刘伯骏,等,2020. 人工智能技术在重庆临近预报业务中的初步研究与应用[J]. 气象,46(10):1286-1296. Gu J F, Zhou G B, Liu B J, et al, 2020. Study on artificial intelligence technology and its application to Chongqing operational nowcasting[J]. Meteor Mon, 46(10):1286-1296(in Chinese).
- 韩雷,王洪庆,林静静,2008. 光流法在强对流天气临近预报中的应用[J]. 北京大学学报(自然科学版),44(5):751-755. Han L, Wang H Q, Lin J J, 2008. Application of optical flow method to nowcasting convective weather[J]. Acta Sci Nat Univ Pekinensis, 44(5):751-755(in Chinese).
- 黄骅文,蔡荣辉,姚蓉,等,2021. 深度学习网络在降水相态判识和预报中的应用[J]. 气象,47(3):317-326. Huang J W, Cai R H, Yao R, et al, 2021. Application of deep learning method to discrimination and forecasting of precipitation type[J]. Meteor Mon, 47(3):317-326(in Chinese).
- 王丹,王改利,刘黎平,等,2014. 基于雷达回波外推和中尺度模式预报的短时降水对比分析[J]. 高原气象,33(3):811-822. Wang D, Wang G L, Liu L P, et al, 2014. Comparisons analysis on short-term precipitation between the radar-based extrapolation and the meso-scale numerical model weather prediction[J]. Plateau Meteor, 33(3):811-822(in Chinese).
- 吴剑坤,陈明轩,秦睿,等,2019. 变分回波跟踪算法及其在对流临近预报中的应用试验[J]. 气象学报,77(6):999-1014. Wu J K, Chen M X, Qin R, et al, 2019. The variational echo tracking method and its application in convective storm nowcasting[J]. Acta Meteor Sin, 77(6):999-1014(in Chinese).
- 吴涛,万玉发,沃伟峰,等,2013. SWAN 系统中雷达反射率因子质量控制算法及其应用[J]. 气象科技,41(5):809-817. Wu T, Wan Y F, Wo W F, et al, 2013. Design and application of radar reflectivity quality control algorithm in SWAN[J]. Meteor Sci Technol, 41(5):809-817(in Chinese).
- 许小峰,2018. 从物理模型到智能分析—降低天气预报不确定性的新探索[J]. 气象,44(3):341-350. Xu X F, 2018. From physical model to intelligent analysis: a new exploration to reduce the uncertainty of weather forecast[J]. Meteor Mon, 44(3):341-350(in Chinese).
- 徐亚钦,翟国庆,黄旋旋,等,2011. 基于 TREC 法以多重动态区域反演风场[J]. 浙江大学学报(工学版),45(10):1738-1745. Xu Y Q, Zhai G Q, Huang X X, et al, 2011. Wind retrieval by using multi-scale dynamic region based on TREC method[J]. J Zhejiang Univ(Eng Sci), 45(10):1738-1745(in Chinese).
- 徐月飞,顾松山,黄兴友,等,2011. 计算几何法在风暴识别中的应用[J]. 热带气象学报,27(2):173-178. Xu Y F, Gu S S, Huang X Y, et al, 2011. Application of computational geometry in storm identification[J]. J Trop Meteor, 27(2):173-178(in Chinese).
- 俞小鼎,周小刚,王秀明,2012. 雷暴与强对流临近天气预报技术进展[J]. 气象学报,70(3):311-337. Yu X D, Zhou X G, Wang X M, 2012. The advances in the nowcasting techniques on thunderstorms and severe convection[J]. Acta Meteor Sin, 70(3):311-337(in Chinese).
- 张蕾,魏鸣,李南,等,2014. 改进的光流法在回波外推预报中的应用[J]. 科学技术与工程,14(32):133-137, 148. Zhang L, Wei M, Li N, et al, 2014. Improved optical flow method application to extrapolate radar echo[J]. Sci Technol Eng, 14(32):133-137, 148(in Chinese).
- 赵悦,陈家华,章建军,等,2007. 基于中值滤波和小波变换的天气雷达回波图像处理[J]. 气象科学,27(1):63-68. Zhao Y, Chen J H, Zhang J J, et al, 2007. Weather radar echo image processing based on median filter and wavelet transform[J]. Sci Meteor Sin, 27(1):63-68(in Chinese).
- 周康辉,郑永光,韩雷,等,2021. 机器学习在强对流监测预报中的应用进展[J]. 气象,47(3):274-289. Zhou K H, Zheng Y G, Han L, et al, 2021. Advances in application of machine learning to severe convective weather monitoring and forecasting[J]. Meteor Mon, 47(3):274-289(in Chinese).
- Austin G L, Bellon A, 1982. Very Short-Range Forecast of Precipitation by the Objective Extrapolation of Radar and Satellite Data [C]//Browning K, Nowcasting. Salt Lake City: Academic Press.
- Bechini R, Chandrasekar V, 2017. An enhanced optical flow technique for radar nowcasting of precipitation and winds[J]. J Atmos Ocean Technol, 34(12):2637-2658.
- Dixon M, Wiener G, 1993. TITAN: Thunderstorm identification, tracking, analysis, and nowcasting—a radar based methodology [J]. J Atmos Ocean Technol, 10(6):785-797.
- Gibson J, 1979. The Ecological Approach to Visual Perception[M]. Boston: Houghton Mifflin.
- Johnson J T, Mackeen P L, Witt A, et al, 1998. The storm cell

- identification and tracking algorithm; an enhanced WSR-88D algorithm[J]. *Wea Forecasting*, 13(2):263-276.
- Le Guen V, Thome N, 2020. Disentangling physical dynamics from unknown factors for unsupervised video prediction[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Seattle:11471-11481.
- Li L, Schmid W, Joss J, 1995. Nowcasting of motion and growth of precipitation with radar over a complex orography[J]. *J Appl Meteor Climatol*, 34(6):1286-1300.
- Lucas B T, Kanade T, 1981. An iterative image registration technique with an application to stereo vision[C]//Proceedings of the 1981 DARPA Imaging Understanding Workshop:121-130.
- Nie D, Trullo R T, Lian J, et al, 2018. Medical image synthesis with deep convolutional adversarial networks[J]. *IEEE Trans Biomed Eng*, 65(12):2720-2730.
- Rinehart R E, 1981. A pattern recognition technique for use with conventional weather radar to determine internal storm motions [J]. *Atmos Technol*, 13:119-134.
- Rinehart R E, Garvey E T, 1978. Three-dimensional storm motion detection by conventional weather radar[J]. *Nature*, 273(5660):287-289.
- Rosenfeld D, 1987. Object method for analysis and tracking of convective cell as seen by radar[J]. *Atmos Res*, 191:1-11.
- Shi X J, Chen Z R, Wang H, et al, 2015. Convolutional LSTM network; a machine learning approach for precipitation nowcasting [C]//Proceedings of the 28th International Conference on Neural Information Processing System. Montreal: MIT Press: 802-810.
- Shi X J, Gao Z H, Lausen L, et al, 2017. Deep learning for precipitation nowcasting; a benchmark and a new model[C]//Proceedings of the 31st International Conference on Neural Information Processing System. Long Beach: Curran Associates Inc. : 5622-5632.
- Wang G L, Yang J, Wang D, et al, 2016. A quantitative comparison of precipitation forecasts between the storm-scale numerical weather prediction model and auto-nowcast system in Jiangsu, China[J]. *Atmos Res*, 181:1-11.
- Wang Y B, Gao Z F, Long M S, et al, 2018. PredRNN++: towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning[C]//Proceedings of the 35th International Conference on Machine Learning. Stockholm:5123-5132.
- Wang Y B, Long M S, Wang J M, et al, 2017. Predrnn: recurrent neural networks for predictive learning using spatiotemporal LSTMs[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems(NeurIPS). Long Beach: Curran Associates Inc. :879-888.
- Wang Y B, Zhang J J, Zhu H Y, et al, 2019. Memory in memory: a predictive neural network for learning higher-order non-stationarity from spatiotemporal[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR). Long Beach: IEEE:9154-9162.
- Wang Z, Bovik A C, Sheikh H R, et al, 2004. Image quality assessment: from error to visibility to structural similarity[J]. *IEEE Trans Image Process*, 13(4):600-612.
- Wilson J W, Crook N A, Mueller C K, et al, 1998. Nowcasting thunderstorms; a status report[J]. *Bull Amer Meteor Soc*, 79(10):2079-2100.
- Wilson J W, Feng Y R, Chen M, et al, 2010. Nowcasting challenges during the Beijing Olympics; Successes, failures and implications for future nowcasting systems[J]. *Wea Forecasting*, 25(6):1691-1714.
- Yu W, Lu Y C, Steve E, et al, 2020. Efficient and information-preserving future frame prediction and beyond[R]//Addis Ababa:8th International Conference on Learning Representations(ICLR).
- Zhao H, Gallo O, Frosio I, et al, 2017. Loss function for image restoration with neural networks[J]. *IEEE Trans Comput Imaging*, 3(1):47-57.

(本文责编:俞卫平)