Scientific Journal of Intelligent Systems Research Volume 8 Issue 6, 2026
ISSN: 2664-9640

Deep Learning for Precipitation Nowcasting

Yifan Tao*
Anhui Wenda University Of Information Engineering, Hefei, China
*Corresponding Author

Abstract

This paper provides a comprehensive review of deep learning approaches for precipitation nowcasting, a challenging spatiotemporal prediction task in meteorology. We systematically categorize existing methods into four major families: recurrent neural network (RNN)-based models, convolutional encoder-decoder architectures, generative adversarial networks (GANs), and transformer-based models. For each category, we analyze its underlying modeling principles, representative architectures, strengths, and limitations. Furthermore, we summarize widely used radar datasets and evaluation metrics to provide a unified benchmark perspective. Finally, we discuss key challenges such as prediction blurriness, class imbalance in extreme precipitation events, and computational inefficiency, and outline promising future directions including multimodal fusion, efficient attention mechanisms, and self-supervised pretraining. This review aims to provide a structured reference for both researchers and practitioners in the field of data-driven weather forecasting.

Keywords

Precipitation Nowcasting; Deep Learning; Convolutional LSTM; Generative Adversarial Networks.

1. Introduction

1.1. Research Background and Significance

Precipitation nowcasting aims to predict how rainfall evolves over short time ranges, usually within 0–2 hours, for specific local areas. It has become important in systems that need fast decisions while still requiring detailed spatial information. Unlike conventional numerical weather prediction (NWP), nowcasting relies more on high-resolution spatiotemporal modeling and is designed to produce results shortly after radar data are collected, under tight time constraints.

This type of forecasting is widely used in fields where weather conditions can directly affect safety and daily operations, such as aviation, farming, and marine transportation[1]. In these situations, short-term rainfall information helps support operational planning, reduce potential risks, and improve preparedness when weather changes rapidly, especially during convective storm development[2].

Even with recent progress in data-driven and physical modeling approaches, it is still difficult to accurately represent rapidly changing or extreme rainfall processes. The 2021 Zhengzhou extreme rainfall event is a typical example, where existing forecasting systems failed to fully capture the intensity and evolution of the storm, leading to serious social and economic consequences and revealing the limitations of current methods[3].

1.2. Evolution and Limitations of Traditional Methods

Before deep learning methods became widely used, precipitation nowcasting was mainly studied through two directions: numerical weather prediction (NWP) and optical-flow-based radar extrapolation.

NWP is built on the numerical solution of atmospheric governing equations, given certain initial and boundary conditions. It combines different types of observations, such as radar, satellite, and surface measurements, to estimate atmospheric states including wind, moisture, and precipitation. While this approach has a solid physical basis, it does not perform well in short-term forecasting. One reason is that atmospheric systems are highly sensitive to initial states, which makes small errors grow quickly over time[4]. In addition, the computational demand is heavy, so model updates cannot be done frequently enough for real-time nowcasting[5].

Another line of methods treats radar images as motion sequences and applies optical flow to estimate the displacement between consecutive frames[6]. These approaches are relatively fast and can keep local spatial patterns in the short term, which makes them useful for very short prediction windows. However, they depend on the assumption that motion is approximately linear. In practice, precipitation systems often evolve in a nonlinear way, such as developing, splitting, or weakening, which these methods cannot represent well. This usually leads to error accumulation as the prediction horizon increases[7].

In general, the two paradigms reflect a trade-off between physical modeling and real-time responsiveness. This gap has encouraged the development of learning-based approaches that rely on data-driven representations.

1.3. Opportunities from Deep Learning

Recent improvements in GPU hardware and the availability of large-scale datasets have noticeably accelerated progress in deep learning research[8]. Within this context, precipitation nowcasting has increasingly been treated as a typical spatiotemporal prediction problem where data-driven approaches can be effectively applied.

Instead of relying on manually designed features, deep learning models can extract useful representations directly from raw radar observations in an end-to-end learning process. This reduces the need for hand-crafted preprocessing steps that were common in earlier approaches. Another advantage is that these models are able to handle nonlinear relationships in the data, which is important for representing complex and evolving atmospheric processes that are difficult to describe with simpler formulations. After training, the inference stage is usually efficient, which makes it feasible to use such models in time-sensitive forecasting systems. In some cases, knowledge learned from one region can also be reused in another through transfer learning, especially when local data are limited.

Taken together, these properties explain why deep learning has become a practical direction for precipitation nowcasting, and they have also led to continuous exploration of different network designs and training strategies.

2. Enhanced Discussion Section

2.1. Unified Perspective

Precipitation nowcasting can be viewed as a spatiotemporal prediction problem built on radar reflectivity observations. In practice, radar measurements form a time-ordered sequence, where each frame describes the spatial distribution of reflectivity values (dBZ), which are commonly used to indicate rainfall intensity.

ISSN: 2664-9640

At each time step, the radar data can be written as a tensor $X_t \in R^{H\times W\times C}$, where H and W represent spatial dimensions and C denotes the number of channels. Given a historical sequence $\{X_{t-m+1},\ldots,X_t\}$, the objective is to estimate the upcoming sequence $\{X_{t+1},\ldots,X_{t+n}\}$.

From a modeling perspective, this problem is often treated as a sequence-to-sequence learning task, in which a parameterized function is used to approximate the conditional evolution of future radar states given past observations. In operational meteorology, radar reflectivity is also commonly transformed into rainfall intensity through empirical Z–R relationships, although the corresponding parameters are not fixed and may change depending on regional characteristics and climatic environments.

2.2. Intrinsic Challenges

Precipitation nowcasting remains a difficult problem for several fundamental reasons. One major difficulty comes from the underlying atmospheric processes themselves. Rainfall evolution is governed by nonlinear interactions in space and time, involving mechanisms such as convection, advection, and thermodynamic exchange. These processes often produce complicated behaviors, including the formation, weakening, splitting, and merging of precipitation cells, which are not well captured by simplified motion assumptions.

Another issue is related to the distribution of observed data. In most radar datasets, light rain occurs frequently, while heavy or extreme events appear only occasionally. This imbalance makes learning difficult, especially when commonly used objectives like mean squared error are applied, since they tend to prioritize frequent low-intensity samples and reduce the model's ability to respond to rare but important extremes.

Generalization across regions is also non-trivial. Precipitation characteristics are not consistent across different climatic zones, and significant variations can be observed between tropical, temperate, and monsoon-dominated areas. As a result, models trained in one region often struggle to perform well when transferred to another without adaptation.

In addition, predicted results often suffer from a smoothing effect. This is partly caused by the fact that future atmospheric evolution can follow multiple possible paths, while deterministic training objectives encourage models to regress toward an average outcome. Although recent generative approaches help reduce this issue to some extent, producing sharp and realistic forecasts is still an open problem.

3. Challenges

3.1. Recurrent-based Methods

Early deep learning studies on precipitation nowcasting largely relied on recurrent architectures. These approaches adapt recurrent neural networks to handle spatiotemporal inputs by introducing convolutional operations, so that spatial patterns and temporal evolution can be processed together within the same framework.

Such models are capable of producing competitive forecasting results, especially for short-term prediction tasks. However, this performance often comes at the cost of increased computational complexity, which can make training and inference relatively expensive in practical settings.

3.1.1. Basic Architecture: ConvLSTM

Before ConvLSTM was introduced, long short-term memory (LSTM) networks were commonly used for sequential data modeling. While LSTM is effective in addressing gradient vanishing through its gating mechanism, it is originally designed for one-dimensional sequences. When directly applied to image sequences, the spatial layout must be flattened, which breaks the original spatial structure and leads to a rapid increase in the number of parameters.

ISSN: 2664-9640

ConvLSTM was proposed to overcome this issue by integrating convolution operations into both the input-to-state and state-to-state transitions[9]. In this way, spatial correlations are preserved while temporal dependencies are still modeled within the recurrent framework. Because the same convolutional kernels are shared across spatial positions, the model becomes more parameter-efficient and naturally maintains translation consistency.

This design laid the groundwork for many later spatiotemporal forecasting architectures, such as encoder–decoder frameworks and autoregressive prediction strategies, and also inspired models like PredRNN and MIM. However, the architecture remains relatively limited in depth, which restricts its ability to capture highly complex dynamics. Over longer prediction intervals, errors tend to accumulate and produce increasingly blurred outputs. Moreover, interactions across different recurrent layers are not explicitly modeled, which limits the flow of temporal information across hierarchical representations.

3.1.2. Hierarchical Memory Modeling

To improve the limited representational ability of ConvLSTM, later research explored deeper architectures as well as more explicit memory design strategies. One representative direction is PredRNN, which introduces a unified memory unit that connects different layers over both space and time[10]. Through a specially designed update route, information is able to propagate not only forward in time but also across hierarchical levels, which helps the model capture longer-range dependencies in spatiotemporal sequences.

Building on this idea, MIM proposes a different perspective by separating motion patterns into components with different temporal characteristics[11]. Inspired by classical differencing techniques in time-series analysis, it introduces a mechanism that highlights temporal changes explicitly rather than modeling raw states directly. Two memory streams are maintained in parallel, one focusing on relatively stable evolution and the other targeting abrupt variations, allowing the model to respond more flexibly to different types of dynamics.

Despite these improvements in accuracy, such architectures tend to be computationally demanding. The increased depth and additional memory units lead to higher training cost and greater memory usage. When applied to medium-resolution radar data, these models often require substantial GPU resources, which makes deployment in real-time operational systems less practical.

3.1.3. Attention-Enhanced Recurrent Models

Conventional convolution operations mainly focus on local neighborhoods, which makes it difficult for them to capture dependencies over large spatial ranges[12]. In precipitation nowcasting, this limitation becomes more pronounced because weather systems often span wide areas, and conditions in one region may be influenced by distant atmospheric activity.

To improve global context modeling, later studies introduced attention mechanisms into recurrent frameworks. For example, SA-ConvLSTM[13] incorporates a self-attention module during the memory update process, allowing the model to aggregate information from a broader spatial context instead of relying only on local convolutional features. In a similar direction, E3D-LSTM extends the modeling capacity by using three-dimensional convolutions to jointly process spatial and temporal information[14], while also introducing attention-based gating to strengthen long-range dependency modeling across multiple time steps.

However, these improvements come with a significant increase in computational cost. In particular, self-attention mechanisms scale poorly with spatial resolution due to their quadratic complexity, which makes training on high-resolution radar data very expensive. In practice, this often forces such models to operate on reduced-resolution inputs or within limited spatial domains.

Overall, recurrent-based approaches enhanced with attention mechanisms improve modeling flexibility, but they still face challenges in efficiency and scalability when applied to large-scale, high-resolution forecasting tasks.

3.2. Convolutional Encoder–Decoder Methods

Convolutional encoder–decoder architectures, with U-Net as a typical example, offer a different way of handling precipitation nowcasting without relying on recurrent computation[15]. The network is built in a symmetric form, where feature representations are gradually compressed and then reconstructed, while skip connections are used to pass fine-scale spatial details directly from earlier layers to later stages.

This type of structure fits precipitation forecasting well, since rainfall systems usually contain patterns at multiple scales, ranging from localized convective cells to broader stratiform regions. By processing information at different levels of resolution, encoder–decoder models can learn hierarchical spatial representations while still producing outputs with consistent spatial dimensions.

A number of improved variants have been developed on this basis, such as RainNet, SE-ResUNet, and SmaAt-UNet[16]. These models introduce modifications like attention over channels, more efficient convolution designs, and structural refinements to improve both performance and computational efficiency.

In comparison with recurrent-based approaches, encoder–decoder networks are generally easier to optimize and require less computation during training and inference. However, they do not explicitly model temporal evolution. To compensate for this, common practices include stacking multiple input frames along the channel dimension or using autoregressive prediction schemes, although these strategies may lead to limited long-term dependency modeling or gradual error accumulation over time.

3.3. Generative Adversarial Networks

Deterministic objectives such as mean squared error often lead to overly smooth forecasts in precipitation nowcasting, mainly because future atmospheric evolution can follow multiple plausible trajectories rather than a single outcome. As a result, models trained under such criteria tend to average these possibilities, which reduces sharpness in the predicted fields.

To address this limitation, generative adversarial networks (GANs) have been introduced into the forecasting framework[17]. Instead of relying only on direct supervision, these methods learn the data distribution through a competition between a generator, which produces future radar sequences, and a discriminator, which attempts to distinguish generated outputs from real observations. In this setting, the generator is encouraged not only to match the ground truth but also to produce more realistic spatial structures.

In precipitation nowcasting applications, variants such as GA-ConvGRU adopt recurrent-based generators to model temporal evolution while producing future radar frames. The discriminator is used to evaluate whether the generated sequences resemble real radar patterns. Training typically combines reconstruction-based objectives with adversarial feedback, which helps improve visual sharpness and structural consistency.

Further developments extend this idea by introducing discriminators operating at multiple spatial scales or across both space and time, allowing the model to better capture fine details and large-scale organization simultaneously. Empirical studies have shown that such designs can improve perceptual quality and produce more realistic precipitation structures under certain settings.

Despite these advantages, GAN-based approaches are difficult to train in practice. Instability during optimization, sensitivity to hyperparameter choices, and issues such as mode collapse are commonly reported. Moreover, compared with deterministic approaches, the overall

ISSN: 2664-9640

training process is more complex and less stable, which can limit reproducibility and deployment in operational forecasting systems.

3.4. Transformer-based Methods

Transformer-based architectures have recently gained attention in spatiotemporal forecasting tasks[18]. Unlike recurrent or convolutional designs, these models rely on self-attention as the core mechanism, which allows dependencies to be modeled across the entire input sequence without relying on step-by-step processing or local receptive fields. This design also makes it possible to process data in parallel, which is attractive for large-scale training.

Early studies such as EarthFormer and RainFormer adapt this idea to geophysical data by introducing attention mechanisms that operate jointly across spatial and temporal dimensions[19,20]. These approaches show that Transformers are capable of modeling long-range interactions within precipitation systems, which are often difficult to capture using traditional architectures.

When compared with CNN- and RNN-based methods, Transformer models differ mainly in how information is aggregated. They do not depend on local convolution or recurrent transitions, which allows information to flow more freely across space and time. This often results in better gradient behavior during training and more flexible dependency modeling over long sequences.

However, these advantages come with a clear computational cost. The self-attention operation scales poorly with both spatial resolution and sequence length, making high-resolution radar forecasting expensive. Because of this limitation, existing Transformer-based models have not yet shown consistent superiority over carefully designed recurrent or adversarial approaches in practical nowcasting scenarios.

3.5. Method Comparison and Selection

Different modeling paradigms in precipitation nowcasting tend to prioritize different aspects of performance, and none of them fully optimizes accuracy, efficiency, and scalability at the same time. Recurrent-based approaches generally achieve strong predictive performance, but this often comes with high computational cost and limited efficiency during training and inference. In contrast, encoder–decoder architectures are easier to deploy and computationally lighter, yet their ability to represent temporal evolution is relatively limited.

Generative adversarial approaches focus more on producing visually realistic outputs, although this benefit is typically accompanied by unstable optimization and increased training complexity. Transformer-based methods, on the other hand, offer a more flexible framework for modeling long-range dependencies, but their computational demand becomes a major constraint when applied to high-resolution radar data.

In practical use, the choice of model is usually driven by the target application rather than purely by model capability. Systems that prioritize accuracy tend to adopt recurrent structures, while operational forecasting systems with strict latency requirements often prefer encoder–decoder designs. Adversarial models are more commonly used when visual realism is important, and Transformer-based methods are still largely explored in research settings where scalability is the main focus.

4. Datasets and Evaluation Metrics

4.1. Open Radar Datasets

Open-access radar datasets have become an essential foundation for research on short-term precipitation nowcasting, as they provide shared benchmarks for training and evaluating different forecasting models. In recent years, meteorological organizations in various regions have released a number of standardized datasets, which has significantly improved the comparability of experimental results across studies. This section briefly reviews several representative datasets that are widely used in the literature.

One of the earliest and most frequently adopted benchmarks is HKO-7[21], released by the Hong Kong Observatory. It consists of radar observations collected between 2009 and 2015, obtained from a single radar system operating at an altitude of around 2 km. The dataset covers a spatial extent of roughly 512 km × 512 km, with each frame represented at a resolution of 480 × 480 pixels and sampled at 6-minute intervals. Because it captures relatively rich short-term motion patterns while remaining manageable in size, HKO-7 has been commonly used in early studies, including models such as TrajGRU, for evaluating spatiotemporal prediction performance.

Another widely used dataset is DWD-12, provided by the German Weather Service (Deutscher Wetterdienst)[22]. Unlike HKO-7, it is constructed as a large-scale radar mosaic covering central Europe, generated from observations of 17 radar stations over the period from 2006 to 2017. It spans approximately 900 km × 900 km with a spatial resolution of 900 × 900 pixels and a temporal resolution of 5 minutes. The dataset reflects a temperate maritime climate with highly variable precipitation structures, making it more challenging for modeling. At the same time, its large scale also leads to substantial computational demands during training.

The CIKM AnalytiCup 2017 dataset comes from a data mining competition organized by the Shenzhen Meteorological Bureau[23]. It focuses on a smaller region of about 101 km × 101 km, with matching spatial resolution and a 6-minute time interval. Although its spatial coverage is limited, it remains popular in the community due to its well-established preprocessing pipelines and available baseline implementations, which make it convenient for quick experimentation and model comparison.

Beyond these radar-only datasets, MeteoNet extends the data sources by including satellite imagery and ground station measurements in addition to radar observations, which makes it suitable for multimodal learning studies. The SEVIR dataset contains over 10,000 precipitation events, each lasting around four hours, and is often used for long-sequence forecasting tasks. In contrast, the NJU-CPOL dataset provides polarimetric radar variables, allowing researchers to explore physically informed and multi-variable modeling approaches.

Overall, these datasets differ in spatial scale, temporal resolution, and data modality. Together, they form a complementary benchmark suite that supports the development and evaluation of precipitation nowcasting methods from different perspectives.

4.2. Evaluation Metrics

Evaluating precipitation nowcasting models is not straightforward, since different aspects of forecast quality cannot be reflected by a single measurement. For this reason, studies usually rely on a set of metrics that capture different properties of the predictions, rather than depending on one indicator alone. These metrics are often discussed from three perspectives: numerical errors at the pixel level, the ability to detect precipitation events, and the visual or structural quality of the generated fields.

4.2.1. Pixel-wise Error Metrics

Pixel-wise metrics are used to measure the difference between predicted precipitation fields and ground-truth observations at individual grid locations. In practice, commonly adopted choices include mean squared error (MSE), mean absolute error (MAE), and root mean squared error (RMSE), which evaluate prediction quality from a purely numerical perspective.

These measures are straightforward to compute and behave consistently during optimization, which explains why they are frequently used as training objectives in many forecasting models. However, their usefulness is limited when it comes to capturing spatial organization in precipitation fields. Since they treat each pixel independently, they do not explicitly account for structural patterns or spatial continuity. In addition, when the underlying distribution of future

ISSN: 2664-9640

states is multi-modal, optimizing MSE often drives models toward an averaged solution, which can suppress fine-scale variations and produce overly smooth results.

4.2.2. Event-based Categorical Metrics

In operational forecasting settings, precipitation fields are often simplified into categorical events by applying a threshold, separating rain and no-rain conditions. Once this conversion is made, model performance is commonly assessed using a confusion-matrix framework that summarizes predictions in terms of correct detections, false alarms, missed events, and correct rejections.

From this perspective, several standard indicators are derived. The probability of detection (POD) reflects how many observed precipitation events are successfully identified by the model, providing a measure of detection capability. The false alarm ratio (FAR), on the other hand, describes how often the model incorrectly predicts rainfall when none is observed, which relates to forecast reliability. Among these metrics, the critical success index (CSI) is often used as a more balanced measure, as it jointly considers hits, misses, and false alarms in a single score.

$$\mathrm{CSI} = \mathrm{TP} / (\mathrm{TP} + \mathrm{FP} + \mathrm{FN}) \tag{1}$$

This metric explicitly accounts for both types of errors, penalizing cases where precipitation is missed as well as situations where rainfall is incorrectly predicted.

Another commonly used indicator is the Heidke Skill Score (HSS), which evaluates how much a forecasting model improves over a random reference[24]. The score ranges from negative values to 1, where a value of 1 corresponds to a perfect forecast, and values around zero indicate no meaningful improvement over random guessing.

Together, these measures provide evaluation results that are more aligned with practical forecasting needs in meteorological applications.

4.2.3. Perceptual Image Quality Metrics

To assess how well spatial structures are preserved in predictions, the Structural Similarity Index (SSIM) is often used as an evaluation measure. Instead of focusing only on pixel-level differences, it compares images in terms of overall brightness, contrast, and structural consistency, which together reflect how similar two fields appear from a perceptual perspective.

The score is bounded between 0 and 1, where larger values indicate closer agreement with the reference image. Compared with purely numerical error measures, SSIM places more emphasis on spatial organization, making it more appropriate for checking whether predicted precipitation fields maintain meaningful patterns such as spatial continuity, boundaries, and localized structures.

4.3. Summary

Overall, evaluating precipitation nowcasting models requires considering several complementary aspects rather than relying on a single indicator. In practice, studies usually report a combination of pixel-level error measures (such as MSE, MAE, and RMSE), event-oriented scores (including POD, FAR, CSI, and HSS), and perceptual metrics like SSIM, so that both numerical accuracy, event detection performance, and structural quality can be examined together.

It is also common to observe that these metrics do not always agree with each other. A model may achieve low numerical error while still performing poorly in event-based evaluation, or vice versa, which reflects the complex and uncertain nature of precipitation evolution. Because of this, conclusions based on a single metric can be misleading. As a result, using a multi-metric evaluation setup has become standard practice in precipitation nowcasting research.

5. Challenges and Future Directions

5.1. Core Challenge: Blurry Predictions

Despite considerable progress in deep learning methods for precipitation nowcasting, producing sharp and realistic forecasts is still a major challenge in practice. One of the main issues is that predicted results often appear overly smooth, which limits their usefulness in representing real weather structures. This limitation can be traced back to several underlying factors.

On the one hand, rainfall evolution is governed by complex atmospheric processes that are highly nonlinear and uncertain over space and time, as discussed earlier in Section 2.2. Because of this complexity, future states of precipitation are not uniquely determined, but instead may follow multiple possible trajectories. This makes deterministic forecasting inherently difficult, since a single prediction is expected to represent an entire range of plausible outcomes.

On the other hand, commonly used optimization objectives such as mean squared error introduce additional constraints. When the future distribution contains multiple possible modes, minimizing such a loss tends to drive the model toward an averaged result, which reduces variability in the output. In practice, this often leads to blurred representations of important structures, such as localized convective cells and clear precipitation boundaries.

From a practical standpoint, another difficulty arises from the imbalance in observed data. Intense rainfall events are relatively rare, yet they are often the most important in real applications. During training, however, learning is dominated by frequent low-intensity or no-rain samples, which can bias the model toward these common cases. As a result, performance on extreme events is often weaker. How to better handle this imbalance and design learning objectives that are more sensitive to rare but critical events remains an open question in this area.

5.2. Future Direction I: New Model Architectures

Transformer-based models, including Vision Transformers (ViTs), have recently shown strong performance in a range of computer vision tasks such as image classification and semantic segmentation. This success has naturally encouraged their exploration in precipitation nowcasting, where learning complex spatial–temporal patterns is a central challenge.

Work such as EarthFormer and RainFormer indicates that attention-based architectures may offer advantages over traditional recurrent models by allowing information to be exchanged across the entire spatial and temporal domain. Instead of relying on local receptive fields or step-by-step state updates, these models compute relationships directly between distant elements, which provides a more flexible way to represent long-range dependencies in radar sequences.

Despite these advantages, applying Transformers to high-resolution meteorological data remains challenging. The main issue comes from the computational cost of self-attention, which increases rapidly with both spatial size and sequence length. This makes straightforward deployment on large radar grids impractical in many cases. As a result, current research has increasingly focused on reducing this cost through more efficient attention designs, such as sparse connectivity patterns, low-rank or linear approximations, and combinations with convolutional structures.

ISSN: 2664-9640

5.3. Future Direction II: Learning Strategies and System Design

A growing line of research has focused on incorporating multiple data sources to improve precipitation forecasting. From a physical perspective, combining radar reflectivity with satellite imagery, surface measurements, and atmospheric variables such as wind can provide a more complete description of weather systems. This richer input information helps models better represent the complexity of precipitation processes.

However, using heterogeneous data in a unified framework is not straightforward. Different modalities often come from different spatial grids and sampling frequencies, which makes direct fusion difficult. In practice, alignment and interpolation are usually required before joint modeling, but these preprocessing steps may also introduce additional uncertainty and potentially affect the final prediction quality.

Another important direction is transfer learning. Instead of training models independently for each region, pretraining on data-rich areas and then adapting to new regions has become a common strategy. This is particularly useful when dealing with newly installed radar systems or regions with limited historical observations, where directly training a model from scratch is often not feasible.

At the same time, reducing computational and memory costs has become an important concern. High-resolution radar sequences and long input windows can quickly exceed hardware limits, so recent work has explored more efficient designs, including patch-based representations, multi-scale feature extraction, and lightweight network structures. Despite these efforts, scaling models to higher resolutions and longer sequences remains challenging in practical deployment.

Finally, improving learning objectives beyond standard pixel-wise losses has also attracted increasing attention. Instead of relying only on mean squared error, recent studies have explored combinations of reconstruction-based, perceptual, and adversarial objectives, as well as reweighted formulations that emphasize rare events. Although these approaches can improve certain aspects of prediction quality, designing a loss function that consistently preserves sharp structures while remaining physically meaningful is still an open problem.

6. Conclusion

This paper reviews recent progress in deep learning methods for short-term precipitation nowcasting, with a focus on four representative modeling families, namely recurrent neural networks, convolutional encoder–decoder structures, generative adversarial networks, and self-attention-based models. Rather than following a single design philosophy, these approaches differ in how they balance prediction accuracy, computational cost, and resource efficiency.

Even with the rapid evolution of model architectures, producing sharp and physically consistent forecasts remains a difficult problem. One persistent issue is the tendency toward overly smooth predictions, which is closely related to both the inherent uncertainty of atmospheric processes and the limitations of commonly used optimization objectives.

Looking forward, further improvements are likely to depend on progress in several complementary directions, including more efficient attention-based architectures, better integration of heterogeneous meteorological data, and modeling strategies that place greater emphasis on rare but high-impact precipitation events. Progress in these areas will likely benefit from closer interaction between data-driven machine learning techniques and traditional meteorological understanding.

ISSN: 2664-9640

References

[1] Wilson, J. W., Crook, N. A., Mueller, C. K., Sun, J., & Dixon, M. (1998). Nowcasting thunderstorms: A status report. *Bulletin of the American Meteorological Society*, *79*(10), 2079–2100. https://doi.org/10.1175/1520-0477(1998)079<2079:NTASR>2.0.CO;2

[2] Sun, J., Li, Q., & Wang, Z. (2014). Use of numerical weather prediction for nowcasting convective precipitation. *Monthly Weather Review*, *142*(10), 3806–3822. https://doi.org/10.1175/MWR-D-14-00056.1

[3] Nie, Y., Sun, J., Zhang, Z., & Chen, X. (2022). The July 2021 Henan extreme rainfall: Characteristics, causes and predictability. *Advances in Atmospheric Sciences*, *39*(8), 1235–1253. https://doi.org/10.1007/s00376-022-1446-9

[4] Lorenz, E. N. (1969). The predictability of a flow which possesses many scales of motion. *Tellus*, *21*(3), 289–307. https://doi.org/10.3402/tellusa.v21i3.10089

[5] Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, *525*(7567), 47–55. https://doi.org/10.1038/nature14956

[6] Wong, W. C., & Woo, W. K. (2017). Optical flow based precipitation nowcasting. In *Proceedings of the 2017 International Conference on Weather Forecasting* (pp. 112–120). Hong Kong.

[7] Bowler, N. E., Pierce, C. E., & Seed, A. W. (2004). STEPS: A probabilistic precipitation nowcasting scheme. *Journal of Hydrometeorology*, *5*(6), 1132–1152. https://doi.org/10.1175/JHM-387.1

[8] Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

[9] Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., & Woo, W. C. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. In *Advances in Neural Information Processing Systems* (Vol. 28, pp. 802–810). Curran Associates.

[10] Wang, Y., Jiang, L., Yang, M., Long, M., & Wang, J. (2017). PredRNN: A recurrent neural network for spatiotemporal predictive learning. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 882–890). Curran Associates.

[11] Wang, Y., Gao, Z., Long, M., Wang, J., & Yu, P. S. (2018). Memory in memory: A predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 9154–9162). IEEE. https://doi.org/10.1109/CVPR.2018.00953

[12] Yu, F., & Koltun, V. (2016). Multi-scale context aggregation by dilated convolutions. In *International Conference on Learning Representations*. https://arxiv.org/abs/1511.07122

[13] Lin, Z., Li, M., Zheng, Z., & Chen, H. (2020). SA-ConvLSTM: Self-attention based ConvLSTM for precipitation nowcasting. In *2020 IEEE International Geoscience and Remote Sensing Symposium* (pp. 2345–2348). IEEE. https://doi.org/10.1109/IGARSS39084.2020.9323274

[14] Wang, Y., Jiang, X., Chang, S., & Wang, J. (2019). E3D-LSTM: Spatiotemporal video prediction with enhanced 3D LSTM. In *International Conference on Learning Representations*. https://arxiv.org/abs/1904.06394

[15] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 234–241). Springer. https://doi.org/10.1007/978-3-319-24574-4_28

[16] Ayzel, S., Heistermann, M., Winterrath, C., Sensoy, A., & Jacobi, S. (2020). RainNet: A convolutional neural network for radar-based precipitation nowcasting. *Hydrology and Earth System Sciences*, *24*(2), 383–398. https://doi.org/10.5194/hess-24-383-2020

[17] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In *Advances in Neural Information Processing Systems* (Vol. 27, pp. 2672–2680). Curran Associates.

[18] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 5998–6008). Curran Associates.

[19] Gao, Z., Shi, X., Wang, H., & Yeung, D. Y. (2022). EarthFormer: Exploring space-time transformers for Earth system forecasting. In *Advances in Neural Information Processing Systems* (Vol. 35, pp. 25390–25403). Curran Associates.

[20] Bai, C., Sun, F., Zhang, J., & Liu, Y. (2022). RainFormer: Features extraction balanced network for radar-based precipitation nowcasting. *IEEE Geoscience and Remote Sensing Letters*, *19*, Article 4023305. https://doi.org/10.1109/LGRS.2022.3141357

[21] Chan, P. W., & Wong, W. K. (2016). *HKO-7 radar dataset for precipitation nowcasting*. Hong Kong Observatory Technical Note.

[22] Deutscher Wetterdienst. (2018). *DWD-12 composite radar dataset for Central Europe*. Deutscher Wetterdienst.

[23] Shenzhen Meteorological Bureau. (2017). *CIKM AnalytiCup 2017 precipitation nowcasting dataset*.

[24] Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*. Academic Press.