M2 Informatique IMA

Report PRAT

Rain nowcasting using deep learning models

Author :
Jacques Zhang

Supervisor:
Dominique Béréziat

Feb. 2024

Contents

1 Introduction 2

1.1 Deep learning methods . . . 2
1.2 Related work . . . 3

2 Problem statement 3

2.1 Data . . . 3

3 Model 3

3.1 First stage . . . 4
3.2 Second stage . . . 6

4 Results 7

4.1 First stage . . . 7
4.2 Second stage . . . 10
4.3 Scores . . . 12

5 Conclusion 14

Bibliography 15

1 Introduction

Precipitation nowcasting is the forecasting precipitation at a very short time horizon, up to two hours into the future, focusing on high spatiotemporal resolutions. This nowcasting is crucial for the decision-making in weather-dependent operations across a diverse variety of sectors, including flood early warning systems, air traffic control, marine services, emergency response, and energy management.

From a data science perspective, precipitation nowcasting is intriguing for distinct reasons. First of all, for nowcasting to be useful in other applications, the forecast must provide highly accurate predictions across multiple spatial and temporal scales. Secondly, the forecasts must account for uncertainty and be verified probabilistically. Additionally, nowcasting must perform well in predicting rare, heavy precipitation events that have significant impacts on both human life and the economy.

The nowcasting models deployed worldwide mostly use physics based numerical weather prediction (NWP), which simulate coupled equations of the atmosphere to generate multiple realistic precipitation forecasts.
Table 1 show some of them that are currently in use around the world.

Nowcasting system	**Input sources**	**Region of application**	**Reference**
AROME	Weather radar, NWP	France	[4]
Short-Term Ensemble Prediction System (STEPS)	Weather radar, NWP	UK	[5]
Korea Nowcasting System (KONOS)	Weather radar, NWP	Korea	/
Radar-Online-Adjustment (RADOLAN)	Weather radar	Germany	[14]
Integrated Nowcasting System through Comprehensive Analysis (INCA)	Weather radar, NWP, satellite, surface station observations	Europe (Alpine regions)	[7]
Short-range Warning of Intense Rainstorms in Localised Systems (SWIRLS)	Weather radar, NWP	China	[13]
Auto-Nowcast System (ANC)	Weather radar, NWP, satellite, surface station observations, wind profiler, atmospheric sounding, lightning detector	US	[9]
Global Synthetic Weather Radar (GSWR)	Satellite, lightning, NWP	US	[11]
McGill Algorithm for Precipitation nowcasting using Lagrangian Extrapolation (MAPLE)	Weather radar	Canada, US	[6]
Spectral-Prognosis (S-PROG)	Weather radar	Australia	[12]

Table 1: Some of the nowcasting systems that are currently in operational use around the world.

1.1 Deep learning methods

In recent years, the application of deep learning methods using radar reflectivity maps, along with the inclusion of additional data such as air temperature, wind speed, and humidity, has significantly enhanced forecast accuracy.

Advancements in data-driven deep learning techniques have demonstrated promising potential in this field, rivaling conventional numerical methods with their advantages of being more skillful, efficient, and scalable.

The first paper to start in the field of radar-based precipitation nowcasting was introduced by a group from Hong-Kong Meteorological observatory with ConvLTSM [2] in 2015 which combines the strengths of convolutional networks and long-short-term-memory processing spatial and temporal data.

1.2 Related work

The paper [1] introduce the U-Net architecture for precipitation forecasting. A U-Net architecture has two defining qualities, the first one is an encoder-decoder network that extracts more features the deeper it goes, and the second one is the presence of skip connections which allows the reintroduction of detailed features into the decoder.

Using a U-Net for classification resulting in a prediction of the amount of rainfall on a given time t produces results that are general and detailed. However, the further we go in time prediction, the more blurry and less detailed are the rain maps.

2 Problem statement

2.1 Data

We will be working with an open weather dataset provided by Météo-France known as MeteoNet. Specifically, we will focus on radar reflectivity maps measured every 5 minutes over a 550×550 km area in the northwest of France.

These maps, essentially images, have dimensions of 565×784 pixels, with each pixel representing the cumulative rainfall (CRF) over 5 minutes. However, the study area has been restricted to 128×128 pixels (roughly 100×150 km).

The dataset has been slip into training, validation, and test sets.

The inputs of the networks are sequences of consecutive images :
if $\boldsymbol{x}$ is an input of size $10 \times 128 \times 128$, it corresponds to a sequence of 10 images collected five minutes apart over 50 minutes).
We will also add an additional dimension corresponding to the channel :
$\boldsymbol{x}$ becomes an input of size $10 \times 1 \times 128 \times 128$.

The outputs are of size $n \times 1 \times 128 \times 12$, where n is predefined and corresponds to the time horizon prediction ($n * 5$ minutes in the future).

During training, the inputs of the network will have an additional dimension corresponding to the batch size :
if $\boldsymbol{X}$ is an input of size $128 \times 10 \times 1 \times 128 \times 128$, there are 128 sequences of 10 consecutive images. However, since PyTorch does not work well with 5D tensors we will usually merge the two first dimensions to get a 4D tensor :
$\boldsymbol{X}$ becomes an input of size $1280 \times 1 \times 128 \times 128$.

3 Model

We have implemented a modified version of the model introduced in [10]. This model proposes a two-stage network approach: a first-stage network for producing initial predictions and a second-stage network to refine these raw predictions further.

The first stage, named the Spatiotemporal Prediction Network (STPNet), is an encoder-decoder structure based on an RNN block. It incorporates convolutional structures in both the input-to-state and state-to-state transitions.
The previous radar echo observation sequence is given as input into the encoder of STPNet, which returns n layers of RNN states. Then, another n layers of RNNs are used to generate the future radar echo predictions based on the encoded states. The prediction is the first-stage result and is an intermediate result.

The second stage, named the Detail Refinement Stage, proposed in this paper, acts as a defuzzification network for the predicted radar echo images. Additionally, it enhances the ability of spatial and temporal feature extraction for echo image detail information. The first-stage prediction is given as input to the detail refinement stage to acquire a refined prediction, which is also the final result, improving image quality and refining precipitation prediction precision.

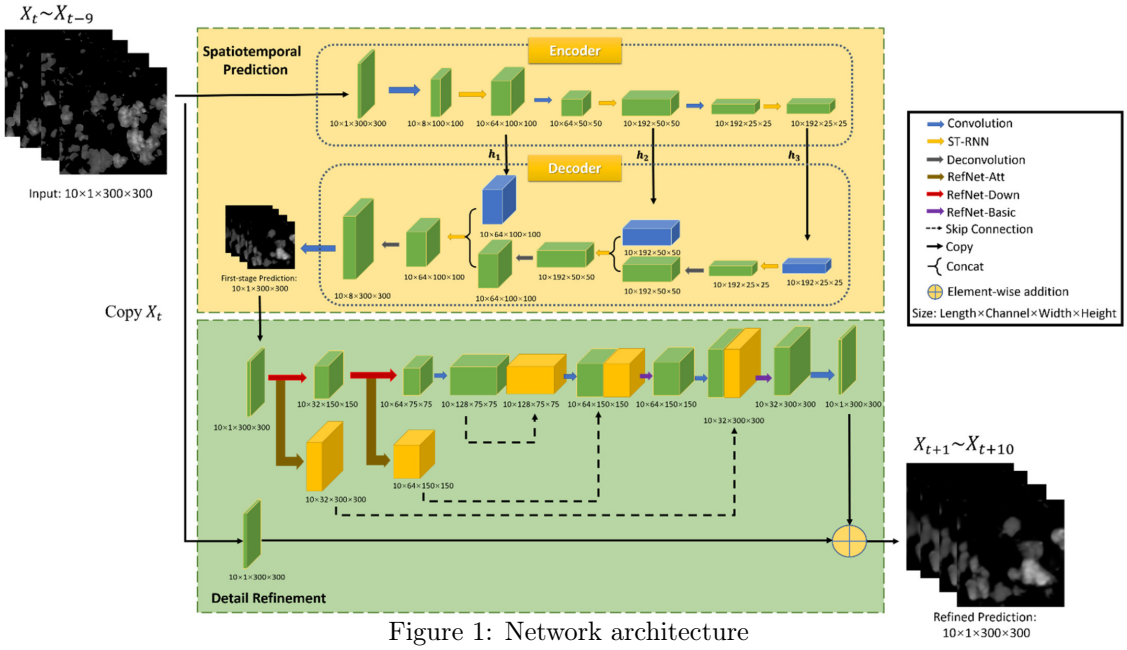

Figure 1: Network architecture

3.1 First stage

The RNN used is the TrajGRU that was first introduced in [3], the detailed equations are :

$$\mathcal{U}_t, \mathcal{V}_t = \gamma(\mathcal{X}_t, \mathcal{H}_{t-1}),$$

$$\mathcal{Z}_t = \sigma\big(\mathcal{W}_{xz} * \mathcal{X}_t + \sum_{l=1}^{L} \mathcal{W}_{hz}^l * \text{warp}(\mathcal{H}_{t-1}, \mathcal{U}_{t,l}, \mathcal{V}_{t,l})\big),$$

$$\mathcal{R}_t = \sigma\big(\mathcal{W}_{xr} * \mathcal{X}_t + \sum_{l=1}^{L} \mathcal{W}_{hr}^l * \text{warp}(\mathcal{H}_{t-1}, \mathcal{U}_{t,l}, \mathcal{V}_{t,l})\big),$$

$$\mathcal{H}'_t = f\Big(\mathcal{W}_{xr} * \mathcal{X}_t + \mathcal{R}_t \circ \big(\sum_{l=1}^{L} \mathcal{W}_{hh}^l * \text{warp}(\mathcal{H}_{t-1}, \mathcal{U}_{t,l}, \mathcal{V}_{t,l})\big)\Big),$$

$$\mathcal{H}_t = (1 - \mathcal{Z}_t) \circ \mathcal{H}'_t + \mathcal{Z}_t \circ \mathcal{H}_{t-1}.$$

Where L is the total number of links, $\mathcal{U}_t, \mathcal{V}_t$ are the flow fields generated by γ.
$\mathcal{W}_{hz}^l, \mathcal{W}_{hr}^l$ and $\mathcal{W}_{hh}^l$ are the weights for projecting the channels, which are implemented by 1×1 convolutions.
And the $\text{warp}(\mathcal{H}_{t-1}, \mathcal{U}_{t,l}, \mathcal{V}_{t,l})$ function selects the positions pointed out by $\mathcal{U}_{t,l}, \mathcal{V}_{t,l})$ from $\mathcal{H}_{t-1}$ using the bilinear sampling kernel.

Encoder

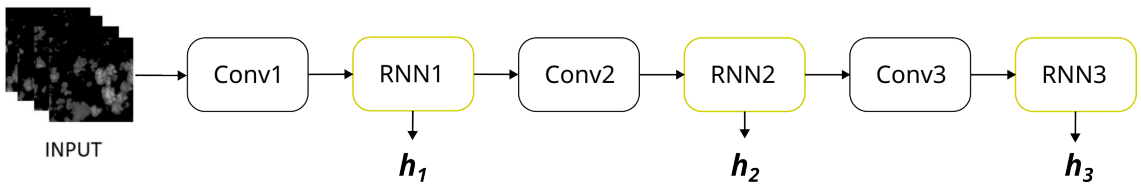

Figure 2: First stage encoder

The structure of the encoder is formed by stacking Convolution, RNN, Down Sampling, RNN, Down Sampling and RNN.
After each down sampling layer n, the spatial features of the echo images is extracted at different scales through the RNN layers, and the hidden state h_n is returned.

Decoder

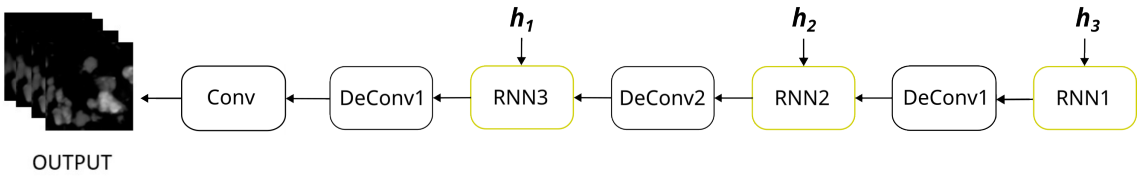

Figure 3: First stage decoder

The structure of the decoder is dual to that of the encoder, including RNN, Up Sampling, RNN, Up Sampling, RNN, DeConvolution, and Convolution. The order of the decoder network is reversed, where the high-level states capturing the global spatiotemporal representation are utilized to influence the update of the low-level states. When the decoder is initialized, all RNN layers receive the hidden states h_1, h_2, and h_3 from the encoder, respectively.

Through the deconvolution layer, by combining features and filling in the small-scale image details, the first-stage radar echo prediction image sequence is returned.

Tables 2 and 3 show the detailed structure settings of the encoder and decoder of our spatiotemporal RNN model.

— Kernel is a matrix that moves over the input data, and performs the dot product with the sub-region of input data.
— Stride defines the step size of the kernel when sliding through the image.
— L is the number of links in the state-to-state transition.

Table 2: Encoder parameters from first stage

Name	Kernel	Stride	L	Channels input/ouput
Conv1	5×5	3×3	—	1/8
RNN1	5×5	1×1	13	8/64
Conv2	5×5	3×3	—	64/192
RNN2	5×5	1×1	13	192/192
Conv3	3×3	2×2	—	192/192
RNN3	3×3	1×1	9	192/192

Table 3: Decoder parameters from first stage

Name	Kernel	Stride	L	Channels input/ouput
RNN1	3×3	1×1	13	192/192
DeConv1	4×4	2×2	—	192/192
RNN2	3×3	1×1	13	192/192
DeConv2	5×5	3×3	—	192/64
RNN3	3×3	1×1	9	64/64
DeConv3	5×5	3×3	—	64/8
Conv	3×3	1×1	—	8/1

3.2 Second stage

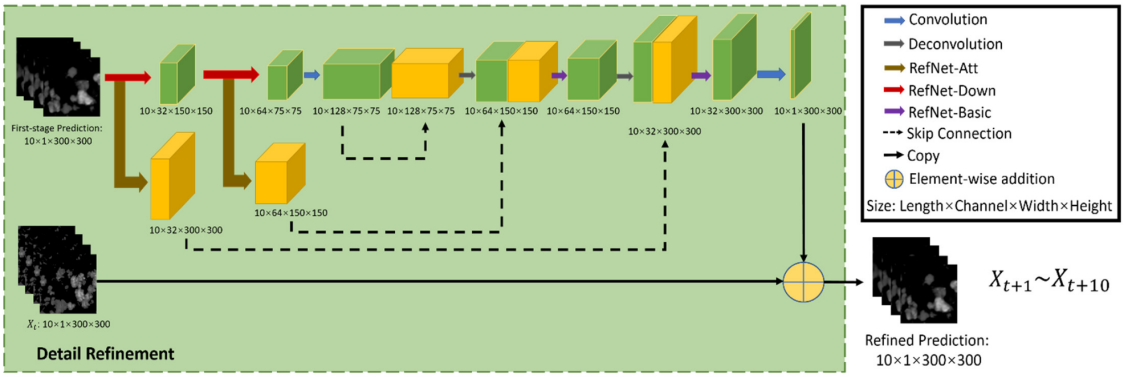

Figure 4: RefNet

In the second stage, an encoder–decoder network is proposed to further extract spatiotemporal features and refine the predicted radar echo sequences.

The encoder-decoder serves as a feature selector for focusing on the locations full of tiny textures.

A multi-level skip connection is employed between different scale features for feature sharing and reuse. Both local and global residual learning is integrated for preserving the low-level features and decreasing the difficulty of training and learning. Different hierarchical features are combined to generate finer features, which favor the reconstruction of high-resolution images.

We have chosen for the channel-wise attention layer, a Squeeze-and-Excitation (SE) block introduced in [8] which allows the model to learn the spatio-temporal feature importance, and producing the importance weight matrix for input feature map sequences.

Three sub-nets are defined to build up the refinement network (RefNet), named RefNet-Basic (5), RefNet-Att (6) and RefNet-Down (7).

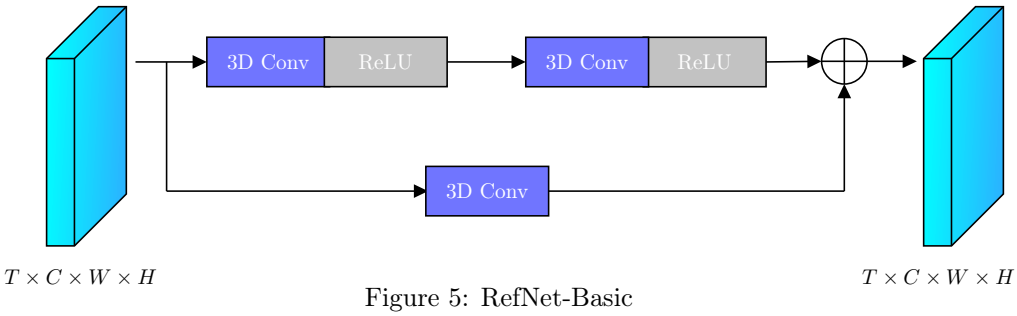

Figure 5: RefNet-Basic

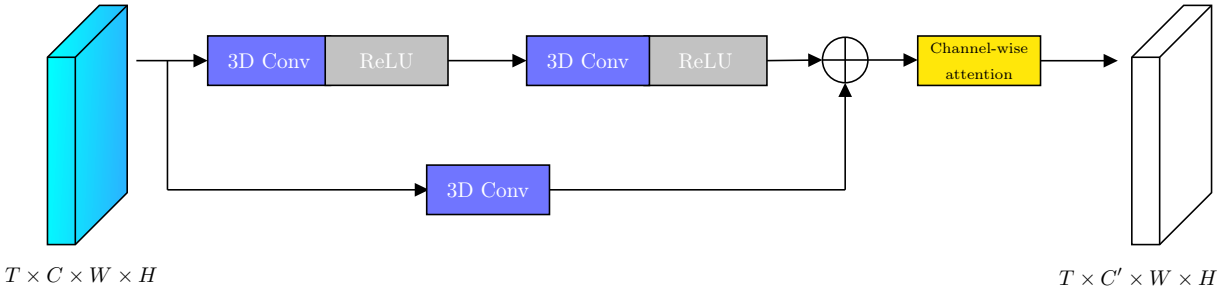

Figure 6: RefNet-Attention

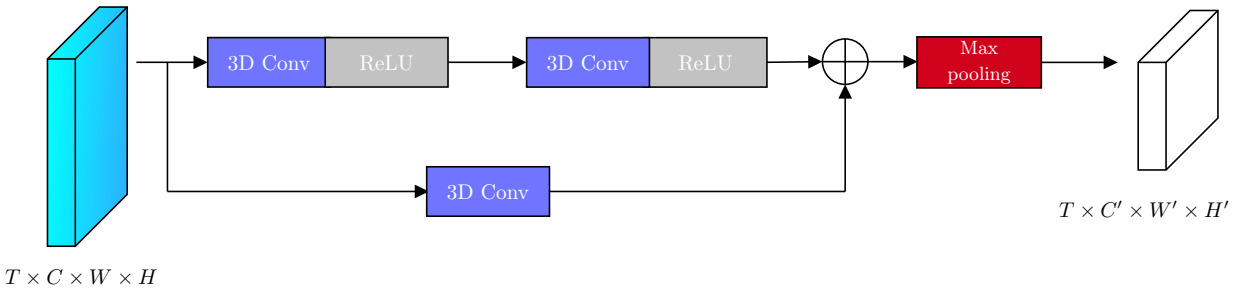

Figure 7: RefNet-Down

As input, this network takes the first-stage prediction and the initial sequences of consecutive images. The operation $\bigoplus$ in Figures 4 and 5,6,7 is performed using a shortcut connection and element-wise addition.

4 Results

4.1 First stage

The first stage of the model has been trained for 25 `epochs` on a GPU with the following parameters :
`batch_size` = 128, `input_length` = 10 and `time_horizon` = [4,5,6,7], `loss` = `MSE`, `Adam` optimizer with `lr` = 0.0001.
The activation function used for both encoder and decoder is a `LeakyReLU` with `negative_slope` = 0.02.

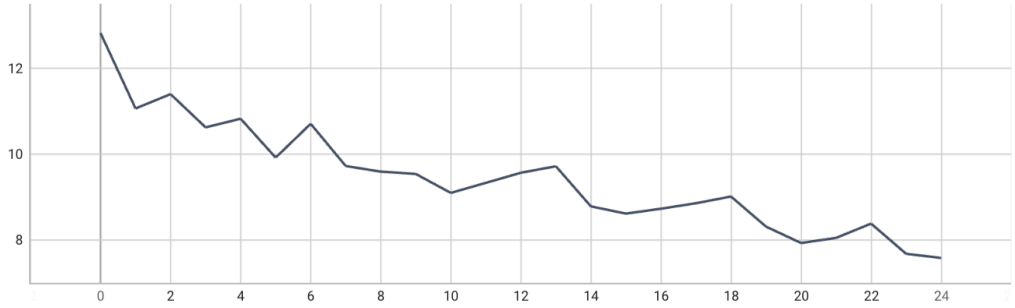

Figure 8: Train loss for `time_horizon` = 5.

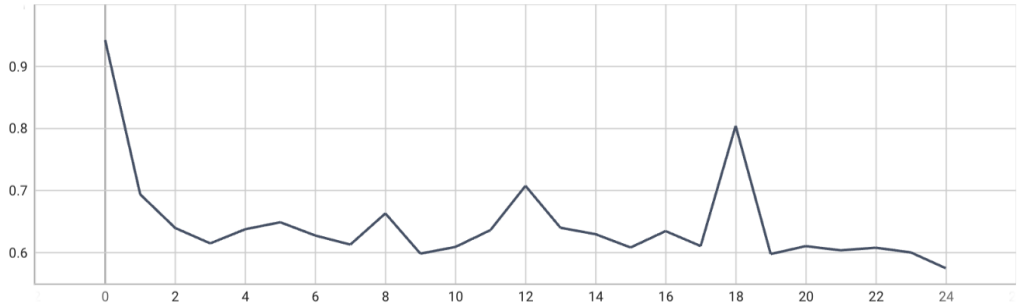

Figure 9: Validation loss for `time_horizon` = 5.

The loss function curve during the training period show that the model can effectively converge.

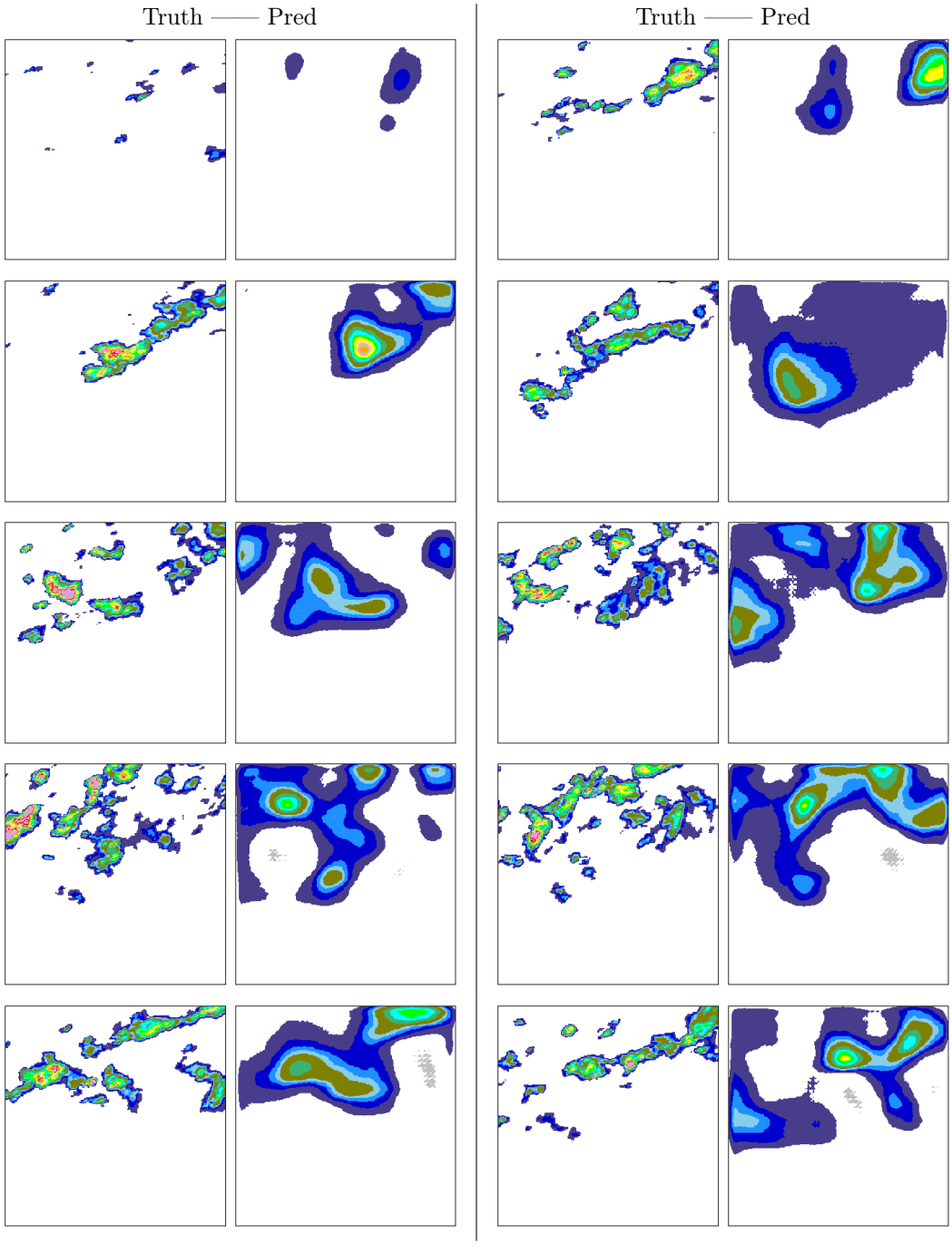

Table 4: `time_horizon` = 4

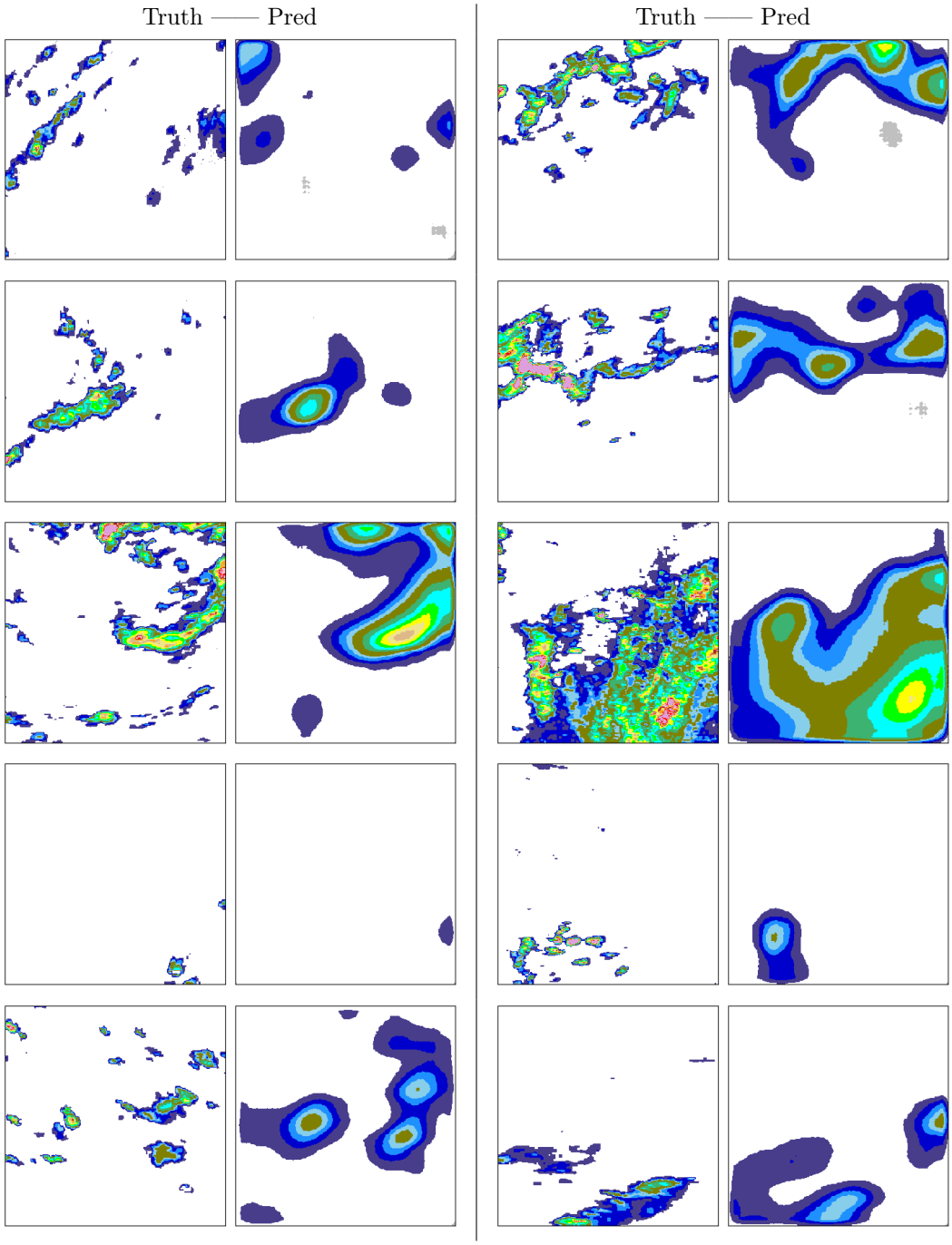

Table 5: `time_horizon` = 5

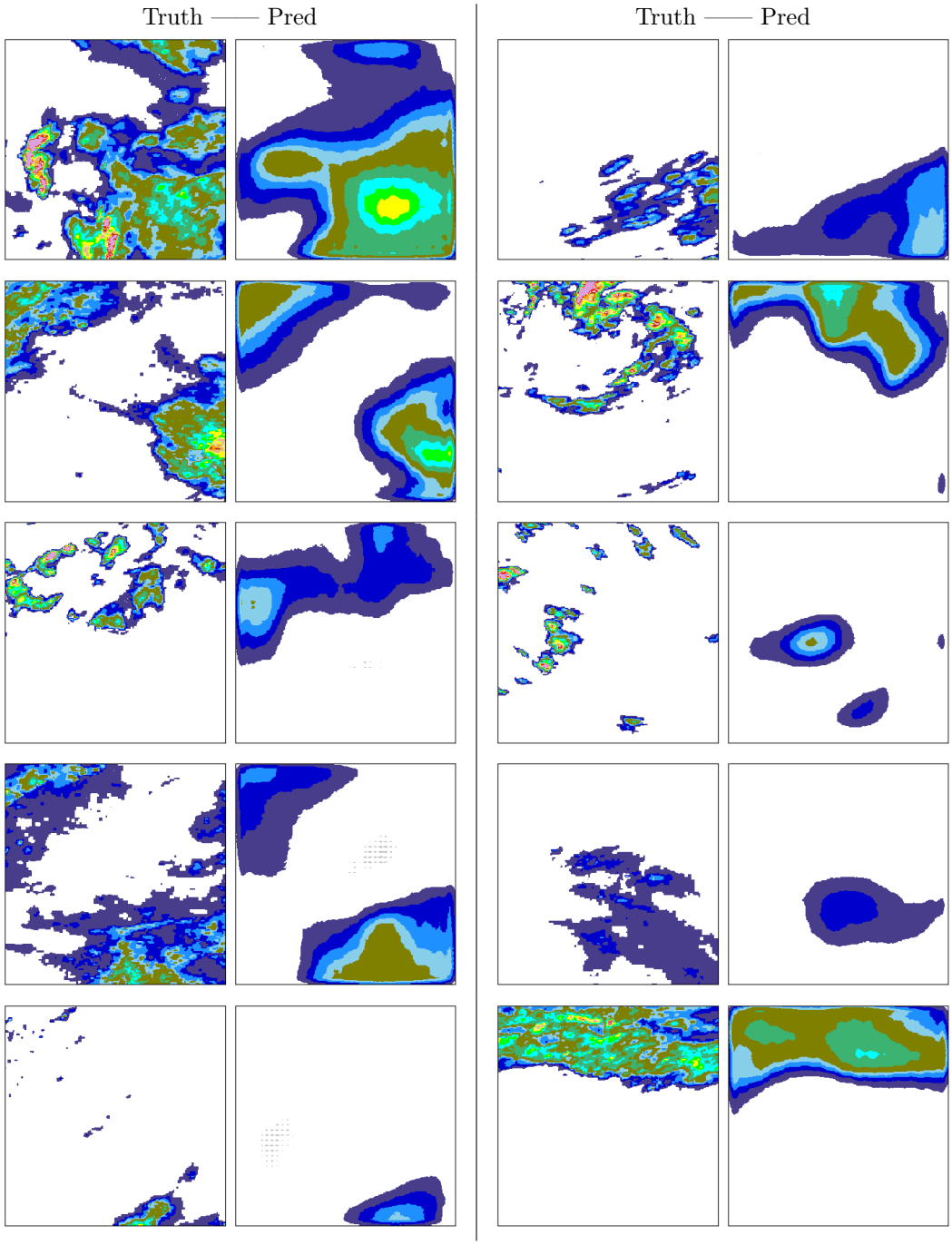

Table 6: `time_horizon` = 6

The predicted images from the first stage demonstrate relatively good performance in terms of spatial features. However, the extrapolation radar echo images still seem to be blurred and lack details.

4.2 Second stage

Due to errors and mistakes in the original paper, we did not manage to perfectly train the second stage (RefNet). Specifically, a convolutional block was omitted, and we did not notice it until after training commenced.
An experimental network was trained for 20 `epochs` on the CPU; however, it did not produce the expected results as shown in the original paper.

Due to the long training time required on the CPU, it was impossible to conduct additional training sessions.

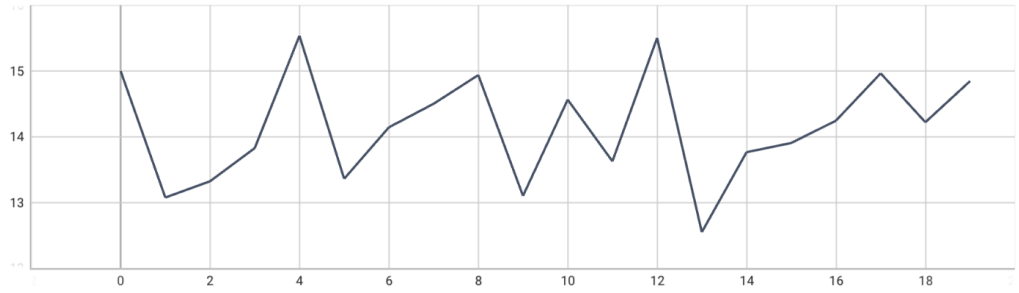

Figure 10: Train loss for `time_horizon` = 5.

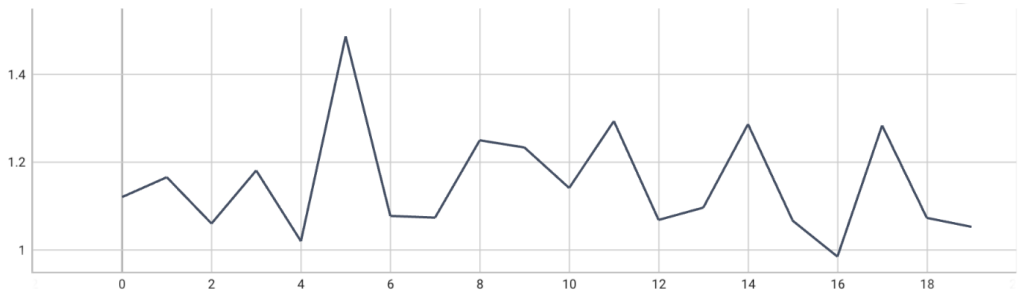

Figure 11: Validation loss for `time_horizon` = 5.

The loss function curve does not show convergence nor decrease, indicating a mistake in the training process.

We show below the binarized output compared to the binarized ground truth for a prediction at a `time_horizon` = 5.

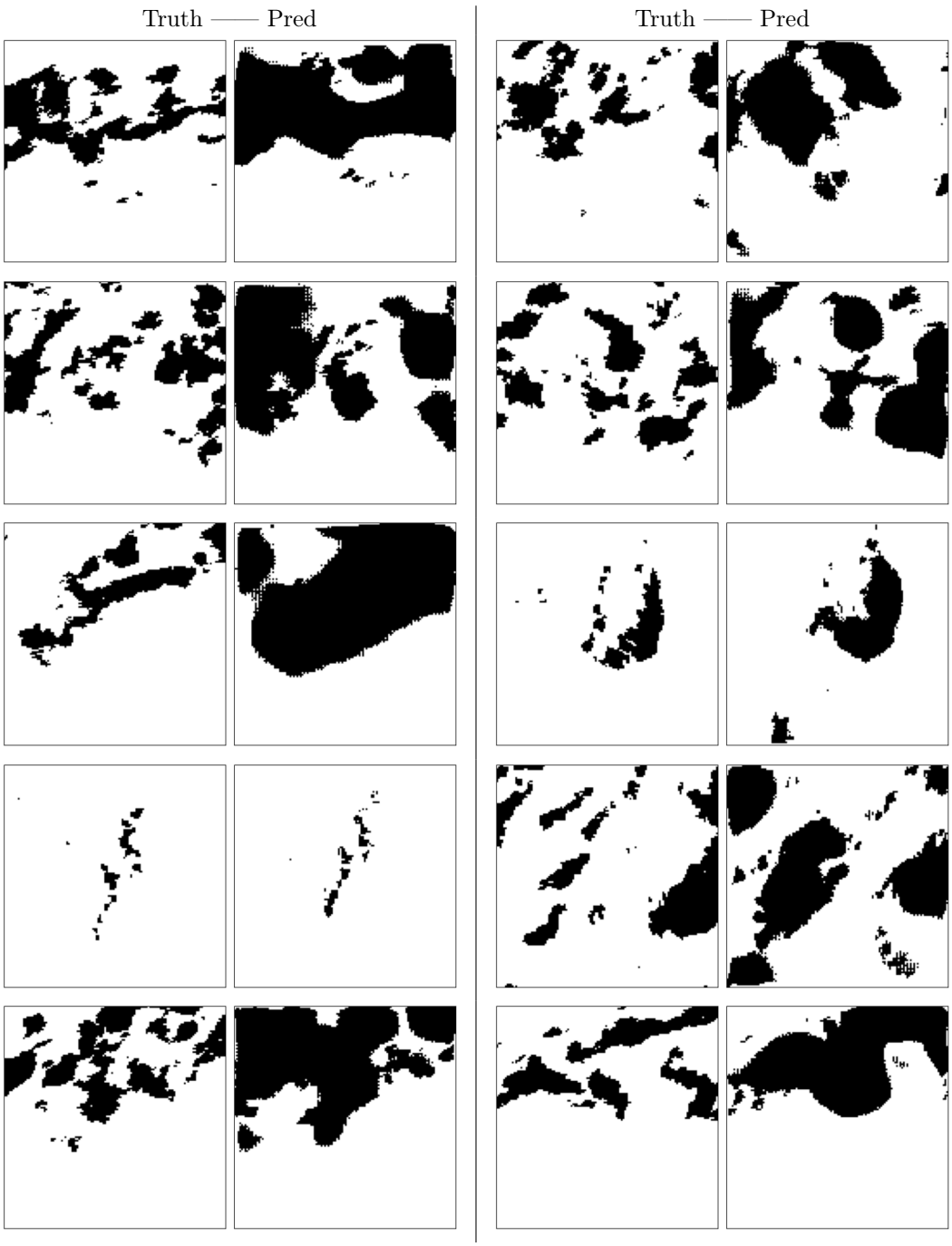

Table 7: `time_horizon` = 5

The results demonstrate sharper predicted images with more details compared to using only the first-stage STPNet method.
Additionally, there is less blurriness present in the predictions.

4.3 Scores

The standard metrics used to summarize model performance consist of the counts of true positive (TP), false positive (FP), false negative (FN), and true negative (TN).

Here, TP represents hits, FP stands for false alarms, FN indicates misses, and TN corresponds to correct negatives.

These metrics are the building blocks of other metrics :

Name	Formula
Precision	$\frac{TP}{TP+FN}$
Recall	$\frac{TP}{TP+FP}$
F1	$2\frac{Prec \times Recall}{Prec + Recall}$
Threat Score (TS, also called CSI)	$\frac{TP}{TP+FP+FN}$
Bias	$\frac{Recall}{Prec}$
False Alarm Rate (FAR)	$\frac{FP}{FP+TN}$
Heidke Skill Score (HSS)	$\frac{2(TP \times TN - FP \times FN)}{(TP+FN)(FN+TN)+(TP+FP)(FP+TN)}$
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Equitable Threat Score (ETS)	$\frac{TP - TP_{random}}{TP+FP+FN-TP_{random}}$
Odds Ratio Skill Score (ORSS)	$\frac{TP \times TN - FP \times FN}{TP \times TN + FP \times FN}$

Table 8: Scores metrics

Where $TP_{random} = \frac{(TP+FN)(TP+FP)}{TP+FP+FN+TN}$.

To compute these metrics, the target and prediction are binarized.

Name	1st stage	2nd stage
Precision ↑	0.115284	**0.437380**
Recall ↑	**0.999379**	0.756923
F1 ↑	0.206721	**0.554404**
TS ↑	0.115276	**0.383512**
BIAS	8.668851	1.730584
HSS ↑	0.000244	**0.478230**
FAR ↓	0.998319	**0.126739**
Accuracy ↑	0.116591	**0.859861**
ETS ↑	0.000122	**0.314259**
ORSS ↑	0.460844	**0.910935**

Table 9: Scores obtained for both networks

In this table, (↑) signifies that higher values are preferable, while (↓) signifies that lower values are preferable. The best results are indicated in bold.

The second column corresponds to the results of the first stage (STPNet) and the third column corresponds to the results produced by the second stage (RefNet) when given the output of the first stage and original sequence.

We can clearly see that the second stage improves the performance. The only area where it shows poor performance is in recall, indicating an increased prediction of false positives compared to the first stage.
This can be explained by the fact that a mistake was made during training where we performed an element-wise addition with the original input sequence instead of first feeding it to a convolutional block, resulting in a lot more unexpected values.

5 Conclusion

We have implemented a first stage encoder-decoder structure model which extracts the dynamic spatial and temporal correlations in a sequence of radar echo images and outputs a first-stage prediction.
A second stage multi-scale feature extraction and fusion residual block is implemented to acquire a better performance on details and prediction accuracy of the nowcasting radar echo images.

The challenges are that sharp and accurate predictions of the whole radar maps in longer-term predictions are quite difficult.

The limitation of such a deep learning method is that:

- The input and output dimensions of the model are fixed, and it does not deal with length or dimension variant input sequences. If different input numbers or input dimensions images are provided, the model must be redesigned (including adjustments to the number of input and output channels) and retrained accordingly.
- The blurring effect of deep learning methods may be caused by the inherent uncertainties of the task. Since sharp and accurate predictions of the whole radar maps in longer-term predictions are quite difficult, blurring the predictions to alleviate the error and decrease the MAE or MSE-based loss caused by this type of uncertainty is utilized. Therefore, exploring and designing more effective loss functions could be beneficial for enhancing the quality of nowcast images in future work.
- For the scores evaluation, we have deliberately chosen a threshold for the binarization (0 if ≤ 0.5, 1 otherwise), but since our predictions are done at the pixel level, it would be better to project them back to radar echo intensities and calculate the rainfall at every cell of the grid before binarizing the pixels.
- The choice of activation function in the first stage (LeakyReLU) can be experimented with and changed (e.g., ReLU, Sigmoid) to explore its impact on performance. Similarly, the same can be done in the second stage.
- In the second stage, the choice of the channel-wise attention block can also be changed to find one that improves performance.

The source code can be found here :

https://github.com/j-zhang19/PRAT

References

[1] Bouget, V., Béréziat, D., Brajard, J., Charantonis, A., and Filoche, A. Fusion of rain radar images and wind forecasts in a deep learning model applied to rain nowcasting. *Remote Sensing 13*, 2 (2021).

[2] Xingjian Shi, Chen, Z., Wang, H., Yeung, D.-Y., kin Wong, W., and chun Woo, W. Convolutional lstm network: A machine learning approach for precipitation nowcasting, 2015.

[3] Xingjian Shi, Gao, Z., Lausen, L., Wang, H., Yeung, D. Y., Wong, W.-K., and chun Woo, W. Deep learning for precipitation nowcasting: A benchmark and a new model. In *Neural Information Processing Systems* (2017).

[4] Auger, L., Dupont, O., Hagelin, S., Brousseau, P., and Brovelli, P. Arome–nwc: a new nowcasting tool based on an operational mesoscale forecasting system. *Quarterly Journal of the Royal Meteorological Society 141*, 690 (2015), 1603–1611.

[5] Bowler, N., Pierce, C., and Seed, A. Steps: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled nwp. *Quarterly Journal of the Royal Meteorological Society 132* (01 2007), 2127 – 2155.

[6] Germann, U., and Zawadzki, I. Scale-dependence of the predictability of precipitation from continental radar images. part i: Description of the methodology. *Monthly Weather Review 130*, 12 (2002), 2859 – 2873.

[7] Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B., and Gruber, C. The integrated nowcasting through comprehensive analysis (inca) system and its validation over the eastern alpine region. *Weather and Forecasting 26*, 2 (2011), 166 – 183.

[8] Hu, J., Shen, L., Albanie, S., Sun, G., and Wu, E. Squeeze-and-excitation networks, 2019.

[9] Mueller, C., Saxen, T., Roberts, R., Wilson, J., Betancourt, T., Dettling, S., Oien, N., and Yee, J. Ncar auto-nowcast system. *Weather and Forecasting 18*, 4 (2003), 545 – 561.

[10] Niu, D., Huang, J., Zang, Z., Liujia, X., Che, H., and Tang, Y. Two-stage spatiotemporal context refinement network for precipitation nowcasting. *Remote Sensing 13* (10 2021), 4285.

[11] Reen BP, Cai H, R. J. Preliminary investigation of assimilating global synthetic weather radar. Tech. rep., CCDC Army Research Laboratory (US), sep 2020.

[12] Seed, A. W. A dynamic and spatial scaling approach to advection forecasting. *Journal of Applied Meteorology 42*, 3 (2003), 381 – 388.

[13] Yeung, L. H., Wong, W., Chan, P. K., and Lai, E. S. Applications of the hong kong observatory nowcasting system swirls-2 in support of the 2008 beijing olympic games. In *WMO Symposium on Nowcasting* (2009), vol. 30.

[14] Dwd climate data center (cdc), aktuelle stündliche radolan-raster der niederschlagshöhe (binär),version. https://www.cen.uni-hamburg.de/en/icdc/data/atmosphere/radolan.html. Accessed: 2023-11-11.