High-Resolution Radar Echo Prediction with ML

Matej Choma[1,2], Jakub Bartel[1], Petr Šimánek[2], Vojtěch Rybář[2]

[1]Meteopress s.r.o.

[2]Faculty of Information Technology, Czech Technical University in Prague

Introduction

One of the up-to-date challenges in meteorology is a radar echo prediction as a part of the precipitation nowcasting - computing short-time predictions delivered to the end-users in real-time.

Even though traditional weather forecasts based on numerical weather prediction (NWP) models are great for prediction for the upcoming days, they struggle in the short term. Some of the reasons are long computation time and insufficient spatial and temporal resolution.

We have been researching the use of machine learning (ML) models for accurate, high-resolution precipitation forecasting. We formulate the problem as a prediction of K future frames $X_{t+1}, \ldots, X_{t+K}$ of a radar echo image sequence based on the last J measured ones $X_{t-J+1}, \ldots, X_t$ as following:

$$X_{t+1}, \ldots, X_{t+K} = \underset{X_{t+1},\ldots,X_{t+K}}{\operatorname{argmax}} P(X_{t+1}, \ldots, X_{t+K} | X_{t-J+1}, \ldots, X_t), \quad (1)$$

where t denotes the time of the last measurement.

Dataset

Weather radar echo images are the primary data source for precipitation nowcasting. These are capable of capturing precipitation fields in real-time and high resolution.

In this work, we use images captured above the Czech Republic obtained through the OPERA programme. The spatial resolution is 1 km^2 and the time step between two images is 10 minutes. Given the used data from 23. 10. 2015 to 21. 7. 2020, we have separated 275 independent precipitation situations. We have defined independence by at least 24-hour gap with no rain between two consecutive situations. The situations were randomly divided into three sets. Training set contains 83798 prediction samples, validation set 13899 and test set contains 17040 samples. The image dimensions are 544 × 352 pixels (Width × Height).

Intermediate results indicate that using other types of data may improve performance. We consider satellite, NWP models and data describing the earth surface of the prediction domain. However, the integration of this data into our ML models needs further research.

Methods

We experiment with two neural network (NN) architectures. Both are built on the process called convolution to capture the spatial information from the 2D radar echo images.

The first one, U-net, is a fully convolutional NN with traditional encoder-decoder design (Figure 1). The convolutional blocks consist of three consecutive convolutional layers with 3×3 filters. The stacking is used to ensure a larger receptive field with fewer parameters [1]. The U-net takes the sequence of J input frames at once as a 3D tensor with shape $[J, H, W]$.

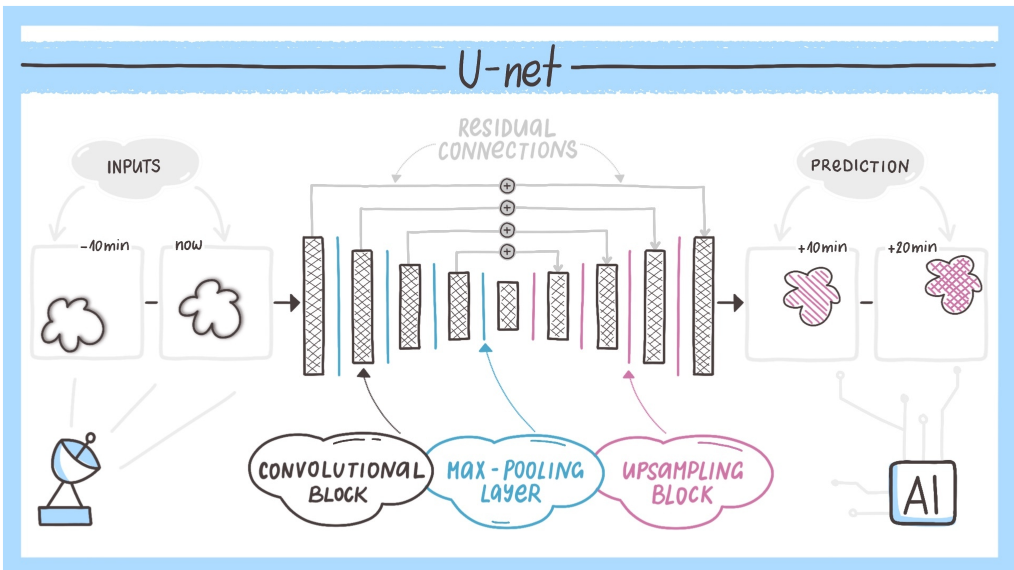

Figure 1:U-net architecture

The second network, PredRNN, is a recurrent neural network (RNN). The RNNs are capable of sequentially processing J multi-channel radar echo images with shape $[C, H, W]$ (Figure 2). This idea behind recurrent architectures is to capture the temporal development of the precipitation better. PredRNN is created by stacking three SpatioTemporalLSTM cells [2]. Moreover, these cells are connected with the skip connections. This emphasizes the fact that the prediction X_{t+1} is created as a change to the previous image X_t:

$$X_{t+1} = X_t + f(X_t), \quad (2)$$

where f denotes the PredRNN cell.

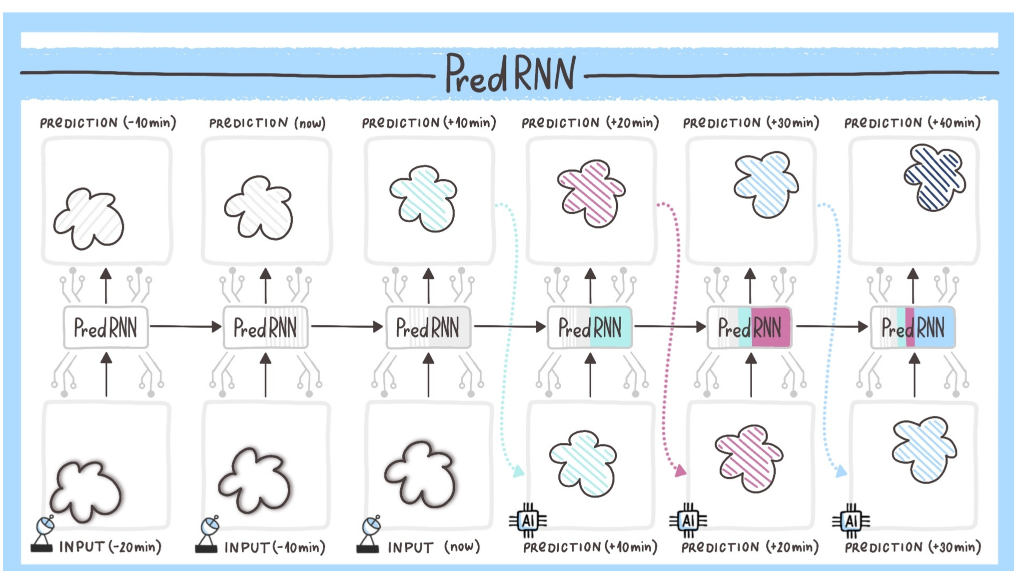

Figure 2:PredRNN architecture wireframe

Experimental Results

In precipitation prediction, the model is predicting both precipitation area and intensity. Thus, one metric often does not give comprehensive information. We use both the Critical Success Index (CSI) and mean absolute error.

Predictions of our models have been validated on the created test set against methods based on the optical flow – pySTEPS Anvil, S-PROG and rainymotion Dense.

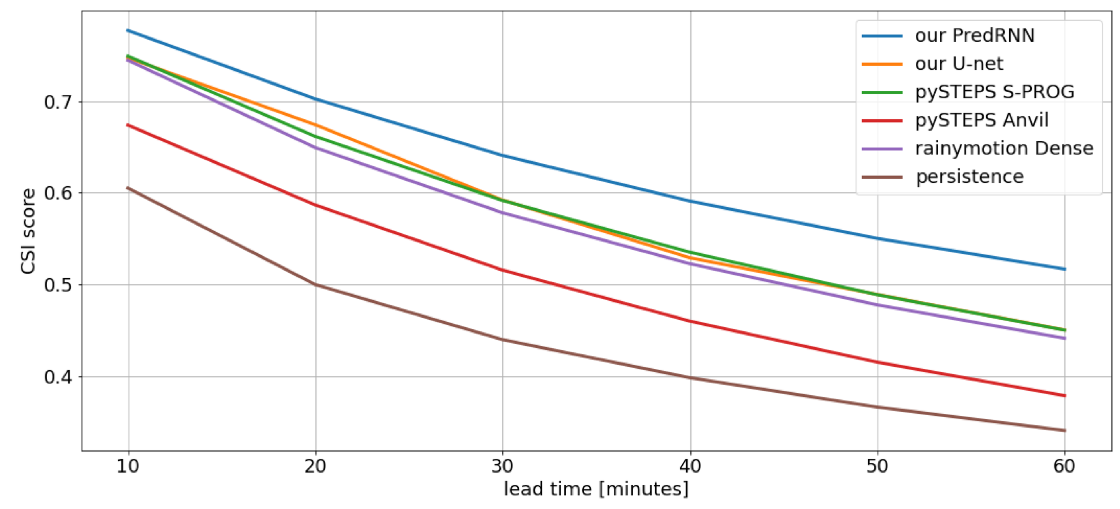

Figure 4:Evaluation of the predictions using CSI with rain threshold 0.5mm/h. Higher is better.

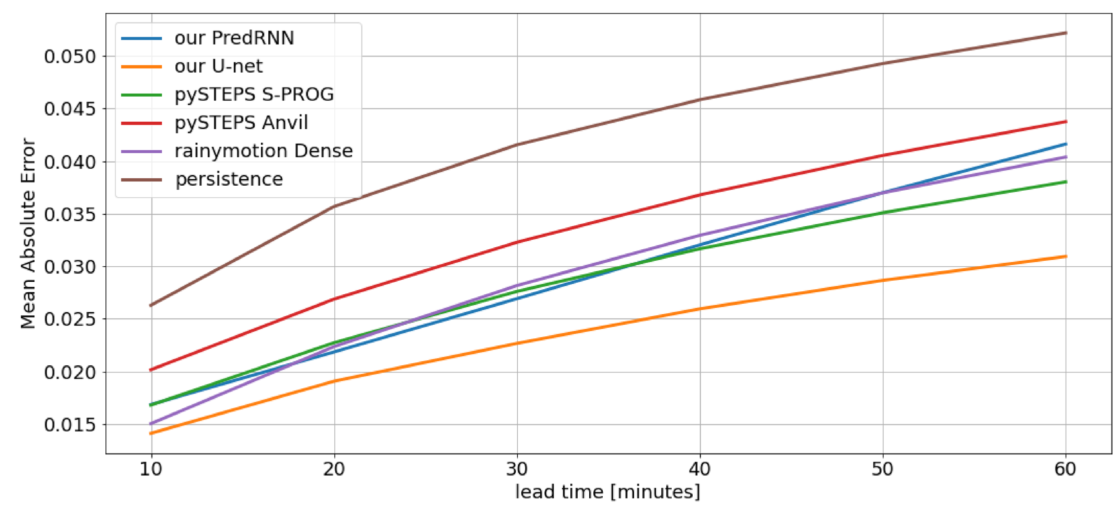

Figure 5:Evaluation of the predictions using mean absolute error. Lower is better.

Qualitatively (Figure 3), our models can not only capture location change but also change of structure and intensity of precipitation cells.

Conclusion

In this work, we have explored two ML approaches to the problem of radar echo prediction. Our models were trained using radar data from above the Czech Republic, but can be used in any other geographical location. The quantitative evaluation showed that our models achieve comparable or better performance than traditionally used optical flow models.

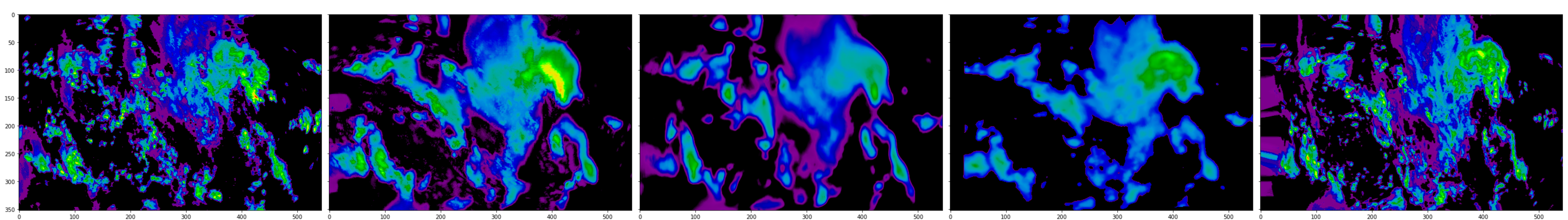

Figure 3:Visual comparison of predictions for 60 minutes lead time. From left: *1.* ground truth, *2.* our PredRNN, *3.* our U-net, *4.* pySTEPS S-PROG, *5.* rainymotion Dense.

References

[1] Choma Matej.
Interpolation and extrapolation of subsequent weather radar images.
Bachelor's thesis, Czech Technical University in Prague, Faculty of Information Technology, 2019.

[2] Yunbo Wang, Mingsheng Long, Jianmin Wang, Zhifeng Gao, and S Yu Philip.
Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms.
In *Advances in Neural Information Processing Systems*, pages 879–888, 2017.

FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE

meteopress