SUPER-RESOLVED RAINFALL PREDICTION WITH PHYSICS-AWARE DEEP LEARNING

S. Moran, B. Demir

Technical University of Berlin
Berlin, Germany

F. Serva, B. Le Saux*

European Space Agency Φ-lab, ESRIN
Frascati, Italy

ABSTRACT

Rainfall prediction at the kilometre-scale up to a few hours in the future is key for planning and safety. But it is challenging given the complex influence of climate change on cloud processes and the limited skill of weather models at this scale. Following the set-up proposed by the *weather4cast* challenge of NeurIPS, we build a two-step deep-learning solution for predicting rainfall occurrence at ground radar high spatial resolution starting from coarser resolution weather satellite images. Our approach is designed to predict future satellite images with a physics-aware ConvLSTM network, which is then converted into precipitation maps through a U-Net. We find that our two-step pipeline outperforms the baseline model and we quantify the benefits of including physical information. We find that local-scale rainfall predictions with good accuracy starting from satellite radiances can be obtained for up to 4 hours in the future.

Index Terms**—** Physics-based machine learning, rainfall prediction, nowcasting

1. INTRODUCTION

Improvements in remote sensing technologies and the growth of computational capabilities in the last decades have allowed us to predict the weather with higher skill compared to the past [2]. At the core of current weather forecast systems are complex weather models referred to as Numerical Weather Prediction (NWP) methods. Providing accurate forecasts of precipitation events is important given their impacts on society and our economy [3].

NWP models are computationally expensive and take a long time to spin up a prediction. Therefore, current research [15, 4, 14] is focusing on the potential for Deep Learning (DL) methods for short-term forecasting (nowcasting, typically up to a few hours in the future, [16]). State-of-the-art DL methods have good accuracy for low rain rates but provide blurry predictions for longer lead times and more impactful medium and heavier rainfall rates [11].

Including physical constraints is a possible solution for more robust predictions while also taking advantage of DL to learn complex relationships from the data [9, 15]. While weather satellites provide high-frequency imagery at a moderate spatial resolution, ground-based radars give precipitation estimates with high spatial resolution but have limited coverage due to their high costs. In [17, 5, 11] it has been shown that combining radar rainfall estimates with weather satellite data for training large DL networks can be competitive with NWP models.

In this work, we tackle the *weather4cast* challenge prediction task by means of a physics-aware DL workflow to predict rainfall with high spatial resolution using only coarser resolution weather satellite imagery as input.

2. DATA AND METHODS

2.1. Experiments setup

The problem formulation follows the *weather4cast*[1] competition (stage 1) presented at NeurIPS 2022: given a one-hour input sequence of Meteosat satellite data [1] across the visible, infrared and water vapour bands, binary precipitation maps (threshold 0.001 mm/hour) at the resolution of OPERA ground-based radar [13] are to be predicted for the next hours every 15 mins. Since the spatial resolution of radar data is **6** times finer than that of satellite images, the problem requires a multi-modal sequence prediction task with super-resolution (SR). The input satellite sequences have the shape: channel, time, height, width (11, 4, 252, 252), while outputs are shaped as time, height, width (16, 252, 252). To inform models about surrounding weather conditions, the satellite context covers 1008 km^2 on the ground, while the target region is 168 km^2 large, i.e. six times smaller.

We considered the three target regions of *weather4cast* stage 1, but given the limited availability of resources we chose to train and test on a northern European region, an area relatively balanced in terms of rain/no-rain distribution. Given the nowcasting focus, predictions are sought up to 4 hours in the future. The evaluation metrics we use are the Recall, Precision, F1-score and Critical Success Index (CSI) or Intersection over Union (IoU).

---

*Now at the National Research Council of Italy

[1] https://www.iarai.ac.at/weather4cast/, 2022 edition

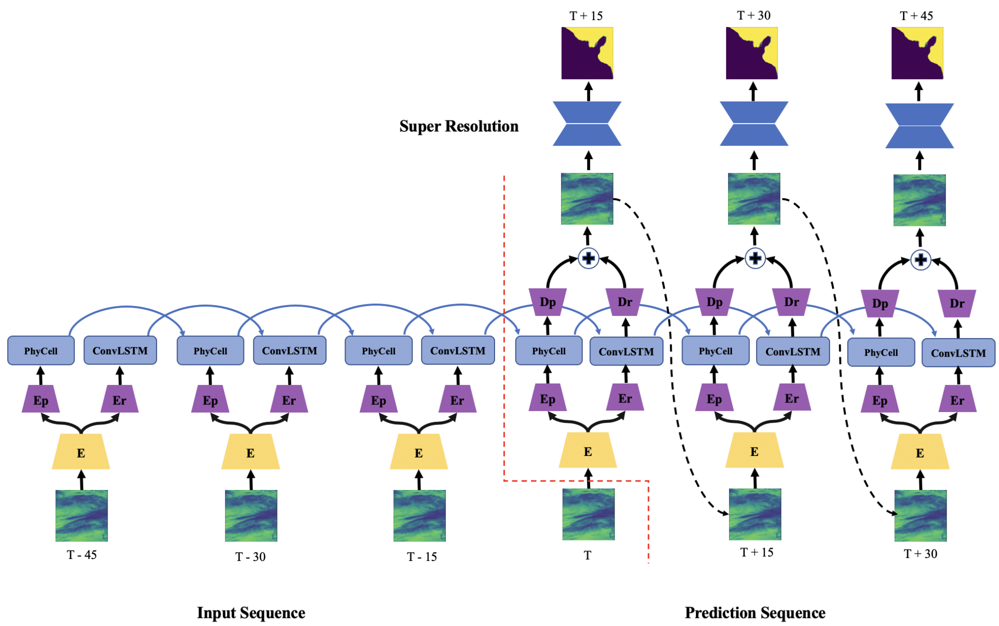

Fig. 1. Overview of the proposed SR-PhyDNet architecture for satellite image sequence prediction, super-resolution and segmentation tasks. Satellite images are the input and a binary (rain in yellow, no rain in purple) mask is the output.

2.2. Baseline 3D convolutional model

A 3D U-Net architecture [12] is used as a baseline that can be trained end-to-end for this challenge. The model has a depth of 5 blocks. In each block, convolutional layers with 3D kernels are used to model spatio-temporal data relationships. Each block uses a Rectified Linear Unit (ReLU) activation function and dropout (0.4 probability) is added to prevent overfitting. The intermediate output size of the network is (32, 4, 252, 252), and a further convolutional layer and two further transpose convolutions are used to reshape outputs as tensors (16, 252, 252).

2.3. A PhyDNet/U-Net architecture

Our physics-based DL solution for this task is the SR-PhyDNet architecture. As shown in Fig. 1, it consists of two sub-networks; 1) PhyDNet, a physics-based architecture described by [7], used to predict the temporal evolution of satellite radiances, and 2) SR U-Net for the purpose of SR and image translation. The two sub-modules were trained independently and their combination makes the SR-PhyDNet a two-step workflow. PhyDNet is a two-branch recurrent architecture where one branch is a convolutional/recurrent ConvLSTM cell that models residual factors and the other is PhyCell which performs partial differential equation (PDE) constrained prediction with convolutional filters. The method is detailed in [8] and learns a latent space **H** where physics and residual factors are disentangled $h = h^p + h^r$ and follows the dynamics: $\partial_t h(t, x) = \partial_t h^p + \partial_t h^r$. The SR task is carried out by a U-Net sub-module, with a depth of 4, that converts PhyDNet satellite predictions into radar binary rain maps, and performs a translation from the satellite context to a smaller target patch.

Our approach was developed without considering the submissions to *weather4cast*, but interestingly the winning team [10] adopted a similar approach. However, the winning solution uses a three-module architecture, with PhyDNet predicting the satellite sequence into the future and a U-Net converting it to a rainfall prediction mask. However, the key difference is that an additional sat2rad U-Net is used to convert the input sequence into a rainfall mask which is provided as additional input channels to the final conversion U-Net. Another difference is that in this approach the full context region is maintained throughout all predictions and only at the end is the image cropped and up-scaled to match the target resolution As a new contribution, in this work we quantify the impact of including physical information through ablation experiments.

3. RESULTS AND ANALYSIS

In this section, we report the overall classification performance of the methods described above. Both the baseline 3D U-Net and the SR-PhyDNet models perform very well for non-rain events. The 3D U-Net shows some artefacts in its predictions but overall the number of FP is small. For

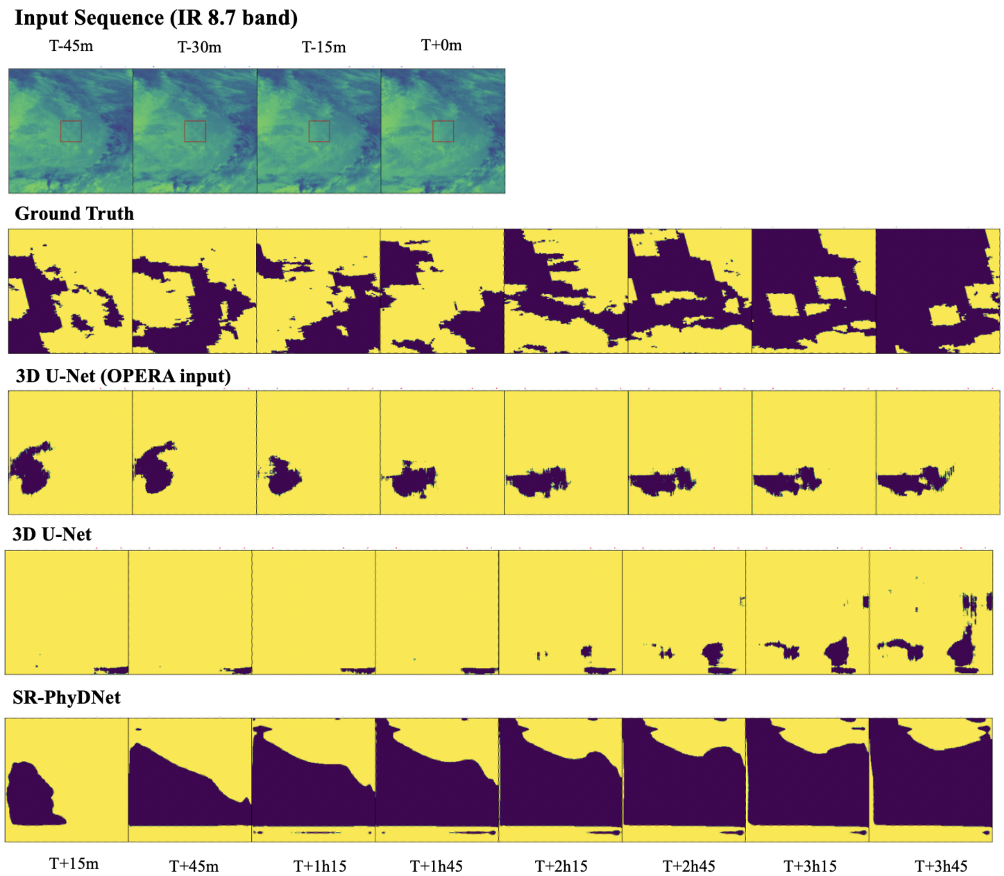

Fig. 2. Input satellite data (top, for one arbitrary infrared band) and comparison of different models with radar ground truth for a heavy rain event dissipating over time (in rows 2-5, rain in yellow, no rain in purple).

rain events, as shown in Table 1, SR-PhyDNet outperforms the 3D U-Net by 3.2% in terms of IoU. The drivers behind these metrics can be seen when the prediction sequences of each method are visualised. SR-PhyDNet achieves better scores and visually has fewer artefacts than the baseline as shown in Fig. 2. For this precipitation event dissipating over 4 hours, the baseline results are generally more static than the physics-aware model. On the other hand, SR-PhyDNet tends to produce rounded-edged patterns likely due to overly strong diffusion at the edges. The predictions of the 3D U-Net contain straight lines and box-like artefacts in many instances. The use of OPERA input (a simpler task) provides more accurate predictions spatially for early lead times and reduced presence of artefacts, but cannot be extended to all zones covered by the satellite.

The metrics in Table 1 support this analysis. While the baseline method has a better test recall and can therefore correctly identify more positive instances, it has a lower precision than that of SR-PhyDNet and is, therefore, more prone to overestimate the extent of rain events. The inclusion of physical information improves predictions, which are up to 20-30% better compared to when the PhyCell is bypassed (NoPhys).

Table 1. Comparison of proposed methods on the classification task. Metrics are averaged across test predictions. Best performing results among the first three are bolded.

Method	IoU/CSI	F1	Recall	Prec
3D U-Net	0.378	0.549	0.808	0.416
SR-PhyDNet	**0.391**	**0.562**	0.771	**0.442**
SR-PhyDNet (NoPhys)	0.300	0.461	**0.832**	0.319
3D U-Net (OPERA)	0.434	0.606	0.890	0.459

As a control experiment, the baseline architecture with only OPERA data as input (3D U-Net-OPERA) is tested to understand better the effect of using satellite input for the task. This method obtains the best overall performances as expected with an IoU of 0.434 (Table 1). However, temporal evolution similarly to the baseline remains too static (Fig. 2).

Fig. 3 shows that using OPERA data as input makes the task easier for shorter lead times (3D U-Net-OPERA). Using satellite-only input however, SR-PhyDNet performs better than 3D U-Net baseline and we show that including physics helps on longer times, beyond 2-3 hours lead times. These results are consistent with the findings of [10], who however did not quantify the specific contribution from the PhyCell. Besides the quantitative metrics, predictions of SR-PhyDNet appear more realistic visually, which is important for applications [11].

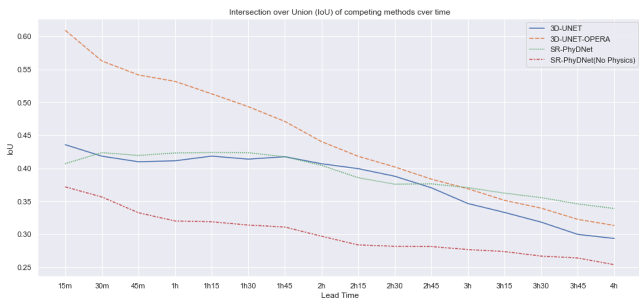

Fig. 3. IoU metrics over a 4-hour forecast period. Higher values indicate better predictions.

4. CONCLUSIONS AND FUTURE WORK

In this work, we propose a physics-aware DL workflow for predicting precipitation events with the spatial resolution of ground-based weather radars starting from a coarser resolution of weather satellite imagery. Following the *weather4cast* problem formulation, we study the advantage given by a combination of ConvLSTM networks with physics-aware components to predict satellite-like images of cloud evolution over time when compared to a baseline 3D U-Net architecture for a region with a balanced rain/no-rain distribution.

Our results show that while not a replacement for weather radar rain measurements, satellite data-driven deep learning models are suitable for predicting precipitation in areas lacking expensive ground-based instruments [6]. This should be tested by performing additional experiments on different regions with different climate conditions. As a novel contribution, we demonstrate the importance of physical constraints through the superior performances and more realistic patterns of the SR-PhyDNet workflow, which outperforms for longer lead times the baseline model even when precipitation is added in input.

REFERENCES

[1] B. Battrick, editor. *Meteosat second generation: the satellite development*. Number 153 in ESA BR. ESA Publications, Noordwijk, 1999. ISBN 978-92-9092-634-4.

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

[3] D. B. Bernet and coauthors. Characterizing precipitation events leading to surface water flood damage over large regions of complex terrain. *Environmental Research Letters*, 14 (6):064010, 2019. doi: 10.1088/1748-9326/ab127c.

[4] M. Bonavita and coauthors. Machine Learning for Earth System Observation and Prediction. *Bulletin of the American Meteorological Society*, 102(4):E710–E716, 2021. doi: 10.1175/BAMS-D-20-0307.1.

[5] L. Espeholt and coauthors. Skillful Twelve Hour Precipitation Forecasts using Large Context Neural Networks. *arXiv:2111.07470 [physics]*, 2021.

[6] P. G. Hill and coauthors. How skilful are nowcasting satellite applications facility products for tropical Africa? *Meteorol Appl.*, 27:e1966, 2020. doi: 10.1002/met.1966.

[7] V. Le Guen and N. Thome. Disentangling Physical Dynamics from Unknown Factors for Unsupervised Video Prediction. *arXiv:2003.01460 [cs]*, 2020.

[8] V. Le Guen and N. Thome. A Deep Physical Model for Solar Irradiance Forecasting with Fisheye Images. In *2020 IEEE/CVF Conf. (CVPRW)*, pages 2685–2688, 2020. doi: 10.1109/CVPRW50498.2020.00323.

[9] A. McGovern and coauthors. Making the Black Box More Transparent: Understanding the Physical Implications of Machine Learning. *Bulletin of the American Meteorological Society*, 100, 2019. doi: 10.1175/BAMS-D-18-0195.1.

[10] J. Pihrt and coauthors. Weatherfusionnet: Predicting precipitation from satellite data. *arXiv:2211.16824 [cs]*, 2022.

[11] S. Ravuri and coauthors. Skillful Precipitation Nowcasting using Deep Generative Models of Radar. *Nature*, 597(7878): 672–677, 2021. doi: 10.1038/s41586-021-03854-z.

[12] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv:1505.04597 [cs]*, 2015.

[13] E. Saltikoff and coauthors. OPERA the Radar Project. *Atmosphere*, 10(6):320, 2019. doi: 10.3390/atmos10060320.

[14] R. Schneider and coauthors. ESA-ECMWF report on recent progress and research directions in machine learning for earth system observation and prediction. *npj Clim Atmos Sci*, (51), 2022. doi: 10.1038/s41612-022-00269-z.

[15] M. G. Schultz and coauthors. Can deep learning beat numerical weather prediction? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2194):20200097, 2021. doi: 10.1098/rsta.2020.0097.

[16] J. Sun and coauthors. Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bulletin of the American Meteorological Society*, 95(3):409 – 426, 2014. doi: https://doi.org/10.1175/BAMS-D-11-00263.1.

[17] C. K. Sønderby and coauthors. MetNet: A Neural Weather Model for Precipitation Forecasting. *arXiv:2003.12140 [physics, stat]*, 2020.