Deep learning models for composite reflectivity prediction

Albu Alexandra
Babeș-Bolyai University

WeADL 2021 Workshop

The workshop is organized under the umbrella of WeaMyL, project funded by the EEA and Norway Grants under the number RO-NO-2019-0133.
Contract: No 26/2020.

Working together for a green, competitive and inclusive Europe

1. Computational approaches for nowcasting
2. AutoNowP
3. NowcastX
4. Conclusions and future directions of research

Computational approaches for nowcasting

- Computational approaches: Numerical Weather Prediction methods, optical flow algorithms
- Deep learning methods
  - provide a data-driven approach:
    - minimal assumptions about the physical system
    - learn patterns from the data itself
  - model nowcasting as a spatio-temporal prediction problem $\rightarrow$ convolutional and recurrent networks
  - learn a mapping from weather states in a geographical region at timestamps $t-k$, $t-k+1$ ..., t to the state at that location at timestamps $t+1$, $t+2$, ... $t+p$, where $k, p \geq 1$

Literature review of deep learning nowcasting models

- Two categories of deep learning approaches:
  - recurrent neural networks: ConvLSTMs [10, 12, 13], TrajGRUs [11, 7]
  - fully convolutional neural networks (i.e. convolutions applied on concatenated timestamps): U-Net [3, 14, 4], 3D convolutions [9], causal convolutions [5]

Literature review of deep learning nowcasting models

Neural networks can be trained by optimizing:

- pixel-wise loss functions (Mean Squared Error, Root Mean Squared Error, Mean Absolute Error)
- similarity losses [13]
- weighted loss functions [11, 7]

Challenges and limitations of current deep learning models

- Underestimation of high values $\leftarrow$ highly imbalanced data sets
- Blurry predictions when training with traditional methods
- Predictions for large areas are difficult to obtain
- Lack of interpretability

Our Approach

The aim of our research is to improve the weather nowcasting solutions using deep learning techniques.
Approaches developed in the project so far:

- AutoNowP
- NowcastX

Radar Data Sources

Radar data available on the MET Norway THREDDS data server

- Composite reflectivity
  https://thredds.met.no/thredds/catalog/remotesensing/reflectivity-nordic/catalog.html
- Reflectivity on multiple elevations, corrected and uncorrected + velocity
  https://thredds.met.no/thredds/catalog/weamyl/Radar/catalog.html

Data analysis

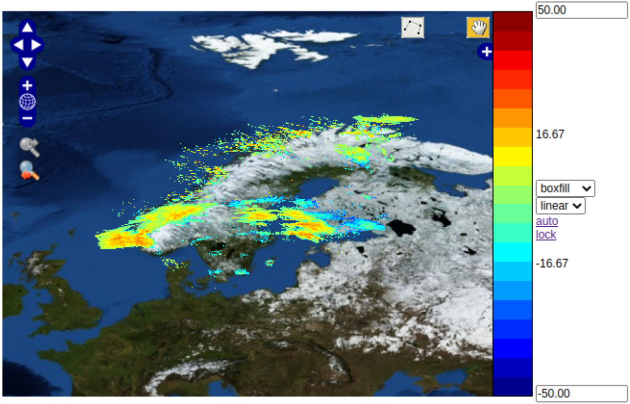


Figure: Visualization of composite reflectivity. From MET Norway THREDDS Data server [1]

AutoNowP classification model

- binary classification model
- predicts whether a point will have a value greater or smaller than a given threshold using the neighbours of that point at a previous timestamp
- uses two convolutional autoencoders - one for each class - trained to learn the characteristics of that class

AutoNowP classification model

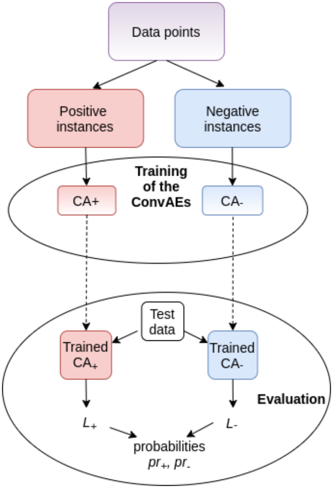


Figure: Overview of the AutoNowP approach.

Loss function

$$MSE_{greater}(X, X') = \frac{1}{d^2} \sum_{\substack{1 \leq i,j \leq d \\ x_{ij} > \tau}} (x_{ij} - x'_{ij})^2$$

$$MSE_{smaller}(X, X') = \frac{1}{d^2} \sum_{\substack{1 \leq i,j \leq d \\ x_{ij} \leq \tau}} (x_{ij} - x'_{ij})^2$$

$$\mathcal{L}(X, X') = \alpha \cdot MSE_{greater}(X, X') + (1 - \alpha) \cdot MSE_{smaller}(X, X')$$

where $X = (x_{ij})_{1 \leq i,j, \leq d}$ is the point neighbourhood,
$X' = (x'_{ij})_{1 \leq i,j, \leq d}$ is the reconstructed neighbourhood

Data set

Product of interest	**# instances**	**% of “+” instances**	**% of “-” instances**	**Entropy**
Composite reflectivity	6,607,836	31.97%	68,03%	0.904

Table: Description of the data set gathered from MET Norway THREDDS data server for a threshold of 10.

Evaluation metrics

- **Critical success index**: $CSI = \frac{TP}{TP+FN+FP}$
- **False alarm rate**: $FAR = \frac{FP}{TP+FP}$
- **Probability of detection**: $POD = \frac{TP}{TP+FN}$
- True skill statistic: $TSS = \frac{TP \cdot TN - FP \cdot FN}{(TP+FN) \cdot (FP+TN)}$
- Positive predictive value: $PV = \frac{TP}{TP+FP}$
- Negative predictive value: $NPV = \frac{TN}{TN+FN}$
- Specificity: $Spec = \frac{TN}{TN+FP}$
- Area Under the ROC Curve
- Area Under the Precision-Recall Curve

Results

τ	*CSI*	*TSS*	*POD*	*PPV*	*NPV*	*Spec*	*AUC*	*AUPRC*
10	**0.681** ± 0.014	**0.740** ± 0.009	**0.872** ± 0.019	**0.757** ± 0.027	**0.936** ± 0.005	**0.867** ± 0.026	**0.870** ± 0.005	**0.814** ± 0.008
15	0.566 ± 0.05	0.626 ± 0.09	0.675 ± 0.12	0.793 ± 0.08	0.920 ± 0.03	0.951 ± 0.03	0.813 ± 0.05	0.734 ± 0.029
20	0.401 ± 0.090	0.500 ± 0.223	0.536 ± 0.269	0.710 ± 0.173	0.947 ± 0.026	0.963 ± 0.046	0.750 ± 0.111	0.623 ± 0.048

Table: Experimental results for a 3-fold cross-validation evaluation procedure. 95% CIs are used for the results.

- performance decreases with the increase of the threshold

Comparison with other classifiers

Model	*CSI*	*TSS*	*POD*	*PPV*	*NPV*	*Spec*	*AUC*	*AUPRC*
AutoNowP	**0.681** ± 0.014	**0.740** ± 0.009	**0.872** ± 0.019	**0.757** ± 0.027	**0.936** ± 0.005	**0.867** ± 0.026	**0.870** ± 0.005	**0.814** ± 0.008
Logistic regression	0.760 ± 0.006	0.796 ± 0.002	0.853 ± 0.001	0.875 ± 0.007	0.932 ± 0.003	0.943 ± 0.002	0.898 ± 0.001	0.864 ± 0.004
Linear SVC	0.761 ± 0.006	0.798 ± 0.002	0.858 ± 0.001	0.870 ± 0.007	0.934 ± 0.003	0.940 ± 0.003	0.899 ± 0.001	0.864 ± 0.004
Decision tree	0.670 ± 0.010	0.710 ± 0.004	0.804 ± 0.005	0.801 ± 0.009	0.908 ± 0.003	0.906 ± 0.002	0.855 ± 0.002	0.803 ± 0.007
Nearest Centroid Classification	0.681 ± 0.009	0.728 ± 0.005	0.831 ± 0.009	0.791 ± 0.007	0.919 ± 0.001	0.897 ± 0.006	0.864 ± 0.003	0.811 ± 0.007

Table: Comparative results between *AutoNowP* and other classifiers. 95% CIs are used for the results.

- highest POD and NPV among all classifiers

NowcastX

- encoder-decoder convolutional neural network
- based on the Xception architecture [6]

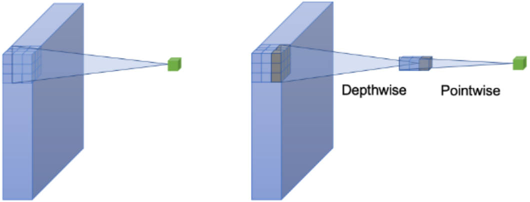
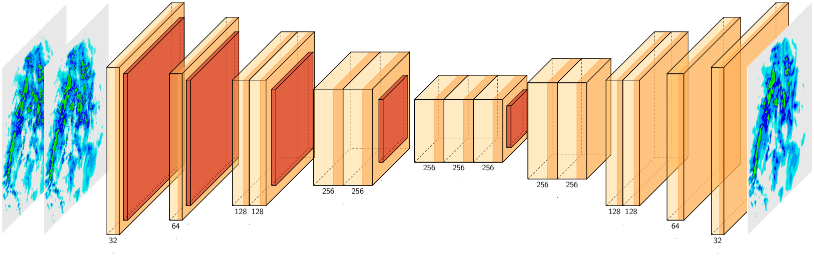


Figure: Convolution versus Depth-wise separable convolution. Picture taken from [8]

NowcastX



- Channel-wise concatenated past timestamps
- Single-step prediction
- Regression problem $\rightarrow$ RMSE loss

Architecture drawn using PlotNeuralNet [2]

Data sets

- **Composite reflectivity**
  - 10 days with meteorological events, selected from CAP warnings available at https://api.met.no/weatherapi/metalerts/1.1?show=all&lang=en
  - 8 days used for training, 1 for validation, 1 for testing
  - time resolution: 5 minutes
  - 200x200 region around Oslo
- **Base Reflectivity**
  Preliminary experiments:
  - uncorrected reflectivity on first level
  - 321 days with no missing timestamps
  - 128 days for training, 33 for validation, 160 for testing
  - time resolution: 10 minutes
  - 400x400 square (center of the radar grid)

Data analysis

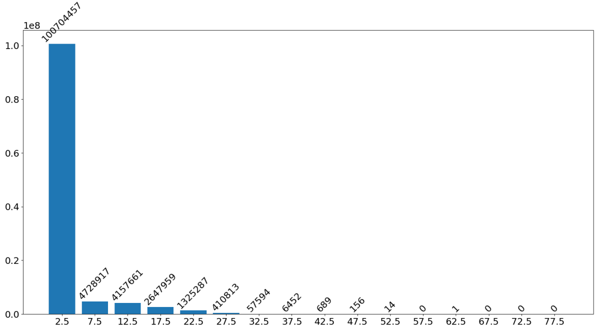


Figure: Histogram of composite reflectivity values in the 10-days dataset.

NowcastX - temporal context analysis

Goal

- evaluate the impact of the temporal context

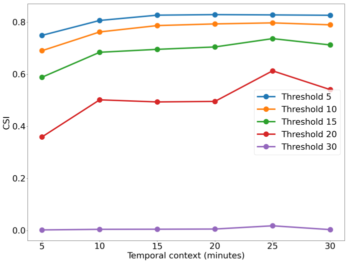
Training configuration

- multiple past timestamps concatenated channel-wise

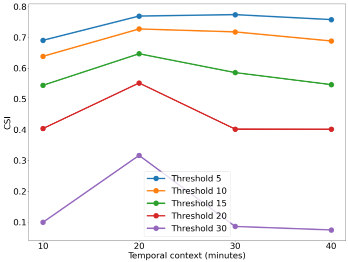
Evaluation measures

- CSI, FAR, POD metrics at multiple thresholds

NowcastX - preliminary results



(a) 10 days data set

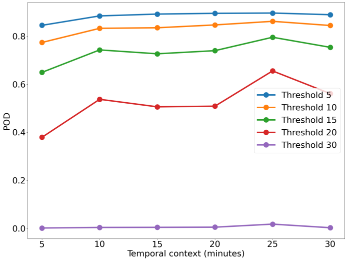


(b) 321 days data set

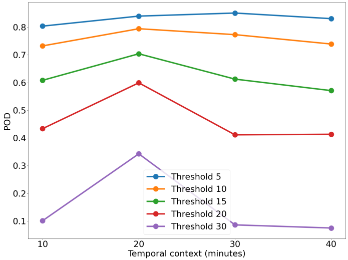
Figure: CSI metric for multiple timestamps and thresholds.

- performance increases up to 20-25 minutes, then stagnates or starts decreasing

NowcastX - preliminary results



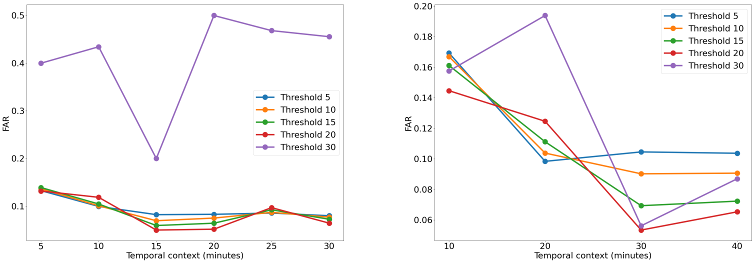
(a) 10 days data set



(b) 321 days data set

Figure: POD metric for multiple timestamps and thresholds.

NowcastX - preliminary results



(a) 10 days data set

(b) 321 days data set

Figure: FAR metric for multiple timestamps and thresholds.

NowcastX - sample predictions

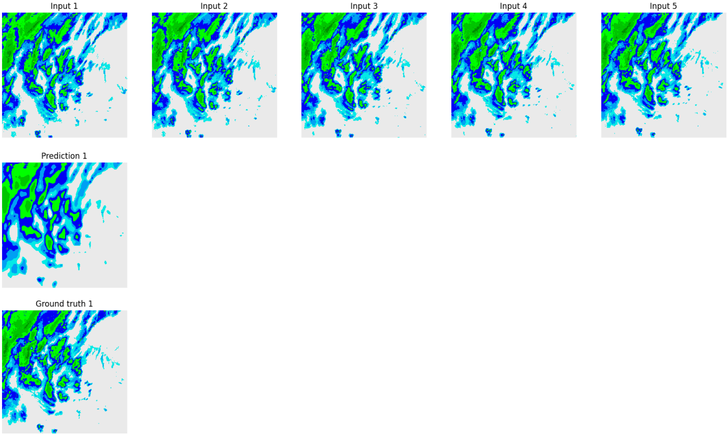


Figure: Predictions using the best model on the 10 days dataset.

NowcastX - sample predictions

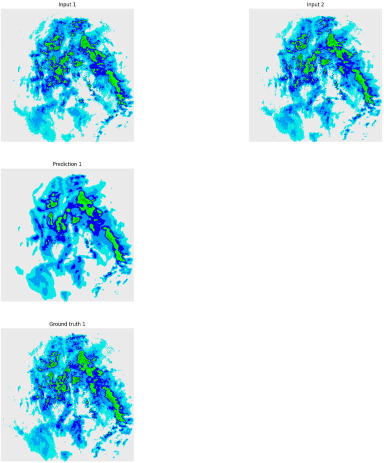


Figure: Predictions using the best model on the 321 days dataset.

Alternative loss function

- Model limitation: performance decreases for larger thresholds $\rightarrow$ the network fails to predict extreme values, which are relevant for nowcasting
- Proposed solution: use a weighted loss which puts more emphasis on errors obtained for high values [11]

Alternative loss function

- Model limitation: performance decreases for larger thresholds $\rightarrow$ the network fails to predict extreme values
- Proposed solution: use a weighted loss which puts more emphasis on errors obtained for high values [11]

$$\mathcal{L}_w(X, X') = \frac{1}{n^2} \sum_{1 \leq i,j \leq n} w(x_{ij}) \cdot (x_{ij} - x'_{ij})^2$$

where $X = (x_{ij})_{1 \leq i,j, \leq n}$ is the ground truth radar image, $X' = (x'_{ij})_{1 \leq i,j, \leq n}$ is the prediction and w is a step function which assigns higher weights to the errors corresponding to higher pixel values.

Alternative loss - preliminary results

Threshold	Loss	CSI	FAR	POD
5	RMSE	0.828	0.086	0.897
	$\mathcal{L}_w$	0.822	0.087	0.891
10	RMSE	0.797	0.087	0.863
	$\mathcal{L}_w$	0.790	0.096	0.863
15	RMSE	0.737	0.092	0.796
	$\mathcal{L}_w$	0.739	0.117	0.812
20	RMSE	0.613	0.097	0.656
	$\mathcal{L}_w$	0.629	0.125	0.691

Table: Comparative results with RMSE and weighted loss function for 5 timestamps using the 10 days data set, obtained using a step function with 5 intervals.

- the weighted loss provided higher CSI and POD than the RMSE for higher thresholds

Conclusions and future directions of research

- Accurate nowcasting of severe events is challenging

Future directions:

- multi-step prediction
- using an adaptive weighted loss
- quantifying uncertainty in our predictions

Thank you!

Bibliography I

MET Norway Thredds data server.
https://thredds.met.no/thredds/catalog.html.
Accessed: 2021-05-07.

PlotNeuralNet.
https://github.com/HarisIqbal88/PlotNeuralNet.
Accessed: 2021-05-25.

S. Agrawal, L. Barrington, C. Bromberg, J. Burge, C. Gazen, and J. Hickey.
Machine learning for precipitation nowcasting from radar images.
arXiv preprint arXiv:1912.12132, 2019.

Bibliography II

G. Ayzel, T. Scheffer, and M. Heistermann.
Rainnet v1. 0: a convolutional neural network for radar-based precipitation nowcasting.
Geoscientific Model Development, 13(6):2631–2644, 2020.

R. Castro, Y. M. Souto, E. Ogasawara, F. Porto, and E. Bezerra.
Stconvs2s: Spatiotemporal convolutional sequence to sequence network for weather forecasting.
Neurocomputing, 426:285–298, 2021.

F. Chollet.
Xception: Deep Learning with Depthwise Separable Convolutions.
In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1800–1807, 2017.

G. Franch, D. Nerini, M. Pendesini, L. Coviello, G. Jurman, and C. Furlanello.
Precipitation nowcasting with orographic enhanced stacked generalization: Improving deep learning predictions on extreme events.
Atmosphere, 11(3), 2020.

Y. Guo, Y. Li, L. Wang, and T. Rosing.
Depthwise convolution is all you need for learning multiple visual domains.
In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 8368–8375, 2019.

Bibliography IV

L. Han, J. Sun, and W. Zhang.
Convolutional neural network for convective storm nowcasting using 3-d doppler weather radar data.
IEEE Transactions on Geoscience and Remote Sensing, 58(2):1487–1495, 2020.

X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo.
Convolutional lstm network: A machine learning approach for precipitation nowcasting.
arXiv preprint arXiv:1506.04214, 2015.

Bibliography V

X. Shi, Z. Gao, L. Lausen, H. Wang, D.-Y. Yeung, W.-k. Wong, and W.-c. Woo.
Deep learning for precipitation nowcasting: A benchmark and a new model.
arXiv preprint arXiv:1706.03458, 2017.

C. K. Sønderby, L. Espeholt, J. Heek, M. Dehghani, A. Oliver, T. Salimans, J. Hickey, S. Agrawal, and N. Kalchbrenner.
Metnet: A neural weather model for precipitation forecasting.
arXiv, pages 1–17, 2020.

Q.-K. Tran and S.-k. Song.
Computer vision in precipitation nowcasting: Applying image quality assessment metrics for training deep neural networks.
Atmosphere, 10(5):244, 2019.

Bibliography VI

K. Trebing, T. Staczyk, and S. Mehrkanoon.
Smaat-unet: Precipitation nowcasting using a small attention-unet architecture.
Pattern Recognition Letters, 145:178–186, 2021.