Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology MeteoSwiss

comerge FAU Friedrich-Alexander-Universität Erlangen-Nürnberg MeteoSwiss
ETH zürich UNIVERSITY OF AMSTERDAM

Photographic Visualization of Weather Forecasts

Colloquium in Climatology, Climate Impact and Remote Sensing
September 28th, 2022 – Universität Bern

Christian Sigg (MeteoSwiss), Flavia Cavallaro (Comerge),
Tobias Günther (FAU) and Martin R. Oswald (ETH Zürich and UvA)

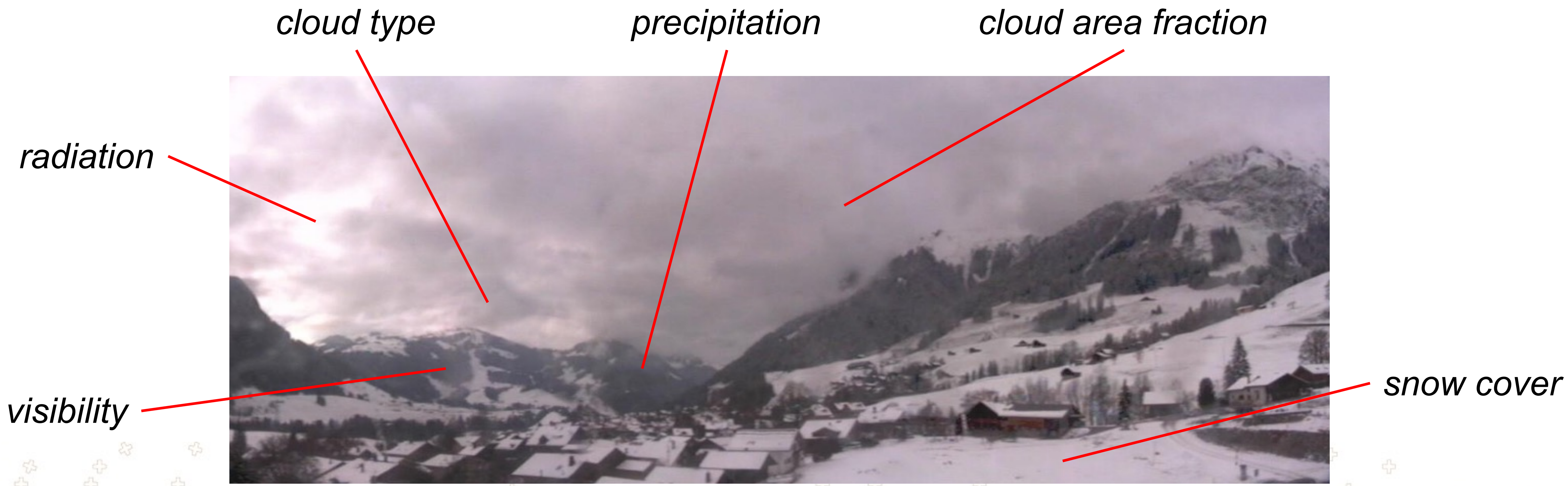
Contact: christian.sigg@meteoswiss.ch

Outline

- Motivation: Why Photographic Images?
- Baseline and Evaluation Criteria
- Method: Conditional GANs
- Results
- Conclusions and Future Work

Outdoor Weather Cameras

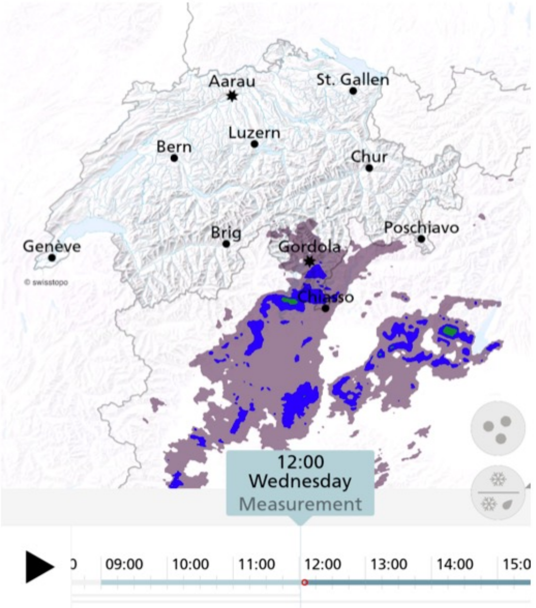
An information-dense yet accessible visualization of the past and present weather:



Useful for experts and the general public

Visualization of Weather Forecasts

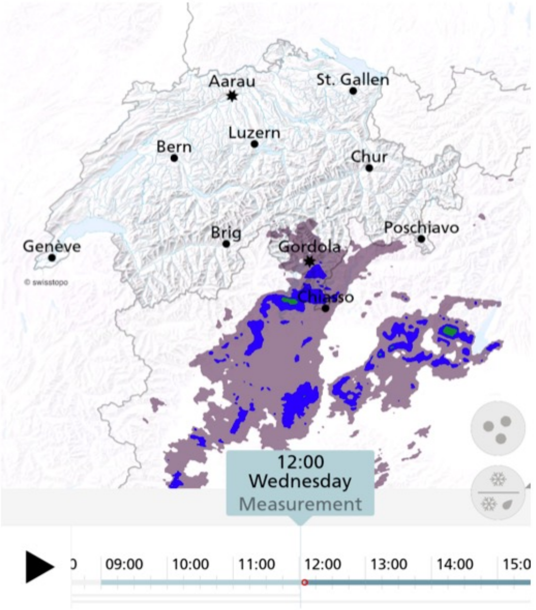
Use multiple visualizations, each focusing on different aspects:



Screenshots of the MeteoSwiss smartphone app

Visualization of Weather Forecasts

Seamless transition from observation to forecast:

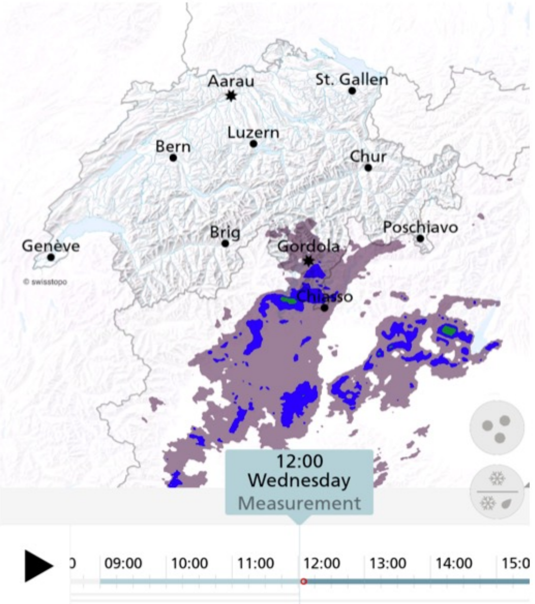


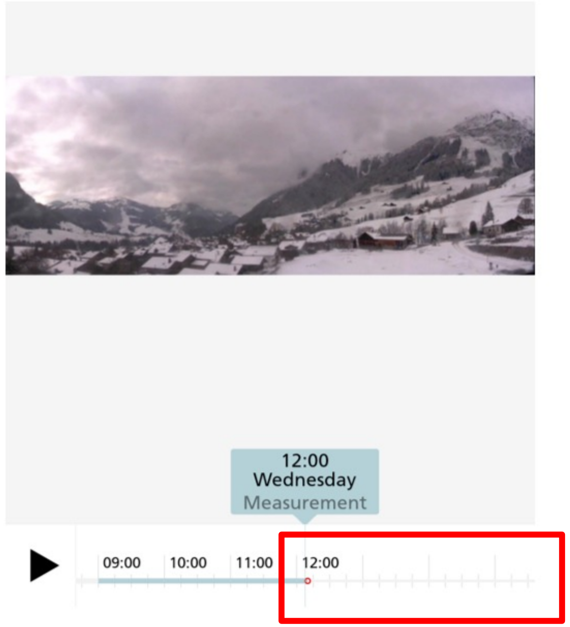
observation *forecast*

observation *forecast*

Visualization of Weather Forecasts

Also use photographic images to visualize future weather conditions!





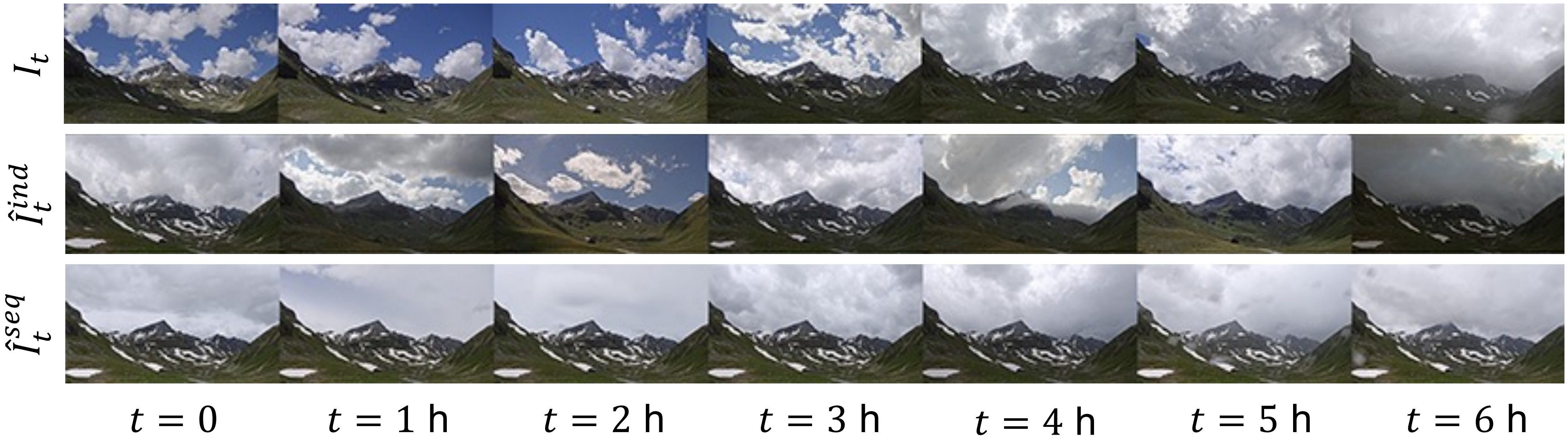
observation *forecast*

observation *forecast*

Outline

- Motivation: Why Photographic Images?
- Baseline and Evaluation Criteria
- Method: Conditional GANs
- Results
- Conclusions and Future Work

Baseline: Analog Retrieval

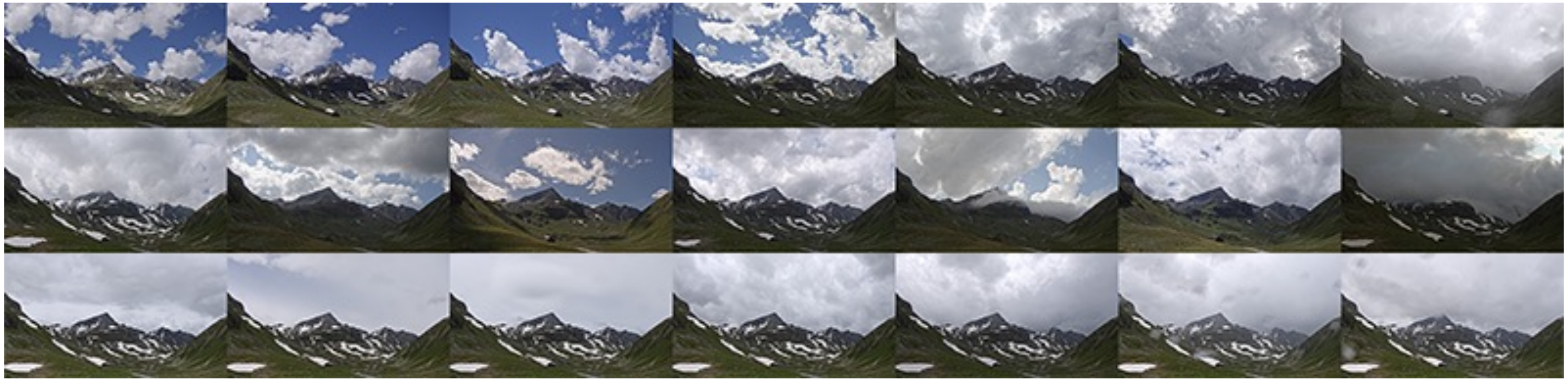


I_t Image sequence taken at Flüela, 10 to 16 UTC on July 2nd, 2020

$\hat{I}_t^{ind}$ Retrieval of best matching individual images from annotated archive

$\hat{I}_t^{seq}$ Retrieval of best matching sequence

Proposed Evaluation Criteria



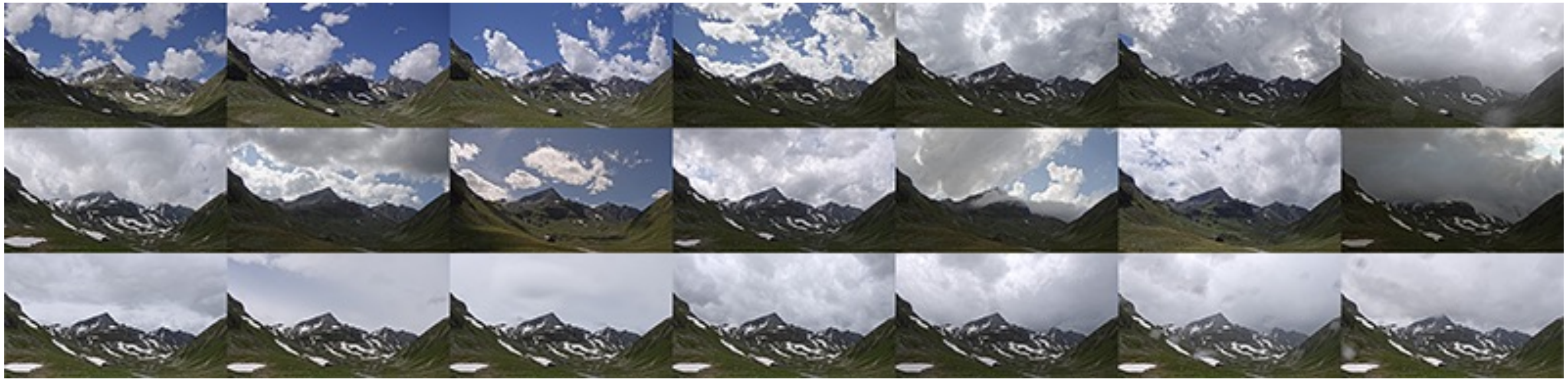
observations

analog images

analog sequence

I. Images should look real, no obvious artifacts
II. Match future atmospheric, ground and illumination conditions
III. Seamless transition from observation to forecast
IV. Visual continuity between consecutive images

Evaluation of Analog Retrieval



observations

analog images

analog sequence

	I. Realism	II. Matching conditions	III. Seamless transition	IV. Visual continuity
Analog images	🙂	😐	🙁	🙁
Analog sequence	🙂	🙁	🙁	🙂

High information density of images → retrieving analogs is not feasible 🤔

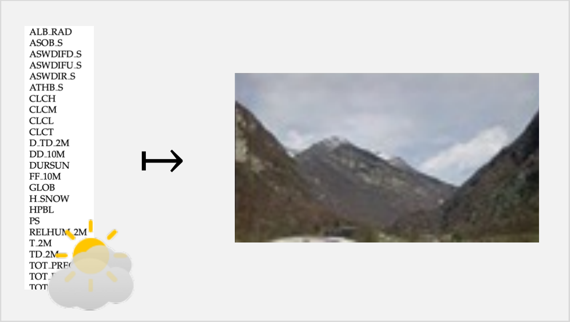
Outline

- Motivation: Why Photographic Images?
- Baseline and Evaluation Criteria
- Method: Conditional GANs
- Results
- Conclusions and Future Work

Image Synthesis: A Regression Problem

Generate photographic image $\hat{I}_t$, given forecast w_t of future weather conditions

$$G: w_t \mapsto \hat{I}_t$$



Generator $G(w;\theta)$ is a neural network, θ trained by minimizing expected loss

$$\operatorname{argmin}_\theta \mathbb{E}_{w_t, I_t}[L(G(w_t;\theta), I_t)]$$

Difficulty of Choosing Loss Function L

$$\text{argmin}_\theta \ \mathbb{E}_{w_t, I_t}[L(G(w_t;\theta), I_t)]$$

Forecast w_t does not determine exact shapes and locations of clouds
→ pixel-wise loss function is not appropriate, results in uniform sky:

$\hat{I}_t$ *for* L_1 *loss*

I_t

Goal: User should not be able to tell whether $\hat{I}_t$ or I_t is the real image, even if they are not identical.

Generative Adversarial Networks Goodfellow et al., 2014

Generator $G: z \mapsto \hat{I}$, creates image $\hat{I}$ from random input z

$z \mapsto$ 

Discriminator $D: I \mapsto [0, 1]$ tries to discriminate between real and generated images

$D(\hat{I}) \approx 0$ 👎

$D(I) \approx 1$ 👍

Adversarial Training of G and D

Generator $G: z \mapsto \hat{I}$, creates image I from random input z

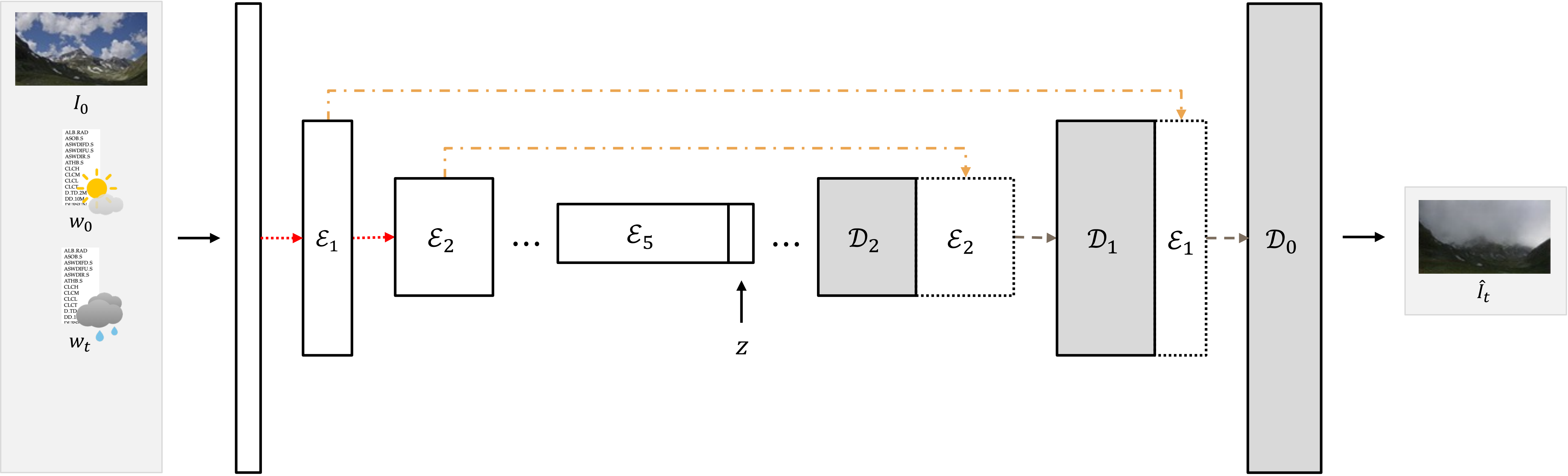
Discriminator $D: I \mapsto [0, 1]$, learns L implicitly through adversarial training

Joint **training objective** for $G(z;\theta)$ and $D(I;\eta)$:

$$\min_\theta \max_\eta \underbrace{\mathbb{E}_I[\log D(I;\eta)]}_{\textit{authenticate real images}} + \underbrace{\mathbb{E}_z[\log\{1 - D(\underbrace{G(z;\theta)}_{\textit{fool discriminator}};\eta\}]}_{\textit{spot fake images}}$$

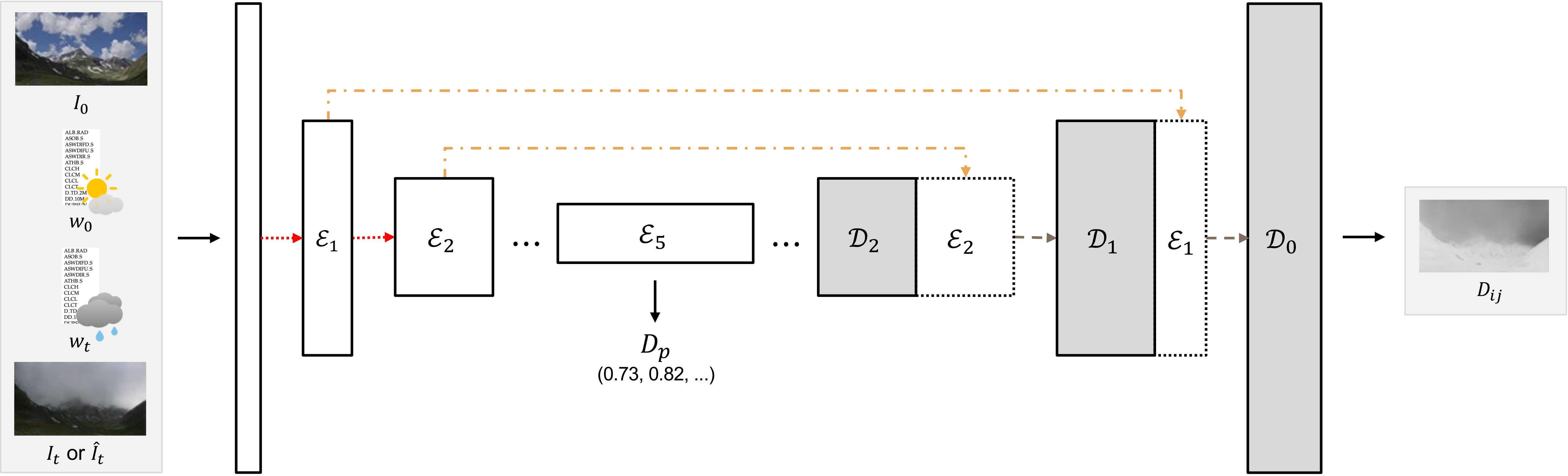
The training is completed successfully if G creates realistic images and the accuracy of D is not better than random guessing.

Our Generator Architecture



- Conditional Generator Mirza and Osindero, 2014 transforms current image I_0
- Encoder-decoder with skip connections Ronneberger et al., 2015
- Spectral normalization applied to all convolution layers Miyato et al., 2018

Our Discriminator Architecture



- Conditional discriminator $D(I|I_0, w_0, w_t)$
- Two output heads: patch-level D_p and pixel-level D_{ij} Schonfeld et al., 2020

Training Progress

One video frame per 1000 mini-batch updates of trainiable weights:

Current image I_0 *Future image* I_t *Two visualizations* $\hat{I}_t(z_1)$ and $\hat{I}_t(z_2)$

Pixel level discriminations

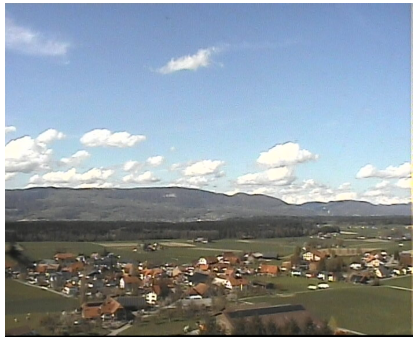
$D_{ij}(I_t)$ $D_{ij}\left(\hat{I}_t(z_1)\right)$ $D_{ij}\left(\hat{I}_t(z_2)\right)$

Outline

- Motivation: Why Photographic Images?
- Baseline and Evaluation Criteria
- Method: Conditional GANs
- Results
- Conclusions and Future Work

Evaluation Data

Cevio (elevation 421 m)



Etziken (elevation 524 m)

Flüela (elevation 2177 m)

Descriptor w: time of day, day of year, 31 COSMO-1 hourly output fields

Training: all pairs (I_0, w_0) and (I_t, w_t), $t \in [0, 10, 20, \ldots, 360 \text{ min}]$ of 2019

Test: Jan to Aug of year 2020 (until decommissioning of COSMO-1 at MCH)

Downscaled to 64 x 128 pixels to speed up training and conserve GPU memory

I. Realism

What is your first impression of the image?

"Looks realistic" vs. "Looks artificially generated"

real

generated

generated

real

generated

real

real

generated

Results of perceptual evaluation

Perceptual evaluation of generative models is the gold standard Zhou et al., 2019
→ Conducted a user study with 5 MeteoSwiss experts on 450 images:

	Judgment	
Actual	Real	Generated
Real	57	18
Generated	43	32

Cevio: 59 % accuracy

	Judgment	
Actual	Real	Generated
Real	52	23
Generated	32	43

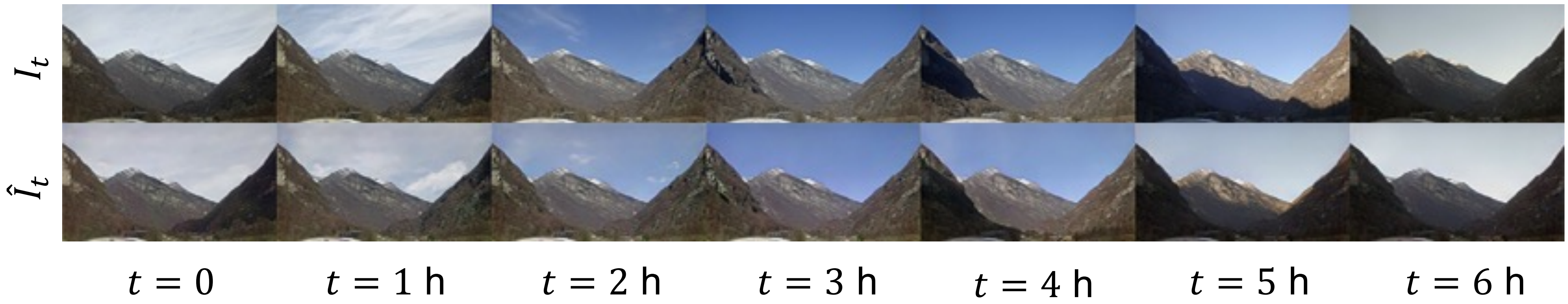
Etziken: 63 % accuracy

	Judgment	
Actual	Real	Generated
Real	57	18
Generated	49	26

Flüela: 55 % accuracy

Overall examiner’s accuracy of 59 % (55 to 64 % bootstrap CI) is not much better than random guessing. 🙂

II. Matching Weather Conditions



A pixel-wise comparison of forecast visualization $\hat{I}_t$ and true future image I_t is not appropriate → use overall descriptive criteria to evaluate accuracy:

- **Atmosphere**: cloud cover, cloud type, visibility
- **Ground**: dry, wet, frost, snow
- **Illumination**: time of day, diffuse or direct

Two possible causes for mismatches

Example: mismatch in observed cloud cover

$\hat{I}_t$

I_t

Inaccurate forecast. w_t does not accurately describe visible conditions

- Evaluating COSMO-1 output fields at camera site can be insufficient
- Spatial or temporal resolution of output fields can be too coarse

Inconsistent visualization. G fails to properly account for the changes from w_0 to w_t in the transformation of I_0 to $\hat{I}_t$.

Matching evaluation results

Evaluation by three experts on 450 image pairs:


$\hat{I}_t$


I_t

	Matching conditions							
	Atmosphere				Illumination			
Camera	C. cover	C. type	Visibility	Ground	Time of day	Sunlight	w_t accurate	$\hat{I}_t$ consistent
Cevio	**71 (63, 77)**	**79 (71, 84)**	97 (91, 99)	**99 (95, 100)**	**100**	94 (89, 97)	61 (53, 68)	**70 (62, 76)**
Etziken	**69 (61, 75)**	**82 (75, 87)**	90 (83, 93)	**99 (95, 100)**	**99 (93, 99)**	90 (83, 93)	60 (51, 67)	**73 (65, 79)**
Flüela	61 (52, 68)	80 (72, 85)	73 (65, 79)	**100**	97 (93, 99)	89 (83, 93)	51 (42, 58)	**57 (49, 65)**
All	**67 (62, 71)**	**80 (76, 84)**	86 (83, 89)	**99 (98, 100)**	**98 (96, 99)**	91 (88, 93)	57 (52, 62)	**67 (62, 71)**

- Values are percentages of the possible maximum 150 matches
- Values in parentheses are 95 % bootstrap CIs
- Bold values: CIs do not overlap with corresponding sequence analog CIs

Analysis of cloud cover mismatches

At Cevio, 44 out 150 cases (29 %) had a mismatch in cloud cover:

	Matching conditions							
	Atmosphere				Illumination			
Camera	C. cover	C. type	Visibility	Ground	Time of day	Sunlight	w_t accurate	$\hat{I}_t$ consistent
Cevio	**71 (63, 77)**	**79 (71, 84)**	97 (91, 99)	**99 (95, 100)**	**100**	94 (89, 97)	61 (53, 68)	**70 (62, 76)**
Etziken	**69 (61, 75)**	**82 (75, 87)**	90 (83, 93)	**99 (95, 100)**	**99 (93, 99)**	90 (83, 93)	60 (51, 67)	**73 (65, 79)**

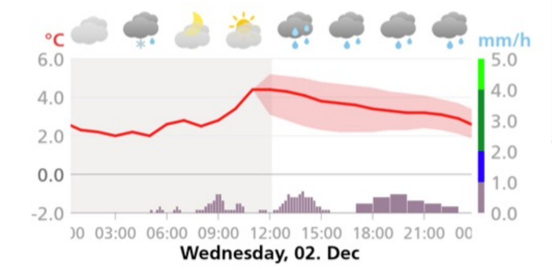
Example: But w_t predicted 100 % cloud area fraction in medium troposphere!

$\hat{I}_t$

I_t

Overall: only in 9 of those 44 cases was the mismatch due to the visualization 🙂

III. Seamless Transition



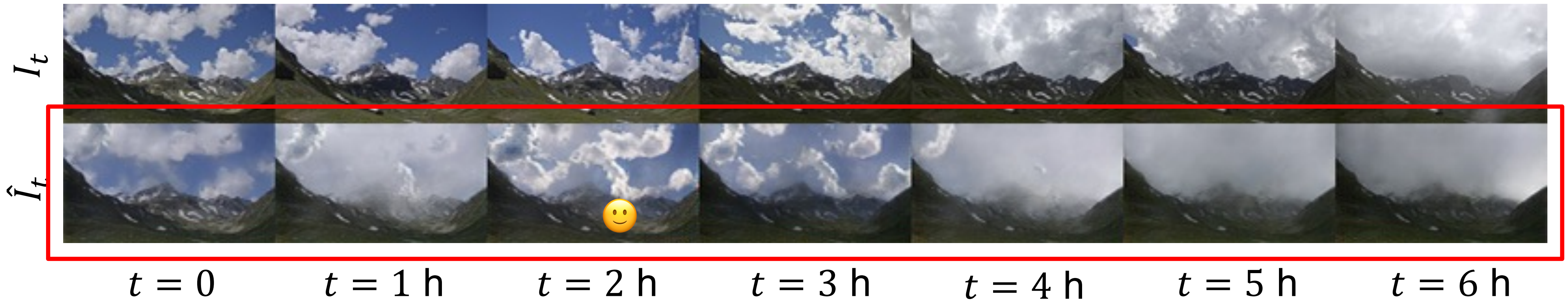
Reproduce current image, $I_0 = \hat{I}_0$:

$t = 0$ | $t = 1$ h | $t = 2$ h | $t = 3$ h | $t = 4$ h | $t = 5$ h | $t = 6$ h

Pixel-wise RMSE$(I_0, \hat{I}_0) = 1.11\times10^{-2}$, compared to 4.64×10^{-2} for sequence analogs $\hat{I}_0^{seq}$ 🙂

III. Seamless Transition

Retain present weather conditions as long as they persist into the future:



Only possible because $G(I_0, z|w_0, w_t)$ transforms I_0 into $\hat{I}_t$

IV. Visual Continuity

Natural looking cloud development and movement of shadows:

I_t

$\hat{I}_t$

$t = 0$ | $t = 1$ h | $t = 2$ h | $t = 3$ h | $t = 4$ h | $t = 5$ h | $t = 6$ h

IV. Visual Continuity

Change of daylight conditions, including the appearance of artifical lights:

I_t

$\hat{I}_t$

$t = 0$ $t = 1$ h $t = 2$ h $t = 3$ h $t = 4$ h $t = 5$ h $t = 6$ h

IV. Visual Continuity

Generate the correct newly visible scenery:

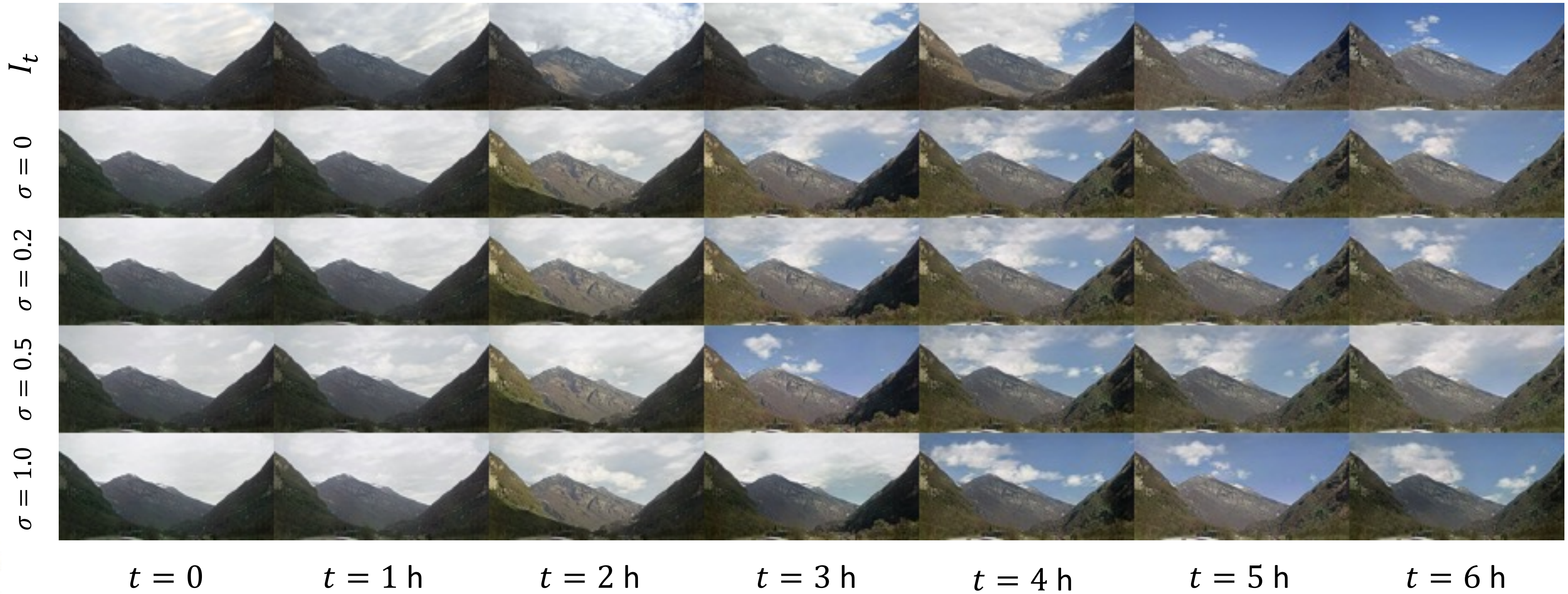
I_t

$\hat{I}_t$

$t = 0$ $t = 1$ h $t = 2$ h $t = 3$ h $t = 4$ h $t = 5$ h $t = 6$ h

Visual Continuity vs. Image Diversity

Increasing σ when sampling $z_i \sim \mathcal{N}(0, \sigma^2)$ leads to a greater image diversity that is deemed consistent with w_t:



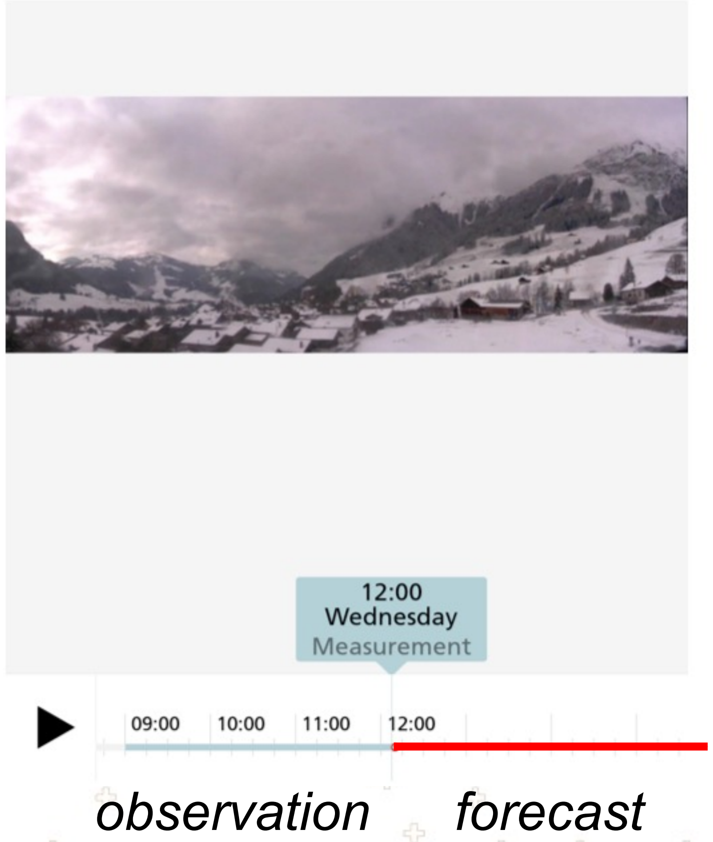
Outline

Conclusions

- Photographic images can also visualize future weather conditions
- Look realistic, match predicted weather conditions, attain seamless transition from observation to forecast and visual continuity

Applications:

- Communicate localized forecasts in webcam feeds, smartphone app
- Provide similar service to communities and tourism organizations



Future Work

- Use more accurate and descriptive weather descriptors, to improve matching the forecast *and* the conditions visible in the future
- Scale image size beyond 64 x 128 pixels e.g. using Karras et al., 2018
- Improve transformations involving translations of isolated clouds:

I_t

$\hat{I}_t$

$t = 0$ $t = 1$ h $t = 2$ h $t = 3$ h $t = 4$ h $t = 5$ h $t = 6$ h

  (Including self-attention layers Zhang et al., 2019 did not help)
- Synthesize whole sequences to improve temporal evolution Wu et al., 2020

Resources

Our paper is currently under review, the pre-print is available at

https://arxiv.org/abs/2203.15601

Tensorflow code, trained models and results are available at

https://github.com/meteoswiss/photocast

Acknowledgments

We thank Rega for giving us the permission to use images from the Cevio camera in this study.

We thank Tanja Weusthoff for the preparation of the COSMO-1 forecast data.

We thank Christian Allemann, Yannick Bernard, Thérèse Obrist, Eliane Thürig, Deborah van Geijtenbeek and Abbès Zerdouk for taking part in the perceptual evaluations.

We thank Daniele Nerini for providing his expertise on nowcasting and post-processing of forecasts.

Bibliography

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. Advances in neural information processing systems, 27.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2017). Progressive growing of gans for improved quality, stability, and variation. arXiv preprint arXiv:1710.10196.

Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.

Miyato, T., Kataoka, T., Koyama, M., & Yoshida, Y. (2018). Spectral normalization for generative adversarial networks. arXiv preprint arXiv:1802.05957.

Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241).

U. Schättler, G. Doms, and C. Schraff. (2021). COSMO- Model Version 6.00: A Description of the Non- hydrostatic Regional COSMO-Model - Part VI: Model Output and Data Formats for I/O.

Schonfeld, E., Schiele, B., & Khoreva, A. (2020). A u-net based discriminator for generative adversarial networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 8207-8216).

Wu, S., Xiao, X., Ding, Q., Zhao, P., Wei, Y., & Huang, J. (2020). Adversarial sparse transformer for time series forecasting. Advances in Neural Information Processing Systems, 33, 17105-17115.

Yun, S., Han, D., Oh, S. J., Chun, S., Choe, J., & Yoo, Y. (2019). Cutmix: Regularization strategy to train strong classifiers with localizable features. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 6023-6032).

Zhang, H., Goodfellow, I., Metaxas, D., & Odena, A. (2019, May). Self-attention generative adversarial networks. In International conference on machine learning (pp. 7354-7363). PMLR.

Zhou, S., M. Gordon, R. Krishna, A. Narcomey, L. F. Fei-Fei, and M. Bernstein, 2019: HYPE: A Benchmark for Human eYe Perceptual Evaluation of Generative Models. *Advances in Neural Information Processing Systems*, Curran Associates, Inc., Vol. 32.

Generator Objectives to be Minimized

How much $G(I_0, z|w_0, w_t)$ struggles to fool the discriminator on the patch level

$$\mathbb{E}_{I_0,w_0,w_t}\mathbb{E}_z\left[\sum_p \log\left[D_p(G(I_0, z|w_0, w_t)|I_0, w_0, w_t)\right]\right]$$

and on the pixel level

$$\mathbb{E}_{I_0,w_0,w_t}\mathbb{E}_z\left[\sum_{ij} \log\left[D_{ij}(G(I_0, z|w_0, w_t)|I_0, w_0, w_t)\right]\right]$$

How similar two generated images look at the pixel level, given different random inputs $z_1, z_2 \sim \mathcal{N}(0, 1)$

$$-\mathbb{E}_{I_0,w_0,w_t}\mathbb{E}_{z_1,z_2}\left[\sum_{ijc} \left|G_{ijc}(I_0, z_1|w_0, w_t) - G_{ijc}(I_0, z_2|w_0, w_t)\right|\right]$$

Discriminator Objectives to be Maximized

How well the patch head D_p authenticates real images

$$\mathbb{E}_{I_0,w_0,I_t w_t}\left[\sum_p \log D_p(I_t|I_0,w_0,w_t)\right]$$

and spots generated images

$$\mathbb{E}_{I_0,w_0,w_t}\mathbb{E}_z\left[\sum_p \log\left[1 - D_p(G(I_0,z|w_0,w_t)|I_0,w_0,w_t)\right]\right]$$

How well the pixel head D_{ij} can distinguish pixels of a cut-mix Yun et al., 2019 composite C

$$\mathbb{E}_C\left[\sum_{ij} M_{ij}D_{ij}(C) + \left(1 - M_{ij}\right)\log(1 - D_{ij}(C))\right]$$

Artifacts Induced by Residual Learning He et al., 2015

Present image I_0

Forecast visualization $\hat{I}_t$

Clouds in I_0 are still partially visible in the clear sky regions of $\hat{I}_t$

→ Residual transformation learned by the generator does not fully cancel their appearance

Subset of COSMO-1 Output Fields Schättler et al., 2021

Abbreviation	Unit	Name
ALB_RAD	%	Surface albedo for visible range, diffuse
ASOB_S	W/m^2	Net short-wave radiation flux at surface
ASWDIFD_S	W/m^2	Diffuse downward short-wave radiation at the surface
ASWDIFU_S	W/m^2	Diffuse upward short-wave radiation at the surface
ASWDIR_S	W/m^2	Direct downward short-wave radiation at the surface
ATHB_S	W/m^2	Net long-wave radiation flux at surface
CLCH	%	Cloud area fraction in high troposphere (pressure below ca. 400 hPa)
CLCM	%	Cloud area fraction in medium troposphere (between ca. 400 and 800 hPa)
CLCL	%	Cloud area fraction in low troposphere (pressure above ca. 800 hPa)
CLCT	%	Total cloud area fraction
D_TD_2M	K	2 m dew point depression
DD_10M	°	10 m wind direction
DURSUN	s	Duration of sunshine
FF_10M	m/s	10 m wind speed
GLOB	W/m^2	Downward shortwave radiation flux at surface
H_SNOW	m	Snow depth
HPBL	m	Height of the planetary boundary layer
PS	Pa	Surface pressure (not reduced)
RELHUM_2M	%	2 m relative humidity (with respect to water)
T_2M	K	2 m air temperature
TD_2M	K	2 m dew point temperature
TOT_PREC	kg/m^2	Total precipitation
TOT_RAIN	kg/m^2	Total precipitation in rain
TOT_SNOW	kg/m^2	Total precipitation in snow
U_10M	m/s	10 m grid eastward wind
V_10M	m/s	10 m grid northward wind
VMAX_10M	m/s	Maximum 10 m wind speed