Predicting ice flow dynamics using machine learning

Yimeng Min, S. Karthik Mukkavilli and Yoshua Bengio

Mila - Quebec AI Institute, Montreal, Canada
University de Montreal, Montreal, Canada

An Important Problem

Though machine learning has achieved notable success in modeling sequential and spatial data for speech recognition and in computer vision, applications to remote sensing and climate science problems are seldom considered. In this paper, we demonstrate techniques from unsupervised learning of future video frame prediction, to increase the accuracy of ice flow tracking in multi-spectral satellite images. As the volume of cryosphere data increases in coming years, this is an interesting and important opportunity for machine learning to address a global challenge for climate change, risk management from floods, and conserving freshwater resources. Future frame prediction of ice melt and tracking the optical flow of ice dynamics presents modeling difficulties, due to uncertainties in global temperature increase, changing precipitation patterns, occlusion from cloud cover, rapid melting and glacier retreat due to black carbon aerosol deposition, from wildfires or human fossil emissions. We show machine learning method helps improve the accuracy of tracking the optical flow of ice dynamics compared to existing methods in climate science.

Model

We use a stochastic video generation with prior for prediction. The prior network observes frames $\mathbf{x}_{1:t-1}$ and output $\mu_\psi(\mathbf{x}_{1:t-1})$ and $\sigma_\psi(\mathbf{x}_{1:t-1})$ of a normal distribution and is trained with by maxing:

$$\mathcal{L}_{\theta,\phi,\psi}(x_{1:T}) = \sum_{t=1}^{T} [\mathbb{E}_{q_\phi(z_{1:t}|x_{1:t})} log p_\theta(x_t|x_{1:t-1}, z_{1:t}) \\ -\beta D_{KL}(q_\phi(z_t|x_{1:t})||p_\psi(z_t|x_{1:t-1}))]$$

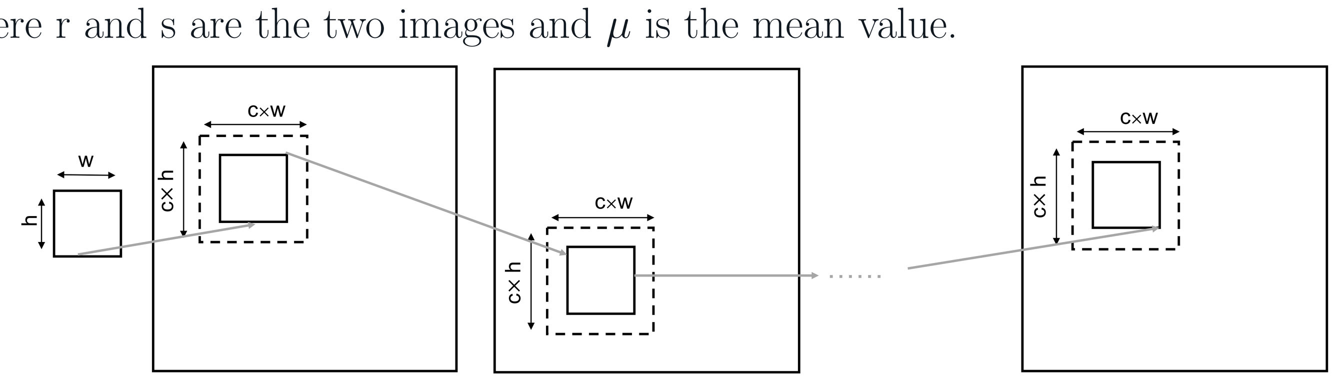
Where p_θ, q_ϕ and p_ψ are generated from convolutional LSTM. q_ϕ and p_ψ denote the normal distribution draw from $\mathbf{x}_t$ and $\mathbf{x}_{t-1}$ and p_θ is generated from encoding the $\mathbf{x}_{t-1}$ together with the $\mathbf{z}_t$. Subscene $\hat{\mathbf{x}}_t$ is generated from a decoder with a deep convolutional GAN architecture a by sampling on a prior $\mathbf{z}_t$ from the latent space drawing from the previous subscenes combined with the last subscene $\mathbf{x}_{t-1}$. After decoding, the predict subscene is passed back to the input of the prediction model and the prior. The latent space $\mathbf{z}_t$ is draw from $p_\psi(\mathbf{z}_t|\mathbf{x}_{1:t-1})$. The details of the model, also refered as stochastic video generation can be found in [1].

Labels

The images are denoted as F_i where i is from 1 to 12 and the frames(subscenes) in each image are $x_i^j \in R^{128\times128}$, where $i \in \{1...12\}$ and $j \in \{1...1525\}$. For finding the next subscene, or chip, that matches the x_{i-1}^j best, we compare the x_{i-1}^j to a range of possible regions by calculating the correlation between two chips, the equation writes as:

$$CI(r,s) = \frac{\sum_{mn}(r_{mn}-\mu_r)(s_{mn}-\mu_s)}{[\sum_{mn}(r_{mn}-\mu_r)^2]^{1/2}[\sum_{mn}(s_{mn}-\mu_s)^2]^{1/2}}$$

where r and s are the two images and μ is the mean value.



Labels

Previous results also show applying high pass filter on both sides of the pairs can be a feasible solution to increase the correlation at certain areas[3, 2].

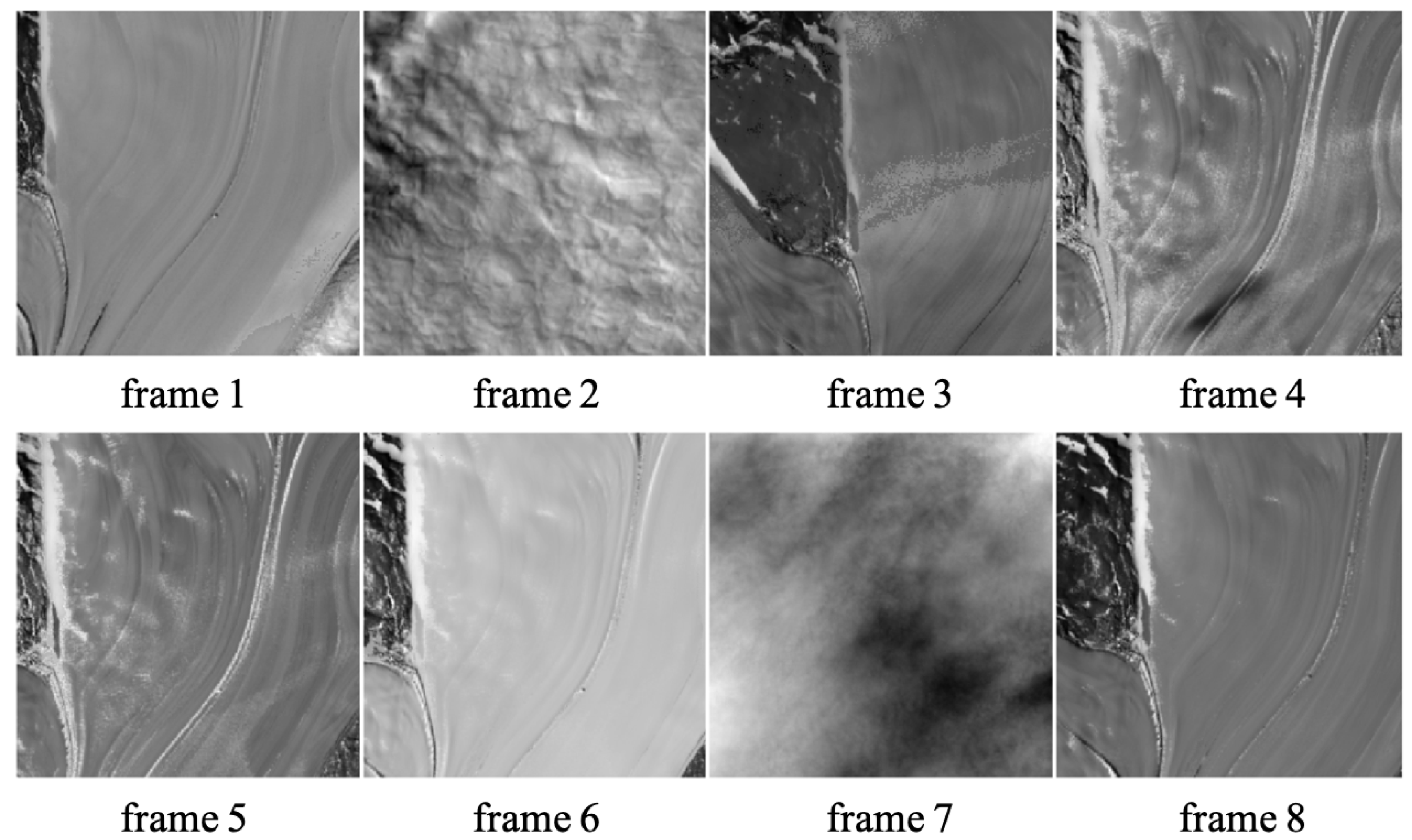


Fig. 2: The subscenes in our dataset, frame 2 and frame 7 are contaminated by the aerosol

Experiment Results and Discussion

Correlation	Mean	Persistence(Last frame)	Hi-pass Filter	Machine learning
		0.237	0.201	0.362
Low	< 0.3	0.699	0.598	0.393
Medium	0.3∼0.7	0.271	0.337	0.557
High	> 0.7	0.0300	0.0651	0.0504

Fig. 3: Results of three models.

We train our model with $\mathbf{z} \in R^{128}$ and 2 LSTM layers, each layer has 128 units. By conditioning on the past eight subscenes, the results of our model on different types of subscenes are shown in Figure 5 and 4.

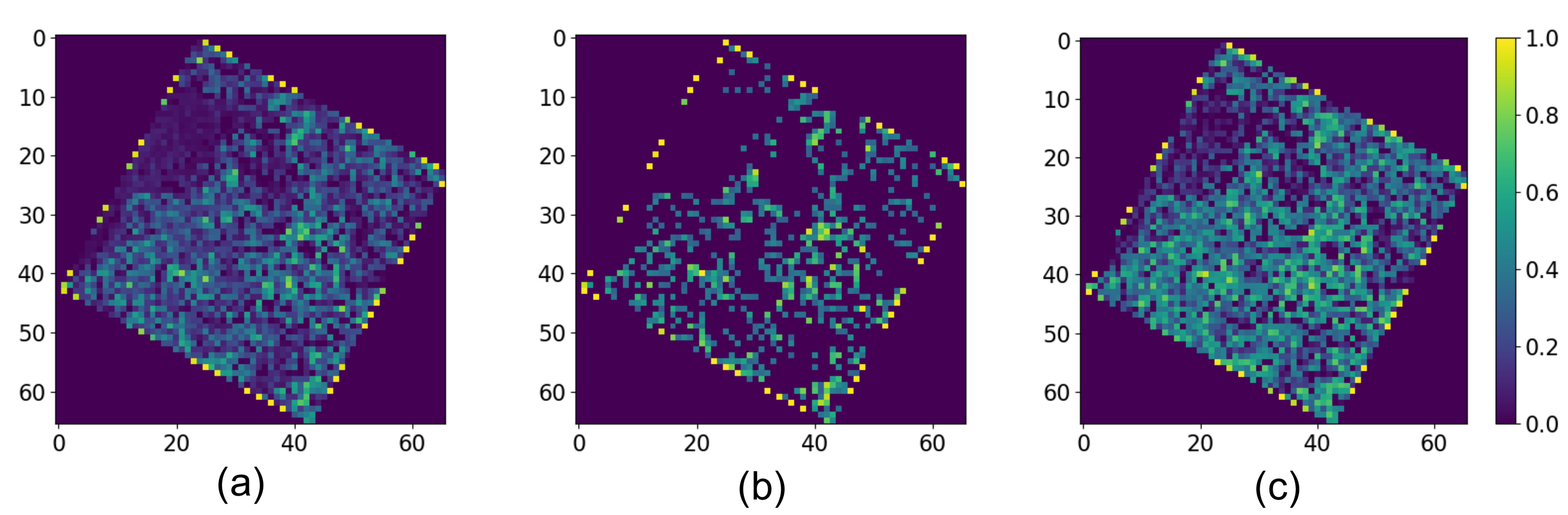


Fig. 4: The correlation map. a) persistence model(correlation between t_0 and t_2); b) high frequency model (correlation between filter$_0$

Experiment Results and Discussion

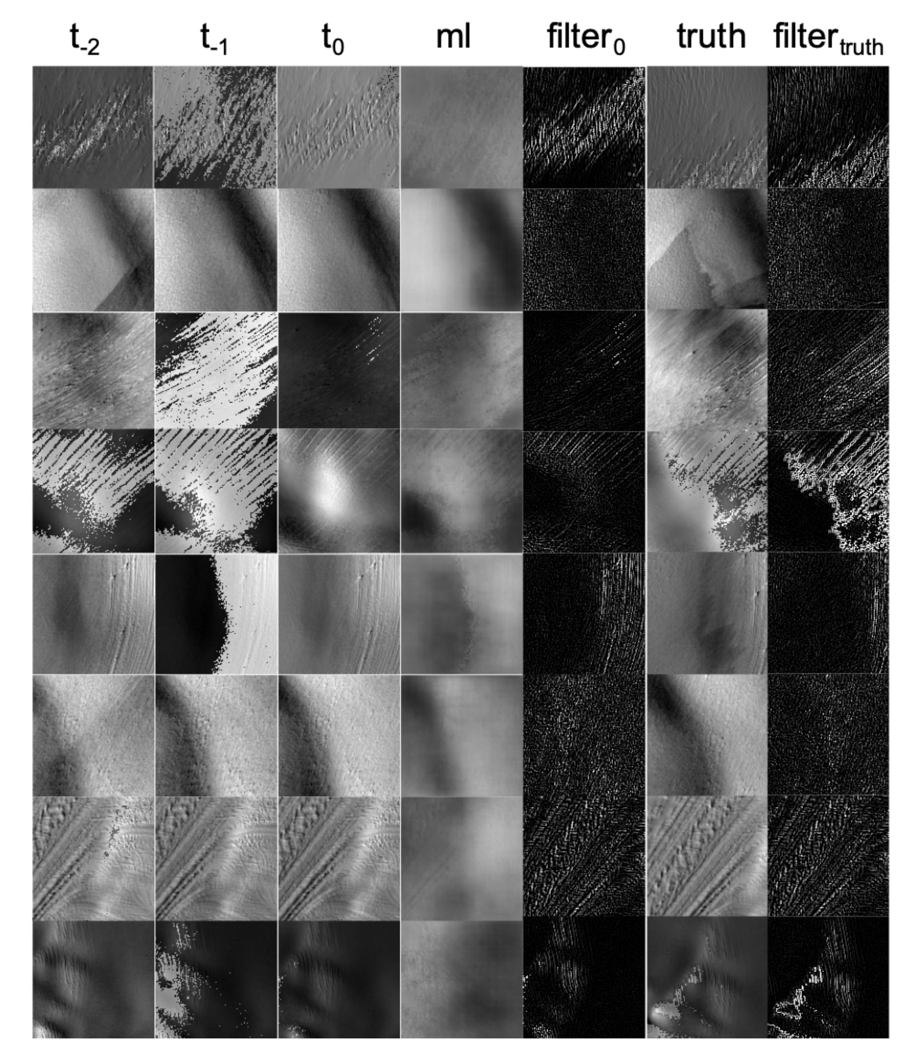


Fig. 5: Subscenes generated with different models, the first three columns: the past three subscenes; the fourth column: machine learning predicted next subscene; fifth column: high pass of t_0; sixth column: the ground truth; last column: high pass of ground truth.

Remarks

Our model can also be improved if more physical and environmental parameters are introduced into the model, for example, the wind speed and the aerosol optical depth components in the atmosphere. The first parameter provides a trend for the ice flow movement and the second parameter gives us a confidence factor about the satellite images' quality, dropout to particular frames can be applied if the aerosol optical depth rises over a threshold. Furthermore, black carbon aerosols were found to accelerate ice loss and glacier retreat in the Himalayas and Arctic from both wildfire soot deposition and fossil fuel emissions.

Acknowledgements

Y.M. thanks Emily Denton for the help in implementing the generative model.

References

[1] Emily Denton and Rob Fergus. "Stochastic video generation with a learned prior". In: *arXiv preprint arXiv:1802.07687* (2018).

[2] Mark Fahnestock et al. "Rapid large-area mapping of ice flow using Landsat 8". In: *Remote Sensing of Environment* 185 (2016), pp. 84–94.

[3] Remko de Lange, Adrian Luckman, and Tavi Murray. "Improvement of satellite radar feature tracking for ice velocity derivation by spatial frequency filtering". In: *IEEE transactions on geoscience and remote sensing* 45.7 (2007), pp. 2309–2318.