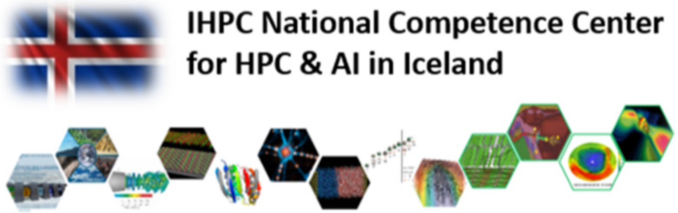
Towards Frequent Update of Land Cover Maps

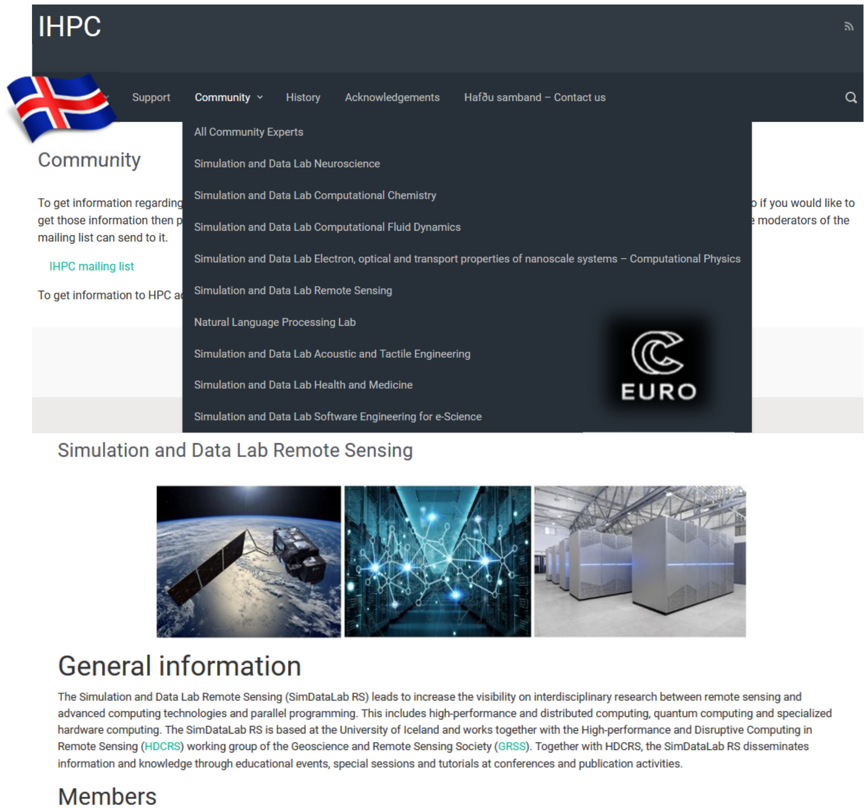
ROCCO SEDONA, PHD STUDENT
HIGH PRODUCTIVITY DATA PROCESSING RESEARCH GROUP

SIMULATION & DATA LAB REMOTE SENSING
JÜLICH SUPERCOMPUTING CENTRE
UNIVERSITY OF ICELAND

EuroCC NCC Iceland – Simulation & Data Lab Remote Sensing



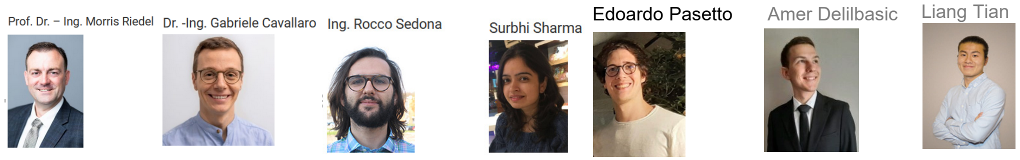
Icelandic HPC Community Web page



- Simulation & Data Lab Communities
  - Experts w.r.t. **HPC in domain-specific topics**
  - Co-organizing the HDCRS Summer school 29th May - 2nd June, 2022 [https://www.hdc-rs.com/summer-schools/2022-iceland/home]

Juelich Supercomputing Centre, Simlabs

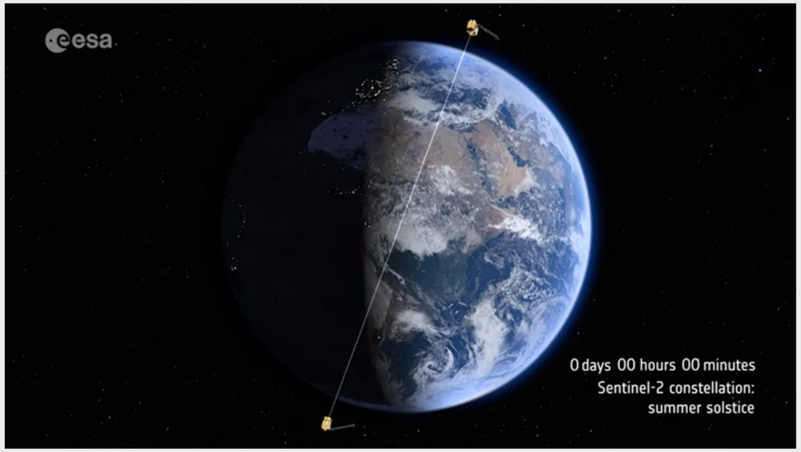
JÜLICH Forschungszentrum | JÜLICH SUPERCOMPUTING CENTRE

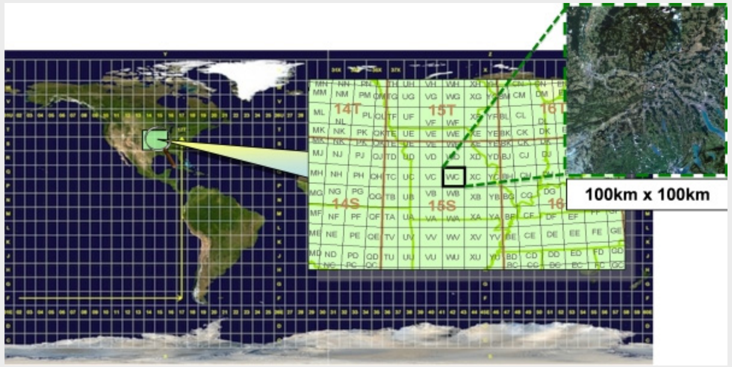


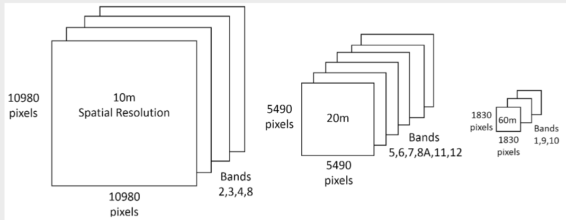
Icelandic HPC National Competence Center for HPC & AI

Sentinel-2 Multispectral Images

- Twin polar-orbiting satellites with temporal resolution of 5 days



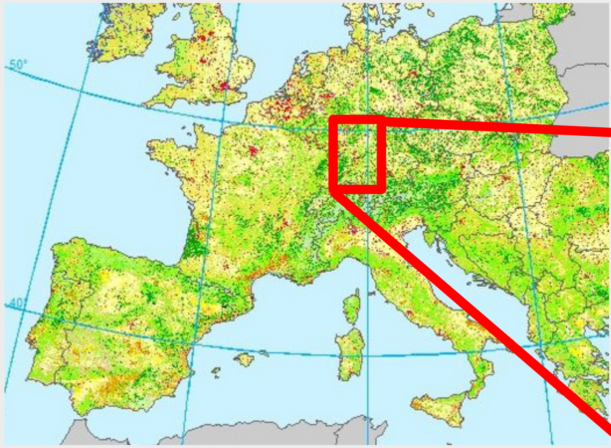




Data are free

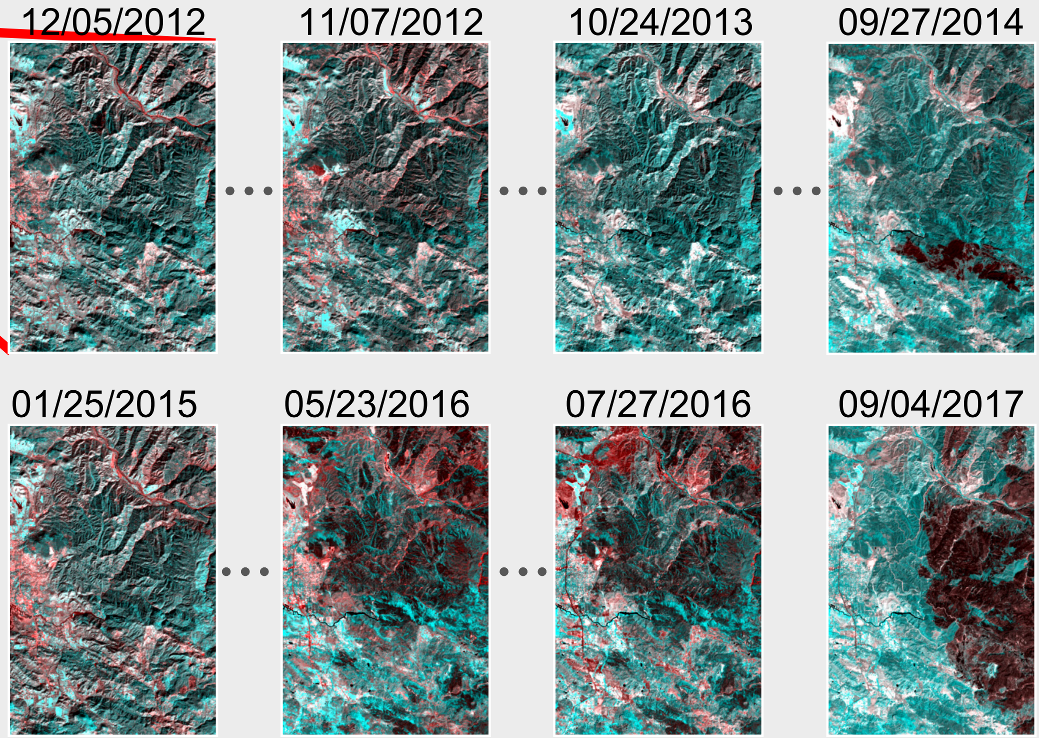
https://sentinels.copernicus.eu/

Processing Workflow for Updating Land Cover Maps



CORINE Land Cover 2018
https://land.copernicus.eu/pan-european/corine-land-cover

Landsat Time Series False Color Composition
(RGB: Red/NIR/NIR)



Problem: long time to create new thematic maps

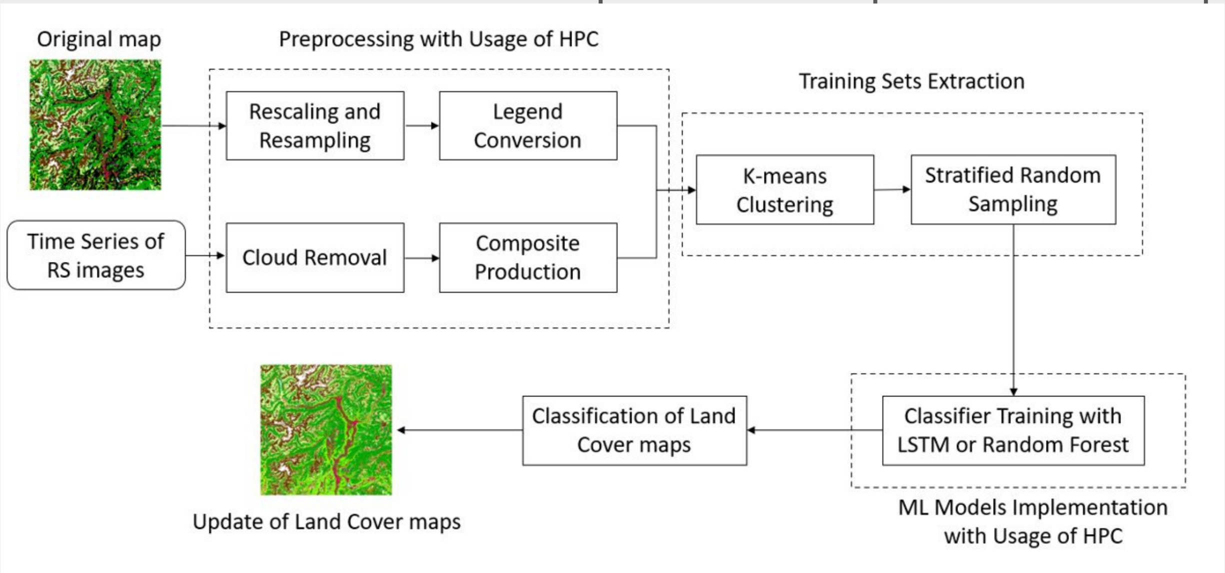
Goal: more frequent update of maps

How: using satellite imagery

Processing Workflow for Updating Land Cover Maps

Fully automated framework

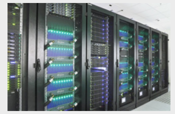
Can extract large number of labelled samples from open thematic products



C. Paris, L. Bruzzone, D. Fernández-Prieto, "A Novel Approach to the Unsupervised Update of Land-Cover Maps by Classification of Time Series of Multispectral Images," IEEE Transactions on Geoscience and Remote Sensing, Vol. 57, No. 7, pp. 4259-4277, 2019,

Rocco Sedona, Claudia Paris, Liang Tian, Morris Riedel, Gabriele Cavallaro, "An automatic approach for the production of a time series of consistent land-cover maps based on long-short term memory", IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2022 (accepted)

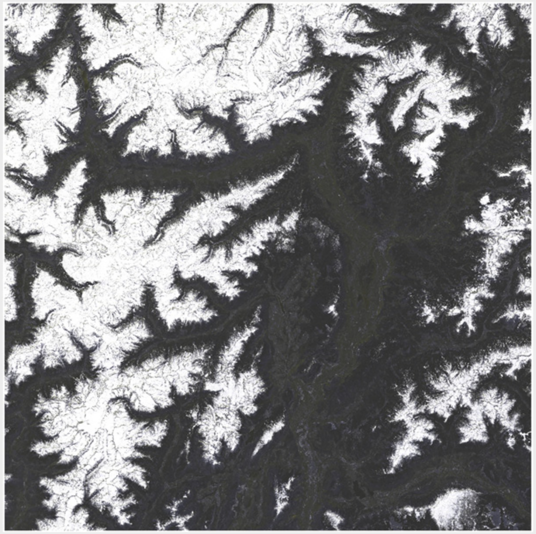
Initial Results



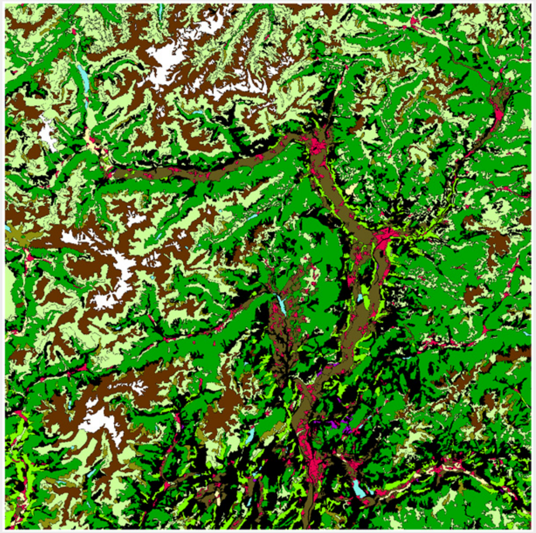
DEEP-EST
prototype system
https://www.deep-projects.eu/

Processing workflow ported to the DEEP-EST prototype system

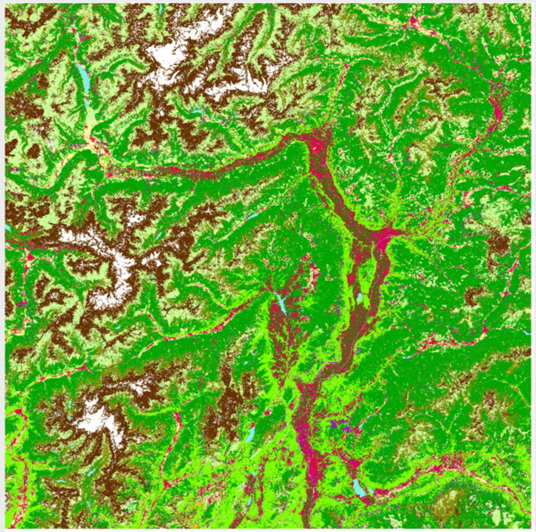
Classification map produced over a single Sentinel-2 tile



Sentinel-2 image



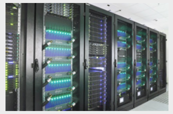
Old classification map
(CORINE 2018)



New Classification map
(SVM)

- Urban
- Grass
- Not irrigated crops
- Annual crops
- Mineral
- Rocks
- Sand
- Broadleaves
- Conifers
- Shrub
- Lake
- Snow

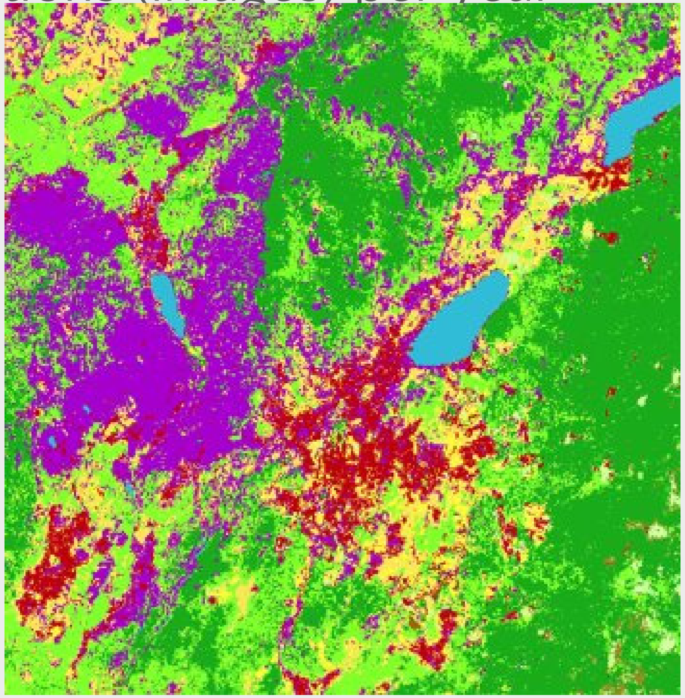
Initial Results



DEEP-EST
prototype system
https://www.deep-projects.eu/

LSTM using the 12 bands of pixels extracted from Sentinel-2 as features, sequences of 10/20 acquisitions (images) per year

Sentinel-2 image (2018)

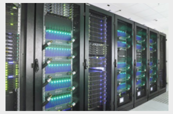


LSTM Classification map (2018)

- Urban
- Grass
- Not irrigated crops
- Annual crops
- Mineral
- Rocks
- Sand
- Broadleaves
- Conifers
- Shrub
- Lake
- Snow

RH0

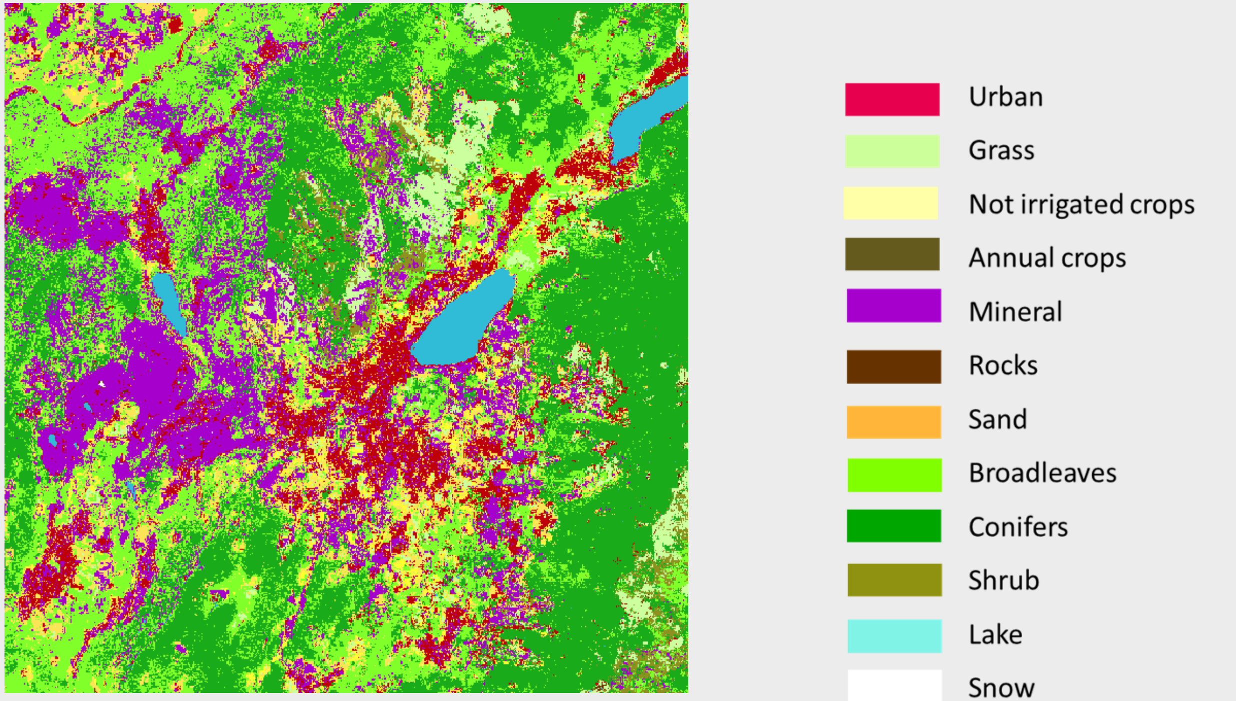
Initial Results



DEEP-EST
prototype system
https://www.deep-projects.eu/

Problem: real change or noise?

Sentinel-2 image (2019)



LSTM Classification map (2019)

'AN AUTOMATIC APPROACH FOR THE PRODUCTION OF A TIME SERIES OF CONSISTENT LAND-COVER MAPS BASED ON LONG-SHORT TERM MEMORY',
R. Sedona, C. Paris, L. Tian, M. Riedel, G. Cavallaro, IGARSS2022 (accepted)

Slide 8

RH0 mention LUCAS

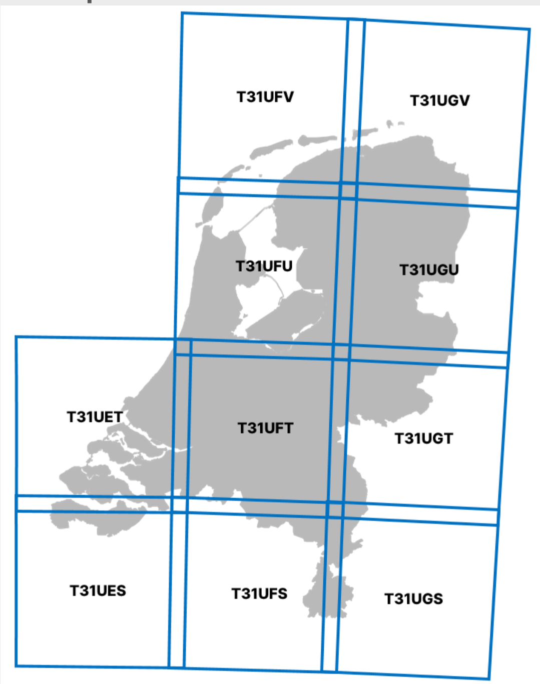
Rocco Sedona - HI; 2022-02-23T18:08:32.585

RH0

Current Work: Data Retrieval

RAISE Center of Excellence

Retrieval of Sentinel-2 time series for the Netherlands and related CORINE thematic maps



tile	num 2018	size 2018 [GB]	num 2019	size 2019 [GB]	num 2020	size 2020 [GB]
31UFT	17	16.16	14	15.33	19	20.91
31UGS	19	21.17	15	16.56	21	23.4
32ULC	20	22.08	18	19.55	20	21.16
31UGU	19	19.97	20	20.38	20	20.92
31UGV	20	18.68	17	15.5	27	21.7
31UFV	19	16.59	17	14.83	22	18.8
31UFU	18	16.65	15	15.24	21	21.31
31UFS	21	23.66	15	16.63	23	25.5
31UES	12	13.56	16	16.82	19	21.11
31UET	16	13.42	14	13.17	20	18.43
total	181	181.94	161	164.01	212	213.24

Slide 9

RH0 add total years

Rocco Sedona - HI; 2022-02-23T18:08:08.632

Current Work: Data Pre-processing

Aim: speed-up pre-processing and training set extraction (clustering)

Ideas:

- Download acquisitions and corresponding land cover maps
- Create composites and harmonize data
- Process multiple tiles in parallel

Next Steps: DL Model

Aim: Exploit time series with long sequences of data acquired over multiple years

Models: Recurrent models or transformers

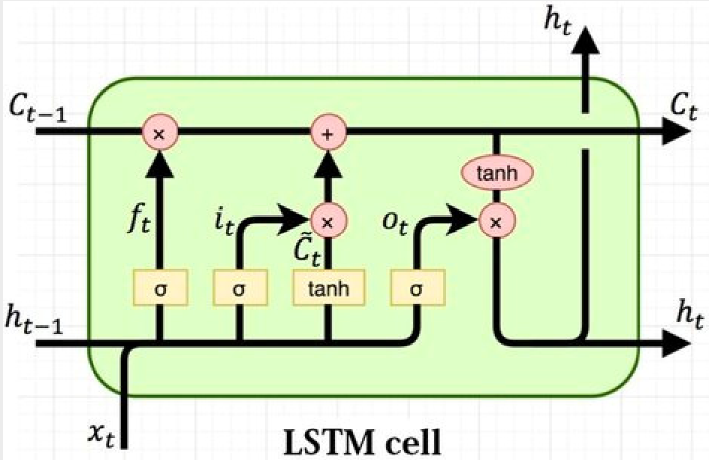
Ideas:

- Keep into consideration the dynamics of the outputs, as we want to get an output year by year
- Make prediction more robust and less noisy (retain cell state at prediction?)
- Use convolution to exploit spatial correlation? Ex. ConvLSTM
- Add regularization term to penalize change of label?
- Continual learning? Avoid catastrophic forgetting

LSTM

Exploit temporal dynamics of the signal

- Recurrency
- ability to remove or add information to the cell state
- Gates are a way to optionally let information through



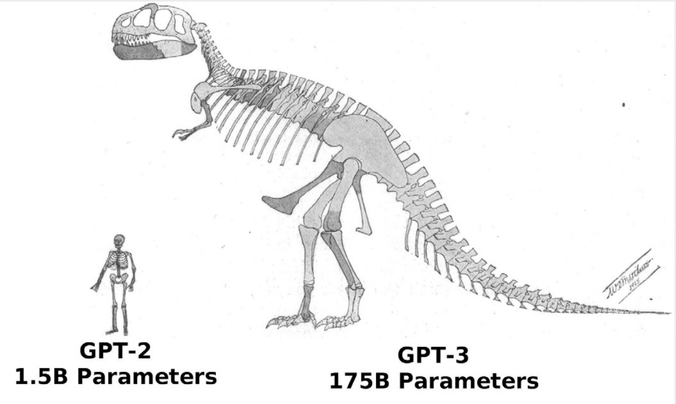
$$i_t = \sigma\left(x_t U^i + h_{t-1} W^i\right)$$
$$f_t = \sigma\left(x_t U^f + h_{t-1} W^f\right)$$
$$o_t = \sigma\left(x_t U^o + h_{t-1} W^o\right)$$
$$\tilde{C}_t = \tanh\left(x_t U^g + h_{t-1} W^g\right)$$
$$C_t = \sigma\left(f_t * C_{t-1} + i_t * \tilde{C}_t\right)$$
$$h_t = \tanh(C_t) * o_t$$

https://www.researchgate.net/publication/329362532_Designing_neural_network_based_decoders_for_surface_codes

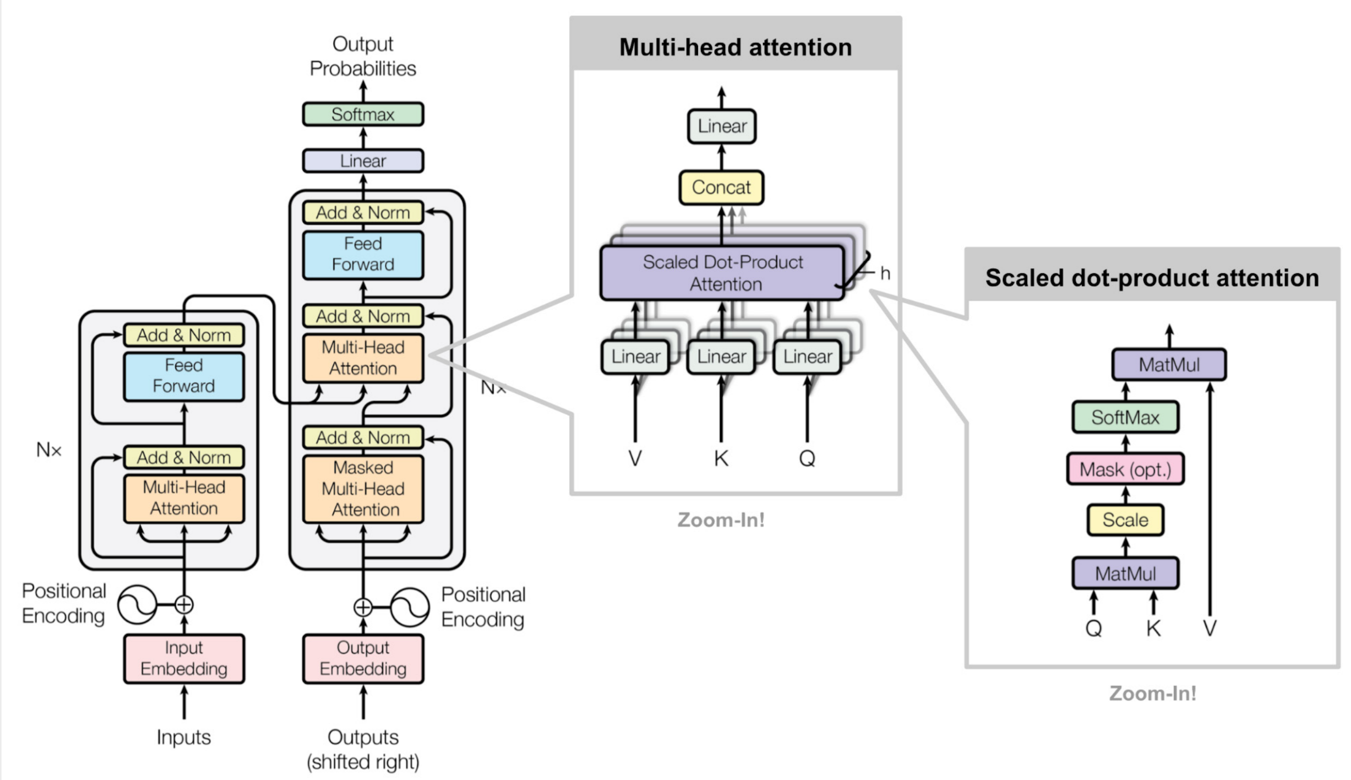
Transformer

Get rid of recurrency

- Highly scalable
- Basic block for new models with billions of parameters like GPT3



https://www.exxactcorp.com/blog/Deep-Learning/what-can-you-do-with-the-openai-gpt-3-language-model

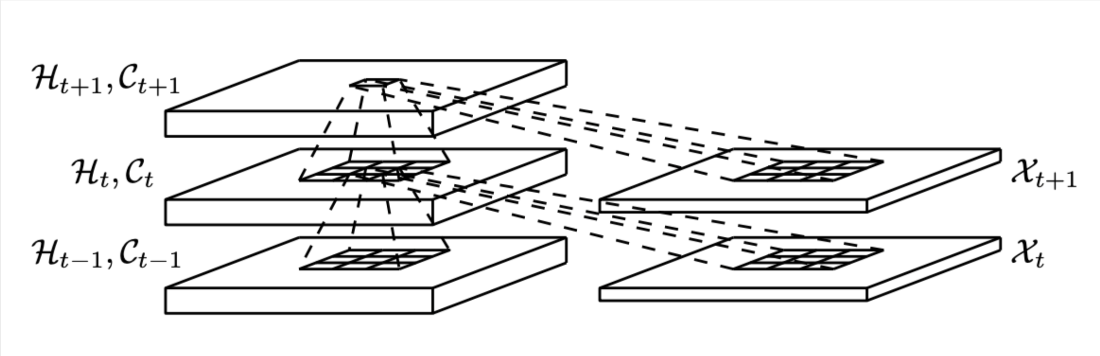


Vaswani, et al., 2017

ConvLSTM

Temporal and spatial correlation

- The ConvLSTM determines the future state of a certain cell in the grid by the inputs and past states of its local neighbors



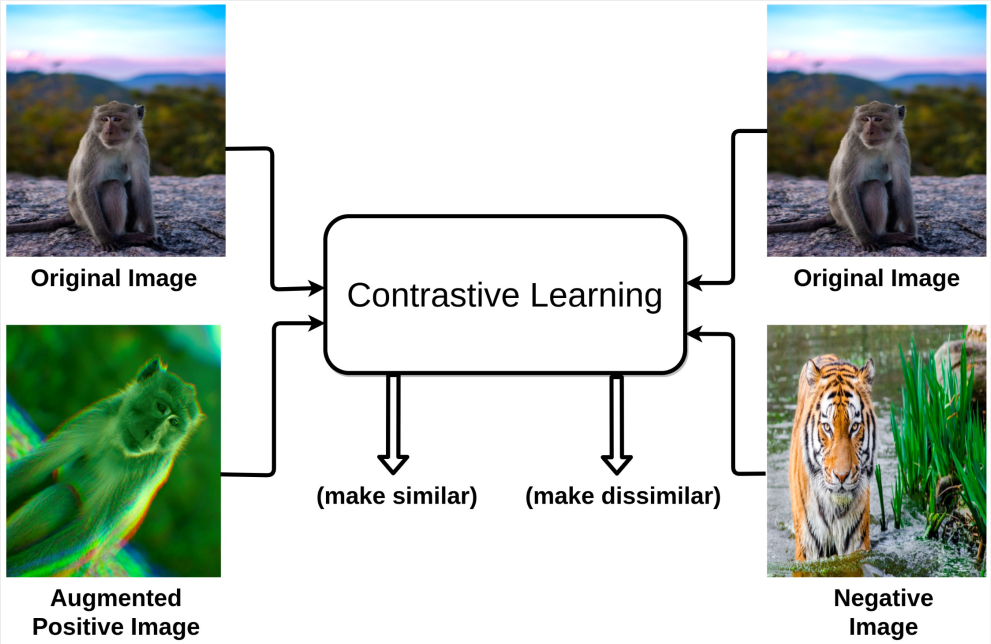
Shi et al. in [Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting]

Contrastive Learning

RAISE Center of Excellence

Exploit unlabeled data

- Train model to maximize agreement between different views of the same image
- What about RS data?

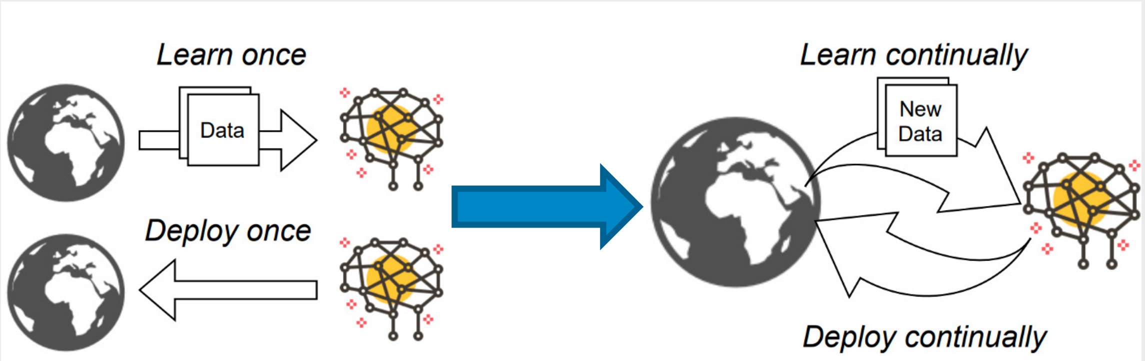


https://www.mdpi.com/2227-7080/9/1/2/htm

Continual Learning

New data is continuously acquired by Sentinel-2

- How to incorporate information while avoiding catastrophic forgetting?
- Loss with decreasing weight for older samples?
- Fine tuning of the previous model?

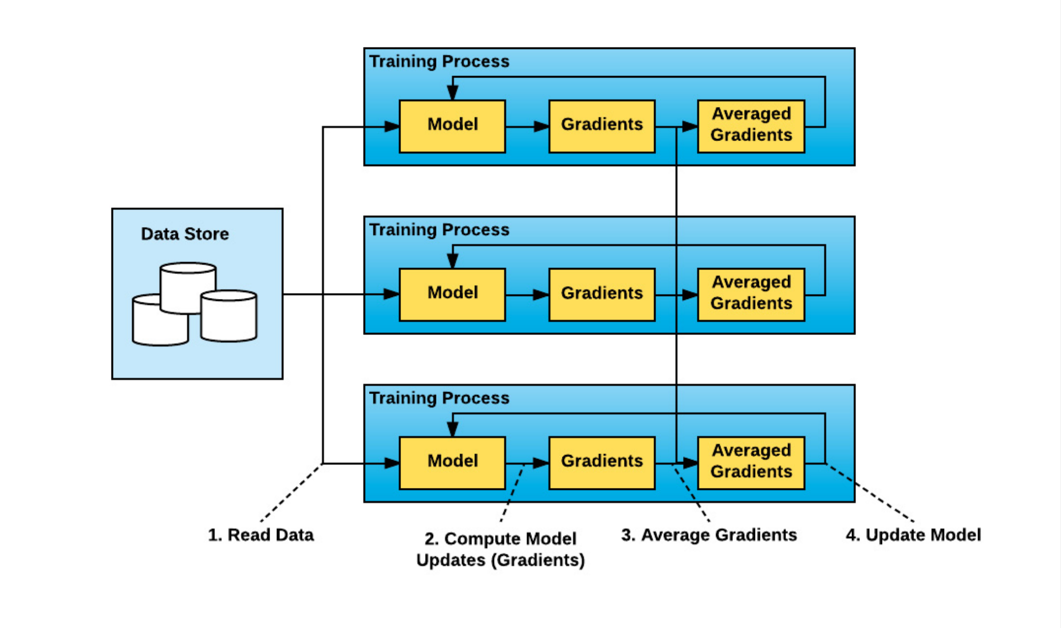


https://medium.com/continual-ai/towards-adaptive-ai-with-continual-learning-f493fd0d698

Distributed Deep Learning: What is it?

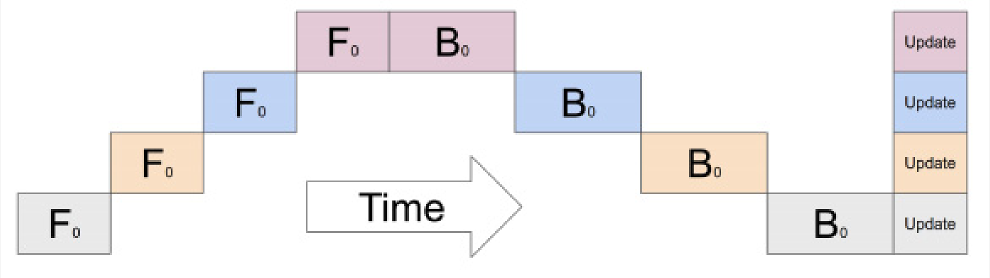
Problem: large models and dataset -> Solution: scale-up!

Data Parallelism



http://www.idris.fr/eng/ia/apprentissage-distribue-eng.html

Model Parallelism



https://pytorch.org/docs/stable/pipeline.html

Distributed Deep Learning: Frameworks

RAISE Center of Excellence

Many frameworks exists

- Horovod offers a very stable solution for Data Parallelism

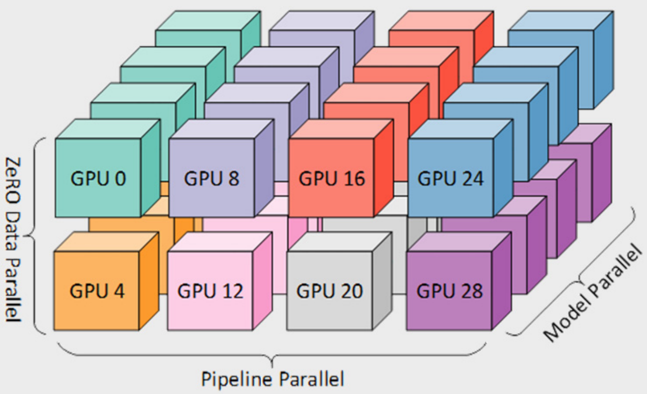
HOROVOD

https://github.com/horovod/horovod

HEAT Helmholtz Analytics Toolkit

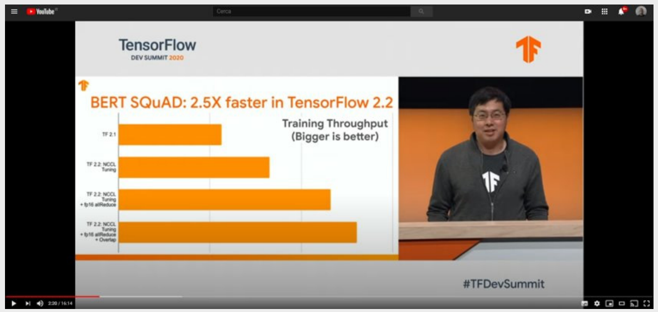
DLR JÜLICH FORSCHUNGSZENTRUM KIT Karlsruher Institut für Technologie

TARANTELLA



https://github.com/microsoft/DeepSpeed

Colossal-AI



https://www.youtube.com/watch?v=jKV53r9-H14&ab_channel=TensorFlow

Hyperparameter Tuning

RAISE Center of Excellence

- Many hyperparameters, such as learning rate, scheduler optimizer, etc.
- Large global batch size require extensive tuning
- Automatic Hyperparameter Tuning
- 'ACCELERATING HYPERPARAMETER TUNING OF A DEEP LEARNING MODEL FOR REMOTE SENSING IMAGE CLASSIFICATION', M. Aach, R. Sedona, A. Lintermann, G. Cavallaro, H. Neukirchen, M. Riedel, IGARSS2022 (accepted)

https://docs.ray.io/en/latest/tune/index.html

Recap & Conclusion

Still setting up the framework:

- Use composites?

Once data pre-processing is stable:

- Experiment with more DL models (exploit also spatial dimension)
- Transition towards a feeding of the data as it becomes available
- Virtual constellations? NASA Harmonized Landsat Sentinel

Big questions:

- How to validate at large scale?

drive. enable. innovate.

RAISE Center of Excellence

The CoE RAISE project has received funding from the European Union's Horizon 2020 – Research and Innovation Framework Programme H2020-INFRAEDI-2019-1 under grant agreement no. 951733

Follow us: