Séminaire DL - AMAP
20/10/2017

Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks [Ienco17]

Dino Ienco, Raffaele Gaetano, Claire Dupaquier and Pierre Maurel

LIRMM Associated

Dino Ienco,
Research Scientist,
IRSTEA – UMR TETIS - Montpellier,
dino.ienco@irstea.fr

[Ienco17] D. Ienco, R. Gaetano, C. Dupaquier, P. Maurel: Land Cover Classification via Multitemporal Spatial Data by Deep Recurrent Neural Networks. IEEE Geosci. Remote Sensing Lett. 14(10): 1685-1689 (2017)

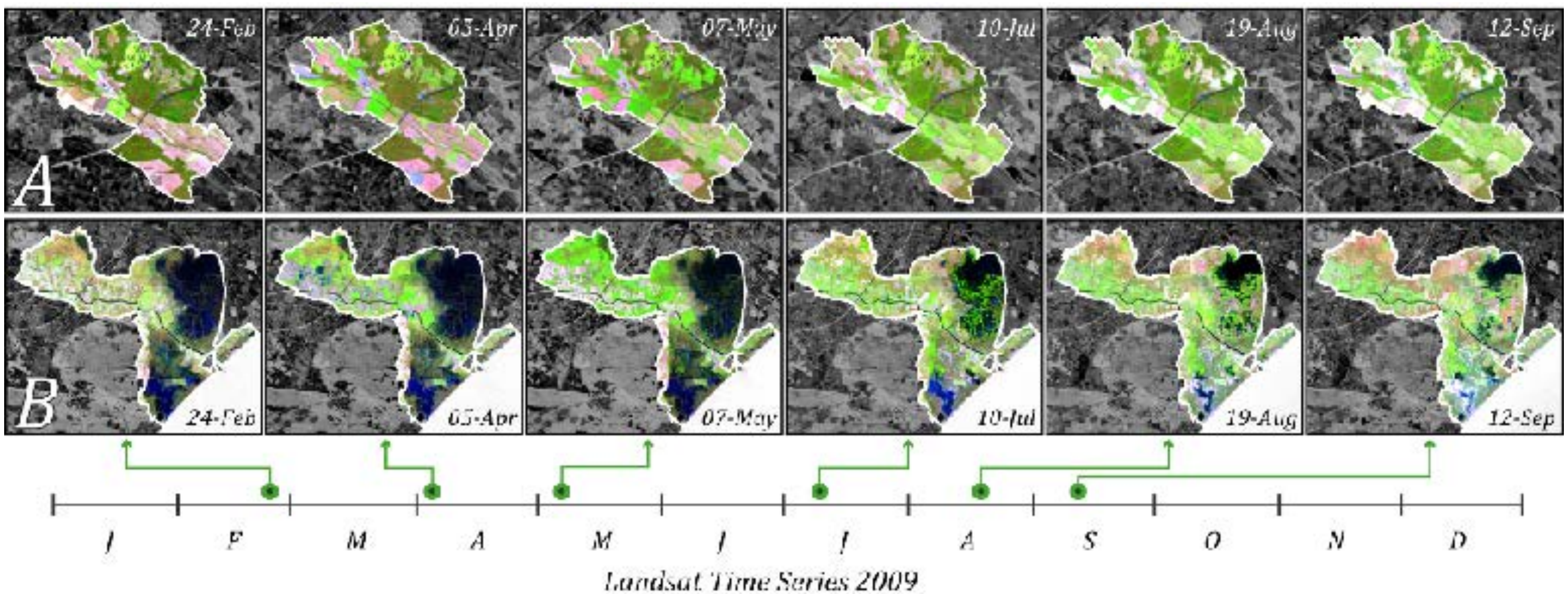
Outline

- Intro to Satellite Image Time Series (SITS) data
- Deep Learning Techniques
- Recurrent Neural Networks (RNNs) & SITS Data
- Case Studies
- Conclusions and On-going Works

Earth Observation Data (EOD)

EOD allows to collect Satellite Image Time Series (SITS)

The same geographical area is observed during the time



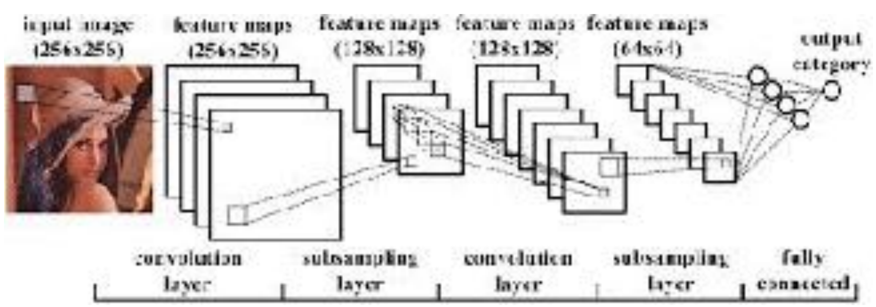
SITS data are useful to analyze spatio-temporal phenomena (trends and changes) over the time

Deep Learning Techniques

CNNs [Zhang16]:

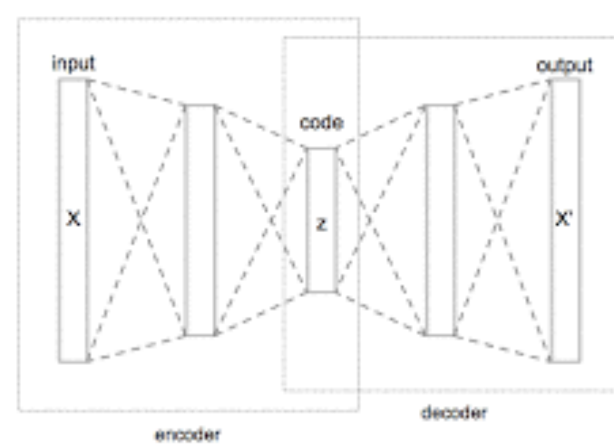
- Convolutional approaches (1-D, 2-D and 3-D)
- Very useful to manage and deal with Spatial Information
- Current best performances in Image Classification

Starting to be known in Remote Sensing



AutoEncoder Networks [LeCun15] :

- Interesting approaches to extract compressed data representation
- Current state-of-the-art for supervised image segmentation

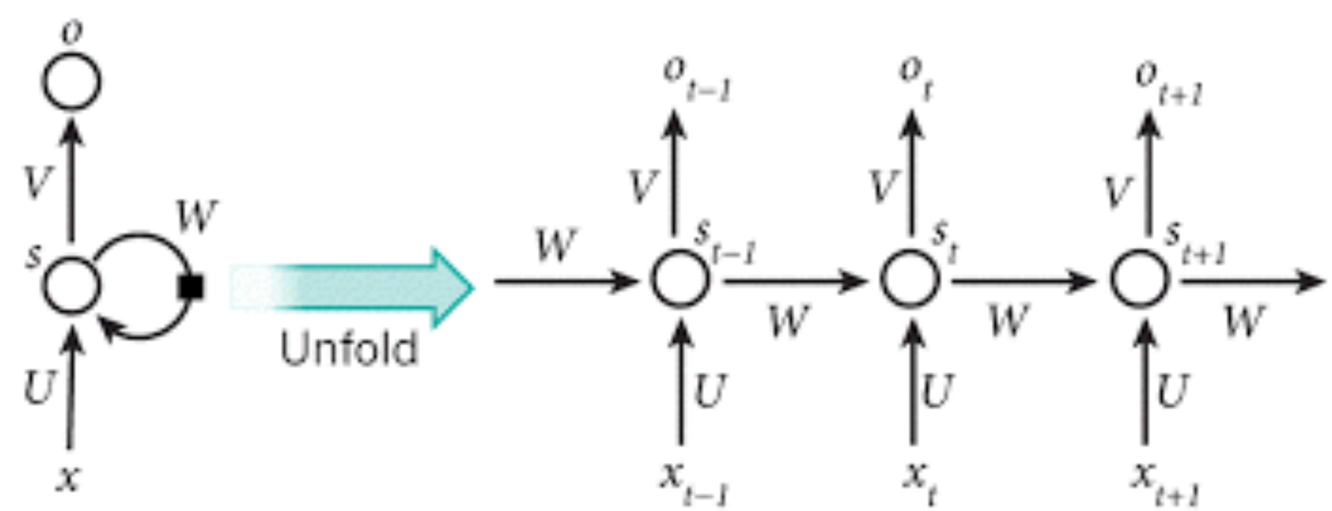


Not widely exploited in Remote Sensing

Recurrent Neural Networks [Greff16]:

- Useful tool to deal with temporal and sequence data
- Mainly explored in Speech Recognition and Natural Language Processing

Not yet Well Known in Remote Sensing

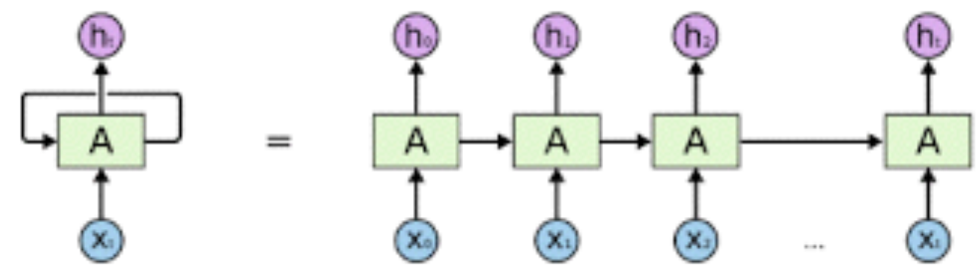


[**Greff16**] Klaus Greff, Rupesh Kumar Srivastava, Jan Koutník, Bas R. Steunebrink, Jürgen Schmidhuber: LSTM: A Search Space Odyssey. CoRR abs/1503.04069 (2015)
[**Zhang16**] L. Zhang, L. Zhang and B. Du, “Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art”. IEEE Geosc. and Rem. Sens. Magazine: 4(2): 22-40 (2016).
[**LeCun15**] Y. LeCun, Y. Bengio and G. Hinton. “Deep Learning” In Nature 52(8): 436-444 (2015)

Recurrent Neural Networks

Domain in which RNNs are successful:

Speech recognition, Natural language processing (i.e. Automatic Translation between Languages), Image completion



The principle of recurrent neural networks is to use the results produced at time t-1 to feed the network at time t

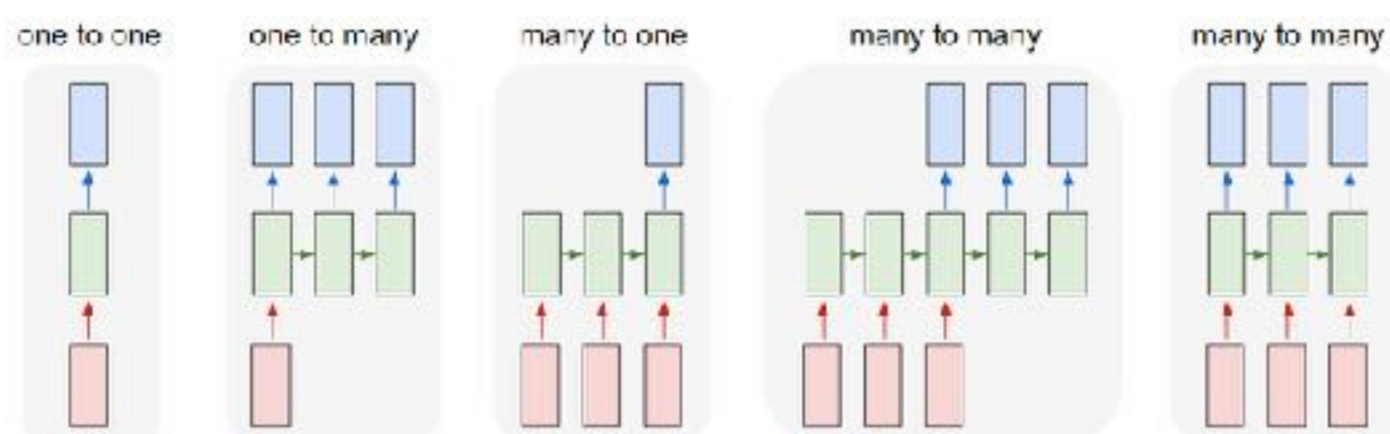


Image classification

Image captioning

Sentiment analysis

Machine translation

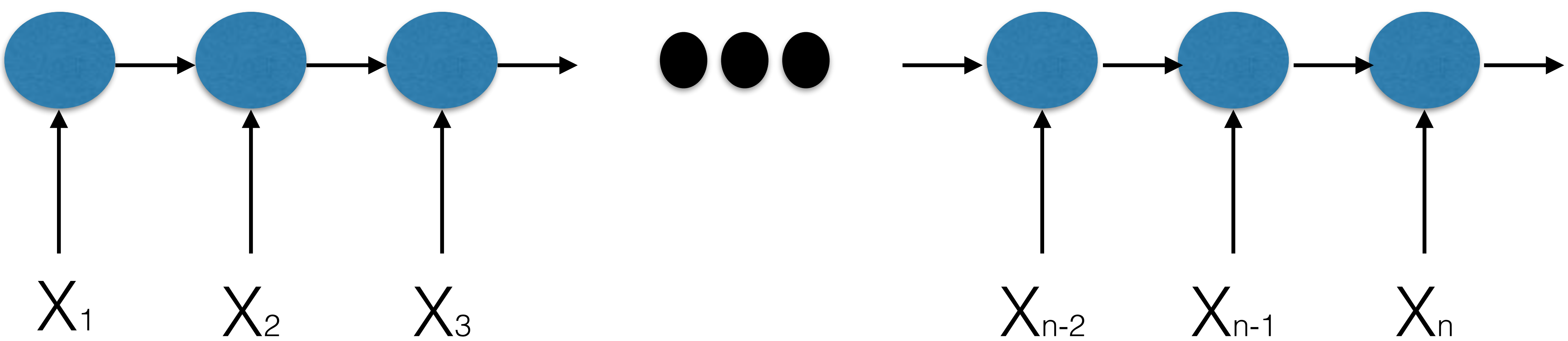
Synced sequence(video classification)

Recurrent Neural Networks

The most famous model is LSTM (Long-Short Term Memory) [Greff16]:

- It explicitly captures temporal correlations by recursion
- It is able to extract the important signal portions forgetting irrelevant information

$X = \langle X_1, X_2, \ldots, X_n \rangle$



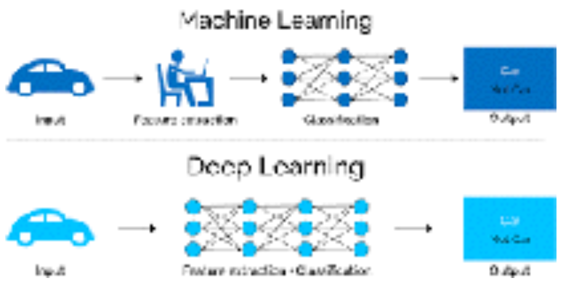
Each X_i is a multi-dimensional vector that contains some information

[Greff16] Klaus Greff, Rupesh Kumar Srivastava, Jan Koutník, Bas R. Steunebrink, Jürgen Schmidhuber: LSTM: A Search Space Odyssey. **CoRR abs/1503.04069** (2015)

Advantages w.r.t. Standard ML Classifiers

Standard ML Classifier (i.e. RF and SVM):

- Do **not model temporal correlations** (Features are managed independently)
- Only **original features** are considered

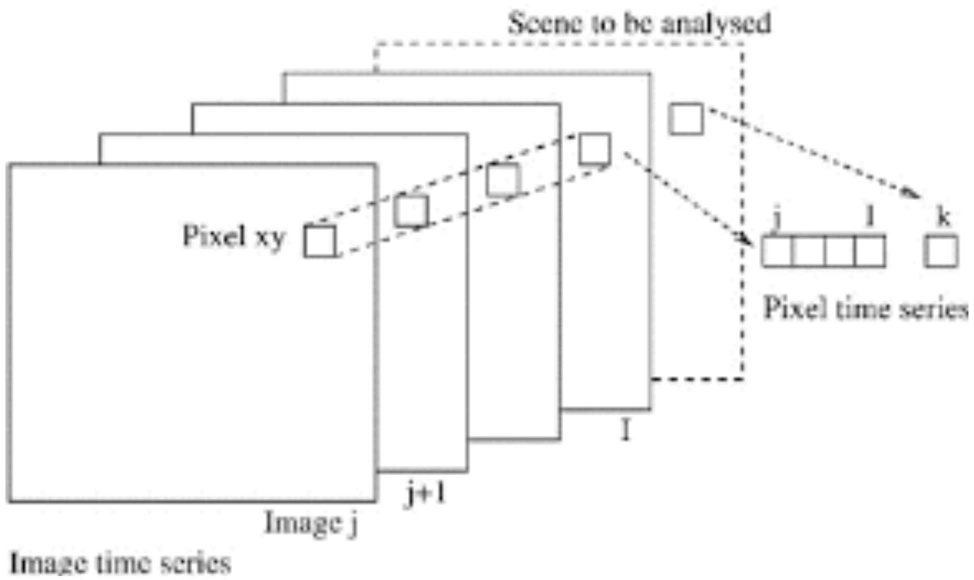


LSTMs (RNNs):

- Explicitly model temporal dependencies
- Learning at once **classifier** and **new data representation** [**Bengio13**] tailored for the discriminative task

Satellite Image Time Series

In the Remote Sensing field (SITS data), the data sequence is the multi-dimensional time series of radiometric information of a pixel along the different images.



[Bengio13] Y. Bengio, A. C. Courville, P. Vincent: Representation Learning: A Review and New Perspectives. **IEEE Trans. Pattern Anal. Mach. Intell**. 35(8): 1798-1828 (2013)

Case studies

Thau Dataset:

3 Pléiades VHSR images (2~m) acquired (July 2012 - March 2013). Object-Oriented Classification (Multi-temporal objects). Information: R,B,G, NIR, NDVI (avg+std). Around 15k objects.

ID	Land Cover Class	N. of Objects
(1)	Tree crops	600
(2)	Forests and woods	2 445
(3)	Water	556
(4)	Summer crops	81
(5)	Winter crops	677
(6)	Grasslands	3 882
(7)	Sclerophyll vegetation	2 457
(8)	Truck farming	227
(9)	Bare soils	299
(10)	Salt marshes	236
(11)	Vineyards	3 735

Reunion Dataset:

23 Landsat8 images acquired in 2014 (30~m), level 2A and preprocessed by MAACS (CNES). Pixel-Oriented Classification. Information: (7 surface reflectances + NDVI, NDWI and brightness index - BI). Around 37K pixels.

ID	Land Cover Class	N. of Pixels
(1)	Urban areas	10 000
(2)	Other built-up surfaces	1 500
(3)	Forests	10 000
(4)	Sparse Vegetation	5 095
(5)	Rocks and bare soil	3 729
(6)	Grassland	1 744
(7)	Sugarcane crops	2 832
(8)	Other crops	1 500
(9)	Water	1 500

[Ienco17] D. Ienco, R. Gaetano, P. Maurel and C. Dupaquier: "Land Cover Classification via Multi- temporal Spatial Data by Recurrent Neural Networks". **IEEE Geoscience and Remote Sensing Letters** - Minor Revision - (2017)

Experimental Settings

Methods:

- *Our Proposal - LSTM (3 stacked neurons) + SoftMax Layer*
  - *Dim. Hidden vector 64 (Thau Dataset) and 512 (Réunion Island)*
- Random Forest (400 trees)
- SVM (RBF - Complexity = 100 and Sigma = 0.01)

Use Deep Model (LSTM) as feature extractor*:*

- *RF(LSTM)*
- *SVM(LSTM)*

Evaluation:

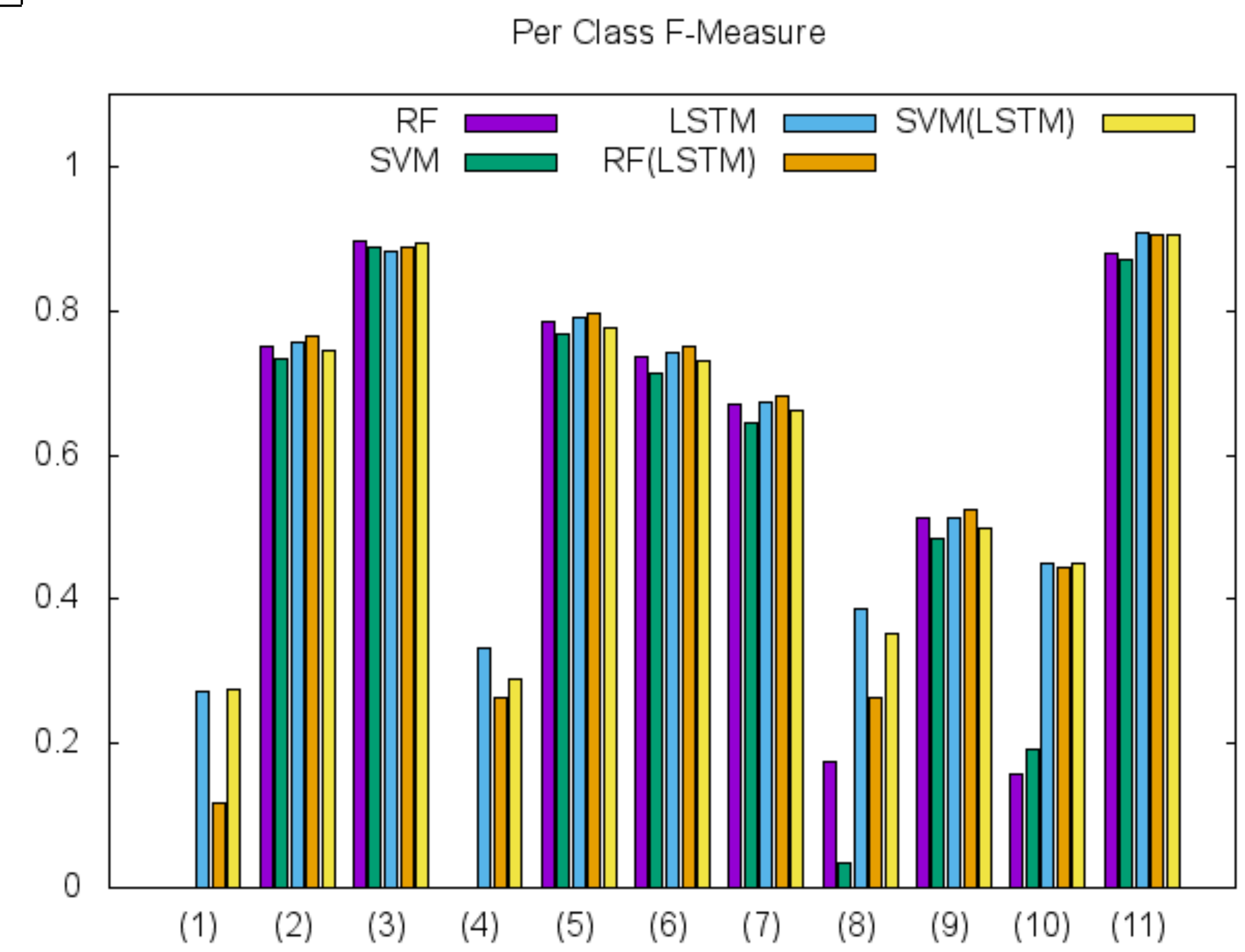
- 5 Fold Cross Validation (Training/Test)
- Different evaluation measures:
  - Accuracy
  - F-Measure (F-Measure x Class)
  - Kappa

Results on Thau Dataset

Method	Accuracy	F-Measure	Kappa
RF	74.28% ± 0.75%	71.67% ± 0.74%	0.68 ± 0
SVM	72.43% ± 0.74%	69.83% ± 0.63%	0.65 ± 0
LSTM	75.18% ± 0.61%	**74.57**% ± 0.59%	0.69 ± 0
RF(LSTM)	**75.80**% ± 0.36%	74.24% ± 0.40%	**0.70** ± 0
SVM(LSTM)	74.10% ± 0.40%	73.72% ± 0.42%	0.68 ± 0

TABLE III
5-FOLD CROSS VALIDATION ON THE *THAU* DATASET

ID	Land Cover Class	N. of Objects
(1)	Tree crops	600
(2)	Forests and woods	2 445
(3)	Water	556
(4)	Summer crops	81
(5)	Winter crops	677
(6)	Grasslands	3 882
(7)	Sclerophyll vegetation	2 457
(8)	Truck farming	227
(9)	Bare soils	299
(10)	Salt marshes	236
(11)	Vineyards	3 735

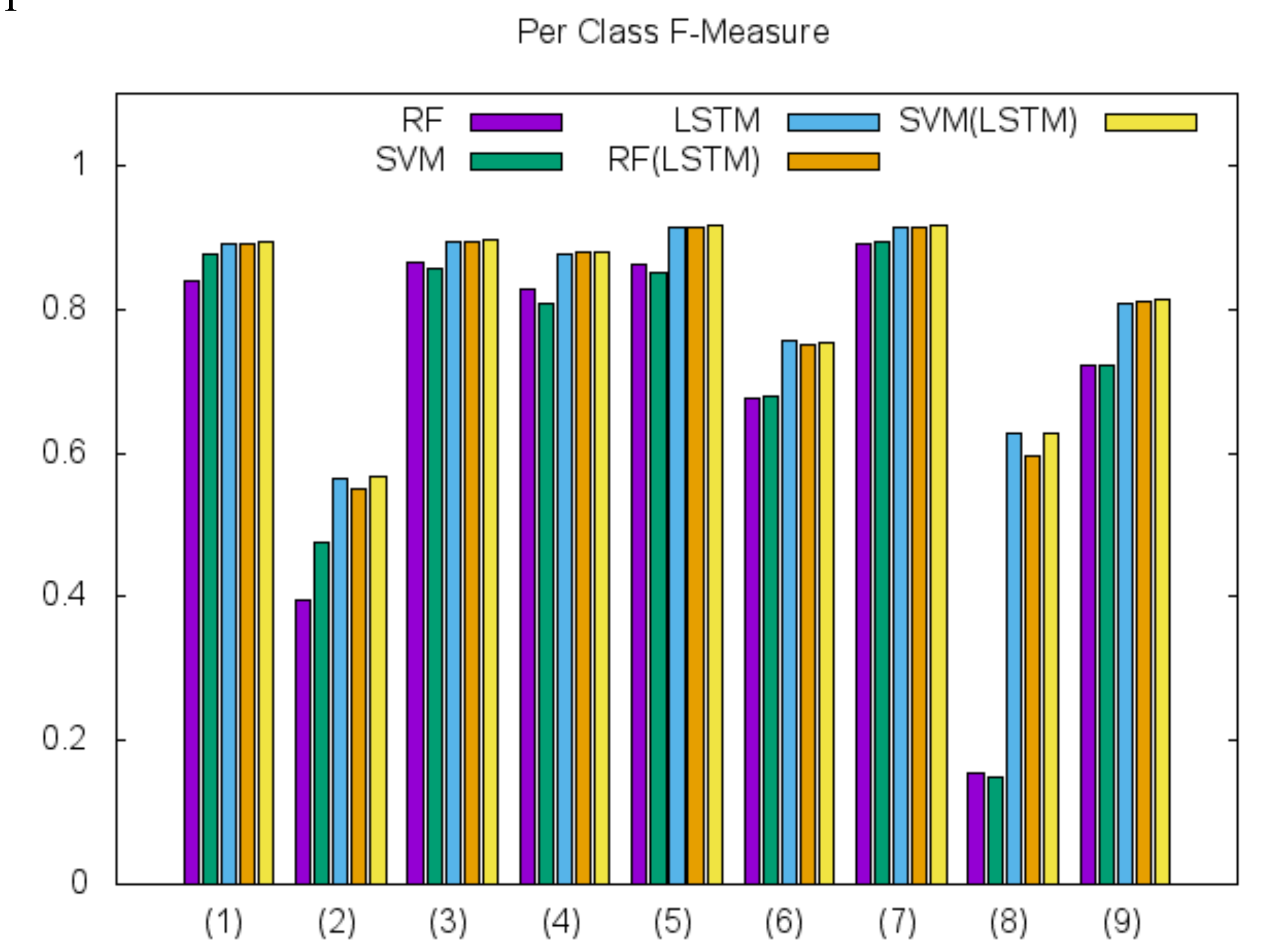


Results on Réunion Dataset

Method	Accuracy	F-Measure	Kappa
RF	81.19% ± 0.72%	79.40% ± 0.75	0.77 ± 0
SVM	81.59% ± 0.47%	80.01% ± 0.43%	0.77 ± 0
LSTM	86.23% ± 0.62%	86.11% ± 0.58%	0.83 ± 0
RF(LSTM)	86.15% ± 0.50%	85.97% ± 0.48%	0.83 ± 0
SVM(LSTM)	**86.41**% ± 0.60%	**86.23**% ± 0.56%	**0.83** ± 0

TABLE IV
5-FOLD CROSS VALIDATION ON THE REUNION ISLAND DATASET

ID	**Land Cover Class**	N. of Pixels
(1)	Urban areas	10 000
(2)	Other built-up surfaces	1 500
(3)	Forests	10 000
(4)	Sparse Vegetation	5 095
(5)	Rocks and bare soil	3 729
(6)	Grassland	1 744
(7)	Sugarcane crops	2 832
(8)	Other crops	1 500
(9)	Water	1 500



Discussion and Preliminary Findings

- LSTM-based classification model **outperforms** SVM and RF in the case studies
- The use of a model that **explicitly consider temporal behaviour** seems important for SITS data classification
- The major gain is supplied w.r.t. **lower represented (and highly mixed) classes**
- **Features extracted** by LSTM seem beneficial for Standard Classifier, RF(LSTM) (resp. SVM(LSTM)) outperforms their counterpart trained on original data
- The two case studies underline that these techniques can be applied to SITS data considering both **pixel-based** and **object-based representation**

Technical and Implementation details on Experiments

- The Deep RNNs classifier are implemented in **KERAS**
- We employ **Theano** as Backend
- Experiments are performed on Server with Intel(R) Xeon(R) CPU E5-2667 v4@3.20Ghz with 256Gb of RAM and **GPU TITAN X Pascal 2**

Bigger Dataset:

- 37900 pixels with 23 * 10 dimensions
- For each fold (5-Folds CV), we took (on average) **204 minutes** to learn a model
- We train the model for **200 epochs** with batch size equals to 20

Smaller Dataset:

- 15196 objects with 3 * 10 dimensions
- For each foot (5-Folds CV), we took (on average) **6 minutes** to learn a model
- We train the model for **200 epochs** with batch size equals to 20

For all the models we have used the **RMSprop** optimisation procedure (variant of SGD)

Conclusion & On-going Works

Considerations:

- The RNNs approaches are **still under exploited** in the Remote Sensing Field, really few works exist as of now (Bitemporal Change Detection [**Lyu16**], Hyperspectral Classification [**Mou17**]).

- **RS data is different** w.r.t. standard Computer Vision data (RGB images) on which DL methods were developed:
  - Multi-band informations
  - Different Resolutions, Different Sensors
  - SITS Data

- **New Applications** and **New Trends** require adapted DL techniques/architectures (i.e. LSTM literature mainly comes from NLP domain)

[Lyu16] Haobo Lyu, Hui Lu, Lichao Mou:
Learning a Transferable Change Rule from a Recurrent Neural Network for Land Cover Change Detection. **Remote Sensing** 8(6): 506 (2016)
[Mou17] L. Mou, P. Ghamisi, X. X. Zhu, "Deep Recurrent Neural Networks for Hyperspectral Image Classification", **IEEE Transactions on Geoscience and Remote Sensing** X(99): 1 - 17 (2017)

On-going Works

On-going works:

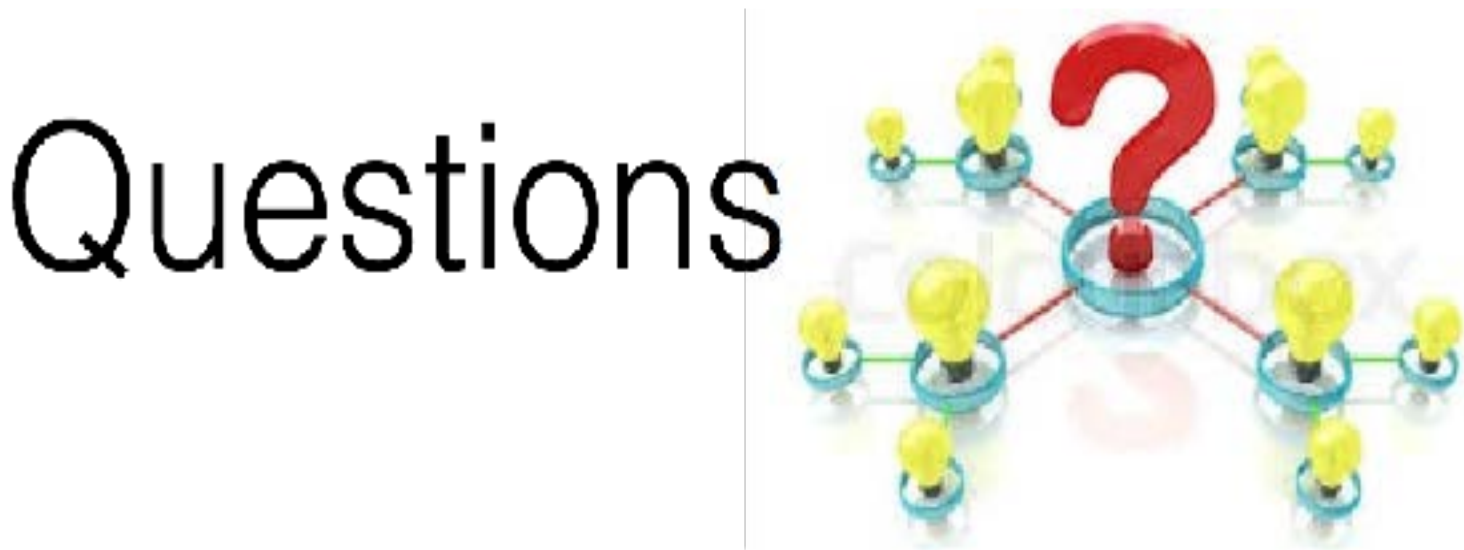
- Considering RNNs to work on **Radar Time Series** (Sentinel 1):
  - We apply **LSTM** and **GRU** in order to produce winter vegetation quality cover [Minh17].
  - Results have shown that LSTM/GRU seems **work better than RF and SVM.**
  - Use LSTM/GRU features still improve quality of the standard classifier.

- On-going internship on (high resolution - S2) **SITS and VHSR** (Very High Spatial Resolution - Pleiades/SPOT6-7) **fusion** for Land Cover mapping. Idea coupling RNN and CNNs to leverage both temporal and spatial info.

- Considering RNNs models (DL architectures) to **fusion/combine Optical/Radar** Time Series Data.

[Minh17] D. Ho Tong Minh, D. Ienco, R. Gaetano, N. Lalande, E. Ndikumana, F. Osman, P. Maurel:
Deep Recurrent Neural Networks for mapping winter vegetation quality coverage via multi-temporal SAR Sentinel-1. CoRR abs/1708.03694 (2017)

Thank you
for your Attention



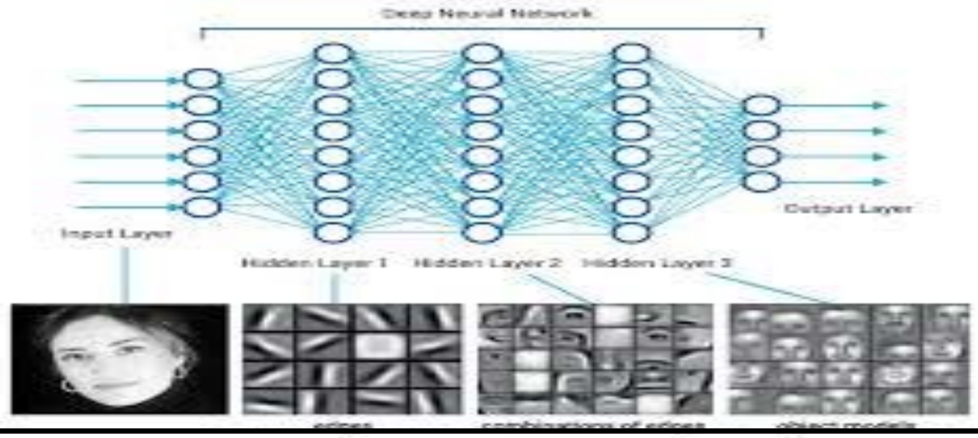
Emerging Trends in EOD

Nowadays, Satellite imagery analytic uses Machine Learning Techniques to:

- Deal with huge amount of data
- Automatically build predictive methods
- Group together similar areas
- Detect Objects of Interest

Recent Trends ‘Deep Learning Methods’ [LeCun15] :

- Inspired by human brain
- Layers architecture
- Applications in different domains:
    + Speech Recognition
    + Image Recognition
    + Natural Language Processing



[LeCun15] Y. LeCun, Y. Bengio and G. Hinton. “Deep Learning” In Nature 52(8): 436-444 (2015).

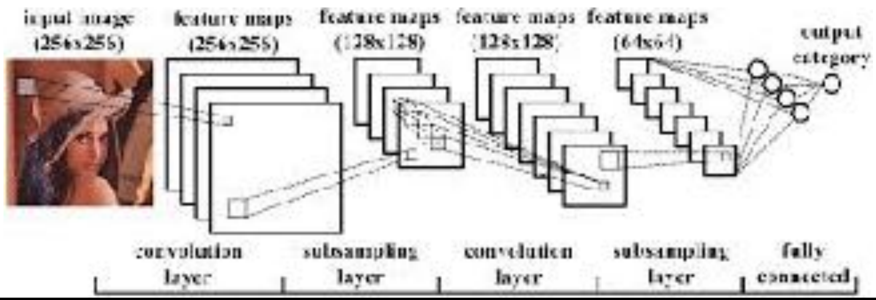
Deep Learning Techniques

State of the art performances in Image and Signal classification

Different Family of approaches exist (Feed-Forward and Recurrent) [LeCun15]

Convolutional Neural Networks (CNNs) [Zhang16]:
- Convolutional approaches (1-D, 2-D and 3-D)
- Very useful to manage and deal with Spatial Information
- Current best performances in Image Classification

Starting to be known in Remote Sensing



[Zhang16] L. Zhang, L. Zhang and B. Du, "Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art". **IEEE Geoscience and Remote Sensing Magazine**: 4(2): 22-40 (2016).
[LeCun15] Y. LeCun, Y. Bengio and G. Hinton. "Deep Learning" In Nature 52(8): 436-444 (2015).