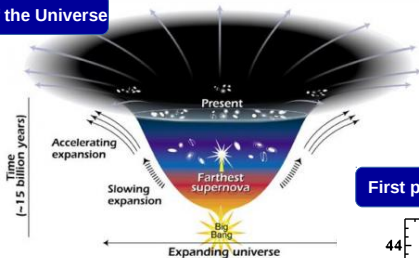# Deep Learning in LSST

Johanna Pasquet
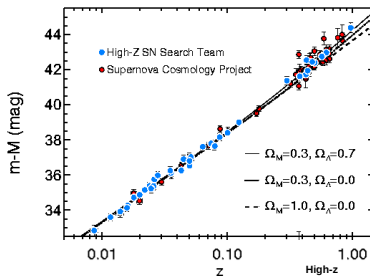
Nature of the dark energy
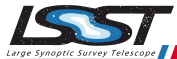
History of the Universe

First proof with supernovae Ia

- Dark energy causes the universal expansion to accelerate

- Recent observations of supernovae have produced a value for an acceleration that implies a universe that is about 70 % dark energy

The Large Synoptic Survey Telescope

- LSST will produce the deepest, widest, image of the Universe :
  - 37 billion stars and galaxies
  - 10 year survey of the sky
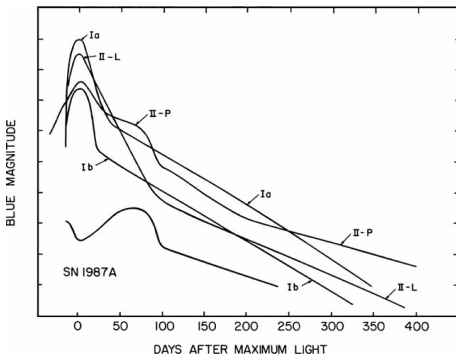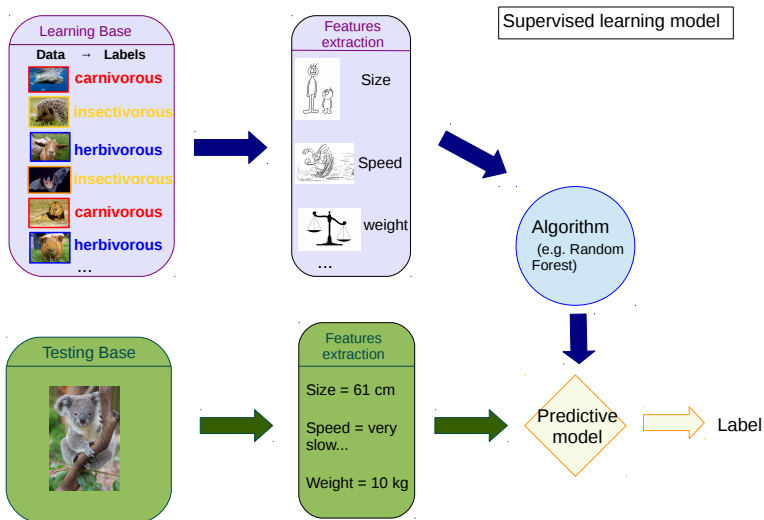 - 15 Terabytes of data ...every night !

Issues :

- LSST will discovery hundreds of thousands of type Ia supernovae

- Be able to automatically identify Sne Ia among all the supernovae with only the photometric information
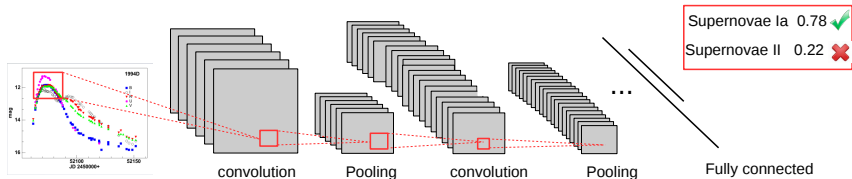
Goals

- Identify automatically the Supernovae Ia from the light curve (variation of the brightness of an object along time)
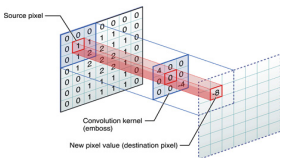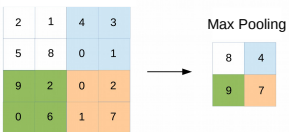
Supervised learning
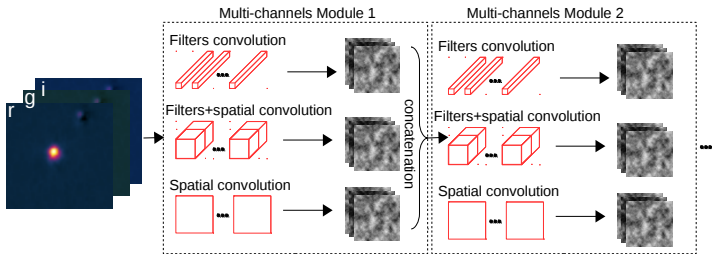
Convolutional Neural Network

| Supernovae Ia 0.78 ✔ |
| Supernovae II   0.22 ✖ |

convolution   Pooling   convolution   Pooling   Fully connected

Convolution

Source pixel

Convolution kernel
(emboss)

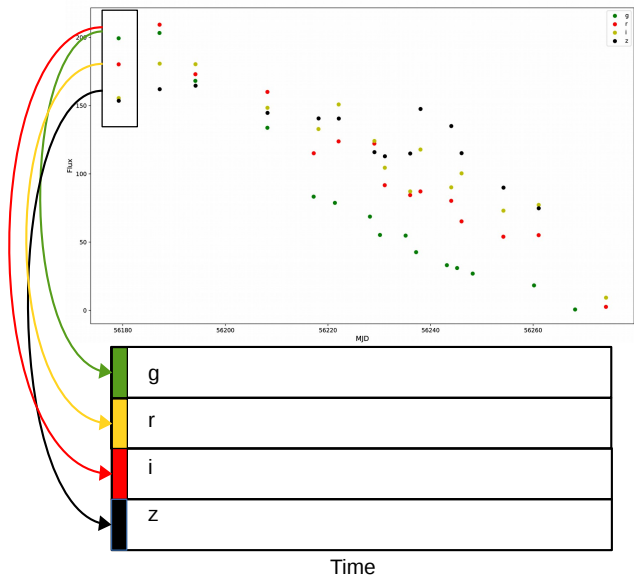New pixel value (destination pixel)

Pooling
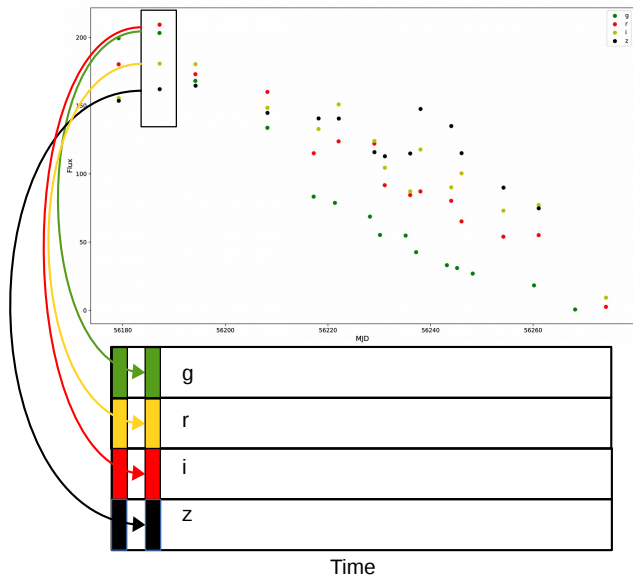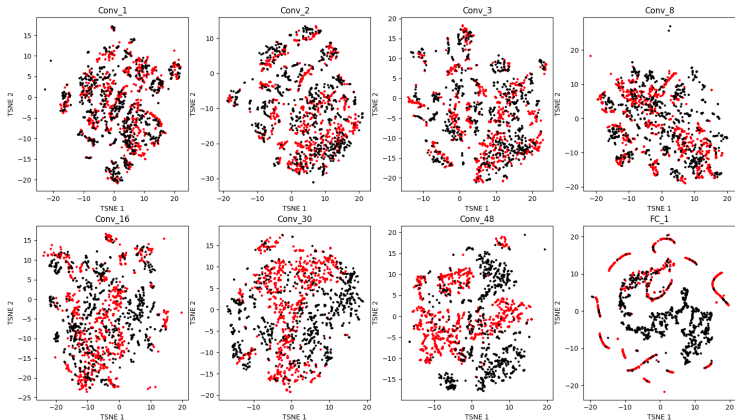
Max Pooling

Input data : images

Input data : light curves images

Time

Input data : light curves images

Time

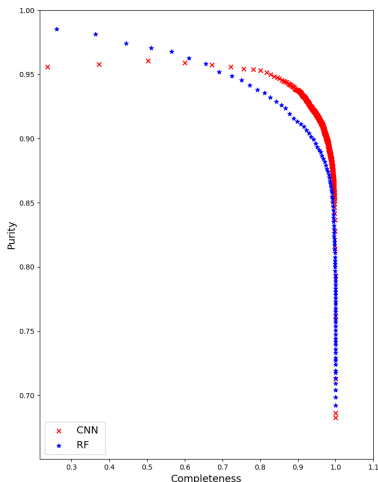# Separation of classes with a t-SNE

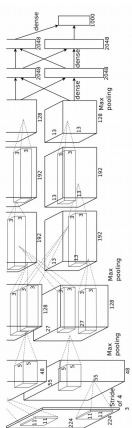# First result

- Bi-metric to evaluate the algorithm :
  - purity $= \frac{TP}{TP+FP}$
  - completeness $= \frac{TP}{TP+FN}$

- The **CNN** performance are compared to an optimized **random forest (RF)** performance with a feature extraction step
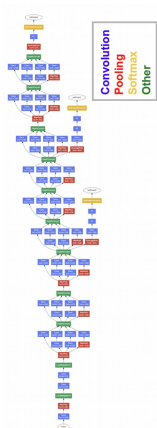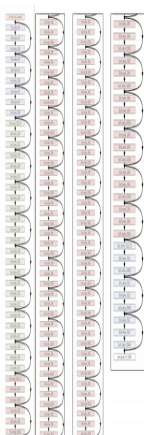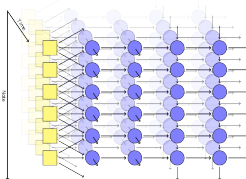
Going deeper

Alexnet
~60 millions
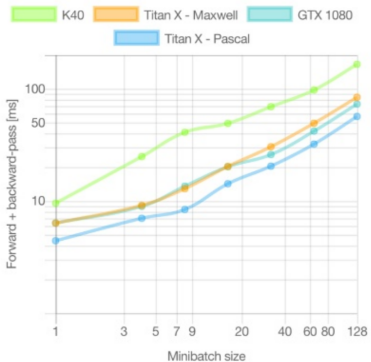
Googlenet
~5 millions

ResNet
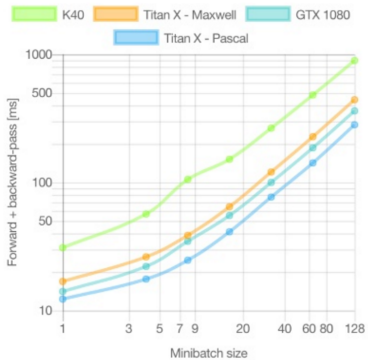~3,8 billions

LTSM / RNN

Nvidia GeForce GTX Titan X

- cores : 3584
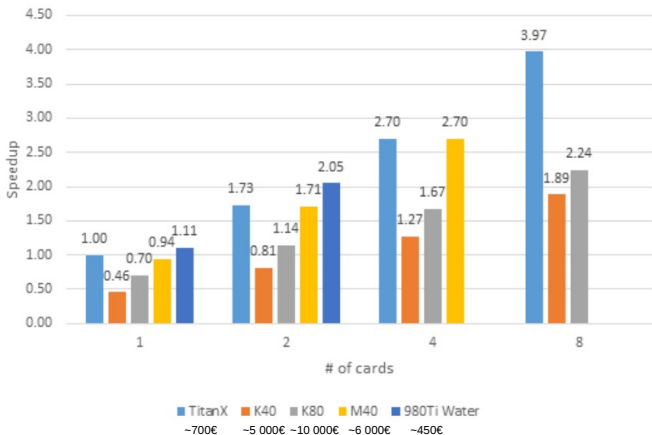- Frequency : 1531 MHz
- VRAM : 12 GB

Benchmark

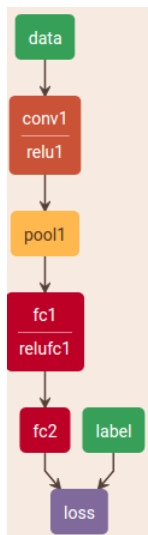# Benchmark

Caffe Benchmark Test Results
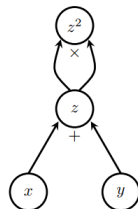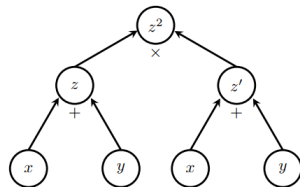
Deep Learning Frameworks

Caffe

- Caffe is a deep learning framework made with expression speed, and modularity in mind. Caffe is developed by the Berkeley Vision and Learning Center as well as community contributors and is popular for computer vision.
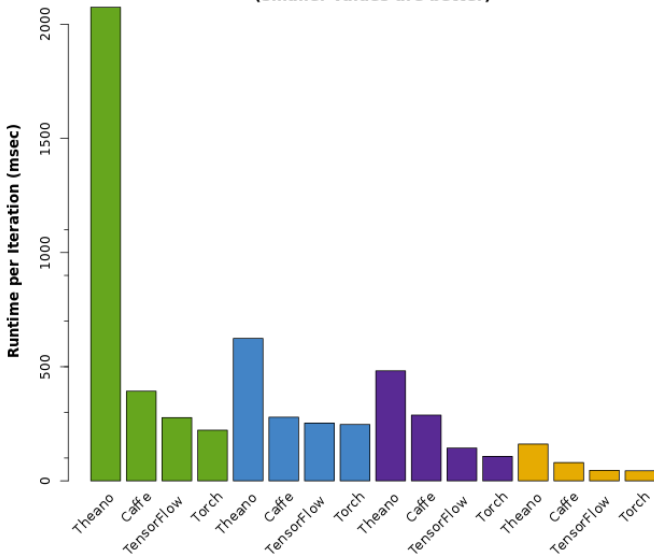
TensorFlow

- TensorFlow is a software library for numerical computation using data flow graphs, developed by Google's Machine Intelligence research organization.
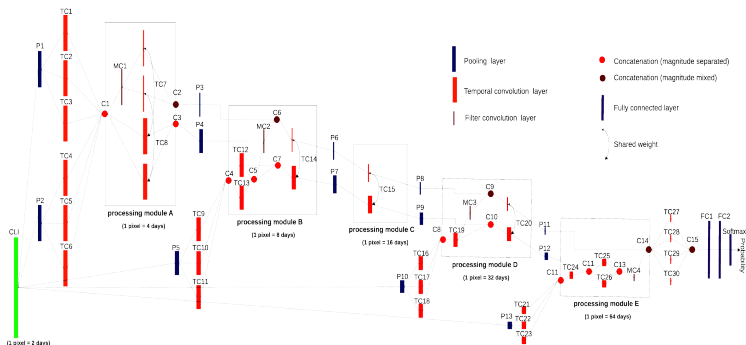
Frameworks benchmarks

Runtime per Training Iteration on the Tesla P100
(smaller values are better)

Conclusion

Titan X (maxwell)   : 3 days
GTX 1080 (Pascal)   : <2,5 days
CPU i7 6700 : 8x3.4GHz : >27 days

(estimation on 2 epochs)

- A problem ?
- The solution :