Multiple face recognition in images

Simone Caldarella

IEEE Student Branch Brescia
University of Brescia

Presentation on Convolutional Neural Networks, 2018

Outline

Outline

Motivation

- Create a ground for recognition problems
- Show the power of retraining

Outline

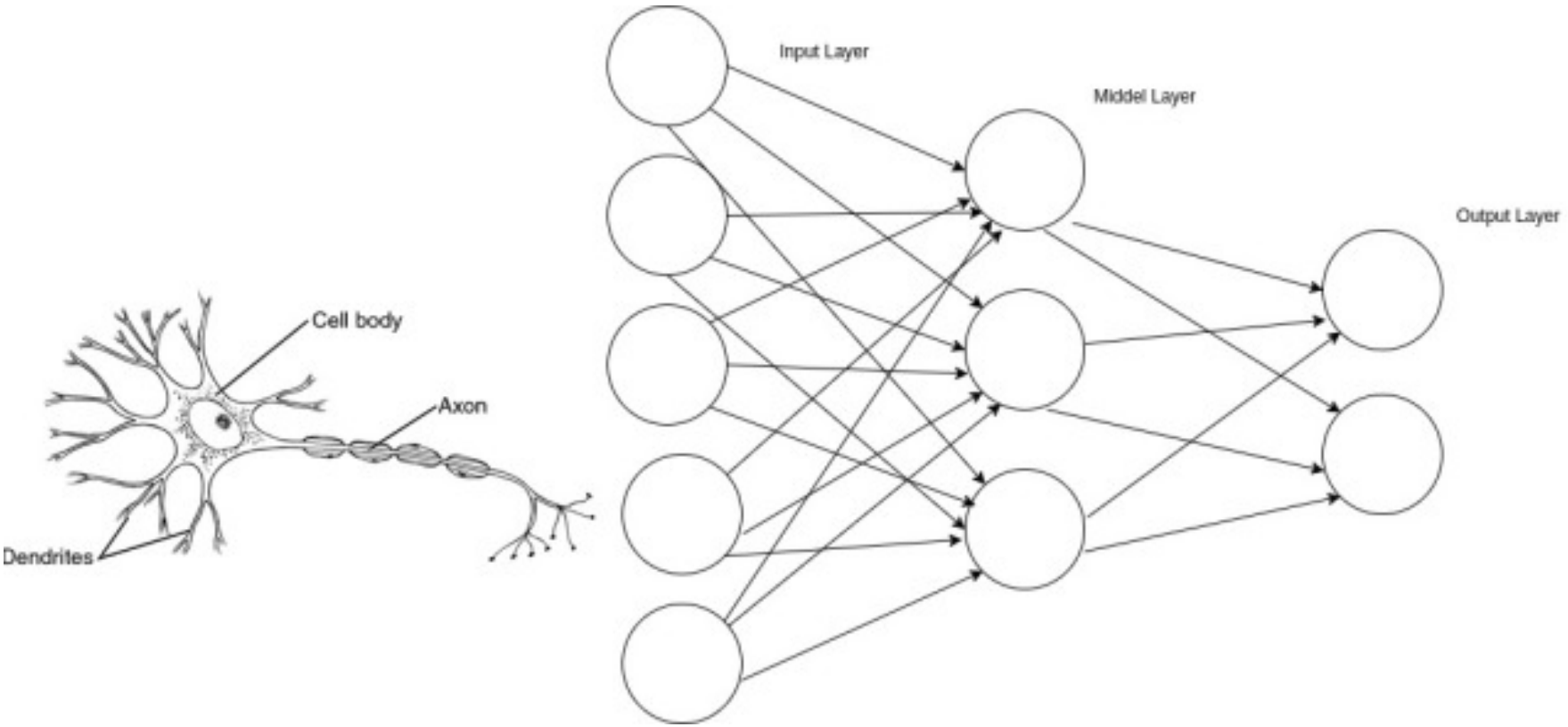
Target

- Detect and recognize different faces in photos
- Use a computer vision's framework (OpenCv) and a machine learning framework (Tensorflow)

Outline

What does machine learning means?

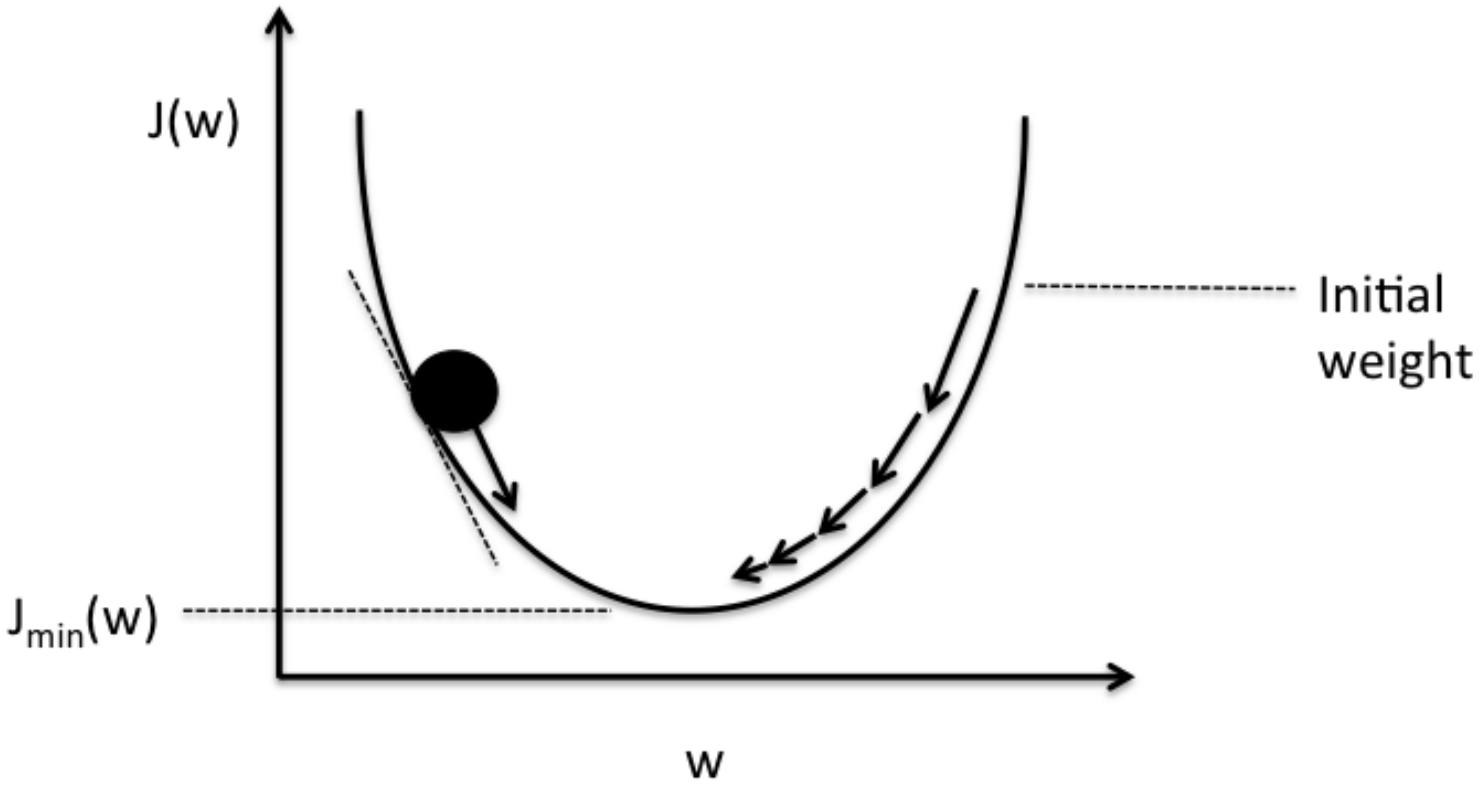
- Is this a **neural network** or a **graph**?



What does machine learning means?

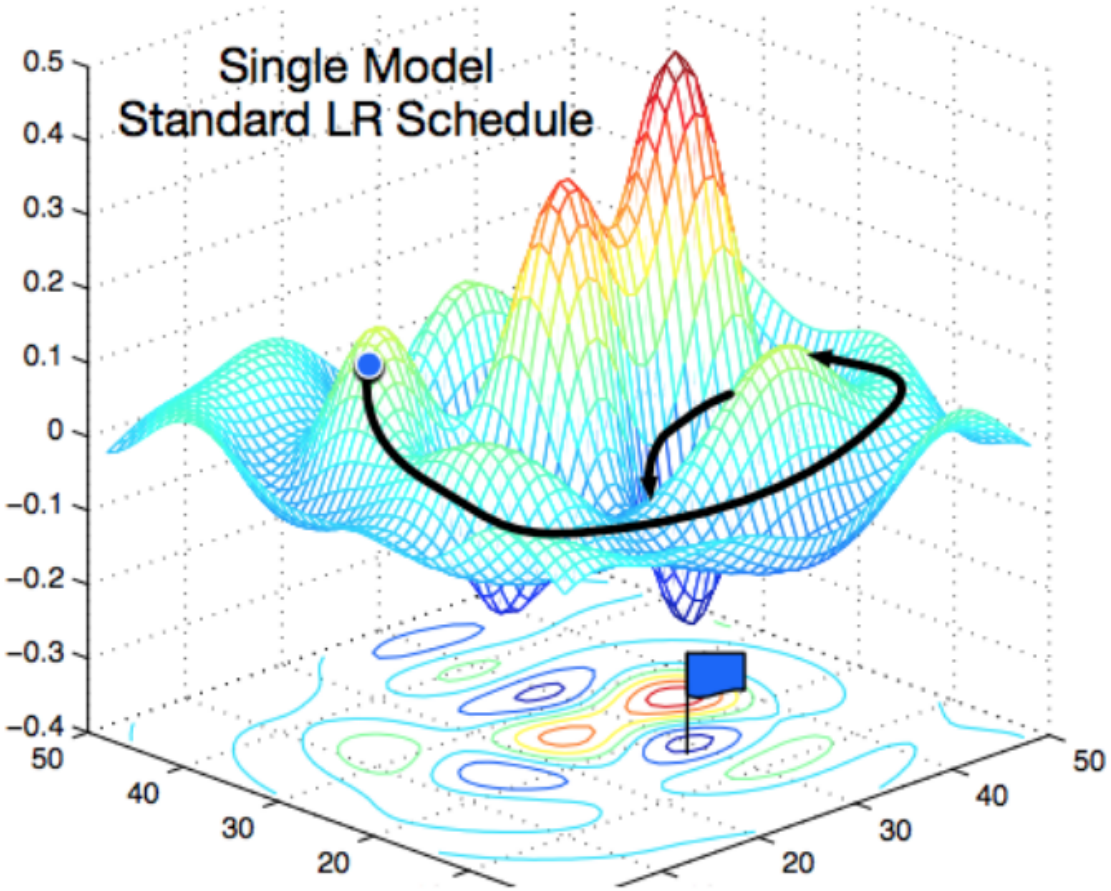
- The concept of **training**:
  - input and output;
  - inference and loss function;
  - gradient descent and weights update.
- The importance of a large and well organized dataset:
  - common problems;
  - cognitive bias.

Training and gradient descent



Schematic of gradient descent.

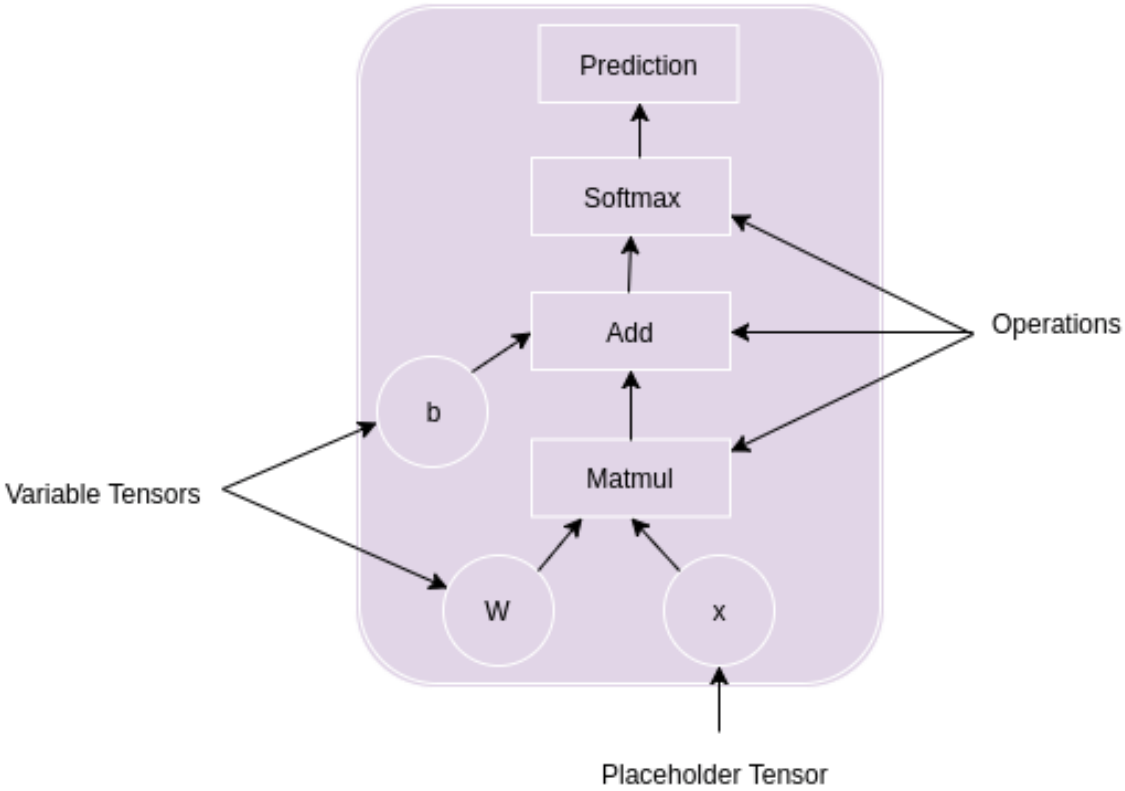
A more complex gradient descent loss



Outline

Tensorflow and OpenCv

- Tensorflow and computational graph concept



Computational Graph

Tensorflow

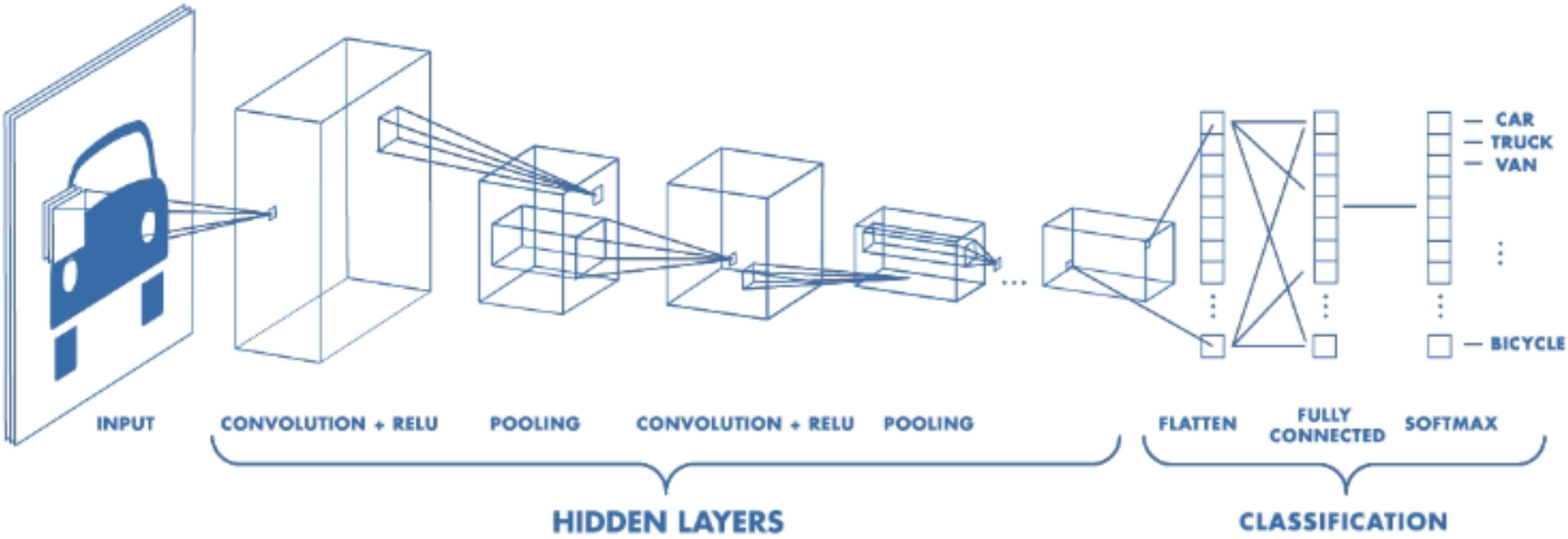
- Why Tensorflow?
  - Tensors and use of computational graph;
  - different levels of abstraction.
- Low level and high level API:
  - Tensorflow functions;
  - Keras and tflearn.

OpenCV

- Why OpenCV?
  - Very fast;
  - lots of library and algorithm at the state-of-art.
- OpenCv detection algorithm:
  - HaarCascadeClassifier;
  - Dlib library for facial features detection.

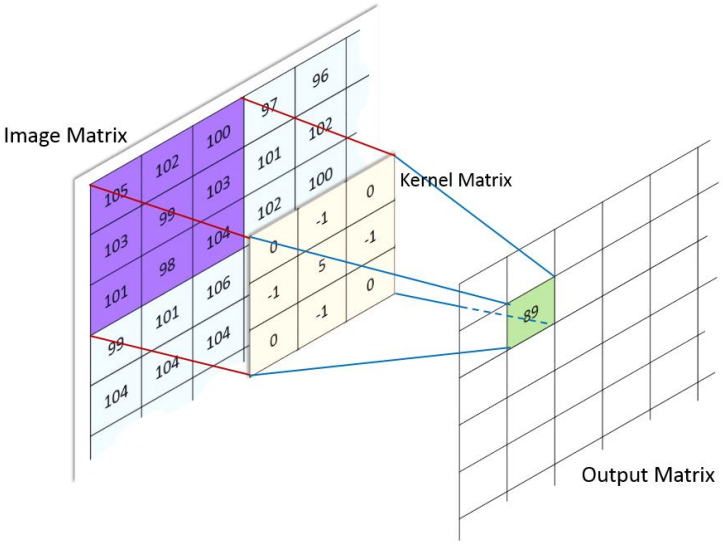
Outline

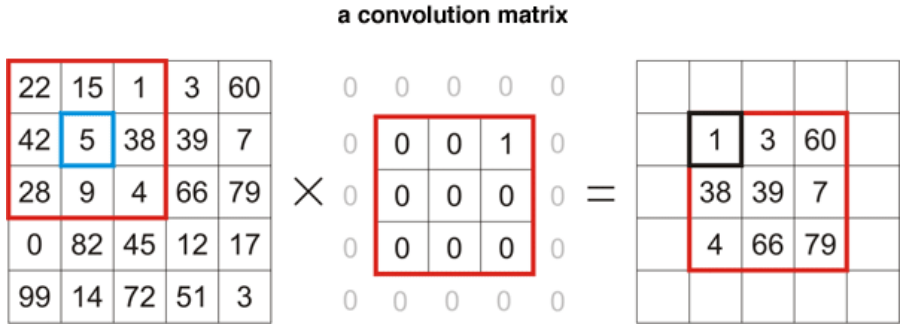
Convolutional Neural Network



Convolutional layers

- Convolutional matrix (Kernel)

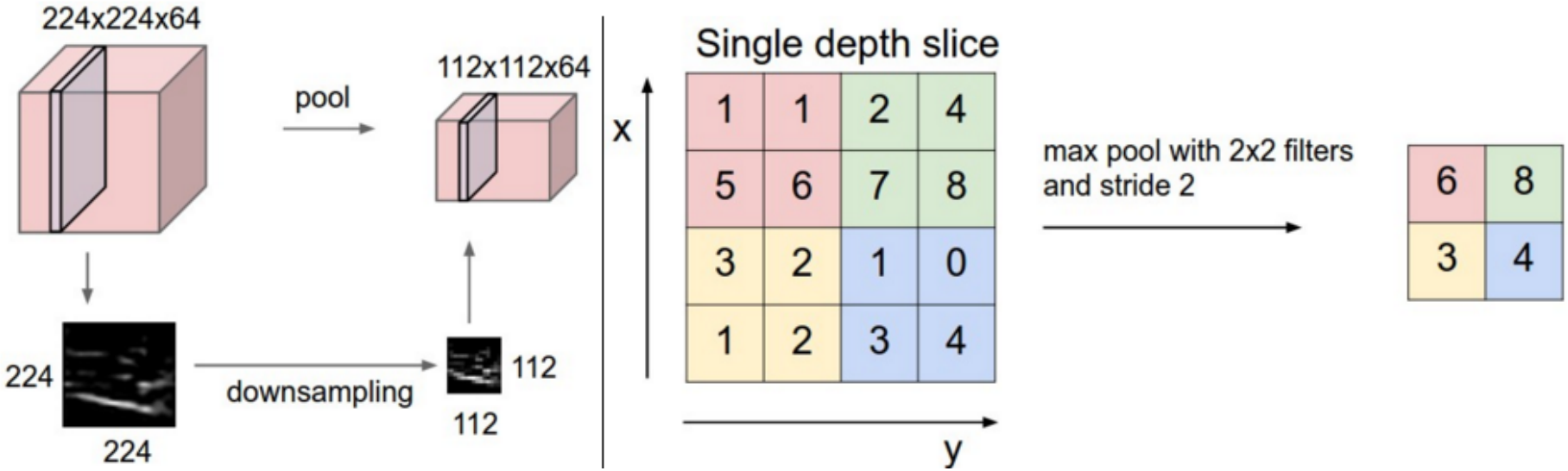




- 3x3, 5x5, or 7x7, why only odd numbers?
- Edge detection:
  - Similarity with human vision;
  - From simple to complex forms.

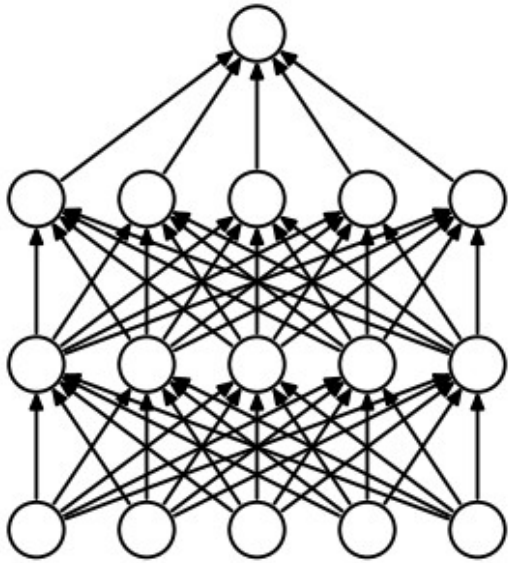
Pooling layers

- Reducing number of information: best way to avoiding **overfitting** and decreasing computation complexity

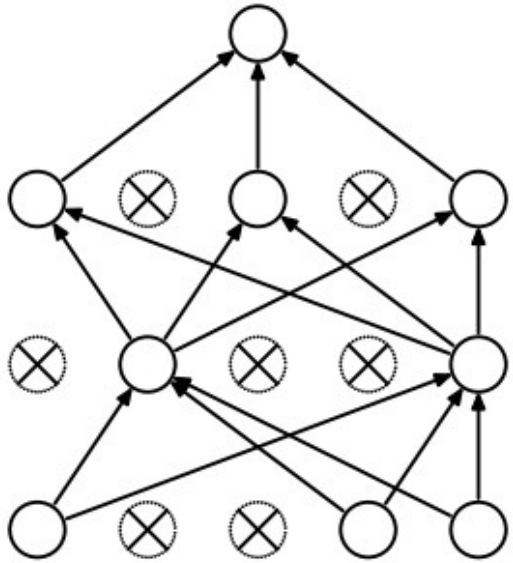


Fully connected layers and dropout

- Fully connected layers are the last layer of the CNN
- Once the high-level features are recognized, they deal with classifications
- Dropout regolarization



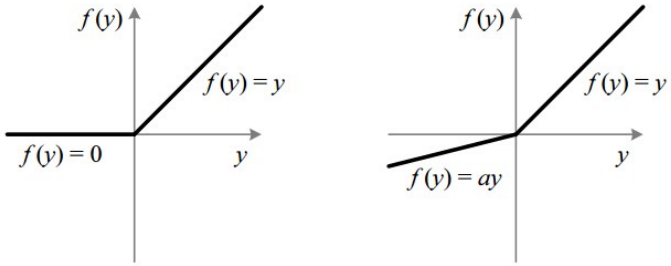
(a) Standard Neural Net



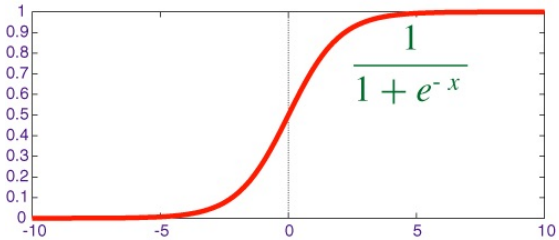
(b) After applying dropout.

Activation functions

- ReLU



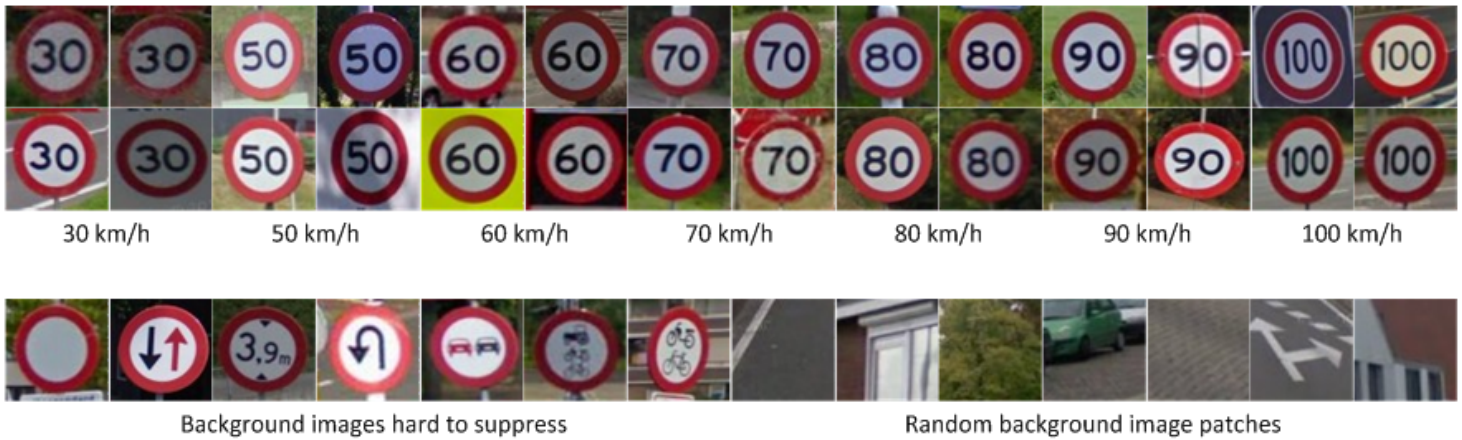
- Softmax



Outline

Dataset

- The perfect dataset should be:
  - made with hundreds of images;
  - different images with different colors to help the network classify them better.



Outline

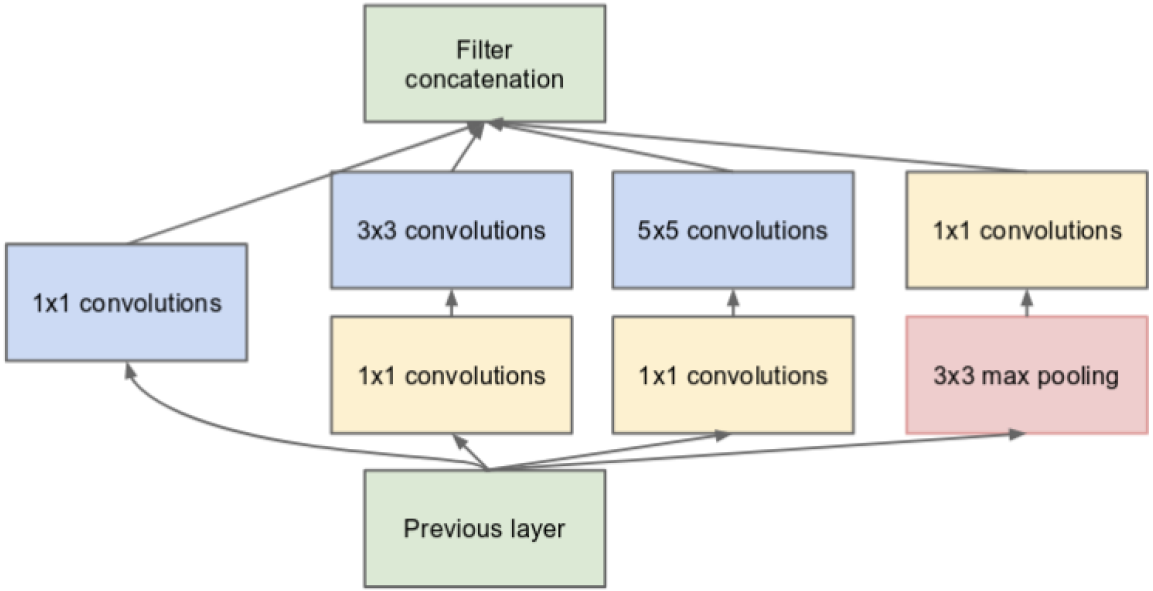
ImageNet Challenge

- The ImageNet project:
  - large visual database designed for visual object recognition software research;
  - it contains over 20 thousand categories("balloon", "strawberry", etc...);
  - all the images are labelled.
- ImageNet Large Scale Visual Recognition Challenge (ILSVRC):
  - is a competition where research teams evaluate their algorithms on the given data set(ImageNet);
  - they compete to achieve higher accuracy on several visual recognition tasks.

Outline

Inception network

- Inception networks analize images with different kernel size (in the same conv layer)



Outline

Tensorflow and retraining

- Concept of retraining
- Tensorflow-hub: the key to create your own classifier with good result and without a Tesla k80

Outline

Obtain images

- It use some OpenCv functions to get hundreds of photos in less than 30 seconds;
- after that, it crops them keeping only the faces(using a face detection algorithm) and it saves them;
- this is made to avoid the recognition of unwanted features as background color without the need of hundreds of images taken in different places.

Outline

Training

- The algorithm takes the cropped pics and use them as dataset for the training;
- in fact only the fully connected layers are trained;
- relatively fast.

Outline

Inference

- Choose a pic with multiple faces that you want to recognize/classify;
- select the graph you want to use for inference;
- you should obtain the original photo with multiple boxes around all the faces and, under each of them, the name of the most probable person.

Outline

Conclusion

- Create your own machine learning program using another pre-trained model can help you to build something useful without the need of a workstation or cloud computing
- This project is only a small example of the potentiality of Tensorflow and the machine learning approach

Outline

Future implementation

- Let the users to choose beetwen more pre-trained models
- Find the best way to recognize an unknown person (someone who does not have photos yet)

Outline

Best bugs

- OpenCv imshow freezing bug on unix like system
- Tensorflow-hub requires a tensorflow version that could not work with lots of processors(precompiled with AVX activation) (Link to issue)

Useful links

- Project repository: Link to repo
- Tensorflow: Link to Tensorflow page
- OpenCv: Link to OpenCv project
- Convolutional Neural Network example: GoogleNet
- Tensorflow retrain: tensorflow retraining