

Multiple face recognition in images

Presentation on Convolutional Neural Networks, 2018

Simone Caldarella

IEEE Student Branch Brescia
University of Brescia

Outline

- 1 Introduction
 - Motivation
 - Target
- 2 Background
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet
- Inception V3 by Google
- Tensorflow and retraining
- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference
- 4 Summary
 - Conclusion
 - Future implementation
 - Best bugs

Outline

1 Introduction

• Motivation

• Target

2 Background

• Machine learning

• Tensorflow and OpenCv

• Convolutional Neural Network

• Dataset

• ImageNet

• Inception V3 by Google

• Tensorflow and retraining

3 Proposal

• Obtain images

• Training (or retraining)

• Inference

4 Summary

• Conclusion

• Future implementation

• Best bugs

Motivation

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

Outline

- 1 Introduction
 - Motivation
 - **Target**
- 2 Background
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- Inception V3 by Google
 - Tensorflow and retraining
- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference
 - 4 Summary
 - Conclusion
 - Future implementation
 - Best bugs

Target

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

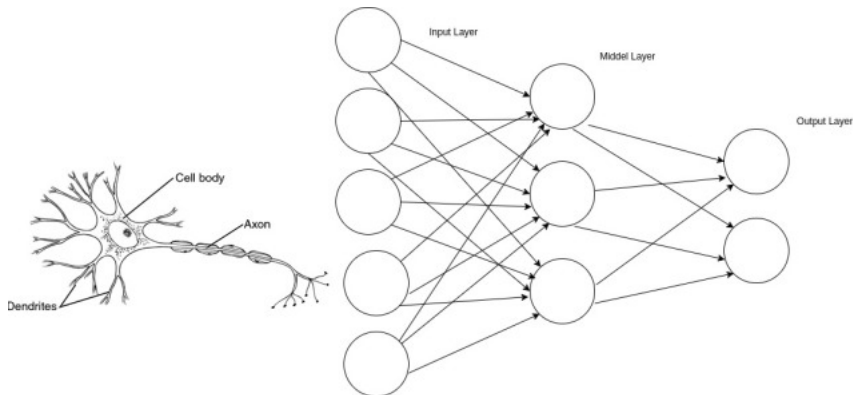
Outline

- 1 Introduction
 - Motivation
 - Target
- 2 **Background**
 - **Machine learning**
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- Inception V3 by Google
 - Tensorflow and retraining
- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference
 - 4 Summary
 - Conclusion
 - Future implementation
 - Best bugs

What does machine learning means?

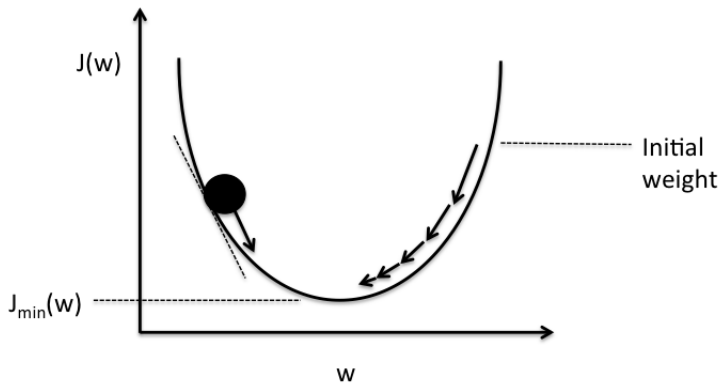
► Is this a **neural network** or a **graph**?



What does machine learning means?

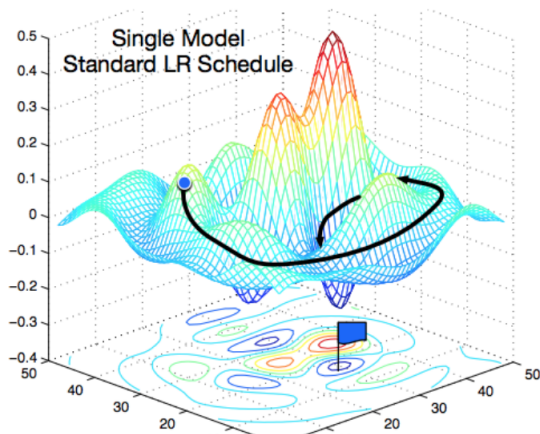
- ▶ The concept of **training**:
 - input and output;
 - inference and loss function;
 - gradient descent and weights update.
- ▶ The importance of a large and well organized dataset:
 - common problems;
 - cognitive bias.

Training and gradient descent



Schematic of gradient descent.

A more complex gradient descent loss



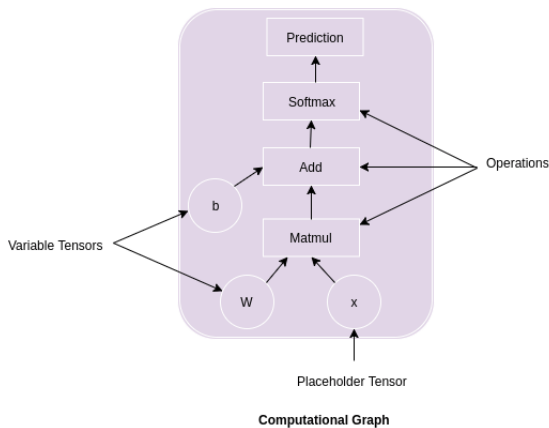
Outline

- 1 Introduction
 - Motivation
 - Target
- 2 **Background**
 - Machine learning
 - **Tensorflow and OpenCv**
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- Inception V3 by Google
 - Tensorflow and retraining
- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference
 - 4 Summary
 - Conclusion
 - Future implementation
 - Best bugs

Tensorflow and OpenCv

► Tensorflow and computational graph concept



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

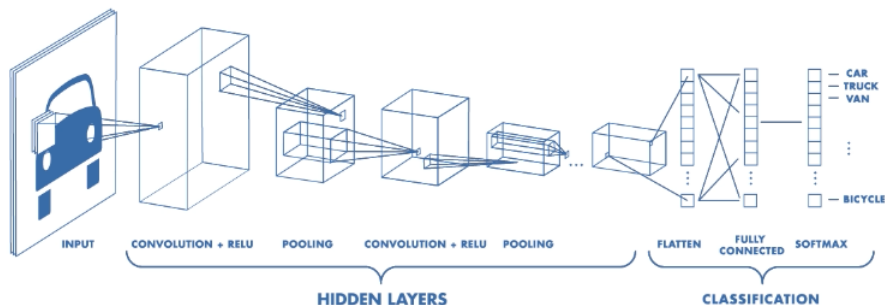
- ▶ Why OpenCV?
 - Very fast;
 - lots of library and algorithm at the state-of-art.
- ▶ OpenCv detection algorithm:
 - HaarCascadeClassifier;
 - Dlib library for facial features detection.

Outline

- 1 Introduction
 - Motivation
 - Target
- 2 **Background**
 - Machine learning
 - Tensorflow and OpenCv
 - **Convolutional Neural Network**
 - Dataset
 - ImageNet

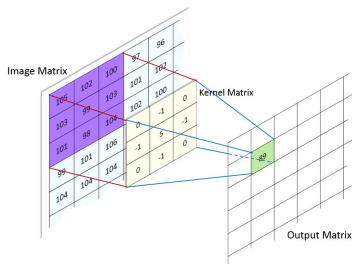
- Inception V3 by Google
 - Tensorflow and retraining
- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference
 - 4 Summary
 - Conclusion
 - Future implementation
 - Best bugs

Convolutional Neural Network



Convolutional layers

► Convolutional matrix (Kernel)



a convolution matrix

22	15	1	3	60
42	5	38	39	7
28	9	4	66	79
0	82	45	12	17
99	14	72	51	3

 \times

0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

 $=$

1	3	60		
38	39	7		
4	66	79		

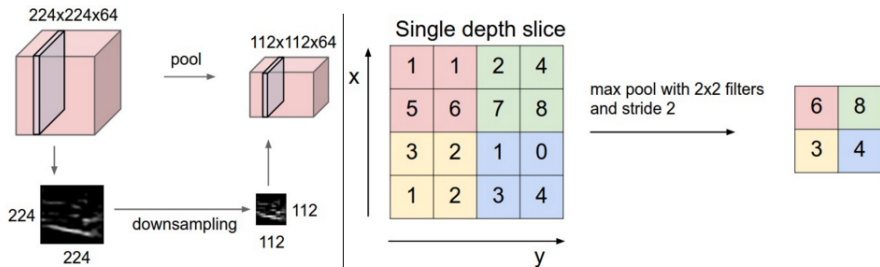
► 3x3, 5x5, or 7x7, why only odd numbers?

► Edge detection:

- Similarity with human vision;
- From simple to complex forms.

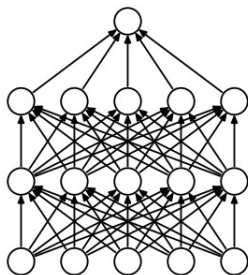
Pooling layers

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

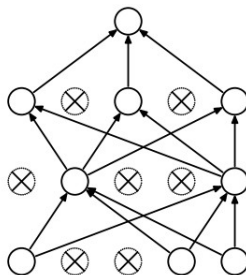


Fully connected layers and dropout

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



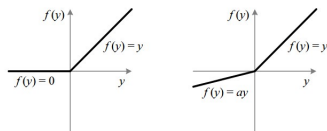
(a) Standard Neural Net



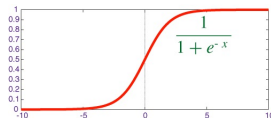
(b) After applying dropout.

Activation functions

► ReLU



► Softmax



Outline

- 1 Introduction
 - Motivation
 - Target
- 2 **Background**
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - **Dataset**
 - ImageNet

- Inception V3 by Google
- Tensorflow and retraining

- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference
- 4 Summary
 - Conclusion
 - Future implementation
 - Best bugs

Dataset

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



Background images hard to suppress

Random background image patches

Outline

- 1 Introduction
 - Motivation
 - Target
- 2 **Background**
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - **ImageNet**

- Inception V3 by Google
 - Tensorflow and retraining
- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference
 - 4 Summary
 - Conclusion
 - Future implementation
 - Best bugs

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.

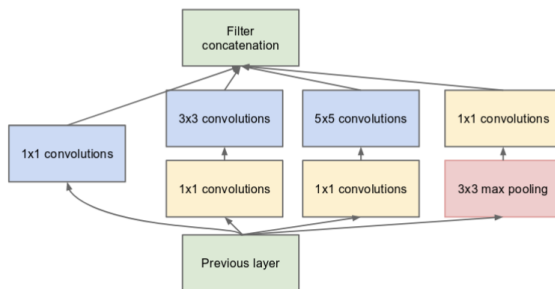
Outline

- 1 Introduction
 - Motivation
 - Target
- 2 **Background**
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- **Inception V3 by Google**
 - Tensorflow and retraining
- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference
- 4 Summary
 - Conclusion
 - Future implementation
 - Best bugs

Inception network

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



Outline

- 1 Introduction
 - Motivation
 - Target
- 2 **Background**
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- Inception V3 by Google
 - **Tensorflow and retraining**
- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference
 - 4 Summary
 - Conclusion
 - Future implementation
 - Best bugs

Tensorflow and retraining

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

Outline

- 1 Introduction
 - Motivation
 - Target
- 2 Background
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- Inception V3 by Google
- Tensorflow and retraining

3

Proposal

- **Obtain images**
- Training (or retraining)
- Inference

4

Summary

- Conclusion
- Future implementation
- Best bugs

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

- 1 Introduction
 - Motivation
 - Target
- 2 Background
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- Inception V3 by Google
- Tensorflow and retraining

3

Proposal

- Obtain images
- **Training (or retraining)**
- Inference

4

Summary

- Conclusion
- Future implementation
- Best bugs

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

- 1 Introduction
 - Motivation
 - Target
- 2 Background
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- Inception V3 by Google
- Tensorflow and retraining

3

Proposal

- Obtain images
- Training (or retraining)
- **Inference**

4

Summary

- Conclusion
- Future implementation
- Best bugs

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

- 1 Introduction
 - Motivation
 - Target
- 2 Background
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- Inception V3 by Google
- Tensorflow and retraining

- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference

- 4 **Summary**
 - **Conclusion**
 - Future implementation
 - Best bugs

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

- 1 Introduction
 - Motivation
 - Target
- 2 Background
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- Inception V3 by Google
 - Tensorflow and retraining
- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference
 - 4 **Summary**
 - Conclusion
 - **Future implementation**
 - Best bugs

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

- 1 Introduction
 - Motivation
 - Target
- 2 Background
 - Machine learning
 - Tensorflow and OpenCv
 - Convolutional Neural Network
 - Dataset
 - ImageNet

- Inception V3 by Google
- Tensorflow and retraining

- 3 Proposal
 - Obtain images
 - Training (or retraining)
 - Inference

- 4 Summary
 - Conclusion
 - Future implementation
 - **Best bugs**

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](#)