+33 (0)1 48 10 09 46
formation@ambient-it.net
140 Avenue Jean Lolive 93500 Pantin - France

Updated 03/15/2024

Sign up

Introduction to Deep Learning for image interpretation & processing

2 days (14 hours)

Presentation

One of the main fields of application for Deep Learning is image processing. Since 2012, this field has focused on classification problems, and all the main application or architecture innovations were initially dedicated to image interpretation or transformation. This course aims to present the main approaches and, for each one, the fundamental algorithms and architectures up to a selection of the state of the art.

Topics covered range from classification and segmentation to image transformation, and include the generation of text-oriented analyses. Two points of attention are proposed concerning the use of convolutional networks on a mobile device, and existing tools for interpreting trained models.

Objectives

- Review of the fundamentals of Deep Learning and convolutional networks
- Mastery of fundamental classification architectures
- Understanding the specifics of approaches dedicated to mobile networks, object detection and image segmentation
- Mastery of the main architectures and applications for image transformations and generation of text summaries of images
- Review of interpretation and security issues for a convolutional neural network

Target audience

Developers, Architects, Big Data Data Analyst / Data Engineer / Data Scientist

Prerequisites

- Knowledge of Python and mathematics

Further information

- As an introduction to Artificial Intelligence, we offer you the following training course
- Complementary technology
  - TensorFlow from Google
  - Pytorch from Facebook

Training program Introduction to **Deep Learning for image interpretation and processing**

[DAY 1]

1. Convolutional Neural Network: the basics

- Presentation of the fundamental architecture of a CNN layer: convolution, stride, pooling.
- Common non-linearity functions.
- Using pooling.
- Classification issues: cost functions, probabilistic approach.
- Back propagation and learning, generalization vs. overfit.
- Basic adjustments to prevent overfitting.
- Feature maps & Features at a CNN level.
- Examples: Visualizing how a CNN works

2. Fundamental convolutional network architectures

- LeNet architecture: calculation parallelization, implementation details and results.
- DropOut presentation for regularization.
- VGG architecture: different models up to VGG 19, limits of layering.
- Network in Network architecture: 1x1 convolution.
- Inception architecture (Google): exploiting low-dimensional convolutions, building an Inception module, using auxiliary cost functions.
- Batch Normalization: presentation of the regularization used in Inception.
- Residual connection: Highway & residual network. Managing an uninterrupted gradient flow, new architecture dimensions.
- Inception v3, XCeption.

3. Classification with performance management (mobile device, embedded)

- MobileNets v1: CNN architecture dedicated to mobile devices. Optimization of a convolutional resource
- ShuffleNet v1: Using group convolutions and channel shuffling
- MobileNets v2: Performance enhancements: Inverted residual structure and linear bottlenecks
- ShuffleNets v2: Practical guidelines for performance-oriented CNN architecture design

4. Object detection and image segmentation

- U-Networks Architecture: Principle architectures and implementations for segmentation. Medical applications/urban scenes (Kaggle Ultrasound Nerve Segmentations or Carvana Image Masking competitions).
- Attention models for a convolutional network and for a U-Network.
- SegNet architecture: resource-efficient approach, implementation details
- DeepLab: "Atrous Convolution", exploitation of Conditional Random Fields
- Yolo and Yolo v3 approaches: details of a performance-oriented architecture for static or real-time object detection
- Region-based approaches: R-CNN and Faster R-CNN: architecture and implementation principles
- Mask Region Based CNN: state-of-the-art architecture for multi-class instance-aware segmentation

[DAY 2]

5. Image captioning & Visual Question Answering: generating image-oriented descriptions or analyses

- Recurrent architecture and LSTM: basic presentation, gradient flow difference, internal states of a recurrent cell.
- Modularization of convolutional and recurrent architectures: design strategies (Recurrent CNN VS Stacking LSTM / CNN)
- Fundamental implementation: "Show and tell, neural image caption generator".
- Attention models for image processing
- Attention-grabbing: "Show, Attend and Tell
- VQA Challenge and state of the art 2017: Bottom-up & top-down attention for Image Captioning.

6. Simple image transformation and interpretation

- Topics: super resolution, image repair (inpainting), neural style
- Notion of high-level feature of a CNN as a latent representation of an image. Neural style applications.
- U-Networks architecture for these topics
- Enhanced Deep Residual Networks: state of the art 2017, block residual architecture
- Deep Image prior: Bayesian approach to one-shot learning

- Generative Adversarial Networks architectures: fundamental principles, minmax convergence rules, latent space
- Pix2Pix HD : Semantic supervised image generation
- CycleGAN: Unsupervised generation of image transformations

7. Examples of specific applications of CNN architectures (examples of software implementations)

- Video optical flow estimation (motion representation)
- Detection and estimation of the pose (skeletal position) of an individual in an image or video.
- Image unscrambling: comparative approaches
- Colorization of black and white images: comparative approaches.

8. Interpretation and safety

- Limits to the interpretability of Deep Learning
- Adversarial attacks, principle and ease of misleading a neural network
- Visualization of kernels and activation maxima
- Hierarchical Contextual Decompositions
- Tensorflow Lucid: activation atlases, activation grids, spatial visualization.

Companies concerned

This training course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

Teaching methods

Practical course: 60% Practical, 40% Theory. Training material distributed in digital format to all participants.

Organization

The course alternates theoretical inputs from the trainer supported by examples and

brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

Sanction

A certificate will be issued to each trainee who completes the course.