

Formation Introduction Deep Learning pour l'interprétation & le traitement d'images

2 jours (14 heures)

Présentation

Un des champs d'application privilégiés du Deep Learning est le traitement de l'image. C'est sur des problèmes de classification que ce domaine s'est révélé depuis 2012, et toutes les principales innovations d'application ou d'architecture ont été dans un premier temps dédiées à l'interprétation ou à la transformation d'images. Cette formation vise à présenter les principales approches et, pour chacune, les algorithmes et architectures fondamentales jusqu'à une sélection d'états de l'art.

Les thèmes abordés vont de la classification ou la segmentation à la transformation d'images en abordant les problématiques de génération d'analyses orientées à partir de texte. Deux points d'attention sont proposés concernant l'utilisation de réseaux convolutionnels sur un périphérique mobile, ou les outils existant en interprétation de modèles entraînés.

Objectifs

- Revue des fondamentaux en Deep Learning et en réseaux convolutionnels
- Maîtrise des architectures fondamentales de classification
- Compréhension des spécificités d'approches dédiées aux réseaux mobiles, détection d'objets et segmentation d'images
- Maîtrise des architectures principales et applications pour les transformations d'images et la génération de résumés textuels d'images
- Revue des problématiques d'interprétation et de sécurité d'un réseau de neurones convolutionnel

Public visé

Développeurs, Architectes, Big Data Data Analyst / Data Engineer / Data Scientist

Pré-requis

- Connaissance de Python et en mathématique

Pour aller plus loin

- ? Notre certification au [Deep Learning](#) ?
- Nous vous proposons en introduction un formation sur l'[Intelligence Artificielle](#)

- En complément la technologie
 - TensorFlow de Google
 - Pytorch de Facebook

Programme de la formation Introduction **Deep Learning pour l'interprétation ou le traitement d'images**

[JOUR 1]

1. Convolutional Neural Network : présentation des bases

- Présentation de l'architecture fondamentale d'un layer CNN : convolution, stride, pooling.
- Fonctions de non-linéarité usuelles.
- Utilisation du pooling.
- Problématiques de classification : fonctions de coût, approche probabiliste.
- Back propagation et apprentissage, notion de généralisation vs. overfit.
- Régularisations basiques contre l'overfitting.
- Feature maps & Features haut niveau d'un CNN.
- Exemples : Visualisation du fonctionnement d'un CNN

2. Architectures fondamentales de réseaux convolutionnels

- Architecture LeNet : parallélisation du calcul, détail d'implémentation et résultats.
- Présentation du DropOut pour régularisation.
- Architecture VGG : différents modèles jusqu'au VGG 19, limites de la superposition de couches.
- Architecture Network in Network : convolution 1x1.
- Architecture Inception (Google) : exploitation des convolutions à faible dimension, construction d'un module Inception, utilisation de fonctions de coût auxiliaires.
- Batch Normalisation : présentation de la régularisation utilisée dans Inception.
- Connexion résiduelle : Highway & residual network. Gestion d'un flux de gradient ininterrompu, nouvelles dimensions d'architecture.
- Inception v3, Xception.

3. Classification avec gestion de la performance (appareil mobile, embarqué)

- MobileNets v1 : Architecture CNN dédiée aux périphériques mobiles. Optimisation d'une ressource convolutionnelle
- ShuffleNet v1 : Utilisation de convolutions de groupes et shuffle des channels
- MobileNets v2 : Améliorations et performances : Structure résiduelle inversée et bottlenecks linéaires
- ShuffleNets v2 : Guidelines pratiques de design d'architecture CNN orienté performances

4. Détection d'objets et segmentation d'image

- Architecture U-Networks : Principe d'architectures et implémentations pour segmentation. Applications médicales/scènes urbaines (concours Kaggle Ultrasound Nerve Segmentations ou Carvana Image Masking).
- Modèles d'attention pour un réseau convolutionnel et pour un U-Net.
- Architecture SegNet : Approche efficace en ressources, détail d'implémentation

- DeepLab : "Atrous Convolution", exploitation de Conditional Random Fields
- Approches Yolo et Yolo v3 : détail d'une architecture orientée performance pour la détection d'objets statique ou temps réel
- Approches "Region based" : R-CNN et Faster R-CNN : principes d'architecture et d'implémentation
- Mask Region Based CNN : architecture d'état de l'art en segmentation "instance aware" multi-classes

[JOUR 2]

5. Image captioning & Visual Question Answering : génération de description ou d'analyse orientée sur une image

- Architecture récurrente et LSTM : présentation fondamentale, différence de flux de gradient, états internes d'une cellule récurrente.
- Modularisation d'architectures convolutionnelles et récurrentes : stratégies de design (Recurrent CNN VS Stacking LSTM / CNN)
- Implémentation fondamentale : "Show and tell, neural image caption generator"
- Modèles d'attention pour le traitement d'image
- Exploitation de l'attention : "Show, Attend and Tell"
- VQA Challenge et état de l'art 2017 : Bottom-up & top-down attention for Image Captioning.

6. Transformation simple et interprétation d'images

- Sujets : super résolution, réparation (inpainting) d'images, neural style
- Notion de feature haut niveau d'un CNN comme représentation latente d'une image. Applications neural style.
- Architecture U-Networks pour ces sujets
- Enhanced Deep Residual Networks : état de l'art 2017, architecture résiduelle par blocs
- Deep Image prior : approche bayésienne en one-shot learning
- Architectures Generative Adversarial Networks : principes fondamentaux, règles de convergence minmax, espace latent
- Pix2Pix HD : Génération d'image supervisée sémantique
- CycleGAN : Génération non supervisée de transformations d'images

7. Exemples d'applications spécifiques d'architectures CNN (exemples d'implémentations logicielles)

- Estimation de l'optical flow d'un vidéo (représentation du mouvement)
- Détection et estimation de la pose (position du squelette) d'un individu sur une image ou une vidéo.
- Défloutage d'images : approches comparées
- Colorisation d'images noir et blanc : approches comparées.

8. Interprétation et sécurité

- Limites de l'interprétabilité du Deep Learning
- Attaques adversariales, principe et facilité de mise en erreur d'un réseau de neurones
- Visualisation de kernels et de maximums d'activation
- Hierarchical Contextual Decompositions
- Tensorflow Lucid : Atlas d'activations, grilles d'activations, visualisation spatiale.

Sociétés concernées

Cette formation s'adresse aux entreprises, petites ou grandes, souhaitant former ses équipes à une nouvelle technologie informatique avancée.

Méthodes pédagogiques

Stage Pratique : 60% Pratique, 40% Théorie. Support de la formation distribué au format numérique à tous les participants.

Organisation

Le cours alterne les apports théoriques du formateur soutenus par des exemples et des séances de réflexions, et de travail en groupe.

Validation

À la fin de la session, un questionnaire à choix multiple permet de vérifier l'acquisition correcte des compétences.

Sanction

Une attestation sera remise à chaque stagiaire qui aura suivi la totalité de la formation.