Building a Visual Understanding Pipeline

02/02/2018

Charles Ollion, CoFounder at Heuritech
OUTLINE

HEURITECH
What we build & why we use image analysis

TECHNOLOGY
Our Visual Understanding Pipeline

FRONTIERS
Our research areas
Heuritech is a ML based company, Mostly **Text & image recognition**

Our vertical business focus: Trends spotting and analysis for **Fashion** and **Beauty**

- **22** people
- **3** Backend Engineers
- **3** Devops Engineers
- **7** People in R&D Team
- **7 PhDs + 1** PhD Candidate
Our AI technology, well-trained by experts, has reached brand new levels of performance, precise enough to grasp the details of fashion & luxury products.

**IMAGE ANALYSIS AT HEURITECH**

- **PEOPLE**
- **CONTEXT**
- **PRODUCT CATEGORIES**
- **SHAPES**
- **DETAILS & FEATURES**
- **COLOR SHADES**
- **PATTERNS & TEXTURES**
- **BRANDS**
- **STYLES**

- #Hat
- #NewsboyCap
- #Black
- #Bomber
- #Jacket
- #Military
- #Neon
- #Kaki
- #Orange
- #Backpack
- #Bag
- #PalmspringsMini
- #LouisVuitton
- #Patches
- #Leather
- #DarkBrown
- #Monogram
- #Top
- # Oversized
- # Shirt
- #Chic
- #Formal
- #Plain
- #White
- #Skirt
- # MidiLength
- # Asymmetric
- #Ruffles
- #Stripes
- #Black
- #Statement
- #Multicolor
- #PearlGrey
- #Red
- #Shoes
- #Heels
- #ShoeLaces
- #LowBoots
- #Camel
- #Black

© getty images / Christian Vierig
FASHION: VERY DIVERSE IMAGES

Source: Tamara L. Berg
VISION IN THE WILD

Difficulties

- Handling false positives
- Numerous
- Very small objects
- Object relations and attributes
OUTLINE

HEURITECH
What we build & why we use image analysis

TECHNOLOGY
Our Visual Understanding Pipeline

FRONTIERS
Our research areas
CONVOLUTIONAL NEURAL NETWORK (CNN)

BEYOND CLASSIFICATION

Classification

Classif + Localisation

single object

multiple objects

Object Detection

Instance Segmentation
TRANSFER LEARNING

SEGMENTATION OF OBJECTS

Mask R-CNN, K He, G Gkioxari, P Dollár, R Girshick, 2017 (arxiv preprint)
TRAINING INDUSTRIALISATION

From R&D training...

**algorithm, architecture, hyperparams, losses, sampling strategies**

...to Industrialized training

Train with new data / classes (Fixed algorithm)

**Models, performance** and **datasets** are versioned

Active Learning

Run current model on **large unlabelled dataset**

**Manually** relabel images that were wrongly tagged
OUTLINE

HEURITECH
What we build & why we use image analysis

TECHNOLOGY
Our Visual Understanding Pipeline

FRONTIERS
Our research areas
WEAKLY ANNOTATED DATA

Millions of e-commerce pictures + description
But description is very noisy

Used as basis for other classifiers / applications
DOMAIN PROBLEMS

Source domain:
E-COMMERCE

Millions of products openly found on the web. Associated with (almost clean) labels

Target domain:
WORN CLOTHES

Millions of images openly found on the web. But unlabeled: no correspondence with product!

Need a method for domain adaptation!
DOMAIN ADAPTATION

Labels from e-commerce

MODEL

Good results on e-commerce

Domain Adaptation

ADAPTED MODEL

Also good results on worn clothes

Source:
e-commerce

Target:
Worn clothes
Learn to classify clothes
(type, shape, style...)

Learn not to classify domain
(e-commerce VS worn)

Products = text + images

Posts (facebook, etc...) = text + images

In Computer Vision, Visual Question answering = benchmark for multimodal (text + image) fusion

[Diagram of GRU and Resnet with text: Is the mustache real?]

[VQA - Mutan 2017]