

# How computer vision is shaping the future in ecommerce



**machinalis**  
Machine Learning Solutions Delivery

## > Who Am I?

Systems Engineer

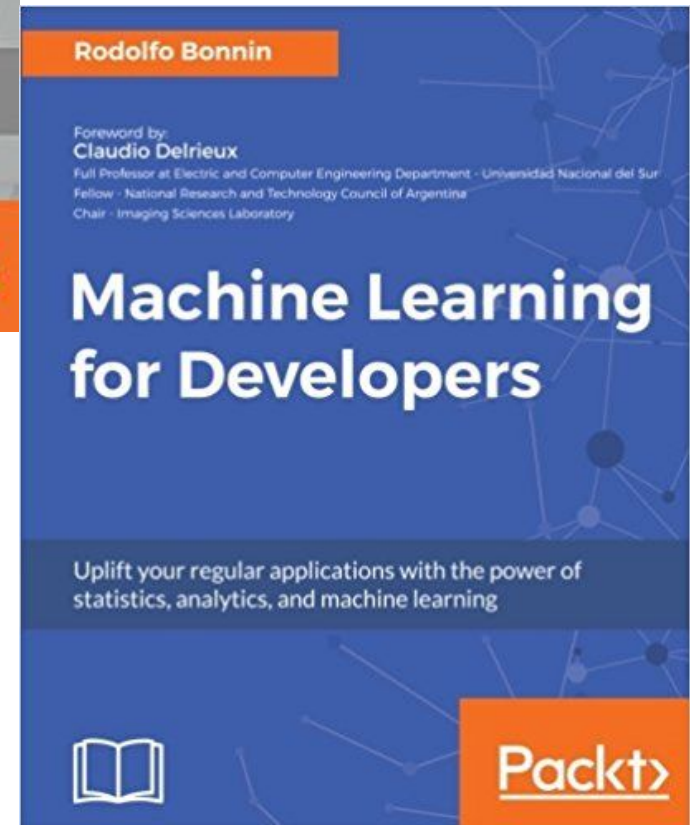
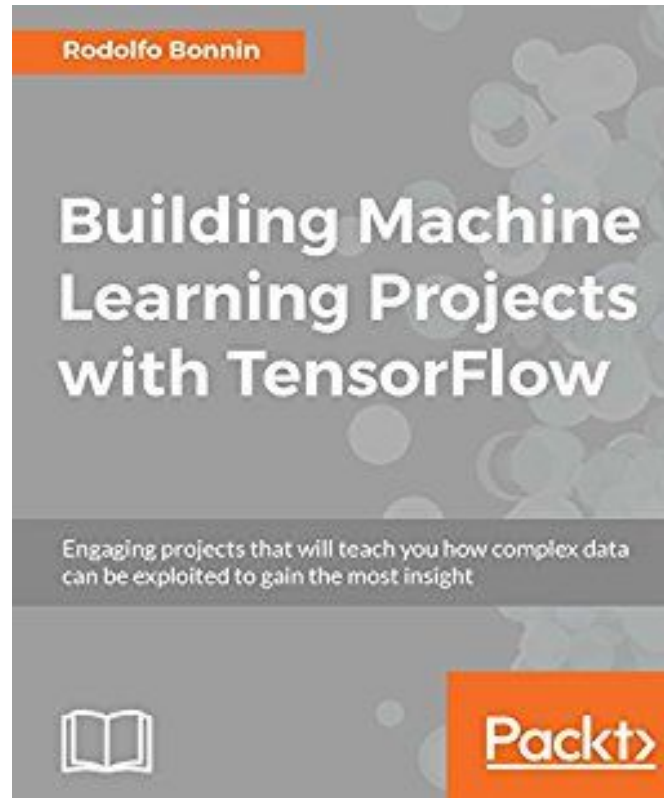
Ph.D Student, UTN

Working on Deep Learning since 2008

Ex ML @ Mercadolibre

CV Specialist @ Machinalis

Author of “Building Machine Learning Projects with Tensorflow” and “Machine Learning for Developers”





About us



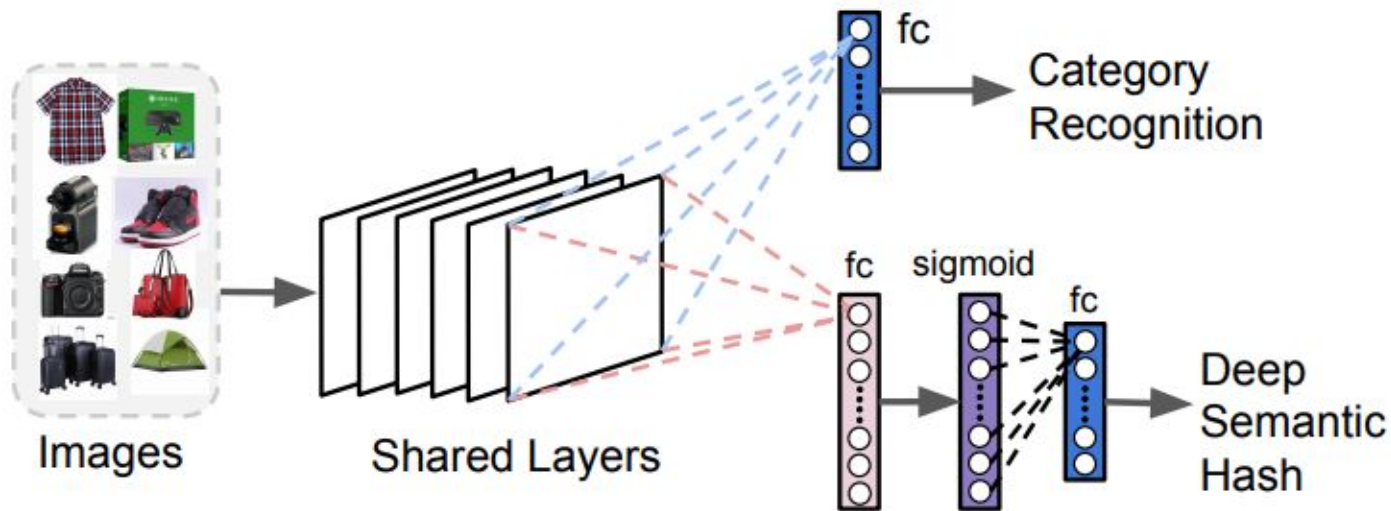
Visual Search

Automatic Categorization

Automatic Product Tagging

Moderation

➤ **Visual Search &  
Categorization @  
Ebay**

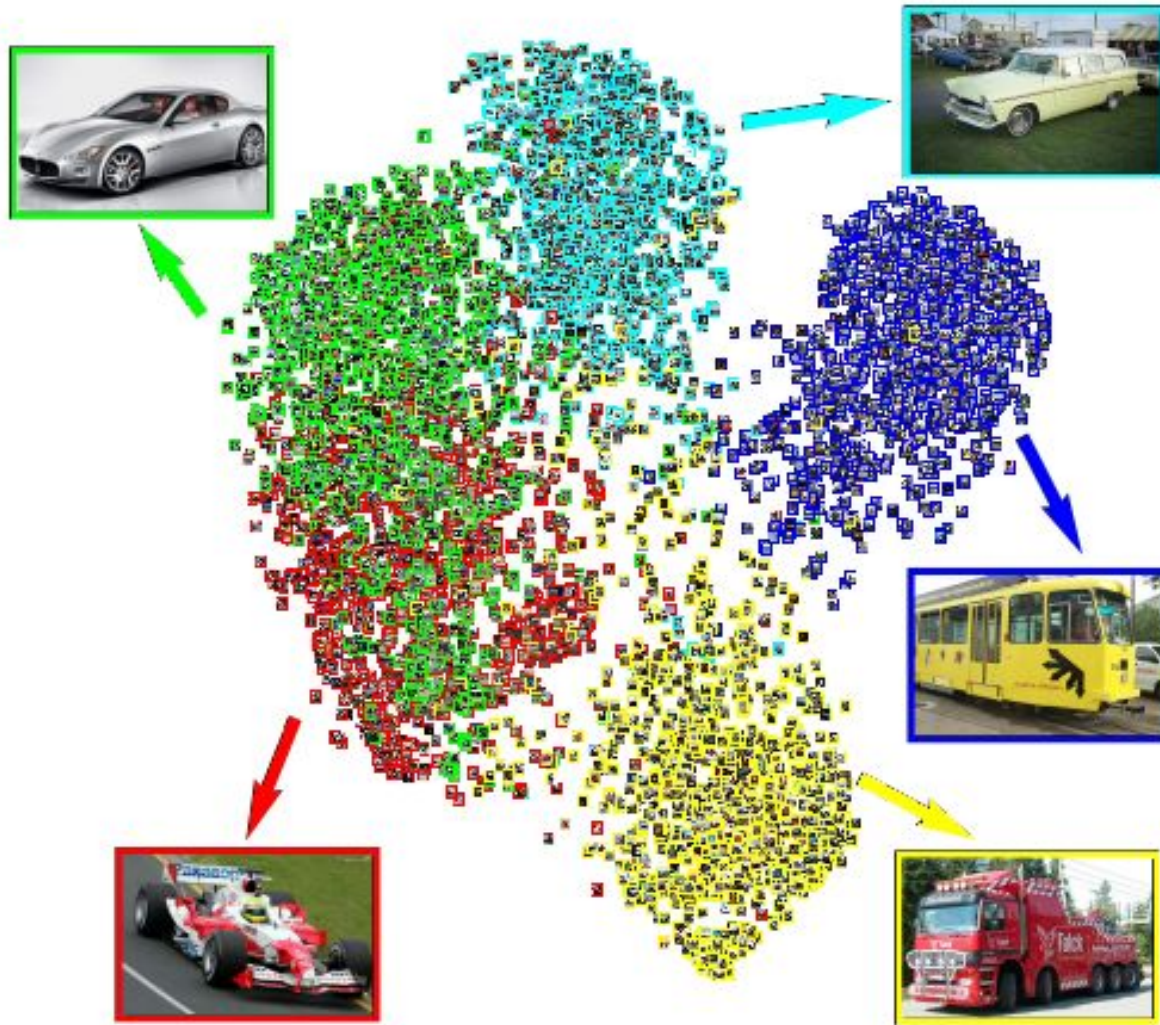


- Represent images as binary signatures instead of real values: reduces storage requirement and computation overhead.
- Hamming Distance

- DNN (ResNet50)
- Search only among top predicted categories and then use semantic binary hash with Hamming distance for fast ranking.
- For speed and low memory footprint, shared topology for both category prediction and binary hash extraction

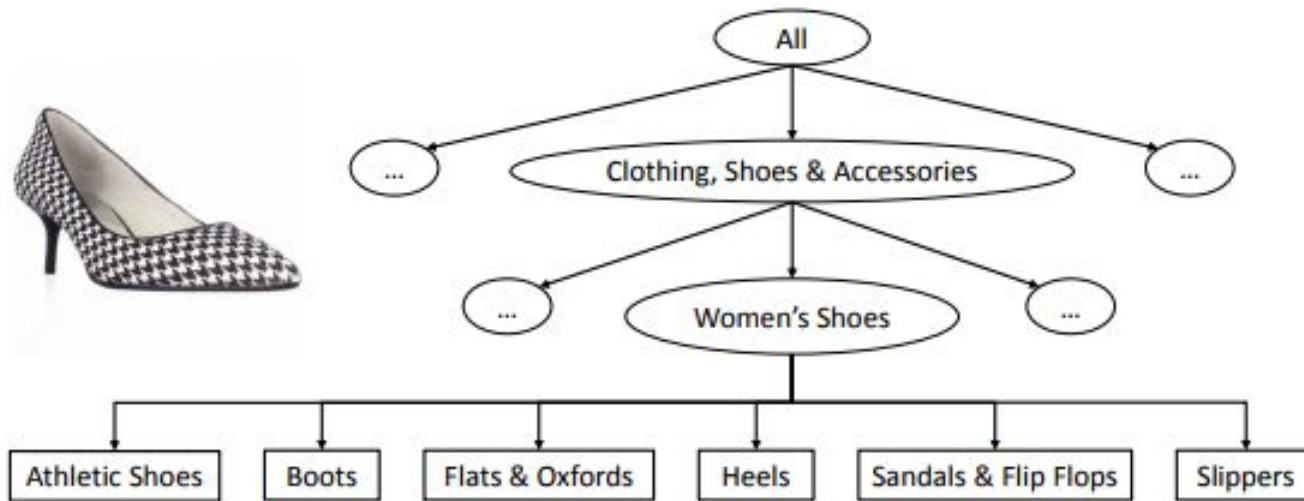


## ➤ Visual detection & Search @ Ebay



Experiments over imagenet:  
T-SNE Over 4096 semantic features (non-parametric mapping!)

## ➤ Visual detection & Search @ Ebay

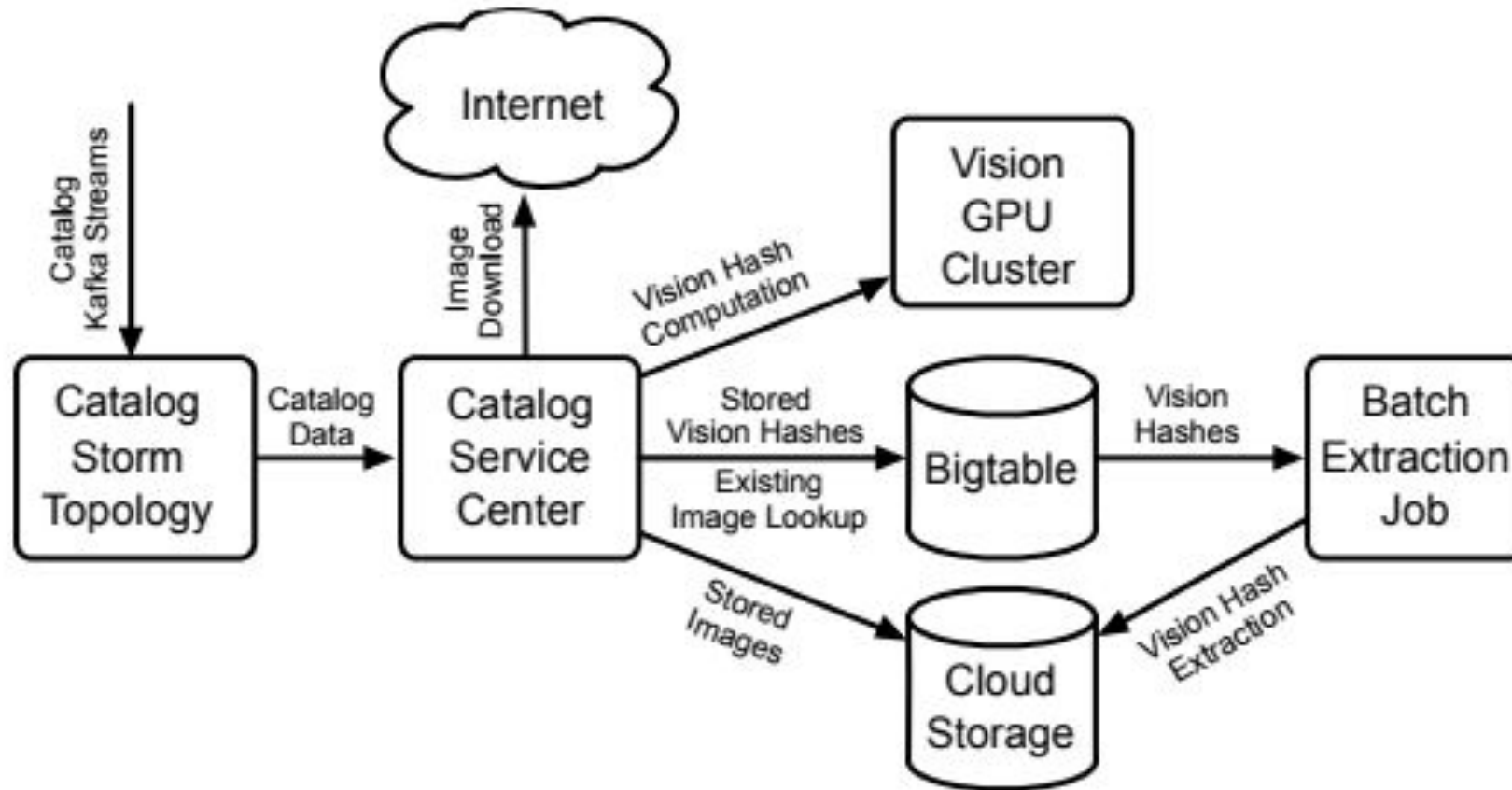


### Item Specifics

Condition:	New with box	Brand:	Michael Kors
Heel Type:	Kitten	Style:	Kitten Heels
Heel Height:	Med (1 ¾ in. to 2 ¾ in.)	US Shoe Size (Women's):	7
Material:	Haircalf	Width:	Medium (B, M)
Pattern:	Houndstooth	Color:	Multi-Colored

Aspects  
(Individual  
classifiers)



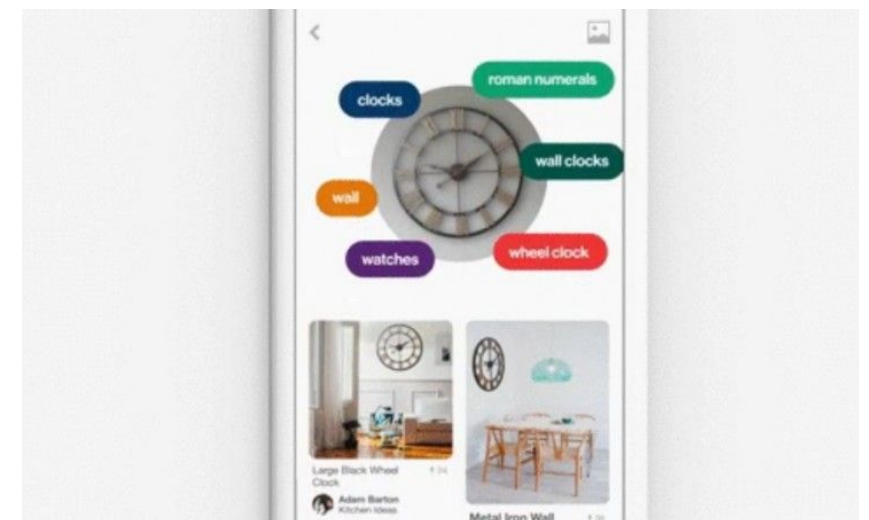
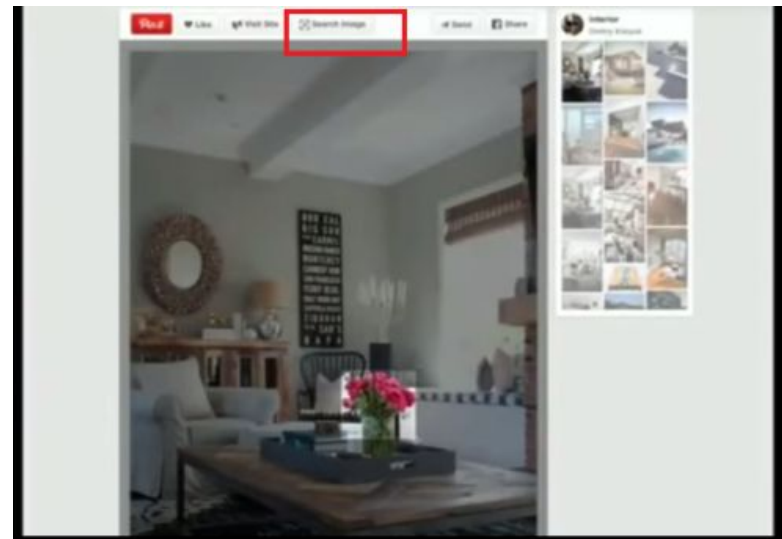
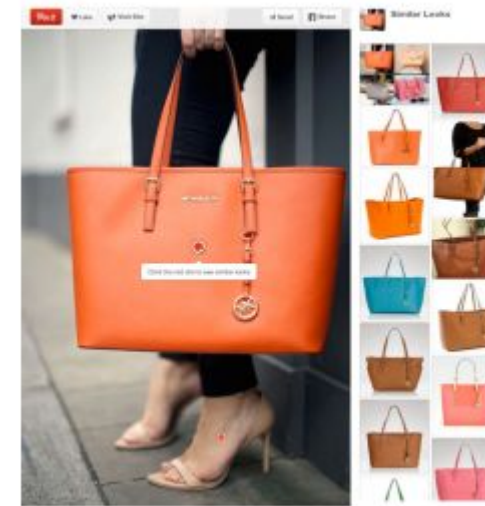
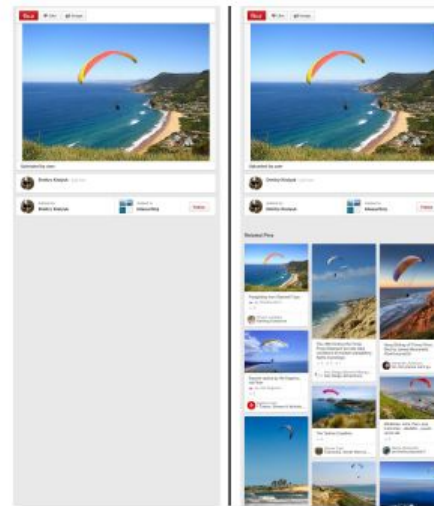


**Figure 6: Image ingestion system architecture**

➤ **Visual Search @  
Pinterest**

## ➤ Visual detection & Search @ Pinterest

- Related Pins (2014),
- Similar Looks (2015),
- Flashlight (2016)
- Lens (2017)



## ➤ Image representation and features

- Extraction of a variety of features from images, including **local features** and “**deep features**” extracted from the activation of **intermediate layers** of deep **convolutional neural networks**.
- Architectures based on **AlexNet** and **VGG** , extracting feature representations from fc6 and fc8 layers [Running on GPU]
- These features are binarized for representation efficiency and compared using **Hamming distance**.
- **Click Through Rate** based on **Visual Features**

## ➤ Two step object detection @ Pinterest

... fashion, **bags** & shoes hand  
picked by MoMo | See more ...



**Figure 3:** Instead of running all object detectors on all images, we first predict the image categories using textual metadata, and then apply object detection modules specific to the predicted category.

Table 3: Object detection performance.

	Faster R-CNN		SSD	
	precision	recall	precision	recall
<b>Fashion</b>	0.449	0.474	0.473	0.387
<b>Home decor</b>	0.413	0.466	0.515	0.360
<b>Vehicles</b>	0.676	0.625	0.775	0.775
<b>Overall</b>	0.426	0.470	0.502	0.371
<b>Latency</b>	272 ms		59 ms	



 **Specialized  
providers**

➤ Specialized providers: Vue.ai by Madstreetden





 **Our Work @  
Machinalis®**

➤ Border detection



➤ Personal data embedded in text



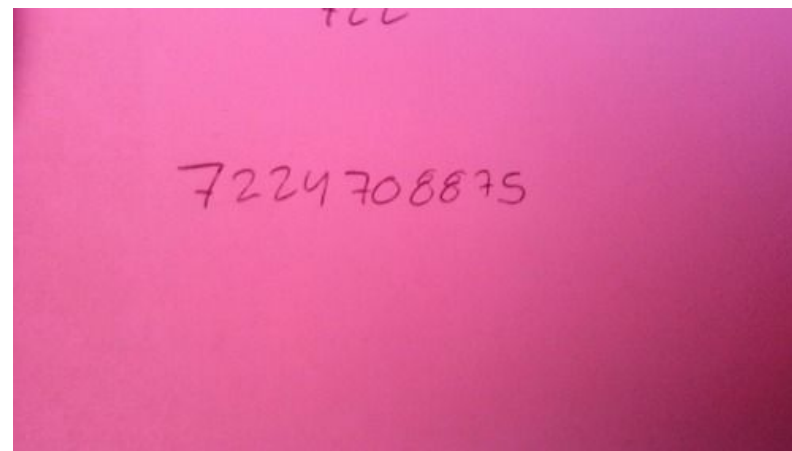
Recognized: ['.', '04142894437']



Recognized: ['0426-3905511', 'CHLOE']



[""] ["o", 'SN', 'whats', '19433422485', 'EL', 'runa', '=', 'a', 'y']



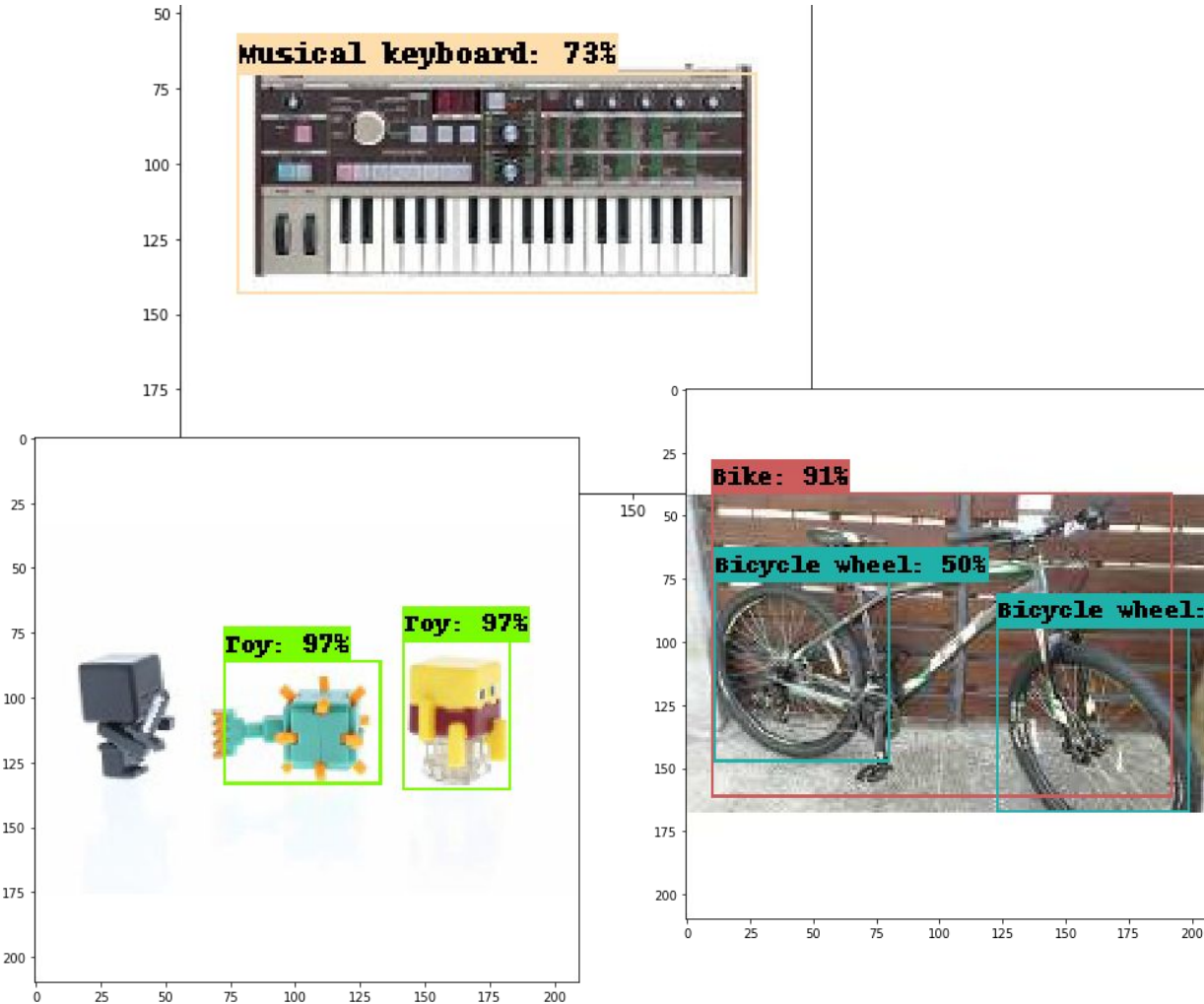
Recognized: ['224208895']



['rd', '9351288065']



# ➤ Logo detection, category automatic detection



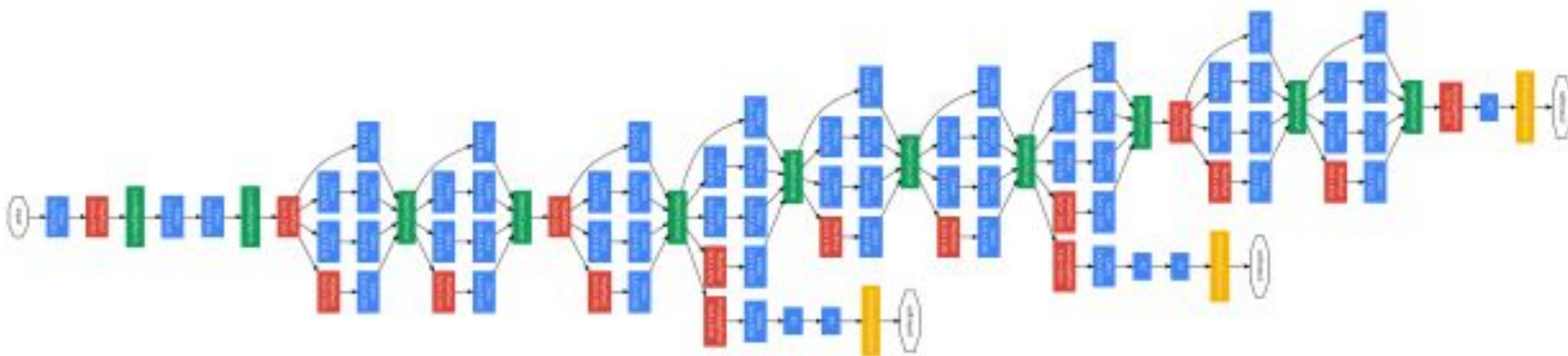
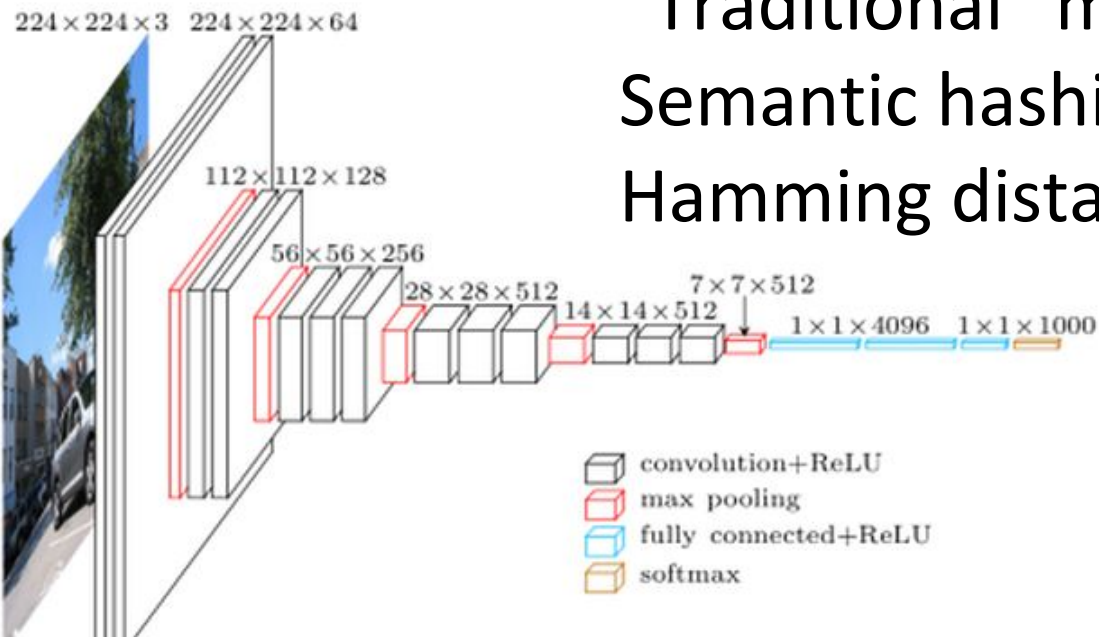
Score : 99.99998807907104



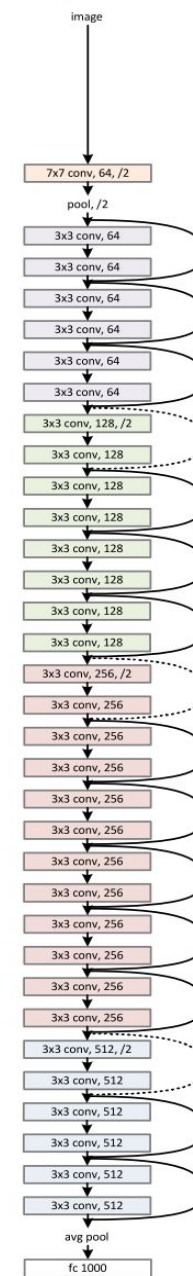
 **Tools of the trade**

## Summarizing

“Traditional” models are still alive  
Semantic hashing works  
Hamming distance enough (for now)



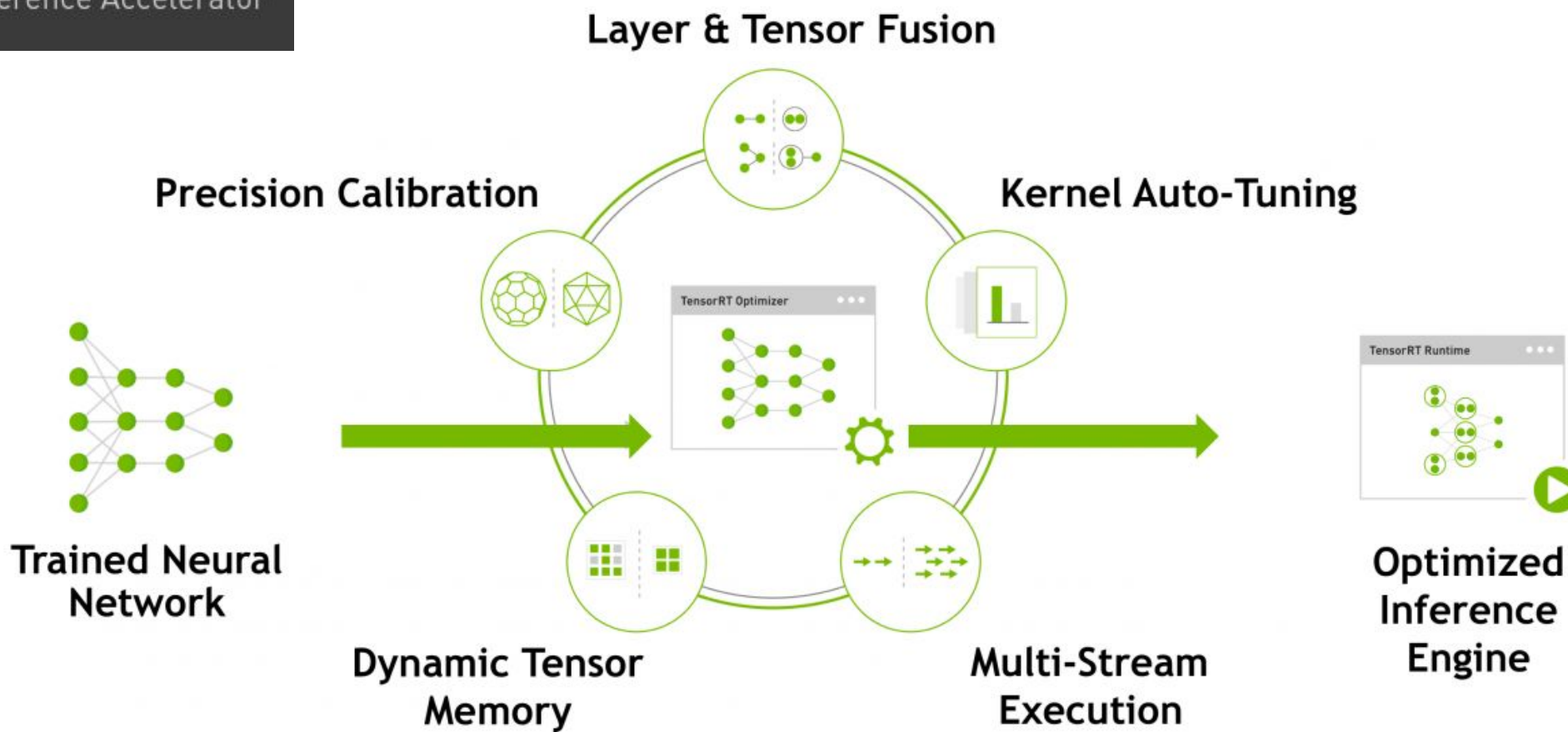
34-layer residual



- Very specialized attributes detectors
- Jump to segmentation from detection
- Vertical marketplaces adopting ML
- Also generalized marketplaces (Shopify?)

NVIDIA TensorRT

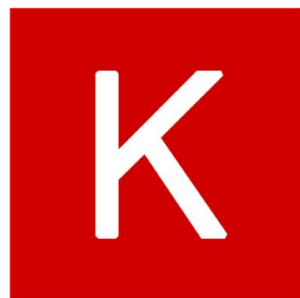
Programmable Inference Accelerator





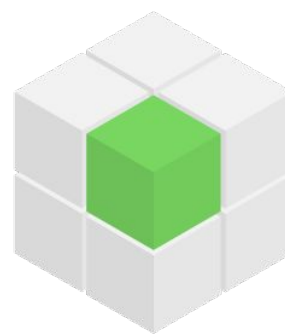
Caffe

TensorFlow™



Keras

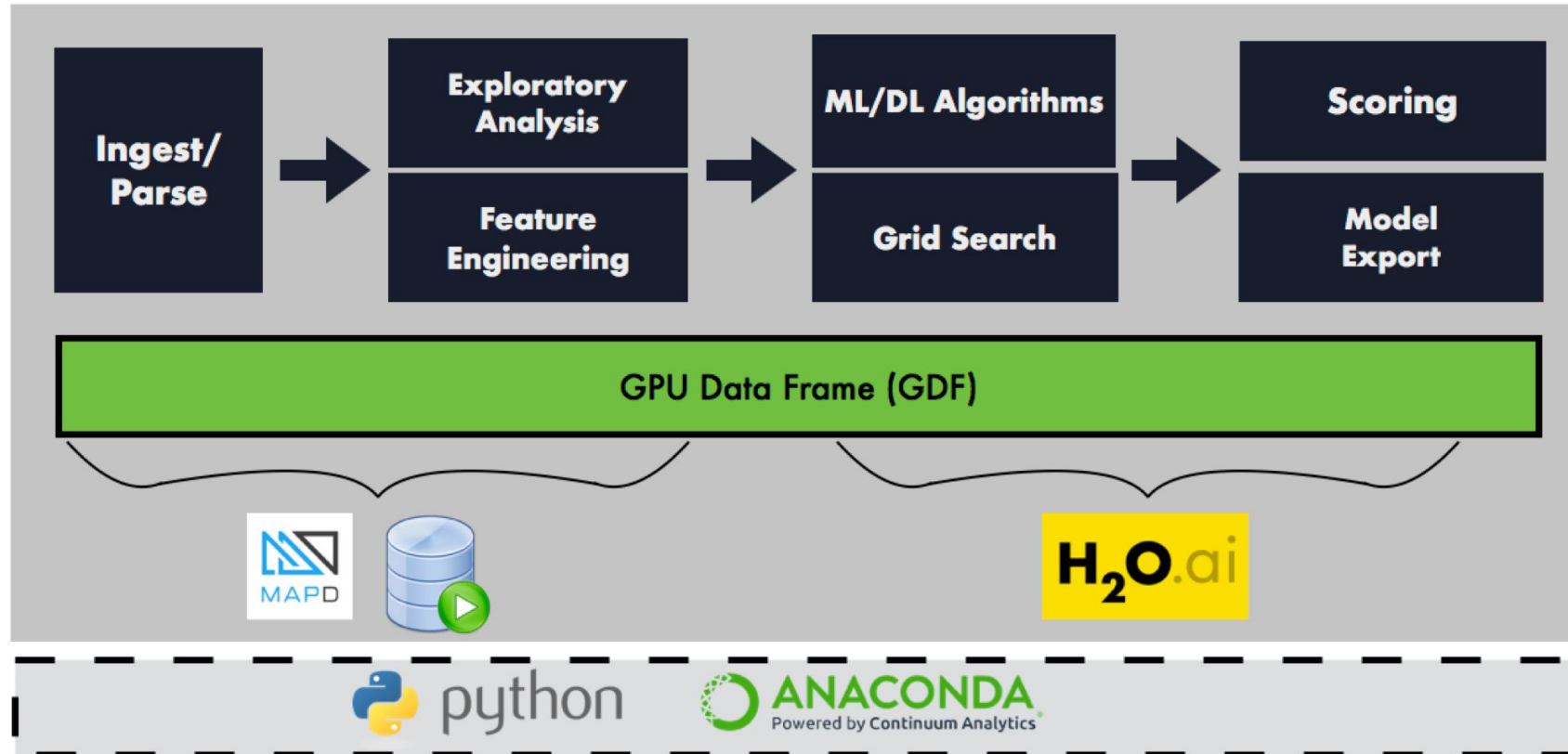
G



Ai

GPU OPEN ANALYTICS INITIATIVE

## > Full GPU Analytics stack



Questions?



**machinalis**  
Machine Learning Solutions Delivery