

# Video Analytics Survey

Author: Dmitry Khizbullin

[Intellivision](#)

# Areas of video analytics (VA) application

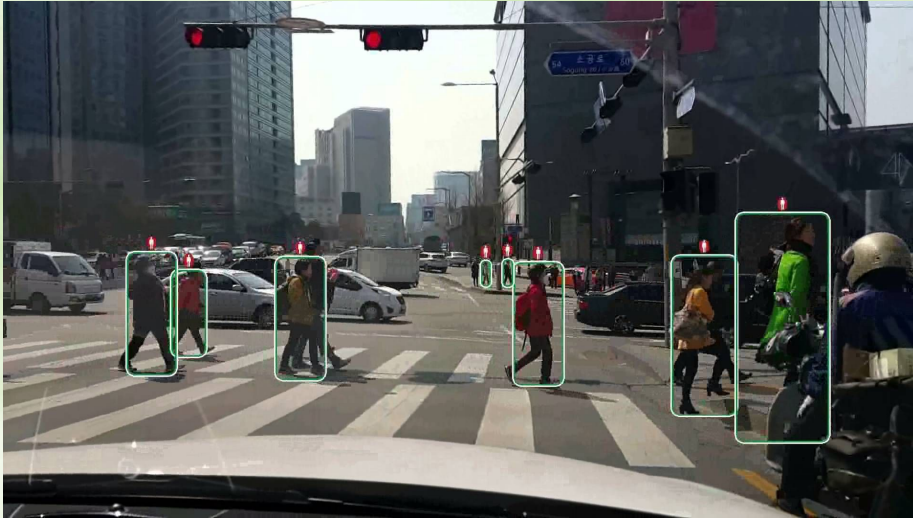
1. Surveillance / security
2. Retail (shops, stores, parks)
3. Automobile: driver assistance / autopilot
4. Mobile apps
5. Military
6. Medical

# Video analytics applications

1. Object detection (images): person/pedestrian, car, License Plate Detection, tumor, Frontal Collision Warning
2. Object recognition (images): face, tumor, age, gender, License Plate Recognition
3. Object tracking (video): single-camera, multi-camera, moving-camera, Lane Departure Warning
4. Special metrics/events extraction: [heartbeat](#), goods theft, people counting for retail, parking lot occupancy, crossroad occupancy, left/abandoned, slip/fall, fever

# Examples

Pedestrian detection



License plate detection/recognition



# Examples

Frontal collision warning



Lane departure warning

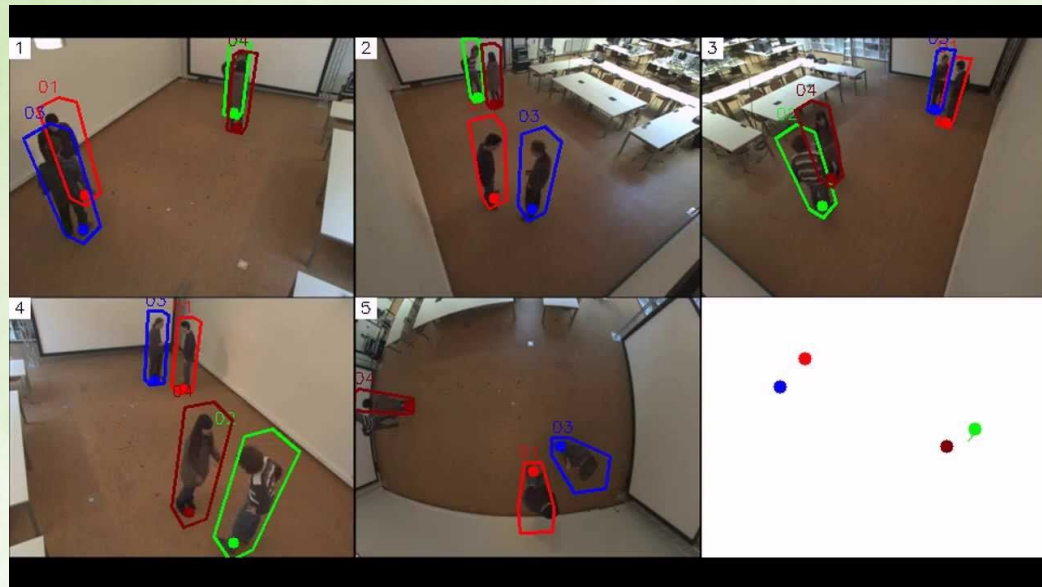


# Examples

Object tracking



Multicamera



# Examples

Parking lot occupancy



Photo: Migra Systems, Inc.

Point of service



# Examples

Abandoned object



Fever detection





# Image processing applications with ML

1. Image search, similar images, image annotation
2. Style transfer: Prisma
3. Face enhancement: FacePlus by itseez
4. Fun overlays: iSwap Faces, Funny Camera

# Eras of ML (my view)

1. Expert systems (past, few ML)
2. Single-knowledge-area systems (nowadays)
3. Future systems (multi-knowledge-area)

A DARPA Perspective on Artificial Intelligence

<https://www.youtube.com/watch?v=-O01G3tSYpU>

# Handcraft vs Machine Learning

1. Pure handcrafted solution
2. Hybrid handcrafted-ML solution
3. End-to-end ML solution

Pre-process handcrafted features	ML stage	Post-process handcrafted decision logic
--	----------	---

# Handcraft

Handcrafted features:

LBP, HOG, Optical flow, edges, frame difference

Handcrafted post-processing:

Spatial filtering (clusterization), temporal filtering

# Machine learning

1. Conventional: statistics, SVM, decision trees, cascades of classifiers
2. Deep learning: classic fully-connected, convolutional (CNN), full-convolutional (FCN), complex architectures (Google Inception, MS Residual/Resnet, Darknet/YOLO etc), LSTMs

# Image public datasets

1. [Imagenet](#) (1000 class, 1.2M images)
2. [MS COCO](#) (80 class, 120k images)
3. [Pascal VOC](#) (20-class, 25k images)

# Transfer techniques

1. **Transfer data** (use the same dataset in a different task)
2. **Transfer learning** (use the same learned model in a different task)
3. **Transfer intuition** (use your knowledge of what is feasible, what is the maximal possible accuracy and what approach is the best)

# Points to care about

1. Source data: public & private
2. Runtime: offline / low-latency / real-time
3. How critical the accuracy is: advisory / statistically-accurate / never-fail
4. Servers for training models



# Legitimacy

	Public data	Private data
Allowed for commercial use	Very limited	Must collect by your own, usually small
Prohibited for commercial use	Rather big, take care to not disclose that you use it	Big, but take care to not disclose occasionally

Thanks for your attention

Questions?