Machine Learning for Video Tracking of Passengers and Divested Objects at the Checkpoint

Henry Medeiros, Marquette University

henry.medeiros@marquette.edu

OPUS College of Engineering
MARQUETTE UNIVERSITY

ALERT
AWARENESS AND LOCALIZATION OF EXPLOSIVES-RELATED THREATS

This material is based upon work supported by the U.S. Department of Homeland Security, Science and Technology Directorate, under Grant Award 70RSAT18FR0000141. . The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security. [12/2013]

So What? Who Cares?

- **Space:** Video-based monitoring of passengers and divested items in airports.
  - Correlating Luggage and Specific Passengers (CLASP) Program, ALERT & DHS S&T Apex Screening at Speed Program.

- **Problem:** Machine learning algorithms are effective but unpredictable

- **Solution:** Uncertainty-aware models

- **Results:** Task-specific improvements
  - e.g., segmentation +4% accuracy,
  - gaze estimation +20% accuracy

- **TRL:** 4/5

- **Contact me**

henry.medeiros@marquette.edu
414-288-6186

What is the CLASP program trying to do?

Automatically identify within the airport security checkpoint:

- The location of passengers and their divested items
- When items are left behind at the checkpoint
- Theft events at the checkpoint
- Bottlenecks within the security checkpoint operation

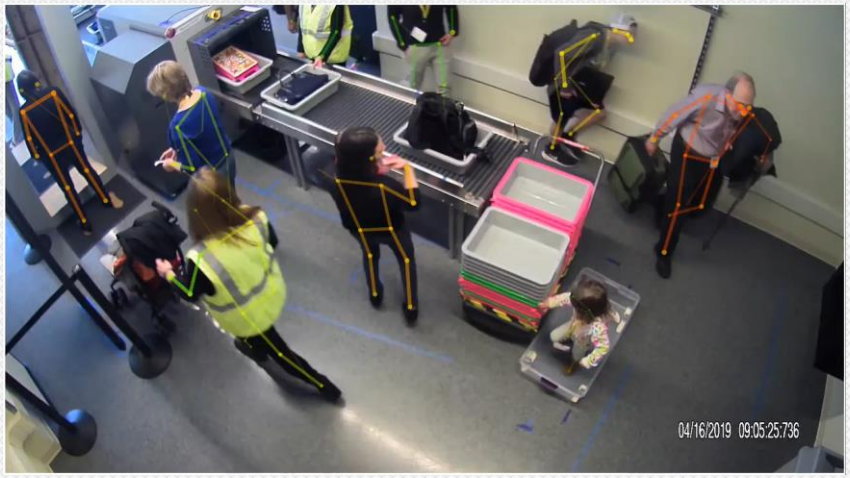
Making the TSA's Job Easier

By maturing algorithms developed under CLASP1, the team strives to make the passenger-baggage tracking capability sufficiently robust so it can **support operational pilots** and support Risk Based Screening in an airport environment.

The CLASP team hopes to **reduce the cognitive load** on Transportation Security Officers (TSOs) at the security checkpoint or command center and reduce operating costs by automating the detection of these events.

Applications of machine learning in CLASP

- Convolutional Pose Machine [1]

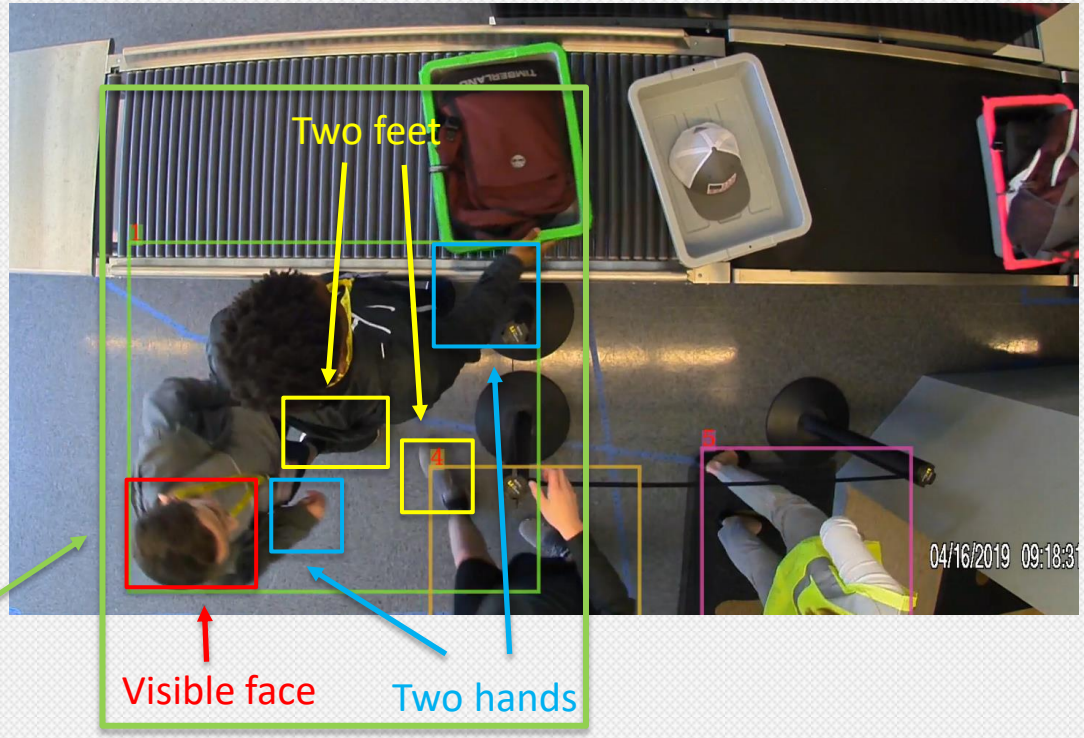


- Mask R-CNN [2]

[1] Shih-En, Ramakrishna, Kanade, Sheikh. "Convolutional pose machines." *IEEE Conference on Computer Vision and Pattern Recognition* 2016
[2] He, Gkioxari, Dollár, Girshick. "Mask R-CNN." *IEEE International Conference On Computer Vision* 2017

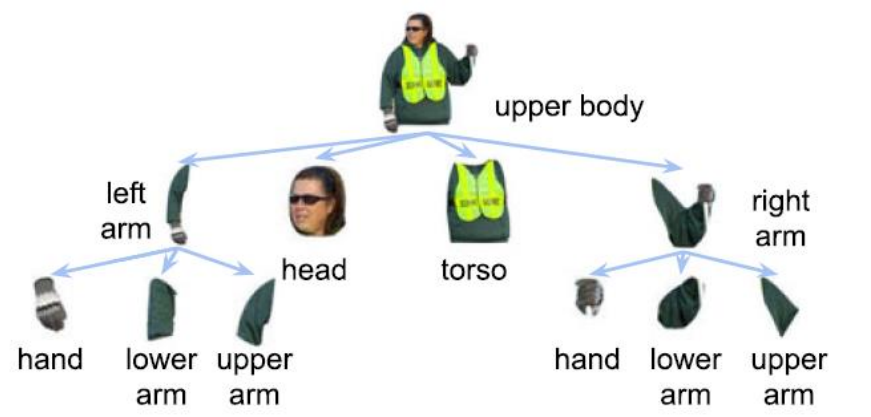
Failure cases of Mask R-CNN

- Compositional models help **explain** failure cases
- They do not help **detect** or **solve** them
- Need mechanisms to estimate network uncertainty



Interpretability of convolutional neural networks

- Image grammars are interpretable [3]



- And can help understand the behavior of CNNs [4]

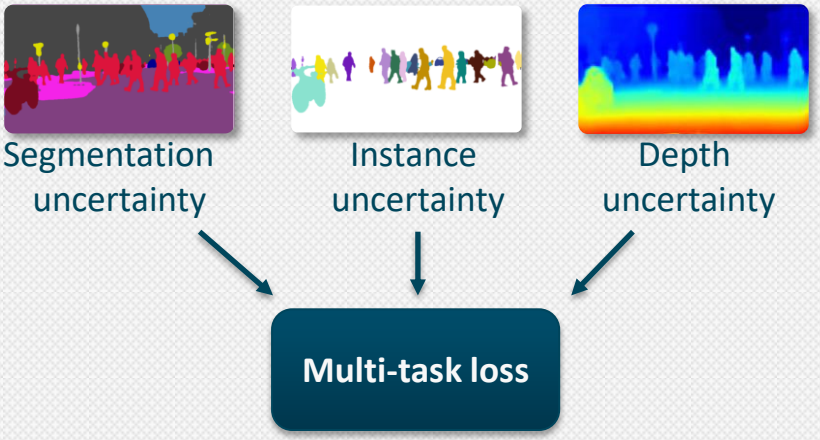
[3] Park, Nie, Zhu "Attribute and-or grammar for joint parsing of human pose, parts and attributes"
IEEE Transactions on Pattern Analysis and Machine Intelligence 2017
[4] Wu, Sun, Li, Song, Li. "Towards Interpretable R-CNN by Unfolding Latent Structures" *arXiv preprint* 2017

One solution to two problems: Uncertainty-aware models

- **Adaptability:** self-supervised, semi-supervised learning, active, reinforcement learning

- 99.3% accuracy in MNIST using 50 labeled samples per class [7]

- **Robustness:** error estimates, outlier detection, resampling, data fusion



- E.g., multi-task learning improves the performance of individuals tasks [6]

[6] Kendall, Gal, Cipolla "Multi-task learning using uncertainty to weigh losses for scene geometry and semantics" *IEEE Conf. on Computer Vision and Pattern Recognition* 2018
[7] Ribeiro, Calivá, Swainson, Gudmundsson, Leontidis, Kollias "Deep Bayesian Self-Training" *Neural Computing and Applications* 2019

Challenges

- Choosing the network architecture
  - Network size, multi-scale processing, fully convolutional, specialization
  - **Sequential models**
  - Effective prior distributions and likelihood models
- Bootstrapping the models
  - Transfer learning strategies
  - **Manual annotation**
- Evaluating the accuracy of the estimated uncertainties
  - How to combine task-specific variances into useful model information
- Reducing computational costs

Thank You

henry.medeiros@marquette.edu

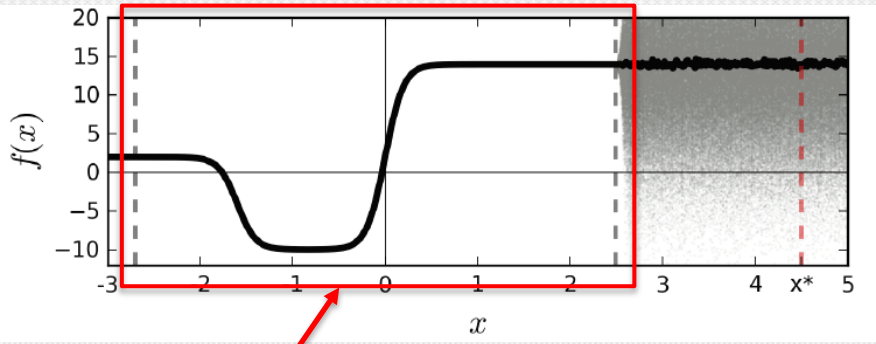
OPUS College of Engineering
MARQUETTE UNIVERSITY

ALERT
AWARENESS AND LOCALIZATION OF EXPLOSIVES-RELATED THREATS

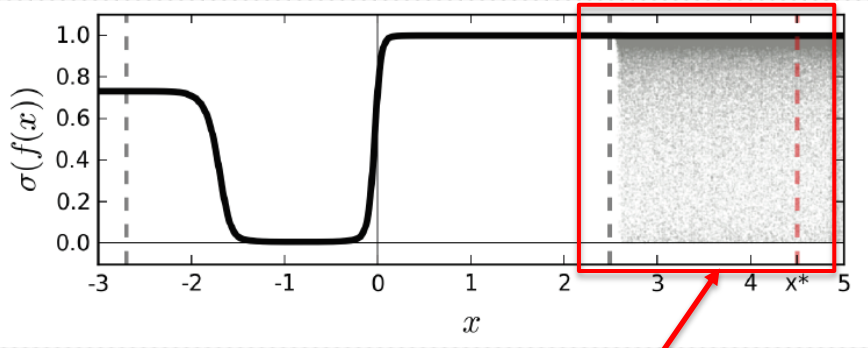
This material is based upon work supported by the U.S. Department of Homeland Security, Science and Technology Directorate, under Grant Award 70RSAT18FR0000141. . The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the U.S. Department of Homeland Security. [12/2013]

Neural networks don't know what they don't know

- **Network predictions are not confidence estimates** [5]



Training data

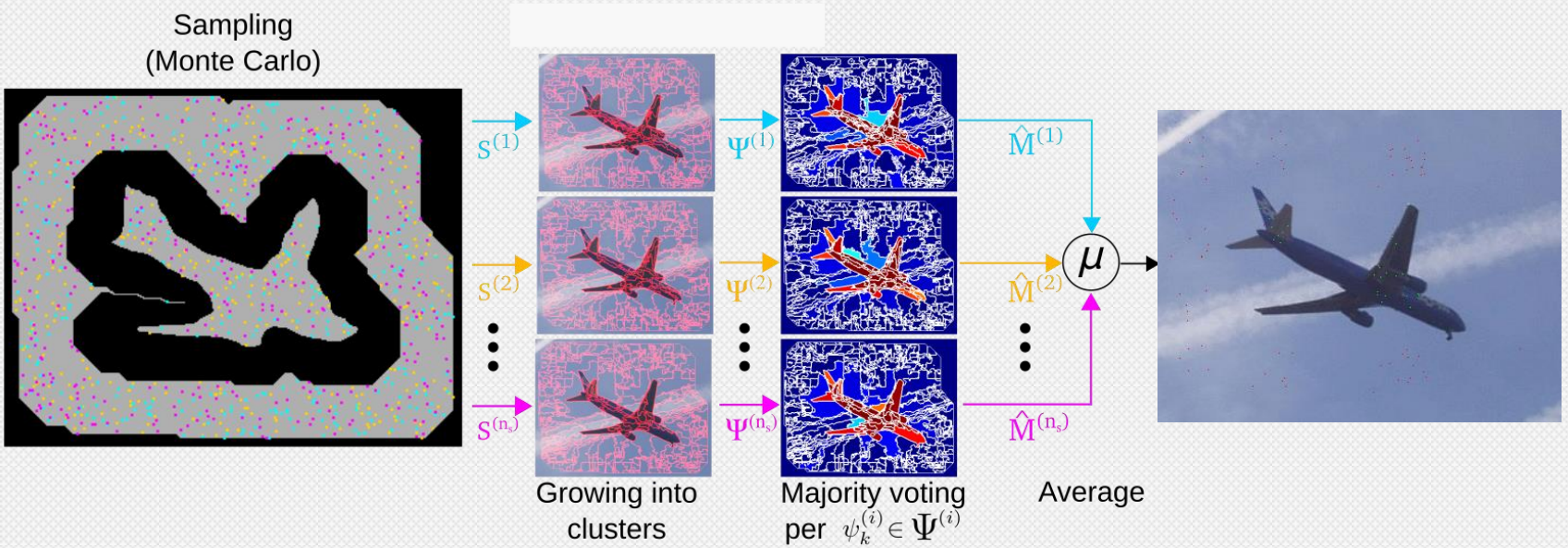


Predictions with unjustified high confidence

[5] Gal, Ghahramani "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning" *International Conference on Machine Learning* 2016

Estimating uncertainty with Monte Carlo sampling

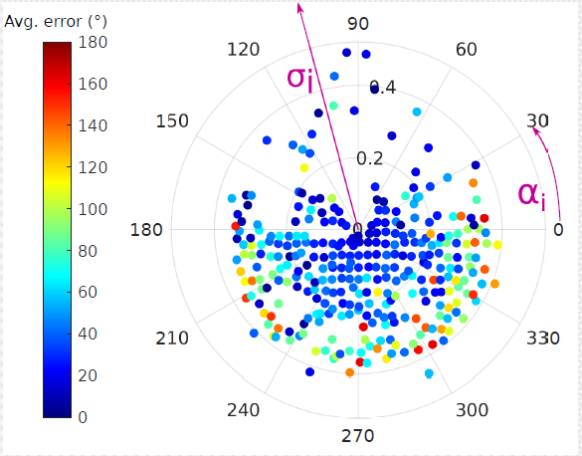
- Use multiple predictions to correct high-confidence mistakes
- E.g., semantic segmentation refinement [8]



[8] Dias, Medeiros "Semantic segmentation refinement by Monte Carlo region growing of high confidence detections" *Asian Conference on Computer Vision* 2018

Uncertainty learning

- Train the network to explicitly predict the uncertainty
- E.g., uncertainty-aware gaze estimation [9]



[9] Dias, Malafronte, Medeiros, Odone "Gaze Estimation for Assisted Living Environments" *IEEE Winter Conference on Applications of Computer Vision* 2019