

RESEARCH REVIEW 2022

Carnegie
Mellon
University
Software
Engineering
Institute

Knowing When You Don't Know:

Quantifying and Reasoning about Uncertainty in Machine Learning Models

NOVEMBER 14, 2022

Eric Heim
Senior Machine Learning Scientist, AI Division

©2022

[DISTRIBUTION STATEMENT A] Approved for public release and unlimited distribution.

Document Markings



Copyright 2022 Carnegie Mellon University.

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8702-15-D-0002 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

The view, opinions, and/or findings contained in this material are those of the author(s) and should not be construed as an official Government position, policy, or decision, unless designated by other documentation.

NO WARRANTY. THIS CARNEGIE MELLON UNIVERSITY AND SOFTWARE ENGINEERING INSTITUTE MATERIAL IS FURNISHED ON AN "AS-IS" BASIS. CARNEGIE MELLON UNIVERSITY MAKES NO WARRANTIES OF ANY KIND, EITHER EXPRESSED OR IMPLIED, AS TO ANY MATTER INCLUDING, BUT NOT LIMITED TO, WARRANTY OF FITNESS FOR PURPOSE OR MERCHANTABILITY, EXCLUSIVITY, OR RESULTS OBTAINED FROM USE OF THE MATERIAL. CARNEGIE MELLON UNIVERSITY DOES NOT MAKE ANY WARRANTY OF ANY KIND WITH RESPECT TO FREEDOM FROM PATENT, TRADEMARK, OR COPYRIGHT INFRINGEMENT.

[DISTRIBUTION STATEMENT A] This material has been approved for public release and unlimited distribution. Please see Copyright notice for non-US Government use and distribution.

This material may be reproduced in its entirety, without modification, and freely distributed in written or electronic form without requesting formal permission. Permission is required for any other use. Requests for permission should be directed to the Software Engineering Institute at permission@sei.cmu.edu.

Carnegie Mellon® and CERT® are registered in the U.S. Patent and Trademark Office by Carnegie Mellon University.

DM22-0893

Team



Eric Heim
Senior ML Researcher
AI Division



John Kirchenbauer
(Former) Machine Learning Engineer
AI Division



Jon Helland
(Former) Machine Learning Researcher
AI Division



Jacob Oaks
(Former) Associate Developer
AI Division



Aarti Singh
Associate Professor
Machine Learning Department



Zachary Lipton
Assistant Professor
Machine Learning Department

Quantifying Uncertainty: A Key Component for **Informative** and Robust AI Systems



Image: South Carolina National Guard, 151st Signal Battalion

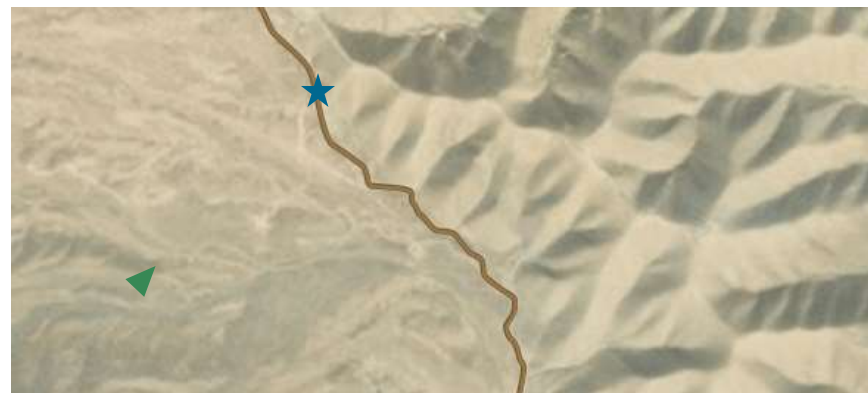
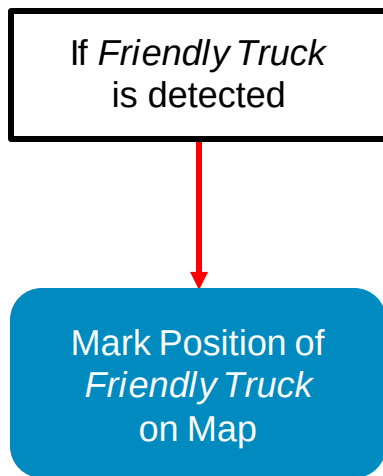
Quantifying Uncertainty: A Key Component for **Informative** and Robust AI Systems



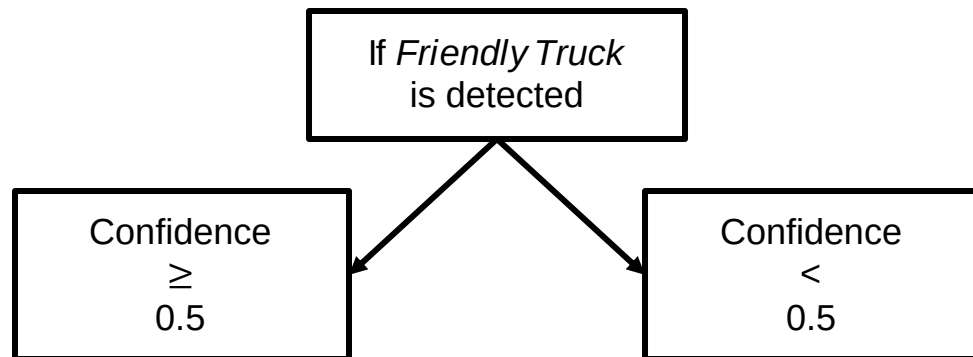
Image: South Carolina National Guard, 151st Signal Battalion

Accurate estimates of uncertainty can lead to better informed **decision making.**

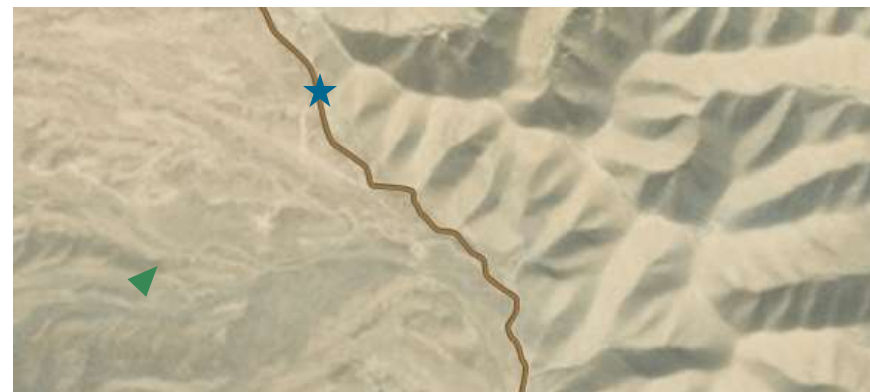
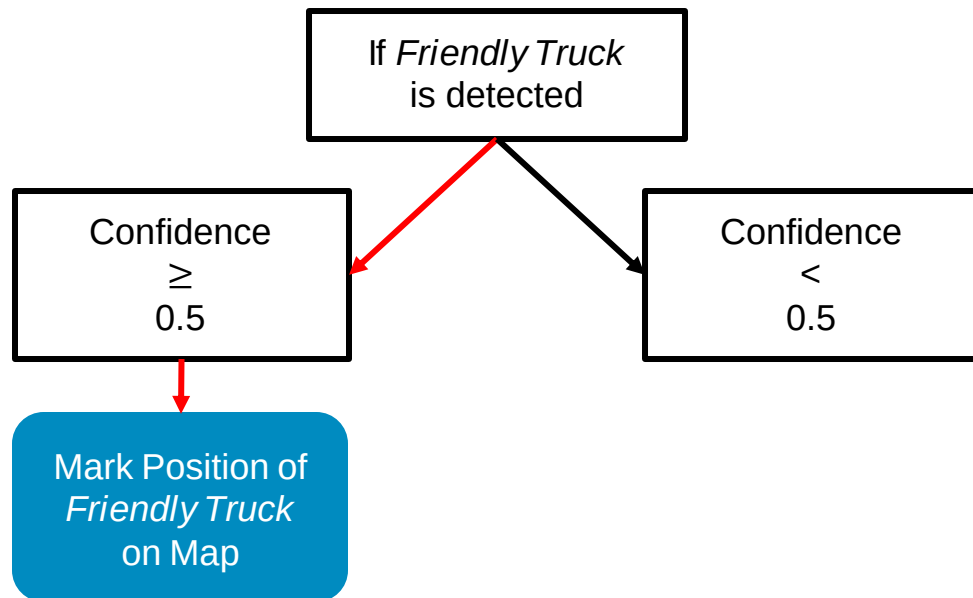
Quantifying Uncertainty: A Key Component for Informative and **Robust** AI Systems



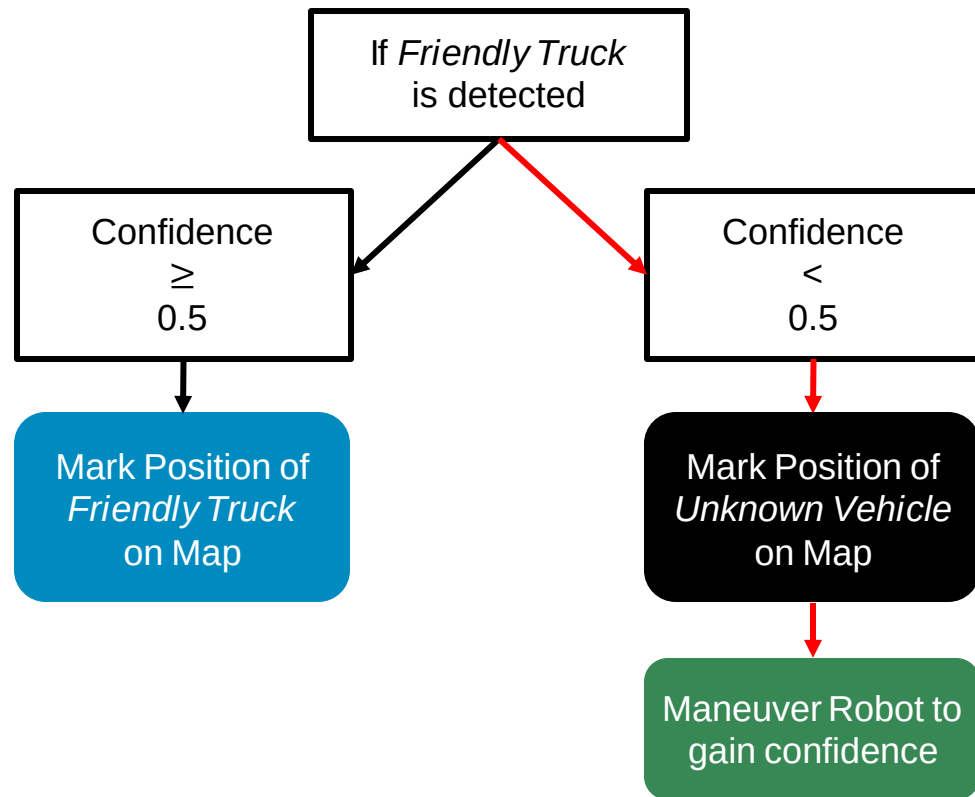
Quantifying Uncertainty: A Key Component for Informative and **Robust** AI Systems



Quantifying Uncertainty: A Key Component for Informative and **Robust** AI Systems



Quantifying Uncertainty: A Key Component for Informative and **Robust** AI Systems

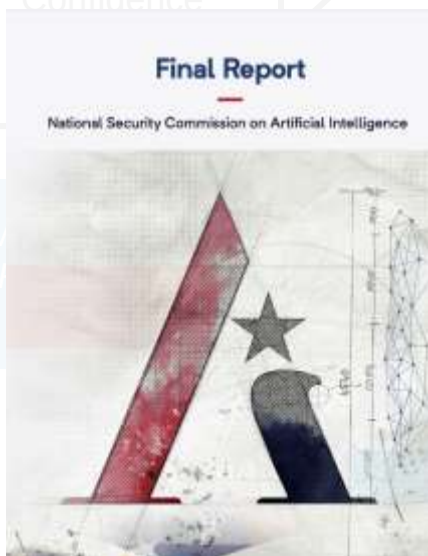


By allowing high-level reasoning to be informed by predictive uncertainty, AI systems can be **more robust** to failures caused by unconfident predictions.

Quantifying Uncertainty: A Key Component for Informative and **Robust** AI Systems

ML models that can accurately express their uncertainty...

1. Can better inform end users, leading to less opaque, more **trustable** AI Systems.
2. Be evaluated, debugged, improved upon, and built around in a more **robust** way.

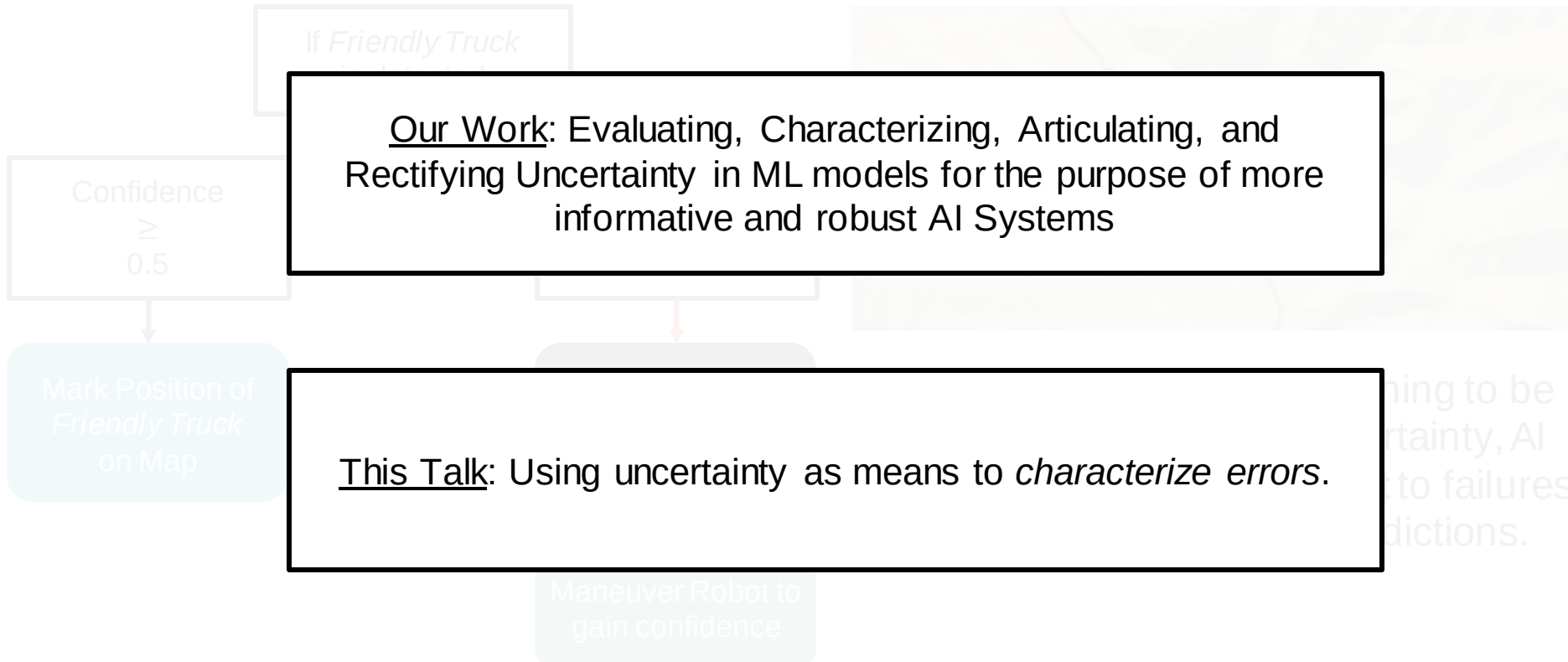


Frontiers of AI Technology.

The next decade of AI research will likely be defined by efforts to incorporate existing knowledge, push forward novel ways of learning, and make systems more **robust, generalizable, and trustworthy**.¹¹ Research on advancing human-machine teaming will be at the forefront, as will improvements in hybrid AI techniques, enhanced training methods, and explainable AI.

National Security Commission on Artificial Intelligence, Final Report

Quantifying Uncertainty: A Key Component for Informative and **Robust** AI Systems



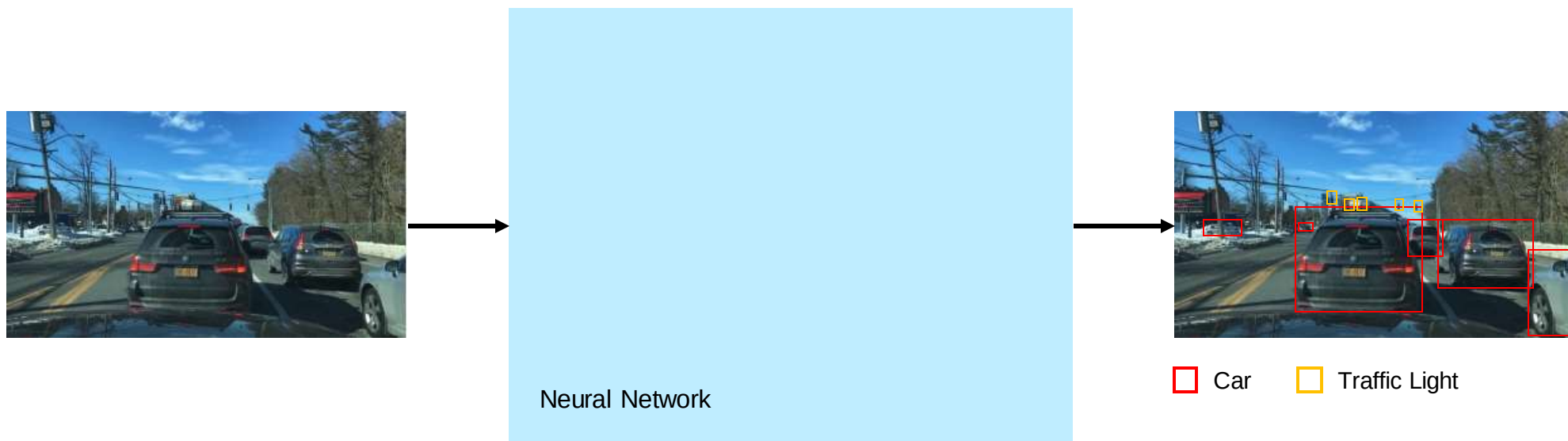
Our Work: Evaluating, Characterizing, Articulating, and Rectifying Uncertainty in ML models for the purpose of more informative and robust AI Systems

This Talk: Using uncertainty as means to *characterize errors*.

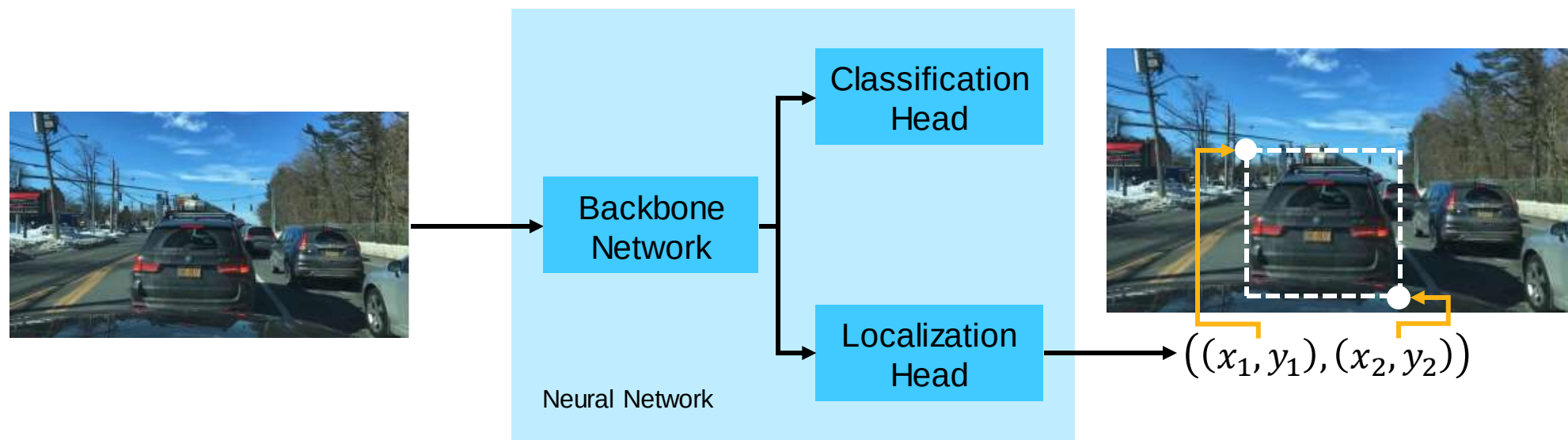
Introduction to Modern Object Detection

Object detection is really two tasks done in tandem:

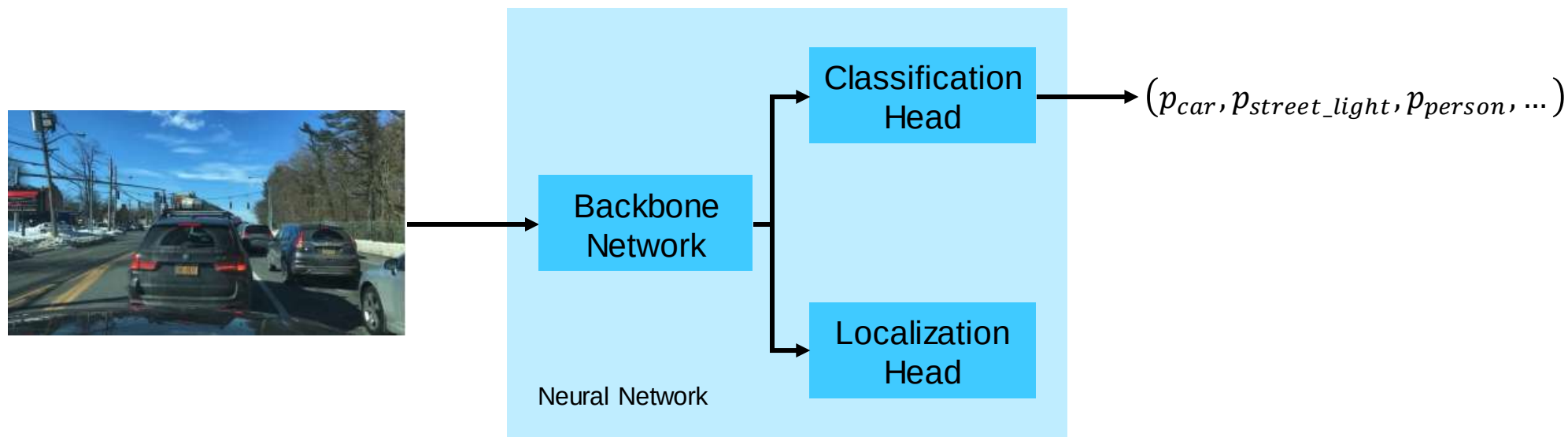
1. **Localization:** Identifying *where* in the image objects are
2. **Classification:** Identifying *what* those objects are



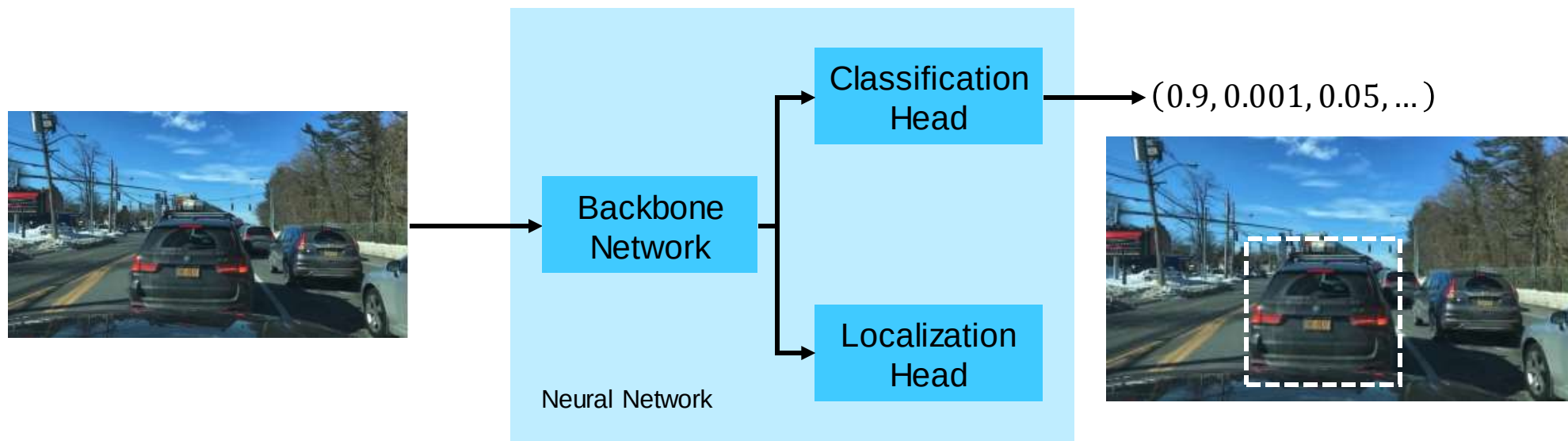
Introduction to Modern Object Detection



Introduction to Modern Object Detection

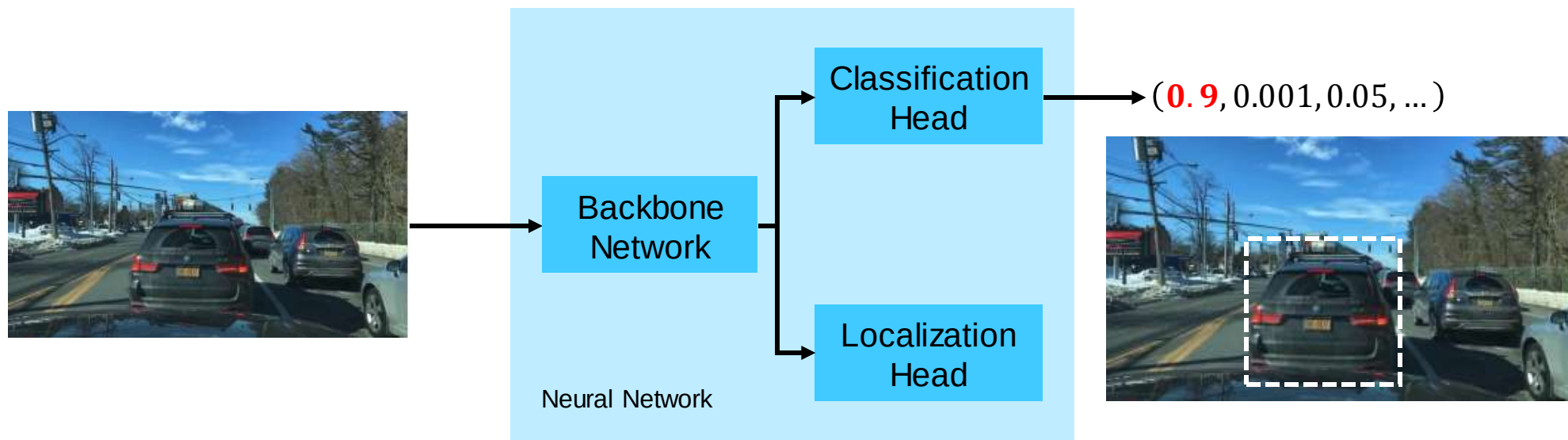


Introduction to Modern Object Detection



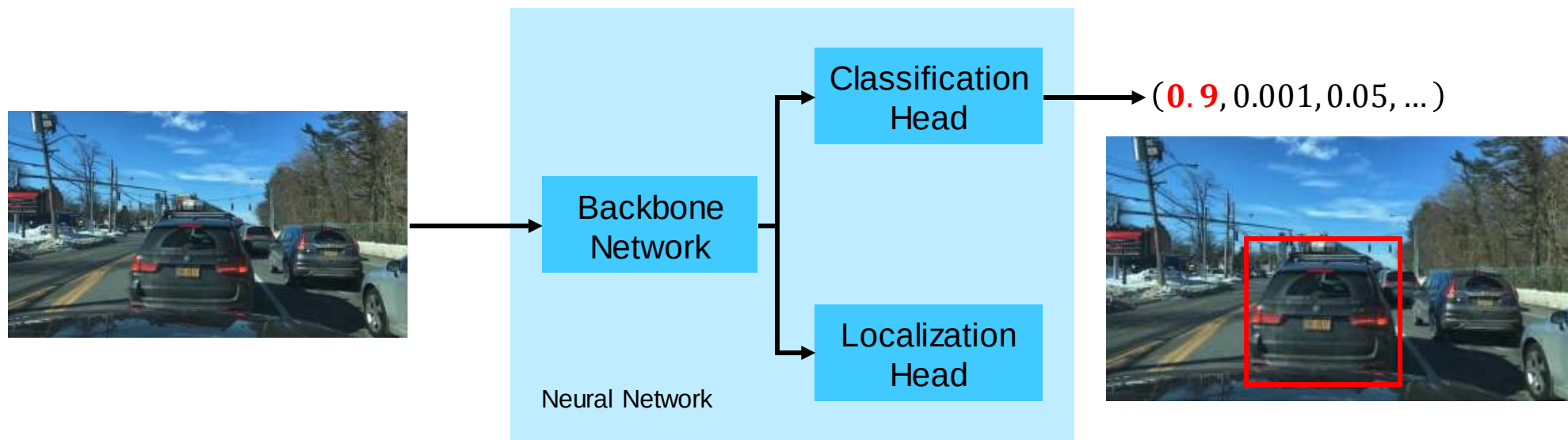
Introduction to Modern Object Detection

Maximum value corresponds
to “Car” class



Introduction to Modern Object Detection

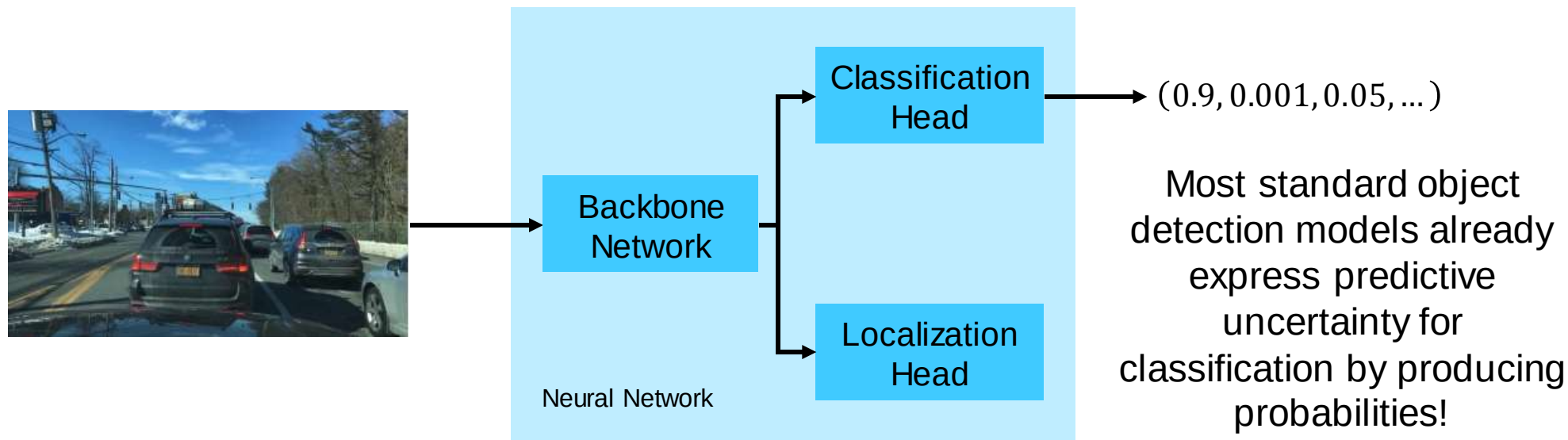
Box is assigned class“ Car”



Uncertainty in Object Detectors

Predictive Uncertainty – Uncertainty in the *output* of the model.

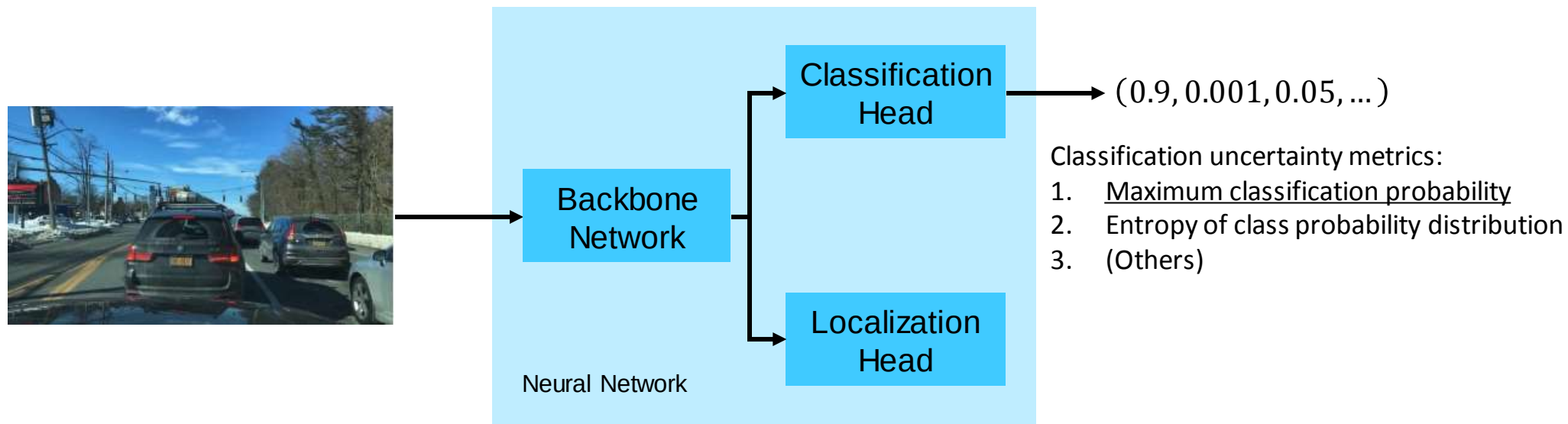
- A combination of *aleatoric* and *epistemic* uncertainty
 - Epistemic: Uncertainty in the parameters of the model. Can be reduced by training on more data
 - Aleatoric: Uncertainty caused by inherent noise in the data. Cannot be reduced by training on more data.
- Uncertainty can be expressed for both classification and localization



Uncertainty in Object Detectors

Predictive Uncertainty – Uncertainty in the *output* of the model.

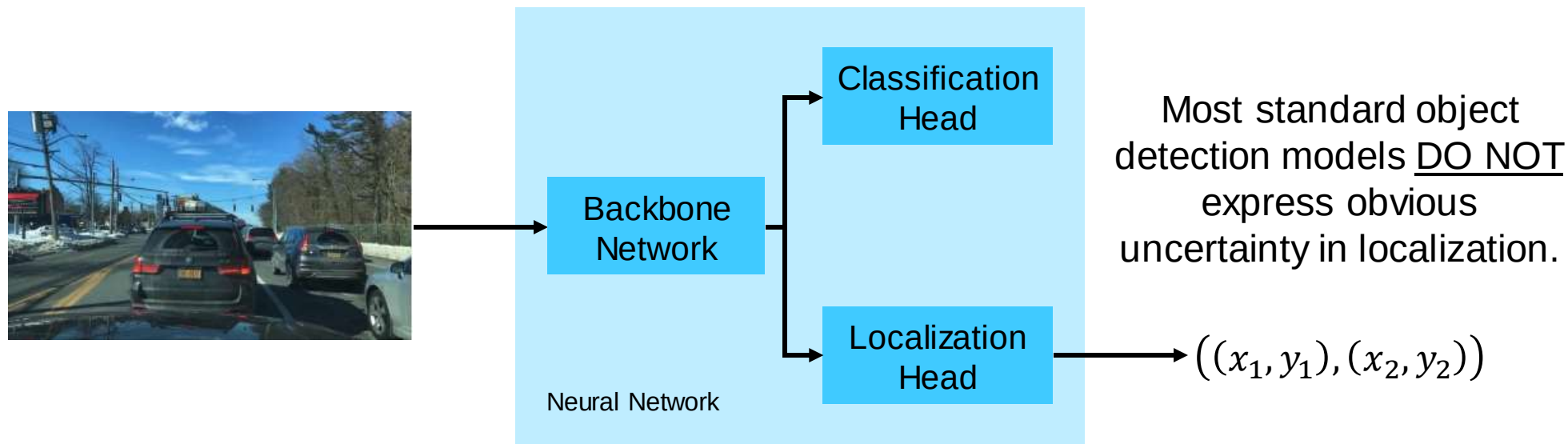
- A combination of *aleatoric* and *epistemic* uncertainty
 - Epistemic: Uncertainty in the parameters of the model. Can be reduced by training on more data
 - Aleatoric: Uncertainty caused by inherent noise in the data. Cannot be reduced by training on more data.
- Uncertainty can be expressed for both classification and localization



Uncertainty in Object Detectors

Predictive Uncertainty – Uncertainty in the *output* of the model.

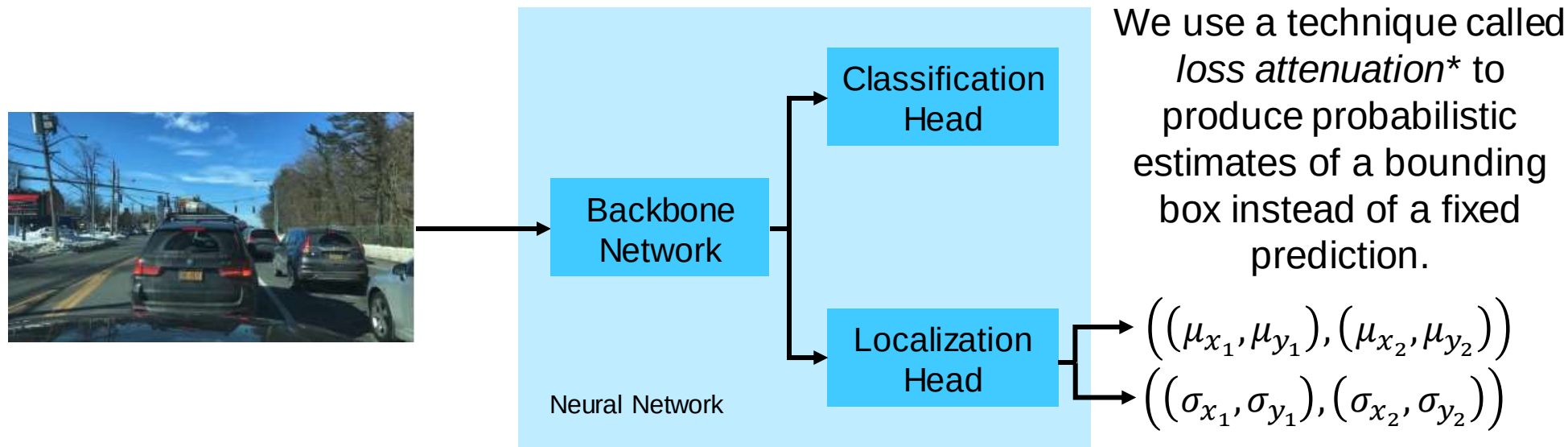
- A combination of *aleatoric* and *epistemic* uncertainty
 - Epistemic: Uncertainty in the parameters of the model. Can be reduced by training on more data
 - Aleatoric: Uncertainty caused by inherent noise in the data. Cannot be reduced by training on more data.
- Uncertainty can be expressed for both classification and localization



Uncertainty in Object Detectors

Predictive Uncertainty – Uncertainty in the *output* of the model.

- A combination of *aleatoric* and *epistemic* uncertainty
 - Epistemic: Uncertainty in the parameters of the model. Can be reduced by training on more data
 - Aleatoric: Uncertainty caused by inherent noise in the data. Cannot be reduced by training on more data.
- Uncertainty can be expressed for both classification and localization

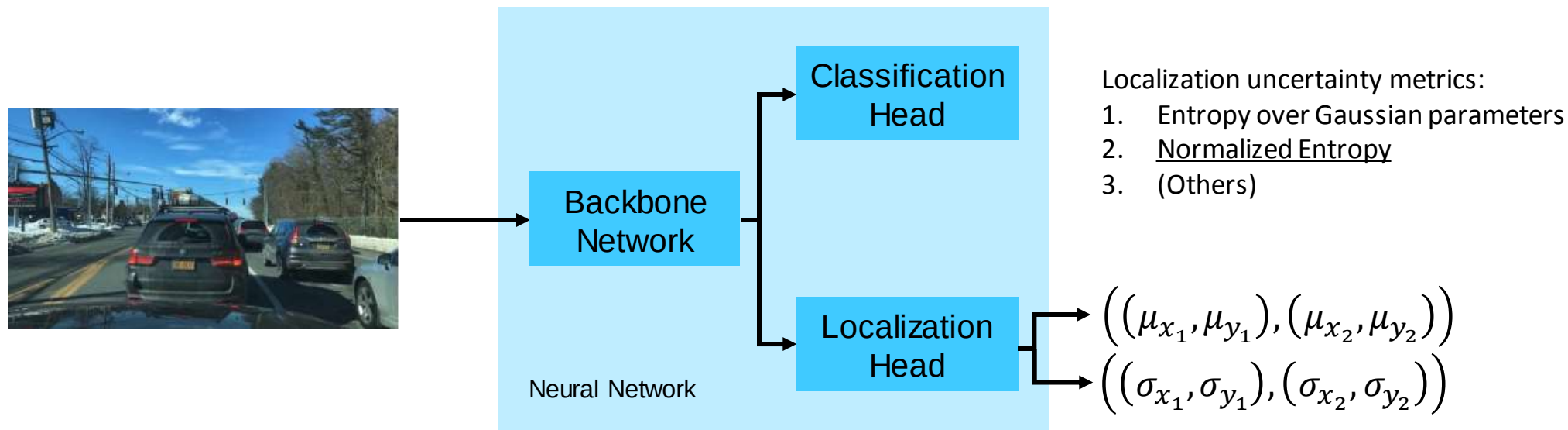


*Kendall, Alex, and Yarin Gal. "What uncertainties do we need in bayesian deep learning for computer vision?." *Advances in neural information processing systems*30 (2017).

Uncertainty in Object Detectors

Predictive Uncertainty – Uncertainty in the *output* of the model.

- A combination of *aleatoric* and *epistemic* uncertainty
 - Epistemic: Uncertainty in the parameters of the model. Can be reduced by training on more data
 - Aleatoric: Uncertainty caused by inherent noise in the data. Cannot be reduced by training on more data.
- Uncertainty can be expressed for both classification and localization

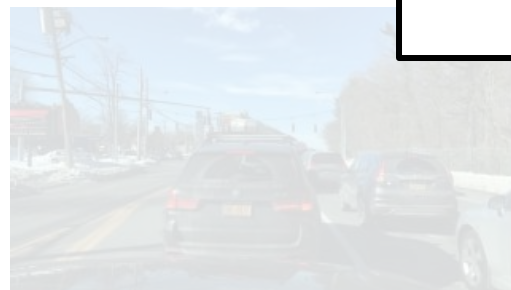


Uncertainty in Object Detectors

Predictive Uncertainty – Uncertainty in the *output* of the model.

- A combination of *aleatoric* and *epistemic* uncertainty
 - Epistemic: Uncertainty in the parameters of the model. Can be reduced by training on more data
 - Aleatoric: Uncertainty caused by inherent noise in the data. Cannot be reduced by training on more data.
- Uncertainty can

What happens when a detector is uncertain?



Backbone
Network

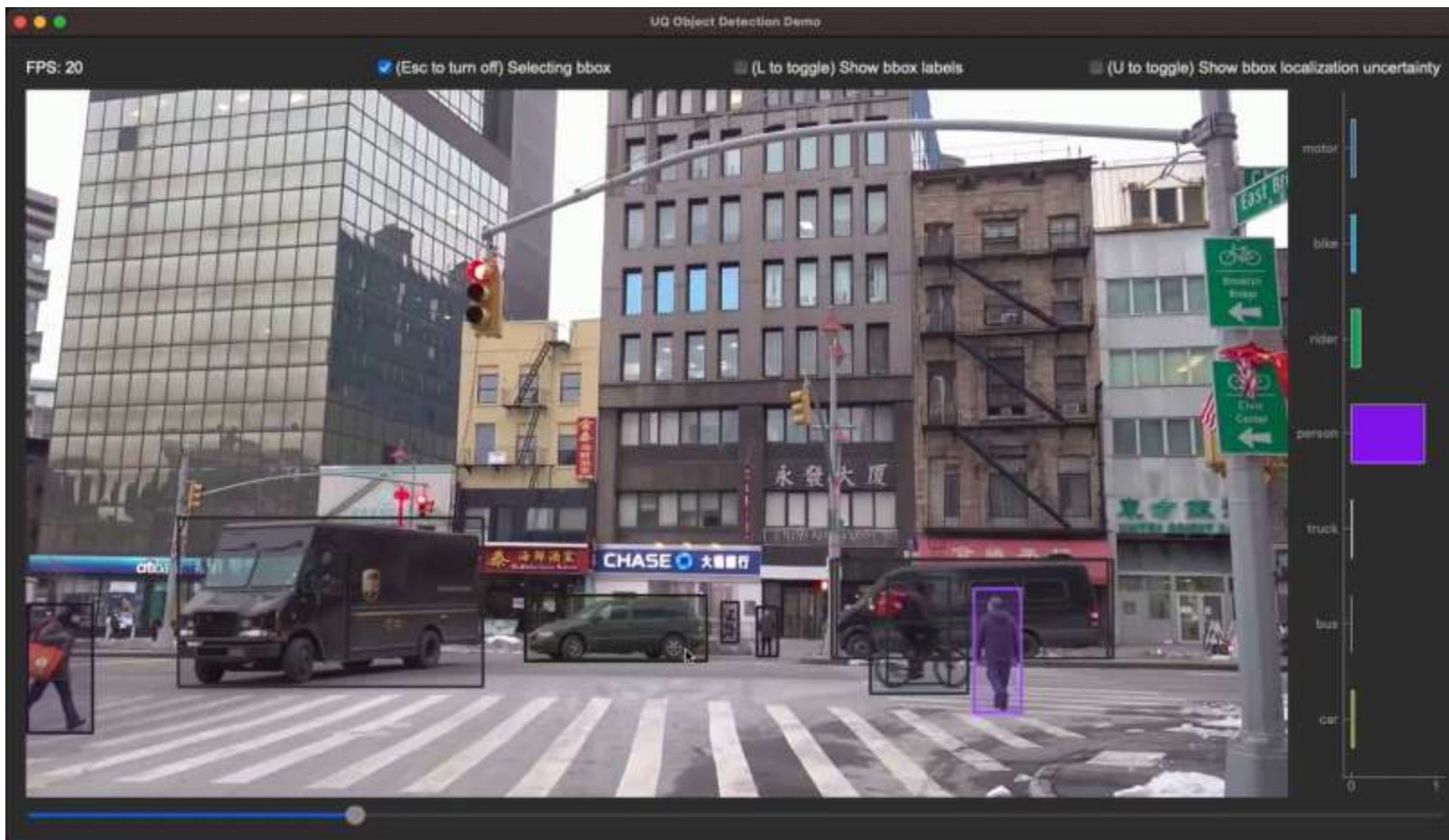
Localization
Head

Neural Network

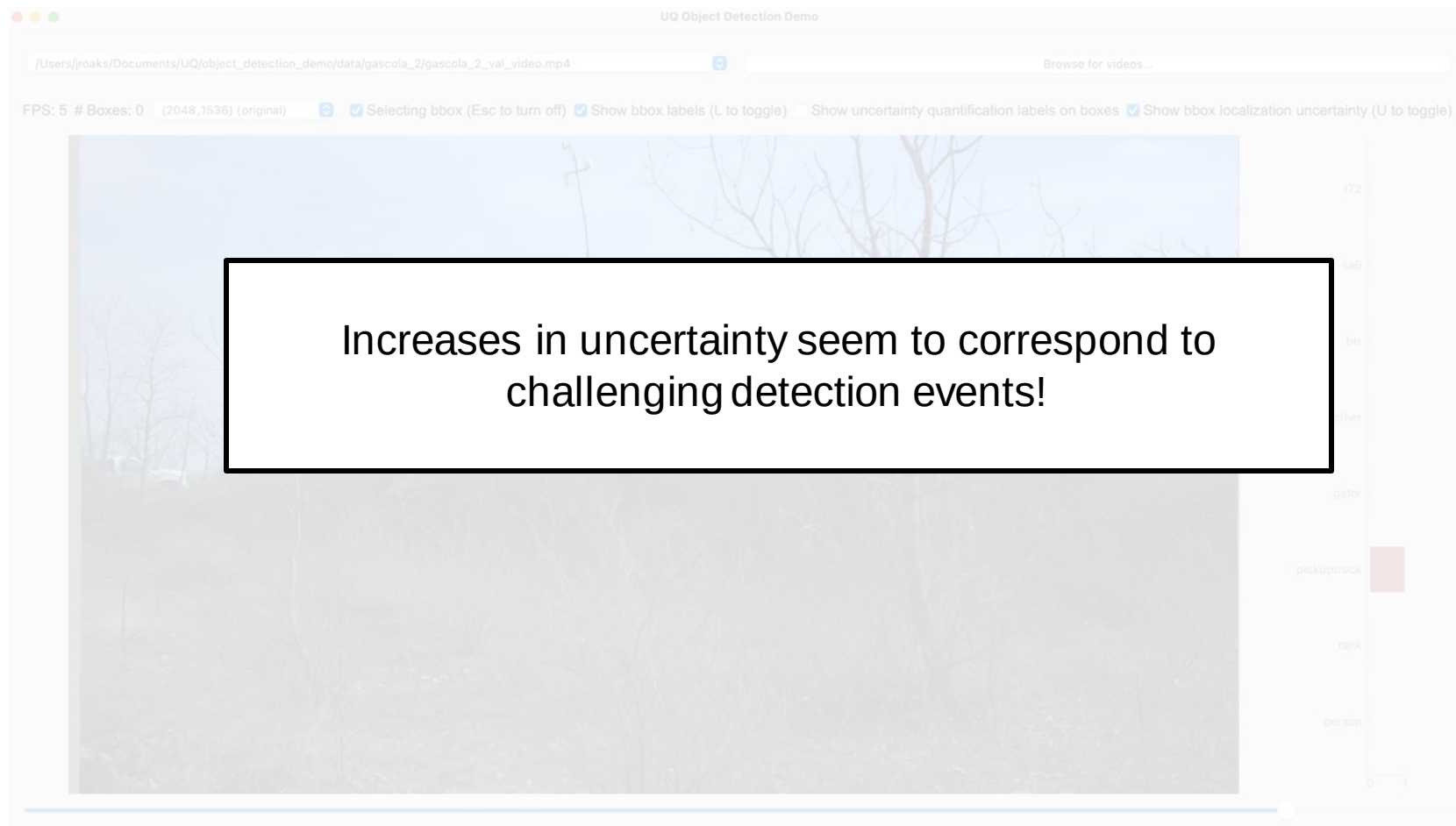
uncertainty metrics:
1. Gaussian parameters
2. Normalized Entropy
3. (Others)

$((\mu_{x_1}, \mu_{y_1}), (\mu_{x_2}, \mu_{y_2}))$
 $((\sigma_{x_1}, \sigma_{y_1}), (\sigma_{x_2}, \sigma_{y_2}))$

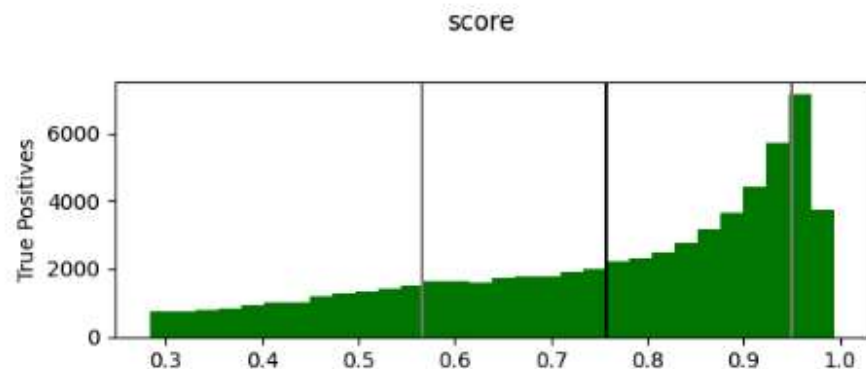
Probabilistic Object Detection Example – Overlapping Objects



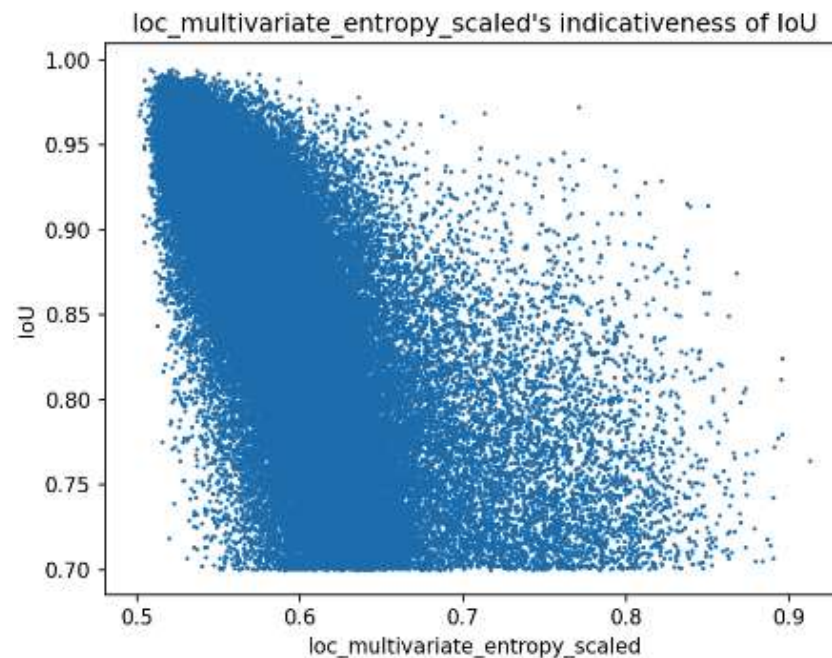
Probabilistic Object Detection Example - Occlusion



Preliminary Quantitative Results



Classification

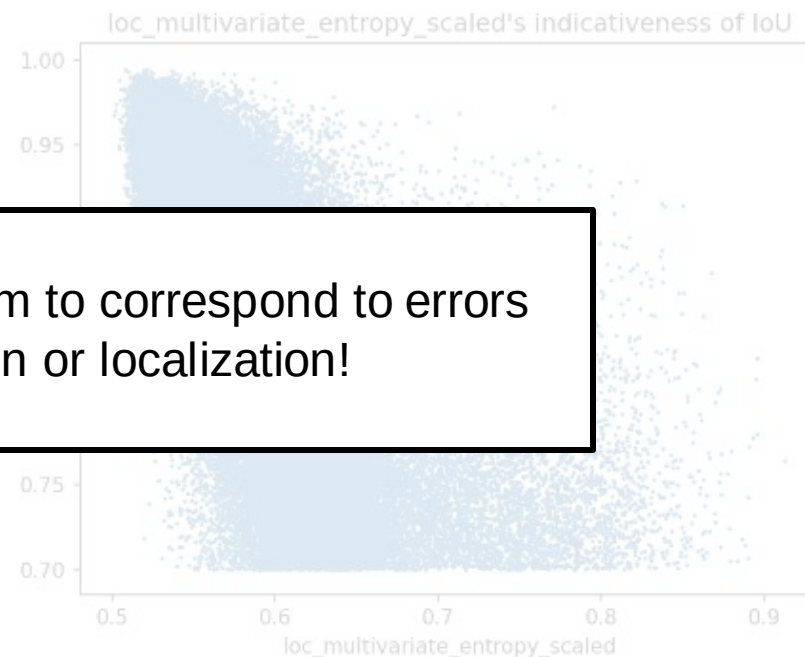


Localization

Preliminary Quantitative Results



Classification



Localization

Increases in uncertainty seem to correspond to errors in either classification or localization!

Bringing It All Together

Observations:

- Qualitative: Increase in detector uncertainty correspond to events.
- Quantitative: Increase in detector uncertainty correspond to errors.

Next Step: Using context and uncertainty values to characterize potential errors.

By using both we can not only predict *when* errors are likely, but *characterize the events that caused them*.

Events like: Occlusions, intersection of objects, objects leaving frame, duplication of predictions, etc.

Even without much context we can differentiate between errors in *localization* versus those in *classification*.

Practical Benefit: End users can reason about events that caused model errors.

Summary

Uncertainty can be a key component to more robust and trustworthy machine learning models.

We showed:

- How uncertainty can be quantified by modern object detectors
- Some qualitative results showing events causing the detector to be uncertain.
- Some preliminary quantitative results showing uncertainty corresponds to error.
- An outline of upcoming work combining the two to use uncertainty to detect and characterize errors in object detection models.

Other work in the project:

- Metrics for evaluating a model's ability to express uncertainty accurately (Kirchenbauer, Oaks, and Heim; 2022)
- Learning from limited sources of information (Garg et al; 2021)(Garg, Balakrishnan, and Lipton; 2022)
- Learning to detect when instances are “out of domain”.