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

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

Document Markings

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.

DM22-0893

Our Team

Carnegie Mellon University Software Engineering Institute

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

Carnegie Mellon University Software Engineering Institute

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

Image: South Carolina National Guard, 151st Signal Battalion

Carnegie Mellon University Software Engineering Institute

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

Carnegie Mellon University
Software Engineering Institute

If *Friendly Truck* is detected

Mark Position of *Friendly Truck* on Map

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

If *Friendly Truck* is detected

Confidence $\geq$ 0.5

Confidence $<$ 0.5

Carnegie Mellon University Software Engineering Institute

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

If *Friendly Truck* is detected

Confidence $\geq$ 0.5

Confidence < 0.5

Mark Position of *Friendly Truck* on Map

Carnegie Mellon University Software Engineering Institute

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

If *Friendly Truck* is detected

Confidence ≥ 0.5

Mark Position of *Friendly Truck* on Map

Confidence < 0.5

Mark Position of *Unknown Vehicle* on Map

Maneuver Robot to gain confidence

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

Carnegie Mellon University Software Engineering Institute

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

Carnegie Mellon University
Software Engineering Institute

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

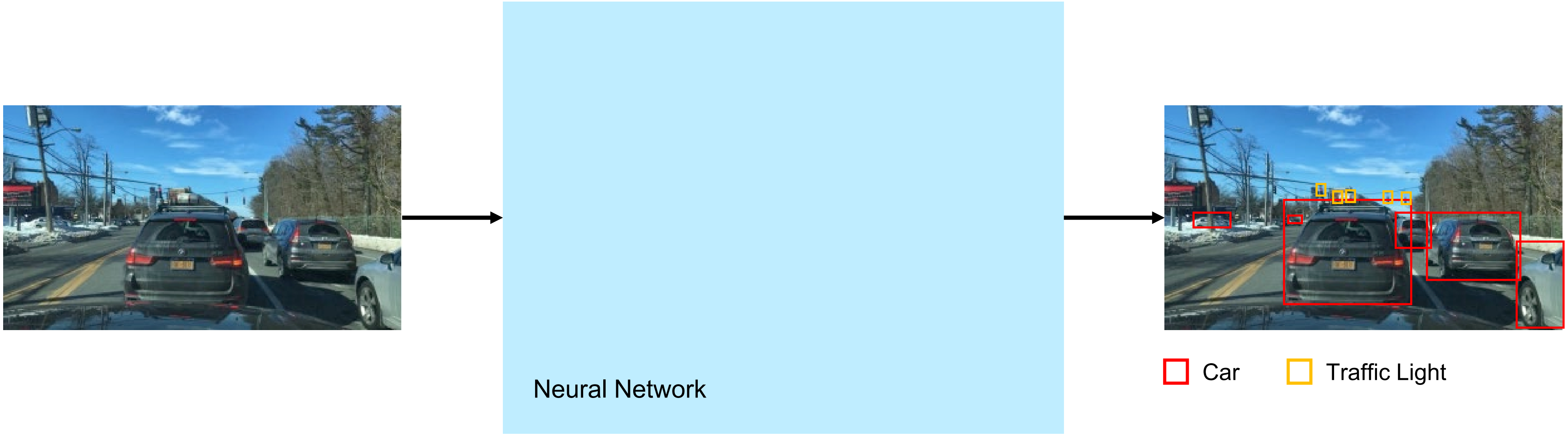
Our Work: Evaluating, Characterizing, Articulating, and Rectifying Uncertainty in ML models for the purpose of more informative and robust AI Systems.

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

Introduction to Modern Object Detection

Object detection is really two tasks done in tandem:

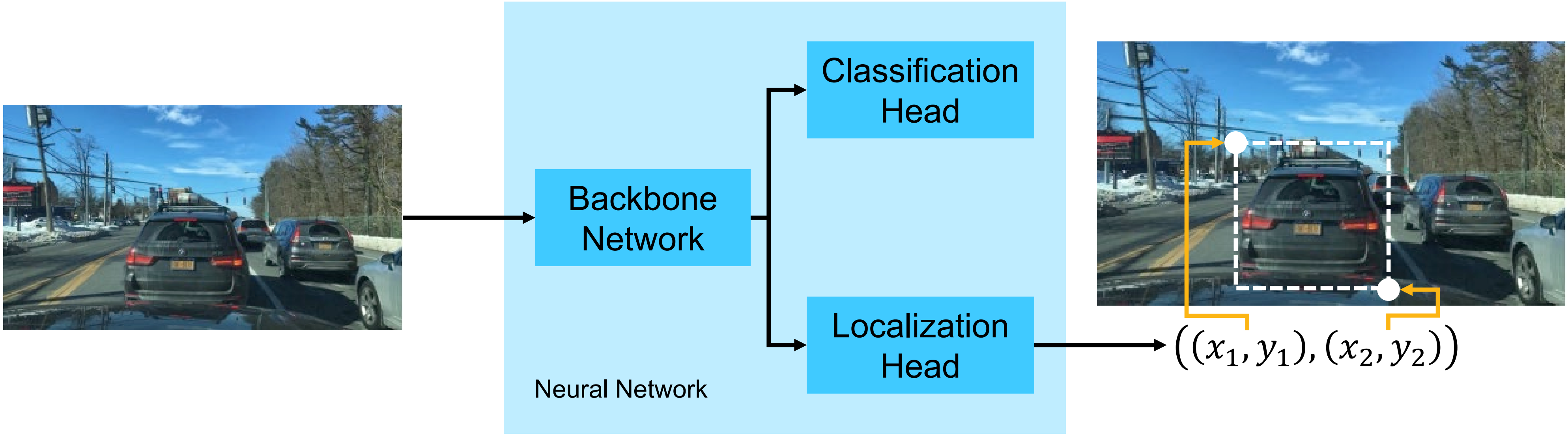
1. **Localization**: Identifying *where* in the image objects are
2. **Classification:** Identifying *what* those objects are



Introduction to Modern Object Detection

Object detection is really two tasks done in tandem:

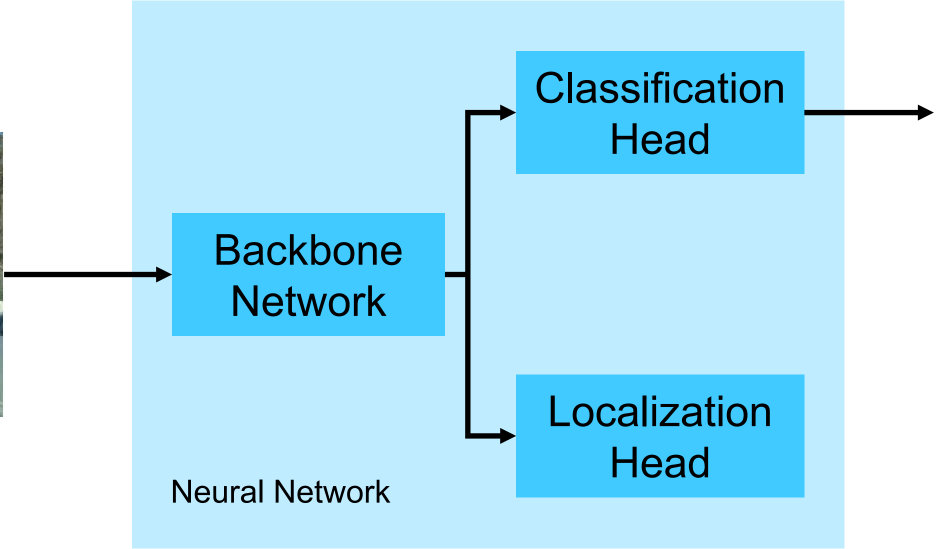
1. **Localization**: Identifying *where* in the image objects are
2. **Classification:** Identifying *what* those objects are



Introduction to Modern Object Detection

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

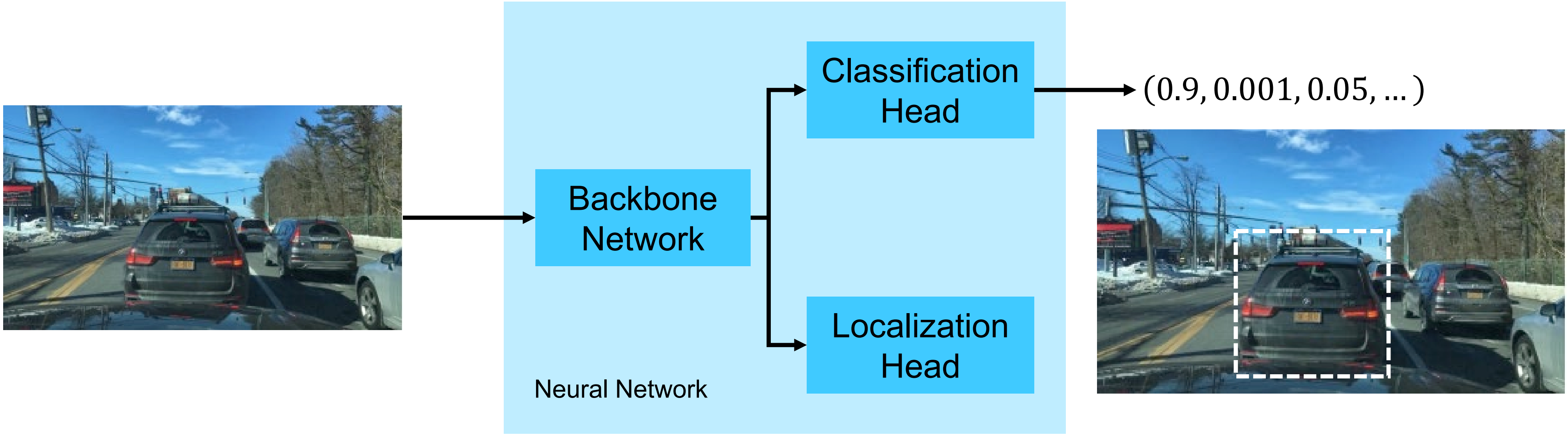


$$(p_{car}, p_{street\_light}, p_{person}, \dots)$$

Introduction to Modern Object Detection

Object detection is really two tasks done in tandem:

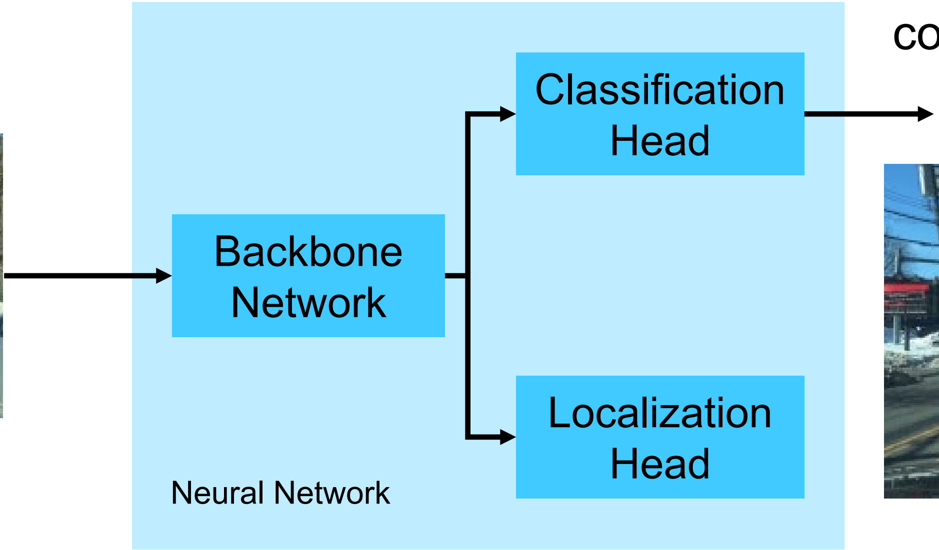
1. **Localization**: Identifying *where* in the image objects are
2. **Classification:** Identifying *what* those objects are



Introduction to Modern Object Detection

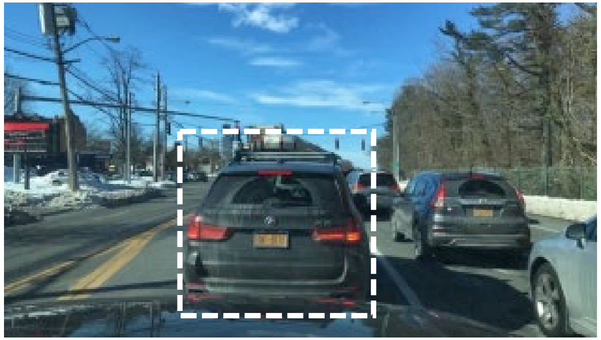
Object detection is really two tasks done in tandem:

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



Maximum value corresponds to “Car” class.

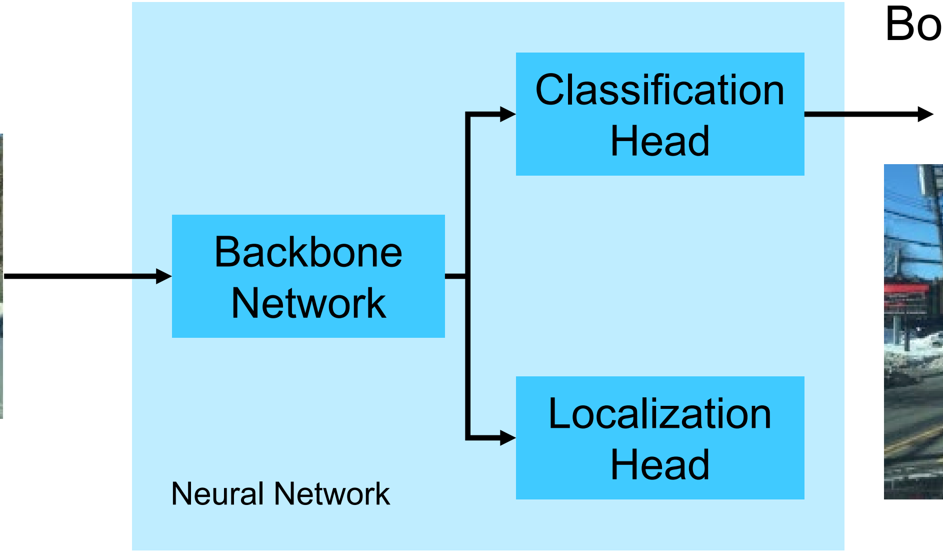
$(\mathbf{0.9}, 0.001, 0.05, \ldots)$



Introduction to Modern Object Detection

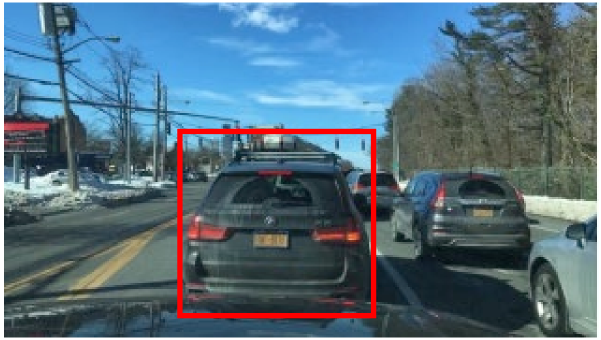
Object detection is really two tasks done in tandem:

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



Box is assigned class “Car”.

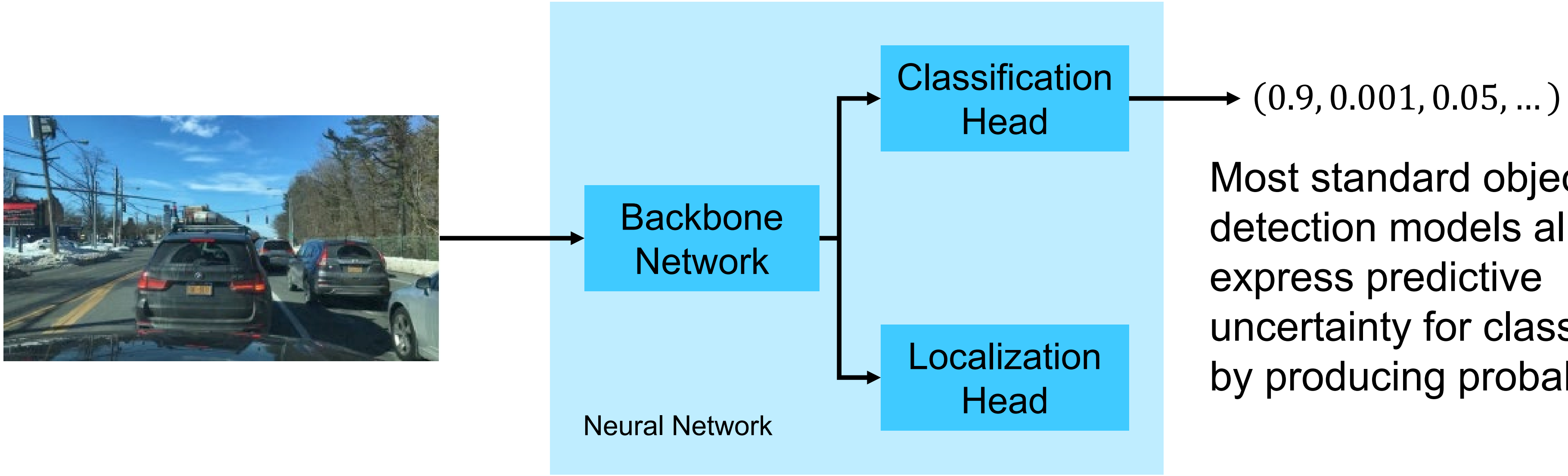
$(\mathbf{0.9}, 0.001, 0.05, \ldots)$



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.



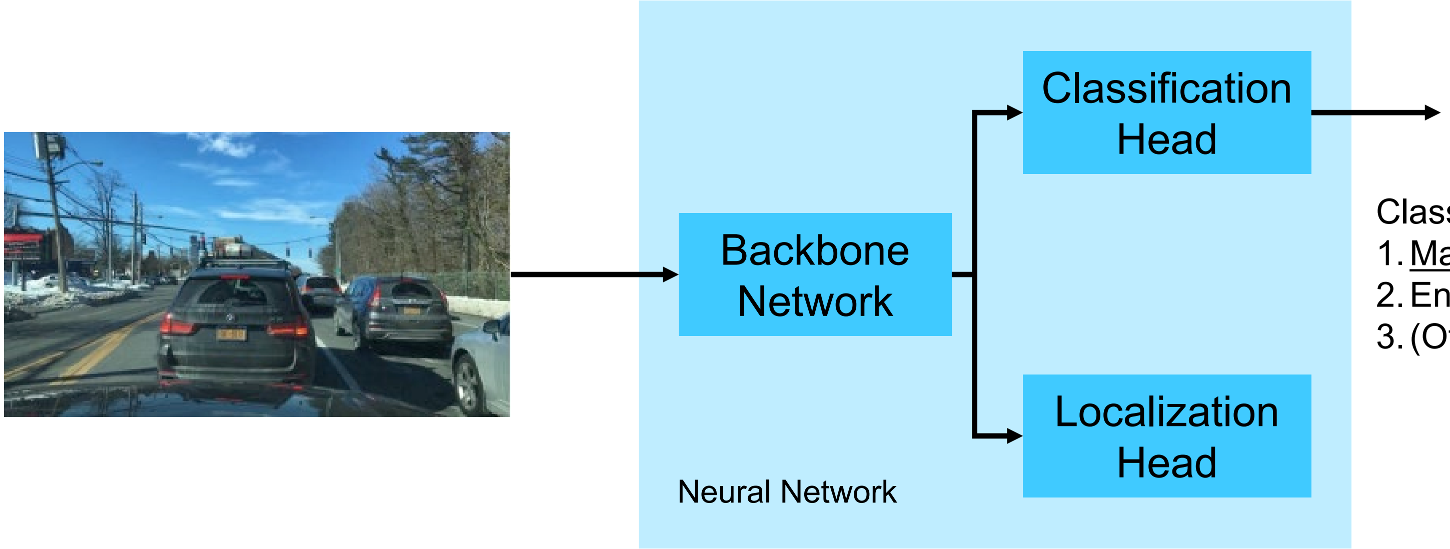
Most standard object detection models already express predictive uncertainty for classification by producing probabilities!

Carnegie Mellon University Software Engineering Institute

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.



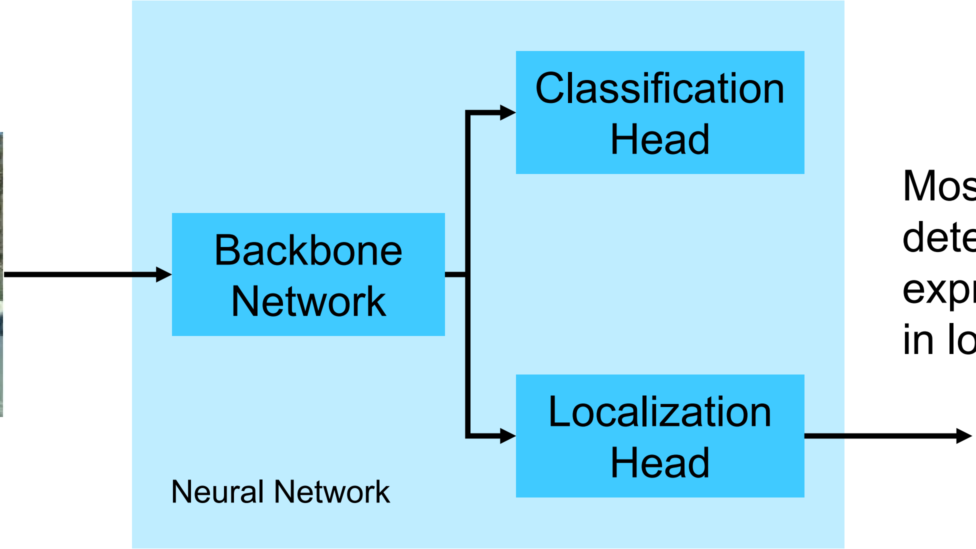
$(0.9, 0.001, 0.05, \ldots)$

Classification uncertainty metrics:
1. Maximum classification probability
2. Entropy of class probability distribution
3. (Others)

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.



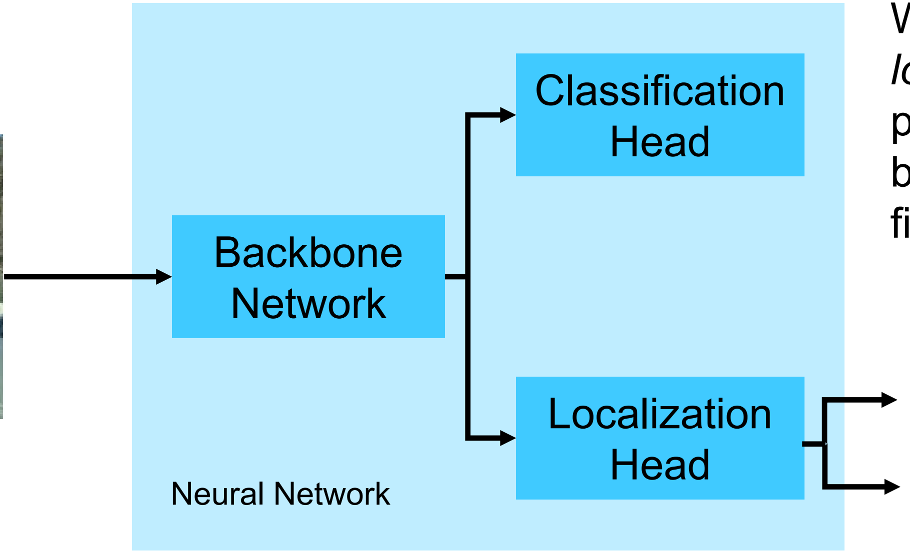
$$((x_1, y_1), (x_2, y_2))$$

Most standard object detection models <u>DO NOT</u> express obvious uncertainty in 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.



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

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

We use a technique called *loss attenuation** to produce probabilistic estimates of a bounding box instead of a fixed prediction.

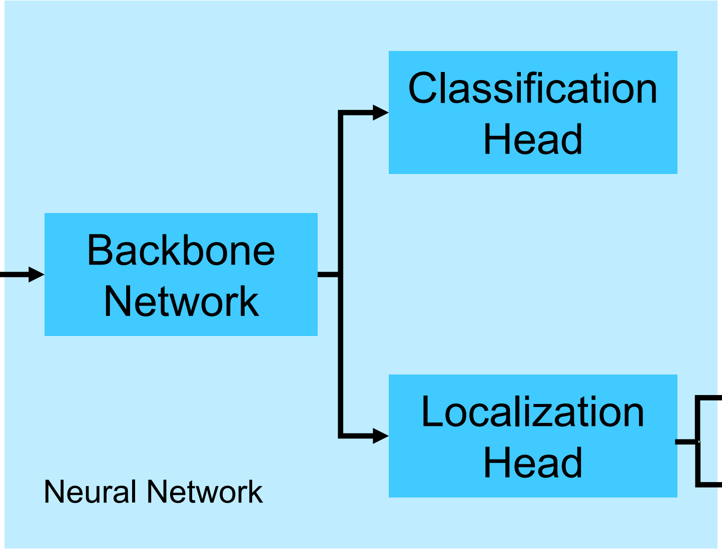
*Kendall, Alex, and Yarin Gal. "What uncertainties do we need in bayesian deep learning for computer vision?." *Advances in neural information processing systems* 30 (2017).

Carnegie Mellon University
Software Engineering Institute

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.



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

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

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

Uncertainty in Object Detectors

Carnegie Mellon University Software Engineering Institute

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.

What happens when a detector is uncertain?

ertainty metrics:

er Gaussian parameters

2. Normalized Entropy
3. (Others)

Backbone Network

Localization Head

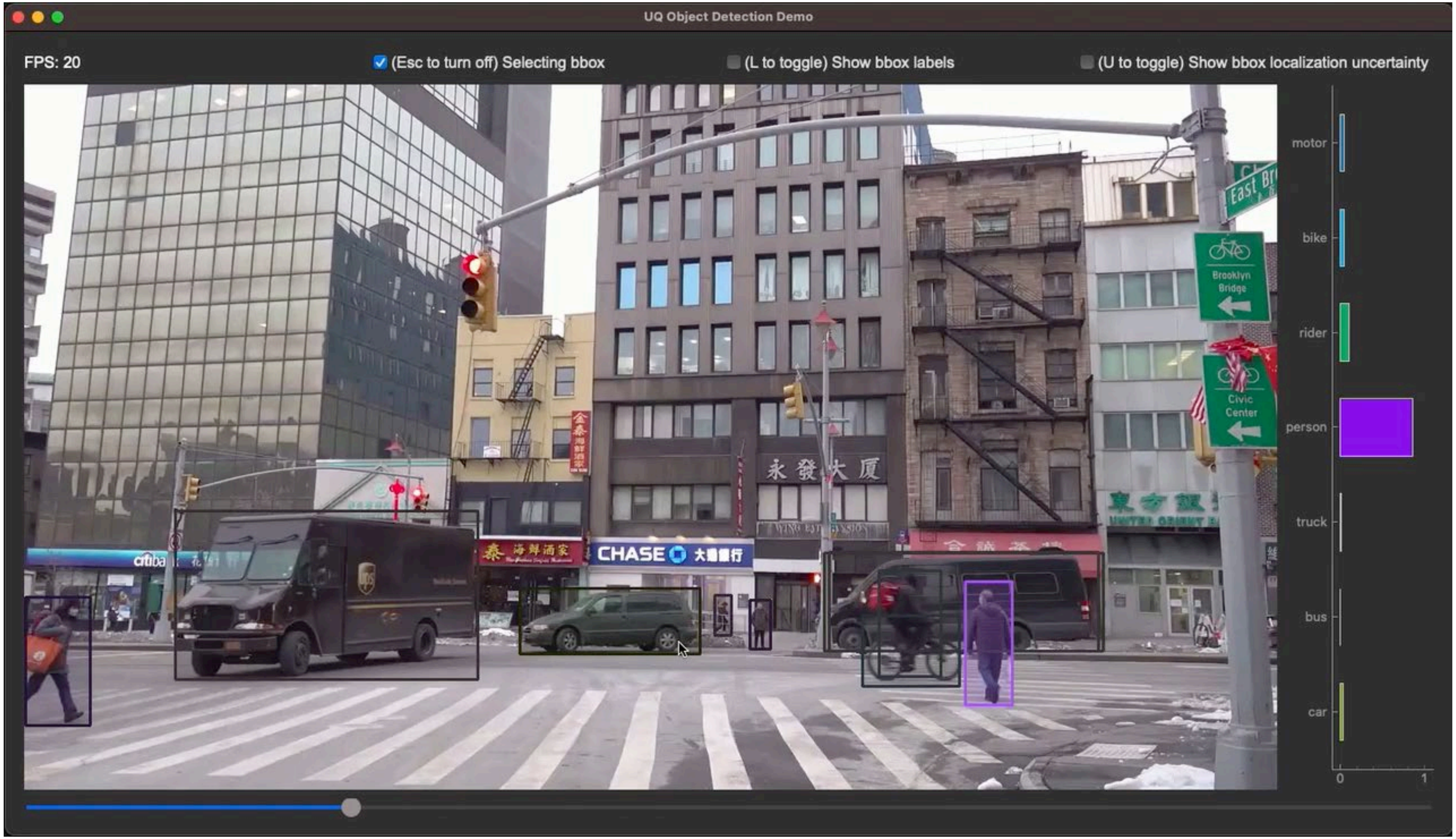
Neural Network

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

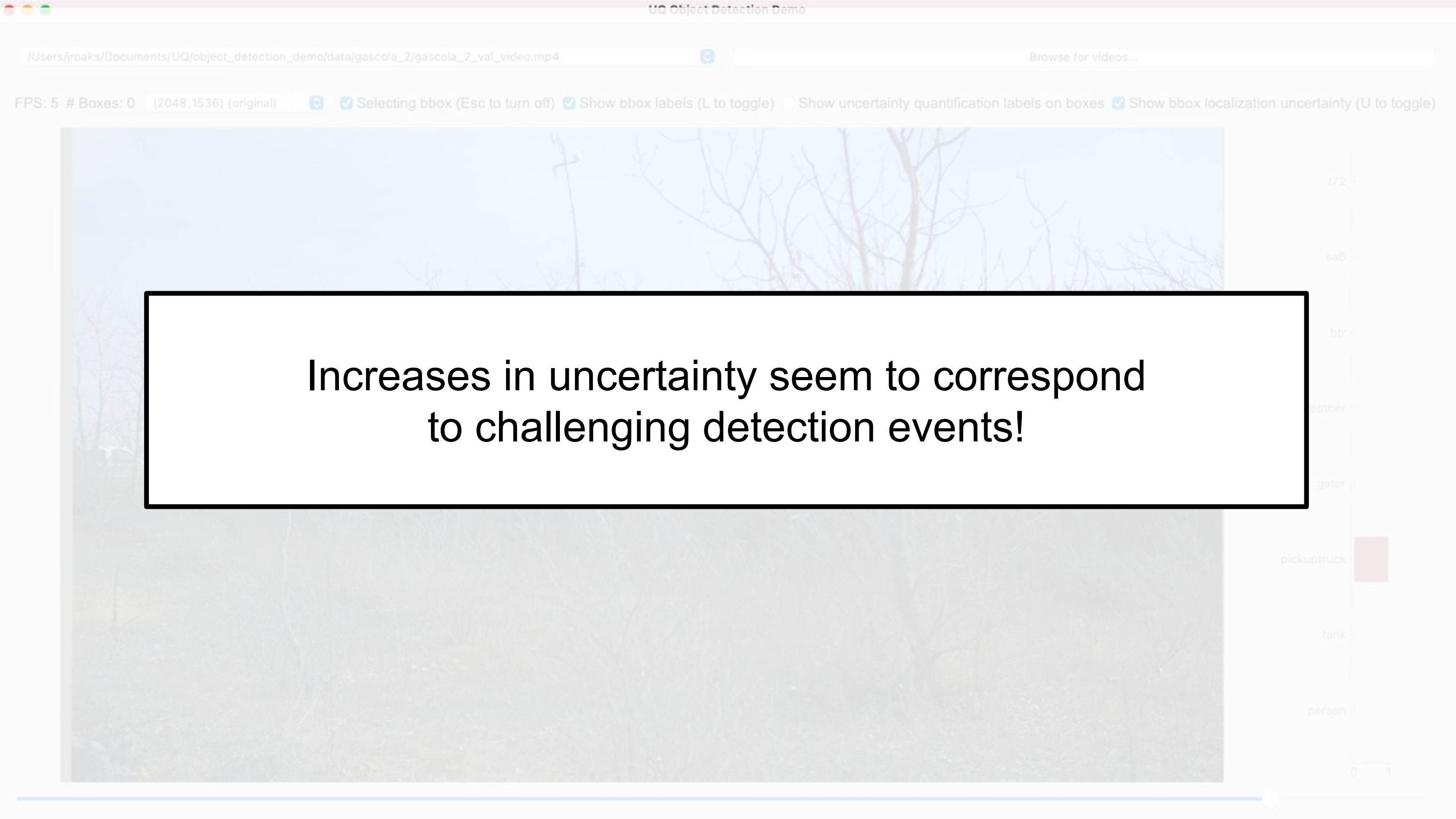
$$\left(\left(\sigma_{x_1},\sigma_{y_1}\right),\left(\sigma_{x_2},\sigma_{y_2}\right)\right)$$

Carnegie Mellon University Software Engineering Institute

Probabilistic Object Detection Example – Overlapping Objects

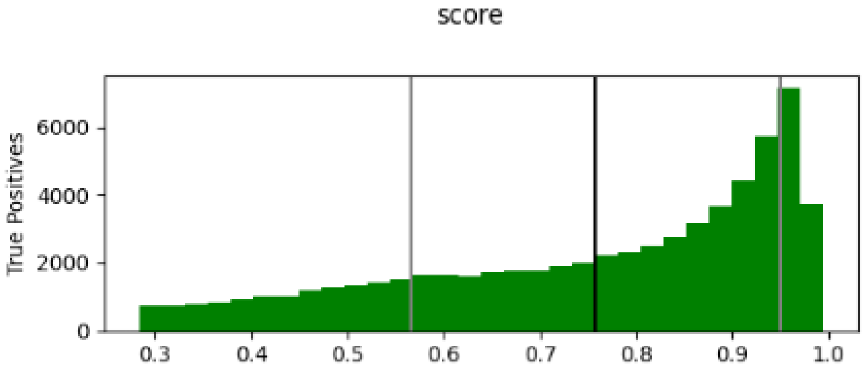


Probabilistic Object Detection Example – Occlusion

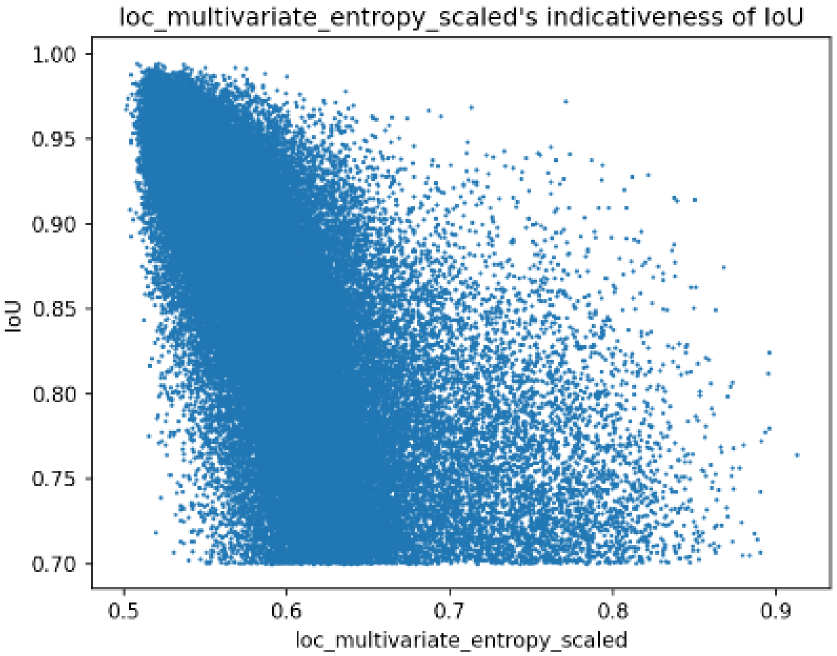


Increases in uncertainty seem to correspond
to challenging detection events!

Preliminary Quantitative Results



Classification

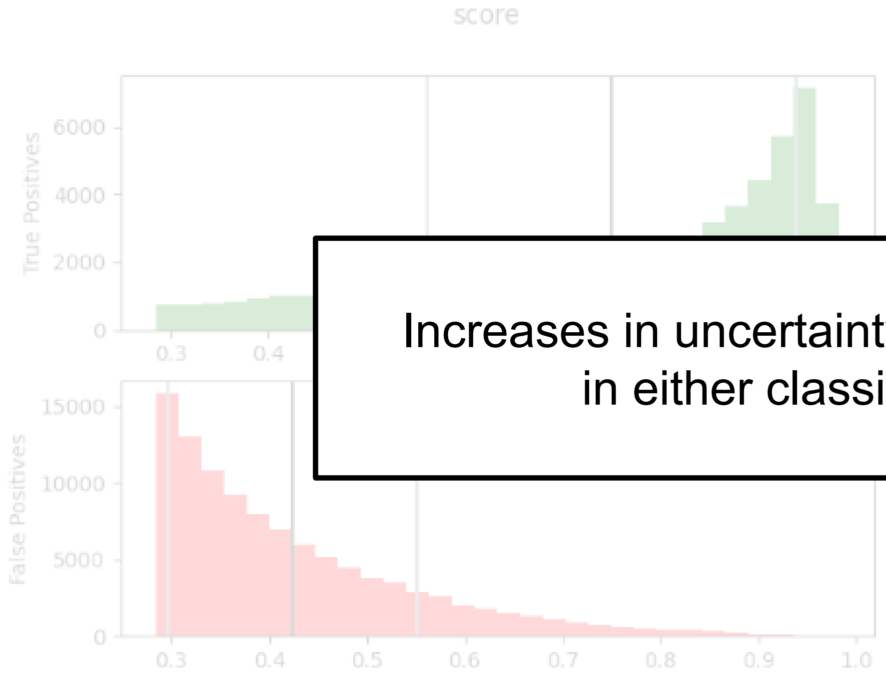


Localization

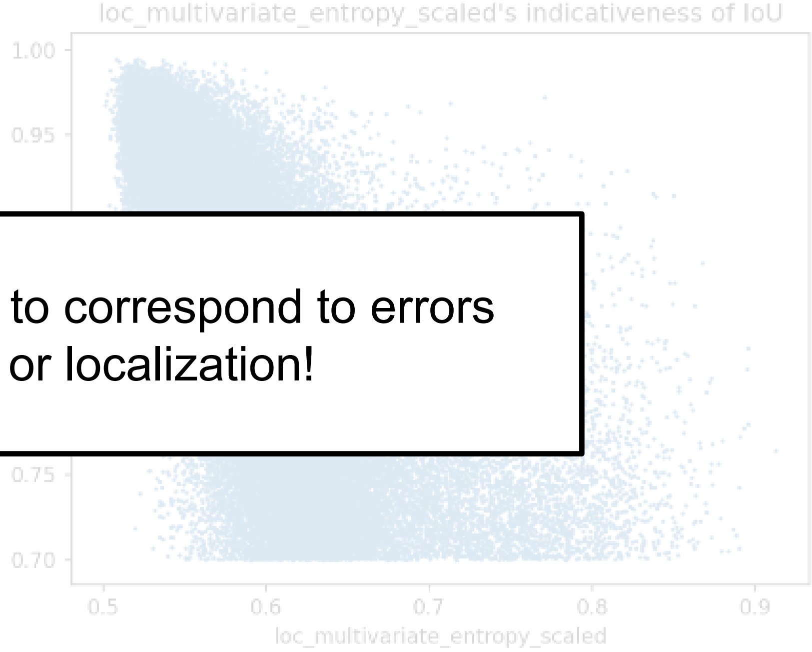
Preliminary Quantitative Results

Carnegie Mellon University Software Engineering Institute

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



Classification



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 also *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"