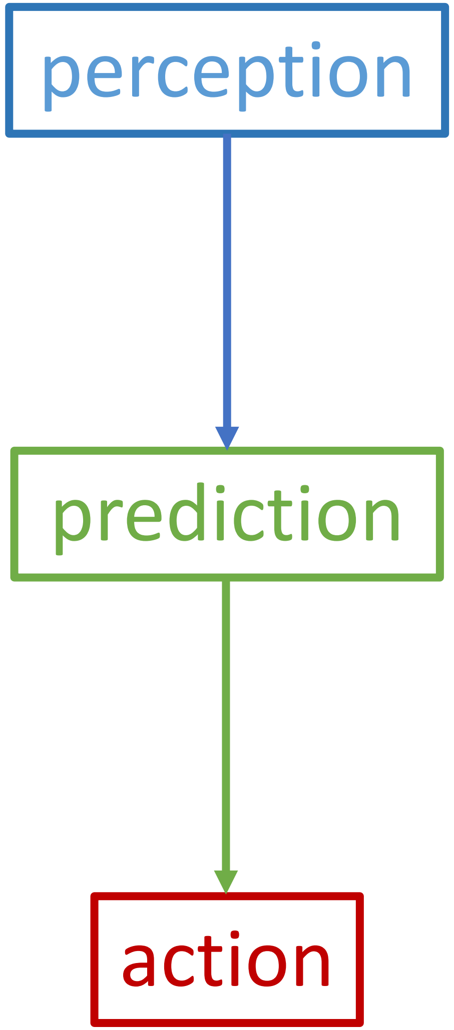
UNIVERSITY OF CAMBRIDGE

IJCAI-17 MELBOURNE

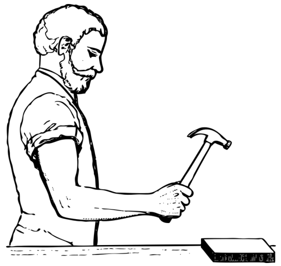
Concrete Problems for Autonomous Vehicle Safety: Advantages of Bayesian Deep Learning

Rowan McAllister, Yarin Gal, Alex Kendall, Mark van der Wilk, Amar Shah, Roberto Cipolla, Adrian Weller

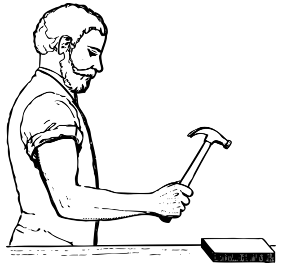
How can we make autonomous vehicles safe?


perception
prediction
action

perception



prediction



action

Object detection

Object tracking

Free-space estimation

Behavior Estimation

Motion Prediction

Motion planning

Control

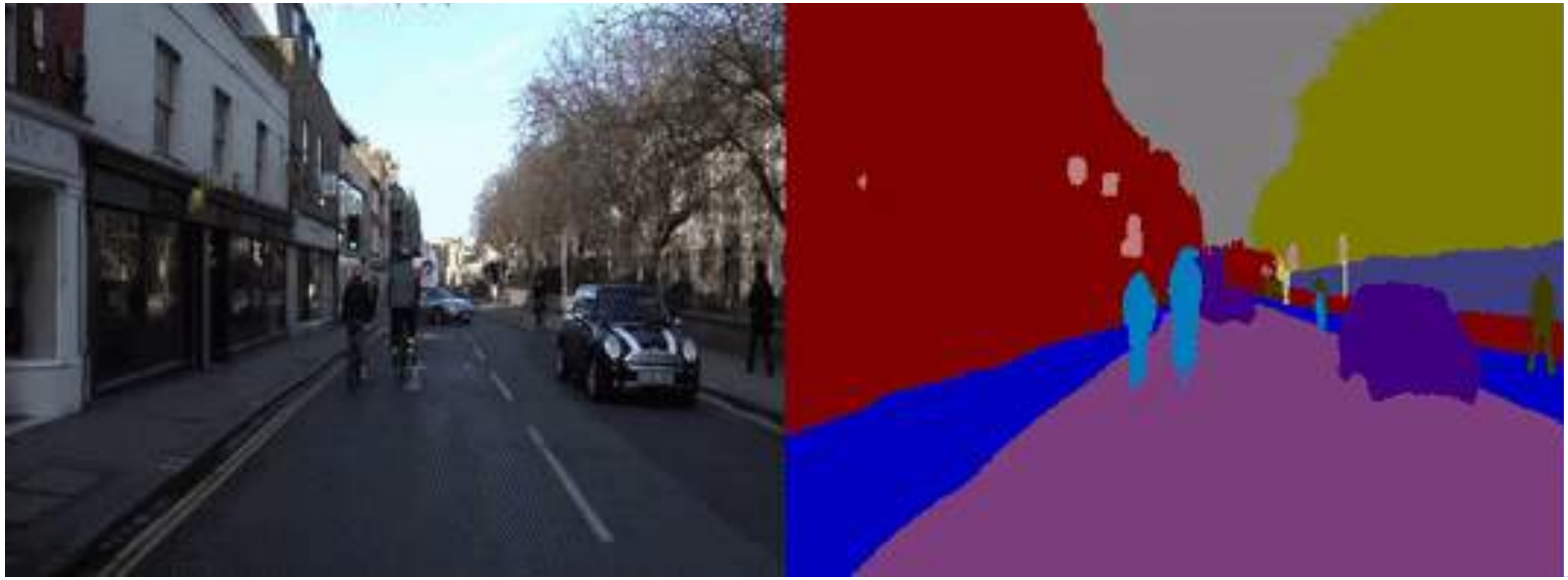
This doesn't scale...

End-to-end learning from perception to action

Mnih et al. "Human-level control through deep reinforcement learning." *Nature* 2015

Bojarski et al. "End to end learning for self-driving cars." arXiv 2016

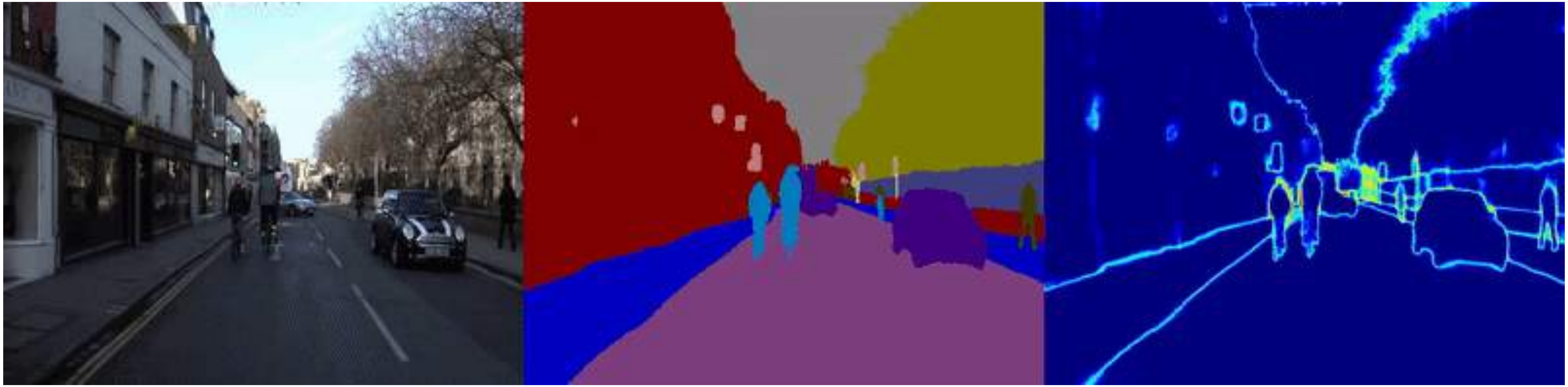
Deep Learning



Input Image Semantic Segmentation

Alex Kendall and Yarin Gal. *"What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?"* arXiv preprint 1703.04977, 2017.

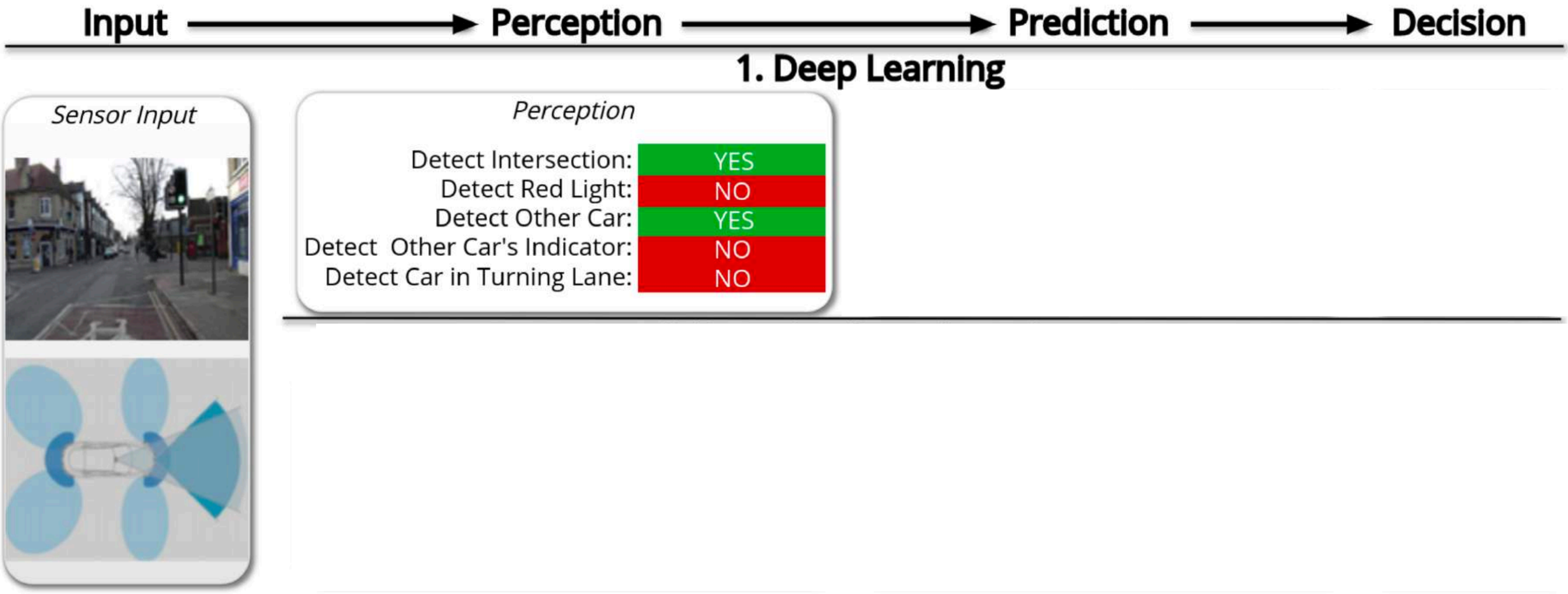
Bayesian Deep Learning



Input Image　　Semantic Segmentation　　Uncertainty

Alex Kendall and Yarin Gal. *"What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?"* arXiv preprint 1703.04977, 2017.

A Concrete Example



Three research themes

1. Safety
2. Interpretability
3. Compliance

Safety Concrete Problems

- Improving uncertainty inference in Bayesian deep learning models
  - Real-time constraint restricts sampling methods
  - Models often underestimate uncertainty
- Propagate uncertainty through all layers
  - Important to account for input uncertainty in modular systems
- Metrics for quantifying uncertainty estimates
- Accurately distinguishing different modes of uncertainty

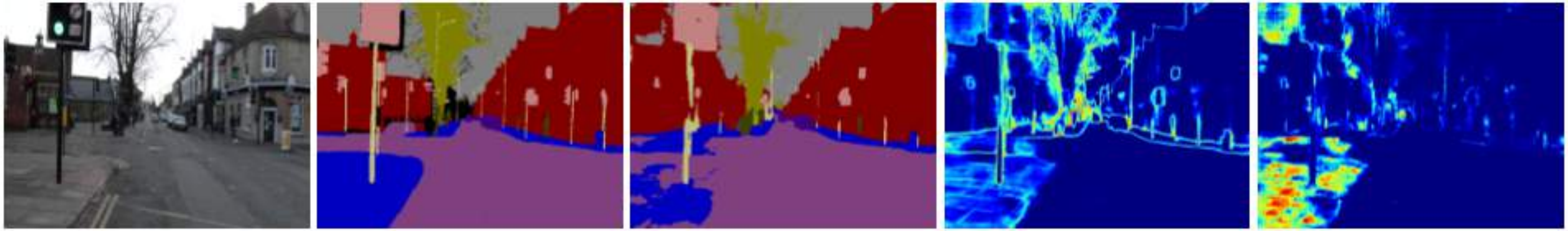
Types of uncertainty

Epistemic uncertainty

- Measures what you're model doesn't know
- Can be explained away by more data

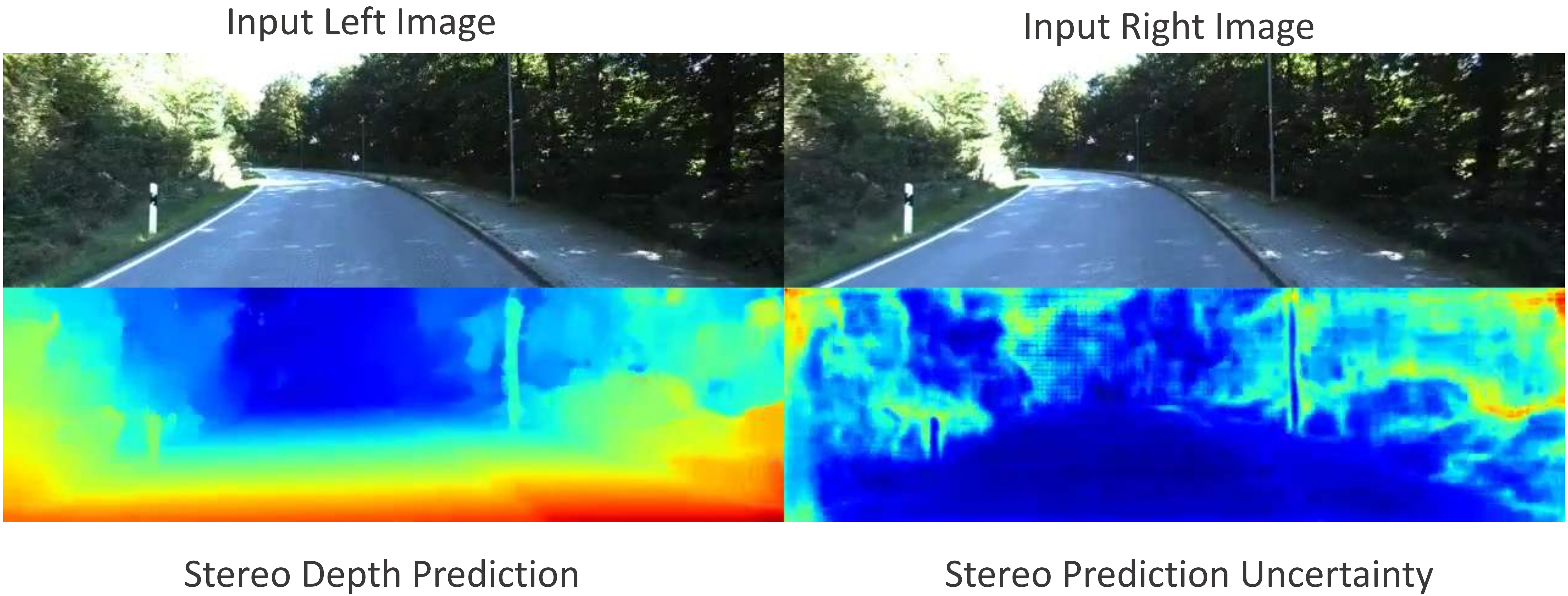
Aleatoric uncertainty

- Measures what you can't understand from data
- Can be explained away by better sensing



(a) Input Image (b) Ground Truth (c) Semantic Segmentation (d) Aleatoric Uncertainty (e) Epistemic Uncertainty

Stereo Depth Estimation



Kendall et al. *"End-to-End Learning of Geometry and Context for Deep Stereo Regression"* ICCV 2017

Interpretability Concrete Problems

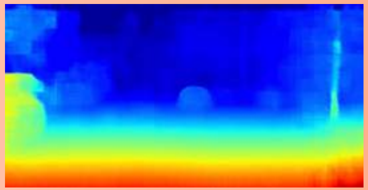
- Model saliency (how models make decisions)
  - inferring causal relationship between input signal and output decision
- Auxiliary outputs
  - human understandable intermediate representations
- Attribution of performance
  - validating individual components in an end-to-end model

Interpretability

End-to-end learning with intermediate outputs

Inputs:
1. Camera video
2. Sat-Nav directions

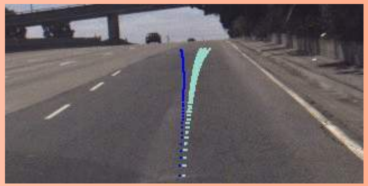
Scene geometry



Scene semantics

Object motion prediction

Outputs:
Driving commands



Alex Kendall, Yarin Gal and Roberto Cipolla. *"Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics."* arxiv preprint 1705.07115, 2017.

Compliance Concrete Problems

- Compliance to the passenger and law
- V2V, V2I, V2U communication
  - ”Back seat driving” within safe operational envelope
- Data efficient and bias-free learning
  - removing algorithmic bias and fairly accounting for rare scenarios
- Leveraging uncertainty for smarter learning curriculum

Conclusions

- Deep learning is not enough, we need Bayesian deep learning for safe autonomous vehicles
- Research problems can be grouped into safety, interpretability and compliance
- In addition to safety, this research is important to:
  - help passengers trust AV technology and explain behavior
  - help society overcome a reasonable fear of the unknown
  - aid engineers validate against safety standards
  - accountability for insurance and legal liability by explaining decisions

More Information

- Rowan McAllister, et al. **Concrete Problems for Autonomous Vehicle Safety: Advantages of Bayesian Deep Learning.** IJCAI, 2017.
- Dario Amodei, et al. **Concrete problems in AI safety.** arXiv, 2016.
- Alex Kendall and Yarin Gal. **What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?** arXiv, 2017.
- Yarin Gal. **Uncertainty in deep learning.** PhD thesis, University of Cambridge, 2016.
- Alex Kendall, Yarin Gal and Roberto Cipolla. **Multi-Task Learning Using Uncertainty to Weigh Losses for Scene Geometry and Semantics.** arXiv, 2017.

alexgkendall.com/blog @alexgkendall agk34@cam.ac.uk