# [Deep Learning Is Not Good Enough,<br>We Need Bayesian Deep Learning for Safe AI]

I. Understanding what a model does not know is a critical part of many machine learning systems. Unfortunately, *today's* *deep learning algorithms are usually unable to understand their uncertainty.* These models are often taken blindly and assumed to be accurate, which is not always the case. For example, in two recent situations this has had disastrous consequences.

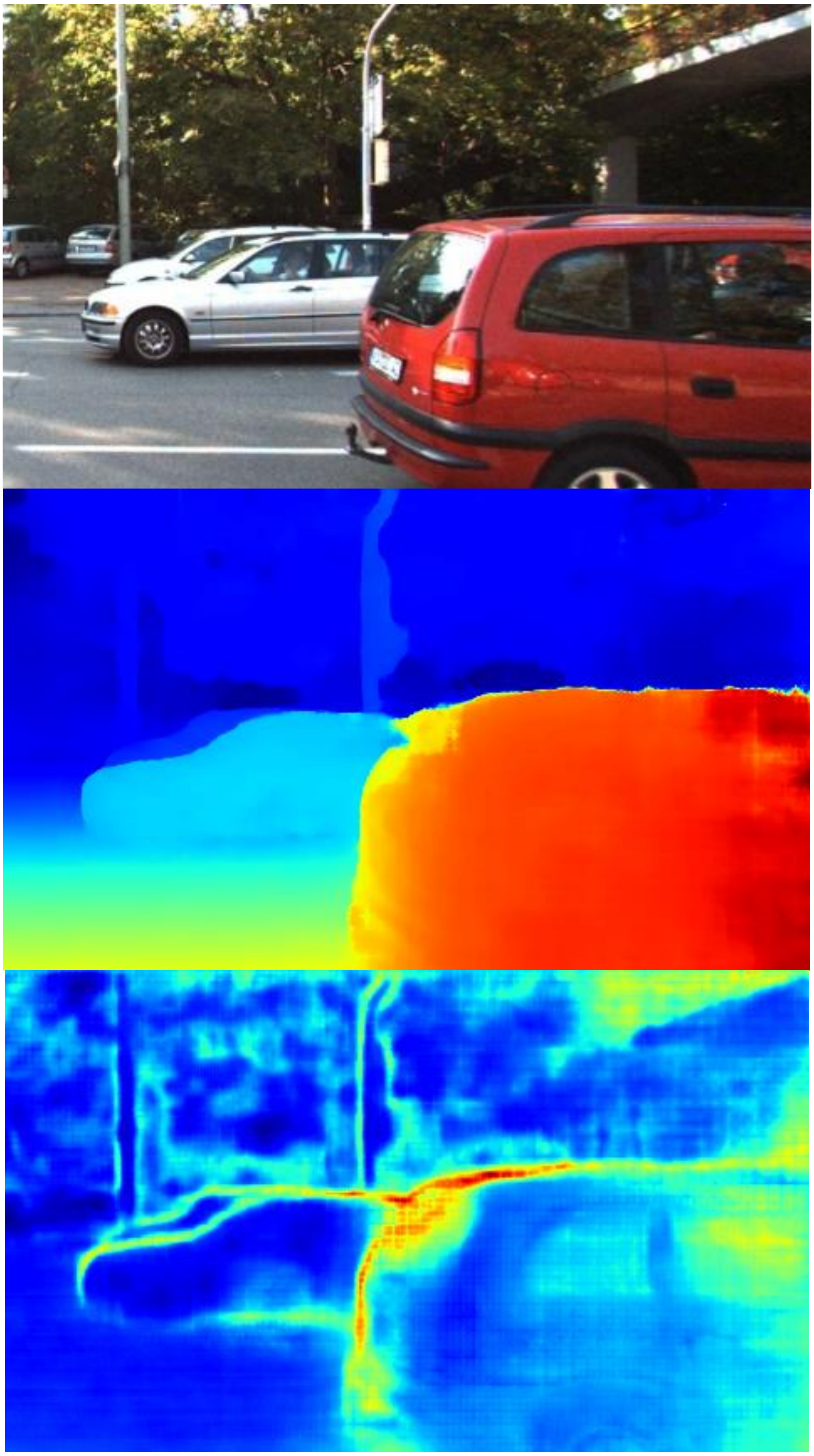
   a. In May 2016 we tragically experienced the first **fatality** from an assisted driving system. According to the manufacturer's blog, "Neither Autopilot nor the driver noticed the white side of the tractor trailer against a brightly lit sky, so the brake was not applied."

   b. In July 2015, an image classification system erroneously identified two African American humans as gorillas, raising concerns of racial discrimination. See the news report here.

II. And I'm sure there are many more interesting cases too! If both these algorithms were able to assign a high level of uncertainty to their **erroneous predictions**, then each system may have been able to make better decisions and likely avoid disaster.

III. It is clear to me that understanding uncertainty is important. So why doesn't everyone do it? The main issue is that traditional machine learning approaches to understanding uncertainty, such as **Gaussian processes**, do not scale to high dimensional inputs like images and videos. To effectively understand this data, we need deep learning. But deep learning struggles to model uncertainty.

IV. In this post I'm going to introduce a field known as Bayesian deep learning (BDL), which provides a deep learning framework which can also model uncertainty. BDL can achieve state-of-the-art results, while also understanding uncertainty. I'm going to explain the different types of uncertainty and show how to model them. Finally, I'll discuss a recent result which shows how to use uncertainty to weight losses for multi-task deep learning.



An example of why it is really important to understand uncertainty for depth estimation. The first image is an example input into a Bayesian neural network which estimates depth, as shown by the second image. The third image shows the estimated uncertainty. You can see the model predicts the wrong depth on difficult surfaces, such as the red car’s reflective and transparent windows. Thankfully, the Bayesian deep learning model is also aware it is wrong and exhibits increased uncertainty.

V. **Types of uncertainty**

The first question I'd like to address is what is uncertainty? There are actually different types of uncertainty and we need to understand which types are required for different applications. I'm going to discuss the two most important types – epistemic and aleatoric uncertainty.

VI. **Epistemic uncertainty**

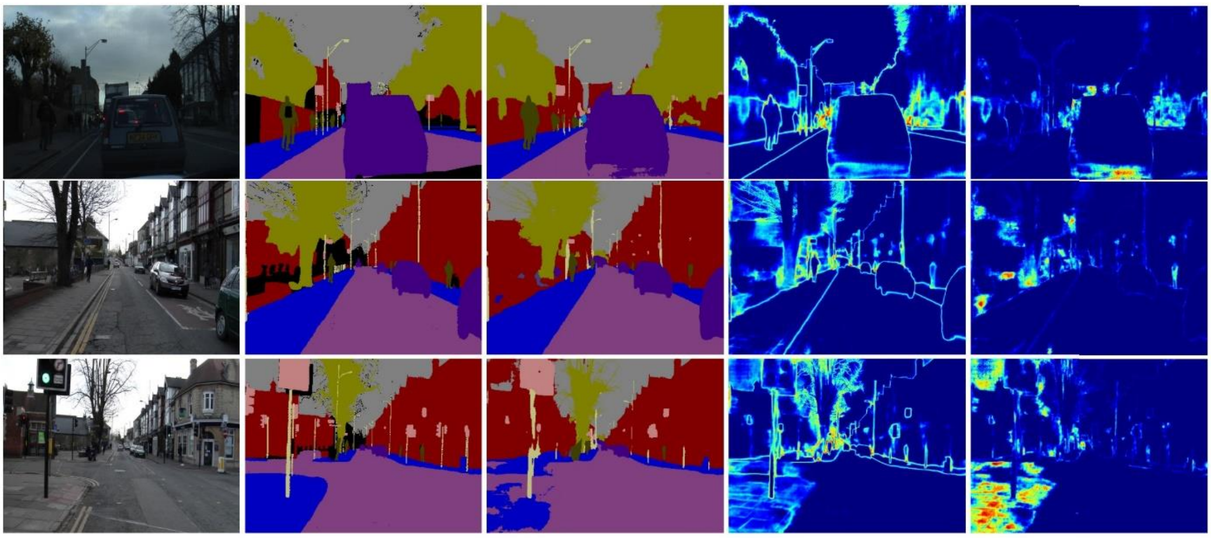
*Epistemic uncertainty* captures our ignorance about which model generated our collected data. This uncertainty can be explained away given enough data, and is often referred to as *model uncertainty*. Epistemic uncertainty is really important to model for:

a. Safety-critical applications, because epistemic uncertainty is required to understand examples which are different from training data,
b. Small datasets where the training data is **sparse**.

VII. **Aleatoric uncertainty**

*Aleatoric uncertainty* captures our uncertainty with respect to information which our data cannot explain. For example, aleatoric uncertainty in images can be attributed to occlusions (because cameras can't see through objects) or lack of visual features or over-exposed regions of an image, etc. It can be explained away with the ability to observe all explanatory variables with increasing precision. Aleatoric uncertainty is very important to model for:

a. Large data situations, where epistemic uncertainty is mostly explained away,
b. Real-time applications, because we can form aleatoric models as a deterministic function of the input data, without expensive **Monte Carlo sampling**.



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

Illustrating the difference between aleatoric and epistemic uncertainty for semantic segmentation. You can notice that aleatoric uncertainty captures object boundaries where labels are noisy. The bottom row shows a failure case of the segmentation model, when the model is unfamiliar with the footpath, and the corresponding increased epistemic uncertainty.

Next, I'm going to show how to form models to capture this uncertainty using Bayesian deep learning.

VIII. **Bayesian deep learning**

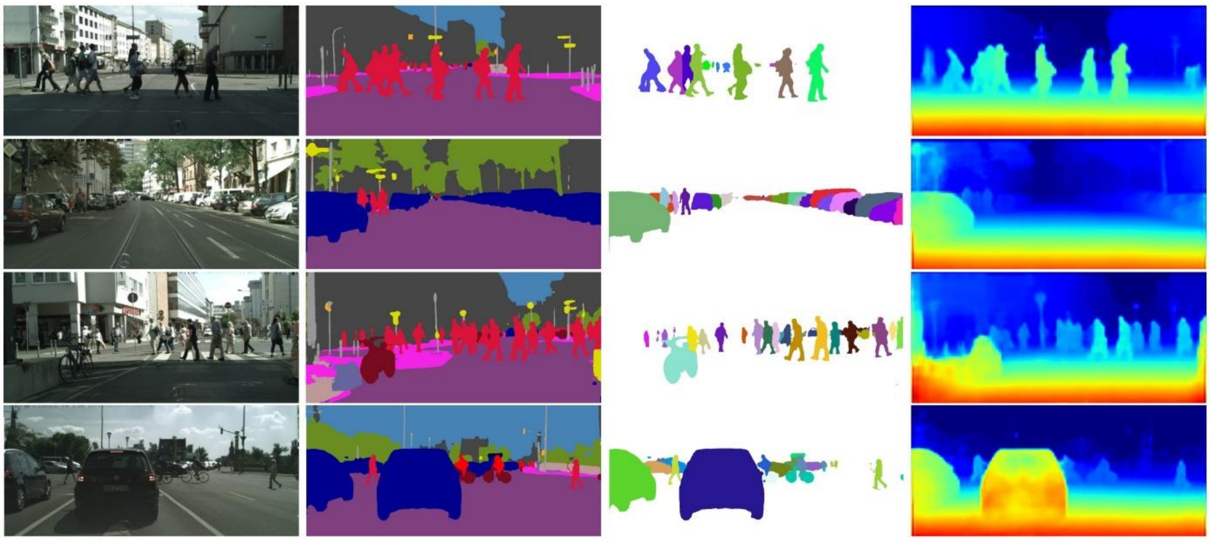
Bayesian deep learning is a field at the intersection between deep learning and Bayesian probability theory. *It offers principled uncertainty estimates from deep learning architectures.* Bayesian deep learning models typically form uncertainty estimates by either placing distributions over model weights, or by learning a direct mapping to probabilistic outputs.

IX. **Uncertainty for multi-task learning**

X. Next I'm going to discuss an interesting application of these ideas for multi-task learning.

XI. Multi-task learning aims to improve **learning efficiency** and **prediction accuracy** by learning multiple objectives from a shared representation. It is prevalent in many areas of machine learning, from **NLP** to speech recognition to computer vision. *Multi-task learning is of crucial importance in systems where long computation run-time is prohibitive, such as the ones used in robotics.* Combining all tasks into a single model reduces computation and allows these systems to run in real-time.

XII. We explore multi-task learning within the setting of visual scene understanding in computer vision. *Scene understanding algorithms must understand both the geometry and semantics of the scene at the same time.* This forms an interesting multi-task learning problem because scene understanding involves joint learning of various regression and classification tasks with different units and scales. Perhaps surprisingly, we show our model can learn multi-task weightings and outperform separate models trained individually on each task.



(a) Input image (b) Segmentation output (c) Instance output (d) Depth output

Multi-task learning improves the smoothness and accuracy for depth perception because it learns a representation that uses cues from other tasks, such as segmentation (and vice versa).

XIII. **Some challenging research questions**

Why doesn't Bayesian deep learning power all of our A.I. systems today? I think they should, but there are a few really tough research questions remaining. To conclude this blog I'm going to mention a few of them:

*a.* *Real-time epistemic uncertainty techniques are preventing epistemic uncertainty models from being deployed in real-time robotics applications. Either increasing sample efficiency, or new methods which don't rely on Monte Carlo inference would be incredibly beneficial.*

b. *Benchmarks for Bayesian deep learning models. It is incredibly important to quantify improvement to rapidly develop models* – look at what benchmarks like ImageNet have done for computer vision. We need benchmark suites to measure the calibration of uncertainty in BDL models too.

*c.* *Better inference techniques to capture multi-modal distributions.*

## Vocabulary

| | | | |
|---|---|---|---|
| fatality | erroneous predictions | Gaussian processes | Monte Carlo sampling |
| learning efficiency | prediction accuracy | NLP | to run in real-time |

## Explain or elaborate upon the following statements taken from the text above:

*a.* *deep learning algorithms are usually unable to understand their uncertainty (Para I)*

*b.* *It offers principled uncertainty estimates from deep learning architectures (Para VIII)*

*c.* *Multi-task learning is of crucial importance in systems where long computation run-time is prohibitive, such as the ones used in robotics (Para XI)*

*d.* *Scene understanding algorithms must understand both the geometry and semantics of the scene at the same time (Para XII)*

[We need…]

*e.* *Real-time epistemic uncertainty techniques are preventing epistemic uncertainty models from being deployed in real-time robotics applications. Either increasing sample efficiency, or new methods which don't rely on Monte Carlo inference would be incredibly beneficial (Para XIII)*

*f.* *Benchmarks for Bayesian deep learning models. It is incredibly important to quantify improvement to rapidly develop models (Para XIII)*

g. *Better inference techniques to capture multi-modal distributions (Para XIII)*

## Question

Explain the what aleatoric and epistemic uncertainty are and their significance for developing and implementing Bayesian inference systems in AI.