What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?

Alex Kendall (agk34@cam.ac.uk) and Yarin Gal (yarin@cs.ox.ac.uk)

1. Types of Uncertainty

In Bayesian modelling, there are two main types of uncertainty we can model [1]:

- **Epistemic uncertainty**: uncertainty in the model, capturing what our model doesn’t know due to lack of training data. Can be explained away with increased training data.
- **Aleatoric uncertainty**: information which our data cannot explain. Can be explained away with increased sensor precision.

2. We jointly model aleatoric and epistemic uncertainty with deep learning. Our model’s uncertainty for pixel output \( y_i \) is given by:

\[
\text{Var}(y_i) \approx \frac{1}{T} \sum_T \sigma(x_t)^2 + \frac{1}{T} \sum_T f(x_t)^2 - \left( \frac{1}{T} \sum_T f(x_t) \right)^2
\]

Using Monte Carlo dropout samples, \( T \), learning aleatoric uncertainty with loss:

\[
\text{Loss}(\theta) = \frac{1}{D} \sum_x \frac{1}{2\sigma(x_i)} \left| y_i - f(x_i) \right|^2 + \log \sigma(x_i)
\]

3. SOTA performance

For semantic segmentation and per-pixel depth regression datasets.

We use a convolutional network based on DenseNet [20] with 103 layers and 9.4M parameters

4. Uncertainty with Distance from Training Data

Experiments training on one dataset and testing on another.

- Aleatoric uncertainty cannot be explained away with more data,
- Aleatoric uncertainty does not increase for out-of-data examples (situations different from training set),
- Epistemic uncertainty increases with decreasing training size,
- Epistemic uncertainty increases with examples out of the training distribution.

5. Conclusions

It is important to model **aleatoric** uncertainty for:

- Large data situations, where epistemic uncertainty is explained away,
- Real-time applications, because we can form aleatoric models without expensive MC samples.
- Noisy data, because we can learn to attenuate erroneous labels.

And **epistemic** uncertainty is important for:

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