Master's Thesis: Laplace Approximation for Computer Vision

With the increasing reliability of deep-learning models, they are being deployed in a growing number of safety-critical applications. In this context, however, it is often crucial to be able to quantify the uncertainty of the model prediction.

Maximum likelihood estimation, which is associated with common objective functions in deep learning, can only capture the uncertainty reflected in the training data (e.g. the measurement noise) because it relies on the assumption that the training data covers the entire scope of the model's application. In applications that process high-dimensional input under real-world conditions, as is the case in many computer-vision tasks, this assumption is often violated.

Bayesian deep learning addresses this limitation as it allows estimating the uncertainty caused by a lack of knowledge---i.e. for inputs that differ from the distribution of the training data---referred to as epistemic uncertainty. However, most Bayesian approaches require a specific objective function and sampling during inference time. Both requirements are generally not suitable for achieving state-of-the-art performance.

An exception to this is the Laplace approximation (LA): It can be applied post-hoc to any model that was trained with a convex and twice differentiable objective function. Recent developments have furthermore led to computationally efficient versions of the LA that can provide closed-form predictive uncertainties without the need for sampling [1].

The goal of this master's thesis is to develop an approach for applying the LA to popular computer-vision tasks, in particular to

1. object detection (bounding-box regression) and
2. pose estimation (keypoint regression),

and ideally to verify this approach on state-of-the-art models, such as recent versions of YOLO [2, 3]. The main challenge here is that the corresponding objective functions, i.e. IoU-based losses [4] and the OKS loss [5], are not necessarily convex or twice differentiable. This does not allow using the LA out of the box. In addition, recent computer-vision models introduce their own degree of complexity that should be taken into account.

The master's thesis is supervised in collaboration with Airbus Central Research and Technology and is directly related to the currently developed vision-landing system.

[1] Overview LA with code base: Laplace Redux -- Effortless Bayesian Deep Learning

[2] YOLOv8

[3] YOLO-NAS

[4] Overview IoU-based losses

[5] Overview OKS loss