netidee STIPENDIEN

Increasing Trustworthiness of Edge AI by Adding Uncertainty Estimation to Object Detection

Zwischenbericht | Call 18 | Stipendium ID 6885

Content

1 Introduction

This interim report highlights the current progress on the Master's thesis *Increasing Trustworthiness of Edge AI by Adding Uncertainty Estimation to Object Detection*. The aim of this thesis is to adapt object detection, a common computer vision task, to include uncertainty estimation in its predictions. Thus, the model can express itself with "I don't know" when it is overwhelmed with irregular input, instead of making overconfident false predictions. Additionally, this process must use as little computational overhead as possible, to be feasible on resource-constrained edge devices, e.g., for smart-traffic applications. The first steps of this thesis include, i) checking state-of-the-art literature, ii) setting up a methodology for evaluation, and iii) gathering insights by implementing a prototype.

The following blog posts provide more details:

- Uncertainty Estimation in Edge AI
- Adding Uncertainty Estimation to Object Detection
- Evaluating the Quality of Uncertainty Estimation

2 Status

2.1 Milestone 1 – Literature Research

Regarding object detection models, there are two main categories currently widely in use, one-stage detectors and two-stage detectors. According to the requirement of low computational overhead in this work, the faster one-stage object detectors, such as the *Single Shot MultiBox Detector (SSD)* [1] or the *You Only Look Once (YOLO)* model [2], will be chosen for the experiments. For uncertainty estimation, *Bayesian* neural networks [3] and network *ensembles* [4] are commonly used techniques. Most interesting for the scope of this thesis, *Evidential Deep Learning (EDL)* [5, 6] as an emerging uncertainty estimation technique is promising faster computation compared to the alternatives.

Combining object detection and uncertainty estimation has gained popularity in recent years. Kraus and Dietmayer [7] use the well-established concept of Bayesian networks for uncertainty estimation, at the cost of some inference overhead due to multiple forward passes for the same input. Other recent works recognized the potential of EDL for fast uncertainty estimation and added it to object detection [8, 9, 10]. However, they do not cover the combined aspects of this thesis, i) per-object uncertainty estimation, ii) comparing different uncertainty estimation methods, iii) focus on a smart-traffic use-case, iv) optimizing models for low inference time on edge devices. This initial literature research helped to define the scope of this work and will guide further progress.

2.2 Milestone 2 – Defining Methodology

The following methodology will be used in this thesis:

1. Literature research of state-of-the-art for object detection and uncertainty estimation.
2. Implementing and testing a prototype model with added uncertainty estimation.
3. According to the limitations of the prototype, adapting various state-of-the-art models and techniques for accurate object detection and fast uncertainty estimation on the edge.
4. Setting up an evaluation framework for models to be evaluated together.
5. Evaluation of model size, inference time, and mean Average Precision, with datasets suiting the targeted use-case of a smart-traffic scenario.
6. Summary of gained insights and writing of the Diploma Thesis.

2.3 Milestone 3 – Prototype Implementation and Evaluation

The prototype was implemented by combining the commonly used SSD object detector [1] and adapting it to use the EDL classification uncertainty estimation technique [5]. Necessary modifications were changing the activation function of the output layer of the network, from *Softmax* to *ReLU*, and changing the classification part of the loss function according to [5]. Due to the simplicity of these modifications, no significant inference time overhead occurred, establishing one partial goal of this thesis. The initial focus was on pedestrian detection, where models were trained and evaluated on the Oxford Town Centre dataset [11]. The results showed that although uncertainty could be estimated, object recognition precision declined noticeably compared to the regular object detection model. This is illustrated in Figure 1, where the mean Average Precision (mAP) went from ~49% (base model) to ~40% (modified EDL uncertainty model). Consequently, uncertainty estimation will take some additional computational effort and cannot be obtained for "free" without compromising the base model performance. Therefore, considering different model architectures is necessary, where the recent works of [8, 9, 10] will serve as inspiration.

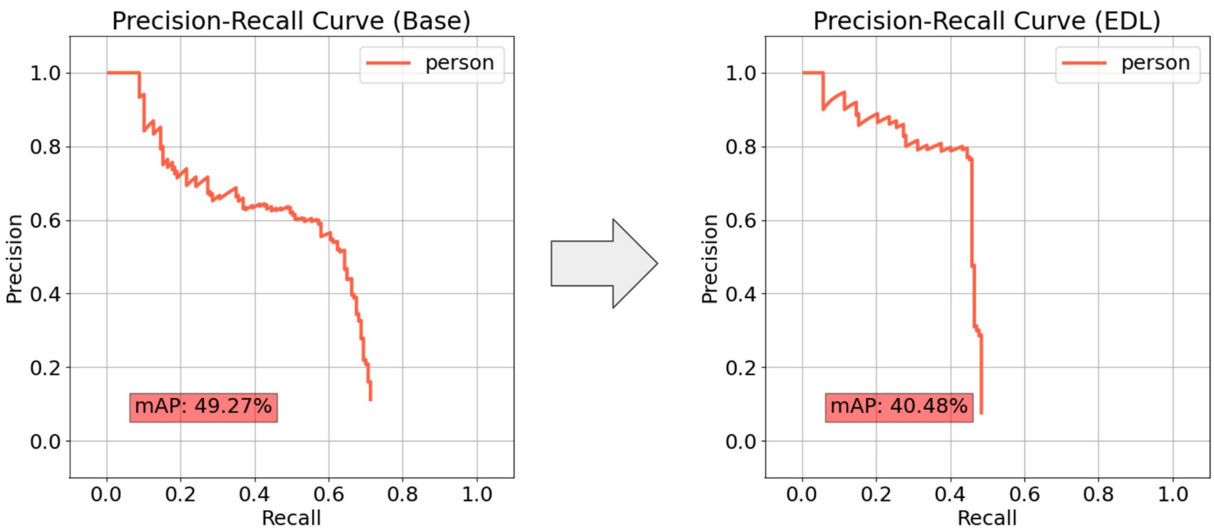

Figure 1: Precision-Recall evaluation plots of base object detection model (left) and model with EDL uncertainty estimation (right).

3 Summary of Changes in Planning

Overall, progress is on track, but the planning was split up into more fine-granular and detailed milestones. One additional milestone was *Defining Methodology* (Section 2.2) , to set up a guideline for evaluating this work. This addition caused most other milestone target dates to shift one month into the future, which was compensated for by reducing the planned period for the writing part of the thesis by one month (September-October 2024). Therefore, the target completion time of this thesis stayed the same with October 2024. Another milestone added for October 2024 was the publication of the evaluation code artifact used in this thesis. The code is currently planned to be published under the *MIT* license, however, this may still change depending on other works that get incorporated, which may cause conflicts between different licenses.

4 References

[1] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In European conference on computer vision, pages 21–37. Springer, 2016.

[2] J Redmon. Yolov3: An incremental improvement. CoRR, 2018.

[3] Yarin Gal and Zoubin Ghahramani. Dropout as a bayesian approximation: Representing model uncertainty in deep learning. In international conference on machine learning, pages 1050–1059. PMLR, 2016.

[4] Balaji Lakshminarayanan, Alexander Pritzel, and Charles Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. In Proceedings of the 31st International Conference on Neural Information Processing Systems, pages 6405–6416, 2017.

[5] Murat Sensoy, Lance Kaplan, and Melih Kandemir. Evidential deep learning to quantify classification uncertainty. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, pages 3183–3193, 2018.

[6] Alexander Amini, Wilko Schwarting, Ava Soleimany, and Daniela Rus. Deep evidential regression. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, Advances in Neural Information Processing Systems, volume 33, pages 14927–14937. Curran Associates, Inc., 2020.

[7] Florian Kraus and Klaus Dietmayer. Uncertainty estimation in one-stage object detection. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC), pages 53–60. IEEE, 2019.

[8] Younghyun Park, Wonjeong Choi, Soyeong Kim, Dong-Jun Han, and Jaekyun Moon. Active learning for object detection with evidential deep learning and hierarchical uncertainty aggregation. In The Eleventh International Conference on Learning Representations, 2022.

[9] Eduardo Aguilar, Bogdan Raducanu, Petia Radeva, and Joost Van de Weijer. Continual evidential deep learning for out-of-distribution detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 3444–3454, 2023.

[10] Ruofan Wang, Rui-Wei Zhao, Xiaobo Zhang, and Rui Feng. Towards evidential and class separable open set object detection. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 38, pages 5572–5580, 2024.

[11] Ben Benfold and I. Reid. Stable multi-target tracking in real-time surveillance video. CVPR 2011, pages 3457–3464, 2011.