Estimating Uncertainty using Generative Models in Object Detection

H.J.M., van Genuchten
ID: 1297333
Artificial Intelligence and Data Engineering Lab
Eindhoven University of Technology
h.j.m.v.genuchten@student.tue.nl

December 18, 2023

Abstract

Provide an informative summary of your proposal (topic, approach and potential importance of the results) in no more than three hundred words. Make sure to provide an informative and relevant abstract, as this is often the first part of your proposal that reviewers will read. The abstract should clearly describe what you are going to investigate, why you are going to investigate this subject and which results you expect to find. The abstract should have no more than 300 words and should contain a single paragraph. This proposal should be self-contained and has a limit of pages per section. **Title page, Abstract, Sections 1 and 2 have a combined limit of 8 pages. Section 3 has a limit of 8 pages. Section 4 has a limit of 2 pages. References and Appendices are unlimited.** You should not change the overall format considerably. Please do not change the margins nor introduce considerable LaTeX code that try to compress the content. You may use the appendices for any further information that you want to provide. While there is no page limit in the appendix, reviewers are not obliged to read all your appendix content.
Wordcount: 190

1 Overall aim and goals

- Create a generative model to model the uncertainty of various scences
  i.e. Detecting weird combinations of objects within a scence
- Use uncertainty to make more efficient use of human labelers.
- Design a system which can make efficient use of uncertainty of previous stages.

Using a generative model we can generate a better explanation of why a scene is uncertain. E.g. if we are able to model the following equation:

$$p(b, c, x) = p(x) * p(b|x) * p(c|x) \tag{1}$$

The separate parts can be used to give a more detailed

1.1 RQ

Current Ideas for possible RQs

- Is uncertainty a good indicator for active learning?
- Is uncertainty a good indicator for relabeling?
-

1.2 Motivation and Challenges

Artificial Intelligence solutions are increasingly put in new and challenging scenarios. Yet, often AI models are only able to give a distinction between a small set of answers even when the actual answer is outside of this set or the model has never seen anything like it before, and will gues an answer . Uncertainty Quantification (UQ) can enable a system to detect when a prediction might be of lesser quality, and allow it to preemptively react to that. Either by stopping or requesting human intervention. Furthermore, understanding when models are uncertain, allows for more effective data sampling and labeling by making use of active learning schemes . The later is especially useful in industries where labeling is expensive or timeconsuming. .

Add citation

Add citation of OoD

Add citations active learning usage

Add citation of label cost

Uncertainty Quantification is especially difficult in the case of Object Detection. It requires both localization (often by using regression) and classification.

1.3 Broad Literature Analysis

This project covers broader research areas, each will be covered separately in subsections 1.3.1 and 1.3.2. Related work in the combination will be described in subsection 1.3.3

1.3.1 Uncertainty Quantification

The ability to distinguish certain and uncertain outputs from machine learning models is useful for various reasons. It can be used for active learning [0, 0, 0], which makes better use of limited labeling capacity.

[0] distinguishes two kinds of uncertainty. Aleatoric uncertainty, the uncertainty that is caused by imprecise input data (i.e. x) and epistemic uncertainty, the uncertainty that is caused by the model.

Neural Networks (NN) have proven to be excellent at a broad range of tasks. However with

Uncertainty Quantification (UQ) is an important factor to increase the trust in automated processes based on machine learning. Their exist a various amount of methods both post-hoc and built-in .

Add post-hoc citations

Add citations

Reliable uncertainty estimation of Artificial Intelligence Models is important in safety critical situations. Expressing explicit uncertainty with Deep Neural Networks (DNN) can

1.3.2 Object Detection

Expand difference between one-stage and multi-stage

There are two main paradigms within Object Detection, One-Stage [0, 0, 0] and Multi-Stage detectors [0, 0**?**]. Single stage detectors directly detect a set of bounding boxes and their respective classes, whereas Multi-Stage detectors often first identify interesting patches which are then classified.

DiffusionDet: [0]

Describe the multi-faceted problem (i.e. both classification and regression)

1.3.3 In combination

The combination of Uncertainty Quantification and Object Detection is especially difficult as Object Detection is both a regression and a classification task.

[0]

Distinction between backbone and head, and some examples

Talk about "bag of freebies", these allow for better accuracy for the same inference speed at the cost of training time

1.4 Formulation of the problem and objectives

Let $x \in \mathbb{R}^{HxWxC}$ be the input image which contains objects O_i, which are a combination of a bounding box coordinates: $B_i = (x_{center}, y_{center}, x_{width}, y_{height})$ and a class $c \in C$. Our goal is to create a probabilistic model which

[0]

Formulate model

References

[] M. Bernhardt, D. C. Castro, R. Tanno, A. Schwaighofer, K. C. Tezcan, M. Monteiro, S. Bannur, M. P. Lungren, A. Nori, B. Glocker, J. Alvarez-Valle, and O. Oktay. Active label cleaning for improved dataset quality under resource constraints. *Nature Communications*, 13(1), Mar. 2022.

[] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao. Yolov4: Optimal speed and accuracy of object detection, 2020.

[] S. Chen, P. Sun, Y. Song, and P. Luo. Diffusiondet: Diffusion model for object detection, 2023.

[] Y. Gal et al. Uncertainty in deep learning. 2016.

[] S. Gasperini, J. Haug, M.-A. N. Mahani, A. Marcos-Ramiro, N. Navab, B. Busam, and F. Tombari. Certainnet: Sampling-free uncertainty estimation for object detection. *IEEE Robotics and Automation Letters*, 7(2):698–705, Apr. 2022.

[] R. Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015.

[] R. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.

[] B. Settles. Active learning literature survey. 2009.

[] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv preprint arXiv:2207.02696*, 2022.

[] B. Yang, J.-T. Sun, T. Wang, and Z. Chen. Effective multi-label active learning for text classification. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 917–926, 2009.

[] X. Zhou, D. Wang, and P. Krähenbühl. Objects as points. In *arXiv preprint arXiv:1904.07850*, 2019.