発表番号16

Study on Reduce of Working Time in Image Data Annotation for Deep Learning through Semi-Automated Active Learning

芝浦工業大学

運転支援システム研究室

NB19108 ヒルマン

指導教員 伊東 敏夫

Shibaura Institute of Technology

Advanced Driver Assistance Systems Lab

NB19108 Mohd Hafiz Hilman bin Mohammad Sofian
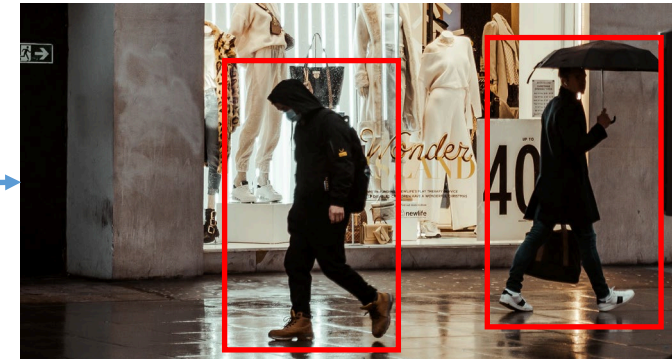
Supervisor Toshio ITO

4 QUALITY EDUCATION

9 INDUSTRY, INNOVATION AND INFRASTRUCTURE

Autonomous Driving Vehicles and Perception System

Stream of images (videos)

Deep Learning

Image + Deep Learning (DL) technology
- Close to human perception
- Cheap
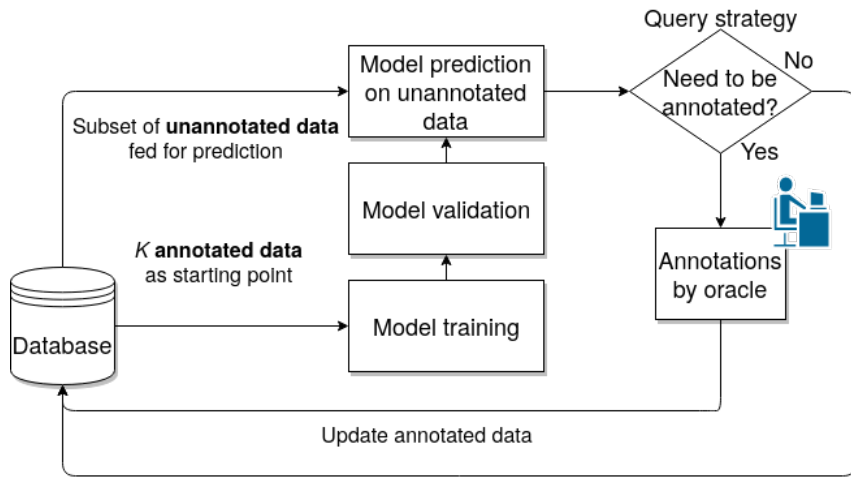- Can take advantage of the big data of the modern world

Problems
- A lot of data need to be labelled by human beings
- This intorduces high costs of money, labor and time

Objective: Solve these problems

Solutions

Existing solution: Active Learning

- A loop system that uses the DL model and "query strategy" to choose only subset of images for labelling and training
- Focus on quality of data rather than quantity of data
- Still needs human power

Proposed solution: Semi-Automated Active Learning (SAAL)

Machine Teachers

Semi-Automated Active Learning
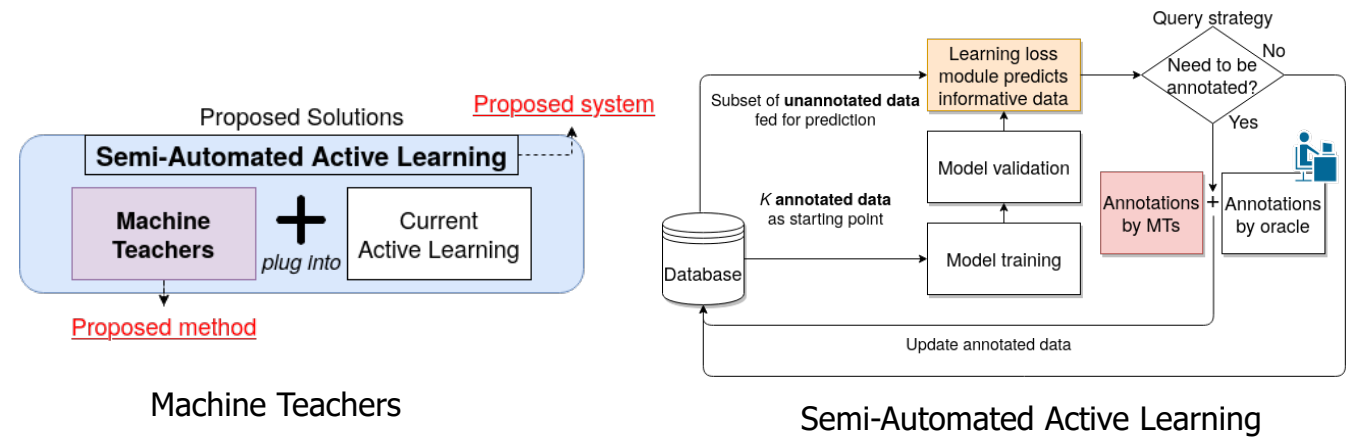
Machine Teachers (MTs)

An algorithm where a pretrained neural network and other image processing algorithms are stacked upon each other to help human in annotating image data from video

Semi-Automated Active Learning (SAAL)

AL system with the MTs plugged in. This allow SAAL to be automatized to further reduce the costs introduced by conventional AL system
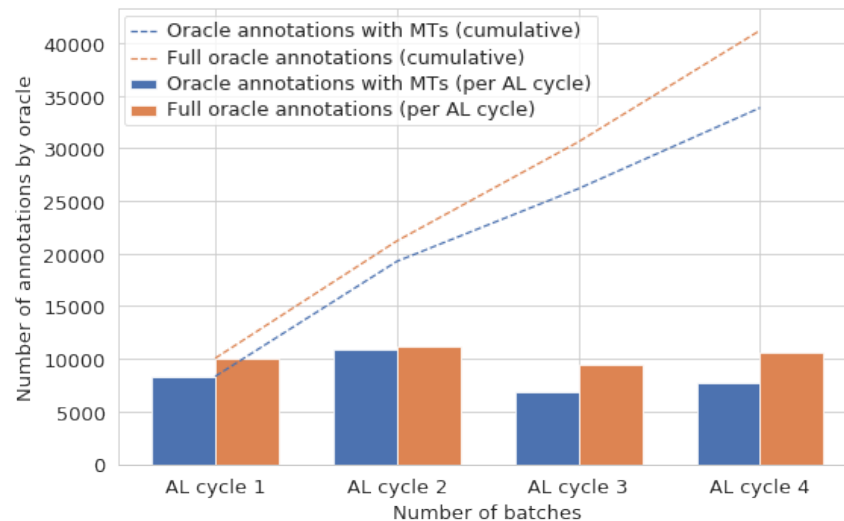
3

Experiment, Results and Discussion

1. Train DL with AL and SAAL frameworks with image data from vehicle dash recorder
2. Compare the accuracy result (mAP result) and number of works done by human being

mAP Results

Comparision of amount of works by human being
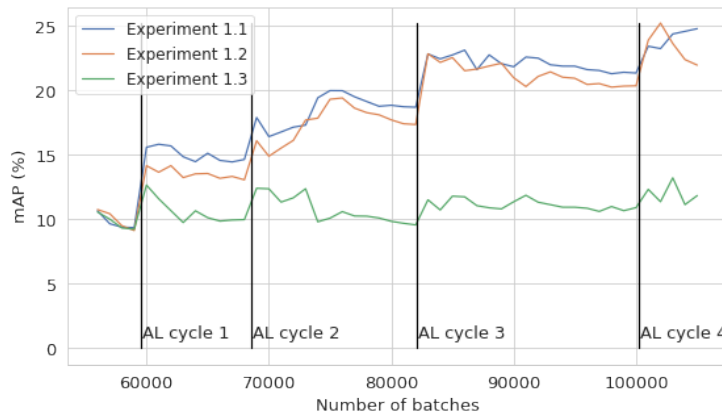
Experiment 1.1: AL
Experiment 1.2: SAAL

Discussion and Conclusion

1. SAAL managed to have accuracy close to AL while managed to reduce the amount of human power by 17.85%
2. Implies SAAL is effective for reducing the need of human power in data annotation

Future Works

1. Improve MTs to be used with other kind of datasets
2. Construct a methodology on how to develop MTs based on one's use

Advanced Driver Assistance Systems Laboratory