Project in Deep Learning (236502)

Lecturer in charge: Prof. Ran El-Yaniv

Instructor: Ido Galil

Prerequisite: Any credited deep learning course offered at the Technion or elsewhere. For example; 236781, 046211, 097200.

Registration: Manual by the instructor. To apply to this course, please fill the following form:

https://docs.google.com/forms/d/e/1FAIpQLSerHFQ-0D4FzZwHXWHFRkZXLh_PrtMoDBpvlNKixgBEo6_XIQ/viewform?usp=sf_link

The number of available slots is limited. Students will be notified if they have been accepted into this course by March 15th.

The course is offered to both undergraduate and graduate students.

Description: In this course, the students will engage in mini-research projects conducted in small teams of 1 to 3 students. Under the guidance of the instructor, and in conjunction with the students' skills and interests, each team will select a project from the current list of projects.

For each project, an initial idea and a direction will be provided to the team, with basic guidelines for the experiments and work needed to achieve a proof of concept (POC). Students are encouraged to expand on this basis and innovate.

Note that the projects will require a strong sense of independence and self-learning abilities. A good background in deep learning is required. All projects will focus on Computer Vision, and experiments will be conducted on image datasets such as CIFAR-10, and ImageNet.

The goal of this course is to deepen the knowledge and understanding of specific topics in deep learning and to provide a taste of research in the field. The projects are designed to include research components including writing a mini research proposal, conducting a brief literature review, designing and defining benchmark tests and writing a summary report. Each team will conduct a weekly or bi-weekly meeting with the instructor. In these meetings the team will report on their progress and present the obstacles and problems encountered. Guided by the instructor, the team will plan their next steps and short term goals.

Current List of Projects

*Many of the projects are related to *deep uncertainty estimation*.

Deep neural networks show great performance in a wide variety of application domains. Successful deployment of these models in risk-sensitive applications (such as medical applications, or autonomous vehicles), however, is critically dependent on providing an effective uncertainty estimation of their predictions in the form of either rejecting some inputs (the model's way to say "I don't know"), or by providing a probabilistic confidence score for their predictions. To get a quick background on uncertainty estimation and how it is measured (selective prediction, ranking, calibration), it is recommended to watch the first 12 minutes of the following seminar talk:

https://www.youtube.com/watch?v=cqkX4H-7RNE&t=236s

1. Knowledge Distillation – Researching why it improves uncertainty estimation

Knowledge distillation is a training regime involving a "teacher" model (an already trained model), and a "student" model (the model being trained). While originally intended for distilling the knowledge learned by large models into smaller models, it has since been discovered to improve the models' generalization and accuracy.

We recently discovered this regime to furthermore be the single best training regime for improving uncertainty estimation performance, by any aspect.

This project will focus on finding out how and why this regime accomplishes this, and attempt to further improve it for the purpose of training models with better uncertainty estimation.

Knowledge distillation and its affects on uncertainty estimation are briefly mentioned in the seminar talk provided above (especially at minute 20:00)

2. Deep Enhancers

A novel idea with only few related works. Enhancers process the inputs to the neural networks with an aim to improve uncertainty estimation performance (and perhaps even accuracy). The suggested direction utilizes "*adversarial attacks*" to achieve this (one special type of adversarial attack is discussed in the seminar talk provided above at minute 25:00).

3. Classification and uncertainty estimation with side information

This project explores the idea of utilizing expert knowledge about the data. It has two potential objectives:

1) Improving the model's performance (accuracy and\or uncertainty estimation).

2) Allowing the model to express its uncertainty about an input in a way humans can understand.

4. ExclusiveNet

This project revolves around constructing an architecture that should be resilient to a new type of an attack. The initial direction involves constructing an architecture inspired by DenseNet.

5. Augmentation:

CutMix is an extremely popular augmentation and is used in many training regimes. Yet, it suffers from some significant drawbacks. This project's objective is to improve the CutMix algorithm by employing a segmentation model.

6. Improving performance by introducing "communication" between models

This project explores the idea of improving an already trained model's accuracy by introducing a novel mechanism allowing for "communication" with other models. The suggested direction utilizes "*adversarial attacks*" to achieve this.

Schedule

Important milestones:

- Week 1:
    - Meeting all teams together and presenting the available topics and technical details (how to use our GPU server).
    - Team meeting: setting up project objectives and initial steps.
- Week 4: Presenting a short mini-research proposal, including a brief literature review and defining benchmark tests.
- Week 7: Mid-semester presentation. Presenting initial results, difficulties, and steps toward completing the project.
- Week 13: Presenting the project and submitting the project's summary report. **Subject to the instructor's approval. This presentation could be moved to after the exams period.*