TDT4265

SOTA Paper Challenge

Computer Vision and Deep Learning

Delivery Deadline: **Friday, April 26th, 23:59PM.**
This assignment counts for **5%** of the final grade
You are allowed to work in groups of 1 or 2 people.

Task:

- Find a vision-based paper published in 2024 (2023 if you must) and try to explain the underlying method with your own words based on what you have learned.
- The focus should be on explaining the method/model itself (illustrations can be very useful). Some information about the background / motivation / CV-task addressed, as well as the performance (accuracy and speed), can be useful, but should not take more than 45 sec. of the max time in total.
- The ability to explain the chosen method / model is the only evaluation criteria. The complexity of the method / model will be taken into account. This is also the case for the number of participants in the group, where we look for the extent of the explanations.
- The model/method/ paper that you choose to explain cannot be the same as the key model you use in your practical project.
- All models/papers with a clear vision-component are ok (for example, in addition to large vision models, large vision-language models and multi-modal foundation models would also be allowed).
- Besides that, you are free to do whatever you want, inc. looking at relevant blog posts about the method, but try to make your exploration your own.

Deliverable:

A video with speech of your slides no longer than 5 or 6 minutes (for 1 or 2 participants respectively). The video presentation should be uploaded to blackboard within the deadline.

Potential Sources:

- **arXiv**:
  - https://arxiv.org
  - https://arxiv.org/list/cs.CV/recent
- **Papers with Code**:
  - https://paperswithcode.com
  - https://paperswithcode.com/methods/area/computer-vision
- **X/Twitter**:
  - https://twitter.com