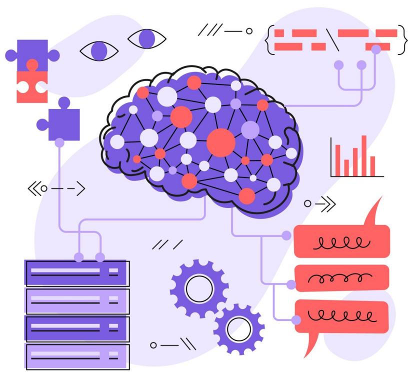
Mini-Project
Visual Intelligence

Delivery Deadline:
Nov 29th, 23:59PM.
This project counts for **45%** of the final grade
You are allowed to work in groups of 1 or 2 people.

1 Introduction

This mini project is supposed to simulate a potential real-world scenario in some months, i.e., your first assignment, in your new ICT-consulting job, helping an external customer with some visual intelligence (VI) related tasks. The information provided is therefore limited to what you can find below. But you are free to solve the task the way you want, using whatever tools you find useful, just remember to report what you have done. Also, here you will be able to choose between the projects proposed below.



Group-sizes can be one or two participants.

It is possible to choose between three projects on the topics: AD (autonomous driving), MIC (medical image computing) and football tracking. However, all projects deal with core visual intelligence (VI) tasks that can be applied in other domains as well.

1.1 Presentation Structure

The projects should be presented using the following structure: Background / Motivation, Approach / Strategy, Data-analysis, Methods/Models, Results, Discussion, Key learning points, References and who did what, if the group contains two participants.

1.2 Sustainability

Sustainability is supposed to be part of all courses now. The way we will do that in this module is as follows; you will make a note of the combined compute time needed to achieve your results (approximately, training mostly, but also inference). The compute-time can be converted to energy, and the needed energy could have alternatives uses, for example, how far, from Trondheim to Oslo, can you get in your brand new fully electric car, e.g., a Tesla Model Y (most sold car in Norway), using the same amount of energy.

1.3 Project Rules

To achieve a fair project (and grading..), we ask you to follow these rules:

1. The code should be developed by yourself, but you are free to use open-source repositories and libraries. If you take code from anywhere else, please attribute the original authors in your source code.

2. You are allowed to use any open-source model architecture from PyTorch or MONAI etc. You may also use pre-trained model weights, given that you fine-tune the model for your project.

3. You are allowed to use YOLO models, but given the low complexity of implementation, we expect more in other areas like exploratory data analysis and your understanding of the model.

1.4 Computing Resources

These projects will typically require a lot of computing power in the form of GPU's. If you don't have access to a modern Nvidia GPU, you are welcome to use one of the computing resources available to you via NTNU.

IDUN Cluster

As attendees of this course, you should now have access to our compute cluster IDUN, where you can use your NTNU username and password to log in. This cluster has a lot of computing power but requires a bit of learning in order to utilize it. We recommend the following guide to get you started: IDUN Guide

More information about the cluster can be found at their home page: https://www.hpc.ntnu.no/idun/

Please note that this cluster uses a scheduling system called Slurm, this means you request computing resources and are put into a queue. Typically, at the end of the semester there are many students / researchers using the cluster, so the waiting time can be quite long at times. If you choose to use this option, you should dedicate some time into learning how to use it, and conduct your experiments in good time before the project deadline to account for long queues.

Cybele Lab

In addition to IDUN you should also have access to our Cybele lab located at: MazeMap link. You can access the lab using your student card and pin code and use your NTNU username and password to log into the computers. The lab has 25 powerful Linux workstations featuring an Nvidia RTX 4090.

Please note that the lab is also being used by other courses, so do **not** use the lab at these hours:

- Mondays 14:15 - 19:00
- Tuesdays 14:15 - 16:00
- Wednesdays 14:15 - 16:00

2 Project options

In this section, the project options are described, allowing you to tailor your learning experience based on your interests. Each option will challenge you to develop specialized computer vision and deep learning expertise. **You should only select one of the following, and the selection should be slightly different from your specialization project!**.

2.1 Option 1: Snow Pole Detection with images from Trondheim

There has been major advancements in Autonomous driving in recent years, attributed to computer vision and Deep Learning (CV & DL). However, such models are mostly focused on driving in ideal conditions, and thus struggles with challenging conditions like snowy roads. One way to localize the road in winter time is by relying on the location of snow poles, which are typically erected on either side of the road in areas prone to snow in the winter. Your task here is to perform real time object detection of snow poles, in order to further develop AD capabilities in winter conditions.

Figure 1: Sample with labels from the Poles2025 dataset

- **Dataset:** You'll utilize data captured in the Trøndelag region. There are 1 bounding box class (snow poles) in the accompanying labels, in the YOLO format. Please note that the test set labels are kept hidden in order to make a leaderboard which will be announced before the project delivery deadline. The RoadPoles-MSJ does not contain any labels, but you may use it for qualitative assessment of your model or as additional training data (AI generated).

  The dataset can be found at these locations:

  - IDUN: */cluster/projects/vc/courses/TDT17/ad/Poles2025*
  - Cybele: *datasets/TDT17/ad/Poles2025*

  Please note that redistribution of the dataset is prohibited!

- **Performance Metrics:** We expect you to calculate the following metrics on the test set:
  - Precision
  - Recall
  - mAP@50
  - mAP@0.5:0.95
- **Model size:** As the model is ultimately supposed to run on an edge device for real time object detection, you should select a model that is suited for this task. Tiny or Small versions of YOLO could be a good fit in this regard.
- **Expectations based on group size**:
  - Groups of 1: Implement at least one architecture.
  - Groups of 2: Implement at least two architectures (where only 1 can be a YOLO variant).

2.2 Option 2: Medical Image Analysis

Breast cancer remains the most common cancer among women. Although mammography is the standard screening tool, Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI) offers greater sensitivity in dense breasts and high-risk populations. However, reading large 3D MRI volumes is time-consuming and variable across readers. The ODELIA challenge aims to catalyze the development of AI methods that can **generalize across multiple centers, scanners, and acquisition protocols** to support clinical decision making.

Task: For each MRI study, **classify each breast (left and right) into one of three classes:** `normal`, `benign`, or `malignant`. Your model should output calibrated **probability distributions (softmax)** per breast, e.g.:

```
{"right": {"normal": 0.001000, "benign": 0.01000, "malignant": 0.988000},
 "left":  {"normal": 0.987000, "benign": 0.02000, "malignant": 0.001000}}
```

- **Dataset:** A curated, multi-centre DCE-MRI dataset spanning multiple vendors and field strengths. Please note that labels for the RSH subset of the dataset will be kept hidden for our leaderboard that will be announced before project deadline.
  - IDUN: */cluster/projects/vc/courses/TDT17/mic/ODELIA2025*
  - Cybele: */datasets/TDT17/mic/ODELIA2025*

  Please note that redistribution of the dataset is prohibited!

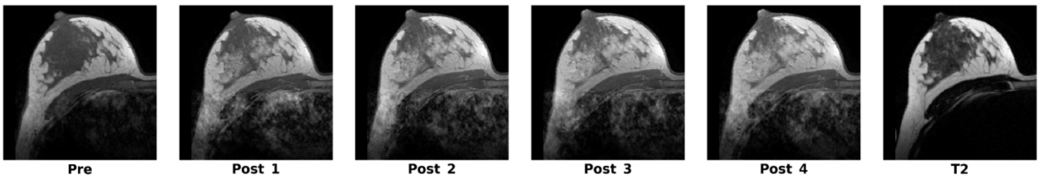


Figure 2: Axial slice of the pre-contrast, first to fourth post-contrast T1w sequences, and the T2w sequence from the "unilateral" dataset.

- **Performance Metrics:**
  - Primary: Area Under Reciever Operating Curve (AUROC) aka AUC for malignant-vs-non-malignant classification (evaluated per breast and aggregated as specified by the challenge).
  - Operating-point metrics: **Specificity at 90% sensitivity** and **Sensitivity at 90% specificity**.
  - Feel free to use the evaluation code in the dataset folder for this purpose.
- **Recommended tools:**
  - MONAI (PyTorch) for 3D/4D preprocessing, model training, and evaluation.
  - 3D Slicer for loading/inspecting DCE series, drawing QA ROIs, and visualizing attribution maps.
- **Expectations based on group size:**
  - Groups of 1: Train at least 1 model for breast cancer classification.
  - Groups of 2: Train at least 2 models for breast cancer classification.
- **Additional information:**

- Official challenge site: ODELIA 2025 (Grand Challenge)
- Background and updates: ODELIA Consortium page
- Dataset Details: A European Multi-Center Breast Cancer MRI Dataset

2.3 Option 3: Football analysis

Harness the power of deep learning to analyze football match footage, extracting valuable insights for RBK (or their opponents). The main tasks here will be to detect and track the players / referee and ball.

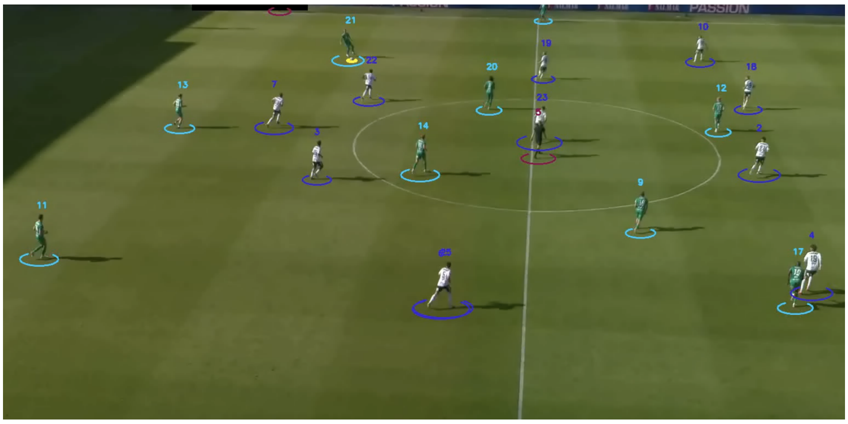


Figure 3: Sample with labels from the dataset

- **Datasets:** The datasets are meticulously annotated with 2D bounding boxes for players, referees, and the ball. Each player and referee has a consistent tracking ID throughout the sequence, even when temporarily out of camera view (marked with a "not visible" property).
  - IDUN: */cluster/projects/vc/courses/TDT17/other/Football2025*
  - Cybele: */datasets/tdt17/other/Football2025*

  Please note that re-distribution of the dataset is prohibited!
- **Main Tasks:**
  - Object detection
  - Tracking
- **Performance Metrics (Detection):**
  - Precision
  - Recall
  - mAP@50
  - mAP@0.5:0.95
- **Performance Metrics (Tracking):**
  - HOTA: A Higher Order Metric for Evaluating Multi-Object Tracking
- **Expectations based on group size**:
  - Groups of 1: Focus on object detection, but feel free to explore tracking as well.
  - Groups of 2: Object detection and tracking.

3 Deliverables

Documenting and reporting your approach is an important part of any deep learning project. This will consist of a video presentation. Students working alone will have maximum 12 minutes to present, groups of two will have 14. Please make sure to follow the presentation structure guidelines in section 1.1, and include a part about sustainability as mentioned in section 1.2.

In addition to the video presentation, we may need to review your source code for clarification. You can either provide a link to your GitHub repository or keep the code locally on your computer, ready to share if requested during grading.

Your presentation should be uploaded to teams (**exact location will be announced later**).

Please note that we may request a teams meeting for additional clarifications.

4 Evaluation

As your submissions will likely be very diverse, we will evaluate the projects based on the overall impression of your work based on your video presentation.

Some of the things we look for:

1. Understanding of the dataset
2. Model development is guided by exploratory data analysis
3. Understanding of the chosen method
4. Complexity of the method
5. Thoroughness in the work
6. The presentation is clear and easy to follow