June-July Exam session Projects List

Objectives

The goal of the project is to tackle one of the proposed topics in the field of Computer Vision, developing non-trivial solutions. Students should explore the problem from an original perspective, applying methods that go beyond conventional solutions and demonstrating critical thinking and problem-solving skills.

Lines of conduct

- **Student groups:** The project can be carried out by a group, each consisting of a maximum of 3 people. Projects can also be completed by individual students, but we suggest to work in team.
- **Notebook format:** The project must be implemented using a notebook (e.g., Google Colab, Kaggle) or with an IDE (e.g., VSCode, PyCharm). The code should be optimized to support GPU usage and run without any error. The delivered code must follow the structure outlined below:
  - *Imports*: all the needed packages (for the notebook format)
  - *Globals*: useful variables on the whole code
  - *Utils*: code support functions
  - *Data*: everything related to data management
  - *Network*: code to structure the neural network
  - *Train*: part containing the training cycle elements
  - *Evaluation*: tests needed for the trained network

  You can find a sample template at this link. Try to maintain as much as possible this conceptual structure.
- **Deep learning framework:** All projects **MUST** be done in Python via the Pytorch framework.
- **Project assignment:** You are required to choose a project through this Google Form. In this form, you will provide information about your team and the project, and include the link to the project's GitHub repository. In this repository you have to upload:
  - Code (or notebook) implementing the project
  - Dataset (or a link to it)
  - Project presentation
  - Detailed README to provide a quick overview of the project and instructions on how to run it
- **Project submission:** The project must be presented on one of the exam dates. It can be presented at a different time than the written exam. Both the written exam and the project MUST be completed within the academic year (i.e., between the June 2025 session and the March 2026 session).

  October and March session are reserved to "categories of students referred to in Article 40, paragraph 6, of the General Study Manifesto, and out-of-school students enrolled for the A.Y. 2024-2025 in the third year of a Bachelor's degree and in the second year of a Master's degree".
- **Plagiarism:** Any attempt to plagiarize, whether by copying other students' work, directly replicating code from online resources, or submitting content highly retrived from generative AI models, will be strictly penalized. This course values originality and personal effort; therefore, students must submit independently developed solutions. On the other hand, it is acceptable to consult external resources for inspiration or guidance.

Project 7: Efficient Anomaly Detection in Industrial Images using Transformers with Dynamic Tanh

Abstract: In industrial environments, detecting anomalies in visual data is crucial for maintaining high standards of quality and operational safety. Traditional image analysis methods often face limitations when dealing with complex, high-dimensional data. Recent advancements in computer vision, particularly in transformer-based architectures such as Vision Transformers (ViTs), have shown great potential in capturing rich spatial features from images. At the same time, new techniques like Dynamic Tanh (DyT) offer promising solutions to improve the computational efficiency of these models. The combination of powerful feature extraction and efficient processing represents a significant opportunity for advancing anomaly detection systems in industrial applications.

Dataset: BTAD, MVTec Anomaly Detection Dataset

Task: The aim of this project is to explore and implement an advanced approach to anomaly detection in industrial images by combining two cutting-edge techniques in computer vision: Vision Transformers (ViTs) for feature extraction and Dynamic Tanh (DyT) for improving transformer model efficiency. This proposal leverages the strengths of ViTs in capturing spatial dependencies alongside the efficiency enhancements offered by DyT as a replacement for traditional normalization layers. Students will focus on both anomaly detection and localization tasks, applying the DyT method to improve network speed and overall model performance on industrial datasets. The project will involve designing and testing an end-to-end system that integrates ViTs for image analysis with DyT-based efficient transformers. Additionally, students will evaluate the efficiency benefits of DyT by comparing runtime and performance metrics against models using traditional normalization layers.

Main objectives:

- *Implement DyT for Improved Efficiency*: Replace traditional normalization layers in transformer models with the Dynamic Tanh method to assess improvements in model efficiency
- *Evaluate Performance*: Evaluate the performance of the model in anomaly detection and localization tasks using appropriate metrics.
- *Baselines comparison*: Benchmark the proposed solution against existing and not efficient transformer models, to highlight the performance advantages of using DyT in Vision Transformers.

References:

1. Mishra, P., Verk, R., Fornasier, D., Piciarelli, C., & Foresti, G. L. (2021, June). VT-ADL: A Vision Transformer Network for Image Anomaly Detection and Localization. 2021 IEEE 30th International Symposium on Industrial Electronics (ISIE), 01–06. doi:10.1109/isie45552.2021.9576231
2. Zhu, J., Chen, X., He, K., LeCun, Y., & Liu, Z. (2025). Transformers without Normalization. arXiv [Cs.LG]. Retrieved from http://arxiv.org/abs/2503.10622
3. Liu, J., Xie, G., Wang, J., Li, S., Wang, C., Zheng, F., & Jin, Y. (2024). Deep Industrial Image Anomaly Detection: A Survey. Machine Intelligence Research, 21(1), 104–135. doi:10.1007/s11633-023-1459-z