# ECE371 Neural Networks and Deep Learning

**Final Project**

**Due Date: 23:59, 6 July, 2025**

This is a group project designed for 2 to 3 members, which will account for 40% of the final grade.

## 1. Project Topics

### 1) Object Detection Project

Object detection is a computer vision task that involves detecting objects of interest within an image or video and localizing them by drawing bounding boxes around them. The development of object detection algorithms has been driven by the need for machines to perceive and understand the world in the same way as humans. Object detection has a wide range of applications, including autonomous vehicles, surveillance systems, robotics, and image retrieval. Object detection is an active area of research, and there are many challenges to overcome, such as detecting objects at different scales and orientations, dealing with occlusion and clutter, and detecting objects in complex scenes. However, with the continued advancement of deep learning techniques, it is expected that object detection algorithms will continue to improve and find more diverse and practical applications

- **Task Description**

You are required to finish the VOC2012 object detection competition. Specifically, there are 20 categories including: Aeroplanes, Bicycles, Birds, Boats, Bottles, Buses, Cars, Cats, Chairs, Cows, Dining tables, Dogs, Horses, Motorbikes, People, Potted plants, Sheep, Sofas, Trains, TV/Monitors. We provide the object detection data in VOC2012_det.zip. There are 5823 validation samples and 5717 training samples, please use the validation set as the test set ( you can use a subset of the validation set ). You can also access the full dataset containing other tasks on the competition home page: http://host.robots.ox.ac.uk/pascal/VOC/voc2012/.

Please note that /VOCdevkit/VOC2012/ImageSets/Main contains the list for object detection task.

The Google Drive link for dataset is as follows:

https://drive.google.com/file/d/1aLnY1PYFFBmR1NUEBdULXFxVx7iJQkzJ/view?usp=sharing

- **Suggestions**

Since then, numerous other algorithms have been developed, such as Fast R-CNN, Faster R-CNN, and YOLO, which have significantly enhanced the accuracy and speed of object detection. Nowadays, many models based on the Transformer architecture have also emerged. The Pascal VOC is a classic object recognition competition, and you can find a wealth of materials online. While it is acceptable to use some state-of-the-art (SOTA) methods, you need to provide proper references, present your own understanding of the methods, and conduct your own attempts (e.g., by changing some settings or performing ablation experiments) to thoroughly investigate the methods you are using.

## 2) Semantic Segmentation Project

Semantic Segmentation is an important task in the field of computer vision, which aims to assign each pixel in an image to its corresponding semantic class. Unlike object detection, which only needs to determine the existence and position of objects in an image, semantic segmentation requires pixel-level classification of the image. Semantic segmentation has applications in various fields. For example, in autonomous driving, semantic segmentation can separate different objects such as roads, vehicles, and pedestrians, which helps the autonomous driving system make more accurate decisions. In medical image analysis, semantic segmentation can separate different tissue structures, which helps doctors make more accurate diagnoses.

- **Task Description**

In this project, you need to use deep models to perform semantic segmentation on humans. Dataset is in Pascal_seg.zip. In specific, you should segment 7 semantic parts including head, torso, upper-arms, lower-arms, upper-legs, lower-legs, and background. The dataset is provided, which contains the training and validation sets.

The Google Drive link for dataset is as follows:

https://drive.google.com/file/d/1dvQ6d7Bsf_esVEp74VIqTANIuYCln-C6/view?usp=sharing

## 3) Machine Translation Project

Machine translation is the process of using computer algorithms to automatically translate text or speech from one language to another. It is a subfield of natural language processing (NLP) and has been an active area of research for several decades. Recently, with the development of deep learning and neural networks, neural machine translation (NMT) has emerged as a powerful approach to machine translation. NMT leverages deep neural networks to learn a mapping from the source language to the target language, and it has been shown to produce better translations than traditional statistical methods. Machine translation has a wide range of applications, including language learning, cross-cultural communication, and international business. However, machine translation is not perfect and can still produce errors and mistranslations, especially for complex sentences and idiomatic expressions. Nonetheless, it continues to improve and has become an increasingly important tool for global communication.

### ➢ Task Description

In this project, we will employ neural networks to perform Chinese-to-English translation. For instance, if a user inputs "祝同学们最后的大作业顺利!", the system should be able to generate a translation like "Good luck with your final project!". The dataset for this project is provided in the form of two JSON files: one for training, named train.json, and the other for validation, named valid.json. You can distinguish between the two based on their file names.

The Google Drive link for dataset is as follows:

https://drive.google.com/file/d/1gGnpS0fQBNcTAwhpYNGjxJJwsHYQzKag/view?usp=sharing

### ➢ Suggestions

(1) The tokenizers from the SpaCy package are highly recommended, including both the Chinese and English versions. (2) You can start by experimenting with some basic models, such as RNNs or transformers, to establish a baseline. (3) Given the current popularity of models like ChatGPT and GPT-4, you might also consider leveraging open-source pre-trained large language models (LLMs), such as PaLM or LLaMA, if you have access to sufficient GPU resources.

### 4) Self-selected Topic

We provide students with the capacity for further learning the space for independent exploration. You may choose a cutting-edge project related to deep learning, which is not related to any of the aforementioned projects. If you choose a self-selected topic, please contact Professor Ruimao Zhang via email to confirm the content of the topic. The relevant topic must not be submitted to other related courses.

## 2. Requirements

- Each project team is required to select one of the aforementioned projects to conduct the experiment. You should integrate the content covered in the lectures of the course with the relevant information available on the Internet in practice, and utilize Python to complete all of the experiment.
- We do not provide any baselines; you are expected to conduct the investigations and compile the report independently. It is mandatory to establish a baseline and undertake sufficient efforts to enhance the results. You have the freedom to experiment with any model architectures, techniques, and even pre-trained models are permitted within the scope of this project. Creativity and novelty are highly encouraged. Ideally, you should modify your baseline based on your understanding of the topic. While off-the-shelf methods are acceptable, you must provide analyses to demonstrate their utility to your results.
- Our assessment of your project is not based on performance alone. What truly matters to us is your comprehensive understanding of the task. We encourage you to delve deeply into the project, conduct thorough experiments, and strive to achieve better results. For specific evaluation criteria, please refer to the detailed metrics outlined on the course website：http://zhangruimao.site/ECE371.html
- When investigating this topic, please pay close attention to the metrics employed by others in the field. It is crucial to select appropriate and reasonable metrics to evaluate performance. Ideally, your experiments section should present both quantitative results, which provide objective numerical data, and qualitative ones, which offer subjective interpretations.

- Write a report with the following template: https://www.overleaf.com/read/vjsjkdcwttqp#ffc59a (see Final_Project.tex). It contains Introduction, Related work, Method, Experiment, Conclusion, and Reference.

## 3. Submission

- You are required to submit your code, report, and models to GitHub Classroom before **July 6, 23:59 (the last day of this semester)**. The submission link for the final project will be provided shortly.
- We will arrange a time slot for students to present their projects to the teaching assistants and lecturer one by one. **You may decide whether to prepare Slides based on your needs**, but you must explain your ideas, analyze your results, and so on. We may ask you some questions. We will not tolerate plagiarism.