Introduction to Deep Learning – Spring 2023

Term Project Proposal

Rodrigo Picinini Méxas

Topic:

Classification (using CNN and Landmarks) and Object Detection for Sign Language

Goal:

The goal of the project is to train one classifier using CNN, one classifier using hand landmarks and one object detector using a sign language dataset. With that done, a comparison and analysis of the obtained results will be made.

Solution Method:

- Utilization of the dataset: 27 Class Sign Language Dataset (https://www.kaggle.com/datasets/ardamavi/27-class-sign-language-dataset). The numbers 0 to 9 and the letters A, B, C, D and E will be used. However, for training the object detector, the dataset from https://datasets.cms.waikato.ac.nz/ufdl/american-sign-language-letters/ will be used, as they contain the annotation of the bounding boxes. In this case, the letters A, B, C, D and E will be used.
- For organizing the dependencies, a conda environment will be used.
- For the classification, libraries such as Matplotlib, Numpy, Pillow and OpenCV will be used.
- When training the classifier with the CNN, Pytorch will be used.
- When training the classifier with landmarks, MediaPipe will be used to extract the landmarks of the hand and then Pytorch will be used to train a simple DNN using as input the hand landmarks.
- For the object detection, YOLO will be used.
- After training and testing, an evaluation will be conducted and relevant information, images and graphs will be presented, such as the loss curves and confusion matrix.

Application:

Through this project it's possible to learn on practice the theory learned on class about classification, Convolutional Neural Networks, pose estimation (landmarks) and object detection. Also, it's an application that could help speech-impaired people communicating with people that don't understand the sign language.