Course DATA586 Project

Project Description

In this project, we will explore the use of advanced machine learning techniques (e.g. Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs)) for computer vision tasks. We will mainly use deep learning platforms (e.g. PyTorch, Tensorflow) to build, train, and test our models.

The scope of this project involves image classification, object detection, and anomaly detection. You are welcome to choose the one that you are most comfortable with as the project topic.

To accomplish the project, generally you will need to implement the following procedures:

- Dataset options: For your convenience, I have designated a specific dataset for each computer vision task. Please conduct your project based on the assigned dataset. However, if you have a customized dataset and you are keen on utilizing it for the project, kindly inform me. Otherwise, I would appreciate it if you utilize one of the provided datasets.
- Preprocessing: Please do preprocessing (e.g. resizing, normalizing, and augmenting, etc.) to create a training set and a validation set.
- Building a model: Please build a proper model and train it on the dataset, and evaluate the model's performance on the validation set and test set.
- Tuning hyperparameters: Please tune the hyperparameters of the models (e.g. the learning rate, weight decay, dropout rate, and batch size, etc.) to improve their performance on the validation set.
- Recording and Visualizing results: Please record the training process and visualize the model results.
- Discussion of Project : Please recap the whole project. Describe the difficulties you encountered and how you solved them. Analyze model performance and the strategies you used for performance improvement.

Deliverables

The final deliverables for the project will include:

1. A report summarizing the project, including the dataset used, the models built, the hyperparameters tuned, the results obtained, and the contribution of group members if applied. Don't forget to acknowledge any sources of inspiration, guidance, or code that you used in your project.

The content of a report includes

- Abstract
- Introduction[0.5page]
- Literature Review[1-2page]andBackground[1-2pages]

- Methodology[2-3pages]
- Experiment and Results[2-3pages] and References[<1page]

The project report will be submitted in the form of a 6-8 page report in IEEE/ACM format. https://www.overleaf.com/latex/templates/acm-conference-proceedings-new-master-template/pnrfvrrdbfwt

2. A Python code repository containing the implementation of the models, along with the scripts used for result replication.

3. A presentation summarizing the dataset information, model architecture, project results, and key findings and insights of the project.

Project Evaluation

In this course project, you will be evaluated based on the performance of your trained model, as well as the thoroughness of your analysis and the presentation. While the performance of your model is an important factor, it is not the only thing that will be considered in the project evaluation.

To receive a high evaluation score, you should carefully analyze your model, including the hyperparameters, training strategy, and evaluation strategy that you used. You should provide detailed explanations of why you made the choices you did, and how they affected the performance of your model.

In addition to the analysis of your model, you should also conduct experiments in a thorough and thoughtful way. This means carefully designing your experiments, including a baseline model and at least one experimental model with different hyperparameters, data augmentation techniques, or other modifications. You should also evaluate the performance of your models using a variety of metrics, such as accuracy, precision, recall, and F1 score.

Finally, you should clearly and concisely communicate your results and analysis in the report and presentation, including visualizations of your results, and providing insights into what worked well and what did not. Your report and presentation should be well-organized, and provide a clear overview of your methods, results, and analysis.

Overall, your evaluation will be based on both the quality of your model and your thoroughness in analyzing and experimenting with different strategies. Good luck with your project!

Expected Outcome

At the end of the project, we expect to have gained theoretical knowledge and hands-on experience in building and training Neural Networks for common computer vision tasks. We will have a good understanding of the key hyperparameters that affect the performance of the models and how to tune them. We will also be able to visualize the results.

Support Materials

Image classification

Dataset: Caltech 101

It is an image dataset commonly used for image classification tasks. It contains 101 object categories, with each category consisting of around 40 to 800 images, for a total of about 9,000 images. The images are of varying sizes and resolutions, and they were collected from the web and Caltech campus. The dataset includes a variety of objects, such as animals, vehicles, household items, and natural scenes. Caltech 101 is a popular dataset for beginners because it is relatively small compared to other datasets, making it easier to work with on a personal computer.

Please find Caltech-101 dataset in https://data.caltech.edu/records/mzrjq-6wc02

It should also be noted that Caltech-101 is also incorporated into the Torchvision library, and you can access to the dataset via https://pytorch.org/vision/main/generated/torchvision.datasets.Caltech101.html

Object detection

Dataset: PASCAL VOC 2012 (Visual Object Classes)

It is a widely used benchmark dataset for object detection. It was first introduced in 2005 and consists of images with annotations for 20 object classes, including person, car, cat, and bicycle. The dataset is widely used in computer vision research and provides a standardized way of evaluating object detection models.

PASCAL VOC 2012 contains about 11,000 images for training and validation, and 20,000 images for testing. Each image in the dataset is annotated with object bounding boxes, indicating the location of the object in the image, as well as the object category label.

PASCAL VOC 2012 for detection is incorporated into the Torchvision library, and you can access to the dataset via https://pytorch.org/vision/stable/generated/torchvision.datasets.VOCDetection.html#torchvision.datasets.VOCDetection

Anomaly detection

Dataset: European Card Data (ECD)

The goal of credit card fraud detection is to identify fraudulent transactions made using stolen or compromised credit card information. The ECD dataset contains two days of transaction data of European cardholders in September 2013. This dataset contains 284,807 samples and 31 features. Out of the given samples, only 492 are fraud cases and account for 0.172% of the dataset.

Please use deep learning models for this task and the dataset can be downloaded from https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud