Question 1 (50 points)

Problem

* A few samples of images for BOX with ground truths are given. Students have to label themselves if there is no ground truth.
* Train an OBJECT DETECTION MODEL that detects the person and box.
* Students can add more data from open source or self collection.

On Submission Day

* The submission date will be on XX November 2024
* Evaluation metrics will be Mean Average Precision (Students have to use this as the objective)
* On submission day, an unseen test dataset (without ground truth) will be provided. Students have to prepare inference code (including data preprocessing, saving results and visualization) ahead to run the model on the unseen dataset.
* The inference must run LIVE, thus students have to make sure your code is working and the outputs are matched with the requirements.
* Specify the public dataset name(s) you added if they are not self-annotated (you don’t need to submit).
* The deliverables are listed below:

1. One brief paragraph about your experiment in pdf (model, data, training, and so on).
2. A set of text files in .txt format (each image has its own prediction file) containing the inference result. The text file name is file1.txt (just change the extension from .jpg to .txt). (The file will be saved and transferred to the examiner LIVE).
3. All the codes, and visualizations have to be submitted in Google Classroom

* The text format of each file should be exactly like as follows (case-sensitive):

Filenames = name1.txt, name2.txt (must be exactly the same as the image names except the .txt extensions)

| 0 0.08359 0.27123 0.01874 0.05709  1 0.95961 0.12464 0.05464 0.06196  2 0.28843 0.23794 0.03569 0.05850 |
| --- |

Between each element is a space

| <class\_name> <x\_center> <y\_center> <width> <height> |
| --- |

Question 2

Problem

* A few samples of images for VEHICLE TYPE with (or without) ground truths are given. Students have to label themselves if there is no ground truth.
* Train a CLASSIFICATION MODEL that predicts the given 10 different types of vehicle types.
* Students can add more data from open source or self collection.

On Submission Day

* The submission date will be on XX November 2024
* Evaluation metrics will be Macro F1 Scores (Students have to use this as the objective)
* On submission day, an unseen test dataset (without ground truth) will be provided. Students have to prepare inference code (including data preprocessing, saving results and visualization) ahead to run the model on the unseen dataset.
* The inference must run LIVE, thus students have to make sure your code is working and the outputs are matched with the requirements.
* Specify the public dataset name(s) you added if they are not self-annotated (you don’t need to submit).
* The deliverables are listed below:

1. One brief paragraph about your experiment in pdf (model, data, training, and so on).
2. A csv file containing the inference result. The csv file name is st123456\_Q1.csv. (The file will be saved and transferred to the examiner LIVE).
3. All the codes, and visualizations for data, validation loss and accuracy

* The csv format should be exactly like as follows (case-sensitive):

| File | prediction | class |
| --- | --- | --- |
| file1.jpg | sedan | 0 |
| file2.jpg | SUV | 1 |

Question 3 (Bonus/ 20 points)

Problem

* From Question 1 DATASET, apart from BOX and PERSON classes, add extra class(es) that are not listed in the ground truth.
* The new classes are according to the student's choice.
* Train an OBJECT DETECTION MODEL that detects the added extra class(es) as well as person and box.
* Students can add more data from open source or self collection but the object from the new class(es) should exist in the provided dataset.
* Students can use labelImg tool for annotations

On Submission Day

* The submission date will be on XX November 2024
* Evaluation metrics will be manually by the examiner.
* 20 unseen images will be provided and the model should predict the newly added class as well as the bounding box including for the person and the box.
* The inference must run LIVE, thus students have to make sure your code is working and the outputs are matched with the requirements.
* The inference also draws the detected bounding boxes on the images with predicted label names.
* Specify the public dataset name(s) you added if they are not self-annotated (you don’t need to submit).
* The deliverables are listed below:

1. One brief paragraph about your experiment in pdf (model, data, training, and so on).
2. A set of text files in .txt format containing the inference result. The text file name is file1.txt (just change the extension from .jpg to .txt). (The file will be saved and transferred to the examiner LIVE).
3. New annotated data
4. All the codes, and visualizations
5. Prediction bounding boxes drawn on the images.

* The text format of each file should be exactly like as follows (case-sensitive):

Filenames = name1.txt, name2.txt (must be exactly the same as the image names except the .txt extensions)

| 0 0.08359 0.27123 0.01874 0.05709  1 0.95961 0.12464 0.05464 0.06196  2 0.28843 0.23794 0.03569 0.05850 |
| --- |

Between each element is a space

| <class\_name> <x\_center> <y\_center> <width> <height> |
| --- |