# Infilect Assignment



The objective of this assignment is to create and deploy an AI pipeline based on the requirements provided below

**AI Pipeline application:**

This AI pipeline application is a service that serves a simple API request and gives back the response in JSON. The application accepts the list of images in a defined format and the output should be the response containing inference through various AI models present as part of the AI pipeline. The goal is to have an application's latency as minimum as possible and scale it to as many users as possible.

**AI Pipeline blocks and its documentation:**

1. Flask Webserver: https://flask.palletsprojects.com/en/2.2.x/
2. Detection Model: Yolo-v5 COCO pretrained object detection model- https://huggingface.co/spaces/nakamura196/yolov5-char/blob/0f967fc973c5b77dbe95cd0cba1d328b14c884a1/ultralytics/yolov5/README.md
3. Classification Model: Pytorch Imagenet pre-trained classification model - https://pytorch.org/vision/stable/models.html
4. Captioning Model: HuggingFace image captioning pre-trained transformer - https://huggingface.co/nlpconnect/vit-gpt2-image-captioning

**AI Pipeline Design:**

This AI pipeline takes in input as an image in a Flask server. The Flask server will send. Once the image is downloaded, the flask should send these images for detection followed by

classification on each detected object and captioning models. In the end, the output would be a result of detection, classification as well as captioning block.

The AI pipeline design and data flow state is defined below.
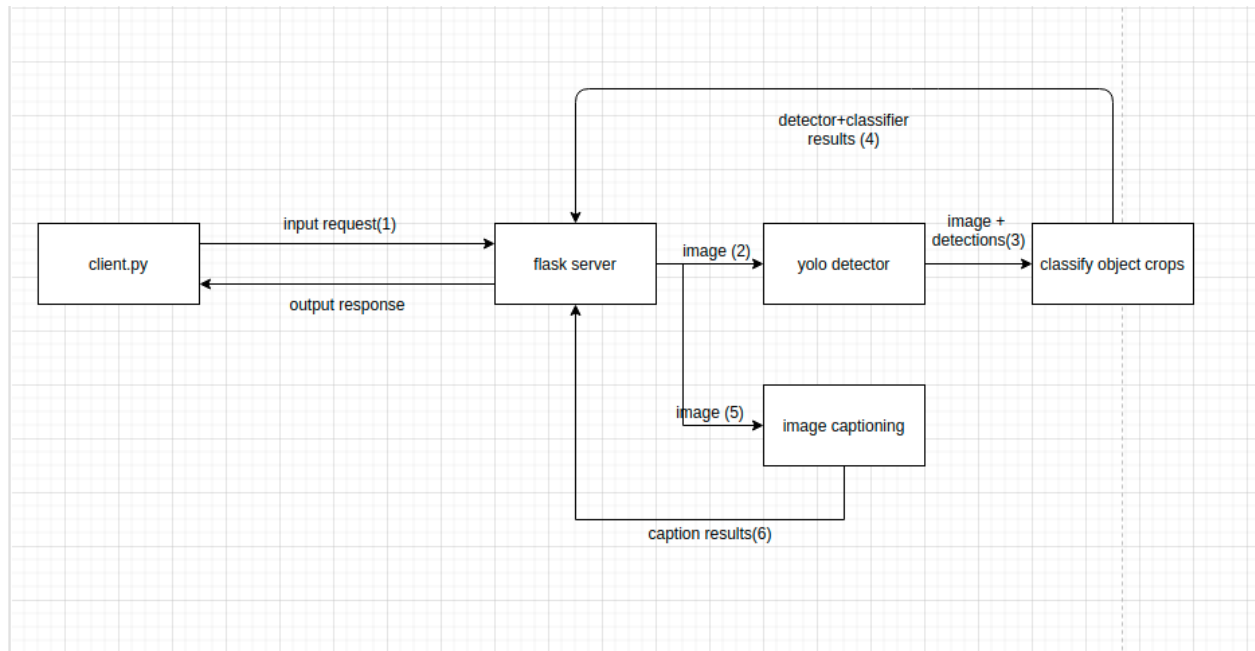


**Note:**
1. Local file paths cannot be used for client-server communication.
2. Class labels from both COCO object detector as well as ImageNet classifier should be saved as part of response.

**Input Format:**

{
"Images": [ "<image1", "<image2>", .... , "<imageN>"]
}

**Output Format:**

Based on the requirements provided above, you can define your output json format for this application which should contain the following for all images.

1. Detection results
2. Classification results
3. Captioning results

## The assignment:

For the assignment, we would like you to do the following:

- Create a Flask server that accepts the input in the given format as above.
- Create microservices/processes for each of the blocks provided in the AI pipeline design block diagram and the documentation provided
- Choose appropriate open source models which works best on given sample images for detection, classification and captioning.

## Deliverables:

- A zip file for the working demo containing all relevant scripts along with the project folder and steps to start this project locally.
- The generic input-output format for the application in the documentation. Define appropriate json formats for input(modify if necessary) and output of each block (flask, detector, classifier and captioning)
- A write-up of the end-to-end steps and documentation for this assignment. A detailed explanation of your solution. It may also contain other approaches you think will work for the problem and some comparisons between them.
- Images with output visualizations(boxes, class labels, image captions)

**Expected workflow:**

1. Imagine below is an input image for the flask server from client



2. The flask server would send this image to the yolo detector which detects all the objects in the image(4 cars). Upon visualization it should look like below



3. The results are sent to image classifier which classifier each of the individual objects and returns labels to each(this case "car" as a label)
4. The final result from detector + classification would have coordinates to all objects from the image and for each object, the output from the classifier is attached.
5. The flask also sends the image to the captioning model which would give a caption to the whole image.

6. Also generate images with visualizations having boxes and labels and captions after processing each image



## Questions

If you have any questions about the assignment or the project setup or are stuck anywhere, feel free to contact us at raghul@infilect.com and raghu@infilect.com.

Finally, you will present your process, technical decisions, and outcome to us. Looking forward to it!

Good luck with the assignment!