YOLO V8 Inference Time:

We tested the YOLO V8 inference on laptop with following specs:

- Memory: 32GB
- Processor: intel Core i7, CPU 2.7GHz
- Graphics: Nvidia 4GB
- Storage: 512GB

Yolo v8 processed the video with an FPS rate of 0.5 seconds (on average).

Yolo v8 processed the same video with FPS rate of 0.3 sec (on average) when we processed it with the GPU parameter.

Hardware Requirements :

YOLO v8 can work real time on the CPU as well. Based on some of our R&D to run YOLO v8, we figured out following optimal hardware specs:

- A Dedicated Nvidia GPU with at least 8Gb VRam
- A CPU with at least 16 cores
- At least 16 GB of Ram
- A local SSD with at least 500 Gb storage

These are minimal recommended requirements for YOLO inference. It generally depends on several factors such as concurrent requests, size of the video and the desired performance

Concurrent Requests Handling :

As per our initial R&D, multiple requests handling mechanism is offered by cloud services, which we are choosing. For example Amazon EC2 and ESC based environments handle multiple requests via parallel processing. These services initiate virtual machines to handle several requests simultaneously.
So the mechanism of handling requests in bulk will be done on the environment, on which we are supposed to deploy the project.

We also tested it on the local system via PostMan by hitting several requests. As it is running on the local system, so there is no parallel processing, like it is available on cloud environments. So it is entertaining one request at time by putting other requests on hold.

Expected Date for Workshop:

As you expected from us to share some tentative date to conduct a workshop to present the implementation of ANN. As we have to work on deployments related tasks, which are supposed to be done by the first week of May.

We have initiated working on this workshop. Some of our team members are actively working on it.

We are targeting to deliver the workshop right after this deployment. We will share the exact date of it as well later on.