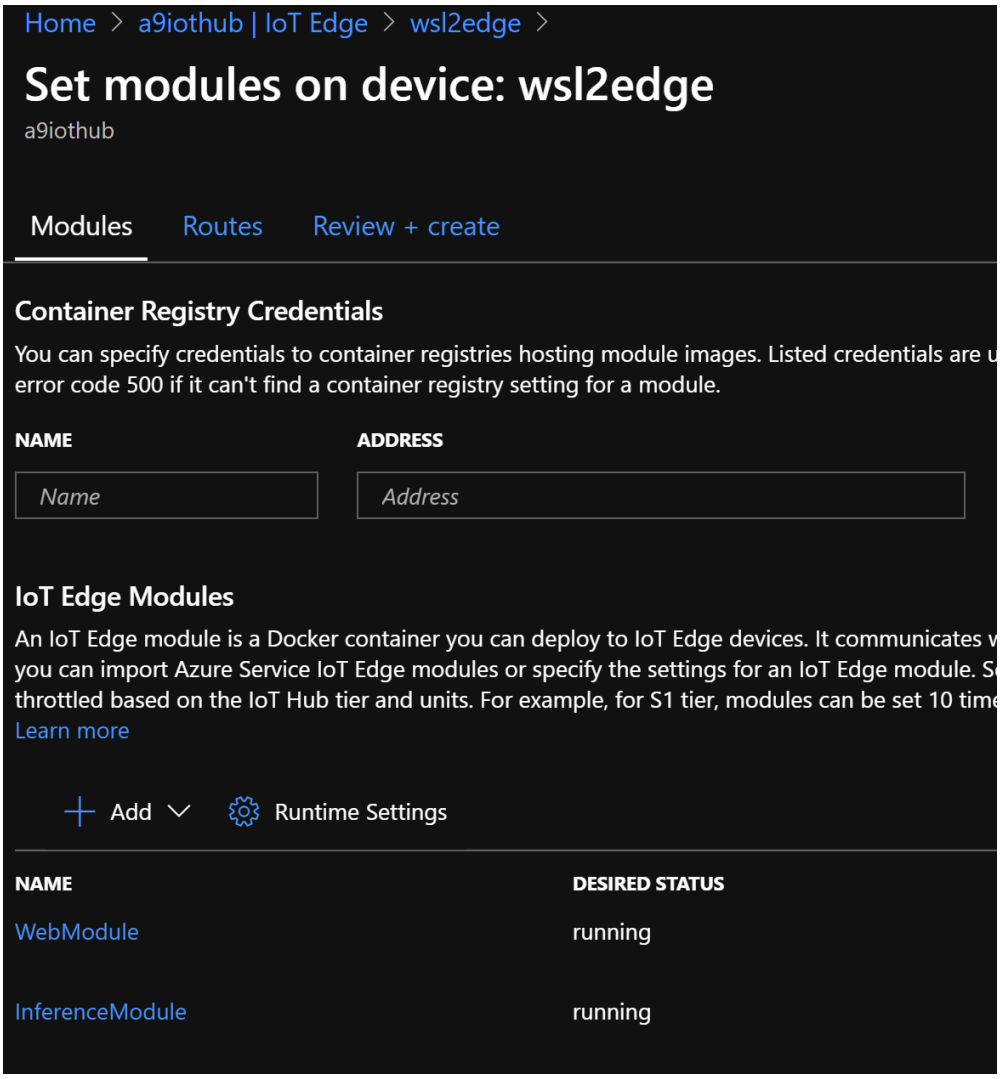
RUNNING FACTORY AI ON WSL2

GPU Acceleration Enabled

Tommy Wu

1. WSL2 Environment Setup :
   A. Surface Book2 with GTX 1060
   B. Windows Insider Preview Build 20150 or higher : register for the Windows Insider Program.
   C. Download Nvidia Driver : https://developer.nvidia.com/45541-gameready-win10-dch-64bit-international
   D. Enable WSL 2 and install a glibc-based distribution (like Ubuntu or Debian). I choose Ubuntu 18.04
   E. Check this guide to install Nvidia container. (Should install this before IoT Edge installation) .

2. Install IoT Edge and Configure your device on IoT Hub by the doc.



3. Steps to run AI workload for Web Module & Inference Module –
   A. Start the previous distro environment

```
PS C:\Users\towu> wsl.exe --list --all --verbose
  NAME              STATE           VERSION
* Ubuntu-18.04      Stopped         2
  ubuntu-iotedge    Stopped         2
PS C:\Users\towu> wsl -d ubuntu-iotedge
root@DESKTOP-5RK65D0:/mnt/c/Users/towu# cd /`
```

B. Clone the start.sh on the root folder.

C. cd ~ & ./start.sh

```
root@DESKTOP-5RK65D0:~# ./start.sh
root@DESKTOP-5RK65D0:~# iotedge list
```

D. Check your module status :

```
root@DESKTOP-5RK65D0:~# iotedge list
NAME             STATUS           DESCRIPTION      CONFIG
InferenceModule  running          Up 5 minutes     intelligentedge/inferencemodule:0.1.10-gpuamd64
WebModule        running          Up 6 minutes     intelligentedge/visionwebmodule:0.1.10-amd64
edgeAgent        running          Up 6 minutes     mcr.microsoft.com/azureiotedge-agent:1.0
edgeHub          running          Up 5 minutes     mcr.microsoft.com/azureiotedge-hub:1.0
root@DESKTOP-5RK65D0:~#
```

E. nano the /etc/docker/daemon.json as

```
{
  "default-runtime": "nvidia",
  "runtimes": {
        "nvidia": {
            "path": "nvidia-container-runtime",
            "runtimeArgs": []
        }
    },
  "dns": ["8.8.8.8", "8.8.4.4"]
}
```

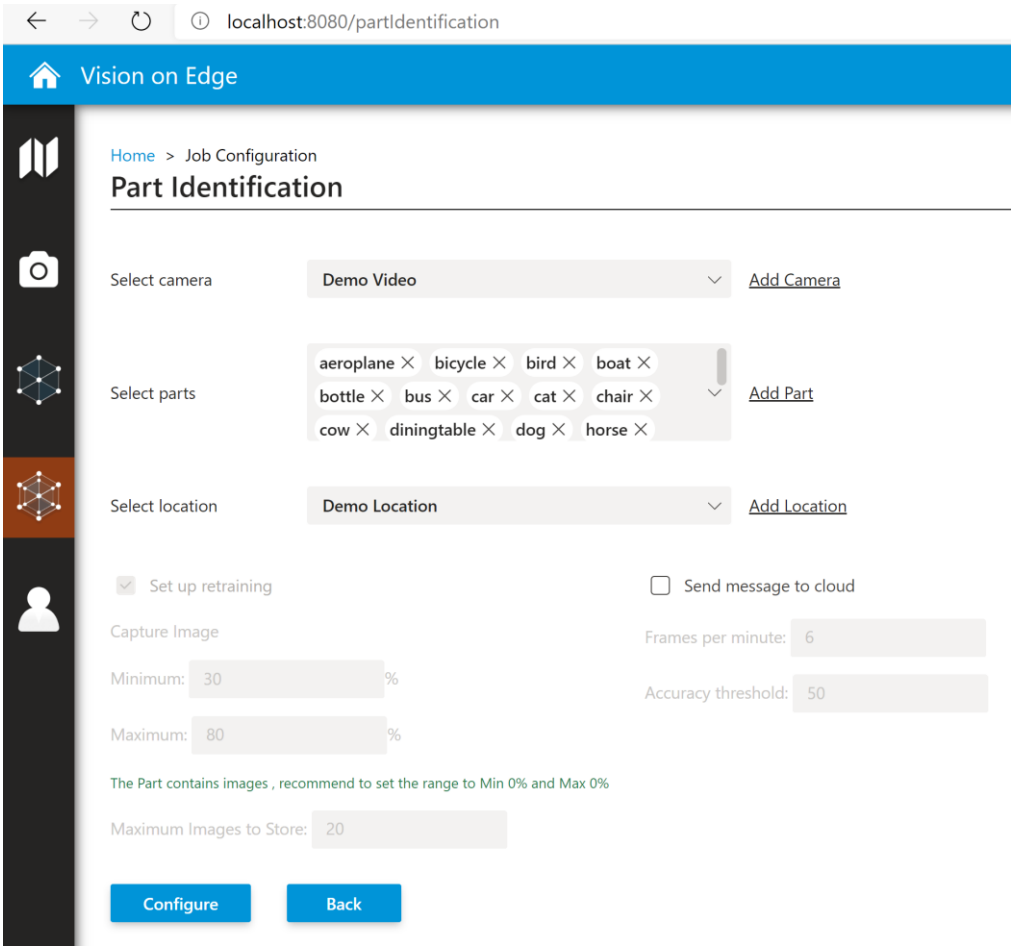
F. systemctl restart docker & iotedge restart

4. Testing Process :

A. Check WebModule is running :

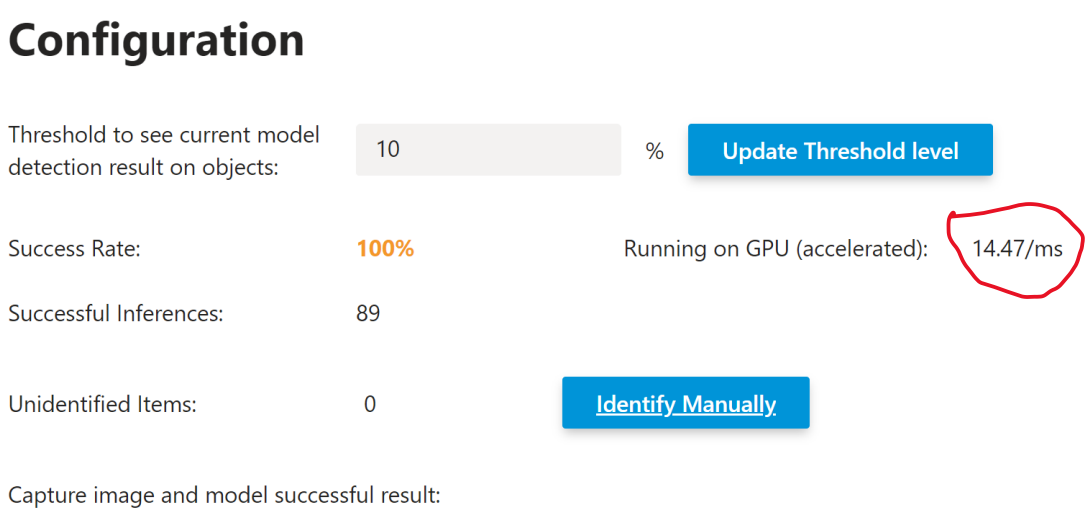
iotedge logs -f WebModule --tail 500

B. Go to http://localhost:8080/ and add a new location.

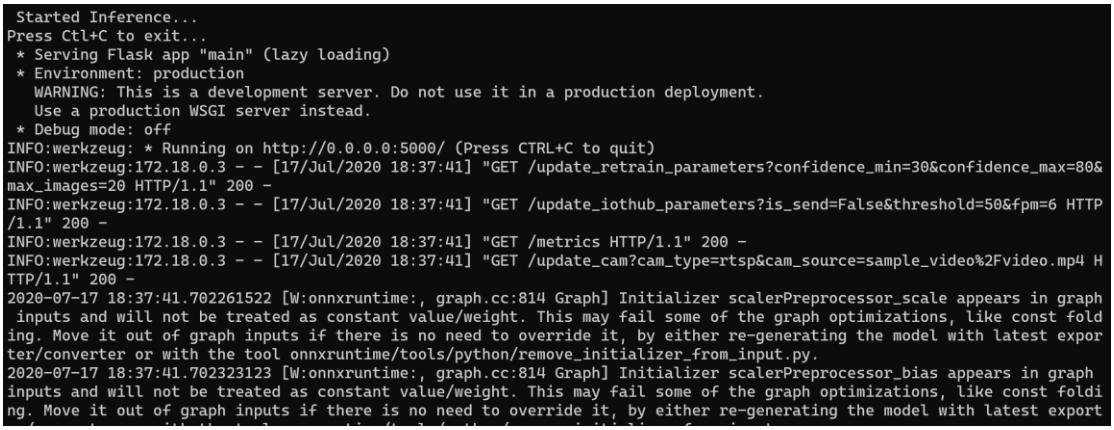


C. Choose “demo Pretrained Detection “ and Configure

D. Check the Right-hand side Configuration about GPU(accelerated) : around 10-15 ms , compared with CPU around 300-500 ms.



E. Verify the logs on inference module : some onnxruntime warning but initializer scalepreprocessor scale.



F. Done and Enjoy your WSL2 !