Real-time Pothole detection using Nvidia TAO and Nvidia DeepStream SDK

Team Tankers:

-Saif Eddine Layouni

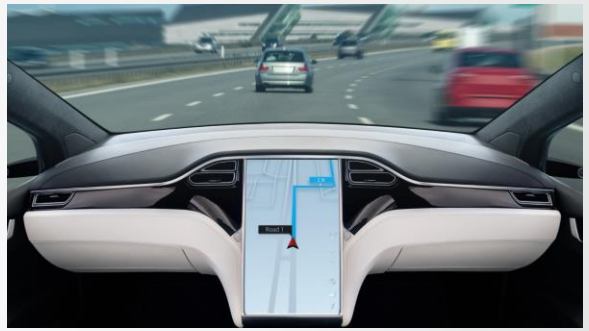
-Mohamed Ali Mimouni

-Melek Elloumi

InstaDeep™ NVIDIA

Problem

- The world is marching toward making all cars self-driven



- The traffic detection models are trained on videos of streets from developed countries.

=> They are biased toward roads in good conditions !! Streets in developing countries of Africa are less maintained and contains more potholes.

Solution

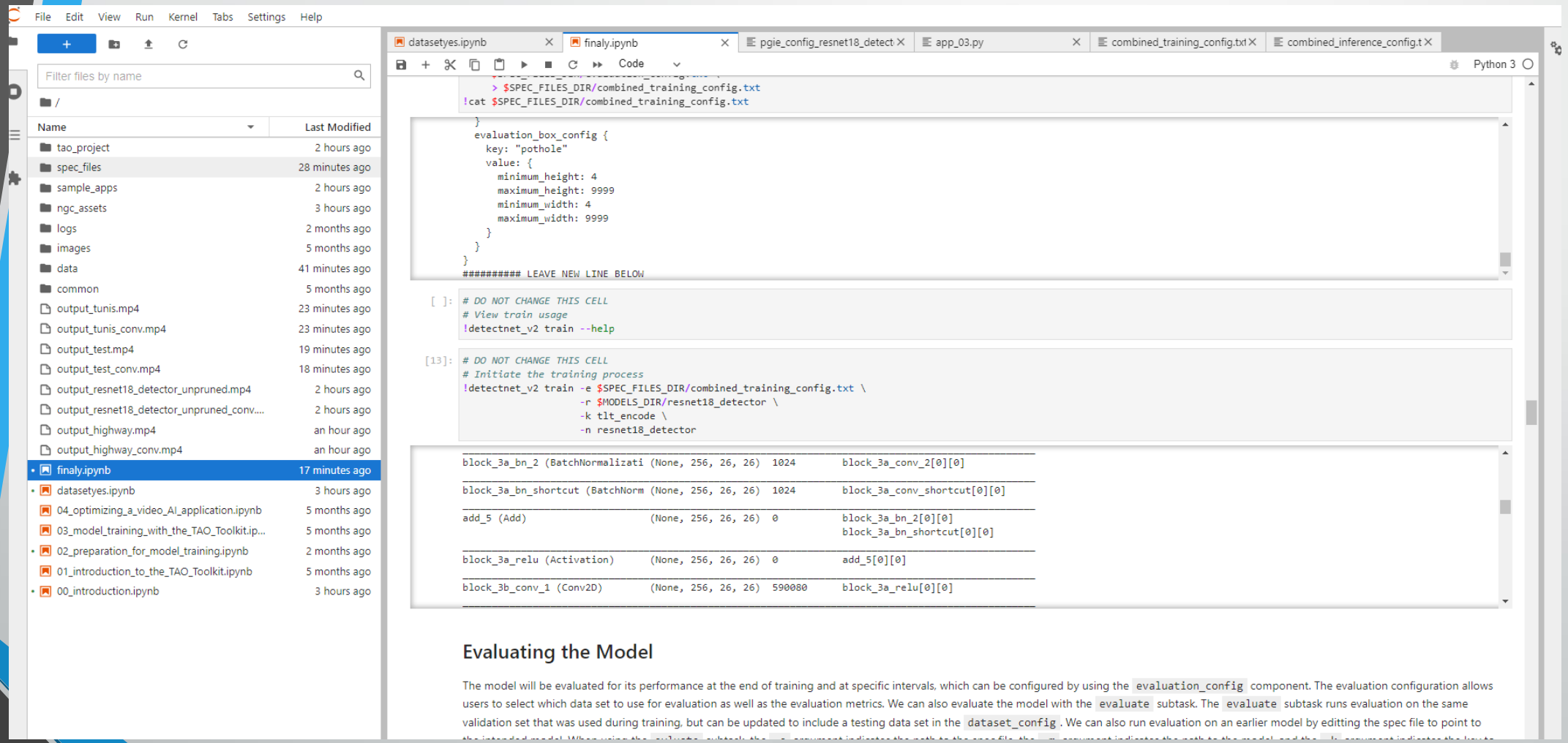
=> Use transfer learning to improve already-built traffic detection models and add pothole detection to them.

The model need to be in real time to stop the car or steer it in case of a pothole in front of it

Technologies

To achieve high performance and faster training Time, we used Nvidia TAO for transfert learning And Nvidia DeepStream for Streaming pipeline.

Notebook Showcase

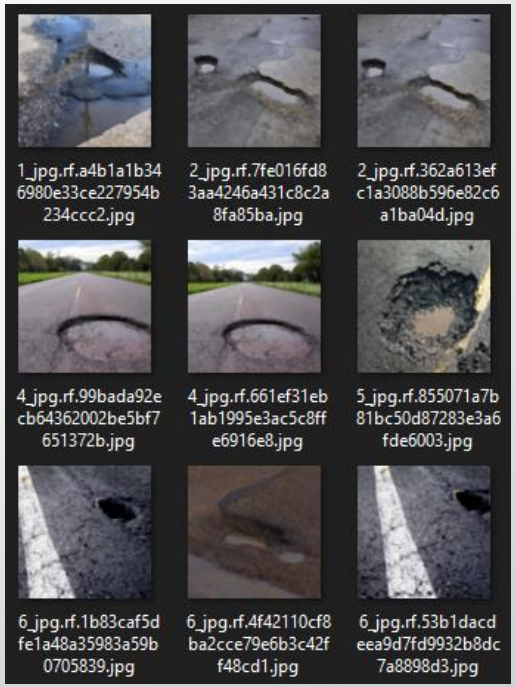


TAO

We used Nvidia's TraficCamNet as the base model. It detects cars, pedestrian traffic signs and motorcycles from images with bounding boxes.

We supplied a pothole dataset from Roboflow and applied data augmentation.

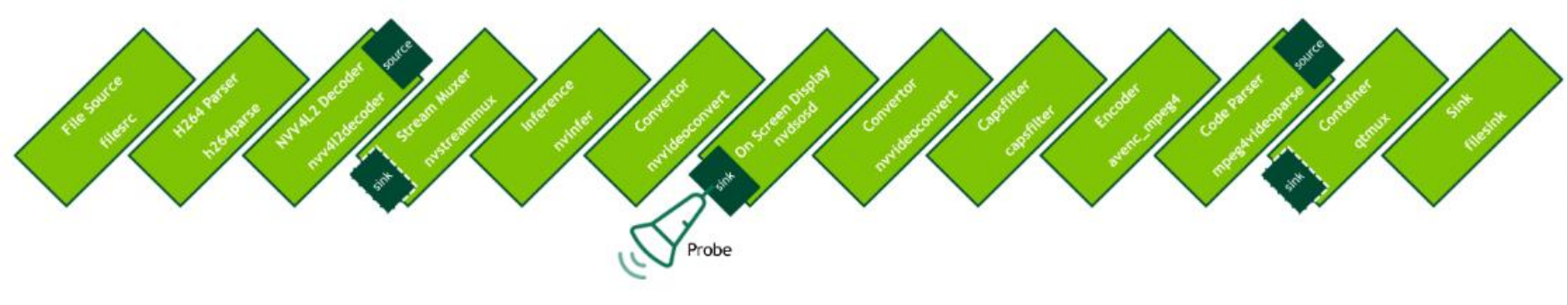
We used the config of the Nvidia DLi course but we accomodated it to our data and added NMS (Non-max suppression)



DeepStream

Self-driven cars use Edge computing for real time computing which is accommodate to using DeepStream with Nvidia Edge Gpus.

So, we used The DLI DeepStream architecture since it suits our needs to decode video source, apply the trained model, encode it for screen display and save it in storage for future use.



Results

We obtained 57% precision accuracy.

The streaming had a stable fps (30 fps)

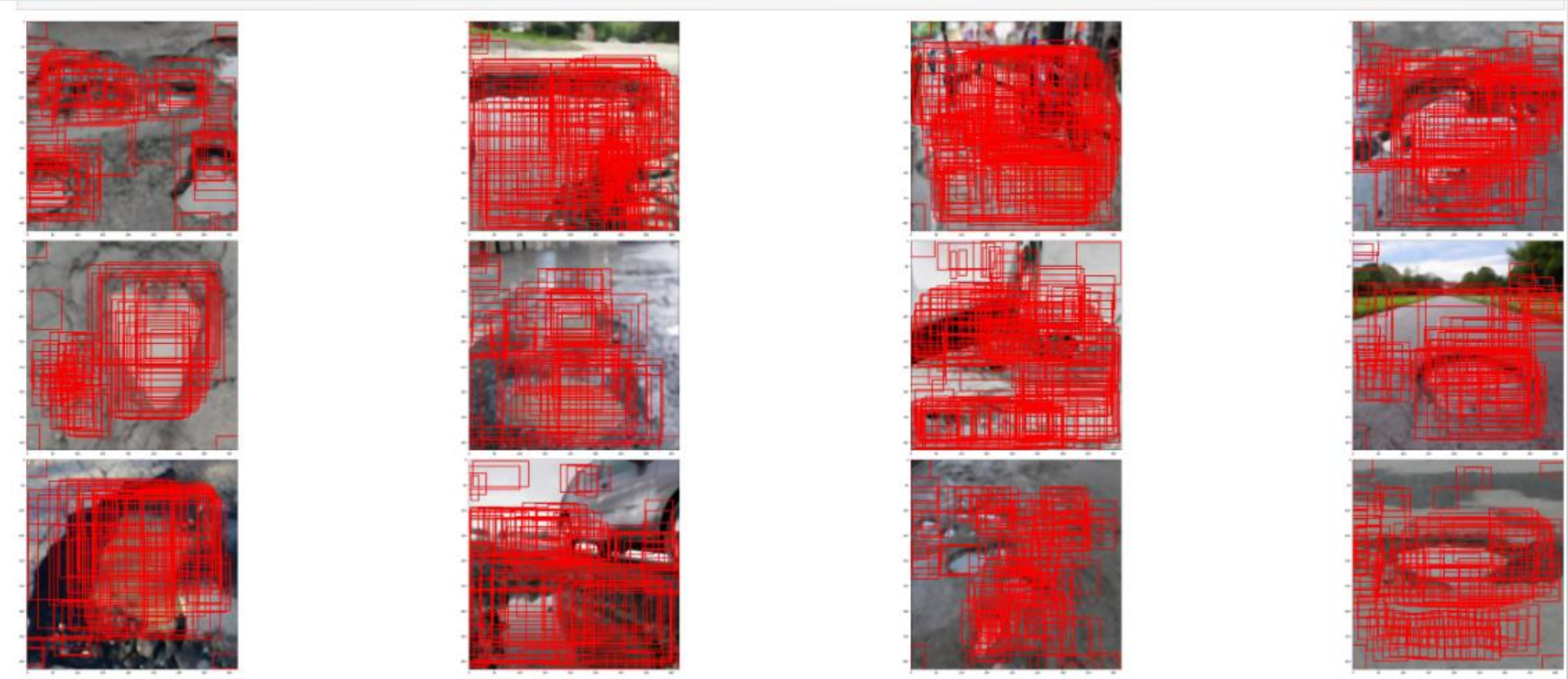
We didn't test latency and GPU usage but they are important metrics to rate our model and streaming.

```
Validation cost: 0.003703
Mean average_precision (in %): 57.2466

class name      average precision (in %)
------------  --------------------------
pothole                         57.2466

Median Inference Time: 0.005544
2022-08-30 05:58:49,512 [INFO] __main__: Evaluation complete.
Time taken to run __main__:main: 0:00:13.665325.
```

```
Creating Pipeline
Adding elements to Pipeline
Linking elements in the Pipeline
Starting pipeline
FPS: 0.14 @ Frame 0.
FPS: 29.94 @ Frame 100.
FPS: 29.68 @ Frame 200.
FPS: 30.28 @ Frame 300.
FPS: 29.9 @ Frame 400.
FPS: 29.68 @ Frame 500.
FPS: 29.9 @ Frame 600.
FPS: 29.72 @ Frame 700.
FPS: 29.96 @ Frame 800.
FPS: 30.25 @ Frame 900.
FPS: 30.0 @ Frame 1000.
FPS: 30.29 @ Frame 1100.
FPS: 29.96 @ Frame 1200.
FPS: 29.64 @ Frame 1300.
FPS: 29.95 @ Frame 1400.
FPS: 29.95 @ Frame 1500.
FPS: 30.28 @ Frame 1600.
FPS: 29.69 @ Frame 1700.
FPS: 29.97 @ Frame 1800.
FPS: 29.68 @ Frame 1900.
FPS: 29.97 @ Frame 2000.
FPS: 29.99 @ Frame 2100.
FPS: 29.68 @ Frame 2200.
```

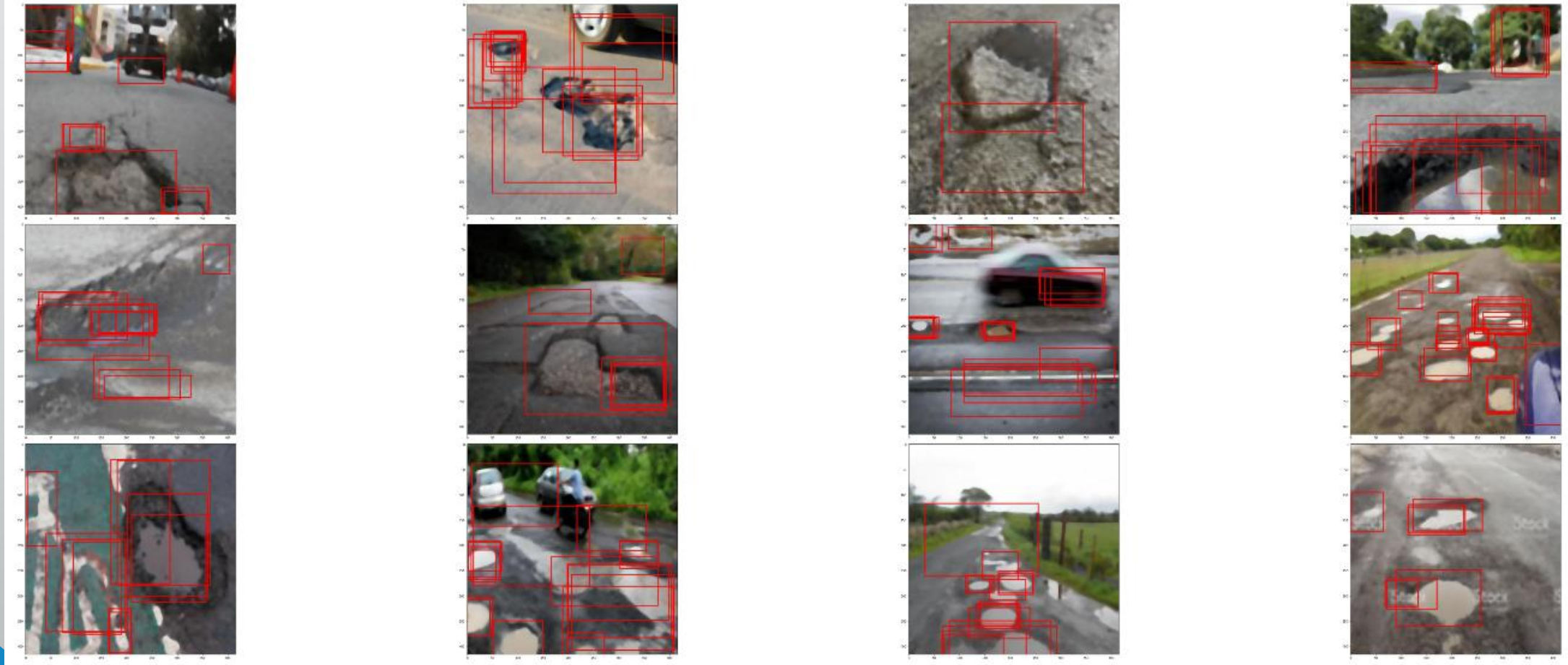
Results

Inference on dataset images when we trained the first time (10% accuracy)



Results

Latest inference results (57% accuracy)



Results

Applied the model on video footage.

Work to do

To improve results, we need to invest more time in the pothole dataset, collect more diversified data (specifically from African countries' streets)

We need to fine tune the hyperparameters and reduce overfitting of the model.

We already applied NMS but it still needs more work to have nearly perfect result of pothole detection.

What we learned

Today we discovered Nvidia real time technologies such as deep stream SDK for Video Analytics and TAO and managed to work with them in a short time.

We feel they are important tools that companies should use them to accelerate their Machine vision solutions.

Nvidia NGC is rich with amazing models that we can't wait to explore them more throughly in the future.

Thank you for your attention