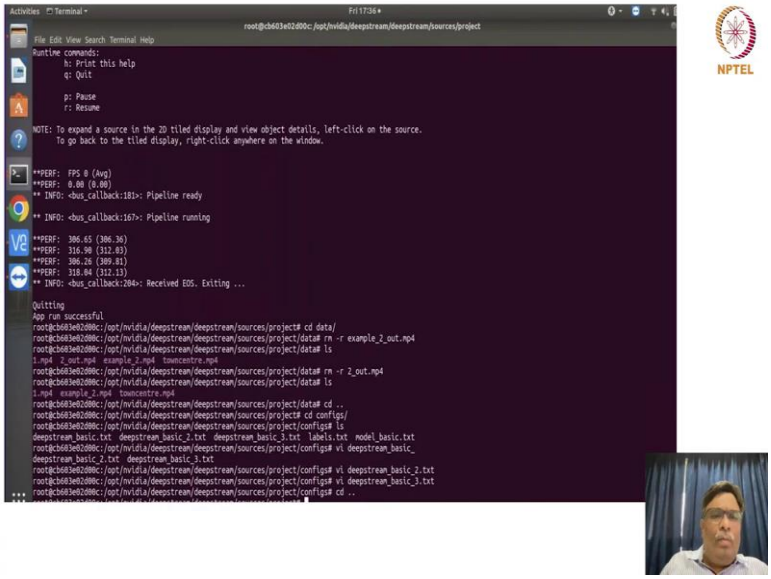
Applied Accelerated Artificial Intelligence
Prof. Satyadhyan Chickerur
Department of Computer Science and Engineering
Indian Institute of Technology, Palakkad

Lecture - 57
Applied AI: Smart City (Intelligent Video Analytics)
Session 2: DeepStream - Part 2

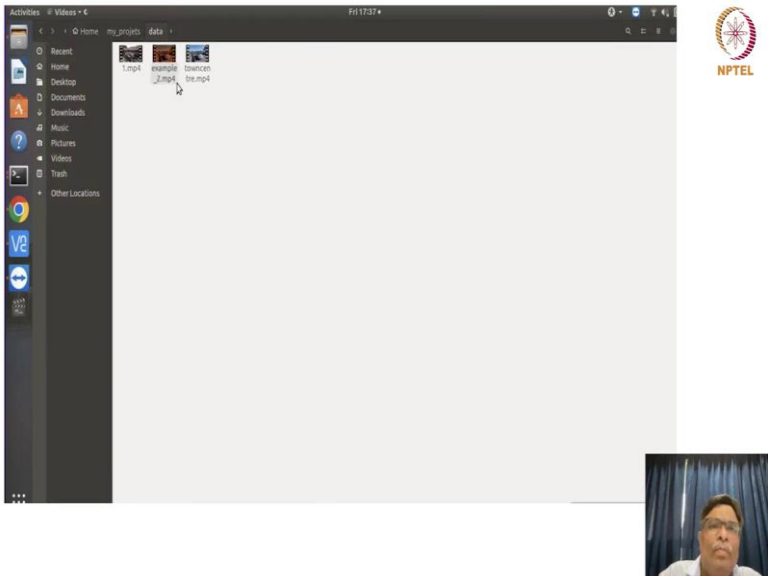
(Refer Slide Time: 00:17)



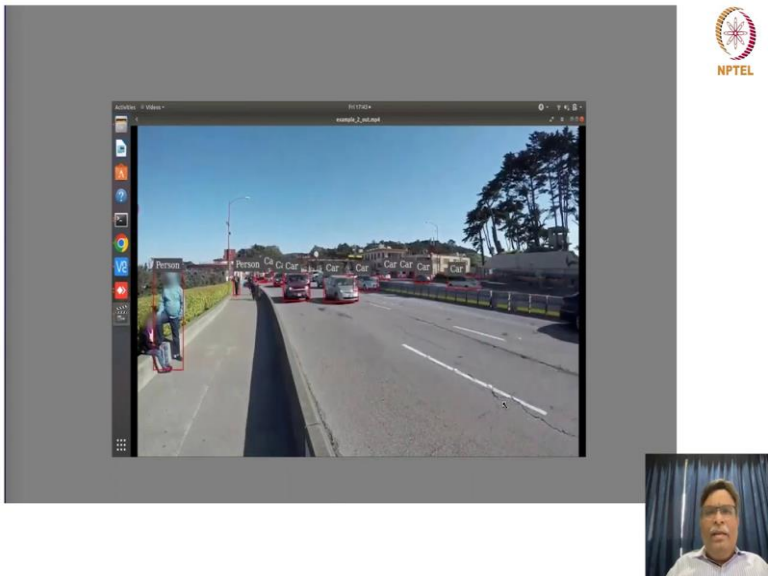
So, now see what we have tried to do here is this project right which you are now seeing is a folder right, now this folder is tagged to the see this is a docker container which we are running ok. Please, understand this. So, this is a docker container which is running and this particular docker containers project folder right is linked to the my project folder of the deep stream folder.

So, you are trying to link both of these first. So, dockers project folder and my project folder ok.

(Refer Slide Time: 01:17)

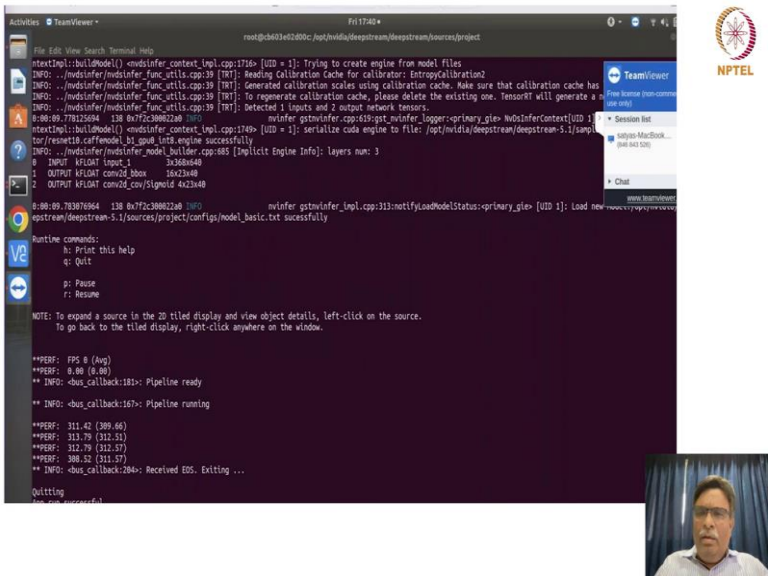


(Refer Slide Time: 01:19)



Now, let us now try to see the video which is before the processing ok. Let us try to see the video, yes. So, this is one video which we are going to use for processing ok.

(Refer Slide Time: 01:41)

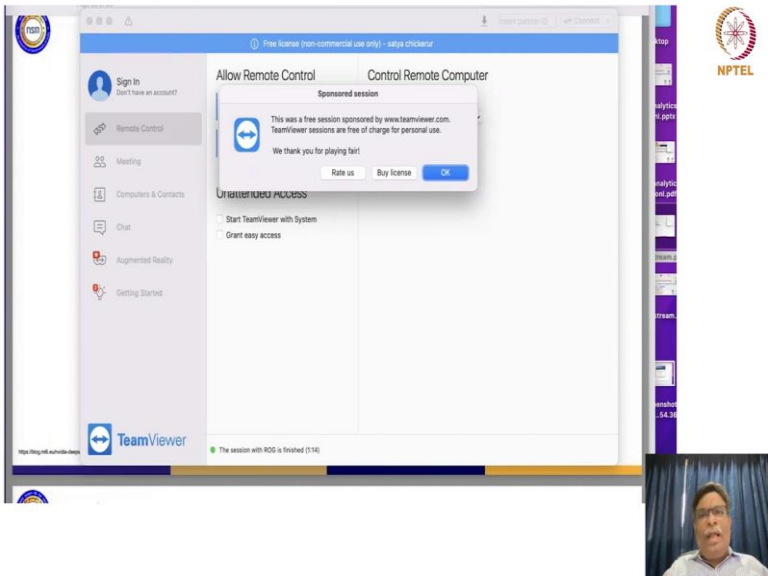


So, now let us start the deep stream app. So, now, let us try to understand this while running yeah. So, this is when you run ok, the deep stream app to actually try to track ok and try to classify and do things on that particular video which just now showed you ok.

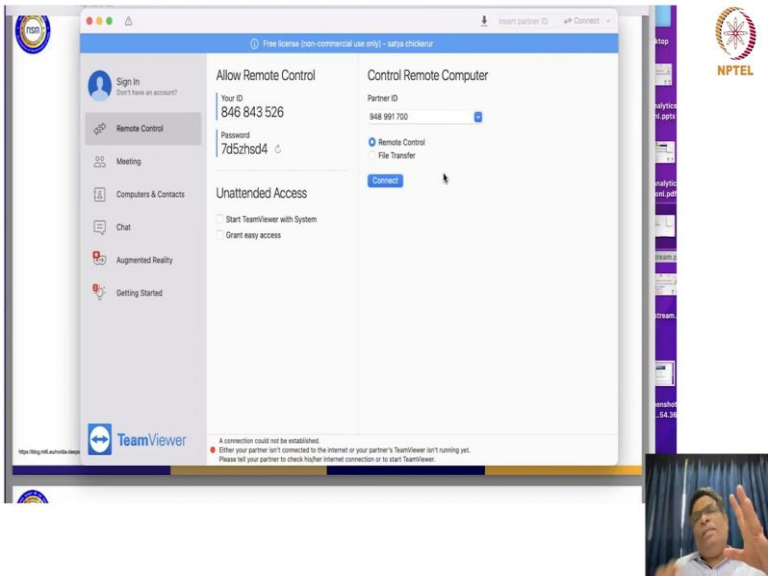
Now, if you see here this particular thing this app is actually using all of this right and nvinfer, gst-nvinfer all of this right and it has done some calibrations because we are using tensorRT here right. So, all of this and now this is getting processed so, we have paused it for some time. Now, see here the runtime commands if you see you can pause this and we will resume it now ok

So, if it is resumed right. So, app has been run successfully and now since we are using dockers ok if you also use these dockers what is going to happen is the immediate stream which we showed you right.

(Refer Slide Time: 02:45)



(Refer Slide Time: 02:51)



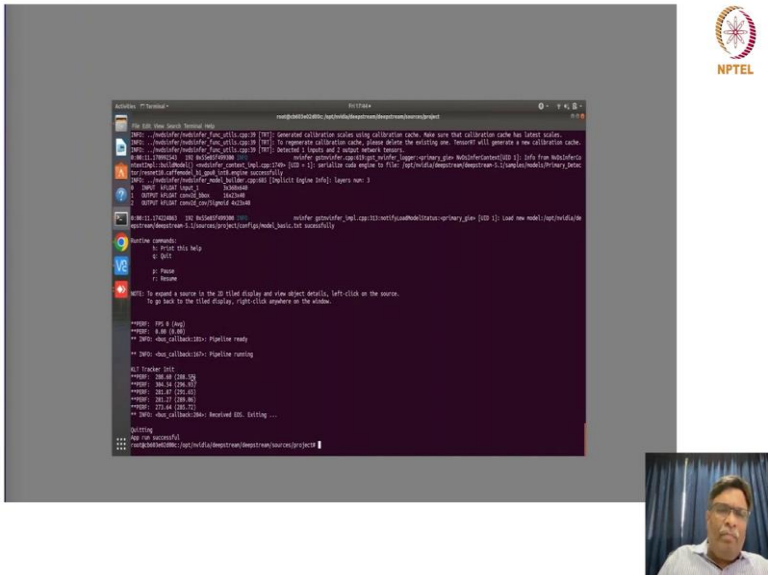
So, what effectively happens is when you pause this or when you process it since I was telling that you are trying to work on dockers the whole live image stream what you are trying to process ok is not visible because of this x server issues and all, but it will get stored ok somewhere in your this thing right in your folder.

So, we will show you that as well now, since it has been processed. So, if you see this um. So, it is processed and the images which are here. Yes. So, this is how it is being done. So, the initial video which we showed you ok it is processed using this deep stream

application for tracking and for classifying ok this and this we are done it on a discrete GPU right.

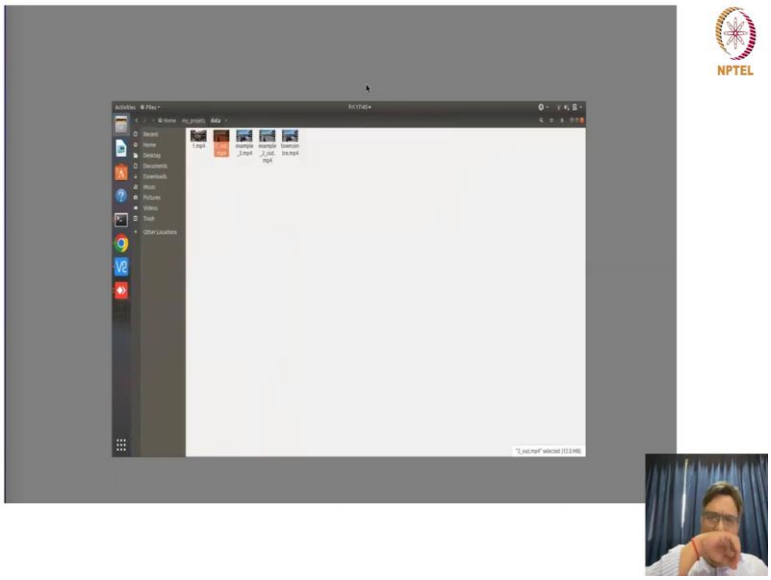
Now, this is only detection right now we will try to do only detection. So, this is the detection and now we will try to do the next example and we will show you the configuration files also afterwards, but for the time being just let us execute it and show you this. So, if you see here gst-nvtracker batch processing is off ok.

(Refer Slide Time: 04:28)



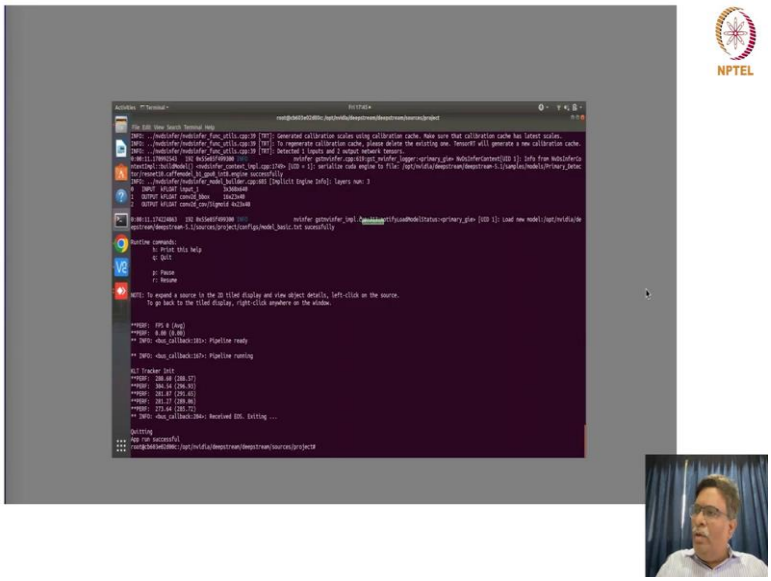
So, again this is doing this processing and if you see here this performance of the tracker tells us about the frames per second being processed and this is the average ok. So, app, yes, so, now no not this one, this. So, this is basically helping us to do counting also, initially it was just detection of cars, detection of people, tracking them, but now we are able to count also right.

(Refer Slide Time: 05:43)

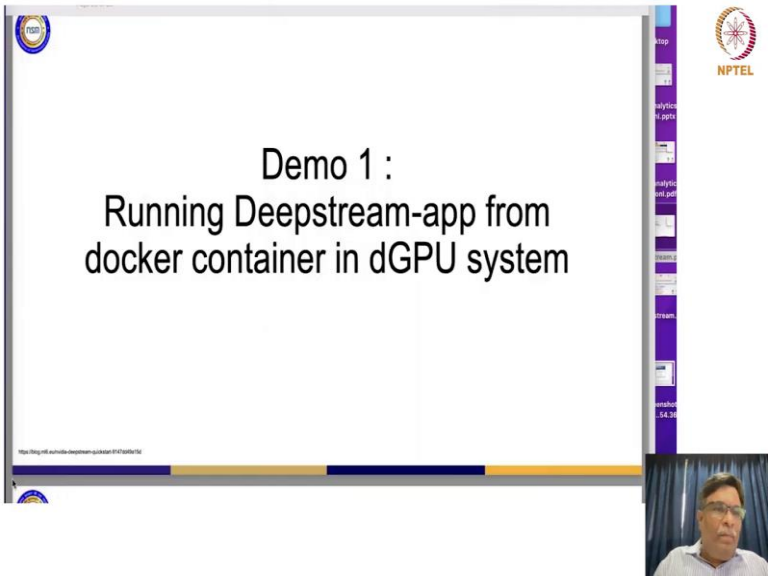


So, this is the next example which actually shows us how many cars have been tracked, how many people have been tracked ok and their count and whatever right. So, this is how it is going to actually be used ok. So, now, let us try to work when it is basically put on the inferencing thing right.

(Refer Slide Time: 05:51)

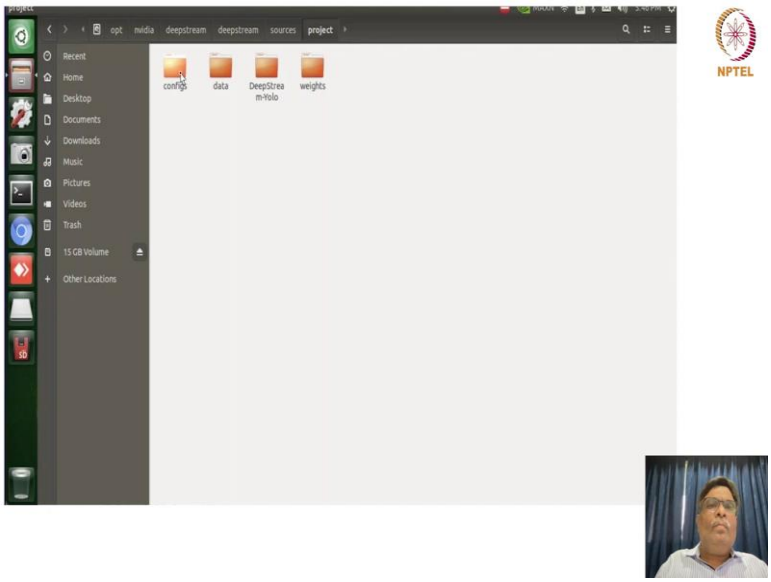


(Refer Slide Time: 06:07)

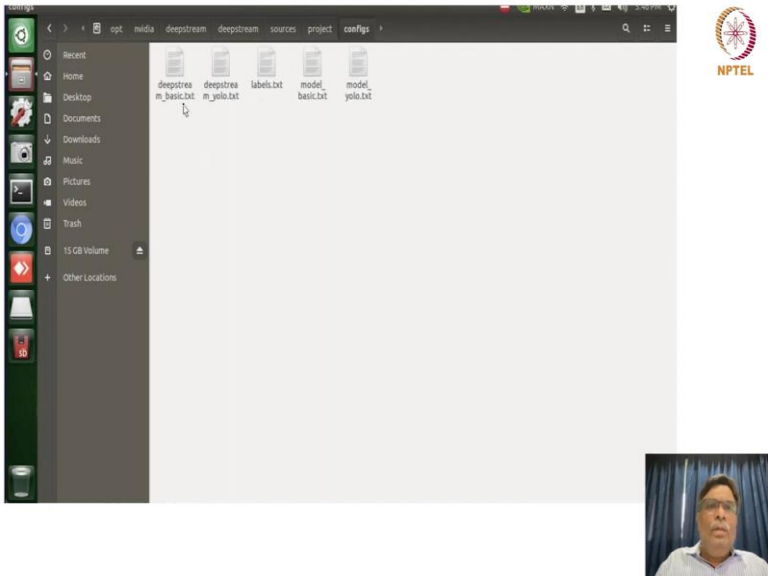


So, now, let us try to do let on Jetson TX 2. So, I will try to do it on Jetson TX 1 not 2 actually yes. So, let us try to do it on this ok. So, before we start let me just show you the config files you know yeah.

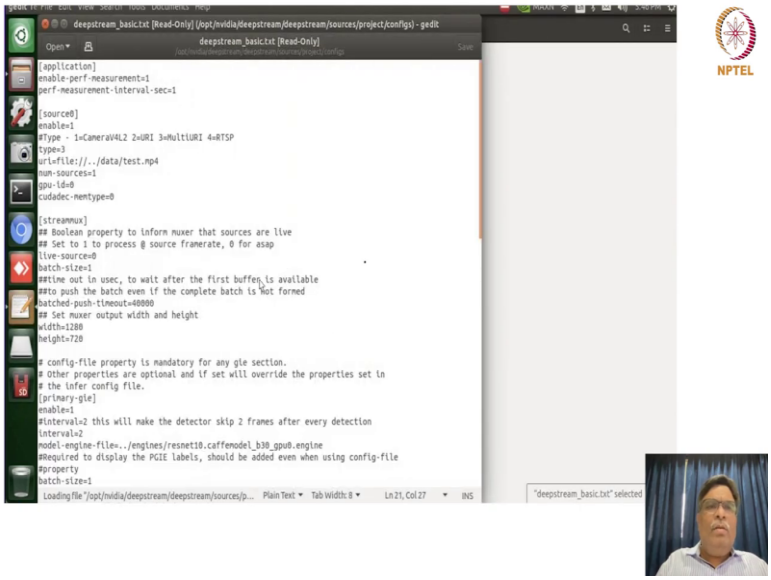
(Refer Slide Time: 06:17)



(Refer Slide Time: 06:24)

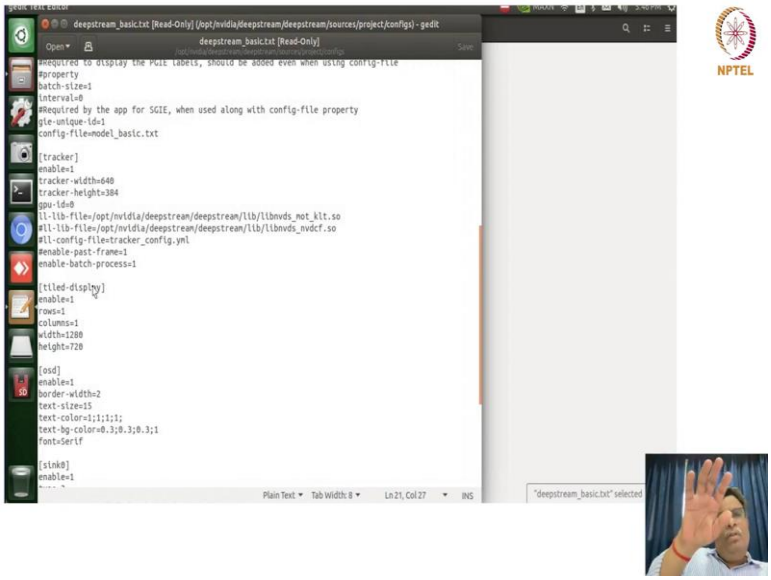


(Refer Slide Time: 06:27)



So, if you see this there is something which is this config file which actually helps us to configure things right. So, see, what are you going to do the source image right, number of sources, then if you talk of streammux right, what would be my batch size, what would be the output of the muxers width and height ok.

(Refer Slide Time: 06:51)

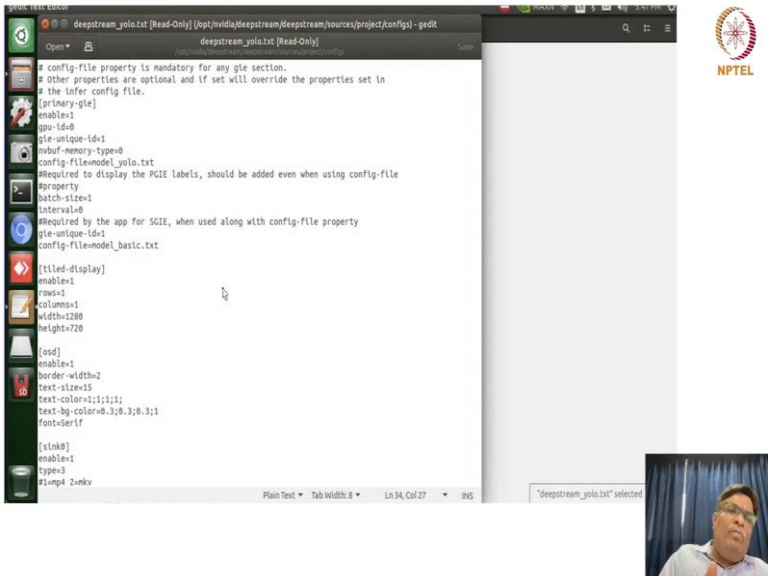


And then what is that model engine, how is that tracker being used, what is the size of the tracker, the width as well as the height ok and whether you are going to use a tiled display or not ok. So, how is it going to be enabled or disabled, what is the size of the OSD right on screen display and then.

So, what we are going to do is, again we are going to see the thing of how we are going to run the deep stream application right. So, this directory structure and all is very simple as to you will have to use the directory structure you will have to create a project folder from various sources and then you basically have to configure all of this right.

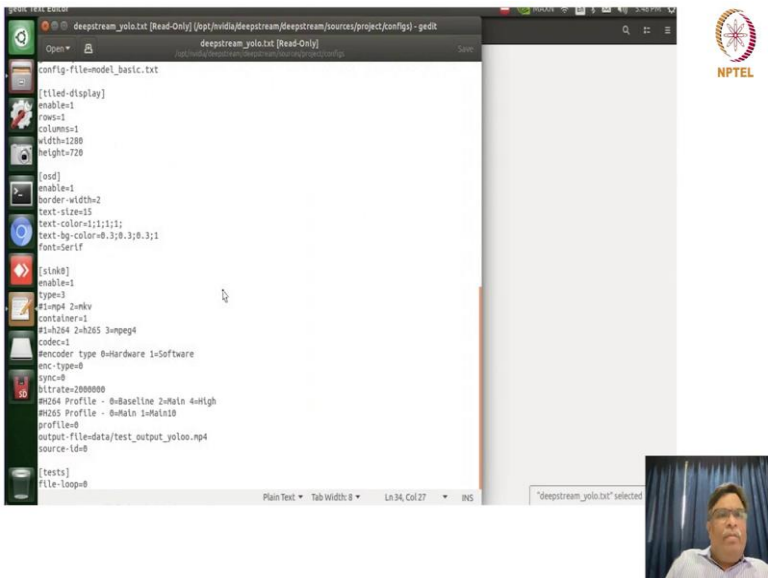
And then you can see the data folder tested and then as we showed in the previous example the same thing is going to happen. So, let us try to run it ok and see what basically happens ok. So, this is the same thing, let me just show you for this model also ok.

(Refer Slide Time: 08:06)



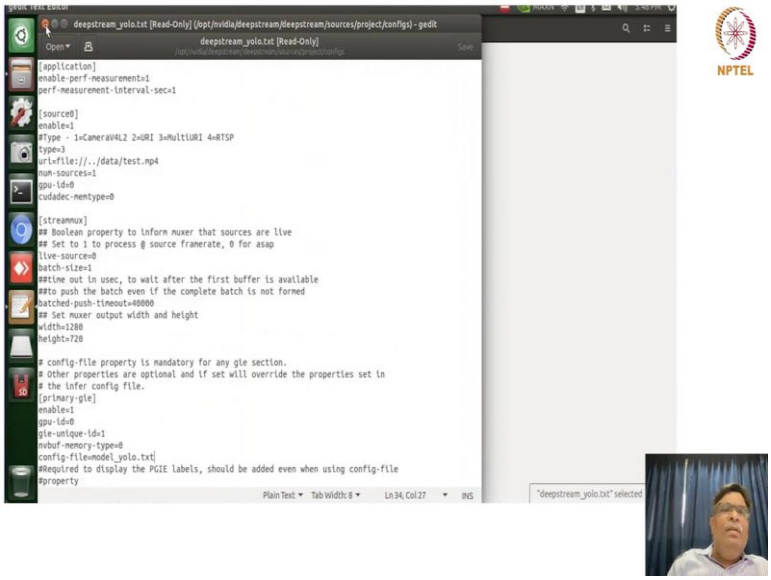
So, this is a YOLO model which we are using for the next demo, there are two demos which we will be showing rather three demos which will be showing now.

(Refer Slide Time: 08:15)



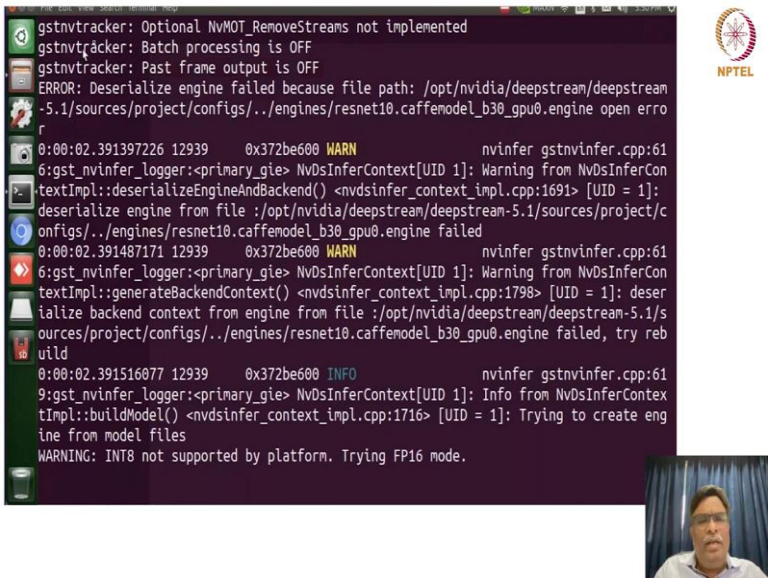
So, this again has basically something which is sink 0, sink 0 in the sense the whole output when it is going to be shown right. So, how is that what are the various things which you can use with ok. So, the H 264, H 265, decoder, encoder and all of that right values, right.

(Refer Slide Time: 08:45)



So, what type MPEG- 4 or whatever those containers, how are you going to use it and all of these parameters you can actually just see ok, not going into the details we will show you the demo of the executions yeah.

(Refer Slide Time: 09:00)

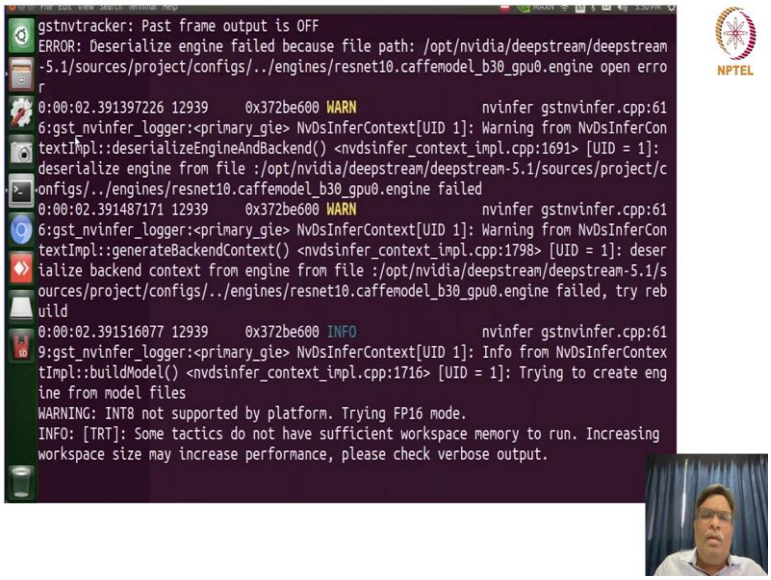


So, what we are trying to do is now let me just increase the size of the, 1 minute, let me just increase the size of this ok.

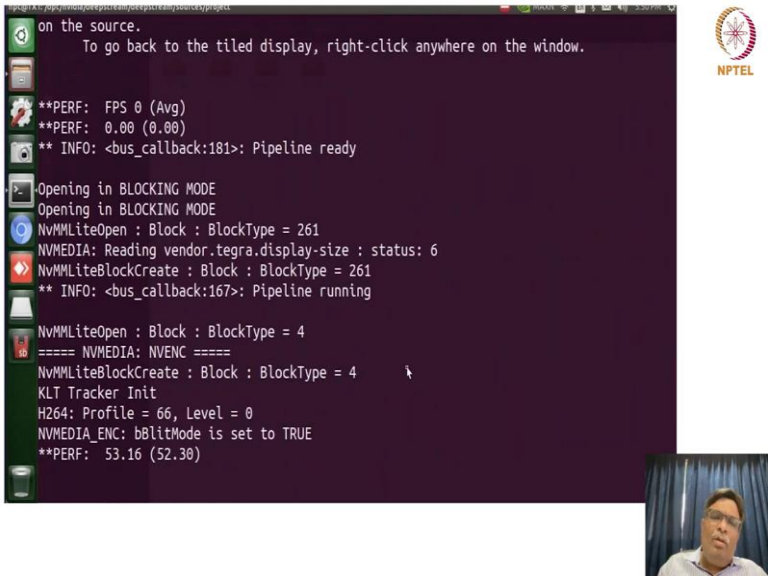
So, now what we will show you is the demo of a object detection ok with basically detect it. So, let me just try to show you that, yes. So, these are all configuration files again. So,

again we are trying to work on FP16 mode INT8 not supported by the platform is a different issue, but yeah this is just to show you deep stream ok.

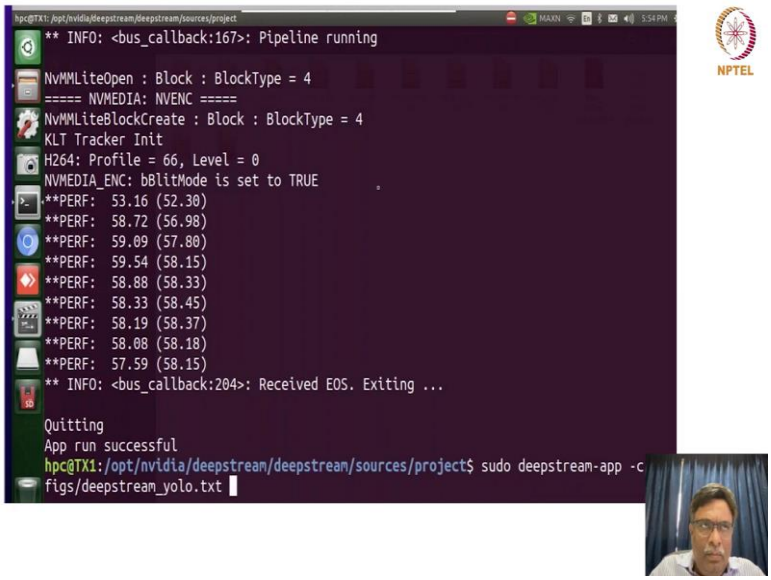
(Refer Slide Time: 10:28)



(Refer Slide Time: 10:41)

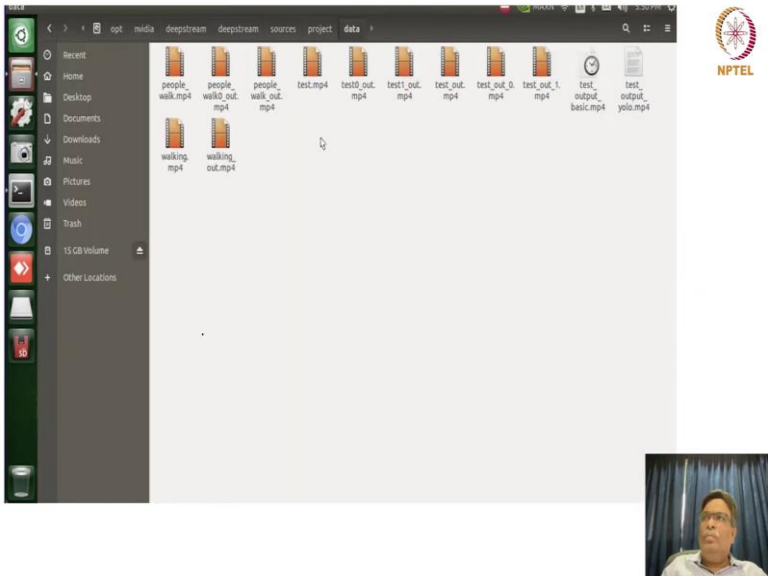


(Refer Slide Time: 10:44)

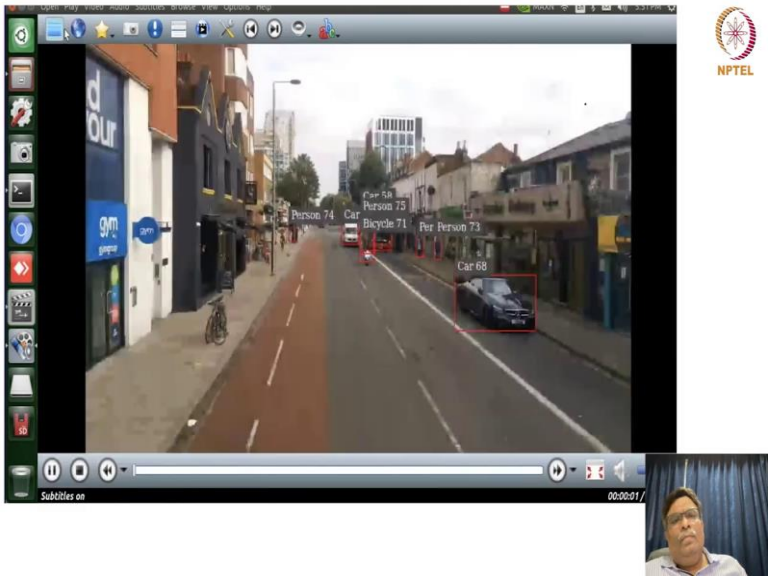


So, there are certain things related to the tensorRT stuff wherein it is going to give you right the info about the memory requirement does not happen ok. So, what is the pipeline running, again if you see the performance here since it was running on a discrete GPU now this embedded GPU. So, this is the processing speed and this is the average processing speed with frames per second, yes.

(Refer Slide Time: 11:11)



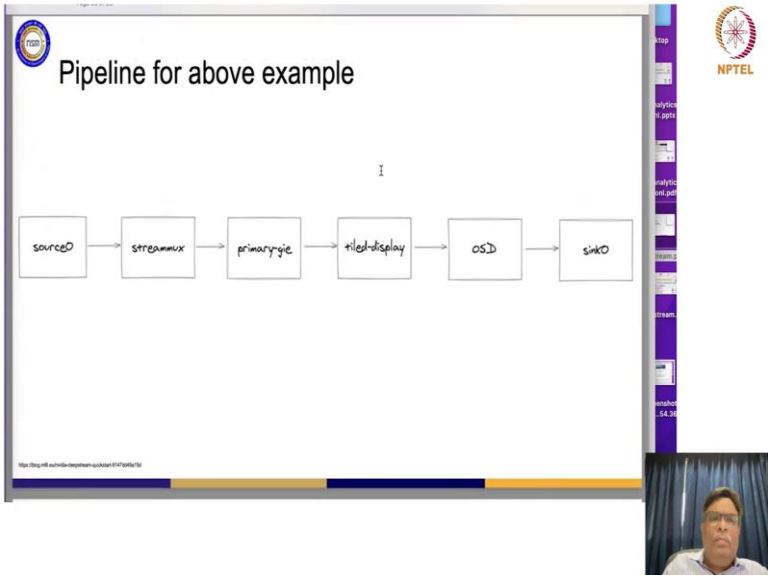
(Refer Slide Time: 11:32)



So, let us now try to see the data or the video. So, this is basically trying to see the output we have tested it so many times we will just try to see. Yes, so, this is captured like this because we are trying to run it through different types of viewers as well as you know this video ok.

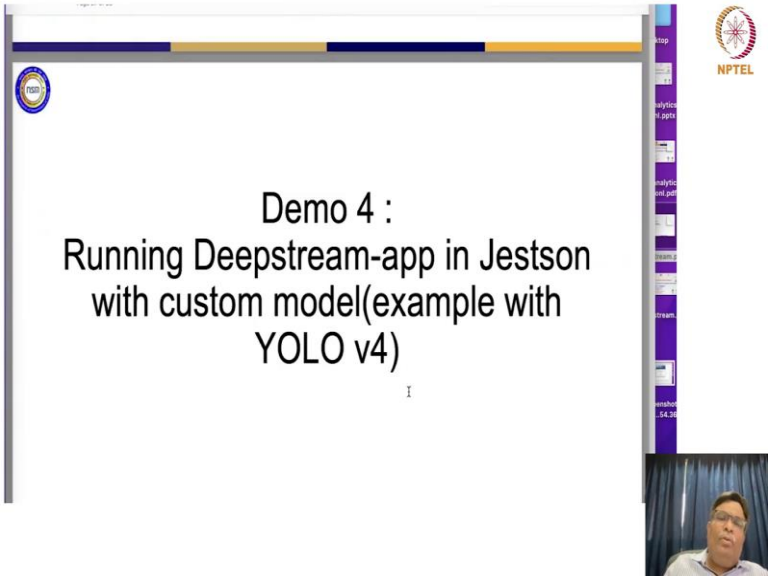
So, let us try to now demo of object detection wherein we can put in our own model also we will try to show you this the third one yeah.

(Refer Slide Time: 12:27)



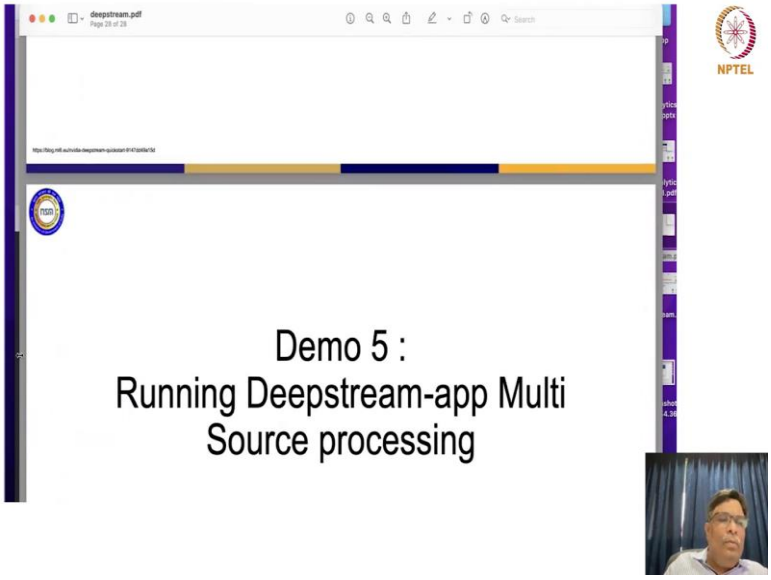
So, this is what is that generally everything of will be of this type the any pipeline will be of this type right, you have source, you have streammux, you have this primary gie, then tiled display, OSD and sinko right.

(Refer Slide Time: 12:47)

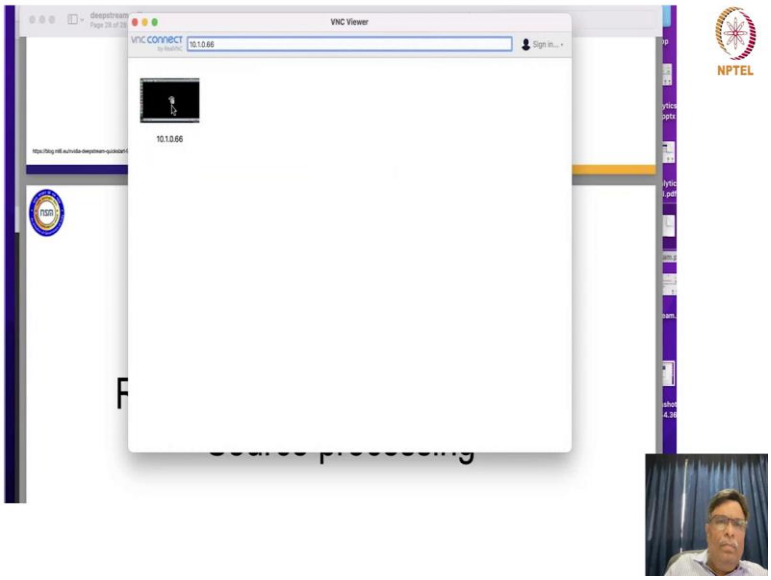


So, yeah. So, we will try to show you this 4th demo wherein we are trying to use our own custom model which is YOLO v4 ok and then let us try to see how best we can do that ok, yes.

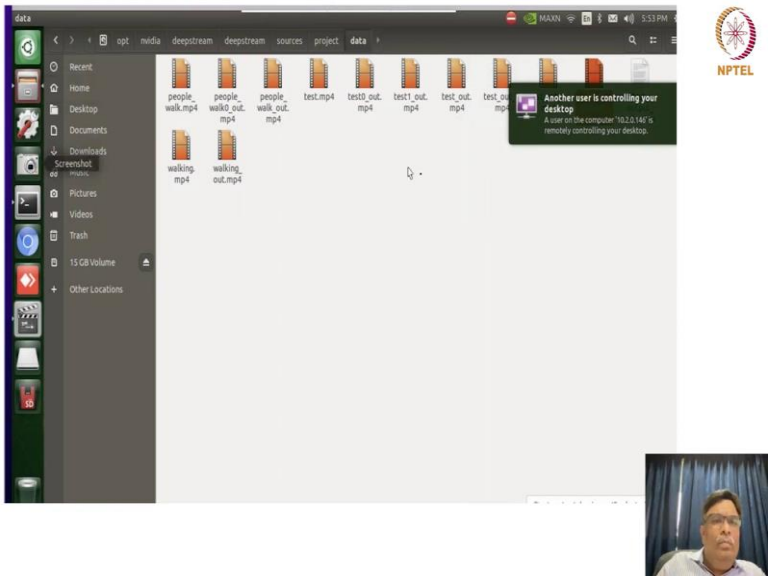
(Refer Slide Time: 13:01)



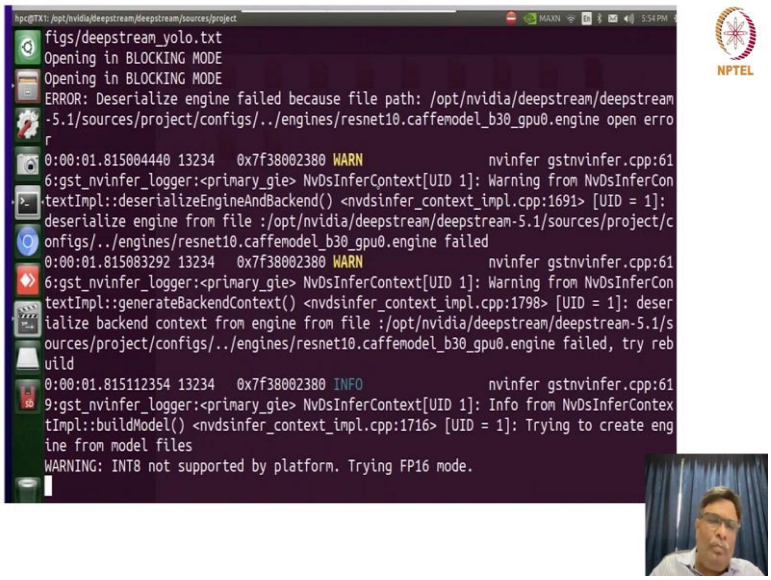
(Refer Slide Time: 13:15)



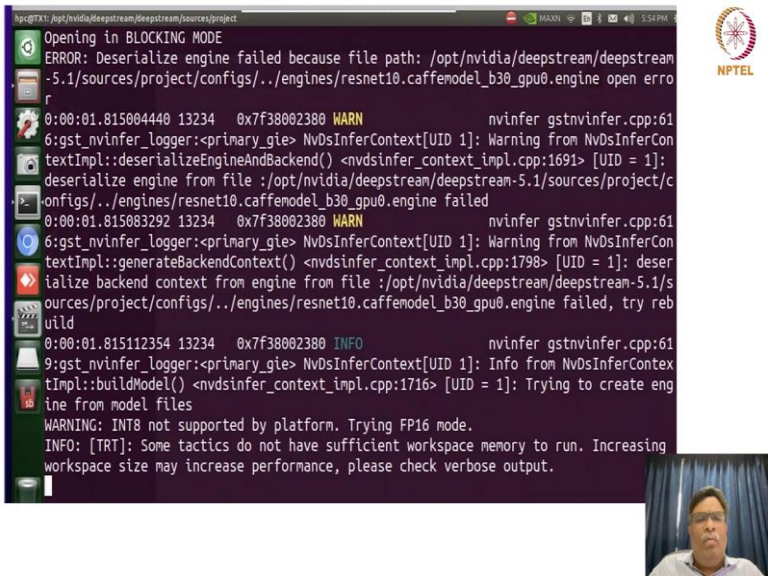
(Refer Slide Time: 13:30)



(Refer Slide Time: 13:58)



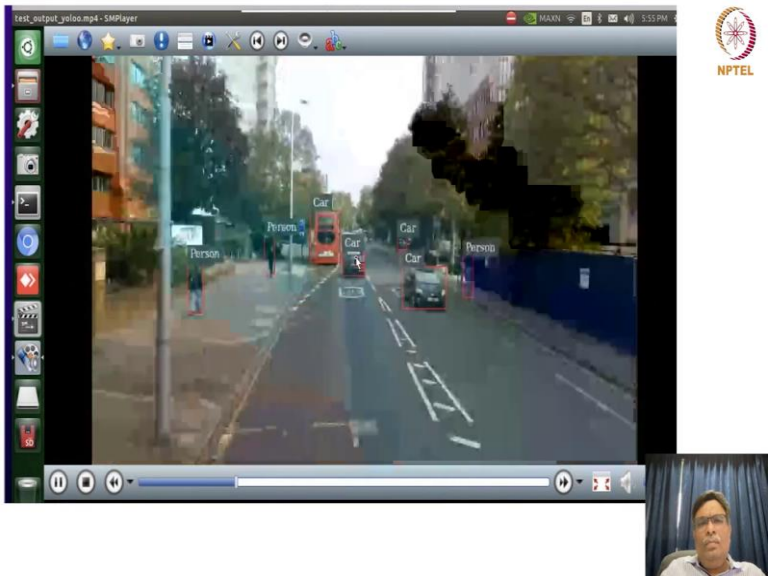
(Refer Slide Time: 14:11)



So, yes so, let me just go there yes. So, again the same thing we have not yet fully optimized it through intake, but yeah we are trying to work within a FP 16 mode.

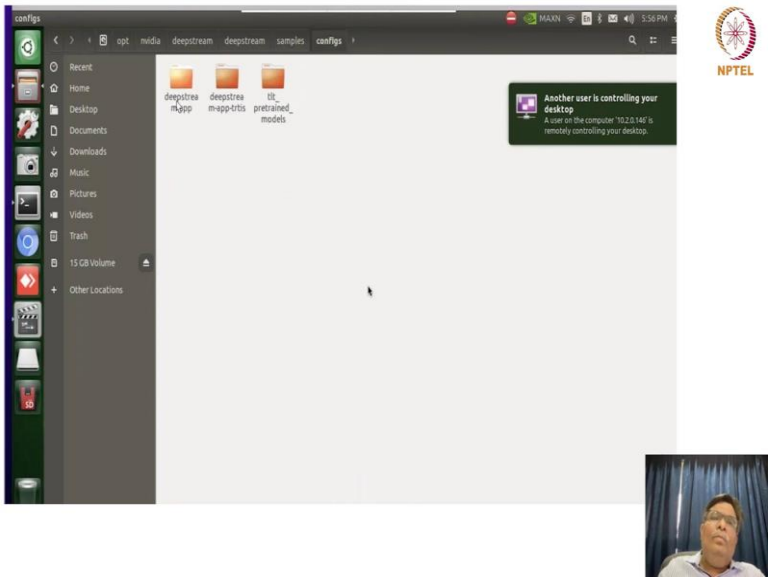
So, we are trying to put in our own model now ok. So, again if there are certain areas wherein the sufficient workspace memory is not available right and it will give you a warning the tensor RT thing that you can increase the workspace size this that everything and again it is going to do this ok. So, now, that application is run I do not know why this gets stuck so much.

(Refer Slide Time: 15:01)

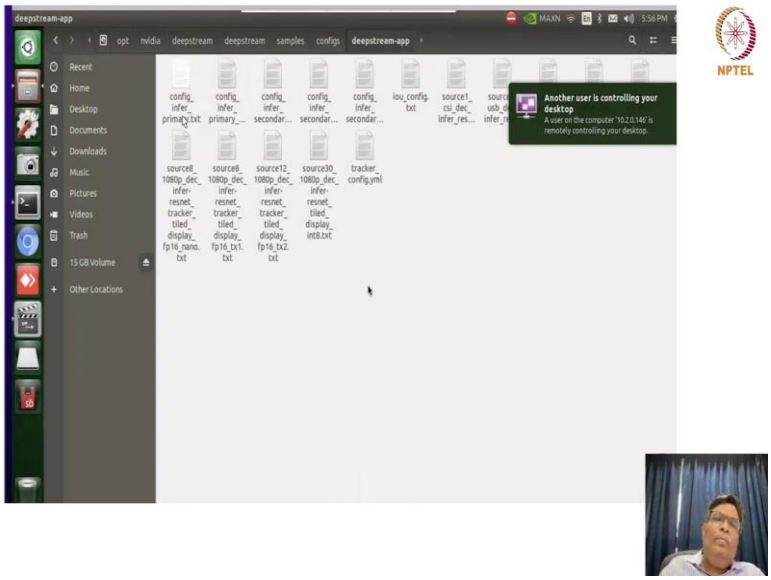


So, now this is with our own model which you have modified instead of the detect.net, we had put in the YOLO.net on this and then do it. So, now, we will show you multi source input processing wherein you will get lot of multiple inputs from various sources yeah. So, 1 minute I do not know there is something you know yeah yes.

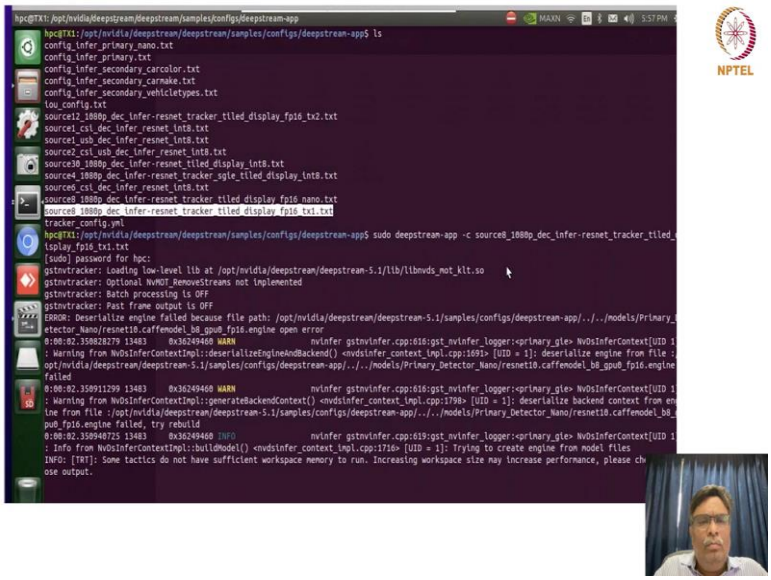
(Refer Slide Time: 15:57)



(Refer Slide Time: 15:59)

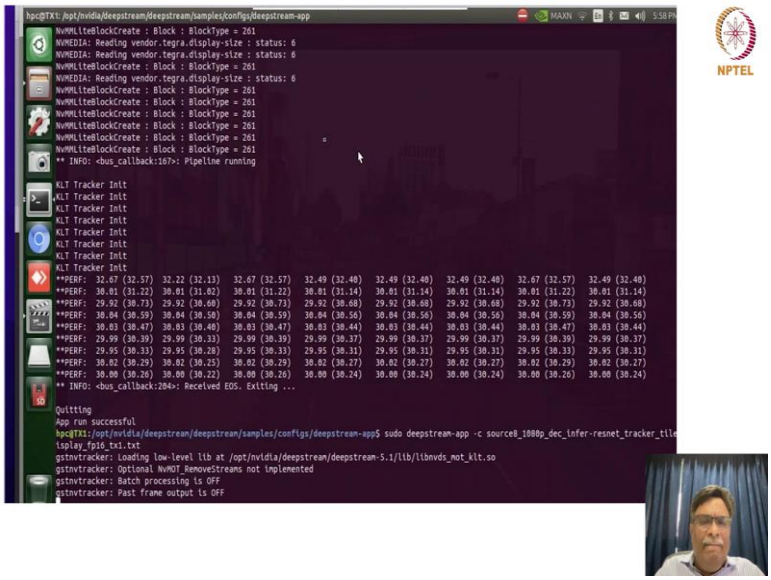


(Refer Slide Time: 16:07)



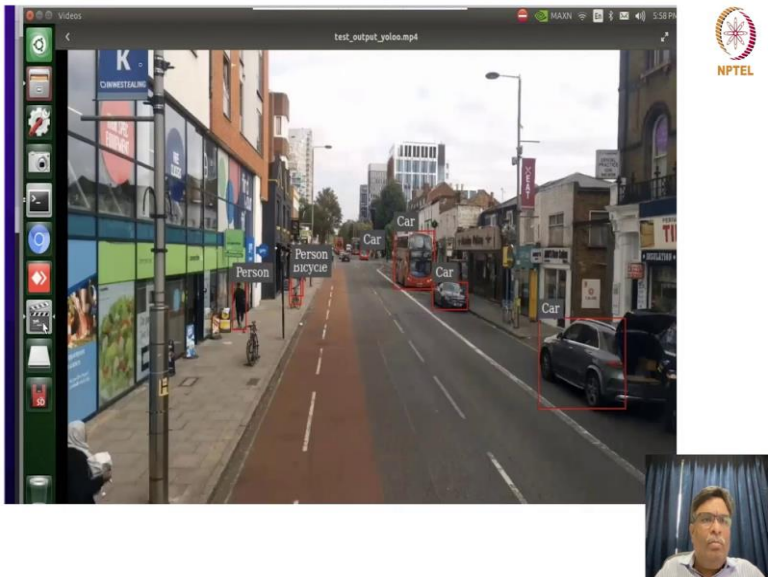
So, yes so, you will be able to appreciate this of multiple listing little bit slow.

(Refer Slide Time: 17:15)

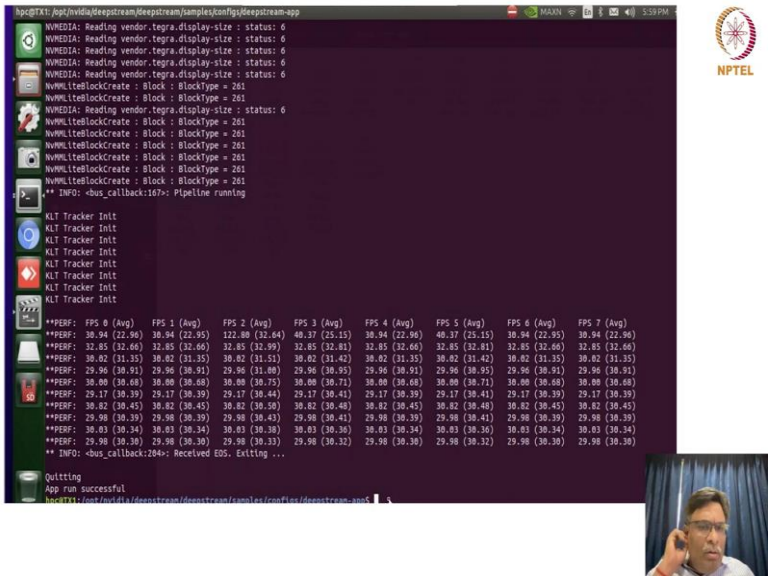


Yes. So, the only thing which I feel will not be able to show you, you can show it as a video or a snapshot because it is multiple images getting processed. So, each of this particular thing which you are seeing here ok, is for a specific tile wherein the input is getting processed. So, I do not think that would be available.

(Refer Slide Time: 18:05)



(Refer Slide Time: 18:26)



1 minute, let me just open that and see if it is possible on screen somehow let me just let me again try it if I can show you that, but we should be able to see that know, why not able to see the GUI here. Just we will try to actually now try to capture some snapshot and put it on the desktop from there. So, I suppose I am done with it and.