



Visionary Course - Energy Al Week 09

Apr. 29, 2022 Seokju Lee





Week 09a – Depth Estimation

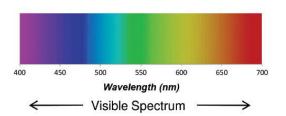
Human Visual System

"About **half** of neocortex in humans is devoted to **vision**." [1]

Geometric perception oo Where is it?

→ How can we perceive the 3D world?

Low-level visual signal



Semantic perception

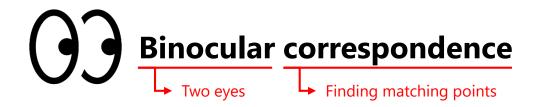
Two independent

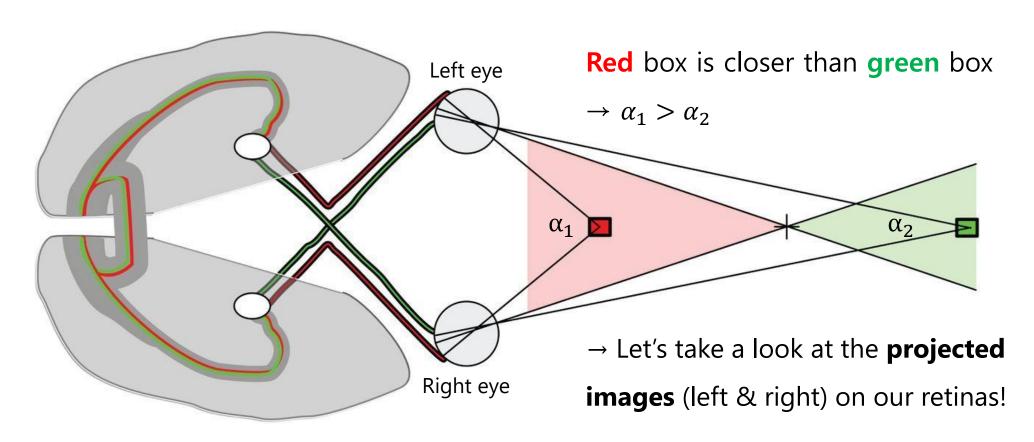


^[1] Barton, Robert A. "Visual specialization and brain evolution in primates." Proceedings of the Royal Society of London (1998).

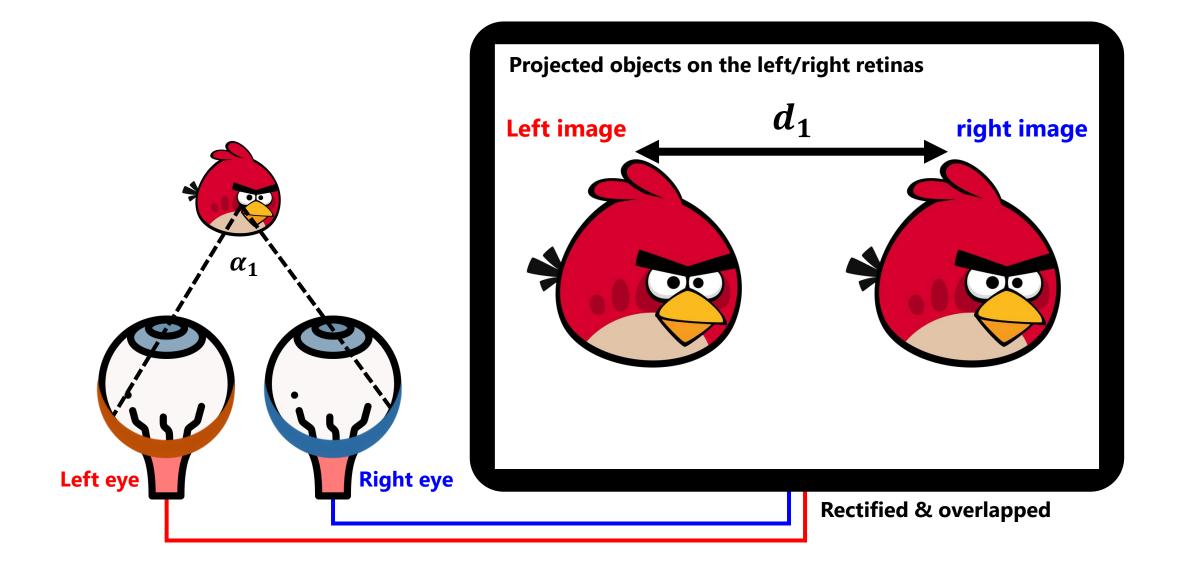
^[2] M. A. Goodale, et al., "Separate visual pathways for perception and action." Trends in Neurosciences (1992).

Human 3D Perception: Binocular System

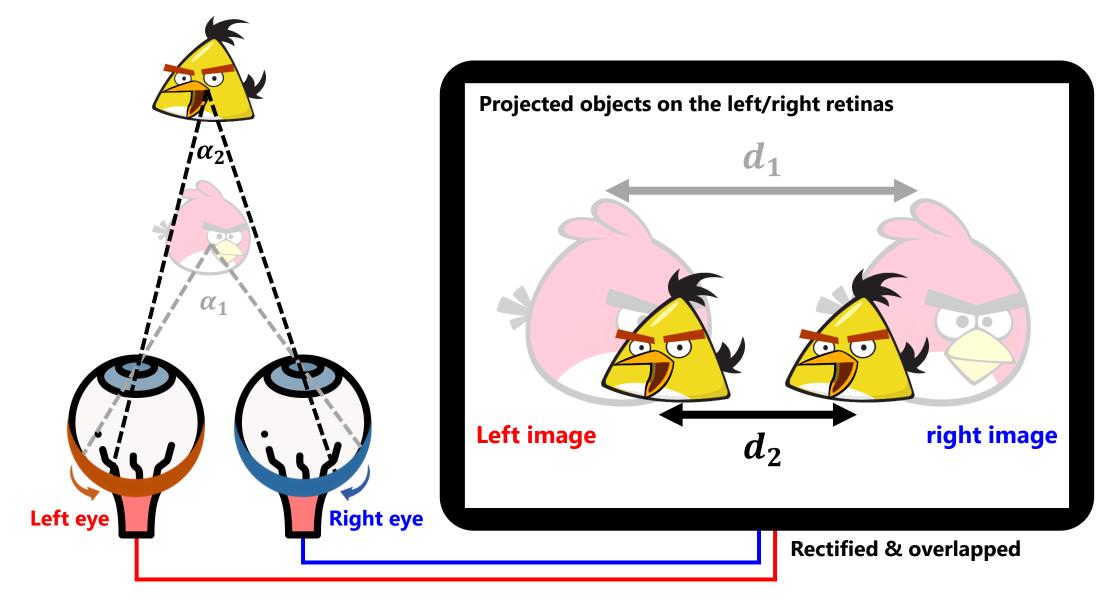




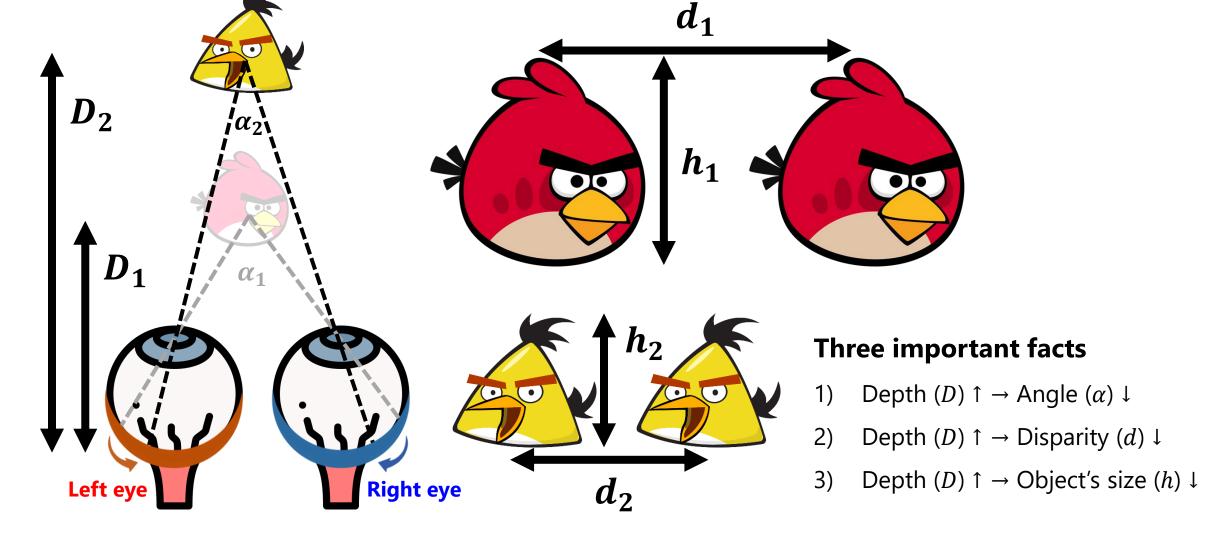
Binocular Correspondence: Near Objects



Binocular Correspondence: Far Objects



Binocular Correspondence: Disparity (d)



Human perception is remarkably flexible!

Other depth cues to perceive 3D world?

Perceiving 3D World with Monocular System



Department of Transportation, USA

Driving With One Eye

Not only is it possible to drive with one eye (assuming that you have good vision in your remaining eye) it's also legal in many states. Though there isn't a federal law dictating whether people with monocular vision can drive, it is up to each state to determine these regulations. [7] Of course, like everything else, it will take some time and practice to get used to driving with one eye, so you may want to look into specialized driving classes in your area.

"한 눈 잘 보이면" 시각장애인도 '1종 보통' 면허 딴다

VIEW 19,534 | 2016.04.18 05:20

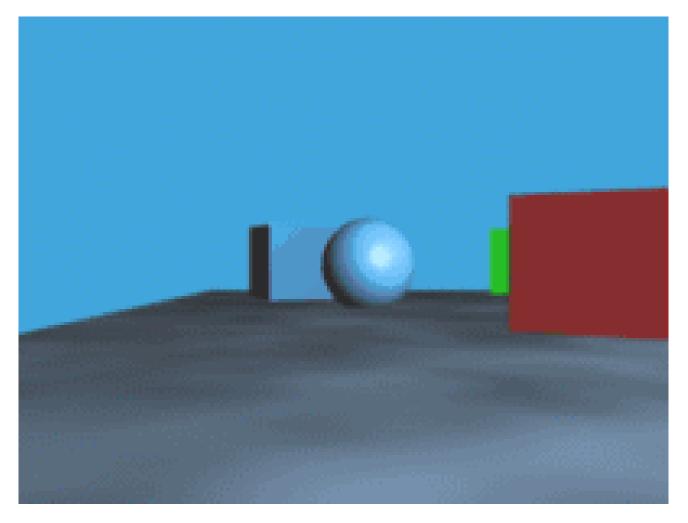


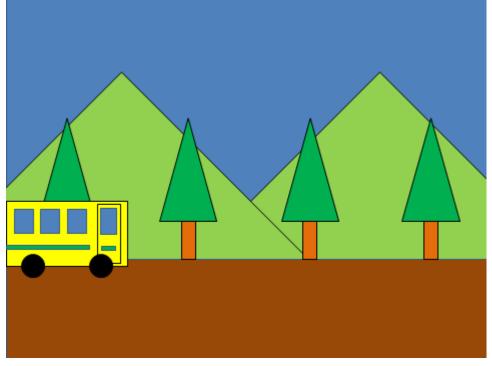




Ⅰ '시력 0.8 일정 수준 이상 시야' 경찰, 단안 시각장애 면허기준 마련

Monocular Depth Cues (1): Motion





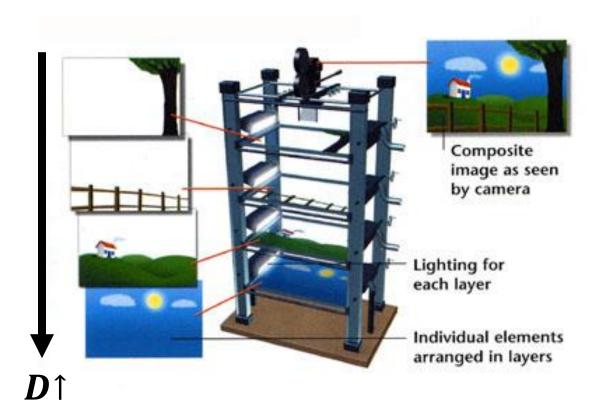
"Motion parallax"

→ Differences in image motions between objects
 at different depths (motion ↑ → depth ↓)

[1] Figure credits: Wiki – motion parallax

Motion Parallax in Film Industry

Walt Disney's Multiplane Camera





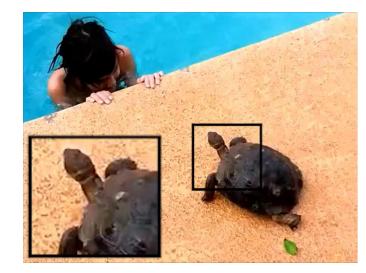
→ It allows us to feel the 3D effect!

Motion Parallax by Animals

Strange behaviors of animals...







[1] YouTube, "Getting down with some biggie!"

[2] YouTube, "Super cute budgie bobbing his head"

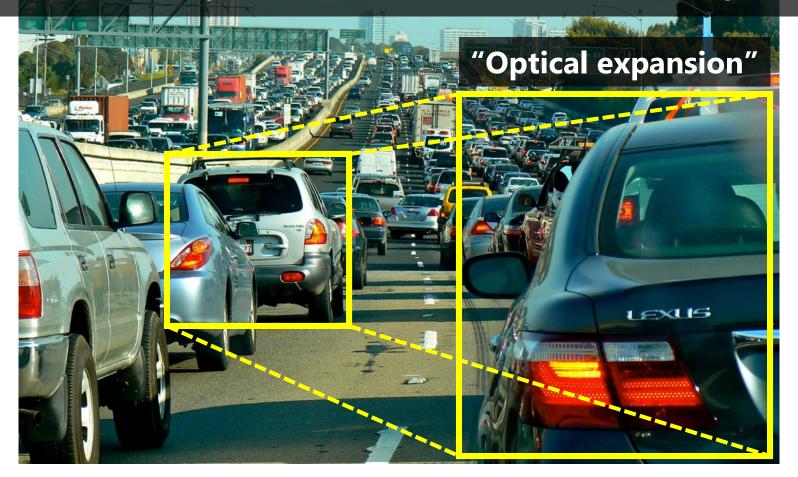
[2] YouTube, "Headbanging/bobbing Turtle"

Have a reason... 3D perception!

Monocular Depth Cues (2): Relative Size

Here is an assumption:

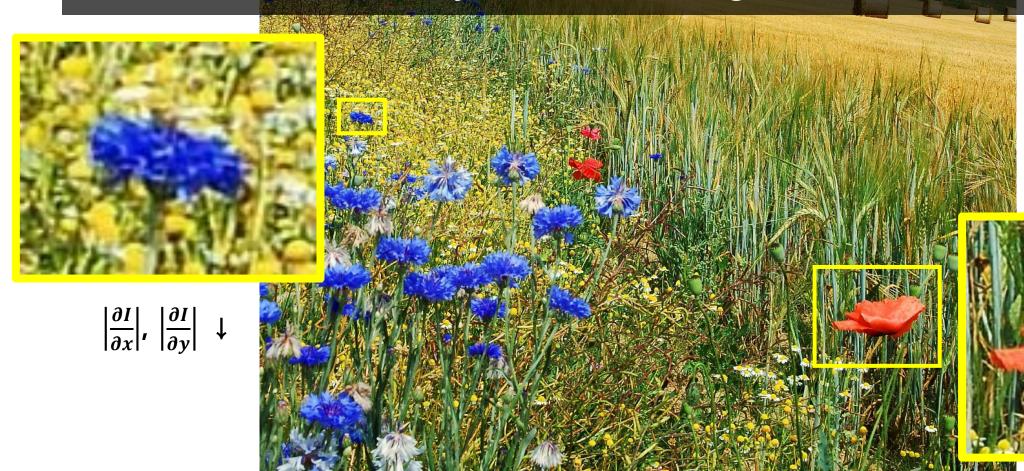
We need to know if both objects are of the same class! (semantic prior knowledge)



[1] Figure credits: Daniel Storage

Monocular Depth Cues (3): Texture Gradient

"The farther the object is, the more vague the texture becomes"

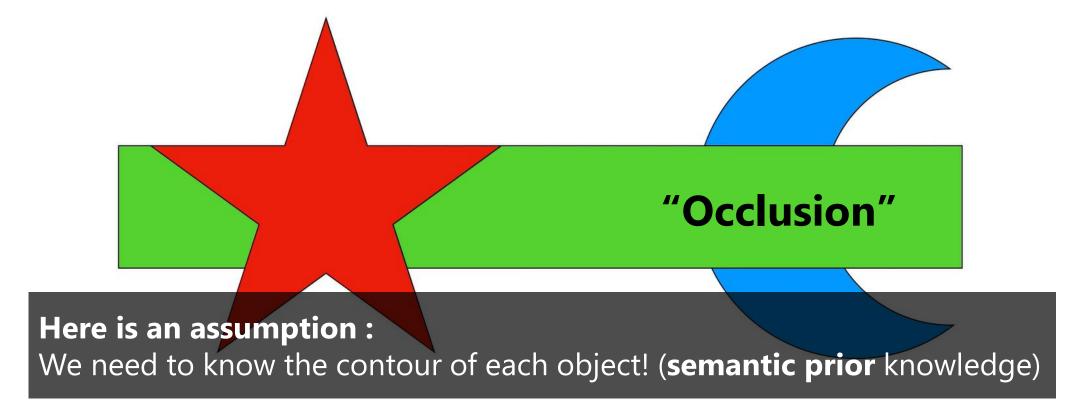


 $\left|\frac{\partial I}{\partial x}\right|, \left|\frac{\partial I}{\partial y}\right| \uparrow$

Monocular Depth Cues (4): Interposition

Object A is blocking our view of Object B

→ Object A is closer to us than Object B



[1] Figure credits: Daniel Storage

Monocular Depth Cues (5): Linear Perspective



Optical Illusion by False Linear Perspective

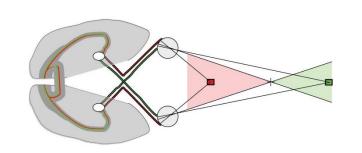


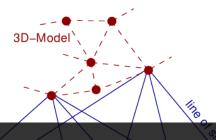
[1] YouTube, "Ames Room (Philip Zimbardo)"

Summary of 3D Perception: from Human to Machine

Binocular correspondence

→ "Stereo matching" in computer vision

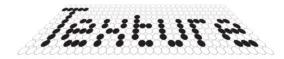




Motion parallax Applied to train deep neural networks *Structure-from Motion (SfM) in computer vision with different learning mechanisms!

Other monocular cues











3D Vision Applications

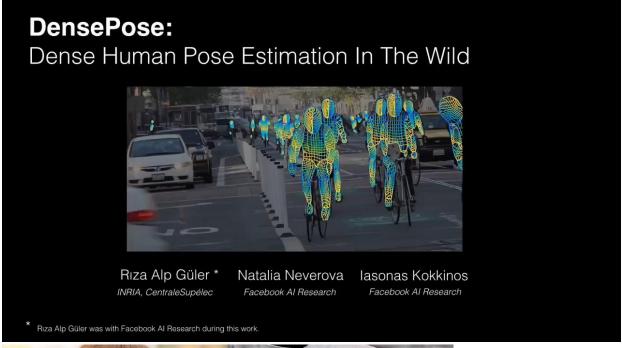
3D Reconstruction



Speed 4x CoEx, IROS'21









DensePose, CVPR'18

3D human pose estimation

EG3D, 2021

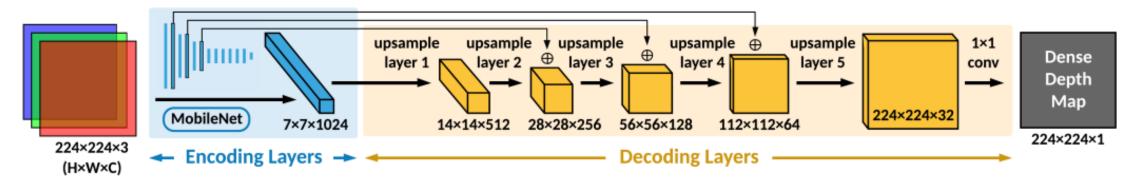
3D face reconstruction



Week 09a – Depth & Pose Estimation on Jetson Nano

References

Moncular depth estimation (https://github.com/dusty-nv/jetson-inference/blob/master/docs/depthnet.md)



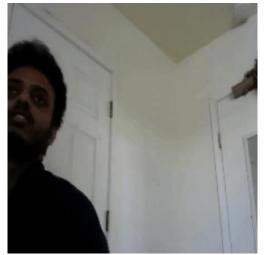
FastDepth, ICRA'19



References

Human pose estimation (https://github.com/dusty-nv/jetson-inference/blob/master/docs/posenet.md)







→ Locating various body parts (human body + hands) that form a **skeletal topology** (keypoints + links).

^[1] Zhe Cao, et al., "Realtime multi-person 2d pose estimation using part affinity fields." CVPR 2017.

^[2] Bin Xiao, et al., "Simple baselines for human pose estimation and tracking." ECCV 2018.

Before starting

*Your basic workspace is here: "cd ~/jetson-inference/build/aarch64/bin" Every code is pre-built in this path.

Run mono-depth with image files

Q1. Check the original image file "images/trail_8.jpg". Run "python3 depthnet.py images/trail_8.jpg images/test/output_trail_8.jpg". Open the result image file, and what is in the file? What is different from "segnet.py" in the previous class? Please guess the meaning of the color in the figure.

Try other models for mono-depth

```
Q2.1. (optional) Check your installed model
"cd ~/jetson-inference/build/aarch64/bin/networks"
"ls"
you can see the model folder installed.
Q2.2. (optional) After Q2.1, follow the command to download models.
"cd ../"
                           move back to ~/jetson-inference/build/aarch64/bin/
"cd ../../"
                          move to ~/jetson-inference/
"cd tools/"
                          move to ~/jetson-inference/tools/
"./download-models.sh"
After download,
"cd ../"
                           move to ~/jetson-inference/
"cd build/aarch64/bin/"
                          move to the basic workspace
```

Q2.3. Run specified model "python depthnet.py --network=monodepth-fcn-resnet50 images/trail_8.jpg images/test/output_trail_8_res50.jpg".

Discuss how the mono-depth of the same object is different for each model.

Try using different options for mono-depth

- Q3.1. Before running the python code, execute the command "cat depthnet.py". Could you find the code about the colormap option? What is the default value?
- Q3.2. Try Q2.3 with different colormap options. Run "python depthnet.py --network=monodepth-fcn-resnet50 --colormap=magma images/trail_8.jpg images/test/output_trail_8_res50_magma.jpg".

 Please try other colormaps (plasma, turbo, etc.), which are listed in the below link:

 https://matplotlib.org/stable/tutorials/colors/colormaps.html
- Q3.3. Select visualization option with live mono-depth.

 Run "python depth.py --flip-method=rotate-180 csi://0 images/test/output_depth.mp4" 10 seconds and quit. Could you find something special in the output screen? How about in output depth file? Could you guess the default option?
- Q3.4. Run "python depth.py --visualize=depth -flip-method=rotate-180" and observe the difference from
- Q3.5. Control depth-size(default depth-size=1). Run "python3 depthnet.py --network=monodepth-fcn-resnet50 --depth-size=0.5 --flip-method=rotate-180".

Try mono-depth with your own objects

- Q4.1. Locate your camera on the angle that includes more than one object. Run "python3 video-viewer.py --flip-method=rotate-180 csi://0 images/my video.mp4"
- Q4.2. Try observing Mono Depth of the video with 3 different colormap options. Please run "python depthnet.py --colormap=plasma images/my_video.mp4 images/test/my_video_plasma.mp4", "python depthnet.py --colormap=turbo images/my_video.mp4 images/test/my_video_turbo.mp4", ...
- Q4.2. Observe the results of mono-depth which are saved in ~/jetson-inference/build/aarch64/bin/images/test folder. Could you determine which object is located closer?

Pose estimation on images

Q5.1. Locate your image files and modify/run "./posenet.py "images/humans_*.jpg" images/test/pose_humans_%i.jpg"

Pose estimation from video

- Q6.1. Run "python posenet.py --flip-method=rotate-180 csi://0"
- Q6.2. Run "python posenet.py --network=resnet18-hand --flip-method=rotate-180 csi://0"
- Q6.3. Try different networks and discuss how each output is different. What does each node (keypoint) and link mean? You can check the models in the below link:

https://github.com/dusty-nv/jetson-inference/blob/master/docs/posenet.md#pre-trained-pose-estimation-models

Some useful tips while debugging

*Sometimes, the python process does not respond. In this case, please terminate the process with ctrl+c. If it still does not respond at all, forcibly stop the process with ctrl+z, and check the running process name with the ps -a command, and then type sudo pkill -9 [name-of-process] command to kill the process. If you don't shut it down, it will remain as a zombie and keep occupying the processor (CPU or GPU) in the background.

*Sometimes, the best solution for resolving an issue is just rebooting the system.

```
"cd" = "change directory"
"ls" = "list segment (files & directories)"
```

Conclusion







Image Classification

Object Detection

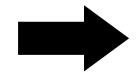
Semantic Segmentation







Mono Depth



Basic computer vision tasks



Let's start driving!

