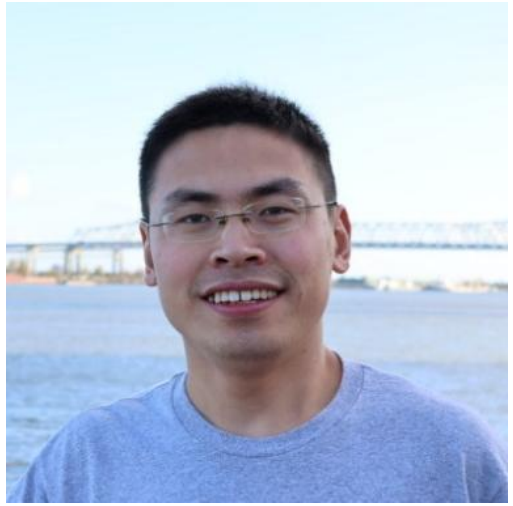
RBE474X/595-B01-ST: Deep Learning For Perception

Class 1: Introduction, Logistics And Sensors

Prof. Wei Xiao

Special thanks to Prof. Nitin Sankit



Prof. Wei Xiao

he/him/his

Assistant Professor

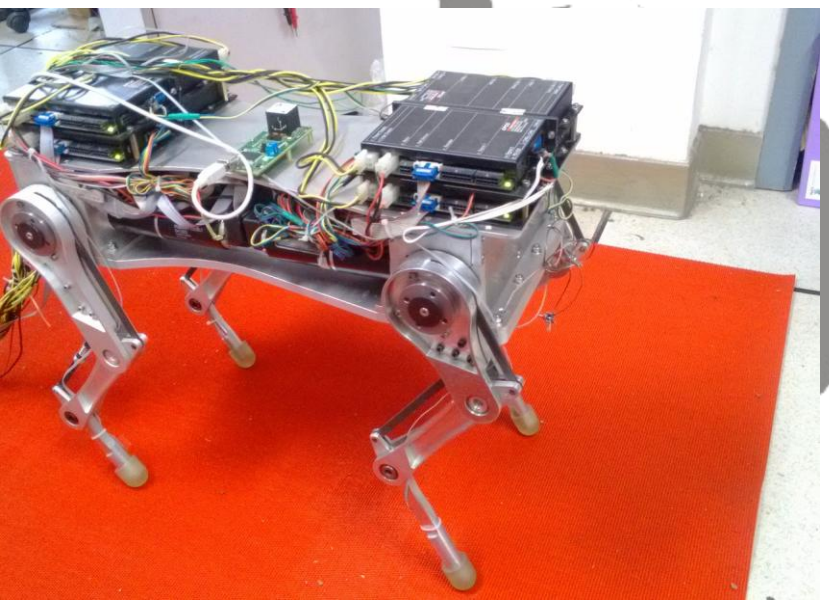
Safe Autonomy and Intelligence Lab (SAIL)

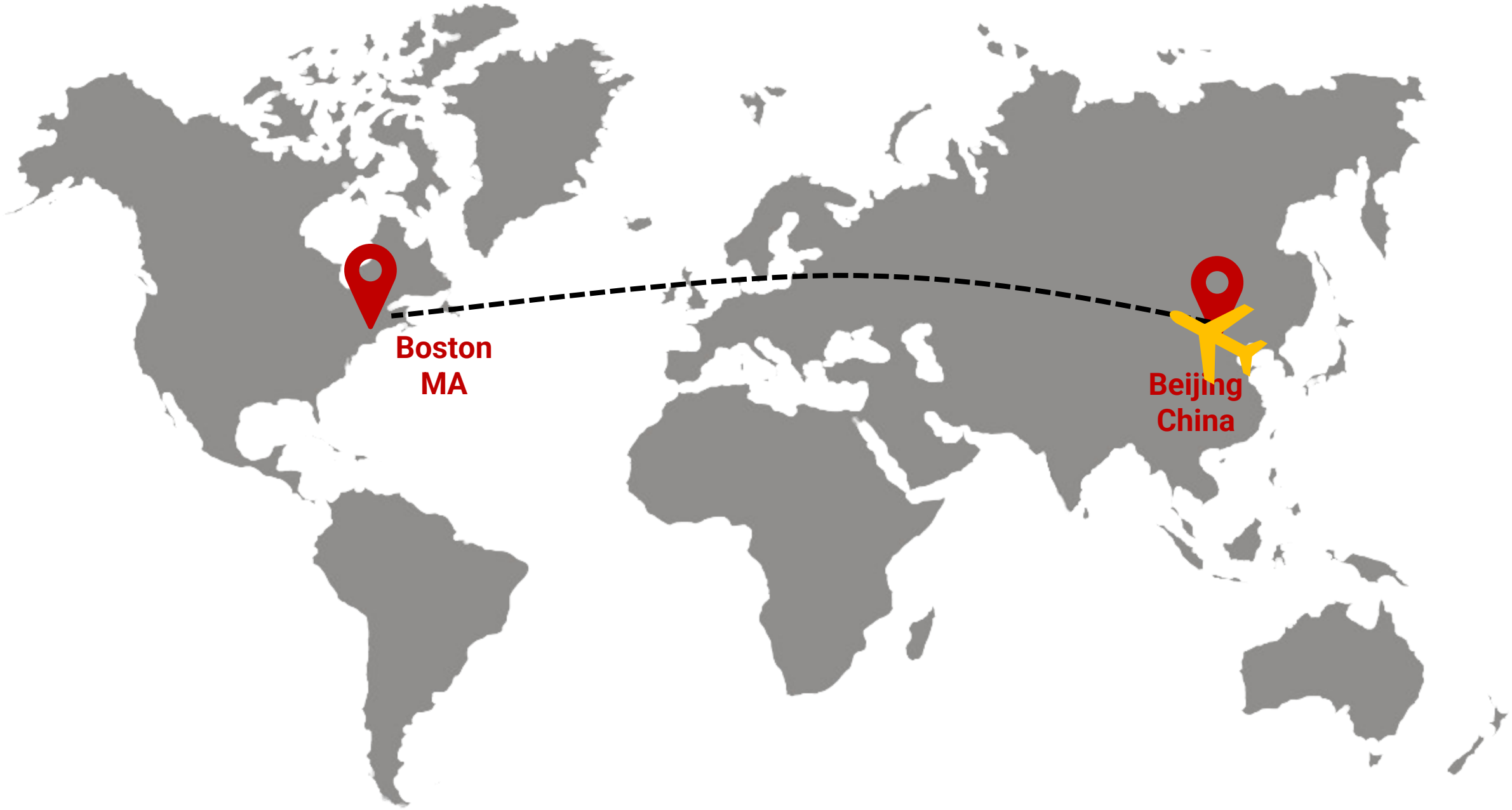
Robotics Engineering

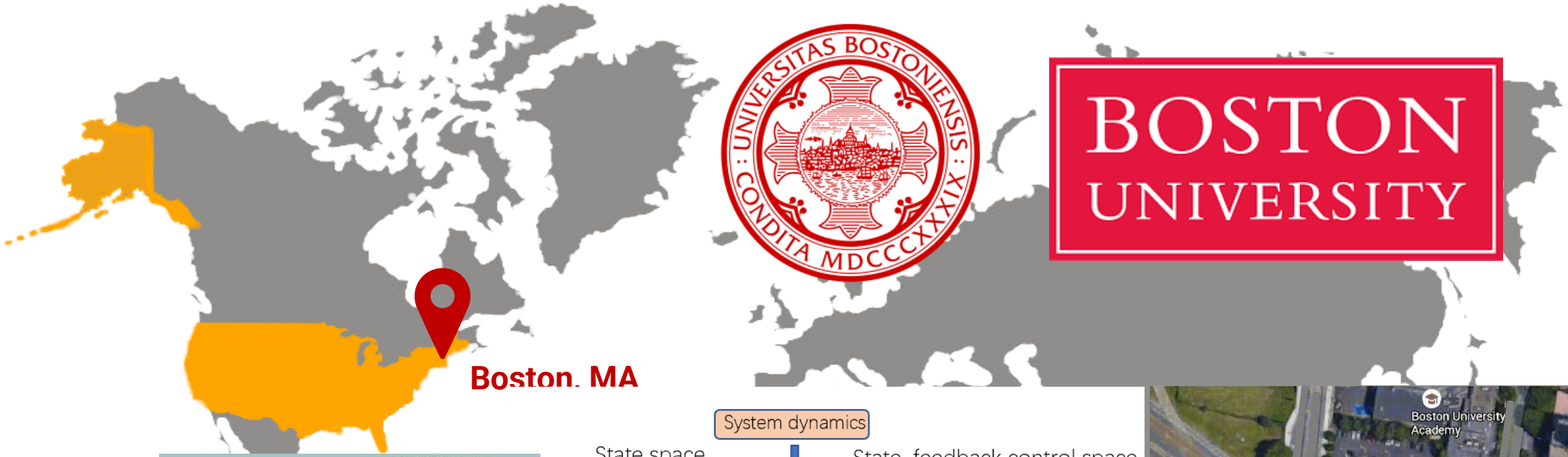

Sichuan,
China

University of Chinese Academy of Sciences

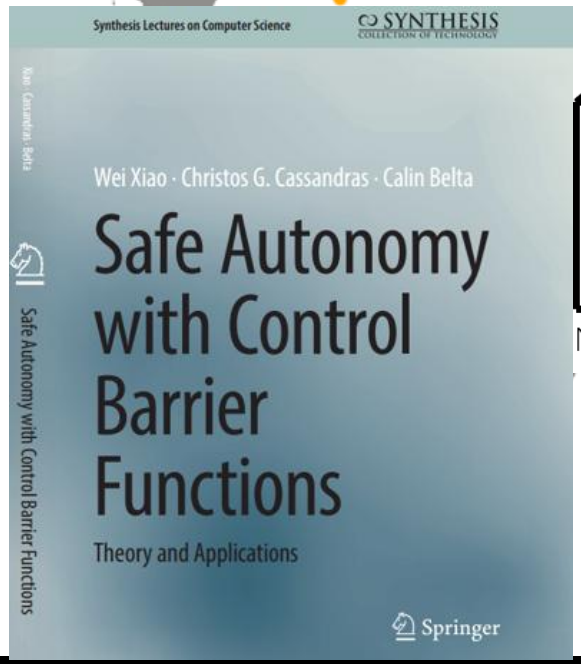
Beijing,
China

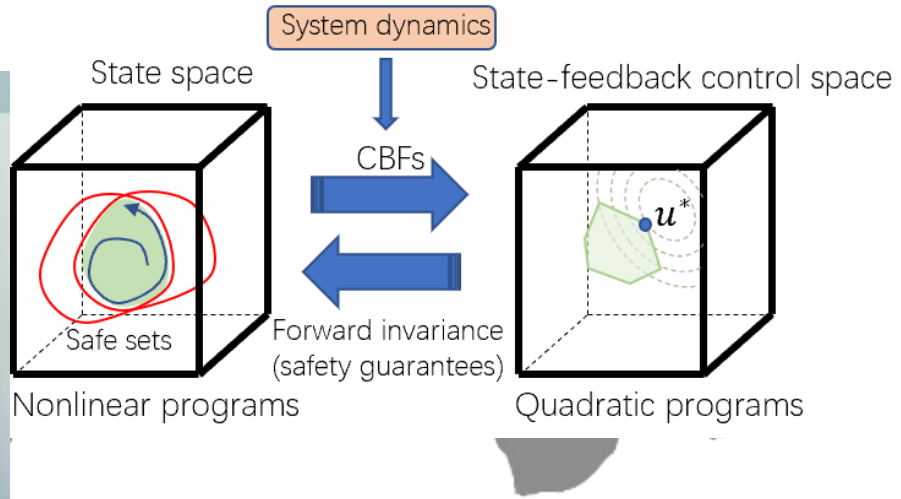


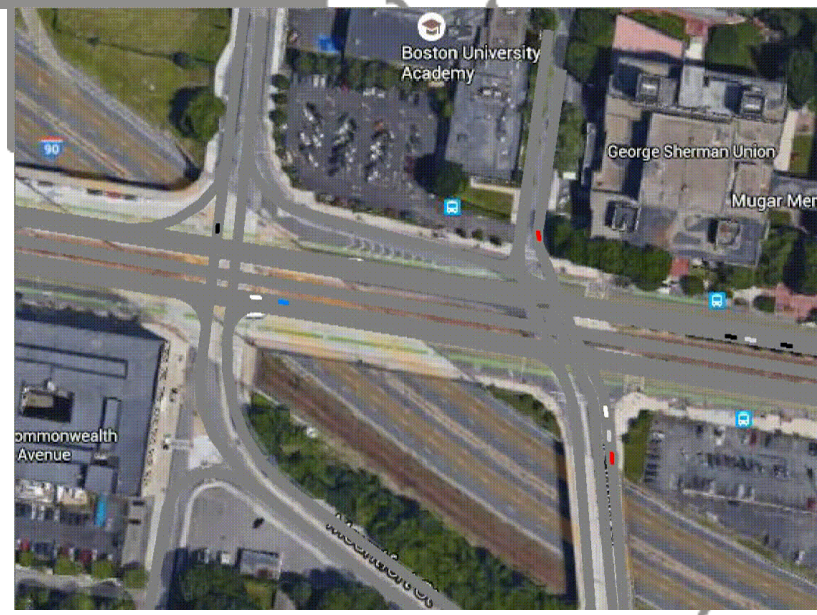

Boston
MA
Beijing
China



BOSTON UNIVERSITY

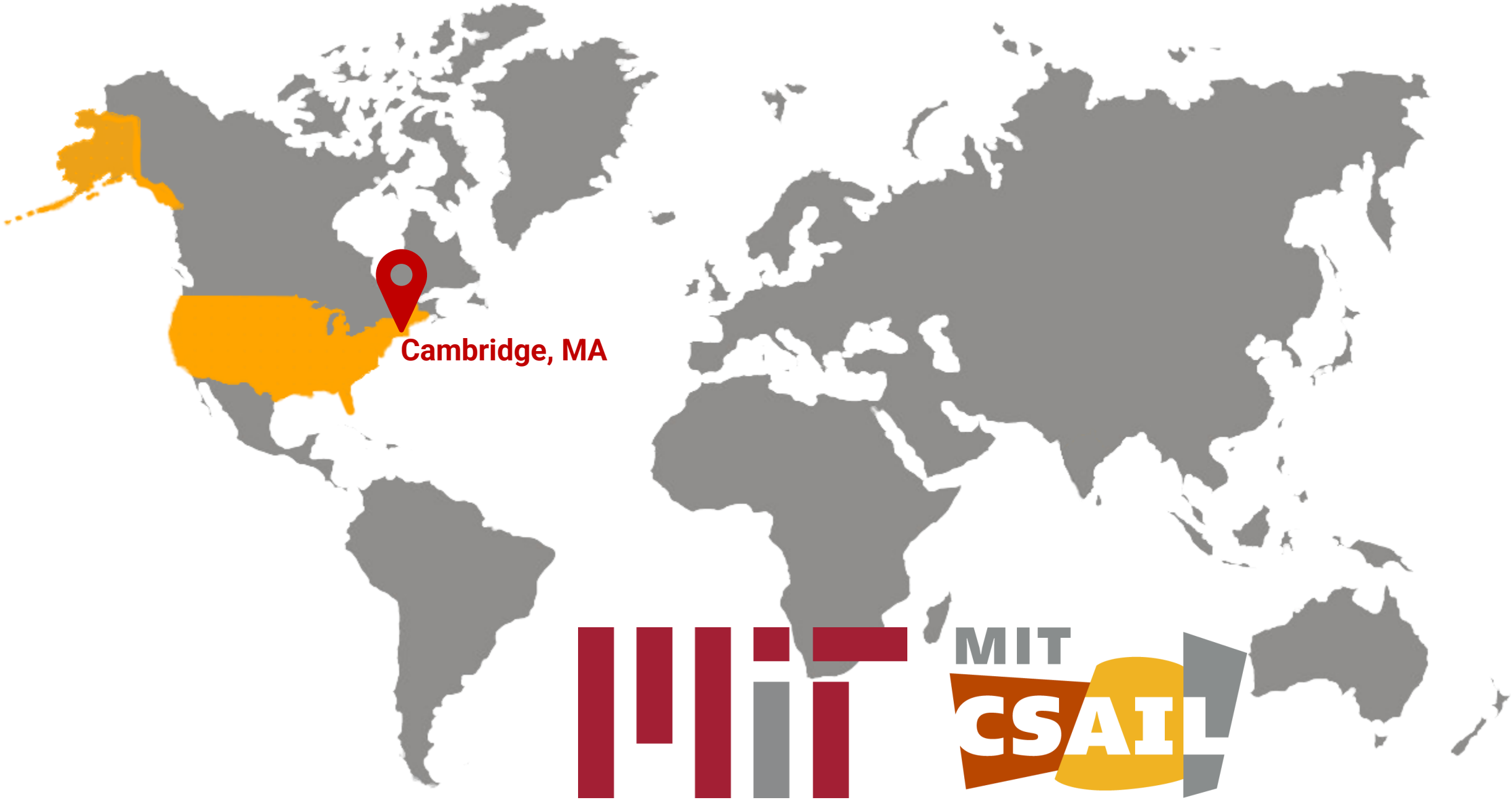


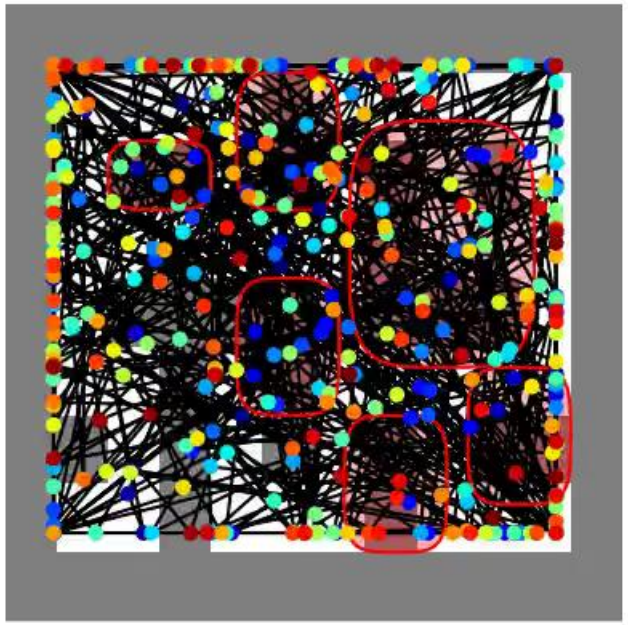


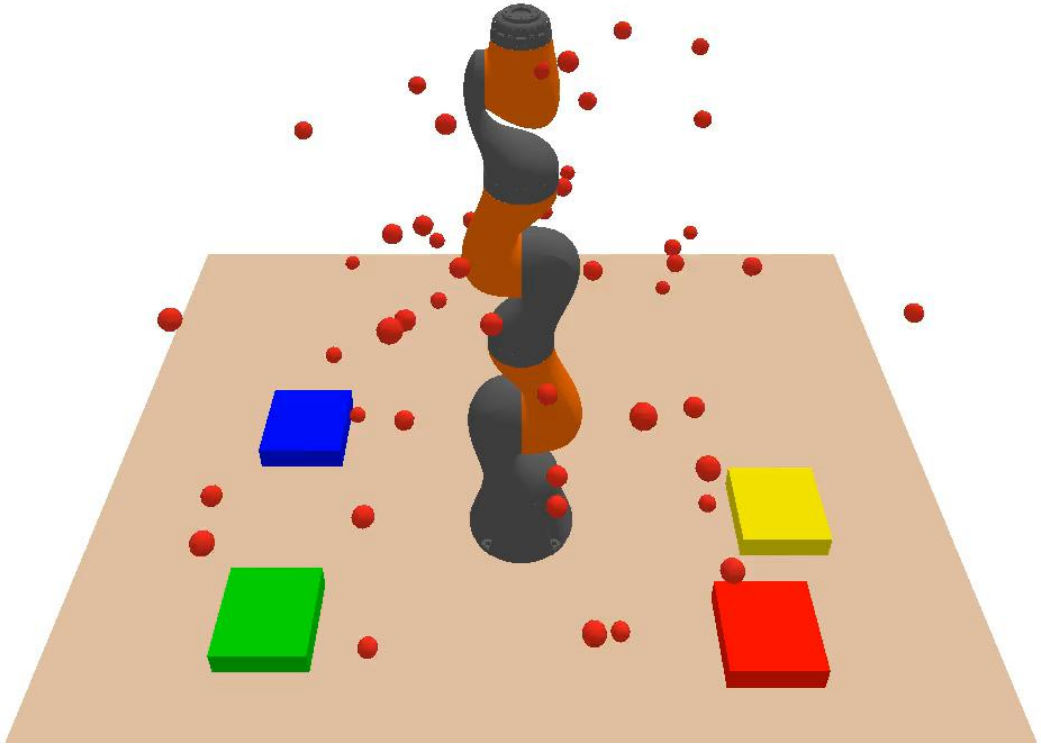


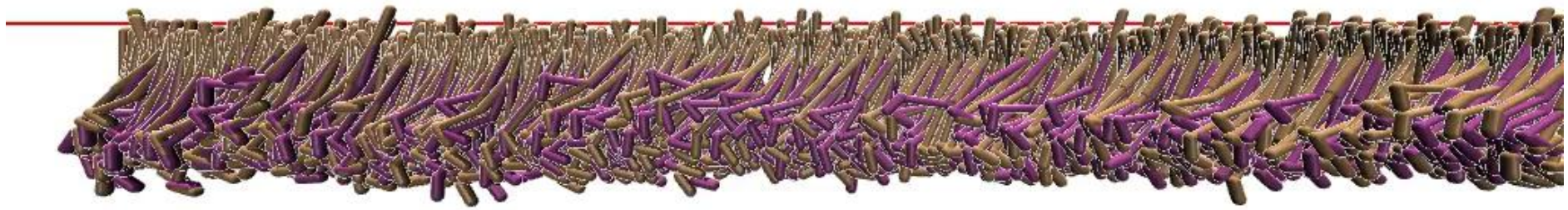
WPI


Boston University Outstanding Dissertation Award
2020 IEEE-CDC Outstanding Student Paper Award


Cambridge, MA
MIT
CSAIL





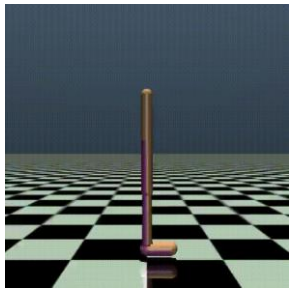


WPI

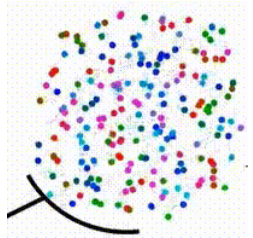
Assistant Professor

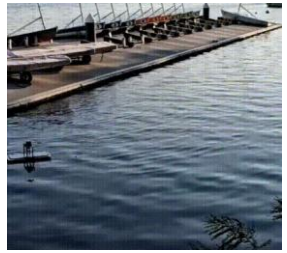
SCAN ME

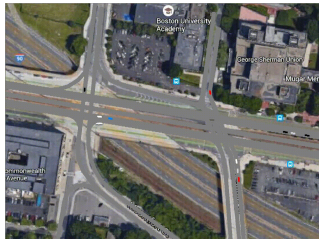
SAIL

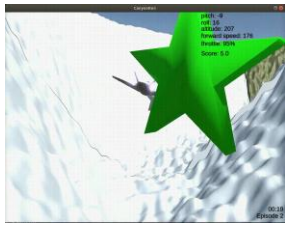


Safe Autonomy and Intelligence Lab



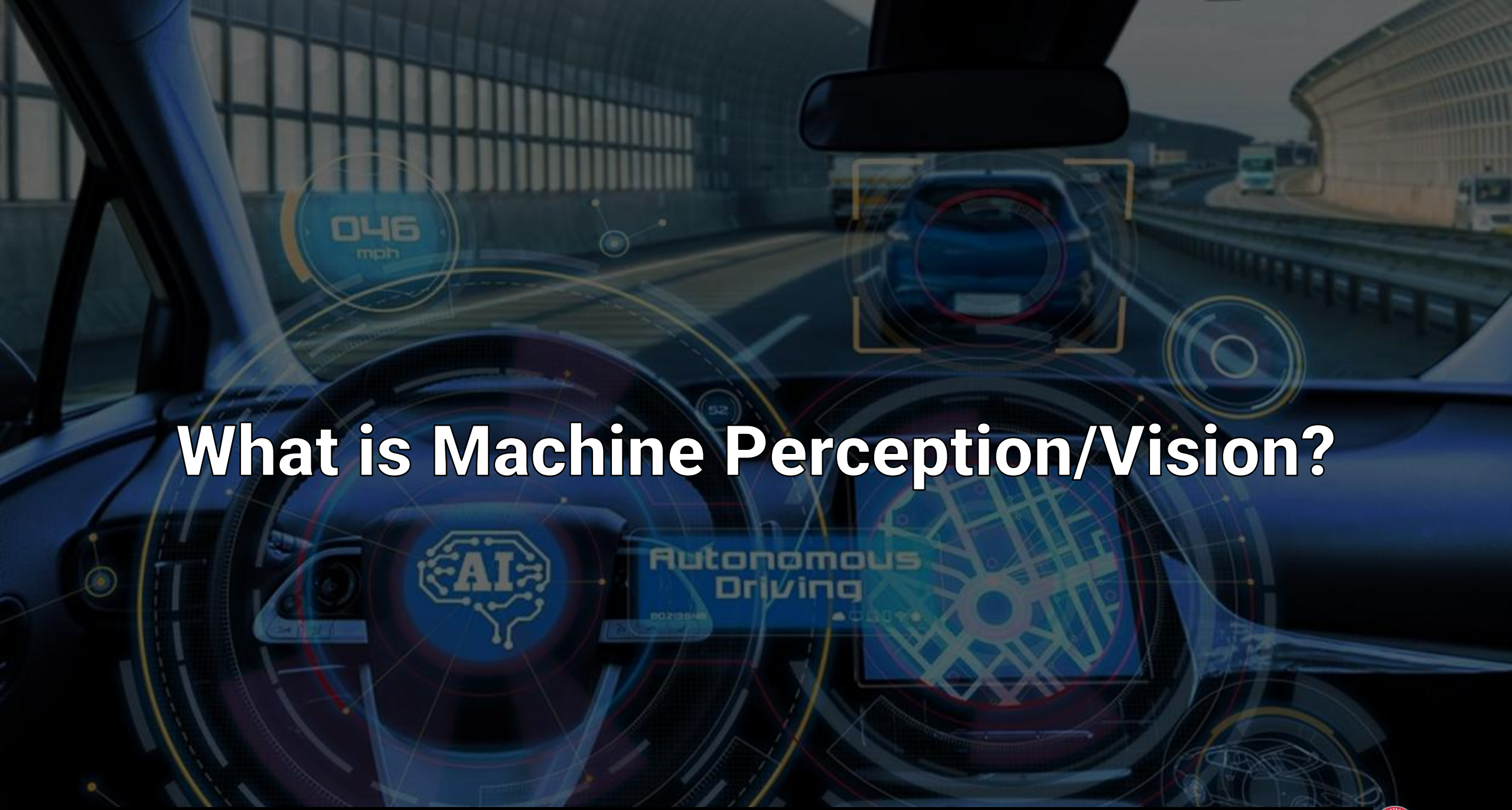






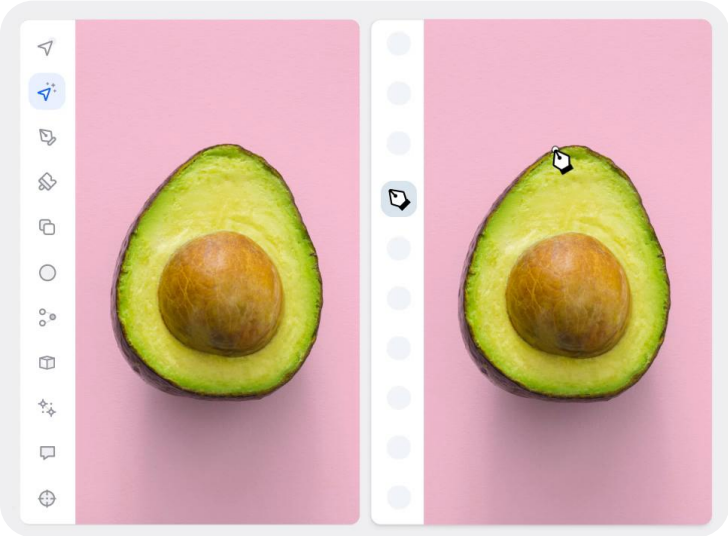
WPI

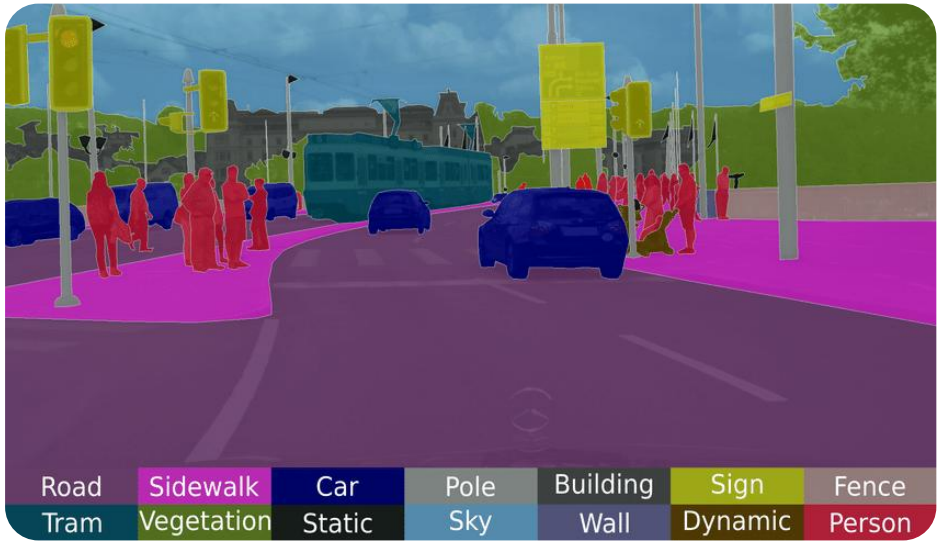
What is Computer Perception/Vision?

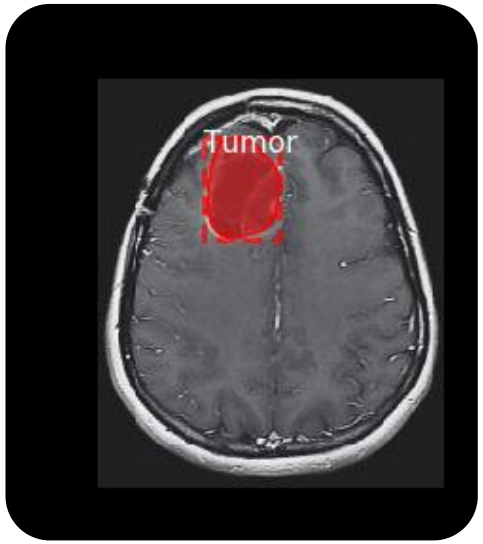


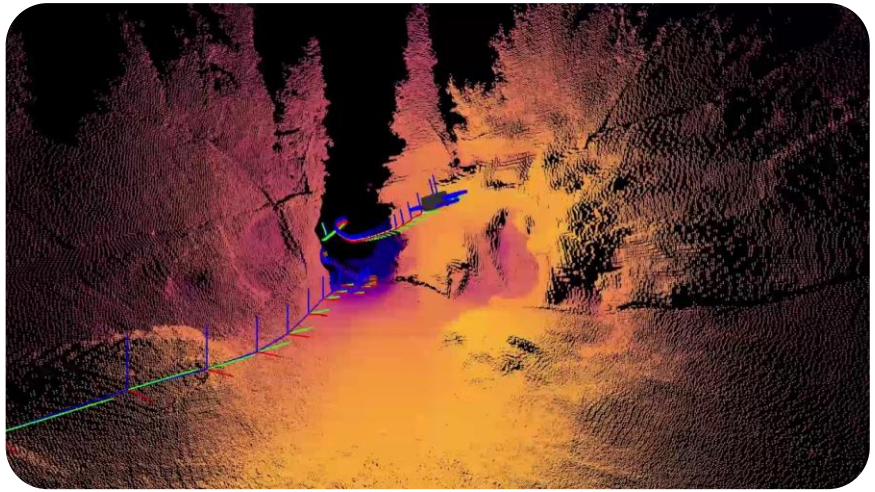
What is Machine Perception/Vision?

Some Applications

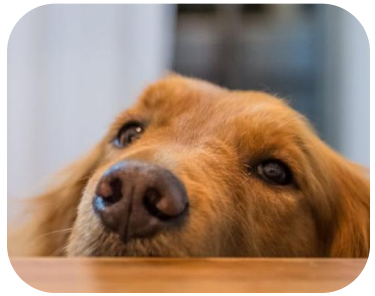








What is this course about?

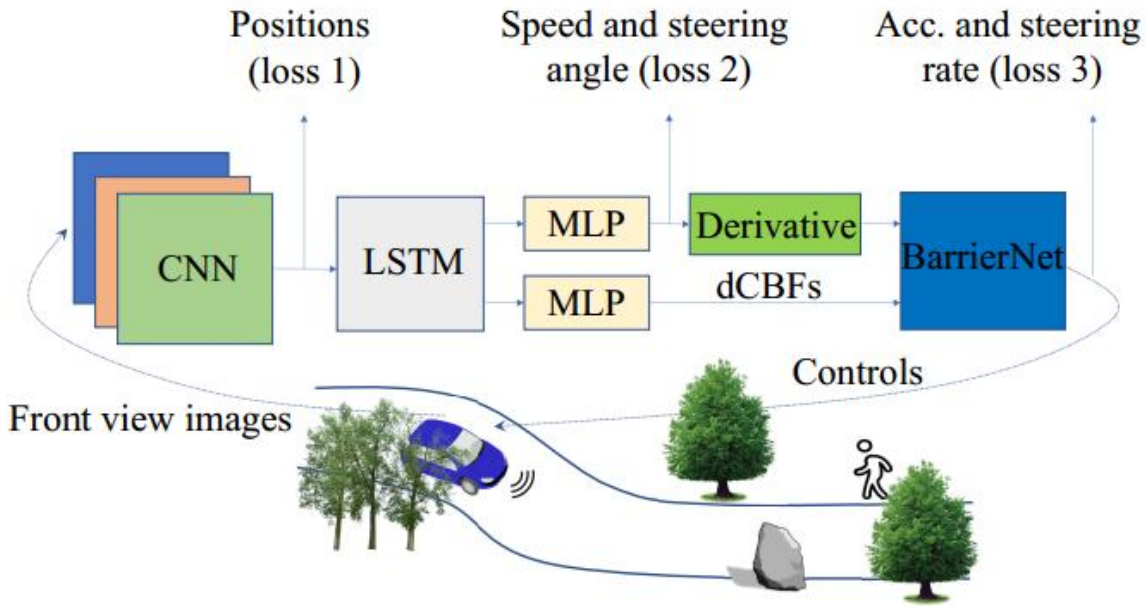


Magical "Box"

$p(dog|I)$

Same as $1 - p(cat|I)$

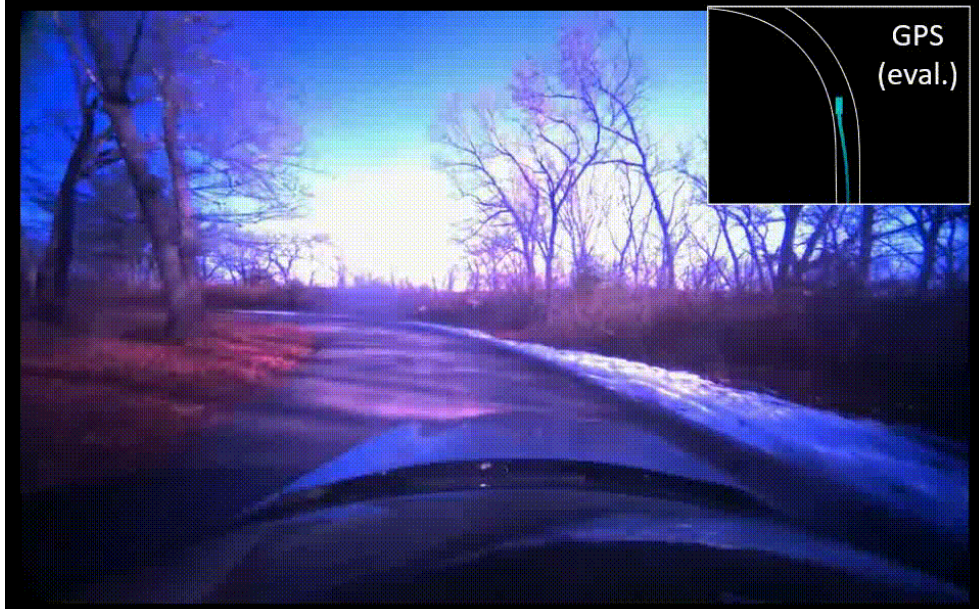
What is this course about?



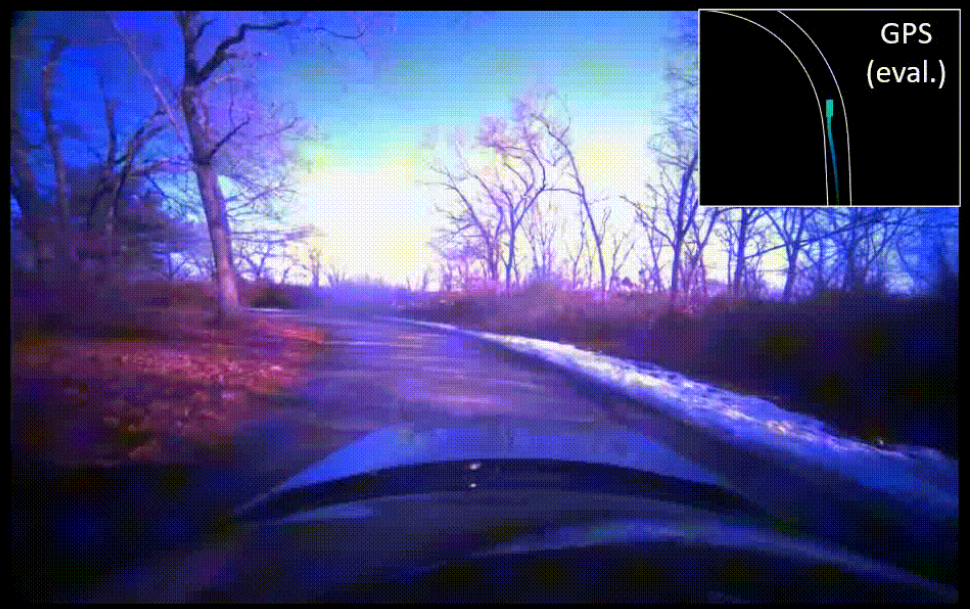
What is this course about?

Challenging Case : Reflection & Shadow

Regular End-to-end learning



BarrierNet



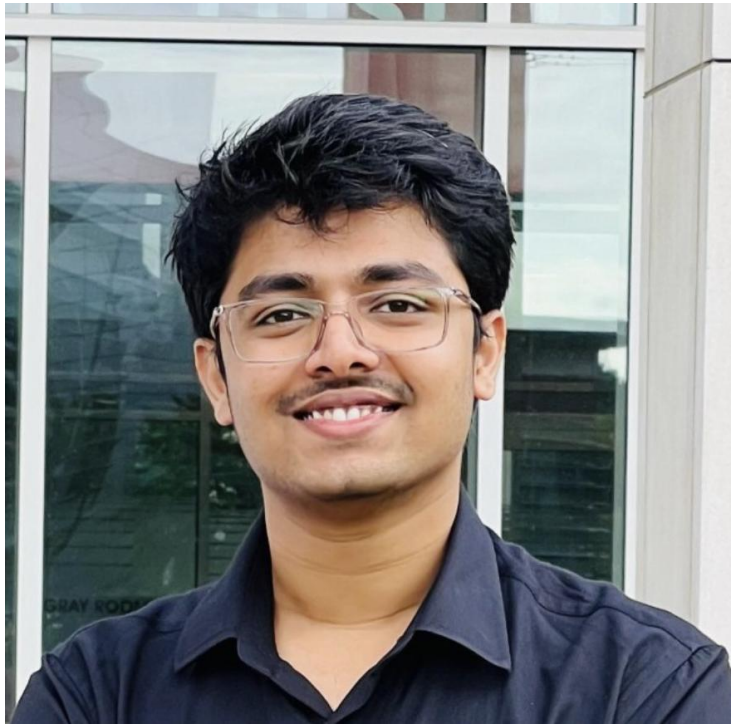
What is this course about?

Slight Street Sign Modifications Can Completely Fool Machine Learning Algorithms - IEEE Spectrum

How Tesla Teaches Cars to Stop (roboflow.com)

What is this course about?

Meet Your Awesome Grader/TA!



Khobragade, Shreyas Devdatta

he/him/his

Class Locations

Washburn 323

T-F | 4:00 PM - 5:50 PM

COVID-19/Flu/Other Virus is Serious!

Be safe and keep everyone safe!

- Please stay back home if you are **sick** (don't come to class), contact me for accommodation
- **Please inform me** if you are sick
- Contact me via **Piazza** if you need extensions and you are out of your late days when you are sick
- Check WPI Resources: https://www.wpi.edu/COVID-19-Resources
- Masks are **optional**

Class Resources

474x

https://canvas.wpi.edu/courses/79479

https://github.com/Weixy21/RBX-474x595

https://piazza.com/wpi/fall2025/rbe474x595

595

https://canvas.wpi.edu/courses/79516

Class Resources

- Assignments, deadlines and other instructions will be released on the **Website (Syllabus)**
- Use **Piazza** for all communication, under extreme urgency such as COVID-19 contact or medical emergency use my email at **wxiao3[at]wpi[dot]edu. Use '[RBE474X]' tag to get through my email filters**
- **Canvas** will be used for all grades

Contacting Me

- Use **Piazza** for all communication, under extreme urgency such as COVID-19 contact or medical emergency use my email at **wxiao3[at]wpi[dot]edu. Use '[RBE474X]' tag to get through my email filters**
- **Do not contact me over email regarding grades: This is against the university policy**
- Feel free to contact me on **Piazza** to fix an appointment to meet me for discussions
- If you are interested in my research, **do email me!**
- My office location is Unity Hall 287. **My office hours are ONLY for Serious Issues with Appointment Over Piazza!**

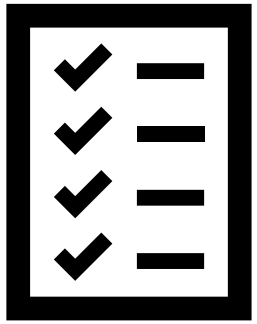
Home Rules!

The entire class is your **family**, treat everyone with respect and help each other.

4 late days
15% late penalty per day.

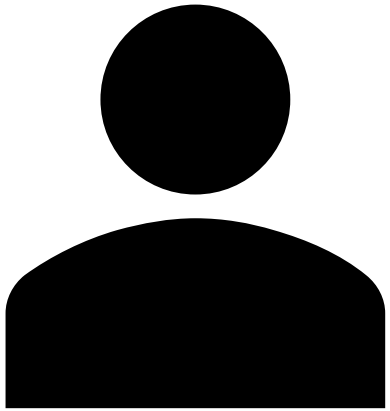
No Cheating

https://www.wpi.edu/about/policies/academic-integrity



Follow Submission Guidelines
25% penalty for not following them!

Home Rules!

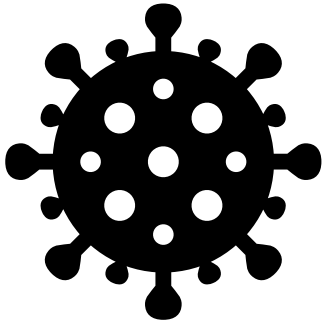


Strictly **In-person**

Pro ChatGPT and other LLMs
But NO CHEATING!

Home Rules!

Zero Tolerance
for plagiarism.

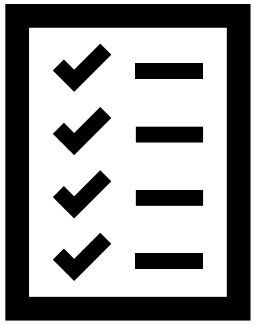


Zero Tolerance
for not following health and safety protocols.

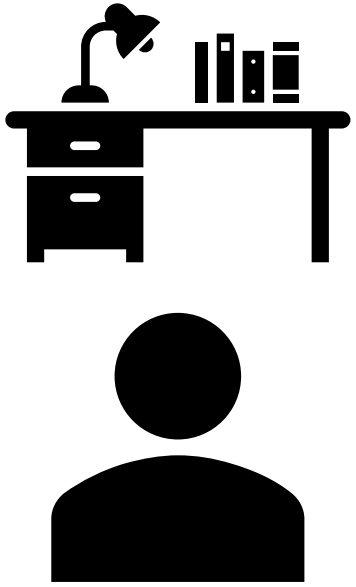
Class Ideology

- Learning how to implement **concepts from scratch**
  - Math to code and code to math
  - Deeper understanding with extra reading outside class
- Failure is **common!**
  - It's not called failure, it's called **"path to success"**
  - **If life is without ups and downs, you have not tried hard enough!**
  - Giving up is **not**; leaving things till the **last minute** is **not** either!
- Ask for **conceptual** help!
  - Instructors **will not debug code**
  - Instructors **will not fix your computer**
- You need to **manage your time** throughout the term
  - It is easy to fall behind (submit assignments on time!)

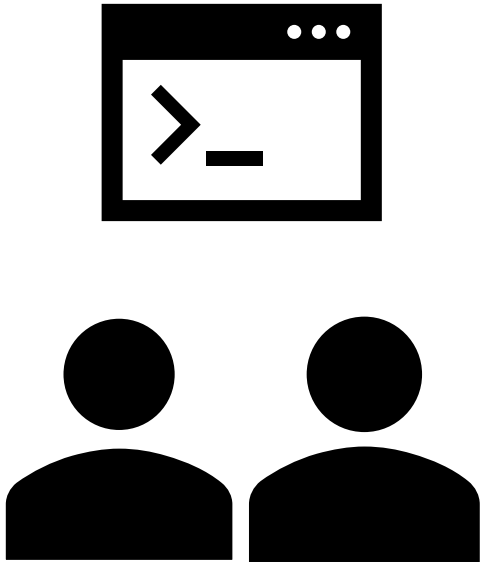
Assignments



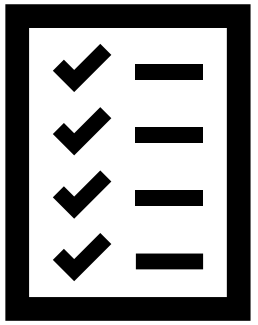
Four
Assignments



One
Individual
Homework

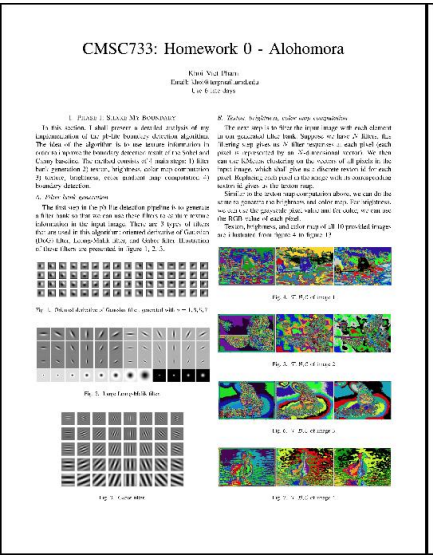


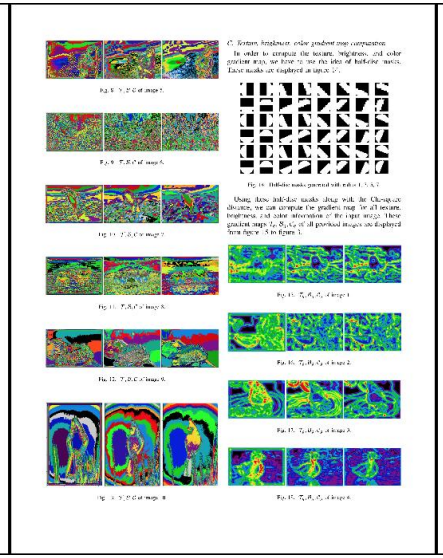
Two
Homework
Groups of 2 or 3

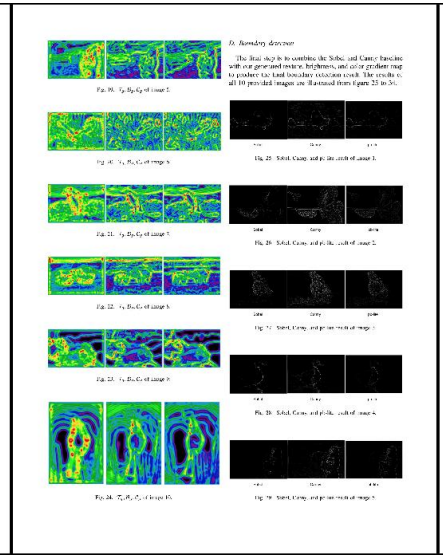


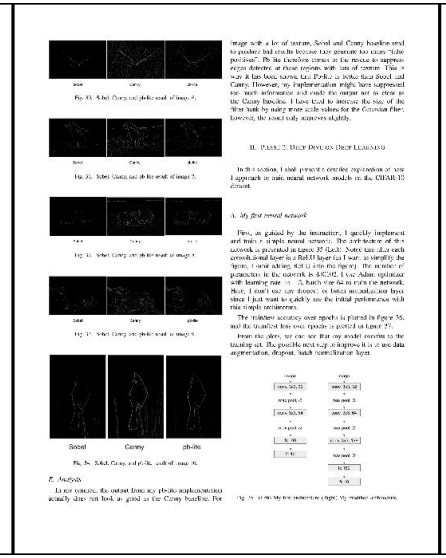
One
Course Project
Groups of 2 or 3
with oral presentation

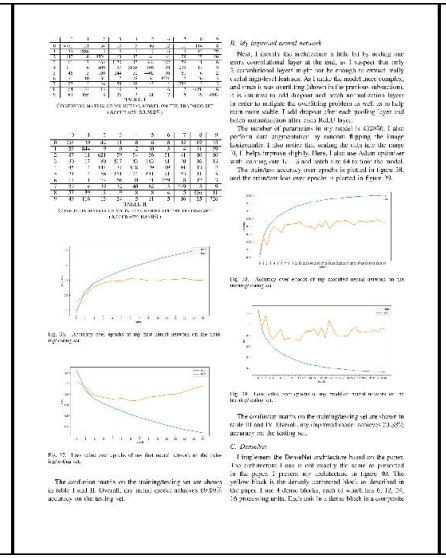
Reports

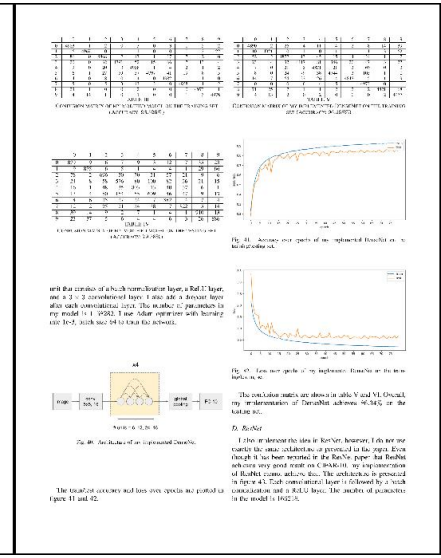

CMSC733: Homework 0 - Alohomora











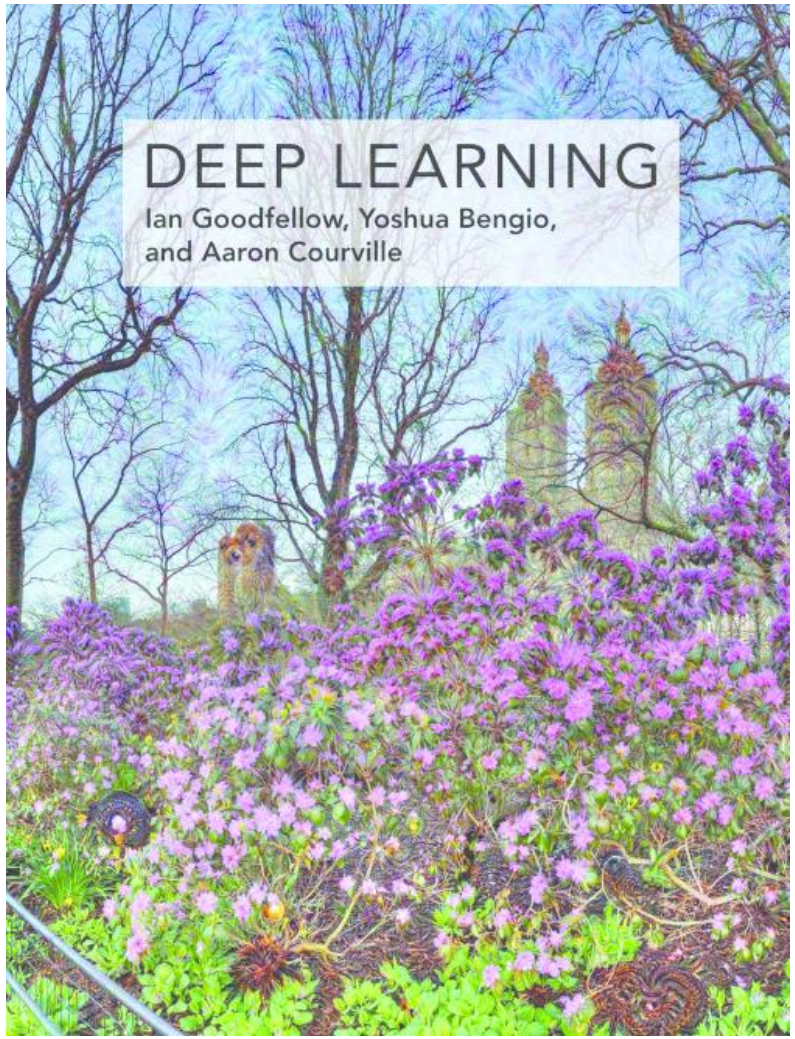
Evaluation

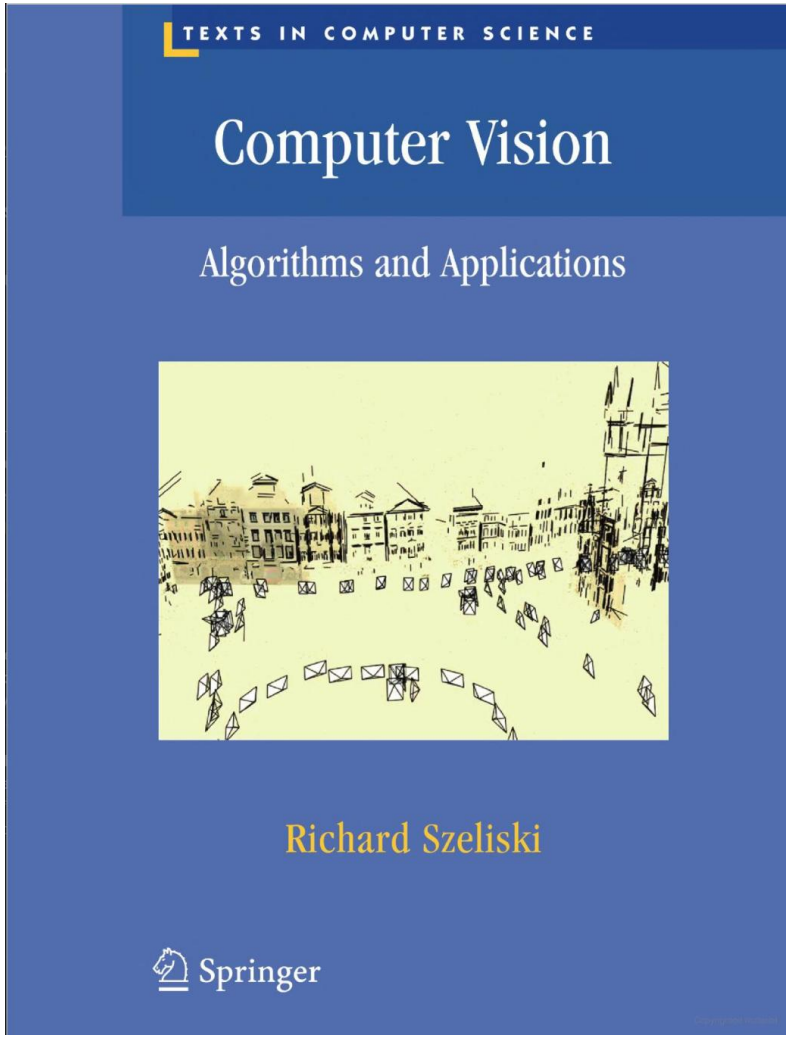
Assignment Name	Grade Percentage	Deadline
In-class quiz	5	
HW0: Alohomora **(Individual)**	10	10/31/2025
HW1: Nifty Neural Networks! **(Group of 2 or 3)**	15	11/14/2025
HW2: Dreaming Data! **(Group of 2 or 3)**	20	12/05/2025
Course Project: Learning for End-to-End Robot Policies: From Perception to Control **(Group of 2 or 3) – coding, report, and oral presentations**	50	12/12/2025

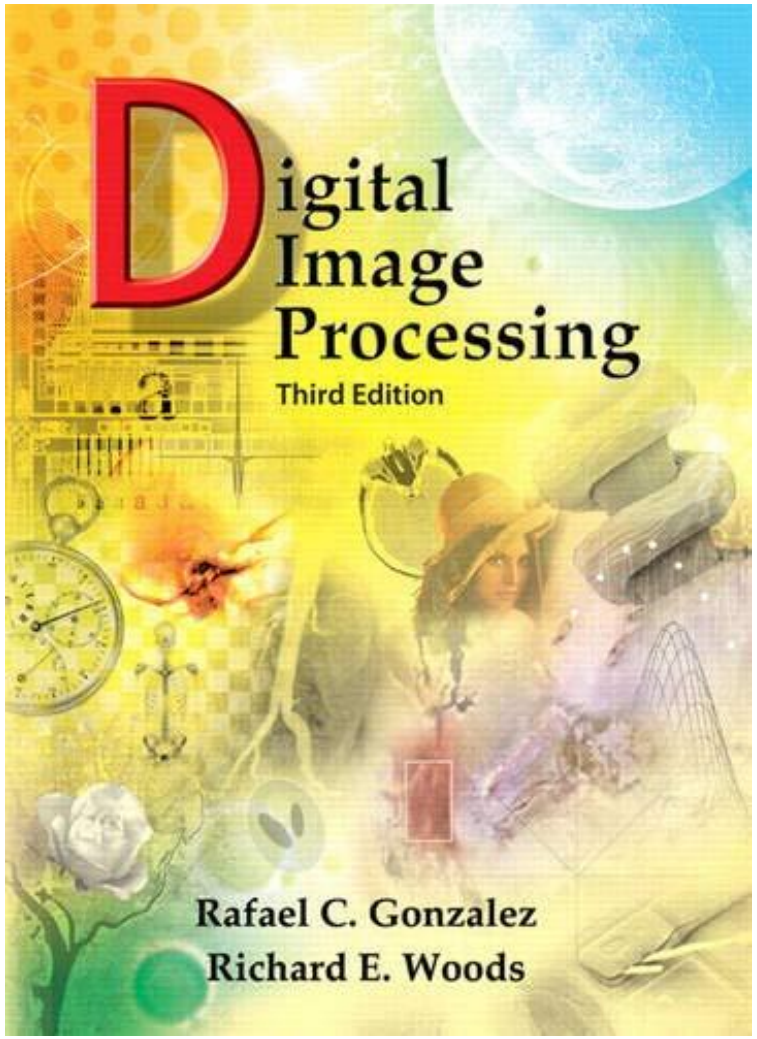
Instructions and Codebase - https://github.com/Weixy21/RBX-474x595

Software Environment

Reference Books





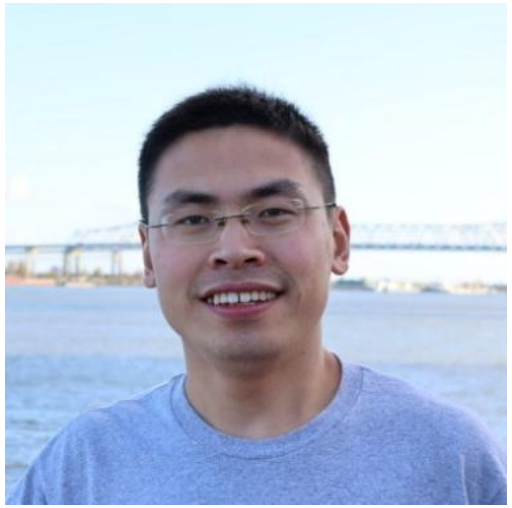


Acknowledgements

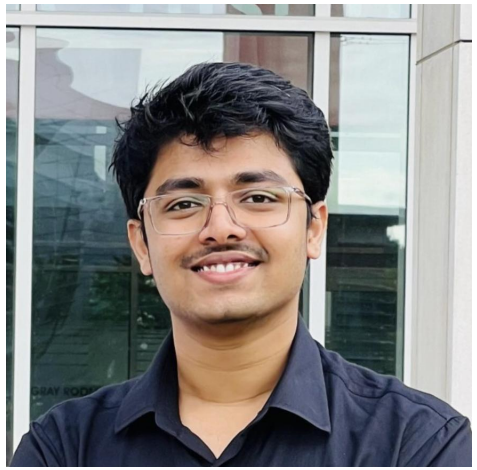
More Resources

- University of Maryland, College Park's CMSC733
- University of Minnesota, Twin Cities' CSCI5563
- Stanford University's CS231n
- University of Maryland, College Park’s CMSC764
- WPI, RBE 474X/595 2024

Office Hours



Prof. Wei Xiao
he/him/his
Unity Hall 287
Office Hours: For Serious Concerns ONLY



Shreyas Devdatta Khobragade
he/him/his
TBD
Office Hours: **TBD**

RBE474X vs RBE595-B01-ST

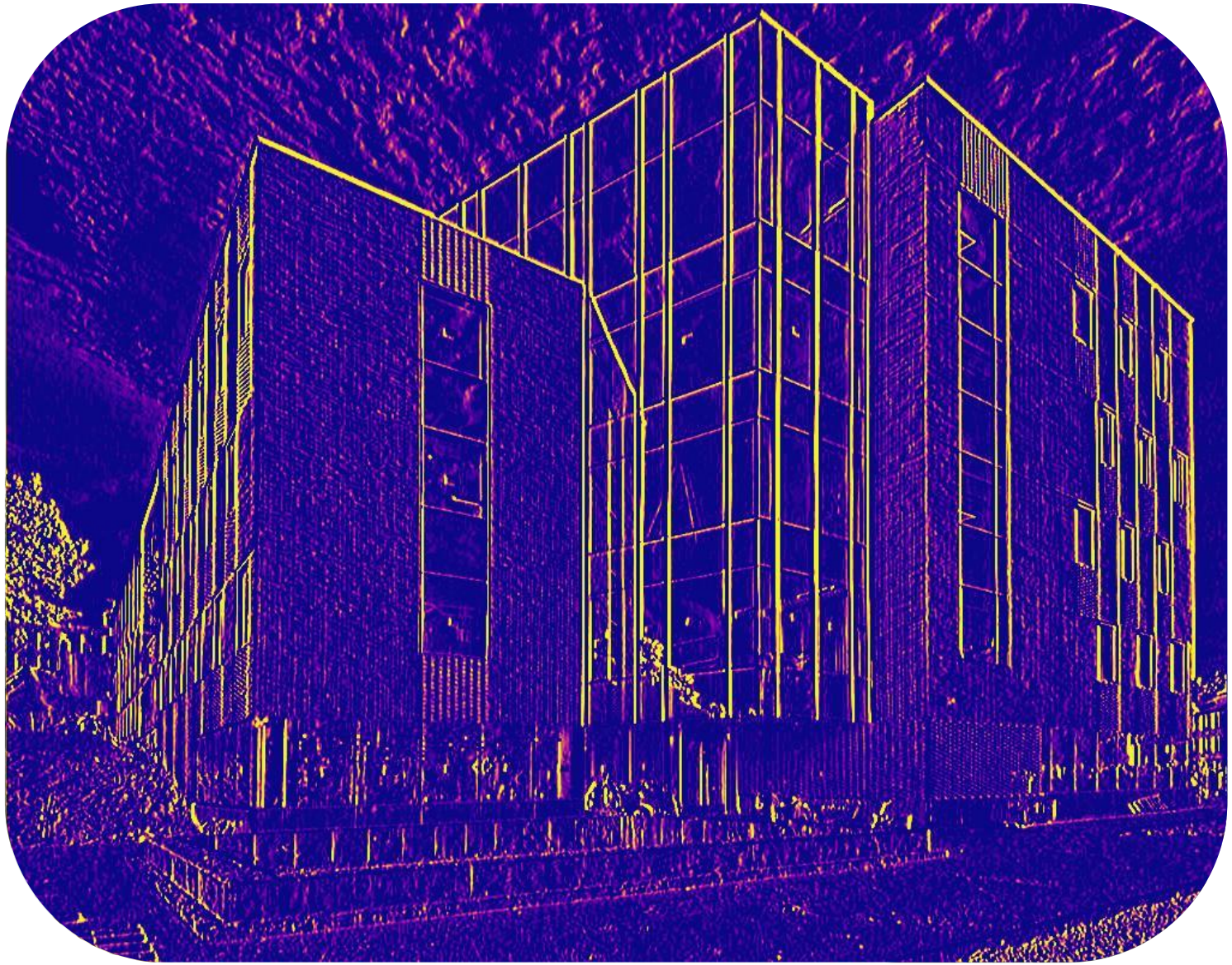
Undergraduate flavor	Graduate flavor
High Expectations	Higher Expectations
Graduate parts of the assignment are optional and extra credit	Graduate parts of the assignment are **REQUIRED**
Gives a good coverage of the basics	Gives a good coverage of the basics + a little extra

HW0: Alohomora

Due on: Oct 31, 2025 at 11:59:59 PM
(Individual Submissions)

Part 1

Convolution

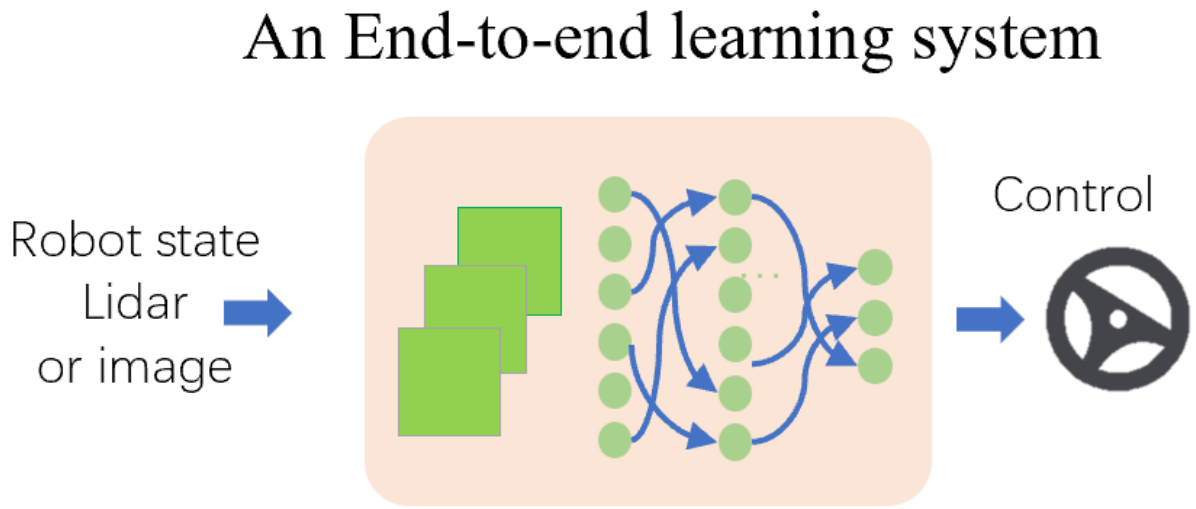


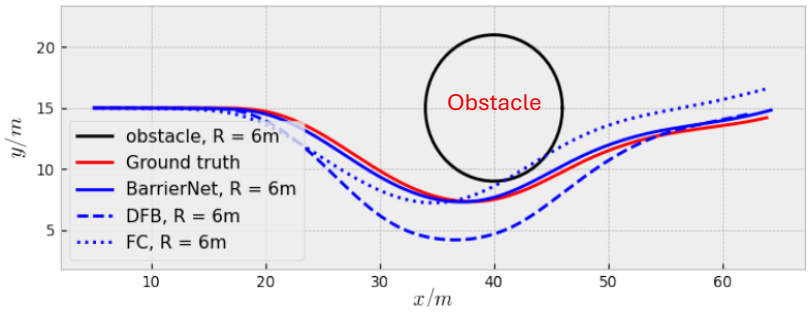
Part 2

Classification based on colors

Course Project

Learning for End-to-End Robot Policies: From Perception to Control (**Group of 2 or 3**)





Robot trajectories under different models

Duration:

From now to 12/12/2025

Course Project

Learning for End-to-End Robot Policies: From Perception to Control **(Group of 2 or 3)**

1. Learning from robot state to control (option 1, **recommended for undergraduates**)
2. Learning from Lidar to control (option 2, **recommended for graduates**)
3. Learning from image to control (option 3, **advanced project**)

Course Project

Learning for End-to-End Robot Policies: From Perception to Control (**Group of 2 or 3**)

1. Complete all tasks (except optional safety filters): 30
2. Project reports with all necessary contents from all tasks: 30
3. Project presentations: 40
4. Add safety filters: 50 (bonus)

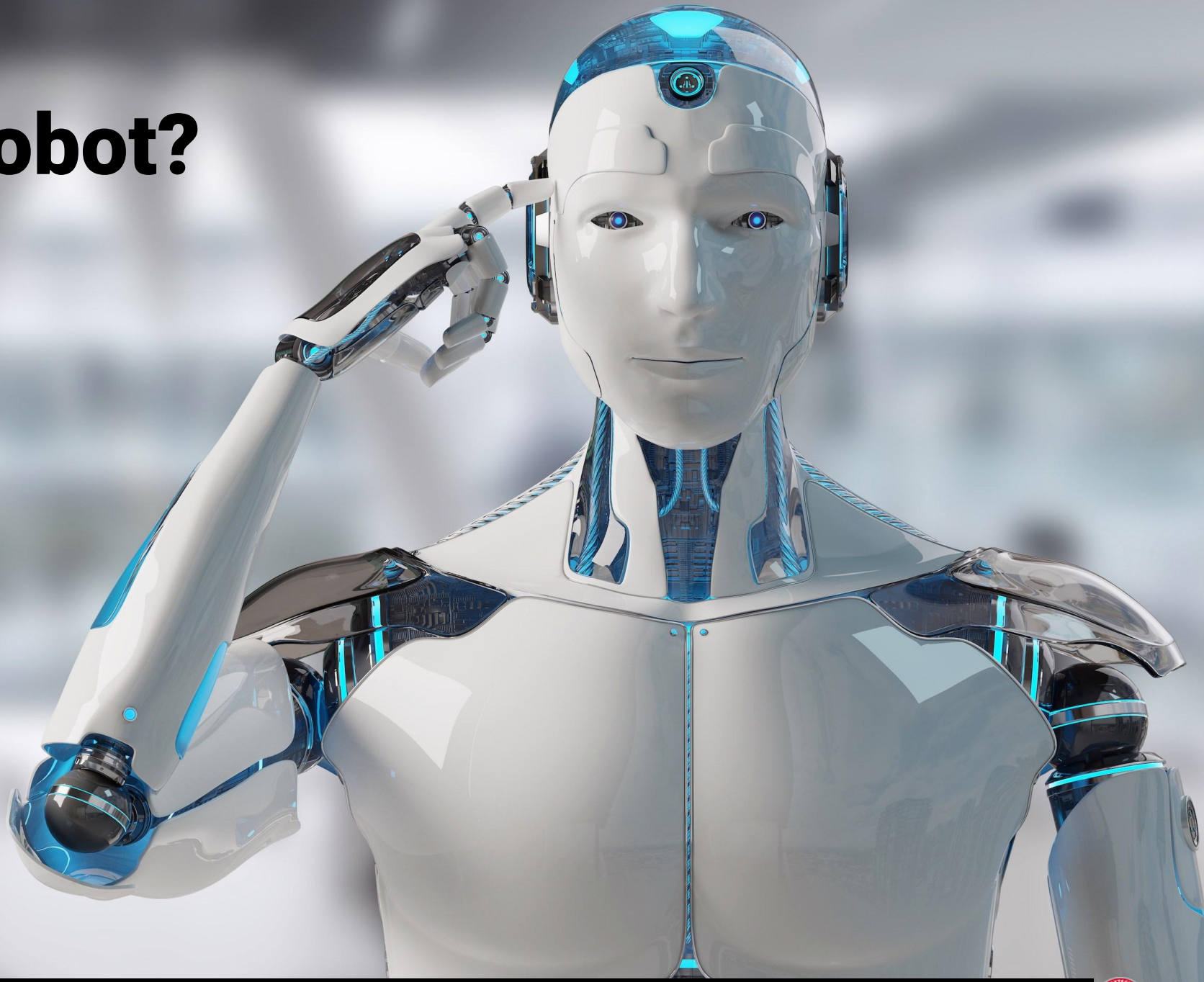
Total: **full credits (100/100), credits with bonus (150/100)**

Note: You may complete both task options 1 and 2, and there will also be a bonus based on the performance (determined by the lecture and TA).

Course Schedule

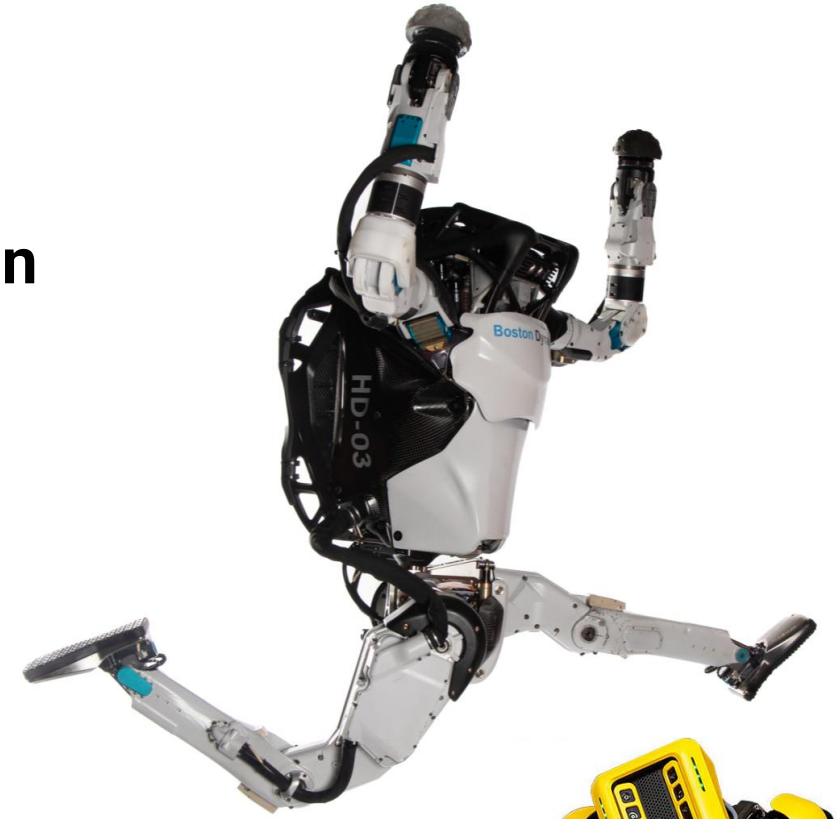
10.21 Class 1: Introduction, Logistics And Sensors
10.24 Class 2: Multi Layer Perceptron And Backpropagation
10.28 Class 3: NN Tuning, Image Filtering And Convolutional Neural Networks
10.31 Class 4: Advanced CNN Architectures And Image Warping
11.4 No class
11.7 Class 5: Simulation for Data Generation And Sim2Real
11.11 Class 6: Object Detection And Segmentation
11.14 Class 7: Learned Depth: Monocular + Stereo
11.18 Class 8: Vision Transformers, Can We Trust Neural Networks?
11.21 Class 9: Single Pixel Attacks, Patch Based Attacks
11.25 Class 10: Generative Models: VAEs, GANs, Attacking GANs
11.28 No class
12.2 Class 11: Advanced Generative Models: Diffusion Models
12.5 Class 12: Advanced Generative Models++: Multi-modal Generative Deep Learning
12.9 Presentation 1 (Group-based)
12.12 Presentation 2 + Course Summary (Group-based)

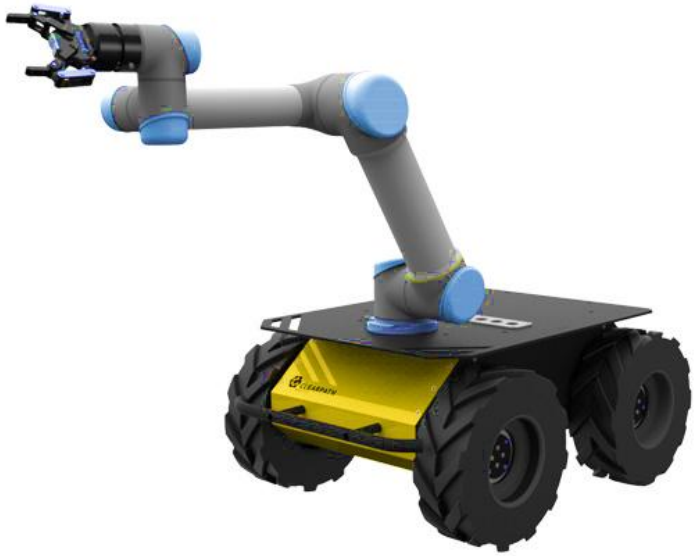
What is a Robot?



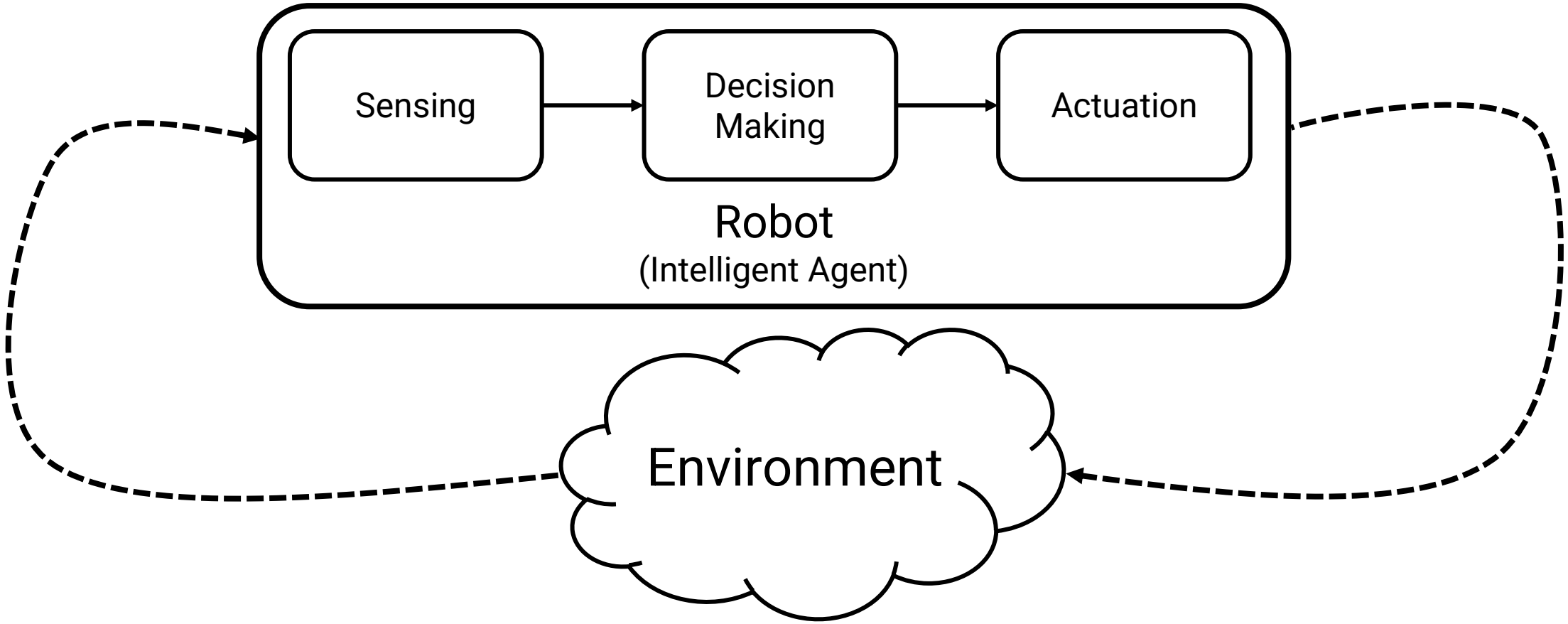
Robot
High-school Version

Robot
(Intelligent Agent)

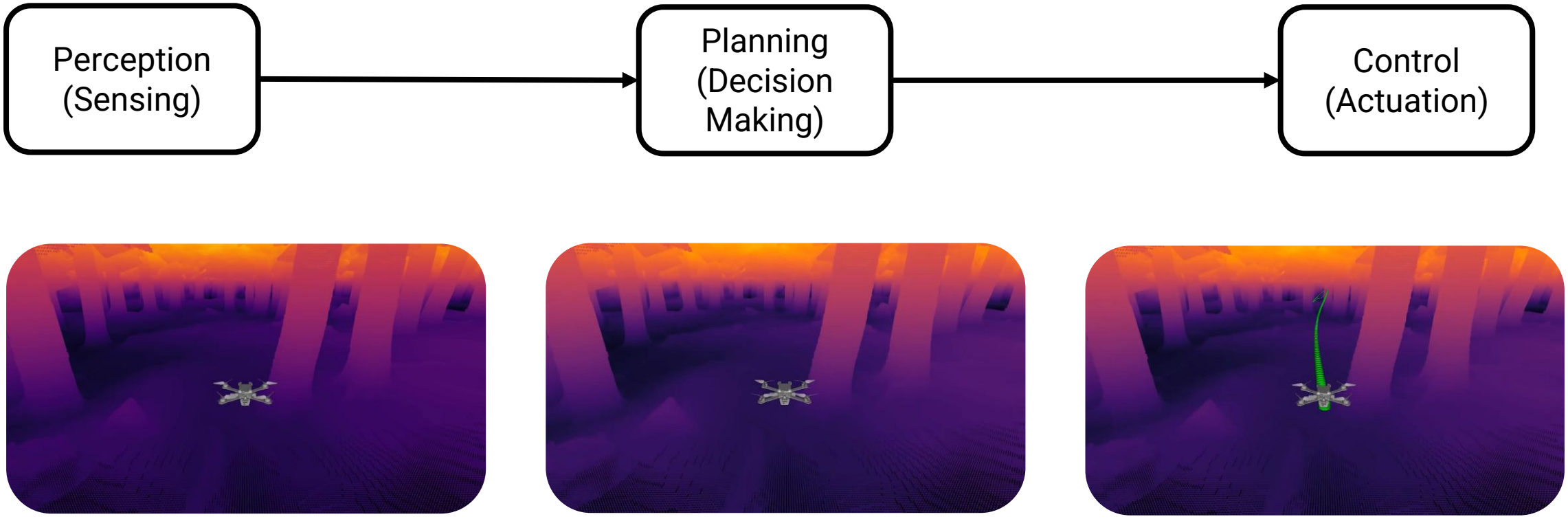




Robot
High-school Version



Robot
Ph.D. Version



Loquercio, Antonio, Elia Kaufmann, René Ranftl, Matthias Müller, Vladlen Koltun, and Davide Scaramuzza. "Learning high-speed flight in the wild." *Science Robotics* 6, no. 59 (2021).

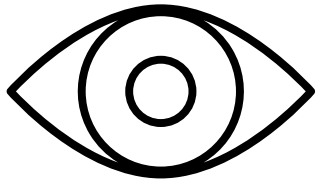
Sensors!

What is a sensor?

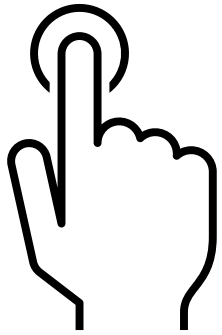
A device which detects or measures a physical property and records, indicates, or otherwise responds to it.

AKA How we perceive the world?

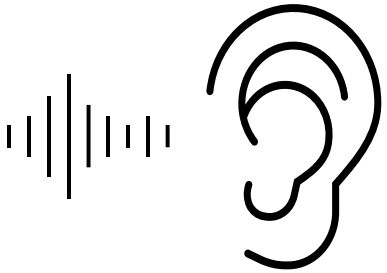
Types of Sensors!



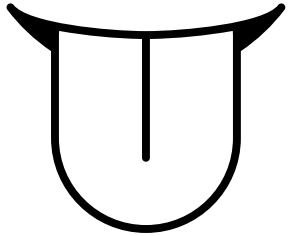
Vision



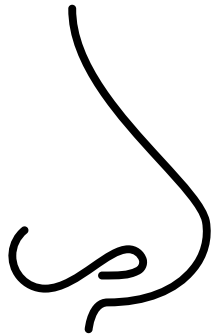
Touch



Sound



Taste

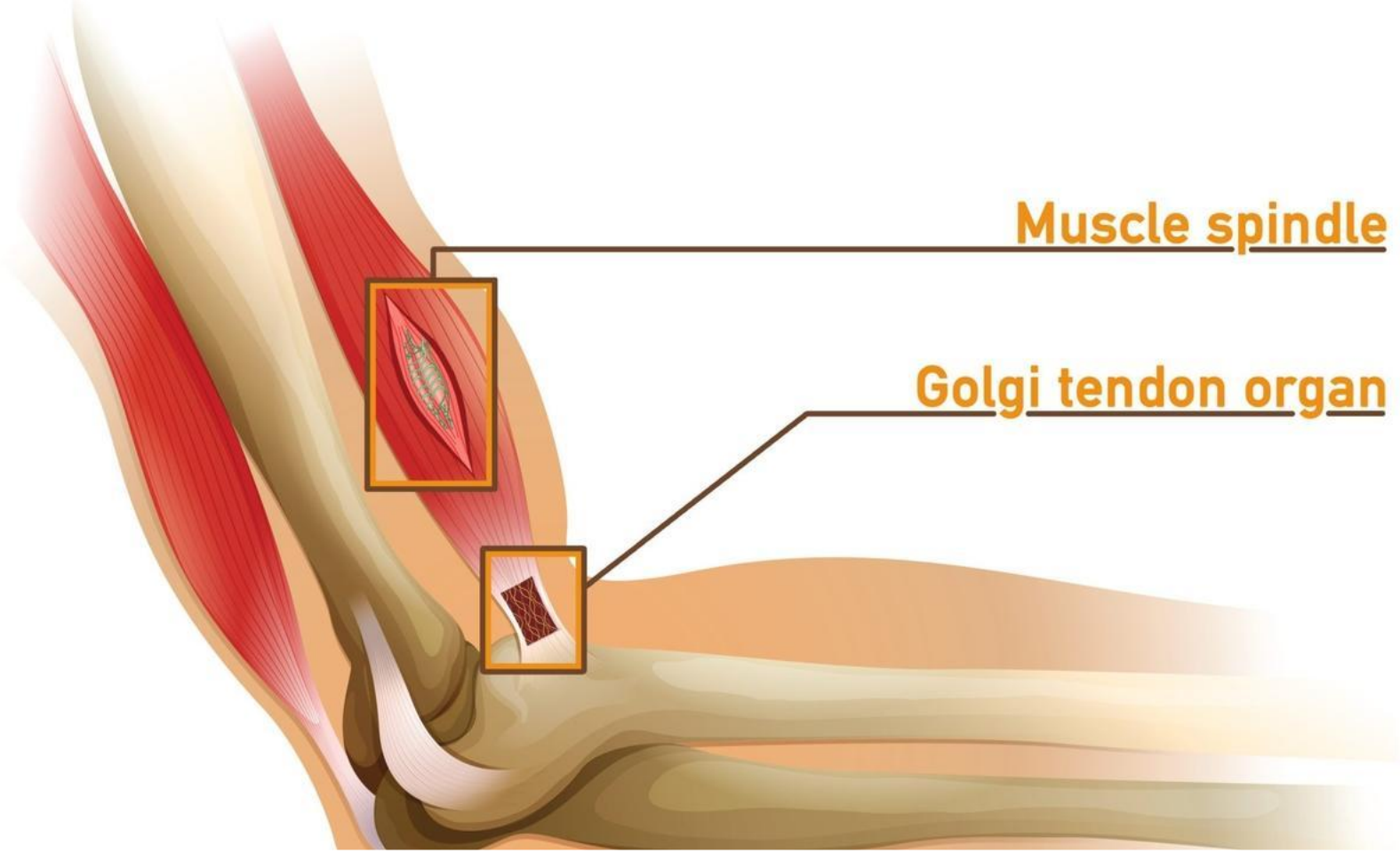


Smell

What kind of sensor enables you to know how much your knee is rotated?

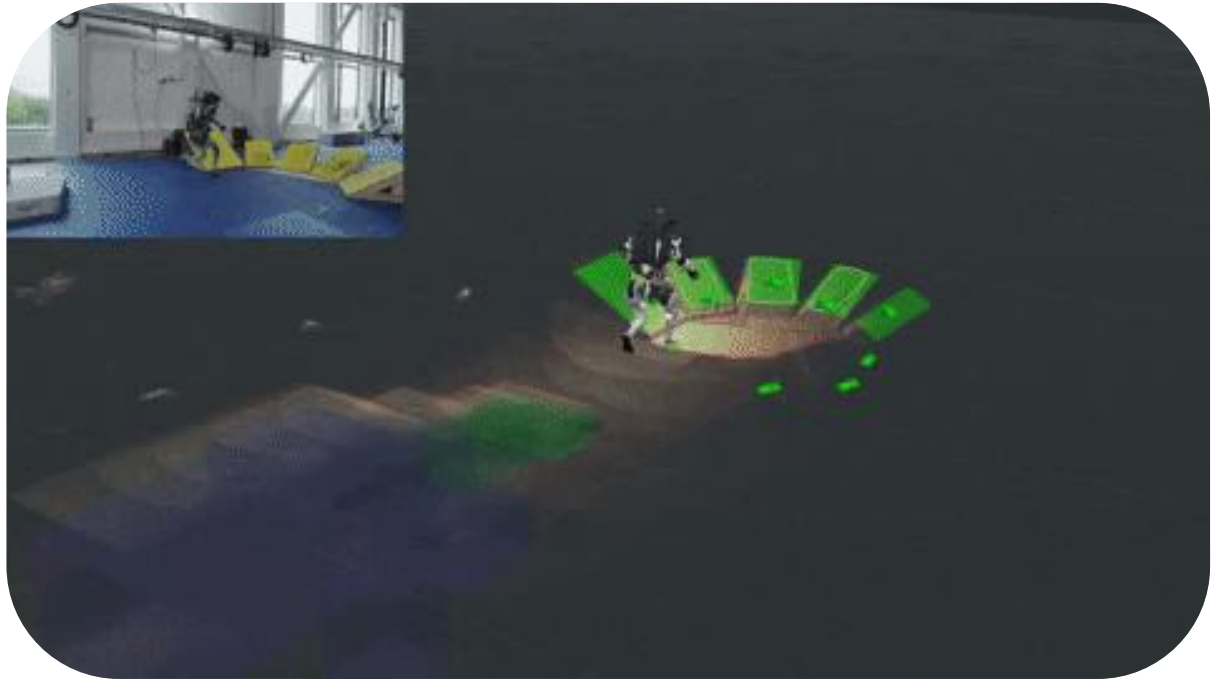
ANOTHER ONE

Internal Perception: Proprioception



Let’s Play A Game!

Boston Dynamics Atlas Humanoid Robot



What sensors does the Atlas have?
What does the perception/vision stack have?
What does the planning stack have?
What does the control stack have?
What does the AI stack have?

Let’s Play A Game!

PRG Husky

The Husky doesn’t perceive depth!
How do you think it’s finding the gap?

Let’s Play A Game!

Boston Dynamics Spot Mini

What considerations should the Spot have to open the door and go in?

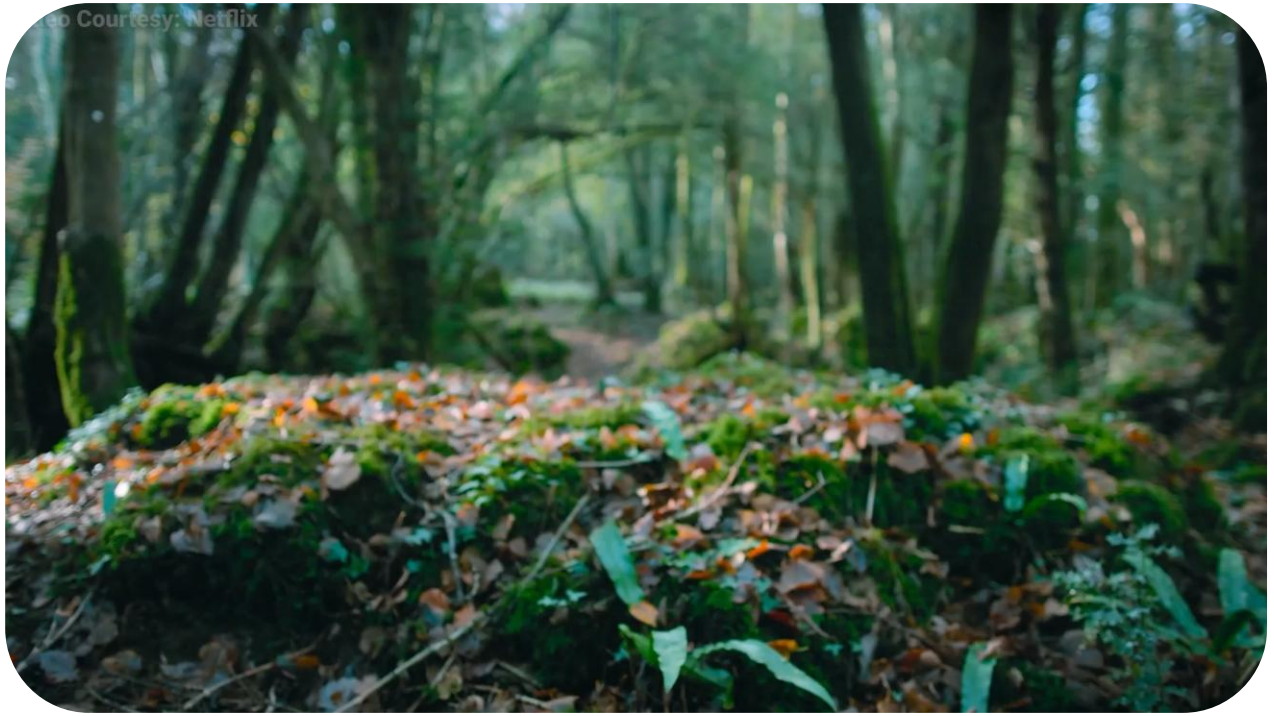
Let’s Play A Game!

Tesla Model S

What design considerations should the car have?
What scenarios should it work in?

Let’s Play A Game!

PRG Labrador 500 α



The labrador doesn’t sense depth!
How do you think it can navigate through various scenarios?

Most Ubiquitous Robotics Sensor!

Camera/Imaging Sensor

- Cheap
- Dense information
- Mimics one of our senses
- Multi-functional

Elon Musk

@elonmusk · **Follow**

Vision became so good that radar actually reduced SNR, so radar was turned off.

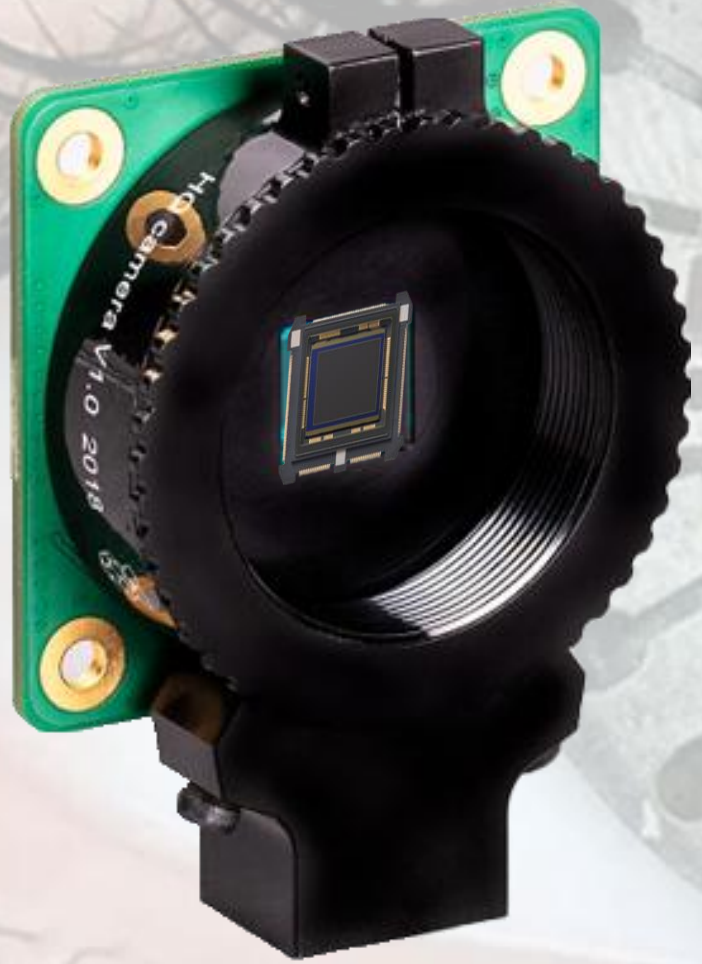
Humans drive with eyes & biological neural nets, so makes sense that cameras & silicon neural nets are only way to achieve generalized solution to self-driving.

11:44 AM · Oct 11, 2021


RADAR had “seen” through the mountain!

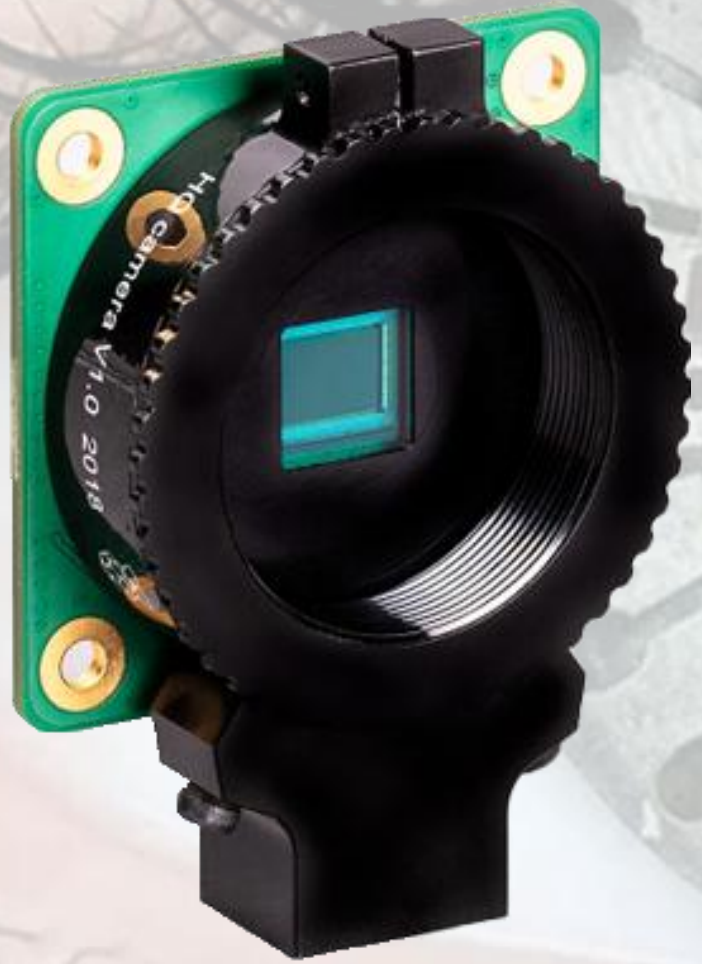
Vision Sensors

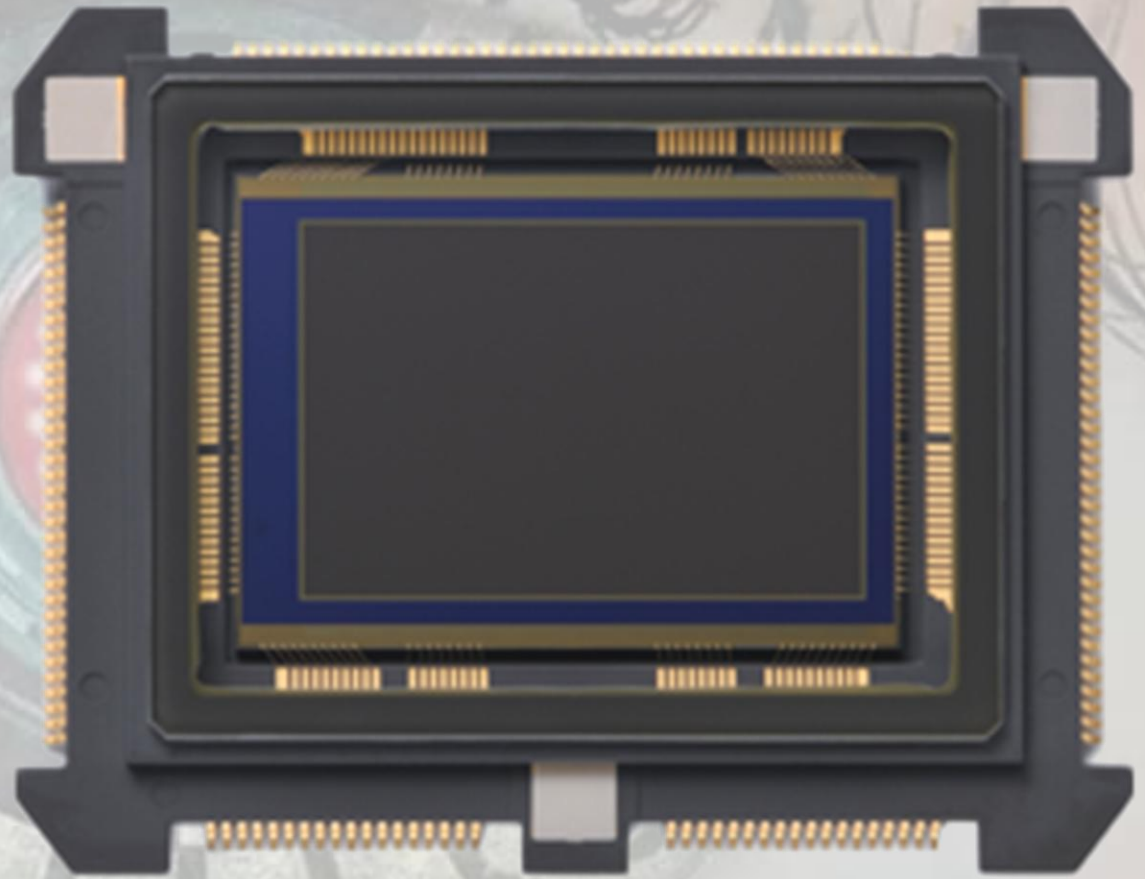
Imaging Sensors



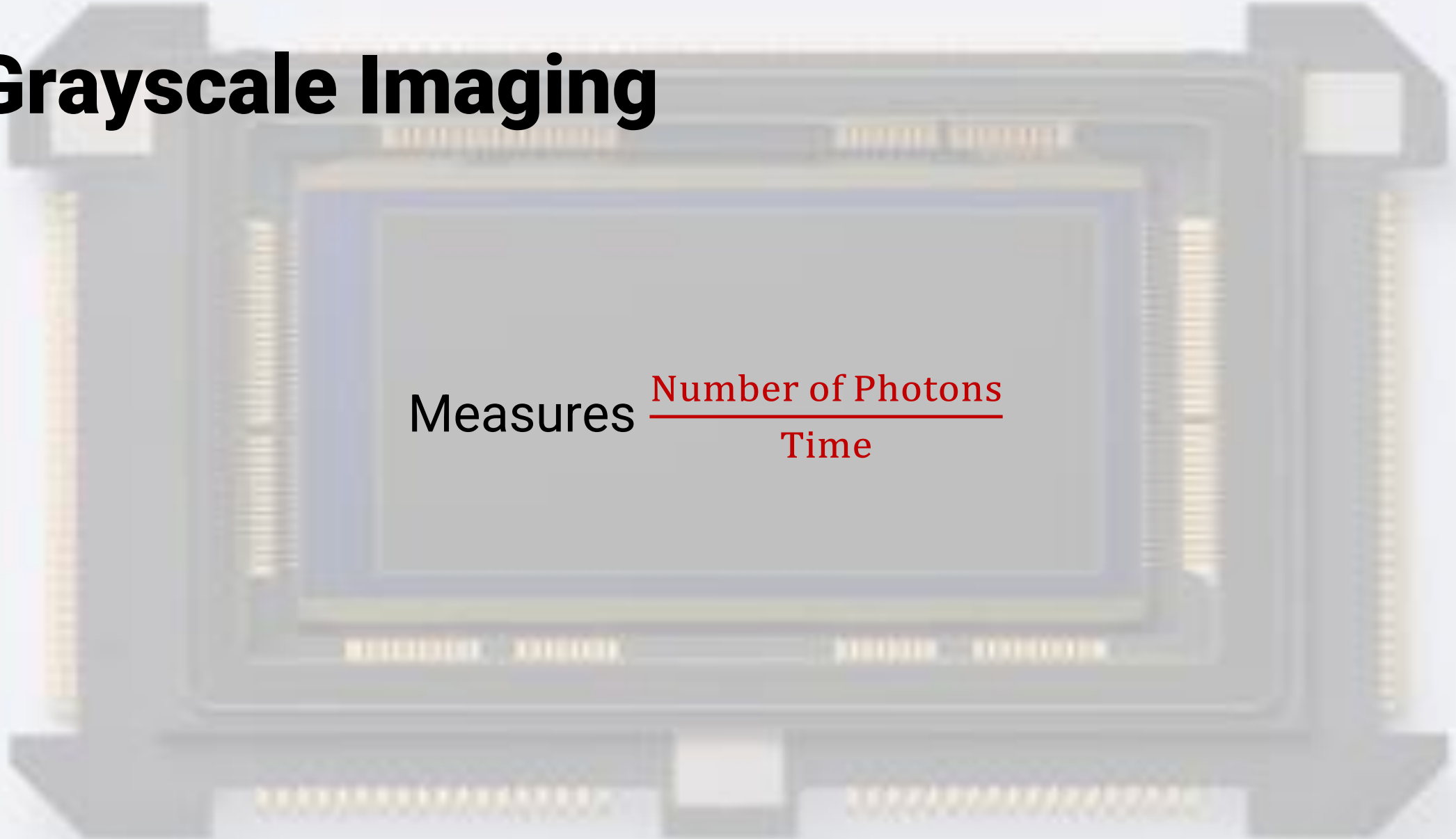
Vision Sensors

Imaging Sensors





Grayscale Imaging

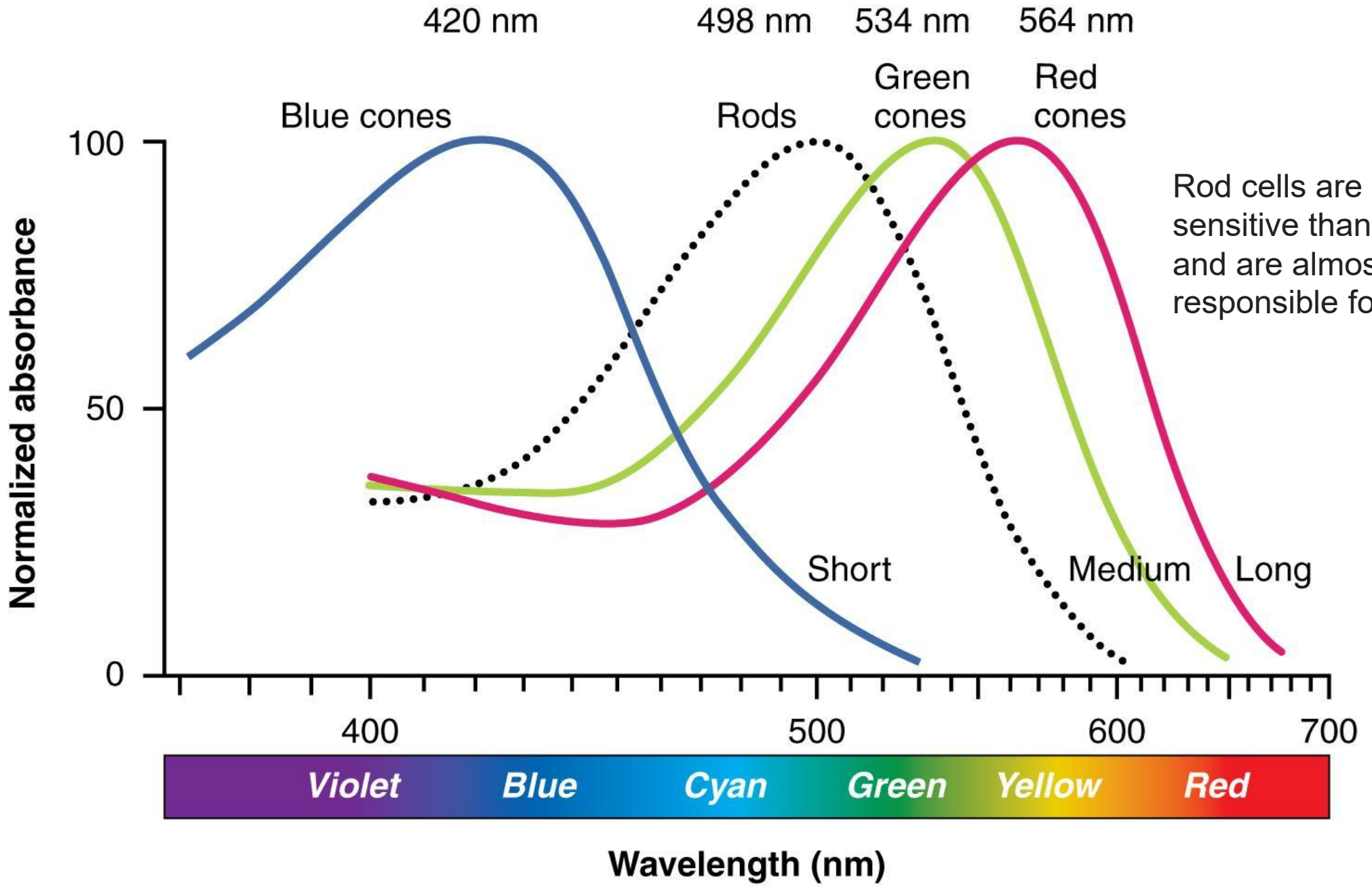


Measures $\frac{\text{Number of Photons}}{\text{Time}}$

RGB Color Imaging

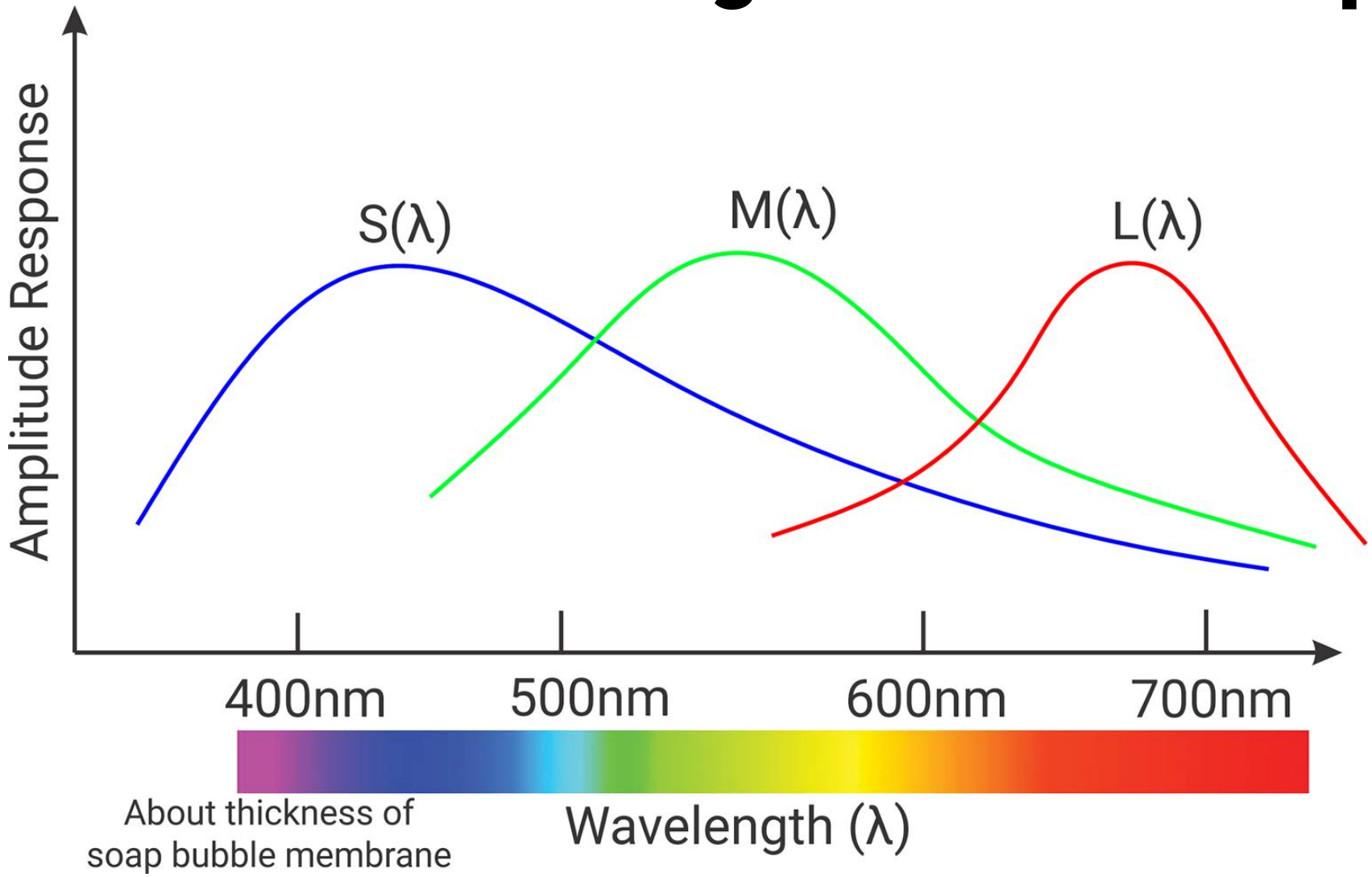
Why **RGB**?


How do we see color?

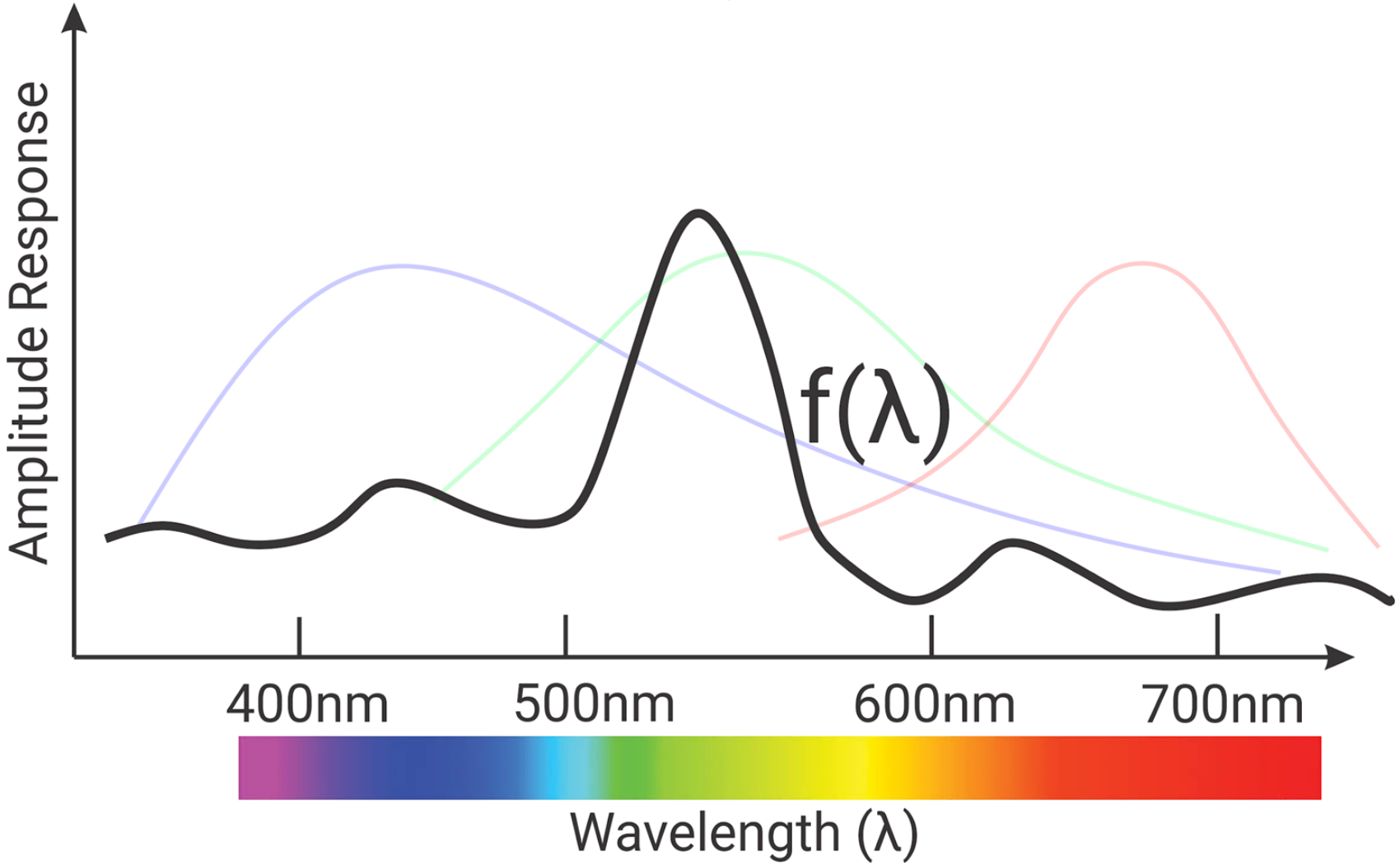


Rod cells are more sensitive than cone cells and are almost entirely responsible for night vision

Mathematical Modelling of Color Perception



Mathematical Modelling of Color Perception



Mathematical Modelling of Color Perception

$$S_{res} = \int S(\lambda) f(\lambda) d\lambda$$

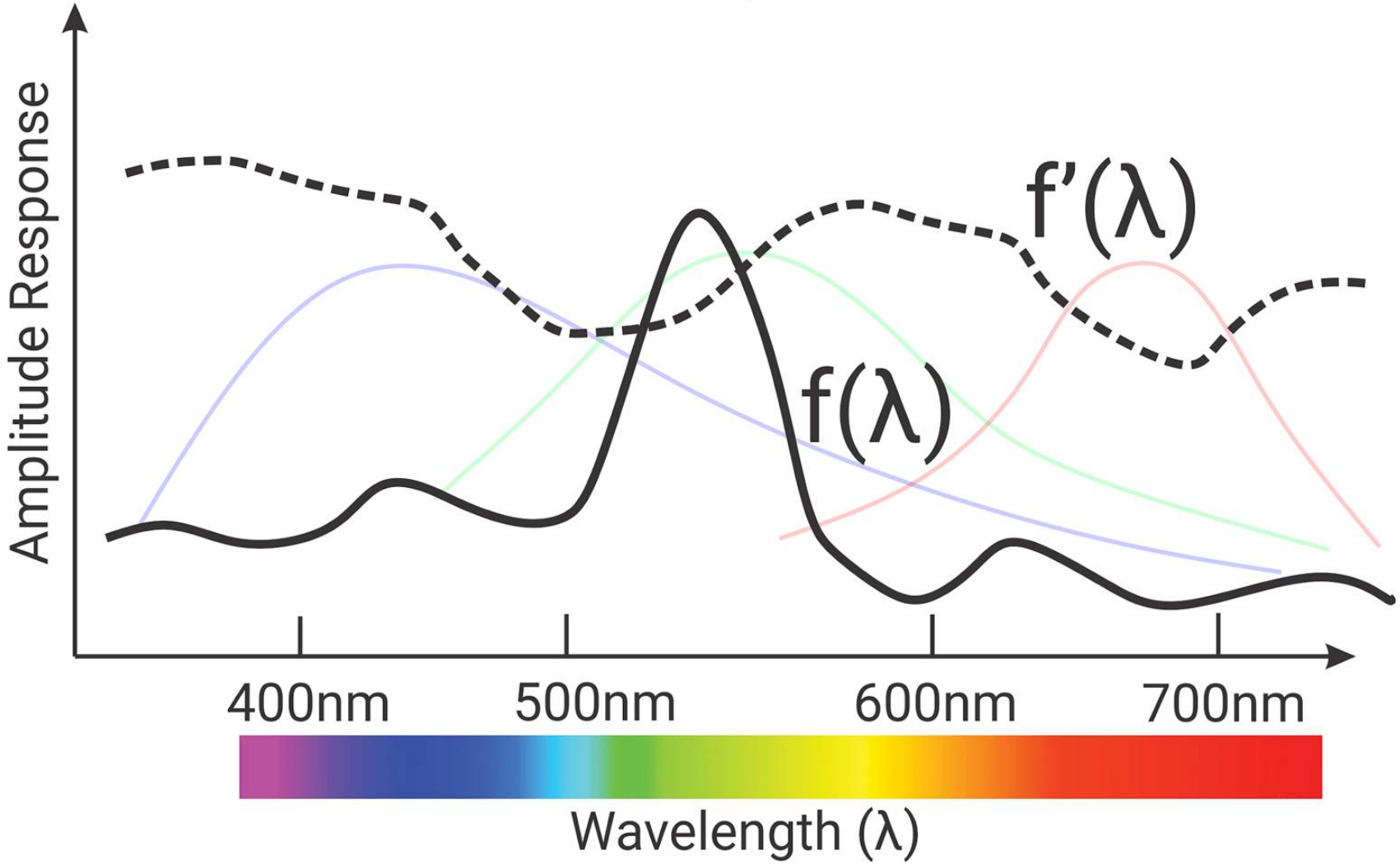
$$M_{res} = \int M(\lambda) f(\lambda) d\lambda$$

$$L_{res} = \int L(\lambda) f(\lambda) d\lambda$$

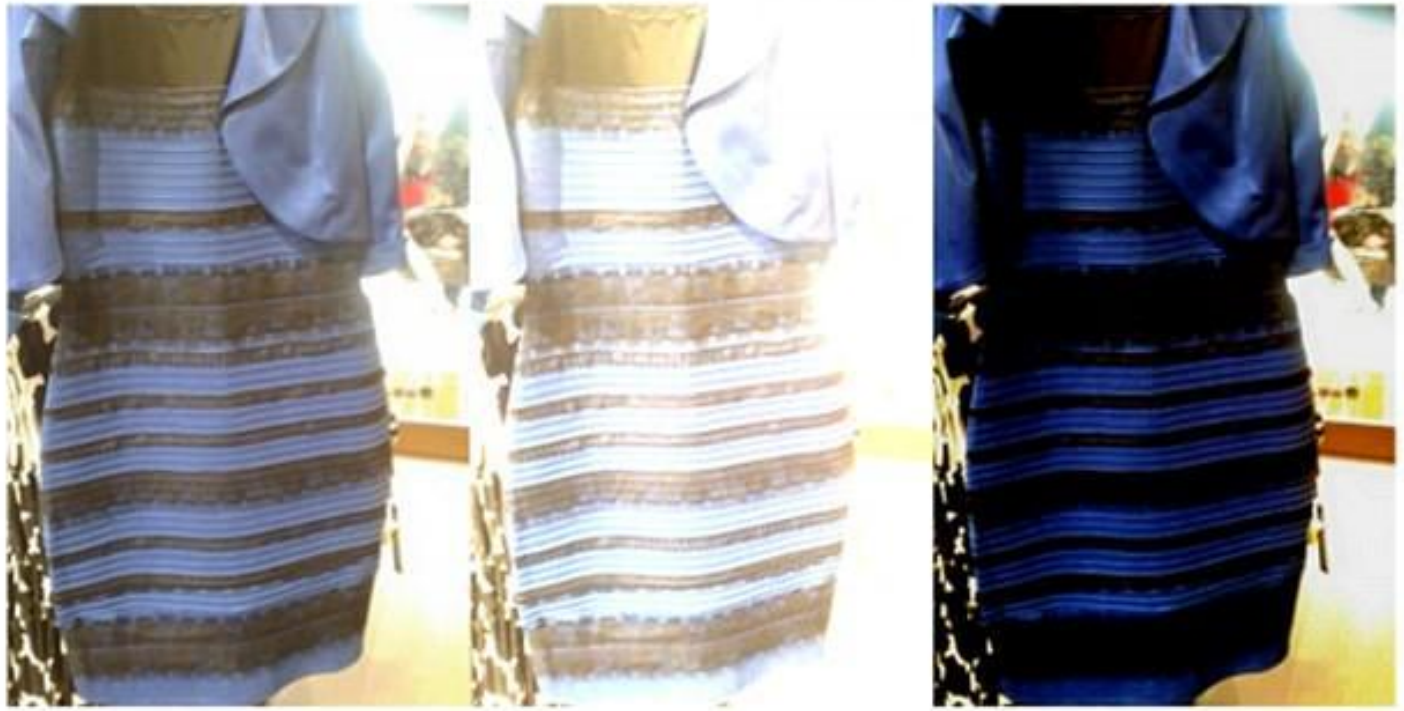
$$\mathbb{R}^{\infty} \rightarrow \mathbb{R}^{3}$$

how a color will be perceived under specific viewing conditions

Mathematical Modelling of Color Perception



Color Illusion



ORIGINAL
(Blue and Gold)
+0% brightness, +0% contrast

BRIGHTER
(White and Gold)
+40% brightness, +40% contrast

DARKER
(Blue and Black)
-30% brightness, +40% contrast

Next Class!

Multi Layer Perceptrons And Backpropagation