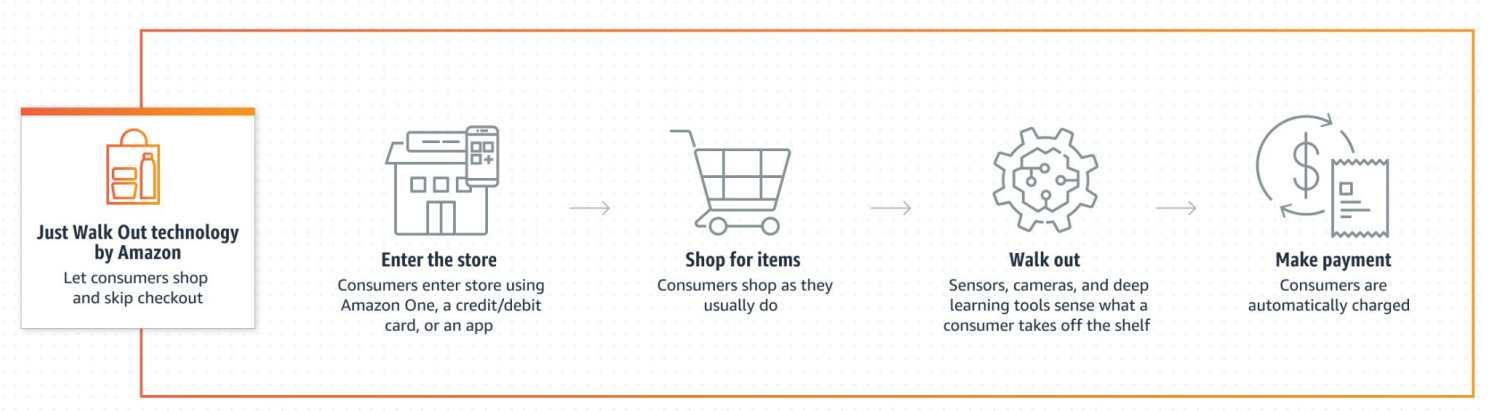
Computer Vision for Pattern Recognition

Presented by:

Duy-Anh Dang
Manjesh Prasad
Nicolas Yuan
Shuzhu Chen

Just Walk Out Technology

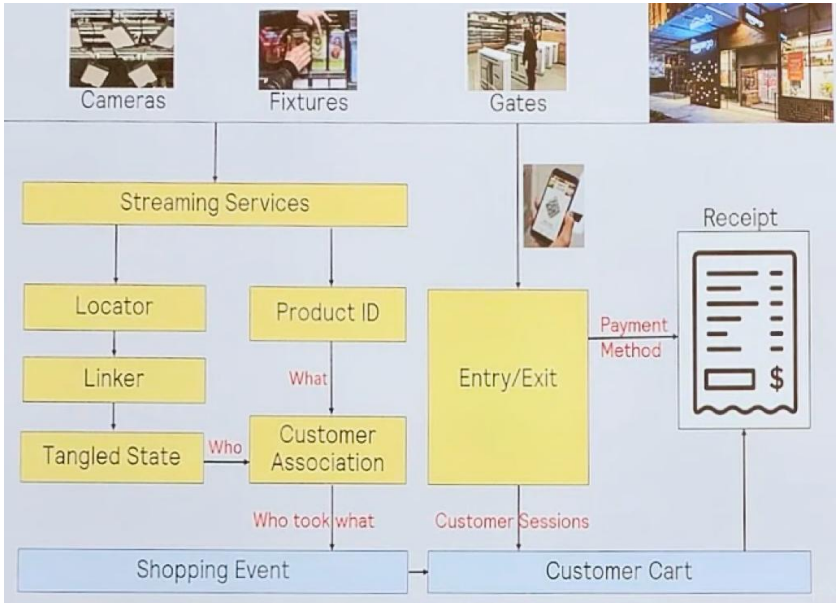
Just Walk Out technology simplifies the shopping experience by removing checkout and helping consumers to get in and out quickly and seamlessly.



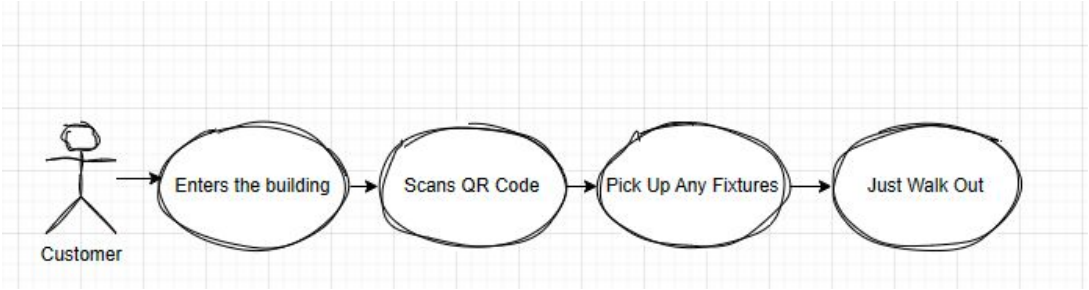
Introduction

Short Recap:

Last time we simply discuss about the system architecture, this presentation, we are going to talk about the functionality requirements for Just Walk Out System.

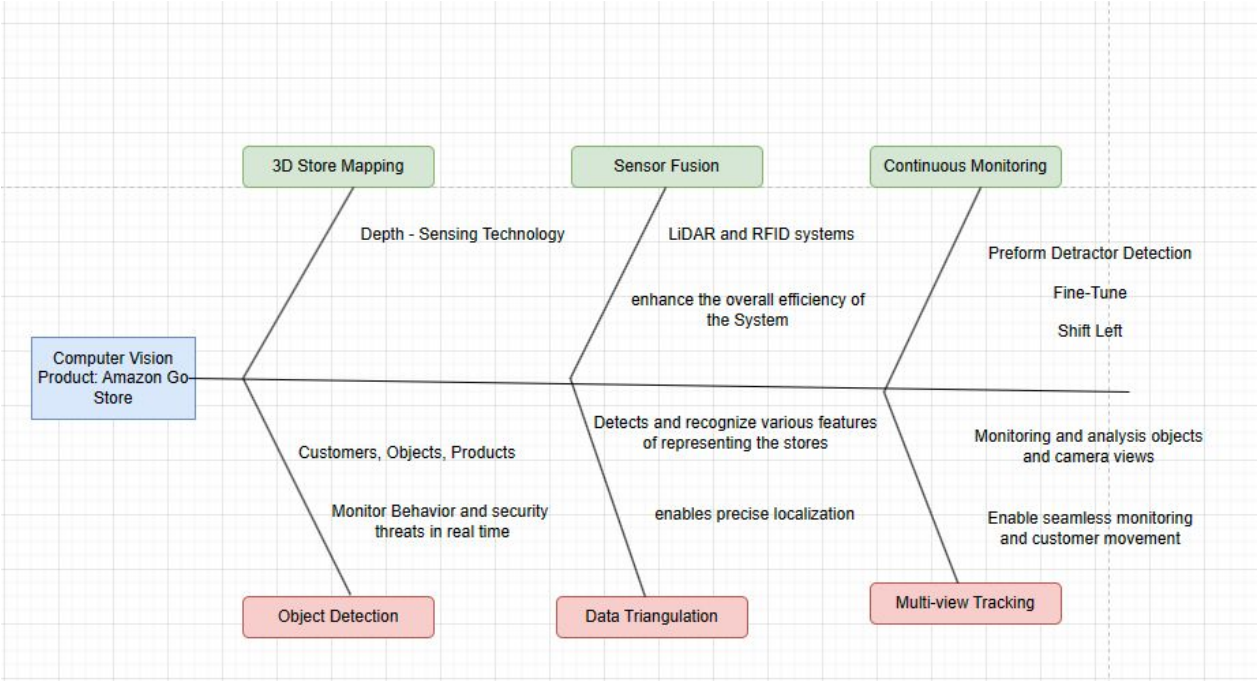


System Requirements - Business Perspective



- **User Requirement #1:** "As an Amazon Prime member, I want to enter an 'Amazon Go Store,' scan my QR code, pick up any items, and simply walk out without any human interactions."
- **User Requirement #2:** "As a Customer, I don't want to continuously feel like I am being watched when I shop by another human,and want every transaction to be accurate."
- **User Requirement #3:** "Surveillance cameras within the retail store must meet the following specifications:
  - High-definition resolution to ensure clear and detailed images and video footage.
  - Proper calibration to maintain accurate color representation and image quality.
  - Integration with knowledge within store model, fusion sensors, and object detection capabilities to enhance surveillance and security monitoring."

Requirement #3 in a CV System



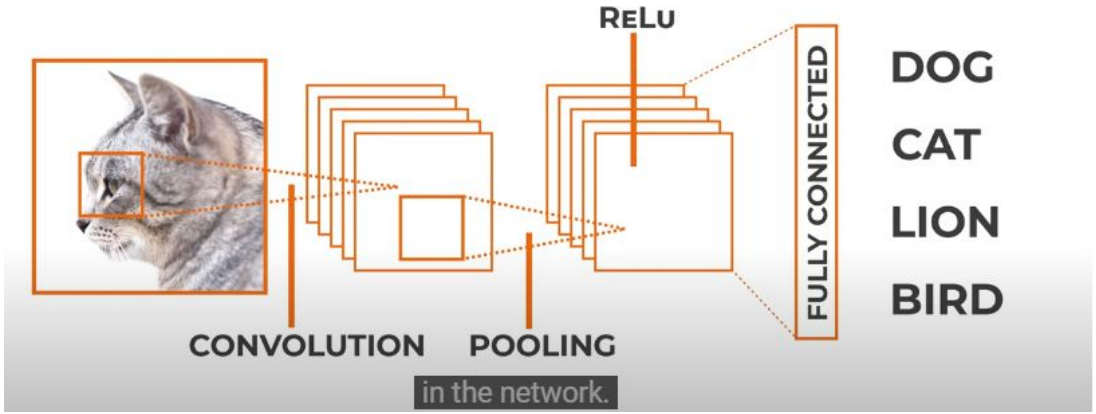
Convolutional Neural Network (CNN)

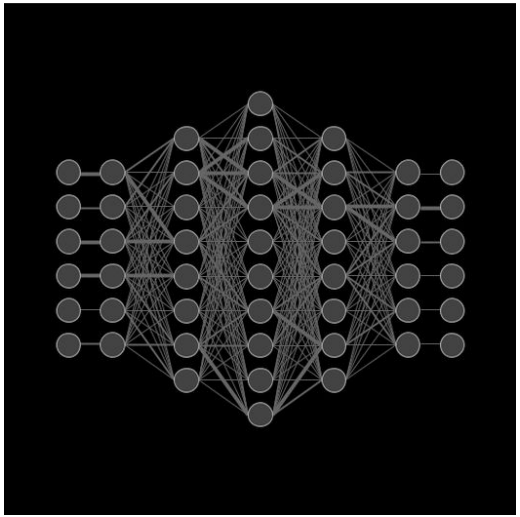
Convolutional Layer: Simplify complex images and objects through a filtering process within the network.

Polling Layer: Reduce the sample size of a particular feature map. Aiding in Feature abstractions.

Rectified Linear Unit Layer (RELU): Serves as an activation function, ensuring non-linearity in the processing of data.

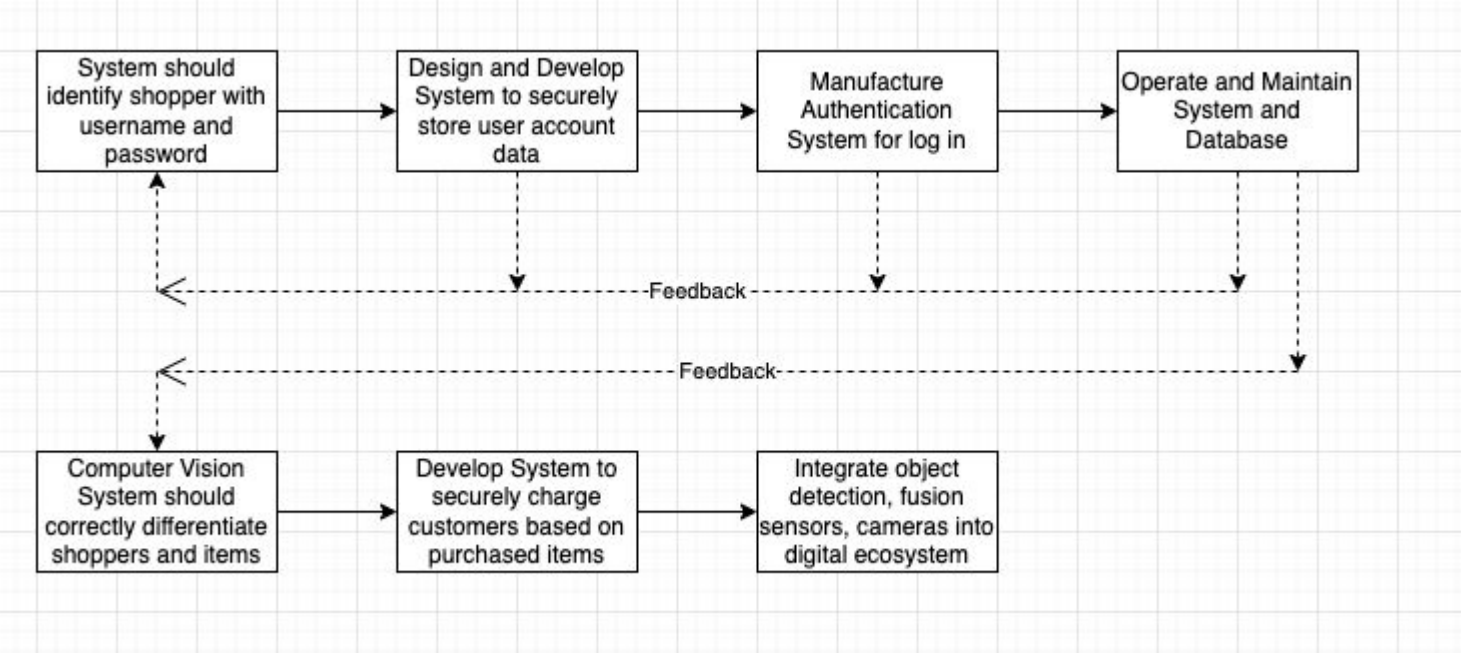
Fully Connected Layer: classification of the image by connecting every neuron to every neuron in the preceding layer.





Functional Analysis

System Requirement: Store's digital ecosystem should identify shoppers with their respective accounts and be able to differentiate items



Trade Off in Computer Vision System

Alternative 1:

Depth Cameras and Load Sensors

Alternative 2:

RGB Cameras with Palm Scanning

Pros and Cons of the Alternatives

Depth Cameras and Load Sensors

- Pros:
  - Privacy: do not capture detailed shopper images
  - Cost-effective: less expensive than advanced RGB cameras
- Cons:
  - Accuracy: hard to differentiate between similar items or handle situations with multiple shoppers in close proximity.
  - Limited functionality: do not provide additional data beyond the items they take, unable to provide targeted promotions or personalized shopping experiences

Pros and Cons of the Alternatives

RGB Cameras with Palm Scanning

- Pros:
  - Higher accuracy: pass the image through CNN, compare feature vectors in database
  - Additional functionality: have unique user identification and enable personalized shopping recommendations
- Cons:
  - Privacy concerns: collect detailed shopper palm scan images
  - Higher cost than traditional depth camera and load sensors

References

- **Performance Characterization in Computer Vision:**
  - https://link.springer.com/chapter/10.1007/978-1-4471-3201-1_1

- **Computer Vision By E.R Davies**
  - https://books.google.com/books?hl=en&lr=&id=mEuZDgAAQBAJ&oi=fnd&pg=PP1&dq=computer+vision+design+process&ots=FxJ8toOq-T&sig=dBSh7SYY11ge9lq2h_QVnKOzhWM#v=onepage&q=computer%20vision%20design%20process&f=false

- **How the Amazon Go Store's AI Works**
  - https://books.google.com/books?hl=en&lr=&id=mEuZDgAAQBAJ&oi=fnd&pg=PP1&dq=computer+vision+design+process&ots=FxJ8toOq-T&sig=dBSh7SYY11ge9lq2h_QVnKOzhWM#v=onepage&q=computer%20vision%20design%20process&f=false