Computer Vision For Pattern Recognition - System Modeling Engineering

Manjesh Prasad, Duy-Anh Dang, Nicolas Yang, and Shuzhu Chen
San José State University, San José, CA

Abstract

Computer Vision (CV) has revolutionized the way machines perceive information through observation. It has rapidly advanced the initial roles cameras had, from capturing, recording, and viewing to mimicking human traits. Similar to how humans collect information through speculation, CV aims to reflect these characteristics. To further extend this point, all these traits are accomplished through Machine Learning (ML) where the cameras are trained to identify patterns within a large big dataset to help understand and make predictions of the world. Although it is a powerful technology that continues to become advanced, it has not yet established itself as a truly reliable source. This study explores the concept of a mode application of System Modeling Language (SysML) for the model-based system engineering (MBSE) of a CV system within a retail atmosphere. By leveraging the four pillars of SysML - Structure, Behavior, Requirements, and Parametrics - we provide a systematic approach that can identify potential bottlenecks and make a powerful implementation of a CV within a retail environment. Overall, this paper will highlight the significance a properly established SysML would have to a MBSE for a CV system, resulting in saving time, money, and staff for the retail industry. With that in mind, let's dig into our approach to making an advanced and reliable CV system.

1 Introduction -Identifying the Role of a CV System - How has CV changed the Industry?

The demand for a sustainable CV system has always been a necessity. For example, areas that require security assurance (airports, schools, and buildings) or inventory management usage can benefit greatly from the innovation this technology can have. Despite a camera system already being installed to monitor specified parameters, alerts and notify another entity, it is without human speculation that this is accomplished. CV aims to eliminate direct human intervention by assisting in a process of continuous monitoring to help capture direct abnormal items and actions. These include, but are not limited to, illegal weapons, wanted suspects, and criminal-like behaviors. By analyzing and tracking these abnormal entities, it can lead to faster actions being taken, eliminating the need to be probed by a human upon first site. For our project, we wanted to inject how a well-established CV system can play a pivotal role in the retail industry.

2 Some examples to get started

2.1 How to create Sections and Subsections

Simply use the section and subsection commands, as in this example document! With Overleaf, all the formatting and numbering is handled automatically according to the template you've chosen. If you're using the Visual Editor, you can also create new section and subsections via the buttons in the editor toolbar.

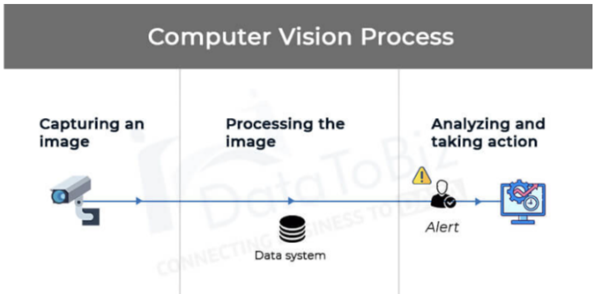


Figure 1: Notice in the image, the cybernetics of this digital economy are circulated around what is being captured, and how it is processed.

Certain alerts and outcomes will be determined based on the process of the image and analysis steps.

- **Real World Applications of a CV system:**
  - **Case #1: Image Recognition:**
    - The CV system will use the Convolutional Neural Network (will discuss later what this means) to empower applications such as facial recognition, medical image analysis for abnormal disease detection, or to help empower autonomous vehicles.
  - **Case #2: Object Detection:**
    - The CV system can help identify and localize objects within images or videos. This is significant for tasks related to security surveillance, Artificial Intelligence (AI), and other content management systems.
  - **Case #3: Natural Language Processing:**
    - Despite CV being primarily utilized for images, it can still be used to scan syntax and apply text summarizing process if given. For example, when cameras scan a person's phone while they are scrolling through their feeds or other forms of content.

2.2 Introduction to Amazon Go Store "Just Walk Out" Policy

The best-known retail company that has attempted to implement a 'Just Walk Out' policy is Amazon Go. This concept offers customers a new shopping experience by removing the need for human interaction during their store visit. Upon written text, the process is simple, an active member is able to enter a building, and scan their prime membership QR code. Once the system properly authenticates the user, the customer is able to walk through the gates, and shop for any fixtures of their choice. The store's interior cameras will monitor the customer's actions, and if the customer picks up an item; the network will automatically process and apply decision making results to determine if the customer is potentially buying the item. The customer can just walk out of the store with the item, and be charged with the item they currently possess.

Despite this being an innovative topic that can help engage in the potential of a CV system; it does not come without a consequence. With the recent headlines and during our studies, we took notice how a credible company, Amazon, failed to properly implement a no-human interaction model within their physical retail. To emphasize this point, Amazon advertised how its ML system is sophisticated, and fully functional, but in reality; it heavily relied on thousands of oversea workers to physically monitor the cameras. This has led to many controversies indicating CV and ML technologies are still in their infant stage. As a team, we analyze various factors that lead to the downfall of the "Just Walk Out" policy; which includes concepts of over-designing over-engineer and releasing a product pre-maturely. Based on our research, we conclude while CV plays a pivotal role in the retail industry, the overall concept needs to be further explored before it is executed. With the downfall of Amazon Go Store, it presents how the understanding of ML techniques is still not understood by a wider demographic, and it will continue to stay that way for many years to come.

2.3 What was the real problem with Amazon Go Store?

The problem for having a no-human intervention atmosphere within a retail industry is relying heavily on the technology performance. For instance, having a QR scanner present, and available before entering the store has to be properly functional. Or by calibrating the cameras to account for all the features present in the store, allowing them to track the actions a customer has performed while in the store is essential. All of these have to meet except performance before being put out in circulation. When a technology is first launched, it usually enters its beta mode for further quality assurance testing. This allows proper monitoring to test out reliability and functionality using participants and being cautious. Certain security measures have to be taken into account which can make the entire system obsolete and unsustainable. To emphasize this point, as a team, we presented a significant amount of trade-offs that can improve the overall quality of the retail industry which we will explain later in this paper.

3 Pillar #1 - Structure

To further understand the overall complexity of the CV system within the retail industry, it is important to understand its overall structure in the SysML and the pillar within the MBSE architecture. As a structure, it sets the precedents of the other 3 pillars (Requirements and Behavioral, and Parametric) on how the system should theoretically function. It is vital to know that the technical systems designed by the industry are increasingly complex and interconnected; that the process where the products are designed, produced, and operated become proportionality complicated as well. To address this concern, the MBSE approach initially starts with a pillar which highlights the important functions to make a successful system.

Figure 2: Block Diagram of the CV integrated store

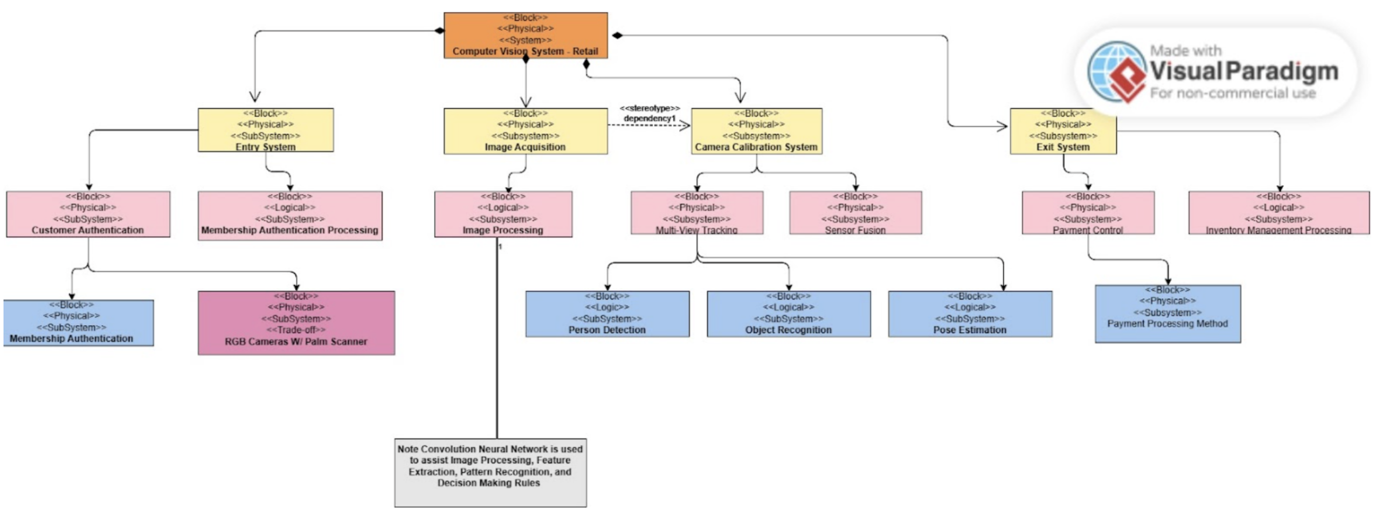


Figure 3: A copy of a Block diagram will be attached

3.1 Data Collection and Feature Extraction Step: Human Attributes vs. CV Hardware vs. Amazon Go Store

Before any behavioral analysis is done, we must indicate the early stage, to mimic how humans spectate when looking in cameras. To better understand the structure, we need to present how these attributes are compared. When working with a SysML diagram, it's important to reduce complexity and enhance requirement traceability. This can lead to better design analysis and efficient documentation which will theoretically eliminate any potential bottlenecks that may arise. On the left side of each block are human attributes. The right side of the blocks represents technological hardware that replicates the same functions as the human attributes. External attributes on the right represent equipment needed for the retail store itself.

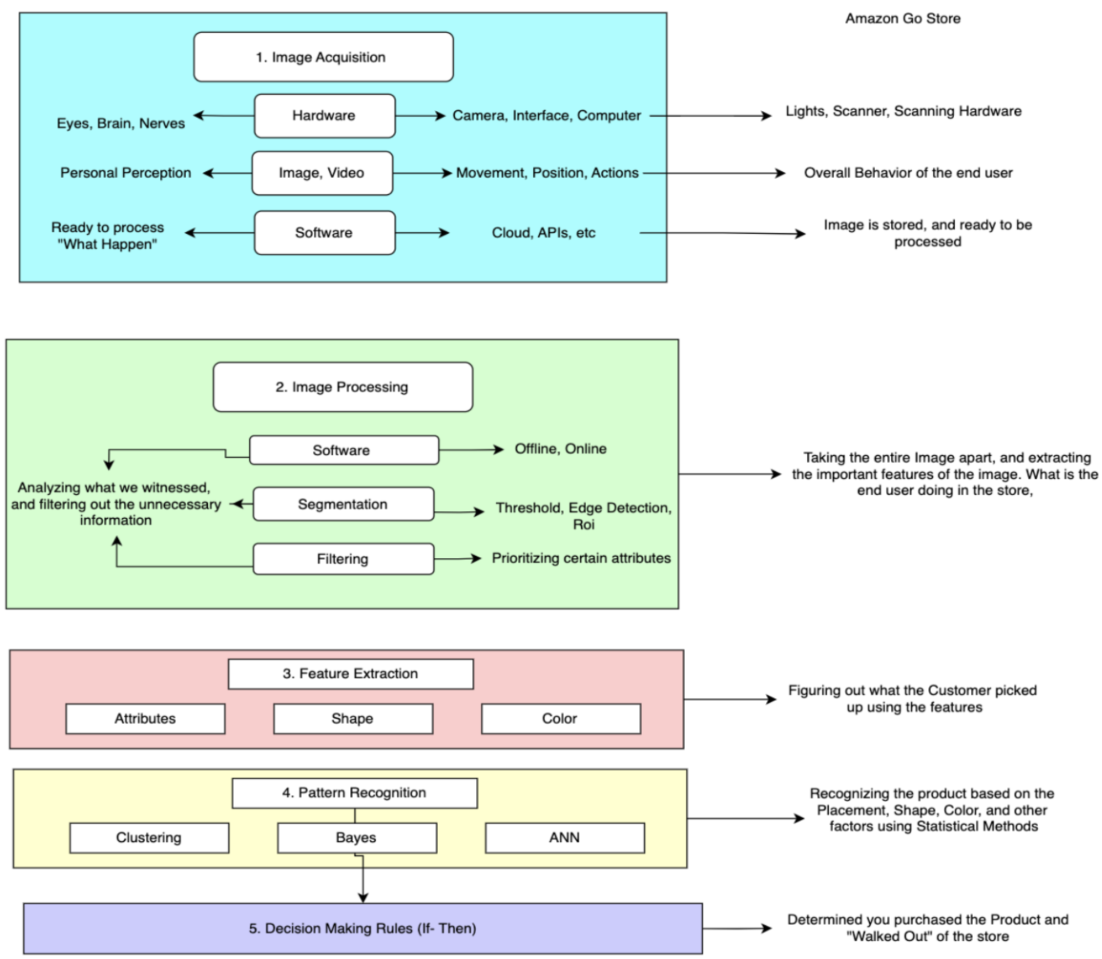
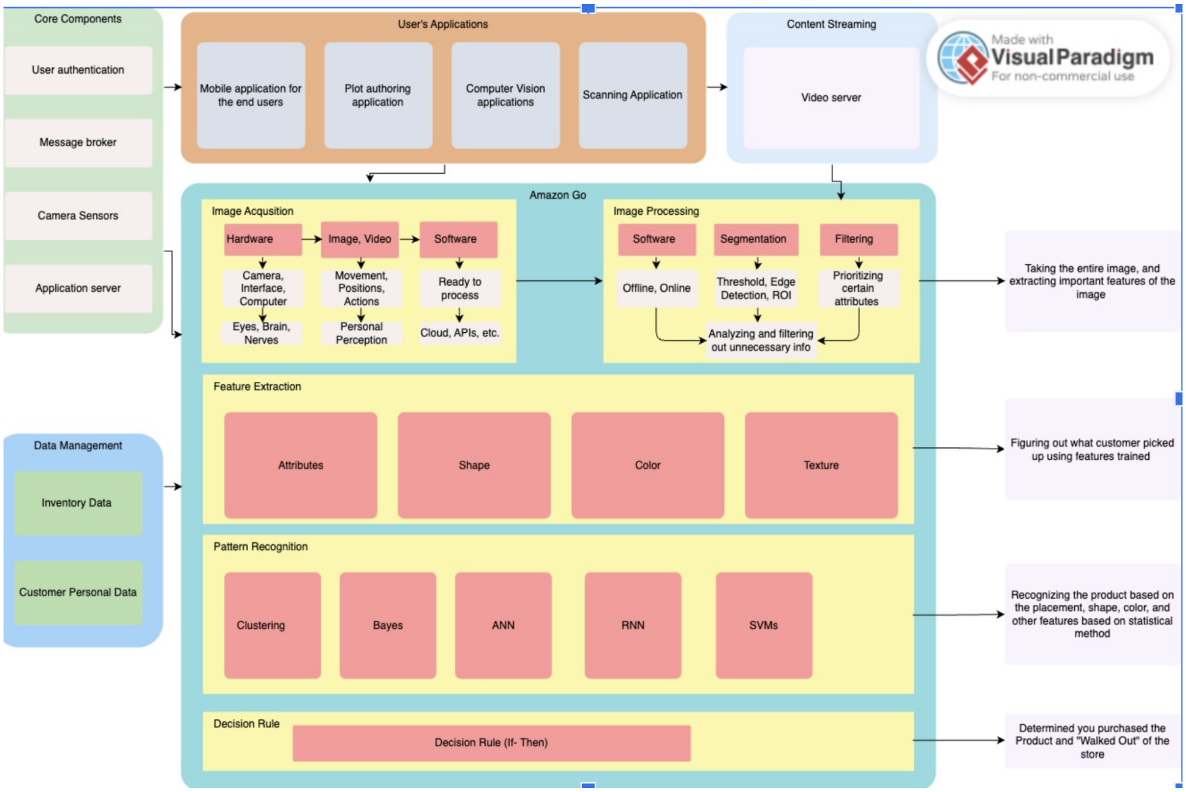


Figure 4: A copy of a Block diagram will be attached

3.2 Image Processing

This image processing diagram showcases how every step is crucial for ensuring valid decision-making rules. An image must first be obtained and stored, typically in secure cloud storage, before processing can begin. Once stored, the image undergoes processing, segmentation, and filtering to reveal specific features. This includes shapes, colors, or any attributes that simplify the image. Convolutional Neural Networks (CNNs) are a traditional model used in the processing step. They analyze the image to identify important features. Finally, trained patterns and layers perform pattern recognition, leading to the decision-making process. The goal of this image processing system is to mimic human perception of objects. Just like our senses, the system interprets what it observes based on its perspective. After processing and analysis, a conclusion is developed. As humans, we must be aware of what we are observing and the potential for inaccuracies. This is why this step is the core of the entire retail digital ecosystem's accuracy.



3.2.1 Pattern Recognition

Notice in the pattern recognition section, we have Algorithms Neural Network, Recurrent Neural Network (RNN), and Support Vector Machines (SVMs); these algorithms can direct any form of classifications that are needed for curing image processing. Providing spatial patterns, all these neural network-based algorithms can help decipher and interpret what is 'happening.' Imagine having a recording of someone purchasing something. We apply weights on the features and identify unique patterns to better understand the fixtures, textures, and properties. Our task is to implement these traits for the machines.

4 Pillar #2 - Requirements

To build a successful CV product, we must determine the requirements with respect to the Block Diagram above. Requirement analysis is the censorious step in implementing any system or products. It indicates the needs and expectations of its end-users, and stakeholders identity, and provides proper clarity that is present within the product itself. Based on the information provided by the stakeholders, we must distinguish the key components, and highlight the core of the entire ecosystem. It is the prerequisite, and the bridge factor that would indicate any constraints, actors, and processes within the entire system itself. Having said that, this process can be done through verbal interviews, surveys, or any other form of communications. It will create an accountable system to eliminate any potential defects, and capture the main intention of the product.

To further extend this point, the requirement analysis is the essential role that would set the precedents to any successful product. The steps are highlighted with respects to the block diagram above:

- Define the problem a CV system has to solve, and the technologies that circulate around it.
  - Understand how certain key features are derived from the origin, and the entire application of the system.

* For example, a sample verbal requirement from a stakeholder would be, "I want the entire system to be secure so active members can shop the entire store without having to deal with staff interaction and can 'Just Walk Out".

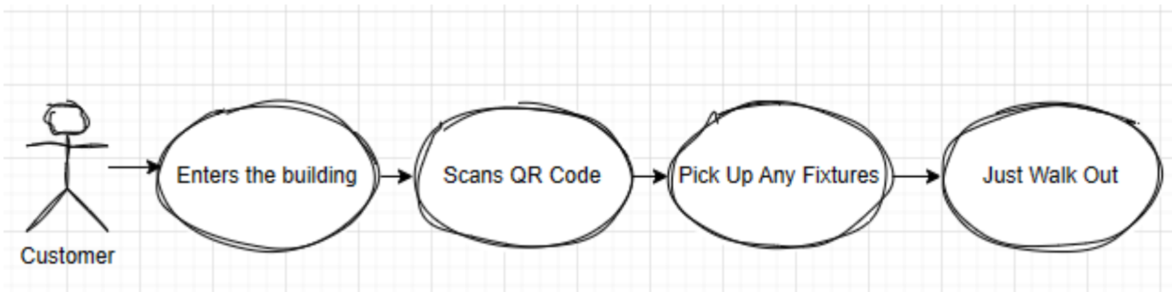


Figure 5: An example of how a stakeholder diagram will potentially look like.

4.1 Requirement Document

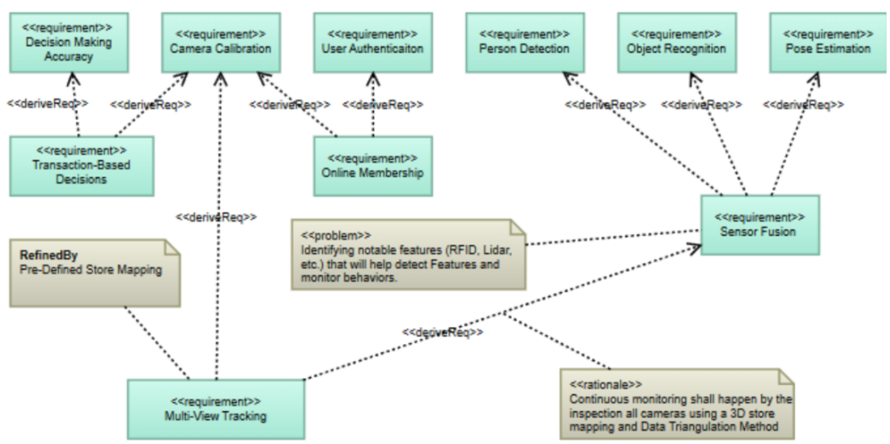


Figure 6: An example of how a stakeholder diagram will potentially look like.

The 6 requirements we have captured during our research are:

1. **Decision-Making Accuracy:** As the name suggests, the entire retail space must construct a decision-making process that will provide validity given any situation. As mentioned earlier, Accuracy is essential for understanding end-user actions and behaviors. Every transaction relies on this principle, requiring the system to perfectly interpret what is happening within it. These actions can include anything from making a purchase to simply walking out of the store.

2. **Camera Calibration:** The most crucial feature where all active cameras must be calibrated and tailored with respect to its environment. This can help in capturing high-quality images that are needed to be processed to make statistical decisions.

3. **User Authentication:** The requirement that facilitates if a member is valid and can enter the gates. This is to ensure proper identifications and can help perform accurate transactions and other forms of payment.

4. **Person Detection:** The cameras must apply ML that is able to detect where a customer is and the behavior they perform in the store.

5. **Object Recognition:** The ML is able to identify all the regions interior and exterior of the building. This includes walls, shelves, fixtures, and other key information (RFID and serial numbers).

6. **Pose Elimination:** Able to point out key body features in humans and videos to understand their position. This can help distinguish between every customer that enters the building.

Notice in the diagram above, Transaction-Based is a derived Requirement from Decision Making Accuracy and Camera Calibration. This is because applying statistical algorithms (Bayes, Clustering, and Artificial Neural Network) is directly correlated with the image that is being captured as well as how well the system is trained. Just like how every transaction is based on the collected information, and every choice being performed is a subjective decision to what is presented. Similar to how Online Member is an imitative requirement of Camera Calibration and User Authentication as its origin. Only active memberships are able to shop within the building. Since every active member has a default contact information, and payment method, this is to eliminate the complication of identification and IAM principles. Once the user enters the store, the ML camera system will automatically link the taken image with the membership the person has scanned into.

Although these are derived from the three origin requirements, the bulk of all these requirements can be traced back to the Camera Calibration process. As presented in the diagram, the Multi-View Tracking requirement arises from camera calibration and sensor fusion, but it can also be traced to Person Detection, Object Recognition, and Pose Estimation. In Multi-View Tracking, the system is refined by incorporating a predefined Store Mapping system, allowing cameras to utilize information from multiple cameras or sensors. This helps capture data from various viewpoints and allows the system to auto scalability when transitioning from a single perspective. Notable applications that use this technology include autonomous vehicles and other security-related surveillance systems. Sensor fusion involves combining data from various sensors, such as LiDAR, radar, and other close capturing television systems. Since each sensor has its strengths and weaknesses, it only makes sense that fusing data from multiple sensors will allow machine learning systems to compensate for limitations and gain a better understanding of Person Detection, Object Recognition, and Pose Estimation.

Despite all of these requirements having a direct relationship to the 6 base requirements; some significant challenges that may arise are potential performance bugs which include limitation to transaction-based decision making results, and high security risk. For the ML product to be successful, continuous thorough audits are needed to ensure the seamless utilization of this powerful technology. We are not satisfying a single group of stakeholders, but building an impactful system that can help revolutionize the retail industry. With that said some of the potential challenges and potential solutions do include:

- **Case Problem #1: QR Code Scanner Identification Theft:**
  - **What:** It's where a user can take other's active membership and use it as their own. Each QR code is equipped with the person's default contact and payment information which can lead to identity theft.
  - **Solution:** We have indicated a potential trade-off where we can use more of a QR code and RGB Palm Scanners for entry. This is to help ensure proper authentication because fingerprints and vein placements are different among every individual.
- **Case Problem #2: Notable Fixtures Attributes:**
  - **What:** The cameras can only process and extract specific characteristics of the items and/or behaviors. While it is easy to determine if a customer picks up or lets go of an item, identification can be a challenge.
  - **Solution:** Rather than identifying Fixtures with various attributes, shapes, colors, and textures tracked by LiDAR and other sensors as a singular solution. We can use more of a uniform tag, RFID to differentiate the products, and have a shopping basket that listens to any unique waves. Since every RFID tag is linked to one single uniform frequency, we could save more time, money, and will provide simplicity to the entire system.

4.2 Camera Calibration Requirements - Convolutional Neural Network

The core of all the camera calibration will be the Convolutional Neural Network (CNN). CNNs are powerful deep learning algorithms that excel at tasks involving images, videos, and even audio. They are inspired by the structure of the human visual cortex, allowing them to process information in a similar way humans do.

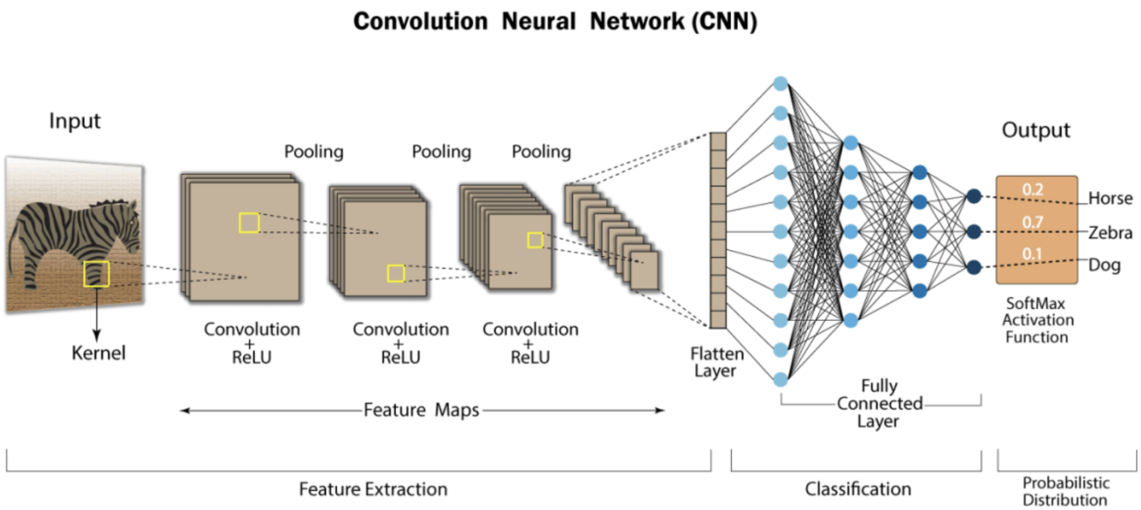


Figure 7: An Example of a CNN

- In CNN, we have 3 layers, each of them has an amazing definitive role:
  - **Convolutional Layer:**
    - Applies simplicity upon complex images and objects through a filtering process within the network itself. They use small filters that slide across the grid of the input image, and help detect specific characteristics such as edges, corners, or shapes.
    - As the filter moves, it generates a "feature map" that highlights these characteristics itself.
  - **Pooling Layer:**
    - Known as the filtering layers. It mitigates the sample size of a particular feature map using a predefined grid and unique weights. This can improve assistance in feature abstractions and help filter out any unnecessary information.
  - **Rectified Linear Unit Layer (ReLU):**
    - Main intention serves as an activation function, ensuring non-linearity in the processing of data. These layers combine the exact features from previous layers to make complex decisions, and help identify what the object is.

When determining the camera calibration, we have to take into account how many customers are within the store itself, how much inventory we are going to have, and in addition to how deep our store is. CNN is not the solution to many potential problems that may arise during the system development stage, but can help wear down many of the dilemmas ML will have when blueprinting the layouts. We will in fact, by default, have many sources collecting information in relation to behaviors and capturing the image, but need to have more vigorous means of processing the image itself.

4.3 How can we incorporate CNN and CV to the retail environment:

In the Amazon Go Store, CV systems are an integral part in security and surveillance to follow suspicious activity in public areas, airports, and other settings, identify faces, and detect intruders by monitoring and analyzing video feeds. Utilizing both software and hardware components, this technology focuses on crime prevention and investigation while also improving security measures. Enhancing productivity, decreasing mistakes, and facilitating improved decision-making based on real-time visual data.

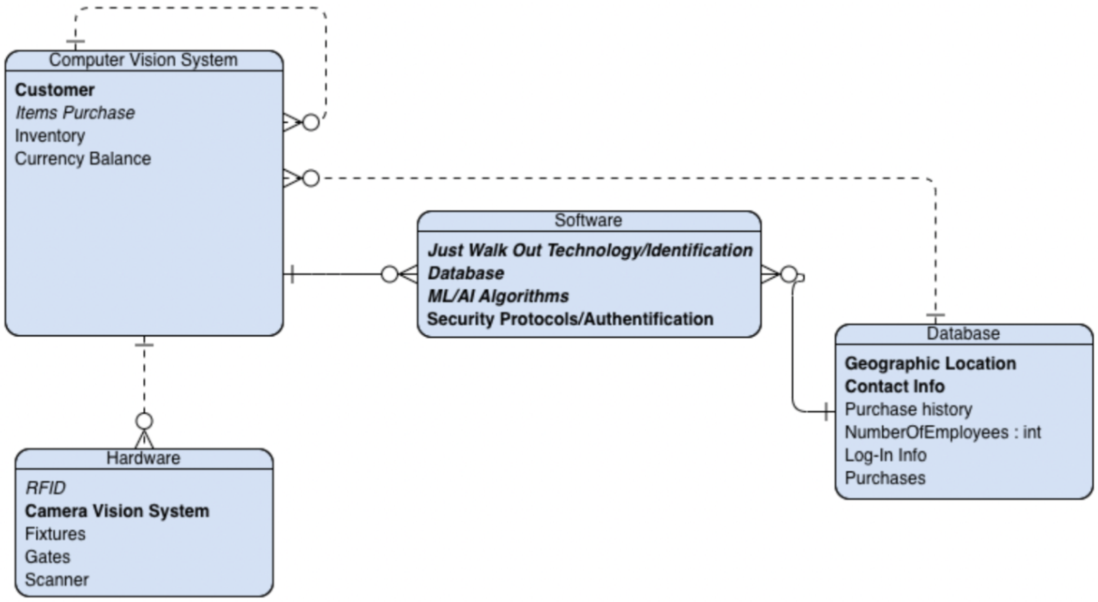


Figure 8: Class Diagram of CNN

Cameras may be placed all throughout the store to keep an eye on the shelves and track which products are available. These video feeds may be analyzed by computer vision algorithms to determine the goods, their numbers, and their positions on shelves. When items run short, the system may automatically issue replenishment orders by linking this data with inventory management software.

Furthermore, checkout systems that automatically identify and detect goods as consumers put them in their carts or baskets can be powered by computer vision. The technology can precisely identify products and their pricing by merging this visual data with product databases, which makes the checkout process easier. Customers may pay online using the CV system instead of interacting with a typical cashier and have their purchases charged to the appropriate customer account.

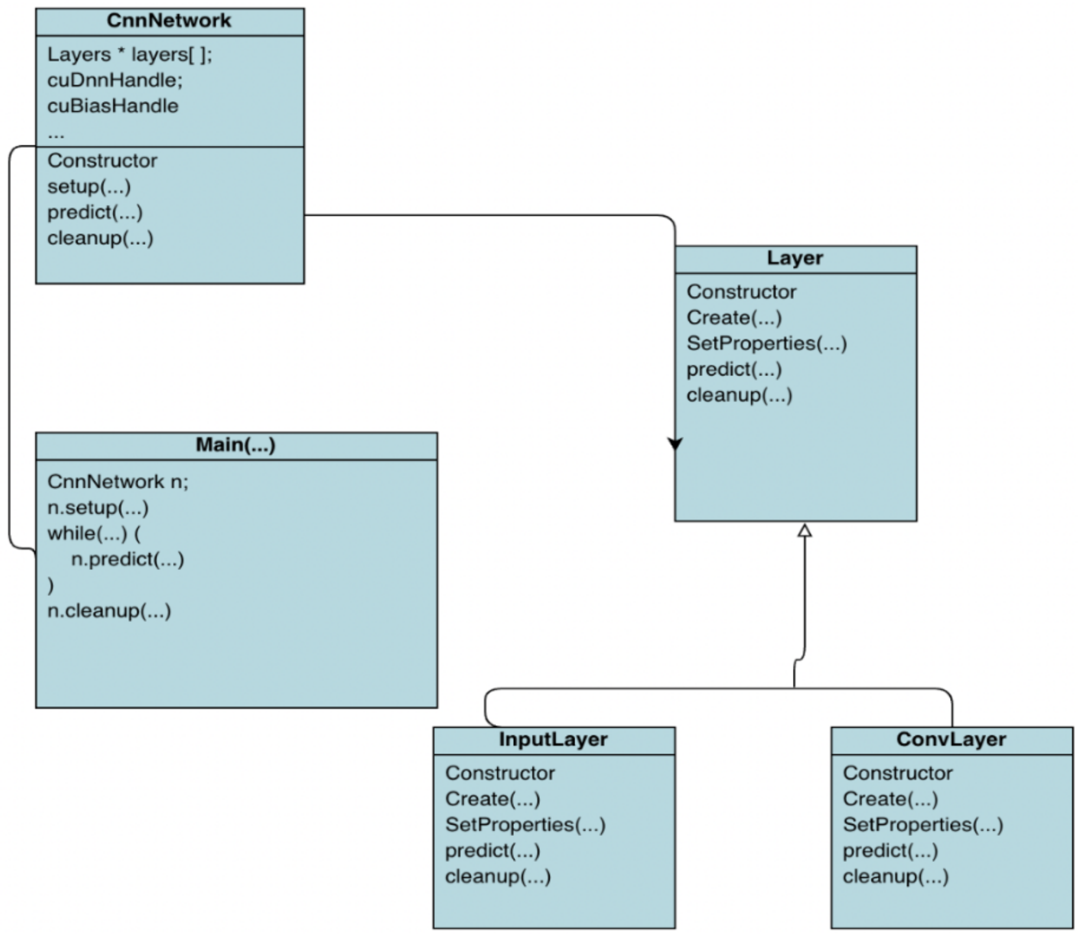


Figure 9: CNN for Automated Checkout in Grocery Stores

By comparing the actual inventory recorded by cameras with the projected inventory (derived from sales data), computer vision can assist in detecting cases of theft or lost products. Inconsistencies can reduce losses from shrinkage or theft by setting off alarms that prompt personnel to look into the matter. By analyzing consumer behavior and offering tailored suggestions based on how they interact with things, computer vision may be employed to improve the entire shopping experience and operational efficiency by analyzing video feeds to get insights on customer movement and behavior throughout the store. This information can then be used to optimize staffing schedules, product placement, and store architecture. Computer vision systems are able to dynamically alter price by assessing real-time data on client traffic, purchase habits, and inventory levels.

CNNs have become a highly effective tool for computer vision applications, including those in grocery shops and other retail settings. It can be used in automated checkout systems to visually recognize and track things that consumers have picked, removing the need for traditional barcode scanning in the process thanks to a comprehensive computer vision system. Contents of consumers' baskets are recorded by strategically positioned cameras and other sensors around the store. Noise reduction and normalization are two prepossessing techniques that get the pictures ready for CNN analysis. Convolutional layers recognize edges and patterns in an image by acting as feature detectors. These layers are trained to identify distinguishing characteristics between distinct objects, such the color of a cereal box or the shape of a banana. Pooling layers in the CNN down sample the extracted features, reducing computational complexity. Fully connected layers then classify the features into specific item categories. The final output layer assigns a probability score to each potential item, allowing the system to identify the most likely product with high accuracy. By recognizing identified items, the system can access the store's inventory database to retrieve product information like pricing, enabling real-time basket value calculation and automating billing for customer checkout.

CNNs do offer a few advantages in this retail application: for one, their ability to learn from large image datasets allows for robust item recognition, handling variations in product packaging, brand logos, and even slight occlusions within the basket. In addition, CNNs can be continuously trained on new data, improving the system's accuracy over time. They present a promising approach for automated checkout systems in grocery stores. Their ability to learn intricate visual features makes them well-suited for identifying a wide range of items, providing a friction-less shopping experience.

5 Pillar #3 - Behavioral

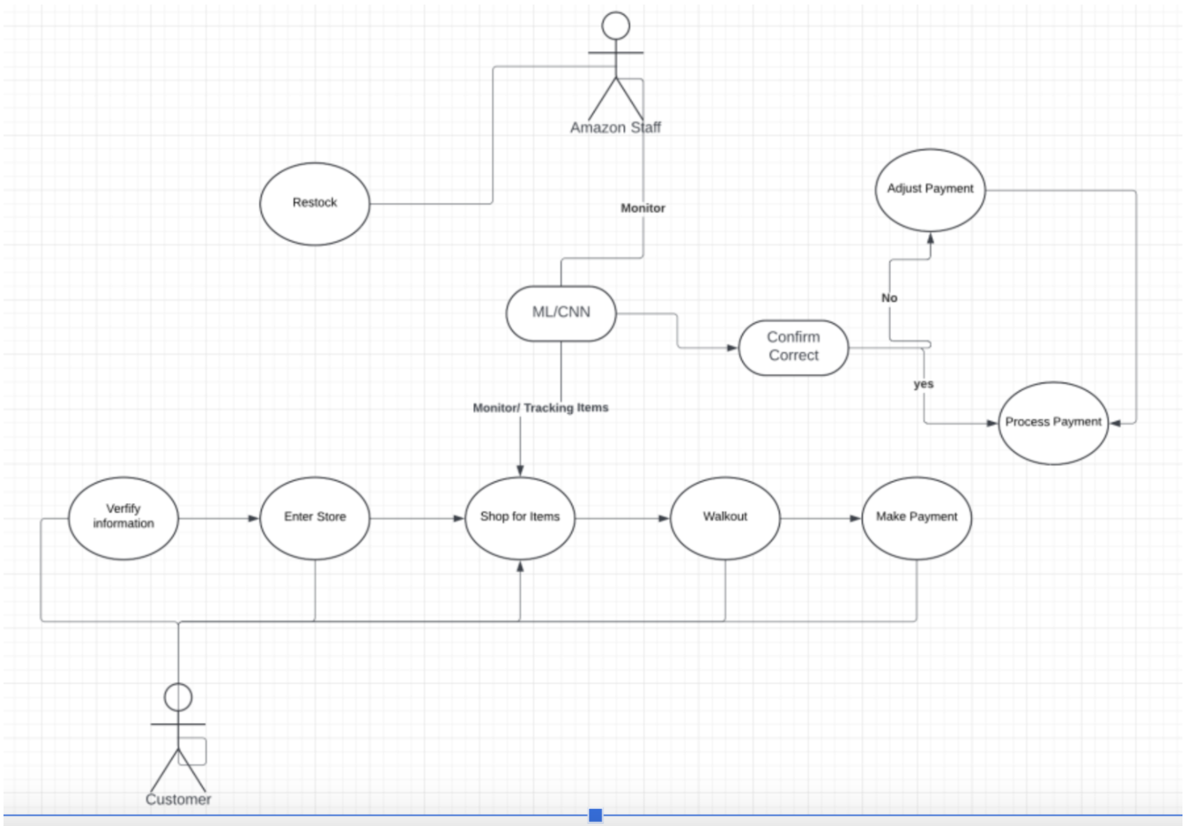


Figure 10: The use case diagram depicts the interaction between two key actors within the system: Customers and Staff.

Customer Use Cases:

(a) **Verify Account Information:** This step involves validating customer identification associated with the account.

(b) **Enter the Store:** Upon successful verification, customers' movement within the store is tracked by computer vision devices.

(c) **Shop for Items:** Customers select desired items and add them to their cart, with their movements analyzed by ML/CNN algorithms.

(d) **Walkout:** Upon leaving the store, payment is automatically processed.

Staff Use Cases:

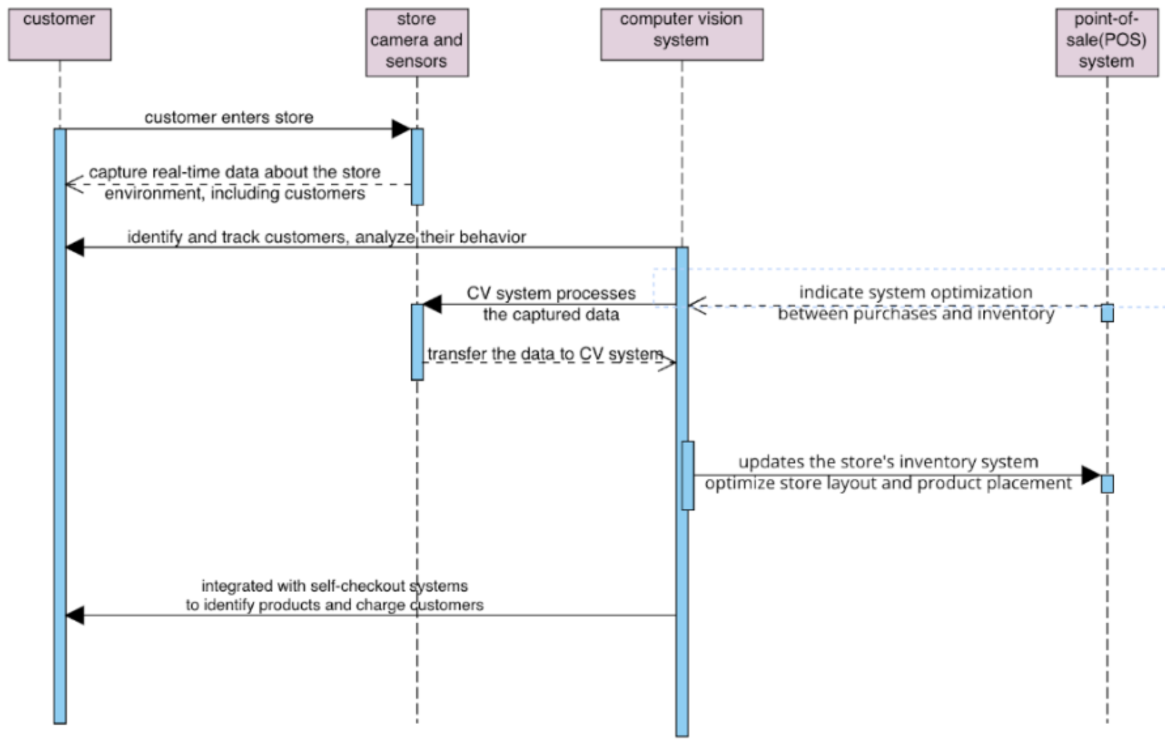
(a) **Restock:** Amazon staff replenish items based on system reminders.

(b) **Enter the Store:** Upon successful verification, customers' movement within the store is tracked by computer vision devices.

5.1 Sequence Diagram

The key to a good CV product lies in understanding how it functions within its environment. This includes how the system interacts with itself, users, and other systems. It is a critical step to analyze these behaviors, acting as a kind of quality check before any system or product is released.

This sequence diagram outlines the communication flow between different parts of the system, and visualizes the step-by-step business activities involved in completing a specific task. It essentially showcases how components interact with each other.



- **Customer:** The central figure in the retail experience. Their movements, product interactions, and shopping behaviors are analyzed by the system.
- **Store Cameras and Sensors:** The eyes and ears of the system, capture real-time data about the store environment.
- **Computer Vision System:** The brains behind the operation, processing data and extracting insights to optimize various aspects of the retail experience.
- **Point-of-Sale (POS) System:** The traditional checkout system, potentially integrated with computer vision for seamless self-checkout options.

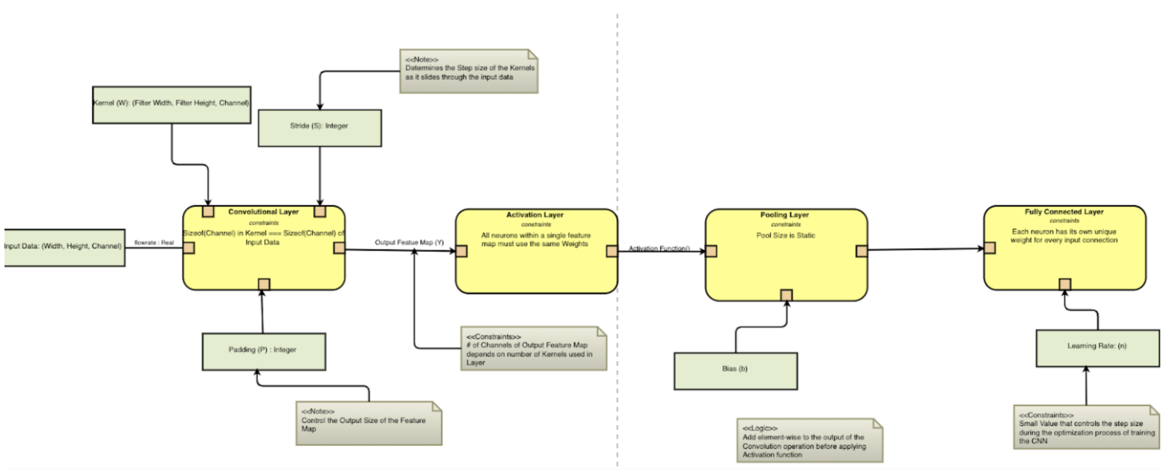
5.1.1 A Step-by-Step Look of Sequence Diagram

- Customer Enters Store: As a customer enters the store, their presence is captured by store cameras.
- Camera and Sensor Data Capture: Cameras and sensors continuously capture real-time data of the store environment, including customer movements, product interactions and potential shoplifting activities.
- Computer Vision Analysis: The captured data is then transferred to the CV system for further processing. The system identifies and tracks customers, analyzes their interactions with products, and detects any suspicious behavior that might require store staff intervention.

- Customer Tracking and Product Recognition: As customers navigate the store, the system keeps track of their interactions with products, and identifies the products they pick up, or add to their shopping carts. This implementation allows retail stores to provide customers with a more personalized shopping experience. For instance, a customer who frequently examines skincare products might receive personalized discounts on related items. In other words, stores can send targeted promotions based on each customer's product interactions and optimize their profits.
- Real-Time Inventory Management: By analyzing product interactions, the CV system can continuously update the store's inventory system in real time. This ensures accurate inventory levels and prevents stockouts, ultimately improving customer satisfaction.
- Self-Checkout and Automated Payment: In our retail settings, the CV system can be integrated with self-checkout systems. This integration streamlines the checkout process, automatically identifying products and adjusting charges accordingly. Imagine a scenario where a customer simply places items in their cart, and the system seamlessly recognizes each product in customer's shopping journey, and the payment process is automatically processed when they are leaving the store.
- Improved Store Layout and Product Placement: The data collected by the CV system offers invaluable insights into customer behavior and store dynamics. By analyzing factors such as customer product popularity and traffic patterns, retail stores can optimize their store layout and product placement.

6 Parametric Diagram

The fourth pillar of MBSE, the Paradigm diagram, is known as the essential component for capturing the mathematical relationships between all identified variables. This diagram further analyzes the reliability, sustainability, and performance of the entire system. To be able to make proper predictions and to understand potential outcomes is why the Parametric is the most important pillar of the system design. Since it's the fourth pillar of the entire design approach, it considers all the components that make an effective convolution system. To further emphasize this point, the most noteworthy aspect of the diagram is its ability to identify complex data as a whole and allows the system to better understand its inputs. This includes considering input data, the filtering process, and bias terms to further analyze the data. The mathematical representation CNN (since its the bulk of camera calibration) is expressed



as :

Notice the convolution only takes into account and filters particular regions within the graph itself, and how it has a direct linear relationship given a defined input matrix, bias term and filter matrix. This is because it aims to apply a constant weight before it is going through the pooling and fully-connected layer. Of course, there are multiple convolutional layers that are stacked on top

Operation	Description	Mathematical Expression
Convolution	Slides a filter across the input data and computes dot product	(X * W) + b
Activation Function	Adds non-linearity to the model's output	f((X * W) + b)
Pooling	Downsamples the data extracted by convolution	-

X = input data (matrix)
W = Filter (kernel) matrix
f(x) = convolution operation
B = bias term

of each other, and each layer identifies certain features based on the previous performance. Having a mathematical representation of the linear relationship will continuously vary depending on the network architecture as well as its real-world applications. In the Activation Function, the output will not be linear, and allows to identify complex data given the newly developed grid. The most common functions include the Rectified Linear Unit layer that will help identify the newly-understood image during the processing stage. With that being said, the Parametric diagram is a powerful and adaptable tool in SysL that incorporates linear and non-linear relationships, identifies constraints, and measures performance metrics. This allows engineers to further dissect how every layer should be performed, and allows fine-tuning among the entire process of CV within the retail environment.

7 Which is better? - Document Centric Model Versus Model-Based System Engineering Approach:

Before we discuss which is better between Model-based system Engineering or Document Centric Model

1. **PROS - Document-Centric Approach:**
   - **Comprehensive Process:** The Document Centric Approach is generally more detailed and provides an insight of the general system analytics. This can ensure a better understanding of the technical system performances and indicates how the system should behave using a well-defined map.
   - **Mitigating Potential Issues:** Allows for better risk management by leveraging and identifying potential bottlenecks early on in the development process. This can help save the cost of rework and time.
2. **CONS - Document-Centric Approach:**
   - **Less Visual Representation:** Since the Document Centric Model is based on a heavily typed collection of records. It is sometimes difficult to understand and construct a graphical representation, unless the individual has past experience.
   - **Limited collaboration:** There is the potential for misinterpretations among the development team and stakeholders. This can create communication gaps and misunderstandings about everyone's goals and priorities if not articulated properly.
3. **PROS - Model Based System Engineering (MBSE):**
   - **Improved Collaboration:** Utilized IEEE standards to provide inspection, transparency, and adaptability between different teams and stakeholders.
   - **Graphical Representation:** In MBSE provides more of a visual representation of the system itself, making it easier to comprehend information for all stakeholders and help understand the overall structures and behavior.
4. **Cons - Model Based System Engineering:**

- **Extensive Focus on Visual Representation:** The intention of creating models can lead to over-engineering of the process rather than meeting the original objective. The goal is to construct plain text into a visual representation, and as we know, there will always be a communication gap between the stakeholders and technical individuals.
- **Incomplete Coverage:** MBSE may not cover all aspects of the system development as much details as the Document centric Approach. This can lead to gaps within system performances and analysis.

7.1 Our Opinions of Which Diagrams are Better:

While the Document Centric Approach emphasizes a systematic process within the design principle; it also directs many more factors that make it a significant approach to any system development. This includes Requirement Allocation, Feasibility Analysis, and other system design considerations. Technical Performance Measurements and potential trade-off Analysis are the best contributing factors that make Document Centric the go-to for any engineer.

With that in mind, Model-Based System Engineering (MBSE) with SysML focuses on four key aspects: Structure, Requirement, Behavioral, and Parametric. Because of this, it offers more advantages that allow the ability to create a visual representation of the system's structure and overall behavior. This encourages many representatives to engage in collaboration and allows more inspections to take place to ensure System Performance.

To create a retail environment empowered by CV technologies, we would prefer using a hybrid approach of both Document Centric and MBSE. The reason why we would do this is to allow various people to analyze the entire system performance in their preferred choice. During both stages, it is important to view the analytics and technical performance measurement while sustaining a visual representation of the requirements itself. This can help eliminate any potential bottlenecks and cost of rework done by the developer. The advantage of Document centric is its potential to apply trade-offs as an alternative route to the powerful system. As you seen in the diagram above, we provided a trade-off from the document centric approach as a visual representation. This eliminates any possible security restrictions and allows a projectile in performance. In conclusion, based on our analysis, we believe a hybrid approach would be more optimal than selecting a preferred approach.

8 Conclusion

Throughout this report, we have explored numerous ideas aimed at crafting an efficient digital retail environment. We have not only identified potential risks but also directed various aspects of system behavior, establishing background information and visual representation to guarantee optimal performance. Our approach has drawn upon insights from comprehensive overviews, integrating the four pillars of the SysML application and leveraging them as an efficient tool. While each pillar serves as a unique yet well-defined approach to the overall CV model, they all can be traced back to the origin. A well-defined pillar means all aspects have been analyzed, and recorded, leading to a scalable approach to any system platform. In this case, applying a well-defined CNN within all the cameras to ensure image acquisition and processing phase while the customer is inside the building.

By providing a well-defined SysML, we can apply more inspection and adaptation to the entire digital ecosystem. Many of these requirements are based on what each individual presents, and all of which are traced by the origin requirements. We believe having a hybrid approach for both Document Centric and Model-Based System Engineering would be efficient because it can encourage many individuals of different backgrounds to engage and collaborate their insights. It will be of benefit to encourage collaboration while highlighting a visual representation and behavioral approach to modern system design.

References

- Haralick, R. M. (1992). Performance characterization in computer vision. In BMVC92: Proceedings of the British Machine Vision Conference, organized by the British Machine Vision Association 22–24 September 1992 Leeds (pp. 1-8). Springer London.

- Davies, E. R. (2017). Computer vision: principles, algorithms, applications, learning. Academic Press. Wankhede, K.,

- Wukkadada, B., Nadar, V. (2018, July). Just walk-out technology and its challenges: A case of Amazon Go. In 2018 International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 254-257).

- IEEE. Polacco, A., Backes, K. (2018). The amazon go concept: Im plications, applications, and sustainability. Journal of Business and Management, 24(1), 79-92