VIETNAM DATATHON 2023

WINDOW MIRROR

Team: The Da-rer

Table of Contents

1 Introduction

2 Problem Statement

3 Solution Overview

4 Methodologies

5 Core Functionality

6 Performance Metrics

7 TImeline and Roadmap

8 UI or Interaction

9 Limitations and Future Enhancements

10 Conclusion

1
INTRODUCTION

Provide an ***overview of the problem***, ***introduce*** the MVP and its ***purpose***, or opportunity the MVP aims to address

1. Introduction

*In the realm of retail, **boosting sales** for fashion stores is an **ongoing challenge...***

Our MVP strides in with a **bold mission** - **Empowering Trends, Elevating Sales**

→ Proposing **real-time and targeted ads** based on Customer

→ **Heatmap** that **tracks** the popularity of displayed items

2

PROBLEM STATEMENTS

- Clearly define the ***challenge*** or ***problem*** that the MVP seeks to solve
- Describe the ***pain points*** or ***inefficiencies*** associated with the problem
- Describe the ***current landscape*** and ***positioning of the competitors*** if any

Problems statement

The MVP aims to **increase fashion store sales** by introducing **a real-time changing window display** screen that adapts content based on customer behavior and foot traffic, addressing the limitations of static traditional displays.

01

Engage

Engage customers with more relevant and interesting content

02

Trigger

Trigger product try-ons and purchases

03

Collect

Collect data on customer behavior to improve the shopping experience

In summary

Fashion store owners need a way to capture the attention of passersby and entice them to enter their store.

The MVP solves this challenge by creating a dynamic and interactive window display that can be tailored to the specific needs and interests of each customer.

The pain points or inefficiencies associated

The pain points or inefficiencies associated with the problem of **traditional window displays** are as follows:

01

Traditional window displays lack adaptability to customer behavior and foot traffic.

The same content is displayed to all customers, regardless of their individual interests or the current situation.

02

Traditional window displays can be expensive and time-consuming to create and maintain.

This can be a significant barrier for small businesses and independent retailers.

03

Traditional window displays are not very effective at collecting data on customer behavior.

This makes it difficult for fashion store owners to understand what their customers are interested in and how they are interacting with their window displays.


WINDOW
MIRROR
Creating a **real-time changing window display** screen that can be tailored to the specific needs and interests of each customer.

Competitive Landscape

01 Current landscape

The global digital signage market size: $21.7 billion (2020) and is projected to reach $39.6 billion by 2028, growing at a CAGR of 8% from 2021-2028

(Verified Market Research)

The AI retail analytics market size was valued at $3.54 billion in 2021 and is expected to expand at a CAGR of 21.5% from 2022 to 2030.

(Grand View Research)

02 Key competitors

NEC Display Solutions

Offers digital signage systems integrated with sensors & computer vision for customer analytics

intel

Grew digital signage segment revenue to $5.5 billion in 2021 through IoT, AI and analytics solutions

aws

Launched retail monitoring cameras and software powered by AI/ML platforms

03 Competive advantage

- First-mover advantage in a new category before others recognize the window opportunity.
- Simpler design allows innovation and ROI focus versus feature-rich systems.

→ Focusing on the unmet needs of window displays provides a niche for disruptive growth versus direct competition with digital signage giants. The low-cost, targeted MVP is well positioned initially

3

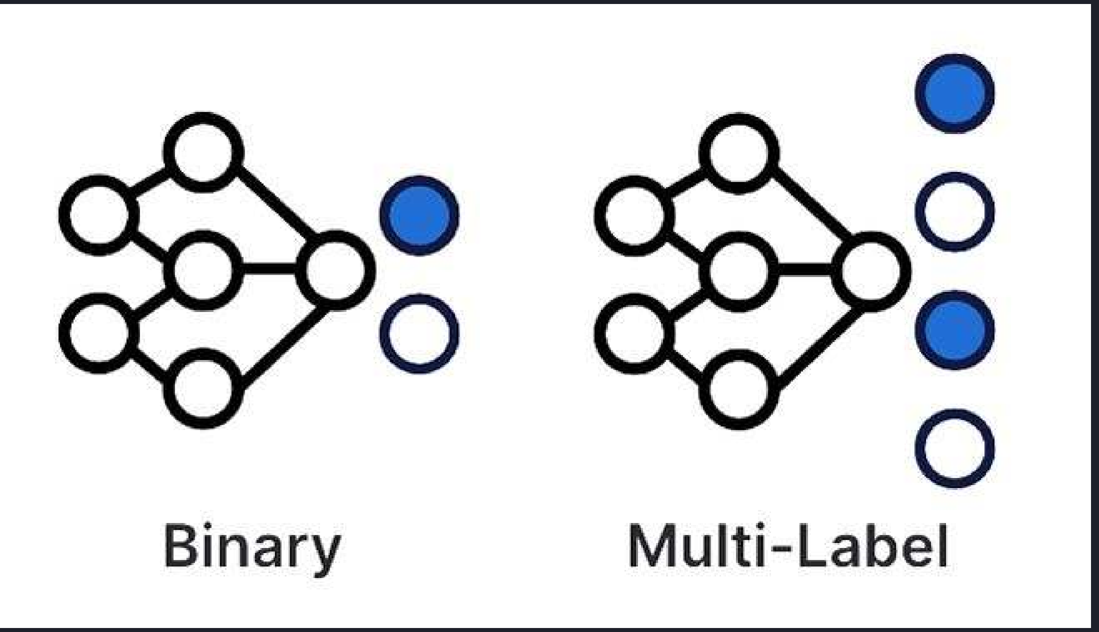
SOLUTION OVERVIEW

- Present a high - level overview ***AI-based solution*** proposed in the MVP
- Explain how the ***solution leverages AI techniques***, algorithms or models
- Explain how this ***solution is innovative***, highlighting ***its novelty*** either in terms of its technical aspects or its fulfillment of business needs

High - level overview AI-based solution

Understand the problem clearly

- In our task, we tackle the challenge of dual classification: identifying individuals as either 'person' or 'non-person' (binary classification), and simultaneously characterizing the behavior of customers through multi-label classification.



- Our objective is to develop a model that excels in both binary and multi-label classification, catering to the dual nature of our recognition task and providing a robust solution for our application.

High - level overview AI-based solution

Faster R-CNN

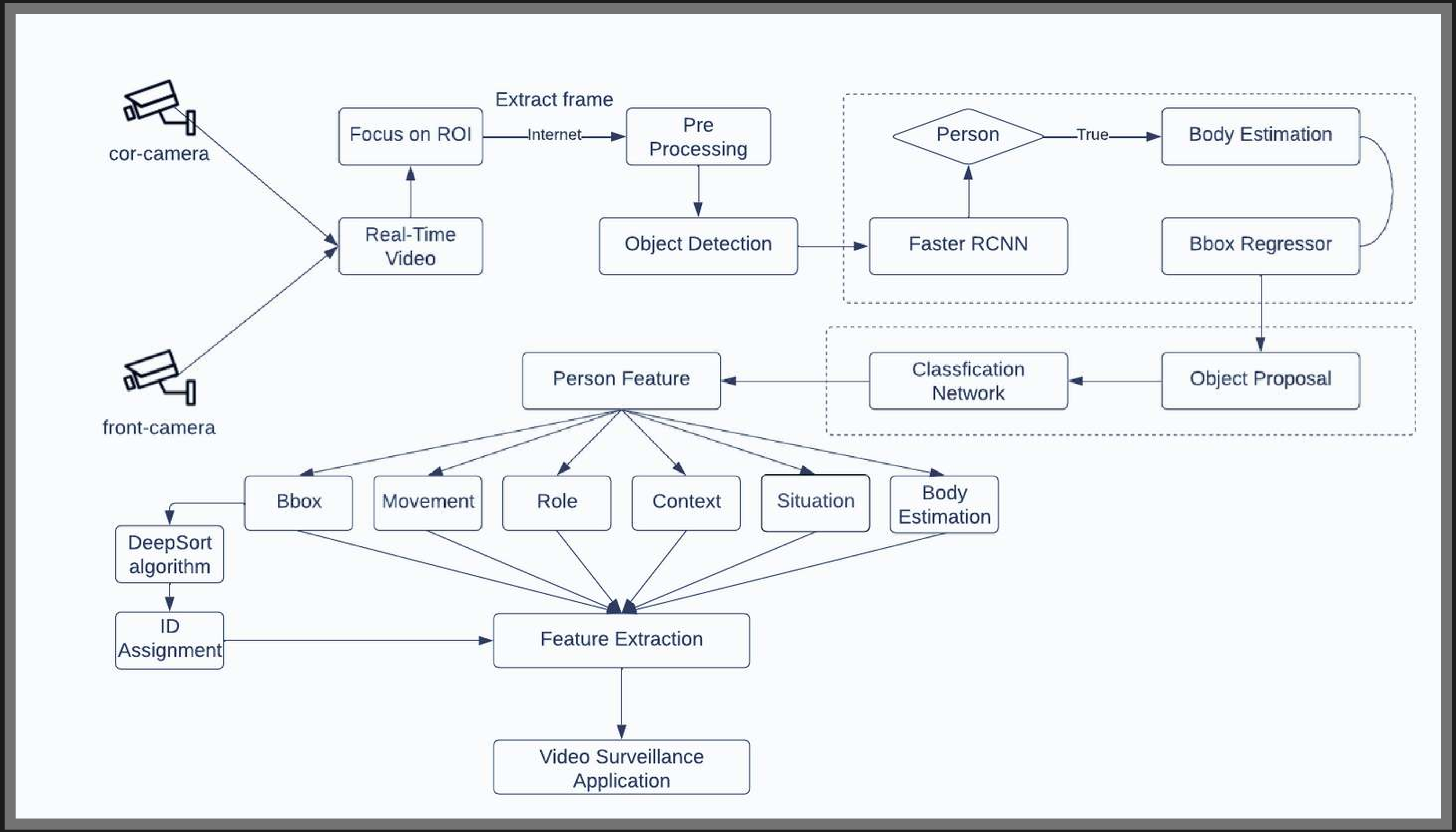
Model

- There are numerous models available for solving object detection tasks, with YOLO being one of the prominent ones.

- After examining the data and requirements of our task, we propose the use of the Faster R-CNN model. This decision is based on our goal, which prioritizes high accuracy without overly sacrificing frame rate speed.

- The dataset we possess, particularly focused on the Region Of Interest, aligns well with the capabilities of Faster R-CNN, contributing to optimizing the camera scale.

- Additionally, our aim is to leverage Faster R-CNN for easy integration with custom-designed models and the flexibility to fine-tune its parameters.



The general system overview diagram

The chart illustrates the most fundamental process of the system that we have conceptualized and agreed upon.

How the solution leverages AI techniques, algorithms or models

01 Faster R-CNN model with the ResNet50 layer for accurate person detection in our system.

02 Additionally, the DeepSort algorithm is employed for tracking and analyzing object movements, providing detailed insights into customer behavior.

03 The PyTorch framework is selected for its flexibility and robust community support, allowing us to optimize the performance of Faster R-CNN and ResNet50.

04 For camera stream extraction, we use the H.264 protocol to ensure efficient data transmission with reduced bandwidth.

05 Implemented MongoDB as the database to store and manage customer journeys. This NoSQL database provides flexibility and scalability, allowing us to efficiently store and retrieve detailed information about customer interactions, movements, and behaviors.

Novelty in fulfillment of business needs

People enjoy shopping in physical stores for the experience it provides, not just out of necessity. Windows in the future must go beyond passive product displays and more actively augment the shopping experience for browsers and passersby. Our website turns store windows into a storytelling gateway to deepen loyalty and word-of-mouth - representing an innovative approach to optimizing the physical retail environment.

01

Personalized window display.

Personalizing content for each individual or group helps capture more targeted customers traditionally ignored by "one-size-fits-all" displays

02

Impulse buying.

Encourages impulse buying by window shoppers through targeted product information and limited-time discounts.

03

Interactive digital experience.

Engages crowds with an interactive digital mini-game to promote the brand and drive interest in offers/gifts during busy periods.

04

Testing.

Allows testing different display content quickly to optimize for conversions without changing physical displays.

Computer vision & dynamic digital display capabilities + customer-centric use case = innovative retail solution that can boost customer engagement, conversion, and sales.

4

METHODOLOGIES

- Provide the ***description of the architecture*** or structure of the AI model that you use
- Explain the ***key components***, ***layers*** or ***modules*** of the model
- Provide the ***technologies*** that you intend to use

Methodologies

As outlined above in the deployment of machine learning models and algorithms, we will now delve deeper into the data processing pipeline, machine learning model construction, storage, security, and deployment onto the application.

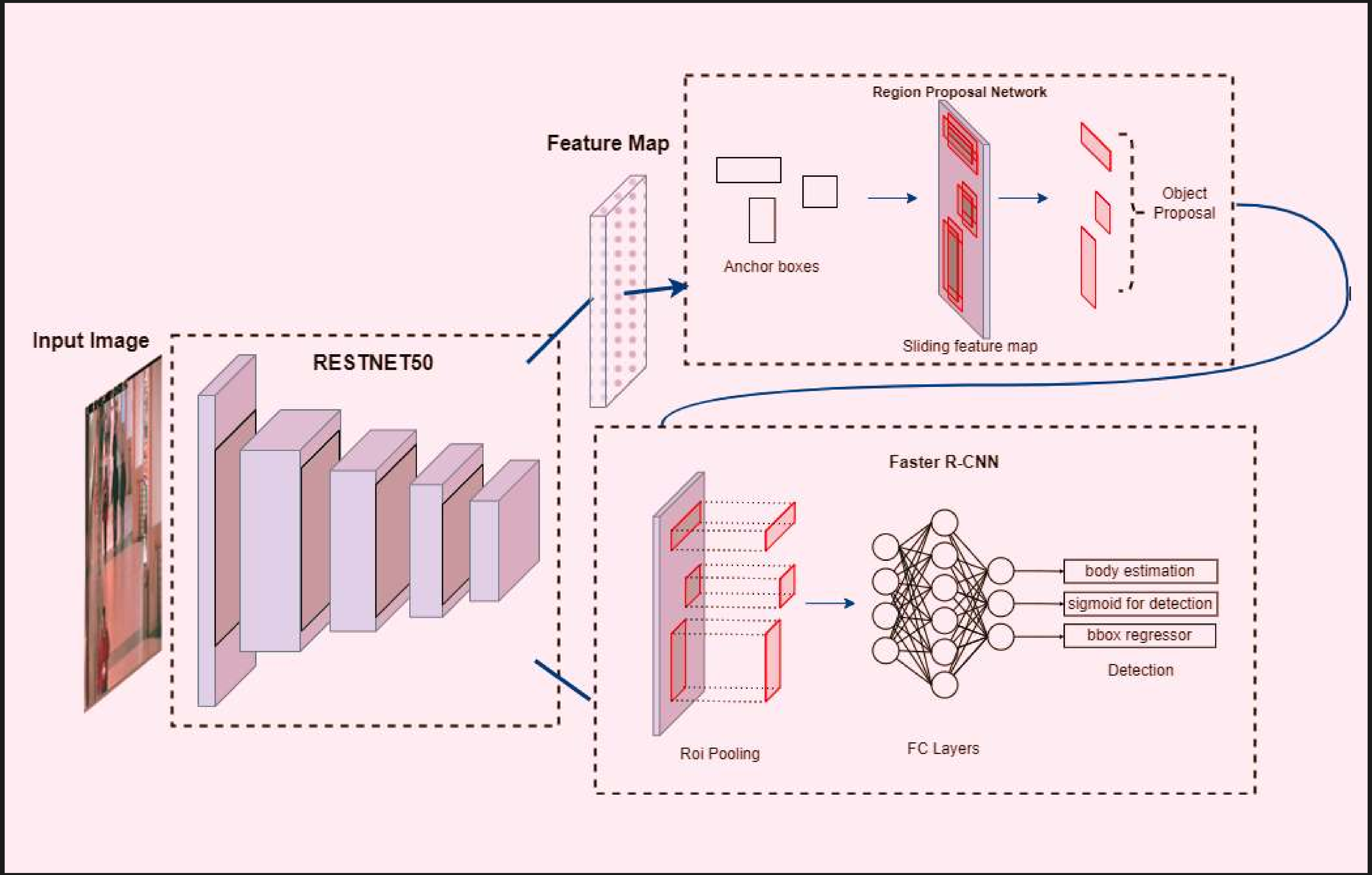
Deep Learning Model & Algorithms

- Faster R-CNN, Restnet50
- Classification Network
- DeepSort (Tracking & ID Assignment)

Data Storage & Application Deployment

- MongoDB Database
- Application Deployment

Deep Learning Model & Algorithms

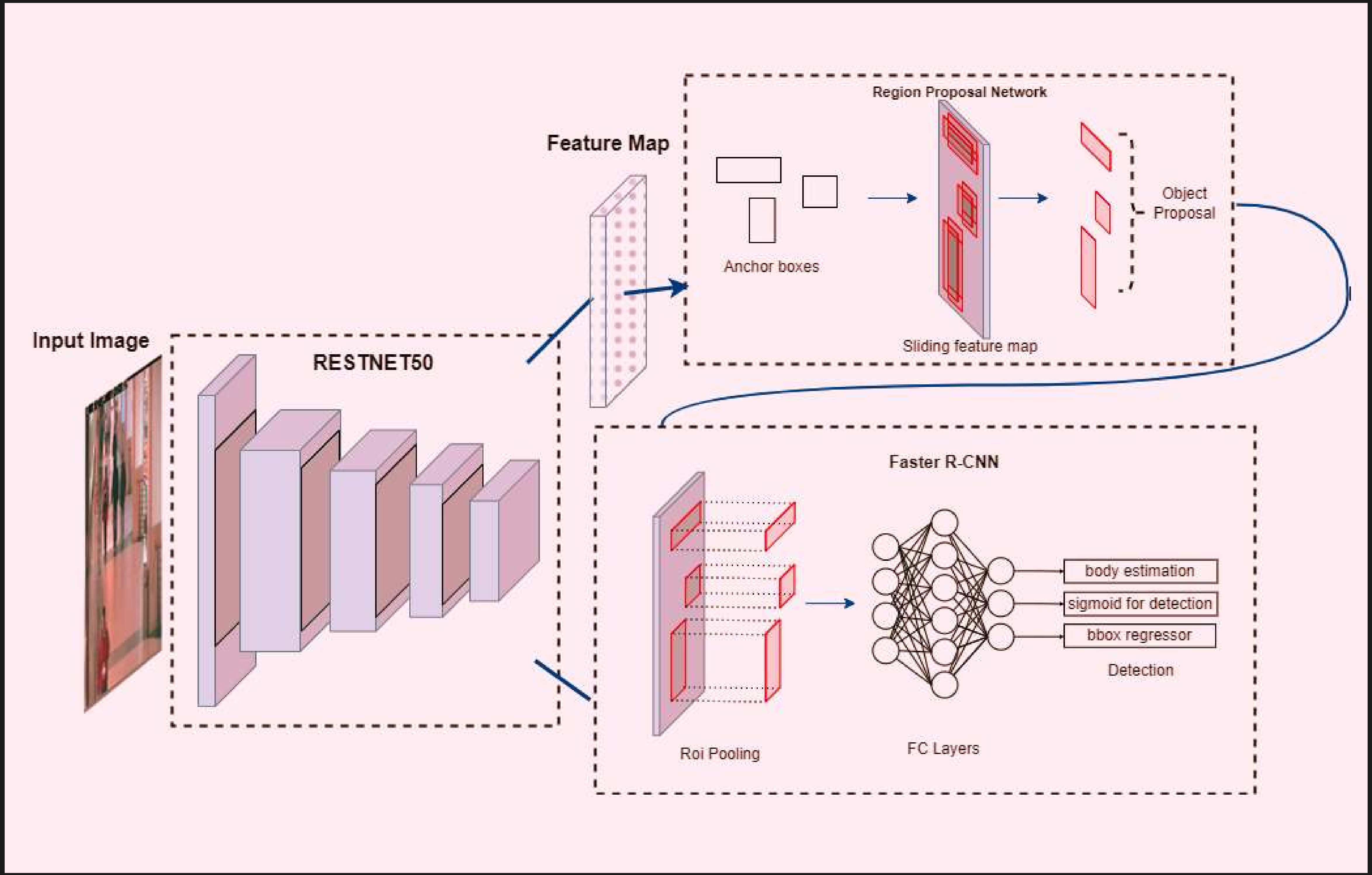


Faster R-CNN & Restnet50 Diagram

- Feature Extraction: The ResNet50 model is used to extract features from the input image. This process helps generate a numerical representation of the image with features at different levels.

- Region Proposal Network (RPN): After obtaining features from ResNet50, Faster R-CNN utilizes the RPN to propose regions containing potential objects. The RPN suggests important regions based on the extracted features.

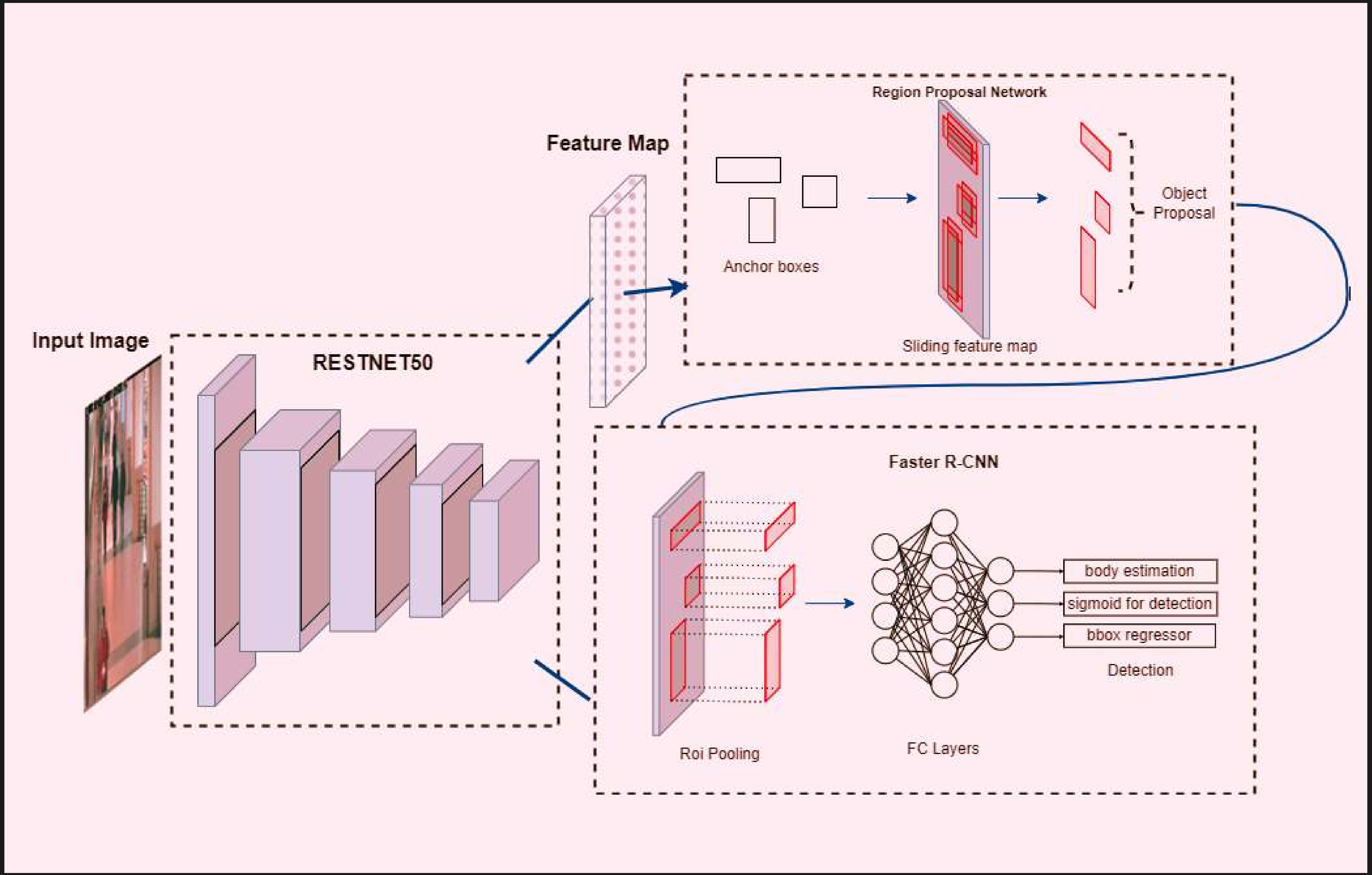
Deep Learning Model & Algorithms



Faster R-CNN & Restnet50 Diagram

- ROI Pooling (Region of Interest Pooling): Following the region proposals from the RPN, Faster R-CNN uses ROI Pooling to transform the proposed regions into fixed-size representations. These representations can be fed into a fully connected network to predict the class and position of the objects.
- Class and Bounding Box Prediction: The model employs a fully connected neural network to predict the class of objects in the proposed regions and adjusts (fine-tunes) the position of the predicted bounding boxes.

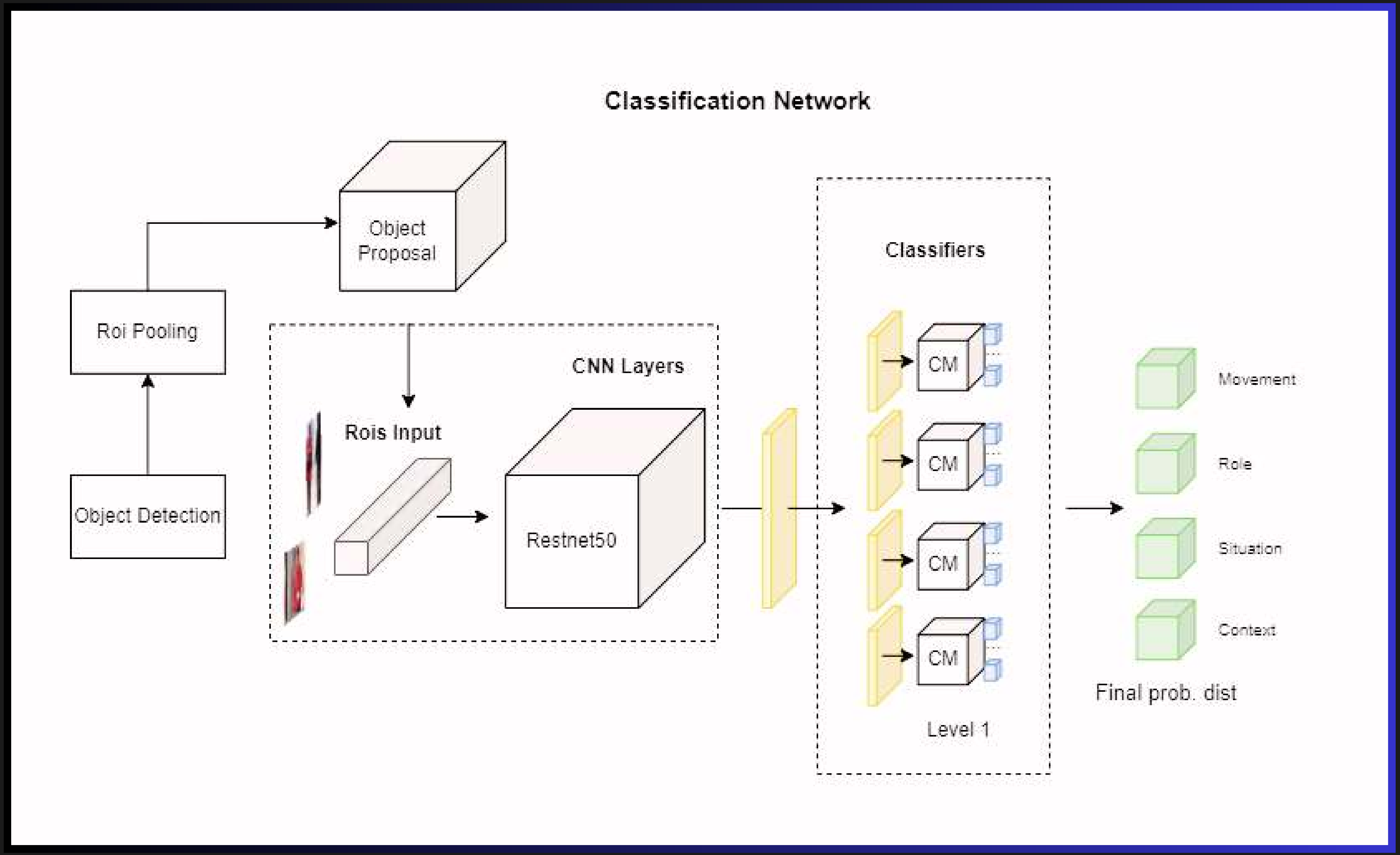
Deep Learning Model & Algorithms



Faster R-CNN & Restnet50 Diagram

- Here, we use a sigmoid function for the output of Faster R-CNN because it is a binary classification problem (0 or 1). This choice is to simplify the interpretation of the problem instead of using the softmax function (for multi-class classification).

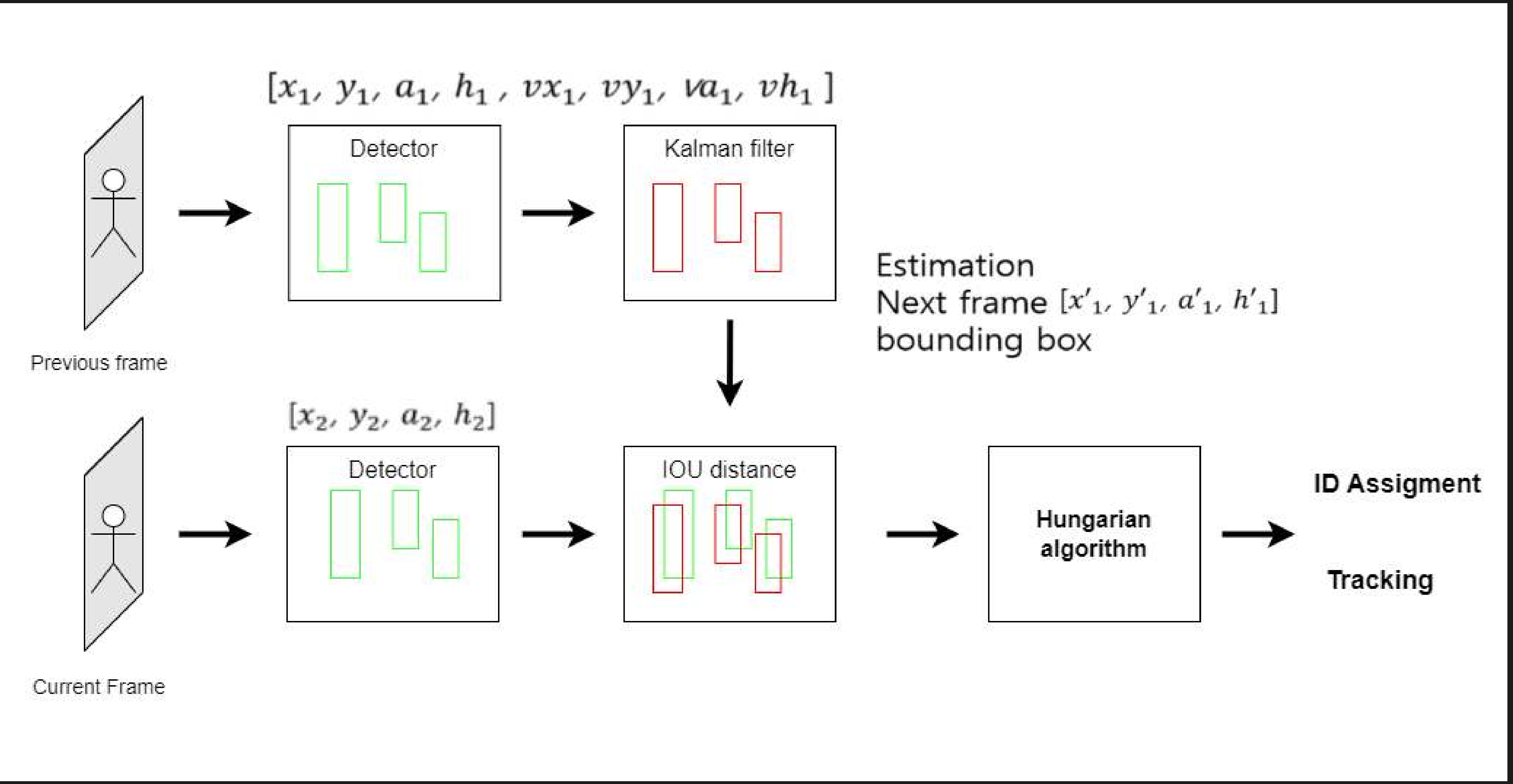
Deep Learning Model & Algorithms



Classification Network Based On ResNet50

After extracting RoIs (Regions of Interest) Proposals from the preceding network, we input these regions into a classification network to predict features related to individuals, with 4 outputs representing the 4 attributes of interest (Movement, Role, Situation, Context).

Deep Learning Model & Algorithms

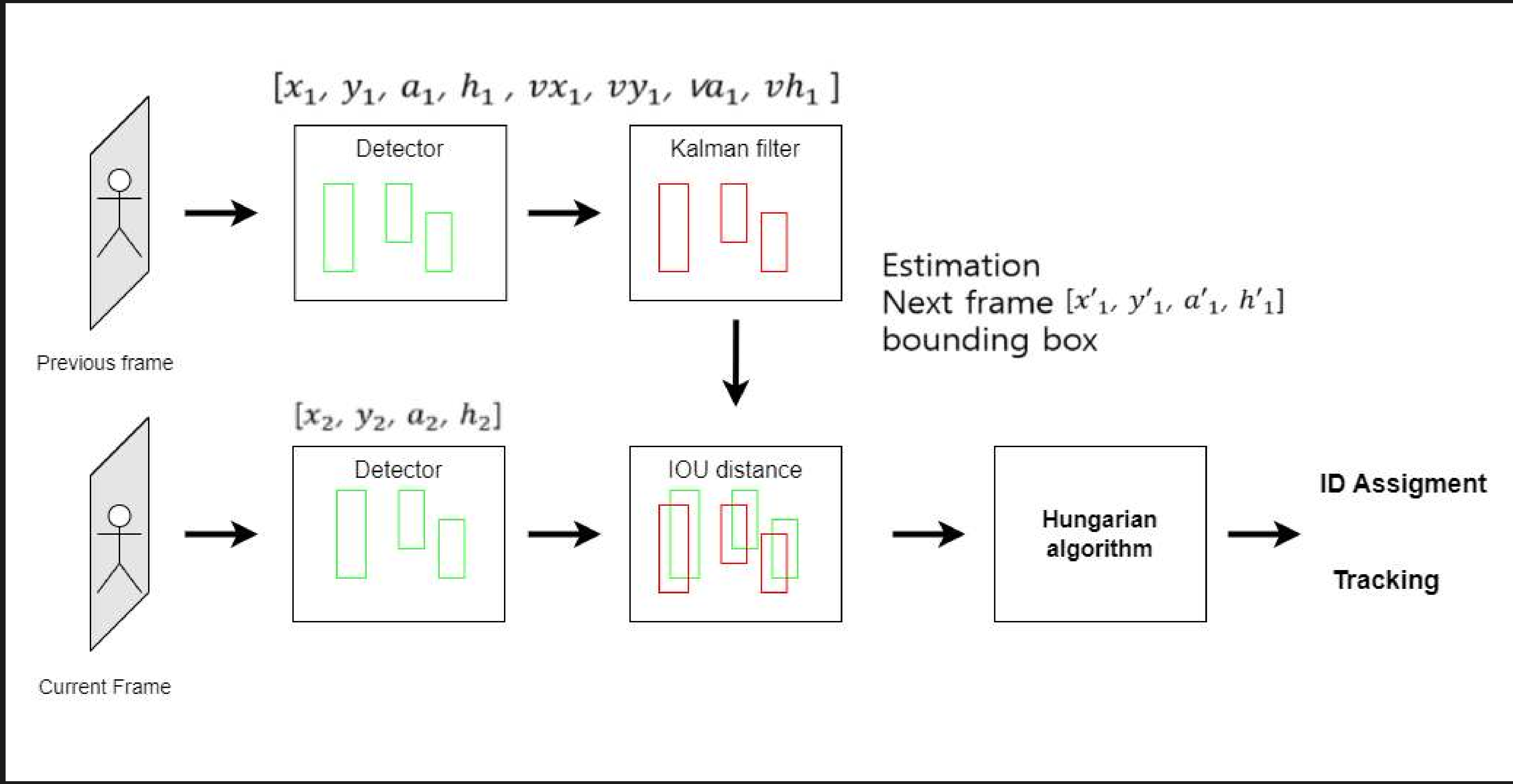


DeepSort Algorithms Diagram

- Previous Frame: In the previous frame, the Faster R-CNN model predicted bounding boxes for the appearing customers.

- Current Frame: In the current frame, the Faster R-CNN model is used again to detect and locate objects. New bounding boxes are predicted for each object in the current frame.

Deep Learning Model & Algorithms

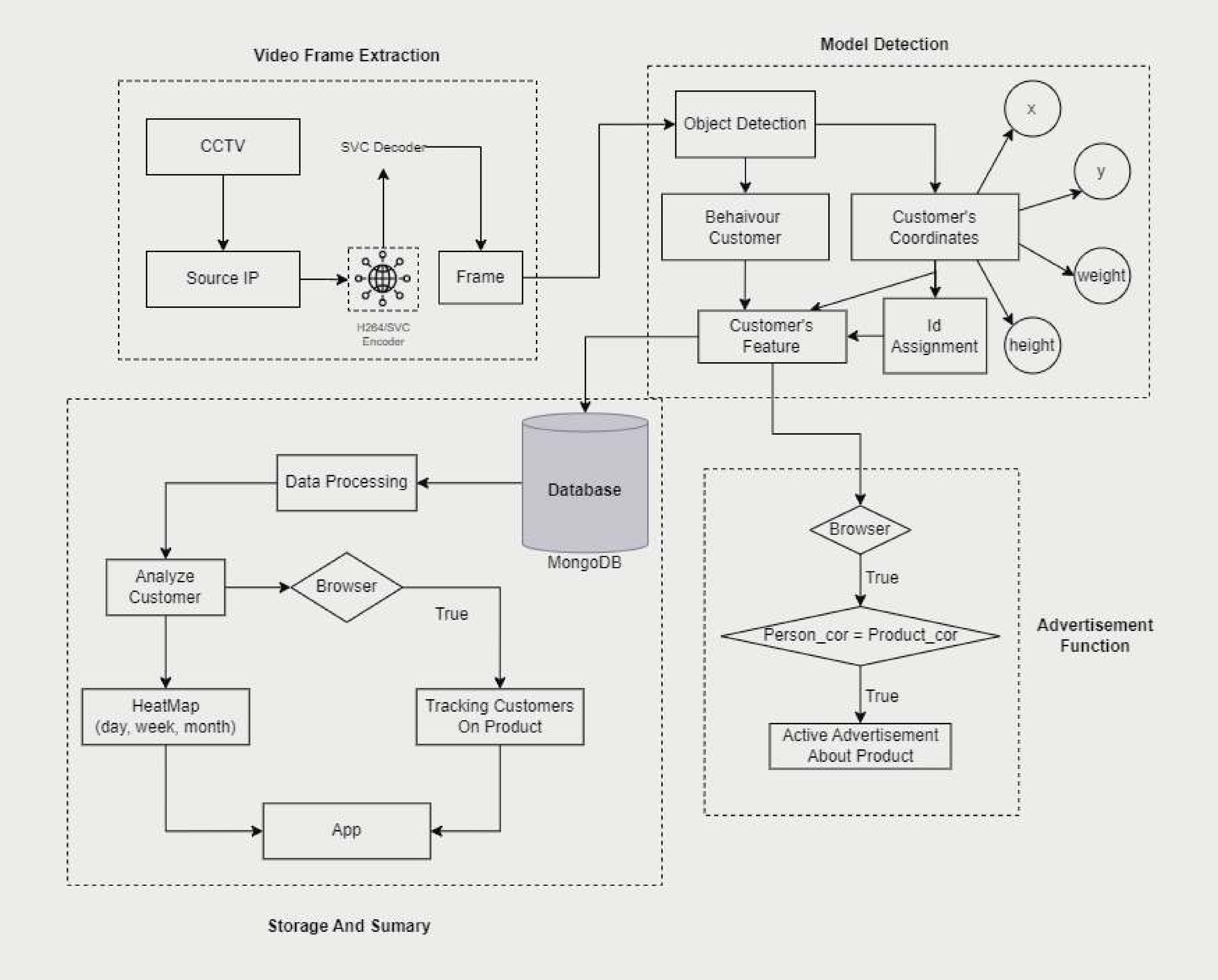


DeepSort Algorithms Diagram

Association and Tracking

- Deep Association: Features are utilized to compute the similarity between objects in the previous and current frames. DeepSORT employs a deep association model to associate objects from the previous frame with their corresponding objects in the current frame.
- Kalman Filter: For tracking purposes, DeepSORT often employs the Kalman filter to predict the positions of objects in the upcoming frames and adjusts predictions based on new information from the current frames.

Data Storage & Application Deployment

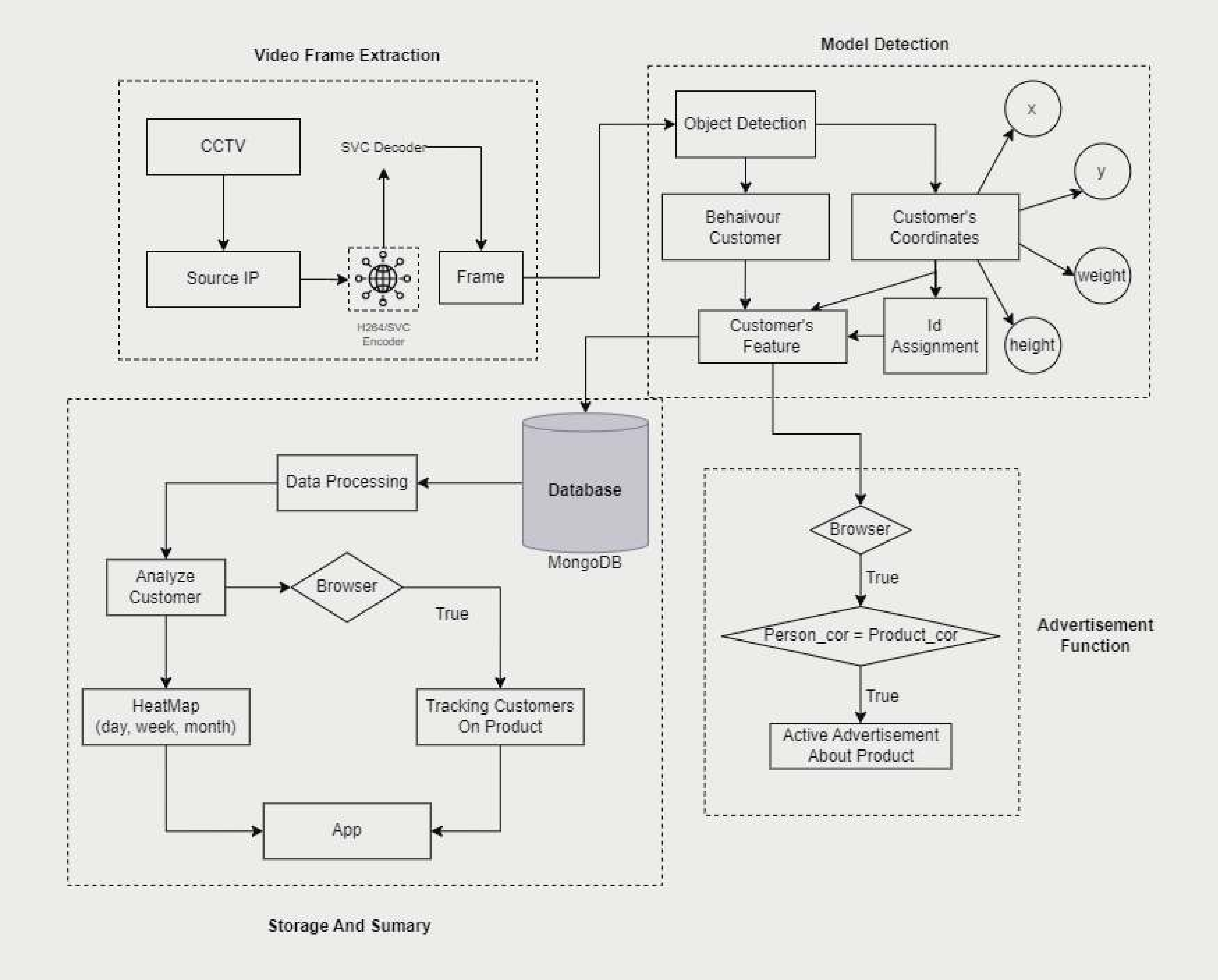


Data Storage & Application Deployment Diagram

H264

- Frame Structure: employs a frame structure consisting of different frame types such as I-frames (Intra), P-frames (Predictive), and B-frames (Bi-directional) to optimize video compression and transmission.
- Bitrate Control: provides bitrate control methods such as CBR, VBR, and RC to adjust video quality and file size based on specific requirements.
- Coder and Decoder Standard: Defines both a coding (coder) and decoding (decoder) standard for compressing and decompressing video.
- Efficient Video Compression Techniques: Employs various efficient video compression techniques such as avoiding data redundancy, block partitioning, prediction, and using transform coding to reduce the amount of data required for transmission.

Data Storage & Application Deployment



Data Storage & Application Deployment Diagram

Database

Using MongoDB to store the coordinates of customers in JSON format, the system will perform daily statistics. Users can optionally choose weekly or monthly statistics based on the dataset size. After that, data preprocessing will be conducted, and a grid will be constructed to draw a heatmap, illustrating the level of customer interest and counting the number of occurrences.

Advertisement Function

After obtaining the coordinates of individuals in a frame, we will sequentially check each set of coordinates to identify if any customer is near the coordinates of the provided products and is in a browsing state. If a match is found, we will execute the campaign as outlined.

5

CORE FUNCTIONALITY

Outline the ***primary features*** and ***functionalities*** of the MVP

Core functionality

Our website analyzes CCTV footage and recommends relevant content to enable real-time changing display screens for a better window shopping experience.

Core Features

- **Computer vision/AI module** to analyze CCTV footage in real-time and **identify customer behaviors** - walking alone, in groups, window shopping, and crowd sizes.
- **Digital display screens** are connected to the website for **personalized, real-time content** based on detected behaviors.
- **Database** to store predefined **video/image creative content** for different customer scenarios.

Additional MVP Features

- **Basic attention heatmaps** from computer vision to highlight browsed products.
- Collection of **store visit duration** and **conversion rate** basics

Key Functionalities

- Automatically detect customer behaviors and switch display content accordingly - **brand messages** for individuals, discounts for groups, and **product information** for attentive window shoppers.
- **Interactive mini-game** to engage crowds and collect customer contact information for rewards/offers.
- QR code generation and scan for **time-bound product discounts** to window shoppers.
- Backend dashboard for store admins to upload **new content** and **view analytics**.

→ The MVP would focus on building the core real-time detection and dynamic display switching along with interactive features. Additional analytics and integrations can be expanded later based on feedback.

6

PERFORMANCE METRICS

- Define the ***key performance metrics*** that will be used to access MVP's success

- Explain how the ***MVP's performance will be measured*** and ***evaluated***

Key performance metrics

Mean Average Precision (mAP)

One of the most crucial metrics for object detection is mAP (Mean Average Precision). mAP measures the model's ability to accurately determine the precise location of objects and the confidence level of predictions. It calculates the average precision values at various Intersection over Union (IoU) thresholds.

$$mAP = \frac{1}{n} \sum_{k=1}^{n} AP_k$$

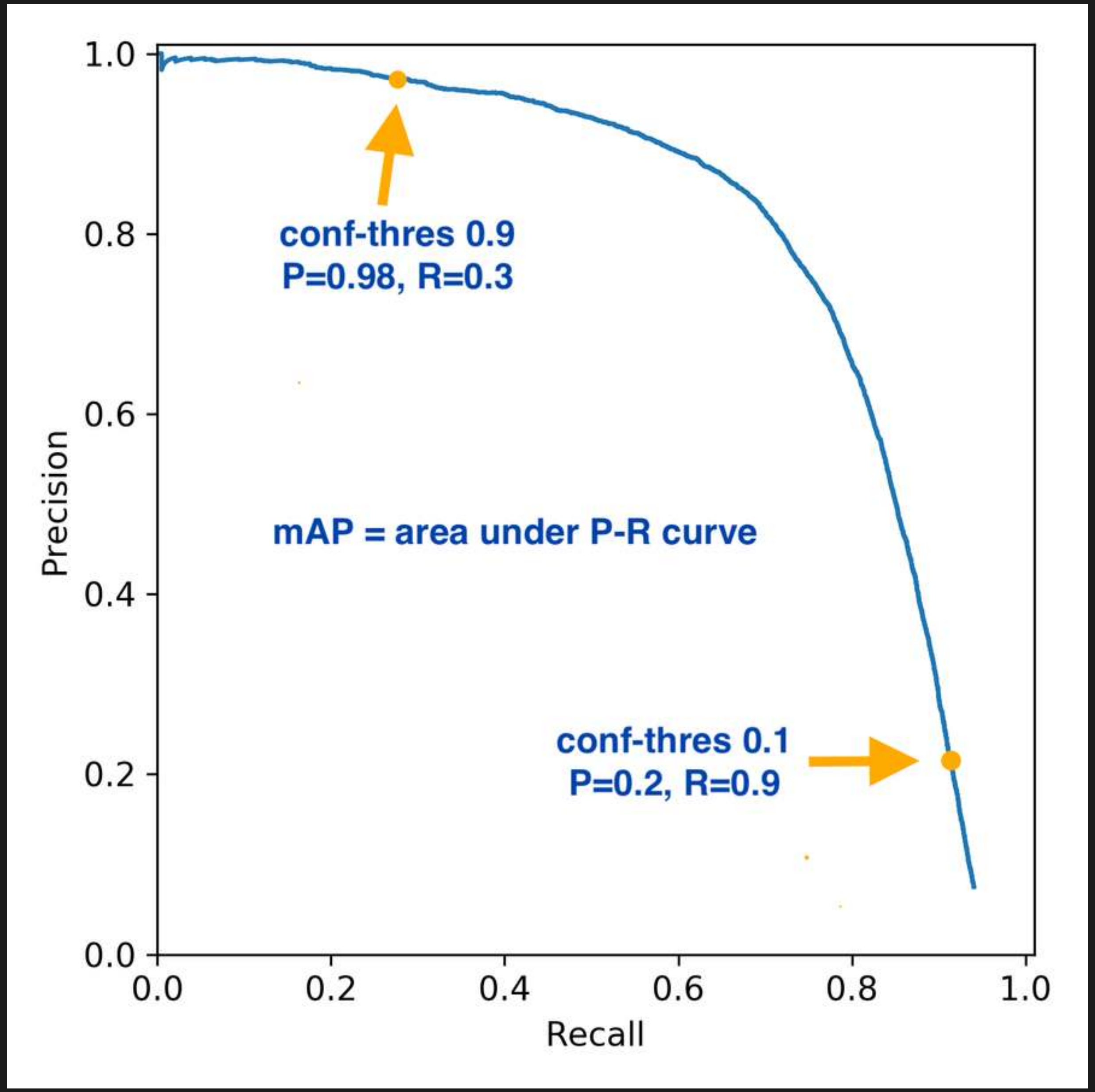
$$AP_k = The\ AP\ of\ class\ k$$

$$n = number\ of\ classes$$

Key performance metrics

Precision-Recall Curve

The Precision-Recall Curve is a commonly used graph to evaluate the performance of a model in binary classification tasks. This graph illustrates the relationship between precision and recall at various confidence thresholds.



Key performance metrics

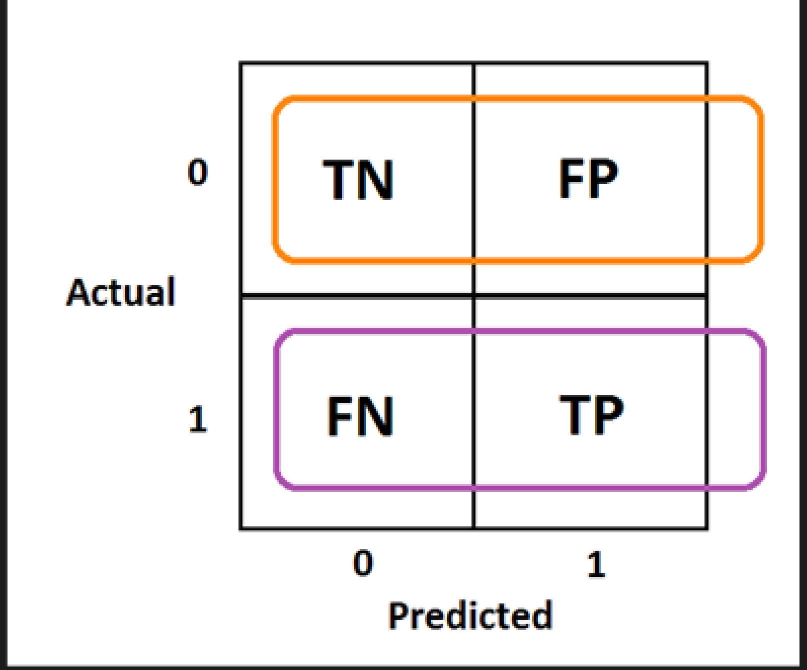
Precision-Recall Curve

- Precision (Positive Predictive Value): Precision is the ratio of true positive predictions to the total number of positive predictions made by the model. It quantifies the accuracy of the positive predictions.

- Recall (Sensitivity or True Positive Rate): Recall is the ratio of true positive predictions to the total number of actual positive instances. It measures the ability of the model to capture all positive instances.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

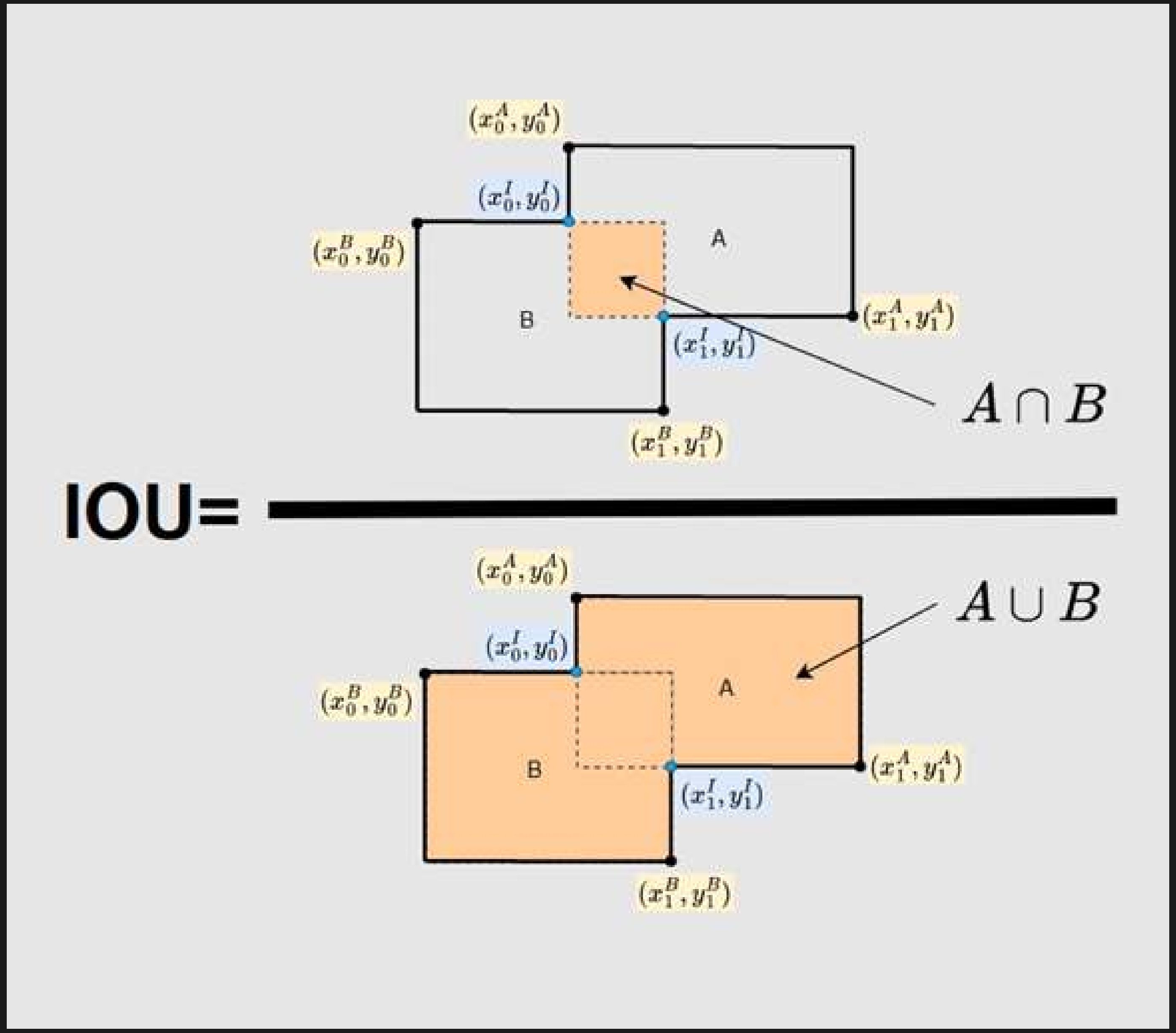


Key performance metrics

IoU (Intersection over Union):

IoU is a metric that measures the overlap between the predicted bounding box and the ground truth bounding box of an object. It is widely used in object detection tasks to evaluate the accuracy of localization. The IoU is calculated as the ratio of the area of intersection to the area of the union of the two bounding boxes.

$$IoU = \frac{Area\ of\ \ Union}{Area\ of\ \ Intersection}$$

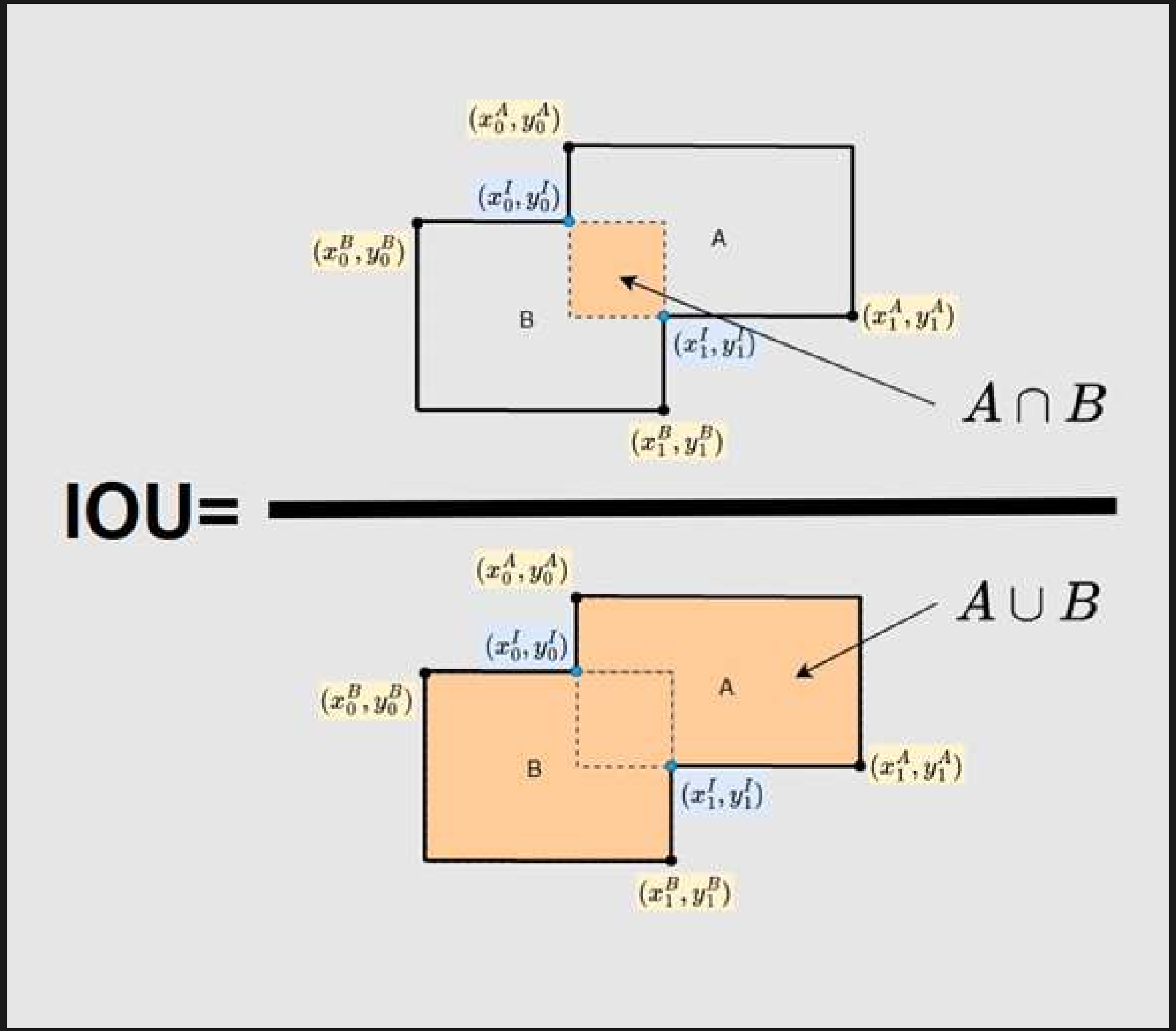


Key performance metrics

IoU (Intersection over Union):

Threshold for Accuracy: A threshold IoU value is often set to determine whether a prediction is accurate. If the IoU between the predicted and ground truth bounding boxes exceeds this threshold, the prediction is considered accurate; otherwise, it is treated as a misclassification.

$$IoU = \frac{Area\ of\ \ Union}{Area\ of\ \ Intersection}$$



Key performance metrics

Speed and Inference Time

Speed and inference time are important metrics, especially in our application that requires average inference times. These metrics focus on evaluating how well a model can make predictions.

$$Average\ Inf\,erence\ Time\ per\ Image = \frac{Total\ Inf\,erence\ Time}{Total\ Number\ of\ \ Image}$$

$$Number\ of\ \ Images\ Processed\ in\ a\ Fixed\ Time\ Interval = \frac{Total\ Number\ of\ \ Image}{Total\ Inf\,erence\ TIme}$$

Key performance metrics

F1-Score

In object detection and customer behavior monitoring, both precision and recall are crucial metrics. Precision measures the accuracy of positive predictions, while recall gauges the model's ability to capture positive instances comprehensively. The F1 score is the harmonic mean of both, creating a balanced metric

The F1 Score ranges from 0 to 1, where 1 indicates perfect precision and recall, and 0 indicates poor performance. It is particularly useful when dealing with imbalanced datasets

$$F1_{Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

7

TIMELINE & ROADMAP

- Present a timeline or roadmap for the development and deployment of the MVP
- Outline the major milestones and deliverables

Objectives

- Develop and deploy an MVP within 6 weeks.
- MVP meets the basic needs of users.
- MVP receives positive feedback from users and stakeholders.

Timeline

Phase	Iteration	Task	Start Date	End Date	Duration
Planning and Design	1	Define MVP requirements	6/11/2023	11/11/2023	5 days
		Identify key user stories and acceptance criteria	6/11/2023	10/11/2023	4 days
Development	2	Develop technical architecture	13/11/2023	18/11/2023	5 days
		Set up development environment	20/11/2023	21/11/2023	1 day
		Install necessary tools, libraries, and code editors	21/11/2023	22/11/2023	1 day
		Configure development environment for efficient collaboration	22/11/2023	23/11/2023	1 day
Development	3	Develop backend systems	20/11/2023	25/11/2023	5 days
		Implement database schema and data access layer	23/11/2023	24/11/2023	1 day
		Build API endpoints to handle user requests	24/11/2023	25/11/2023	1 day
		Integrate authentication mechanisms for secure user access	25/11/2023	27/11/2023	2 days
Development	4	Develop front-end UI	27/11/2023	02/12/2023	5 days
		Translate UI mockups into HTML, CSS, and JavaScript code	27/11/2023	29/11/2023	2 days
		Ensure responsive design for optimal viewing across devices	30/11/2023	01/12/2023	2 days
		Implement user interactions and data retrieval from backend	01/12/2023	02/12/2023	1 day
Testing and Deployment	5	Conduct user testing sessions to gather feedback on usability and satisfaction	04/12/2023	09/12/2023	5 days
		Analyze feedback to identify areas for improvement	04/12/2023	09/12/2023	5 days
		Make necessary adjustments	04/12/2023	09/12/2023	5 days
		Deploy MVP	11/12/2023	11/12/2023	1 day
		Prepare documentation	11/12/2023	16/12/2023	5 days
		Present MVP to judges and participants	17/12/2023	17/12/2023	1 day

Milestone

Milestone	Deliverables	Status
Planning and Design (Week 1)	MVP requirements document	Complete
	User experience design	Complete
	Technical architecture	Complete
Coding (Weeks 2-3)	Backend systems	In progress
	Front-end UI	In progress
	Integration of backend systems and front-end UI	Planned
Testing (Weeks 4-5)	Test cases	Planned
	Test results	Pending
	Any necessary adjustments to the MVP based on feedback from users and other stakeholders	Pending
Deployment (Week 6)	Deployed MVP	Pending
	Documentation	Planned
	Hackathon day presentation	Planned
	Feedback on the MVP	Pending

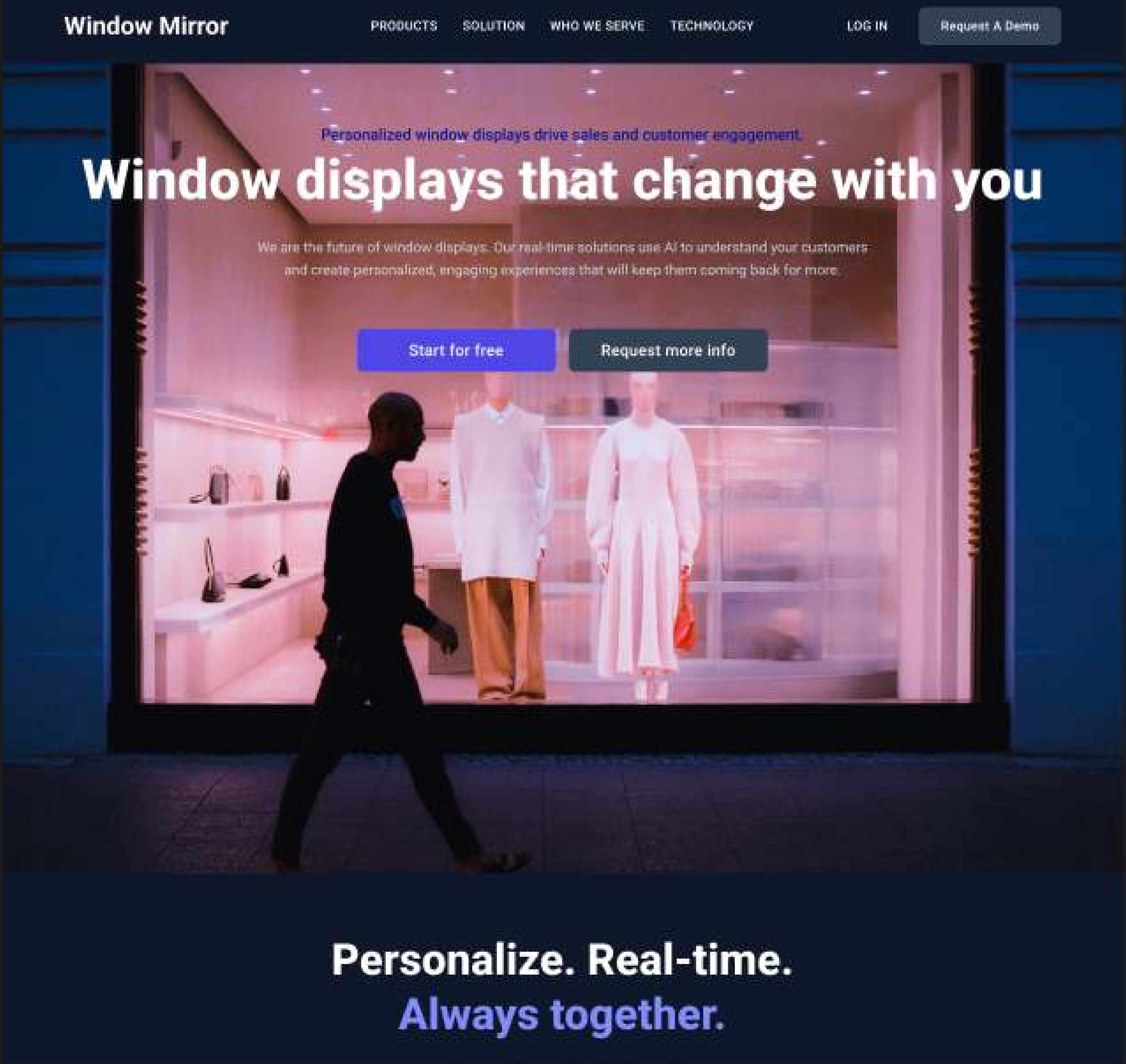
8

USER INTERFACE

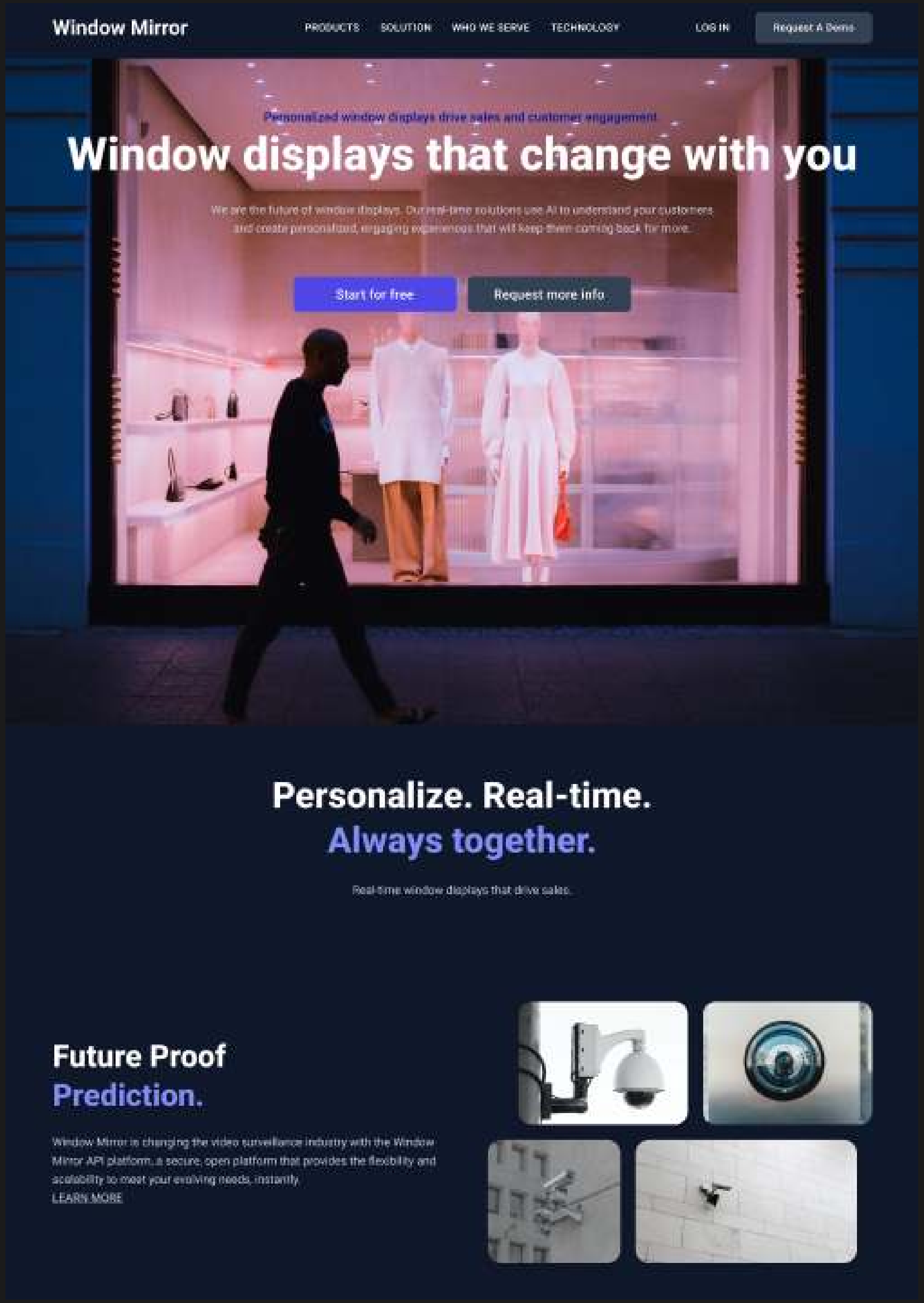
- ***Describe the user interface*** or interaction components of the MVP
- ***Explain how users will interac***t with the AI-powered features

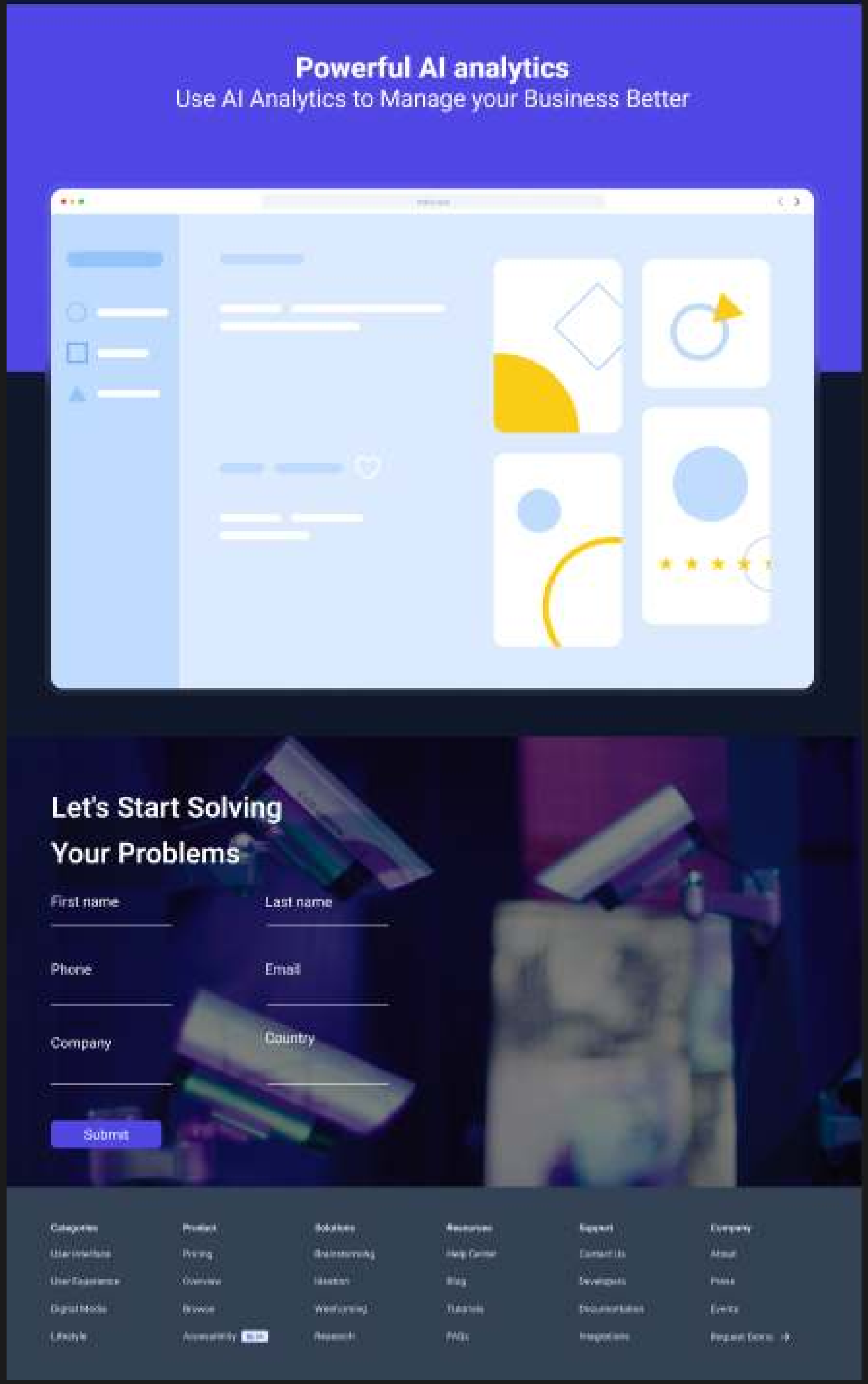
User Interface

Link to the demo UI

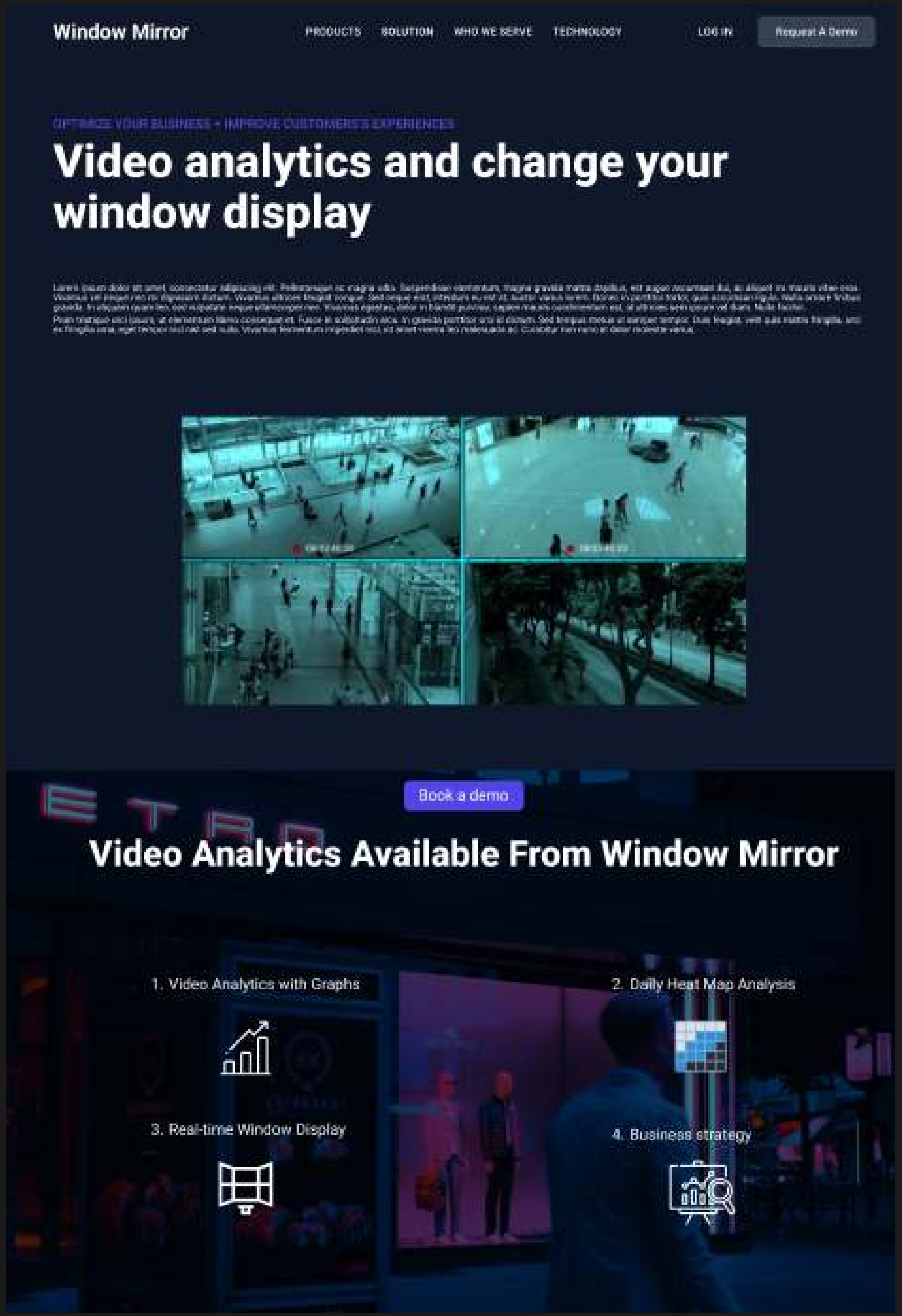


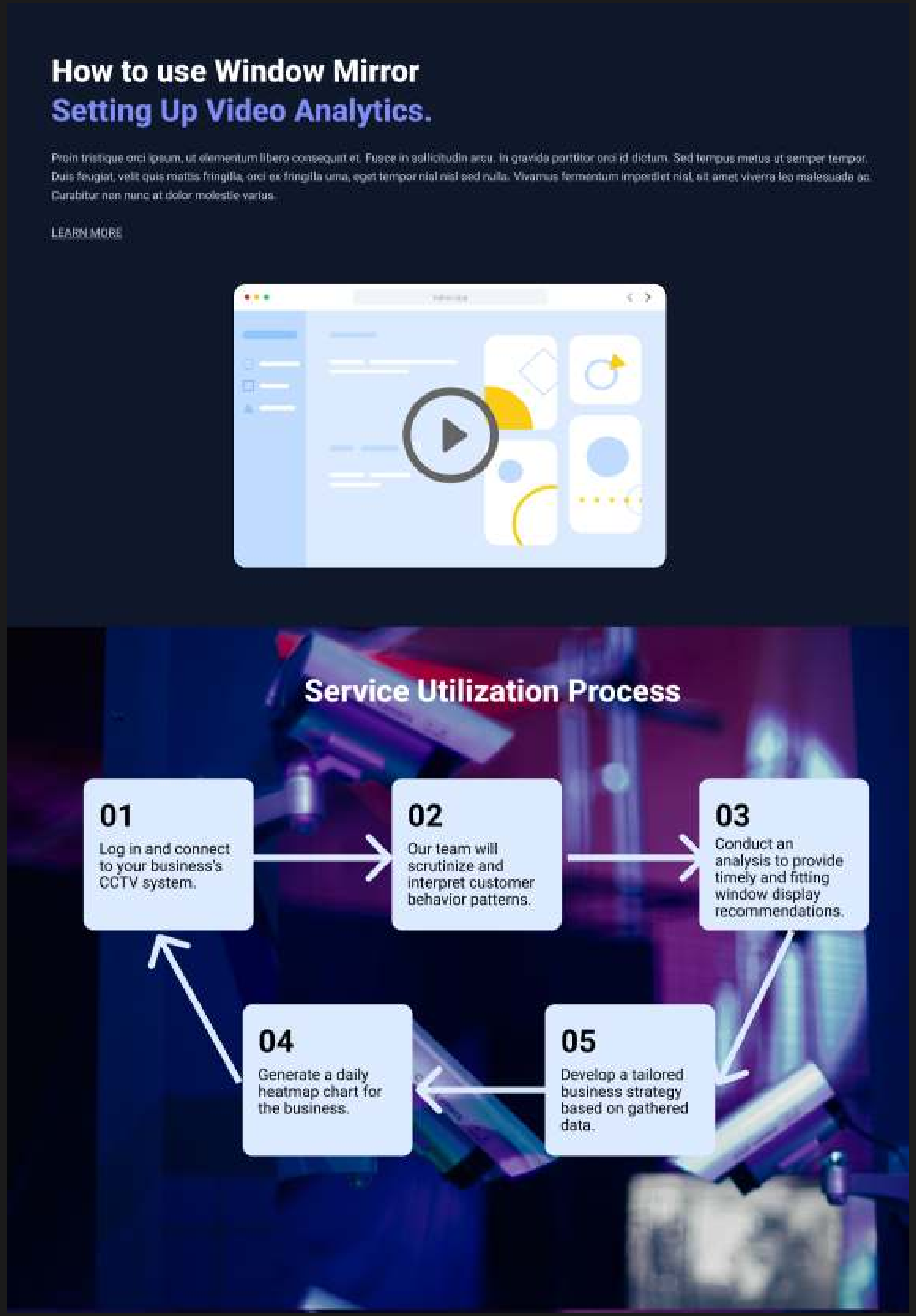
Landing page

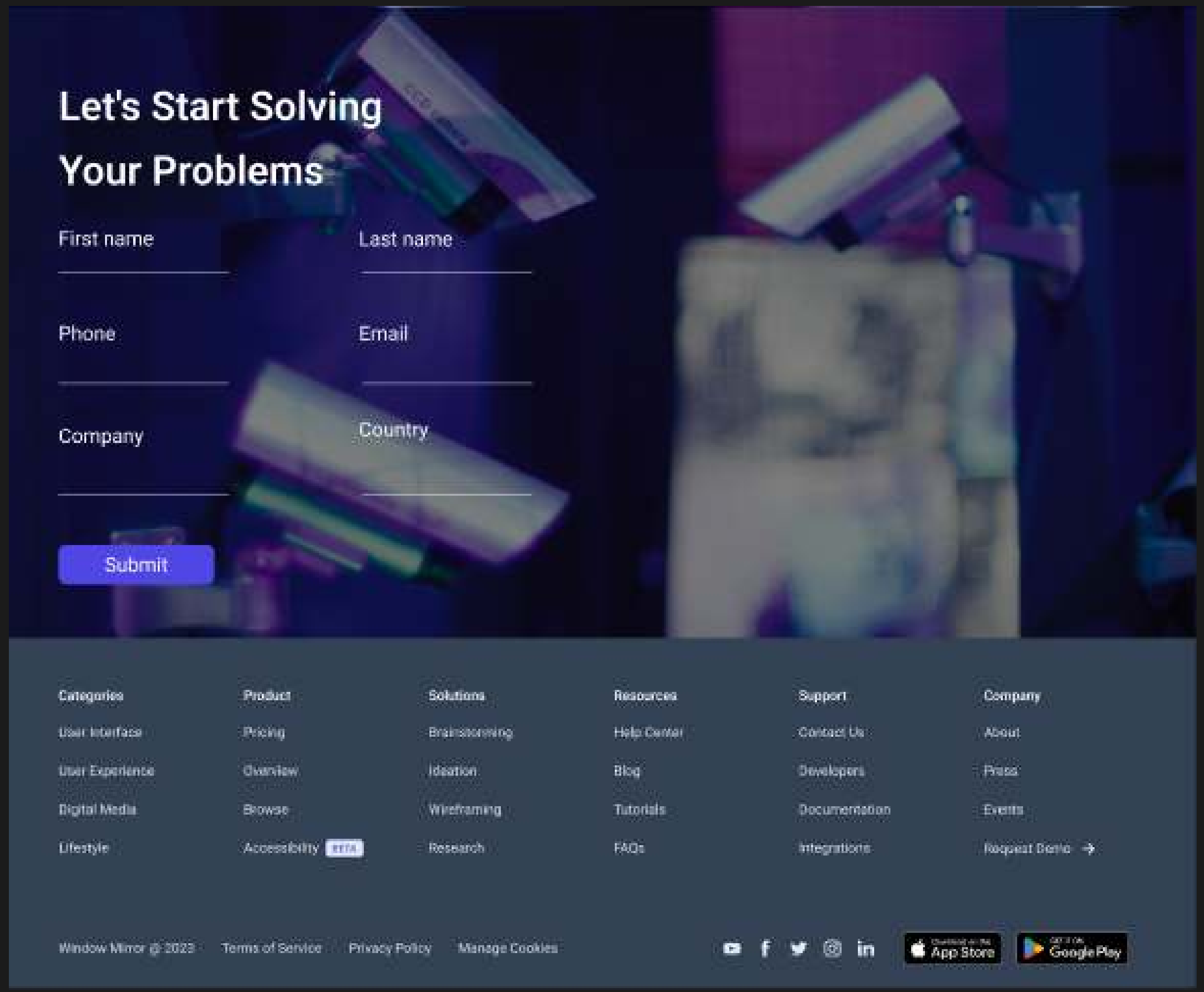




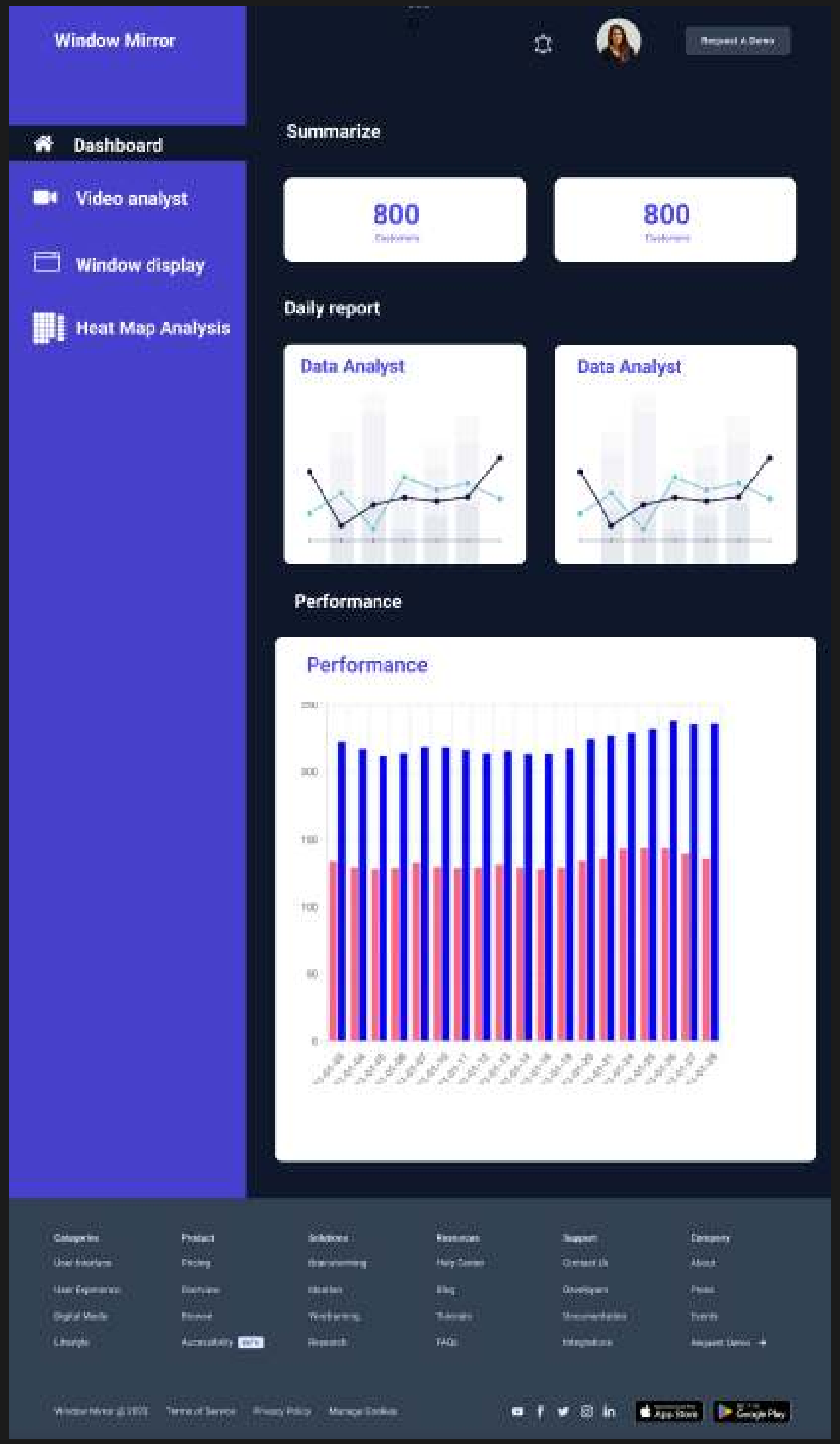
Video analytics



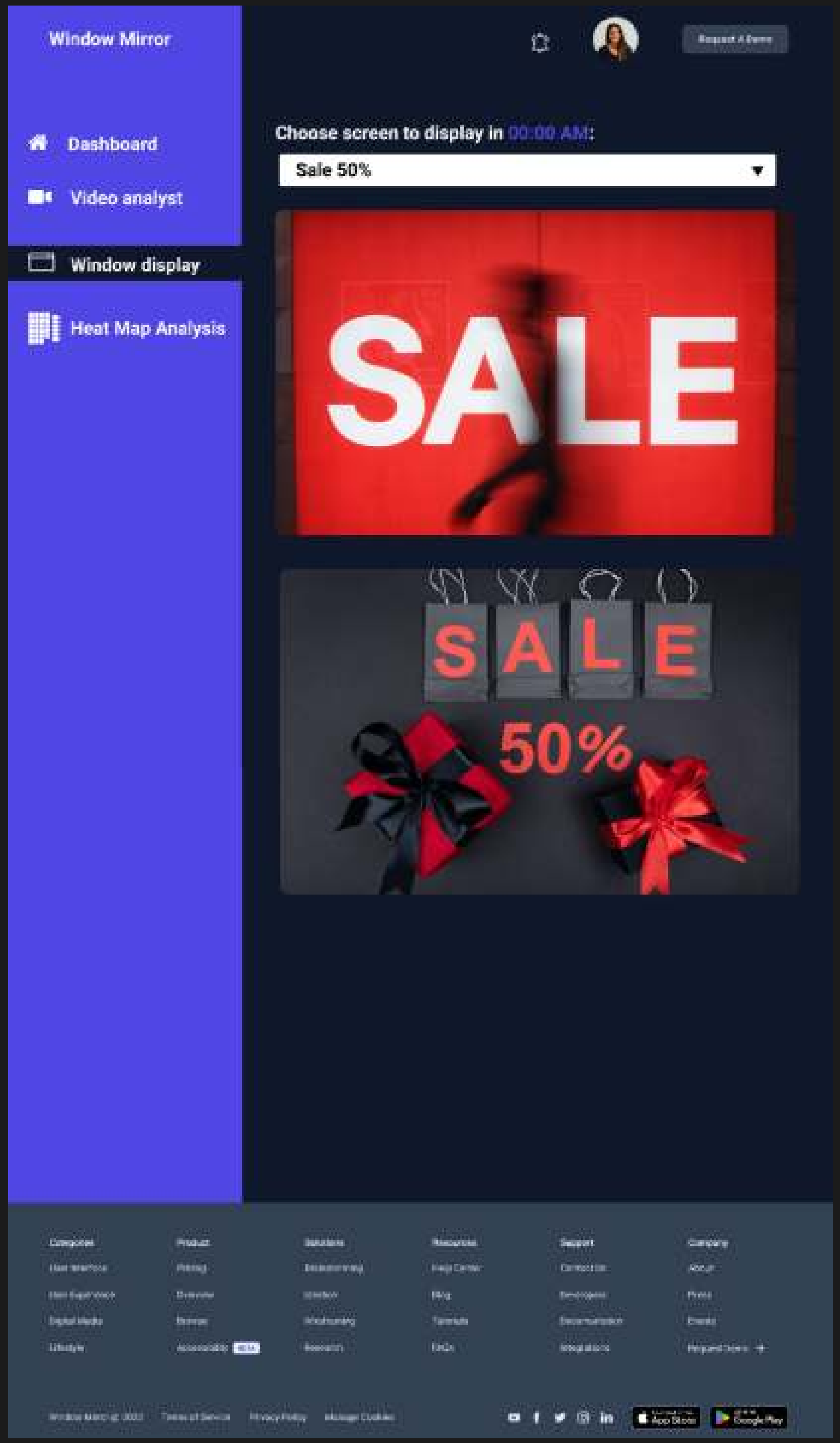




Service page



Service page



9

LIMITATIONS & FUTURE ENHANCEMENTS

- Acknowledge any ***limitations*** or ***contraints*** of the MVP
- Discuss ***potential future enhancements*** of additional features that could be incorporated

Limitations

Funding

- The lack of funding lead to an unstable and poor performance database system
- It’s also a burden on maintaining the server to keep our website running

Not A All-Round Solution

- Our solution can only do a couple of things in it’s capability
- Still require the users to manually do tasks beforehand

Security

- To use our solution effectively, users have to connect to the CCTV in front of their store through the IP address and that alone would cause a lot of paperwork with the mall. Even if our customers are not mall-based retailers there will still be a concern about the IP address of their camera leaked out

→ With some limitations here and there, our solution might not be perfect. But that also leave some room for future enhancements

Future enhancements

A Proper Funding

- With a better funding, the problem of database and server is no longer a burden to our solution
- This also help us to expand our team

Trying To Go For The “Full-Stack” Route

- With the expansion of our team, we can have some marketing experts, business analysts to help with the process of creating advertisements or analyze the performance of the product.

Security

- Constantly update the newest encryption method to keep the IP address of the CCTV safe
- Launch our own brand of CCTV that comes compatible with an app version of our solution.

10

CONCLUSION

- Summarize the ***key points*** and reiterate the ***value*** proposition of the MVP
- Highlight the ***potential impact*** and ***benefits*** of the AI solution

Conclusion

- With the implementation of Faster-RCNN, Deep neural network-based model to ***detect and track people walking*** by the store in order to ***dynamically display an advertisement*** based on customers behavior.
- Through the process of coming up with this solution we hope to bring a better way to ***enhance the efficiency of advertisements*** and ***boost the profit*** of our targeted users, which are small and medium fashion retailers.
- ***The knowledge gained from this project sets the stage for future solutions that try to push the boundaries of AI capabilities and bring positive impacts to society.***