A REVIEW OF ACTION ANTICIPATION IN HUMAN-ROBOT COLLABORATION

Pedro Amaral nMec: 93283

Seminário - MRSI

INTRODUCTION

The concept of Human-Robot Collaboration (HRC) involves the research of mechanisms that allow humans and robots to work together to achieve a shared goal.

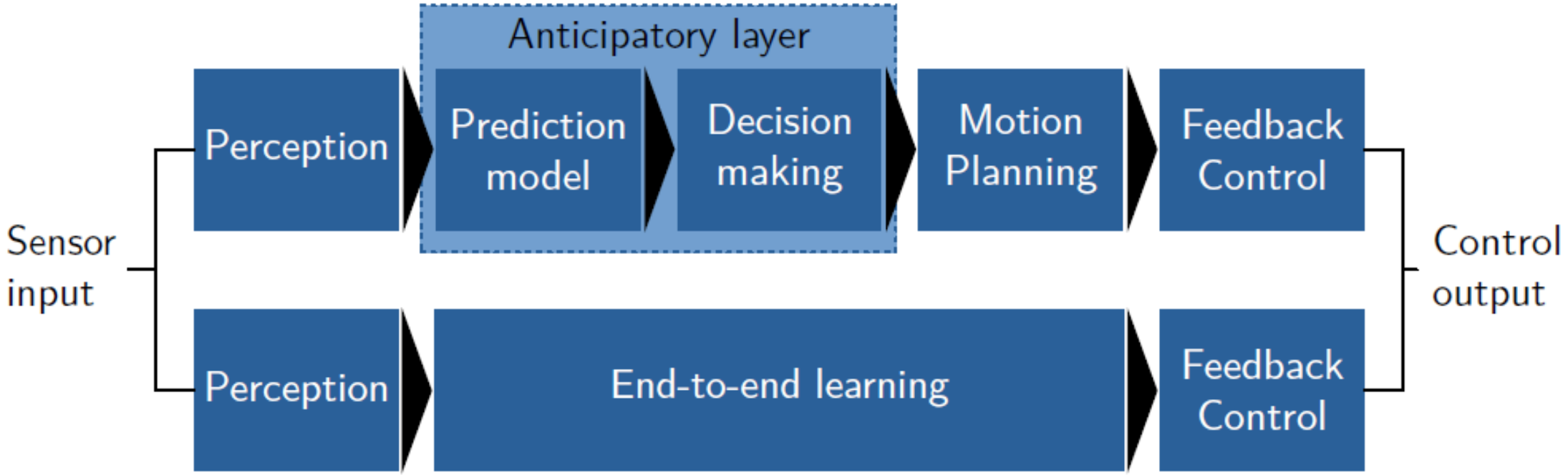
The presented article aims at reviewing previous work relevant to the topic of action anticipation to enhance human-robot collaboration in industrial settings.

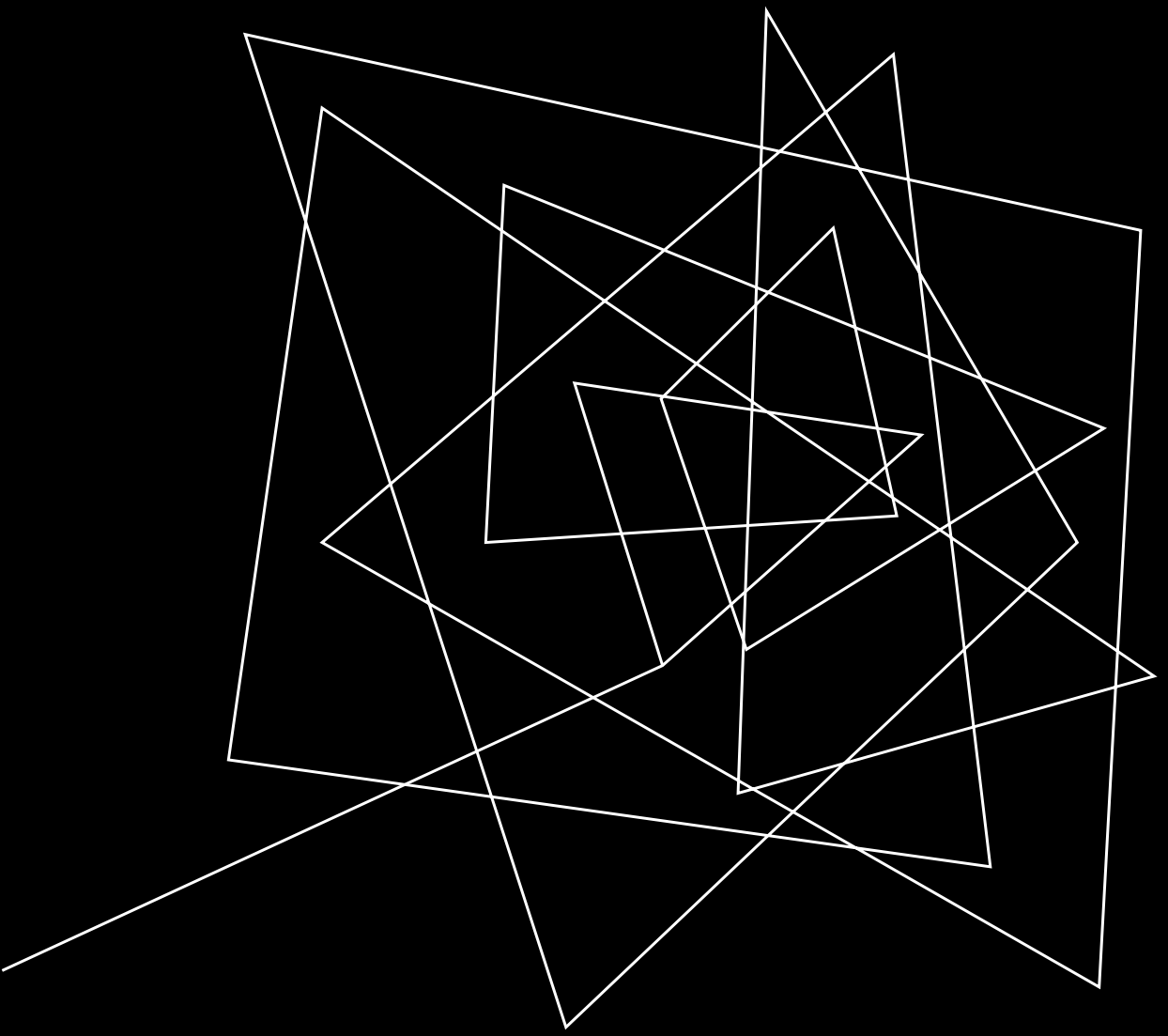
WHAT IS ANTICIPATION?

“An anticipatory system is a system containing a predictive model of itself and/or its environment, which allows it to change state at an instant in accord with the model’s predictions pertaining to a later instant.”

Robert Rosen, Anticipatory Systems: Philosophical, Mathematical and Methodological Foundations., 1985

HOW TO IMPLEMENT ANTICIPATION?





COLLABORATIVE ROBOTICS

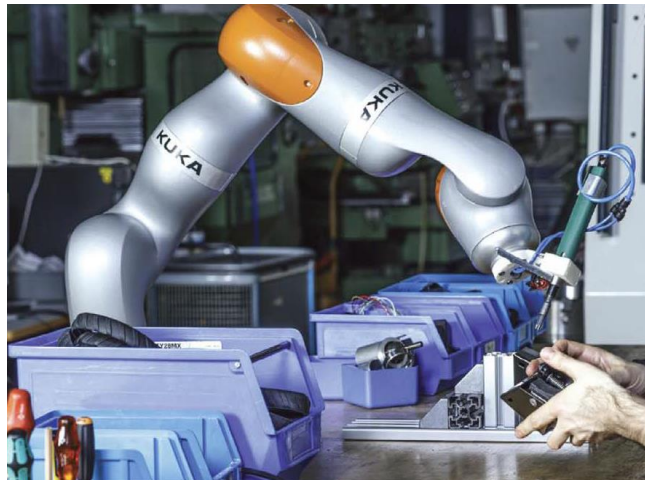
Collaborative Robots and Safety in Collaboration

COLLABORATIVE ROBOTS

- Compact
- Easy to install
- Easy to program
- Flexible
- Mobile
- Consistent
- Precise
- Take care of the most monotonous and dangerous actions
- Reduce the production cost

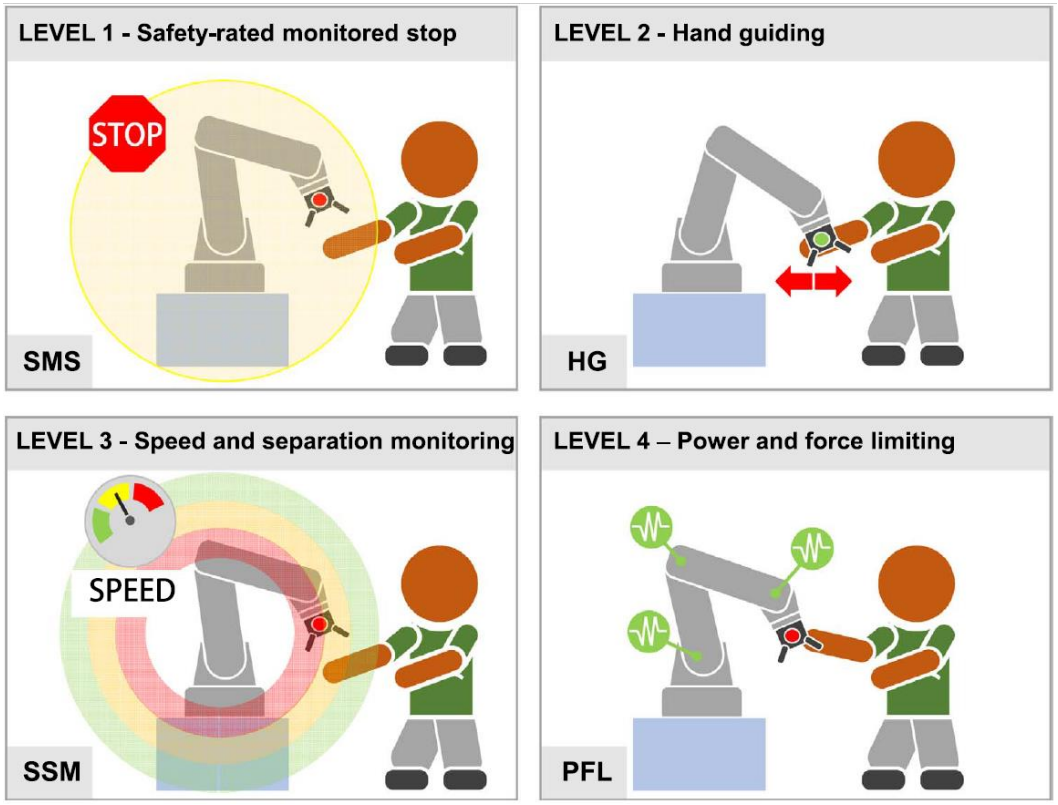
SAFETY IN COLLABORATION

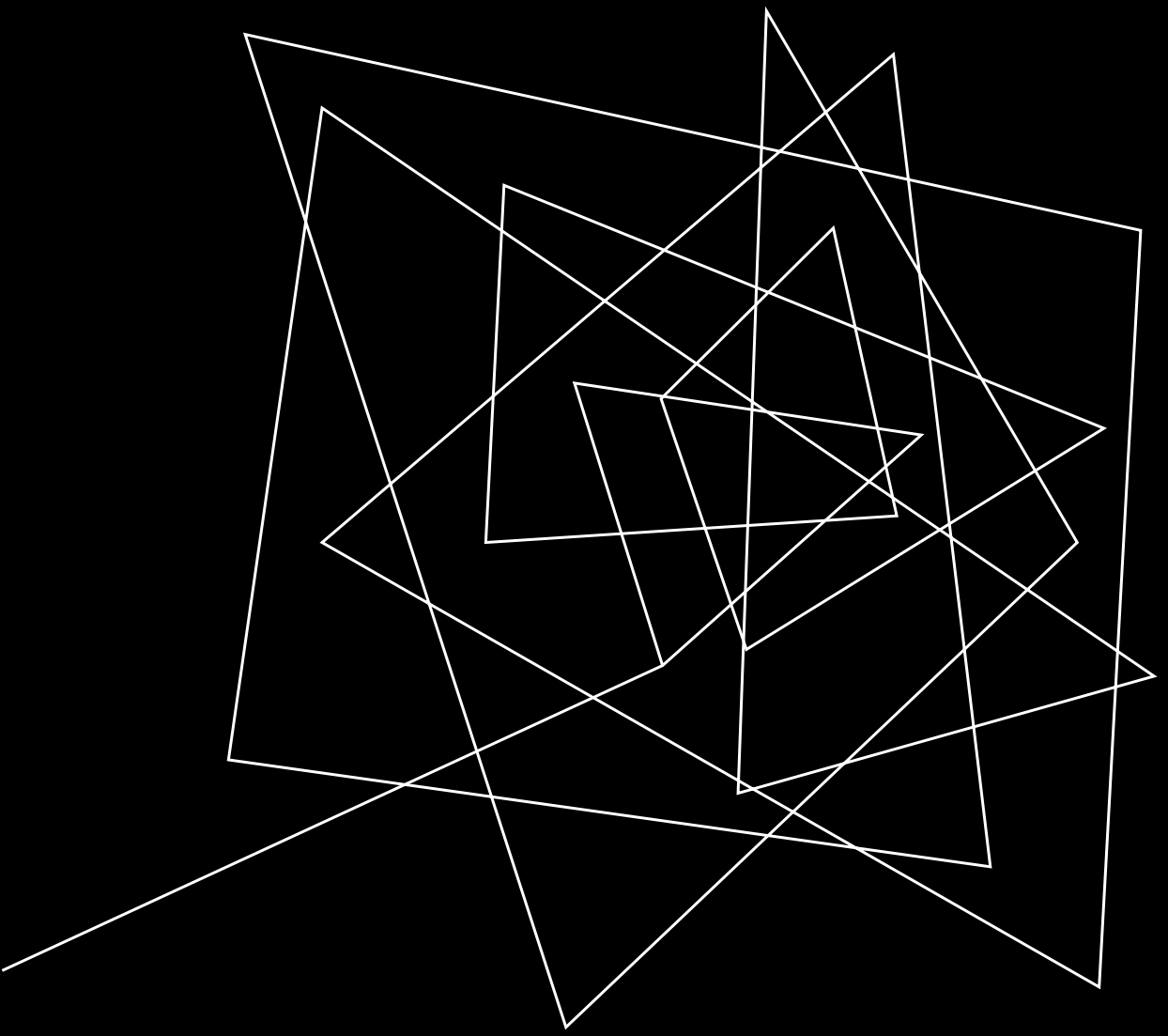
- Safety is a critical topic in Human-Robot Collaboration;
- Traditional robots are in an isolated workspace so that they do not injure humans;
- Collaborative robots need to be able to work with humans without injuring them;
- ISO 10218-1 and 10218-2 are two norms created for this.



ISO 10218-1 AND 10218-2 NORMS

1. Safety-rated monitored stop: when a human enters the cobot's workspace, it completely stops;
2. Hand guiding: when an operator manually moves the cobot, it is compliant;
3. Speed and separation monitoring: as the human moves closer to the cobot, it becomes gradually slower;
4. Power and force limiting: the cobot has its operation restricted in terms of force and torque.





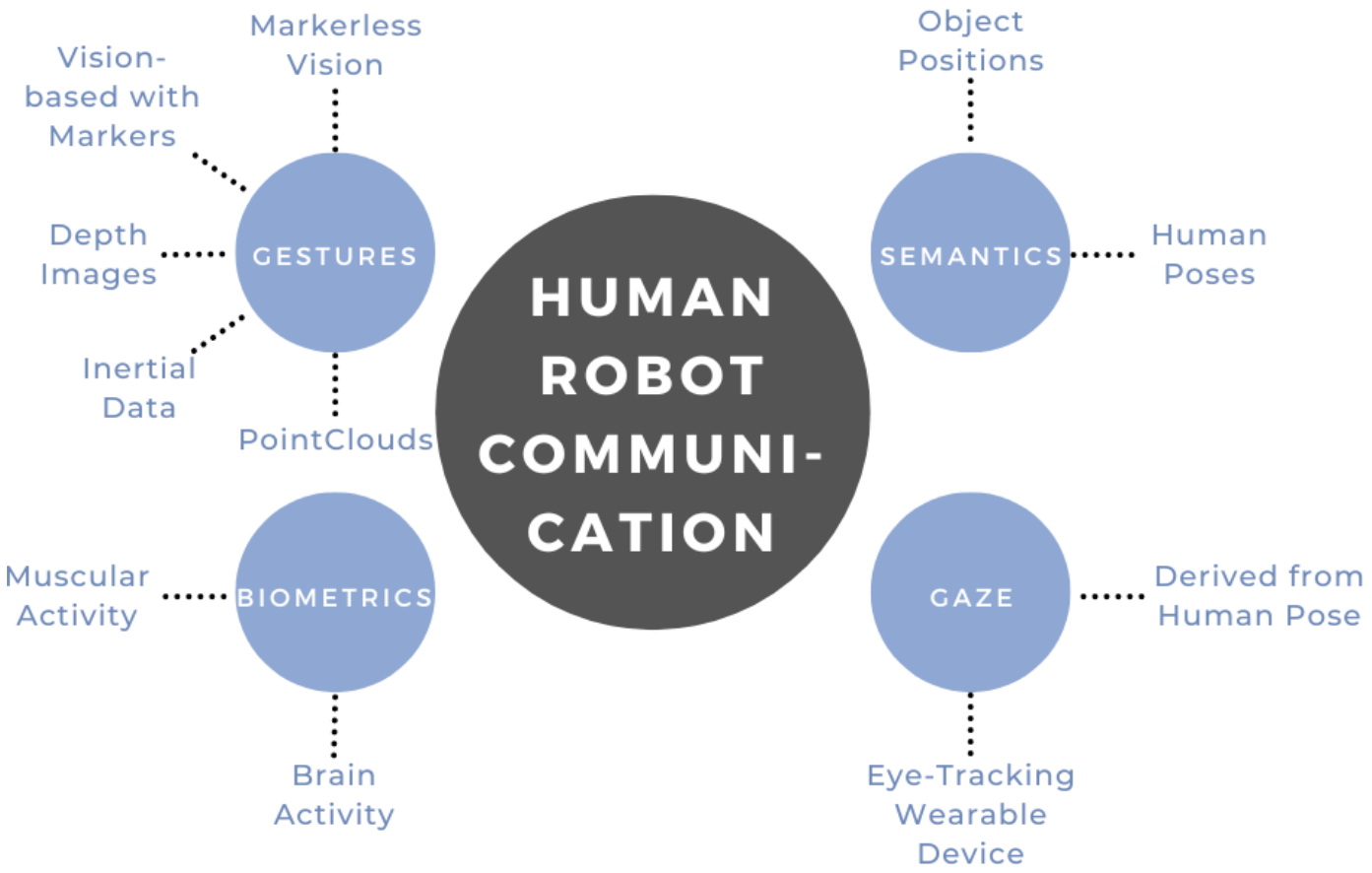
DATA SOURCES AND SENSORS

Perception Block

WHICH DATA SHOULD WE CAPTURE FROM THE ENVIRONMENT?

- Humans and robots can communicate through several methods;
- In action anticipation, passive and indirect communication methods are used;
- The user should not need to do anything for the robot to act, the robot must be able to understand the worker's body language, such as his involuntary pose, gestures or gaze.

DATA SOURCES

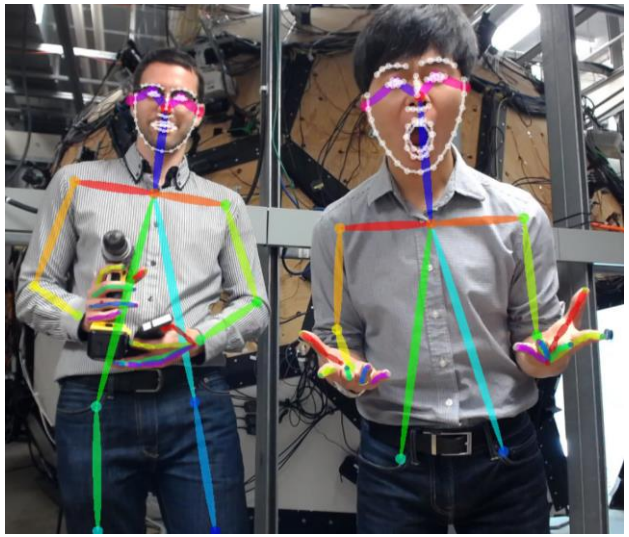


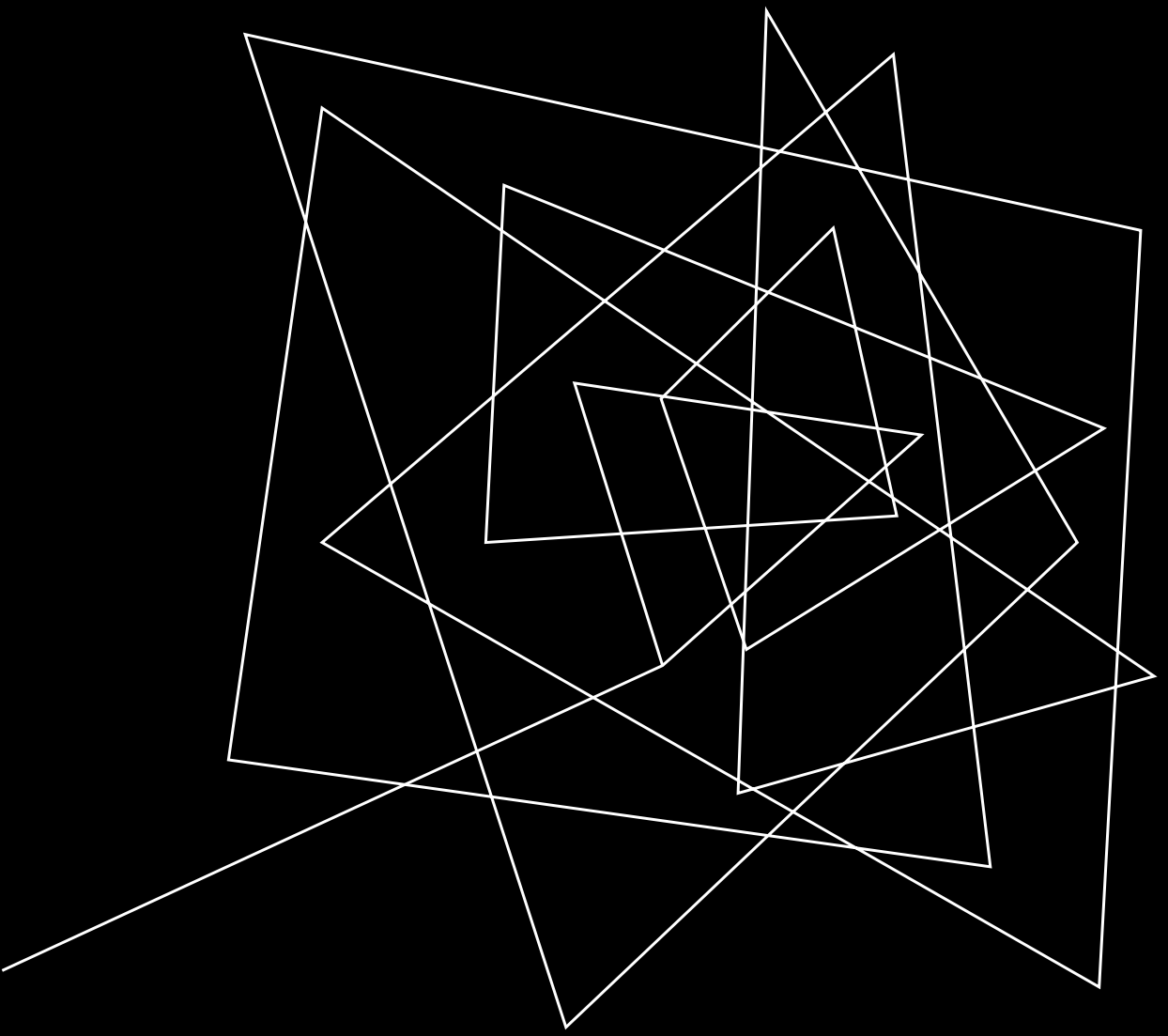
SENSORS

Sensor	Nº Occurrences
RGB Camera	9
RGB-D Camera	1
Inertial Measurement Unit (IMU)	1
Electromyography (EMG) Sensors	1
Wearable Gaze Detector	1

RGB Images can be used in different ways:

- Directly in the model;
- Process optical flow;
- Process marker positions;
- Input to key point detection frameworks.

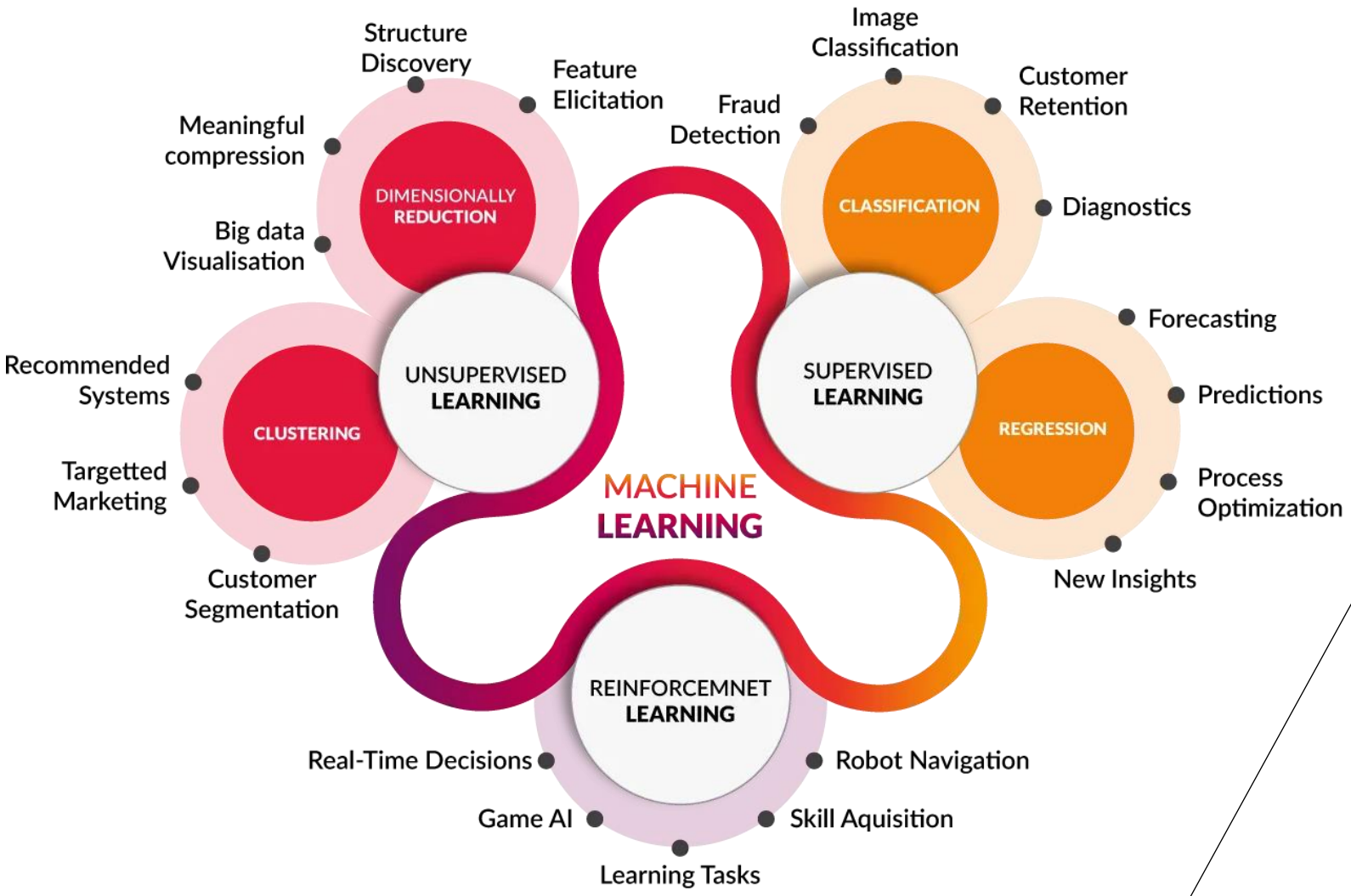




PREDICTIVE MODELS

Prediction Model Block

TYPES OF MACHINE LEARNING



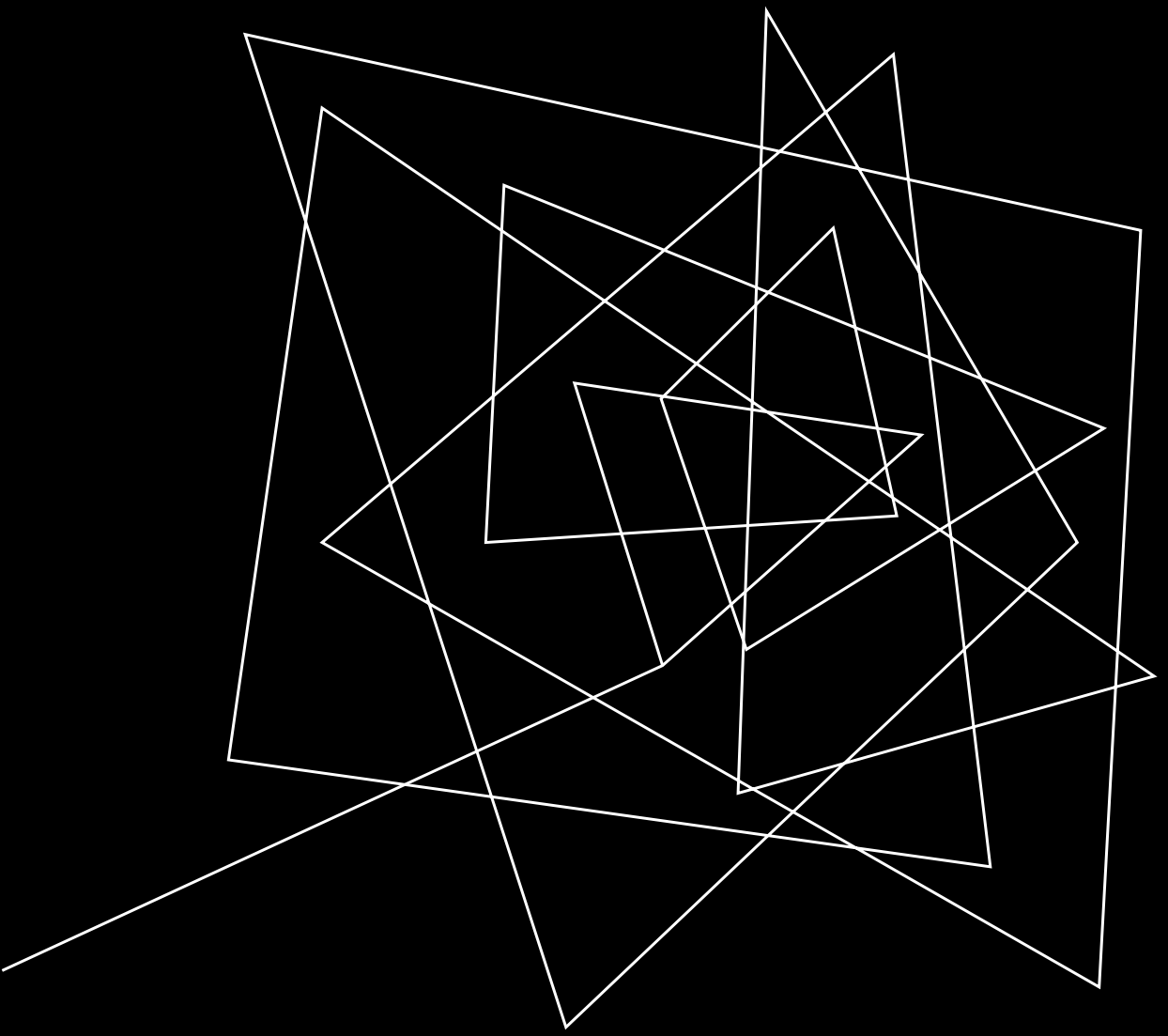
PREDICTIVE MODELS

- Predicting the next action of the worker can be represented as a classification problem;
- It is possible to use a sequence of images that must be classified as a particular future action class.

PREDICTIVE MODELS

Model Name	Nº Occurrences
LSTM	7
CNN	3
TSN	1
Nearest Neighbor	1
SVM	1

Model Name	Nº Occurrences
ResNet-34	1
ResNet-50	1
VGG-16	1
TS	1
ConvNet	1



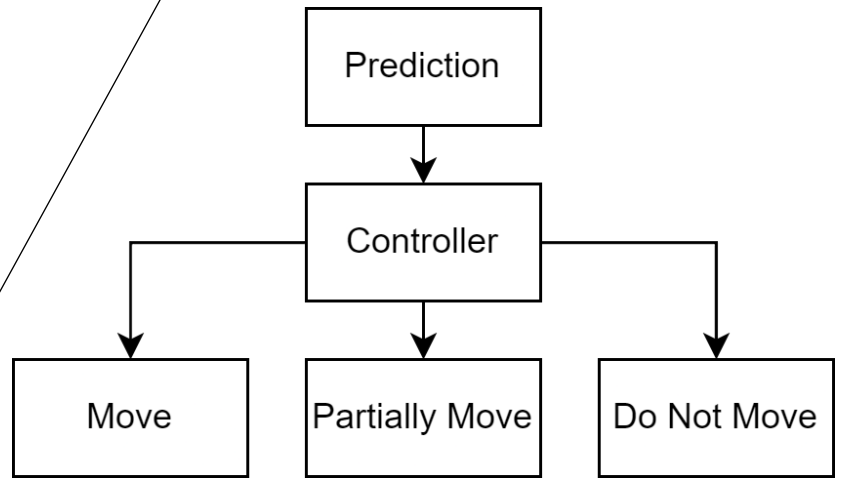
FROM PREDICTION TO PLANNING

Decision Making and Motion Planning Blocks

DECISION-MAKING

Two options were found to decide when to move:

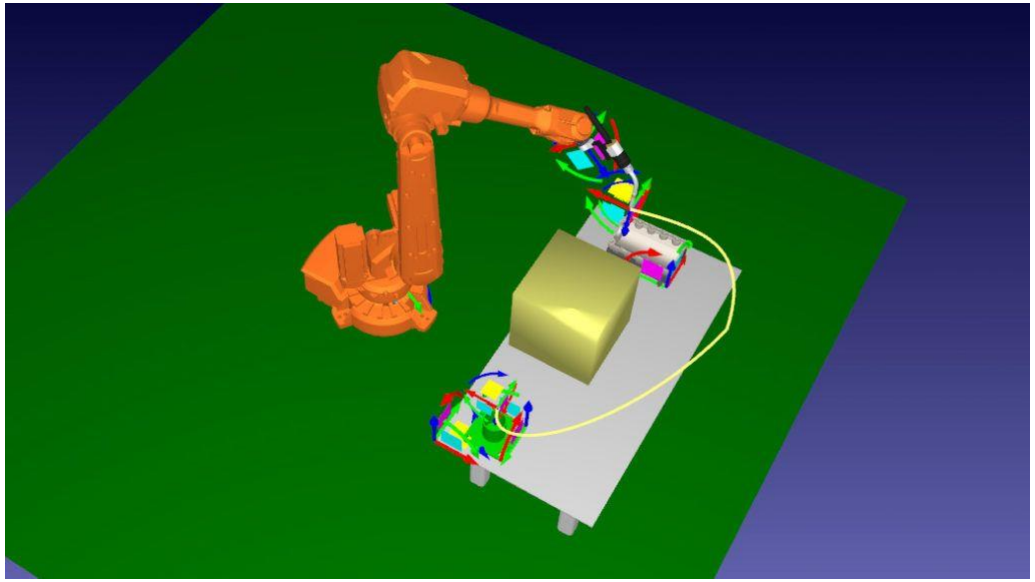
- Wait until a certain level of certainty about the prediction to make the needed action;
- As soon as it makes a prediction with a certain certainty threshold start slowly moving towards the target, do the entire action after it reaches a second threshold.

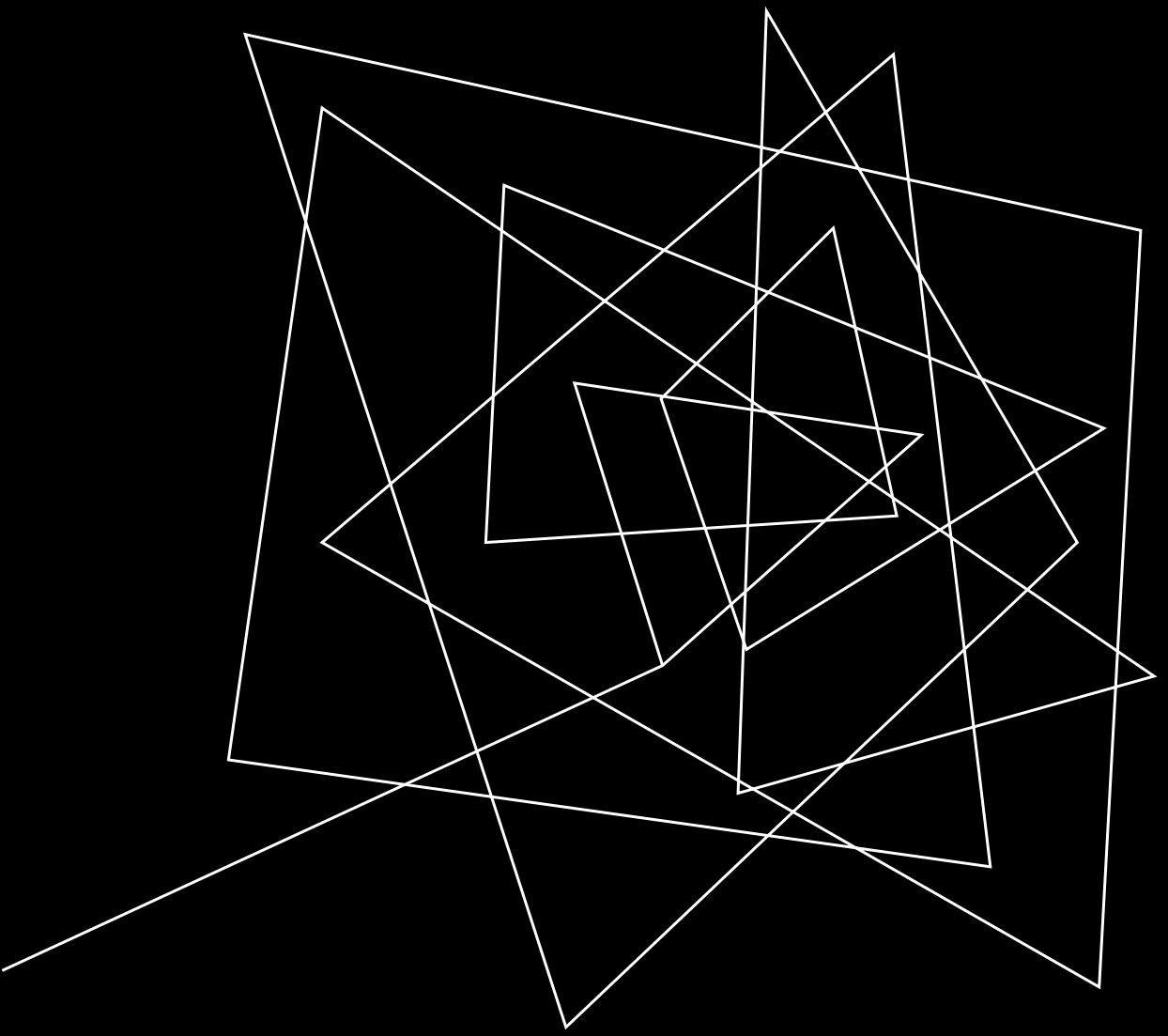


MOTION PLANNING

Two options were found to decide when to move:

- Use ROS Open Motion Planning Library (OMPL) and MoveIt! to handle the trajectory planning jobs;
- Plan trajectories considering previous poses when doing this interaction to avoid collisions.

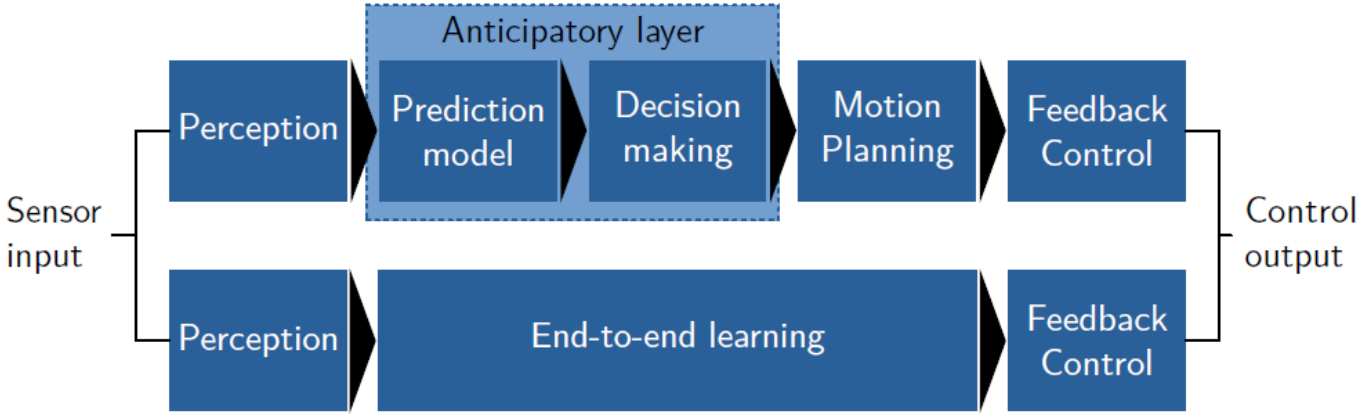


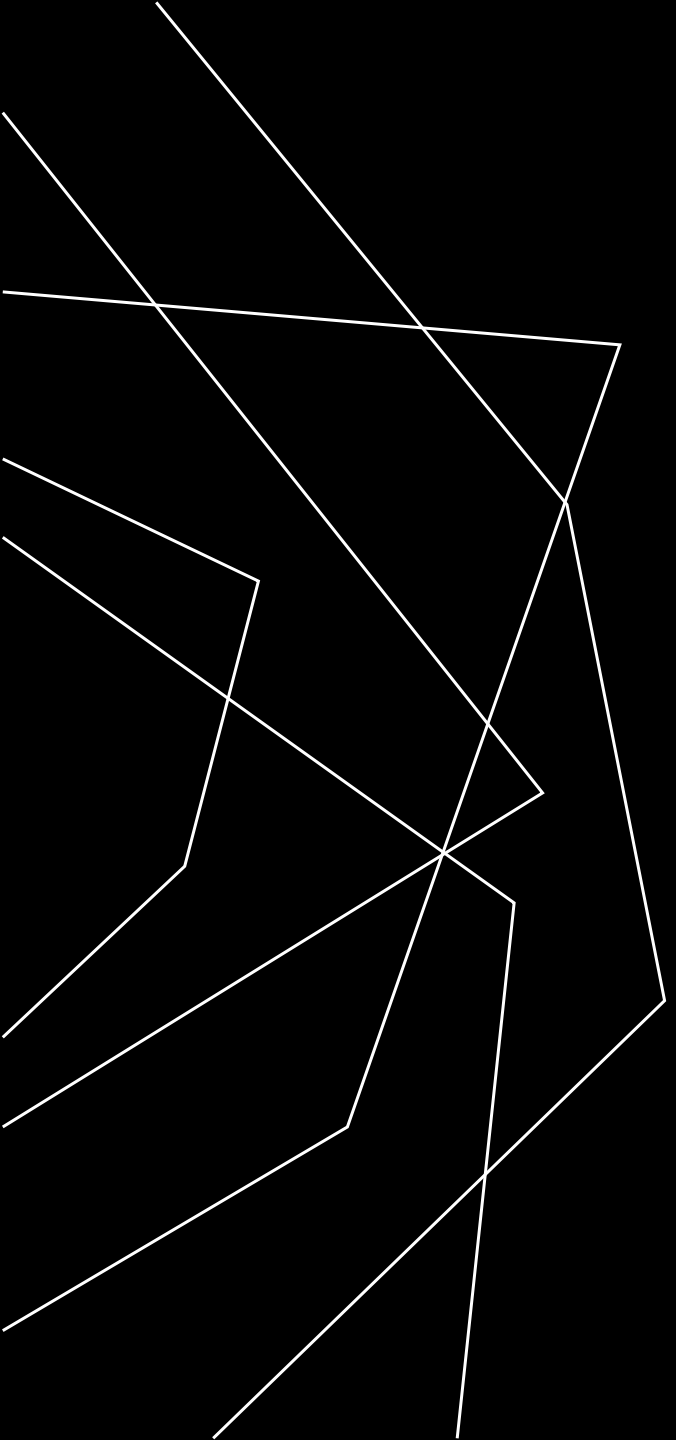


SUMMARY

SUMMARY

- There was a clear predominance of RGB cameras for perception;
- LSTMs and CNNs were the most common models;
- There were not many options for decision-making and motion planning;
- There was not work found for an end-to-end learning approach.





THANK YOU