Neural-based navigation system for 4x4 Jaguar Platform

Thesis Adivsors: Prof. Dr. Ing. Corneliu BURILEANU,
As. Univ. Drd. Ing. Ana Neacșu
Student: Alexandru-Mădălin Costea

Facultatea de Electornică, Telecomunicaţii şi
Tehnologia Informaţiei

July 2020

SpeeD
Speech & Dialogue Research Laboratory

Schedule

SpeeD
Speech & Dialogue Research Laboratory

Overview

Purpose

The purpose of the project is to $\rightsquigarrow$ create an autonomous navigation system with a top-down view, using neural networks.

SpeeD
Speech & Dialogue Research Laboratory

Objectives

1. Create a dataset;

Objectives

1. Create a dataset;
2. Train and compare different neural networks for object detection;

SpeeD
Speech & Dialogue Research Laboratory

Objectives

1. Create a dataset;
2. Train and compare different neural networks for object detection;
3. Implement a pathfinding algorithm;

SpeeD
Speech & Dialogue Research Laboratory

Objectives

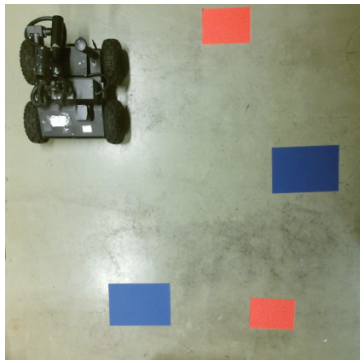
1. Create a dataset;
2. Train and compare different neural networks for object detection;
3. Implement a pathfinding algorithm;
4. Control the 4×4 Jaguar platform to move according to the output of the algorithm.

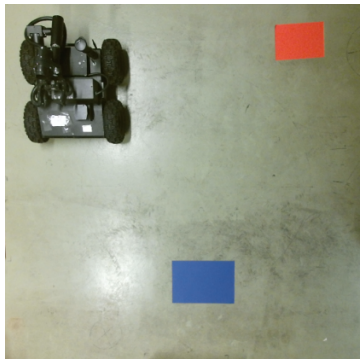
SpeeD
Speech & Dialogue Research Laboratory

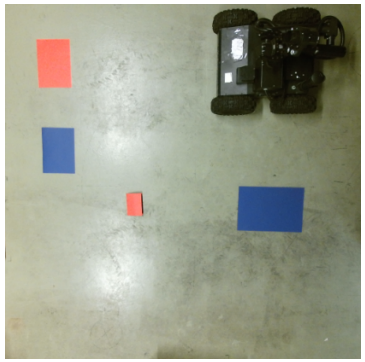
Image Dataset

Acquisition details:

- Total no. of images: 100
- Classes:
  - Jaguar – Robot
  - Red – Obstacle
  - Blue – Pass-through object







SpeeD
Speech & Dialogue Research Laboratory

General Overview

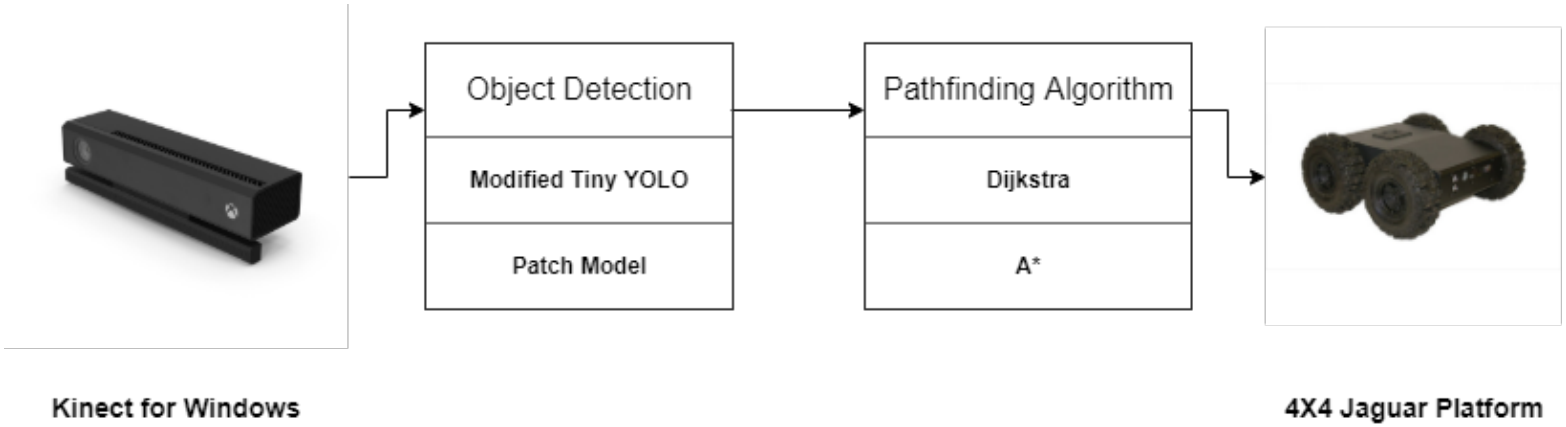


Figure: General Overview of the System

Detection algorithm

Tiny Yolo NN

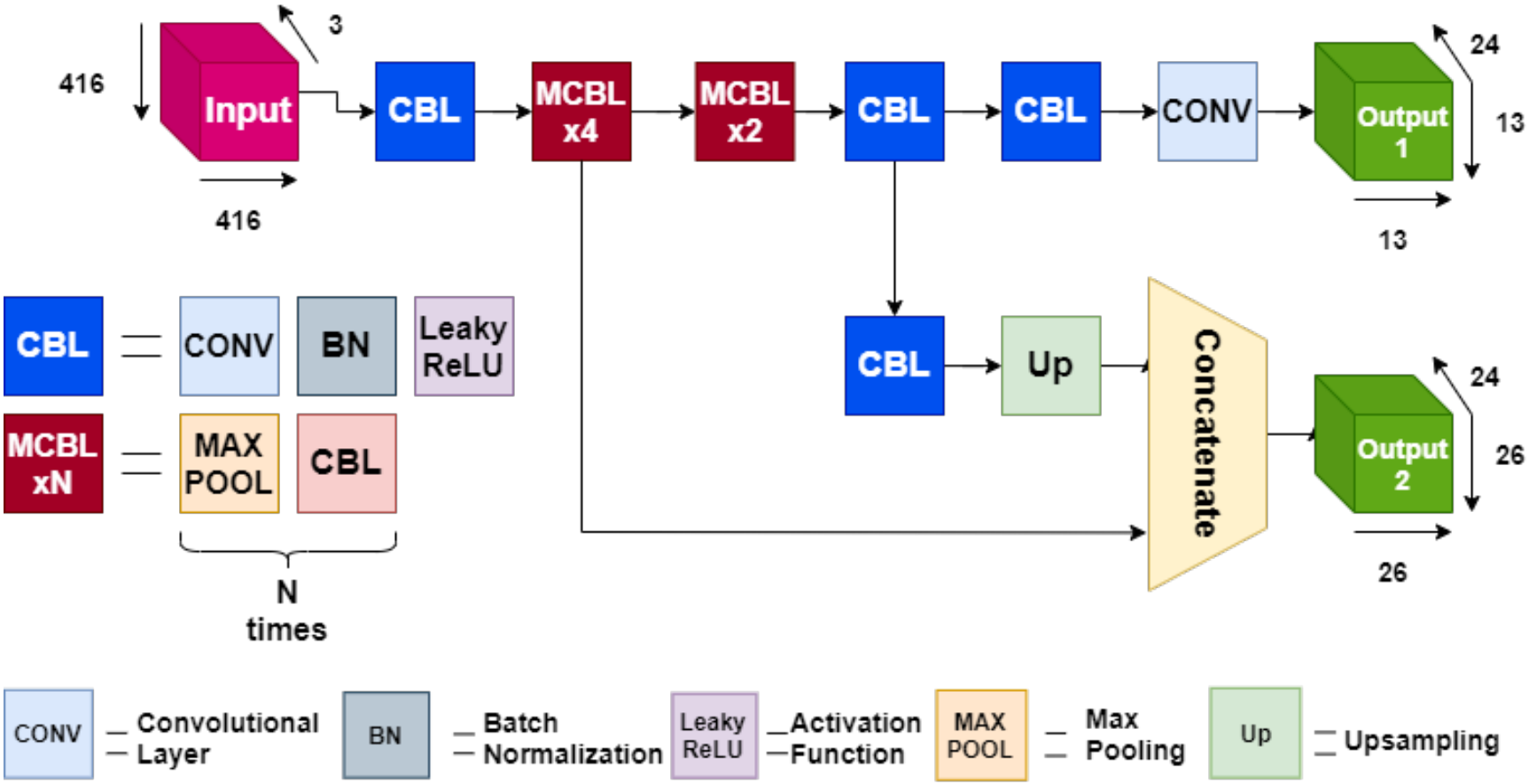
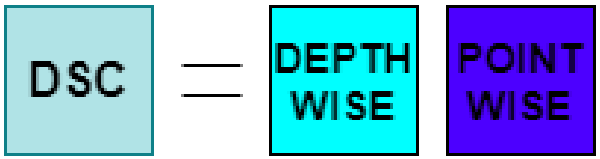
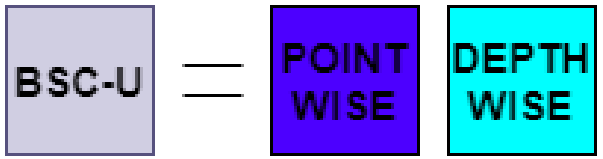


Figure: Tiny Yolo Architecture

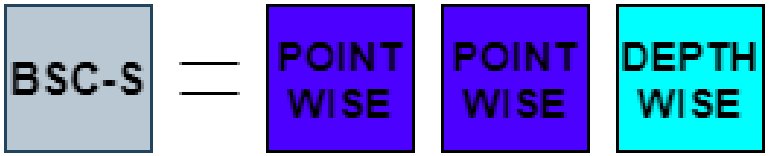
Separable Convolutions



(a) Depthwise Separable Convolution



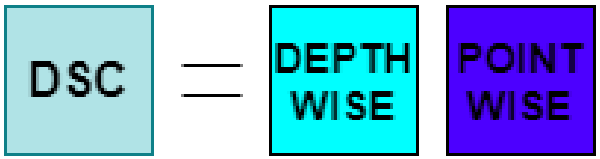
(b) Unconstrained Blueprint Separable Convolution



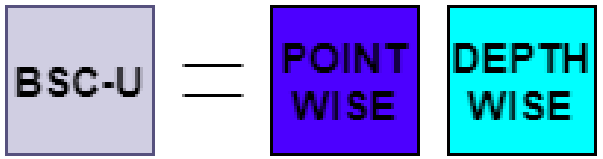
(c) Subspace Blueprint Separable Convolution

SpeeD
Speech & Dialogue Research Laboratory

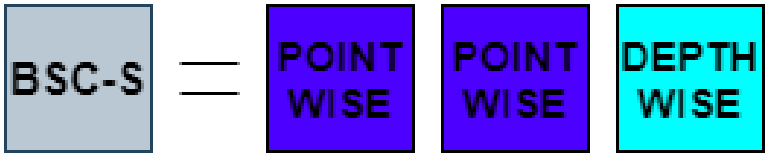
Separable Convolutions



(a) Depthwise Separable Convolution



(b) Unconstrained Blueprint Separable Convolution



(c) Subspace Blueprint Separable Convolution

Separable convolutions split standard convolutional layers $\rightsquigarrow$ same output, reduced computational cost [4].

SpeeD
Speech & Dialogue Research Laboratory

Subspace Blueprint Separable Convolution

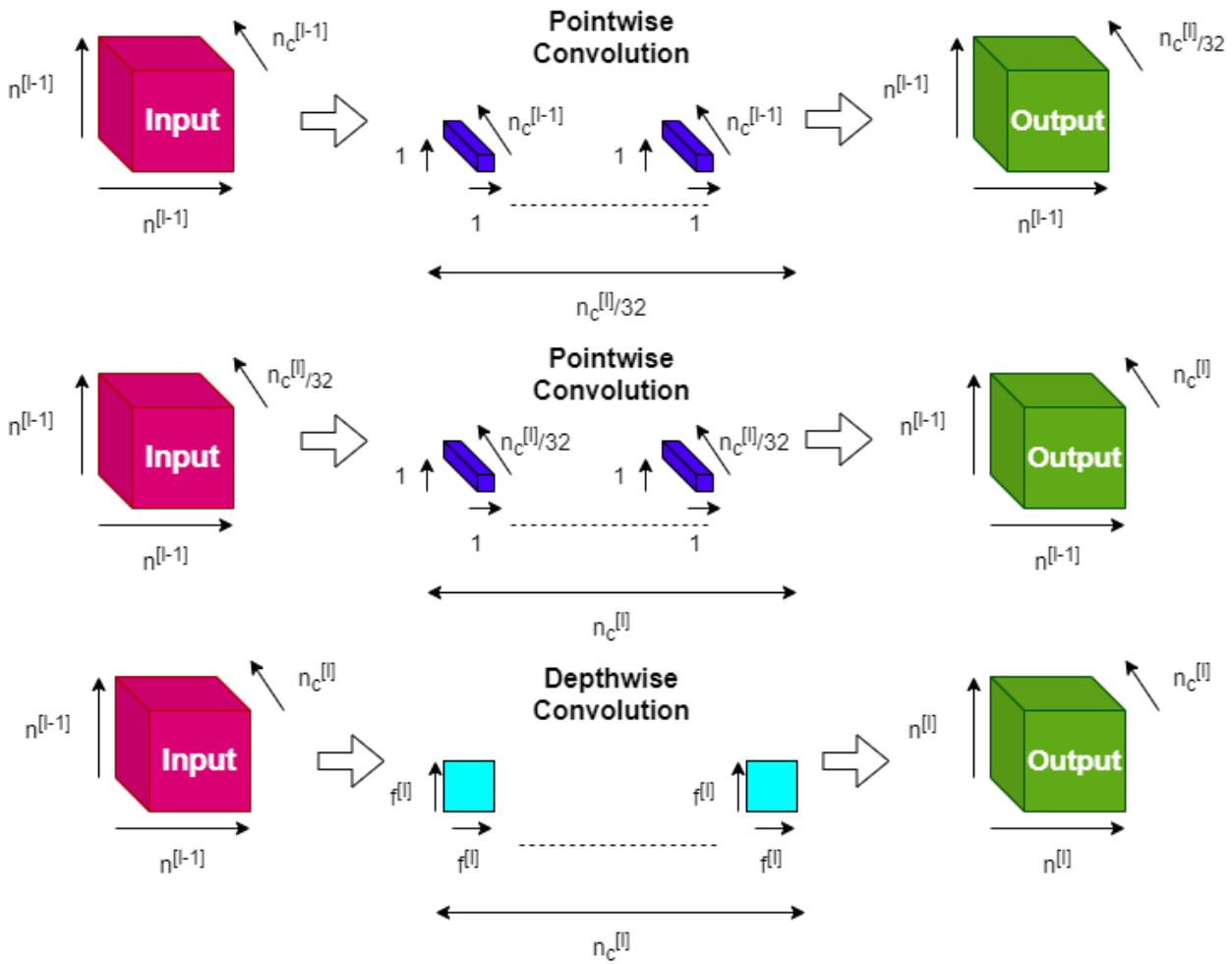


Figure: BSC-S Operations

Computational Cost

Standard Computational Cost:

$$kernel\_size \times input\_depth \times output\_depth \times output\_size$$

Computational Cost

Standard Computational Cost:

$$kernel\_size \times input\_depth \times output\_depth \times output\_size$$

BSC-S Reduction in Computational Cost:

$$\frac{input\_size}{kernel\_size \times output\_size \times \alpha} + \frac{input\_size \times \frac{output\_depth}{\alpha}}{kernel\_size \times output\_size \times input\_depth} + \frac{1}{input\_depth}$$

SpeeD
Speech & Dialogue Research Laboratory

Computational Cost Example

BSC-S Reduction in Computational Cost:

$$\frac{input\_size}{kernel\_size \times output\_size \times \alpha} + \frac{input\_size \times \frac{output\_depth}{\alpha}}{kernel\_size \times output\_size \times input\_depth} + \frac{1}{input\_depth}$$

Computational Cost Example

BSC-S Reduction in Computational Cost:

$$\frac{input\_size}{kernel\_size \times output\_size \times \alpha} + \frac{input\_size \times \frac{output\_depth}{\alpha}}{kernel\_size \times output\_size \times input\_depth} + \frac{1}{input\_depth}$$

Example:

$$\frac{416 \times 416}{3 \times 3 \times 416 \times 416 \times 32} + \frac{416 \times 416 \times \frac{64}{32}}{3 \times 3 \times 416 \times 416 \times 32} + \frac{1}{32} = 0.0415$$

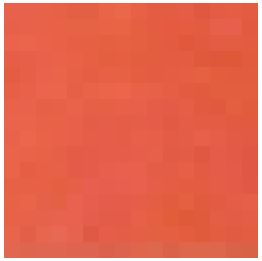
A BSC-S block has a reduction of computational cost of 100% - 4.15% = **95.85%** compared to the standard convolutional layer.

SpeeD
Speech & Dialogue Research Laboratory

Patch Model Dataset

- 16×16 pixels patches;
- Total no. of images:
  - 423 background;
  - 1096 jaguar;
  - 261 red;
  - 441 blue.

(a) Background



(c) Red

(b) Jaguar

(d) Blue

- Scripts:
  - patch annotation;
  - patch split;
  - patch encoding – `pickle` library.

SpeeD
Speech & Dialogue Research Laboratory

Patch Model NN

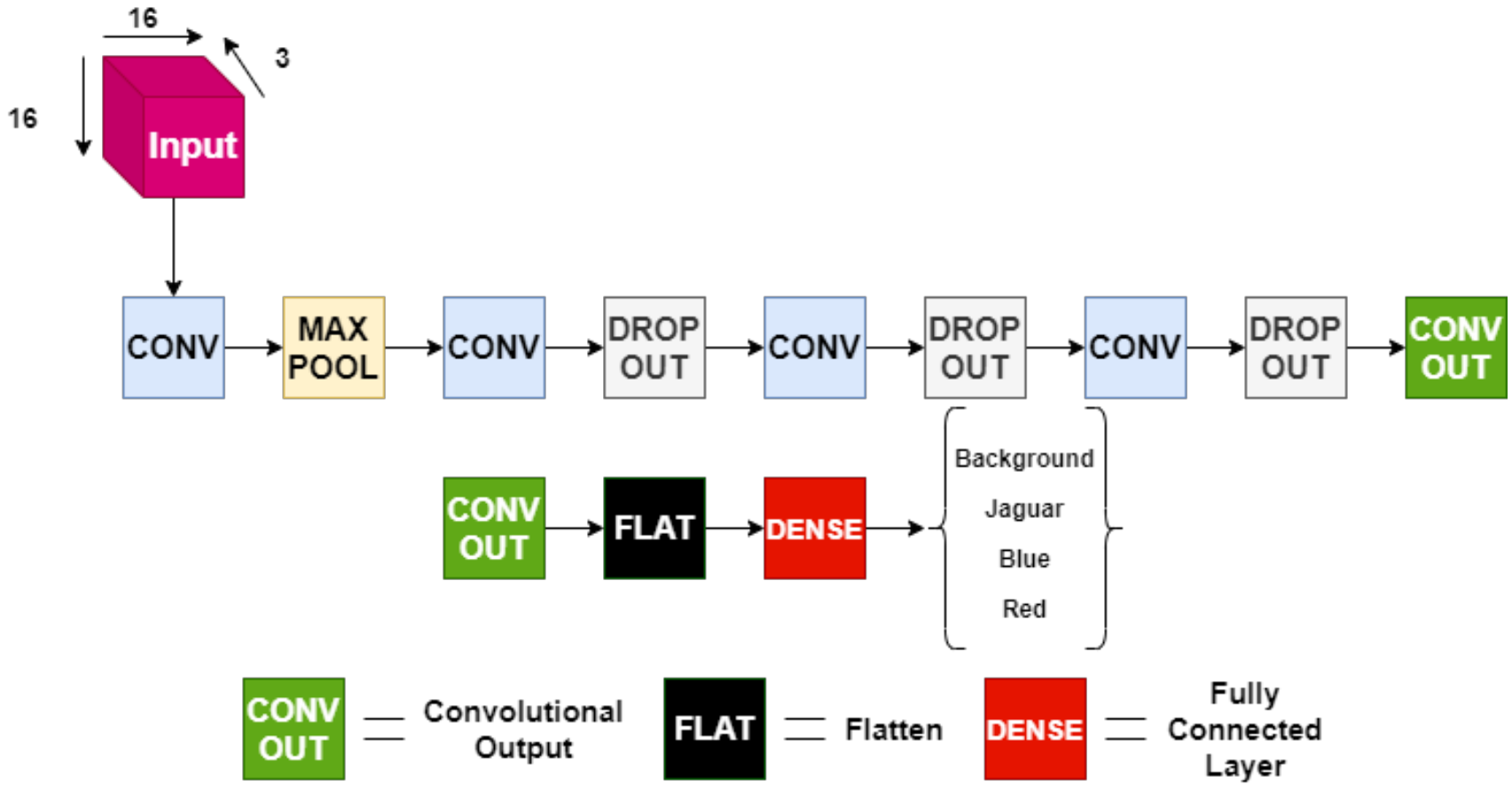


Figure: Patch Model Architecture

Results

- Modified Tiny YOLO:
  - Overall Confidence (C):

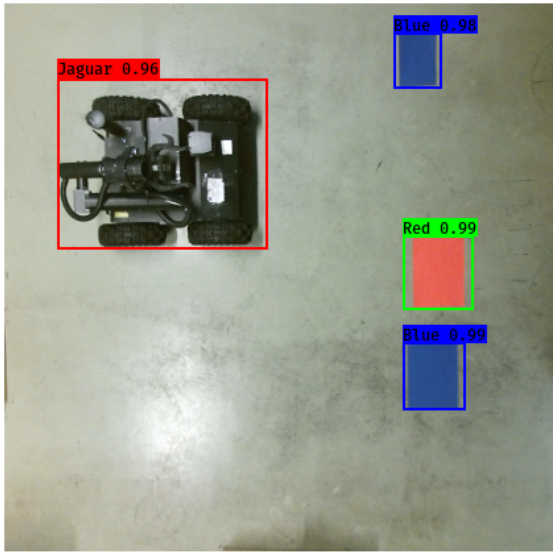
$$C = \text{Objectness} \times \text{Class Confidence}$$

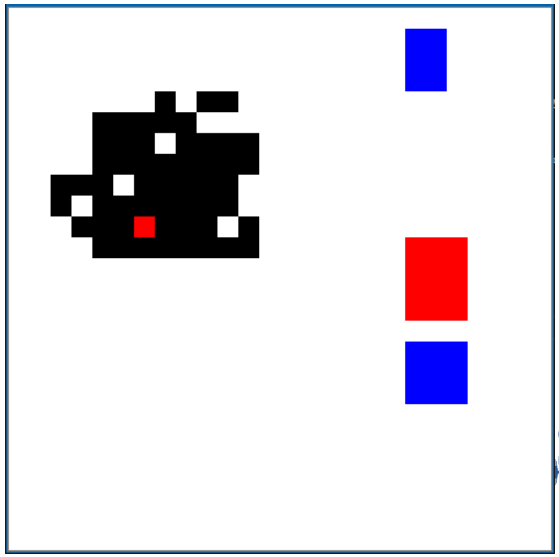
  - Non-max suppression which uses IoU:

$$IoU = \frac{\text{Intersection}}{\text{Union}}$$

- Patch Model:
  - SoftMax Activation S turns the last outputs into probabilities.

$$S = \frac{e^{z^{[l]}}}{\sum_{i=1}^{c} e^{z_i^{[l]}}}$$

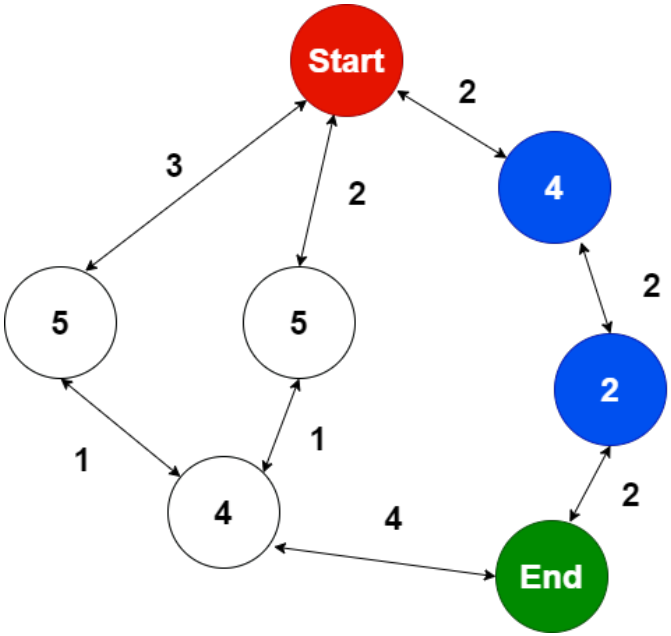




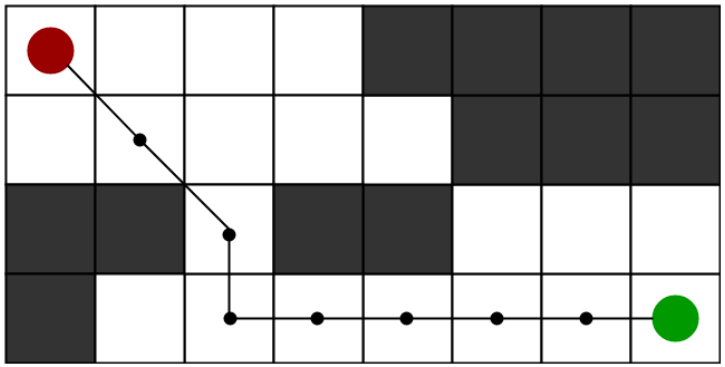
SpeeD
Speech & Dialogue Research Laboratory

Path finding algorithm

A* Algorithm



(a) A* Graph

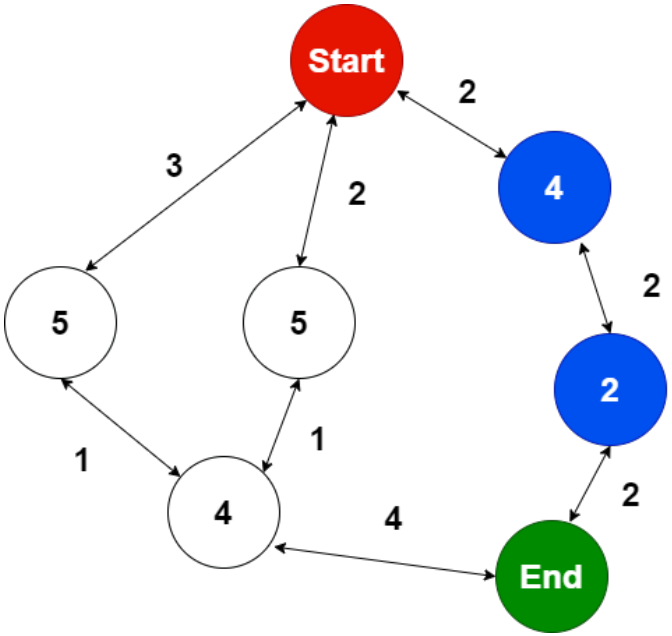


(b) A* Cell Path [5]

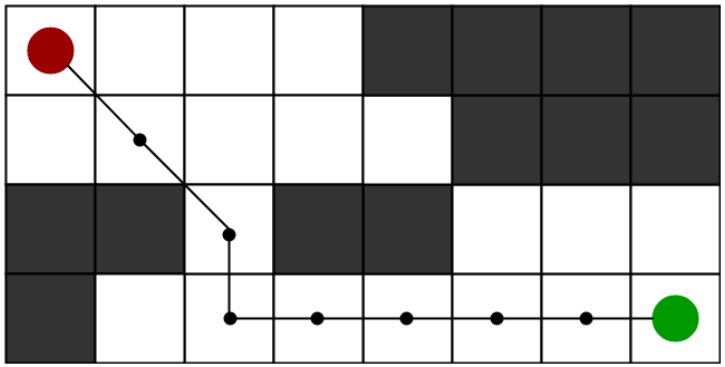
The **A* algorithm** makes use of a cost to choose the path.

SpeeD
Speech & Dialogue Research Laboratory

A* Algorithm



(a) A* Graph

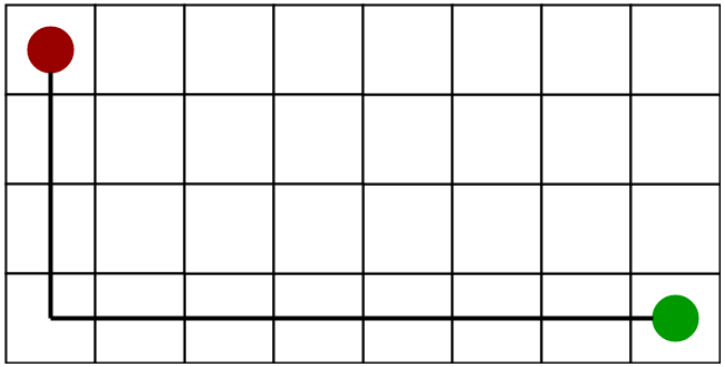


(b) A* Cell Path [5]

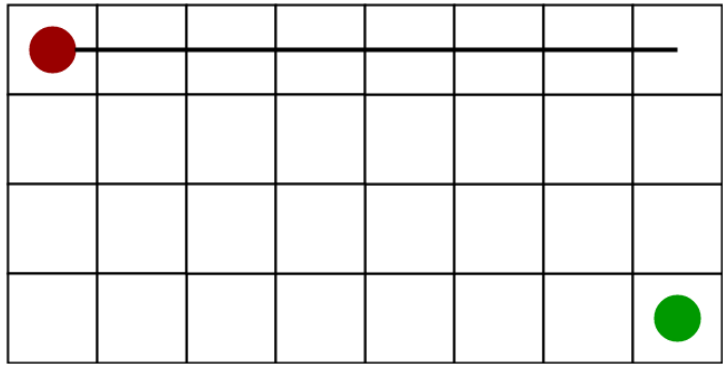
The **A* algorithm** makes use of a cost to choose the path.
Cost $\rightsquigarrow$ sum of the cost to the child node and the heuristic (distance) from the child node to the end node.

SpeeD
Speech & Dialogue Research Laboratory

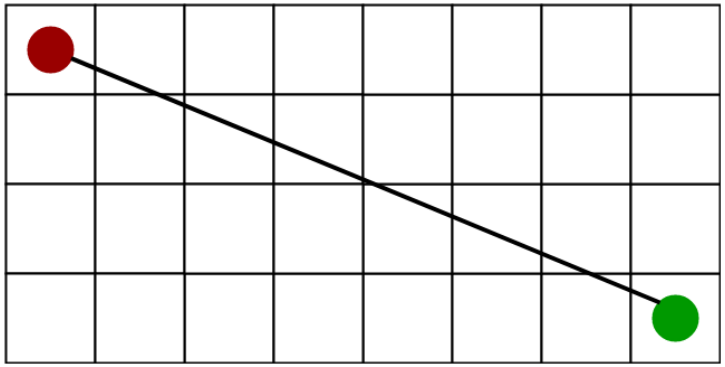
Heuristics



(a) **Manhattan Distance** [5]



(b) Diagonal Distance [5]



(c) Euclidean Distance [5]

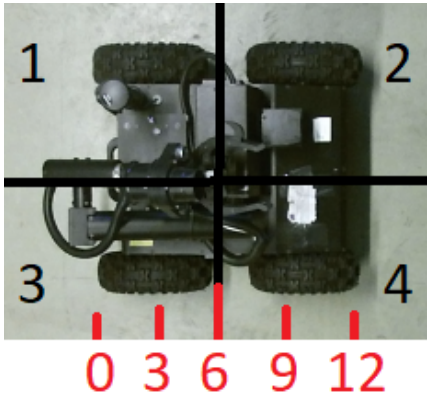
Figure: Types of Distances

SpeeD Speech & Dialogue Research Laboratory

Helpful Functions

Jaguar Direction

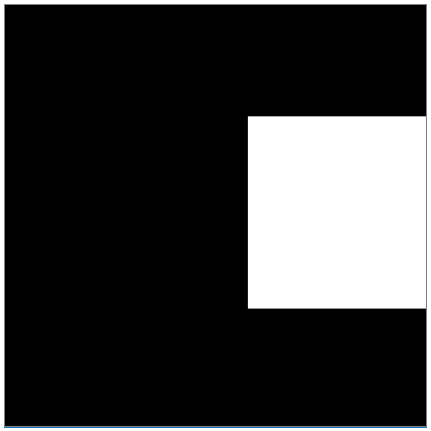
- find robot front;
- objective direction compared to front of the robot;
- robot turning command towards objective.



(a) Jaguar Direction

White Objects

- non-obstacles painted black;
- obstacles painted white;
- white block expanded for clearance.



(b) White Objects

Jaguar Movement

Translating **pixels path** into $\rightsquigarrow$ physical distances and commands.

```
Directions:  ['bottom-right', 'bottom', 'bottom-right', 'right']

Angles:  [45, 90, 45, 0]

Distances:  [27, 60, 33, 13]
```

```
1 if path[i][0] > path[i + 1][0] and path[i][1] == path[i + 1][1]:  # west
2     vector.append("left")
3 if path[i][0] < path[i + 1][0] and path[i][1] < path[i + 1][1]:  # south-east
4     vector.append("bottom-right")
```

SpeeD
Speech & Dialogue Research Laboratory

Final output of the proposed system

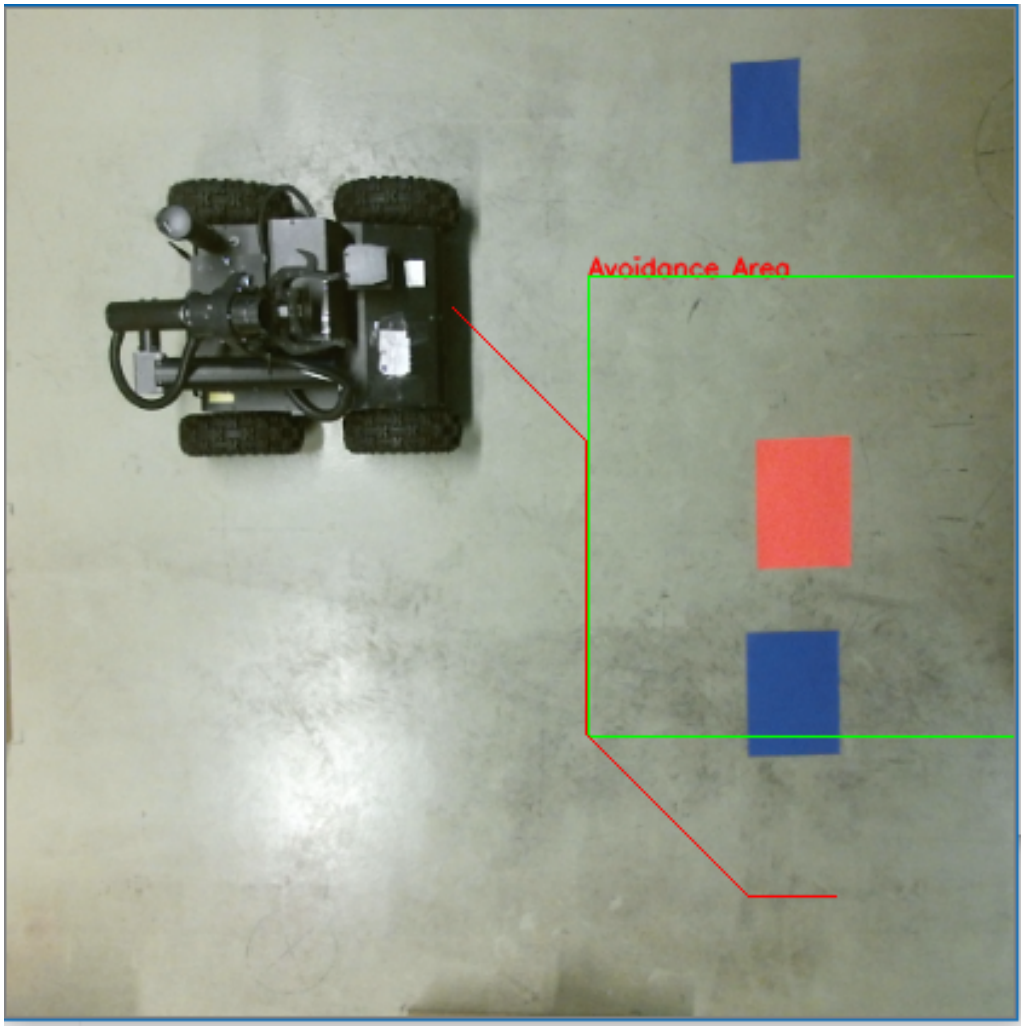


Figure: Final output of the proposed system

References

References


Andrew G. Howard, Menglong Zhu, Bo Chen, and et. al.

Mobilenets: Efficient convolutional neural networks for mobile vision applications, 2017.


Joseph Redmon and Ali Farhadi

Yolov3: An incremental improvement, 2018


Tudoroiu Mihai-Cristian

Sistem de evitare automata a obstacolelor cu robotul Jaguar 4X4 pe baza procesarii unor secvente video, 2019


Daniel Haase and Manuel Amthor

Rethinking depthwise separable convolutions: how intra-kernel correlations lead to improved mobilenets, 2020


Geeks for Geeks

A search algorithm*

SpeeD
Speech & Dialogue Research Laboratory

Contributions

Contributions

- Modified Tiny YOLO
  - Dataset;
  - Separable convolutions implementation;
  - Architecture fine-tuning;
- Patch Model
  - Dataset;
  - Architecture;
  - Bounding box coordinates.
- Pathfinding System
  - A* algorithm implementation;
  - Direction script;
  - Clearence script;
  - Robot commands script.

SpeeD
Speech & Dialogue Research Laboratory

Thank you!

SpeeD
Speech & Dialogue Research Laboratory