FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

U.PORTO

FEUP FACULDADE DE ENGENHARIA UNIVERSIDADE DO PORTO

Vehicle's Interior Presence Detection and Notification System

Jóni Ramiro Pires Gonçalves

Mestrado Integrado em Engenharia Eletrotécnica e de Computadores

Supervisor: Professor Doutor Jaime dos Santos Cardoso

Second Supervisor: Professora Doutora Maria Luísa Castro Guedes

July 25, 2018

Resumo

Esta tese tem como objetivo apresentar uma solução que previna a morte de crianças dentro de automóveis, tirando partido de tecnologias *low-cost* que podem ser facilmente integradas, ou até já fazer parte de um veículo.

As circunstâncias que conduzem a estas consequências podem restringir-se à falta de atenção dos cuidadores, por exemplo, quer por deixarem a criança desacompanhada quando foram fazer compras e demoraram mais que o esperado, quer simplesmente porque se esquecerem dela depois de levarem as compras para casa.

Diversos sistemas foram considerados aquando do estudo do estado de arte, culminando na idealização de um sistema que incorpora dois tipos de *input* – um sensor de movimento e uma câmara – enquanto o condutor é notificado via SMS.

A dissertação foi realizada tendo em conta duas perspetivas. Uma abordagem *low-cost*, que utiliza o sensor de movimento, e outro, que se antecipa ser mais seguro, consistindo em detetar humanos através de visão. Enquanto o sistema integrado é idealizado como a solução ótima para reconhecer a presença de crianças dentro do veículo, os estudos realizados sugerem que uma aplicação independente dos dois sistemas pode ser viável.

Ao longo deste estudo podem verificar-se limitações inerentes à implementação *low-cost*, uma vez que é necessário que haja movimento para que a deteção seja efetuada, ao passo que as câmaras podem reconhecer crianças mesmo que estejam a dormir. Ao mesmo tempo, os algoritmos de visão tornam-se mais viáveis a longo prazo, devido tanto ao seu constante desenvolvimento, como ao aprimoramento da eletrónica dos automóveis, que cada vez mais utilizam câmaras para a resolução de diversos problemas, por exemplo, sonolência do condutor, acabando por mitigar a limitação dos custos.

É esperado que esta dissertação sirva tanto como uma solução, bem como catalisador de soluções, uma vez que não está limitada a esta aplicação e sendo capaz de prevenir roubos de automóveis ou arrombamentos para surpreender o condutor sem que este se aperceba, ainda que o principal foco seja apresentar um sistema que detete e notifique de forma confiante e que permita reduzir o número de acidentes com crianças em automóveis.

Palavras chave - Algoritmos de visão, sensores de movimento PIR, tecnologia GSM, YOLO

Abstract

This thesis aims to present a solution to prevent in-car infants' deaths, capitalizing on low-cost technologies that can easily be integrated on the vehicle or are already a part of it.

The circumstances that led to this outcome can be narrowed down to parent's carelessness, for example, whether they left the child unattended because of a short time stop to do groceries that turned out too long, or simply because the child was forgotten after bringing such groceries home first.

As part of the state of art, several possible systems were considered, culminating in a system idealization incorporating two approaches as inputs – a motion sensor and a camera – while notifying the driver via SMS.

The dissertation was conducted in two perspectives. A low-cost approach, which utilizes the motion sensor, and other, that is anticipated to be more reliable, consisting in human detection through vision algorithms. While the integrated system is idealized to optimally recognize the presence of children inside the vehicle, the performed studies suggest a stand-alone application may be viable.

Throughout this study it is noted the low-cost implementation limitations due to the sensor requiring movement to acknowledge that there is someone present, while the use of a camera can recognize infants even if they are asleep. At the same time, vision algorithm becomes even more reliable overtime due to their constant improvements as well as the car electronics enhancements, with cameras providing solutions for several other issues such as driver drowsiness detection, mitigating cost constraints.

It is hoped that this dissertation serves as both a solution and solution catalyst since it is not limited to this application, being able to prevent cars from being stolen or burglar's break-in to surprise an unaware driver, although it's main focus is to present a reliable detection and notification system that allows in-car infant casualties decrease.

Keywords - GSM technology, PIR motion sensor, vision algorithms, YOLO

Acknowledgements

I would like to start by expressing my deepest gratitude to both my supervisors, professors Jaime dos Santos Cardoso and Maria Luísa Castro Guedes, for all the expertise and support throughout this endeavor, without their guidance it would have been a much harder path to follow.

To my parents, Lina Gonçalves and Ramiro Gonçalves, I am profoundly grateful for all the support and for believing in me since the beginning.

To my wonderful girlfriend Cristiana Teles, who stood by my side no matter the circumstance, always had an encouraging, kind and wise word whenever I felt insecure, and for being my emotional cornerstone, I can't express my gratitude enough.

To my grandparents Otília Rosa, Fátima Martins, José dos Ramos and José Gonçalves, my godmother Regina Pires, my aunt Dilma Santos and my godfather Jorge Santos, thanks for always being there whenever I needed.

I also want to especially express my gratitude to Diogo Sebe, Diogo Dias, Francisco Alpoim and Nuno Pires who accompanied me throughout my entire journey since I first arrived in Porto, helping me grow as a person and as an engineer.

To all my friends from Bragança, I want to thank all the good moments and for enduring my everlasting absence, especially to my "brother" Ricardo Afonso, who has always supported my decisions and has never let me down.

I also would like to thank Filipe "7k" Ferreira and Francisco "7k" Carneiro for all the never-ending distractions that kept me from burning out throughout these last couple of years, as well as Gil "7k" Valente, who also contributed to the edition of some images in this dissertation, adding some minor value to the Arts course.

A big thank you to my friends from FEUP Alexandre Pires, Ricardo Pereira, Pedro Dinis, Fábio Vasconcelos and Joaquim Ribeiro for integrating me in their restrict group, Miguel Rosa, Jorge Pinto and professor Agostinho Rocha for giving me the honor of accompanying them for lunch and also Pedro Galvão for his unceasing good mood and witty personality.

Thank you.

Jóni Ramiro Pires Gonçalves

“Sorry losers and haters, but my I.Q. is one of the highest -and you all know it! Please don’t feel so stupid or insecure,it’s not your fault”

Donald J. Trump

Contents

List of Figures

List of Tables

Acronyms

CNN	Convolutional Neural Networks
FEUP	Faculdade de Engenharia da Universidade do Porto
FPS	Frames Per Second
GPU	Graphics Processing Unit
GSM	Global System for Mobile Communications, originally Groupe Spécial Mobilet
GPS	Global Positioning System
GPU	Graphics Processing Unit
IEEE	Institute of Electrical and Electronics Engineers
LED	Light-Emitting Diode
MCU	Microcontroller Unit
MPU	Microprocessor Unit
PIC	Peripheral Interface Controller
PIR	Passive InfraRed
R-CNN	Regional Convolutional Neural Network
RAM	Random Access Memory
ROC	Receiver Operating Characteristic
RoI	Region of Interest
RPN	Regional Proposal Network
SMS	Short Message Service
USB	Universal Serial Bus
UWB	Ultra Wide Band
YOLO	You Only Look Once

Chapter 1

Introduction

1.1 Context and Motivation

The automobile industry is presently in a mature state, where thousands of cars are commercialized per day and every year new models are presented that surpass previous ones. In order to remain competitive, automotive corporations explore new ways to improve their vehicles taking advantage of technological advances. The integration of electronics systems, in detriment of mechanical ones, provided enhancements both at comfort and security levels. The Electrical Engineering advancements allowed the development of several technologies which became standard in every automobile such as electronic injection and ignition, or as extras, such as GPS or automatic parking. This dissertation inserts itself in this branch of technological improvements, where a system is envisioned to solve a life-threatening behavior.

The main purpose is the development of a system that recognizes human presence inside a vehicle whenever the driver leaves the vehicle and posterior notification, as this dissertation title suggests. The necessity of such solution has as main problematic children left unattended inside the automobile, which in some cases might lead to the end of their life.

According to a San Jose State University's study from Jan Null [2], an average of thirty-seven kids die per year in the USA, victims of heatstroke. Statistics show that 54% of these casualties are due to being forgotten and 28% due to being unattended while playing. The study comprises kids age up to 14 years, yet 96% of the deaths are infants up to 5 years old. This reality indicates negligence after parking the vehicle, either by leaving the infants unattended or even forgetting about them.

In this dissertation we propose a solution whose purpose is recognizing the presence of infants, regardless of their age, whenever the caregiver is distracted by its daily routine and simply forgets the children in the rear in the back seat.

In the following, two different systems are proposed, a low-cost approach whose implementation is documented and evaluated as well as a higher-cost alternative whose reliability is studied.

The most desirable solution would compound both these concepts in a single, more reliable system.

Both standalone proposed solutions should provide reliable detection and notification on their own and be able to mitigate in-car infant fatalities.

1.2 Objectives

As previously state, in this dissertation we propose a solution to the in-car infant death. In order to achieve the optimal solution several steps are taken to decide which path should be chosen, such as study of different detection and notification techniques, analyze and compare each ones' pros and cons.

Along with these steps several objectives are drawn, namely:

- System's architecture choice;
- System's architecture and algorithm implementation;
- System's validation through test.

It is also required that the developed system is:

- Generic - able to port to any automobile brand;
- Low-cost - in order to be affordable for virtually everyone;
- Low power consumption - since the energy is drawn from the car's battery.

Following these lines, it is expected that the developed system is reliably capable of in-car detection and subsequent driver warning. Lastly, conclusions of the results will be drawn, as well as suggestions for future improvements.

In short, we intend to propose a generic low-cost system, whose selected architecture has been carefully considered and its reliability and limitations tested and acknowledged.

1.3 Structure

The present document is structured in a way that showcases the dissertation agenda, meaning that, more or less, the chapters accompany the time frame of the realized work.

Chapter 1 serves as an introduction to this dissertation thematic, its technological context, the catalysts that propelled its proposition and its fundamental objectives.

Chapter 2 presents the state of art, several publications of this dissertation problematic are reviewed, as well as other solutions that are also related. It is also briefly reviewed some vision algorithms state of art.

Chapter 3 extends the previous chapter into an evaluation of the potential technologies that might be chosen to the final system's implementation. For both detection, communication and control choices pros and cons are pondered and considered.

Chapter 4 introduces the system idealization and further expands the implemented solution, the options that were adopted, architecture, realized tests and their results.

In chapter 5, an overall conclusion is made about the several different points discussed along the document. In this chapter, future work possibilities are also discussed as well.

Chapter 2

State of the Art

Throughout this chapter it is presented the state of the art, where several scientific papers whose contents address this dissertation problematic or have solutions that can be redesigned for this system's application, are analyzed. The technologies pros and cons are expanded in the following (chapter 3).

2.1 Potential Conceptual Systems

The first two papers analyzed were authored by Fairuz Rashidi and Ikhwan Muhamad [3] [4], both aiming to mitigate in-car infants death, due to caregivers' neglection, fundamentally the same as this dissertation.

In their first paper [3] it is conceptualized a system that alerts the driver whether he is nearby or not. The system consists of:

- A passive infrared (PIR) sensor;
- A GSM module;
- A PIC microcontroller.

The algorithm kicks off when the vehicle is immobilized, checking, for example, if the ignition is turned off. Since there is no need for a full time running system, this allows a decreased current consumption. After starting up, and whenever the PIR sensor detects motion, the PIC microcontroller interprets its signal and generates an alert message which is sent through the GSM module. According to the authors, the developed prototype behaves as intended, suggesting a potential integration within the vehicle electronics, allowing the employment of the vehicle's alarm as an alternative notification system.

Their second paper's premise is the same, in-car infant detection, however using a different type of approach.

This other system consists of two cameras, one for the back seats, the other for the front ones, and a Raspberry Pi. It follows the same rule to start up, requiring the vehicle to be immobilized and turned off. It is also worth noting that it turns OFF thirty minutes without positive detection.

Since this solution has higher current consumption than the sensor-based one, this guideline, turning OFF after thirty minutes, allows a better energy management. There is also the temperature inside the vehicle which is bound to rise risking heatstroke fatalities, especially during sunny days. For this reason, it is demanded fast recognition and, therefore, it is required a low false negative (FN) rate.

As expected, this system's input is a video, so it requires a higher processing power, therefore the necessity for a Raspberry Pi. The image processing is based on a simple algorithm of motion detection.

The authors explain that the testing procedures were performed in controlled light conditions, admitting that further development and investigations could improve the system reliability. It is also showcased the algorithm's performance, claiming that it can detect human presence in less than three minutes.

On a slightly different premise, Haibin Cai [5] proposes a vision-based system that aims to reduce car theft.

The presented system can be broken down in two parts, one consisting of an accelerator and the other of a camera. It is assumed that the accelerator is able to detect evident theft attempts, such as breaking a window. The camera is only used if the accelerator recognizes an attempt but is not evident enough. The *modus operandis* can be visualized in Figure 2.1. The system starts wit a window damage detection, immediately activating the alarm if it was indeed damaged. After that, the accelerator is turned on, and will be generating two possible outputs, the assault detection, which will immediately activate the alarm, or a possible disturbance detection, which will then activate the vision system. The vision system will then confirm the assault attempt, and turned OFF right after.

Whenever the camera is activated, the image processing algorithm distinguishes intrusion movements from, for example, people crossing by.

It is claimed that this system was tested in several locations under different conditions, simulating both intrusion movements and people crossing by, presenting a 91.7% success. The author also provides the current consumption rates, that varies between $2mA$ while the accelerator is turned on, to $160mA$ when the camera is activated.

The paper entitled "RF-Based child occupation detection in the vehicle interior" [6] presents yet another solution to reduce in-car casualties. In this article it is documented the development of a radio frequency sensor capable of detecting respiration as well as the heartbeat of sleeping babies or children.

The main challenge of this approach is to differentiate background noise from a heartbeat or respiration. In Figure 2.2 it is shown the distinction between a rain signal and a child using a

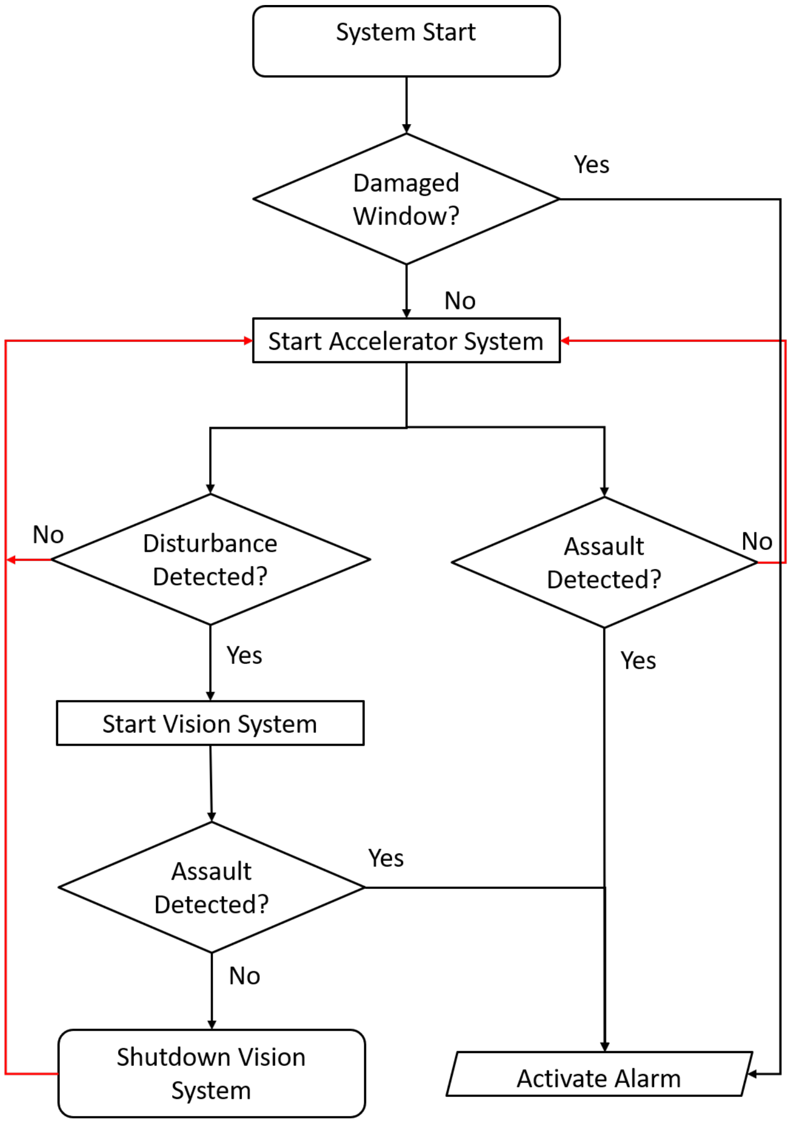

Figure 2.1: Haibin Cai functional system's flowchart [5]

pattern recognition algorithm.

This article's authors also declare that tests were run with fifty different children, being able to distinguish them apart from traffic noises or rain without any false positives. It is also referred that a dummy was created in order to simulate the worst-case scenery, allowing further systematic experiences regarding the system's reliability.

Contrasting with previous papers, the self-regulated automatic ventilation of vehicle interior [7] premise is not about preventing casualties or even thefts, instead its objective is opening the cars windows if the vehicle temperature rises above a certain threshold.

Although it does not fit this dissertation theme, it showcases an idea that can be considered relevant, for instance, if a child is still inside the vehicle, the ability to create an airflow can reduce

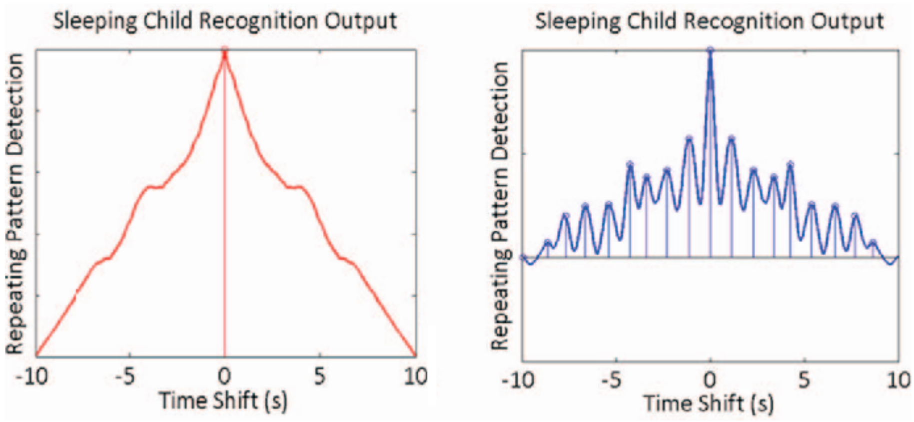

Figure 2.2: Comparison of analyzed patterns after algorithm process [6] of rain (left) and child heartbeat (right)

the inherent danger, although not being able to mitigate it.

The idealized system consists of a temperature sensor as well as motion sensor that detect movement around the vehicle and a precipitation sensor in case of rain.

The last solution analyzed was published by Norizam Sulaiman and Kamarul Hawari Ghazali entitled "Development of comprehensive unattended child warning and feedback system in vehicle" [8]. It consists in a system that contains some of the referred technologies such as the PIR motion and temperature sensors, GSM module, car alarm integration and window opening, including also an *EasyVR Shield* for voice recognition and an *eNose* for odor recognition.

The *modus operandis* is similar to the previous systems, for example it is activated only once the car is immobilized. The detection blocks are turned ON in sequential order, as follows:

- Temperature sensor;
- PIR sensor;
- Voice recognition;
- Odor recognition.

If any of these modules gets triggered, the alarm is set ON generating a notification that is transmitted to the vehicle's owner via SMS. If this message is ignored, the car alarm is turned OFF and the windows are lowered to create an airflow.

2.2 Vision Algorithms

Some of the previously reported papers present a vision-based solution, which demands image processing. Throughout this section a brief history and evolution of some of the approaches are presented.

2.2.1 Viola-Jones Object Detection

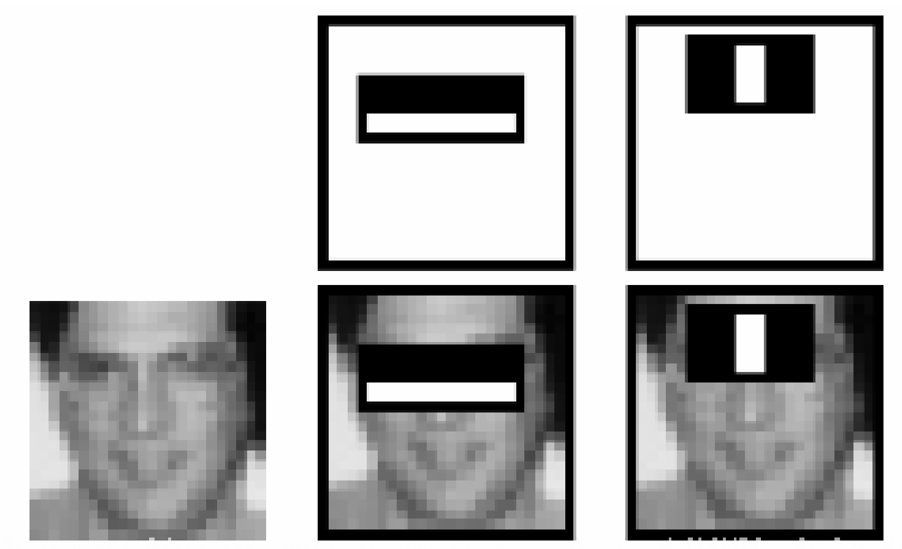

Figure 2.3: Viola-Jones first and second selected features for face detection [9]

In 2001 Paul Viola and Michael Jones [9] presented a machine learning approach capable of fast object detection with high detection rates.

The presented algorithm had as cornerstones three key contributions:

- Integral Image - which allows fast computation;
- AdaBoost based algorithm - which selects a small number of visual characteristics;
- Combining complex classifiers in a cascade which concedes a fast dispose of the background.

This algorithm allowed real-time applications running at 15 fps without resorting to image difference or skin color, as previous state of art solutions.

In Figure 2.3 it is shown two features that helps the algorithm perceive human faces.

The left feature measures the light difference between the eyes and upper cheeks taking advantage of the eye region being darker. The right feature works the same way, comparing the eyes region with the nose's one.

2.2.2 R-CNN Family

Several other approaches were proposed, until in 2012 Alex Krizhevsky *et al.* [10] turned CNN (Convolutional Neural Networks) into the image classification standard, winning the annual Olympics of computer vision - ImageNet - with a CNN known as AlexNet.

The CNN itself was not a novelty, CNNs were around since the 90's, however two reasons were holding it back, and only then was possible to reliably adopt CNN at image classification:

- The GPU improvement - increasing substantially the computing power;
- The amount of data - creating far wider training sets.

With this revolutionary tool, one of the most popular approaches was introduced by Ross Girshick *et al.* - the R-CNN (Regional CNN) [11].

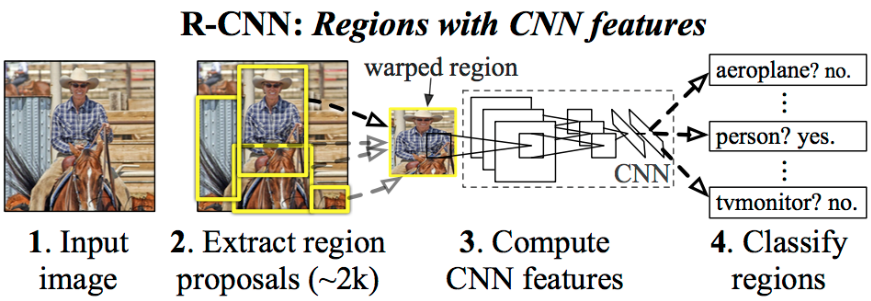

Figure 2.4: RCNN classification process [11]

With R-CNN the main goal was to correctly identify where the objects were in the picture. The process is enhanced due to running a *selective search* to extract around two thousand region proposals, which are then fed to a CNN version of the AlexNet [10] as pictured in Figure 2.4. The classification layer operates with linear regression that allows the creation of a tighten bounding box.

Carrying on with this evolutionary chain, Fast R-CNN was created in 2015 also by Ross Girshick [12], envisioning the improvement of the training while speeding the computational process. R-CNN could be considered slow due to the process of each of the two thousand regions going through the AlexNet CNN. On top of that, R-CNN also has a three-stage training, the CNN image feature, the class prediction classifier and the bounding box linear regression.

One of the improvements was the RoI (Region of Interest), since several of the regions overlapped, instead of running the two thousand region proposal through the CNN, the image goes through the CNN only once, outputting a RoI projected in a $h \times w$ grid.

The second improvement consisted in the junction of the training stages both the CNN, classifier and the bounding box.

Figure 2.5 represents the Fast R-CNN architecture, the image and the RoI are fed as inputs to the CNN. The RoI pooling layer gets a fixed-size feature map for all RoI, and the FCs (Fully Connected layers) map a feature vector, outputting the softmax probabilities and the bounding boxes regression offsets.

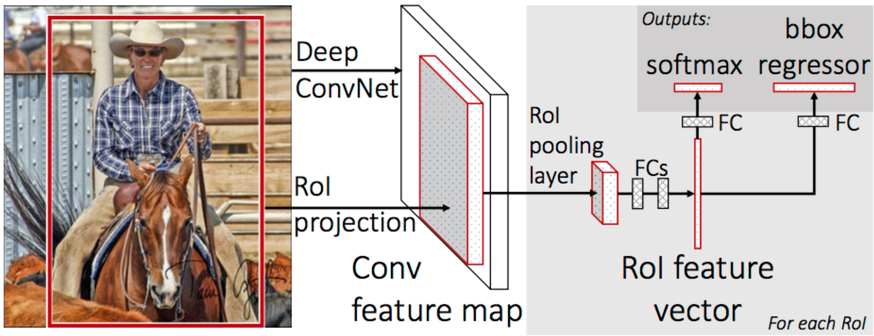

Figure 2.5: Fast R-CNN architecture [12]

2.2.3 YOLO family

The improvement series kept going, and in 2016 Shaoqing Ren *et al.* published the Faster R-CNN [13], a Fast R-CNN upgrade. After previous improvements, the region proposal computation was exposed as a bottleneck, which previously resorted to the selective search.

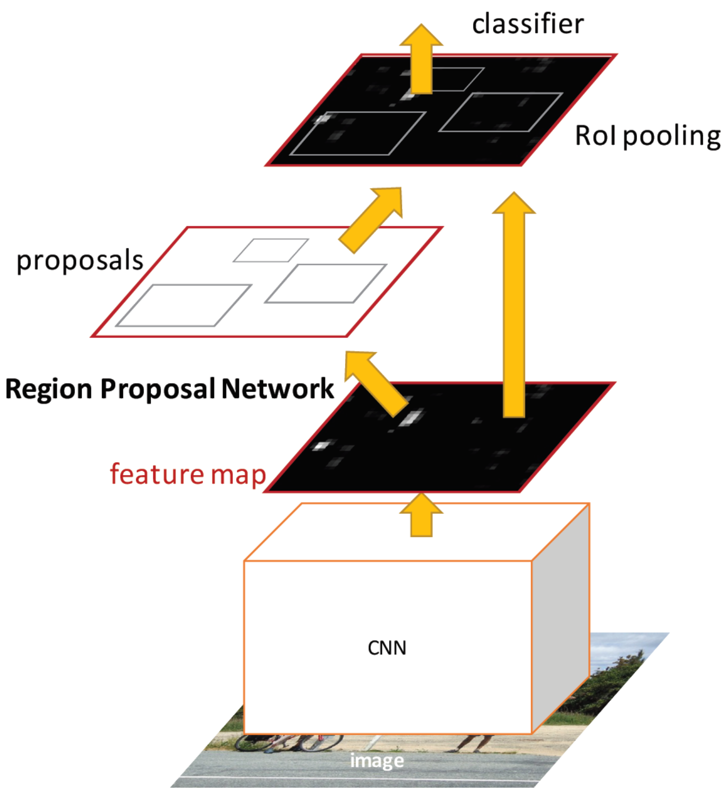

Figure 2.6: Faster R-CNN architecture [13]

It was noted that the regional proposals depended of the features already calculated, so instead of using the selective search, an RPN (Regional Proposal Network) was created. The RPN works as an *attention* mechanism and permits that only one CNN needs to be trained, allowing nearly cost-free region proposals by telling the Fast R-CNN module where to look.

In Figure 2.6 it is shown the Faster R-CNN architecture, where the only CNN input is the image. Feature map then feeds both the RPN that creates the region proposals as well as the RoI pooling.

In 2017 Kaiming He *et al.* published the last (to date) R-CNN development - Mask R-CNN [14]. Mask R-CNN extends Faster R-CNN by adding an extra branch to predict an object mask, which outputs a binary mask predicting if a given pixel is part of an object.

The pixel-to-pixel behavior employs the RoI features, which are small feature maps that need to be well aligned to faithfully preserve the spatial correspondence, which motivated the development of "RoIAlign".

The feature map of an image is not the same size, so the pixel translation from the feature map back to the image may originate in decimal pixels that would be rounded down with RoIPool. With RoIAlign a bilinear interpolation is adopted to precisely compute the exact values of the input features.

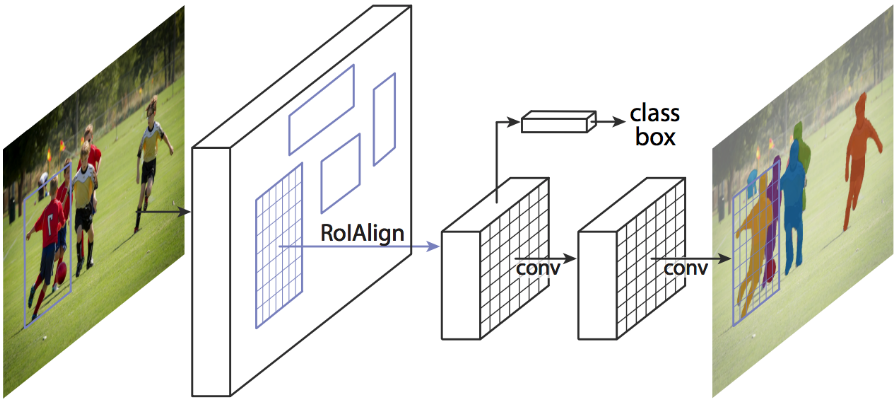

Figure 2.7: Mask R-CNN architecture [14]

In Figure 2.7 it is shown the faster R-CNN branch which outputs the class label and the bounding box parallel to the FCN (Fully Convolutional Network) that predicts the mask segmentation pixel-to-pixel.

The previous models, the R-CNN "family", try to recognize the objects through their classification, however in 2016 Joseph Redmon *et al.* created a model that used unified regression instead, published as YOLO (You Only Look Once) [15]. YOLO image processing can be described in three steps:

- Resize the input image;
- Run CNN;

- Apply confidence bounding boxes.

The YOLO model is depicted in Figure 2.8.

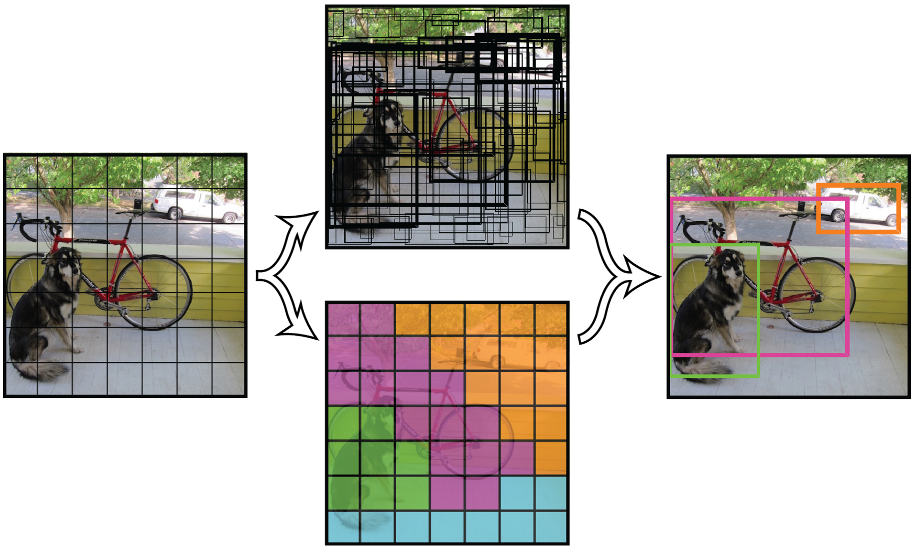

Figure 2.8: YOLO model [15]

This model can be narrowed down to:

- Dividing the image into an $S \times S$ grid - leftmost image in Figure 2.8;
- For each grid cell it is predicted B bounding boxes - uppermost image - which are represented through thickness, meaning that the thicker the bounding box, the higher the confidence of the prediction;
- C class probabilities are predicted - lowermost image - depicted as different colors in different cells;
- The result - rightmost image - is given by setting the confidence threshold, eliminating every other bounding box.

The first published YOLO paper presented two versions of this algorithm, the YOLO and the fast YOLO, the latter meant to push the fast object detection boundaries. This consists in less convolutional layers (nine instead of twenty-four) and was able to run at $155fps$ (frames per second) compared to the $45fps$ of the original version. Its precision [1] is 52.7% mAP (Mean Average Precision) while the original YOLO pushes further to 63.4%. The main struggle, acknowledged by the authors, is the localization of errors.

Still in 2016, Joseph Redmon and Ali Farhadi publish the article "YOLO9000 Better, Faster, Stronger" [16], with two new directions, the YOLO9000 that is able to detect over nine thousand

[1] Precision - The probability that a decision is correct - $\#TruePositives/\#PredictedPositives$

object categories and the YOLOv2 that is capable of running at $67fps$ with 76.8% mAP in the VOC 2007 dataset.

The YOLOv2 improvement comes from several designed decisions such as:

- Batch normalization;
- High resolution classifier;
- Convolutional network with anchor boxes;
- Dimension clusters;
- Direct location prediction;
- Fine grained features;
- Multi-scale training.

Each of these improvements led to significant mAP increases, excluding the convolutional network with anchor boxes, which saw a small decrease of 0.3% mAP but displaying a higher recall[2], meaning that there is still room for improvement.

Lastly, in April 2018 Joseph Redmon and Ali Farhadi published the last (to date) YOLO upgrade, the YOLOv3 [17].

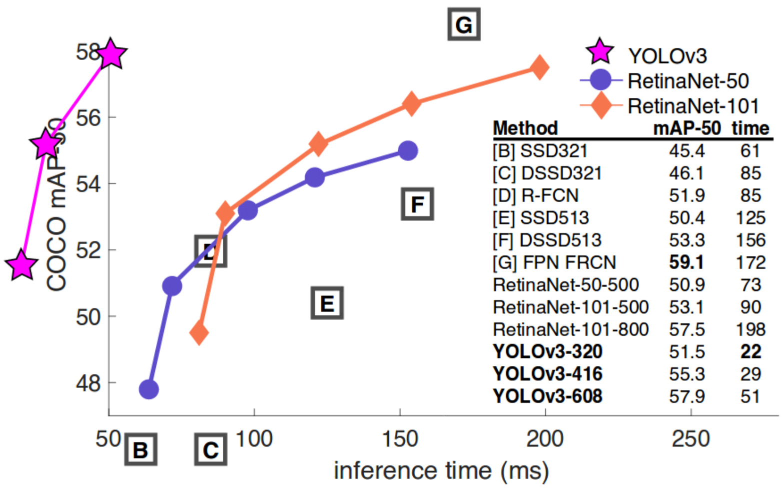

Figure 2.9: YOLOv3 performance [17]

In this "tech report" some small improvements are made to YOLO and some approaches that did not improve it are also listed, such as:

[2] Recall - The probability that a prediction is correct - $\#TruePositives/\#ActualPositives$

- Anchor box *x,y* offset predictions - this formulation decreased the model stability;
- Linear *x,y* predictions instead of logistic - this decreased a couple of mAP points;
- Focal loss - also decreased a couple of mAP points;
- Dual IOU thresholds and truth assignment - also decreased the mAP.

It is acknowledged that these techniques eventually might produce better results with fine tuning.

In this paper there is a metric switch to measure the detectors accuracy and Figure 2.9 displays the precision versus inference time of YOLOv3 comparatively to RetinaNet.

2.3 Solutions Limitations

All the analyzed papers provide a solution that can decrease the fatalities. However it is also possible to pin point some limitations inherent to the proposed technologies.

Starting with PIR sensor, it is only possible to detect motion if the vehicle temperature differs from the child's, otherwise the reliability decreases, especially during Summer.

The radio-frequency sensor has another type of constraints related with the children's health. According to the IEEE standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, $3kHz$ to $300gHz$ [18] the maximum power density level for human exposure is $10mW/cm^2$.

Even though cameras have a higher energy need comparatively to other options, it is expected to become more reliable over time following the algorithms natural evolution. An elegant solution to the consumption problem, as previously analyzed, is the integration with another detection system and activating the cameras only when there is doubt in the detection.

In order to employ cameras there is also the need for a vision algorithm. Some potential approaches were also analyzed, all of them relying in machine learning techniques. There are also several other approaches that were not reviewed such as the background subtraction whose operation principal is close to the motion sensor as it requires motion to effectively detect the target.

As for the facial recognition algorithms, it is expected that its reliability increase over time, as the several successive upgrades to the R-CNN and YOLO suggest.

2.4 Conclusions

In this chapter several papers with a relevant basis were analyzed in order to present potential starting points as well as provide an overview of their limitations.

Several different types of implementations were analyzed and some were proven to be clearly worse due to the dissertation thematic, for instance, there is little to gain from a voice pattern recognition to detect infants cries as it requires more processing power and will not work if the child is silent. However, this implementation was integrated with three more detection technologies, and showcasing that an integration of different options might upgrade the systems reliability and usability.

It is also possible to see a trend in the components that the majority of the articles analyzed follow, which are cameras and PIR motion sensors. It is deduced that these choices seem to be more reliable than their counterparts, and further review, in chapter 3, shall provide a better picture of pros and cons associated with each technology.

Lastly, a small introduction of some vision algorithms was presented, in order to provide a better understanding of the vision solutions and if it is indeed a reliable option. As the articles suggest, both the last versions of Mask R-CNN and YOLO provide an increasing accuracy rate as the algorithms evolve, inducing that over time their reliability will keep increasing. As it stands, it is already considered a reliable solution to object detection while running in real-time.

Chapter 3

Technologies

Throughout the previous chapter several documents were presented that aim to detect presence inside the vehicle. Any of the previous solutions could be chosen as a starting point for this dissertation thematic, however, the selection of the system components needs to consider their inherent pros and cons.

In this chapter we discuss the technologies that can be potentially used, weighting their advantages. Three distinct groups are analyzed:

- Sensors;
- Communication techniques;
- Control Units.

This evaluation will lead to the components that were chosen. More detailed analysis of the selected components is presented with the system development, in chapter 4.

3.1 Sensors

3.1.1 Weight Sensor

There are several different sensors that can be applied to in-car human detection. A simple way to know if a seat is occupied is by measuring the weight. A weight sensor installed on the seat will be able to sense whenever someone is sitting there. There are several topologies of this sensor, the most common being the strain gauge, coupling the gauges to a Wheatstone bridge, as depicted in Figure 3.1 half bridge.

This sensor working principle simply consists in generating a variable electric resistance which unbalances the Wheatstone bridge translating in a voltage output deviation that is proportional to the gauge deformation. Since this deformation proportionately depends on the weight applied to the gauge the weight measure is enabled. The strain gauge is the most used weight sensor as a result of its reliability along with long lifetime.

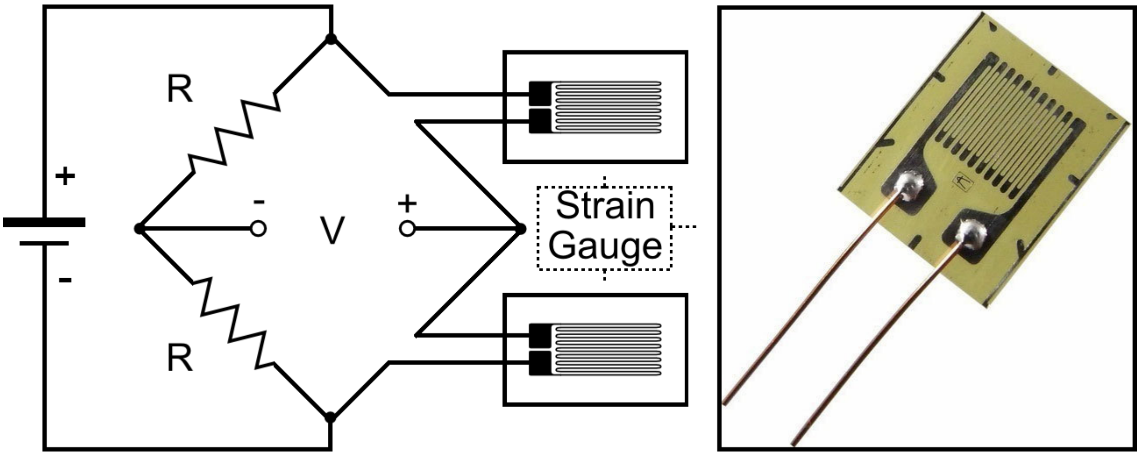

Figure 3.1: Half Wheatstone bridge, left, and strain gauge, right

The recognition algorithm could rely on just a couple of measurements, the first whenever the car starts moving, which guarantees that no one is going to enter the car afterwards and the second after the car is locked, assuring that no one is going to leave the car. It is also possible to set the value for the empty seat and comparing to the last measure, requiring a measure only after the car is locked. The difference of these measurements provides information on whether there is someone inside.

According to the sensor precision it is possible to detect a small child's weight, however there is no sensibility to whether the pressure is done by a person or some other object left in car. This system is already present in several vehicles with the goal of checking if everyone inside is wearing their seat-belts. However if a heavier object is on top of the seat, the alert will be equally activated.

3.1.2 Microphone

In section 2.1, the last paper analyzed [8] introduces a voice recognition system. This solution requires that every noise needs to be distinguished from the children cry, requiring a voice pattern recognition algorithm that potentially needs the same processing power as a camera, with more disadvantages. This system is not reliable when the baby is sleeping, and even if he is awake it needs the baby to cry. If the child is older, then it might not even cry, making this solution ineffective.

3.1.3 Continuous Wave Sensor

Also analyzed in previous chapter, a continuous wave (CW) sensor [6], might be the most efficient one. A potential architecture is depicted in Figure 3.2, from İsmail Şişman et al. publication "Micro-doppler radar for human breathing and heart-beat detection" [19].

This type of sensor allows to detect heart-beats, and with a pattern recognition algorithm it is possible to distinguish it from other type of noises, such as rain. The working principle relies in reflecting a signal on the target, resorting in the Doppler effect, and filtering the received wave eliminating its noise, obtaining an observable pattern, as shown in Figure 3.2.

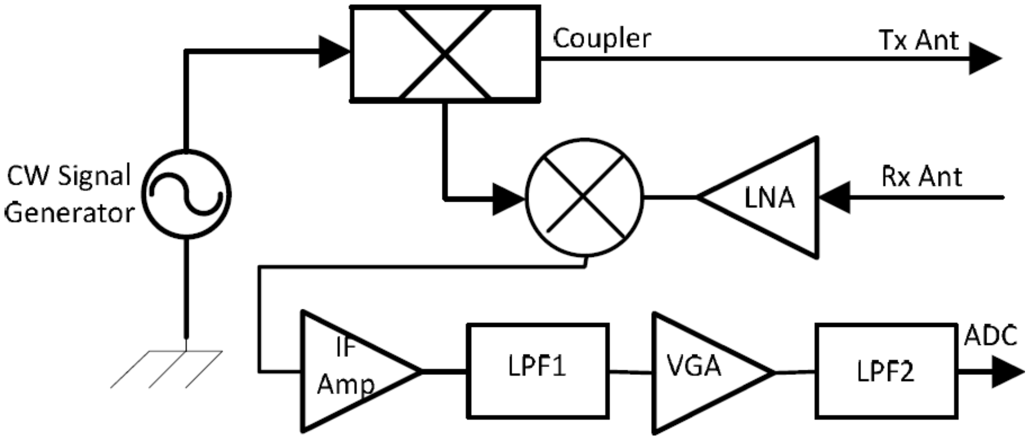

Figure 3.2: Doppler radar architecture [19]

Contrary to the other sensors, this approach is immune to light and temperature conditions as well as target shape or steadiness, precisely detecting human beings (or animals) [19] [20]. This technology has, however, a health constraint. Since this type of sensors resort in emitting energy, there is a concern regarding the emitted signal's power which cannot be over $10mW/cm^2$ [21] [18]. An exposure to a high power radiation for a long period of time might originate in radiation poisoning.

3.1.4 Passive InfraRed Sensor

Throughout chapter 2.1 several investigators choose to detect in-car presence resorting to motion PIR sensors. This type of sensors is sensitive to the body temperature, in contrast with the surroundings.

Every object with greater temperature than absolute zero emits a thermal radiation, which is not observable with naked eye - infrared radiation. The PIR sensor captures the energy transmitted by the target objects although no energy is transmitted from the sensor, hence the passive nomenclature. If the target temperature becomes too close to the surroundings, the sensor's reliability drastically decreases, not being able to distinguish one from the other. The temperature inside a vehicle rapidly rises due to sun's radiation and when it gets in the range of human body temperature, detection is no longer possible.

In the presence of temperature differences, the infrared sensor will only be activated if motion is detected. This means that a still body, like a sleeping baby, might not be detected.

3.1.5 Cameras

On pair with PIR sensors, also cameras seem to be a good option for in-car presence detection. The integration of cameras in vehicles has increased over the last decade, benefiting of both algorithm improvements, automobile industry evolution and brands competition. Since the cameras range of applications is not restricted to the automobile industry, its improvement benefits from every field of application breakthroughs.

Newer cars already have several camera applications installed such as automated parking system as well as applications being researched such as driver drowsiness or interior presence detection.

In order to detect infants inside the car several options can be chosen, for example, a motion detection. This approach is very similar to the motion sensors, since it requires motion to activate, without the temperature gap requirement downside. It has, however, the need for target motion which associated with car vibration's sensibility might translate in a false positive triggers escalation.

The more relevant approach relies in object recognition, such as the algorithms enumerated in chapter 2. This method reliability has been enhanced over the last decade due to different CNN algorithms, such as R-CNN or YOLO, series of improvements.

Setting the reliability of these algorithms in detecting human beings aside, cameras require high processing power which therefore leads to an overall current consumption increase.

3.1.6 Comparison

In table 3.1 a crude overall comparison is made, where the PIR sensor can be highlighted as the best option. In this sense, cameras seem to be unnecessary, however, as a second layer of detection the system's reliability is expected to improve.

In the CW sensor case it has already been discussed the health complications, so even if this is the most reliable option, it is still the most dangerous.

Table 3.1: Different sensors overall comparison

	Power consumption	Cost	Reliability
Weight sensor	++	++	-
Microphone	-	-	–
Continuous wave sensor	-	-	++
PIR sensor	++	++	+
Cameras	-	-	+

3.2 Control Units

The previously discussed sensing components require a control unit in order to accurately process their output signals. To properly interpret such signals, different control units might be required, and since the presented sensors have different processing requirements, different control units should be considered.

In the previous section, several times was mentioned that certain sensors required higher processing power, specifically the pattern recognition and vision-based solutions. The rationale that distinguishes a high processing power from a low process requirement can be described as the amount of information the control unit needs to handle in order to acknowledge, in this application's case, if there is in fact someone inside the vehicle. The information can be as simple as read a digital input where zero value translates to no motion and one to motion detected, or complex as evaluate the facial features on video frame.

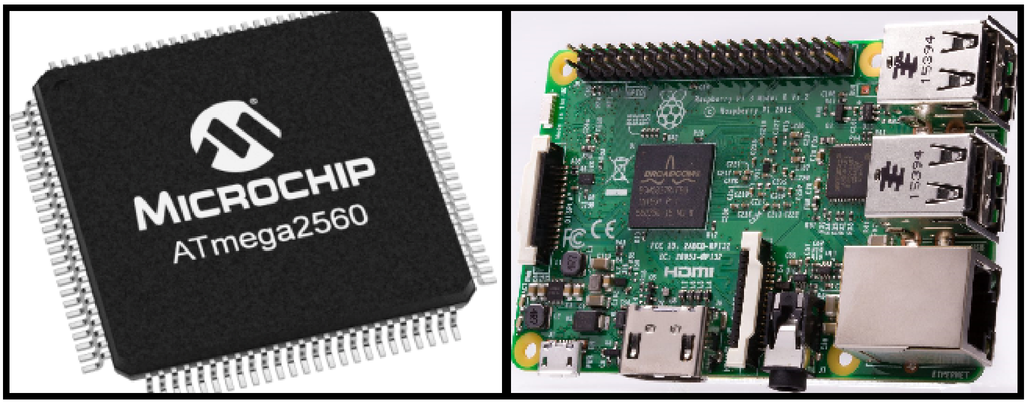

Figure 3.3: ATmega2560 [22], left, and Raspberry Pi 3 [23], right

The PIR or weight sensors output can be read as a digital or analog input, not needing any other processing consideration. On the other hand, voice or vision detection relies on pattern or features recognition. To recognize the first type of signals a microcontroller unit (MCU), such as an ATmega [22], is enough, which allows for a total cost of the system reduction while also reducing the current consumption. The second type of detection requires a microprocessor unit (MPU), such as a Raspberry Pi [23]. The choice of an MPU increases the number of features, like RAM, that are needed in order to be able to run vision or pattern recognition algorithms, however the overall cost of the system is increased, as well as the current consumption. Figure 3.3 depicts both an ATmega and a Raspberry.

3.3 Communication Techniques

The recognition of presence inside the vehicle is meaningless if the driver is not warned. In order to accomplish the proposed system purpose, there is a need to communicate alerts whenever an individual is detected inside the car.

3.3.1 Car Alarm

Considering an integration of the developed system in the vehicle electronics, there are several opportunities to manage the alert communication. The first potential way is triggering the car's alarm. If the system detects presence inside the vehicle immediately after the car is locked, the driver will most likely still be nearby. However, there are several problems inherent to this approach, such as:

- Considering that the system turns ON whenever the doors are locked, if the driver forgets to lock them the system will not start up;
- Considering that the system turns ON whenever the ignition is turned off, if presence is immediately detected the alarm will trigger before the driver is out of the car;
- Considering that the system turns on, for example, ten minutes after the ignition is turned OFF the driver might already be too far away to hear the alarm.

In addition, every time the system falsely detects interior presence the alarm will be turned on. For all these reasons, the driver might decide to entirely turn OFF this system.

3.3.2 Text Message Notification - GSM

It is possible to employ a GSM module that sends a text message to the driver phone. This solution does not rely on the driver distance from the car, it will send the message regardless of the owners position. There are, however, a couple of issues associated with this solution:

- If the driver forgets the phone, this notification is meaningless;
- For this solution to work, it is required a monthly plan or a prepaid plan.

If the phone ends up being forgotten, the driver most likely will miss it and go back to get it since it became such an important item from the everyday life. On the other hand, a periodical payment for the GSM module's SIM card is not an appealing feature that the buyer looks forward to.

While the driver might come back to get his phone, the system might become unappealing since it requires a periodical payment for the GSM service.

3.3.3 Wireless Technologies

Comparatively to activating the car's alarm other options are subtler. For instance, the Bluetooth protocol is a proven technology that is integrated in a large number of car models, which permits the integration of a quieter alarm in the car keys. It is also not restricted to car integration, requiring that the driver carries the alarm alongside the car keys.

Bluetooth is a PAN (Personal Area Network) - an IEEE standard, namely 802.15.1 [24] - but there are several other PAN protocols that could be used instead, such as the standard 802.15.3a High Rate WPAN [25] or the 802.15.4 Low Rate WPAN which requires low power, but it can also be considered WLAN (Wireless Local Area Network) protocols, such as the 802.11a/b/g Wi-Fi [26].

Table 3.2: Wireless technologies for automotive systems comparison [1]

Standard	Bluetooth IEEE 802.15.1	ZigBee IEEE 802.15.4	UWB IEEE 802.15.3a	Wi-Fi IEEE 802.11a/b/g
Freq. band	2.4 Ghz 2.5 Ghz (ver 1.2)	2.4 Ghz	3.1-10.6 Ghz	2.4 Ghz (b/g) 5 Ghz (a)
Network	P2P	Mesh	P2P	P2P
Modulation technique	Frequency Hopping Spread Spectrum (FHSS)	Direct Sequence Spread Spectrum (DSSS)	Orthogonal Frequency Division Multiplexing (OFDM) or Direct- -Sequence UWB (DSUWB)	OFDM or DSSS with ComplementaryCode Keying(CCK)
Maximum network speed	1 Mbps (ver 1.0) 3 Mbps (ver 1.2) 12 Mbps (ver 2.0)	250 Kbps	50-100 Mbps (480 Mbps within short ranges expected).	54 Mbps (802.11a) 11 Mbps (802.11b) 54 Mbps (802.11g
Network range	Up to 100 meters, depending~on radio class (effective 10 meters).	Up to 70 meters (effective 20 meters).	Up to 20 meters (effective 10 meters).	Up to 100 meters (effective50 meters).
Main usage	Voice applications. Eliminating short-distance cabling on radio class	Sensors/control applications. Grand-scale automation. Remote control.	Multimedia applications. Healthcare applications.	Office/home networks. WLAN. Replace Ethernet cables
Strong points	Dominating PAN tech. In vehicles today. Easy synchronization of mobile devices. Frequency hopping tolerant to harsh environments.	Static network. Control/sensor. Many devices/nodes. Small data packets. Low duty cycle. Low power	Easy and cheap to build. Consume very little power. Provides high bandwidth. Broad spectrum of frequencies (robustness).	Dominating WLAN tech. Know-how.
Weak points	Interference with WiFi. Consume medium power	Low bandwidth	Short range. Interference.	Traditionally consume highpower
Automotive usage (potential)	Portable devices. Diagnostics tools. Real-time communications. Device connectivity.	In-vehicle communications. Mobile/static sensor networks.	Robust vehicle communications. High bandwidth communications.	Inter-vehicle communications. Vehicle-to-vehicle. Vehicle-to-roadside.

Table 3.2 was taken from Thomas Nolte and Hans Hansson published article entitled "Wireless Automotive Communications" [1] and shows all these protocols strong and weak points, alongside with their main characteristics.

There are some issues with wireless technologies, such as the maximum range of operation, requiring that the system recognizes interior presence before the driver gets too far away.

Another point to keep in mind is the IoT (Internet of Things) evolution. There is an increasing interest in this field, and the automobile industry is bound to keep up. Tesla, for example, already provides an android/iOS [27] app that allows the driver to control and monitor his/her vehicle.

3.4 Conclusion

Throughout this chapter several technologies that could be chosen to integrate the final implementation are presented. They were arranged in three sections where their strong and weak points were evaluated. It is possible to point out that the motion sensor might be the safest pick, comparatively to the other analyzed options:

- it has low current consumption in comparison with vision or pattern recognition;
- is safer than continuous wave sensor;

- seems to be more practical when compared to the weight sensor, which will be triggered if a slightly bigger object is left on top of the seat.

On the other hand, the coupling of two systems can be a better approach, as was already expected prior to the state of art analysis, since it allows for better current consumption management, using the higher consumption system only occasionally. At the same time higher reliability is also provided, where each technology tries to complement the other. Discarding the continuous wave sensor for its potential health issues, as well as the voice recognition for its low usability, cameras appear to be the most reliable option.

In the communication case, a wireless solution, with an integrated alarm in the car keys seems the best option. However, as a plugin-in solution the GSM module is preferred so that the driver has only to carry the mobile phone.

Chapter 4

System Implementation and Vision Study

4.1 Integrated Solution

The solution presented in this dissertation ultimately aims to allow both low-cost mass production as well as low consumption rates while reliably detecting if someone is inside the car. In order to reach this goal, we propose a two-level detection system that systematically detects in-car presences. This architecture is represented in Figure 4.1.

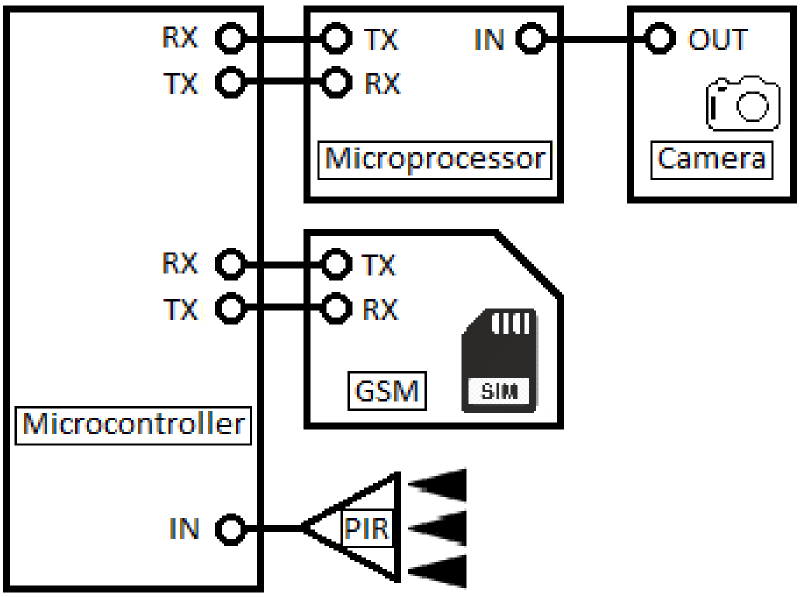

Figure 4.1: Ideal system's electrical circuit

As depicted in Figure 4.1, two inputs provide data regarding the presence of individuals inside the vehicle, through both a PIR sensor and a camera. This system echoes the one present in the paper considered in chapter 2 by Haibin Cai [5], exchanging the accelerometer with a PIR sensor, which is preferable since it is intended to detect motion, and employing a GSM module to generate notifications.

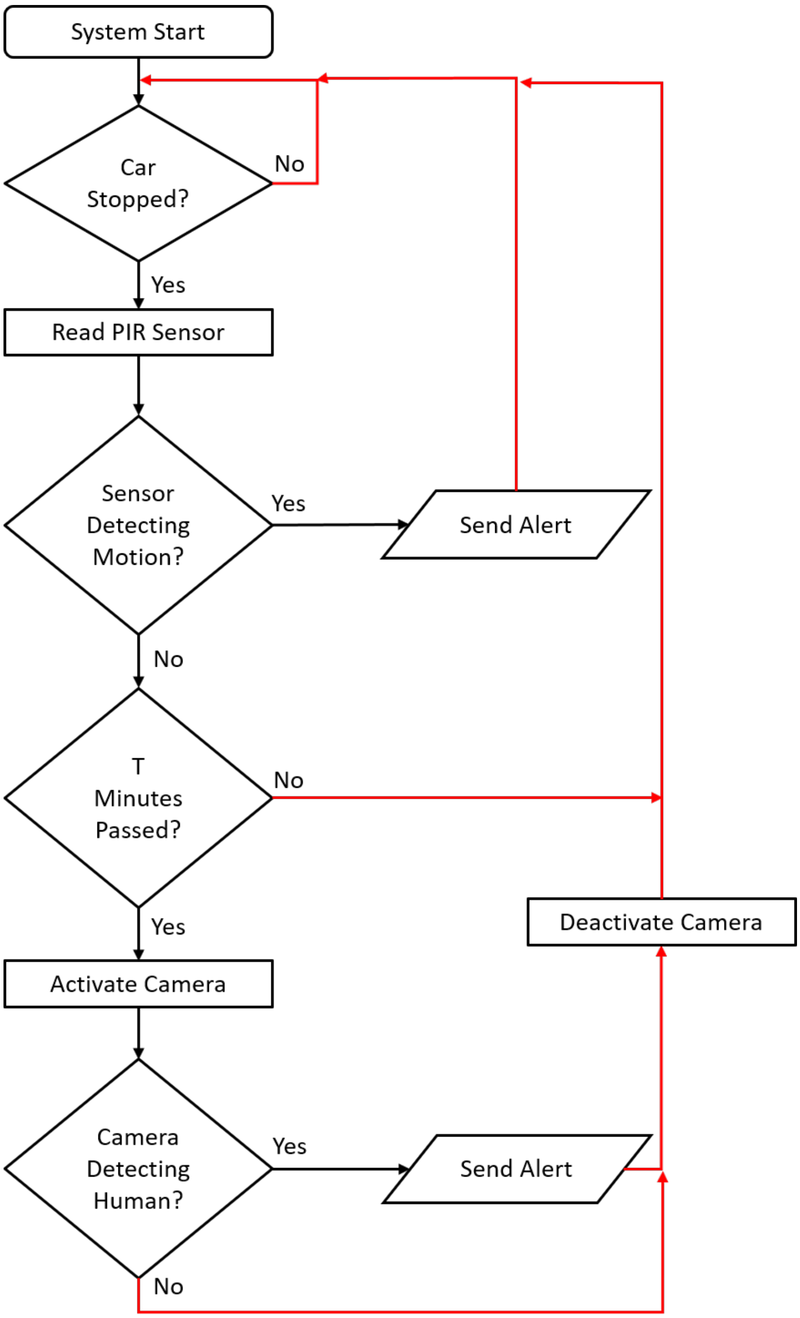

Figure 4.2: Ideal system flowchart

The camera is connected to a microprocessor unit, which will treat the vision information provided by the camera. The PIR sensor, on the other hand, is connected to a microcontroller unit, which will analyze the sensor's signal.

Lastly, the GSM module is connected to the microcontroller unit. This is not mandatory, and the GSM module could also be connected to the microprocessor unit, however, there are advantages in connecting it to the microcontroller. Since the camera consumes more energy than the PIR sensor, and since a single frame should be enough to detect if there is anyone inside the vehicle, there is no need for the camera to be turned ON all the time. Therefore, it is also preferred that the microprocessor only turns ON when the image is going to be processed. By having the GSM module connected to the microcontroller unit, it is possible to send alerts whenever the PIR sensor detects motion without needing the microprocessor branch to be turned on.

The thought-out process of such system is depicted in the flowchart in Figure 4.2 . The primary detection level consists on the PIR sensor, which is constantly receiving information, generating alerts whenever motion is detected.

The second level consists of a camera that captures an image every *10 min*, for example, which is afterwards processed. A single image should suffice to prove if there is in fact a human inside the vehicle, however light issues might make it harder to detect and different time frames might have different light levels.

The logic behind a two-level detection is to get the lowest current consumption while getting the highest detection rate. It is anticipated that both levels, as standalone systems, are capable of reliably detecting infant presence. However the PIR sensor might not be enough when the baby is sleeping, while the camera will consume $160mA$, taking Haibin Cais' [5] system as example, against the $0.3mA$ of some PIR sensors [28] [29].

In this dissertation we propose the first step to the realization of such system, where a low-cost solution is implemented. This low-cost approach consists on the sensor-based solution integrating a PIR sensor and a GSM module as well as a preliminary study on vision detection, where different algorithms are analyzed and compared. The coupling of both the developed system and the vision algorithm requires the acquisition of a camera and a microprocessor unit.

4.2 Low-cost Implementation

The first level of detection idealized in the previous section 4.1 consists in a motion sensor-based solution, much like Fairuz Rashidi and Ikhwan Muhama [3].

Although the approach is the same, their published article presents few conclusive results and no limitations are even discussed. The tests are also omitted which leaves no answer to whether the system was actually experimented in a real environment and is not even mentioned what could possibly further improve it.

For this implementation the following components were selected:

- A microcontroller - Arduino Mega 2560;
- A GSM module - Adafruit Fona 800;
- A motion sensor - Both the analog AMN22111 [28] or the digital EKMC1693112 [29].

4.2.1 Components Selection Rationale

The selection of all these components had a reasoning behind it. The microcontroller does not need to be too powerful, there was no memory or port requirement that would narrow down its choice, so the Arduino Mega 2560 was selected because it was already owned. This means that, in mass production, the control unit selected could be both cheaper and with lower consumption. For example, an ATtiny85 [30], which only has eight pins (minus two for power supply) could be selected for this application, since only four pins are required, one for the sensor and three for the GSM module.

In terms of price, an ATmega2560 unit cost is 10.01 €, with a bulk price for more than 5000 pcs of 7.27 € [22] while the ATtiny85 cost is 0.80 € and with a bulk of 0.58 € for 5000 pcs [30], both in the Microchip store.

Regarding the PIR sensors, two of them were selected - one digital and one analog. Several reasons were taken into consideration for this choice. For instance, the sensor is expected to be installed in the front seat pointing towards the baby seat, so the detecting range should be at least around $100cm$. Both sensors fulfill this requirement, the analog, AMN22111, with $200cm$ and the digital, EKMC1693112, with $220cm$ range.

It is also required that the system has low consumption rate, which both sensors datasheet claim to be between $0.17-0.3mA$ at $5V$.

Since it is intended to be a low-cost implementation, the sensors should be cheap. The analog sensor, AMN22111, cost was 19.47 € [31] while the digital, EKMC1693112, was 10.37 € [32], bought together with 18.00 € shipping cost. As of 29-05-2018 the unit cost is 28.04 € and 9.90 € while a bulk purchase of 100 pcs or more lowers the cost down to 19.12 € and 4.87 €, respectively. After comparing both sensors performances there will be a better term of comparison, assuming both of them to be equally reliable the cost of the digital sensor is considerably lower.

Other than these requirements, both sensors can operate at $5V$, which is helpful since the microcontroller, and the GSM module can both operate at this voltage value.

Lastly, the GSM module was selected keeping in mind that enough library support would reduce both the learning and implementation curve, the Adafruit FONA was chosen, which costs 59.95 $ in the official website [33]. This means that the GSM module could also be cheaper, but at the same time, although this particular set is currently out of stock, there is also a price decrease while buying in bulk.

4.2.2 System Circuit

As previous section system's architecture, the low-cost architecture is similar to its motion sensor branch. The PIR sensor is responsible for detecting in-car movements and will signal it through its output to the microcontroller unit, an Arduino ATmega2560.

The microcontroller in its turn will interpret the motion sensor signal in order to generate the alert messages whenever a positive read is made.

Whenever the alert is generated, the message will be forwarded through the Fona800 GSM module, also connected to the microcontroller.

In 4.3 the electrical circuit is depicted, as well as its electrical connections. The digital PIR sensor output is connected to pin 32 - PD7 - while the analog sensor is connected to the analog input pin 51 - ADC0. The GSM module is connected to port B, since the Fona libraries [34] allow to define whichever pins are desired.

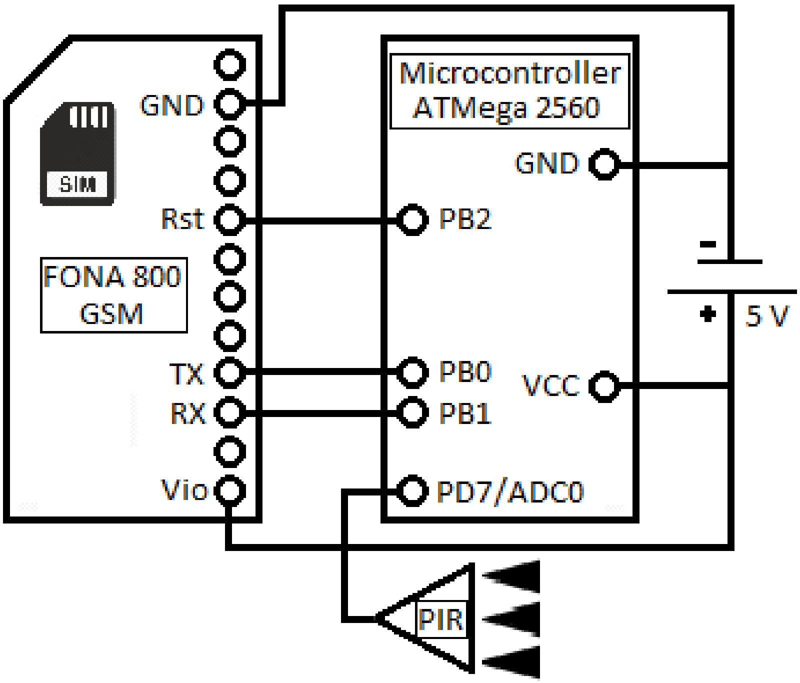

Figure 4.3: Low-cost implementation electrical circuit

To power the system, it was used a 2600mAh power bank. The GSM module requires a specific battery to work, however it cannot be used to power up the whole system. The power bank has a charging rate of 1A and should be connected to the car's USB port or the cigarette lighter socket in order to charge while the car is running.

If the battery goes lower than a predetermined value, the driver should be notified so that the system can be fully charged at home.

4.2.2.1 PIR Sensors

Both selected sensors have some particularities regarding their characteristics and functional operation that are almost the same in both of them.

The analog sensor selected, AMN22111, has several other models with different purposes, for example, it is possible to choose a digital sensor. One of the specifications that varies when choosing a digital model is the current consumption, where two options are provided. The selected model has a standard current consumption, however there are also models with low current consumption as, for example, the AMN43121. This sensor current rate is expected to be around $0.046-0.060mA$. The reason for not selecting such model lays in the fact that the operating voltage is between $2.2-3V$, and we were looking for a $5V$ operating sensor.

Both sensors need time to stabilize after the system is turned ON, the analog AMN22111 time is less than $45s$ while the digital EKMC1693112 is up to $30s$.

The detection range of the selected sensors, both slight motion detector, is around two meters, while other models range can go up to ten meters. The detection range is also defined by a conical shape, as partially represented in Figure 4.4. For the analog sensor, the conical shape has an amplitude of $91°$ ($\pm45.5°$) while the digital has only $44°$ ($\pm22°$).

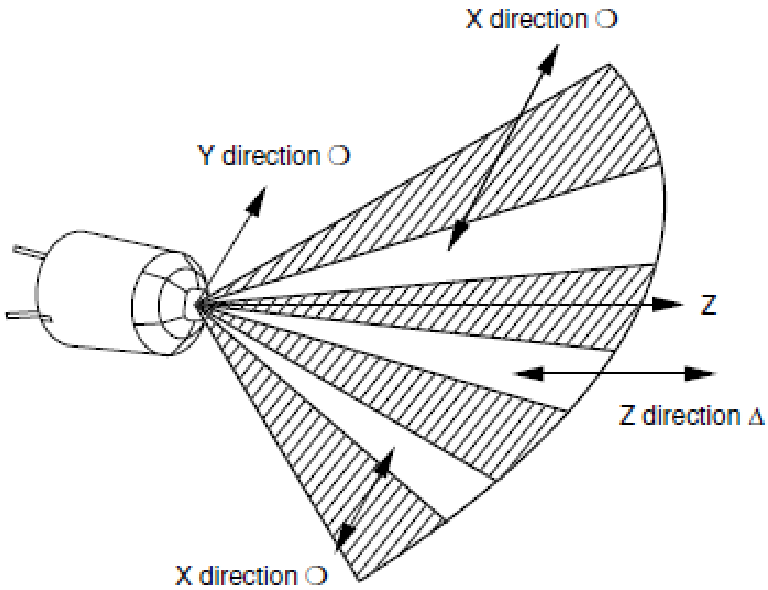

Figure 4.4: PIR sensor detection area [28]

The PIR sensors' performance depends heavily on the differences of temperature between the ambient and targeted object. In order to be able to detect an object, its temperature needs to be different from the ambient. For example, during sunny days, where the temperature rapidly rises inside the car, this sensor's reliability decreases, reaching its minimum when both temperatures are equal, on account of the sensor not distinguishing the object from the surroundings.

The selected analog sensor does not document any information about temperature differences in the datasheet. The digital sensor expresses its conditions for characteristics measurements, one of them being the temperature differences during tests. For the slight motion detector, the temperature difference was $4°C$.

Lastly, there are two other relevant conditions regarding the target: the movement speed and size. Both sensors' datasheets state that they can measure gestures of objects sized $200x200mm$, for example a human head. The digital sensor claimed target motion speeds vary between $0.3-1.0m/s$. The analog sensor indicates that the measurement conditions dwell on $0.5m/s$ motions.

4.2.2.2 Fona800

Fona 800 requires several peripherals in order to work, for this reason it was bought the pack which come with all necessary accessories instead of the Fona alone.

The module needs to be connected to a microcontroller which the datasheet recommends Arduino, because of the supported libraries, which can be found in their GitHub repository [34]. However any microcontroller with an UART interface is enough.

The most important module pins are the Vio, Tx, Rx and Rst. The Fona 800 Vio pin has to be connected to the microcontroller's logic voltage, since it is the pin that defines the Fona logic level, either $3V$ or $5V$. Rx and Tx pins are used for communication to the GSM module and from it, respectively. The Rst pin is the module hard reset.

It is also required a Lipoly battery of $500mAh$ or larger to withstand $2A$ spikes. The module has a micro USB port that charges the battery, however its design does not work without the battery. It is possible to keep the module charging through the micro USB port, however it does not work without the battery.

An antenna is also required. Two types of module builds are available, one with an SMA input, present in the chosen module, and another with an uFL input. The later requires an uFL to SMA adapter.

The module has also several LEDs for easy information about the module status:

- A blue LED that signals that the module is running;
- An orange LED that signals that the module is charging;
- A green LED that signals that the module is fully charged;
- A red LED that signals the network status.

The red LED is not always lit, it blinks at different intervals signaling in which mode the module is on, making it easy to check if the SIM is registered to the network or not.

The communication with the GSM module is done through *AT commands*. There are several commands for all the specific GSM control. For the current application only two At commands are actually relevant:

- *AT+CMGS* - command for sending messages;
- *AT+CBC* - command for battery voltage check.

Both these features are of utmost importance, the notification has to be sent and knowing the module's energy permits warning in case the battery is excessively low.

There are other commands that could be potentially relevant, especially the command *AT+CMGR* which permits messages to be read by the module and therefore will allow the user to interact with the device sending structured SMS texts.

During the module setup it was encountered a problem that prevented the SIM card to register in the network. This problem was eventually solved by disconnecting the battery while connecting the module to charge, a non-intuitive solution that was not found acknowledged by Adafruit.

4.2.3 System Algorithm

The algorithm that was developed can be seen in image 4.5 flowchart. Whenever the system starts, the first task is checking if it is charging. The system will be connected to the car's lighter charger or USB port. Whenever the car is turned OFF, these ports are shut down, meaning the system will no longer be charging. This is the moment when the system should get online.

The next task to execute is checking the battery levels, if the value is too low an SMS is sent to the car's owner, notifying that the vehicle charge no longer suffices and that it should be fully charged. The battery levels check occurs every T minutes.

Finally, and continuously, the PIR sensor signal is read, and whenever a motion is detected an alert is sent to the owner's phone, notifying the presence of individuals inside the vehicle. The alert is not made, however, every time the sensor signals the detection.

Since there is a sensor reading every $150us$, approximately 6667 readings per second, the alert is only generated if in a two second time frame the sensor is activated, at least, for one second. This allows a false positive rate reduction.

The tests that were made, discussed further ahead, show that other vehicles crossing near the parked car do not generate enough trepidation to trigger the SMS alert, so further investigation might show that this limitation, not responding to every impulse, might not be necessary and even limiting the system's reliability.

4.2.4 Tests

Several tests were made to evaluate this system's performance. The first consisted in measuring the current consumption of the system. Each sensor current consumption was verified, and it was noted that both sensors are in range of the datasheet specification of $0.17 - 0.3mA$.

While charging, whenever the car is running, the GSM battery will draw up to $500mA$, its measured current was $300mA$. The power bank specifies a current consumption up to $1A$, the measured value was $700mA$. This means that the system batteries will draw up to $1.5A$, and the measured average is $1A$.

The Fona has a current draw of $25mA$ average, that can go up to $200mA$ while sending or receiving messages. This current is drawn from its own battery.

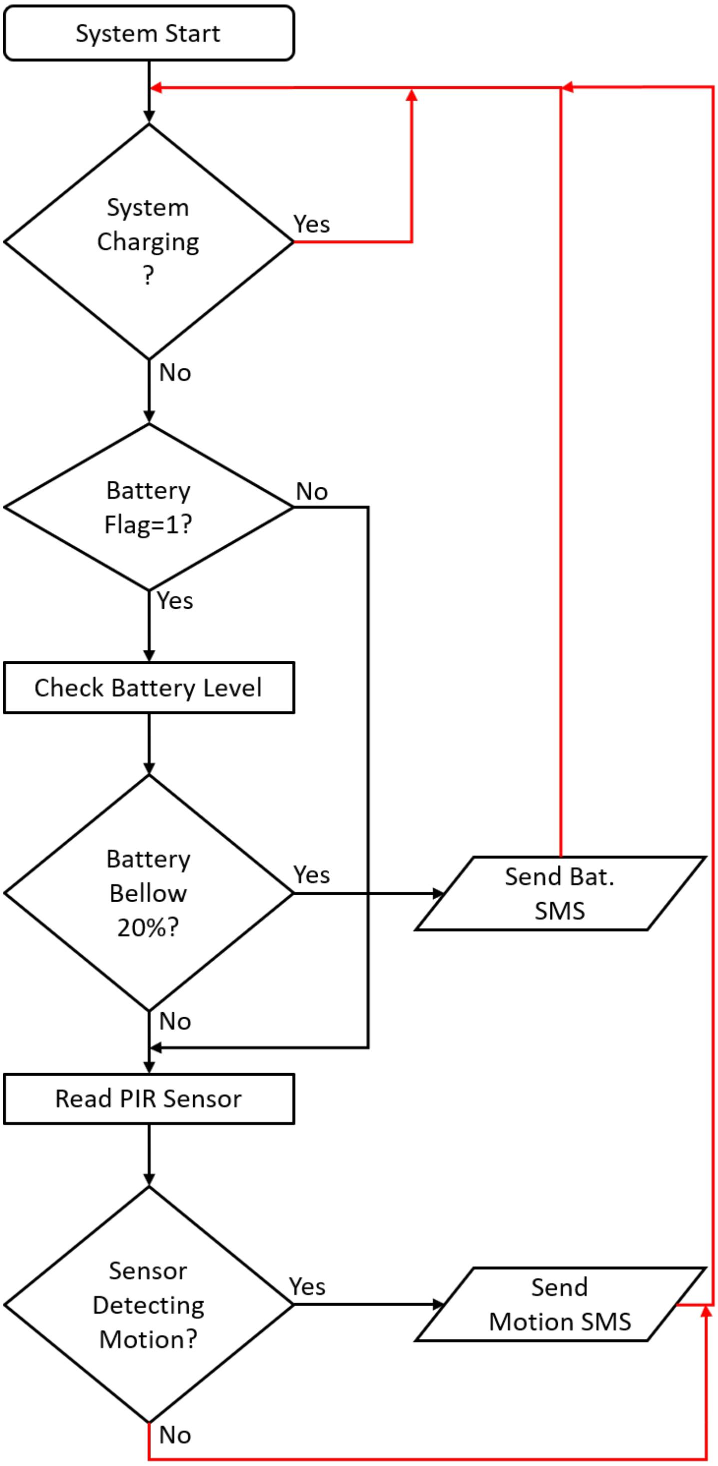

Figure 4.5: Low-cost implementation flowchart

Each sensor draws up to $0.3mA$ while detecting motion. The Fona also has a small impact in the power bank consumes, drawing around $3.7mA$. The ATmega 2560 has a current consumption of $10mA$ at $5V$ and $8MHz$. The expected current draw of the system is $39mA$.

Since the power bank has $2600mAh$, supplying $14mA$ it should last up to $185h$. The Fona battery, on the other hand, has $1200mAh$, and supplying $20mAh$ should last up to $60h$.

This is considering that no losses exist, meaning that the actual duration should be lower.

To further test this implementation, it was necessary to investigate how reliable the system would be in a real environment. In order to verify its robustness, the system was activated in a parked car to see if other vehicles passing by could trigger the sensor. The car was parked in three different environments, and in neither of them the alert message was generated:

- Underground park, Norte Shopping from 20 h to 24 h;
- Tar road, FEUP student parking lots from 12 h to 13.30h;
- Cobblestone road, Cathedral of Bragança in a semi-controlled environment.

To showcase the prior tests, we recorded a couple of clips of the tests while the car was parked in cobblestone. In one of the clips three cars cross at approximately 40 km/h without triggering the sensor.

A scenario with a small doll was also tested, in order to verify if the sensor was able to detect small gestures, and that it is even activated. When the motion of the doll is detected, the SMS alert is received in a smartphone, also present in the video.

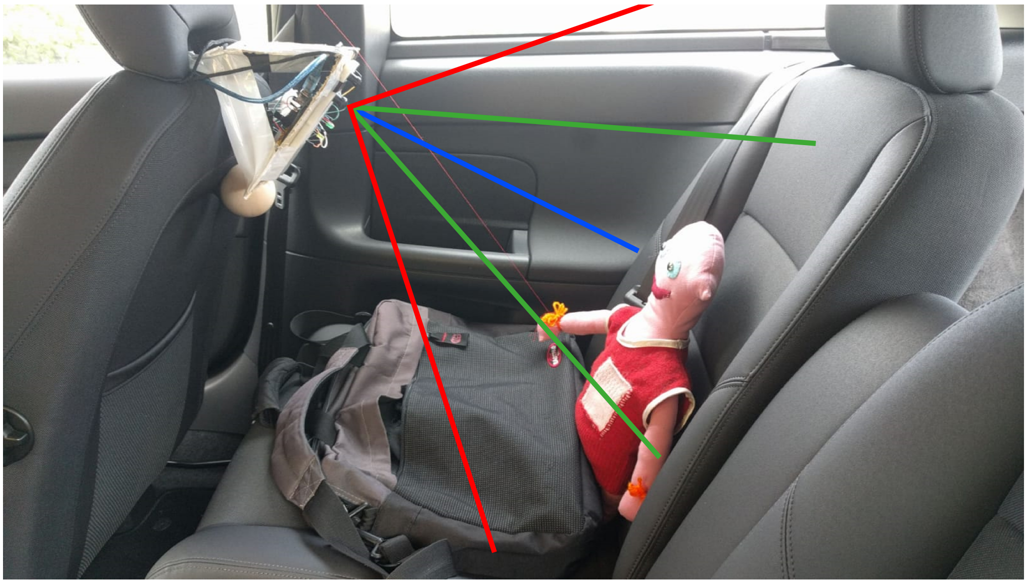

Figure 4.6: Experiment illustration

In this clip the system is not mounted in the front seat, which might mean that the trepidation is less felt. However, in a posterior test, with the system fully mounted on the seat, the experiment was repeated with the same results.

Although it was not verified with trucks crossing it is expected that the system also ignores the generated trepidation.

There is, however, a false negative problem, where the system does not trigger under motion. It was also recorded a clip showcasing this behavior. The failed detection occurs when the doll's arm moves in the sensor's direction.

These tests, and particularly the latter, require an actual baby to further provide correct information of the sensor's reliability. It is necessary to note that the doll's temperature might be too close the vehicle's one, which might merge into the surroundings. To try to contradict this test lack of faithfulness, we tried to heat a small portion of tin that was placed in the doll's arm, however the test was also not positive and not entirely reliable.

Figure 4.6 illustrates the setup used for the tests. The red lines indicate (approximately) the analog sensor detection range ($91\,^{\circ}$) and the green represent the digital one ($44\,^{\circ}$). The approximate distance from the sensor to the doll is $55cm$.

In this figure the system is mounted on the front seat, it is expected to be the optimal place to mount it, because of both the the angle formed and the smaller distance, while also being easier to mount. However further testing with the system coupled to the ceiling should also be experimented.

4.3 High-cost Implementation

Alternately to the sensor detection, the adoption of cameras seems to be a better option going forward, as was stated when reviewing its state of art. The employment of cameras for real time applications with acceptable accuracy is already possible, and as the vision algorithms improve, so does its reliability.

Currently both Mask R-CNN and YOLO are some of the best vision algorithm choices, so in this chapter we describe the tests performed in order to decide which is more reliable.

This section also displays other algorithms that were experimented.

4.3.1 Background Subtraction

Even though YOLO and Mask R-CNN are the state of the art in what vision detection concerns, there are a couple of other approaches that were experimented. The first one was *background subtraction* which consists in subtracting the numeric value of an image from another. This method allows to observe changes between a succession of frames from a video capture, which implies that an object moved. It is also possible to generate a background and compare it to a frame of the video, which enables the detection of objects not present in the background.

This method [35] implies several assumptions, mainly that both the background and camera are static, while the foreground objects are moving.

The process can be described as:

- Building a background model;
- Creating a mask of foreground pixels based on the distance between current frame pixels values and the background;

- Updating background model.

This method was tested resorting to the Chris Dahms algorithm [36] and allowed a better perception of this approach. Several obstacles are associated to this specific method, in this dissertation scope the main problem is when the infant is sleeping, therefore static. In this situation the child will eventually be considered part of the background. Even using a frame subtraction approach, not relying in the background, this problem persists, in addition to others such as illumination changes or camera shaking.

4.3.2 Haar-cascades

The other method that was tested prior to YOLO and Mask R-CNN was also facial detection, resorting in the Viola-Jones [9] haar-cascades method, that was briefly tackled in chapter 2.2.

The employment of facial recognition methods can mitigate background subtraction problems since it doesn't rely on static environments nor cares if the object is moving or not.

The base algorithm that was tested [37] handled the detection of two features, namely the face as a whole [38] and both eyes [39].

This method was tested using stock images of adults and seemed to be a reliable approach. However, it was also created a small dataset of small children, and in this specific case the results were way more disappointing. One of the problems arises from closed eyes. As explained in 2.2 there is a comparison of light between the eyes and the nose or upper cheeks, so whenever the eyes are closed the light difference is almost nonexistent. This problem was also visible with adults to a lesser extent. However, even if the infants were with both eyes open there were still several images where no detection was achieved. The solution that was thought-out was an haar-cascade training with an infant dataset. This solution was not able to solve this problem, which could be due to the small training dataset, as well as the image discrepancy, since there was no norm to the babies' facial expression or position.

4.3.3 YOLO - You Only Look Once

As stated in chapter 2.2, since 2012, CNN has become the state of art of object detection. Out of the presented approaches, two of them were selected and analyzed as potential solutions for infant detection - Mask R-CNN and YOLO - since these are the most recent developments and, therefore, the most reliable ones.

The first algorithm tested was YOLO. However instead of using the original implementation [40], we employed a translation [41] that runs with TensorFlow [42], "an open source machine learning framework for everyone".

In order to set everything up a small dataset was gathered and a simple script was developed to check both the bounding boxes and confidence factors. One of the most relevant top-level parameters was the threshold, which defined if the confidence factor in each image was enough to create the expected bounding boxes. This parameter was set to 0.3, the recommended default value, which proved enough for the majority of the dataset – but not for all of it.

4.3.3.1 Procedure

After the first contact with the algorithm there was a necessity of augmenting the data, by both gathering more images online as well as a selection and manually edition of some of the pictures with the following objectives:

- Darkening the picture, simulating a lightless environment, like an underground parking or in the night;
- Brightening the picture, simulating an intensive daylight;
- Adding Gaussian blur to the picture;
- Adding general noise to the picture;
- Creating black boxes obstructing the eyes, mouth or nose of the infant, simulating, for example, a blanket – these boxes were also colored with pink and blue.

This script was running in a GeForce GT 740M GPU, taking up to two to three minutes to run the script and processing each image almost instantly. As expected, the majority of the edited images got their confidence levels lowered, however some of the dark edited pictures tended to present a higher confidence value and the brighter a lower one.

Figure 4.7: YOLO image brightness variation confidence, left to right - 0.71; 0,77; 0.47

In Figure 4.7 an original image is presented, as well as its darker and brighter edition. As stated before, the darker picture presents a higher confidence factor, while the brighter picture's confidence is drastically decreased by 0.25 points.

In the presented example, this value is meaningless, however, some of the selected images present a confidence factor of 0.3 or less. With such a confidence decrease not even a 0.1 threshold could recognize the child.

Also, worth noting, ten pictures of infants were edited with boxes in certain facial parts in three different colors - black, pink and light blue - to check if it would influence the confidence evaluation in a perceptive way, as depicted in Figure 4.8. Out of the ten edited images, six had higher confidence with black boxes, and four with the pink box. The majority of the editions did not get a relevant difference of confidence across each color. However, one of them turned impossible to detect, with a blue box edition, and three presented differences above 0.1 points.

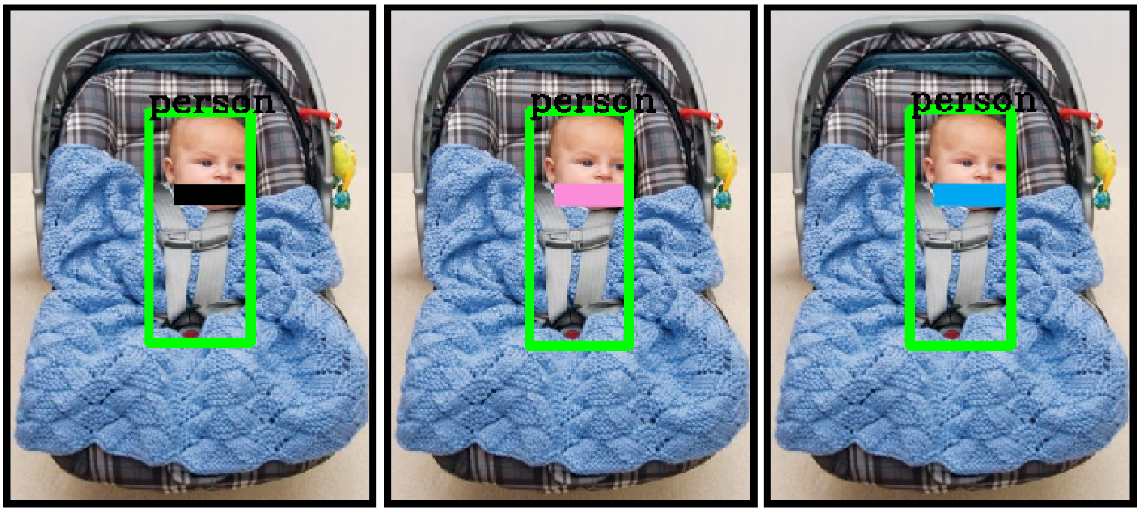

Figure 4.8: YOLO colored boxes confidence, left to right - 0.33; 0,28; 0.19

Also, worth noting that none of the black box editions got bellow 0.2 confidence value, while one pink and two blues did.

As for the Gaussian blur and general noise editions, the results are depicted in Figure 4.9. Out of the ten edited images the results were varied. As Figure 4.9 suggests, the majority of the edition had lower confidence values than original, with three exceptions: one from the Gaussian edition with higher confidence value and the other two from the general noise edition.

The results are not very conclusive, aside from two of the images all the other edited counterparts had confidence levels discrepancies higher than 0.2 points. There are also mixed results of which edition has the higher confidence values. From these results it is possible to deduce that this edition largely depends of the original photo.

The last notable result that can be drawn, is that out of the ten original pictures none had a confidence level lower than 0.2 and out of the Gaussian editions four became lower than this threshold, and for the general noise only two got this low.

4.3.3.2 YOLO Reliability

To better perceive this algorithm's reliability for infant detection, the script was modified to create a CSV file gathering the confidence value for every picture in the set.

Figure 4.9: YOLO normalized, Gaussian and general noise edition confidence, left to right - 0.64; 0,55; 0.69

Table 4.1: FP, TN, FN and TP results of the YOLO algorithm for four levels of confidence (0.1, 0.2, 0.3, 0.4) with G gamma manipulation

	Confidence level	0.1		0.2		0.3		0.4	
G=0.5	FP	327	57.6 %	209	36.8 %	133	23.4 %	81	14.3 %
	TN	241	42.4 %	359	63.2 %	435	76.6 %	487	85.7 %
	FN	9	3.1 %	37	12.8 %	67	23.2 %	98	33.9 %
	TP	280	96.9 %	252	87.2 %	222	76.8 %	191	66.1 %
G=0.8	FP	272	47.9 %	164	28.9 %	112	19.7 %	72	12.7 %
	TN	296	52.1 %	404	71.1 %	456	80.3 %	496	87.3 %
	FN	7	2.4 %	34	11.8 %	61	21.1 %	89	30.8 %
	TP	282	97.6 %	255	88.2 %	228	78.9 %	200	69.2 %
G=1.0	FP	258	45.4 %	153	26.9 %	96	16.9 %	52	9.2 %
	TN	310	54.6 %	415	73.1 %	472	83.1 %	516	90.8 %
	FN	10	3.5 %	42	14.5 %	64	22.1 %	90	31.1 %
	TP	279	96.5 %	247	95.5 %	225	77.9 %	199	68.9 %
G=1.2	FP	245	43.1 %	136	23.9 %	78	13.7 %	48	8.5 %
	TN	323	56.9 %	432	76.1 %	490	86.3 %	520	91.5 %
	FN	10	3.5 %	42	14.5 %	72	24.9 %	97	33.6 %
	TP	279	96.5 %	247	85.5 %	217	75.1 %	192	66.4 %
G=1.5	FP	226	39.8 %	119	21.0 %	68	12.0 %	40	7.0 %
	TN	342	60.2 %	449	79.0 %	500	88.0 %	528	93.0 %
	FN	13	4.5 %	45	15.6 %	76	26.3 %	111	38.4 %
	TP	276	95.5 %	244	84.4 %	213	73.7 %	178	61.6 %

Table 4.1 contains the false positive (FP) and false negative (FN) results obtained, as well as true positive (TP) and true negative (TN), for a total of 568 pictures of car seats (or child seats –

relevant background) and 289 pictures of infants (edited pictures included).

As previously stated, the darker edited pictures presented a tendency to have higher confidence, while the opposite was also observable, with brighter ones presenting a smaller confidence, forwarding the next round of test to an image light manipulation within the script. This experiment was run with a gamma variation of both 1.2 and 1.5 of the original images, which results are also present in the same table.

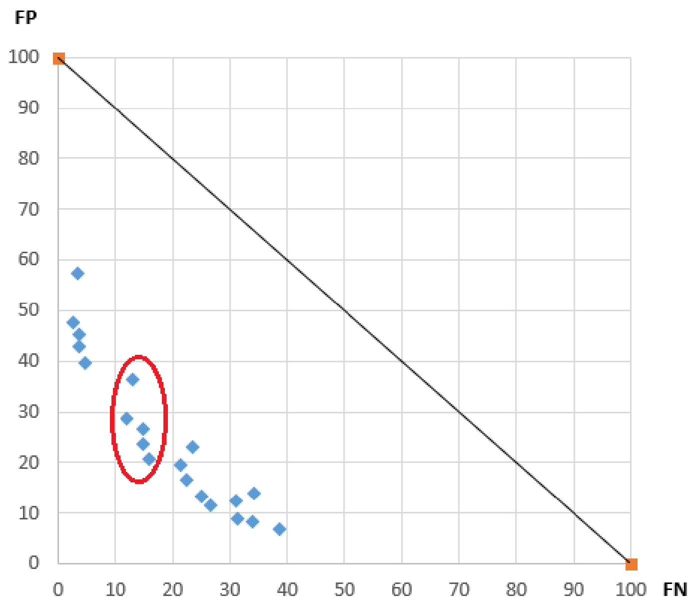

Figure 4.10: YOLO ROC Results

Since some improvement was noticed while darkening the input images, the opposite was also tested, with gamma values lower than the original images, of 0.5 and 0.8, that are also displayed in table 4.1.

From all the represented experiments there is a high result variance, from 2.4% to 38.4% FN rate as well as 7% to 57.6% FP rate (noting that the better FP cases have almost the worst FN results).

Figure 4.10 illustrates the results with the FP and FN ROC distribution. For this specific application it is particularly important that the children present inside the vehicle are detected, therefore having as little FN results as possible, meaning that values closer to the left axis are preferred.

Assuming that the FP results can be dismissed, the best result obtained would be with a 0.8 gamma variation and 0.1 confidence level. However, this means that in 47.9%, almost half the cases, there will be an incorrect positive read.

In table 4.2 are present the precision, recall and accuracy values, to better perceive which results are ideal.

It is possible to understand the trend that with precision gains, the recall will decrease, as the amount of total positive detection increases, the higher the amount of false positive reads. The

Table 4.2: Precision, recall and accuracy of the YOLO algorithm for four levels of confidence (0.1, 0.2, 0.3, 0.4) with G gamma manipulation

	Confidence	0.1	0.2	0.3	0.4
G=0.5	Precision (%)	46.1	54.7	62.5	70.0
	Recall (%)	96.9	87.2	76.8	66.1
	Accuracy (%)	60.8	71.3	76.7	79.1
G=0.8	Precision (%)	50.1	60.9	67.1	73.5
	Recall (%)	97.6	88.2	78.9	69.2
	Accuracy (%)	67.4	76.9	79.8	81.2
G=1.0	Precision (%)	52.0	61.8	70.1	79.3
	Recall (%)	96.5	85.5	77.9	68.9
	Accuracy (%)	68.7	77.2	81.3	83.4
G=1.2	Precision (%)	53.2	64.5	73.6	80.0
	Recall (%)	96.5	85.5	75.1	66.4
	Accuracy (%)	70.2	79.2	82.5	83.1
G=1.5	Precision (%)	55.0	67.2	75.8	81.7
	Recall (%)	95.5	84.4	73.7	61.6
	Accuracy (%)	72.1	80.9	83.2	82.4

accuracy simply reflects the hit rate.

It is possible to obtain the following range of rates:

- Precision - 46.1% to 81.7%;
- Recall - 61.6% to 97.6%;
- Accuracy - 60.8% to 83.4%.

The previous best FN case, with 0.1 confidence level and 0.8 gamma manipulation, achieves 50.1% precision, 97.6% recall and 67.4% accuracy. This case has the second worst precision levels while having the best recall, as should be expected, while achieving almost an intermediate level of accuracy.

Even if these TP results are the best, and therefore preferred, the amount of FP is too high for a commercial application. The following results, referred in Figure 4.10 with a red circle, present results with approximately 20% FP reduction for a 10% FN increase.

From these values, the best candidates are:

- *A* - 0.8 gamma and 0.2 confidence, with 11.8% FN rate and 28.9% FP;
- *B* - 1.5 gamma and 0.2 confidence, with 15.6% FN rate and 21% FP.

Both *A* and *B* candidates achieve higher accuracy and precision values than the previous considered option, with an accuracy of 76.9% and 80.9% respectively. As for the precision and recall, *A* has 60.9% and 88.2% while *B* has 67.2% and 84.4%.

As expected, the recall is decreased, but A and B still achieve a TP rate of 88.2% and 84.4% respectively.

4.3.4 Mask R-CNN

The Mask R-CNN algorithm [43] was also tested as a term of comparison.

The GPU that was being used for the YOLO algorithm was not enough to perform this task and running it in the CPU is a much time consuming process, so we used another computer that possessed a GTX 550Ti. Even though it was possible to run this algorithm, the processing time for each image took longer than with YOLO.

Table 4.3: FP and FN results of the Mask R-CNN algorithm for two levels of confidence (0.7, 0.8) with G gamma manipulation

	Confidence level	0.7		0.8	
G=1.0	FP	99	17.5 %	80	14.1 %
	FN	91	33.8 %	142	52.8 %
G=0.8	FP	99	17.5 %	72	12.7 %
	FN	93	34.6 %	128	47.6 %
G=1.5	FP	79	13.9 %	64	11.3 %
	FN	104	38.7 %	151	56.1 %

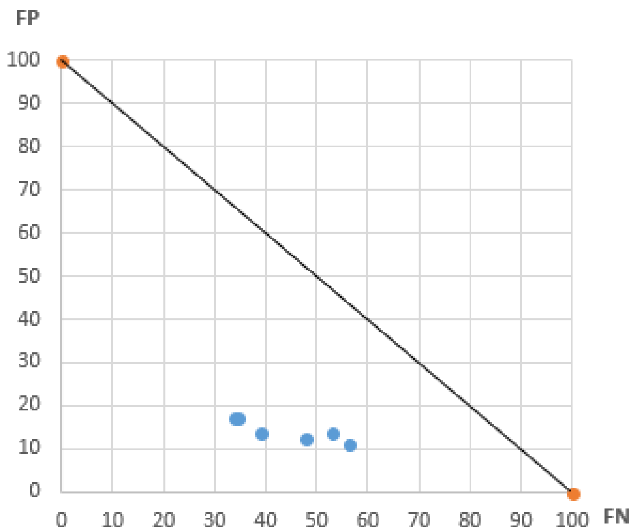

Figure 4.11: Mask R-CNN ROC Results

The adopted approach was similar to the YOLO one. After setting up Mask, a script was developed that generated a confidence CSV file of the same dataset. The results obtained are presented in table 4.3. It was also created a brightness variation which provided the following results, as shown in the same table.

A brief observation of the tables results easily give the perception of being inferior to the YOLO ones'. The discrepancy can be noted especially with the FN rate results being superior to 30%. Figure 4.11 better illustrates the results with the FP and FN ROC distribution, allowing a better perspective of the tables data.

4.3.5 CNN Approaches Comparison

Since a non-detected child might cause a casualty, it is considered primarily preferred to have a higher number of FN than FP. For instance, with a gamma variation of 0.8 and 0.2 threshold, the parameters for the second lowest FN rate, it is still possible to miss around 1/10 infants inside the car, while detecting them in 1/3 of the situations where no one is there.

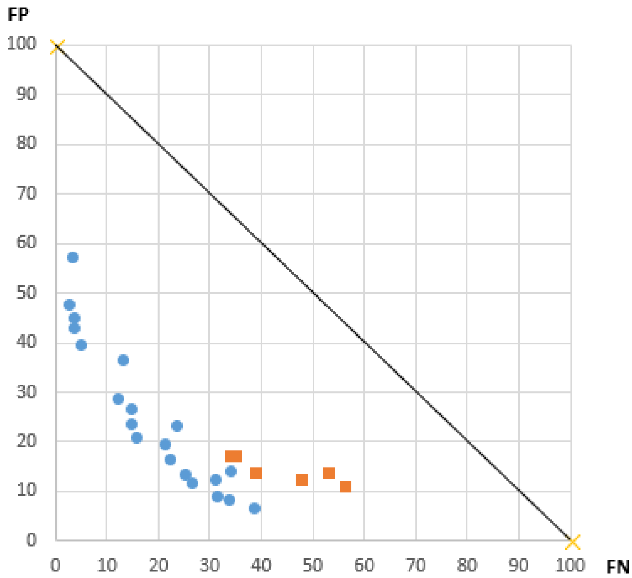

Figure 4.12: Mask R-CNN (square orange) and YOLO (circle blue) ROC comparison

As Figure 4.12 illustrates, the FN values are much higher with Mask R-CNN, especially considering that it is more critical for this specific application than the FP.

Both of these algorithms were tested with the pre-trained models for the “object” person, which may contain a small number of – or even none – infants. Noting that the FN and FP ratios should be lower, it was anticipated that a specific training of the neural network could potentially improve these values substantially.

4.3.5.1 YOLO Training

As the results provided by both algorithms accused higher accuracy from YOLO, it was decided to perform a CNN training in order to try to increase the system’s reliability for detecting children.

To accomplish this task, neither of the previously used GPUs was powerful enough since the training with these would takes weeks to perform, and so another solution was required.

Google Cloud Platform is an online platform that allows users to rent GPUs for several different purposes and therefore it was selected to realize the training.

With this tool it was possible to create a virtual machine renting a Nvidia K80. The same YOLO algorithm was again selected, employing once again the darkflow algorithm [41] with TensorFlow [42].

Table 4.4: FP and FN results of the tiny YOLO algorithm for four levels of confidence (0.1, 0.2, 0.3 and 0.4) with G gamma manipulation

	Confidence level	0.1		0.2		0.3		0.4	
G=0.5	FP	202	35.6 %	163	28.7 %	112	19.7 %	68	12.0 %
	FN	53	18.3 %	100	34.6 %	148	51.2 %	183	63.3 %
G=0.8	FP	200	35.2 %	151	26.6 %	102	18.0 %	70	12.3 %
	FN	39	13.5 %	74	25.6 %	118	40.8 %	164	56.7 %
G=1.0	FP	199	35.0 %	149	26.2 %	102	18.0 %	64	11.3 %
	FN	44	15.2 %	73	25.3 %	115	39.8 %	164	56.7 %
G=1.2	FP	214	37.7 %	155	27.3 %	107	18.8 %	66	11.6 %
	FN	44	15.2 %	74	25.6 %	117	40.5 %	158	54.7 %
G=1.5	FP	220	38.7 %	156	27.5 %	102	18.0 %	56	10.0 %
	FN	48	16.6 %	78	27.0 %	117	40.5 %	157	54.3 %

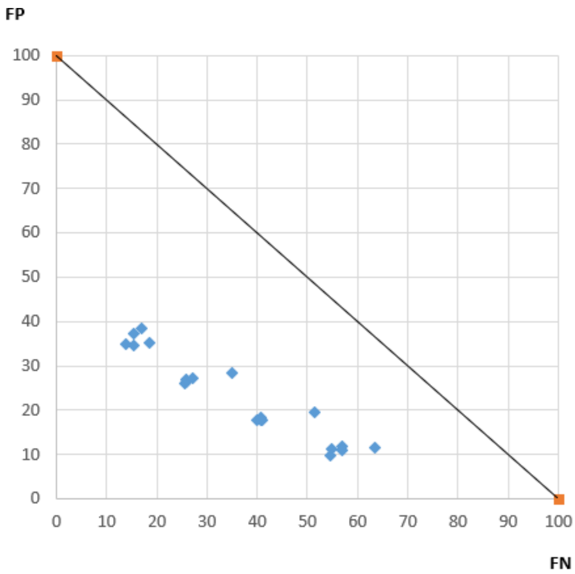

Figure 4.13: Tiny YOLO ROC Results

Two trains were realized, the first one resorting with the original YOLO weights and the second with the tiny YOLO one's. The reason for the second train lies in the fact that training on the original YOLO requires a lot more time, since the CNN has more layers. For this reason, the first train, although it ran for roughly a day, was barely capable of detecting the original unedited images.

With tiny YOLO, the original reliability was already worst than it's original counterpart, however it better illustrates what was expected of a full fine tune of the original weights. In table 4.4 are presented the tiny YOLO FP and FN rates to later get a better term of comparison.

Figure 4.13 represents the FN and FP table values. It is easily observable that the results are worse than with either Mask or YOLO. It is also possible to admit that the most interesting result, the most left one, is for a confidence level of 0.1 and 0.8 gamma variation.

In image 4.14 the results are displayed.

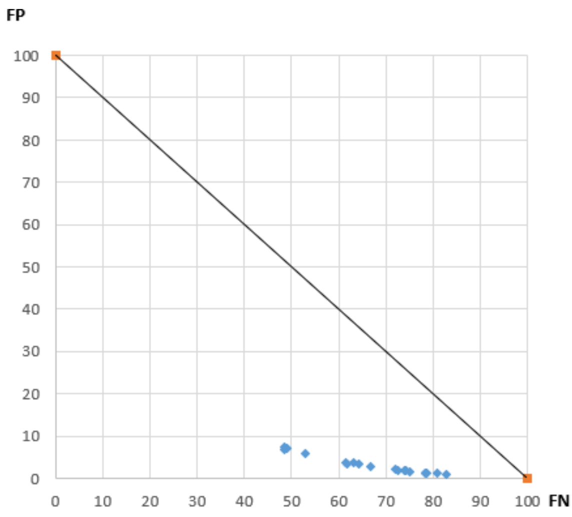

Figure 4.14: Tiny YOLO ROC Results

These results are the culmination of several training approaches. As we can see they are not what is expected, since they don't present an improvement to the original weights.

To obtain these results it was made a fine tuning of the CNN while only training the last layer, following darkflow's issue 486 [44]. With this approach, it was expected that the model gave a response similar to the tiny YOLO. It was also not expected a big improve, due to several reasons:

- Small training dataset, six hundred images is not enough data for a meaningful train;
- Poor image systematization, the images do not follow any sort of pattern, they were taken as both frontal and lateral, awake and asleep;
- Poor image quality, some images resolution might be too low for a meaningful learning.

The aggregation of all these points is a constraint for substantial improvement, however should not have led to such performance decrease, especially considering that only the last layer is being trained.

4.4 Conclusion

The low-cost solution, which was implemented, proved to be a reliable solution. From the experiments it is possible to conclude that falsely detecting in-car movements is highly unlikely.

Even though the reliability of this solution was not tested in the presence of a small sleeping children, it is expected that the system reliably triggers. However, further testing regarding this issue is mandatory.

As for the high-cost solution, the study also provides some degree of reliability. However, from the obtained results, where one in ten children will not be detected, while three in ten triggers will be from false recognition, it is possible to conclude that there is still margin for improvement.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

Throughout this dissertation small conclusions have been established. In this chapter, a consolidation of all the work is presented.

First, it is necessary to note that the developed system will not be integrated in the vehicle's electronics and it is intended to be a generic plug-in technology that can be installed in any car at any time. However, a fully integration of this system in the vehicle is not to be discarded.

It is also worth retaining that the system will draw energy from the vehicle's battery, it will be charged through the USB or the cigarette lighter while the car is driving, which might not allow to fully charge the system in small trips. On the other hand, an integration of the system in the vehicle's electronics is able to charge the system even after turned off, meaning that a lower current consumption system is preferred, preventing the battery's energy to be fully drawn. From this point of view, a low consumption option is desired, which happens to also be the lower-cost solution. Even though a lower cost is indeed preferred, several automobile brands have opted for cameras instead of sensor, for example, in parking support systems, which is a cost latter financed by the consumer.

An integration of the system in the vehicle's electronics could also provide a better alert solution with an alarm in the car keys. As a plug-in application, the GSM module is preferred in order to fully provide notifications independently of the driver's distance since it is not needed to carry any other device other than the phone. On the other hand, the integrated solution could have a wireless car keys alarm, which the driver will always have to carry.

These are the main aspects taken into consideration throughout this dissertation, culminating in the final presented implementation. As for the possibly better approaches, section 5.2 slightly tackles this issue.

For the actual implemented system, the low-cost approach seemed to be a reliable solution for in-car infant detection, as the Figure 5.1 taken from one of the recorded clips suggest. However, as pointed out before, further experimentation, especially with newborn children is of utmost importance.

Figure 5.1: SMS received upon low-cost system motion detection

As for the vision detection results it seems that it still might not be a stand-alone solution specially because of its high false positive - false negative trade off. Further CNN training with a large dataset specifically dedicated to small children inside vehicles would, most likely, highly improve the algorithms accuracy.

Finally, it is expected that the coupling of these two systems will improve the system's reliability.

5.2 Future Work

To further improve and achieve a final commercial product, several aspects should be considered, for instance:

- Further testing of the low-cost approach - as is, the implementation lacks the real target experimentation. There is a necessity to observe the system's behavior with a newborn child, specially while asleep;

- Further low-cost approach - obtain cheaper components (for example, as referred in chapter 4.2, ATtiny85) with the lowest power consumption (for example, as referred in chapter 4.2 the PIR sensor AMN43121) and check if the system's reliability is maintained;

- Vision algorithm implementation - with the acquisition of both a camera and a microprocessor the vision approach should be physically tested in order to observe how the camera position, for example, influences the system's reliability;

- Systems coupling - as was initially idealized, the coupling of these two solutions is expected to be the better reliability versus low power consumption approach with the two-level detection, and the low-impact consumption from the microprocessor branch;
- Further testing of the integrated solution.

The accomplishment of the previous points should come together into the development of the best infant recognition system, but there are also other options that could be, once again, taken into consideration. A potential incorporation to the vehicles electronics, particularly wireless notification system. In chapter 3.3 it is referred that the incorporation of an alarm in the car keys might be the optimal notification solution in contrast with the GSM module.

The accomplishment of all these conditions should culminate in a fully viable commercial option.

References

[1] Tomas Nolte, Hans Hansson, and Lucia Lo Bello. Wireless automotive communications. In *Euromicro Conference on Real-Time Systems*, volume 6, pages 35–38, 2005.

[2] Jan Null. No heat stroke. http://noheatstroke.org/, MAY 2018. (Accessed on 23/05/2018).

[3] Fairuz RM Rashidi and Ikhwan H Muhamad. Vehicle's interior movement detection and notification system. *Recent advances in automatic control, modelling and simulation*, pages 139–144, 2013.

[4] Ikhwan H Muhamad and Fairuz RM Rashidi. In-car suffocating prevention using image motion detection. *Recent Advances in Electrical Engineering Series*, pages 145–150, 2013.

[5] H. Cai, D. Lee, H. Joonkoo, Y. Fang, S. Li, and H. Liu. Embedded vision based automotive interior intrusion detection system. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2909–2914, Oct 2017. doi:10.1109/SMC.2017.8123069.

[6] A. R. Diewald, J. Landwehr, D. Tatarinov, P. Di Mario Cola, C. Watgen, C. Mica, M. Lu-Dac, P. Larsen, O. Gomez, and T. Goniva. Rf-based child occupation detection in the vehicle interior. In *2016 17th International Radar Symposium (IRS)*, pages 1–4, May 2016. doi:10.1109/IRS.2016.7497352.

[7] V Pavani and R Srinivas. Self-regulated automatic ventilation of vehicle interior. *Self*, 3(7), 2016.

[8] Norizam Sulaiman, Kamarul Hawari Ghazali, Mohd Shawal Jadin, Amran Abdul Hadi, Muhammad Sharfi Najib, Mohd Salmizan Mohd Zain, Fatimah Abdul Halim, Suhaimi Mohd Daud, Nurdiyana Zahed, and Abdul Adam Abdullah. Development of comprehensive unattended child warning and feedback system in vehicle. In *MATEC Web of Conferences*, volume 90, page 01008. EDP Sciences, 2017.

[9] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 1, pages I–511–I–518 vol.1, 2001. doi:10.1109/CVPR.2001.990517.

[10] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012. URL: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf.

[11] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.

[12] Ross Girshick. Fast r-cnn. *arXiv preprint arXiv:1504.08083*, 2015.

[13] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6):1137–1149, 2017.

[14] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Computer Vision (ICCV), 2017 IEEE International Conference on*, pages 2980–2988. IEEE, 2017.

[15] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.

[16] Joseph Redmon and Ali Farhadi. Yolo9000: better, faster, stronger. *arXiv preprint*, 2017.

[17] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.

[18] IEEE. Ieee standard for safety levels with respect to human exposure to radio frequency electromagnetic fields, 3 khz to 300 ghz. *IEEE Std C95.1-2005 (Revision of IEEE Std C95.1-1991)*, pages 1–238, April 2006. doi:10.1109/IEEESTD.2006.99501.

[19] İ. Şişman, A. O. Canbaz, and K. Yeğin. Micro-doppler radar for human breathing and heartbeat detection. In *2015 Computational Electromagnetics International Workshop (CEM)*, pages 1–2, July 2015. doi:10.1109/CEM.2015.7237422.

[20] D. Obeid, S. Sadek, G. Zaharia, and G. E. Zein. Doppler radar for heartbeat rate and heart rate variability extraction. In *2011 E-Health and Bioengineering Conference (EHB)*, pages 1–4, Nov 2011.

[21] Yue Tian Xijing Jing Jianqi Wang Guohua Lu, Fang Yang. Contact-free measurement of heart rate variability via a microwave sensor. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3267188/#b33-sensors-09-09572, "7" 2009. (Accessed on 01/21/2018).

[22] Microchip. Atmega2560 - 8-bit avr microcontrollers. http://www.microchip.com/wwwproducts/en/ATmega2560. (Accessed on 06/09/2018).

[23] Raspberry. Raspberry pi 3 model b - raspberry pi. https://www.raspberrypi.org/products/raspberry-pi-3-model-b/. (Accessed on 06/09/2018).

[24] IEEE. Ieee standard for telecommunications and information exchange between systems - lan/man - specific requirements - part 15: Wireless medium access control (mac) and physical layer (phy) specifications for wireless personal area networks (wpans). *IEEE Std 802.15.1-2002*, pages 1–473, June 2002. doi:10.1109/IEEESTD.2002.93621.

[25] IEEE. Ieee standard for high data rate wireless multi-media networks. *IEEE Std 802.15.3-2016 (Revision of IEEE Std 802.15.3-2003*, pages 1–510, July 2016. doi:10.1109/IEEESTD.2016.7524656.

[26] IEEE. Ieee standard for information technology- telecommunications and information exchange between systems- local and metropolitan area networks- specific requirements part ii: Wireless lan medium access control (mac) and physical layer (phy) specifications. *IEEE Std 802.11g-2003 (Amendment to IEEE Std 802.11, 1999 Edn. (Reaff 2003) as amended by IEEE Stds 802.11a-1999, 802.11b-1999, 802.11b-1999/Cor 1-2001, and 802.11d-2001)*, pages i–67, 2003. doi:10.1109/IEEESTD.2003.94282.

[27] Tesla. Android and iphone app | tesla. https://www.tesla.com/support/android-and-iphone-app?redirect=no. (Accessed on 06/09/2018).

[28] Panasonic. *MOTION SENSOR (PASSIVE INFRARED TYPE)*, 3 2005.

[29] Panasonic. *MOTION SENSOR (PASSIVE INFRARED TYPE)*, 3 2016.

[30] Microchip. Attiny85 - 8-bit avr microcontrollers - microcontrollers and processors. https://www.microchip.com/wwwproducts/en/ATtiny85. (Accessed on 06/10/2018).

[31] Digikey. Amn22111 panasonic electric works. https://www.digikey.pt/products/en?keywords=amn22111. (Accessed on 29/05/2018).

[32] Digikey. Ekmc1693112 panasonic electric works. https://www.digikey.pt/products/en?keywords=ekmc1693112. (Accessed on 29/05/2018).

[33] Adafruit. Adafruit fona 800 breakout board starter pack - sma version. https://www.adafruit.com/product/2522. (Accessed on 29/05/2018).

[34] adafruit. Github - adafruit/adafruit_fona: Arduino library for the adafruit fona. https://github.com/adafruit/Adafruit_FONA, 6 2015. (Accessed on 06/08/2018).

[35] Gabriele Bleser Didier Stricker. Background subtraction. https://ags.cs.uni-kl.de/fileadmin/inf_ags/opt-ss14/OPT_SS2014_lec05.pdf. (Accessed on 29/05/2018).

[36] Chris Dahms and Nic Williams. Github - microcontrollersand-more/opencv_3_image_subtraction_cpp. https://github.com/MicrocontrollersAndMore/OpenCV_3_Image_Subtraction_Cpp, 2 2016. (Accessed on 01/25/2018).

[37] Python Programming. Haar cascade object detection face & eye opencv python tutorial. https://pythonprogramming.net/haar-cascade-face-eye-detection-python-opencv-tutorial/. (Accessed on 01/25/2018).

[38] Rainer Lienhart. opencv/haarcascade_frontalface_default.xml at master · opencv/opencv · github. https://github.com/opencv/opencv/blob/master/data/haarcascades/haarcascade_frontalface_default.xml, 12 2013. (Accessed on 01/25/2018).

[39] Shameem Hameed. opencv/haarcascade_eye.xml at master · opencv/opencv · github. https://github.com/opencv/opencv/blob/master/data/haarcascades/haarcascade_eye.xml, 12 2013. (Accessed on 01/25/2018).

[40] Github - pjreddie/darknet: Convolutional neural networks. https://github.com/pjreddie/darknet, 6 2015. (Accessed on 06/08/2018).

[41] Trieu. Github - thtrieu/darkflow: Translate darknet to tensorflow. load trained weights, retrain/fine-tune using tensorflow, export constant graph def to mobile devices. https://github.com/thtrieu/darkflow, 8 2016. (Accessed on 06/08/2018).

[42] Tensorflow. https://www.tensorflow.org/. (Accessed on 06/08/2018).

[43] Matterport. Github - matterport/mask_rcnn: Mask r-cnn for object detection and instance segmentation on keras and tensorflow. https://github.com/matterport/Mask_RCNN, 11 2017. (Accessed on 06/08/2018).

[44] thtrieu. Train the last n layers · issue #486 · thtrieu/darkflow · github. https://github.com/thtrieu/darkflow/issues/486. (Accessed on 06/18/2018).