Politecnico di Torino

Master’s Degree in Engineering and Management
A.y. 2022/2023
Sessione di Laurea Ottobre 2023

Analysis of People Counting IoT solutions and possible expansions into the Industry 4.0 sector for Dropper s.r.l.

Supervisor:

Prof.ssa Giulia Bruno

Candidate:

Lorenzo Gelli

Table of Contents

Introduction

Modern entrepreneurship, in any sector, is characterized by a strong emphasis on the need for a continuous and accurate stream of data. This is not related solely to software-based products and companies, as the need for data can also refer to a number of Key Performance Indicators (KPI) and a lot of different information fundamental to the optimization of, for instance, manufacturing productive workflows.

Different types of data can be provided by different instruments, but any company will involve a human resource, in the form of operators and/or users and customers, which means that people counters will always have an array of possible applications. Whether for the management and optimization of employees activities and paths, the validation and definition of marketing campaigns, or the study and analysis of crowd behaviours in congested areas, the People Counting and Crowd Management Systems industry has been expanding for a few years now, and has promise to become an ever-larger sector with encouraging growth margins (CAGR estimates between 10 and 13% [1] in the next five years) reaching a $ 2 billion market by 2028, a growth and success that has only been improved and accelerated by the years of the pandemic.

In this innovative and dynamic context, Dropper s.r.l. is a relative newcomer, a Turin-based startup formally founded in 2021 and currently hosted in the Incubatore delle Imprese Innovative del Politecnico di Torino (I3P), sporting a wide range of people counting and crowd management systems for various purposes for both private and public customers.

This work will analyse some of the key aspects of this industry, starting with a literature review of the state of the art in terms of both the central technologies instrumental to the people counting solutions, and then the actual solutions themselves, in their different generations, that have been developed and perfected over the last few years. This will be instrumental context to the analysis of the different devices and software offered in the specific instance of the Dropper manufacturer company, which will be displayed in the next chapter, and that will culminate in the final chapter, illustrating the current state, and possible developments, of application of the solutions to the optimization and improvement of a number of different industrial environments. Some concluding remarks will follow as a summary of the work done and the conclusions reached, and to comment on the potential for future expansions of the Dropper core business into said industrial environments.

Chapter One

Fourth Industrial Revolution and Crowd Management

The people counting and crowd management industry is a very innovation-intensive, high-tech sector. As a relatively new industry, innovation is centred on improving the accuracy levels of the automated counting, but also on the analysis of the data obtained by the devices in order to provide as many Key Performance Indicators (KPI) as possible that can be useful to the specific needs of the customer. As such, the devices and software themselves are actually based on a well-established, yet ever-renovating body of scientific literature.

The following chapter will be divided into two sections in order to explore such literature more effectively. First, the technological framework necessary to operate the people counting and crowd management devices will be illustrated, as well as the main points of the state of the art of the Internet of Things (IoT) and Industry 4.0 structure supporting these solutions. Then, we will address the state of the art for the devices themselves, divided into the different generations that have characterised this industry up to this day.

1.1 Industry 4.0 framework and Internet of Things technology

The expression "Industry 4.0" refers to the set of practices that develops a manufacturing environment towards integrating and putting into communication different equipment and machinery, employed in different phases of the production process. Rather than to a single technology, Industry

4.0 is a term that refers to several distinct but integrated technologies, and they are all changing the way we understand society so dramatically that they are believed to be giving shape to a **Fourth Industrial Revolution**. Among these pillar technologies are cited: **Big Data**, Autonomous Robots, Simulation, Horizontal and Vertical System Integration, **Internet of Things**, **Cloud and Cloud Computing**, **Additive Manufacturing,** Augmented Reality, Cyber Security [2].

While all of these technologies are, in some way or another, involved in the production of modern People Counting systems, and People Counting systems technologies in turn all find applications in the Industry 4.0, I will analyse more in depth here the ones that are relevant in a more specific and material way.

1.1.1 Big Data

The gathering and management of data has become a central feature of modern enterprise, especially in the high-tech sector, as data can achieve a competitive advantage for a company in a variety of ways. More specifically, it has been theorised to be divided into three categories depending on the use and purpose the company gives to data: data as a tool, data as an industry, and data as a strategy.

Essentially, all of these uses apply to the case study company Dropper. For starters, while the company does not trade in raw personal data[1], it is also true that customers come to Dropper and other similar companies to receive primarily information that they don’t have regarding their own business. Therefore, data can be considered to be the very object traded in this case, as sometimes this is what can be required by the customer rather than the processed information; data

---

[1] Data as an industry (or as product) is most commonly associated with the buying and selling of personal data for commercial uses and other purposes, as, for example, clarified by the language and main arguments in Sunitha and Rajeshwari [3] as well as by the concerns raised in Martin [4].

could also be seen here to be used strategically, as the gathering of data is, as the definition states, functional to the development of a business model through the information obtained. Finally, we can indicate “data as a tool” as another fundamental purpose of big data in this context: as established in the literature, this means that data can be elaborated into information, and used to “solve traditional value chain problems by existing capabilities” [2]. This is precisely what the customers plan to do with the precious information (or data) they obtain through people counting and crowd management systems provided by companies such as Dropper.

Since the Age of Information has supplied a near-unlimited amount of data, the current challenge is that of wading through it and analysing it, which is why recent years have seen a proliferation of Data Analysis Algorithms for an array of different uses, including the proprietary algorithms developed by Dropper for the production of their solutions.

1.1.2 Internet of Things

The (Industrial) Internet of Things is particularly important, probably the most crucial technology for the development of these solutions. The idea of the Internet of Things is, in theory, almost as old as the Internet itself, but it is only with the last wave of technological advancement, what we now call the Fourth Industrial Revolution, that it has really developed to its full potential, with applications that range from manufacturing to public transportation, from domotics (home automation) systems to environmental management.

It consists of “networked devices [able to] propagate their information about physical world objects through the web” [5], meaning that the focus of the interaction through the web moves away from people and straight to the devices, which are then capable of (partial or total) independent interaction with one another. The equation that is often used to represent this interaction is:

IoT environment = Internet + WSN

Where "WSN" stands for Wireless Sensor Network, meaning the sensors themselves, interconnected in a network in spite of their physically dispersed distribution, are used to observe and gather the aforementioned information; then the information is propagated through the internet. This is precisely the idea of how the people counting and crowd management systems work, through a network of devices, whose algorithms and processing can be more or less centralized, that communicate with each other and with a central server to provide the final information to the user.

1.1.3 Cloud

The Cloud is a well-established means of storing data in a secure, non-hardware-relying way, allowing people and companies to access information at any time and from anywhere. This technology is also employed in a company such as Dropper for a variety of purposes, e.g., to communicate updates to the algorithms or exchange material internally. However, the latest technological advancements have brought the Cloud from simple storage utility to the fundamental work tool that is Cloud Computing.

Cloud Computing allows users to transfer resources and services in a virtual environment, where a company can utilize both storage space and processing capabilities, that with Cloud Computing are usually rented, allowing them to reduce the investment necessary to the daily management of activities. More specifically, Dropper employs Amazon Web Services, which some literature has used in the recent past to define a whole "style" of cloud computing [6] the **server virtualization**.

This kind of cloud computing is characterised by high flexibility, scalability, and compatibility for integration with existing software and application: specifically for this case, the people counting algorithms and the front-end customizable dashboard for information visualization. These are the aspects that have drawn the administrators of Dropper to choose Amazon Web Services for the needs of the company. While exclusively software-based companies can benefit from cloud computing in a more all-encompassing way, to the point that they could move very high percentages of their value chain on the cloud, hardware-software integrated companies such as Dropper can also thrive on the cloud to simplify the production process.

1.1.4 Additive Manufacturing

3D printing, or Additive Manufacturing, is the process of manufacturing a product through the successive application of layer upon layer of material, usually plastic materials such as PLA, in the form of filaments. In addition, blends can be created with a growing number of raw materials to characterise the mechanical properties of the final product, including carbon.

While this type of manufacturing process can face some serious economic sustainability issues on larger batches, companies can benefit from numerous advantages on smaller lots of products. Additive manufacturing can considerably reduce lead times, as it is entirely automated once the input has been given, and of course it can allow for a high level of customization [2].

These are the main advantages that have also brought the Dropper company production team to choose, on occasion, 3D printing for the case of some of the products, especially the need for customization and the strategic choice of creating Dropper’s own design for the devices.

Technology	**Description**	**Application**
Big Data	Large, complex datasets that affect the decision making of companies	Development of data into useful information for the customer through data analysis algorithms
IoT	Connection of the physical objects and systems	Interconnection of the devices and their effective management and functionality in a smart network
Cloud	Shared platforms that serve to the multiple users	Updates sharing, data elaboration through cloud computing
Additive Manufacturing	3D printing technology, producing in mass customization	Product case manufacturing

Source: partially from Erboz [2]

Table 1 - The Industry 4.0 Technologies in the People Counting Systems production.

1.2 The People Counting and Crowd Management Systems industry: a state of the art

The people counting systems have had a long history, throughout the last few decades, of adapting to both technological and social changes. As a matter of fact, evolution in this industry has not merely been the result of substantial technological advancements, but also the reflection of socio-economic changes. The renewed attention to the theme of privacy, especially in the context of the European Union, has produced a political momentum to legislate on the matter, culminated in the issuance of the **General Data Protection Regulation**[2] (GDPR) in 2016.

---

[2] Full text available in every language of the Union at [7].

The GDPR is only the last and most comprehensive in a long tradition of Data and Privacy protection legislation put forward by the European Union through the ad-hoc European Data Protection Board (EDPB), a bureau that tackles the necessary task of keeping the personal data of the European citizens protected.

This task requires continuous updating of the privacy regulations, all the more so in today’s dynamic technological environment, and this industry demonstrates this quite effectively, as it has seen a great number of technologies employed for the same – or similar – purposes. The following solutions, all of which can still be found on the market, were deemed by the Dropper company to be either too imprecise, or not compliant enough with the privacy regulations, which is why when the company had to consider the privacy issues in people counting technology, it reached different answers from the ones listed here. Those will be displayed in the next chapter.

1.2.1 Beam sensor people counter

The beam sensor (or break-beam) people counter is one of the most rudimentary systems, it consists of an infrared beam emitter placed on the points of access to an indoor space; the device counts based on the times that the infrared beam is interrupted.

While this is a long-established technique for people counting, it is also exposed to a number of critical issues. First of all, while some companies advertise this kind of device for any environment, the obvious interference of the lighting conditions should be considered, as it can seriously affect the final accuracy of the counting. Moreover, a single beam is unable to render a sense of direction, meaning that it is impossible to discern people entering a space from people exiting, and restricting the use of this kind of device to accesses that are exclusively entrances or exits, not both. More recently, a development in beam sensors has allowed to mitigate this problem, as now some sensors involve up to three different beams, allowing to determine direction of passage from the order in which beams are interrupted. Lastly, and maybe most prohibitively, beam sensor people counters are only able to notice the first (closest to the emitter) object (not necessarily a person) interrupting the beam, they are therefore incapable of identifying multiple passages at the same time, and they can also be misled by shopping carts, strollers and other objects that further restrict the scope of their applicability.

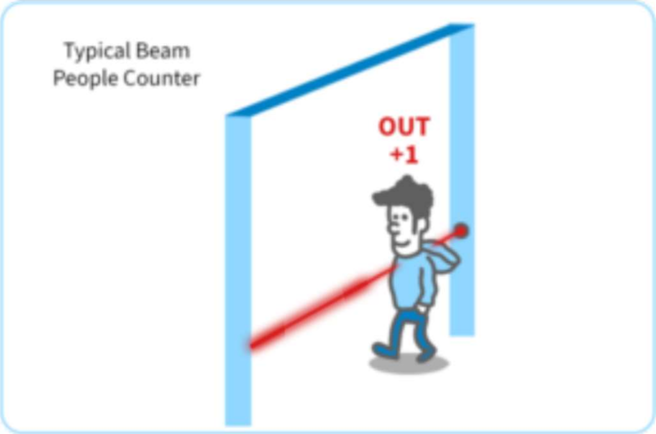

Figure 1 - Infrared beam sensor Source: V-count

While this technology is widely regarded as established and, on some respects, even outdated, it still maintains a limited market, which demonstrates the existence of applications for which the break-beam infrared people counter is still an interesting option. In [8], for instance, the whole project revolves around the use of break-beam people counters, which are employed at the access points of rooms with a double purpose. While the sensor counts people present within the room, it is also combined with a microcontroller to automatically control the lights of said room: the counter gives the number of people in the room, and when said number goes from zero to one, it turns the lights on, vice versa turning them off when everyone exits, and the count ends up going from one to zero.

Break-beam solutions are also cited in Mohammadmoradi and Munir [9], a paper on infrared thermal solutions application that will be further expanded upon in the next paragraph, and there they are mentioned as the most inexpensive solutions on the market for people counting, with the possibility to use them also with energy efficiency improvement applications. However, it is also mentioned that their installation, with a number of modules required, can be tricky and easily prone to failure.

1.2.2 Thermal camera

The technology of the thermal camera for people counting purposes is based on identifying the presence of one or more persons thanks to the natural body heat emitted by humans.

As clarified, for instance, in Gade and Moeslund [10] everything at a temperature above the absolute zero emits a radiation in the infrared spectrum, called the thermal radiation. These cameras can detect this radiation and transpose it into visible images, in a technique that is called **thermal imaging**.

Thermal imaging has been invented and spread during the 40s and 50s, ca 150 years after the discovery of the infrared, at first for military purposes. From there it has rapidly spread to a variety of different uses, from domestic animals farming and wild fauna detection and management, to building inspection and gas detection, to, of course, people counting. While it does have slightly better accuracy than the infrared beam sensor for slightly higher costs, thermal imaging is still exposed to functionality flaws that have considerably reduced its spread. The most crucial issue with thermal cameras is that their effectiveness is strictly related to the external conditions. These cameras could be, theoretically, used in any environment, but especially for outdoor spaces and for entrances from the outside, high external temperatures can seriously affect the ability of the camera to distinguish the heat coming off a person. This also overlaps with the general difficulty of these devices to distinguish individuals in a group of people passing at the same time [11], especially when in close proximity to each other. The resulting lower accuracy of these devices, combined with (however slightly on average) higher costs compared to the beam sensors, makes this technology much less appealing to develop.

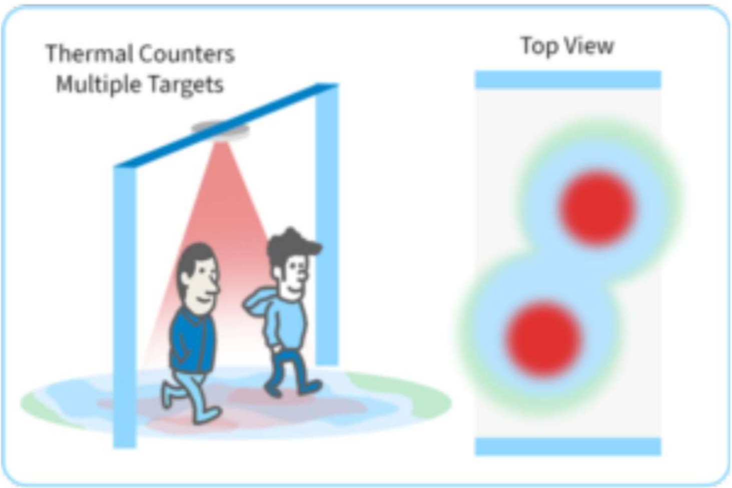

Figure 2 - Thermal counters view *source: V-count*

As mentioned, an application of this technology can be found also in [9]. This use case employs infrared-based technology for thermal people counting and applies it to the purpose of improving the energy efficiency of a building. The scenario involves beam sensor placed and every accessway, counting entries and exits events, then the data gathered and elaborated does not simply return a total number of presences, (with high declared accuracy and adaptive to different walking speeds) but also a rather accurate estimation of the subjects' skin temperature, from which information on their thermal comfort can be derived and, consequently, adjustments to the Heating, Ventilation and Air Conditioning systems can be made.

This experimental paper really highlights the advantages of such a solution, such as its inexpensiveness and its ease of installation, including very limited needs for structural and infrastructural modification of the area of installation. However, the paper itself highlights some of the faults of this technology, such as its lower effectiveness when the environmental and meteorological conditions bring the temperature of the spaces closer to that of a human body.

1.2.3 2D cameras

2D cameras are monocular (single lens) cameras installed on the accesses to an area to monitor the passage of people underneath the doorway. Being an image-based sensor, they rely on systems that are similar to CCTV cameras.

This, however, creates two main problems for the use of this technology. Firstly, the video-based cameras were not originally thought for people counting, and any adjustment to their software has to proceed with this lasting disadvantage. Second, as these cameras lack depth perception, they will count any moving object passing in front of the lens, with a consequently high chance of false positives, thus potentially overestimating the occupancy levels of the area they monitor. Lastly, cameras are more strictly regulated, both by European regulations such as the aforementioned GDPR [12] and by national regulations such as those issued by the Italian bureau of Garante per la Protezione dei Dati Personali (Data Protection Authority) [13], as cameras of this kind could very well be classified as video surveillance tools, even when employed for people counting.

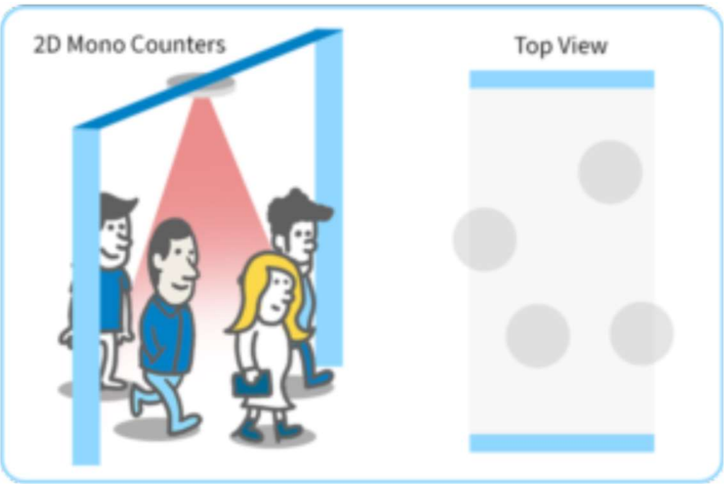

Figure 3 - 2D camera counters view *source: V-count*

A most recent application of this method used in the context of people counting can be found in [14]. This paper presents a rather typical approach to people counting, specifically for the 2D monocular camera technology. This entails a device placed on top of an entryway and the people counting process is more precisely divided in the following steps:

- The image capturing, both through an IP camera and a video streaming recording, which is used as input and does capture the features of the subjects passing underneath the system.
- A first algorithm, needed for people detection, which through background subtraction, morphological operation and blob detection reaches the conclusion that a person in present underneath the camera.
- And finally, a second algorithm executes the actual counting, which effectively happens by dividing the area of interest into regions, following the passage of the person in one direction from on to the other, and then adding or subtracting to the total count.

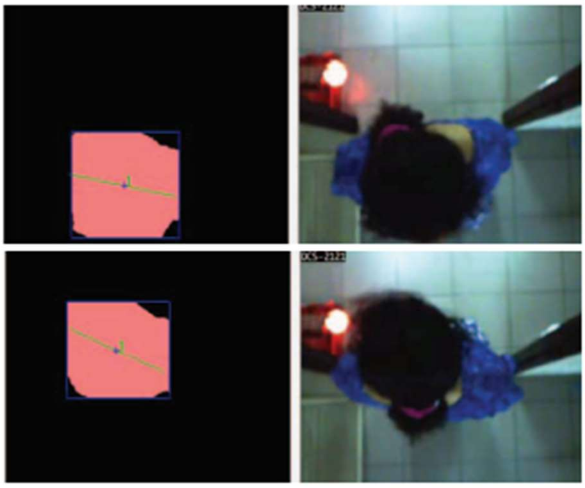

Figure 4 - The blob detection demo, including the video feed upon which the imaging is based on

With various fine-tuning steps and adaptations to the specific application scenario, this is considered to be an effective and reliable system for people counting. However, as mentioned before, its applicability is severely reduced by the violation of privacy policy in recording the faces and personal features of anyone who walks underneath this system, particularly in countries, such as those that are part of the EU, where privacy regulation has recently become considerably stricter.

1.2.4 3D cameras

Finally, the latest advancements in the industry of image-based sensors for people counting resulted in the 3D camera. These devices gather data in the form of images with the use of a double lens that, similarly to human eye vision, is able to perceive depth in the images it detects.

The technology itself can be traced as far back as the late 80s. For example, Yamaguchi et al. published [15], for the IEEE 1988 International Conference on Consumer Electronics Digest of Technical Papers, a short paper detailing one of the first "3D-CAM", certainly one of the first ones to move towards a standardisation of this technology. As it can be read in the paper, the device is provided with two CCD cameras; the two signals acquired through the two cameras are then merged and cross-referenced, in a way not dissimilar to how human eyes see, to create a coherent stereoscopic image, i.e., an image capable of conveying the sense of depth. More recently, the applications of this particular technique have expanded to people counting and crowd management. For instance, Beymer [16] discusses in detail the functioning of a 3D camera for the purpose of keeping count of the customers of a shopping mall. The author includes comparisons with a monocular, 2D system, in particular the novelties introduced by a 3D system, such as:

> […] (1) remapping the stereo disparities to an orthographic "occupancy map", which simplifies person modeling, and (2) tracking people using a Gaussian mixture model. On a test set of 900 enter/exit events in 4 hours of video, our system has achieved a net counting error rate of just 1.4%.

These would allow, as it is said, for a very high level of accuracy. However, in the later years, this technique would meet high levels of hostility due to privacy issues (which, for the scope of this work, will be illustrated briefly).

These images are then processed, including a reduction of the definition for privacy purposes, and sent through the counting algorithm that then identifies moving objects as people and counts them. In addition, this technology can be used not only vertically, on entrances, but also by mounting the camera on a wall and using it to monitor the paths taken by, for instance, customers of a store, with the advantages for the retailers that will be further analysed later on. The main issue that this technology is faced with relates to privacy. Even though they are later processed and anonymised, sources images gathered by the cameras could, theoretically, be used to identify a specific individual, just as, if not better than, with a two-dimensional camera.

Since this is a newer technique for people counting than the aforementioned ones, the recent literature on 3D cameras is especially abundant. One such example con be found in Burbano, Bouaziz and Vasiliu [17], who design a distributed system for people tracking and counting. It their paper, the proposed infrastructure involves a network of downward view[3] RGBD depth cameras. Each node of this network is capable of detecting, counting, and tracking the people passing underneath it, and of communicating with the adjacent nodes on the network. The detection and counting algorithms work very similarly to the functioning of the 2D camera process. In the network of smart

---

[3] It is a matter of debate, as the paper itself also mentions, whether 3D cameras are better suited in downward view or side view. Some works are also cited both arguing for the advantages of the former [18] [19], such as the avoidance of complete occlusion, and of the latter [20] [21], like the possibility to gather more information on the human body.

cameras, each detects the objects passing through the field of view, then the software separates and identifies the objects as people according to the given constraints, and counts them as present in its area, then sending all relevant information to the network.

The organisation into a network allows not only for the cameras to verify each other's data though the superimposition, but also to account for possible node failures through automatic management of the loss and re-adjustment, and track most common paths in the general area, which can have useful applications for marketing purposes in retail settings, for instance.

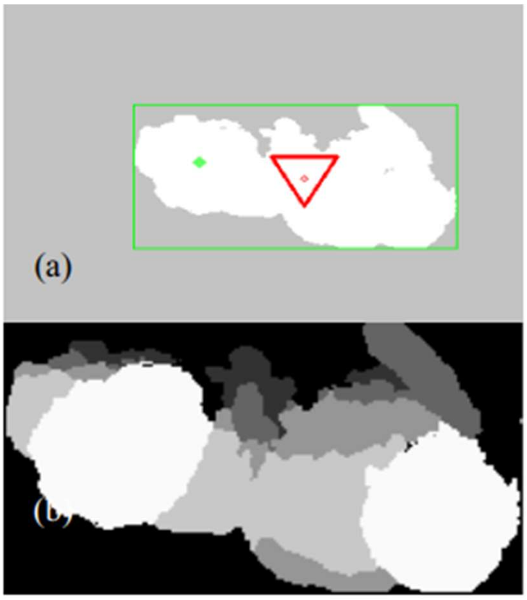

Figure 5 - The blob detection view (a), compared to the segmentation of the blob into layers, the main characterizing feature of the 3D camera

Figure 6 - The four generations of people counters

These are the generations of people counting sensors that have dominated the market especially in the past, each of them with their defects and issues. Said issues are the reason why Dropper has decided to go in a different direction, developing and producing the devices that will be presented in the next chapter.

1.2.5 Privacy issues and concerns

As a very last point, I would like to add a few observations as a form of state-of-the-art of the current situation in privacy policymaking, mentioning briefly what political junctures have led to the current state of the public debate on the matter.

As mentioned at the beginning of this paragraph, the public debate on privacy and data management had had a central place in political arenas before, during the Digitalization Era and the so-called Third Industrial Revolution (second half of the XX Century). With the arrival of the Digital Revolution, in fact, there are plenty of examples of privacy being the central topic of public debate, as exemplified time and again by Gandy, Jr. [22]. Not only was privacy the centre of attention for so long, during the 90s especially, but the issue was also at times purposefully misconstrued by the media to manipulate public opinion through the use of polling as a political instrument. Aside from this, the paper by Gandy, Jr. also highlights how emblematically the case of post-9/11 United States population shows a shift in public attention from privacy towards other values, such as security, considered to be more urgent.

We could say that for the following ten to fifteen years, the situation has remained basically unvaried, but then data usage has met with a newfound attention due to social media platform diffusion, and in particular during the so-called "Cambridge Analytica scandal". The irrefutable proof, thanks in part to the crucial intervention of whistleblowers, that a private corporation was offering the service of swaying major political elections and referendums through the sheer power of data acquisition combined with some targeted advertising, to whomever was willing to pay was simply too much for public opinion to ignore.

Finally, this brings us (whether we believe it to be correlated to CA or not) to the latest and most comprehensive piece of policy on the privacy and data protection issues for twenty years, the GDPR. While its importance has already been stressed, it bears also remembering that this is a monumental piece of legislation not only for the protection of European citizens' data, but it is already being observed to be a determinant influence on several more countries and their own privacy policies: whether by force of imitation or in order to be able to make business with the EU, countries in the hundreds are, in their own way, adapting to the new standard of data protection [23].

We will see how this has impacted the strategic choices of Dropper when it comes to what kind of device to produce, in the next chapter.

Chapter Two

The Dropper s.r.l. company and its solutions

This chapter will serve as a presentation of the case study company, Dropper s.r.l. The analysis of the competitive environment surrounding the startup, as well as the insight on its various IoT solutions have all been gathered in the occasion of a curricular internship that took place between the months of May and September of 2022. During the internship, I have developed the skills and knowledge that have then led me to carry out the foundational studies for this work.

2.1 Competitive context and company overview

The people counting and crowd management market has known a consistent global growth in the last few years, and yet it is still considerable as a *blue ocean* in Italy. A blue ocean is a market condition wherefore said market is basically yet unexplored, an "unknown market space, untainted by competition" [24]. As such, competitive rules in the crowd management market in Italy have not yet been precisely established, and the very few competitors currently active (mostly locally to their own small market) have been enjoying the possibility to establish prices in a non-competitive, quasi-monopolistic way. Internal company studies attribute this lack of competition mainly to a general delay in the Italian industry and tertiary sector in keeping up with the most recent developments in crowd management and its capability to gather fundamental production and sales KPI.

Abroad, a few large players in the people counting and crowd monitoring market have been sharing the main segments of the market, for the most part each specialising in one or a few of the many applications of this sector.

More specifically, the market of people counting is represented by both companies that specialise on the production of software for this purpose, and companies, like Dropper, that manufacture the hardware as well as writing the algorithms. Some of the main competitors in this market include:

- **V-Count** [25]: this company founded in Turkey in 2015 is a main competitor for reasons both geographical and of application. They are a large corporation, with a few branches around the world, and with several solutions, showing their keen understanding of the versatility of the people counting and crowd monitoring sector. However, their approach to the issue, relying heavily on the use of 3D cameras, necessarily meets with some difficulties in expanding on the European market, as the legal position of such technologies could be considered unsure, with a present risk of it becoming stigmatised by the GDPR.
- **Density** [26]: the American corporation Density has been operating since 2014, with the main business goal of assisting in the management of office spaces. To this aim, they propose solutions based on either millimetre-waves radar or time-of-flight technology, with the former being a recent addition to their catalogue. They deploy what the team at Dropper believes to be the best technologies, in a very specific market, and this gives them high specialisation factors and a long-standing reputation on the other side of the Atlantic.
- **Affluences** [27]: Affluences is a French company first founded in 2014. It is an example of software-only product, as they basically offer access to an online platform to monitor occupancy of public places. Said data can be gathered both with proprietary hardware or simply entered into the

platform by the management of said places, as often happens, and there even is an option involving only a manual counting solution where management can physically enter and store the historical data about the number of people occupying the area. They are particularly popular for libraries, study rooms, museums, and local administrative buildings, for both people counting and queue and crowd management.

- **Visionarea** [28]: finally, the only Italian competition that actually aligns to the Dropper market segment, Visionarea has been operating since 2008 and has quite the head start in terms of current partnerships and clientele. They have been distributing their 3D camera-based devices and accompanying software to retail store chains, workspaces and local administrations alike, exploring the versatility of these solutions much like Dropper is starting to do. They are the one larger Italian player in the market, and that gives them both the advantages of scale and a widespread connection network, and the disadvantages of less of a possibility to support the customer and, if need be, to pivot to new areas of interest. The fact that they are almost in a monopolistic condition in the Italian market also leaves them open to undercutting in prices.

It is in this context that founding partners Domenico Galdiero and Alessandro Severini decided to start Dropper s.r.l. in 2021, at a time when this industry was even more topical due to the need to collect data on occupancy levels and people gatherings during the COVID-19 pandemic. In the short time of its existence, the team at Dropper has already obtained partnerships with corporations in the retail sector (Basko s.p.a., Norauto Italia s.p.a.), education institutions (Politecnico di Torino, Università degli Studi di Torino), museums, and local public administrations (Città di Torino, Edinburgh Police).

2.2 The Dropper solutions

The following paragraph will present the devices and technologies, different from the ones seen in the last chapter, that the Dropper company has chosen to adopt as its core business, due to their balance of high accuracy of the results and privacy protection. The three main solutions are: the **Wi-Fi Sniffer**, the **3D Time-of-Flight sensor**, and the **mm-waves Radar**.

2.2.1 The Wi-Fi Sniffer

This sensor was developed thanks to so-called **Wi-Fi sniffing** technology, developed for people counting purposes in this device by Dropper. Wi-Fi sniffing is a system used for many purposes, ranging from cyber security testing to crowd management and people counting both for indoor and outdoor areas. While their Wi-Fi or geolocation is on, devices such as smartphones emit probe request packages, and Wi-Fi sniffers intercept these packages, and use them for counting and monitoring the movement of people in a defined area.

What is really being counted is the MAC address contained within the probe request packages, a code that identifies the Wi-Fi port of the device that is emitting the package. In recent years, phone producers have changed the protection protocols of the devices to increase security, including a randomizing mechanic on the MAC addresses. The proprietary software developed by Dropper involves a de-randomizer to mitigate this issue, achieving a precision rate of 70% in instances involving a randomizing algorithm.

In an Industry 4.0 outlook, the software has been developed with Machine Learning to further refine and improve the accuracy of the count. Machine Learning algorithms are employed also for de-randomizing purposes: to be more precise, this term can be misleading, since the aim is not to return to the original MAC address, but to be able to group together – to **cluster** – the similar enough randomized addresses and identify those packages as emitted by the same device.

Machine Learning is implemented, to this end, through both a supervised and an unsupervised algorithm. A supervised algorithm is a Machine Learning application in which the human user is allowed to input data that contribute to the training of the algorithm itself: in the case of the Wi-Fi Sniffer, it is possible to input the manual count and the algorithm will consequently self-correct to increase accuracy. The unsupervised algorithm, on the other hand, only requires the human developer to input a variable that is programmed to control the degree of similarity between randomized MAC addresses in order for the algorithm to recognize them as emitted by the same device.

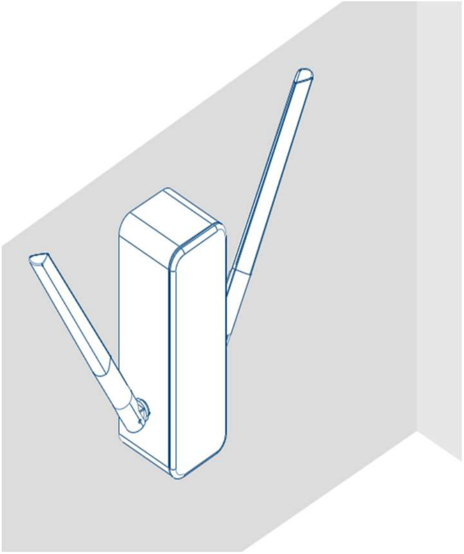

Figure 7 - Wi-Fi Sniffer representation

Additionally, these sensors can – through the use of Dropper's proprietary algorithms – monitor the frequency with which a device is detected, or produce heat maps to highlight the most frequently visited areas, consequently helping store managers in optimizing merchandise layout in order to captivate with promotional offers and maximise conversion rate.

Many recent studies support the methodologies put in place by the company, both for the people counting algorithms and for the de-randomization. In Vattapparamban [29], for instance, a few variations on the application of Wi-Fi probe requests tracking for people counting and crowd management have been analysed. As it is pointed out in the paper, and as already discussed here, any device capable of connecting through a Wi-Fi network continuously sends out "probe requests" packages, which can be passively collected and decoded with a Wi-Fi sniffer device, "without connecting to a particular network or transmitting any signal". Specific to this paper are two main characterisations that make the application interesting for further development, as well as original with respect to other, similar use cases.

First, the idea to combine the Wi-Fi sniffing technology with the use of Unmanned Arial Vehicles (UAVs, or drones). This allows us also to imagine this solution applied in the field for critical search and rescue tasks and, more generally, adds to the versatility and mobility of the IoT solution. In addition to and in combination with this, the proposed solution adds a layer to the use of probe requests, in that the algorithm was designed to not only detect and decode said packages, but also to measure the intensity of the signal received, and draw from it conclusions on the distance to the device transmitting them, thus increasing the accuracy and range of applications of the Wi-Fi sniffing techniques for people tracking. This paper also mentions the issues related to MAC randomization and proposes some solutions around them not unlike the ones used by the Dropper company and already discussed here.

More extensive on the problem of MAC address randomization, however, is the paper by Uras et al. [30]. Aside from another instance of the explanation of how probe requests detection works, which I will not be discussing again, this paper focuses on proposing an algorithm to identify randomised MAC addresses. Covering another one of the fields of application explored by Dropper, that of city services and local administrations, this paper explains, as already discussed here, the inner workings of the Wi-Fi sniffing technique through active probe request detection. In addition, however, it focuses on the MAC address randomization, defining it and its functioning:

> Randomization of MAC addresses is the process of generating virtual MAC addresses by end-devices during the phase of active scanning for Access Point in the WiFi context. Such activity is designed and performed to guarantee devices that their real MAC address remain unknown, preventing tracking issues. When the AP and the device find themselves, they set up the connection and only after that, the device uses its real MAC address due the fact which only starting from that moment the entire communication is encrypted.

In the central part of the paper, then, the authors propose an idea for a derandomizing algorithm, and this proposal is not so far from what the developers at Dropper have applied for the Wi-Fi Sniffer solution. The premise of said algorithm is that there are small differences in how manufacturers implement the randomisation of the addresses, such as differences in their Information Elements (IEs). When analysed, these reveal discrepancies that pretty much univocally identify a device from the others: while it would never be possible to return to the actual MAC address of said device, for the purpose of accurate people counting it is enough to be able to tell that the randomisation has happened, and to identify each device from the others in spite of it. Based on this line of reasoning, the algorithm follows the steps intuitively summarised in the following graph:

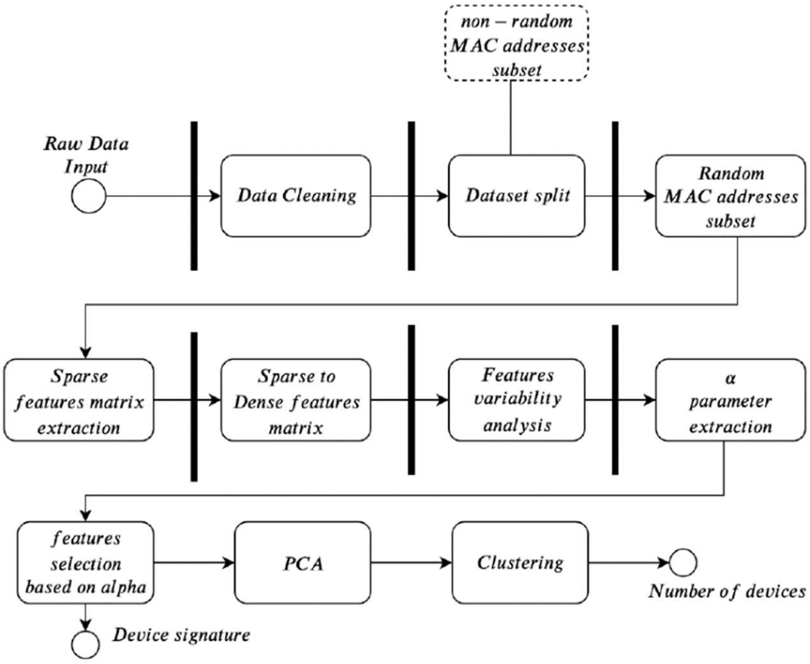

Figure 8 - Data Analysis pipeline

Reviewing some of the fundamental steps shown in Figure 8, at first the data is cleaned and prepared, checking for errors and corrupted packages in the raw data. After cleaning and preparation, of course the first important division is aimed at identifying the subset of MAC addresses that have been randomised (this is recognisable from one specific bit in the address itself). At that point, the "features matrices" identify the aforementioned IEs regarding all the packets received, at different levels of analysis. Lastly, the clustering is possible by setting a minimum level of features variability, α, defined as "the minimum unique values that the column j of matrix E_{ies} needs to have to be taken in account". Belonging to one of the columns that respect parameter α allows to assign a "signature" to the randomised MAC address that sent the cluster of packages, though which the address is now identifiable.

The fundamental ideas, premises and reasoning behind the algorithm proposed by Uras et al. are a strong base for the functioning of the derandomizing algorithm used by the Dropper company, and its results have been more than satisfying for the company standards, allowing very high precision of the counts even when in the presence of address randomisation.

2.2.2 The 3D Time-of-Flight sensor

One of the very first technologies explored by Dropper, the ToF device is based on a depth camera, or RGBD. Despite its name, this device doesn't actually capture images that would make the subject recognizable. It merely measures the distance between the sensor and an object as the time taken by an infrared beam to complete a round trip from the sensor and back to it after it has been reflected by the object. The data on the time of flight is then translated into greyscale, with the closer layers in the foreground in a darker hue and the farther ones lighter. As represented in the following image, this allows the software to profile the subjects counted without putting their privacy at risk, as they become completely unrecognizable in the greyscale.

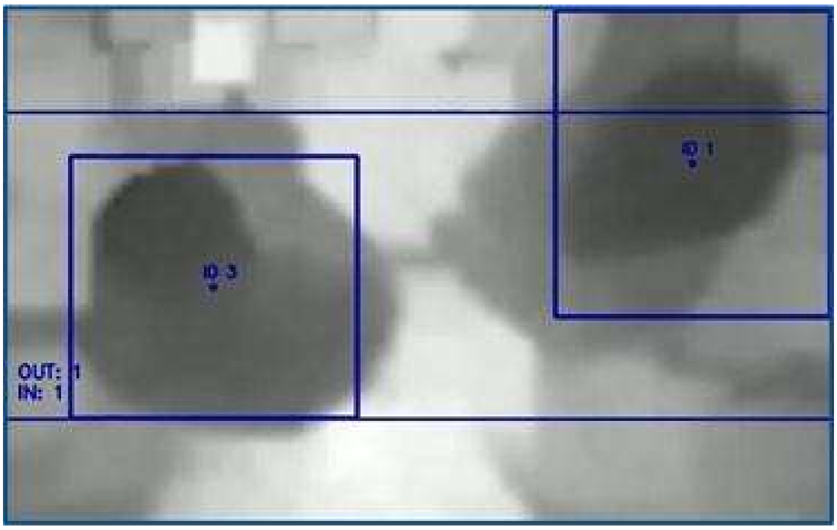

Figure 9 - 3D ToF sensor demo shot

The actual counting happens with the software taking the following steps:

- Recognition of a subject entering the frame from one of the sides, which are known to the algorithm to be either the inside or the outside of the monitored room.

- Identification of the subject within a bounding box and assignment of an ID.
- Tracking of the identified subject across the frame, all the way until it disappears on the opposite side from the one they entered.
- Addition (or subtraction) of 1 to the total count to mark the entrance (or exit) of the subject.

This algorithm ensures that the ID is counted in or out only once they are actually, physically all the way across the access, allowing the device to reduce wrong counts in many varied use-cases such as people going back and forth but not crossing the threshold, two people entering or exiting together, in opposite directions and so on and so forth.

The device can be programmed to adapt to different scenarios: meant to be placed at access points to a room or a building, the software can then be set according to the specific width of the door to reduce the effect of false-positive counting. Moreover, the code also defines limit values for the bounding boxes including (lower) bound height or minimum width of an object to assign an ID to it and count it. For instance, some applicative scenarios might involve a particularly high number of children entering an establishment (say, a toys shop), and it might be in the interest of the customer of Dropper, the manager of the establishment, to decide whether they would like to count all visitors entering and exiting, regardless of their ability to convert merchandise (thus including the children by lowering the minimum height from the ground / maximum distance from the sensor needed to be detected), or only the possiby paying customers (vice versa, increasing the limit and excluding the children).

Other specific applications might include pets or trolleys in a supermarket, and all these particular circumstances can be managed through personalised modifications to the software that the Dropper team refines for every customer at the time of setup. It is also thanks to this precise setting-up phase that this device is extremely error-free, with a tested accuracy rating reaching as high as 98%.

Additionally, the Dropper team is currently in the process of developing a smaller, parallel device to the Time-of-Flight, called a Mini-ToF. The Mini-ToF is characterised by reduced costs and size, will have a battery for mobile deployment, more versatile use, for temporary use or on alternative access points, while maintaining high levels of precision as it is based on the same technology as the 3D ToF.

Figure 10 - The Mini-ToF sensor

Examples of recent papers on the use of such technology for people counting purposes are also very abundant in scientific literature. A highly representative example, with similar application parameters and yet interesting operational differences from the common Dropper use cases is that offered by Stec, Herrmann and Stabernack [31]. Just like for the Dropper application cases, this paper envisions a vertical, looking down, single depth camera, but it also experiments with alternative perspectives, testing the capabilities of this method as shown in the following images of a frontal and angled view.

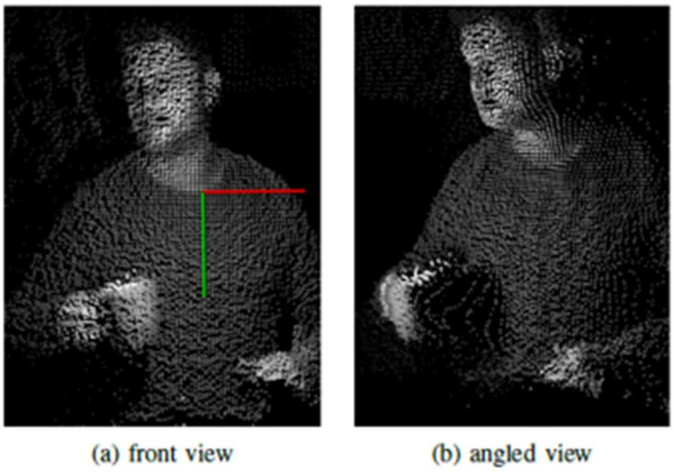
(a) front view (b) angled view

Figure 11 - Example point cloud mapping: Person with cup

The view angle is rounded up to 60° and the area covered by the ToF camera is calculated by a simple formula dependent on the height of installation of the sensor, and only one dimension is calculated, the length of the area, since the width is not necessary, constrained as it is by the width of the access. These parameters are very similar to the variables usually considered by the Dropper company, and the elaborated algorithm is also quite alike, as represented by this block diagram.

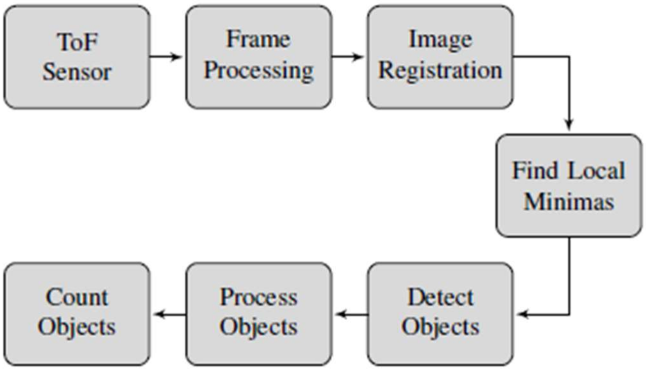

Figure 12 - Block diagram of the proposed algorithm

Other specified parameters are the assumed velocity and the assumed area of the items to be recognised as people. The end result of the authors' experiment is a set of shots that are very similar to what the Dropper company currently produces through its own Time-of-Flight devices. As for the algorithm implemented, schematised in the next Figure, the authors have envisioned the peaks and valleys detected by the sensor, that in reality represent heads and shoulders of the subjects, as if they were geographical features, that are then relatively "filled" with water. This *Waterfilling* algorithm, thus used and with some personal fine-tuning by the authors of the paper, allows to individuate points of local minima directly in the depth map created by the sensor, imagining it as it was a natural environment with water flowing to go and fill those spaces, and it is usable even at lower resolution levels.

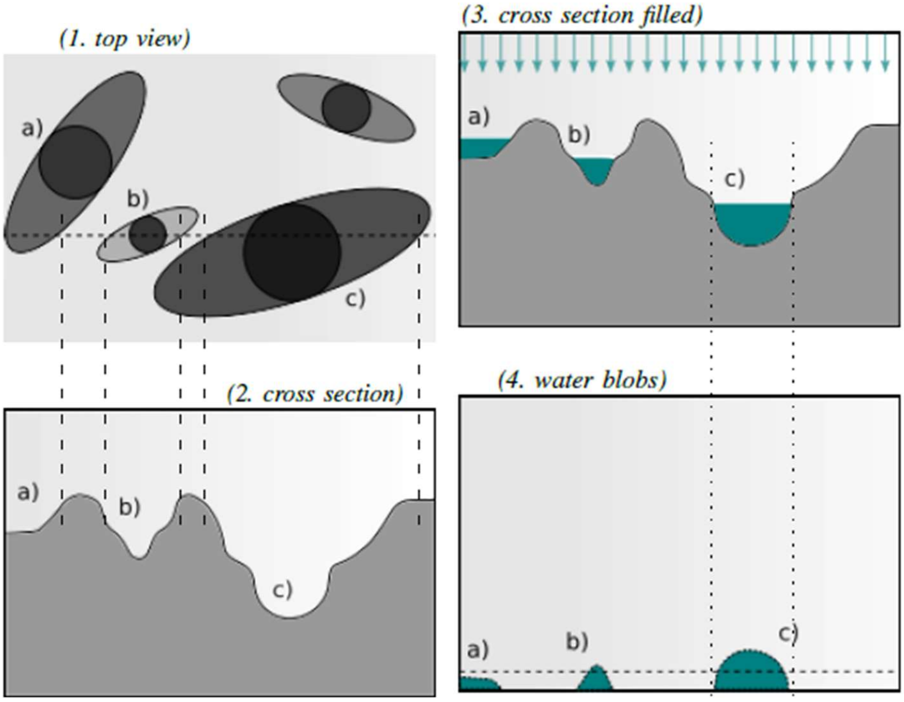

Figure 13 - Waterfiller: top view with marked cross section, shape of the cross section, cross section filled with water, water blobs

The results reached by this study are consistent with the operations carried out by Dropper: an accurate count in all applicative scenarios, with single subjects, or with multiple people both walking in the same and in opposite directions, or with larger groups, as large as the passage allows, walking under the sensor.

2.2.3 The mm-waves Radar

The latest addition to the Dropper offer, this sensor is based on radar technology, more specifically short-wavelength radar. Essentially, the device irradiates the area of interest with millimetre waves, so every moving object in the area produces a reading of a cloud of points: obviously, a point cloud is completely anonymous, preserving entirely the privacy of the subjects traversing the area. The identification “as a person” happens through the use of deletion algorithms putting limits on the volume and velocity of the items detected by the sensor.

This method – which has already found a variety of other applications that will be examined shortly – allows for the detection and counting of people down to a centimetre accuracy, which is useful for safety and security applications such as major public events: the mm-waves Radar sensor can in fact detect panic or other emergency situations in a crowd, shortening response times.

Its ability to monitor people positioning and movements make it useful to track occupancy levels of indoor environments, as well as movement patterns in a production line or to monitor public places of encounter in the city to gain statistics on their frequentation, together with a plethora of application both in the public services and Industry 4.0 environments.

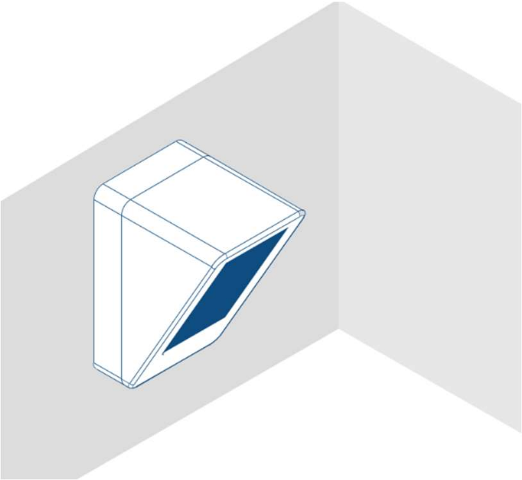

Figure 14 - mm-waves Radar

Just as it is for the Dropper solutions, mm-waves Radar devices in people counting and crowd monitoring are an emerging technique for the scientific community, as well. For this reason, the examples found in academic papers are all very recent, and yet the applications encountered by Dropper are quite typical and represented. For instance, in Cui and Dahnoun [32] the question of privacy is immediately raised, pointing out that cameras are too invasive, and more privacy-compliant techniques can and should be used. One such technique involves the use of radio frequency signal transceivers, more specifically millimetre-wave radars. The proposed solution requires defining several variables, which are then determined mathematically, such as the angular and distance position of the object, and the formula for the calculation of its velocity, as well as the configuration of the antenna for the set up depending on the technical characteristics of the radar hardware itself.

For the particular application of this paper, a set of two radar devices is used at a short range, but similar results could also be obtained with a single device and at a greater distance. Shown below is the representation of the proposed setup. The use of two devices and its advantages are clear once one understands the detection and tracking process, divided into three parts, that is used here.

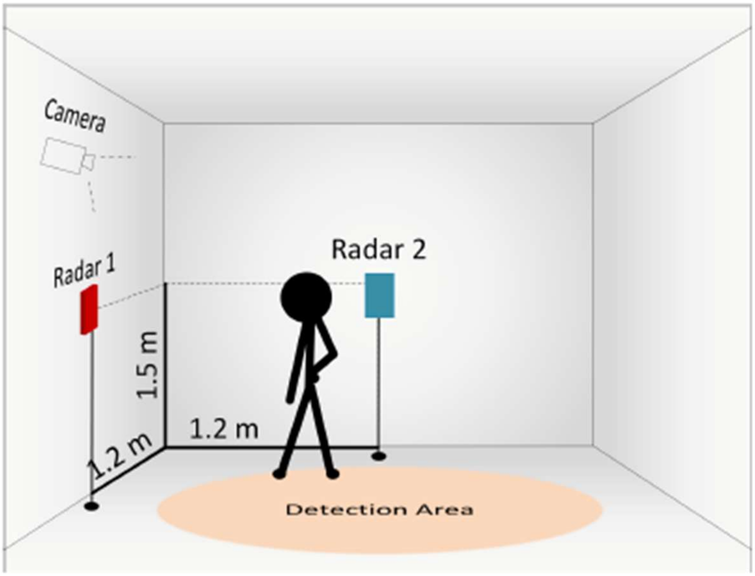

Figure 15 - Hardware setup of the two radars for human detection

The first step is the "Individual Detection", which involves, as the name suggests, the two radars separately detecting the subject plus any noise they might suffer from in their field of view. However, in the second step, "Data Fusion", both radars communicate the results of their reading of the area to a CPU called Central Frame Processor. The processor then combines the two coordinate systems and the two results, and from the analysis of the overlap it will determine which detected values represent a "candidate human", possessing certain estimated properties.

Finally, the "Tracking" module has the purpose of following the detected subject as it moves through the target area. In order to do that, it records all the "candidate humans" at each timestamp and the uses the temporal relationship between them to estimate the probability distribution of the candidate's new position, then verifying it against the new reading.

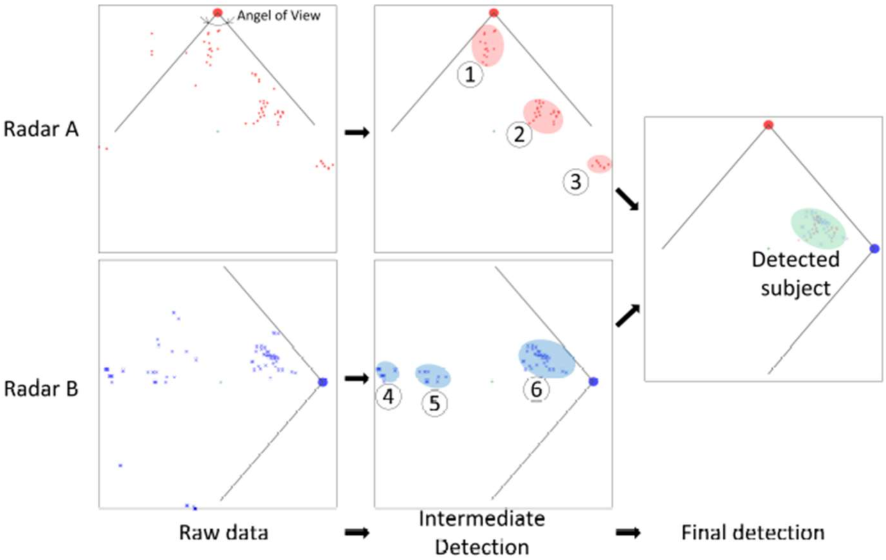

Figure 16 - Workflow of the human detection system, with one person presenting in the area (top-down view)

The end result is something that resembles almost exactly what the Dropper mm-waves Radar sensors are capable of achieving but on a larger scale, like for a whole room or even a street or a square, as has been done for the City of Turin, both for law enforcement and over the course of experimental programs such as the CTE Next call for testing. While it does present some issues, such as when subjects are in very close proximity, it still allows for effective and precise anonymous people tracking through lightweight algorithms over relatively inexpensive hardware.

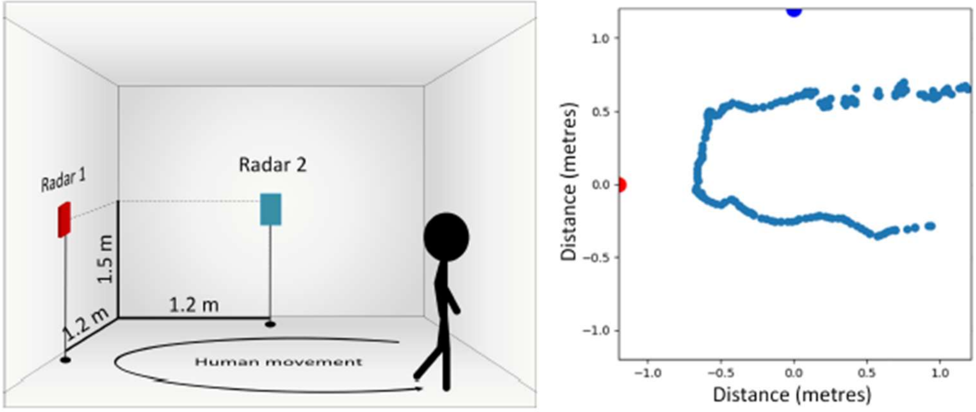

Figure 17 - Example of human tracking. Left: room setup and the movement of the human. Right: human location tracked by the radars

For all of the aforementioned devices, the gathered and elaborated data are uploaded via the cloud to Amazon Web Services servers, so that they may be monitored through the Dropper visualization dashboard, both by the customer and, for technical and maintenance reasons, by the company, too.

Chapter Three

The Dropper solutions technologies: applications to industrial settings

The technologies illustrated in the last two chapters, both older and newer, both those included and those excluded from the Dropper product offer, have all found a number of varied applications. While the current Dropper business model is centred around people counting and crowd management systems, their current experience could be applied to exploring other sources of business, with minimal cost and effort to repurpose the already existing and working hardware and software. In the present chapter, I will explore a few of the most promising and studied instances of millimetre-waves radar and time-of-flight sensors used in industrial settings, in an Industry 4.0 perspective.

3.1 Millimetre-waves radar applications

3.1.1 Millimetre-waves imaging

As mentioned, millimetre-waves radar is a versatile technology with plenty of applications. However, while use cases may vary greatly throughout different industrial sectors, all the ones that pertain to this work are reliant on a specific application of mm-waves, which is **imaging**. This technique, that we have seen here used for the purpose of people counting already, has been studied for a few decades now, with developments in the last few years that have been groundbreaking enough to hypothesize, and in some cases realise, some amazing ideas that could once only be imagined.

One emblematic example, among others, on the possibilities of this technology is offered in the article by Ahmed et al., [33], which illustrates the functioning as well as the potential of microwave imaging in general. For starters, we define the millimetre-waves as ranging between 30 and 300 GHz, with a corresponding wavelength from 1 to 10 mm.

Secondly, it is also necessary to make a first fundamental differentiation between active and passive radar imaging. Passive imaging detects the radiation that an object naturally emits and/or reflects, when compared to the neutral background, and thus measures such characteristics of said objects (quantified in reflectivity, emissivity), a method most effective in outdoors settings due to the low background radiation temperature of the sky, which gives a good contrast with the object; on the other hand, active imaging systems comprise some method to illuminate the area of interest with radiation, analysing the consequently reflected or transmitted field. This latter method is more demanding, as it requires more complicated hardware capable of both receiving and transmitting radar waves, but it is also the most effective and the one most commonly used for the high-grade applications discussed in this work, therefore it is what I will reference from now on when mentioning mm-waves imaging.

Then, we can look to Guan et al. [34], to bring us more specifically into radar microwaves imaging. Through the proposed use case of achieving fully autonomous vehicles (so-called *level 5*, an application that will be further expanded upon in a later section), more precisely of how to tackle issues with meteorological conditions such as dense fog, the paper gets into the technological specifics of millimetre-waves radar imaging, its effectiveness also when compared to alternative systems for object detection such as LiDAR sensors or cameras which have similar uses but with different properties and behaviours. Most importantly, radars are able to penetrate, although not without increased noise, through dense fog and other severe weather conditions, unlike the aforementioned optical sensors, which can struggle is such situations. There are reasons why they are rarely preferred, and that is because optical sensors can achieve a better resolution, and radar signals are more prone to reflections and noise-inducing phenomena, in an application where a low resolution or a shape accidentally mistaken for noise by the software could cause serious jeopardy to human beings.

However, the proposed solution of this article integrates an additional software to enhance performance of radar imaging through dense fog. A machine learning algorithm is trained on 3D models of cars, first through an artificial data synthesiser and in actual experiments in outside environments with clear weather second. From there, the algorithm can basically interpolate the partial data collected in adverse conditions of the rough shape of the subject (cars) with input 3D models of cars through the recognition of some basic features in the former (like wheels, orientation etc.) to synthesise said data into formed shapes that enormously reduces (up to a tenth) the failure rate.

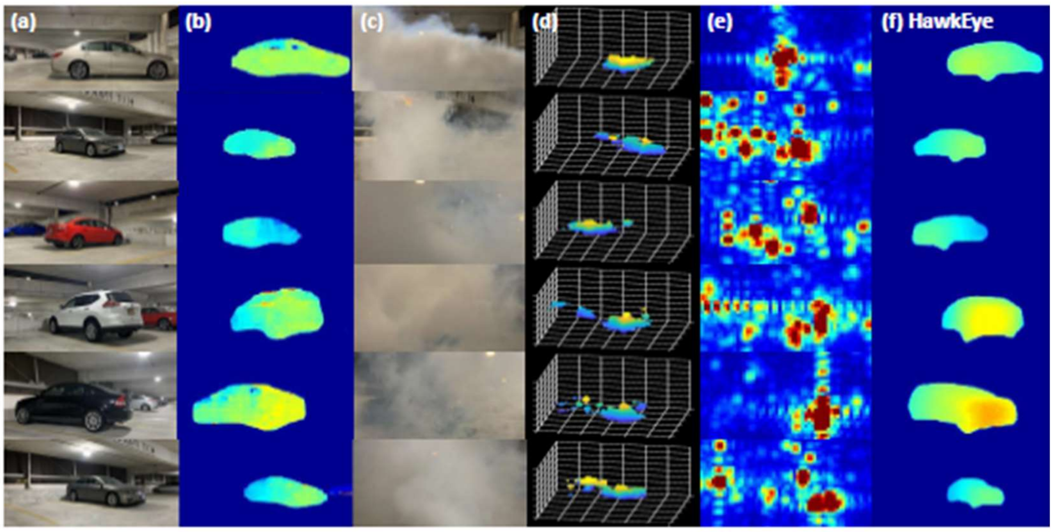

Figure 18 – The algorithm's performance with fog in scene. Column (a) and (b) show the original scene and corresponding stereo depth map. Column (c) shows the scene filled with fog. Column (d) and (e) show the radar heatmap in the fog scene as 3D point cloud and 2D front view projection. Column (f) shows the output resulting from the additional algorithm implementation

Errors on orientation, reflection recognition and car surface missed are all considerably improved by the algorithm when compared to other baselines, and this makes millimetre-waves radar imaging an even more promising method to achieve level 5 automated vehicles. The illustrated use-cases employ technologies, both regarding the hardware expertise and the algorithms, that are already core to the Dropper solution as explained previously in the last chapter, and would require minimal adjustment and pivot effort should they desire to open to a new market.

3.1.2 mm-waves radar in the automotive sector

The automotive industry has long been interested in radar technology, which has been studied and utilised for a variety of uses. One of the very first of many examples is offered in Wenger [35]. In this paper, written at a time when millimetre-waves radars were rapidly moving from research and development to pre-production phase, the author outlines the current state of the possible applications, especially relating to driver's safety and comfort which were in many cases since implemented and are now common in most vehicles.

The then recent evolution of research and development on antennas, devices and circuitry all around allowed to envision compact, lightweight and relatively cheap radar sensors, together with the huge improvements to computational capabilities brought on by the Third Industrial Revolution. The first application envisioned is the forward-looking radar (FLR) for collision avoidance systems. The maturity of the radar techniques allowed to use it to improve safety in the vehicle by designing and implementing radar-based systems to avoid accidents that already at the time could be reduced by 25% when experimented on overland busses. They are simple, single beam devices, allowing to limit additional weight while offering remarkable safety improvements, most relevantly because, as already discussed and as was already clear at the time, they are impervious to adverse weather conditions. On the other hand, using more beams was already under study, and was eventually developed for various purposes of object detection, including intelligent cruise control systems, but also to increase resolution thus improving precision and, through data analysis algorithms, recognition.

As we mentioned, since then many applications that had been imagined involving radar in the safety and comfort of those who use personal vehicles have been implemented or are still under study. The idea to improve safety and comfort in the car via the use of radar has also been expanded in Bloecher, Dickmann and Andres [36]. With an eye to the past and the other to the future, the authors highlight the potential of radar technology in helping recognise dangerous situations early on, in time to avoid serious damage and injury. They illustrate the development of, at first, the so-called *passive safety systems*, such as the invention of the crumple zone, to begin with, but then also the standardisation of the three-point seat belt: everything, that is, that is implemented with the purpose of reducing the severity of the accident event. Then came the *active safety systems*. With these, the new (and additional, not substitutional) aim is to prevent the accident event altogether. Some archetypal examples of this are the anti-lock braking system (ABS), which allows to retain control even during sharp steering; and then, several years later, the development of the electronic stability program (ESP), which allows to control all wheels separately to detect and prevent skid.

This briefly summarises the evolution of automotive security systems in the past. Then, turning our eyes to present and future developments, current and prospective security systems are an example of combination of the two aforementioned types, having the aim of both reducing the probability of an accident and its severity in cases where prevention is not enough. One very representative example in this paper is offered by the numerous security systems of a well-known German automotive sector leader, realised through the use of two short-range radars (SRR) in the front, two in the rear bumper, one mid-range radar (MRR) and one long-range radar (LRR)

The first is the Brake Assist (BAS), a program that analyses the braking habits of the driver, thus preparing to interpret a situation as of “panic braking” and, in that occasion, build up the full braking force. This system, implemented in 1996 and since installed in about 58% of European cars currently in circulation, has more recently improved its performance thanks to the inclusion of radar technology. The so-called BAS PLUS offers sometimes minimal differences in reaction times, but in these situations sometimes a short time of seconds or even fraction of seconds can prevent a disaster, and this is achievable thanks to the two short-range radars in the front bumper. While the previous system analysed braking behaviour directly from the compression rate of the braking system, the short-range and long-range radar systems monitor the traffic situation, identifying critical situations and consequently adjusting braking power and warning the driver first with a red light on the instrument cluster, and then with a sharp warning tone, sometimes even before they can notice. Tests have proven this system, allowing controlled, targeted braking with adjustable force, reduces accident event rate in a rear-end collision situation from 11 to 42 percent.

Another example is the DISTRONIC proximity system, presented in 1999. By measuring the distance to the car in front of the one in which it is installed through the use of a radar system, and comparing it to the speed of the car itself, its aim is to maintain the car at a constant speed while monitoring for potential risky situation of rear-end collision. In 2005, this was followed by the development of DISTRONIC PLUS. This renewed security and comfort system integrates the previous functionalities, adding the possibility to manage stop & go traffic situations, with the ability to maintain a constant distance from the car in front to the point of being able to bring the car it is installed upon to a complete stop. For this purpose, the previous sole radar sensor is accompanied by long-range radar (up to 200 m), as well as a medium range radar (60° Field of View, up to 60 m) and a short range one (80° FoV, 30 m range), with the result of preventing 20% of all head-to-tail accidents in Germany, with the rate rising up to 36% on motorways, and it is estimated that it could have an even greater impact in countries with a more serious situation, such as the United States.

Finally for the evolution in the rear-end collision case study security systems, radar sensors employed for the PRE-SAFE brake system aim to assist the driver's reflexes and in case of a possible distraction on their part. Basically, the system, released on the market in 2006, recognises a rear-end collision trajectory using a front facing radar sensor, estimating impact from current velocity and distance, much like other systems. During the timeline of the projected impact, then, the PRE-SAFE warns the driver at two increasingly close points in the approach to the other car, both visually and acoustically. Finally, the car will start an autonomous partial braking process: the acceleration reduction can get up to 4 m/s^2, or 0.4 g, equivalent to about 40% of the total braking potential.

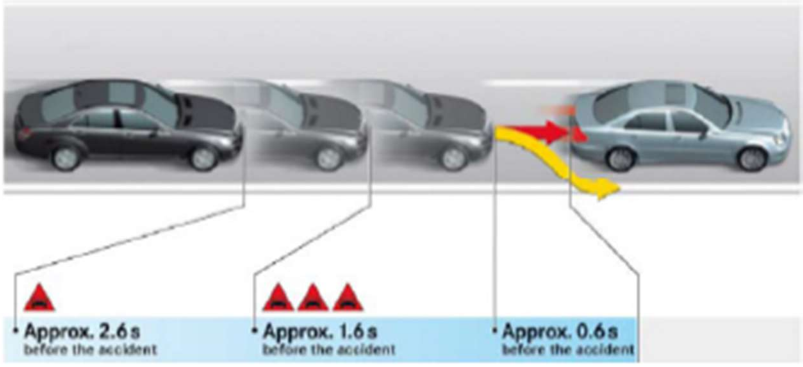

Figure 19 - Timeline of a Typical Rear-End Collision using PRE-SAFE brake system

This approach, developed specifically for this use case (the PRE-SAFE ignores events of sudden appearance of obstacles, or oncoming cars and so on for other similar instances) has a double purpose.

On the one hand, it warns the driver, possibly gathering their attention from any distraction, with both increasingly evident red lights on the dashboard and increasingly loud sound alerts. On the other hand, it adds upon previous similar such systems by allowing the car to autonomously brake and, at the very last fraction of a second, even bring the car to a complete halt if the warnings are not heeded. This double aim speaks to the ability of this security system to bridge the gap between the active and passive security systems, such as they were previously illustrated.

In combination with ESP to prevent skidding and maintain stability of the vehicle upon sudden steering, this radar-based system can significantly reduce both occurrence and force of impact, being determinant in avoiding serious or even fatal consequences of rear-end collision.

In the analysis of the evolution of rear-end collision security systems, the many advantages of adopting radar-based sensors of various ranges, much like the ones employed by Dropper for people counting and crowd management purposes, were rather clear, and they are cited in this paper as such:

- Direct distance and speed measurement
- Robust against weather influences and pollution
- Unaffected by light
- Measurement of stationary and moving objects on and in the vicinity of the road
- Invisible integration behind electromagnetically transparent materials (for example, bumpers).

In addition to this, ultra-wide band bandwidth complements high range resolution, reliable object imaging, detecting even small objects and calculating a response in near-zero time: they are necessary as much as the long-range radar are to begin early assessment of the possible impact trajectory, and only through the precise use of the integration of different range radars can these security systems work best. Only a radar-based system can offer this combination of pros, a combination that is creating an ever-growing market for this kind of technology.

In this market, the team of Dropper would have easy entry and possibly ample room to grow, having already contributed many man-hours to the study of the best hardware, as well as to the development of algorithms that do much more than simply recognise the presence of an object and calculate its distance and vector of approach.

Millimetre-waves radar imaging and object detection come together as the two centrepiece technologies in Abdu et al. [37]. This recent paper is a perfect example of radar technology used to bring the automotive industry a step closer to its next generational revolution: complete autonomous driving and the set of Advanced Driver Assistance Systems (ADAS) using deep learning that will be instrumental to said revolution. The various radar sensors considered in the frame of this paper are applied to the estimation of the physical dimension of range, velocity and angle. They differ in waveform scheme significantly, as they may be characterized by a continuous rather than a pulse wave, with or without modulation. These parameters set apart radars according to which situations they are best suited to, for instance, Frequency Modulated Continuous Wave (FMCW) is a particularly well-established radar for the aforementioned purposes, both in the academic literature and in empirical (although, sometimes, just at the experimental level) applications. Where a continuous wave transmits a constant, unmodulated frequency measuring the target radial velocity, but without range information, and a pulse radar waveform transmits short pulse signals in sequence that can establish both range and radial velocity of a moving target with the obvious downside of an inconstant information, especially if a longer pulses is employed to improve the signal-to-noise ratio, a linear frequency modulated continuous waveform has the advantages of consistent result from the signal, but is modulated so as to be able to also provide information on the range of the target, as well as its velocity, simultaneously and with high resolution. These benefits have already been explored in applications such as adaptive cruise control, and are an important part of development of autonomous driving.

At the other end of the spectrum, we have systems such as the Frequency Shift Keying (FSK). This waveform is transmitted through discrete pulses, at two different frequencies separated by a step, with the advantage of being able to achieve the same resolution and variety of information as the previous one, except this system is robust against ghost target ambiguity, a perk of discontinuous waveforms due to the fact that the reading is refreshed (and, therefore, "clean") at each iteration of the signal.

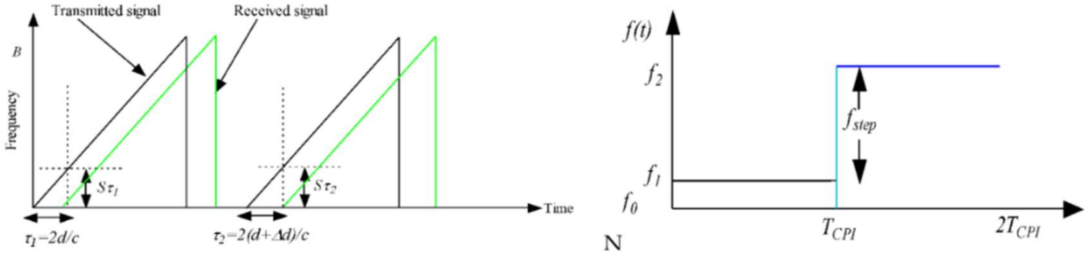

Figure 20 - Graphical comparison between continuous (FMCW) and discrete (FSK) radar waveforms

On the other hand, FSK waveforms are unable to distinguish multiple subjects with similar speed and in proximity to each other, including multiple static targets, as phase information cannot separate them. As this section has hopefully conveyed, the choice of an appropriate waveform is rarely an obvious one, and can depend on many variables including role, purpose and mission of the radar application. This is clearly beyond the scope of this work, but still is an important discussion to have when realising radar solutions such as Dropper has, and in that case specifically the team landed on a consensus to use FMCW radar waveform, as it was thought to be the most versatile and convenient to approach the varied applicative instances the company faces.

In the presented instance, of course, the radar signals are used to collect data in the form of point clouds, much like Dropper devices do in order to recognise people passing through an area. A few algorithms are then presented, implementing machine learning and deep learning techniques, with the purpose of analysing said point clouds and achieve object detection and recognition with high levels of accuracy. An issue regarding the use of machine learning algorithms is represented by the fact that, as this is a relatively new field of application for radar technology and machine learning techniques, there is a very conspicuous lack of large and public radar datasets, which are essential to the training phase of any artificial intelligence implementation process. Surely a more relevant issue here than it is in the people counting industry, as on an effective and accurate autonomous driving algorithm might depend the life of people, this issue will probably be the object of further investigation in the future. For now, the authors had to limit their findings to the few datasets on radar reading that are publicly available, whereas the team at Dropper builds their own dataset (of relatively much smaller size) for their own applications.

Finally, applications of radar technology in the automotive industry can include some instances that are less obvious, and it only goes to show the versatility of these kinds of solutions. An example of this can be found in Islam et al. [38], expanding on a previous study with similar aim but narrower scope. In the last few decades, continuous wave radar technology has been developed for application for the monitoring of human life signs and general physiological condition, with satisfactory results in the healthcare sector especially for sleep apnoea, sudden infant death syndrome, etc. Then, in recent years, a similar use of millimetre waves radar is being studied for development in an automotive setting. While radar may be fundamental in protecting human lives as a part safety and comfort systems, they can additionally be a life-saving technology by monitoring the physiological conditions of the driver and the other occupants of the vehicle, signalling possible situations of emergency that might help create dangerous driving situations. For this purpose, the authors propose an application that, thanks to the use of FMCW radar waveform, would be able to follow the respiratory patterns of multiple subjects without concerns regarding the possibility of confusion of the various signals.

The basic concept of this system involves the installation of a radar sensor pointed at the chest are of the subjects from a very short range (ca. .5 to 1 m). When the radar signal is reflected on the chest surfaces and sent back to the receiver, it arrives with a phase shift. The extreme precision of a millimetre-waves radar allows to recognise the very small chest compressions (ca. 1 cm), that are due to respiration, as well as its vibration, connected to the heart rate, by deducing it from said phase shift. Now, aside from the mathematical details of the systems of equations and the technical specifications put in place to create a suitable model, technical specifications that, to the extent that befits the needs and scope of the present work, have already been discussed, the empirical testing of this solution should be mentioned and illustrated.

At this stage, the experimentation was carried out in an office setting, but using the same automotive radars employed in the industry for the purposes already discussed. While the settings, distances and hardware were not dissimilar from the hypothesised application, the implications of the additional vibration given from the moving vehicle are not discussed or considered as possibly problematic. The end result is as shown in Figure 21.

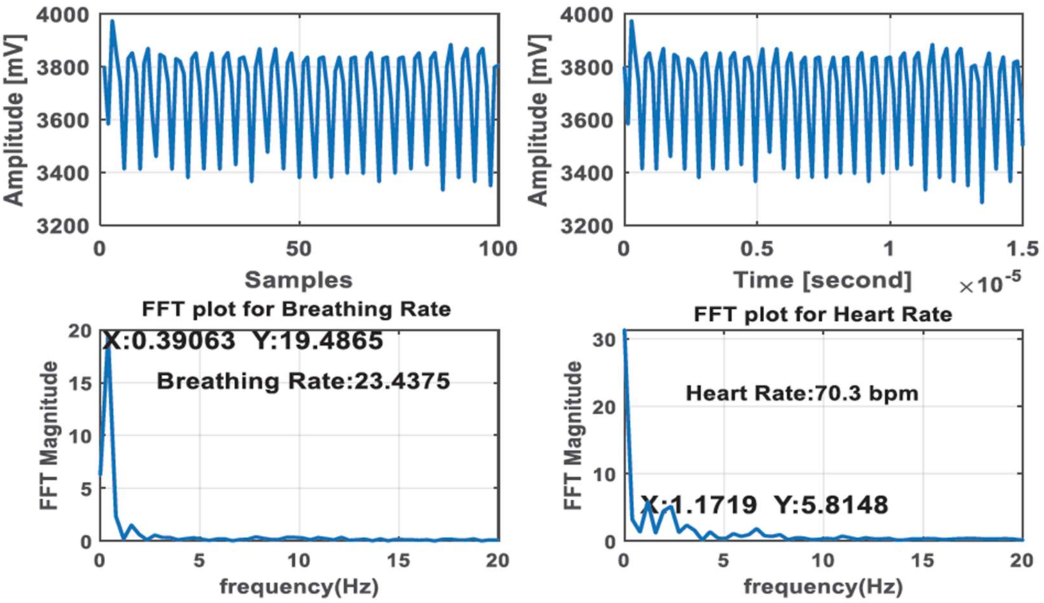

Figure 21 - Radar capture of 256 samples

The radar captured a total of 256 samples in conditions reasonably comparable to the scenario of application, and the elaboration algorithm subsequently determined a breathing rate of .39 Hz and a heart rate of 1.179 Hz. While it can be considered to be slightly more *niche*, this application for the automotive industry could represent an even more suitable fit for a possible expansion of Dropper, as it is more similar both in distances and in subjects to what the company currently employs this technology for, with respect to the object detection and collision avoidance systems.

3.2 Time-of-Flight sensors applications

As seen, radar technology can demonstrate great adaptability to varied scenarios. However, all of the technologies developed at Dropper are characterised by their versatility, and the time-of-flight technology is yet another good example of this. This particular kind of depth sensors has found applicability in many fields, including precision robotics, computer graphics and assembly assistance.

3.2.1 ToF technology in Computer Graphics

The main advantage of time-of-flight technology is the possibility to obtain accurate and fast three-dimensional readings of a target area. This can be instrumental in many fields, one chief one being **Computer Graphics**. In Kolb et al. [39], the authors give us a good rendition of the functioning of these sensors, as well as a thorough overview of how exactly they are used in this sector.

For starters, to understand how this technology functions precisely, we have to understand that there exists an inverse correlation between the maximum distance and the field of view angle. In a typical setting, you can obtain unambiguous readings as far as 7.5 m away with an angle of 30°. Increasing the distance would require a smaller angle of the field of view, and moreover illumination would have to be focused through an additional special illumination unit for the light to precisely impact the target surface and come back to the sensor with a definite reading. Lastly, the reading would be exposed to the risk of added error due to interfering light. In any event, calibration to the specific position of the sensor in the application field is always necessary to compensate for lateral distortion and similar shifting phenomena. Other issues can instead be mitigated through other systems like Suppression of Background Intensity (SBI), which most ToF cameras support nowadays. Moreover, low resolution is a factor that remains specific to this kind of sensors, and while this was mentioned as a perk in the field of people counting, as it helps maintain the privacy of the subjects being shot, for any other field where anonymity is not an issue, this represents an obvious drawback in this instance. As for the experimentation itself, it was partly carried out thanks to the possibility to simulate a ToF camera. While this is entirely possible, in order to evaluate preciseness of sensor parameters in a setup, it can also result inefficient, serving very little purpose except for when on-field calibration is extremely inconvenient.

Nevertheless, these simulators are feasible and often used in development phases.

The calibrations are central to a successful implementation of this kind of system. For instance, these sensors are subject to a systematic error that needs to be checked in comparison to some kind of reference data. This is called a “ground truth”, a 100% validated data that is gathered in a way that is independent from the device, including at times by manual count, and that can even be used as input to a supervised training algorithm in the case of Artificial Intelligence implementations of the software. The means to ascertain a ground truth can be varied and dependant on the application. For example, robotics-related applications can refer to the known position of the robot’s tool centre point, and said real position can be validated also by visual check with the aid of a modelisation of the plane upon which the objects move, such as a checkboard. However, some errors require more complex approaches, namely that some errors can be examined to find patterns, like a sinusoidal function, and be consequently corrected by compensation, or even get as detailed as to correct single outlier pixels. These corrections can reduce the actual error in the final reading of a position to the point where it is irrelevant, in the centimetre order of magnitude, more than enough precision to obtain satisfactory readings of an area meant to be then transposed into computer graphics.

In addition to outlier elimination and systematic error compensation, ToF camera data needs further elaboration before it is ready to be used in any computer graphics instance, as well as in basically any other. As mentioned, low resolution is only an advantage if the ToF sensor is used in a privacy sensitive application, but if the final aim is to obtain imaging of an area, as detailed as possible and as accurate as possible, sometimes ToF depth images can be integrated with the use of traditional 2D cameras. This process can not only increase detail level of the final product, but it is also able to assign colouring to each pixel, thus attaining colour pictures as an end result. The two (or more, as a ToF can be assisted by more than one conventional camera) image readings are fused and combined by a third CPU, using a few different approaches:

> In some approaches a rather simple data fusion scheme is implemented by mapping the ToF pixel as 3D point onto the 2D image plane, resulting in a single color respectively grayscale value per ToF pixel […]. A more sophisticated approach […] projects the portion of the RGB image corresponding to a representative 3D ToF pixel geometry, e.g. a quad, using texture mapping techniques. Furthermore, occlusion artifacts in the near range of the binocular camera rig are detected. Huhle et al. […] present a range data smoothing based on Markov Random Fields (MRFs). This idea was adopted from Diebel et al. […], who improved the resolution of a low-resolution range maps not acquired by ToF-cameras by fusion with high-resolution color images. These methods exploit the fact that depth discontinuities often co-occur with color or brightness discontinuities. Huhle et al. […] proposed a fusion scheme which incorporates an outlier removal and range data smoothing based on the combined color and depth data in their non-local denoising scheme. Yang et al. […] combine a high-resolution one or several color image with a depth image by upscaling depth to color resolution. They apply bilateral filtering and sub-pixel smoothing on the depth data with good results.

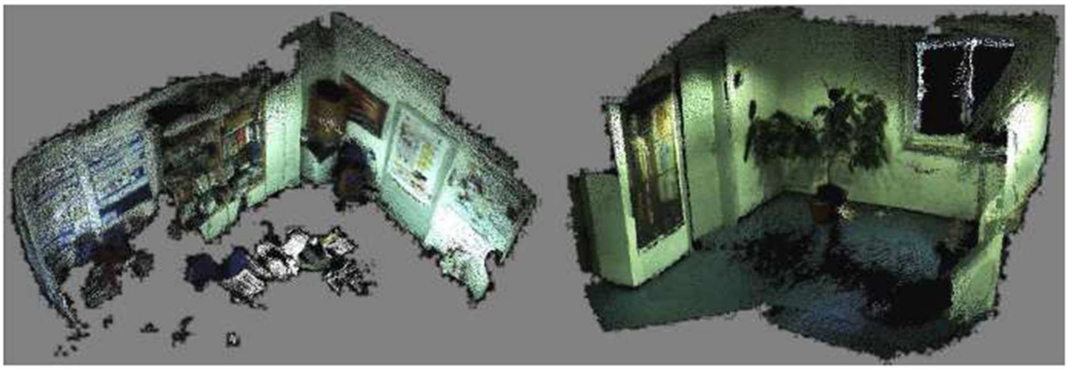

Figure 22 - Two office scenes acquired using a 2D/3D camera combination (seen from a third person view)

As seen, the options are many, but the advantages are similar: depth reading with colour images that can be easily utilized in a computer graphics application for a variety of uses, ranging from the videogame industry to the robotics field in heavy industry to the point where an intensity image, based exclusively on contrast differences, is able to approximate even the features of a human face. This is only possible, of course, if the data collection is carried out in the correct illumination conditions that satisfy the shading constraint, by which the reflectance properties need to be known and they are what allows for a certain surface to be interpreted in three dimensions: the use of already defined models can facilitate the surface analysis.

Figure 23 - From left: Intensity image; lateral view of raw measured surface; and surface reconstructed using the shading constraint in lateral and frontal views

Given the effectiveness, as well as all the limits, of such system, taken into consideration the necessary corrections to achieve the best results, and considered the potential of this technology especially when combined with complementary technologies in cases where privacy issues are not an item, the almost newborn market for ToF sensors is already growing, with the first cameras for commercial sale now available. Dropper is, in its own way, a pioneer of this technology, this being the first ever kind of sensor developed by the company, and this puts it in a unique position to evaluate the pros and cons of a possible pivot towards computer graphics solutions, especially since that is probably the industrial application requiring the least adaptation costs and efforts on their part. This allows them also to consider the possibility of an integration with additional cameras, as those are the combinations that appear to achieve the best accuracy and definition such that they are usable in augmented and mixed reality applications, where real and virtual objects integrate.

3.2.2 ToF technology in Robotics for obstacle avoidance

Mentioned in the previous source as well, the robotics and automation sector more in general can represent a sound and interesting sector for an industry expansion, with a lot of potential. The methods and systems detailed in Bascetta et al. [40] are both well established and of relatively easy access for the company, should it wish to do so. The current state of robot-human interaction in industrial settings is at a fundamental juncture. The paradigm itself of how that interaction functions in the workplace is on the verge of a shift: this will surely involve the existence of both human workers and robots working in close proximity (the expected *coexistence*), and hopefully reach the point where their interaction is common practice (the hoped-for *cooperation*). In order to prepare for these evolutions of the workplace, several safety concerns will need to be dealt with, to ensure that workers interacting or even just existing in proximity of automated machines come to no harm.

While some actions can be taken pre-emptively, in a passive way, such as designing inherently lightweight robots, thus minimising the seriousness of potential impacts, active action could instead involve additional and improved sensor equipment. This can at times be necessary, as lightweight robots are not always the appropriate choice for some tasks, at least not to the point where they can remain safe for the workers they have to coexist with, and LiDAR or time-of-flight as well as camera-based vision systems are considered to be the best application fit, often mounted near the robot or above rather than on the robot itself. Of course, this approach opens to possible occlusions of the signal, possibly undetected, thus remaining unable to guarantee the safety of the workers beyond a certain proximity. On the other hand, systems that mount a set of devices directly onto the robot are subjected to all kinds of effects that either reduce their reliability, or require too many specific adjustments and calibration to make economic and operational sense.

As a matter of fact, environmental conditions (e.g., materials and their properties) create issues that can be mitigated through the use of time-of-flight sensors, which can be employed as distance sensors with two different approaches. In order to understand both approaches, we need to introduce the instrumental "Virtual Wall", which requires the definition of a protective area[4] around an obstacle. Upon reaching the virtual wall, the robot is programmed to perceive an artificial repulsive force and, being the perception of repulsive forces the very base for the movement algorithms in robots, this in turn makes it move away from the "wall", and thus the obstacle. The virtual wall does not simply repel the robot the minute they touch, but it impresses a virtual force that is proportional to the degree of penetration of the robot into the wall itself, in a way akin to an elastic force.

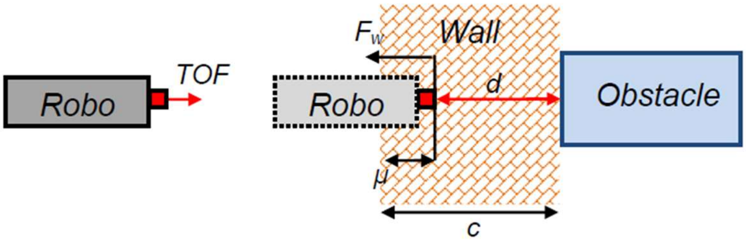

Figure 24 - Graphic representation of the "Virtual Wall" approach

In the graphic representation above, μ is the aforementioned penetration, whenever c (the wall thickness) is greater than d (distance between the robot and the obstacle). As can be seen in the system of equations below, the virtual force impressed on the robot is dependent (differentially) on said penetration (and the relative speed), so long as it is present

$$F_w = \begin{cases} K_w \mu + D_w \dot{\mu} & \text{if } \mu > 0 \\ 0 & \text{if } \mu < 0 \end{cases}$$

---

[4] In the presented paper, the protective "area", is in reality a protective distance, as the system is modelled in a 1D setting. With a similar principle, the 2D extension is presented later in the paper, and will be disregarded for the purpose of this work.

as well as the gains of the virtual wall K_w and D_w, determined from the fact that when the distance reaches its possible minimum (d_{min}), the force applied is at its maximum (F_{max}) according to the following relations depending on the mass of the robot as well (M).

$$K_w = \frac{F_{\max}}{c - d_{\min}}, \quad D_w = \sqrt{4MK_w}$$

In addition to this, the model evolves a way for the robot to carry out a trajectory once it impacts the virtual wall, dependant on the applied force, and that adjusts after each time step according to it. Following the proposed model, a threshold force F_{stop} is selected (for example, at one half of the maximum allowable force F_w), so that the trajectory will begin reverting to get away from the obstacle as soon as F_{stop} is reached[5].

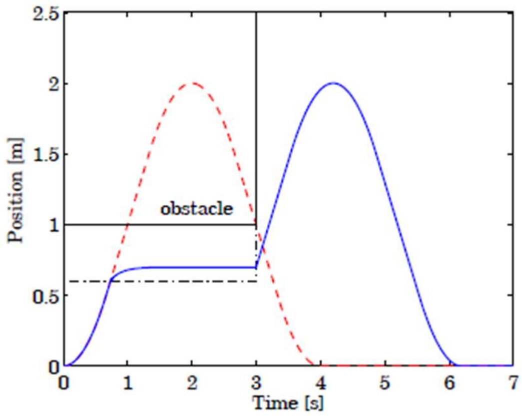

Figure 25 - Graphic representation of trajectory change after the application of a virtual wall (dashdot line), from the dashed line to the solid one

The discussed approach is called GCT, Geometry Consistent Trajectory, as it revolves around the geometrical properties of the trajectory, which are preserved, while the time-related properties are released. In other words, the trajectory will be completed coherently with what the robot was meant to do in the first place, but the time this requires might vary depending on how many

---

[5] That is, assuming that the robot is moving towards the obstacle. In case it isn't, the trajectory velocity will have the opposite sign and, of course, no adjustment will be necessary to the trajectory.

obstacles the robot encounters and how long it takes to avoid them, through a trial-and-error feedback loop of attempts and corrections. The opposite approach takes the name of Time Consistent Trajectory (TCT), and its implications can be inferred from the name. This second anti-collision system enforces the desired dynamic behaviour of the robot, that is that the machine is able to track a reference point in its intended trajectory, including when meeting with the virtual wall. Upon "impact", however, it does not immediately revert the trajectory after detecting a certain threshold force is applied to it, like in the GCT; instead, the robot applies an opposite and equal force, an impedance that results in a system of two contrasting springs. The exact selection of an elastic parameter K to assign to the opposing force from the robot is a delicate procedure, having to balance the need for a big enough coefficient that the movement does not appear sluggish and inconsistent, and a small enough coefficient so that the reaction is not too stiff, too sudden.

Letting that that question is settled upon, the resulting system will have a final trajectory that is, indeed, time consistent, but does not follow the same geometry: in other words, the trajectory will be modified, and its furthest point will not be reached, but it will complete in the intended time and with the same end point it was supposed to reach to begin with.

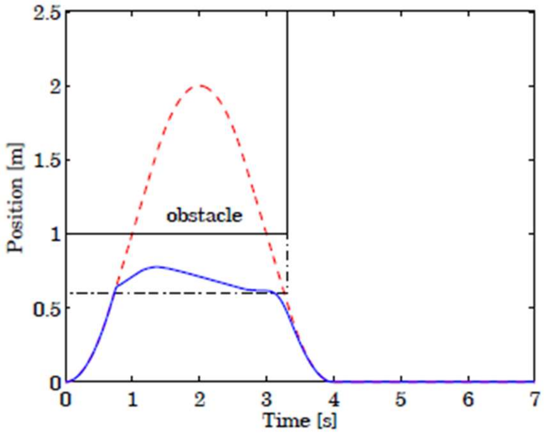

Figure 26 - The change in trajectory after a TCT virtual wall (dashdot line) is applied, from the dashed line to the solid line

In the experimental setting proposed by this paper, a two-head linear motor plays the part of the robot, and the commercial ToF sensor used is mounted on top of one of the two heads. The sensor detects the distance to the obstacle, thus creating the virtual wall as the set of points that are at a certain fixed distance from the robot and the obstacle, in-between. While the maximum range of the commercial model used is rather limited (vision limited to 16 cm distant objects), as per product specifications, its response time is short enough that the overall model allows for satisfactory results of safety and security of the imagined worker. The sensor was tested with the sudden appearance of an obstacle and had a step reaction that was in the order of magnitude of 2 ms. From this testing phase we can surmise the potential of the Time-of-Flight sensor as a quick and relatively simple method to measure distances, in a way that is reliable enough to have the safety of human workers depend on it.

As a matter of fact, the end result of both experimented approaches, while determining different trajectories, are characterised by an ample margin of safety with respect to the d_{min} of the virtual wall. This solidifies the ToF sensors as a suitable solution for obstacle detection in robotics, through a method that can be relatively easily implemented.

Compared to other applications, it would require Dropper a higher investment, both financial and in manhours, to adapt their current solution to something of this kind. The hardware used by the company is thought for longer ranges and lower definition, so that integration would add to the obvious need to develop brand new software suitable to this purpose. However, considering the constant rise and expansion of automation both as a social and cultural phenomenon and as an industry, evaluating the possible economic advantages brought by, among other things, a rising trend of public incentives for the development of this industry, the cost benefit ratio of such a pivot, especially at the startup’s young age, could appear to be a favourable long-term investment.

3.2.3 ToF technology in Assembly Line Assistance

In a previous section of this paragraph, we have seen how industrial or commercial grade Time-of-Flight sensors can be deployed effectively to aid in mixed-reality computer graphics applications, including for use in the gaming industry. Now we come full circle, and we examine a proposed solution that does exactly the opposite.

Microsoft first developed and released the *Microsoft Kinect*, an innovative gaming system that detects and recognises body movements for gaming, in late 2010, for integration with the Seattle-based IT giant's gaming consoles (at the time, the *Xbox 360* specifically), and has since interrupted production for a few years now. Aside from the controversial and criticised commercial success in the gaming industry, the Kinect has left a perhaps even more relevant mark with its cognate meant for industrial applications, the *Azure Kinect*. The Azure Kinect is indeed still on the market, and it is equipped with a time-of-flight depth sensor in the near infrared light range, as well as microphones for sound detection. It has found a home in several distinct industrial applications, which are examined partly in Niedermayr and Wolfartsberger [41], particularly with an experimental focus on assembly line assistance for human operators.

The basic concept of the experimentation, in simple words, is as follows. The paper proposes a solution that involves the use of a depth camera in concert with a traditional color camera, in order to acquire at a higher definition and with coloured pixels, in a way not dissimilar to other systems already discussed here. The double system, conjoined in the Azure Kinect, is placed above the assembly line, at various checkpoints located in between the automated operations. By obtaining a reading of the items below, the aim is to program the system to recognise whether an automated operation was executed correctly or not. Of course, this can be done by comparison, so an image of the correctly assembled item is fed to the system before it starts evaluating the new piece undergoing the process. Based on this comparison, the algorithm will iteratively evaluate the deviation from the reference to determine the correctness of the assembly.

This system is faced with a few issues. First of all, like almost any ToF sensor, the system has to deal with problems related to the physical phenomena that have to do with the lighting conditions that affect this kind of measurement. More specifically, bright lights, including sunlight, can interfere with the near infrared beams projected and return an imprecise reading, the same can happen with very reflective surfaces. When faced with particularly IR-reflective[6] materials, it is intrinsic in the functioning of a time-of-flight sensor that said reflections create a multipath error, that is a kind of misrepresentation where the normal path of the light from the sensor to the object and back is actually split and diverted, rendering an imprecise, noisy end result with unexpected variations. It is appropriate to specify, that the error here discussed is in the order of magnitude of the millimetre, reaching a worst-case of 15-20 mm only where the experimenters have purposefully increased reflection with their own proximity to the target area, and even went as far as to create disturbances with paper sheets. In most other applications, including people counting, this would have been a more than acceptable margin of error, however the presented system could cause issues in the face of this reflectivity, as 2 centimetres can be more than enough to falsely flag a piece as incorrectly assembled.

The wavelength choice is also partially problematic, as choosing a wavelength that is outside the spectrum of the human vision can delay the detection of issues with the projected beams to the examination of the final reading, bringing to unexpected errors. Then we have to consider the contextual errors. For instance, a correct recognition requires that the pieces be always fixed

---

[6] It bears reminding that not all surfaces that are reflective in the visible spectrum of light are also reflective in the infrared, and vice versa. For instance, most plastics are reflective to IR light, but not to the visible. This can also depend on the colour of the item.

in the exact same position, and that the camera is never moved, not even slightly, two constraints whose respect is hard to ensure with complete certainty in a real-life factory setting; however, they have been assumed to be satisfied for the purpose of the presented experimentation. In the experimentation phase, a machine learning like approach was followed in handling the algorithm, with a training set of comparison fed to the program, equally divided between correctly and incorrectly assembled pieces.

Considering and examining a number of different results that can be gathered by the sensor system, the authors deem the system problematic in some regards, therefore refusing to come to definitive conclusions. The system does present issues, such as the ones already discussed, relating to the physical properties of a light-based sensor technology, so the authors offer some guidelines that can be followed in order to mitigate some of those limitations:

- Operating mode: The recommended sensor mode is NFOV unbinned to get the best resolution, if the extended range of the binned mode and the wider field of view of the WFOV mode are not needed.
- Assembly mount: The object to be assembled should be clamped in a fixed and reproducible position and the Kinect is mounted rigidly above, so the assembly process can be repeated without capturing new reference images.
- Distance: The distance to the Kinect should be minimal (min. 0.5 m) without invalid pixels in the assembly region. If there are too many invalid pixels in the relevant areas, the distance should be increased, until most pixels become valid.
- Assembly position: If the object to be assembled does no[t] cover the whole field of view, it should be placed preferably in the center of the image (more reflected light, best precision). If there is a spot with invalid pixels in the center due to overexposure (on a highly reflective surface), either increase the distance or move the object slightly out of the center.
- Filtered measurements: Several consecutive depth images should be averaged to achieve higher precision. For the comparison images, we calculated the average over only ten frames to get a faster response. Because the reference images are captured only once, more images can be used for those to obtain more accurate reference images.

Following these few, simple rules allows us to envision the results of this method more clearly. We could then hypothesize the model to achieve results that go from a best to a worst case, as a rule. If we suppose that disturbances are not present at all or, at worst, negligible, that the objects are completely opaque, causing no reflections in the IR spectrum, that the distance is ideal and so is the area occupied by the object, also in relation to the overall target area footprint, then we can expect the best accuracy from the environmental conditions. On the other hand, conditions can pretty much always worsen, making it difficult to determine what "rock bottom" is for this experimental setting. We can, however, imagine, in a *realistic worst case* sort of thought process, a highly reflective material which is theorised to cause a maximum offset of 2 cm, that is of course unless the object is transparent, which would make it undetectable to this system.

In conclusion, the authors mention ostensibly the technical characteristics of the Kinect Azure, as well as the physical properties of the materials that shift through under the lenses of the ToF camera, as the main limitations of this model. While little can be done to foresee, or even mitigate, the issues relating to the environmental conditions (distance, lighting, etc.), the ones that refer to the hardware can instead be turned to an advantage for Dropper, in the case of a possible expansion into this industry. As a matter of fact, the high level of expertise of the company in this particular technology has already been invested in the research of more performing equipment, and the definition of algorithms that resemble more the ones needed for this application (the use of bounding boxes for the subject, the technical need to distinguish and identify a whole subject rather than single movements, the optimization for longer ranges than the Kinect) puts Dropper ahead, too. In addition to this, we should consider that the Kinect was not originally designed for industrial applications, and while this could be said of the product offered by Dropper as well, the latter's relatively reduced size and scope of operation allow it a possibility of pivot that is not available to a company the size of Microsoft.

3.3 Wi-Fi Sniffing devices for Cyber Security

Lastly, we consider what application the Wi-Fi Sniffing technology can have in relation to Industry 4.0. One of the main enabling technologies of Industry 4.0, as illustrated at the beginning of this work, is Cyber Security. Unfortunately, Wi-Fi sniffers are often on the "wrong side" in that domain, in that they are used as a precursor to a hacking attack. However, the other side of the coin is that, much like hackers that can sometimes be hired to test the security of a protected network, so can Wi-Fi sniffers be employed in stress tests for similar purposes, as we can see in Sontowski et al. [42]. This application is a slight variation from the previous ones, as it is not, strictly speaking, an industrial setting, since the paper pertains to the agricultural application of such a technology. However, the environment of a smart farming infrastructure, heavily reliant on IoT features, is tightly related to the principles and enabling technologies of Industry 4.0, and the need to protect from cyber-attacks is a common point with any modern industry. "Smart farming" is the name given to a recent and ever more popular trend of modernising agriculture in the face of such great demands as today's market can make. In the United States, one of the largest crop producers worldwide, this phenomenon has begun sooner than in other countries, modernising the agricultural production process which contributed $1.053 trillion to the US Gross Domestic Product in 2017, according to the logics and techniques of the Fourth Industrial Revolution. This modernisation has included control on the water supply, recording of the soil moisture levels, pesticide spraying through the use of drones, etc., all aimed at increasing both quantity and quality of the crops yielded.

Clearly, including information systems in the production process, and making those process depend on the IT, exposes farms to the risk of so-called *Cyber-Agroterrorism* attacks. The risk is great, as a malicious attack on these systems can cause any kind of grave damage both to the farmers’ livelihood and to the wellbeing of the soil and environment itself, some examples involving flooded fields, excessive use of pesticide and crop destruction.

One accurate example of how a possible cyber agroterrorism attack might work, used as a representative instance in the mentioned paper, is through a DoS attack. Generally speaking, a Denial of Service (DoS) attack consists of a third party inserting themselves in a system, in order to interrupt the normal, established connection, usually by flooding a functioning server with requests in order to exhaust its resources, so it stops working. In this particular case, the attack was executed slightly differently, although it falls under the same definition, as it was carried out in two steps. First, with the technique of Wi-Fi sniffing, packets were located[7] to ensure the connectivity of the IoT device (in the context of this experimentation, exemplified by a Raspberry Pi) to the network, as well as to monitor traffic on that network. Then, in the second step, the attack proper takes place. With the help of a deauther tool, the interested access point is located, selected (the Raspberry Pi, in this case) and then a “poisoned” packet with a deauthentication frame is sent to it, disconnecting it from its server.

Seemingly a simple and unimportant thing, this action can actually have disastrous consequences. As mentioned at the beginning of this thesis, the Internet and a wireless connection between the devices are fundamental steppingstones to any IoT system. Disconnecting this kind of device from the Internet is about everything you need to do to make them useless and, in the proposed scenario, an IoT device losing connection to the cloud will not only be

---

[7] Incidentally, Wireshark was used to sniff packets, the same software that the Dropper Wi-Fi Sniffer is based on.

unable to upload data on readings, if that is its purpose at all, it will most importantly be barred from obtaining any form of feedback or order necessary to its correct functioning, meaning fields could go unwatered, pesticides unsprayed (or the opposite), and so on and so forth, with the aforementioned consequences both to the farmers and to the environment.

Although the illustrated application represents more of an experimental case for what concerns the use of Wi-Fi sniffers, it should also be considered how the smart farming environment is but an example of just how many fields this kind of devices can be of instrumental help in, to train against malicious cyber activity. A useful tool to proxy for something quite larger, smart farming is actually joined by any IoT infrastructure that relies on Wi-Fi technology, as they are all vulnerable to DoS attacks as well. All of them then represent a need for prevention and, therefore, a possible new market to explore for the Dropper Wi-Fi Sniffers, with minimal to no adaptation efforts.

Concluding remarks

The startup environment is, by definition, a hotbed of innovation in continuous expansion and redefinition, where companies grow by changing, learn by selling, and keep always on the move, otherwise successful outcomes become very unlikely, sometimes hopefully receiving some guidance. In some regard, such is the case for Dropper s.r.l., thanks to the incubation program at I3P which, if it does not guarantee success, it has certainly given a few extra helping hands to develop the initial ideas of the two founders.

It is also because of this, that the path of Dropper has been, already in its short life, made of many turns and changes of direction: starting from the sole idea of a time-of-flight people counter sensor, the company has time and again added and modified both technical characteristics and business goals. For starters, the offer was expanded to include both the Wi-Fi Sniffer and the mm-waves radar sensor, opening up to new markets. Of course, the company consequently started chasing those new markets providing their expanded offer to an expanded demand. Originally intended mainly for large-scale retail trade, the Dropper solutions have already reached stores, museums, education institutions and local administration, and are now further expanding by trying to conclude an agreement with corporate clients in space management for layout optimization and energy efficiency.

Some of this has been summarized in this work, which started with an overview of all the enabling technologies that are crucial to development of the Dropper solutions, making the company an example of Industry 4.0. Moreover, some older or less suitable approaches to people counting have been illustrated in that same chapter. This usefully showed a comparison, in the second chapter, with the Dropper solutions, the most modern and discreet technologies for people counting currently developed, together with a more in-depth elaboration on the aforementioned history of the company. These solutions are believed to be the most promising by the Dropper team, also by virtue of the versatility of their technologies, how easily they can integrate, both with each other and with other technologies, whether pre-existing or specifically developed. All this becomes evident in the last chapter, as it concludes this work with very tangible examples of how those same technologies that today support the Dropper solutions can be applied to industrial production processes in many different fields, from automotive to robotics, and from life signs detection to cybersecurity.

In light of this, considering that some applications especially would require extremely little adaptation from the current Dropper solutions, the company might consider a pivot towards a more industrial approach of these technologies. While the people counting and crowd management market is quite consistent, bound to reach $2 billion in the next five years, markets like that of the sensors for automotive application and autonomous driving represent an astonishing $21.4 billion, poised to reach $47.5 billion by 2026 and $55 billion by 2032 (CAGR: 10.1%). And that is only considering radar sensors applied to the automotive sector. Of course, the automotive industry is also famous for its barriers to entry. While these apply mainly to the manufacturing of the vehicles themselves, it is reasonable to expect that, because of those barriers, producers will be reluctant to rely on suppliers that are less than well known, established and thoroughly verified by at least a few competitors. Be that as it may, and while the decision will certainly fall upon the management team at Dropper, it is hard to ignore such a large pie, even if the potential piece available to the company is quite smaller, and hopefully this thesis will have provided some limited insight on that decision.

References

[1] *MarketsandMarkets.com*, URL = https://www.marketsandmarkets.com/Market-Reports/people-counting-system-market-206905443.html .

[2] ERBOZ, G., *How To Define Industry 4.0: Main Pillars Of Industry 4.0*, «Managerial Trends in the Development of Enterprises in Globalization Era», (2017).

[3] SUNITHA, S. and JANTHAKAL, R., *Designing and development of a new consumption model from big data to form Data-as-a-Product (DaaP)*, «2017 International Conference on Innovative Mechanisms for Industry Applications» (ICIMIA), (2017), pp. 633-636.

[4] MARTIN, K., *Ethical Issues in the Big Data Industry*, «MIS Quarterly Executive», (14)2 (2015), pp. 67-85.

[5] ALI, Z. H., ALI, H. A. and BADAWY, M. M., *Internet of Things (IoT): definitions, challenges and recent research directions*, «International Journal of Computer Applications», (128)1 (2015), pp. 37-47.

[6] QIAN, L., LUO, Z., DU, Y. and GUO, L., *Cloud Computing: An Overview*, «IEEE International Conference on Cloud Computing», (2009).

[7] *Eur-lex.europa.eu*, URL = https://eur-lex.europa.eu/eli/reg/2016/679/oj .

[8] RAJ, M. A. M., *Automatic Room Light Controller With Bidirectional Visitor Counter* (Doctoral dissertation, Mahatma Gandhi University, Kottayam), (2013).

[9] MOHAMMADMORADI, H., MUNIR, S., GNAWALI, O. and SHELTON, C., *Measuring People-Flow through Doorways Using Easy-to-Install IR Array Sensors*, «2017 13th International Conference on Distributed Computing in Sensor Systems (DCOSS)», (2017), pp. 35-43.

[10] GADE R. and MOESLUND, T. B., *Thermal Cameras and Applications: A Survey*. «Machine Vision & Applications», (2014), Jan;25(1), pp. 245-262.

[11] *V-count.com*, URL = https://v-count.com/people-counting-technologies-a-comprehensive-guide/ .

[12] *European Data Protection Supervisor*, URL = https://edps.europa.eu/data-protection/data-protection/reference-library/video-surveillance_en .

[13] *Garante per la Protezione dei Dati Personali*, URL = https://www.garanteprivacy.it/temi/videosorveglianza

[14] BAMRUNGTHAI, P. and PUENGSAWAD, S., *Robust people counting using a region-based approach for a monocular vision system*, «2015 International Conference on Science and Technology (TICST)», (2015).

[15] YAMAGUCHI, S., KIMURA, M., HOSOKAWA, J. and TAKEMURA, Y., *Stereoscopic video movie camera "3D-CAM"*, «Consumer Products Engineering Lab. Toshiba Corporation», (1988), pp. 178-179.

[16] BEYMER, D., *Person counting using stereo*, «Proceedings Workshop on Human Motion», (2000), pp. 127-133.

[17] BURBANO, A., VASILIU, M. and BOUAZIZ, S., *3D-Sensing Distributed Embedded System for People Tracking and Counting*, «Proceedings of the 2015 International Conference on Computational Science and Computational Intelligence (CSCI) (CSCI '15)», (2015), pp. 470–475.

[18] TERADA, K., YOSHIDA, D., OE, S. and YAMAGUCHI, J., *A method of counting the passing people by using the stereo images*, «Image Processing, Procs. Int.l Conf. on», (1999) 2, pp. 338–342.

[19] RAUTER, M., *Reliable human detection and tracking in top-view depth images*, «Conference on Computer Vision and Pattern Recognition Workshops», (2013), pp. 529–534.

[20] KIRCHNER, N., ALEMPIJEVIC, A., DAI, X., PLÖGER, P. and VENKAT, R. K., *A robust people detection, tracking, and counting system*, «International Conference on Robotics and Automation», (2014).

[21] Fu, H., MA, H. and XIAO, H., *Real-time accurate crowd counting based on RGB-D information"*, «19th IEEE International Conference on Image Processing (ICIP)», (2012), pp. 2685–2688.

[22] GANDY, O. H., Jr., *Public Opinion Surveys and the Formation of Privacy Policy*, «Journal of Social Issues», 59 (2003), pp. 283-299.

[23] GREENLEAF, G. and COTTIER, B., *Data Privacy Laws and Bills: Growth in Africa, GDPR Influence*, «152 Privacy Laws & Business International Report», (2018), pp. 11-13.

[24] *Blueoceanstrategy.com*, URL = https://www.blueoceanstrategy.com/what-is-blue-ocean-strategy/ .

[25] *V-Count.com*, URL = https://it.v-count.com/ .

[26] *Density.io*, URL = https://www.density.io/ .

[27] *Affluences.com*, URL = https://www.pro.affluences.com/ .

[28] *Visionarea.net*, URL = https://www.visionarea.net/ .

[29] VATTAPPARAMBAN, E. G. *People counting and occupancy monitoring using wifi probe requests and unmanned aerial vehicles*, «FIU Electronic Theses and Dissertations» (2016).

[30] URAS, M., COSSU, R., FERRARA, E., BAGDASAR, O., LIOTTA, A. and ATZORI, L., *Wifi probes sniffing: an artificial intelligence based approach for mac addresses de-randomization*, «2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)», (2020), pp. 1-6.

[31] STEC, M., HERRMANN, V. and STABERNACK, B., *Using time-of-flight sensors for people counting applications*, «2019 Conference on Design and Architectures for Signal and Image Processing (DASIP)», (2019), pp. 59-64.

[32] CUI, H. and DAHNOUN, N., *High precision human detection and tracking using millimeter-wave radars*, «IEEE Aerospace and Electronic Systems Magazine», 36 (2021) 1, pp. 22-32.

[33] AHMED, S. S., SCHIESSL, A., GUMBMANN, F., TIEBOUT, M., METHFESSEL, S. and SCHMIDT, L.-P., *Advanced Microwave Imaging*, «IEEE Microwave Magazine», 13 (2012) 6, pp. 26-43.

[34] GUAN, J., MADANI, S., JOG, S., GUPTA, S. and HASSANIEH, H., *Through fog high-resolution imaging using millimeter wave radar*, «Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition», (2020), pp. 11464-11473.

[35] WENGER, J., *Automotive mm-wave radar: status and trends in system design and technology*, «IET Conference Proceedings», (1998), pp. 1-7.

[36] BLOECHER, H. - L., DICKMANN, J. and ANDRES, M., *Automotive active safety & comfort functions using radar*, «2009 IEEE International Conference on Ultra-Wideband», (2009), pp. 490-494.

[37] ABDU, F. J., ZHANG, Y., FU, M., LI, Y. and DENG, Z., *Application of deep learning on millimeter-wave radar signals: A review*, «Sensors», 21 (2021) 6.

[38] ISLAM, S. M. M., MOTOYAMA, N., PACHECO, S. and LUBECKE V. M., *Non-Contact Vital Signs Monitoring for Multiple Subjects Using a Millimeter Wave FMCW Automotive Radar*, «IEEE/MTT-S International Microwave Symposium (IMS)», (2020), pp. 783-786.

[39] KOLB, A., BARTH, E., KOCH, R. and LARSEN, R., *Time-of-Flight Cameras in Computer Graphics*, «Computer Graphics Forum», 29 (2010), pp. 141-159.

[40] BASCETTA, L., MAGNANI, G., ROCCO, P., MIGLIORINI, R. and PELAGATTI, M., *Anti-collision systems for robotic applications based on laser time-of-flight sensors*, «2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics», (2010), pp. 278-284.

[41] NIEDERMAYR, D. and WOLFARTSBERGER, J., *Analyzing the potential of a time-of-flight depth sensor for assembly assistance*, «Procedia Computer Science», 200 (2022), pp. 648-659.

[42] SONTOWSKI, S., GUPTA, M., CHUKKAPALLI, S. S. L., ABDELSALAM, M., MITTAL, S., JOSHI, A., and SANDHU, R., *Cyber attacks on smart farming infrastructure*, «2020 IEEE 6th International Conference on Collaboration and Internet Computing (CIC)», (2020), pp. 135-143.