Czech University of Life Sciences Prague

Faculty of Economics and Management

Department of Information Technologies

Diploma Thesis

Application of artificial intelligence in crisis management

Margarita Gicheva

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Economics and Management

DIPLOMA THESIS ASSIGNMENT

Bc. Margarita Gicheva

Systems Engineering and Informatics
Informatics

Thesis title

Application of artificial intelligence in crisis management

Objectives of thesis

The main objective of this thesis is to compare performance of object recognition by AI and humans during a crisis situation.

The partial goals of the thesis are:
- to make an overview of common problems of AI and how they affect detection of behavioral and social patterns in public spaces;
- to review and compare existing software that can be used in crisis management;
- to conduct a user study in order to compare the ability of humans and AI to track an object and propose possible improvements.

Methodology

To achieve the set objectives, there is a need to review and analyze the literature on the applications of artificial intelligence in public spaces. As one of the goals of this thesis is an overview of existing software, some of existing most promising systems will be studied. For practical part of the research AI and human intelligence will be analyzed in a user study in a usability lab. The performance of the human participants will be measured, statistically analyzed and contrasted with the performance of AI-based object recognition. The results of the analysis will be interpreted, and conclusions will be formulated.

The proposed extent of the thesis

60 – 80 pages

Keywords

Artificial intelligence, crisis management, security, analysis, object recognition, threat detection.

Recommended information sources

BEDAGKAR-GALA, Apurva; SHAH, Shishir K. A survey of approaches and trends in person re-identification. Image and Vision Computing, 2014, 32.4: 270-286.

CARROZZO, G., et al. Interoperation of IoT Platforms in Confined Smart Spaces: The SymbIoTe Smart Space Architecture. In: 2018 Fifth International Conference on Internet of Things: Systems, Management and Security. IEEE, 2018. p. 38-45.

GARCÍA, Cristian González, et al. Midgar: Detection of people through computer vision in the Internet of Things scenarios to improve the security in Smart Cities, Smart Towns, and Smart Homes. Future Generation Computer Systems, 2017, 76: 301-313.

HADDADPAJOUH, Hamed; PARIZI, Reza. A Survey on Internet of Things Security: Requirements, Challenges, and Solutions. Internet of Things, 2019, 100129.

KO, Kwang-Eun; SIM, Kwee-Bo. Deep convolutional framework for abnormal behavior detection in a smart surveillance system. Engineering Applications of Artificial Intelligence, 2018, 67: 226-234.

YOGAMEENA, B.; NAGANANTHINI, C. Computer vision based crowd disaster avoidance system: A survey. International journal of disaster risk reduction, 2017, 22: 95-129.

Expected date of thesis defence

2020/21 WS – FEM (February 2021)

The Diploma Thesis Supervisor

Ing. Miloš Ulman, Ph.D.

Supervising department

Department of Information Technologies

Advisor of thesis

Akshay Pottathil, Ph.D.

Electronic approval: 20. 7. 2020

Ing. Jiří Vaněk, Ph.D.

Head of department

Electronic approval: 5. 11. 2020

Ing. Martin Pelikán, Ph.D.

Dean

Prague on 30. 11. 2020

Declaration

I declare that I have worked on my diploma thesis titled "Application of artificial intelligence in crisis management" by myself and I have used only the sources mentioned at the end of the thesis. As the author of the diploma thesis, I declare that the thesis does not break copyrights of any their person.

In Prague on ______________________________

Acknowledgement

I would like to acknowledge everyone who played a role in my academic accomplishments. First of all, I would like to express my gratitude to my parents Irina Gicheva and Valeriy Gichev, who supported me with love and understanding. Without you, I could never have reached this current level of success.

Secondly, I would like to thank my research supervisor Miloš Ulman, Ph.D. for providing guidance and feedback throughout this project and my research advisor Akshay Pottathil, Ph.D. for encouragement and useful advice during this study.

I would also like to thank Darvis company, especially its Founder and CEO Jan-Philipp Mohr for providing the materials and useful advice regarding analysis. Additional thanks to Czech University of Life Sciences for providing a laboratory and needed tools for the research.

And finally, I would like to thank all my close friends. You have all encouraged and believed in me. You have all helped me to focus on what has been a hugely rewarding and enriching process

Thank you all for your unwavering support.

Application of artificial intelligence in crisis management

Abstract

This study was done for understanding how artificial intelligence can be used for crisis management and security of public places. The research topic was chosen because of the relevance of smart security solutions installed in public spaces nowadays. The theoretical part is based on the study literature about artificial intelligence, crisis management, smart surveillance crisis prediction and intelligent security systems. The practical part is an experimental research study which was made to compare artificial computer intelligence with human intelligence in terms of time and accuracy needed for object recognition and tracking. Research was made on 10 participants who were watching 5 videos with Tobii screen-based eye tracking device. Data for study was collected and analyzed in Tobii Pro Studio, Excel and IBM SPSS Statistical tool Obtained results from the study have demonstrated that people are faster in single object recognition and tracking single object over the time and that AI is better in multiple objects tracking at the same time, since people are switching their attention and gaze from one object to another. This study showed that currently AI cannot fully replace people in crisis management and that for the most efficient work it is better to combine people and AI.

Keywords: artificial intelligence, crisis management, security, analysis, object recognition, threat detection.

Table of content

List of pictures

List of abbreviations

AI – Artificial Intelligence

AOI – Area of Interest

CCTV - Closed-circuit television

CNN – Convolutional Neural Networks

YOLO – You Only Look Once

1 Introduction

Crises are happening every day. They include natural and man-made disasters, terrorist attacks, virus spreading, committed crime and others. Crises may differ in reasons, affected areas and duration, but their consequences are usually common. They lead to loss of people and animal lives, damage of the facilities and destruction of nature. To minimize the residuals of such negative events crisis management is used.

Crisis management is a complex system which includes possible threats detection, continuous situation monitoring and quick decision making if some event has already taken place. It refers to protecting and saving as many lives and property as it is possible during some crisis. Crisis management starts with identifying possible threats and creating detailed plans how to deal with them. Based on data collected in previous periods, the system tries to predict the probability of some event occurring in the future, the best- and worst-case outcome and the damage that it is potentially causing.

Crisis management systems benefit from the digital revolution. Smart security systems use digital infrastructure to predict and prevent security issues and to reduce localized crime that today requires a lot of human resources to patrol and respond to incidents. Artificial intelligence (AI) behind such systems helps to make a decision and take an action quickly based on real time information. AI based security systems are saving thousands of lives every day.

In this thesis we will make an overview of technologies used for crisis management nowadays. Current research starts with literature review where we will speak about algorithms used for object recognition; common crisis management techniques; people privacy and threats while using smart surveillance. In one of the sections, we will speak about crisis prediction techniques and how AI can help in that. We will also explain which software solutions are used for security of public places nowadays and how they work.

In the practical part of the research, we will compare artificial computer intelligence with human intelligence in terms of time and accuracy needed for object recognition and tracking. We will discuss two research questions which say that people are faster in single object recognition and tracking it over time and that AI is better in multiple objects tracking at the same time, since people are switching from one object to another.

2 Objectives and Methodology

2.1 Objectives

The main objective of this thesis is to compare performance of object recognition by AI and humans during a crisis situation. It is necessary to make an overview of common problems of AI and how those problems affect overall performance of object recognition in a crisis scenario. Another goal is to make an overview of existing software that can be used in crisis management. Time, accuracy and the ability to track an object by AI and humans will be compared by conducting a user study.

2.2 Methodology

To achieve the set objectives, literature review on the applications of artificial intelligence suitable for tackling crisis situations in public spaces will be done. A summary of crisis management techniques used nowadays will also be done. As one of the goals of this thesis is an overview of existing software, some of existing most promising systems will be studied. For the practical part of the research AI and human intelligence will be analyzed in a user study in a usability lab. The performance of the human participants will be measured, statistically analyzed and contrasted with the performance of AI-based object recognition. The results of the analysis will be interpreted, and conclusions will be formulated.

3 Literature Review

3.1 Artificial Intelligence

Intelligence in common sense is the ability to learn knowledge and apply it on practice. In computer science, artificial intelligence, sometimes called machine intelligence, it is the ability of a digital computers or computer-controlled robots to perform tasks commonly associated with intelligent beings. Artificial intelligence makes computers to imitate human way of thinking through machine learning algorithms. Due to human nature people have several limitations which machine learning tools do not possess. Among them are number of tasks that can be monitored at the same time, duration of attentive concentration on a specific event, response time if some suspicious event was detected.

Artificial intelligence (AI) was founded as an academic discipline in 1955, after it has experienced several waves of optimism in the beginning, after was disappointment and the loss of funding, known as an "AI winter", followed by new approaches, success and renewed funding. The first AI research, dating back to the 1950s, focused on problem solving and the development of symbolic computing systems. In the 60s, this direction attracted the interest of the US Department of Defense: the US military began to train computers to simulate human mental activity. For example, the US Department of Defense Advanced Research Projects Agency (DARPA) carried out several projects in the 1970s to create virtual street maps. And DARPA was able to create intelligent personal assistants in 2003, long before Siri, Alexa, and Cortana came along. These works became the basis for the principles of automation and formal logic of reasoning, which are used in modern computers, in particular, in decision support systems and smart search engines designed to complement and enhance human capabilities. (Norvig 2003)

AI allows computers to learn from their own experience, adapt given parameters and perform tasks that were previously only possible for human intelligence. In most AI implementations, from computer chess players to self-driving cars, deep learning and natural language processing is critical. With these technologies, computers can be “taught” to perform specific tasks by processing large amounts of data and identifying patterns in them.

The traditional problems solved by AI include reasoning, knowledge representation, planning, learning, natural language processing, object recognition, building trajectory of moving item, perception and the ability to move and manipulate objects. This list is constantly updated with a new task. AI approaches include mathematical models, statistical methods, computational intelligence, and traditional symbolic AI. Many tools are used in AI, including search and mathematical optimization, neural networks, and methods based on statistics, probability and economics. The AI field based on computer science, information engineering, mathematics, geometry, psychology, linguistics, sociology, philosophy, and many other fields.

Due to current development of information technology, Artificial intelligence has become an independent discipline. It involves elements from computer science, mathematics, cybernetics, information theory, algorithms, logic, neurophysiology, linguistics and many other subjects. AI perform data analysis to discover important patterns in real time. It has many practical applications in various fields like healthcare, finances, education, transportation and security. The traditional problems of AI research are natural language processing and understanding, problem solving, logical reasoning, object identification, object recognition, object tracking, automatic programming and robotics.

Artificial intelligence is usually divided into two categories: weak and strong AI. Weak AI is used for one particular task, for example voice to text transformation. Strong AI is used for complex tasks, where human-like logic is needed, it can be foe example self-driving car or intelligent security system.

AI allows to automate repetitive learning and searching processes through the use of data. However, AI is different from robotization, which is based on the use of hardware. The goal of AI is not to automate manual labor, but to reliably and continuously perform numerous large-scale computerized tasks. This kind of automation requires human intervention to initialize the system and ask questions correctly.

Generally, AI technology is not implemented as a standalone application. AI functionality is being integrated into existing products, allowing you to improve them, just like Siri was added to Apple's next generation of devices. Automation, communication platforms, bots and smart computers, combined with large amounts of data, can improve a variety of technologies used in homes and offices, from security data analysis systems to investment analysis tools. So, AI makes existing products intelligent and makes peoples everyday life easier.

AI adapts through progressive learning algorithms so that further programming is done based on data. AI discovers structures and patterns in data that allow an algorithm to learn a certain skill: the algorithm becomes a classifier or predictor. Thus, by the same principle by which the algorithm masters the game of chess, it can learn to offer suitable products online. At the same time, the models adapt as new data becomes available. Backpropagation is a technique that corrects the model by learning from new data if the original answer turns out to be wrong.

Also, AI performs deeper analysis of large amounts of data using neural networks with many hidden layers. A few years ago, creating a fraud detection system with five hidden layers was nearly impossible. That all changed with the tremendous growth in computing power and the advent of big data. Deep learning models require a huge amount of data, since this is the basis on which they are trained. Therefore, the more data, the more accurate the model.

Deep neural networks enable AI to achieve unprecedented levels of precision. For example, working with Alexa, Google Search and Google Photos is deep learning, and the more we use these tools, the more effective they become. In the healthcare field, the diagnosis of cancerous tumors on MRI images using AI technologies (deep learning, image classification, object recognition) is not inferior in accuracy to the conclusions of highly qualified radiologists.

AI allows to get the most out of your data. With the advent of self-learning algorithms, the data itself becomes an object of intellectual property. The data contains the answers you need - you just need to find them using AI technologies. As data plays a much more important role now than ever before, it can provide a competitive advantage. When using the same technology in a competitive environment, the one with the most accurate data wins.

The main limitation of AI is that learning is only possible from data. This means that any inaccuracies in the data will affect the results. And new levels of forecasting or analysis must be added separately.

Modern AI systems are geared towards performing well-defined tasks. These systems are characterized by a very narrow specialization. They are designed to perform one specific task, and they are far from human multitasking. For example system configured to detect fraud will not be able to drive a car or provide legal assistance. Moreover, an AI system designed to detect healthcare fraud will not be able to detect tax fraud or warranty claims with the same degree of accuracy.

In addition, self-learning systems are not self-contained. However, computers can analyze complex data to learn and improve specific skills are no longer uncommon. During crises the abilities of people are not enough to effectively handle the situation without computer programs. With current technologies huge amount of information can be easily received in a few seconds, but without smart analysis this information is useless. Artificial intelligence algorithms are effectively used nowadays for data analysis and future events prediction. When crisis happens, it is necessary to receive spatial information about the event in real time where exactly it occurred, and territory involved. For solving this AI also can be useful.(Ivić 2019)

3.1.1 **People recognition**

AI algorithms help in monitoring people movement and people identification using measurable biological characteristics. Among those characteristics are fingerprints, face scan, eye recognition, palm vein detection, voice identification and others. Some characteristics can be verified quickly and cheap like face recognition, but they are less reliable than expensive and quite long DNA or iris verification. Biological characteristics remain the same during long period of time, so they can ensure that the asset is used by true owner and individual is what he or she claims to be.

Fingerprint recognition allows a person to be identified through the analysis and comparison of his finger dermal ridges. Fingerprint recognition was one of the first techniques used for automatically identifying people and today is still one of the most popular and effective biometric techniques. Fingerprint identification is quick and reliable method of human identification. It is suited for applications where a key, access card, or password is normally used. It is already in use in doors, tool-management systems, online services, fitness centers, cell phones and many others (Gu a Kanade 2015)

Facial recognition is one of the techniques used for people identification. It is the process of identifying or verifying the identity of a person using their face. It captures, analyzes and compares patterns based on the person's facial details. Face recognition allows devices to recognize users quickly and more reliable than using usernames and passwords which can be easily changed. Face recognition is definitely not so accurate as DNA analysis, but for most applications of this technology it is enough. The main advantage of it that it can be used on a distance even without people knowing that they were identified.(Gu a Kanade 2015)

Finger veins are hidden under the people’s skin where blood cells are flowing. In biometrics, the term vein does not entirely correspond to the terminology of medical science. Its patterns are used for authenticating the identity of a person, in which the approximately 0.3–1.0 mm thick vein is visible by near-infrared rays. It is one of the latest biometric technologies which is not widely used nowadays but is uniqueness and high accuracy was proven from medical and statistical point of view. So it has many possibilities to be widely used. (Gu a Kanade 2015)

Algorithms can calculate unique mathematical models of people's movements, for example detect the same gait of a person who changed clothes or was missed by one of the cameras.(Venture et al. 2014) Using object recognition techniques system also can track people's belongings and detect that one bag was carried by different people in same location on a specific time frame.(Ko 2008)

3.1.2 **Overview AI algorithms for object recognition**

Object recognition algorithms are a method of detecting objects in images or video. They're a popular field of research in computer vision and they are used in self-driving cars, facial recognition, disease detection and crisis prevention systems.

Image classification process involves predicting the class of one object in some image. Object localization refers to identifying the location of the objects in an image and drawing border box around their extent. Object recognition combines these two tasks and localizes and classifies one or more objects in an image.(Choudhury 2020)

There are three major computer vision tasks:

- Image Classification: Predict the type or class of an object in an image

  Input: image with a single object

  Output: class label
- Object Localization: Locate the presence of objects in an image and indicate their location with border box

  Input: image with one or more objects

  Output: one or more border boxes
- Object Recognition: Locate the presence of objects with a border box and types or classes of the located objects in an image.

  Input: image with one or more objects

  Output: one or more bounding boxes and a class label for each border box

There are a number of algorithms used for object recognition so we will make an overview of some recent top-performing deep learning models. There are two main algorithm families which are used for object recognition they are: Region-Based Convolutional Neural Networks (R-CNN) and You Only Look Once (YOLO)(Brownlee 2019).

Convolutional Neural Networks (CNNs) are the basic architecture through which an AI system recognizes objects in an image. The CNN use different filters to transform the image across convolutional layers. Using this transformed image result as a feature, the neural network will search for characteristics the image has in common with particular object.

During this trial and error training process, the network will begin to understand which characteristics are most essential to determining object from images. It does this by learning parameters to help it filter the data through each network layer. Through repetitions of this process, from input to output, the neural network can learn to distinguish different objects.

The number of parameters used to detect an object varies with the algorithm. Because there can be as many as millions or even tens of millions of parameters, it is often difficult for humans to understand exactly which characteristics a system uses to make assessments. This in turn can make understanding classification errors difficult, too.

The R-CNN family of methods refers to the R-CNN, which stand for "Regions with CNN Features" or "Region-Based Convolutional Neural Network," developed by Ross Girshick.(Girshick et al. 2012) This includes the techniques R-CNN, Fast R-CNN, and Faster-RCNN designed and demonstrated for object localization and object recognition.

Another popular family of object recognition models is referred to collectively as YOLO or "You Only Look Once," developed by Joseph Redmon (Redmon et al. 2015). The

R-CNN models may be generally more accurate, but YOLO family of models are faster than R-CNN, achieving object detection in real-time.

YOLO approach involves a single neural network trained end to end that takes an image as input and after predicts border boxes and class labels for each box directly. YOLO has a unified architecture which is extremely fast The technique offers lower predictive accuracy but operates at 45 frames per second and up to 155 frames per second for a speed-optimized version of the model. (Redmon et al. 2015)

The model works by first splitting the input image into a grid of cells, where each cell is responsible for predicting a border box if the center of the box falls within it. Each grid cell predicts a border box involving the x, y coordinate and the width and height and the confidence. A class prediction is also works based on each cell separately.

3.1.3 **Social media**

People use different social media all day. They post photos to Instagram, stream videos on YouTube and share their thoughts on Twitter. During the crisis shared posts can be helpful. According to (Saroj & Pal, 2020) often the first-hand experience first comes in the social media before any other media reaches and reports from the spot. Of course, posted information can be fake or a peaceful strike can be classified as a mass riot based on some posts, but due to the large number of people sharing the same event those errors can be minimized.

Based on geolocation of the posts it is possible to detect the affected region and calculate the approximate number of people involved. This information can be used by authorities to send needed numbers of people and equipment. From real-time videos it is possible to understand in which areas it is safe to be and evaluate situations from different angles. Authorities can also understand society's feelings about some events and improve the situation.

For example, the accidental killing of black American man George Floyd in Minneapolis, USA in May 2020 by police officer Darek Chauvin was recorded by security cameras. (BBC 2020) The next day this video became public and was shared by thousands of people in their social media. It leads to mass protests all around the world. People organized a movement under the motto Black Lives Matters against police brutality and racism in the United States. Over 2000 cities and 60 countries supported this movement. Some of the protests were peaceful, but others tend into riots, looting and clashes with the police. Over 14000 people were arrested by the end of June 2020. This incident helped the federal authorities to understand people’s opinion about racism and police misconduct and initiate several legislative proposals on different levels.

We see that people react to events quickly and instantly, so social media posts analysis can be a helpful tool in crisis management.

3.2 **Crisis management**

Crisis is any event that is lead to an unstable and dangerous situation affecting an individual, group, community, or whole society. Crises usually have negative consequences like people's lives loss and damage of the facilities. There are three common elements for every crisis: a threat to people, facilities or nature, surprise and rapid development and a short time for decision making.(Bundy et al. 2017) Regardless of the source of crisis, a natural or human made one, time of happening, sudden or stable any crisis has several stages. Simply we can divide them to before, during and after.

Crisis management is the strategy of dealing with crises at all stages and on different levels. It includes the early detection of possible threats. Techniques which should be applied during the crisis and methods used if something has already happened. Due to rapid and unexpected development crisis management requires quick decision making to react on the situation immediately. The main goal of crisis management is to prevent the effects which are negative for people. (Bundy et al. 2017)

In order to reduce uncertainty in the event of a crisis, organizations and public authorities need to create a crisis management plan with different potential threats. This plan usually starts with risk analysis. Risk analysis is the process of identifying events that may occur and the consequences of them. Based on data collected in previous periods or on special random sets risk managers are creating risk models and scenario tables. They try to predict the probability of a possible risk occurring in the future, the best- and worst-case outcome of any negative event, and the damage that it is potentially causing. After understanding terms of probable risks and its impact risk manager develops a plan which contains an emergency if it will happen.

Crisis management is not the same thing as risk management. Unlike risk management, which involves planning for events that might occur in the future, crisis management also involves reacting to negative events during and after they have occurred.

The principal stages of managing crisis and risk situations are the following: providing information for the management process; setting up mechanisms for early detection of signals and warnings about the crisis; prognosis of potential catastrophes and possible consequences from them; planning of crisis scenarios and crisis management; preparation of specialists, teams and technological means for crisis management; strategic planning of restorative activities.

Prevention and early warning for crises can be done through different methods. For example, by constantly providing actual information about the current situation; analysis and evaluation of changes and deviations; setting up intelligent systems for early detection and warning about crisis; identification and reduction of expected risks.

For management of crisis situations can be used: prognosis of potential events and the consequences of them; development of possible scenarios for different crisis situations; planning of responsibilities and resources in advance; solving setting up and training of crisis teams,

Strategic planning of restorative activities includes define the main functions done before; selecting of resources necessary for the fulfilment of crucial functions and priorities; development of strategy for restorative activities; checking and controlling when functions were restored.

3.2.1 **Public safety**

According to the research, by 2030 nearly 5 billion of people will live in cities. (Seto et al. 2012) Every year the urban population grows. As cities become more attractive places to live and work there are many consequences and threats. People become dependent on city infrastructure and technologies need to provide care about them.

Security for cities is a top priority. Modern cities require protection 24 hours a day and 365 days a year. Video surveillance is an effective way to protect the city. It allows them to receive valuable information from public places, roads and buildings about incidents and response to them as quickly as possible.

People constantly interact with city infrastructure using smart devices. It helps to understand what is happening around. Data collected from citizens, devices and assets is processed and analyzed to manage whole city infrastructure from traffic control to early crime prediction. Smart cities allow us to constantly monitor the situation in unstable regions and ensure that events can be identified in real time and predicted before they occur and cause damage.

Safety is closely linked with economic development and as cities seek to drive economic growth and provide a good quality of life for their citizens, security is at the forefront of city planning. (Batty et al. 2012)

3.2.2 Natural disasters

Natural disasters like earthquakes, flooding, extreme temperature weather, bushfires and droughts kill thousands of people every year. Digital detection technology can be used for improving peoples and facilities safety in a case of disaster. AI based systems can monitor the situation and send an early alarm via SMS or email. Good example of SMS alerting is provided in a research for rural areas of Malaysia.(Ayobami a Rabi'u 2012) Access to disaster state and development can be provided in real time via internet to inform airline industries, energy suppliers and other enterprises of upcoming event and allow them to shut down sensitive facilities. Alerts also can help to evacuate people from dangerous areas. Disaster management refers to protecting and saving as many lives and property as it is possible during the occurrence of natural or man-made disasters.

3.2.3 Human made disasters

Human made disasters like atomic explosion, virus spreading, stampedes, fires, transport accidents, oil spills, terrorist attacks and war are the result of human activity or inactivity. Like any disaster they affect people, infrastructure and nature, but usually they are under human control and can be avoided or identified more easily than natural ones. Societal hazards like criminality, civil disorders, terrorism and war usually solved by police or special authorities. Industrial and technological hazards are solved using the machines and technological approaches. Environmental hazards affect nature and ecosystems, their consequences are often irreparable. (CRED 2015)

3.2.4 Crisis management techniques

To use people's posts in social media, like Facebook, Instagram and Twitter to take measures and minimize loss when some emergency occurs. (Saroj a Pal 2020a) The Ebola crisis and other recent major news events have all confirmed that social media is one of the most important channels of communications. During crisises people use social media more actively. Be sure to establish a social media team to monitor, post and react to social media activity throughout the crisis. (Hossain et al. 2016)

To use pre-installed apps which are tracking people, without collecting any personal information according to GDPR. AI based programs can check individual contacts using Bluetooth and NFC and recent places using GPS and Glonass. These applications might be helpful in situations like virus spreading, when COVID-19 pandemic started in Wuhan city Hubei province, China in December 2019 swept the whole world and was detected in more than 180 countries by April 2020. (WHO 2020) It was widely agreed that to slow virus expansion every country needs to impose social distancing and isolate people who were in contact with confirmed cases. But based on bank transactions information and cell phone location detection it was found out that not all people are taking responsibility and comply with the measures. That's why many countries decided to introduce and use legal surveillance. For example, Israel intends to use the location data of mobile phones, which help the intelligence service find terrorists in order to track the movements of people in contact with carriers of the virus. (Verge 2020) Singapore carefully monitors the movements of its citizens and publishes detailed information for each person, except perhaps a person's name.(CNBC 2020) As it was mentioned with the Israel example the same system also can be used for detecting criminal connections including terrorists, drug dealers, human traffic and many others. It can help not only to solve the crime but to find a whole chain of people involved and stop the whole criminal group at once.

AI based intelligent systems will be useful in crowded places like airports when before3.2check in for a flight, they will have to provide data from a mobile application that tracks their movements. The airline will not find out where exactly person was but will receive a warning if he contacted the carriers of the virus, visited the epidemic area or suspect to be a part of a crime chain.

To detect a number of cell phones in a disaster area using data from mobile operators. Usually disasters are happening fast and hard to predict. In order to respond quickly and effectively when it happened rescue services must know the scale of the affected area, including the approximate number of people involved. In the modern world almost, everybody is having a cell phone connected to some mobile operator. To transmit signal and data phones use antennas and when the phone user moves from one cell area to another, the system automatically detects it and commands the mobile phone and a cell tower with a stronger signal. For each place a number of cell areas and connected devices can be calculated. Received number will have a significant error. To make calculation more accurate this count should be repeated several times with short time delay. Some people might quickly leave the affected area by themselves without any help and some other can come to help with the situation, those ones do not need to be counted. Device duplication should be deleted but marked, because if a number of phones were detected by several towers in different cell areas near the disaster very probably these people were involved.

Uncrewed aerial vehicles or drones can be an effective tool for disaster management. The traditional method for surveying disaster affected areas is to use helicopters, but drones can entirely replace them. They can provide high resolution real-time videos of even inaccessible locations. Figure 1 shows a scene from drone recorded video. With intensive R&D in drone technology, drones can now offer a 3D and 360-degree view of locations. These views can then be used to create accurate disaster maps so measures for reducing disaster consequences and risks are planned accordingly. Drones can deliver food and water to the victims and needed equipment for rescue teams. One of the most breakthrough innovations in drone technology can be to make drones capable of airlifting the disaster victims from affected locations. This can help save a huge number of lives and eventually lead to extremely successful disaster rescue operations.(Allerin 2019)

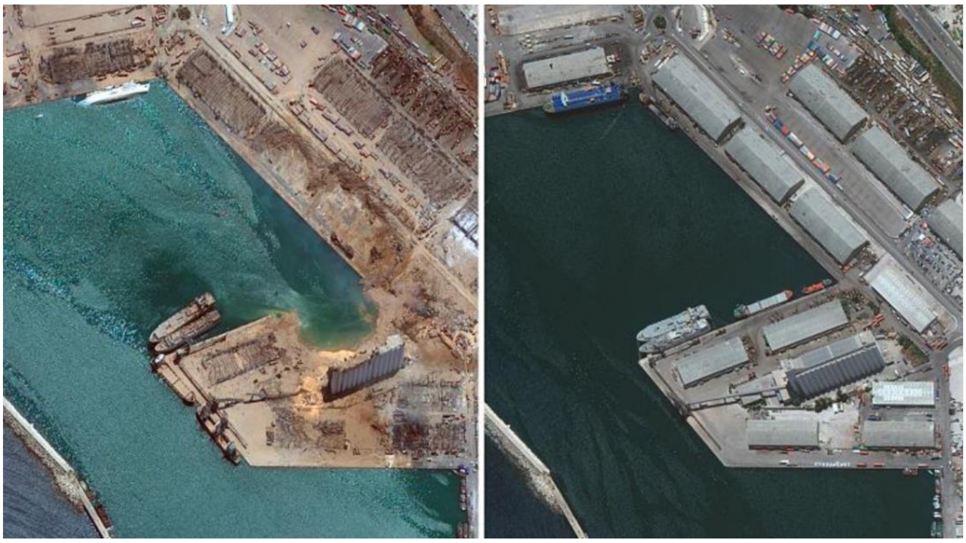

Figure 1 Before and after photos from Beirut's port reveal scale of explosion. Source: (Euronews 2020)

3.3 Smart surveillance

Video analysis is very important for the security of any public place. Smart surveillance systems nowadays use automatic video analysis technologies in real-time recording footage. With the help if AI they can also isolate individuals or objects, track them and detect suspicious patterns.

According to research about social group theory (Hernando et al. 2010), a number of people with whom one person can effectively communicate and have some relation varies from 100 to 230, but the average number is 150, it is called Dunbar number. In a group of those people we usually consist of friends, relatives, colleagues, previous classmates and neighbors and we think about them. Other people who are outside this group are not concerned as a people inside our mind, they are more like machines or a collective image of some group. Let's imagine a friendly cleaner who works in your office building. He is doing a great job allowing you to work in a nice office without full trash bins and dust on the shelves. But when you are going to work every day you don't think about him, you are more worried about tasks which you need to do today or about a meeting with your friends later in the evening. For you he is just a machine-to-clean-the-office not more. If you stop and think about him for a few minutes you will understand that he is also a valuable person for somebody, he also has friends, hobbies and food preferences. And that is why most people do not think about doing some bad things like stealing or murder to people from their close group, but can do it to others, who are outside.

The conclusion of this is, that in a big crowd of people one person becomes invisible. Now we live in a world which is much more complex than small groups of 150 people, we constantly interact with thousands of people every day. For such complex modern life, we need complex solutions in every sphere. This work focuses mainly on security and new approaches for modern life. Traditional surveillance systems based on inputs of hundreds of cameras and sensors, visualized on a single screen were monitored by one person are ineffective nowadays.

Traditional CCTV (closed-circuit television) cameras have been used for decades. Firstly, they were quite expensive low-resolution and installed only on important points. Businesses or city authorities deployed them to film a small area of interest. Few cameras were placed in public, and the power to track people was limited: If police wanted to pursue a person of interest, they had to spend hours collecting footage by foot from nearby locations. CCTV signal from cameras was transferred through security channels to special rooms with a number of monitors and was manually viewed by people. The main problem was that one person can attentively view only one screen, but there were much more of them. In a case of manual control of CCTV screens, many areas remain unmonitored and can be far more easily infiltrated with possible catastrophic consequences. Most of the crimes were not discovered in time, so video recordings were analyzed after the crime had happened. All work was done manually by combining videos from multiple cameras and finding needed attributes. It took a lot of time and human resources but helped to find criminals and arrest them.

After manual screen control came automated systems. They made suspicious situation analyzing easier by extracting photos with better resolution, combining videos from different areas and storing data in databases. It was intermediate step to intelligent systems, because it still required a lot of time and result accuracy was quite low.

With current development of AI technologies video surveillance reached a completely new level. Automated surveillance became an intelligent one. Now CCTV enables not only to view and record events but analyze them in real time and even predict future ones. Video analytics automatically captures a wide variety of data about the areas covered by CCTV cameras. The captured information can be used by AI to recognize crime patterns in real time videos and transfer this information to police.

CCTV became a powerful tool for many types of crime and crises. Currently AI tries to find anomalies in real time and stop a crime before it happens. But even if some event had already happened AI helps emergency services to send needed equipment and people to efficiently solve the incident.

While CCTV surveillance began as a simple tool for criminal justice, it has grown into a multibillion-dollar industry that covers multiple industry verticals. From policing and smart cities to schools, health care facilities, and retail, society is moving toward near-complete visual surveillance of commercial and urban spaces.

3.3.1 Threats: privacy, discrimination

The rise of smart camera networks threatens civil rights and liberties of people all around the world. For example, smart video surveillance can use facial recognition software to identify people and track their public move. According to EU GDPR and GDPR-like laws, data for people identification is personal one and should be protected. Many people feel uncomfortable because of constantly being watched. More cameras installed around us, more privacy we lose. For security reasons can be good, if something will happen, police has an opportunity to solve the situation quicker, but for privacy it is not, some people do not accept their tracking and recognition and they have an official right for that.

Another threat is that security systems can inform city authorities whether people are walking, running, riding a bike, or doing anything “suspicious.” With video analytics, artificial intelligence smart surveillance can identify people sex, age, and type of clothes, and could potentially be used to discriminate people by race, sex or religion. For example, crime prediction programs consider black people to be more likely to do some crime than white people. Or programs consider places where people with low income live to be more dangerous than places of living middle class families. These conclusions, which some people can find discriminative, are made based on previously collected data and calculated statistics, so it can be not a discrimination, but behavior pattern.

In 1998, US courts began using the Compas program, which, based on 137 information sources, should predict the likelihood of recidivism. With this program, judges decide who to extend the sentence to, who can be released or released on bail, and for what amount. The system has been used more than a million times, but it causes a lot of complaints. Researchers from Dartmouth College recently conducted an experiment that proved that people still predict relapse of criminals more accurately than algorithms. For the experiment, they hired 462 people without a law degree, provided them with information about the criminal past of criminals and asked them to guess if they would commit a second crime. It turned out that the predictions of randomly selected people were correct in 67% of cases, and the predictions of Compas in 65%. It was also proved that, despite the same criminal history of white and black criminals, Compas is twice as likely to assume that an African American will become a recidivist. (Dressel a Farid 2018)

The issue of racial bias is one of the most obvious and frequently discussed flaws in crime prediction systems. In the United States, such systems are especially suspicious of African Americans, in Israel to Palestinians, and in China to Uyghurs. Once on the list of suspicious persons, a person can no longer get rid of surveillance and suspicion from the system. If whole nations and ethnic groups can become suspects, it is logical to assume that further the attention of machines can be drawn to representatives of entire economic classes, residents of certain regions and even cities, and ultimately has the potential to turn into a weapon of discrimination on any convenient basis.

3.3.2 **EU GDPR**

The General Data Protection Regulation (GDPR) is a regulation in EU law on data protection and privacy in the European Union (EU) and the European Economic Area (EEA). It also addresses the transfer of personal data outside the EU and EEA areas. The GDPR's primary aim is to give control to individuals over their personal data and to simplify the regulatory environment for international business by unifying the regulation within the EU. (Council of the European Union, 2015)

Images are considered personal data and are protected under the GDPR. The footage or a picture of an individual captured through the video surveillance, which may be used to identify that person (directly or indirectly) is considered to be personal data as well, and in those cases, GDPR requirements for personal data processing need to be put in place. It is important that pictures and video footage is personal data, but not biometric one. Because biometric data is considered as sensitive personal data and processing of sensitive data is restricted. Processing of biometric data is prohibited unless the data subject has given explicit consent, or there are special circumstances allowing the processing.

For video surveillance to be legal, it needs to be based on one of the 6 lawful bases for processing personal data: consent, contract, legal obligation, protection of vital interests, public task or legitimate interests. Video surveillance in public places has to comply with the transparency principle and provide information about the surveillance with contact information about the data controller and reason for surveillance. The notice with information should be easily visible, with the appropriate camera symbol informing everyone entering the area about video surveillance.

People have a right to obtain information from the data controller about whether their data was processed, access to the personal data and the information described in Article 15 of the GDPR- Right of access: what is the purpose of the processing, what are the categories of processed personal data (including recipients or categories of recipients in third countries or international organizations), who are the recipients to whom the personal data have been or will be disclosed and for what period of time will data be stored. When providing that information to somebody, the controller should take all necessary measures to protect the identity of other people on the footage .Important notice is that since the footage is stored for a limited amount of time, this will affect the ability of people to access their footage.

The video footage should not be kept for longer than it is strictly necessary for the purpose that wants to be achieved. In practice, the footage material is usually retained for a short amount of time and deleted afterwards. Taking into consideration the data minimization and storage limitation principles, the personal data should in most cases be deleted automatically, after a few days. (Barnoviciu et al. 2019)

3.3.3 USA CCPA and BIPA

Data privacy legislation is appearing in more and more economies across the globe, USA is inside this list. In June 2018 California became one of the first U.S. states who introduced privacy regulation act CCPA (California Consumer Privacy Act) and previously in October 2008 BIPA (Biometric Information Privacy Act) was created. Those are two US laws that put restrictions on what type of consumer information a company can collect. They have many similarities to the EU's GDPR regulation.

In simple terms, according to CCPA (California a Material 2020) companies must notify users if they want to collect their data and also give them the right to reject that. The basic intentions of the act for consumers are: know what personal data is being collected about them; know whether their personal data is sold or disclosed and to whom; say no to the sale of personal data; access their personal data; request a business to delete any personal data; not be discriminated against for exercising their privacy rights. It includes the ability to give a full report on all data that system has about the person who has requested the information. It also gives people the ability to delete all information about them. (Baik 2020)

BIPA applies to companies that collect and store biometric information (including fingerprints, voiceprints, and scans of the hand or face geometry). The requirements for those companies include obtaining consent from individuals if the company intends to collect or disclose their personal biometric identifiers; destroy biometric identifiers in a timely manner and securely store biometric identifiers.

3.4 Crisis prediction

Video analytics automatically captures a wide variety of data about the areas covered by smart camera networks. The information captured is now being proposed for predictive policing: the use of data to predict and police crime before it happens. Real time videos used to analyze crime patterns for police. A public safety intelligent system will centralize data pulled from disparate databases such as social media, personal documents, police databases and internet activity. An AI-enabled analytics unit will let police assess "anomalies in real time and interrupt a crime before it is committed."

To predict a crisis, we need several parameters: determine the place, time and type of future event, find the victims of it and specify the identities of the criminals.

In time of big data and powerful neural networks prediction of future events becomes a purely mathematical problem. (Qi a Majda 2020) Statistical analysis helps to calculate patterns of human behavior and their relationships, algorithms help to determine the likelihood of events based on previously collected data. Such predictions can help people to avoid future economic crises, diagnose and treat diseases that have not yet appeared and alarm about natural disasters before they happen.

The most advanced developments in the predictive potential of neural networks are aimed at security systems, helping the military to wage war more efficiently, and for police to prevent future crimes and catch unrealized criminals. By themselves, these methods of statistical analysis are not new: crime mapping, for example, has been used by police and security agencies for more than a century. But now the collection of information and its availability has become much easier and cheaper, and the processing of this data is no longer done by people, but completely by computer systems. Predictive policing promises to identify where and when future crimes will happen and stop them before happening using big data and mathematical algorithms. (Bigo a Negru 2014)

3.4.1 **Data for crisis prediction**

The basic rule for gathering information for crisis prediction is that the more the better. Several types of data can be collected about each of us: personal information (age, gender, race, nationality, religion), demographic data (address of residence and place of work, salary, position), information about our activity (from favorite vacation destinations to participating in political movements). Data received from SIM card, biometric passport, medical history, data from surveillance cameras, social media activity history, credit cards and digital payments are sources of information that can be legally collected without permission. The development of technology allows to collect more and more information from those “traces” that we leave without hesitation in the course of everyday life.

European laws on the personal data protection (GDPR) are the most loyal in terms of protecting the privacy of citizens, therefore, the European Union does not intensify surveillance, but improve the collection of data from open sources.

In countries with military conflicts, like Israel, or in more authoritarian regimes, like China), or just in countries where the balance in privacy matters is not in favor of citizens, like the USA, data collection on citizens is much larger.

The data obtained is processed using artificial intelligence using machine learning. AI classifies the collected data, simultaneously getting rid of information noise, develops rules and patterns by which the outcome of future situations will be determined, then compares the data and draws a conclusion about the probability of certain events. All calculations are based on mathematical models and theories. (Council of the European Union 2015)

3.4.2 **Crime prediction models**

3.4.2.1 Theory of Broken Windows

Broken windows theory is a criminological theory that says that visible signs of crime and antisocial behavior create an urban environment that encourages people for further and more serious crime. The broken windows are a metaphor for any visible sign of disorder in some area that wasn’t restored. The theory suggests that policing minor crimes such as vandalism, loitering, public drinking and graffiti will create an atmosphere lawfulness, thereby preventing more serious crimes like rape, robbery and murder. The theory was introduced in a 1982 article by social scientists James Q. Wilson and George L. Kelling. Broken window theory effectiveness has been repeatedly questioned by human rights defenders. It has become associated with zero tolerance form police side, but Bratton and

Keeling answered that this method requires careful training and supervision. (Belleflamm a Peitz 2017)

3.4.2.2 Hawkes Process model

One of the basic mathematical models used in predicting crime is called the Hawkes Process. The basis of this model is the idea of a close relationship of events: one event significantly increases the likelihood that similar events will soon follow, and over time their probability returns to its original value. Hawkes developed this theory by studying earthquakes in the 1970s and has since been used successfully to describe the work of the stock exchange, the spread of epidemics, electrical impulses in the brain, and even when describing the flow of emails within companies. The same model turned out to be effective for predicting the time and place of robberies and shootings between criminals and to this day it has been successfully used in machine prediction of crimes. (Mohler et al. 2019)

3.4.2.3 Theory of optimal foraging states

Another model which is used for crime prediction was borrowed by criminal experts from observations of the animal world. According to the theory of optimal foraging states: even though food provides animals with energy, its search and extraction requires both energy and time, therefore animals tend to choose strategies for finding food that require minimal cost. Simply it means that to feed themselves it is sometimes too expensive for a robber to travel between different areas, so after some time he chooses the same location. This combination of human natural instincts with statistics and machine learning allows the police to effectively prevent robberies and catch criminals.(Sorg et al. 2017)

3.4.3 **AI model training**

Nowadays computer games have feature high-end graphics, realistic locations and characters with whom we can interact. With variety of software tools and game engines computers can render complex worlds. Recently a team from the Computer Science and Artificial Intelligence Lab at Massachusetts Institute of Technology and Nvidia company demonstrated how it is possible to generate synthetic 3D gaming environment using a neural network that has been trained on real videos of cityscapes. (Wang et al. 2018)

According to research, team created a hybrid approach which involves using deep learning artificial intelligence, along with a traditional game engine, to generate visuals synthesized from video footage of the real thing. This process, called video-to-video synthesis, involves getting the AI model to “learn” how to best translate input source video into video output that looks as photorealistic as the original video content.

Research team based their approach on previous work Pix2Pix, an open-source image-to-image translation tool that uses neural networks. They utilized unsupervised deep learning algorithm called generative adversarial networks (GANs), which designates one neural network as a “generator” and another neural network as a “discriminator.” These two networks work together, the generator network aiming to produce a synthesized video that the discriminator network cannot ultimately determine as fake.

Figure 2 Training video

Figure 3 AI-rendered video

Training data was taken from video of driving sequences taken from autonomous vehicle research data in various cities. Objects on videos were segmented into various categories, such as buildings, cars, trees and so on. Real video example is showed on a Figure 2 Training video The GAN is then fed these data segments so that it can then synthesize a variety of fresh and different iterations of these objects as it shown on a Figure 3 above.

This technology can be used in opposite direction. Darvis company is training AI models based on videos, recorded from computer games Figure 4. Neural networks inside their AI-powered system than able to detect objects and differentiate them according to types. Example is shown on a Figure 5.

Figure 4 Training video

Figure 5 AI-based recognition

3.4.4 Future development

The next step after predicting future crimes should be a system to prevent them. One of example of using artificial intelligence to automatically prevent crime has recently been introduced in China. These electronic sensors using lasers and cameras recognize a person who is about to cross the road by the wrong rules, and spray water on him. (Reuters 2018)

Another example is developed by the US Department of Homeland Security it is called Future Attribute Screening Technology (FAST) system that monitors breathing, cardiovascular activity, eye movement, heat, and sudden body movements. This system works like a lie detector, but unlike it, it works from a distance to “remotely (and therefore more safely) recognize intentions to do harm”. It can be used in the near future for automatic systems to block of digital devices or deny access to public places if suspicious activity occurs on them. (Dhs et al. 2011)

Controlling thoughts and emotions can also play an important role in crime prevention. In 2014, Facebook published the results of an experiment that showed how through the control of a news feed you can control the positive and negative emotions of users. With the development of such algorithms, it will become possible to create a reality around users that minimizes the likelihood of their criminal inclinations being realized in the form of robberies, violence or unauthorized political protests. Maybe by reading brain waves, we will be able to transmit potentially unexpressed thoughts and moods between users. (Guardian 2014)

Among the negative consequences are the violation of the privacy of life, potential surveillance and a decrease in the security of user data. However, in the matter of predicting and preventing crime, such a technology can become a revolutionary means by which we will see the fight against thought crimes. (Saroj a Pal 2020b)

3.5 Intelligent Security Systems

With more than 350 million video security cameras worldwide the amount of video recorded is overwhelming. There are simply not enough human eyes to review and analyze them all. Intelligent security systems received additional attention because of the increasing demand on security and safety for public places. Intelligent security systems are able to automatically analyze recorded image, video, audio or other types of surveillance data with limited human intervention or even without it. The recent innovations in sensor technologies, computer vision, and machine learning have an important role in faster development of such intelligent systems. (Ibrahim 2016)

Video surveillance systems are an important necessity for today's daily life. In public places they can monitor the suspicious behavior of the people with computer algorithms based on mathematical models. Intelligent systems continuously monitor video streams, automatically analyze it and select useful information without any operator. With those systems threads can be effectively identified in a short period of time.

Many AI techniques used inside those systems. For example, object identification and tracking. System identifies not only static characteristics like color, shape and size, but also information about how the object got there and by whom was moved. Another widely used and important task for AI is people recognition. Some of the techniques are based on biometrical characteristics like face and eye recognition, palm vein detection or fingerprint scanner. The accuracy of those techniques is extremely high, but the main disadvantage is that people should be close to scanners and it's almost impossible to do it on distance. Other techniques like gait recognition, behavior and emotional patterns detection can be captured from video streams, but their accuracy is lower.

It is essential to intelligent security systems to be wireless and constantly send data to the cloud in order to access images and critical information when wired communication networks are damaged.

3.5.1 Darvis

Company called Darvis was established in early 2015 by Jan-Philipp Mohr, Ingo Nadler and Jan Schlüter in San Francisco. Darvis was founded as Hashplay (a virtual reality company), and in 2017 pivoted towards a computer vision company. (Darvis 2015)

Darvis has developed an intelligent system that addresses the issue of early detection. Darvis is an AI-powered real-world analytics platform allowing its customers to network the surveillance of multiple cameras, continually scanning what is seen by the cameras and sensors, and utilizing data analytics and deep learning to process, identify and track potential threats. Darvis combines data from multiple camera angles, anonymizes recognized humans in real-time, and detects patterns by analyzing various factors (e.g., timing, operation plans or activities like fighting and group building). Darvis does not need to initialize facial recognition if a specific attribute "Person with Gun in Hand" has not been detected. Darvis disregards video feeds immediately after analyzing them. No private data is stored. Consequently, Darvis is 100% GDPR and labor union compliant.

The current practice of placing cameras to protect an area and requiring a person to watch monitors is inefficient and ultimately ineffective. When weapons or threats are detected, Darvis will alert security personnel and law enforcement. Darvis will track and send exact location notifications via text or app to security. The app can stream the footage live so law enforcement will have the relevant information they need to respond appropriately.

Darvis algorithms are based on convolutional neural networks, discussed in 3.1.2 and computed directly on the devices. This approach works faster because there is no time losses while data transmission between the device and server.

After COVID-19 pandemic Darvis company changed their vision to create intelligent solutions for hospitals (Figure 6).

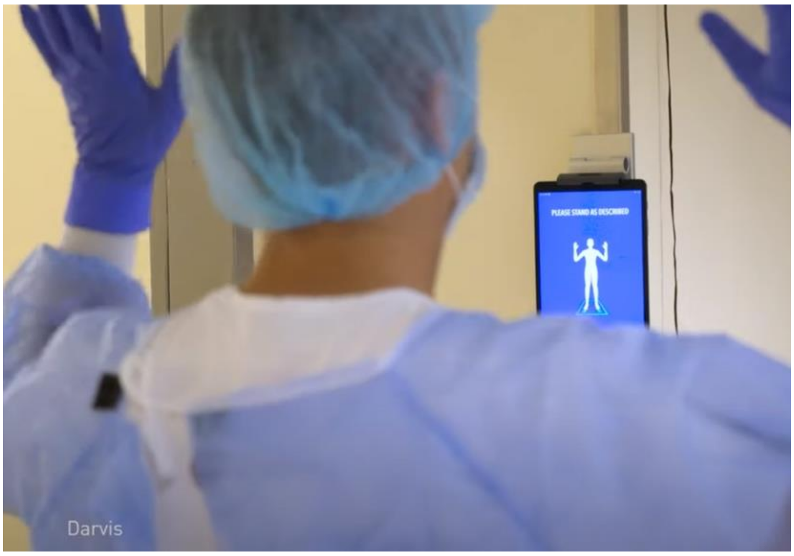

Figure 6 Darvis smart hospital solution. Source: (Darvis 2015)

3.5.2 **BriefCam**

BriefCam was founded in 2008 based on technology developed at the Hebrew University of Jerusalem by Prof. Shmuel Peleg, a renowned computer vision researcher.(BriefCam 2008) This proprietary technology, Video Synopsis, is an award-winning innovation that summarizes hours of events into a “brief” to extract maximum value from video for safety, security and productivity.

BriefCam is a complete video content analytics platform which drives exponential value from surveillance system investments by making video searchable, actionable and quantifiable. The unique fusion of Video Synopsis and Deep Learning solutions enable rapid video review and search, face recognition, real-time alerting and quantitative video analysis insights.

Through BriefCam's unique fusion of Deep Learning and VIDEO SYNOPSIS, organizations can now rapidly detect, track, identify, search and view thousands of objects in every scene, totally transforming how organizations use video (Figure 7 BriefCam video analysis). As the industry's leading provider of Video Content Analytics solutions, BriefCam enables customers to: review hours of video in minutes, rapidly search and identify people and objects of interest, dynamically analyze key performance indicators to optimize operations and receive real-time notifications of critical events The depth of BriefCam's video analytics and cross functional applications maximize the ROI of video surveillance and help transform how organizations operate.

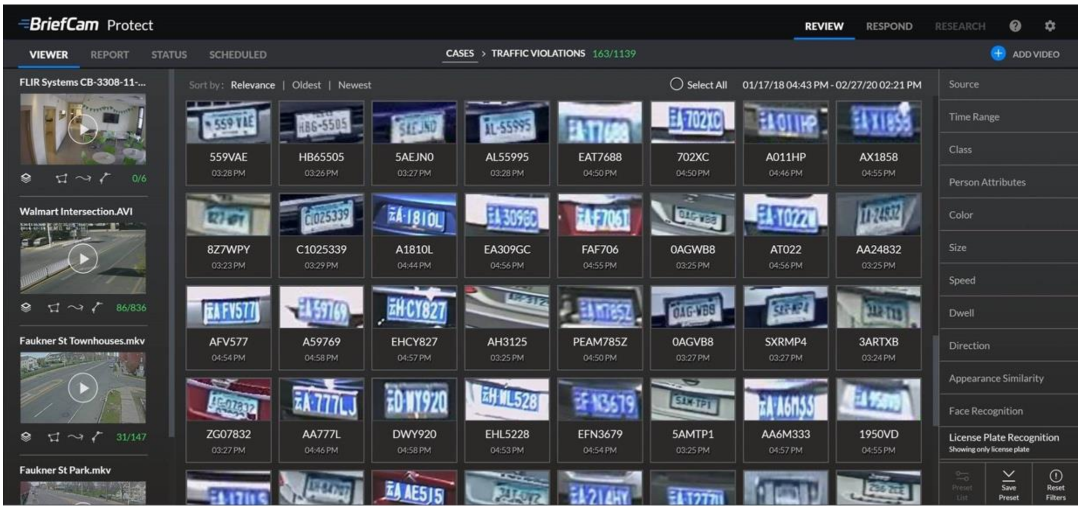

Figure 7 BriefCam video analysis

3.5.3 **PredPol**

One of the most reputable crime prediction programs is PredPol, short for Predictive Policing. (PredPol 2010) PredPol uses predictive analytics to support law enforcement. Firstly, PredPol was a research project between Police Department and University of California in Los Angeles. Bill Bratton was a chief of the police and he wanted to use data from CompStat (Compare Statistics) program for crime prediction. CompStat is program for collection and processing of statistical data. The goal of PredPol was to understand if collected data could provide any future recommendations about where and when some crimes could occur. PredPol team wanted to identify probable crime locations and their time which will allow to send police officers to exact location and help to prevent crimes.

Together with mathematicians and behavioral scientists from University of California and Santa Clara University, the PredPol team evaluated a wide variety of data types, behavioral and forecasting models. The models were further checked with crime analysts and officers from Los Angeles and the Santa Cruz Police Department. They pointed that the three most objective data points collected by police departments provided the most accurate input data for forecasting: crime type, location and datetime.

The current PredPol platform represents a significant investment of over 70 research-years of PhD-level analysis, modeling and development. It is used in many police departments around the world.

Everyday police officers update city maps on their tablets. The map has red squares indicating the time, place and type of crime that could possibly happen there. This square is called High hazard zones and usually they are about 50 square meters. (Figure 8) Hazards zones are noted by special algorithms that use a database of all crimes in the city for the past 11 years for predictions. Police officers must patrol these areas to either catch the criminals red-handed or to prevent the crime just by their presence.

Half of the crime in Seattle over a fourteen-year period could be isolated to only 4.5 percent of city streets. Similarly, researchers in Minneapolis, Minnesota found that 3.3 percent of street addresses and intersections in Minneapolis generated 50.4 percent of all dispatched police calls for service. Researchers in Boston found that only 8 percent of street segments accounted for 66 percent of all street robberies over a twenty-eight-year period.

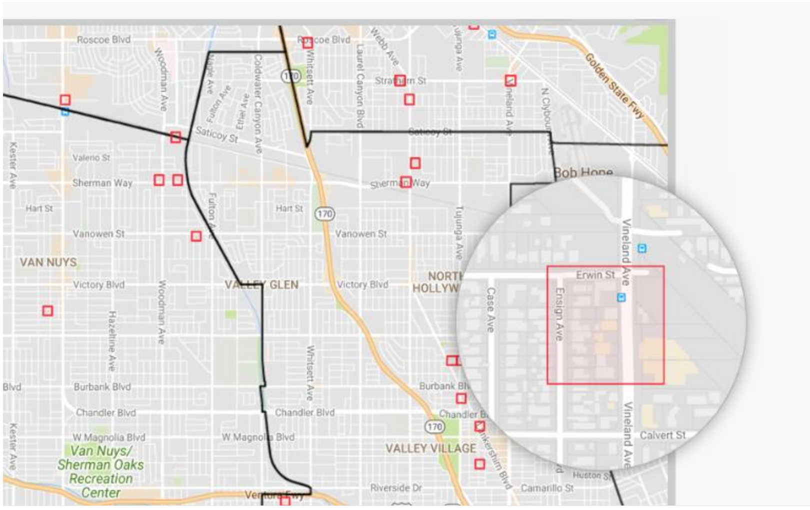

Figure 8 Predpol crime prediction area

3.5.4 **Palantir**

In 2009, at the initiative from Los Angeles became a testing ground for another giant in the field of crime prediction the company Palantir (Palantir 2009). Palantir creates software solutions for data-driven decisions and operations. The company was established in 2004. It was the private company Palantir, which at first specialized in cooperation with the US military and the CIA during military operations in Afghanistan and Iraq.

Palantir was not directly involved in intelligence activities and did not collect data - but analyzed them and arranged them in a user-friendly interface. With the help of Palantir, the military tracked down spies, terrorists, revealed secret locations and were able to find and kill Osama bin Laden.

After military trials, developers from Palantir began to apply their knowledge already in civilian life, including in Los Angeles. Their work is largely similar to the work of PredPol: basically, it is a prediction of the time and place of future crimes, but with additional features. For example, if the police stopped a car with a broken light and there was a former criminal in the passenger seat, this car instantly enters the Palantir database. If a crime occurs a year later and within the radius of 320 kilometers there is the same car that the sensors reading license plates noticed, then the cops are immediately informed that the chronic offender was not far from the crime scene.

Another Palantir program, Gotham, is used by the police in Operation Laser, designed to recognize and detain future criminals. Information from the protocols of detentions, police interrogations and criminal case materials is uploaded to a single database, which compiles a list of people considered to be repeat offenders and chronic troublemakers. This list is distributed to the police: if a person from him is nearby, then they should stop him and interrogate him under any pretext, even if he has not committed any wrongdoing. The colleagues of the Los Angeles police from Chicago went even further, where a list of more than 400 people who were more than 500 times more likely to be violent and commit crimes in the future was compiled. The list includes not only people who have already committed crimes: often those who have 2 or more acquaintances from this list also get there. The cops visit young people from this list and remind them that they can be rattled behind bars under the supervision and at the first suspicion. This turned to be an effective way for crime prevention.

Palantir systems are also equipped with a secondary surveillance network - a web of contacts of alleged criminals, which includes their family, friends, friends and sexual partners. This data is available to any person using the Palantir system, but not publicly disclosed. In addition, Los Angeles police buys personal data from private companies: information from social media, mortgages, travel information, video from surveillance cameras in hospitals, private parking lots, at universities, even information about pizza delivery to your home.

With a help of Palantir a large-scale police raid was organized 103 people were arrested in it. The basis for the arrest was two murders, as well as millions of analyzed Facebook posts on which the police managed to identify not only the perpetrators of the killings, but also their indirect accomplices, as well as people just about to commit crimes. Dozens of young people received 15 years of imprisonment for planning to commit crimes that they did not manage to carry out.

Figure 9 Palantir Gotham screenshot. Source: (Palantir 2015)

3.6 Research questions

Based on literature review made in chapter 3, it was decided to compare artificial intelligence object recognition techniques with human intelligence abilities during crisis events. For the research following research questions were formulated:

1. People are faster in single object recognition and tracking it over the time
2. AI is better in multiple objects tracking at the same time, since people are switching from one object to another.

Those questions will be addressed in the practical part by conducting an empirical study. The results will be then discussed and contrasted with findings in literature.

4 Practical Part

In current research we decided to compare human and artificial intelligence. Our assumption is that people quicker detect needed objects and recognize them even if video has low resolution but cannot track multiple objects in a same time. AI in opposite can analyze the whole scene, record it and create a footprint of all the objects. Those assumptions were successfully proved based on collected data

The practical part of the thesis is focused on comparing the speed and relative effectiveness of detecting objects on a video footage and tracking them by artificial intelligence and human eye. The study was made on a group of 10 people who were asked to watch 5 different videos and identify people and objects. The observed objects were guns carried by people in the videos. For tracking where human participants look, how quickly they notice the objects and how long they can concentrate on them a screen-based Tobii Eye Tracker X2-60 device was used. The aim of the practical part of the thesis is to collect and analyze empirical data in order to compare human and AI abilities in detecting objects and analyzing the situation.

This section starts with an introduction to the case study and justification why it was selected. After it provides a detailed explanation of methods used for data collection and explains the variables which were analyzed. Also, a step by step explanation of the study is presented. In the end of the section, the results are summarized and interpreted.

4.1 Design of the study

To study the visual attention of individuals in this research Tobii X2-60 screen-based eye tracker was used. This device consists of multiple cameras, illuminations and algorithms inside. The illuminators create a pattern of near-infrared light on the eyes; cameras take a high-resolution image of the user’s eye and the patterns; image processing algorithms find specific details in the user’s eyes and reflection patterns. Based on these details, the eyes position and gaze point are calculated, for instance on a computer monitor, using 3D eye model algorithm.

Video recording with eye tracking was made in a Tobii Pro Studio application. Tobii Pro Studio is an all-round eye tracking software that can be used in a wide range of eye tracking studies - from web user experience to common research paradigms in experimental psychology and market research. It covers complete workflows from study design to data analysis and export. (Tobii 2016)

For this research we took 5 different videos. Each video had a number of people and guns on them. Three of these videos were previously analyzed by AI, the other two were not. People were asked to watch videos and click on guns and people on a time of their first appearing. Click action time is considered to be the first moment when the object was noticed by a human participant. This action time was later compared with the time when the same object was firstly spotted by AI. This allowed for comparison between humans and AI.

First video was recorded from a computer game based on computer graphics. Video has a high resolution and a good quality. Objects on this video were labeled by AI with a probability of recognized object type. On this video (Figure 10) we have two women figures and two guns which are constantly moving, hiding behind other objects and changing their location.

Figure 10 Video 1 sample computer game video with AI tagged objects. Source: (Darvis 2015)

Second video was recorded indoors in a gun shop with daylight and high resolution (Figure 11). This video was not analysed by AI before, so people were detecting objects without any hints. On this video there are three people and three guns which they carry. People are constantly moving with a „shake" trajectory between the shelves. As there are many objects on shelves sometimes it is difficult to choose a proper object and track it. People always saccade their gaze from one object to another, which proves the second assumption of the research.

Figure 11 Video 2 sample gun shop video without tagged objects. Source: (Darvis 2015)

Third video was also recorded indoor in a meeting room with a daylight and high resolution. This video was analyzed and labeled by the Darvis AI before (Figure 12). In the third video, we have seven areas of interest: three men, three guns and a mobile phone, which was holded by one of the participants in the beginning. AI didnt make a mistake in recognizing a phone as a gun, but two out of ten participants clicked on it. On this video people were constantly moving while carrying a gun, so people needed to saccade their gaze between multiple objects.

Figure 12 Video 3 sample meeting room video with AI tagged objects. Source: (Darvis 2015)

Next, a fourth video was recorded outdoors under the daylight in a training location staged as a ruined city (Figure 13). People on a video are moving in a distance from the camera and all objects have shades. This video has high resolution and 6 areas of interest: three men and three guns. This video was analysed by AI in real time and labels on objects appeared gradually one by one. This video was chosen for testing the first assumption of practical part of the thesis. Which is that people are faster in single object recognition and tracking them over the time. This assumption was statistically proved during the research based on collected observations.

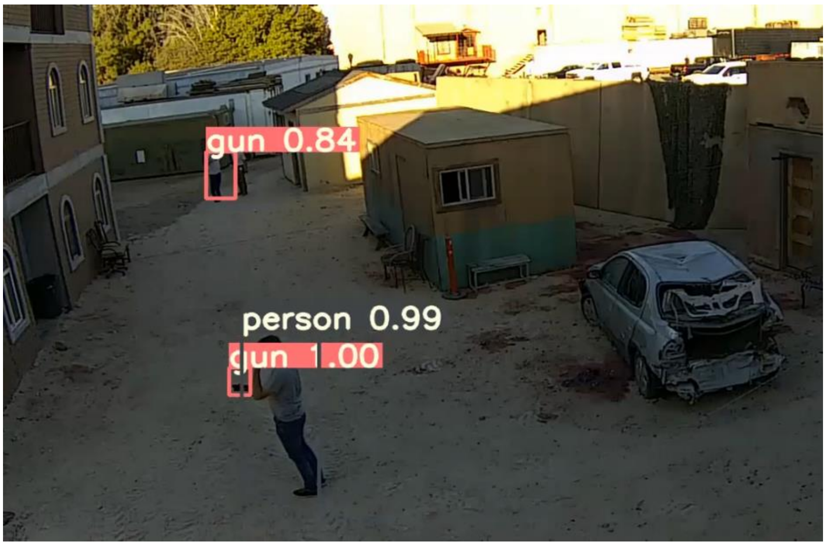

Figure 13 Video 4 sample staged video with AI tagged objects. Source: (Darvis 2015)

Last, the fifth video was a real crime scene, recorded in Cleveland, USA on 25 January 2017 (Figure 14). The video was taken from an open police source only for study purposes. Video was recorded during nighttime by video surveillance cameras installed on building nearby and it has a lower resolution than previous ones. In that video we marked three areas of interest: a car with shooters and two running men. One of them was carrying a gun, but because of the low resolution none of the participants didn't notice it. This video has multiple moving objects, so people need to switch their attention from one object to another.

Figure 14 Video 5 - Crime Scene Cleveland, USA Source: (Youtube 2017)

4.2 Step by step description of the study

Due to COVID-19 crisis, situation current research was made on a small group of 10 participants. This number of participants is considered to be minimal for statistical proof of the assumption of current research. All participants were young people aged between 24-34 years old having a university background. Two out of ten participants were women. Three participants were wearing glasses during the recording of the eye movements, which can cause a small inaccuracy. One of the participants had a large number of unidentified eyes movements, these results were not counted during the research. Data for all participants were anonymized and summarized. No personal data of any participant was used. An eye tracking screen based Tobii X2-60 device was used. All recordings and data preparation were done in Tobii Studio Pro application.

We took 10 people and showed them 5 different videos. Each video had a scene with people and guns and the participants had to click on each gun they noticed at least once. In further analysis people and guns were marked as areas of interest in the Tobii Studio program. Some of the videos were previously analyzed by AI and had labeled marks of the detected objects near to object type probability. Other videos were not analyzed and had no marks. One of the videos without AI tagging was a real crime scene recorded in Cleveland, USA on 25.01.2017.

People were asked to find all people and all guns on videos and click on them on the first appearance. Time of the click was considered to be a moment of noticing in current research. Participants were also trying to track these objects as long as they can. We made 5 separate recordings for each participant. Every recording starts with eye calibration, where participants had to follow the red dot on a screen. The dot was moving to different parts of a screen and it allows the eye tracker to calibrate and provide more accurate measuring results. After completing the calibration process, the tracking device recorded the movements of participants' eyes capturing all fixation points and saccades.

After the sessions with participants were done, we needed to define valuable time segments on each video for further analysis. Time segments are used for analytical purposes in Tobii Pro Studio. They allow carve out periods of time during the test recording over which meaningful behaviors and events take place. Time segments are set up in the Recording tab and allow to generate visualizations, draw AOIs and calculate gaze metrics.

Time segments are used for identifying and aggregating portions of a stimulus presentation for analysis. Creating a segment marks portions of existing recordings so no new media items generated. Segments will map gaze properly even if a web page is scrolled during the segment. And also, segments can be exported as video clips.(Tobii 2016)

Time segment marking had to be done manually in Tobii Studio. Only the data from marked segments was analyzed. We decided to take short scenes of 15-20 sec where the main action happens (e.g. a gun becomes visible or shooting happens). Next step was to find appropriate areas of interest in each video.

The area of interest (AOI) is a concept in Tobii Pro Studio that allows the analyst to calculate quantitative eye movement measures. These include fixation counts and durations. Using this tool, analyst needs to simply draw a boundary around a feature or element of the eye tracking stimulus. Area of Interest is a good and precise tool for setting up quantitative analyses of gaze behavior. They help to structure data analysis. Pro Studio then calculates the desired metrics within the boundary over the time interval of interest. Pro Studio creates AOI as separate from all others, even if they share the same name. For this research it was advantageous to aggregate several AOIs into one logical group we used Group feature for that. (Tobii 2016)

We marked all people and guns separately as areas of interest. On the third video we also marked a mobile phone to check the difference between the objects. After AOI defining we created groups for all areas of interest with the same tracked object. These groups were useful for statistical verification of our hypotheses and unification of obtained data.

Video preparation part was one of the most time-consuming parts of the research. Video segmentation and areas of interest labeling process took approximately 12 hours of work. After the video preparation, in the Data Export tab of Tobii Studio, we chose columns, which need to be exported from Tobii Studio Pro into Excel table for further analysis. Columns selected for data export are:

- *ParticipantName* – name of Participant associated with Tobii Studio recording
- *RecordingName* – name of the recording
- *RecordingDuration*
- *SegmentName* – name of the segment in the recording selected for export in the Select Segments for Export option.
- *SegmentStart* – the start time for each segment in each recording. The time value is calculated relative to the recording timeline
- *SegmentEnd* – the end time for each segment in each recording. The time value is calculated relative to the recording timeline
- *SegmentDuration* – duration of each segment in the recording
- *RecordingTimestamp* – timestamp counted from the start of the recording
- *MouseEvent* – mouse click type: Left button and Right button
- *FixationIndex* – represents the order in which a fixation (i.e. fixations and saccades). The index is autoincrement number starting with 1 (first fixation detected)
- *GazeEventType* – type of eye movement event classified by the fixation filter settings applied during the gaze data export: Fixation, Saccade and Unclassified
- *GazeEventDuration* – duration of eye movement event in miliseconds.

- *AreasOfInterest (AOI)* – represents whether the AOI is active and whether the fixation is located inside of the AOI: -1 = AOI is not active; 0 = AOI is active, the fixation is not located in the AOI; 1 = AOI ative and the fixation is located inside of the AOI. Enabling this column generates one column per AOI.

After defining columns data from Tobii Pro Studio was exported as an Excel sheet in .xlsl format. Obtained table had a raw data, so we needed to define which points will be analyzed during the research.

4.2.1 **Object recognition and tracking by people**

First question of the research is that people are faster in single object recognition and tracking single object over time. For verification of this statement, we have chosen video 4 (Figure 13). Video 4 was recorded outdoors under the daylight; on a video we can see shades of the objects and people are moving in a distance from the camera. Video 4 has a good quality and high resolution. Objects on the video appeared on gradually so it was possible to measure the time. On the video, we took a 14 sec time segment and marked six areas of interest: three men and three guns which they are carrying with them. After that we measured a time when all these objects appeared for the first time. Next step was to find a timestamp when it was noticed by AI and by participants separately. As we had raw data, time was measured in seconds. For data unification we converted all values to milliseconds to have more precise values. After we calculated differences between the time when the object appeared for the first time and when it was recognized by people and AI. Here we were interested only in AOI with guns, not people.

We analyzed videos from 8 participants which had a good quality of the recordings and recorded eye data. Out of 8 participants and 3 different tracked objects we received 24 cases for research analysis. In 17 cases people were faster in object recognition, in 4 cases objects were firstly noticed by AI and in 3 cases objects were not recognized at all. On average AI needs 9146 ms for object recognition while people need 3311 ms. Maximum time needed for object recognition was 14307 ms for AI and 14686 for human; minimum is 3985 ms for AI and only 77 ms for human. These calculations proved the first assumption of our work that people are faster in single object recognition and tracking it over time. All the calculations are shown in a table (Figure 32) in Appendix.

4.2.2 **Object recognition and tracking by AI**

Second question of the research is that AI is better in multiple objects tracking at the same time, since people are switching their attention and gaze from one object to another. For verification of this statement we used GazePlot visualization made in Tobii Studio Pro. GazePlot visualizations were done for each participant and can be found in section 7 Appendix of the research (Figure 24 - Figure 31). For verification we have chosen Video 3. Video 3 was recorded indoor in a full light meeting room and on high resolution camera. This video was analyzed and labeled by the Darvis AI. In this video, we marked seven areas of interest: three men, three guns and a mobile phone.

Gaze plots (Figure 24 - Figure 31) show the location, order, and time participants spent looking at locations on the video. The primary function of the gaze plot is to reveal the time sequence of looking or where participant look and when he looks there. Time spent looking, most commonly expressed as fixation duration, is shown by the diameter of the fixation circles. The longer the look, the larger the circle.

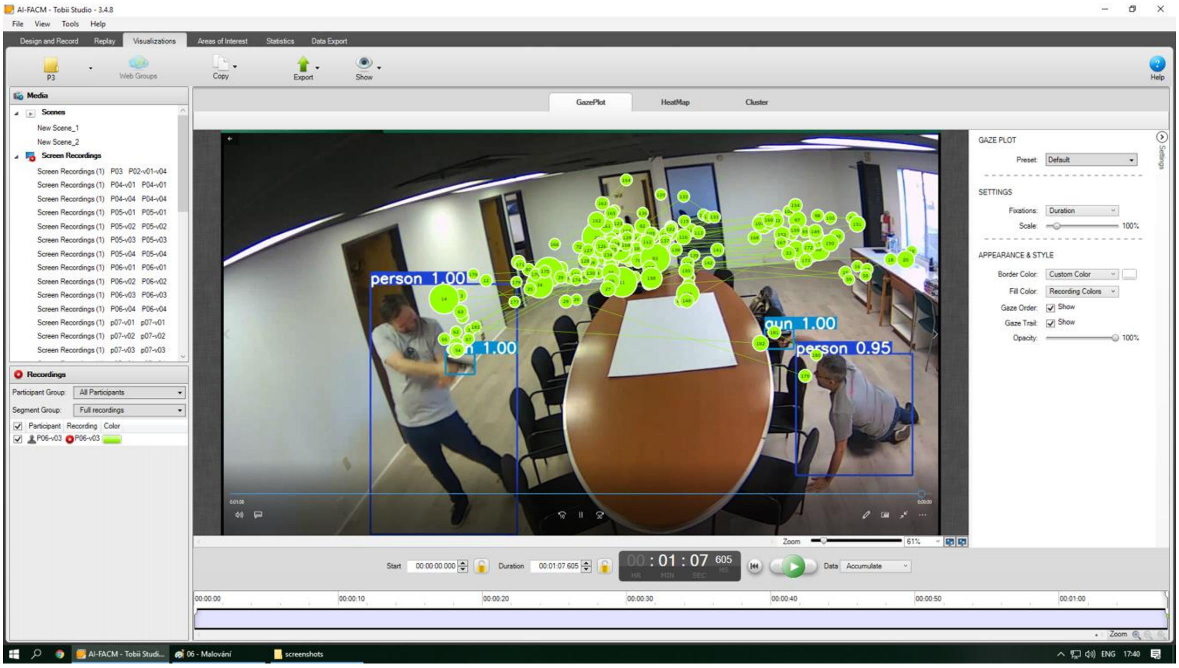

Figure 15 GazePlot

From the GazePlot example above (Figure 15) it is visible that each participant has more than hundred fixation points and saccades between them during the 28 sec time segment interval. Based on a fixation circle diameter we can understand how much time participant spent on a point. As people were asked to track all guns on the video, they needed to switch their attention between the objects. By trying to track down multiple objects at a time, participants were not able to focus.

At the same time, AI kept labels on objects which were recognized and classified. It is possible to see that labels were kept for multiple objects at the same time on a screen which means that AI was able to keep track of those objects simultaneously. Labels were also kept during the movements of objects and classification probabilities were recalculated in real time.

These observations have provided needed evidence to address the second research. Therefore, we can argue that AI is better in multiple objects tracking at the same time, since people are switching their attention from one object to another.

4.3 Statistical analysis

Data analysis for current research was done in IBM SPSS Statistics application, statistical software platform developed by IBM. SPSS is a full-featured data analysis program that offers a variety of applications such as statistical analysis, graphics, reporting, and database management. It is one of the most popular statistical packages which can perform highly complex data manipulation and analysis with simple instruction. SPSS is widely used for statistical analysis and data management by researchers and scientists.(Lal 2012). Data for statistical verification was exported in .xlsx format for Excel from Tobii Pro Studio.

4.3.1 Average time needed for participants and AI to recognize objects

For verification of the first research question which says that people are faster in single object recognition and tracking it over the time it was decided to choose Video 4 because objects appeared gradually one by one and it was possible to measure the time for participants and for AI.

Video 4 was recorded outdoor under the daylight in an training location staged as a ruined city (Figure 13). People on a video are moving in a distance from the camera and all objects have shades. The video has 3 areas of interest with guns, but AI was not able to detect Gun 3, so AOI Gun 3 it will not be analyzed during the research.

The video was previously analyzed by Darvis AI so it was possible to measure the time needed for AI to recognize and label an object. We selected 28 sec time interval and measured a time needed for AI and for participants to recognize on object in milliseconds This was done by analyzing data from the area of interests and time segment that were previously created in the Tobii Studio program. This data is shown in a Figure 16 below. For the AOI Person Gun2 there are several missing values shown as "." values in the table. For the calculation purposes they were automatically replaced by SPSS.

AIgun1	AIgun2	AIgun3	PersonGun1	PersonGun2	PersonGun3
3985	14307	.	4672	1014	7759
3985	14307	.	2279	1507	2412
3985	14307	.	7448	5046	.
3985	14307	.	14686	.	.
3985	14307	.	3177	3567	.
3985	14307	.	190	621	123
3985	14307	.	1970	2930	2770
3985	14307	.	666	.	77

Figure 16 Average time of looking on AOI in milliseconds - Video 4

Data analysis again starts with descriptive statistics (Figure 17) for our data shown on .Figure 16. The first column shows the numbers of valid data and missing data. From the table, we could conclude that there are 8 valid data for AI Gun 1 and 8 for AI Gun 2, 8 for Person Gun 1 and only 6 for Person Gun 2. Gun 3 data will not be analyzed, because it was not detected by AI.

From the second and third column we can see minimum and maximim values. We see that the lowest time of 621 ms and highest of 14,686 ms was recorded on participants, not AI.

Third column shows mean, or average value for each column. For our data the lowest mean 2,447.50 ms has Person Gun 2 and the highest 14,307 ms has AI Gun 2. There is a significant difference between mean values. That means that participants were almost 6.5 times faster than AI in recognition of Gun 2 object on video in average.

The last column standard deviation is a measure of the amount of variation or dispersion of a set of values. For our data std. deviation values are high for participants which means that participant data is not distributed near the mean value. AI standard deviation equals 0, that means that every data value is equal to the mean. The sample standard deviation of a data set is zero if and only if all of its values are identical.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
AIgun1	8	3985	3985	3985.00	.000
AIgun2	8	14307	14307	14307.00	.000
AIgun3	0				
PersonGun1	8	190	14686	4386.00	4761.912
PersonGun2	6	621	5046	2447.50	1703.752
PersonGun3	5	77	7759	2628.20	3129.561
Valid N (listwise)	0				

Figure 17 Descriptive Statistics - Video 4

SPSS offers two statistical tests of normality – the Kolmogorov-Smirnov test and the Shapiro-Wilk test. If the significance value is greater than the alpha value ($p < 0.05$), then there is no reason to think that our data differs significantly from a normal distribution, so we can reject the null hypothesis that the values have non-normal distribution. A complication that can arise here occurs when the results of the two tests don't agree – that is, when one test shows a significant result, and the other doesn't. In this situation, use the Shapiro-Wilk result – in most circumstances, it is more reliable.

As our data for AI has identical values and standard deviation equals 0, normality tests can not be done for them. Kolmogorov-Smirnov test indicated that observations for people in Figure 16 are not normally distributed ($p < 0.05$). In the case of small data sets, a test of significance for normality may lack the power to detect the deviation of the variable from normality (Figure 18). A dependent samples t-test can not be used for this case, as the distribution does not approximate a normal distribution. So the equivalent non-parametric Wilcoxon signed-rank test for statistical hypothesis testing was used.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
AIgun1	.	4	.	.	4	.
AIgun2	.	4	.	.	4	.
PersonGun1	.250	4	.	.963	4	.797
PersonGun2	.254	4	.	.912	4	.495
PersonGun3	.311	4	.	.907	4	.469

a. Lilliefors Significance Correction

Figure 18 Kolmogorov-Smirnov normality test - Video 4

The Wilcoxon sign test is a statistical comparison of average of two dependent samples. The Wilcoxon sign test works with metric (interval or ratio) data that is not multivariate normal, or with ranked/ordinal data. Generally it the non-parametric alternative to the dependent samples t-test. (Rosner et al. 2006). The Wilcoxon sign test tests the **null hypothesis** that there is no significant difference in average time needed for people and AI to recognize object on a video.

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
PersonGun1 - AIgun1	Negative Ranks	5[a]	4.20	21.00
	Positive Ranks	3[b]	5.00	15.00
	Ties	0[c]		
	Total	8		
PersonGun2 - AIgun2	Negative Ranks	6[d]	3.50	21.00
	Positive Ranks	0[e]	.00	.00
	Ties	0[f]		
	Total	6		

a. PersonGun1 < AIgun1

b. PersonGun1 > AIgun1

c. PersonGun1 = AIgun1

d. PersonGun2 < AIgun2

e. PersonGun2 > AIgun2

f. PersonGun2 = AIgun2

Test Statistics[a]

	PersonGun1 - AIgun1	PersonGun2 - AIgun2
Z	-.420[b]	-2.201[b]
Asymp. Sig. (2-tailed)	.674	.028

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

Figure 19 Wilcoxon Signed Ranks test Video 4

Based on Wilcoxon Signed Ranks Test ($p < .05$). shown on a Figure 19 we see that there is no significant difference between time needed for people and AI to recognize a Gun 1 object and there is a significant difference in recognizing Gun 2. So our null hypothesis partially proved for the first case: Gun 1 and rejected for second case: Gun 2.

From Figure 32 AI an human reaction comparising table we can see that with concrete values people are faster than AI, but from statistical point of view there is no significant difference for recognition of Gun 1 and there is statistical difference for Gun 2.

4.3.2 Average time spent on objects by participants

For verification of the second research question which says that AI is better in multiple objects tracking at the same time, since people are switching from one object to another. it was decided to choose Video 3.

Video 3 had 4 areas of interest: 3 guns and 1 phone. By analyzing data from these AOIs, we were able to check the assumption. We selected 18 sec time interval and measured an average time that participants spent looking on AOIs.

For statistical data analysis of Video 3 we took an average time in milliseconds for each participant when he was looking on a particular object. This was detected by analyzing data from the area of interests that were previously created in the Tobii Studio program. This data is shown in a Figure 20 below. For the AOI Gun2 and Phone there are several missing values shown as 0 values in the table, after they are automatically replaced by the SPSS program during calculations.

Participant	Gun1	Gun2	Gun3	Phone
2	168	207	301	226
3	1250	1169	632	0
5	930	0	289	0
6	2766	341	580	0
7	527	294	498	392
8	522	253	1089	212
9	540	557	439	859
10	266	558	217	267

Figure 20 Average time spent in the AOIs in milliseconds - Video 3

Data analysis starts with descriptive statistics presented on Figure 21 . The first column shows the numbers of valid observations and missing observations. From the table, we could conclude that there are 8 valid observations for gun 1, 7 for gun 2, 8 for gun 3 and only 5 for phone. People were asked to look at guns and it explains why we have 3 missing values for the phone.

From the second and third column we can see minimum and maximum values. We see that the lowest time of 168 ms was needed to one of the participants to notice the first gun and in opposite the highest time of 2766 was also detected on the same object. Third column shows the mean value for each column. In the sample, the lowest mean has the phone (391.27 ms) and the highest has the gun 1 (871.16 ms). It means that participants were looking at a gun 1 in average two times longer than on other objects.

The last column standard deviation is a measure of the amount of variation or dispersion of a set of values, for our data the standard deviation values are high which means that the data is not distributed near the mean value.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Participant	8	2	10	6.25	2.816
Gun1	8	168	2766	871.16	840.737
Gun2	7	207	1169	482.96	333.261
Gun3	8	217	1089	505.64	277.308
Phone	5	212	859	391.27	271.038
Valid N (listwise)	5				

Figure 21 Descriptive Statistics - Video 3

Kolmogorov-Smirnov test indicated that observations in Figure 22 are not normally distributed ($p < 0.05$). In the case of small data sets, a test of significance for normality may lack the power to detect the deviation of the variable from normality (Figure 22). The research data violate the normality assumption on which the t-test relies, for this reason t-test cannot be used here. Therefore, an equivalent non-parametric Wilcoxon signed-rank test for statistical hypothesis testing was used.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Gun1	.350	5	.045	.787	5	.063
Gun2	.280	5	.200*	.809	5	.096
Gun3	.313	5	.124	.837	5	.156
Phone	.299	5	.165	.753	5	.032

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure 22 Kolmogorov-Smirnov normality test Video 3

The Wilcoxon sign test is a statistical comparison of average of two dependent samples. This test tests the **null hypothesis** that there is no significant difference in average time which people spend looking on some AOI.

Test Statistics[a]

	Gun2 - Gun1	Gun3 - Gun1	Phone - Gun1	Gun3 - Gun2	Phone - Gun2	Phone - Gun3
Z	-.845[b]	-1.260[b]	-.135[c]	-.169[c]	-.405[c]	-.674[b]
Asymp. Sig. (2-tailed)	.398	.208	.893	.866	.686	.500

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

c. Based on negative ranks.

Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
Gun2 - Gun1	Negative Ranks	4[a]	4.75	19.00
	Positive Ranks	3[b]	3.00	9.00
	Ties	0[c]		
	Total	7		
Gun3 - Gun1	Negative Ranks	6[d]	4.50	27.00
	Positive Ranks	2[e]	4.50	9.00
	Ties	0[f]		
	Total	8		
Phone - Gun1	Negative Ranks	2[g]	3.50	7.00
	Positive Ranks	3[h]	2.67	8.00
	Ties	0[i]		
	Total	5		
Gun3 - Gun2	Negative Ranks	3[j]	4.33	13.00
	Positive Ranks	4[k]	3.75	15.00
	Ties	0[l]		
	Total	7		
Phone - Gun2	Negative Ranks	2[m]	3.00	6.00
	Positive Ranks	3[n]	3.00	9.00
	Ties	0[o]		
	Total	5		
Phone - Gun3	Negative Ranks	3[p]	3.33	10.00
	Positive Ranks	2[q]	2.50	5.00
	Ties	0[r]		
	Total	5		

a. Gun2 < Gun1
b. Gun2 > Gun1
c. Gun2 = Gun1
d. Gun3 < Gun1
e. Gun3 > Gun1
f. Gun3 = Gun1
g. Phone < Gun1
h. Phone > Gun1
i. Phone = Gun1
j. Gun3 < Gun2
k. Gun3 > Gun2
l. Gun3 = Gun2
m. Phone < Gun2
n. Phone > Gun2
o. Phone = Gun2
p. Phone < Gun3
q. Phone > Gun3
r. Phone = Gun3

Figure 23 Wilcoxon Signed Ranks Test - Video 3

The Wilcoxon Signed Ranks test output contains only two tables. (Figure 23) The first table contains all statistics that are required to calculate the Wilcoxon Signed Ranks test. These are the sample size and the sum of ranks. It also includes the mean rank, which is not necessary to calculate but it helps with the interpretation of the data.

The sample contains the total number of observations which were made for each pair of objects. The Wilcoxon Signed Ranks Test answers the question if the difference is significantly different from zero, and thus if the observed difference in Mean Rank was calculated.

The second table Test Statistics contains the test of significance statistics. The SPSS output contains the z-value for all our pairs. The test value z is approximately normally distributed for small samples that are $n<10$, so that $p = 0.05$.

To sum up statistical results, the Wilcoxon signed rank test shows that the observed difference between all pairs of measurements is not significant. Thus, we can accept the null hypothesis that there is no significant difference in average time which people spend looking on some AOI.

4.4 **Summary of statistical analysis**

In section 3.6 we formulated two research questions. Based on data collected in Tobii Studio Pro and exported to Excel we were able to make a statistical verification of them in SPSS statistical application.

For verification of the first research question which says that people are faster in single object recognition and tracking it over time it was decided to choose Video 4 because objects appeared gradually one by one and it was possible to measure the time for participants and for AI. Based on the Wilcoxon Signed Ranks Test it is possible to conclude that there is no significant difference between time needed for people and AI to recognize a Gun 1 object and there is a significant difference in recognizing Gun 2. So, our null hypothesis partially proved for the first case: Gun 1 and rejected for second case: Gun 2. On a Figure 32 AI an human reaction comparising table we see that with concrete values people are faster than AI, so in general the first research question was proved from both points of view.

For verification of the second research question which says that AI is better in multiple objects tracking at the same time, since people are switching from one object to another. it was decided to choose Video 3, where we had multiple AOIs. From the Wilcoxon signed rank test it was possible to accept that there is no significant difference in average time which people spend looking at some AOI. Also from Gaze plots analysis we conclude that people were constantly switching their attention between AOIs, so the second research question was also proved from both points of view.

5 Results and Discussion

Based on literature review made in the first part of the thesis it is possible to conclude that AI can be effectively used for crisis management. Intelligent security systems are a powerful tool for securing our daily life and controlling what is happening around.

With current technology development intelligent security systems can perform video analysis on real time. They are able to detect suspicious events and alarm it for the people. Security systems are using powerful hardware which is now not very expensive and are able to do calculations on a device itself without transferring data to the server and back. For example Darvis software solutions use this technique (Darvis 2015).

AI makes automatic security systems intelligent. It performs data analysis to discover important patterns in real time. It uses object recognition algorithms for detecting objects and people on videos and images. (Ben Mabrouk a Zagrouba 2018) AI can also help in analysis not real-time streams but also previously recorded footage. Like BriefCam does, people put features that should be found in a recorded footage, AI analyses and shows the results. (BriefCam 2008) It can be useful in a case if some crisis or crime already happened and people need to understand what has happened, why and the main question is such situations could it be possible to prevent such an event in future. AI is able not only recognize patterns but also remember them and use them for future analysis.

5.1 AI in smart surveillance

Smart surveillance is based on video analysis. It is very important for the security of any public place. Smart surveillance systems nowadays use automatic video analysis technologies in real-time recording footage. With the help if AI they can also isolate individuals or objects, track them and detect suspicious patterns

With current development of AI technologies video surveillance reached a completely new level. Automated surveillance became an intelligent one. Now CCTV enables not only to view and record events but analyze them in real time and even predict future ones. Video analytics automatically captures a wide variety of data about the areas covered by cameras. The captured information can be used by AI to recognize crime patterns in real time videos and transfer this information to police or other authorities.

Important question for smart surveillance is data privacy. These systems need to comply with GDPR regulation in Europe; CCPA act in California and other GDPR-like regulations in different countries. Security cameras cannot just be installed in public places they need to have special approval for that and reason why to be installed. Because of that smart security systems may face the problem that they do not have access to needed data and cannot access it. In some cases, surveillance violates people's privacy and cannot perform the analysis. This can be avoided if all used data will be anonymized and checked only for patterns. In this case police or authorities will receive an alarm but without concrete personal data. Also, data used for smart surveillance should not be stored after performing analysis or deleted after a needed amount of time if it was agreed. So, it is not possible to implement AI recognition everywhere it should comply with regulations for different areas.

5.2 AI for crisis prediction

Video analytics automatically captures a wide variety of data about the areas covered by smart camera networks. The information captured is now used for predictive policing.

Based on the collected data, intelligent systems try to predict and police crime before it happens. Real time video analysis is used to detect crime patterns for the police.

To predict a crisis, we need to determine several parameters, the place, time and type of future event, find the victims of it and specify the identities of the criminals. Many crimes can be predicted based on statistical data, but if it is a fast scene happening or bad recording conditions the practical experiment showed that it is not optimal. In this case people also need to take part in situation solving. As we understood from the practical part of the research people are better at recognizing objects when the camera is far from the objects, when it is dark and the recording quality is not good.

Based on meteorological information, intelligent systems can try to predict natural disasters, with information from the machines in a factory they can understand technological breakdown which can possibly lead to crisis. For each type of occurring event we need special data, it cannot be the same for all cases.

For precise crisis prediction systems require huge amounts of data to be trained on and tested. Only after that they can try to work with real-time happening situations. So, in the case that some event occurs rarely it is almost impossible to predict it due to a lack of data. It can be possible only if it has some very specific patterns which can be easily detected. Precise recognition requires precise data and sometimes it is not possible to obtain it. (Kang a Kang 2017)

5.3 Intelligent security systems

Intelligent security used for safety for public places nowadays. Intelligent security systems are able to automatically analyze recorded image, video, audio or other types of surveillance data with limited human intervention or even without it. The recent innovations in sensor technologies, computer vision, and machine learning have an important role in faster development of such intelligent systems. These systems can use many types of recorded data like image, video, audio, temperature and pressure from sensors for more complex and precise analysis.

In this study we made an overview of four most promising software solutions used for public security. Each of them has different purposes and unique features. For example, Darvis is able to make video content analysis in real-time making the calculations on the device where it runs, without data transferring. BriefCam can be extremely useful for recorded video analysis to understand what happened before some crisis happened. PredPol and Palantir are used for future crime prediction based on previously collected statistical data. PredPol experience can be broadened from crime prediction to general crisis prediction tool based on the same previously collected data.

One of the advantages of Intelligent Security Systems is that they can monitor the entire situation in a selected region, not only crimes, but they can try to predict natural disasters or technical breakdowns. If those systems will be provided with needed data, they can be trained in advance and expanded.

One of the promising features of such systems can be to detect behavioral patterns and stop people from doing something unwanted before. Recent research showed that such systems under development now and possibly can make a huge change in understanding people’s emotions and gestures (Ben Mabrouk a Zagrouba 2018)

5.4 **Object recognition by humans and AI**

The purpose of this research was to compare artificial computer intelligence with human intelligence in terms of time and accuracy needed for object recognition and tracking. It was an experimental research study that tracked people's eye movements using special screen based Tobii device and compared these results with AI.

In this study participated 10 people who were watching 5 different videos. This number of research participants was minimal for statistical verification of research assumptions. Due to the current COVID-19 crisis situation in the world it was decided to involve the smallest possible number of people to take measures and prevent possible virus spreading. On each video it was a number of people and guns and participants were asked to click on them at the time of first notice. Once people clicked on them a relative timestamp of the event was recorded.

In section 3.6 we formulated two main questions for the practical part of this research. The first research question says that people are faster in single object recognition and tracking it over time. The second question says that AI is better in multiple objects tracking at the same time, since people are switching from one object to another.

These two questions were discussed in detail in the practical part of the research. The research has analyzed the speed and accuracy of artificial intelligence and human intelligence in object detection, object recognition and object tracking. For verification of each research question, we applied two different approaches. We tested them from a statistical point of view and made and analysis of absolute values for the first case and Gaze plot analysis for the second one.

Based on the Wilcoxon‘s signed rank test it is possible to conclude that there is no significant difference between time needed for people and AI to recognize a Gun 1 object and there is a significant difference in recognizing Gun 2. So, our hypothesis was partially proved for the first case (Gun 1) and rejected for the second case (Gun 2). But at the same time, we also analyzed absolute values of needed time for people and AI. Out of 8 participants and 3 different tracked objects we received 24 cases for research analysis. In 17 cases people were faster in object recognition, in 4 cases objects were firstly noticed by AI and in 3 cases objects were not recognized at all not by people, not by AI. This result has proved the first question of the research that people are faster in single object recognition and tracking it over time.

From statistical analysis made for the second research case we conclude that there is no significant difference between average time that people spend focusing on some AOI. It can happen because observing one object interferes with observing another object. As all objects were on a screen at the same time, people were constantly switching between them and changing their gaze from one to another. Based on the gaze plots in the section 7 Appendix and an animated video attached to the thesis document, it is possible to see that people moved their gaze more than hundred times during the 28-second long recording. It is clear that people cannot focus their attention on multiple objects at the same time, while AI is keeping the object label during at all times when the object is visible. From this point, it is possible to say that AI can be more effective in multiple objects tracking than people.

According to the average time participants spent on some objects it is possible to conclude that people were mainly focused on the object they noticed first. Gun 1 appeared before Gun 2 and Gun 3, therefore people were following it for a longer period of time than other objects. This is an interesting finding because people were asked to follow all the objects equally.

To sum up, the major conclusion of this thesis is that the obtained results have demonstrated that people are faster in single object recognition and tracking single objects over the time than AI, while AI is better in multiple objects tracking at the same time. The reason for this is that people are switching their attention and gaze from one object to another. We also came to the conclusion that AI currently cannot fully replace people. The most efficient way to how to use AI for crisis management is an automatic system that works together with people. If an intelligent system detects some suspicious event, a human operator should receive an alarm and based on brief analysis provided by the system, decide what to do next. Based on collected information during some time, intelligent systems can also suggest which steps will be most efficient for current situations, but the last word and decision should be on people.

6 Conclusion

The main objective of this thesis was to understand how to use AI in a crisis situation in public places. For this purpose, the literature review of the applications of artificial intelligence for solving crisis situations in public spaces was conducted. Theoretical part of the research is focused on next topics: artificial intelligence, crisis management, smart surveillance and prediction of future crimes. As a term of crisis management includes many techniques, it was decided to study in detail smart surveillance, which is powered by AI algorithms and can be done both on recorded video footage and in real time. Collected surveillance data can also be used for future crime prediction. A conclusion which was made based on the theoretical part is that with a smart security system installed in public places people can continue their daily life in the knowledge that every care is being taken to protect their safety. The way how data is collected and used now will provide many opportunities for innovation and deployment of advanced technology in future. With current technology development it has become obvious that digital revolution will transform the way how people live.

One of the partial objectives was to make an overview of common problems of AI and how those problems affect detection of behavioral and social patterns in public spaces. A brief summary of crisis management techniques used nowadays was done in the section 3.1 and 3.2. It includes social media analysis, tracking applications, videos recorded from unmanned aerial vehicles and others.

Another goal of this thesis was to review and compare existing software solutions used for crisis management. Four existing, most promising systems were studied: Darvis, BriefCam, PredPol and Palantir. All of those systems have a common application domain, they are used for security of public places and use AI algorithms. On the other hand, these systems have different use cases, so it was not possible to compare them according to the same terms. For example, Darvis is an intelligent system for early threat detection. BriefCam is a complete video content analytics platform. PredPol uses predictive analytics to support law enforcement in the US. Finally, Palantir provides a database for future crime prediction. All these solutions are useful for complex crisis management.

A comparison of artificial intelligence and human reactions in terms of time and accuracy to identify, recognize and track the object was conducted and described in the practical part. The major conclusion from the practical part of the research is that people are faster in single object recognition and tracking single objects over the time and that AI is better in multiple objects tracking at the same time, since people are switching their attention and gaze from one object to another.

Due to COVID-19 crisis situation, current research was made under a number of limitations. Among them, the research was made on a small group of 10 participants, 3 out of 5 recorded videos were labeled by AI in advance, only 2 scenes were analyzed for verification of the research assumptions. However, the small dataset could be used for statistical analysis with a decent level of reliability.

The findings of the research have a practical implication. For the use of AI in crisis management, such as security of public places, object recognition cannot be fully deployed in suboptimal conditions such as low resolution, low light, distant moving objects, etc. In these scenarios, human operators have much better capability to detect and recognize a potentially dangerous object.

Future development of this study can be the creation of complex AI based solutions used for the crisis management system of the whole city. It can include crime detection, disaster monitoring, management of traffic, control of people behavior connected together. In the case of detection some abnormal event system will alarm, and people will decide how to solve it in a most efficient and quick way. When security personnel will receive a notification, they will still have enough time to minimize the consequences or prevent crisis happening. Such a system can be useful for any modern city.

References

ALLERIN, 2019. *Drones for disaster response and management* [online] [vid. 2020-11-11]. Dostupné z: https://www.allerin.com/blog/drones-for-disaster-response-and-management
AYOBAMI, Akanmu Semiu a Shamsuddeen RABI'U, 2012. SMS as a Rural Disaster Notification System in Malaysia: A Feasibility Study. *International Conference on Communication and Media* [online]. (1997). Dostupné z: doi:10.13140/2.1.2727.3925
BAIK, Jeeyun (Sophia), 2020. Data privacy against innovation or against discrimination?: The case of the California Consumer Privacy Act (CCPA). *Telematics and Informatics* [online]. **52**(May), 101431. ISSN 07365853. Dostupné z: doi:10.1016/j.tele.2020.101431
BARNOVICIU, Eduard, Veta GHENESCU, Serban Vasile CARATA, Marian GHENESCU, Roxana MIHAESCU a Mihai CHINDEA, 2019. GDPR compliance in video surveillance and video processing application. *2019 10th International Conference on Speech Technology and Human-Computer Dialogue, SpeD 2019* [online]. Dostupné z: doi:10.1109/SPED.2019.8906553
BATTY, M., K. W. AXHAUSEN, F. GIANNOTTI, A. POZDNOUKHOV, A. BAZZANI, M. WACHOWICZ, G. OUZOUNIS a Y. PORTUGALI, 2012. Smart cities of the future. *European Physical Journal: Special Topics* [online]. **214**(1), 481–518. ISSN 19516355. Dostupné z: doi:10.1140/epjst/e2012-01703-3
BBC, 2020. *George Floyd: What happened in the final moments of his life - BBC News* [online] [vid. 2020-11-23]. Dostupné z: https://www.bbc.com/news/world-us-canada-52861726
BELLEFLAMM, Paul; a Martin PEITZ, 2017. Administrative Law Encyclopedia ofLaw and Economics. *Encyclopedia of Law and Economics* [online]. **52**(473), 1–11. Dostupné z: doi:10.1007/978-1-4614-7883-6
BEN MABROUK, Amira a Ezzeddine ZAGROUBA, 2018. Abnormal behavior recognition for intelligent video surveillance systems: A review. *Expert Systems with Applications* [online]. **91**, 480–491. ISSN 09574174. Dostupné z: doi:10.1016/j.eswa.2017.09.029
BIGO, Vinca a Ioana NEGRU, 2014. Mathematical modelling in the wake of the crisis: A blessing or a curse? What does the economics profession say? *Cambridge Journal of Economics* [online]. **38**(2), 329–347. ISSN 14643545. Dostupné z: doi:10.1093/cje/bet063
BRIEFCAM, 2008. *Rapid review, search and analysis of video content | BriefCam* [online] [vid. 2020-11-11]. Dostupné z: https://www.briefcam.com/company/about/
BROWNLEE, Jason, 2019. *A Gentle Introduction to Object Recognition With Deep Learning* [online] [vid. 2020-11-27]. Dostupné z: https://machinelearningmastery.com/object-recognition-with-deep-learning/
BUNDY, Jonathan, Michael D. PFARRER, Cole E. SHORT a W. Timothy COOMBS, 2017. *Crises and Crisis Management: Integration, Interpretation, and Research Development* [online]. ISBN 0149206316680. Dostupné z: doi:10.1177/0149206316680030
CALIFORNIA, State O F a Electronic Legal MATERIAL, 2020. Senate Bill No. 1121. (1121).
CHOUDHURY, Ambika, 2020. *Top 8 Algorithms For Object Detection One Must Know* [online] [vid. 2020-11-27]. Dostupné z: https://analyticsindiamag.com/top-8-algorithms-for-object-detection/
CNBC, 2020. *Coronavirus: Singapore to make contact tracing tech open source* [online]

[vid. 2020-11-19]. Dostupné z: https://www.cnbc.com/2020/03/25/coronavirus-singapore-to-make-contact-tracing-tech-open-source.html
COUNCIL OF THE EUROPEAN UNION, 2015. Interinstitutional File: 2012/0011 (COD) [online]. **2015**(June), 201. ISSN 19454589. Dostupné z: doi:10.1017/CBO9781107415324.004
CRED, 2015. The Human Cost of Natural Disasters. *Centre for Research on the Epidemiology of Disasters CRED*. 58.
DARVIS, 2015. *About Us : Darvis - We make hospitals intelligent. Digitalization of clinical objects, processes and live analysis of situations. - Darvis* [online] [vid. 2020-11-11]. Dostupné z: https://www.darvis.com/about/
DHS, S, T PIA-, Robert D MIDDLETON a Mary Ellen CALLAHAN, 2011. Future Attribute Screening Technology (FAST)/ Passive Methods for Precision Behavioral Screening. (703).
DRESSEL, Julia a Hany FARID, 2018. The accuracy, fairness, and limits of predicting recidivism. *Science Advances* [online]. **4**(1), 1–6. ISSN 23752548. Dostupné z: doi:10.1126/sciadv.aao5580
EURONEWS, 2020. Before and after photos from Beirut’s port reveal scale of disaster | Euronews. *2020* [online] [vid. 2020-11-26]. Dostupné z: https://www.euronews.com/2020/08/06/before-and-after-photos-from-beirut-s-port-reveal-scale-of-disaster
GIRSHICK, Ross, Jeff DONAHUE, Trevor DARRELL, Jitendra MALIK a U C BERKELEY, 2012. Rich feature hierarchies for accurate object detection and semantic segmentation.
GU, Leon a Takeo KANADE, 2015. *Encyclopedia of Biometrics* [online]. ISBN 9781489974884. Dostupné z: doi:10.1007/978-1-4899-7488-4
GUARDIAN, The, 2014. *Facebook reveals news feed experiment to control emotions | Technology | The Guardian* [online] [vid. 2020-11-11]. Dostupné z: https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds
HERNANDO, A., D. VILLUENDAS, C. VESPERINAS, M. ABAD a A. PLASTINO, 2010. Unravelling the size distribution of social groups with information theory in complex networks. *European Physical Journal B* [online]. **76**(1), 87–97. ISSN 14346028. Dostupné z: doi:10.1140/epjb/e2010-00216-1
HOSSAIN, L., D. KAM, F. KONG, R. T. WIGAND a T. BOSSOMAIER, 2016. Social media in Ebola outbreak. *Epidemiology and Infection* [online]. **144**(10), 2136–2143. ISSN 14694409. Dostupné z: doi:10.1017/S095026881600039X
IBRAHIM, Sutrisno, 2016. A comprehensive review on intelligent surveillance systems. **1**, 7–14.
IVIĆ, M, 2019. Artificial intelligence and geospatial analysis in disaster management. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* [online]. **42**(3/W8), 161–166. ISSN 16821750. Dostupné z: doi:10.5194/isprs-archives-XLII-3-W8-161-2019
KANG, Hyeon-Woo a Hang-Bong KANG, 2017. Prediction of crime occurrence from multi-modal data using deep learning. *Plos One* [online]. **12**(4), e0176244. ISSN 1932-6203. Dostupné z: doi:10.1371/journal.pone.0176244.t005
KO, Teddy, 2008. A survey on behavior analysis in video surveillance for homeland security applications. *Proceedings - Applied Imagery Pattern Recognition Workshop* [online]. ISSN 15505219. Dostupné z: doi:10.1109/AIPR.2008.4906450

LAL, Krishan, 2012. An overview to spss.
MOHLER, George, Rajeev RAJE, Jeremy CARTER, Matthew VALASIK a Jeffrey BRANTINGHAM, 2019. A Penalized Likelihood Method for Balancing Accuracy and Fairness in Predictive Policing. *Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018* [online]. 2454–2459. Dostupné z: doi:10.1109/SMC.2018.00421
NORVIG, Peter, 2003. *Artificial Intelligence*. ISBN 0137903952.
PALANTIR, 2009. *Home | Palantir* [online] [vid. 2020-11-11]. Dostupné z: https://www.palantir.com/
PALANTIR, 2015. *Palantir Gotham screenshot - Awesome post - Imgur* [online] [vid. 2020-11-24]. Dostupné z: https://imgur.com/gallery/b85ds42
PREDPOL, 2010. *PredPol – The Predictive Policing Company* [online] [vid. 2020-11-11]. Dostupné z: https://www.predpol.com/
QI, Di a Andrew J. MAJDA, 2020. Using machine learning to predict extreme events in complex systems. *Proceedings of the National Academy of Sciences of the United States of America* [online]. **117**(1), 52–59. ISSN 10916490. Dostupné z: doi:10.1073/pnas.1917285117
REDMON, Joseph, Santosh DIVVALA, Ross GIRSHICK a Ali FARHADI, 2015. You Only Look Once: Unified, Real-Time Object Detection.
REUTERS, 2018. *China using water spray, facial recognition technology to stop jaywalkers - National | Globalnews.ca* [online] [vid. 2020-11-11]. Dostupné z: https://globalnews.ca/news/4169368/china-jaywalkers-water/
ROSNER, Bernard, Robert J. GLYNN a Mei Ling T. LEE, 2006. The Wilcoxon signed rank test for paired comparisons of clustered data. *Biometrics* [online]. **62**(1), 185–192. ISSN 0006341X. Dostupné z: doi:10.1111/j.1541-0420.2005.00389.x
SAROJ, Anita a Sukomal PAL, 2020a. Use of social media in crisis management: A survey. *International Journal of Disaster Risk Reduction* [online]. **48**, 101584. ISSN 22124209. Dostupné z: doi:10.1016/j.ijdrr.2020.101584
SAROJ, Anita a Sukomal PAL, 2020b. Use of social media in crisis management: A survey. *International Journal of Disaster Risk Reduction* [online]. **48**(April), 101584. ISSN 22124209. Dostupné z: doi:10.1016/j.ijdrr.2020.101584
SETO, Karen C., Burak GÜNERALP a Lucy R. HUTYRA, 2012. Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences of the United States of America* [online]. **109**(40), 16083–16088. ISSN 00278424. Dostupné z: doi:10.1073/pnas.1211658109
SORG, Evan T., Jennifer D. WOOD, Elizabeth R. GROFF a Jerry H. RATCLIFFE, 2017. Explaining Dosage Diffusion During Hot Spot Patrols: An Application of Optimal Foraging Theory to Police Officer Behavior. *Justice Quarterly* [online]. **34**(6), 1044–1068. ISSN 17459109. Dostupné z: doi:10.1080/07418825.2016.1244286
TOBII, 2016. Tobii Pro Studio eye tracking software, dedicated for UX. *Tobii* [online] [vid. 2020-11-11]. Dostupné z: https://www.tobiipro.com/product-listing/tobii-pro-studio/
VENTURE, Gentiane, Hideki KADONE, Tianxiang ZHANG, Julie GRÈZES, Alain BERTHOZ a Halim HICHEUR, 2014. Recognizing Emotions Conveyed by Human Gait. *International Journal of Social Robotics* [online]. **6**(4), 621–632. ISSN 18754805. Dostupné z: doi:10.1007/s12369-014-0243-1
VERGE, 2020. *Israel is using cellphone data to track the coronavirus - The Verge* [online] [vid. 2020-11-19]. Dostupné z: https://www.theverge.com/2020/3/17/21183716/coronavirus-covid-19-israel-natanyahu-

cellphone-data-tracking
WANG, Ting Chun, Ming Yu LIU, Jun Yan ZHU, Guilin LIU, Andrew TAO, Jan KAUTZ a Bryan CATANZARO, 2018. Video-to-video synthesis. *Advances in Neural Information Processing Systems*. **2018-Decem**, 1144–1156. ISSN 10495258.
WHO, 2020. *Listings of WHO's response to COVID-19* [online] [vid. 2020-11-19]. Dostupné z: https://www.who.int/news/item/29-06-2020-covidtimeline
YOUTUBE, 2017. *Cleveland gang-related shootout captured on video - YouTube* [online] [vid. 2020-11-11]. Dostupné z: https://www.youtube.com/watch?v=BR4KyIE3Mrg&list=WL&index=3&t=91s

7 Appendix

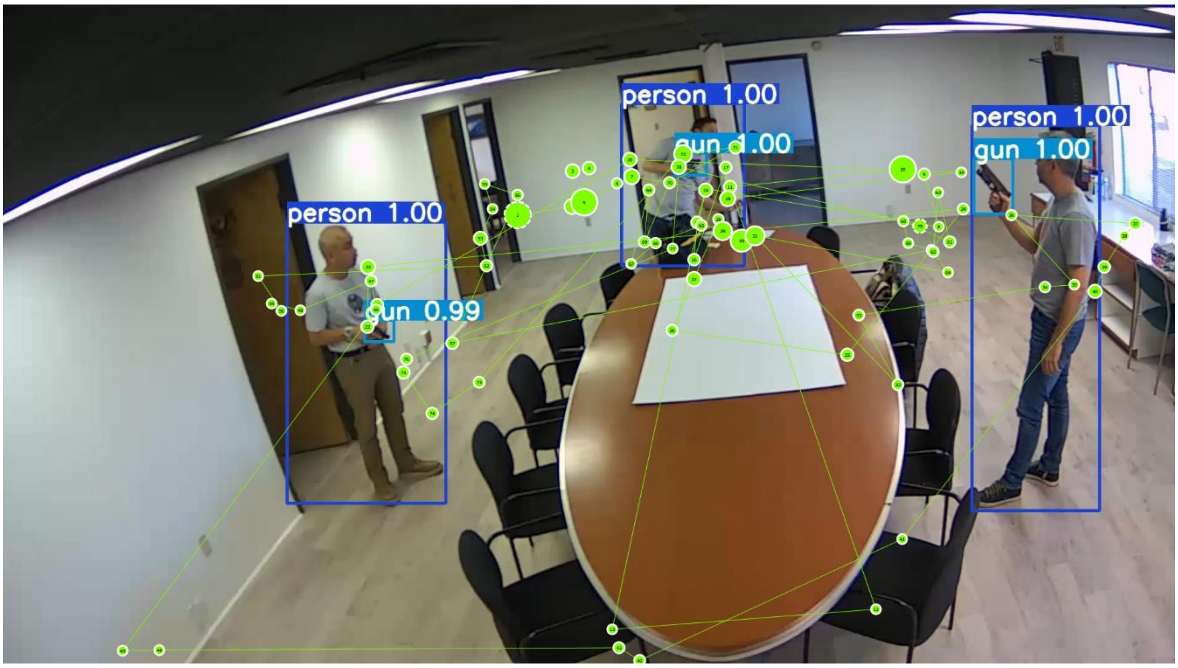

Figure 24 Participant 2 Video 3 Gaze Plot

Figure 25 Participant 3 Video 3 Gaze Plot

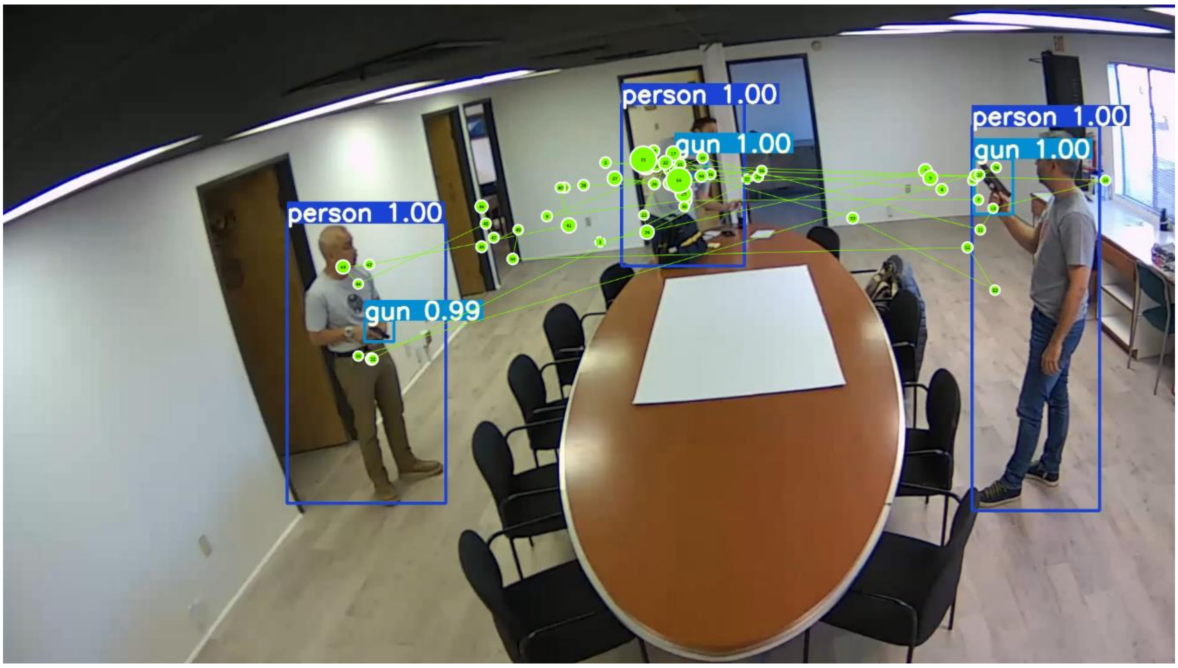

Figure 26 Participant 5 Video 3 Gaze Plot

Figure 27 Participant 6 Video 3 Gaze Plot

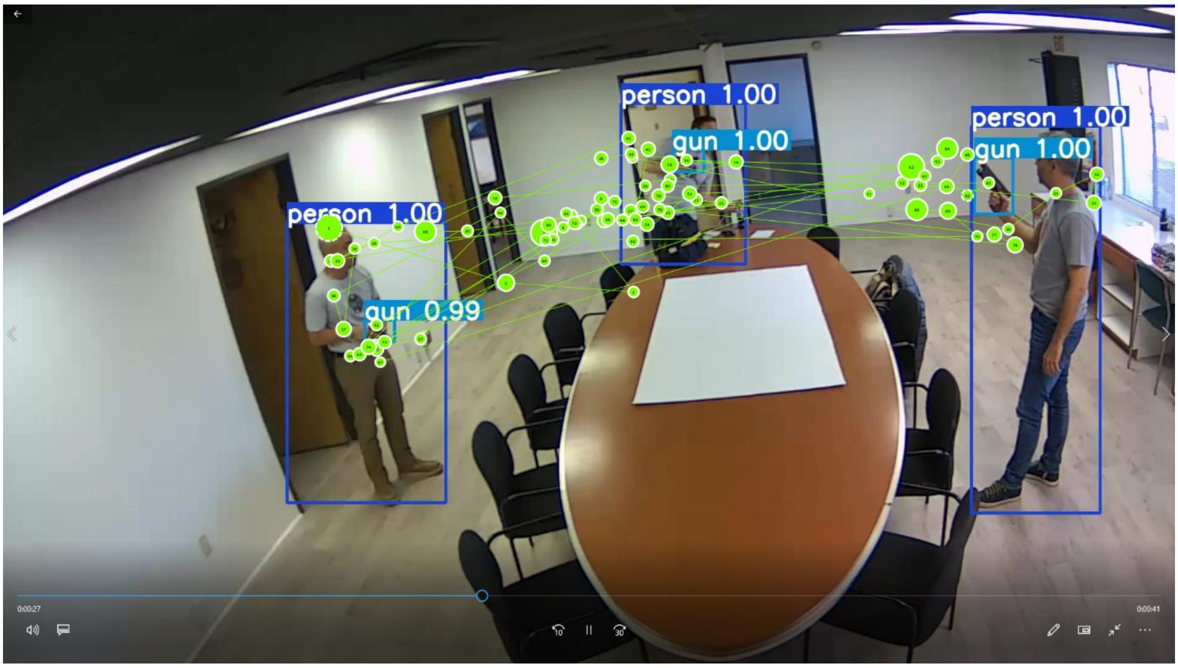

Figure 28 Participant 7 Video 3 Gaze Plot

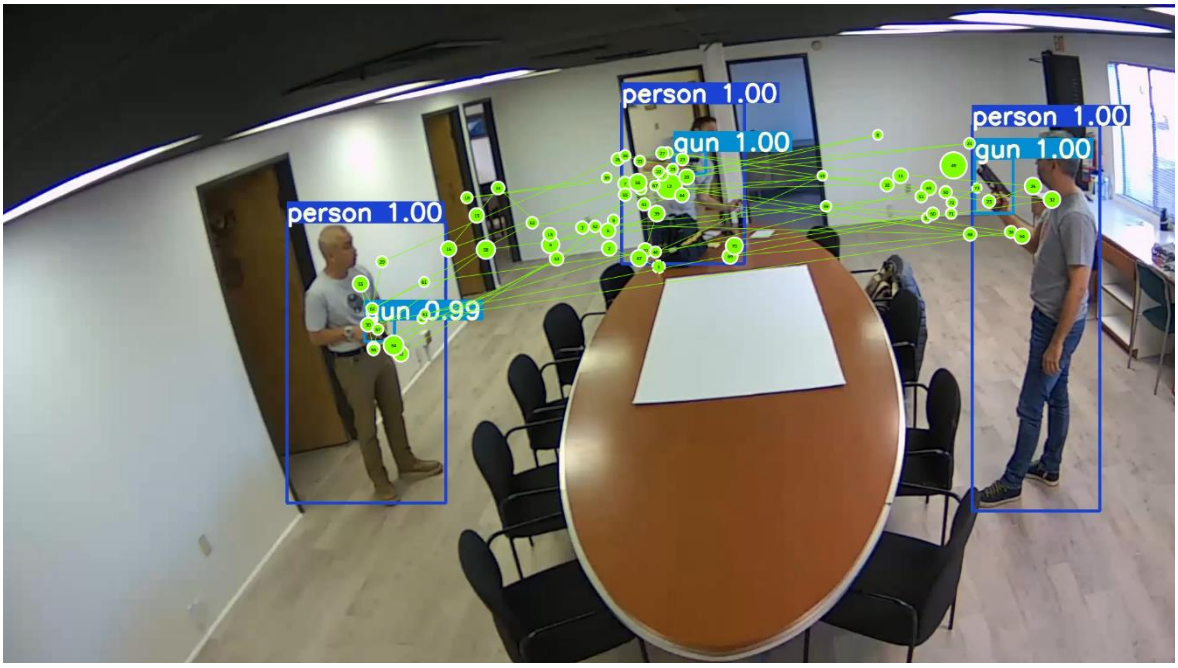

Figure 29 Participant 8 Video 3 Gaze Plot

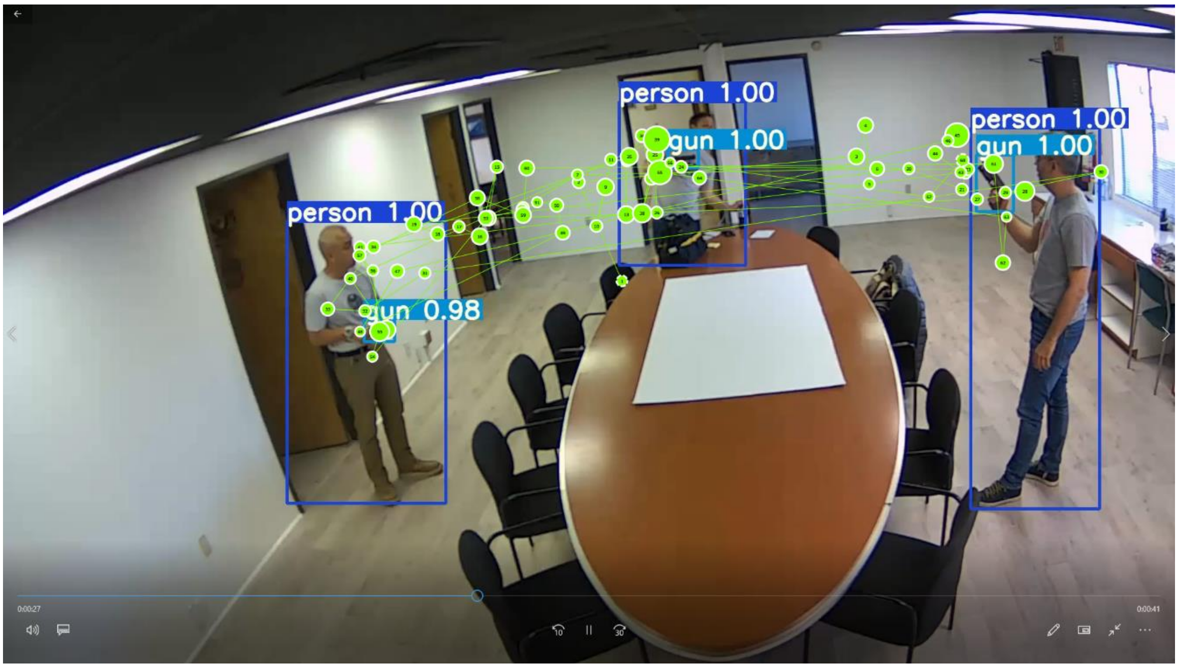

Figure 30 Participant 9 Video 3 Gaze Plot

Figure 31 Participant 10 Video 3 Gaze Plot

	A	B	C	D	E	F	G	H	I	J	K	L	M
1			**Time in seconds**										
2	**Video**	**Object**	**1st appeareance**	**AI noticed**	**P2**	**1st appearea**	**AI**	**P4**	**1st appea**	**AI**	**P5**	**1st appea**	**AI**
3	4	gun 1	2 53 197	2 57 182	2 57 869	00 03 436	00 08 641	00 05 715	00 26 279		00 33 727	00 06 149	
4	4	gun 2	2 54 343	3 08 650	2 55 357	00 06 049	*	00 07 556			00 32 471		
5	4	gun 3	3 01 816	*	3 09 575	00 13 617	*	00 16 029			*		
6													
7													
8		1 min	60000		1146								
9		1 sec	1000		7473								
10			**Time in Millisec**										
11			173197	177182	177869	3436		5715	26279		33727	6149	
12			174343	188650	175357	6049		7556	27425		32471	7295	
13			181816	*	189575	13617		16029	34898		*	14768	
14				**Difference**									
15				3985	4672		3985	2279		3985	7448		3985
16				14307	1014		14307	1507		14307	5046		14307
17				*	7759		*	2412		*	*		*
18													
19			Results										
20				8 participants				Average	Max	Min			
21				24 cases	17 human		AI	9146	14307	3985			
22					4 AI		Human	3311.263158	14686	77			
23					3 no result								
24													
25													

	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	**P6**	**1st appea**	**AI**	**P7**	**1st appea**	**AI**	**P8**	**1st appea**	**AI**	**P9**	**1st appea**	**AI**	**P10**
3	00 20 835	00 05 939		00 09 116	00 07 086		00 07 276	00 02 611		00 04 581	00 08 222		00 08 888
4	*			00 10 652			00 08 853			00 06 687	00 08 222		*
5	*			*			00 15 828			00 14 000			00 16 918
11	20835	5939		9116	7086		7276	2611		4581	8222		8888
12		7085		10652	8232		8853	3757		6687	9368		*
13				*	15705		15828	11230		14000	16841		16918
15	14686		3985	3177		3985	190		3985	1970		3985	666
16	*		14307	3567		14307	621		14307	2930		14307	*
17	*		*	*		*	123		*	2770		*	77

Figure 32 AI an human reaction comparising table